

Grupo B. Sistemas y Tecnologías de la información

- Tema 1. Sistemas de Información. Análisis y desarrollo de sistemas de información. Metodologías de análisis y desarrollo. Desarrollo estructurado. Desarrollo orientado a objetos. Metricav3. Técnicas para el análisis y desarrollo de sistemas de información. UML.
- Tema 2. Lenguajes de Programación. Tipos de lenguajes de programación. Programación estructurada. Programación orientada a objetos.
- Tema 3. Arquitectura SOA. Entornos de desarrollo Web: Microsoft .NET y J2EE.
- Tema 4. Almacenamiento estructurado. Sistemas de gestión de bases de datos SGBD. Bases de datos relacionales. Bases de datos orientadas a objetos. Bases de datos espaciales.
- Tema 5. Normas y estándares para la interoperabilidad entre gestores de bases de datos relacionales. Lenguaje SQL. Componentes de SQL: LDD, LMD.
- Tema 6. La descripción e intercambio de datos e información. XML. Estructura de un documento. DTD y XSD. Tecnologías XML. XSL, XLINK.
- Tema 7. Sistemas de Información Geográfica. Definición. Componentes de un SIG. Tipos de SIG. Organización e implementación de un SIG. Aplicaciones.
- Tema 8. Modelado de datos. Modelos vectoriales: espagueti, topología, modelos en red, topología completa. Modelo ráster. Métodos de compresión. Ventajas e inconvenientes de los modelos vector/ráster. Orientación a objeto.
- Tema 9. Formatos de intercambio de datos geográficos. El estándar Geographic Markup Language (GML).
- Tema 10. Captura de datos geográficos. Métodos de captura: Digitalización. Escaneado. Fotogrametría. GPS. Trabajos de campo. Importación de datos. Captura de atributos.
- Tema 11. Edición de datos. Corrección de errores. Métodos de indexación espacial. Actualización de la información.
- Tema 12. Calidad de datos. Fuentes de error en un SIG. Propagación de errores. Descripción de la calidad: exactitud posicional, temática y temporal, compleción, consistencia lógica, propósito, linaje y uso. Métodos estadísticos de determinación de la calidad. Calidad de procesos.
- Tema 13. Metadatos. Definición. Normas de metadatos y sus elementos: ISO19115, Núcleo Español de Metadatos y Dublín Core.
- Tema 14. Análisis vectorial. Análisis de redes. Teoría de Grafos. Camino mínimo. Algoritmos. Problema del viajante. Análisis de polígonos.
- Tema 15. Análisis ráster. Autocorrelación espacial, reclasificación y superposición (overlay) de información geográfica.
- Tema 16. Normas para la información geográfica. Normas ISO 19100. Contenido y campo de aplicación. Relación con otros organismos de normalización de información geográfica (CEN, AENOR, etc). El Open Geospatial Consortium. Especificaciones de interoperabilidad.
- Tema 17. Infraestructuras de Datos Espaciales. Definición y componentes. Geoportales. Arquitectura de una IDE. La Directiva INSPIRE. Iniciativas y proyectos IDE.
- Tema 18. Aplicaciones cliente y servidor en una IDE. Clientes ligeros y pesados. Tipos de servicios: Servicios de Mapas, de Catálogo, de Nomenclátor y otros. Encadenamiento de servicios. Recomendación WMS-C. Peticiones teseladas y gestión de teselas. Mashups geográficos.
- Tema 19. Modelos Digitales del Terreno: Análisis, formatos y aplicaciones (análisis de pendientes y orientaciones, mapas de visibilidad, perfiles). Algoritmos y métodos.
- Tema 20. Comunicaciones, redes e Internet. Medios de transmisión y codificación de datos. Modelo de referencia OSI. Familia de protocolos TCP/IP. Redes de área personal, PAN, Redes de área local, LAN. Redes de área metropolitana, MAN. Internet.

Tema 1. Sistemas de Información. Análisis y desarrollo de sistemas de información. Desarrollo estructurado. Desarrollo orientado a objetos. Metodologías de desarrollo de software. Metricav3. Técnicas para el análisis y desarrollo de sistemas de información. UML.

1.1. Sistema de Información

1.1.1. Definición

Puede definirse como sistema de información (SI) el conjunto de procedimientos, manuales y automatizados, y de funciones dirigidas a la recogida, elaboración, evaluación, almacenamiento, recuperación, condensación y distribución de información, dentro de una organización, orientadas a promover el flujo de las mismas desde el punto en el que se generan hasta el destinatario final de las mismas.

1.1.2. Funciones básicas de tratamiento de la información

Dentro de la complejidad general del SI , las funciones realizadas dentro de cada subsistema tienden a ser conceptualmente claras. Así los datos entran en el sistema y luego son transmitidos, almacenados, manipulados y presentados. Veamos ahora brevemente los principales aspectos de las funciones básicas de tratamiento de la información dentro del SI

- Ingreso de datos. Los datos entran al SI en forma de transacciones que describen sucesos del mundo real.
- Almacenamiento de datos. El SI debe mantener grandes ficheros de datos destinados a suministrar la información para el tratamiento de transacciones y para la toma de decisiones. Los principales aspectos a considerar son el papel de la base de datos y la organización de los mismos en dicho sistema.
- Cálculo. La forma habitual de cálculo implícita en la mayoría de los SI, se refiere a cálculos matemáticos, manipulación de datos y ejecución de diversas acciones de acuerdo con los resultados obtenidos. Mediante los cálculos el SI transforma los datos brutos en información utilizable por el propio sistema o de forma ajena al mismo.
- Presentación de la información. La función de presentación de un SI proporciona una conexión esencial, o interfaz, entre el sistema y el usuario. Su finalidad es presentar la

información de tal modo que mejore la capacidad del usuario para percibir y actuar sobre los hechos reflejados por la información.

- **Comunicaciones.** Los Sistemas de Información actuales se diferencian muy notablemente de los del pasado en su creciente apoyo en las comunicaciones. Los avances experimentados en los Sistemas de Información están estrechamente relacionados con los avances realizados en el mundo de las telecomunicaciones. Ahora asistimos a la implantación de sistemas informáticos distribuidos en los que los ordenadores a través de la organización están conectados por medio de una red de telecomunicaciones. Cada ordenador remoto sobre la red tiene, generalmente, capacidades de cálculo autónomo significativas para servir a las necesidades especializadas de sus usuarios locales proporcionando también acceso a recursos mantenidos en otras localizaciones, eventualmente en un ordenador central con una potencia de procesamiento considerable y gran capacidad de almacenamiento en línea sobre disco. Los recursos necesarios no disponibles sobre un terminal de trabajo personal (estación de trabajo) pueden ser alcanzados a través de la red a nivel departamental si la comparación es rápida, conveniente y eficiente o a nivel corporativo si se trata de recursos demasiado costosos para estar duplicados a nivel departamental. Los diseñadores de la red de telecomunicaciones deberán seleccionar una combinación específica de topología de red, anchos de banda, protocolos de comunicaciones, equipo terminal y proveedores de comunicaciones.

1.2. Metodologías

1.2.1. Definición

Como su nombre indica las metodologías son un conjunto de métodos, de reglas, que por una parte sirven de guía para realizar los trabajos que van dando forma a los desarrollos y que por otra obligan a la dirección del proyecto y a los componentes de los equipos a realizar ciertas comprobaciones sistemáticas de modo que el resultado final, al menos desde un punto de vista formal, no presente incoherencias y esté dirigido a un objetivo claro y prefijado.

1.2.2. Historia

Desde el principio del uso de los ordenadores, al trabajar sobre el desarrollo de los primeros programas, se siguieron una serie de pautas o métodos para llevar a buen fin el proyecto. Bien es verdad que en esta situación la metodología era simple, era el típico proceso de abajo a arriba, con análisis insuficientes, ya que el problema era comprendido fácilmente en su totalidad. Además sólo preocupaba el proceso ya que los datos no solían ser importantes en sí y

lo que realmente importaba era resolver algún tipo de problema físico cuya formulación matemática, y sobre todo su resolución, era complicada. Por lo tanto en el principio de la informática los métodos eran de tipo ascendente y orientado a procesos.

Al introducirse el ordenador en las empresas para la gestión de sus necesidades se partió de la mentalidad antes citada y se trató de resolver aquellos problemas cuya complejidad, en cuanto a procedimiento de resolución, era grande. A finales de la década de los sesenta en España empezaron a aparecer ordenadores en las empresas para resolver problemas del tipo de cálculo de nómina, no de gestión de personal, debido a que era una labor que exige hacer muchos cálculos, en general repetitivos, y por tanto con alta probabilidad de error. Al resolverlos manualmente son trabajos en los cuales lo fundamental es comprender los procesos que se deben seguir para hacer correctamente el cálculo de los haberes y en donde los datos, aún no informatizados masivamente, se consideran secundarios.

En la década de los setenta empezó a tomar cuerpo la idea de que si bien los procesos son importantes, y una incorrección en su tratamiento puede causar notables problemas e incomodidades, más importantes son los datos. Los procesos son, similares en todas las organizaciones y, en algunos casos, son hasta relativamente fácil de transportar. Sin embargo los datos son algo propio de la organización, algo totalmente diferente de los de las demás organización, y que la caracterizan.

Ante esta situación no cabe otra alternativa que el análisis por partes, para lo cual hay que partir de la idea más general posible sobre el sistema a estudiar e ir descendiendo por pasos sucesivos hasta llegar a los detalles menores. Se está ante procesos de arriba a abajo, únicos, que permiten que haya coherencia entre todos los componentes del complicado sistema.

La década de los ochenta es la época marcada por las metodologías dirigida a datos cuya importancia va tomando cuerpo en las organizaciones. A mediados de la década el estado de la técnica permite a considerar entidades más complejas y con personalidad más acusada. Se empiezan a estudiar los objetos en sí como unidades de información.

Por otra parte las metodologías de desarrollo comienzan a interesarse no sólo en lograr que el proyecto sea puesto en funcionamiento sino en minimizar costos durante su desarrollo y sobre todo durante su mantenimiento. Los nuevos métodos van buscando minimizar riesgos y, puesto que los errores más perjudiciales se producen en los primeros pasos, se comienza ya desde la fase más general del estudio por analizar los riesgos que significa seguir con las siguientes fases del desarrollo. Si los riesgos son superiores a lo que se consideran permisibles, se vuelve al paso

anterior o se abandona el desarrollo. No sólo se realizan desarrollos lineales, en cascada, sino también desarrollos y métodos en espiral que son notablemente más complejos.

1.2.3. Características más comunes de las metodologías

- Uso de técnicas gráficas
- Diccionario de datos. Cada uno de ellos tiene asignado un código que dificulta su duplicación a la par que facilita su localización dentro del complejo esquema resultante indicando en qué etapa aparece, qué tipo de objeto es y otra serie de información respecto al objeto.
- Documentación
- Prototipado
- Análisis del riesgo. Obliga a que, tras cada avance en el ciclo de vida, se estime y valore el riesgo de la continuación con las condiciones del momento dando lugar a una toma de decisión sobre si continuar con el siguiente paso o bien retroceder hasta mejorar el conocimiento del sistema. De este modo se asegura, con una probabilidad de fracaso aceptada, el buen éxito del proyecto. Se generan los denominados modelos en espiral que son los que actualmente se consideran más adecuados para grandes proyectos.
- Análisis estructurado. Persigue elaborar algoritmos que cumplan la propiedad de modularidad, para ello, dado un problema que se pretende resolver mediante la elaboración de un programa de ordenador, se busca dividir dicho programa en módulos siguiendo los principios de diseño de Descomposición por refinamientos sucesivos, creación de una Jerarquía modular y elaboración de módulos Independientes.
- Análisis orientado a objetos. Se piensa en términos de objetos, en vez de procedimientos, cuando se planifica un código. Un objeto agrupa datos encapsulados y procedimientos para representar una entidad. Un programa orientado a objetos es descrito por la interacción de esos objetos, para resolver un problema de la organización.

Existen muchas metodologías en el mercado como JSD/JSP, METHOD 1 o ITIL. Este documento se centrará en la metodología MÉTRICA v3 del Consejo Superior de Administración Electrónica, que es la metodología recomendada por este organismo para utilizar en los desarrollos de sistemas e información en la Administración General del Estado.

1.3. MÉTRICA v3 del Consejo Superior de Administración Electrónica (CSAE)

MÉTRICA es una Metodología de Planificación, Desarrollo y Mantenimiento de Sistemas de información. Promovida por el Ministerio de Administraciones Públicas del gobierno español para la sistematización de actividades del ciclo de vida de los proyectos software en el ámbito de las administraciones públicas. Esta metodología propia está basada en el Modelo de Procesos del Ciclo de vida de desarrollo ISO/IEC 12207 (Information Technology - Software Life Cycle Processes) así como en la norma ISO/IEC 15504 SPICE (Software Process Improvement And Assurance Standards Capability Determination).

En una única estructura la metodología MÉTRICA Versión 3 cubre distintos tipos de desarrollo: estructurado y orientado a objetos, facilitando a través de interfaces la realización de los procesos de apoyo u organizativos: Gestión de Proyectos, Gestión de Configuración, Aseguramiento de Calidad y Seguridad.

La metodología contiene una serie de elementos fundamentales que se describen a continuación.

1.3.1. Participantes

MÉTRICA Versión 3 ha sido concebida para abarcar el desarrollo completo de Sistemas de Información sea cual sea su complejidad y magnitud, por lo cual su estructura y los perfiles de los participantes que intervienen deberán adaptarse y dimensionarse en cada momento de acuerdo a las características particulares de cada proyecto.

Los perfiles establecidos son:

- Directivo
- Jefe de Proyecto
- Consultor
- Analista
- Programador

Para cada uno de estos perfiles se analizan una serie de características importantes a la hora de delimitar su participación en el proyecto:

- Correspondencia con Participantes de MÉTRICA Versión 3.
- Responsabilidades o funciones a desempeñar en cada uno de los procesos.
- Perfil o características propias de cada uno de los participantes.

Perfil Directivo

El perfil requerido para este grupo de participantes incluye a personas con un nivel alto en la dirección de la organización, conocimiento de los objetivos estratégicos y de negocio que se persiguen y autoridad para validar y aprobar cada uno de los procesos realizados durante el desarrollo del Sistema de Información. Además deben tener un conocimiento del entorno y de la organización suficiente para proporcionar, a lo largo de todo el proyecto, unos requisitos del Sistema adecuados, completos y suficientemente importantes como para considerarse en el catálogo definitivo de requisitos.

Perfil Jefe de Proyecto

Son los responsables de las distintas áreas del proyecto. Ejercen labores de coordinación y dirección de equipos humanos especializados en la realización de actividades propias de un proceso o interfaz de MÉTRICA Versión 3. La figura principal es el Jefe de Proyecto, el cual recibe el apoyo de los distintos responsables durante la realización de procesos o determinadas actividades a lo largo del proyecto.

El Jefe de Proyecto realiza la estimación del esfuerzo necesario para llevar a cabo el proyecto, selecciona la estrategia de desarrollo, determina la estructura del mismo seleccionando los procesos principales de MÉTRICA Versión 3 que lo integran, fija el calendario de hitos y entregas y establece la planificación del proyecto. Es el encargado de dirigir el proyecto, realizando las labores de seguimiento y control del mismo, revisión y evaluación de resultados y coordinación del equipo de proyecto. Se ocupa también de la gestión y resolución de incidencias que puedan surgir durante el desarrollo del proyecto así como de la actualización de la planificación inicial. Entre sus funciones se encuentran la elaboración de los informes de seguimiento y el archivo de la documentación de gestión del proyecto una vez que este ha finalizado.

Perfil Consultor

La principal función de los Consultores es asesorar en las cuestiones sobre las que tienen un conocimiento especializado.

Perfil Analista

La responsabilidad de los Analistas es elaborar un catálogo detallado de requisitos que permita describir con precisión el sistema de información, para lo cual mantendrán entrevistas y sesiones

de trabajo con los responsables de la organización y usuarios, actuando de interlocutor entre estos y el equipo de proyecto en lo que a requerimientos se refiere. Estos requisitos permiten a los analistas elaborar los distintos modelos que sirven de base para el diseño, obteniendo los modelos de datos y de procesos en el caso del análisis estructurado y los modelos de clases e interacción de objetos en análisis orientado a objeto. Así mismo realizan la especificación de las interfaces entre el sistema y el usuario. Si existe un sistema de gestión de la configuración, se encargan de identificar y registrar en el mismo los productos seleccionados como elementos de configuración en el plan de gestión de configuración.

El Administrador de Bases de Datos participa en la obtención del diseño físico de datos, definiendo la estructura física de datos que utilizará el sistema a partir del modelo lógico de datos normalizado o del modelo de clases, teniendo presentes las características específicas del sistema de gestión de base de datos concreto a utilizar, los requisitos establecidos para el sistema de información, y las particularidades del entorno tecnológico, se consiga una mayor eficiencia en el tratamiento de los datos.

Perfil Programador

La función del programador, miembro del equipo de proyecto, es construir el código que dará lugar al producto resultante en base al diseño técnico realizado por el analista o analista programador, generando también el código asociado a los procedimientos de migración y carga inicial de datos. Igualmente se encarga de la realización de las pruebas unitarias y participa en las pruebas de conjunto de la aplicación.

1.3.2. Procesos Principales de MÉTRICA Versión 3

MÉTRICA Versión 3 tiene un enfoque orientado al proceso, ya que la tendencia general en los estándares se encamina en este sentido y por ello, como ya se ha dicho, se ha enmarcado dentro de la norma ISO 12.207, que se centra en la clasificación y definición de los procesos del ciclo de vida del software.

La metodología descompone cada uno de los procesos en actividades, y éstas a su vez en tareas. Para cada tarea se describe su contenido haciendo referencia a sus principales acciones, productos, técnicas, prácticas y participantes. El orden asignado a las actividades no debe interpretarse como secuencia en su realización, ya que éstas pueden realizarse en orden diferente a su numeración o bien en paralelo, como se muestra en los gráficos de cada proceso. Sin embargo, no se dará por acabado un proceso hasta no haber finalizado todas las actividades del mismo determinadas al inicio del proyecto.

Así los procesos de la estructura principal de MÉTRICA Versión 3 son los siguientes:

- Planificación de Sistemas de Información
- Desarrollo de Sistemas de Información
- Mantenimiento de Sistemas de Información

En este documento se centrará la exposición en el proceso de Desarrollo de Sistemas de Información, remitiendo al usuario a las especificaciones de MÉTRICA Versión 3, para obtener más información.

1.3.3. Desarrollo de Sistemas de Información

El proceso de Desarrollo de MÉTRICA Versión 3 contiene todas las actividades y tareas que se deben llevar a cabo para desarrollar un sistema, cubriendo desde el análisis de requisitos hasta la instalación del software. Además de las tareas relativas al análisis, incluye dos partes en el diseño de sistemas: arquitectónico y detallado. También cubre las pruebas unitarias y de integración del sistema, aunque siguiendo la norma ISO 12.207 no propone ninguna técnica específica y destaca la importancia de la evolución de los requisitos. Este proceso es, sin duda, el más importante de los identificados en el ciclo de vida de un sistema y se relaciona con todos los demás.

Para facilitar la comprensión y dada su amplitud y complejidad se ha subdividido el proceso de desarrollo en cinco subprocesos:

- Estudio de Viabilidad del Sistema (EVS).
- Análisis del Sistema de Información (ASI).
- Diseño del Sistema de Información (DSI).
- Construcción del Sistema de Información (CSI).
- Implantación y Aceptación del Sistema (IAS).

Estudio de Viabilidad del Sistema (EVS)

El propósito de este proceso es analizar un conjunto concreto de necesidades, con la idea de proponer una solución a corto plazo. Los criterios con los que se hace esta propuesta no serán estratégicos sino tácticos y relacionados con aspectos económicos, técnicos, legales y operativos.

Los resultados del Estudio de Viabilidad del Sistema constituirán la base para tomar la decisión de seguir adelante o abandonar. Si se decide seguir adelante pueden surgir uno o varios proyectos que afecten a uno o varios sistemas de información. Dichos sistemas se desarrollarán

según el resultado obtenido en el estudio de viabilidad y teniendo en cuenta la cartera de proyectos para la estrategia de implantación del sistema global.

Análisis del Sistema de Información (ASI)

El propósito de este proceso es conseguir la especificación detallada del sistema de información, a través de un catálogo de requisitos y una serie de modelos que cubran las necesidades de información de los usuarios para los que se desarrollará el sistema de información y que serán la entrada para el proceso de Diseño del Sistema de Información.

Se recogen de forma detallada los requisitos funcionales que el sistema de información debe cubrir, catalogándolos, lo que permite hacer la traza a lo largo de los procesos de desarrollo. Además, se identifican los requisitos no funcionales del sistema, es decir, las facilidades que ha de proporcionar el sistema, y las restricciones a que estará sometido, en cuanto a rendimiento, frecuencia de tratamiento, seguridad, etc.

Para facilitar el análisis del sistema se identifican los subsistemas de análisis, y se elaboran los modelos de Casos de Uso y de Clases, en desarrollos orientados a objetos, y de Datos y Procesos en desarrollos estructurados.

Los productos resultantes del Análisis del Sistema de Información, dependen del tipo de desarrollo de que se trate y se detallan a continuación especificando los que son distintos, según los dos tipos de desarrollo a los que da respuesta MÉTRICA Versión 3:

En Análisis Estructurado:

- Modelo de procesos.
- Modelo lógico de datos normalizado.

En Análisis Orientado a Objetos:

- Modelo de clases.
- Análisis de la realización de los casos de uso.

Diseño del Sistema de Información (DSI)

El propósito del Diseño del Sistema de Información (DSI) es obtener la definición de la arquitectura del sistema y del entorno tecnológico que le va a dar soporte, junto con la especificación detallada de los componentes del sistema de información. A partir de dicha información, se generan todas las especificaciones de construcción relativas al propio sistema.

Construcción del Sistema de Información (CSI)

La construcción del Sistema de Información (CSI) tiene como objetivo final la construcción y prueba de los distintos componentes del sistema de información, a partir del conjunto de especificaciones lógicas y físicas del mismo, obtenido en el Proceso de Diseño del Sistema de Información (DSI). Se desarrollan los procedimientos de operación y seguridad y se elaboran los manuales de usuario final y de explotación, estos últimos cuando proceda.

Para conseguir dicho objetivo, se recoge la información relativa al producto del diseño Especificaciones de construcción del sistema de información, se prepara el entorno de construcción, se genera el código de cada uno de los componentes del sistema de información y se van realizando, a medida que se vaya finalizando la construcción, las pruebas unitarias de cada uno de ellos y las de integración entre subsistemas.

Producto software:

- Código fuente de los componentes.
- Manuales de usuario.

Implantación y Aceptación del Sistema (IAS)

Este proceso tiene como objetivo principal, la entrega y aceptación del sistema en su totalidad, que puede comprender varios sistemas de información desarrollados de manera independiente, según se haya establecido en el proceso de Estudio de Viabilidad del Sistema (EVS), y un segundo objetivo que es llevar a cabo las actividades oportunas para el paso a producción del sistema.

Al finalizar esta tarea, se considera terminado el sistema de información y funcionando en un entorno de producción.

1.3.4. Interfaces de MÉTRICA Versión 3

La estructura de MÉTRICA Versión 3 incluye también un conjunto de interfaces que definen una serie de actividades de tipo organizativo o de soporte al proceso de desarrollo y a los productos, que en el caso de existir en la organización se deberán aplicar para enriquecer o influir en la ejecución de las actividades de los procesos principales de la metodología y que si no existen habrá que realizar para complementar y garantizar el éxito del proyecto desarrollado con MÉTRICA Versión 3.

Las interfaces descritas en la metodología son:

- Gestión de Proyectos (GP).
- Seguridad (SEG)
- Aseguramiento de la Calidad (CAL)
- Gestión de la Configuración (GC)

Se remite al lector a las especificaciones de MÉTRICA Versión 3 si desea profundizar en estos aspectos.

1.3.5. Técnicas y prácticas

La metodología MÉTRICA Versión 3 proporciona un conjunto de métodos y técnicas que guía a los distintos profesionales de Sistemas y Tecnologías de la Información y Comunicaciones en la obtención de los diversos productos de los procesos del ciclo de vida de un proyecto informático.

En este documento se mostrará una introducción a las técnicas de desarrollo de sistemas de información y, dentro de ellas, las que se consideran de mayor relevancia.

Las técnicas de desarrollo son un conjunto de procedimientos que se basan en reglas y notaciones específicas en términos de sintaxis, semántica y gráficos, orientadas a la obtención de productos en el desarrollo de un sistema de información. En desarrollos del tipo estructurado o de orientación a objetos merecen especial atención las técnicas gráficas, que proponen símbolos y notaciones estándares para una mejor comprensión de los sistemas o sus componentes.

En el caso de desarrollos orientados a objetos se ha seguido la notación de UML. Lenguaje Unificado de Modelado (UML, por sus siglas en inglés, Unified Modeling Language) es el lenguaje de modelado de sistemas de software más conocido y utilizado en la actualidad; aún cuando todavía no es un estándar oficial, está respaldado por el OMG (Object Management Group). Es un lenguaje gráfico para visualizar, especificar, construir y documentar un sistema de software. UML ofrece un estándar para describir un "plano" del sistema (modelo), incluyendo aspectos conceptuales tales como procesos de negocios y funciones del sistema, y aspectos concretos como expresiones de lenguajes de programación, esquemas de bases de datos y componentes de software reutilizables.

En ningún caso, esta notación se considera obligatoria. Cada organización podrá utilizar la notación que desee, la que suele utilizar o la que ofrecen sus herramientas de desarrollo, respetando las reglas y restricciones específicas de las distintas técnicas.

Casos de Uso

Los objetivos de los casos de uso son los siguientes:

- Capturar los requisitos funcionales del sistema y expresarlos desde el punto de vista del usuario.
- Guiar todo el proceso de desarrollo del sistema de información.

Estos diagramas presentan dos tipos de elementos fundamentales:

- Actores. En general, los actores serán los usuarios del sistema y los sistemas externos al que se esté desarrollando.
- Casos de uso. Un caso de uso representa el comportamiento que ofrece el sistema de información desde el punto de vista del usuario. Típicamente será un conjunto de transacciones ejecutadas entre el sistema y los actores.

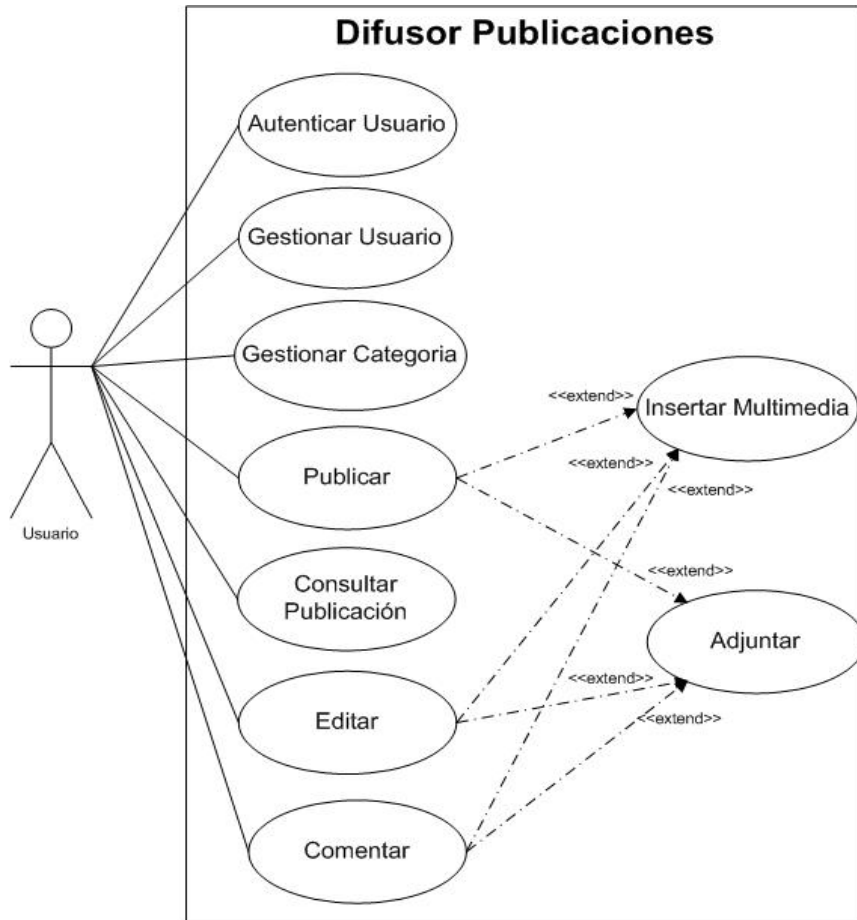


Figura 1. Ejemplo de caso de uso, según Métrica v3. Fuente Internet.

Diagrama de Clases

El objetivo principal de este modelo es la representación de los aspectos estáticos del sistema, utilizando diversos mecanismos de abstracción (clasificación, generalización, agregación).

El diagrama de clases recoge las clases de objetos y sus asociaciones. En este diagrama se representa la estructura y el comportamiento de cada uno de los objetos del sistema y sus relaciones con los demás objetos, pero no muestra información temporal.

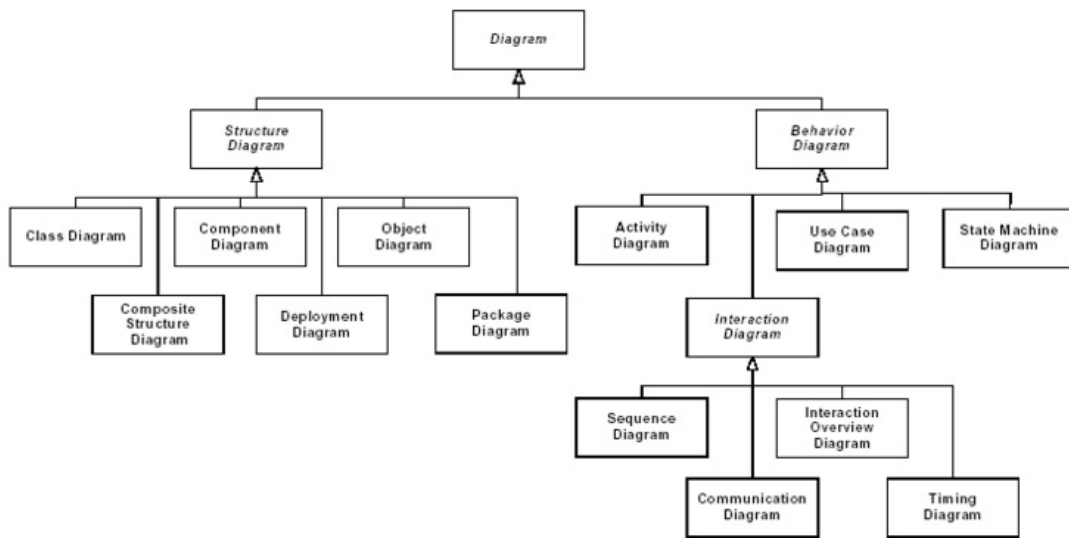


Figura 2. Ejemplo de diagrama de clases. Fuente: Wikipedia.

Diagrama de Flujo de Datos (DFD)

El objetivo del diagrama de flujo de datos es la obtención de un modelo lógico de procesos que represente el sistema, con independencia de las restricciones físicas del entorno. Así se facilita su comprensión por los usuarios y los miembros del equipo de desarrollo.

Un diagrama de flujo de datos es una técnica muy apropiada para reflejar de una forma clara y precisa los procesos que conforman el sistema de información. Permite representar gráficamente los límites del sistema y la lógica de los procesos, estableciendo qué funciones hay que desarrollar. Además, muestra el flujo o movimiento de los datos a través del sistema y sus transformaciones como resultado de la ejecución de los procesos. Esta técnica consiste en la descomposición sucesiva de los procesos, desde un nivel general, hasta llegar al nivel de detalle necesario para reflejar toda la semántica que debe soportar el sistema en estudio.

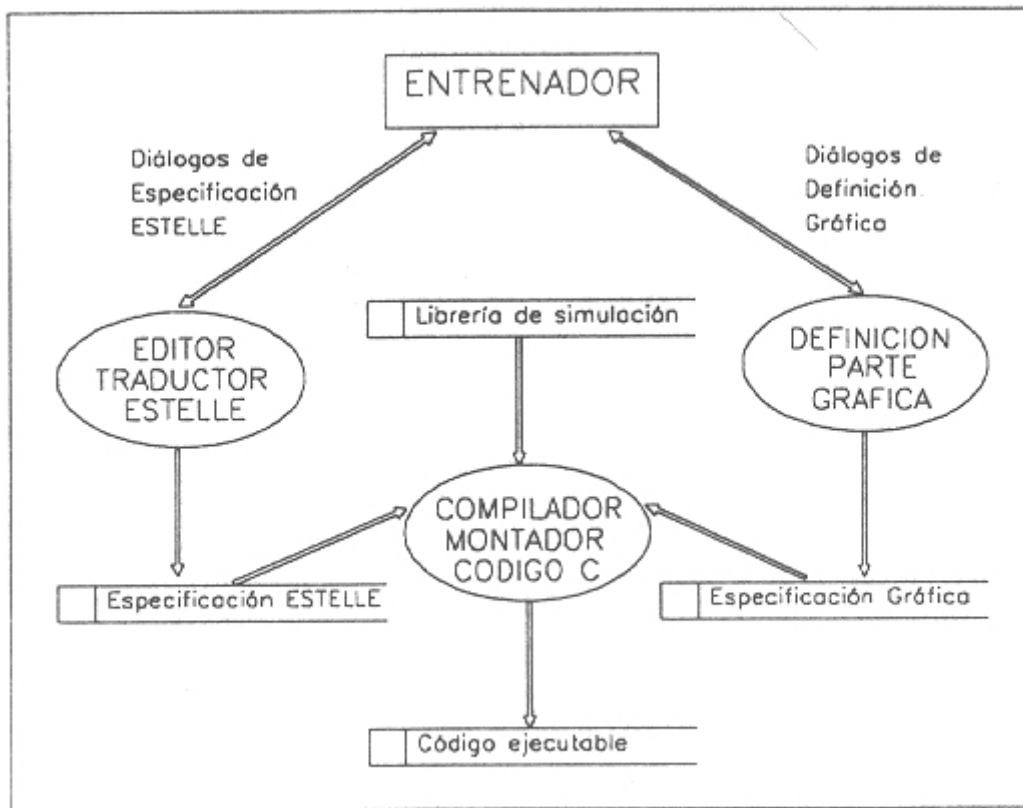


Figura 3. Ejemplo de diagrama de flujo de datos. Fuente <http://www.niee.ufrgs.br>.

Modelo entidad / relación extendido.

Se trata de una técnica cuyo objetivo es la representación y definición de todos los datos que se introducen, almacenan, transforman y producen dentro de un sistema de información, sin tener en cuenta las necesidades de la tecnología existente, ni otras restricciones. Dado que el modelo de datos es un medio para comunicar el significado de los datos, las relaciones entre ellos y las reglas de negocio de un sistema de información, una organización puede obtener numerosos beneficios de la aplicación de esta técnica, pues la definición de los datos y la manera en que éstos operan son compartidos por todos los usuarios.

Este diagrama se centra en los datos, independientemente del procesamiento que los transforma y sin entrar en consideraciones de eficiencia.

El modelo E/R extendido describe con un alto nivel de abstracción la distribución de datos almacenados en un sistema. Existen dos elementos principales: las entidades y las relaciones.

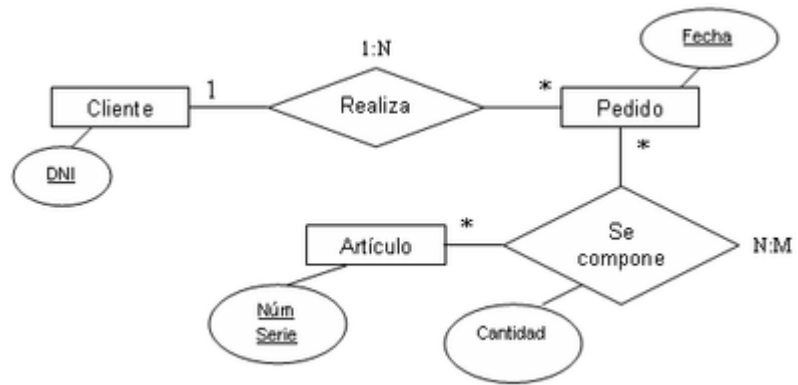


Figura 4. Ejemplo de diagrama de Entidad/Relación. Fuente: Wikipedia.

Bibliografía

- [1] Ian Sommerville. Addison Wesley. “Ingeniería de Software”. 6ª Edición
- [2] Metodología de Planificación y Desarrollo de Sistemas de Información. MÉTRICA Versión 3
- [3] Pressman, Roger S. Editorial. “Ingeniería del Software. Un enfoque práctico”. Mc. Graw Hill. 6ª Edición
- [4] Temario de la Asociación Profesional del cuerpo superior de Sistemas y Tecnologías de la Administración del Estado. Edición 2007.

Tema 2. Lenguajes de Programación. Tipos de lenguajes de programación. Programación estructurada. Programación orientada a objetos.

2.1. Introducción

Con la llegada de las computadoras aparecen las secuencias de posiciones de llaves eléctricas que debían conectarse para obtener una acción determinada, una llave conectada era un 1 y una llave desconectada era un 0. Una sucesión de llaves en cualquiera de sus dos posiciones definía una secuencia de ceros y unos (por ejemplo: 0100011010011101...) que venía a representar una instrucción o un conjunto de instrucciones (programa) para el ordenador (o computador) en el que se estaba trabajando. A esta primera forma de especificar programas para una computadora se la denomina *lenguaje máquina* o *código máquina*.

La necesidad de recordar secuencias de programación para las acciones usuales llevó a denominarlas con nombres fáciles de memorizar y asociar: ADD (sumar), SUB (restar), MUL (multiplicar), CALL (ejecutar subrutina), etc. A esta secuencia de posiciones se le denominó *instrucciones*, y a este conjunto de instrucciones se le llamó lenguaje ensamblador.

Posteriormente aparecieron diferentes lenguajes de programación, los cuales reciben su denominación porque tienen una estructura sintáctica similar a los lenguajes escritos por los seres humanos.

2.2. Lenguajes de programación

Un lenguaje de programación es un lenguaje que puede ser utilizado para controlar el comportamiento de una máquina, particularmente una computadora. Consiste en un conjunto de símbolos y reglas sintácticas y semánticas que definen su estructura y el significado de sus elementos y expresiones.

Aunque muchas veces se usa *lenguaje de programación* y *lenguaje informático* como si fuesen sinónimos, no tiene por qué ser así, ya que los lenguajes informáticos engloban a los lenguajes de programación y a otros más, como, por ejemplo, el HTML (lenguaje para el marcado de páginas web).

Un lenguaje de programación permite a uno o más programadores especificar de manera precisa, sobre qué datos una computadora debe operar, cómo deben ser éstos almacenados y transmitidos y qué acciones debe tomar bajo una variada gama de circunstancias. Todo esto, a través de un lenguaje que intenta estar relativamente próximo al lenguaje humano o natural, tal como sucede con el lenguaje léxico. Una característica relevante de los lenguajes de programación es precisamente que más de un programador puedan tener un conjunto común de instrucciones que puedan ser comprendidas entre ellos para realizar la construcción del programa de forma colaborativa.

Los procesadores usados en las computadoras son capaces de entender y actuar según lo indican programas escritos en un lenguaje fijo llamado lenguaje de máquina. Todo programa escrito en otro lenguaje puede ser ejecutado de dos maneras:

- Mediante un programa que va adaptando las instrucciones conforme son encontradas. A este proceso se lo llama interpretar y a los programas que lo hacen se los conoce como intérpretes.

Traduciendo este programa al programa equivalente escrito en lenguaje de máquina. A ese proceso se lo llama compilar y al traductor se lo conoce como compilador. Clasificación de los lenguajes de programación

Los lenguajes de programación se determinan según el nivel de abstracción, según la forma de ejecución y según el paradigma de programación que poseen cada uno de ellos y esos pueden ser:

2.3. Tipos de lenguajes de programación

2.3.1. Según su nivel de abstracción

Lenguajes de bajo nivel

Los lenguajes de bajo nivel son lenguajes de programación que se acercan al funcionamiento de una computadora. El lenguaje de más bajo nivel es, por excelencia, el código máquina (unos y ceros). A éste le sigue el lenguaje ensamblador, ya que al programar en ensamblador se trabaja con los registros de memoria física de la computadora de forma directa.

Lenguajes de medio nivel

Hay lenguajes de programación que son considerados por algunos expertos como lenguajes de medio nivel (como es el caso del lenguaje C) al tener ciertas características que los acercan a los

lenguajes de bajo nivel pero teniendo, al mismo tiempo, ciertas cualidades que lo hacen un lenguaje más cercano al humano y, por tanto, de alto nivel.

Lenguajes de alto nivel

Los lenguajes de alto nivel son normalmente fáciles de aprender porque están formados por elementos de lenguajes naturales, como el inglés. En este nivel se encuentran lenguajes de programación como el famoso *BASIC* u otros más modernos como *Java* o *Visual .Net*. En estos lenguajes se pueden escribir sentencias como:

```
IF CONTADOR = 10 THEN STOP
```

Esta sentencia, pide a la computadora que pare si *CONTADOR* es igual a *10*. Por desgracia para muchas personas esta forma de trabajar es un poco frustrante, dado que a pesar de que las computadoras parecen comprender un lenguaje natural, lo hacen en realidad de una forma rígida y sistemática.

2.3.2. Según la forma de ejecución

Lenguajes compilados

Naturalmente, un programa que se escribe en un lenguaje de alto nivel también tiene que traducirse a un código que pueda utilizar la máquina. Los programas traductores que pueden realizar esta operación se llaman compiladores. Éstos, como los programas ensambladores avanzados, pueden generar muchas líneas de código de máquina por cada proposición del programa fuente. Se requiere una ejecución de la compilación antes de procesar los datos de un problema.

Los *compiladores* son aquellos cuya función es traducir un programa escrito en un determinado lenguaje a un idioma que la computadora entienda (lenguaje máquina con código binario).

Al usar un lenguaje compilado (como lo son los lenguajes del popular Visual Studio de Microsoft), el programa desarrollado nunca se ejecuta mientras haya errores, sino hasta que luego de haber compilado el programa, ya no aparecen errores en el código.

Lenguajes interpretados

Se puede también utilizar una alternativa diferente de los compiladores para traducir lenguajes de alto nivel. En vez de traducir el programa fuente y grabar de forma permanente el código objeto que se produce durante la compilación para utilizarlo posteriormente en la ejecución del

programa, el programador sólo carga el programa fuente en la computadora junto con los datos que se van a procesar. A continuación, un programa intérprete, almacenado en el sistema operativo del disco, o incluido de manera permanente dentro de la máquina, convierte cada proposición del programa fuente en lenguaje de máquina conforme vaya siendo necesario durante el proceso de los datos. No se graba el código objeto para utilizarlo posteriormente.

La siguiente vez que se utilice una instrucción, se le debe interpretar otra vez y traducir a lenguaje máquina. Por ejemplo, durante el procesamiento repetitivo de los pasos de un ciclo, cada instrucción del ciclo tendrá que volver a ser interpretada cada vez que se ejecute el ciclo, lo cual hace que el programa sea más lento en tiempo de ejecución (porque se va revisando el código en tiempo de ejecución) pero más rápido en tiempo de diseño (porque no se tiene que estar compilando a cada momento el código completo). El intérprete elimina la necesidad de realizar una compilación después de cada modificación del programa cuando se quiere agregar funciones o corregir errores; pero es obvio que un programa objeto compilado con antelación deberá ejecutarse con mucha mayor rapidez que uno que se debe interpretar a cada paso durante su ejecución.

2.3.3. Según el paradigma de programación

Un paradigma de programación representa un enfoque particular o filosofía para la construcción del software. No es mejor uno que otro sino que cada uno tiene ventajas y desventajas. También hay situaciones donde un paradigma resulta más apropiado que otro.

Atendiendo al paradigma de programación, se pueden clasificar los lenguajes en:

Lenguajes imperativos

La programación imperativa, en contraposición a la programación declarativa es un paradigma de programación que describe la programación en términos del estado del programa y sentencias que cambian dicho estado. Los programas imperativos son un conjunto de instrucciones que le indican al computador cómo realizar una tarea.

La implementación de hardware de la mayoría de computadores es imperativa; prácticamente todo el hardware de los computadores está diseñado para ejecutar código de máquina, que es nativo al computador, escrito en una forma imperativa. Desde esta perspectiva de bajo nivel, el estilo del programa está definido por los contenidos de la memoria, y las sentencias son instrucciones en el lenguaje de máquina nativo del computador (por ejemplo el lenguaje ensamblador).

Los lenguajes imperativos de alto nivel usan variables y sentencias más complejas, pero aún siguen el mismo paradigma. Las recetas y las listas de revisión de procesos, a pesar de no ser programas de computadora, son también conceptos familiares similares en estilo a la programación imperativa; cada paso es una instrucción, y el mundo físico guarda el estado.

Los primeros lenguajes imperativos fueron los lenguajes de máquina de los computadores originales. En estos lenguajes, las instrucciones fueron muy simples, lo cual hizo la implementación de hardware fácil, pero obstruyendo la creación de programas complejos. Fortran, cuyo desarrollo fue iniciado en 1954 por John Backus en IBM, fue el primer gran lenguaje de programación en superar los obstáculos presentados por el código de máquina en la creación de programas complejos.

Lenguajes Funcionales

La programación funcional es un paradigma de programación declarativa basado en la utilización de funciones matemáticas.

El objetivo es conseguir lenguajes expresivos y matemáticamente elegantes, en los que no sea necesario bajar al nivel de la máquina para describir el proceso llevado a cabo por el programa, y evitando el concepto de estado del cómputo. La secuencia de computaciones llevadas a cabo por el programa se regiría única y exclusivamente por la reescritura de definiciones más amplias a otras cada vez más concretas y definidas, usando lo que se denominan "definiciones dirigidas".

Los programas escritos en un lenguaje funcional están constituidos únicamente por definiciones de funciones, entendiendo éstas no como subprogramas clásicos de un lenguaje imperativo, sino como funciones puramente matemáticas, en las que se verifican ciertas propiedades como la transparencia referencial (el significado de una expresión depende únicamente del significado de sus subexpresiones), y por tanto, la carencia total de efectos laterales.

Otras características propias de estos lenguajes son la no existencia de asignaciones de variables y la falta de construcciones estructuradas como la secuencia o la iteración (lo que obliga en la práctica a que todas las repeticiones de instrucciones se lleven a cabo por medio de funciones recursivas).

Entre los lenguajes funcionales, cabe destacar a Haskell, Miranda y Lisp.

Lenguajes Lógicos

La programación lógica consiste en la aplicación del conocimiento sobre lógica para el diseño de lenguajes de programación; no debe confundirse con la disciplina de la lógica computacional.

Históricamente, los ordenadores se han programado utilizando lenguajes muy cercanos a las peculiaridades de la propia máquina: operaciones aritméticas simples, instrucciones de acceso a memoria, etc. Un programa escrito de esta manera puede ocultar totalmente su propósito a la comprensión de un ser humano, incluso uno entrenado. Hoy día, estos lenguajes pertenecientes al paradigma de la Programación imperativa han evolucionado de manera que ya no son tan crípticos.

En cambio, la lógica matemática es la manera más sencilla, para el intelecto humano, de expresar formalmente problemas complejos y de resolverlos mediante la aplicación de reglas, hipótesis y teoremas. De ahí que el concepto de "programación lógica" resulte atractivo en diversos campos donde la programación tradicional es un fracaso.

La programación lógica encuentra su hábitat natural en aplicaciones de inteligencia artificial o relacionada (sistemas expertos, demostración automática de teoremas, reconocimiento de lenguaje natural).

El programa lógico no solamente es capaz de responder si una determinada hipótesis es verdadera o falsa. También es capaz de determinar qué valores de la incógnita hacen cierta la hipótesis.

El lenguaje de programación lógica por excelencia es Prolog.

Los lenguajes correspondientes a los paradigmas de programación estructurada y orientada objetos, se tratarán en un apartado aparte.

2.4. Programación estructurada

La programación estructurada es una forma de programación que utiliza únicamente tres estructuras:

- Secuencia
- Instrucción condicional.
- Iteración (bucle de instrucciones) con condición al principio.

Con este estilo se obtienen varias ventajas como las siguientes:

- Los programas son más fáciles de entender, ya que pueden ser leído de forma secuencial.
- La estructura del programa es clara puesto que las instrucciones están más ligadas o relacionadas entre sí.
- Reducción del esfuerzo en las pruebas. El seguimiento de los fallos o errores del programa se facilita debido a la estructura más visible, por lo que los errores se pueden detectar y corregir más fácilmente.
- Programas más sencillos y más rápidos.

El principal inconveniente de este método de programación, es que se obtiene un único bloque de programa, que cuando se hace demasiado grande puede resultar problemático su manejo, esto se resuelve empleando la programación modular, definiendo módulos interdependientes programados y compilados por separado.

Hoy en día las aplicaciones informáticas son mucho más ambiciosas, principalmente debido a las aplicaciones gráficas, por lo que las técnicas de programación estructurada no son suficientes lo que ha llevado al desarrollo de nuevas técnicas tales como la programación orientada a objetos y el desarrollo de entornos de programación que facilitan la construcción de grandes aplicaciones.

2.5. Programación Orientada a Objetos

2.5.1. Introducción

La Programación Orientada a Objetos (POO u OOP según sus siglas en inglés) es un paradigma de programación que usa objetos y sus interacciones para diseñar aplicaciones y programas de computadora. Su uso se popularizó a principios de la década de 1990. Actualmente varios lenguajes de programación soportan la orientación a objetos.

Los objetos son entidades que combinan *estado*, *comportamiento* e *identidad*. El estado está compuesto de datos, y el comportamiento por procedimientos o *métodos*. La *identidad* es una propiedad de un objeto que lo diferencia del resto. La programación orientada a objetos expresa un programa como un conjunto de estos objetos, que colaboran entre ellos para realizar tareas. Esto permite hacer los programas y módulos más fáciles de escribir, mantener y reutilizar.

2.5.2. Origen

La programación orientada a objetos tomó posición como el estilo de programación dominante a mediados de los años ochenta, en gran parte debido a la influencia del lenguaje C++ , una extensión del lenguaje de programación C. Su dominación fue consolidada gracias al auge de las interfaces gráficas de usuario, para los cuales la programación orientada a objetos está particularmente bien adaptada.

Las características de orientación a objetos fueron agregadas a muchos lenguajes existentes durante ese tiempo, incluyendo Ada, BASIC, Lisp, Pascal, entre otros. La adición de estas características a los lenguajes que no fueron diseñados inicialmente para ellas condujo a menudo a problemas de compatibilidad y a la capacidad de mantenimiento del código. Los lenguajes orientados a objetos "puros", por otra parte, carecían de las características de las cuales muchos programadores habían venido a depender. Para saltar este obstáculo, se hicieron muchas tentativas para crear nuevos lenguajes basados en métodos orientados a objetos, pero permitiendo algunas características imperativas de maneras "seguras". *Eiffel* fue un temprano y moderadamente acertado lenguaje con esos objetivos pero ahora ha sido esencialmente reemplazado por *Java*, en gran parte debido a la aparición de Internet, y a la implementación de la máquina virtual de Java en la mayoría de navegadores. *PHP* se ha ido modificando y soporta una orientación completa a objetos, cumpliendo todas las características propias de este paradigma.

Las principales diferencias entre la programación estructurada y la orientada a objetos son:

- La programación orientada a objetos es más moderna, es una evolución de la programación estructurada que plasma en el diseño de una familia de lenguajes conceptos que existían previamente con algunos nuevos.
- La programación orientada a objetos se basa en lenguajes que soportan sintáctica y semánticamente la unión entre los tipos abstractos de datos y sus operaciones (a esta unión se la suele llamar clase).
- La programación orientada a objetos incorpora en su entorno de ejecución mecanismos tales como el polimorfismo y el envío de mensajes entre objetos.

2.5.3. Conceptos fundamentales

La programación orientada a objetos es una nueva forma de programar que trata de encontrar una solución a los problemas al trabajar con lenguajes estructurados. Introduce nuevos

conceptos, que superan y amplían conceptos antiguos ya conocidos. Entre ellos destacan los siguientes:

- **Objeto**: entidad provista de un conjunto de propiedades o atributos (datos) y de comportamiento o funcionalidad (métodos). Corresponden a los objetos reales del mundo que nos rodea, o a objetos internos del sistema (del programa). Es una instancia a una clase.
- **Clase**: definiciones de las propiedades y comportamiento de un tipo de objeto concreto. La instanciación es la lectura de estas definiciones y la creación de un objeto a partir de ellas.
- **Método**: algoritmo asociado a un objeto (o a una clase de objetos), cuya ejecución se desencadena tras la recepción de un "mensaje". Desde el punto de vista del comportamiento, es lo que el objeto puede hacer. Un método puede producir un cambio en las propiedades del objeto, o la generación de un "evento" con un nuevo mensaje para otro objeto del sistema.
- **Evento**: un suceso en el sistema (tal como una interacción del usuario con la máquina, o un mensaje enviado por un objeto). El sistema maneja el evento enviando el mensaje adecuado al objeto pertinente. También se puede definir como evento, a la reacción que puede desencadenar un objeto, es decir la acción que genera.
- **Mensaje**: una comunicación dirigida a un objeto, que le ordena que ejecute uno de sus métodos con ciertos parámetros asociados al evento que lo generó.
- **Propiedad o atributo**: contenedor de un tipo de datos asociados a un objeto (o a una clase de objetos), que hace los datos visibles desde fuera del objeto y esto se define como sus características predeterminadas, y cuyo valor puede ser alterado por la ejecución de algún método.
- **Estado interno**: es una propiedad invisible de los objetos, que puede ser únicamente accedida y alterada por un método del objeto, y que se utiliza para indicar distintas situaciones posibles para el objeto (o clase de objetos).

2.5.4. Características de la POO

Hay un cierto desacuerdo sobre exactamente qué características de un método de programación o lenguaje le definen como "orientado a objetos", pero hay un consenso general en que las características siguientes son las más importantes:

- Abstracción: Cada objeto en el sistema sirve como modelo de un "agente" abstracto que puede realizar trabajo, informar y cambiar su estado, y "comunicarse" con otros objetos en el sistema sin revelar cómo se implementan estas características. Los procesos, las funciones o los métodos pueden también ser abstraídos y cuando lo están, una variedad de técnicas son requeridas para ampliar una abstracción.
- Encapsulamiento: Significa reunir a todos los elementos que pueden considerarse pertenecientes a una misma entidad, al mismo nivel de abstracción. Esto permite aumentar la cohesión de los componentes del sistema.
- Principio de ocultación: Cada objeto está aislado del exterior, es un módulo natural, y cada tipo de objeto expone una *interfaz* a otros objetos que especifica cómo pueden interactuar con los objetos de la clase. El aislamiento protege a las propiedades de un objeto contra su modificación por quien no tenga derecho a acceder a ellas, solamente los propios métodos internos del objeto pueden acceder a su estado.
- Polimorfismo: comportamientos diferentes, asociados a objetos distintos, pueden compartir el mismo nombre, al llamarlos por ese nombre se utilizará el comportamiento correspondiente al objeto que se esté usando. O dicho de otro modo, las referencias y las colecciones de objetos pueden contener objetos de diferentes tipos, y la invocación de un comportamiento en una referencia producirá el comportamiento correcto para el tipo real del objeto referenciado.
- Herencia: las clases no están aisladas, sino que se relacionan entre sí, formando una jerarquía de clasificación. Los objetos heredan las propiedades y el comportamiento de todas las clases a las que pertenecen. La herencia organiza y facilita el polimorfismo y el encapsulamiento permitiendo a los objetos ser definidos y creados como tipos especializados de objetos preexistentes. Estos pueden compartir (y extender) su comportamiento sin tener que reimplementar su comportamiento. Esto suele hacerse habitualmente agrupando los objetos en *clases* y estas en *árboles* que reflejan un comportamiento común.

2.6. Algunos lenguajes de programación

A continuación se exponen brevemente algunos lenguajes de programación que son más relevantes por su popularidad.

- C/C++: C es un lenguaje de programación creado en 1969 en los Laboratorios Bell. Es un lenguaje orientado a la implementación de Sistemas Operativos, concretamente Unix. C es

apreciado por la eficiencia del código que produce y es el lenguaje de programación más popular para crear software de sistemas, aunque también se utiliza para crear aplicaciones. El C++ es un lenguaje de programación, diseñado como extensión del lenguaje de programación C para soportar el paradigma de la orientación a objetos.

- C#: Es un lenguaje de programación orientado a objetos desarrollado y estandarizado por Microsoft como parte de su plataforma .NET, que después fue aprobado como un estándar por la ECMA e ISO. Su sintaxis básica deriva de C/C++ y utiliza el modelo de objetos de la plataforma .NET el cual es similar al de Java. C# fue diseñado para combinar el control de lenguajes de bajo nivel como C y la velocidad de programación de lenguajes de alto nivel como Visual Basic.
- Java: es un lenguaje de programación orientado a objetos desarrollado por Sun Microsystems a principios de los años 1990. El lenguaje en sí mismo toma mucha de su sintaxis de C y C++, pero tiene un modelo de objetos más simple y elimina herramientas de bajo nivel como punteros. La segunda característica, la *independencia de la plataforma*, significa que programas escritos en el lenguaje Java pueden ejecutarse igualmente en cualquier tipo de hardware y sistema operativo. Es lo que significa ser capaz de escribir un programa una vez y que pueda ejecutarse en cualquier dispositivo, tal como reza el axioma de Java, “write once, run everywhere”.
- Perl: Es un lenguaje de programación muy utilizado para construir aplicaciones para Internet. Se trata de un lenguaje de programación muy práctico para extraer información de archivos de texto y generar informes a partir del contenido de los ficheros. Es de libre uso y gratuito. Es un lenguaje de programación interpretado.
- Visual Basic: Visual Basic es un lenguaje de programación desarrollado por Alan Cooper para Microsoft. Su intención es la de simplificar la programación utilizando un ambiente de desarrollo completamente gráfico que facilitara la creación de interfaces gráficas y en cierta medida también la programación misma. Visual Basic fue descontinuado por Microsoft hace ya varios años. Muchos programadores están migrando a lenguajes similares como Delphi (que actualmente también ha sido descontinuado por Borland). Microsoft propone abandonar el desarrollo basado en la API Win32 y pasar a trabajar sobre un framework o marco común de librerías independiente de la versión del sistema operativo, .NET Framework, a través de Visual Basic .NET (y otros lenguajes como C-Sharp (C#) de fácil transición de código entre ellos) que presenta serias incompatibilidades con el código Visual Basic existente.

Bibliografía

- [1] Ambler, A. et al.: "Operational Versus Definitional: A Perspective on Programming". IEEE Computer. Sept. 1992.
- [2] Booch, G.: "Object-Oriented Analysis and Design with Applications". Benjamin Cumming Pub., Santa Clara (California), 1994.

Wikipedia

Tema 3. Arquitectura SOA. Entornos de desarrollo Web: Microsoft .NET y J2EE.

3.1. Arquitectura SOA

El acrónimo SOA proviene del inglés *Service-Oriented Architecture*. Se trata de un modelo de arquitectura que caracteriza el procedimiento para crear y usar los diversos procesos, reunidos en forma de servicios, que configuran un determinado Proceso de Negocio¹. Esta arquitectura define y proporciona la infraestructura necesaria para que el intercambio de información y la participación en los procesos de negocio se lleve a cabo con total independencia de la plataforma hardware-software sobre la que trabajan: sistema operativo, lenguaje de programación, características de los equipos, etc.

3.1.1. Antecedentes

En las últimas décadas los departamentos de Tecnologías de la Información de las empresas han construido una infraestructura que soporta en gran medida la operación de sus empresas y sus clientes. El resultado de este proceso ha sido la creación y mantenimiento de un número considerable de aplicaciones de uso interno, cada una responsable de sus propias tareas.

Los negocios exigen crear aplicaciones cada vez más complejas, en menos tiempo y con menor presupuesto. En muchos casos crear estas aplicaciones requiere de funcionalidades ya antes implementadas como parte de otros sistemas. Ante esta situación los arquitectos de software se enfrentan a dos opciones:

- Tratar de reutilizar la funcionalidad ya implementada en otros sistemas. Una labor difícil de realizar, debido a que estos no fueron diseñados para integrarse o se elaboraron para plataformas y/o tecnologías incompatibles entre ellas.
- Re-implementar la funcionalidad requerida. Aunque implica más tiempo de desarrollo, es en la mayoría de los casos la más fácil y segura.

A pesar de que no sea la más acertada a largo plazo, la segunda opción es la más escogida. Esto trae como resultado:

- Funcionalidad replicada en varias aplicaciones.
- Dificultad de migración de los sistemas internos, al haber múltiples conexiones desde sistemas que dependen de estos para su funcionamiento.
- Al no haber una estrategia de integración de aplicaciones, se generan múltiples puntos de fallo, que pueden detener la operación de todos los sistemas muy fácilmente.
- Un modelo así, por lo general no escala muy bien.

¹ Un proceso de negocio se puede ver como un conjunto estructurado de tareas, que contribuyen colectivamente a lograr los objetivos de una organización.

- El inconveniente final es una pobre respuesta al cambio. Las aplicaciones siguen siendo concebidas desde un principio como islas independientes.

3.1.2. Aspectos básicos

En la arquitectura SOA la funcionalidad deseada se descompone en unidades (servicios) que pueden ser distribuidos en diferentes nodos conectados a través de una red y que, asimismo, son combinados entre sí para alcanzar el resultado deseado. Estos servicios pueden proporcionar datos a otros o llevar a cabo actividades de coordinación entre uno o varios servicios.

Las aplicaciones necesarias para obtener los correspondientes procesos de negocio se logran mediante la combinación de colecciones de pequeños módulos llamados servicios. Estos módulos pueden ser empleados por grupos de usuarios provenientes de la propia organización o ajenos a la misma y las nuevas aplicaciones creadas del aprovechamiento de servicios presentes en un repositorio global muestran mayor flexibilidad y uniformidad. De este modo se consigue un ahorro en el esfuerzo de desarrollo pues se reaprovechan las funcionalidades comunes a las distintas aplicaciones además de favorecer la interacción entre organizaciones dado que se logra la homogeneización de la apariencia y del nivel y tipo de datos de entrada para la validación de los usuarios.

En este entorno de trabajo, las unidades básicas son los servicios. Los servicios son unidades de funcionalidad que desarrollan su actividad de forma independiente y que se aproxima al concepto que los humanos asocian a los mismos como puede ser la visualización del estado de una cuenta bancaria, o la emisión de una petición de un billete de avión o de tren. En lugar de que los servicios contengan en su código fuente llamadas a otros, se definen protocolos que describen cómo pueden comunicarse entre sí.

3.1.3. Colaboración entre servicios

La concepción de los servicios web, como otros muchos entornos dependientes del estado de la tecnología y del uso que de ella hacen los usuarios, ha ido cambiando a lo largo del tiempo. Inicialmente éstos fueron creados para implementar interacciones simples e independientes, sin embargo, su adecuación para ser utilizadas en entornos empresariales obliga a que los servicios se coordinen y colaboren entre sí.

La colaboración entre servicios consiste en la determinación de la secuencia de operaciones que debe acontecer en la interacción entre clientes y servidores. La secuencia debe respetar el orden establecido para que pueda ser válida y, con el objeto de que esto resulte viable, se define un protocolo de coordinación que será, precisamente, el encargado de describir el conjunto de secuencias válidas. Existen dos modelos de colaboración entre servicios: orquestación y coreografía.

- El modelo de orquestación se fundamenta en la existencia de un mecanismo de control centralizado que es el encargado de dirigir las actividades, siendo cada una de ellas una interacción entre servicios. Permite la definición de un modelo de interacción pero sólo desde el punto de vista del controlador. La orquestación define el comportamiento y la forma de llevarlo a cabo, y todos los eventos se supervisan de forma centralizada.

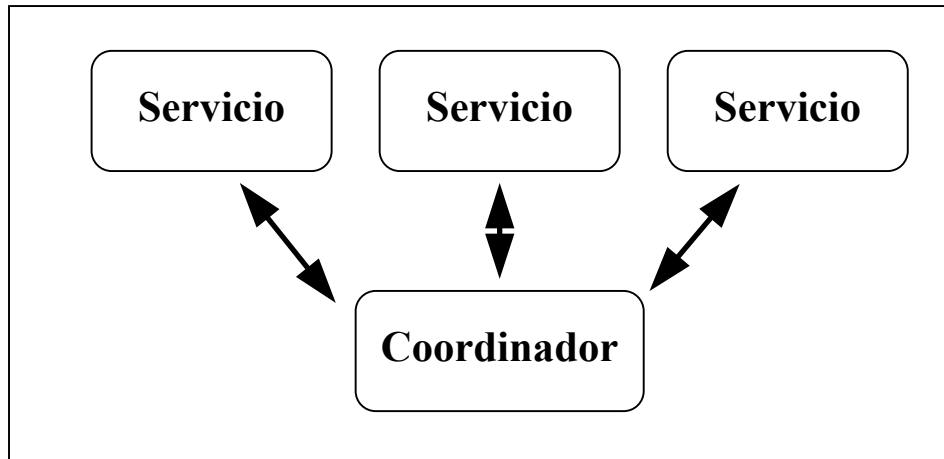


Figura 1. Colaboración entre servicios en el modelo de orquestación

- El modelo de coreografía describe el comportamiento que debe observarse entre las partes que interactúan. Cada una de las organizaciones implicadas en este modelo desarrolla independientemente el papel que desea jugar en la colaboración, la única condición es que respeten el “contrato global” que supone la coreografía descrita. La ejecución y el control es responsabilidad de los propios participantes.

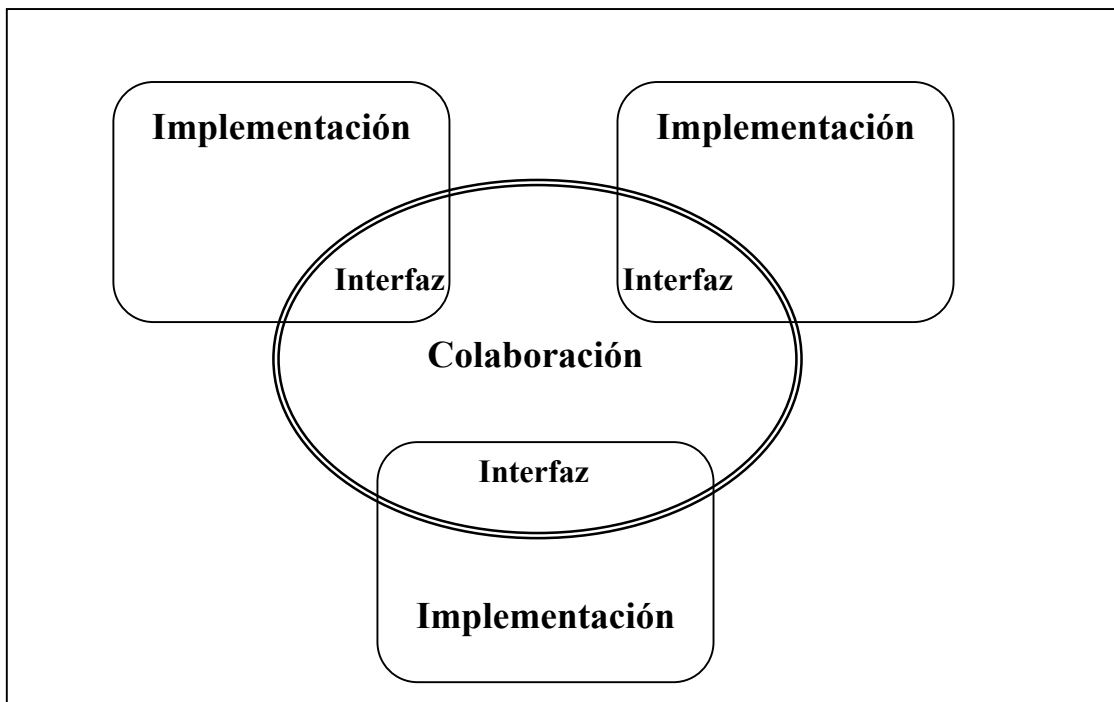


Figura 2. Colaboración entre servicios en el modelo de coreografía

3.1.4. Metadatos

Los metadatos son elementos imprescindibles dentro de esta arquitectura. Deben presentar información suficiente para describir tanto las características de los servicios como los datos que emplean.

El lenguaje XML suele ser el utilizado para crear los datos que conformarán el documento de descripción. Análogamente, los servicios suelen ser descritos mediante WSDL (*Web Service Description Language*) y los protocolos de comunicación mediante SOAP (*Simple Object Access Protocol*). Que estos lenguajes de descripción sean los más adecuados para desempeñar estas funciones y se consolide su uso todavía no está claro. Lo que sí resulta evidente es que una Arquitectura Orientada al Servicio depende completamente de los datos y servicios descritos a través de los correspondientes metadatos, independientemente de cuál sea el mecanismo para su obtención, y que deben reunir 2 requisitos: los metadatos tienen que entregarse de forma que puedan ser utilizados directamente por los propios sistemas y, simultáneamente, permitir que los diseñadores puedan entenderlos y usarlos de la manera adecuada.

3.1.5. Objetivo de la arquitectura SOA

La finalidad de la arquitectura SOA es conseguir combinar distintos módulos funcionales para generar aplicaciones de carácter específico, proviniendo todos de servicios preexistentes.

Cuanto mayor sea la funcionalidad proporcionada por estos módulos menor será el número de interfaces necesarios para alcanzar el objetivo deseado y cada interfaz conlleva un gasto de procesamiento adicional; sin embargo, cuando los módulos son excesivamente grandes resulta complicada su reutilización. Consecuentemente es necesario alcanzar el nivel de granulación adecuado. La expectativa creada por esta nueva arquitectura es que el coste marginal de creación de la *n*-ésima aplicación sea cero dado que el software requerido se encontrará ya disponible para satisfacer los requisitos; tan sólo se requerirá la coordinación entre los elementos.

Una de las claves de SOA es que no existen interacciones entre los módulos que se encuentren embebidas. La interacción entre los servicios, los cuales son independientes, es especificada por los usuarios. De ahí la necesidad de que los servicios presenten mayor funcionalidad que las tradicionales funciones o clases de los lenguajes de programación evitando la desbordante complejidad que la existencia de miles de objetos granulares supondría para el diseñador. Los servicios en sí mismos se desarrollan mediante el uso de lenguajes de programación como C, C++, Java, etc.

Los servicios SOA presentan un acoplamiento bajo², en oposición a las funciones que componen un fichero ejecutable. Los servicios se ejecutan en el interior de un envoltorio (*wrapper*) como Java o .NET, que se encarga de la gestión de la memoria, las invocaciones de los servicios, y el uso de tipos de datos.

Cada vez existe un mayor número de compañías de software que están ofreciendo servicios a cambio del pago de alguna tasa. En el futuro, los sistemas basados en la arquitectura SOA podrán estar compuestos de los servicios ofertados por estas empresas en combinación con otros creados por la propia organización o institución. La gran ventaja es que se reparte el coste de desarrollo entre los participantes, y éstos disfrutan y promueven la estandarización tanto en la propia empresa como entre empresas. En particular, el sector vinculado a los viajes dispone de un amplio número de servicios y datos bien definidos y documentados, suficientes para que cualquier desarrollador pueda crear software para agencias de viajes usando únicamente servicios software ya existentes. Otros sectores como el financiero, están realizando grandes progresos en esta dirección.

La arquitectura SOA se erige sobre un principio fundamental, la orientación al servicio. En un entorno SOA, los servicios pueden ser utilizados sin conocimiento alguno de las características de la plataforma sobre la que se despliegan.

² Cuando existe un acoplamiento bajo, los servicios tienen que ser independientes los unos de los otros. Posteriormente, se desarrolla este concepto en el apartado "Requisitos de una Arquitectura Orientada a Servicios"

3.1.6. Elementos de SOA

Esta arquitectura presenta un modelo de construcción sistemas distribuidos en el que la funcionalidad demandada será entregada a la aplicación a través de servicios. En la figura 3 se muestra el esquema de la arquitectura y los elementos que podrían observarse.

Como puede observarse, el esquema se encuentra dividido en 2 zonas; una que abarca el ámbito funcional de la arquitectura y otra vinculada a la calidad de servicio.

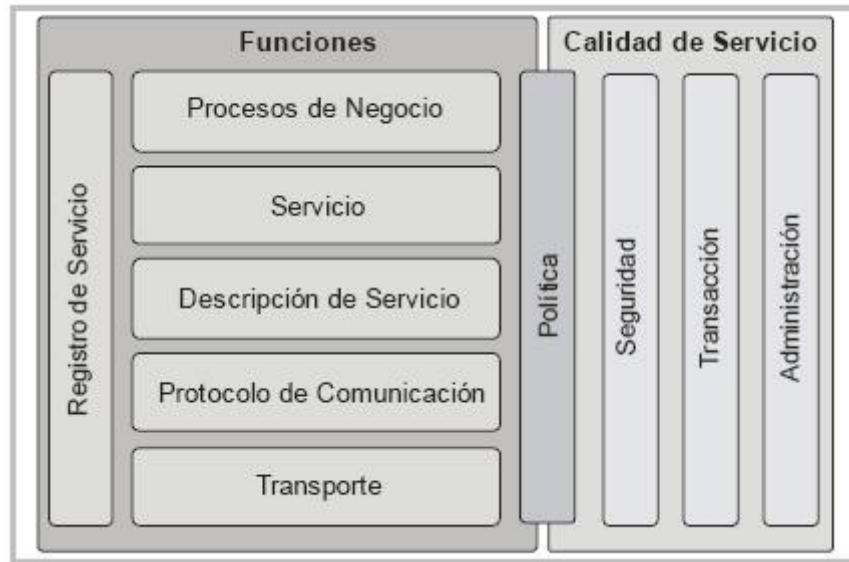


Figura 3. Elementos de la arquitectura SOA

A continuación se describen brevemente los elementos representados en la figura anterior:

- Funciones:
 - Transporte: es el mecanismo utilizado para llevar las demandas de servicio desde un consumidor de servicio hacia un proveedor de servicio, y las respuestas desde el proveedor hacia el consumidor.
 - Protocolo de comunicación de servicios: es un mecanismo acordado a través del cual un proveedor de servicios y un consumidor de servicios comunican qué está siendo solicitado y qué está siendo respondido.
 - Descripción de servicio: es un esquema acordado para describir qué es el servicio, cómo debe invocarse, y qué datos requiere el servicio para invocarse con éxito.
 - Servicio: describe un servicio actual que está disponible para utilizar.
 - Procesos de Negocio: es una colección de servicios, invocados en una secuencia particular con un conjunto específico de reglas, para satisfacer un requisito de negocio.
 - Registro de Servicios: es un repositorio de descripciones de servicios y datos que pueden utilizar los proveedores de servicios para publicar sus servicios, así como los consumidores de servicios para descubrir o hallar servicios disponibles.

- Calidad de Servicio:

- Política: es un conjunto de condiciones o reglas bajo las cuales un proveedor de servicio hace el servicio disponible para consumidores.
- Seguridad: es un conjunto de reglas que pueden aplicarse para la identificación, autorización y control de acceso a consumidores de servicios.
- Transacciones: es el conjunto de atributos que podrían aplicarse a un grupo de servicios para entregar un resultado consistente.
- Administración: es el conjunto de atributos que podrían aplicarse para manejar los servicios proporcionados o consumidos.

Las colaboraciones en SOA siguen el paradigma descubrir, ligar e invocar³, donde un consumidor de servicios realiza la localización dinámica de un servicio consultando el registro de servicios para hallar uno que cumpla con un determinado criterio. Si el servicio existe, el registro proporciona al consumidor la interfaz de contrato y la dirección del servicio proveedor. El siguiente diagrama ilustra las entidades (roles, operaciones y artefactos) en una arquitectura orientada a servicios donde éstas colaboran.



Figura 4. Relación entre las entidades

Cada entidad puede tomar uno de los tres roles posibles correspondientes a consumidor, proveedor y/o registro:

- Un consumidor de servicios es una aplicación, un módulo de software u otro servicio que demanda la funcionalidad proporcionada por un servicio, y la ejecuta de acuerdo a un contrato de interfaz.
- Un proveedor de servicios es una entidad accesible a través de la red que acepta y ejecuta consultas de consumidores, y publica sus servicios y su contrato de interfaces en el registro de servicios para que el consumidor de servicios pueda descubrir y acceder al servicio.

³ En inglés *find, bind and invoke*.

- Un registro de servicios es el encargado de hacer posible el descubrimiento de servicios, conteniendo un repositorio de servicios disponibles y permitiendo visualizar las interfaces de los proveedores de servicios a los consumidores interesados.

Las operaciones que pueden llevar a cabo las entidades son:

- Publicar. Para poder acceder a un servicio se debe publicar su descripción para que un consumidor pueda descubrirlo e invocarlo.
- Descubrir. Un consumidor de servicios localiza un servicio que cumpla con un cierto criterio consultando el registro de servicios.
- Ligar e Invocar. Una vez obtenida la descripción de un servicio por parte de un consumidor, éste lo invoca haciendo uso de la información presente en la descripción del servicio.

Finalmente, los artefactos en una arquitectura orientada a servicios son:

- Servicio. Un servicio que está disponible para el uso a través de una interfaz publicada y que permite ser invocado por un consumidor de servicios.
- Descripción de servicio. Una descripción de servicio especifica la forma en que un consumidor de servicio interactuará con el proveedor de servicio, especificando el formato de consultas y respuestas desde el servicio. Esta descripción también puede especificar el conjunto de precondiciones, pos-condiciones y/o niveles de calidad de servicio (QoS⁴).

3.1.7. Requisitos de una Arquitectura Orientada a Servicios

- No existe una definición estándar de cuales son los Principios de la Orientación a Servicios, por lo tanto, lo único que se puede proporcionar es un conjunto de Principios que estén muy asociados con la Orientación a Servicios. Estos Principios según Thomas Erl son:
- Los Servicios deben ser reutilizables: Todo servicio debe ser diseñado y construido pensando en su reutilización dentro de la misma aplicación, dentro del dominio de aplicaciones de la empresa o incluso dentro del dominio público para su uso masivo.
- Los Servicios deben proporcionar un contrato formal: Todo servicio desarrollado, debe proporcionar un contrato en el cual figuren: el nombre del servicio, su forma de acceso, las funcionales que ofrece, los datos de entrada de cada una de las funcionalidades y los datos de salida. De esta manera, todo consumidor del servicio, accederá a este mediante el contrato, logrando así la independencia entre el consumidor y la implementación del propio servicio. En el caso de los Servicios Web, se logrará mediante la definición de interfaces con WSDL.
- Los Servicios deben tener bajo acoplamiento: Los servicios tienen que ser independientes los unos de los otros. Para lograr ese bajo acoplamiento, lo que se hará es que cada vez que se vaya a ejecutar un servicio, se accederá a él a través del contrato, logrando así la

⁴ QoS, *Quality of Service*, en castellano Calidad de Servicio.

independencia entre el servicio que se va a ejecutar y el que lo llama. Si se consigue este bajo acoplamiento, entonces los servicios podrán ser totalmente reutilizables.

- Los Servicios deben permitir la composición: Todo servicio debe ser construido de tal manera que pueda ser utilizado para construir servicios genéricos de alto nivel a partir de servicios de bajo nivel. En el caso de los Servicios Web, esto se logrará mediante el uso de los protocolos para orquestación (WS-BPEL⁵) y coreografía (WS-CDL⁶).
- Los Servicios deben ser autónomos: Todo servicio debe tener su propio entorno de ejecución. De esta manera el servicio es totalmente independiente y nos podemos asegurar que así podrá ser reutilizable desde el punto de vista de la plataforma de ejecución.
- Los Servicios no deben tener estado: Un servicio no debe guardar ningún tipo de información. Esto es así porque una aplicación está formada por un conjunto de servicios, lo que implica que si un servicio almacena algún tipo de información, se pueden producir problemas de inconsistencia de datos. La solución, es que un servicio sólo contenga lógica, y que toda información esté almacenada en algún sistema de información sea del tipo que sea.
- Los Servicios deben poder ser descubiertos: Todo servicio debe poder ser descubierto de alguna forma para que pueda ser utilizado, consiguiendo así evitar la creación accidental de servicios que proporcionen las mismas funcionalidades. En el caso de los Servicios Web, el descubrimiento se logrará publicando sus interfaces en registros UDDI.

3.1.8. SOA y los Servicios Web

La arquitectura SOA es independiente de la tecnología utilizada. Puede implementarse mediante el uso de un amplio elenco de tecnologías entre las que se incluyen DCOM⁷, CORBA⁸ o Servicios Web. SOA puede ser implementado mediante el uso de uno o más de estos protocolos. La clave es, tal y como se ha comentado anteriormente, la existencia de servicios independientes con interfaces perfectamente definidas y que pueden ser invocados para efectuar sus tareas de una forma estándar sin que sea necesario conocer el funcionamiento interno de los mismos.

Ha sido precisamente el gran auge de los Servicios Web el que ha hecho resurgir el concepto de SOA.

⁵ WS-BPEL, *Web Services-Business Process Execution Language*, en castellano Lenguaje de Ejecución de Procesos de Negocio de los Servicios Web. Es un lenguaje que se desarrolló a partir de WSDL y XLANG, ambos lenguajes basados en XML orientados a la descripción de servicios Web. Consiste en un lenguaje de orquestación, no coreográfico, diseñado para el control centralizado de la invocación de diferentes servicios Web, con cierta lógica de negocio añadida que ayudan a la programación en gran escala

⁶ WS-CDL, *Web Services-Coreography Description Language*, en castellano Lenguaje de Descripción de la Coreografía de los Servicios Web. Es un lenguaje basado en XML que describe colaboraciones *peer to peer* de los participantes definiendo, desde un punto de vista global, un comportamiento observable común y complementario; donde ordenado el mensaje, intercambia el resultado de acuerdo a un objetivo de negocios común.

⁷ Distributed Component Object Model (DCOM), en español Modelo de Objetos de Componentes Distribuidos, es una tecnología propietaria de Microsoft para desarrollar componentes software distribuidos sobre varios ordenadores y que se comunican entre sí. Ha sido abandonada por .NET

⁸ Estándar definido por el Object Management Group (OMG) que establece una plataforma de desarrollo de sistemas distribuidos facilitando la invocación de métodos remotos bajo un paradigma orientado a objetos.

Los Servicios Web se han convertido en el estandarte de SOA, ya que esta tecnología posee un conjunto de características que permiten cubrir todos los principios de la orientación a servicios, no como otras posibles tecnologías de implementación como colas de mensajes o CORBA.

Cuando la arquitectura SOA se implementa mediante el uso de Servicios Web se utilizan los principios y tecnologías básicos de éstos; SOAP como lenguaje de intercambio, WSDL como lenguaje para la descripción de los servicios y UDDI para la publicación o registro de los mismos. En el dibujo que se muestra a continuación, se puede ver la estructura básica de funcionamiento.

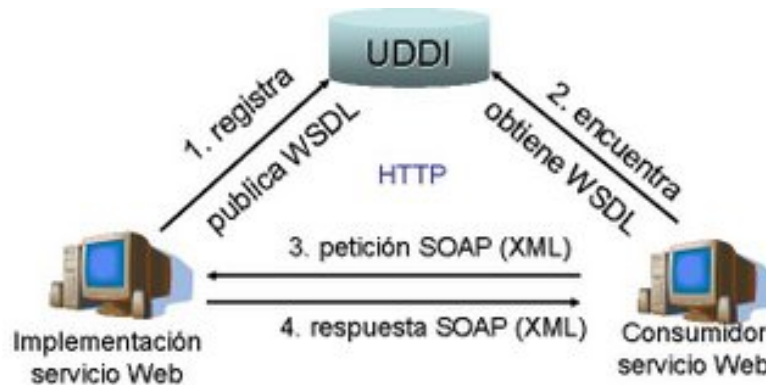


Figura 5. Esquema de intercambio de mensajes en los Servicios Web

Cuatro son los elementos fundamentales a la hora de desarrollar servicios web: XML, SOAP, WSDL y UDDI.

- XML, *eXtensible Markup Language*. Metalenguaje extensible de etiquetas desarrollado por el World Wide Web Consortium (W3C). Es una simplificación y adaptación del SGML⁹ y permite definir la gramática de lenguajes específicos. Tiene un papel muy importante ya que permite la compatibilidad entre sistemas para compartir la información de una manera segura, fiable y fácil.
- SOAP, *Simple Object Access Protocol*. La especificación SOAP indica cómo se deben codificar los mensajes que circularán entre las dos aplicaciones. La especificación SOAP define dos modelos de mensajes:
 - Un mensaje que se enviará desde la aplicación cliente a la aplicación servidor, solicitando la ejecución de un método al que se pasan una serie de parámetros.
 - Un mensaje que se enviará desde la aplicación servidor al cliente, y que contendrá datos XML con los resultados de la ejecución del método solicitado.
- WSDL, *Web Services Description Language*. Permite describir los distintos métodos o funciones que están disponibles en un servicio web, así como su signatura, es decir, el

⁹ SGML, *Standard Generalized Markup Language*. Sirve para especificar las reglas de etiquetado de documentos y no impone en sí ningún conjunto de etiquetas en especial

número de argumentos o parámetros que se les debe pasar, y el tipo de dato que devolverá la función como resultado.

Se establece una equivalencia entre el documento *WSDL* y un “contrato” que especifica los servicios que el servidor se compromete a ofrecer al cliente, siempre que éste los solicite de la forma adecuada.

Los documentos *WSDL* deben estar disponibles en el servidor web que ofrece los servicios.

- **UDDI**, *Universal Description, Discovery and Integration*. Se trata de un sistema de referencia encargado de posibilitar la localización de servicios. Éste proyecto adquiere mayor importancia conforme aumentan el número de servicios disponibles, si bien no es uno de los componentes básicos de la tecnología sobre la que se construye el paradigma de los servicios web.

En la figura 5, se puede observar claramente la existencia de tres roles claramente diferenciados:

- **Cliente del servicio**: Es el que solicita la ejecución del servicio web, y por lo tanto el que lo consume.
- **Proveedor del servicio**: Es el encargado de implementar el servicio web y ofrecerlo a los clientes.
- **Registro del servicio**: Es un repositorio donde se almacenan las descripciones de los servicios, para que así los clientes puedan buscar el servicio web que mejor se adapte a sus necesidades.

La secuencia de ejecución es la siguiente:

1. El proveedor del servicio da de alta el servicio web en el registro. Para ello, el proveedor almacena en el registro el documento de descripción.
2. El solicitante del servicio busca en el registro un servicio web que pueda adaptarse a sus necesidades.
3. Una vez seleccionado el servicio, el solicitante lo invoca mediante el envío de un mensaje SOAP, en el cual se indica la acción a realizar y los datos de entrada.
4. El servicio web recibe la petición y ejecuta la funcionalidad. Para finalizar envía un mensaje SOAP al solicitante con los resultados obtenidos.

3.2. Entornos de desarrollo Web: Microsoft .NET y J2EE

3.2.1. Arquitectura de aplicaciones de .NET: Diseño de aplicaciones y servicios

Todas las soluciones de software contienen tipos de componentes similares, independientemente de las necesidades empresariales que deban cubrir. Por ejemplo, la mayoría de las aplicaciones contienen componentes que tienen acceso a datos, encapsulan reglas empresariales y controlan la interacción con el usuario, entre otros. La identificación de los tipos de componentes que se

encuentran normalmente en las soluciones de software distribuidas facilitará la elaboración de un plano técnico para el diseño de aplicaciones o servicios.

3.2.2. Tipos de componentes

El análisis de la mayoría de las soluciones empresariales basadas en modelos de componentes por capas muestra que existen varios tipos de componentes habituales. En la figura 6 se muestra una ilustración completa en la que se indican estos tipos de componentes.

Aunque la lista que se muestra en la figura 6 no es completa, representa los tipos de componentes de software más comunes encontrados en la mayoría de las soluciones distribuidas. A lo largo de este tema describiremos en profundidad cada uno de estos tipos.

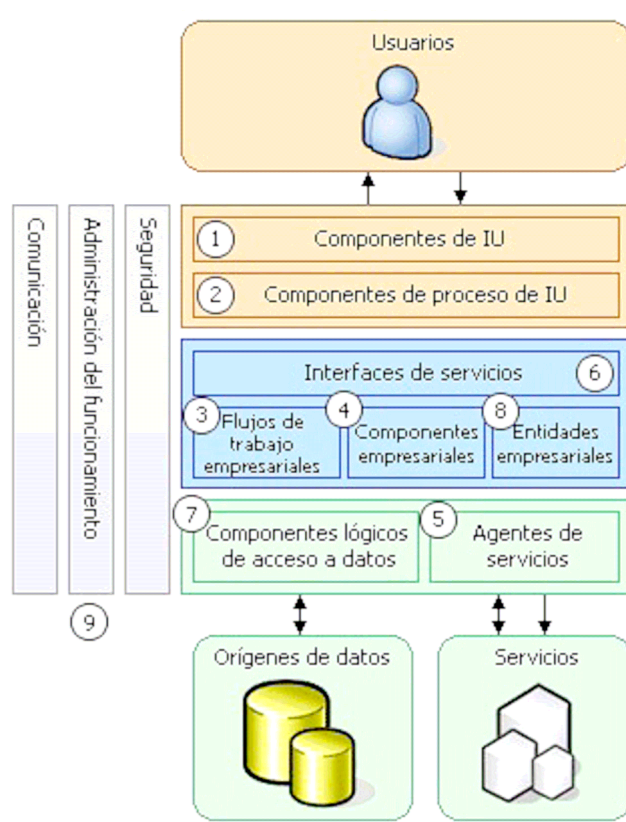


Figura 6. Tipos de componentes utilizados en el escenario comercial de ejemplo

Los tipos de componentes identificados en el escenario de diseño de ejemplo son:

1. Componentes de interfaz de usuario (IU).

La mayor parte de las soluciones necesitan ofrecer al usuario un modo de interactuar con la aplicación. Una aplicación comercial, un sitio Web permite al cliente ver productos y realizar pedidos, y una aplicación basada en el entorno operativo Microsoft Windows® permite a los representantes de ventas escribir los datos de los pedidos de los clientes que han telefonado a la empresa. Las interfaces de usuario se implementan utilizando fundamentalmente páginas

Microsoft ASP.NET, que permiten procesar y dar formato a los datos de los usuarios, así como adquirir y validar los datos entrantes procedentes de éstos.

2. Componentes de proceso de usuario.

En un gran número de casos, la interacción del usuario con el sistema se realiza de acuerdo a un proceso predecible. Por ejemplo, podríamos implementar un procedimiento que permita ver los datos de un producto. De este modo, el usuario puede seleccionar una categoría de una lista de categorías de productos disponibles y, a continuación, elegir uno de los productos de la categoría seleccionada para ver los detalles correspondientes. Del mismo modo, cuando el usuario realiza una compra, la interacción sigue un proceso predecible de recolección de datos por parte del usuario, por el cual éste en primer lugar proporciona los detalles de los productos que desea adquirir, a continuación los detalles de pago y, por último, la información para el envío. Para facilitar la sincronización y organización de las interacciones con el usuario, resulta útil utilizar componentes de proceso de usuario individuales. De este modo, el flujo del proceso y la lógica de administración de estado no se incluyen en el código de los elementos de la interfaz de usuario, por lo que varias interfaces podrán utilizar el mismo "motor" de interacción básica.

Los componentes de proceso de usuario se implementan normalmente como clases .NET que exponen métodos a los cuales pueden llamar las interfaces de usuario. Cada método encapsula la lógica necesaria para realizar una acción específica en el proceso de usuario. La interfaz de usuario crea una instancia del componente del proceso de usuario y la utiliza para efectuar la transición a través de los pasos del proceso. Los nombres de las páginas ASP.NET que se deben visualizar para cada uno de los pasos del proceso se pueden insertar en el código del componente de proceso de usuario (enlazándolo estrechamente por tanto a implementaciones específicas de la interfaz de usuario) o se pueden recuperar de un almacén de metadatos, como un archivo de configuración (facilitando la reutilización del componente de proceso de usuario por parte de varias implementaciones de interfaz de usuario). El diseño de componentes de proceso de usuario para su uso por parte de varias interfaces da lugar a una implementación más compleja, debido al aislamiento de los problemas específicos de los dispositivos. No obstante, puede facilitar la distribución del trabajo de desarrollo de la interfaz de usuario entre varios equipos, cada uno de ellos utilizando el mismo componente de proceso de usuario.

3. Flujos de trabajo empresariales.

Una vez que el proceso de usuario ha recopilado los datos necesarios, éstos se pueden utilizar para realizar un proceso empresarial. Por ejemplo, tras enviar los detalles del producto, el pago y el envío a la aplicación comercial, puede comenzar el proceso de cobro y preparación del envío. Gran parte de los procesos empresariales conllevan la realización de varios pasos, los cuales se deben organizar y llevar a cabo en un orden determinado. Por

ejemplo, el sistema empresarial necesita calcular el valor total del pedido, validar la información de la tarjeta de crédito, procesar el pago de la misma y preparar el envío del producto. El tiempo que este proceso puede tardar en completarse es indeterminado, por lo que sería preciso administrar las tareas necesarias, así como los datos requeridos para llevarlas a cabo. Los flujos de trabajo empresariales definen y coordinan los procesos empresariales de varios pasos de ejecución larga y se pueden implementar utilizando herramientas de administración de procesos empresariales, como BizTalk Server Orchestration.

4. Componentes empresariales.

Independientemente de si el proceso empresarial consta de un único paso o de un flujo de trabajo organizado, la aplicación requerirá probablemente el uso de componentes que implementen reglas empresariales y realicen tareas empresariales. Por ejemplo, en una aplicación comercial, deberá implementar una funcionalidad que calcule el precio total del pedido y agregue el costo adicional correspondiente por el envío del mismo. Los componentes empresariales implementan la lógica empresarial de la aplicación.

Es posible crear componentes que encapsulen la lógica empresarial utilizando .NET Framework. El código administrado puede beneficiarse de Enterprise Services (COM+) para las transacciones distribuidas y otros servicios habitualmente necesarios en las aplicaciones distribuidas.

5. Agentes de servicios.

Cuando un componente empresarial requiere el uso de la funcionalidad proporcionada por un servicio externo, tal vez sea necesario hacer uso de código para administrar la semántica de la comunicación con dicho servicio. Por ejemplo, los componentes empresariales de una aplicación comercial podrían utilizar un agente de servicios para administrar la comunicación con el servicio de autorización de tarjetas de crédito y utilizar un segundo agente de servicios para controlar las conversaciones con el servicio de mensajería. Los agentes de servicios permiten aislar las idiosincrasias de las llamadas a varios servicios desde la aplicación y pueden proporcionar servicios adicionales, como la asignación básica del formato de los datos que expone el servicio al formato que requiere la aplicación.

6. Interfaces de servicios.

Para exponer lógica empresarial como un servicio, es necesario crear interfaces de servicios que admitan los contratos de comunicación (comunicación basada en mensajes, formatos, protocolos, seguridad y excepciones, entre otros) que requieren los clientes. Por ejemplo, un servicio de autorización de tarjetas de crédito debe exponer una interfaz de servicios que describa la funcionalidad que ofrece el servicio, así como la semántica de comunicación requerida para llamar al mismo. Las interfaces de servicios también se denominan fachadas empresariales.

7. Componentes lógicos de acceso a datos.

La mayoría de las aplicaciones y servicios necesitan obtener acceso a un almacén de datos en un momento determinado del proceso empresarial. Por ejemplo, una aplicación empresarial necesita recuperar los datos de los productos de una base de datos para mostrar al usuario los detalles de los mismos, así como insertar dicha información en la base de datos cuando un usuario realiza un pedido. Por tanto, es razonable abstraer la lógica necesaria para obtener acceso a los datos en una capa independiente de componentes lógicos de acceso a datos, ya que de este modo se centraliza la funcionalidad de acceso a datos y se facilita la configuración y el mantenimiento de la misma.

Independientemente del almacén de datos utilizado, la aplicación o el servicio utilizará componentes lógicos de acceso a datos para obtener acceso a los datos. Estos componentes abstraen la semántica del almacén de datos subyacente y la tecnología de acceso a datos (como ADO.NET) y proporcionan una interfaz simple de programación para la recuperación y realización de operaciones con datos.

8. Componentes de entidad empresarial.

La mayoría de las aplicaciones requieren el paso de datos entre distintos componentes. Por ejemplo, en una aplicación comercial puede ser necesario pasar una lista de productos de los componentes lógicos de acceso a datos a los componentes de la interfaz de usuario para que éste pueda visualizar dicha lista. Los datos se utilizan para representar entidades empresariales del mundo real, como productos o pedidos. Las entidades empresariales que se utilizan de forma interna en la aplicación suelen ser estructuras de datos, como conjuntos de datos, DataReader o secuencias de lenguaje de marcado extensible (XML), aunque también se pueden implementar utilizando clases orientadas a objetos personalizadas que representan entidades del mundo real necesarias para la aplicación, como productos o pedidos.

Enterprise Services (COM+) son una clara opción para constituir el entorno host para los componentes empresariales. Enterprise Services ofrecen a los componentes seguridad basada en funciones, control de transacciones heterogéneo, agrupación de objetos e interfaces basadas en mensajes a través de componentes en cola (entre otros). Puede no utilizar Enterprise Services en una aplicación. Sin embargo, para llevar a cabo operaciones más complejas que las realizadas en un único origen de datos, necesitará sus servicios, y el uso del modelo que proporciona Enterprise Services en la etapa inicial ofrece una pauta de crecimiento más adecuada para el sistema.

9. Componentes de seguridad, administración operativa y comunicación.

La aplicación probablemente utilice también componentes para realizar la administración de excepciones, autorizar a los usuarios a que realicen tareas determinadas y comunicarse con otros servicios y aplicaciones.

3.2.3. Java 2 Enterprise Edition

El lenguaje Java, que empezó siendo un lenguaje para sistemas empotrados y que saltó a la fama como lenguaje para el desarrollo de applets, ha pasado a ser en estos momentos un lenguaje maduro, moderno, para el que se han construido diferentes entornos de desarrollo dependiendo de los objetivos perseguidos. En estos momentos Java dispone de 3 plataformas de desarrollo:

1. J2ME (Java 2 Micro Edition). Es la plataforma dirigida al desarrollo de aplicaciones para dispositivos como teléfonos móviles o PDA's. Como el resto de las plataformas está formada por un conjunto de APIs definidas por grupos de trabajo en los que han colaborado empresas del sector de los dispositivos y empresas de software.
2. J2SE (Java 2 Standard Edition). Consiste en el entorno de ejecución estándar de Java. Como el resto, está formado por un conjunto de APIs para crear distintos tipos de aplicaciones como applets que se ejecutan en el entorno del navegador y aplicaciones cliente que se ejecutan sobre una máquina virtual que se distribuye tanto con el entorno de desarrollo como con el entorno de ejecución creado para la plataforma en la que la aplicación debe funcionar.
3. J2EE (Java 2 Enterprise Edition). Consiste en un conjunto de especificaciones dirigidas a facilitar el desarrollo de sistemas empresariales en arquitecturas multicapa.

El resto de este tema estudiaremos las características más importantes de J2EE como las especificaciones líder en las que han participado empresas en todos los ámbitos del software.

J2EE aparece con el objetivo de minimizar los problemas asociados al desarrollo de aplicaciones multicapa en Internet (entornos heterogéneos, integración de tecnologías, complejidad, servidores complejos y propietarios). Sun Microsystems, con el apoyo de otras grandes empresas del software, impulsó una iniciativa de recopilación y elaboración de una serie de especificaciones y recomendaciones para crear una arquitectura estándar que permitiera reducir el coste y la complejidad de estos desarrollos. Las especificaciones Java 2 Enterprise Edition recogen:

- Especificaciones básicas para los servidores de aplicaciones Java que permiten la movilidad entre distintos proveedores, y por lo tanto, creación de un mercado competitivo entre las empresas de software que provoca un abaratamiento de sus productos.
- Un conjunto extenso de librerías sobre las que poder desarrollar sistemas complejos.
- Un conjunto de manuales de buenas prácticas que permitan una formación sólida entre los profesionales dedicados al desarrollo de este tipo de sistemas.

Las especificaciones básicas forman lo que se conoce como la plataforma J2EE, que se ha convertido en la referencia sobre la que la práctica totalidad de fabricantes de servidores de aplicaciones desarrollan sus sistemas, y con la que el primer objetivo de crear un gran mercado competitivo se está consiguiendo.

Arquitectura J2EE

La arquitectura propuesta es una arquitectura multicapa en la que se definen 4 grandes bloques identificados por el contenedor en el que se ejecutan. Se definen dos tipos de contenedores que se encuentran en las máquinas cliente y dos tipos de contenedores que se encuentran en los servidores.

Haciendo uso del patrón MVC¹⁰ podríamos decir que los contenedores de cliente y parte de los componentes del contenedor web son la vista, en el contenedor web está el controlador y en el contenedor EJB está el modelo.

Evidentemente esta es una aproximación ya que puede existir parte del modelo en el contenedor web o incluso en el contenedor de aplicaciones cliente.

A partir de este punto vamos a repasar los elementos que aparecen en el diagrama, es decir: componentes, contenedores y servicios.

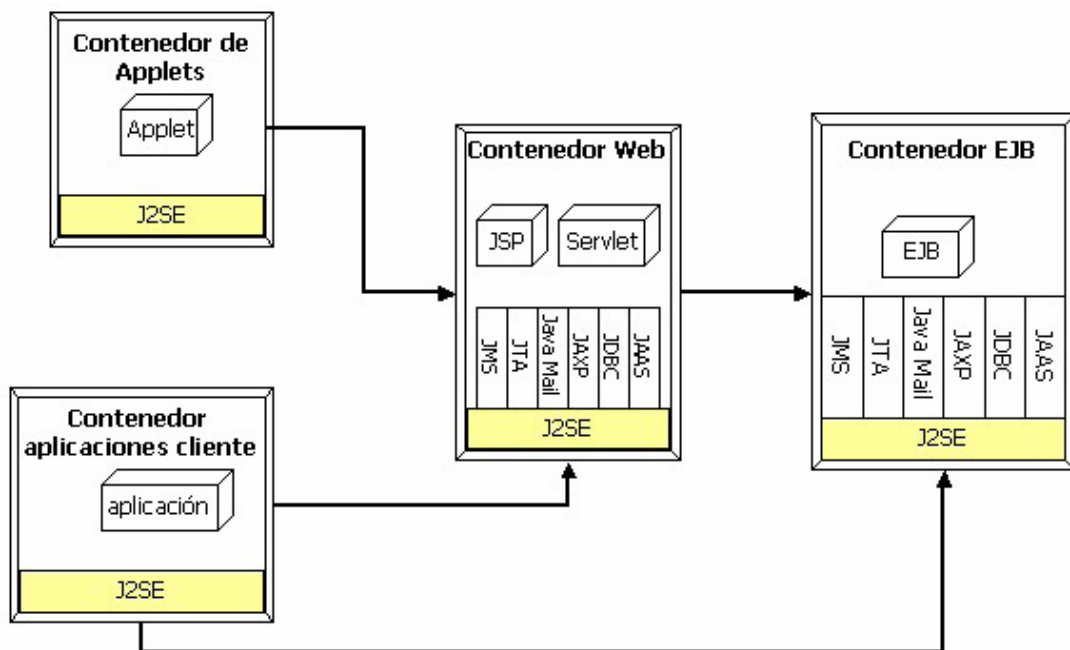


Figura 7. Arquitectura J2EE

Componentes Java para el desarrollo web

Servlets

De forma resumida podemos decir que los servlets son componentes web desarrollados en Java y utilizados para generar contenido dinámico: clases Java que extienden la interfaz HttpServlet

y que son invocadas a través de una URL pudiendo generar dinámicamente diversos tipos de contenido como texto, imágenes y sonido, aunque normalmente es texto con formato HTML o XML. Este contenido es devuelto al cliente mediante el protocolo HTTP.

Un servlet se instala, gestiona y ejecuta en un componente del sistema denominado contenedor web. Más adelante en este tema veremos qué es un contenedor dentro de la plataforma J2EE, pero de momento lo que nos interesa es que este contenedor va a proporcionar al servlet independencia de la plataforma, ya que el desarrollador se podrá encargar sólo de cumplir con la especificación mientras que el contenedor se ocupará de su ejecución independientemente del entorno.

Para entender cómo funciona un servlet vamos a analizar el ciclo de vida del servlet desde su instalación hasta su desinstalación o parada.

1. El servlet es instalado en el contenedor. Básicamente esta instalación consiste en indicar en el fichero de configuración del contenedor la URL a la que responderá y la clase que implementará, y por supuesto copiar la clase en el directorio del servidor.
2. Cuando el contenedor recibe por primera vez una petición de un servlet, el contenedor crea un objeto de la clase Java que lo implementa. Al ser Java un entorno multi-hilo, los servlets están diseñados para permitir dar respuesta a un alto número de solicitudes concurrentes.
3. Para cada petición:
 - El cliente (normalmente un navegador web) envía la petición HTTP al servidor web.
 - El servidor web identifica la petición como una URL que está asociada con algo que se debe ejecutar en el contenedor web y reenvía la petición al contenedor.
 - El contenedor (o motor) de servlets tiene una configuración en la que se indica para cada URL solicitada qué objeto debe crear. Por lo tanto construye o, si ya estuviese creado por una petición anterior utiliza, el objeto de la clase correspondiente a la petición de la URL.
 - Una vez creado el objeto, ejecuta el método doGet() o doPost() (según se indique en la petición HTTP) del objeto. Todo servlet tendrá estos métodos puesto que lo obliga la interfaz HttpServlet.
 - El servlet analiza la información que viene en el request de la petición HTTP y realiza las llamadas pertinentes a otros componentes Java, como son los objetos de negocio, de cara a obtener la información necesaria.
 - Finalmente el servlet genera la respuesta en el response construyendo la página HTML, XML o lo que sea que el servidor web vaya a enviar al cliente.

Los servlets se han mostrado adecuados para determinadas funcionalidades dentro de las aplicaciones corporativas, ya que tienen toda la versatilidad de un lenguaje de propósito general como Java que permite estructurar secuencias de control complejas sin obtener un código

¹⁰ MVC, Modelo, Vista y Controlador. Patrón utilizado en el diseño de sistemas.

excesivamente difícil de mantener. Funcionalidades como la implementación de controladores dentro de arquitecturas del tipo MVC. Así como para la generación de salidas binarias (imágenes, sonido, video, etc.)

Sin embargo, desde los inicios del desarrollo con servlets se comprobó que, aunque los servlets eran un excelente sistema para gestionar las peticiones web de contenido dinámico, no eran un sistema cómodo para generar la interfaz HTML de la respuesta ya que había que codificar mucho para conseguir un nivel aceptable de presentación. Por ello veremos que se han ideado otras tecnologías Java como JSP que aligeran esta fase.

JSP

Como hemos explicado en el apartado anterior los servlets demostraron ser un excelente sistema de gestionar las peticiones de contenido dinámico, pero no era un sistema cómodo para generar la respuesta ya que había que codificar mucho para generar un buen nivel de presentación en HTML. Sun Microsystems intentó resolver el problema formalizando un nuevo estándar llamado *JavaServer Pages* (JSP) en 1999. Siguiendo la misma filosofía de otras tecnologías de generación dinámica de páginas como Microsoft's Active Server Pages, una página JSP es un documento de texto que contiene HTML y bloques de código embebido que se ejecuta en el servidor para generar el contenido dinámico de esa página HTML.

Así, se puede decir que la tecnología JSP facilita el desarrollo de páginas web con contenido estático y dinámico. Como veremos un JSP es en realidad un servlet orientado a la presentación pues, como veremos, antes de ser ejecutado el motor que maneja las JSP preprocesa su contenido y las convierte en auténticos servlets. Esta tecnología no es más que una aproximación a la tecnología servlet de una forma más natural para la creación de la parte estática de una página web.

Una página JSP puede contener código en diferentes lenguajes:

- Código para interpretar en la máquina cliente:
 1. HTML
 2. Un lenguaje interpretado por navegadores como JavaScript o VBScript.
- Código para ejecutar en el servidor:
 1. Directivas JSP
 2. Código Java insertado entre las etiquetas “<%” y “%>”.
 3. Expresiones de Java entre las etiquetas “<%= ... %>” que se ejecutan y devuelven un valor que se escribe directamente en la salida.
 4. Etiquetas JSP estándar. Son etiquetas que realizan funciones básicas del manejo de las páginas JSP como, por ejemplo, la carga de librerías, la iteración de colecciones, acceso a beans, etc. Ejemplo: `<jsp:useBean id="Bean" scope="request" class="Clase"/>`

5. Etiquetas personalizadas (*custom tags*). Estas etiquetas son en realidad funcionalidades complejas encapsuladas en clases y publicadas en unos ficheros “.tld” para que se puedan invocar desde las páginas JSP con una sintaxis muy similar a la de las etiquetas HTML. Ejemplo: `<milibreria:mitag nombre=”ejemplo” />`

La diversidad de lenguajes que se pueden agrupar en un JSP ha sido el origen de la principal crítica a esta tecnología.

Siempre se argumenta que, aunque evidentemente hacen que el contenido estático se pueda desarrollar más fácilmente, cuando existe mucho contenido dinámico y las aplicaciones tienen funciones avanzadas de configuración, personalización, etc. las páginas JSP acaban siendo una mezcla de muchos lenguajes muy difíciles de seguir por los propios desarrolladores y, por lo tanto, tienen un alto coste, no sólo de desarrollo sino sobre todo de mantenimiento.

Una página JSP se puede ver como una página HTML con trozos de código Java que se ejecutan cuando se solicita la página. Pero lo que es importante entender es lo que sucede cuando se solicita una página JSP ya que, como se ha dicho más arriba, un JSP acaba siendo un servlet.

Existen dos fases claramente diferenciadas durante la ejecución de un JSP: La fase de generación de código y compilación del servlet generado y la fase de carga de la clase obtenida por la compilación del servlet:

1. El cliente (normalmente un navegador web) hace una petición HTTP al servidor web de una página JSP.
2. El servidor web identifica la petición como una URL que está asociada con algo que se debe ejecutar en el contenedor web y reenvía la petición al contenedor.
3. El contenedor “sabe” que la petición de una página JSP, en realidad, es la petición del servlet generado a partir de la página, por lo que busca la clase compilada y generada a partir del JSP. Si no existe esa clase, el contenedor interpreta la JSP para generar un servlet que es, en ese momento, compilado.
4. Una vez que encontrado ese servlet, el contenedor crea un objeto de la clase correspondiente a dicho servlet generado y compilado y, a partir de ese instante, continúa la ejecución como si fuera un servlet normal.

Inicialmente podríamos pensar que la ejecución de un JSP puede llegar a resultar muy lenta, ya que tenemos que realizar una generación de código y una compilación durante el tiempo que transcurre desde que se hace la petición hasta que se inicia la respuesta del sistema, pero, en realidad, el contenedor de JSP se encarga de optimizar este proceso realizando esas tareas solo cuando es necesario; algunos, incluso, permiten realizar esta generación y compilación en el mismo momento de la instalación con lo que incluso el rendimiento de la primera petición del usuario es exactamente igual al que tendría un servlet.

Enterprise JavaBeans

Los *Enterprise JavaBeans* son componentes de servidor que encapsulan la lógica de negocio de una aplicación y se pueden ejecutar de forma distribuida. Debido a que los *beans* se ejecutan en contenedores EJB, aprovechan los servicios de bajo nivel que éstos proporcionan permitiendo al desarrollador centrarse en resolver los problemas del negocio.

Los componentes EJB pueden encontrarse físicamente ubicados en máquinas distintas de las que albergan los componentes que les invocan. Para gestionar esta comunicación los componentes hacen uso del protocolo de llamada remota RMI/IIOP. Estas operaciones de invocación remota son bastante pesadas por lo que habrá que evaluar si realmente nos interesa pagar este precio para obtener la mejora de rendimiento en otros aspectos.

Existen tres tipos de *Enterprise Beans*:

- *Session beans*. Representan la transacción de un cliente en el servidor de aplicaciones J2EE. El bean desarrolla una tarea específica del negocio en el servidor.
- *Entity beans*. Representan las entidades del negocio que perduran en el tiempo, ya sea en una base de datos o cualquier otro sistema de almacenamiento persistente. Reflejan en memoria los datos que maneja la aplicación.
- *Message driven beans*. Los *message driven beans* actúan como *session beans* pero de forma asíncrona. En realidad, son una especie de componentes de tipo *listener* que se emplean para el procesamiento asíncrono de mensajes.

Contenedores

Las especificaciones J2EE definen un contenedor como el software responsable de ofrecer una serie de servicios de ejecución a los componentes J2EE de aplicación. Estos servicios los prestan a través del conjunto de APIs que define la especificación J2EE. Sin embargo estas llamadas a las APIs nunca se realizan directamente: se hacen siempre a través del contenedor; de este modo, el contenedor puede incluir la funcionalidad declarada en los ficheros de configuración proporcionados en la instalación de los componentes tales como gestión de transacciones, seguridad, gestión de estados y conexiones.

Contenedor Web

Un contenedor web es el encargado de ejecutar los componentes web (servlets y JSPs) y de proporcionarles los servicios necesarios de seguridad, concurrencia, transacciones, implantación, etc.

Es importante no confundir un contenedor web, con un servidor web: este último está optimizado fundamentalmente para servir páginas HTML estáticas. Un contenedor de web actúa como una capa intermedia de software entre el servidor web y el servlet. El servidor web transmite la petición HTTP al contenedor, éste crea un objeto que implementa la interfaz `HttpServletRequest` y pasa la petición al servlet. El servlet ejecuta su lógica, haciendo las llamadas oportunas a la lógica de negocio y genera una respuesta que el contenedor transmite al servidor web. Por cada petición HTTP el contenedor es responsable de crear y gestionar un nuevo hilo de ejecución que sirva la petición.

Contenedor EJB

El contenedor de EJB es el lugar donde residen y se ejecutan los EJBs instalados en un servidor de aplicaciones J2EE.

Proporciona a los EJB un conjunto de servicios como cache, concurrencia, persistencia de los objetos, seguridad, gestión de transacciones, bloqueos, etc. Un contenedor de EJBs se arranca automáticamente cuando se arranca el servidor de aplicaciones. En un mismo contenedor pueden residir varios beans.

Bibliografía

- [1] Eric Newcomer; Greg Lomow (2005). *Understanding SOA with Web Services*. Addison Wesley. ISBN 0-321-18086-0.
- [2] Thomas Erl (2005). *Service-Oriented Architecture: Concepts, Technology, and Design*. Upper Saddle River: Prentice Hall PTR. ISBN 0-13-185858-0.
- [3] Channabasavaiah, Holley and Tuggle, Migrating to a service-oriented architecture, *IBM DeveloperWorks*, 16 Dec 2003
- [4] Christopher Koch A New Blueprint For The Enterprise, *CIO Magazine*, Mar 1 2005
- [5] Yvonne Balzer Improve your SOA project plans, *IBM*, 16 July 2004
- [6] Thomas Erl Serviceorientation.org - About the Principles, 2005-2006
- [7] SOA Practitioners Guide Part 2: SOA Reference Architecture
- [8] Dion Hinchcliffe Is Web 2.0 The Global SOA?, *SOA Web Services Journal*, 28 October 2005
- [9] Jason Bloomberg Mashups and SOBAs: Which is the Tail and Which is the Dog?, *Zapthink*

- [10] Yefim Natis & Roy Schulte Advanced SOA for Advanced Enterprise Projects, *Gartner*, July 13, 2006
- [11] Sandra Rogers Patience Grasshopper...It's a Learning Process, *Business Trends Quarterly*, Q4, 2006
- [12] Joe McKendrick Anti-SOA 2.0 petition nears 400, *ZDNet.com*, June 29, 2006
- [13] Steve Jones Enterprise SOA Adoption Strategies, *InfoQ*, May 17, 2006, ISBN 978-1-84728-398-6
- [14] Agustín Cernuda y Jose Emilio Labra Validación estática de compatibilidad en procesos de negocio en coreografías de servicios web
- [15] José David Parra. *Hacia una Arquitectura Empresarial basada en Servicios*. Microsoft.
- [16] .Net Framework 1.0 Redistributable Requirements, MSDN
- [17] .Net Framework 1.1 Redistributable Requirements, MSDN
- [18] .Net Framework 2.0 Redistributable Requirements, MSDN
- [19] Perrone, Paul J.; *Chaganti, Krishna (2003)*. J2EE Developer's Handbook. *Indianapolis, Indiana: Sam's Publishing*. ISBN 0-672-32348-6.
- [20] *Bodoff, Stephanie (2004)*. The J2EE Tutorial. *Boston: Addison-Wesley*. ISBN 0-321-24575-X.
- [21] Solveig Haugland, Mark Cade, Anthony Orapallo: J2EE 1.4: The Big Picture, Prentice Hall, ISBN 0-13-148010-3
- [22] Alan Monnox: Rapid J2EE Development: An Adaptive Foundation for Enterprise Applications, Prentice Hall, ISBN 0-13-147220-8
- [23] Renaud Pawlak, Lionel Seinturier, Jean-Philippe Retraillé: Foundations of AOP for J2EE Development, ISBN 1-59059-507-6
- [24] Christopher Judd, Hakeem Shittu: Pro Eclipse JST: Plug-ins for J2EE Development, ISBN 1-59059-493-2

Tema 4. Almacenamiento estructurado. Sistemas de gestión de bases de datos SGBD. Bases de datos relacionales. Bases de datos orientadas a objetos. Bases de datos espaciales.

4.1. Almacenamiento estructurado

El Sistema de Información de una Empresa o de cualquier Organización es un sistema más al servicio de ésta y de sus objetivos, en ningún caso independiente de su estrategia y, como algunos autores afirman, resulta ser parte de su Infraestructura, ya que hoy no se considera perteneciente a ninguna unidad organizativa ni actividad determinada sino interdependiente con los demás sistemas componentes de la organización.

La implantación de un Sistema de Información utilizando las Tecnologías de la Información no ha sido un camino fácil de recorrer por diversas razones, entre las que queremos destacar la siguiente y quizás más obvia:

“No es posible hacer lo mismo que antes y de la misma manera, sólo que utilizando unas tecnologías diferentes, porque la tecnología no es neutral y aporta su propia idiosincrasia”.

Una de las aportaciones más relevantes que ha supuesto la informática para los Sistemas de Información es el concepto de Base de Datos (B.D.); al menos por las siguientes razones:

1. Incorpora una visión global del conjunto de datos de una organización en cuyo diseño todos sus departamentos deben estar de acuerdo.
2. Estructura el S.I. alrededor de la B.D. considerada como Núcleo del Sistema.

En los últimos años la evolución de los S.I. ha sido realmente notable, haciéndose cada vez mayores las exigencias de prestaciones de información almacenada en el Sistema, lo que ha supuesto un avance continuado de la tecnología de las B.D. y un fuerte impacto económico sobre las organizaciones que han aplicado estas tecnologías.

4.1.1. Ventajas de las Bases de Datos

El uso de las B.D. ha crecido vertiginosamente debido a las ventajas que presentan:

- En relación con los datos:
 - Independencia de éstos respecto a los tratamientos y viceversa
 - Disminución de las redundancias
 - Mayor disponibilidad de los mismos
 - Protección de los datos
 - Mayor eficiencia en la recogida, codificación y entrada en el sistema
- En relación con los resultados:
 - Mayor coherencia

- Mayor valor informativo
- Mayor eficiencia
- Mejor y más normalizada documentación de la información
- En relación con los usuarios:
 - Acceso más rápido y sencillo de los usuarios finales
 - Facilidades para compartir los datos

4.2. Sistemas de Gestión de Base de Datos

Las B.D. Requieren básica y fundamentalmente un Software de Gestión que facilite las operaciones y las interfaces con los usuarios. Este logical, es el Sistema de Gestión de Bases de Datos (S.G.B.D.)

Según, de Miguel, 1985, el S.G.B.D. es un conjunto coordinado de programas, procedimientos, lenguajes... que suministra, tanto a los usuarios no informáticos, como a los analistas programadores, o al administrador, los medios necesarios para describir y manipular los datos integrados en la base, manteniendo su integridad, confidencialidad y seguridad.

Este software capaz de soportar B.D. proporciona facilidades para definir Esquemas Conceptuales a través del llamado L.D.D.¹ y además suministra operaciones para consultar y actualizar (altas, bajas y modificaciones) la información residente en la B.D. A este segundo tipo de lenguaje se le conoce por L.M.D.² (Lenguaje de Manipulación de Datos).

La arquitectura de los S.G.B.D. quedó establecida en 1975.

El Comité ANSI/X3/SPARC (grupo de estudio del Standards Planning and Requirements Committee del Organismo Estadounidense de Normalización - American National Standards Institute -) viene ocupándose desde principios de los años 70 de la normalización de los SGBD. Después de una serie de informes parciales, publica en 1975 un informe provisional ANSI (1975), donde propone la arquitectura a 3 niveles. Un informe posterior ANSI (1977) revisa y detalla esta arquitectura. Prosiguen los trabajos y en Mayo de 1985 aparece un nuevo informe del DAFTG (Database Architecture Framework TaskGroup) subgrupo del DBSSG (Database System Study Group) en donde se presenta el Modelo de Referencia para la estandarización de los SGBD (ANSI-1986)

¹ L.D.D., Lenguaje de Definición de Datos

² L.M.D., Lenguaje de Manipulación de Datos

4.2.1. Evolución de la gestión de los datos

Datos orientados hacia el proceso. Dependencia Datos-Aplicaciones

En los comienzos de la historia de la informática los datos estaban integrados en los programas como constantes.

Con la aparición de los ficheros como colección de datos homogénea, es decir, integrados por elementos estructurados de la misma manera, en el que cada elemento contiene el mismo tipo de información y almacenada en soporte informático determinado, se diferencian de manera incipiente la estructura lógica que representa el punto de vista del usuario y la estructura física de los datos (en general cinta magnética y por lo tanto organización secuencial de los datos).

Posteriormente, la definición de ficheros independientes del resto del programa podrá llevarse a cabo por medio del lenguaje de programación. De esta manera, se facilita el acceso y actualización de los ficheros, y se evita la programación de muchas tareas repetitivas. Como consecuencia de la aparición de subsistemas de gestión de datos integrados en los sistemas operativos las estructuras lógica y física se empiezan a diferenciar.

El nivel de diferenciación entre las estructuras lógica y física alcanzado hasta este momento no es suficiente para evitar la dependencia casi total de los datos respecto a los programas y viceversa, y de ambos respecto a la máquina.

Con el fin de atenuar estas dependencias (entre datos y aplicaciones) se pasa a una arquitectura que diferencia claramente la representación de los datos orientados hacia el problema (estructura lógica de los datos) de la representación orientada hacia la máquina (estructura física de los datos), siendo necesaria una transformación (mapping) de una en otra.

Aparición de las Bases de Datos. Arquitectura a dos niveles

En la década de los 60 nacen los primeros S.G.B.D. Los sistemas dejan de estar orientados al proceso o tratamiento y se orientan hacia las B.D.

Los datos y sus interrelaciones se integran en las B.D. y se aíslan de las aplicaciones. La estructura lógica se hace más flexible y sencilla y la estructura física se complica buscando mejorar el rendimiento.

Los S.G.B.D. facilitan la descripción y el almacenamiento de las interrelaciones permitidas, que a su vez caracterizan los distintos Modelos de Datos Jerárquico, red, relacionar.

El tipo de Arquitectura de los primeros S.G.B.D. era a dos niveles:

1. Estructura Global (características lógicas y físicas): Esquema
2. Vistas Lógicas Externas de los usuarios: Subesquemas

Arquitectura a tres niveles

En 1975 el Organismo de Estandarización de Los Estados Unidos (ANSI) publica un informe clave para la posterior evolución de los S.G.B.D. Insiste en la necesidad de conseguir total independencia entre datos y aplicaciones; propone una arquitectura a 3 niveles y define el modelo conceptual para conseguir tales objetivos (figura 1)

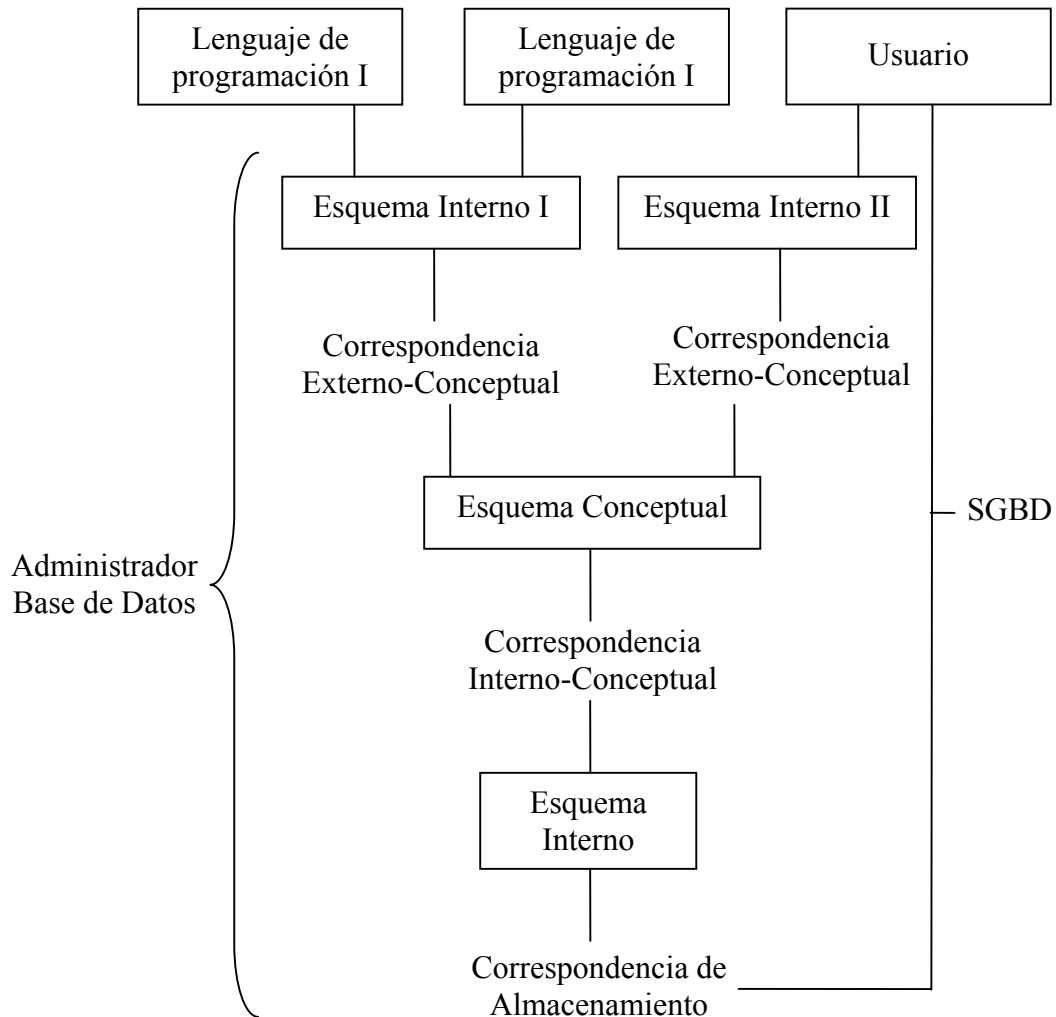


Figura 1. Arquitectura de una Base de Datos de 3 niveles

La Estructura global de la arquitectura a 2 niveles se divide en dos, de forma que en una estructura quedan los aspectos lógicos y el Esquema Conceptual y en otra los aspectos físicos: Esquema Interno.

La arquitectura interna de SGBD que sigue esta normativa presenta en su interior tres niveles perfectamente diferenciados:

1. Nivel Interno o Físico
2. Nivel Conceptual
3. Nivel Externo o Lógico

El Nivel Físico se encarga de engranar con el software más interno de cada máquina (Sistema Operativo y Sistema de Gestión de ficheros). El Esquema Interno especifica qué y cómo son almacenados los datos. Describe la estructura de la B.D. en forma de Modelo Conceptual de almacenamiento.

El Nivel Conceptual materializa el lugar donde definir el resultado del diseño de la B.D.. El Esquema Conceptual debe captar y almacenar el “universo del discurso” que describe a la organización o empresa y que es necesaria para su funcionamiento (capacidad semántica) sirve de punto de control para futuros desarrollos de la B.D., aísla la representación de la información de los requerimientos de la máquina y de las exigencias de cada usuario en particular e independiza la definición de la información de los SGBD en concreto.

El Nivel Lógico o externo de descripción, contiene las vistas externas de la B.D. que están asociadas cada una a un Esquema Externo y permite ver a cada tipo de usuario de la B.D. sólo aquella parte del esquema que es de su interés.

De una B.D. se pueden derivar tantas vistas como haga falta. El propósito principal de esta arquitectura a 3 niveles es conseguir que el Esquema Conceptual sea una descripción estable de la organización e independiente de las “vistas” y de la forma de almacenamiento de los datos.

Debido a la independencia de niveles, las B.D. son flexibles y adaptables a los cambios.

Modelos de referencia de ANSI

En 1985 el DAFTG (Database Architecture Frame-work Task Group) del ANSI / X3 / SPARC Database System Study Group presenta el estudio final “Modelo de Referencia para la estandarización de los SGBD”.

El Modelo de Referencia MR no es por si mismo un estándar sino un marco conceptual cuyo objetivo principal es simplificar el trabajo de estandarización de los distintos componentes y sus relaciones de un SGBD.

Objetivos y beneficios del Modelo de Referencia. Requisitos de su definición

Los objetivos que el Modelo de Referencia quiere conseguir son:

- Servir de herramienta para el desarrollo y coordinación de estándar en el área de SGBD.

- Describir las interacciones entre el SGBD y otros componentes lógicos del Sistema de Información, tales como el Sistema de Diccionario de Datos.
- Facilitar la formación del personal dando un marco común para la descripción de los SGBD.
- Permitir la clasificación de los productos comerciales.
- Ayudar a los usuarios a analizar, cambiar e introducir SGBD en su organización.

Los beneficios que puede traer la estandarización de los SGBD también son múltiples:

- Portabilidad de las aplicaciones
- Mejora de la productividad y disminución de gastos de formación
- Simplificación de los procesos de evaluación y selección de los SGBD
- Reducción de costes
- Posibilidad de intercambio de datos entre distintos SGBD

Para alcanzar los objetivos del MR se deben cumplir los siguientes requisitos:

- La definición del Modelo debe adaptarse a los nuevos desarrollos tecnológicos (Base de Datos Distribuidos, máquinas de BD, etc.)
- La simplificación de la arquitectura ANSI/SPARC (ANSI77)
- La Unificación de los modelos de Datos
- La compatibilidad con otros Modelos

Modelo de Referencia

El Modelo de Referencia representa el marco común para la descripción de los SGBD, facilitando su estudio e investigación. Simplifica su análisis, evaluación y selección. Conduce a la estandarización y a la compatibilidad de los distintos elementos constituyentes de los SGBD. Aumenta la portabilidad de las aplicaciones y como consecuencia de esto, la conectividad total imprescindible en la implantación de Sistemas Abiertos, hoy tan invocados y necesarios en el sector.

El MR está basado en la Arquitectura de tres niveles de ANSI (1975), aunque la revisa, la simplifica y describe las interacciones del SGBD.

Los datos pasan por los distintos niveles de la arquitectura, separados y aislados por interfaces consiguiendo de esta forma la independencia entre niveles.

Niveles de Descripción de Datos

Estos 3 tipos de esquemas conceptual, externo e interno suponen tres tipos de funciones de administración:

1. Administración de la empresa: diseña Esquema Conceptual

2. Administración de la BD: especifica el Esquema Interno más eficiente para el esquema conceptual ya almacenado en los Metadatos.
3. Administración de Aplicaciones: construye Esquemas Externos a partir del conceptual y controla los distintos programas de aplicación que utiliza la BD.

En esta estructura lógica de tres tipos de esquemas se diferencian 2 partes:

- La parte superior para la Definición de la BD
- La parte inferior para la Manipulación de la BD

Aspectos relativos a la Definición de los Datos

La definición se realiza por funciones de programa y sus correspondientes interfaces, que dan como resultado los Metadatos (datos sobre los datos). Estos se almacenan en el Diccionario de Datos que constituye el núcleo de la Arquitectura.

Creación de los tres tipos de esquemas:

1. Se especifica el esquema conceptual mediante un lenguaje de Definición.
2. El procesador del esquema conceptual compila el esquema conceptual.
3. Se almacena el esquema conceptual en el diccionario.
4. El procesador del esquema conceptual facilita información sobre este esquema, necesaria para la definición de los esquemas externo e interno.
5. Los procesadores de los esquemas externo e interno controlan éstos y los almacenan en la BD.

Aspectos relativos a la Manipulación de los Datos

Una vez definidas las Bases de Datos, el usuario mediante el lenguaje de manipulación de Datos puede realizar cualquier función de manipulación, insertar, borrar, modificar.

Una petición del usuario es ejecutada por los transformadores externo\conceptual, conceptual\interno e interno\almacenamiento, que utilizan los metadatos almacenados en el Diccionario. (Ver figura 2)

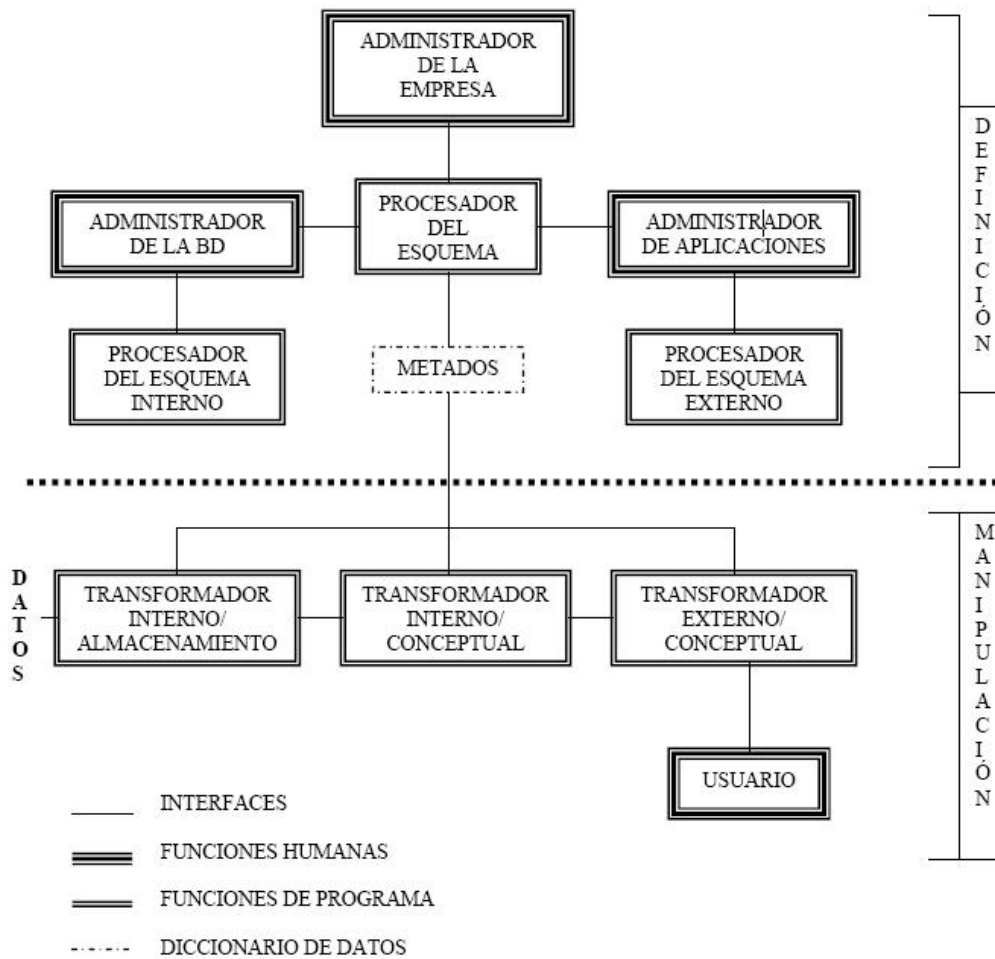


Figura 2. Esquema simplificado arquitectura ANSI/SPARC

Entorno y componentes de los SGBD

La descripción de un SGBD requiere la especificación de sus interfaces y componentes o subsistemas que interaccionan con él. Una interfaz es un lenguaje aceptado por dos o más procesos haciendo posible su comunicación.

En el entorno de un SGBD encontramos:

- Programas de aplicación y procesadores de lenguaje de aplicación
- Herramientas de Gestión
- Sistema de Diccionario de Datos
- Sistemas Operativo y de Gestión de ficheros
- Protocolos y Sistemas Distribuidos

El MR describe las interacciones del SGBD pero no indica nada acerca de su instrumentación.

Los componentes del SGBD son:

1. Sistema de Control de Transformación de Datos (SCTD). Es el núcleo del sistema. Mediante sus operadores es posible definir y manipular los datos. Está basado en un modelo de Datos capaz de soportar la descripción y manipulación de otros modelos de datos.
2. Interfaz LD (lenguaje de Datos) es el lenguaje de manipulación de datos del Modelo de Datos del SCTD.
3. Herramientas de Gestión de Datos. HGD Componentes lógicas de alto nivel que se comunican o interaccionan con el SCTD a través de la interfaz LD que tanto los usuarios como los procesadores utilizan para realizar sus demandas al núcleo del sistema (SCTD). Están orientadas hacia aplicaciones o funciones más específicas que las del LD. Entre ellas se encuentran lenguajes de 4 Generación, Soporte para la ayuda a la Decisión, Diccionario de Recursos de información...
4. Interfaz LD-I(lenguaje de Datos interno) que permite el uso de procesadores que soportan el funcionamiento de los SGBD, especialmente los del Sistema Operativo.
5. Interfaces LDU (lenguaje de Datos de Usuario) y LDU-I (lenguaje de Datos de Usuario interno). Estas interfaces son incluidas en el MR en 1988 por otro subgrupo del DBSSG, el UFTG (User facility Task Group) que propuso un Modelo de Referencia para Facilidades de Usuario, como prolongación del MR que estábamos describiendo.

Entre el SGBD y el usuario se interponen unas componentes denominadas Facilidades de Usuario que mediante las interfaces LDU y LDU-I transforman una demanda del usuario en una petición funcional a las HGD y transforman la salida de éstas en un formato de presentación al usuario. (Ver figuras 3 y 4).

LDU	Lenguaje de Datos de Usuario
LDU-I	Lenguaje Interno de Datos de Usuario
FU	Facilidades de Usuario
SCTD	Sistema de Control de Transformación de Datos
LD	Lenguaje de Datos
LD-I	Lenguaje Interno de Datos
MGD	Herramientas de Gestión de Datos
SCTD	Sistema de Control de Transformación de Datos

Tabla 1. Acrónimos utilizados

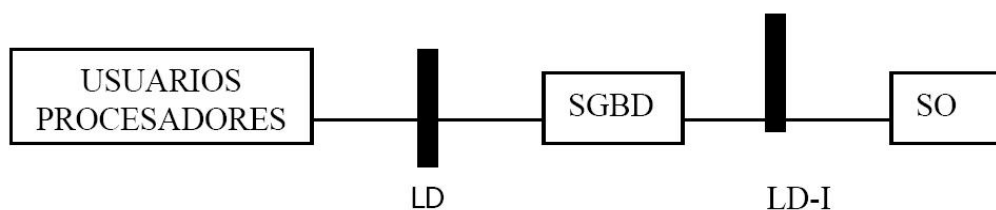


Figura 3. Interfaces entre el SGBD y su entorno

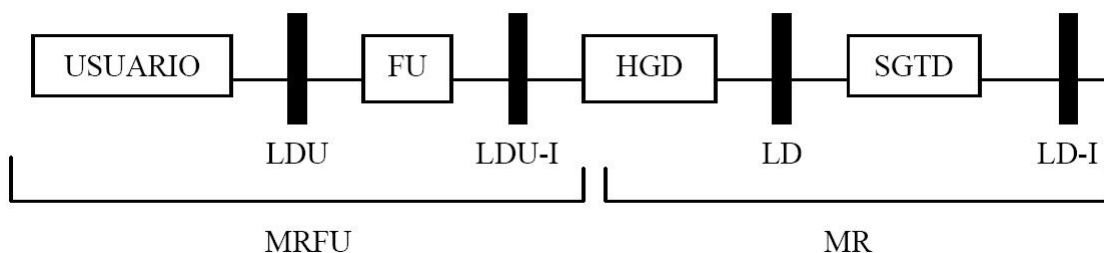


Figura 4. Modelo de Referencia de Facilidades de Usuario

Modelo de Datos

Según el MR de ANSI el SCTD está basado en un Modelo de Datos (MD).

“Conjunto de conceptos, reglas y convenciones que nos permiten describir los datos de una parcela del mundo real (Universo del discurso)”.

“Crear un modelo consiste en definir un mundo abstracto y teórico tal que las conclusiones que se pueden deducir de él coincidan con las manifestaciones aparentes del mundo real” (Flory, 1982).

En el proceso de abstracción que es el diseño de una base de datos se trata de crear un modelo del mundo real para lo cual nos apoyamos en un MD, considerado así como una herramienta intelectual que facilita la interpretación de nuestro universo de discurso y su representación en forma de datos en nuestro Sistema de Información.

El MD debe recoger las propiedades del Universo del Discurso que son de dos tipos: Estáticas y Dinámicas.

- Estáticas. No varían con el tiempo. Se corresponden con la Componente Estática del MD, que es el conjunto de reglas que nos permiten generar la estructura y se definen mediante el Lenguaje de Definición de Datos:
 - Conjunto de objetos (entidades con sus atributos)
 - Conjunto de asociaciones entre objetos (interrelaciones)
 - Conjunto de restricciones (pueden ser inherentes u opcionales)
- Dinámicas. Varían con el tiempo. Se corresponden con la Componente Dinámica del MD que es el conjunto de operadores que se pueden aplicar sobre la estructura y se definen mediante el Lenguaje de Manipulación de Datos.

Existen diversos tipos de MD pero falta unicidad de criterios en relación con su caracterización. Asimismo, no todos los autores hacen distinción entre Modelos Conceptuales y Modelos Instrumentados en los SGBD (M. Jerárquico, En Red y M. Relacional).

Fue ANSI en 1975 quien introduce la expresión de Modelo Conceptual pero no llega a definirla. Entre los MD convencionales (instrumentados en SGBD) establece tres familias que cronológicamente son:

1. M. Jerárquico. Responde a una estructura arborescente a varios niveles. Cada nivel de la jerarquía está compuesto de uno o varios grupos de datos (nodos), de cada uno de los cuales pueden depender otros nodos quedando unidos por ramas. Los nodos y las ramas determinan una relación del tipo 1:n. La forma de recuperar los datos es recorriendo los distintos niveles según el camino definido por la sucesión de nodos en el árbol. Este modelo asume que ciertos datos son más importantes que otros.
2. M. en Red. Es bastante más flexible que el jerárquico, pues permite establecer múltiples conexiones, combinando varias jerarquías arborescentes. Se obtienen relaciones n:m. Estas relaciones permiten al usuario acceder a un dato sin tener que recorrer todas las jerarquías.
3. M. Relacional. Propuesto teóricamente por Codd en 1970. Es reconocida su superioridad frente a los anteriores.

4.3. Bases de datos relacionales

Las bases de datos relacionales se han ido imponiendo a lo largo de la década de los 80 debido a la universal aceptación de la superioridad que ofrece el modelo de datos en el que se basan, el modelo relacional, frente a sus antecesores: modelo jerárquico (estructuras en árbol) y modelo Codasyl (estructuras en red).

Un sistema de base de datos relacional (BDR) se caracteriza por presentar sus datos externamente como tablas, aunque internamente se sigan manejando éstos de forma convencional, por medio de índices, páginas, etc. Otra de sus características es la disponibilidad de un lenguaje para operar con estas tablas, con funciones, al menos de recuperación y modificación. Adicionalmente debe proporcionar funciones para seleccionar subconjuntos de tablas sin predefinición de caminos de acceso. El lenguaje también incluye en determinados gestores las funciones necesarias para definición de datos, obtención de estadísticas básicas (mínimos, máximos, medias...), ordenación en determinada secuencia, agrupamiento, etc. Igualmente debe proporcionar medios para controlar la integridad, seguridad y consistencia de los datos.

Por último, un sistema relacional debe disponer de interfaces que permitan el acceso concurrente desde terminales interactivos y programas de aplicación, así como herramientas estándar para controlar la operación y facilitar, además, los procesos de respaldo y recuperación. Codd estableció en 1985 doce principios, de los cuales al menos seis deben satisfacerse para que una base de datos (BD) pueda considerarse totalmente relacional. Estos fueron precedidos de una regla general global, llamada "Regla Cero". Estos principios pueden resumirse de la siguiente forma:

- Regla 0: Gestión de una BDR. Un sistema de gestión de base de datos relacional (SGBDR), debe ser capaz de manejar la BD exclusivamente con sus capacidades relacionales.

- Regla 1: Representación de la información. Toda la información en una BDR se representa explícitamente a nivel lógico y de una manera única, por medio de valores en tablas.
- Regla 2: Acceso garantizado. Todos y cada uno de los datos elementales en una BDR, deben ser accesibles lógicamente mediante el recurso a una combinación de: nombre de tabla, valor de clave primaria y nombre de columna.
- Regla 3: Representación sistemática de la información que falta. Los valores nulos deben ser soportados por un sistema de gestión de BD (SGBD) completamente relacional para representar, de modo sistemático, la información desconocida o inaplicable.
- Regla 4: Catálogo dinámico. La descripción de la BD se representa, a nivel lógico, en la misma forma que los datos ordinarios, de modo que los usuarios autorizados puedan aplicar el mismo lenguaje relacional para consultarlo.
- Regla 5: Sublenguaje global de datos. Debe existir, al menos, un lenguaje cuyas sentencias sean expresadas por medio de una sintaxis bien definida, como cadena de caracteres, y capaz de soportar definición de datos, definición de vistas, manipulación de datos, restricciones de integridad, autorizaciones y manejo de transacciones.
- Regla 6: Actualización de vistas. Todas las vistas teóricamente actualizables deberán ser también actualizables por el sistema.
- Regla 7: Inserciones, actualizaciones y eliminaciones de alto nivel. La capacidad para manejar, como un solo operando, la relación base o una relación derivada se aplica no sólo a las recuperaciones de datos, sino también, a sus inserciones, actualizaciones y eliminaciones.
- Regla 8: Independencia física de los datos. Los programas de aplicaciones y las actividades terminales permanecerán lógicamente inalterados siempre que se realicen cambios en las representaciones de almacenamiento o en los métodos de acceso.
- Regla 9: Independencia lógica de los datos. Cuando se efectúe en las tablas cualquier tipo de cambio que preserve la información, los programas de aplicación permanecerán intactos.
- Regla 10: Independencia de la integridad. Las reglas de integridad de una BD particular deben ser definibles por medio del sublenguaje de datos relacional y almacenadas en el catálogo, no en los programas de aplicaciones.
- Regla 11: Independencia de la distribución. Un sistema relacional debe tener un sublenguaje de datos que pueda soportar bases de datos distribuidas (BDD) sin alterar los programas de aplicaciones o actividades finales.
- Regla 12: Regla de la no inversión. Si un sistema relacional tiene un lenguaje de bajo nivel, éste no puede ser utilizado para pasar por alto las reglas de integridad y las restricciones expresadas por medio del lenguaje relacional de más alto nivel.

El modelo relacional es el modelo lógico en el que se basan la mayoría de los SGBD comerciales en uso hoy en día. En primer lugar, se trata la descripción de los principios básicos del modelo relacional: la estructura de datos relacional y las reglas de integridad.

A continuación, se presenta un tratamiento detallado del álgebra relacional, que es un conjunto de operaciones para manipular la estructura de datos relacional y especificar consultas de datos. El álgebra relacional es un lenguaje procedural, mientras que el cálculo relacional es un lenguaje equivalente no procedural.

El modelo relacional se basa en dos ramas de las matemáticas: la teoría de conjuntos y la lógica de predicados de primer orden. El hecho de que el modelo relacional esté basado en la teoría de las matemáticas es lo que lo hace tan seguro y robusto. Al mismo tiempo, estas ramas de las matemáticas proporcionan los elementos básicos necesarios para crear una base de datos relacional con una buena estructura, y proporcionan las líneas que se utilizan para formular buenas metodologías de diseño.

Hay quien ofrece una cierta resistencia a estudiar complicados conceptos matemáticos para tan sólo llevar a cabo una tarea bastante concreta. Es habitual escuchar quejas sobre que las teorías matemáticas en las que se basa el modelo relacional y sus metodologías de diseño, no tienen relevancia en el mundo real o que no son prácticas. No es cierto: las matemáticas son básicas en el modelo relacional. Pero, por fortuna, no hay que aprender teoría de conjuntos o lógica de predicados de primer orden para utilizar el modelo relacional. Sería como decir que hay que saber todos los detalles de la aerodinámica para poder conducir un automóvil. Las teorías de la aerodinámica ayudan a entender cómo un automóvil puede ahorrar combustible, pero desde luego no son necesarias para manejarlo.

La teoría matemática proporciona la base para el modelo relacional y, por lo tanto, hace que el modelo sea predecible, fiable y seguro. La teoría describe los elementos básicos que se utilizan para crear una base de datos relacional y proporciona las líneas a seguir para construirla. El organizar estos elementos para conseguir el resultado deseado es lo que se denomina diseño.

4.4. Bases de Datos Orientadas a Objetos

Hasta la aparición de las BD Orientadas a Objetos (BDOO), las BD tradicionales no estaban diseñadas para almacenar objetos, con lo que al guardar los datos de un programa OO se incrementaba significativamente la complejidad del programa, dando lugar a más código y más esfuerzos de programación.

Las BDOO están diseñadas para simplificar la POO³. Almacenan los objetos directamente en la BD, y emplean las mismas estructuras y relaciones que los lenguajes de POO. Las BDOO surgen de la combinación de las BD tradicionales y la POO.

Un SGBDOO⁴ es un Sistema de Objetos y un SGBD. Se puede decir que un SGBDOO es un SGBD que almacena objetos incorporando así todas las ventajas de la OO. Para los usuarios tradicionales de BD, esto quiere decir que pueden tratar directamente con objetos, no teniendo que hacer la traducción a tablas o registros. Para los programadores de aplicaciones, esto quiere decir que sus objetos se conservan, pueden ser gestionados aunque su tamaño sea muy grande, pueden ser compartidos entre múltiples usuarios, y se mantienen tanto su integridad como sus relaciones.

Una clase después de programada es transportada a la BD tal como es, al contrario de lo que sucede en los SGBD relacionales donde el modelo de datos se distribuye en tablas.

Sin las BDOO, hay que hacer un complicado proceso de traducción de los objetos a registros o tablas de BD tradicionales. El problema de la traducción a tablas implica:

- Mayor tiempo de desarrollo. El tiempo empleado en generar el código para la traducción de objetos a tablas y viceversa.
- Errores debidos precisamente a esa traducción.
- Inconsistencias debidas a que el ensamblaje / desensamblaje puede realizarse de forma diferente en las distintas aplicaciones.
- Mayor tiempo de ejecución empleado para el ensamblaje / desensamblaje.

La utilización de una BDOO simplifica la conceptualización ya que la utilización de objetos permite representar de una manera más natural la información que se quiere guardar.

Mejora el flujo de comunicación entre los usuarios, los diseñadores y los analistas.

Las BDOO permiten implementar los tres componentes de un modelo de datos:

1. Propiedades estáticas (objetos, atributos y relaciones)
2. Reglas de integridad de los objetos y operaciones
3. Propiedades dinámicas

Antes de los SGBDOO, los dos primeros componentes de un modelo de datos eran implementados en la BD dejando las propiedades dinámicas a cargo de la aplicación. Los SGBDOO implementan las tres características en la BD permitiendo, por ejemplo, aplicar reglas uniformes a objetos sin necesidad de desarrollar otra aplicación nueva cuando surge un cambio en el modo de trabajar con los datos.

³ POO, Programación Orientada a Objetos

⁴ SGBDOO, Sistema Gestor de Base de Datos Orientada a Objetos

4.4.1. Características Básicas de un SGBDOO

Un SGBDOO debe satisfacer dos criterios:

- Ser un SGBD lo que se traduce en 5 características principales: Persistencia, Concurrencia, Recuperación ante fallos, Gestión del almacenamiento secundario y facilidad de Consultas.
- Ser un Sistema OO6: todo sistema OO debe cumplir algunas características como son: Encapsulación, Identidad, Herencia y Polimorfismo. Además existen otras características deseables como Control de tipos y Persistencia.

4.4.2. Manifiesto de las BDOO

El conocido “Manifiesto de los sistemas de BDOO”, que es resultado de la colaboración de un cierto número de expertos, ha tenido mucha influencia al definir los objetivos del movimiento de las BDOO. Los aspectos principales de dicho manifiesto son los siguientes:

1. Objetos complejos: deben permitir construir objetos complejos aplicando constructores sobre objetos básicos.
2. Identidad de los objetos: todos los objetos deben tener un identificador que sea independiente de los valores de sus atributos.
3. Encapsulación: los programadores sólo tendrán acceso a la interfaz de los métodos, de modo que sus datos e implementación estén ocultos.
4. Tipos o clases: el esquema de una BDOO incluye un conjunto de clases o un conjunto de tipos.
5. Herencia: un subtipo o una subclase heredará los atributos y métodos de su supertipo o superclase, respectivamente.
6. Ligadura dinámica: los métodos deben poder aplicarse a diferentes tipos (sobrecarga). La implementación de un método dependerá del tipo de objeto al que se aplique. Para proporcionar esta funcionalidad, el sistema deberá asociar los métodos en tiempo de ejecución.
7. Su LMD debe ser completo
8. El conjunto de tipos de datos debe ser extensible: Además, no habrá distinción en el uso de tipos definidos por el sistema y tipos definidos por el usuario.
9. Persistencia de datos: los datos deben mantenerse después de que la aplicación que los creó haya finalizado. El usuario no tiene que hacer ningún movimiento o copia de datos explícita para ello.
10. Debe ser capaz de manejar grandes BD: debe disponer de mecanismos transparentes al usuario, que proporcionen independencia entre los niveles lógico y físico del sistema.

11. Concurrencia: debe poseer un mecanismo de control de concurrencia similar al de los sistemas convencionales.
12. Recuperación: debe poseer un mecanismo de recuperación ante fallos similar al de los sistemas convencionales
13. Método de consulta sencillo: debe poseer un sistema de consulta *ad-hoc* de alto nivel, eficiente e independiente de la aplicación.

4.5. Base de datos espaciales

Base de datos espacial es un sistema gestor de bases de datos que maneja datos existentes en un espacio o datos espaciales. El espacio establece un marco de referencia para definir la localización y relación entre objetos. El que normalmente se utiliza es el espacio físico que es un dominio manipulable, perceptible y que sirve de referencia. La construcción de una base de datos geográfica implica un proceso de abstracción para pasar de la complejidad del mundo real a una representación simplificada que pueda ser procesada por el lenguaje de las computadoras actuales. Este proceso de abstracción tiene diversos niveles y normalmente comienza con la concepción de la estructura de la base de datos, generalmente en capas; en esta fase, y dependiendo de la utilidad que se vaya a dar a la información a compilar, se seleccionan las capas temáticas a incluir.

La estructuración de la información espacial procedente del mundo real en capas conlleva cierto nivel de dificultad. En primer lugar, la necesidad de abstracción que requieren los computadores implica trabajar con primitivas básicas de dibujo, de tal forma que toda la complejidad de la realidad ha de ser reducida a puntos, líneas o polígonos. En segundo lugar, existen relaciones espaciales entre los objetos geográficos que el sistema no puede obviar; la topología, que en realidad es el método matemático-lógico usado para definir las relaciones espaciales entre los objetos geográficos puede llegar a ser muy compleja, ya que son muchos los elementos que interactúan sobre cada aspecto de la realidad.

4.5.1. Datos espaciales

Un modelo de datos geográfico es una abstracción del mundo real que emplea un conjunto de objetos dato, para soportar el despliegue de mapas, consultas, edición y análisis. Los datos geográficos, presentan la información en representaciones subjetivas a través de mapas y símbolos, que representan la geografía como formas geométricas, redes, superficies, ubicaciones e imágenes, a los cuales se les asignan sus respectivos atributos que los definen y describen.

Un dato espacial es una variable asociada a una localización del espacio. Normalmente se utilizan datos vectoriales, los cuales pueden ser expresados mediante tres tipos de objetos espaciales: puntos, líneas y polígonos

4.5.2. Álgebra

El álgebra utilizada se denomina álgebra ROSE (RObust Spatial Extension) que está basada en los tipos de datos espaciales reales (STD - spatial data types), pero en este caso, los objetos no están definidos en el espacio Euclidiano continuo sino que en términos de la malla que discretiza el espacio, esto debido a que los cálculos computacionales son discretos. Los operadores espaciales se definen de la siguiente forma.

4.5.3. Métodos de acceso espacial

Para evitar la revisión exhaustiva de los datos en una base de datos, se crean índices que reducen el número de elementos a visitar en la base de datos en un procesamiento de consulta. La clásica indexación por B-tree no es aplicable en el caso espacial donde no existe un orden único de los valores de claves. Es por este motivo que existen tres categorías de métodos de acceso espacial; PAM (Point Access Method), R-Tree y SAM (Spatial Access Method) El método a emplear dependerá del tipo de dato en el que esté la base de datos espacial ya sea raster o vectorial. Aunque se han creado los benchmarks que comparan diferentes métodos, los resultados no son concluyentes, pero se recomienda utilizar cualquiera de ellos.

Un índice R-tree aproxima cada geometría en un único rectángulo que la acota minimizando los espacios llamado MBR (Minimal Bounding Rectangle) y organiza una colección de objetos espaciales en una jerárquica donde las hojas contienen punteros a los datos y los nodos intermedios contienen el rectángulo mínimo que contiene a sus sub-hojas. Todas las hojas aparecen al mismo nivel. Cada entrada a una hoja es una tupla (R,O), donde R es el MBR y O es el objeto. Cada nodo intermedio es un tupla (R,P), donde R es el MBR que contiene los rectángulos hijos apuntados por P.

4.5.4. Lenguajes de consulta espacial

Las bases de datos espaciales no tienen un conjunto de operadores que sirvan como elementos básicos para la evaluación de consultas ya que estas manejan un volumen extremadamente grande de objetos complejos no ordenados en una dimensión. Es por esto que existen algoritmos complejos para evaluar predicados espaciales. Las consultas son realizadas generalmente en SSQL (Spatial SQL), que introduce, mediante extensiones, los distintos conceptos del álgebra

ROSE dentro del lenguaje SQL estándar; es decir, utiliza las cláusulas SELECT-FROM-WHERE para las tres operaciones en el álgebra relacional (proyección algebraica, producto cartesiano y selección).

Las tres categorías fundamentales de consultas en un sistema de información espacial son:

- Consultas exclusivamente de propiedades espaciales. Ejemplo: "Traer todos los pueblos que son cruzados por un río".
- Consultas sobre propiedades no espaciales. Ejemplo: "Cuantas personas viven en Valdivia".
- Consultas que combinan propiedades espaciales con no espaciales. Ej: "Traer todos los vecinos de un cuadro localizada en Los Angeles"

En el lenguaje SSQL, el ejemplo del segundo punto se escribiría de la siguiente forma.

```
SELECT poblacion FROM ciudades WHERE nombre= "Valdivia"
```

El otro tipo de consultas, para los datos obtenidos mediante rasterización, es llamado PSQL (Pictorial SQL) donde cada objeto espacial se extiende mediante un atributo *loc* (localización) el cual es referenciado en la cláusula SELECT para una salida gráfica y una cláusula específica para tratar relaciones espaciales. También se destaca en los lenguajes de modelado de la información espacial a GML que es una estructura para almacenar y compartir datos geográficos. Es una codificación del modelo geométrico de rasgo simple del OGC⁵ usando XML.

Un rasgo geográfico (geographic feature) es definido por el OGC como "una abstracción del fenómeno del mundo real, si éste está asociado con una posición relativa a la Tierra". Por tanto, es posible hacer una representación del mundo real con un conjunto de rasgos. La especificación de un rasgo viene dada por sus propiedades, las que pueden pensarse definidas como un triple (nombre, tipo, valor). Si este rasgo es geográfico entonces la propiedad tendrá un valor geométrico. Por tanto, un rasgo simple del OGC es aquel cuya propiedad geométrica está restringida a una geometría simple en la que sus coordenadas estén definidas en dos dimensiones y en el caso de existir una curva, ésta es sujeta a una interpolación lineal.

4.5.5. Bases de datos con extensiones para bases de datos espaciales

Son sistemas de bases de datos normales a los cuales se les agrega una capa para el manejo de la geometría y hacer el "traspaso" desde datos comunes a datos espaciales transparente al usuario.

⁵ OGC, Open Geospatial Consortium

Bibliografía

- [1] Codd, E.F. (1970). "A Relational Model of Data for Large Shared Data Banks".
- [2] Association for Computing Machinery SIGIR Forum archive Volume 7 , Issue 4
- [3] The origins of the data base concept, early DBMS systems including IDS and IMS, the Data Base Task Group, and the hierarchical, network and relational data models are discussed in Thomas Haigh, "'A Veritable Bucket of Facts:' Origins of the Data Base Management System," ACM SIGMOD Record 35:2 (June 2006).
- [4] Kim, Won. *Introduction to Object-Oriented Databases*. The MIT Press, 1990. ISBN 0-262-11124-1
- [5] Bancilhon, Francois; Delobel, Claude; and Kanellakis, Paris. *Building an Object-Oriented Database System: The Story of O₂*. Morgan Kaufmann Publishers, 1992. ISBN 1-55860-169-4.
- [6] Barry, Douglas and Duhl, Joshua. *Object Storage Fact Books: Object DBMSs and Object-Relational Mapping*. Barry & Associates, Inc., 2001.
- [7] R.H. Güting, M. Schneider, Realm-Based Spatial Data Types: The ROSE Algebra.
- [8] Maz J. Egenhofer, Spatial SQL: A Query and Presentation Language.

Tema 5. Normas y estándares para la interoperabilidad entre gestores de bases de datos relacionales. Lenguaje SQL. Componentes de SQL: LDD, LMD.

5.1. El modelo relacional

En 1970, el modo en que se veían las bases de datos cambió por completo cuando E. F. Codd introdujo el modelo relacional. En aquel momento, el enfoque existente para la estructura de las bases de datos utilizaba punteros físicos (direcciones de disco) para relacionar registros de distintos ficheros. Si, por ejemplo, se quería relacionar un registro *A* con un registro *B*, se debía añadir al registro *A* un campo conteniendo la dirección en disco del registro *B*. Este campo añadido, un puntero físico, siempre señalaría desde el registro *A* al registro *B*. Codd demostró que estas bases de datos limitaban en gran medida los tipos de operaciones que los usuarios podían realizar sobre los datos. Además, estas bases de datos eran muy vulnerables a cambios en el entorno físico. Si se añadían los controladores de un nuevo disco al sistema y los datos se movían de una localización física a otra, se requería una conversión de los ficheros de datos. Estos sistemas se basaban en el modelo de red y el modelo jerárquico, los dos modelos lógicos que constituyeron la primera generación de los SGBD.

El modelo relacional representa la segunda generación de los SGBD. En él, todos los datos están estructurados a nivel lógico como tablas formadas por filas y columnas, aunque a nivel físico pueden tener una estructura completamente distinta. Un punto fuerte del modelo relacional es la sencillez de su estructura lógica, pero detrás de esa simple estructura hay un fundamento teórico importante del que carecen los SGBD de la primera generación, lo que constituye otro punto a su favor.

Dada la popularidad del modelo relacional, muchos sistemas de la primera generación se han modificado para proporcionar una interfaz de usuario relacional, con independencia del modelo lógico que soportan (de red o jerárquico). Por ejemplo, el sistema de red IDMS ha evolucionado a IDMS/R e IDMS/SQL, ofreciendo una visión relacional de los datos.

En los últimos años, se han propuesto algunas extensiones al modelo relacional para capturar mejor el significado de los datos, para disponer de los conceptos de la orientación a objetos y para disponer de capacidad deductiva.

El modelo relacional, como todo modelo de datos, tiene que ver con tres aspectos de los datos:

- Estructura de datos.
- Integridad de datos.
- Manejo de datos.

5.2. Lenguajes de los sistemas de gestión de bases de datos

Los SGBD deben ofrecer lenguajes e interfaces apropiadas para cada tipo de usuario: administradores de la base de datos, diseñadores, programadores de aplicaciones y usuarios finales.

5.2.1. Lenguaje de definición de datos

Una vez finalizado el diseño de una base de datos y escogido un SGBD para su implementación, el primer paso consiste en especificar el esquema conceptual y el esquema interno de la base de datos, y la correspondencia entre ambos. En muchos SGBD no se mantiene una separación estricta de niveles, por lo que el administrador de la base de datos y los diseñadores utilizan el mismo lenguaje para definir ambos esquemas, el lenguaje de definición de datos (LDD). El SGBD posee un compilador de LDD cuya función consiste en procesar las sentencias del lenguaje para identificar las descripciones de los distintos elementos de los esquemas y almacenar la descripción del esquema en el catálogo o diccionario de datos. Se dice que el diccionario contiene metadatos: describe los objetos de la base de datos.

Cuando en un SGBD hay una clara separación entre los niveles conceptual e interno, el LDD sólo sirve para especificar el esquema conceptual. Para especificar el esquema interno se utiliza un lenguaje de definición de almacenamiento (LDA). Las correspondencias entre ambos esquemas se pueden especificar en cualquiera de los dos lenguajes. Para tener una verdadera arquitectura de tres niveles sería necesario disponer de un tercer lenguaje, el lenguaje de definición de vistas (LDV), que se utilizaría para especificar las vistas de los usuarios y su correspondencia con el esquema conceptual.

5.2.2. Lenguaje de manipulación de datos

Una vez creados los esquemas de la base de datos, los usuarios necesitan un lenguaje que les permita manipular los datos de la base de datos: realizar consultas, inserciones, eliminaciones y modificaciones. Este lenguaje es el que se denomina lenguaje de manejo de datos (LMD).

Hay dos tipos de LMD: los procedurales y los no procedurales. Con un LMD procedural el usuario (normalmente será un programador) especifica qué datos se necesitan y cómo hay que obtenerlos. Esto quiere decir que el usuario debe especificar todas las operaciones de acceso a datos llamando a los procedimientos necesarios para obtener la información requerida. Estos lenguajes acceden a un registro, lo procesan y basándose en los resultados obtenidos, acceden a otro registro, que también deben procesar. Así se va accediendo a registros y se van procesando

hasta que se obtienen los datos deseados. Las sentencias de un LMD procedural deben estar embebidas en un lenguaje de alto nivel, ya que se necesitan sus estructuras (bucles, condicionales, etc.) para obtener y procesar cada registro individual. A este lenguaje se le denomina lenguaje anfitrión. Las bases de datos jerárquicas y de red utilizan LMD procedurales. Un LMD no procedural se puede utilizar de manera independiente para especificar operaciones complejas sobre la base de datos de forma concisa. En muchos SGBD se pueden introducir interactivamente instrucciones del LMD desde un terminal o bien embeberlas en un lenguaje de programación de alto nivel. Los LMD no procedurales permiten especificar los datos a obtener en una consulta o los datos que se deben actualizar, mediante una sola y sencilla sentencia. El usuario o programador especifica qué datos quiere obtener sin decir cómo se debe acceder a ellos. El SGBD traduce las sentencias del LMD en uno o varios procedimientos que manipulan los conjuntos de registros necesarios. Esto libera al usuario de tener que conocer cuál es la estructura física de los datos y qué algoritmos se deben utilizar para acceder a ellos. A los LMD no procedurales también se les denomina declarativos. Las bases de datos relacionales utilizan LMD no procedurales, como SQL (Structured Query Language) o QBE (Query-By-Example). Los lenguajes no procedurales son más fáciles de aprender y de usar que los procedurales, y el usuario debe realizar menos trabajo, siendo el SGBD quien hace la mayor parte.

La parte de los LMD no procedurales que realiza la obtención de datos es lo que se denomina un lenguaje de consultas. En general, las órdenes tanto de obtención como de actualización de datos de un LMD no procedural se pueden utilizar interactivamente, por lo que al conjunto completo de sentencias del LMD se le denomina lenguaje de consultas, aunque es técnicamente incorrecto.

5.3. El lenguaje SQL

5.3.1. Elementos básicos de SQL

Entre los lenguajes comerciales de manejo de bases de datos existen diferentes aproximaciones a los lenguaje formales descritos en el modelo relacional teórico presentado por Codd.

EL SQL (Structured Query Language) es un lenguaje estandarizado de peticiones (query) a bases de datos. Es el lenguaje de manipulación de bases de datos relacionales más extendido, habiéndose convertido en un estándar de facto.

Permite realizar consultas utilizando los recursos del álgebra relacional combinados con el cálculo relacional de tuplas. Esto hace que el lenguaje tenga un comportamiento mixto entre declarativo y operacional, aunque con mucha mayor parte declarativa.

Otros lenguajes presentan enfoques diferentes como QBE (Query By Example) que se basa en el cálculo relacional de dominios o el lenguaje Quel (Query Language) basado en el cálculo relacional de tuplas.

La base teórica de SQL es bastante fuerte. Las operaciones funcionan en términos de conjuntos y no de registros individuales. Además no incluye ninguna especificación de localización de los datos o ruta de acceso dejando esta tarea al intérprete del lenguaje.

5.3.2. Descripción de datos con SQL

Elementos de una base de datos SQL

Los datos en bases de datos relacionales se almacenan en tablas. Una tabla se compone de filas y columnas o registros y campos correspondientes a entidades y atributos de una relación. A todas las estructuras lógicas de datos se les llama de una forma genérica objetos de las bases de datos.

Normalmente en un SGBD existen varios usuarios que deberán disponer en primer lugar de acceso al sistema operativo de la máquina. Cada usuario para el SGBD estará identificado por un nombre una palabra clave y unas propiedades. Todo ello es fijado y administrado por el administrador de la base de datos que es un usuario especial con derechos sobre todos los objetos.

Lo normal es que las estructuras de datos que van a ser compartidas sean creadas por el administrador pero esto no es un requisito imprescindible dado que puede haber usuarios para los que sea de interés mantener sus propias bases de datos que compartirán con parte o toda la organización.

El creador o propietario de una tabla puede permitir el acceso a su tabla a otros usuarios del sistema pero puede que le interese también permitir el acceso a una parte de la tabla. Para estos casos se creará una vista (view). En una vista se seleccionan algunos atributos o algunas tuplas de la tabla y se permite que usuarios seleccionados las puedan manejar.

La búsqueda de un registro en una tabla es una búsqueda secuencial. Una forma de agilizar la búsqueda consiste en ordenar los registros según algún criterio y preguntar por los registros según ese orden. Si existe una tabla en la que se anota el valor característico de cada registro y su dirección, la búsqueda se llama indexada y esta tabla auxiliar se llama índice.

En un sistema multiusuario cada usuario tiene un conjunto de objetos que le pertenecen de diferentes tipos como tablas, índices o vistas. A este conjunto se le suele llamar esquema.

Todos los objetos citados hasta ahora deben crearse y configurarse antes de pasar al uso de la base de datos. Estas tareas se realizan con LDD.

Se llama creador de un objeto al usuario que lo crea mediante la correspondiente sentencia CREATE y es el que define su nombre, estructura y propiedades.

Normalmente el creador es también el propietario del objeto y es el que puede modificar su esquema. El nombre del propietario se utiliza para formar el nombre completo del objeto.

Cada objeto de una base de datos tiene un nombre que sirve para localizarlo a lo largo de su vida. Este nombre se llama en SQL identificador.

Estos nombres tienen un ámbito donde son reconocidos. Cada nombre debe ser único en su ámbito. Los objetos que pertenecen a un esquema (tablas, índices, etc.) tienen ese esquema como ámbito máximo. Por ello un nombre de tabla, por ejemplo, puede repetirse en distintos esquemas de usuarios diferentes pero no dentro del mismo esquema.

Además de estas normas básicas las diferentes implementaciones tienen sus propias reglas sobre ámbitos.

Descripción de tablas

Las tablas son creadas por el administrador de la base de datos para su utilización por un grupo de usuarios o por un usuario individual. Para esto se utiliza la sentencia CREATE y se debe especificar el nombre y las características de la tabla.

Las características que se deben especificar son:

- Definición de las columnas. Para definir una columna habrá que decir el nombre y tipo de datos que se van a almacenar en ese campo de la tabla.
- Restricciones de integridad. Se denomina restricción de integridad a una propiedad que debe cumplirse para mantener en una tabla algún criterio de diseño, Entre estas propiedades se cuenta el restringir el rango de posibles valores contenidos en un atributo o el relacionar los valores de dos atributos. En SQL las restricciones de integridad se fijan mediante una sentencia CONSTRAINT.
- Características de almacenamiento
- Datos iniciales

Descripción de vistas

Una vista es un subconjunto de los contenidos de uno o varios objetos de datos que serán tablas o vistas. La vista tiene la misma estructura que una tabla pero no está almacenada en un soporte físico sino que es una tabla lógica.

Las vistas tienen varias utilidades:

- Mostrar a los usuarios los datos que les interesan.
- Proteger los datos.
- Reestructurar datos que pueden estar distribuidos en diferente soporte de manera que aparezcan como una tabla.

- Crear interfaces para aplicaciones que esperan una estructura de tablas distinta a la de creación. Mediante las vistas las aplicaciones se independizan de la estructuración real de los datos.

Descripción de índices

En una base de datos un índice es un medio de acceder a los registros de una forma más rápida que con el simple recorrido secuencial de una tabla. El índice es un objeto de la base de datos que contendrá una entrada para cada valor de las columnas indexadas, con la dirección del registro donde debe buscarse ese valor.

Otro posible uso de los índices es el mantenimiento de una tabla ordenada por distintos criterios y no sólo por su orden natural. Por ejemplo en una tabla de personas ordenadas por apellidos un índice puede servir para reordenarla por edades.

El índice ocupa un espacio adicional al ocupado por los datos. Además un índice debe ser actualizado cada vez que se borra, actualiza o inserta un elemento en una tabla lo que hará que estas operaciones sean más lentas.

5.3.3. Manipulación de datos con SQL

Se llaman manipulaciones a aquéllas operaciones sobre una base de datos que no afectan a la estructura de ésta sino a su contenido. Estas operaciones se realizan con LMD (Data Manipulation Language). Las manipulaciones posibles sobre una base de datos son las siguientes:

- Insertar valores en tuplas (INSERT)
- Eliminar una tupla (DELETE)
- Modificar el valor de un campo en una o varias tuplas (UPDATE)
- Consultar o listar todos o algunos campos de un grupo de tuplas (SELECT)

Todas las operaciones de manipulación se pueden hacer tanto en tablas como en vistas ya que éstas no son sino tablas lógicas.

5.4. Estándares de conectividad

Para la relación de los clientes con los servidores de bases de datos solamente las especificaciones de SQL (de vital importancia en las relaciones Cliente/Servidor) han sobrevivido.

Cuando un desarrollador enfrenta un proyecto enmarcado en Cliente/Servidor de Bases de Datos debe enfrentar una serie de inconvenientes que debe resolver, en particular la conexión y manipulación de la información en la base de datos a través del SGBD o DBMS. El esquema siguiente muestra esta situación:

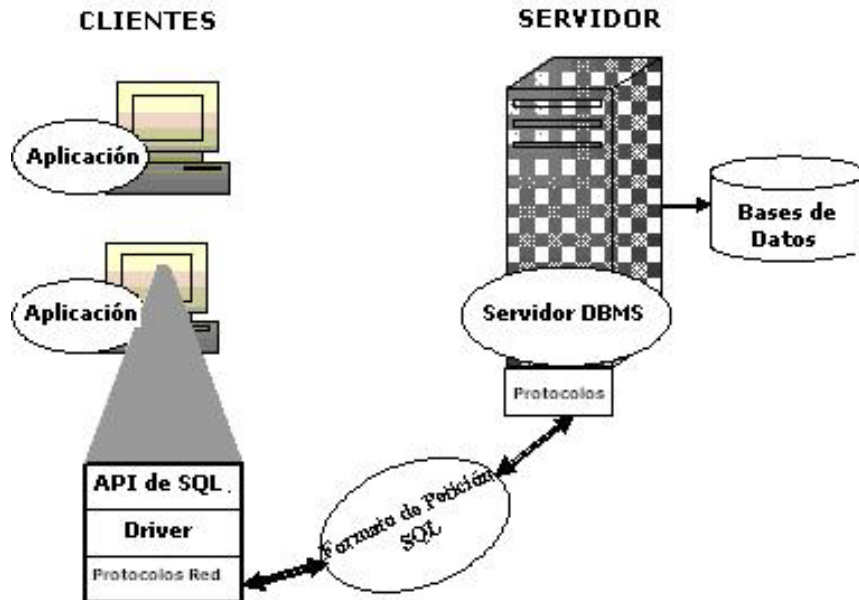


Figura 1.

El programador debe utilizar APIs (Interfaz para el Programador de Aplicaciones). Las APIs son funciones que proveen los fabricantes de hardware y/o software. Cada compañía ha creado sus APIs específicas. Estas funciones actúan sobre los controladores o drivers (similares a los drivers para impresoras, cámaras, etc). Los drivers se comunican con el sistema operativo y con la Base de Datos. Los drivers no son estándar.

En caso de estar en red, se debe tener protocolos comunes de comunicación entre clientes y servidores o traductores de los mismos. En esto los estándares están muy avanzados y son respetados, por lo cual pocas veces hay problema con ello.

En el caso en el cual una aplicación tiene que comunicarse con diferentes SGBD la situación es más compleja.

- Controladores o Drivers. Los controladores reciben la llamada de los interfaces de nivel de llamada (CLI) y la traducen al lenguaje de acceso nativo del servidor de bases de datos. Se requiere un controlador por cada base de datos con la que se realiza conexión.
- APIs de SQL. Dos métodos prevalecen para el soporte de SQL con lenguajes de programación:
 1. SQL embebido ESQL (tratado por la especificación SQL-92. Sirve para incluir instrucciones SQL (realmente no es una API) como si fuesen instrucciones propias del lenguaje de programación ordinario. Deben enmarcarse entre etiquetas. Esto les permite

distinguirse de las demás instrucciones. Estos bloques son precompilados. Tiene como inconvenientes que se debe saber en el momento del desarrollo de la aplicación cuál va a ser la base de datos utilizada, ya que las instrucciones son diferentes para cada producto.

2. Las interfaces del nivel de llamada (CLI) son APIs de SQL para acceso a bases de datos, por lo cual no se requiere precompilación ni conocer de antemano la base de datos. Esto facilita la independencia del producto en particular, pero los estándares no han triunfado.

5.4.1. ODBC

ODBC (Open DataBase Connectivity) es un método estándar de acceso a bases de datos desarrollado por Microsoft Corporation. La finalidad de ODBC es hacer que posible que cualquier aplicación pueda acceder a cualquier dato sin tener en cuenta qué Sistema de Gestión de Base de Datos (SGBD) gestiona los datos. ODBC gestiona esto insertando una capa intermedia llamada un driver de base de datos entre la aplicación y el SGBD.

El propósito de esta capa es traducir las peticiones de datos de las aplicaciones a comandos que el SGBD comprenda. Para que esto funcione tanto la aplicación como el SGBD deben cumplir el estándar ODBC, es decir, la aplicación debe ser capaz de generar comandos ODBC y el SGBD debe ser capaz de responder a ellos. Desde la versión 2.0, el estándar soporta SQL.

SQL/CLI es un interfaz de programación diseñado para soportar acceso SQL a bases de datos desde aplicaciones encapsuladas. CLI fue creado originalmente por el SAG (SQL Access Group). La especificación SAG/CLI fue publicada como la especificación ODBC (Microsoft Open DataBase Connectivity) en 1992. Microsoft era un miembro del SAG y fue la primera compañía en publicar un producto comercial (bajo Microsoft Windows) pero ODBC no es un estándar de Microsoft. En 1993, SAG submitió el CLI a los comités de SQL de ANSI e ISO. ODBC (Open Database Connectivity) es una API de acceso a bases de datos. Está basado en las especificaciones CLI de X/Open e ISO/IEC para APIs de bases de datos y utiliza SQL (uses Structured Query Language) como su lenguaje de acceso a las bases de datos.

SQL/CLI proporciona un interfaz de nivel de llamada (CLI) estándar internacional independiente de la implementación para acceso a bases de datos SQL. Las herramientas cliente-servidor pueden acceder fácilmente a las bases de datos mediante librerías de enlace dinámico (dynamic link libraries). Soporta un conjunto de herramienta clienteservidor.

SQL/CLI es un añadido al estándar SQL de 1992(SQL-92). Fue completado como estándar de ISO ISO/IEC 9075-3:1995 Information technology -- Database languages -- SQL -- Part 3: Call-Level Interface (SQL/CLI). El esfuerzo actual del SQL/CLI se centra en añadir soporte para SQL3.

ODBC tiene interfaces para Visual Basic, Visual C++, SQL y el paquete de drivers ODBC contiene drivers para las bases de datos Access, Paradox, dBase, Text, Excel y Btrieve.

Hay drivers ODBC y herramientas de desarrollo disponibles para Microsoft Windows, Unix, OS/2, y Macintosh.

ODBC está diseñado para obtener la máxima interoperabilidad, es decir, la habilidad de que una misma aplicación pueda acceder a distintos SGBDs con el mismo código fuente. Las aplicaciones de base de datos llaman a funciones en el interfaz ODBC las cuales están implementadas en módulos específicos llamados drivers. El uso de drivers aísla las aplicaciones de llamadas específicas de base de datos. Como los drivers se cargan en tiempo de ejecución un usuario sólo necesita añadir un driver nuevo para acceder a un nuevo SGBD. No es necesario recompilar o reenlazar la aplicación.

Microsoft ofrece un Kit de desarrollo software ODBC (SDK) para ordenadores corriendo Windows NT Server/Windows 2000 Server, Windows NT Workstation/Windows 2000 Professional, y Windows 95/98. Microsoft se ha aliado con Visigenic Software para portar estos SDKs a la plataforma Macintosh y a unas cuantas plataformas UNIX.

- ODBC es un CLI (interfaz de nivel de llamada o call-level interface).
- ODBC define una gramática SQL estándar.
- ODBC proporciona un Driver Manager para gestionar accesos simultáneos a múltiples SGBDs. El Driver Manager administra el acceso a las fuentes de datos.

Arquitectura ODBC:

La arquitectura ODBC tiene 4 capas:

- Aplicación. Realiza el procesamiento y llama a las funciones ODBC para lanzar sentencias SQL y recuperar resultados.
- Driver Manager. Carga y libera los drivers de parte de una aplicación. Procesa las llamadas a las funciones ODBC o se las pasa al driver.
- Driver. Procesa las llamadas a las funciones ODBC, submite peticiones SQL a una fuente específica de datos y devuelve los resultados a la aplicación. Si es necesario, el driver modifica la petición de la aplicación de forma que se ajuste a la sintaxis soportada por el SGBD (DBMS) asociado.
- Data source. Está compuesto de los datos a los que el usuario quiere acceder y su sistema operativo asociado, SGBD y plataforma de red (si la hubiese) utilizada para acceder al SGBD.

La siguiente figura ilustra la relación entre los componentes:

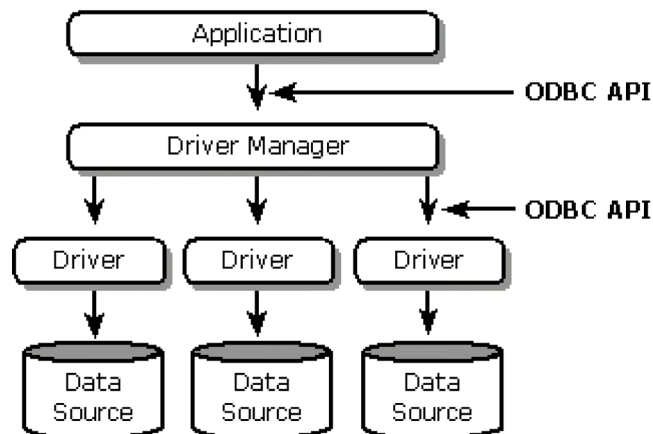


Figura 2. Relación entre componentes ODBC

Pueden existir múltiples drivers y fuentes de datos (data source), lo que permite a la aplicación acceder de forma simultánea a más de una. La API ODBC es utilizada en dos sitios:

- Entre la aplicación y el Driver Manager
- Entre el driver manager y cada driver. Al interfaz entre el Driver Manager y los drivers se le denomina algunas veces SPI (service provider interface o interfaz de proveedor de servicios). Para ODBC, el API (application programming interface) y el SPI son el mismo, es decir, el Driver Manager y cada driver tienen el mismo interfaz a las mismas funciones.

5.4.2. JDBC

JDBC (Java Database Connectivity) es una API de Java que hace que los programas de Java puedan ejecutar sentencias SQL. Esto permite que los programas Java puedan interactuar con cualquier base de datos que cumpla el estándar SQL. Como prácticamente todos los sistemas de gestión de bases de datos relacionales soportan SQL y debido a que Java corre en la mayoría de las plataformas JDBC hace posible escribir una única aplicación de base de datos que puede ejecutarse en diferentes plataformas e interactuar con los distintos SGBDs.

JDBC es similar a ODBC solo que está diseñado específicamente para software escrito en Java mientras que ODBC es independiente del lenguaje. JDBC fue desarrollado por JavaSoft una empresa subsidiaria de Sun Microsystems.

JDBC existe desde 1996, y fue incluido desde la versión 1.1 del JDK. Desde el JDK 1.1 JDBC viene incluido como el paquete java.sql.

JDBC v/s ODBC y otras API's

ODBC no es apropiado para usar directamente con Java por que usa una interfaz en C. Una traducción literal de la API ODBC en C, no es deseable. Por ejemplo Java no tiene punteros y

ODBC hace gran uso de ellos. Una API como JDBC es necesaria para poder desarrollar una solución puramente de Java.

Drivers JDBC

Para usar la API JDBC API con un SGBD en concreto es necesario utilizar un driver basado en tecnología JDBC para mediar entre la tecnología JDBC y el SGBD. Dependiendo de varios factores, un driver puede estar escrito puramente en Java o en una mezcla de Java y métodos nativos de JNI (Java Native Interface).

El último SDK incluye un puente JDBC-ODBC que convierte en disponibles para los programadores de Java a la mayoría de los drivers ODCB. Este driver está aún en fase experimental.

El puente JDBC-ODBC permite a los drivers de ODBC ser usados como drivers de JDBC. Esto, a largo plazo, proporcionará una manera de acceder a alguno de los SGBDs menos conocidas si no se crean drivers específicos JDBC para ellos.

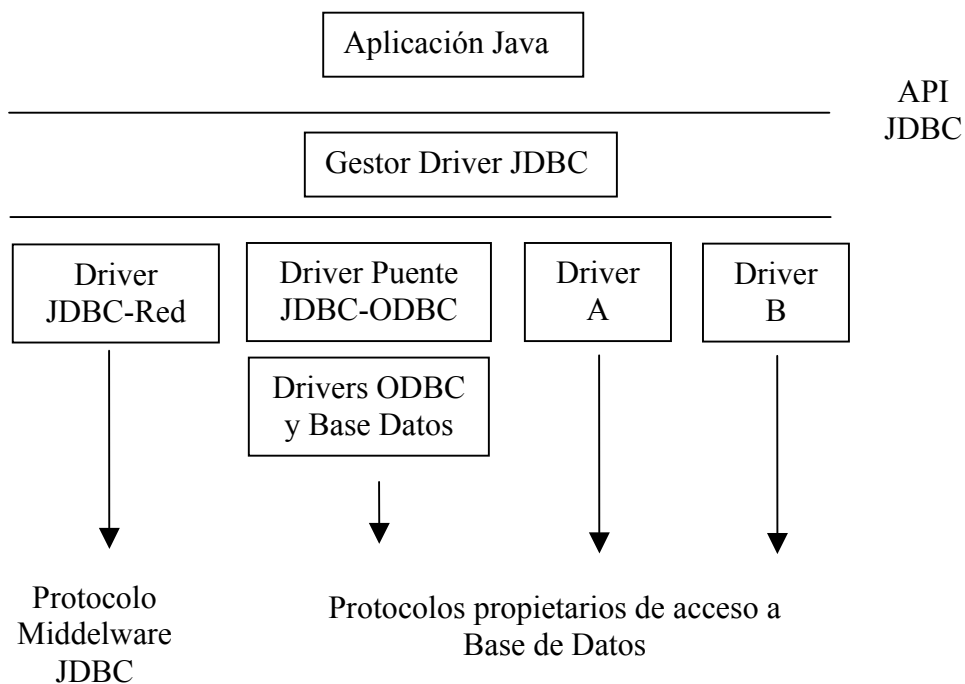


Figura 3. Relación entre componentes JDBC

La API JDBC proporciona acceso universal a datos desde software escrito en Java. Utilizando la API JDBC 3.0 se puede acceder prácticamente a cualquier tipo de datos, desde bases de datos relacionales hasta páginas excel y ficheros planos.

Bibliografía

- [1] Codd, E.F. (June 1970). "A Relational Model of Data for Large Shared Data Banks". *Communications of the ACM* **13** (No. 6): pp. 377-387. Association for Computing Machinery.
- [2] Chamberlin, Donald D.; Boyce, Raymond F. (1974). "SEQUEL: A Structured English Query Language". *Proceedings of the 1974 ACM SIGFIDET Workshop on Data Description, Access and Control*: pp. 249-264. Association for Computing Machinery.
- [3] Opper, Andy (March 1, 2004). *Databases Demystified*. San Francisco, CA: McGraw-Hill Osborne Media, pp. 90-91. ISBN 0-07-225364-9.
- [4] Doll, Shelley (June 19, 2002). Is SQL a Standard Anymore? (HTML). *TechRepublic's Builder.com*. TechRepublic.
- [5] *ISO/IEC 9075-11:2003: Information and Definition Schemas (SQL/Schemata)*, 2003
- [6] ANSI/ISO/IEC International Standard (IS). Database Language SQL—Part 2: Foundation (SQL/Foundation). 1999
- [7] INTO Clause (Transact-SQL) (HTML). *SQL Server 2005 Books Online*. Microsoft (2007).

Tema 6. La descripción e intercambio de datos e información. XML. Estructura de un documento. DTD y XSD. Tecnologías XML. XSL, XLINK.

6.1. La descripción e intercambio de datos e información

El estado actual de la tecnología demanda la existencia de sistemas capaces de trabajar de forma conjunta. Un objetivo como éste exige mecanismos de carácter estándar que permitan el intercambio de información para así explotarla y difundirla de un modo eficiente y, por supuesto, eficaz.

La solución para la descripción e intercambio de datos en el marco tecnológico existente consiste en expresar la información de una manera estructurada. Que la información sea estructurada conlleva que se encuentre compuesta de partes bien definidas, y que esas partes estén formadas a su vez por otras partes, como por ejemplo un tema musical, que se constituye de compases, que son el resultado de la agregación de notas.

La tecnología más adecuada y difundida en respuesta a todas estas necesidades es XML.

6.2. XML

6.2.1. Definición

El Lenguaje Extensible de Marcas, abreviado XML, describe una clase de objetos de datos llamados **documentos XML** y describe parcialmente el comportamiento de los programas de computadora que los procesan. XML es un "perfil de aplicación" o una forma restringida de SGML, el Lenguaje Estándar Generalizado de Marcación [ISO 8879]. Por construcción, los documentos XML son documentos SGML conformados.

Los documentos XML están compuestos por unidades de almacenamiento llamadas *entidades*, que contienen tanto datos analizados como no analizados. Los datos analizados están compuestos de *caracteres*, algunos de los cuales, de la forma datos carácter, y otros de la forma *marca*. Las marcas codifican una descripción de la estructura de almacenamiento del documento y su estructura lógica. XML proporciona un mecanismo para imponer restricciones al almacenamiento y a la estructura lógica.

Se utiliza un módulo software llamado *procesador XML* para leer documentos XML y proporcionar acceso a su contenido y estructura. Se asume que un procesador XML hace su trabajo dentro de otro módulo, llamado *aplicación*. Esta especificación describe el comportamiento requerido de un procesador XML en términos de cómo leer datos XML y la información que debe proporcionar a la aplicación.

6.2.2. Historia y Objetivos

XML fue creado al amparo del Word Wide Web Consortium (W3C) organismo que vela por el desarrollo de WWW partiendo de las amplias especificaciones de SGML.

Su desarrollo se comenzó en 1996 y la primera versión salió a la luz el 10 de febrero de 1998.

La primera definición que apareció fue: *Sistema para definir validar y compartir formatos de documentos en la web*.

Durante el año 1998 XML tuvo un crecimiento exponencial, y con ello me refiero a sus apariciones en medios de comunicación, menciones en páginas web, soporte software, etc.

Los objetivos de diseño de XML son:

- XML debe ser directamente utilizable sobre Internet.
- XML debe soportar una amplia variedad de aplicaciones.
- XML debe ser compatible con SGML.
- Debe ser fácil la escritura de programas que procesen documentos XML.
- El número de características opcionales en XML debe ser absolutamente mínima, idealmente cero.
- Los documentos XML deben ser legibles por humanos y razonablemente claros.
- El diseño de XML debe ser preparado rápidamente.
- El diseño de XML debe ser formal y conciso.
- Los documentos XML deben ser fácilmente creables.
- La concisión en las marcas XML es de mínima importancia.

Esta especificación, junto con los estándares asociados (Unicode e ISO/IEC 10646 para caracteres, Internet RFC 1766 para identificación de lenguajes, ISO 639 para códigos de nombres de lenguajes, e ISO 3166 para códigos de nombres de países), proporciona toda la información necesaria para entender la Versión 1.0 de XML y construir programas de computador que los procesen.

6.3. Estructura de un documento

Un documento XML contiene texto, nunca datos binarios. Un documento XML está compuesto de:

- Elementos (“etiquetas”)
- Texto (cadena de caracteres)
- Referencias a Entidades
- Secciones CDATA
- Comentarios
- Instrucciones de Procesamiento
- La declaración XML

Un objeto de datos es un documento XML si esta bien-formado, tal y como se define en esta especificación. Un documento XML bien-formado puede además ser válido si cumple una serie de restricciones.

Todo documento XML posee una estructura lógica y una física. Físicamente, el documento está compuesto de unidades llamadas entidades. Una entidad puede hacer referencia a otras entidades para que éstas sean incluidas en el documento. Un documento comienza en una "raíz" o entidad documento. Lógicamente, el documento está compuesto de declaraciones, elementos, comentarios, referencias a caracteres e instrucciones de procesamiento, todos los cuales se indican en el documento mediante marcas explícitas. Las estructuras lógica y física deben anidarse de manera correcta.

6.3.1. Documentos XML Bien-Formados

Un objeto de texto es un documento XML bien-formado si:

1. Tomado como un todo, cumple la regla denominada *document*.
2. Respeta todas las restricciones de buena-formación dadas en esta especificación.
3. Cada una de las entidades analizadas que se referencia directa o indirectamente en el documento está bien-formada.

Cumplir la producción *document* implica que:

1. Contiene uno o más *elementos*.
2. Hay exactamente un elemento, llamado **raíz**, o elemento documento, del cual ninguna parte aparece en el contenido de ningún otro elemento. Para el resto de elementos, si la etiqueta de comienzo está en el contenido de algún otro elemento, la etiqueta de fin está

en el contenido del mismo elemento. Es decir, los elementos delimitados por etiquetas de comienzo y final, se anidan adecuadamente mutuamente.

Como consecuencia de esto, para cada elemento no raíz H en el documento, existe otro elemento P tal que H está contenido en P, pero no es el contenido de ningún otro elemento que esté en el contenido de P. P es denominado padre de H, y H es hijo de P.

```
<head>
  <item> Dirección de vuelo</item>
  <item2> Velocidad del avión </item2>
</head>
```

Algunas reglas básicas para que un documento este bien formado:

- Toda etiqueta de comienzo debe tener una etiqueta de fin.
- Las etiquetas no deben solaparse.
- Debe haber una única etiqueta en la raíz del documento (**<correos>**).
- Los valores de los atributos deben ir entre comillas (dobles o simples).
- Una etiqueta no puede tener dos atributos con el mismo nombre.
- Tener en cuenta que:
 - Las mayúsculas y minúsculas no son intercambiables
 - Cada elemento puede contener:
 - Datos de carácter
 - Elementos
 - Ambas cosas a la vez
 - Puede estar vacío
 - Se pueden utilizar comentarios que el analizador no tratará.
 - Los comentarios no pueden ir dentro de las declaraciones, etiquetas y otros comentarios
 - Las secciones CDATA sirven para introducir texto que el analizador tendrá en cuenta como datos de carácter, sin interpretarlo como XML
 - Las entidades predefinidas (&_) permiten incluir ciertos caracteres sin que sean tomadas como XML
 - Podemos definir nuestras propias entidades (DTD)

6.3.2. Caracteres

Una entidad analizada contiene texto, una secuencia de [caracteres](#), que pueden representar marcas o datos carácter. Un carácter es una unidad atómica de texto como se especifica en la norma ISO/IEC 10646 [\[ISO/IEC 10646\]](#). Los caracteres legales son los tabuladores, retornos de carro, finales de línea y los caracteres gráficos legales del estándar Unicode y de la norma ISO/IEC 10646.

Los mecanismos para codificar los códigos de caracteres en patrones de bits pueden variar de una entidad a otra. Todos los procesadores XML deben aceptar las codificaciones UTF-8 y UTF-16 de la norma ISO/IEC 10646.

6.3.3. Construcciones Sintácticas Comunes

A continuación definiremos algunos símbolos utilizados en la gramática.

- El símbolo [S](#) (espacio en blanco) está compuesto de uno o más caracteres espacio, retornos de carro, finales de línea o tabuladores.
- Los caracteres se clasifican, por conveniencia, en letras, dígitos u otros caracteres. Las letras se componen de caracteres de base alfabética o silábica, seguidos de uno o más caracteres combinatorios, o caracteres ideográficos. Las declaraciones completas de cada clase de caracteres se da en "[B. Clases de Caracteres](#)".
- Un Nombre es un "token" que comienza con una letra o uno o más caracteres de puntuación, y que continúa con letras, dígitos, guiones, subrayados, comas y/o puntos. Los Nombres que comienzan por la cadena "xml", o cualquiera que cumpla la regla (('X'|'x') ('M'|'m') ('L'|'l')), quedan reservados para estandarización en esta o futuras versiones de esta especificación.

Nota: El carácter dos puntos en los nombres XML está reservado para la experimentación con espacios de nombres. Se prevé que se estandarizará su significado en un futuro, en el que los documentos que utilicen este carácter con propósitos experimentales deberán ser actualizados. (No existe garantía de que se adopte el mecanismo de espacios de nombres para XML). En la práctica, esto significa que los autores no deberían utilizar los dos puntos en XML, excepto como parte de espacios de nombres experimentales. Sin embargo los procesadores XML deben aceptar este carácter en los Nombres.

- Los datos literales consisten en cualquier cadena entrecomillada que no contenga el tipo de comillas utilizadas como delimitadores en la propia cadena. Los literales se

utilizan para especificar el contenido de entidades internas ("EntityValue"), los valores de los atributos ("AttValue") y los identificadores externos ("SystemLiteral"). Hay que tener en cuenta que el símbolo "SystemLiteral" puede ser analizado sin tomar en cuenta la presencia de marcas.

6.3.4. Datos Carácter y Marcas

El **Texto** consiste en datos carácter y marcas entremezcladas. Las Marcas toman la forma de etiquetas de comienzo, etiquetas de fin, etiquetas de elementos vacíos, referencias a entidades, referencias a caracteres, comentarios, secciones CDATA delimitadores, declaraciones de tipo de documento e instrucciones de procesamiento.

Todo texto que no sea una marca constituye los denominados datos carácter del documento.

El carácter "ampersand" (&) y el paréntesis angular de apertura (<) pueden aparecer de forma literal *sólo* cuando se utilicen como delimitadores de marcas, o dentro de los comentarios, en instrucciones de procesamiento, o en secciones CDATA.

No	<pre><texto> La distancia entre Madrid y Toledo es < que entre Madrid y Málaga </texto></pre>
Si	<pre><texto> <![CDATA[La distancia entre Madrid y Toledo es < que entre Madrid y Málaga]]> </texto></pre>
Si	<pre><texto> La distancia entre Madrid y Toledo es &lt; que entre Madrid y Málaga </texto></pre>

También pueden utilizarse en los valores literales de entidad de una declaración de entidad interna. Si son necesarios en cualquier otro punto, deben ser "escapados" mediante la utilización de referencias numéricas a caracteres, o mediante las cadenas "&" y "<" respectivamente. El paréntesis angular de cierre (>) puede ser representado utilizando la cadena ">", y debe, por compatibilidad, ser "escapado" utilizando ">" o la referencia a carácter cuando aparece como contenido de la cadena "]]>", siempre y cuando la cadena no marque el final de una sección CDATA.

En el contenido de los elementos, los datos carácter son cualquier cadena de caracteres que no contenga el delimitador de comienzo de marca. En las secciones CDATA, los datos carácter son aquellos que no se incluyan en el delimitador de fin de sección CDATA, "]]>".

Para permitir que los valores de los atributos puedan contener tanto las comillas simples como las dobles, el carácter apóstrofe o comilla simple (') puede ser representado como "'", y el comilla doble (") como """.

En resume, en XML se encuentran predefinidas las siguientes referencias a entidades:

- < → Less Than (menor que, <)
- & → Ampersand &
- > → Greater Than (mayor que, >)
- " → Quotation Mark (comilla doble, “)
- ' → Apostrophe (comilla simple, ‘)

6.3.5. Comentarios

Los **Comentarios** pueden aparecer en cualquier punto del documento, fuera del resto de marcas; además, pueden aparecer en la declaración de tipo de documento, en los lugares permitidos por la gramática. No son parte de los datos carácter del documento. Un procesador XML puede, pero no debe, permitir a la aplicación obtener el texto de los comentarios. Por compatibilidad, la cadena "--" (dos guiones) no puede aparecer dentro de un comentario.

Un ejemplo de comentario:

```
<!--declaraciones de <head> y <body>-->
```

6.3.6. Instrucciones de Procesamiento

Las Instrucciones de Procesamiento (PIs) permiten a los documentos contener instrucciones para las aplicaciones.

Las PIs no son parte de los datos carácter del documento, pero deben ser pasadas a la aplicación. Las PIs empiezan con un destino (PITarget) utilizado para identificar a la aplicación a la que está dirigida dicha instrucción. Los nombres de destino "XML", "xml", y las combinaciones de mayúsculas y minúsculas a partir de éstas, están reservados para estandarización en esta o futuras versiones de esta especificación. El mecanismo de Notación XML puede ser utilizado para la declaración formal de destinos de PIs.

6.3.7. Secciones CDATA

Las secciones CDATA pueden aparecer allá donde pueden aparecer los datos carácter. Se utilizan para "escapar" bloques de texto que contengan caracteres que de otra forma se reconocerían como marcas. Las secciones CDATA comienzan por la cadena "<![CDATA[" y terminan con la cadena "]]>":

En una sección CDATA sólo la cadena CDEnd se reconoce como marca, por lo tanto, el paréntesis angular de apertura y el carácter "ampersand" pueden aparecer de forma literal, es decir, no necesitan (y no pueden) ser "escapados" mediante "<" ni "&". No se pueden anidar las secciones CDATA.

Un ejemplo de sección CDATA, en el cual "<saludo>" y "</saludo>" son reconocidos como datos carácter, y no como marcas:

```
<![CDATA[<saludo>Hola,mundo!</saludo>]]>
```

6.3.8. Prologo

Los documentos XML pueden, y deberían, comenzar con una declaración XML que especifica la versión de XML utilizada. Por ejemplo, el siguiente es un documento XML completo, bien-formado pero no válido:

```
<?xml version="1.0"?>  
<saludo>Hola, mundo!</saludo>
```

y también lo es:

```
<saludo>Hola, mundo!</saludo>
```

El número de versión "1.0" debería utilizarse para indicar la conformidad a esta versión de la especificación. Se produce un error si un documento utiliza el valor "1.0" y éste no conforma la especificación de esta versión. Dado que no se han regulado futuras versiones, esta construcción se ha permitido para posibilitar el reconocimiento automáticamente de la versión, si fuese necesario. Los procesadores deben indicar que existe un error si reciben un documento etiquetado con una versión que no soportan.

6.3.9. Instrucciones de Proceso y declaración XML

Las instrucciones de proceso le proporcionan información al parseador XML. Van delimitadas por: “<?” Y “?”>”

Por ejemplo:

```
<?xsl-stylesheet ref="estilo.css"?>
```

(le indica al parseador qué hoja de estilo tiene que utilizar)

La declaración XML es opcional y parece una instrucción de proceso (aunque no lo es)

```
<?xml version="..." encoding="..." standalone="..."?>
```

Si aparece, tiene que ser la primera línea del documento.

La declaración XML tiene 3 atributos:

- **version:** Indica la versión del estándar XML a la cual se ajusta el documento. Actualmente, el único valor válido es “1.0”
- **encoding:** Indica el juego de caracteres del documento. P.ej: “ISO-8859_1” para caracteres de lenguajes europeos (ö, Ç, ñ,...)
- **standalone:** Tiene dos posibles valores.
 - no → Es necesario acceder al DTD del documento para obtener ciertos datos sobre éste (por ejemplo, los valores por defecto de atributos).
 - yes → El documento no depende del DTD.
 - El valor por defecto es no.

6.4. DTD (Document Type Definition) y XSD (XML Schema Definition)

Antes de iniciar la exposición de las diferencias entre las DTDs y los XML Schemas, también referidos como XSD, es importante echar un rápido vistazo a de dónde y cómo surgieron las dos principales soluciones que aquí tratamos.

Inicialmente el uso de las DTDs se debe a SGML (Standard Generalized Markup Language o "Lenguaje de Marcación Generalizado), en el cual describían no sólo el vocabulario necesario para identificar todos los elementos de que iba a constar nuestro documento, si no que también expresaba la estructura que dichos elementos debía respetar. Con la creación de XML, que recordemos que no es más que un subconjunto de SGML, fue necesario mantener la posibilidad de describir los elementos necesarios, al igual que en SGML, por ello se importaron las DTDs y todo su comportamiento.

Posteriormente, y debido al uso principalmente, se vio la necesidad de emplear otros métodos para describir esas necesidades inherentes a XML, con esta idea se creó XML Schema, se pretendía mejorar y ampliar la utilidad de las DTDs, para ello se necesitaban ciertas mejoras que pasaremos a comentar más adelante.

6.4.1. Document Type Definition (DTD)

La **definición de tipo de documento (DTD)** es una descripción de estructura y sintaxis de un documento XML o SGML. Su función básica es la descripción del formato de datos, para usar un formato común y mantener la consistencia entre todos los documentos que utilicen la misma DTD. De esta forma, dichos documentos, pueden ser validados, conocen la estructura de los elementos y la descripción de los datos que trae consigo cada documento, y pueden además compartir la misma descripción y forma de validación dentro de un grupo de trabajo que usa el mismo tipo de información.

Ejemplo de DTD

```
<!ELEMENT Articulo (Cabecera, Contenido, Final)>
<!ELEMENT Cabecera (Titulo, Autor)>
<!ELEMENT Titulo ANY>
<!ELEMENT Autor ANY>
<!ELEMENT Contenido (Extracto, Cuerpo)>
.....
```

La función de las marcas en los documentos XML reside en describir su almacenamiento y estructura lógica, así como asociar pares atributo-valor con dicha estructura lógica. XML proporciona un mecanismo, la **declaración de tipo de documento**, para definir restricciones sobre las estructuras lógicas y para contemplar el uso de unidades de almacenamiento predefinidas. Un documento XML es válido si tiene asociada una declaración de tipo de documento y si el documento cumple las restricciones que allí se expresan.

Los DTD se emplean para determinar la estructura de un documento mediante etiquetas XML o SGML. Un DTD describe:

- Elementos: indican qué etiquetas son permitidas y el contenido de dichas etiquetas.
- Estructura: indica el orden en que van las etiquetas en el documento.
- Anidamiento: indica qué etiquetas van dentro de otras.

Por tanto los DTD no son más que:

- definiciones de los elementos que puede incluir un documento XML,

- de la forma en que deben hacerlo (qué elementos van dentro de otros)
- y los atributos que se les puede dar

Modos de referenciar un DTD en un documento XML:

1. Incluir dentro del documento una referencia al documento DTD en forma de URI (Universal Resource Identifier, o identificador universal de recursos) y mediante la siguiente sintáxis:

```
<!DOCTYPE ficha SYSTEM "http://www.dat.etsit.upm.es/~abarbero/DTD/ficha.dtd">
```

En este caso la palabra SYSTEM indica que el DTD se obtendrá a partir de un elemento externo al documento e indicado por el URI que lo sigue, por supuesto entrecomillado. En este caso, la declaración de tipo documento debe aparecer antes que el primer elemento en el documento.

2. O bien incluir dentro del propio documento el DTD tras la declaración <!DOCTYPE y después del nombre del tipo de documento, en vez de la URI del DTD, el propio DTD entre los símbolos '[' y ']'. Todo lo que hay entre ellos será considerado parte del DTD :

```
<?xml version="1.0"?>
<!DOCTYPE ficha [
<!ELEMENT ficha (nombre+, apellido+, direccion+, foto?)>
<!ELEMENT nombre (#PCDATA)>
<!ATTLIST nombre sexo (masculino|femenino) #IMPLIED>
<!ELEMENT apellido (#PCDATA)>
<!ELEMENT direccion (#PCDATA)>
<!ELEMENT foto EMPTY> ]>
<ficha>
<nombre>Angel</nombre>
<apellido>Barbero</apellido>
<direccion>c/Ulises, 36</direccion>
</ficha>
```

Las **limitaciones** de la DTD son las siguientes:

- Posee un lenguaje propio de escritura lo que ocasiona problemas a la hora del aprendizaje, pues no sólo hay que aprender XML, sino también conocer el lenguaje de las DTDs.
- Para el procesado del documento, las herramientas y analizadores (*parsers*) empleados para tratar los documentos XML deben ser capaces de procesar también las DTDs .

- No permite el uso de *namespaces* y estos son muy útiles ya que permiten definir elementos con igual nombre dentro del mismo contexto, siempre y cuando se anteponga un prefijo al nombre del elemento.
- Tiene una tipología para los datos del documento extremadamente limitada, pues no permite definir el que un elemento pueda ser de un tipo número, fecha, etc. sino que sólo presenta variaciones limitadas sobre cadenas.

Nota: Namespaces

Los espacios de nombres, *namespaces*, permiten eliminar las ambigüedades y solucionar los problemas de homonimia que se producen en los documentos, ya que en un mismo documento existen palabras con el mismo nombre (p.e. "capital"), pero con diferentes significados y espacios semánticos (término geográfico/término económico-financiero).

Por ejemplo, tenemos la ambigüedad del nombre capital para capitales del mundo e inversiones y surgió la pregunta de cómo combinar en un mismo documento varios vocabularios.

```
<inversiones>
<país nombre="España">
<capital>Madrid</capital>
<capital>2000€</capital>
```

Solución: asignar un nombre único a cada etiqueta creando una autoridad mundial que asignara dicho nombre, por eso se pensó en los URIs. Un URI es un identificador global único. Aquí no se trata de utilizar un URI como enlace, ni tiene por qué tener contenido, los URI sólo se utilizan para que el nombre sea único. Ejemplo: <http://www.hipertexto.info>

Los nombres de espacio (namespaces) establecidos en la especificación del W3 Consortium Namespaces in XML (<http://www.w3.org/TR/REC-xml-names/>) sirven para proveer un método simple para calificar los nombres que se usan en los documentos XML por asociación con los namespaces identificados por URI (Uniform Resource Identifier). Este Identificador de Recursos Uniforme es un método que combina URNs y URLs, esto es, nombres/direcciones y que sirve para identificar de forma universal recursos de todo tipo que existen la World Wide Web.

En nuestro ejemplo, la solución estaría en asociar a cada etiqueta un URI que indicara a qué espacio de nombres pertenece:

```
[http:www.bolsa.es]:capital
[http:www.geograf.com]:capital
```

Por ejemplo, de la siguiente forma:

```
<[http://www.bolsa.es]:inversiones>
<[http://www.geograf.com]:país
```



```

[http://www.geograf.com]:nombre="España">
<[http://www.geograf.com]:capital>Madrid
</[http://www.geograf.com]:capital>
<[http://www.bolsa.es]:capital>2000€
</[http://www.bolsa.es]:capital>
<[http://www.bolsa.es]:país>
. . .
</[http://www.bolsa.es]:inversiones>

```

Asignando alias: **xmlns:alias** define *alias* en el ámbito de un elemento.

```

<bolsa:inversiones
xmlns:bolsa=http://www.bolsa.es
xmlns:geograf="http://www.geograf.com">
<geograf:país geograf:nombre="España">
<geograf:capital>Madrid</geograf:capital>
<bolsa:capital>2000€</bolsa:capital>
</geograf:país>
. . .
</bolsa:inversiones>

```

Actualmente, los **esquemas XML** han sustituido a los **DTD** como instrumento de validación de documentos XML, debido a que soporta tipos de datos, está escrito en XML y es extensible

6.4.2. XML Schema (XSD)

Al igual que las DTDs, los Schemas describen la estructura de la información. El motivo de la creación de este nuevo estándar para realizar la labor de las DTDs es, básicamente, la utilidad. Durante un tiempo, y a falta de otra solución más ajustada, se emplearon los mecanismos que proporcionaba SGML para modelizar la información en XML. Pero el descubrimiento de nuevas aplicaciones de XML al margen de la estructuración de documentos forzó la creación de otras soluciones que ayudasen a solventar los nuevos problemas a los que se enfrentaba el mercado. En el apartado siguiente veremos las principales diferencias que existen entre las DTDs y los XML Schemas.

Ejemplo de XML Schema:

```

schema targetNamespace="http://www.aqs.es/Schema_y_DTDs"
xmlns="http://www.w3.org/2001/XMLSchema"
xmlns:AQS="http://www.aqs.es/Schema_y_DTDs" >
<element name="Articulo" type="AQS:tArticulo" />

```

```

<complexType name="tArticulo">
  <element name="Cabecera" type="AQS:tCabecera"/>
  <element name="Cuerpo" type="AQS:tCuerpo"/>
  <element name="Final" type="AQS:tFinal"/>f
  </complexType>
  <complexType name="tCabecera">
    <element name="Titulo" type="string"/>
    <element name="Autor" type="string"/>
  </complexType>
  .....

```

Podemos definir **XML Schema** como un *lenguaje de esquema* utilizado para describir la estructura y las restricciones de los contenidos de los documentos XML de una forma muy precisa, más allá de las normas sintácticas impuestas por el propio lenguaje XML. Se consigue así, una percepción del tipo de documento con un nivel alto de abstracción. Fue desarrollado por el World Wide Web Consortium (W3C) y alcanzó el nivel de recomendación en mayo de 2001.

Un **esquema XML** permite describir la estructura de un documento XML, con el fin de poder validar si su estructura es válida para un contexto determinado.

El propósito de un **esquema XML** es definir los componentes válidos de un documento XML:

- elementos que pueden aparecer
- atributos
- elementos hijos
- orden y número de los elementos
- tipos de datos
- valores por defecto y fijos

A continuación un ejemplo de un documento XML y de su esquema XML correspondiente:

Documento XML
<pre> <?xml version="1.0" encoding="UTF-8"?> <correo xmlns="http://www.ideo.es" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="http://www.ideo.es correo.xsd"> <origen>Paco@yahoo.es</origen> <destino>Marta@google.es</destino> <asunto>Deportes</asunto> <texto>Tengo unas entradas para un partidotexto> </correo> </pre>

Esquema XML

```
<?xml version="1.0"?>
<xs:schema
xmlns:xs="http://www.w3.org/2001/XMLSchema" targetNamespace="http://www.ideo.es"
xmlns=http://www.ideo.es" elementFormDefault="qualified">
<xs:element name="correo">
  <xs:complexType>
    <xs:sequence>
      <xs:element name="origen" type="xs:string"/>
      <xs:element name="destino" type="xs:string"/>
      <xs:element name="asunto" type="xs:string"/>
      <xs:element name="cuerpo" type="xs:string"/>
    </xs:sequence>
  </xs:complexType>
</xs:element></xs:schema>
```

El propósito del estándar **XML Schema** es definir:

- la **estructura** de los documentos XML que estén asignados a tal esquema
- y los **tipos de datos** válidos para cada elemento y atributo.

En este sentido las posibilidades de control sobre la estructura y los tipos de datos son muy amplias.

- Se puede controlar la **cardinalidad** de los elementos (nº max y min de apariciones).
- Se pueden dar **valores fijos** o **por defecto** a los elementos y atributos. Los atributos pueden ser opcionales u obligatorios.
- Se pueden **enumerar** los valores posibles de un elemento o de un atributo.
- Siguiendo la filosofía de la programación orientada a objetos, se pueden declarar elementos **abstractos** que pueden ser sustituidos por otros en el fichero XML.
- Se pueden **agrupar** elementos de cara a la reutilización de código con tres opciones:
 - **Sequence**: Tienen que aparecer todos los elementos en el fichero XML y en el mismo orden en que han sido introducidos en el esquema.
 - **Choice**: Se realiza una elección entre varios elementos.
 - **All**: Si unos elementos se agrupan dentro de esta opción tienen que aparecer todos o ninguno en el fichero XML.
- Se pueden **importar** declaraciones de otros esquemas, lo cual facilita la reutilización de código.
- Los elementos importados pueden ser **redefinidos**.

- Al restringir el contenido de los ficheros XML es posible *intercambiar* información entre aplicaciones con gran seguridad. Disminuye el trabajo de comprobar la estructura de los ficheros y el tipo de los datos.
- XML Schema tiene un *enfoque modular* que recuerda a la programación orientada a objetos y que facilita la reutilización de código

El XML schema es una recomendación del World Wide Web Consortium (W3C) que empezó a trabajar en XML Schema en 1998. La primera versión se convirtió en una recomendación oficial en mayo de 2001.

Esta recomendación está desarrollada en tres partes:

- XML Schema Parte 0 *Primer*: es una introducción no normativa al lenguaje, que proporciona una gran cantidad de ejemplos y explicaciones detalladas para una primera aproximación a XML Schema.
- XML Schema Parte 1 *Structures*: es una extensa descripción de los componentes del lenguaje.
- XML Schema Parte 2 *Datatypes*: complementa la Parte 1 con la definición de los tipos de datos incorporados en XML Schema y sus restricciones.

6.5. Tecnologías XML. XSL, XLINK

XSL (Extensible Stylesheet Language)

El objetivo inicial de XML era proporcionar un medio de describir datos separándolos de su presentación, especialmente en el contexto de la World Wide web. XML Versión 1.0 trata la descripción de los datos. Una tecnología asociada, llamada **XSL** fue la encargada de tratar el punto de vista de la presentación.

XSL (Extensible Stylesheet Language), expresión inglesa traducible como "lenguaje extensible de hojas de estilo") es una familia de lenguajes basados en el estándar XML que permite describir cómo la información contenida en un documento XML cualquiera debe ser transformada o formateada para su presentación en un medio.

Esta familia está formada por tres lenguajes:

- XSLT (siglas de Extensible Stylesheet Language Transformations, lenguaje de hojas extensibles de transformación), que permite convertir documentos XML de una sintaxis a otra (por ejemplo, de un XML a otro o a un documento HTML).
- XSL-FO (lenguaje de hojas extensibles de formateo de objetos), que permite especificar el formato visual con el cual se quiere presentar un documento XML, es usado principalmente para generar documentos PDF.

- XPath, o XML Path Language, es una sintaxis (no basada en XML) para acceder o referirse a porciones de un documento XML.

Estas tres especificaciones son recomendaciones oficiales del W3C.

Desde el 2005 ya son soportadas por algunos navegadores, por ejemplo Mozilla o Internet Explorer, aunque, en su lugar, se pueden usar las CSS que son 100% compatibles aunque con una codificación diferente.

XSL es un lenguaje bastante simple. Proporciona una rica sintaxis para expresar patrones de emparejamiento y reemplazamiento. Es declarativo. Se puede leer fácilmente lo que XSLT dice hacer. Usando sus especificaciones asociadas (XPath y XQL) se puede especificar algunas consultas muy potentes sobre un documento XML. Además XSLT incorpora la capacidad de llamar funciones en otros lenguajes de programación tales como VBScript o Java a través del uso de las funciones extendidas. Esto significa que XSL puede ser usado para realizar consultas y selecciones, y llamar a Java u otro lenguaje para implementar la computación necesaria o la manipulación de cadenas de caracteres. Para tareas simples, XSLT proporciona prestaciones para manejar cadenas y operaciones aritméticas.

XLink (XML Linking Language)

Con la tecnología HTML actual es posible construir conjuntos de datos geográficos enlazados. Uno puede construir mapas que son enlazadas a otros mapas. El mecanismo de enlazado HTML tiene, de todas formas, muchas limitaciones, y como resultado de esto, no es práctico construir conjuntos de datos distribuidos que sean grandes y complejos como ocurre en los sistemas del mundo real.

La mayor limitación significativa es que los enlaces HTML son codificados en el código tanto en la fuente () como en el documento objetivo, un hecho que hace a un sistema frágil e imposible de escalar. **XLink** evita estos problemas al permitir enlaces “fuera de línea”. En los enlaces en línea y fuera de línea, las fuentes apuntan sólo a bases de datos enlazadas y es la base de datos enlazada la que proporciona el puntero al elemento específico del documento direccionado. El enlace no está codificado en ningún documento.

XLink o Lenguaje de vínculos XML es una recomendación del World Wide Web Consortium (W3C) que permite crear elementos de XML que describen relaciones cruzadas entre documentos, imágenes y archivos de Internet u otras redes. De esta forma, **XLink** permite:

- Crear una relación de vínculos entre varios documentos.
- Agregar a un vínculo información acerca del mismo (metadatos).

- Crear y describir vínculos a documentos en multitud de ubicaciones.

Es decir, es un el lenguaje, definido en términos de marcas XML, que nos va a permitir introducir enlaces en nuestros ficheros XML, de modo que podamos relacionar unos ficheros con otros.

Por ejemplo, el XLink tiene una gran importancia para GML al hacer posible construir conjuntos de datos geográficos distribuidos y escalables. Incluso más importante es que XLink y XPointer permiten construir aplicaciones específicas indexadas por fechas. ¿Se necesita tener un grupo de edificios organizados por su dirección postal o se quiere crear un índice de parcelas de una granja basadas en el tipo de cosecha? Con XLink y XPointer, estos y otros muchos esquemas de índices pueden ser construidos, y todos sin alterar la fuente de datos.

Bibliografía

- [1] Extensible Markup Language (XML): <http://www.w3.org/XML/>
- [2] XML.com: <http://www.xml.com>
- [3] <http://es.wikipedia.org/wiki>

Tema 7. Sistemas de Información Geográfica. Definición. Componentes de un SIG. Tipos de SIG. Organización e implementación de un SIG. Aplicaciones.

7.1. Sistemas de Información Geográfica

Los Sistemas de Información Geográfica (SIG) son el resultado de la aplicación de las llamadas Tecnologías de la Información (TI) a la gestión de la Información Geográfica (IG).

En un primer momento, en los años 60, la Informática se utilizó para la producción de mapas, dando lugar a la Cartografía Asistida por Ordenador. Las técnicas se habían modernizado y automatizado en parte para producir el producto clásico, el mapa en papel.

El SYMAP (*Symagraphic Mapping Technique*) de la Universidad de Harvard (USA), aparecido en 1964 fue la aplicación pionera; se trataba de un conjunto de programas capaz de producir mapas por impresora superponiendo los caracteres habituales.

El siguiente paso fue aplicar las posibilidades de los métodos automatizados a interpretar y consultar la información representada en el mapa, así nacieron los SIG. El CGIS (Canadian GIS), esencialmente dedicado a gestión forestal, concebido y desarrollado por Tomlinson a partir de 1964, parece haber sido el primer SIG y, durante bastantes años ha sido el más voluminoso por el área cubierta y el número de capas gestionadas.

Durante los años 70 y 80 se produjo la explosión de los SIG, con grandes desarrollos y aplicaciones prácticamente a todas las áreas de actividad humana, que comenzaron como proyectos de representación y análisis de la realidad y que se han orientado progresivamente a la planificación y la toma de decisiones

7.1.1. Definición

En primer lugar, hay que decir que existen tres acepciones en uso para este acrónimo:

- SIG como disciplina. Se ocupa de la aplicación de las TI a la gestión de la IG; es una ciencia aplicada que se ocupa de todo lo referente a los SIG como proyecto, segunda acepción. En este sentido se habla de experiencia en SIG, cursos y seminarios sobre SIG, especialistas en SIG, libros sobre SIG, etc.

- SIG como proyecto. Cada una de las realizaciones prácticas de la disciplina SIG, cada instalación existente. En una primera aproximación, es un sistema capaz de proporcionar cierta información, ya procesada, sobre elementos de los que se ocupa la Geografía. Ésta es la acepción principal, luego nos detendremos para dar una definición más precisa.

- SIG como “software”. Las casas comerciales llaman SIG al programa, o programas integrados, que ofertan para el establecimiento de un SIG proyecto. Se trata de un caso de sinécdoque, en el que se nombra a la parte por el todo. Un SIG proyecto se compone de *hardware*, *software*, datos, personal y organización. El único problema estriba en que hay que ser consciente de la sinécdoque y no caer en la ingenuidad de creer que por el mero hecho de adquirir un SIG *software* ya se tiene un SIG proyecto.

Un SIG es algo más que un sistema de almacenamiento de mapas digitales, en realidad se trata más de una herramienta informatizada de resolución de problemas con una componente geográfica.

La definición más extendida de SIG, con pequeñas variaciones, es la establecida por el Departamento de Medio Ambiente (DoE), Burrough, Goodchild, Rhin y otros. Entresacando lo mejor de cada formulación y refundiéndolo a nuestro criterio, podemos sintetizarla diciendo que un SIG es un:

“Conjunto integrado de medios y métodos informáticos, capaz de recoger, verificar, almacenar, gestionar, actualizar, manipular, recuperar, transformar, analizar , mostrar y transferir datos espacialmente referidos a la Tierra”

Sin embargo, nos parece mucho más ilustrativa y significativa la definición de Bouillé (1978), que concibe un SIG como un:

“Modelo informatizado del mundo real, descrito en un sistema de referencia ligado a la Tierra, establecido para satisfacer unas necesidades de información específicas respondiendo, del mejor modo posible, a un conjunto de preguntas concreto”

Esta definición implica:

- Una capacidad de registro selectiva ante los fenómenos del entorno, un proceso de abstracción para generar un modelo simplificado, y la elección de una estructura conceptual para los entes considerados, sus propiedades y los sucesos en los que se ven implicados. Un SIG contiene una versión simplificada, abstraída y estructurada de una parte del mundo real.
- Que los datos tienen que estar descritos en un Sistema de Referencia ligado a la Tierra, condición indispensable para que estemos hablando de Información Geográfica, con todas sus características y su problemática específica.

- Una finalidad que orienta todo el proyecto: las consultas a responder, es decir los requisitos que definen de qué SIG se trata.
- Que un SIG es un proyecto de ingeniería que busca una solución óptima para resolver un problema.

7.2. Componentes de un SIG

7.2.1. Hardware

Grupo de componentes materiales utilizados en un sistema informático. Es la parte tangible, física y real de un Sistema de Información.

Los equipos de *hardware* más populares son los ordenadores personales. En SIG, también se usan las estaciones de trabajo o *Workstation*.

Es el componente que ofrece cada vez más prestaciones a un coste cada vez más reducido. Evoluciona a una gran velocidad y se puede decir que es el componente mejor resuelto con diferencia.

7.2.2. Software

Es el conjunto de programas que corre en un sistema informático, la parte no física. La palabra nace por oposición a *hardware*. Depende de él, evoluciona mucho más lentamente, y esto llega a ser un problema, porque usando lenguajes de alto nivel no se aprovechan al máximo las posibilidades de la máquina.

El tiempo medio de vida de una versión de una aplicación es de 2 años.

Un software SIG suele incluir los siguientes grupos de funciones:

- Funciones de Entrada: Los procedimientos de entrada de datos en un SIG permiten:
 - Introducir datos posicionales digitalizando entidades cartográficas de forma georreferenciada, con sus coordenadas x, y, z en base a un sistema de referencia convencional.
 - Introducir datos alfanuméricos. Los más importantes son los códigos individuales y únicos de cada entidad cartográfica, que servirán como identificador y nexo de unión entre las entidades cartográficas y sus atributos alfanuméricos.
 - Importar datos existentes procedentes de otras fuentes y formatos de datos.
- Funciones de Gestión: Las funciones de gestión son llevadas a cabo por un subsistema del SIG: el sistema gestor de la base de datos (SGBD o DBMS- *Database Management System*). Su finalidad es permitir la independencia entre la organización física y la organización lógica

de los datos.

Las funciones principales de un SGBD de un SIG son controlar:

- la organización físico-lógica de los datos
- el almacenamiento
- la recuperación
- y la actualización

• Funciones de Manipulación: Por funciones de manipulación entendemos los procedimientos utilizados para la:

- estructuración topológica
- transformación
- superposición
- e integración de los datos geográficos, cartográficos y temáticos

La función más importante es la de estructuración topológica de los datos, que consiste en la definición digital explícita de las interrelaciones geométricas de los fenómenos (intersección, conectividad, etc.) representados en la base de datos. Este concepto se explicará más adelante.

• Funciones de Análisis:

Las funciones de análisis espacial del SIG son, sin duda, las más representativas y las que le diferencian claramente de otros sistemas de información espacial.

Pueden ser:

- Funciones de recuperación
- Funciones de superposición
- Funciones de vecindad
- Funciones de conectividad

• Funciones de Representación:

El quinto tipo de funciones del *software* SIG son las de representación de los resultados obtenidos del análisis y del tratamiento de los datos.

Los resultados expresan tanto la información contenida en la base de datos geográficos como las operaciones y análisis realizados,

Los procedimientos de representación de resultados incluyen: tratamiento de textos, Simbolización, transformación de los resultados numéricos en resultados gráficos, etc.

Este componente ha evolucionado razonablemente bien y cada vez hay soluciones más eficaces y fáciles de utilizar. Hay programas muy sofisticados y avanzados, sin embargo el proceso de selección de software es complicado y es una de las fases críticas de todo proyecto SIG.

7.2.3. Datos

La información geográfica es intrínsecamente muy compleja y presenta características peculiares que hacen que su gestión constituya una especialidad diferente de la mera gestión de SI: es muy voluminosa, fractal, borrosa, dinámica y polimorfa.

Para abstraer adecuadamente un modelo que represente el mundo real, o al menos la parte que nos interesa, hay que definir un Modelo Conceptual, fase a la que no se le presta la suficiente atención y que resulta ser crítica.

Los datos constituyen uno de los componentes más problemáticos: si son vectoriales por su escasez (todavía hay grandes zonas de la superficie terrestre sin cartografiar a escalas grandes) y si son ráster por su coste, que puede llegar al 60 % o más del coste total del proyecto.

7.2.4. Personal

Conjunto de personas que trabajan en un proyecto SIG y que debe estar especializado tanto en tecnologías de la información como en las disciplinas relacionadas con la cartografía (Geodesia, Fotogrametría, Topografía, Cartografía, Geografía, etc.). El mayor problema de este componente es, por lo tanto, la formación.

7.2.5. Organización

Es de vital importancia, jerárquicamente lo primero, y a menudo no se le presta la atención debida. Teniendo el resto de componentes de gran calidad y adecuados al proyecto, si falla la organización, no sólo no está garantizado el que se complete el trabajo con los plazos y dinero previstos, sino que es muy probable que no se llegue nunca a finalizar.

La organización incluye:

- Definición de objetivos claros y concretos.
- Presupuesto, medios y rentabilidad asegurados.
- Estudio detallado de viabilidad.
- Planificación realista y precisa.
- Coordinación de las distintas fases del proceso.
- Estructura jerárquica clara y eficaz, con reparto de responsabilidades y de las correspondientes capacidades de decisión.

- Cuidadosa labor de selección de *hardware*, *software* y personal.
- Elaborar planes de formación continua del personal y de renovación de *hardware*, *software* y datos.
- Controles de calidad.
- Previsión de todos los supuestos.
- Sentido común.

Un sistema en el que alguno de los componentes citados presente un desequilibrio frente a los demás, no puede funcionar satisfactoriamente y por lo tanto fracasará. El más importante de ellos es la organización, porque en realidad engloba a los demás.

Un aspecto básico de la organización es la planificación, teniendo en cuenta la previsión de todas las situaciones que se puedan presentar por improbables que parezcan, de cada una de las fases de un proyecto SIG.

7.3. Tipos de SIG

7.3.1. Por la arquitectura

- SIG duales

También se le llama el modelo híbrido o dual. Es la solución más extendida. Se mantiene la información gráfica gestionada por un CAD, con una clave identificadora única para cada elemento. Esa clave sirve para realizar búsquedas en el GBD que contiene todos los atributos e información alfanumérica.

Inconvenientes: la parte gráfica no se beneficia de todas las ventajas de estar bajo un GBD, en particular la indexación espacial y el bloqueo espacial admiten soluciones parcialmente satisfactorias.

- SIG orientados a objetos (OO)

Los dos mundos están integrados, primero consiguiendo que un GBD maneje ficheros gráficos (el problema es la longitud variable), y segundo aplicando una concepción orientada al objeto, en la que los elementos contengan información sobre sus propiedades, comportamiento, mensajes de error, etc.

Inconveniente: aparece un límite en el volumen de información que se puede gestionar. En unos casos porque se tomó un GBD que se quedó anclado, sin evolucionar, al no ser relacional y aceptar registros variables. En otros porque se trabaja en memoria real, y siempre hay un límite para la RAM de un equipo. No está totalmente implantada ni extendida, constituye un enfoque algo inmaduro todavía, pero es una concepción filosófica y conceptualmente más avanzada.

7.3.2. Por el tipo de respuesta

- SIG transaccionales

Incluyen consultas, altas, bajas, modificaciones, informes, listados, etc.

- SIG con simulación

Contienen un modelo de comportamiento en el tiempo para algún aspecto de la realidad. Hay modelos atmosféricos, oceanográficos, de difusión en ríos, tráfico en carreteras, muy desarrollados.

- SIG decisionales

Los llamados “*Spatial Decision Support System*” (SDSS), que con técnicas de Inteligencia Artificial, sustituyen total o parcialmente la capacidad de decisión humana.

7.3.3. Por los datos

- SIG 2D 2D1/2 y 3D

- Los SIG 2D son los que utilizan un sistema de coordenadas plano, basado en sólo dos coordenadas, (x, y) o (φ, λ) , por lo que se representa el mundo real como si fuera plano.

- En los SIG 2D1/2, se almacena la tercera coordenada, pero tan sólo como un atributo, nunca como una coordenada más, sin utilizarla para la generación de topología o el cálculo de distancias, áreas, etc.

- En los SIG 3D, se almacenan tres coordenadas y se tratan las tres de igual manera.

- SIG ráster y vector

- SIG ráster

Maneja datos ráster, que consisten en una serie de planos conceptuales superponibles, divididos en celdillas (píxeles o teselas) mínimas que pueden almacenar un valor numérico para un atributo, o un dato binario (0/1).

Los datos de satélite son de este tipo. En un SIG ráster, el volumen de almacenamiento es enorme; el análisis superficial y el cruce de información es más fácil, y la gestión de redes más difícil.

Existen extensiones del modelo ráster a tres dimensiones, para describir vetas de mineral en tres dimensiones, capas geológicas, etcétera, en las que se manejan los conceptos de voxel, área y volumen.

- SIG vector

Maneja datos vectoriales, en los que los elementos (puntuales, lineales y superficiales) vienen descritos por las coordenadas de los vértices que definen las poligonales o curvas con que se los representa. Aquí la gestión de redes es más cómoda, y el análisis espacial más difícil.

- SIG temporales

Son los SIG que almacenan la coordenada tiempo de alguna manera, o bien como una coordenada más o como un atributo que permite contestar consultas temporales.

7.4. Organización e implementación de un SIG

Las principales fases de un proyecto SIG son las siguientes:

7.4.1. Organización y Planificación.

Arranca en primer lugar con una clara definición de objetivos y requisitos, que va a guiar todo el proyecto SIG. El siguiente paso será conseguir los recursos, tanto humanos como financieros, necesarios para la buena marcha del proyecto, y asegurar su continuidad en el tiempo. Por último, hay que incluir aquí, la coordinación entre el resto de las fases y la planificación previa de todas ellas.

7.4.2. Diseño del SIG

Comprende el diseño de todas las fases que siguen a ésta, por lo que incluye un diseño del SIG en su estado final para explotación y actualización y un conjunto de especificaciones y requerimientos muy concretos para cada una de las fases previas.

7.4.3. Captura

Conversión de la información de partida a formato digital. A menudo es necesario desarrollar alguna aplicación de captura que efectúe la conversión y realiza algunos chequeos sobre los datos recogidos. Esta fase, obviamente, no es necesaria si los datos fuente se adquieren ya en forma digital.

7.4.4. Control de calidad

Cada una de las fases debe llevar aparejada otra de detección y corrección de errores, cuyo fin es asegurar que las desviaciones del modelo frente a la realidad se mantienen dentro de los límites permitidos. Hay que incluir también aquí el estudio teórico de estimación de errores y su propagación, y la descripción final de calidad de los datos y resultados finales producidos por

el sistema. El control de la calidad debe iniciarse ya desde la captura y acompañar a todas las fases siguientes.

7.4.5. Tratamiento

Transformación de la estructura y propiedades de los datos capturados hasta que puedan ser gestionados por el sistema para obtener los resultados perseguidos. Suele conllevar una parte automática y otra interactiva. En la mayoría de los casos, los programas de tratamiento han de ser desarrollados a la medida, ya que las necesidades de cada proyecto suelen ser bien diferentes.

7.4.6. Almacenamiento y gestión

Carga en el sistema y mantenimiento del mismo dentro de un nivel de rendimiento aceptable, bajo el impacto de las aplicaciones de Explotación y Actualización. Incluye la optimización del sistema, el mantenimiento, la seguridad, estudios de ampliación de recursos informáticos, etc.

7.4.7. Explotación

Obtención de las respuestas deseadas, en la forma y bajo las circunstancias previstas, ante las consultas previstas en la definición de los objetivos del sistema. Los resultados que se esperan del sistema, pueden evolucionar en el tiempo, tanto en contenido como en forma, evolución que también debe ser prevista y planificada en la medida de lo posible. En ocasiones, esta fase incluye un desarrollo previo, más o menos costoso, de las aplicaciones de explotación.

7.4.8. Actualización

Puesta al día periódica de los datos almacenados para que reflejen la evolución temporal que sufre el mundo real de modo continuo. En algún caso extraordinario, esta fase no es necesaria, si se trata de un proyecto que persigue tan sólo la toma de una única decisión, como por ejemplo el emplazamiento de un acelerador de partículas, o el trazado ideal de un túnel intercontinental.

7.5. Aplicaciones

Aparte de los MDT, como SIG de relieve, y las aplicaciones meteorológicas y oceanográficas.

7.5.1. Agricultura y usos del suelo

Información y planificación de la agricultura y ganadería (a través de los pastos).

Hay compañías privadas que ofrecen servicios de monitorización de cosechas, es decir seguimiento en tiempo real de su evolución, como el servicio CROPCAST (pronóstico de cosechas) de la *Earth Satellite Corporation*.

Son aplicaciones de gran importancia económica y política. Se aplica análisis espacial, y en ocasiones se simula el crecimiento de las cosechas (arroz, trigo, canela, patata, cacao, café) con datos del terreno, meteorológicos y de técnicas de aprovechamiento y laboreo. Cada año se comparan las predicciones con los resultados reales y se mejoran los modelos.

Como el CORINE (Coordinación de Información del Entorno Europeo), que entre otras cosas elabora un mapa digital de Europa a 1:100.000 de usos del suelo, conteniendo superficies con topología y unos 80 códigos, acabado en 1994 y cuya actualización y corrección se está abordando ahora, en los años 200-2001, junto con un estudio de cómo ha evolucionado el uso del suelo europeo en el período 1990-2000.

Otra aplicación muy extendida es el cálculo de la pérdida anual de suelo fértil teniendo en cuenta el tipo de suelo, la pendiente, las prácticas de laboreo, el régimen de lluvias, etc.

7.5.2. Forestales y de espacios protegidos

Hay una gran tradición en Canadá y USA de SIG aplicados a explotación forestal, para determinar por ejemplo cantidades máximas permisibles de tala. El primer SIG se estableció en Canadá en 1964 (Tomlinson) para inventario de recursos forestales.

Existe un programa de seguimiento del Caribú por satélite (el NOAA) desde 1985. Se ha colocado un transmisor en el cuello de 20 animales y se establece su posición geográfica con un error de 500m varias veces al día. Se estudian movimientos estacionales (2.000 km), hábitos de conducta, impacto de oleoductos y grandes obras en sus rutas, etc.

Por último, las aplicaciones orientadas a la protección del medio ambiente están desarrollándose espectacularmente:

- proyecto SAICA (Sistema de Calidad de las Aguas)
- proyecto SINAMBA, sistema de información del medio ambiente, de la Agencia de Medio Ambiente de la Junta de Andalucía (España).

7.5.3. Arqueología

Por un lado, los SIG se utilizan en Arqueología para almacenar y analizar yacimientos arqueológicos. Por ejemplo, se tiene:

- Un fichero general de situación y acceso 1:400.000
- Un levantamiento del yacimiento a gran escala 1:100 a 1:1000

- Levantamientos de fotogrametría terrestre de objeto próximo de cuevas, fachadas, estatuas, etc.
- Atributos, descripciones literales cargados en los tres niveles.

Es muy útil cuando se tiene superposición de yacimientos, cosa muy frecuente en ciudades (Troya I a VI).

Por otro, se ha usado para predecir nuevos posibles emplazamientos, conjugando una serie de variables como: datos geológicos, altitud, pendiente, aspecto, protecciones naturales, proximidad de ríos, etc. El hombre siempre ha buscado: agua, comida, un microclima agradable, buenas comunicaciones y seguridad. Por lo tanto, se aplica un modelo empírico que permite localizar áreas candidatas a poseer nuevos yacimientos.

7.5.4. Prospección

Existen aplicaciones geológicas, que incluyen descripciones multicapas del subsuelo con atributos para consulta, análisis y producción de mapas.

Aplicaciones de prospección petrolífera, metales escasos, combinan datos geológicos, con otros de patrones de configuración del paisaje, relieve, composición superficial, análisis de aguas, levantamientos de anomalías geomagnéticas y gravimétricas. Es uno de los campos de más clara rentabilidad.

Existe un programa de prospección de Wolframio, en el Yukon (Canadá). El Wolframio es un metal escaso y muy útil en fabricación de aleaciones. Está asociado a plutones, rocas de magma que sube y se enfría, que han de estar enterrados para que el mineral no sea atacado por erosión. El sistema se basa en un método empírico: combina datos de un mapa geológico 1:25.000, imágenes Landsat, análisis de aguas y sedimentos en la zona. Se identifica un patrón de parámetros común a todos los yacimientos conocidos y se busca dónde se repite dicho patrón.

7.5.5. Catastro

El catastro necesita manejar la información geométrica de cada parcela, y tenerla relacionada con todos los datos necesarios del propietario correspondiente. La única forma eficaz de gestionarlo todo y mantenerlo al día es con un SIG.

La Dirección General del Catastro (Ministerio de Economía y Hacienda, España) tiene un ambicioso proyecto, el SIGCA. Basado en contratación de trabajos: digitalización y ortofotos. Existe un formato de intercambio desde principios de los 90 (Norma de Intercambio de Cartografía Catastral, NICCA), bien definido y descrito que, debido al volumen de contratación, se ha convertido hasta cierto punto en uno de los standards de facto en España. Luego se pasan

macro procesos de chequeo y se comprueba la calidad de los datos en una muestra representativa. Se manejan las siguientes escalas:

- en rústica 1:2.000 y 1:5.000 hecho el 60% aproximadamente
- en urbana 1:500 y 1:1.000 hecho el 65% aproximadamente, lo que incluye todas las poblaciones de más de 25.000 hab.

Según datos de 2002.

7.5.6. Cartografía

Por cercanía tecnológica, una de las aplicaciones inmediatas de los SIG, es la producción de cartografía. Si el sistema está orientado exclusivamente a producción de mapas, se trata de un sistema de CAO, no de un SIG. Pero hay sistemas que permiten análisis espacial de los datos y entre otras aplicaciones, permiten la obtención de mapas en una variedad de escalas, proyecciones y presentaciones considerable.

7.5.7. Redes de infraestructura

Mantenimiento de redes de gas, agua, electricidad, teléfono, alcantarillado. Son sistemas que permiten la localización de averías, situación exacta del lugar de actuación de las brigadas de reparación, conocer que parte de la red se queda fuera de servicio como consecuencia de una eventualidad en un punto, planificación de la red óptima, etc.

En un principio se hicieron SIG específicos y redundantes para una misma ciudad o región, lo lógico es tender al SIG polivalente, o al menos a que los datos se compartan, pero todavía no siempre es así.

7.5.8. Gestión urbana

Toda la gestión que un ayuntamiento realiza en una ciudad de pequeño o gran tamaño, puede ser realizada con la ayuda de un SIG corporativo: planeamiento urbanístico, concesión de licencias de obra, toma de decisiones, gestión ambiental (contaminación, calor, ruido); planeamiento de servicios (agua, gas, metro, bomberos); simulación de impactos (fábricas, canalización de ríos, túneles, puentes).

7.5.9. Planificación global

Son aplicaciones típicas de análisis espacial en las que se estudian emplazamientos ideales de nuevos equipamientos para una ciudad o una región, por ejemplo, nuevas industrias.

La universidad de Columbia ha desarrollado un SIG para emplazar polígonos industriales teniendo en cuenta : altitud, pendiente, distancia mínima a una carretera principal, distancia a otra industria contaminante, tendido eléctrico existente, no tener acuíferos en un cierto radio, estar a cierta distancia de las poblaciones de 10.000 habitantes, no estar cerca de una zona sensible (parque, reserva, yacimiento) , tipo de suelo, distancia a centros de distribución de materias primas y de productos acabados. Se encuentra qué zonas de la superficie requerida cumplen dichas condiciones.

7.5.10. Análisis de fenómenos sociales

Hay una serie de aplicaciones encaminadas a analizar cómo evoluciona un fenómeno social cualquiera en el espacio y tiempo, como: una epidemia determinada, un tipo de delito en relación con los índices de desempleo y escolarización, etcétera.

Son bastante frecuentes las aplicaciones en el campo de la criminología, orientadas fundamentalmente a la prevención de delitos.

Las Naciones Unidas está utilizando un SIG para estudiar la evolución de un Plan de Erradicación de la filaria en Ghana.

7.5.11. Tráfico

- Control de flota

Se basan en conocer en qué posiciones se hallan los vehículos de una flota (coches de policía, autobuses, coches de bomberos, camiones de transporte), para que una persona delante de una pantalla de ordenador, tome decisiones en función de su distribución.

- Envío de unidades

Envío de unidades (*Dispatching*). Respuesta en tiempo real mediante el envío de unidades móviles ante emergencias, con rutas óptimas, seguimiento, alternativas y coordinación de efectivos. Incluye un paso de sofisticación frente al punto anterior, aquí el sistema calcula qué

unidad móvil es la ideal para atender una emergencia teniendo en cuenta todos los factores, y puede incluso guiar a dicha unidad siguiendo el camino óptimo hasta su destino. Ejemplos: bomberos, policía, ambulancias y hospitales, radiotaxis, mensajería, operaciones retorno.

- Caminos óptimos
 - 1 vehículo y 1 destino

Sistemas de guiado (Proyecto Prometeus)

Coches inteligentes. Equipados con un microordenador, con la red viaria cargada en *floppy*, que permite el que marcando origen y destino, el ordenador marque el camino más corto.

Hay varios sistemas de este tipo, por ejemplo, el sistema CARIN (*Car Information Navigation*) de Philips admite CD-Roms intercambiable de cartografía urbana según la ciudad de que se trate; utiliza un sistema GPS para situar la ubicación del coche, complementado con un giróscopo para mantener la ruta y un cuentakilómetros de precisión para los recorridos por calles estrechas que impiden recibir señal del número mínimo de satélites; proporciona rutas mínimas y caminos alternativos en una pantalla de 5 pulgadas y mediante mensajes sonoros (“gire a la derecha”, “siga recto”, etc.).

Otros sistemas, ya relativamente antiguos, se centran en la red de carreteras y no utilizan el posicionamiento del vehículo. El OS británico vende un sistema de pilotaje llamado OSCAR (OS Car), con datos digitales de todas las carreteras del Reino Unido a 1:625.000, basado en un PC que con varias condiciones (velocidad en cada tipo de carretera, condiciones sobre gasolineras y hoteles, rutas cortadas) elige el camino mas corto y el más rápido.

También existen sistemas para los pilotos automáticos de barcos, basados en cartas náuticas digitales y sistemas GPS de posicionamiento.

- 1 vehículo y n destinos

Rutas estáticas: Ante una distribución de estaciones por las que tiene que pasar un transporte, cálculo de rutas óptimas (problema del viajante). Es un problema matemáticamente muy complejo, que puede incluir varios factores, y muy rentable de resolver. Ejemplos: transportes de seguridad, paquetería, correos, reparto de mercancías.

- n vehículos y n destinos

Las mejoras en la regulación del tráfico global suponen: mejora del transporte (público, privado y de urgencia), ahorro de combustible y disminución de la contaminación.

Históricamente, existen varios sistemas:

1) Planes de semáforos. Existen desde los años 70, un ejemplo es el TRANSIT británico.

- En tiempo fijo. Un ordenador elige entre una biblioteca de planes de semáforos, el óptimo según datos del tráfico. Ahorro global del 15% en tiempo para los conductores.

- En tiempo real. Se calcula el mejor plan de semáforos en función del tráfico. Sistema Siemens. Ahorro 22% en tiempo.

2) Sistemas de guiado individual con optimización según tráfico. Si un ordenador central recoge datos de tráfico y los transmite por radio al coche, el ordenador de éste puede elegir el camino más rápido y evitar calles cortadas.

3) Optimización individual y global (sistema LISB, ensayado 88). El coche sólo lleva un terminal no inteligente, el ordenador central halla una ruta de compromiso entre optimizar el camino del auto y el tráfico general. Dirige la circulación en tiempo real. El problema es el coste (económico, en personal, etc.) del sistema centralizado, es difícil o casi imposible hoy por hoy disponer de un ordenador que dé respuesta optimizada en función del tráfico en tiempo real a una buena parte de los coches de una gran ciudad. En horas punta, tendría que resolver del orden de 1.000.000 de consultas, analizando los datos de posición y velocidad de 1.000.000 de vehículos aproximadamente, con un tiempo de respuesta de muy pocos minutos.

7.5.12. SIG aplicado a los negocios

Otras aplicaciones muy interesantes son las de análisis estadístico con referencias geográficas. Es lo que se está empezando a llamar *geomarketing*.

Por ejemplo:

- estudios de mercado por zonas para determinar en qué barrio de una ciudad es más conveniente abrir un establecimiento comercial
- análisis de la evolución espacial en el tiempo de la renta per cápita para introducir un nuevo seguro
- intención de voto para planificar una campaña electoral.
- planificación de sucursales, oficinas o concesionarios de una empresa.

La General Motors y Nissan Motor Ibérica, han utilizado datos de los municipios españoles a escala 1:1.000.000 acompañados de información sobre población, nivel económico y grado de motorización para diseñar mediante un programa específico su red de concesionarios óptima.

7.5.13. Formación

Menos importante en volumen de instalaciones que las anteriores, pero muy interesantes por las innovaciones tecnológicas que aportan, los sistemas de producción de realidad virtual para entrenamiento de pilotos de avión en condiciones cuasi reales. Utilizan datos de MDT, usos de suelo e imágenes de satélite, en aplicaciones muy espectaculares.

Las aplicaciones orientadas a estudiantes de enseñanza secundaria que utilizan multimedia y datos a pequeñas escalas (1:500.000 y menores) están creciendo rápidamente.

Uno de los intentos más interesantes de acercarlos SIG al público ha sido el proyecto Doomsday, llevado a cabo por la BBC británica en 1986. Basado en un PC, con unos programas de consulta de manejo sencillo, incluye 21 000 mapas o ficheros en un disco óptico (similar a un CD) con datos geológicos, de suelos, población, empleo, estadísticas, etcétera... El sistema se puso a disposición de las escuelas a 3 000 libras esterlinas.

7.5.14. Ocio

Por un lado, se empieza a ver los SIG como una herramienta útil en el campo del turismo, que puede permitir: elegir al cliente qué hoteles cumplen una serie de requisitos, distancia máxima al mar o a la montaña, un clima determinado, comunicaciones por carretera, proximidad de ciudades importantes, cercanía de centros de interés turístico, arqueológico, etc. y además guiarle en su viaje por carretera hasta el destino deseado.

Por otro, aunque menos desarrollado, los juegos basados en la creación de una realidad virtual geográfica, como simuladores de vuelo, programas de respuesta a catástrofes naturales, etcétera.

7.5.15. Bélicas

Poco conocidas por razones obvias de seguridad, constituyen un poderoso motor de desarrollo de la tecnología debido a las grandes inversiones que manejan. Incluyen una gran variedad de aplicaciones: simulación de conflictos bélicos en escenarios realistas (juegos de guerra), simulación de conflagraciones nucleares, sistemas automáticos de guiado de misiles (el misil busca su objetivo utilizando GPS y reconocimiento de formas en imágenes de satélite), etc.

7.6. SIG básico

SIG de propósito general, de referencia, heredero de la idea de Catastro Polivalente, análogo a cartografía básica, que busca dotar a un país de una infraestructura general en lo que a datos geográficos digitales se refiere.

Problemas:

- sólo a escalas medias (1:500.000 a 1:100.000) es fácilmente abordable el proyecto en un país de mediano tamaño como España. El volumen de datos en un SIG general es tal, que los problemas se agudizan y aparecen otros nuevos respecto a los SIG monotemáticos.

- el volumen de trabajo es enorme, y la amortización muy lenta y difícil

- sólo puede ser abordado por la Administración Pública, lo que conlleva problemas bien conocidos: dificultad de convencer a los responsables políticos de la oportunidad de inversiones muy grandes a largo plazo (eterno problema del catastro polivalente); la Administración está encorsetada por una burocracia muy rígida que dificulta la amortización de tales inversiones; falta continuidad en dichas inversiones;etc.

- el estado previo de la cartografía existente es determinante.

Ejemplos:

- Holanda: SIG a 1:250.000. Acabado en el 1986 y revisado cada 4 años

- Francia: SIG topológico 1:200.000 acabado en 1996

 - SIG topológico 1:25.000 fecha prevista de finalización 2020

- España: SIG 1:200.000 topológico acabado 1992 actualizado cada año

 - SIG 1:25.000 datos depurados sin topología finalización prevista 2003

Resumen y perspectivas

1) Aplicación típica SIG:

- Análisis y cruce de gran cantidad de datos y factores

- Fenómenos en los que la situación geográfica es significativa

- El tiempo de respuesta es importante

Prácticamente casi cualquier actividad humana, hecho social o evento natural, puede ser planificado, coordinado, estudiado o previsto mediante un SIG.

Por ejemplo un sistema que informe a un cinéfilo, para cada una de las películas de estreno, a que sala debe ir, en función de las distancia y comunicaciones, y de que la fila siete esté o no ocupada. La cuestión clave es si vale la pena o no en cada caso, si sería rentable.

2) Los SIG monotemáticos gozan de muy buena salud en Canadá (pioneros), USA, Japón y algo menos en Europa. En cambio en algunos países europeos (Francia, Alemania, Holanda, norte de Europa, España) si existen SIG básicos, de carácter general.

3) En los países latinoamericanos, los SIG empiezan a despegar con realizaciones muy interesantes. Las aplicaciones más frecuentes son Catastro (Argentina, Colombia, Brasil), Medio Ambiente, Redes de Infraestructura (Chile) y Sistemas de Gestión Municipal (Colombia, Brasil).

4) En algunos países (Reino Unido), los organismos oficiales prefieren producir y comercializar mapas digitales o bien datos de calidad para su posible integración en un SIG.

Bibliografía

[1] Bourrough, P. A. "Principles of Geographical Information System for Land Resources Assesment" 1986, OXFORD SCIENCE PUBLICATIONS (UK) 1ª edición.

[2] Bernhardsen, Tor "Geographical Information Systems" TOR BERNHARSDSEN 1992, VIAK IT (Noruega).

[3] Longley, P. A., M.F. Goodchild, D.J. Maguire y D.W.Rhind. "Geographic Information Systems and Science", 2001, JOHN WILEY and SONS.

[4] NCGIA Core Curriculum in GIS Science
<http://www.ncgia.ucsb.edu/giscc>

[5] GISWiki
<http://en.giswiki.net/wiki/Special:Allpages>

Tema 8. Modelado de datos. Modelos vectoriales: espagueti, topología, modelos en red, topología completa. Modelo ráster. Métodos de compresión. Ventajas e inconvenientes de los modelos vector/ráster. Orientación a objeto.

8.1. Modelado de datos

Los Sistemas de Información Geográfica, procesos desarrollados en su totalidad por equipos informáticos, manejan datos geográficos en **formato digital** procedentes de la discretización de modelos de información continua.

Estos modelos de información continua constituyen el denominado **formato analógico**. Un claro ejemplo de información geográfica en formato analógico es el mapa tradicional en soporte papel. Una imagen aérea obtenida con una cámara fotográfica analógica, la voz humana, o el sonido directo de un instrumento son otros ejemplos de información analógica.

La teoría de muestreo y los principios de discretización permiten obtener modelos discretos de la información continua cuyo resultado será el formato digital, apto para ser procesados por sistemas informáticos de información geográfica. La imagen captada por un satélite, el plano de una ciudad representado en Google Earth, un fichero de música en formato MP3 son ejemplos de información discreta en formato digital.

El problema de la conversión analógico/digital de la información geográfica se centra en la representación de la componente espacial de los datos. La correcta representación digital de los datos espaciales necesita la resolución de dos cuestiones: la geocodificación de los datos y la descripción en términos digitales de las características espaciales.

La geocodificación de los datos consiste en un procedimiento mediante el cual un objeto geográfico recibe directa o indirectamente una etiqueta que identifica su posición espacial con respecto a algún punto común o marco de referencia.

En segundo lugar se debe realizar una descripción de la posición geométrica de cada objeto y de las relaciones espaciales (topología) que mantiene con los restantes objetos geográficos existentes. Esta descripción se lleva a cabo mediante la abstracción o simplificación de todas las entidades existentes, creando un modelo de datos de los objetos a representar digitalmente. Un modelo de datos es, por lo tanto, “un conjunto de directrices para la representación lógica de los

datos en una base de datos, consistente en los nombres de las unidades lógicas de los datos y de las relaciones entre ellos”, en definitiva, “una abstracción del mundo real que incorpora sólo aquellas propiedades que son relevantes a la aplicación de interés en cada caso”.

Existen varios tipos de modelos digitales de datos geográficos: el modelo vectorial, el modelo ráster y el jerárquico recursivo. Todos ellos son perfectamente válidos para la representación de los mapas planos formados por puntos, líneas y áreas, pero cada uno encaja mejor en una tarea concreta. Unos resultan más útiles para el análisis de los datos espaciales y otros para la cartografía de la información espacial. El modelo vectorial está formado por un conjunto de coordenadas y el modelo ráster por imágenes digitales. Ambos sistemas son complementarios y desempeñan papeles específicos e intercambiables.

8.2. Modelo vectorial de datos espaciales

8.2.1. Conceptos generales

El modelo vectorial representa los elementos espaciales mediante la codificando explícita del límite o perímetro que separa el elemento de su entorno. Las líneas que definen esta frontera se representan vectorialmente mediante las coordenadas de los puntos o vértices que delimitan los segmentos rectos que las forman. (Figura 1).

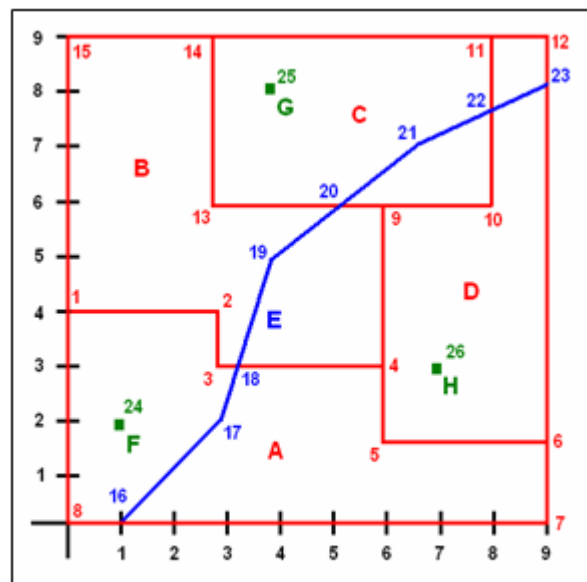


Figura 1. Representación Vectorial.

En la representación vectorial los **elementos puntuales** se representan mediante el par de coordenadas (X,Y) que definen su posición. Los **elementos lineales** se representan mediante las coordenadas (X,Y) de los vértices donde se cortan los segmentos lineales que se aproximan a la

forma del elemento. Por último, los **elementos superficiales** se codifican mediante las coordenadas (X,Y) de los vértices donde se cortan los segmentos lineales que formando un recinto cerrado se aproximan a la forma poligonal del elemento.

Según esto, en la representación vectorial el segmento lineal es el elemento esencial siendo capaz de constituir cualquier tipo de objeto espacial. El segmento lineal queda delimitado por dos vértices, codificándose por las coordenadas (X,Y) de sendos vértices. El elemento puntual es un caso concreto de segmento lineal de longitud cero que empieza y termina en el mismo vértice.

8.2.2. Estructura de datos en el modelo vectorial

Existen diferentes estructuras de datos que se ajustan al modelo vectorial definido anteriormente. Se puede diferenciar entre estructuras que no recogen la topología: espagueti, diccionario de vértices y las que permiten conocer la relación topológica de los elementos geométricos: ficheros DIME, modelo cadena/nodo, modelo en red.

Las más importantes son: la lista de coordenadas o estructura espagueti, el diccionarios de vértices, la organización DIME y la estructura arco/nodo.

- **Espagueti: Lista de coordenadas**

Esta estructura almacena únicamente las coordenadas que definen la entidad. De cada elemento geográfico existente se registra su nombre, el número de vértices que definen su frontera y las coordenadas (X,Y) de cada vértice. Los elementos puntuales están formados por un único vértice. En los elementos superficiales el primer y el último vértice son exactamente el mismo punto, indicando que se trata de una figura poligonal cerrada. (Tabla 1).

Esta estructura es la manera más sencilla y directa de estructurar datos geográficos en formato vectorial, pero presenta serias desventajas, registra la geometría espacial del objeto pero no su topología, despreciando la relación del elemento con los objetos que le rodean. Otra de las desventajas es la duplicidad innecesaria de datos (vértices) con el riesgo de que el mismo vértice sea registrado de forma diferente en cada segmento recto al que pertenezca, provocando problemas en la representación gráfica y visual de los datos espaciales.

LISTADO DE COORDENADAS											
A	11	B	11	C	8	D	10	E	8	F	1
0	4	0	4	3	6	3	6	3	6	1	2
3	4	0	9	3	9	6	3	3	2	G	1
3	3	3	9	8	9	6	6	3,5	3	4	8
4	3	3	6	8	7,5	8	6	4	5	H	1
6	3	5	6	8	6	8	8	5	6	7	3
6	2	6	6	6	6	8	9	7	7		
9	2	6	3	5	6	9	9	8	8		
9	0	3,5	3	3	6	9	8	9	8		
3	6	3	3			9	2				
0	0	3	4			3	6				
0	4	0	4								

Tabla 1. Listado de Coordenadas.

- **Diccionario de Vértices**

Esta estructura mejora la lista de vértices registrando una sola vez las coordenadas de cada vértice, cada uno de los cuales ha sido previamente identificado con un nombre o etiqueta. En primer lugar se crea un listado de coordenadas de vértices en el que figuran el nombre o etiqueta del vértice y las coordenadas (X,Y) de su localización (Tabla 2). A continuación se crea el diccionario de vértices que constituyen cada objeto espacial identificable en el mapa. (Tabla 3).

COORDENADAS DE VERTICES					
Etiqueta	X	Y	Etiqueta	X	Y
1	0	4	14	3	9
2	3	4	15	0	9
3	3	3	16	3	6
4	6	3	17	3	2
5	6	2	18	3,5	3
6	9	2	19	4	5
7	9	0	20	5	6
8	0	0	21	7	7
9	6	6	22	8	8
10	8	6	23	9	8
11	8	9	24	1	2
12	9	9	25	4	8
13	3	6	26	7	3

Tabla 2. Fichero 1: Coordenadas de Vértices.

DICCIONARIO DE VERTICES											
Polígono A	1	2	3	18	4	5	6	7	16	8	1
Polígono B	1	15	14	13	20	9	4	18	3	2	1
Polígono C	13	14	11	22	10	9	20	13			
Polígono D	5	4	9	10	22	11	12	23	6	5	
Línea E	16	17	18	19	20	21	22	23			
Punto F	24										
Punto G	25										
Punto H	26										

Tabla 3. Fichero 2: Diccionario de Vértices.

Esta estructura elimina los problemas causados por la duplicidad de los vértices pero la topología sigue sin estar suficientemente detallada.

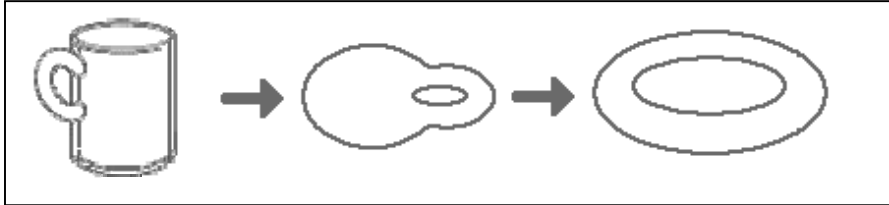
▪ **Topología. Conceptos.**

Una de las ventajas de los modelos vectoriales es la capacidad de almacenar, además de la información geométrica de las entidades, la topología y sus relaciones entre dichas entidades.

Las estructuras de almacenamiento de información vectorial hasta ahora descritas no son capaces de conservar la información topológica. Los modelos que a continuación se detallan estructuran la información geométrica de tal manera que añaden detalles topológicos de los elementos.

La topología es una rama de las Matemáticas, quizá la más reciente datada en el S.XVII, que analiza la posición de objetos, ocupándose de aquellas propiedades de las figuras que permanecen invariantes, cuando dichas figuras son plegadas, dilatadas, contraídas o deformadas, de modo que no aparezcan nuevos puntos, o se hagan coincidir puntos diferentes.

Si la rama de la geometría se fija en distancias, ángulos, volúmenes; la topología se centra en proximidad, agujeros, intersecciones, etc. En términos topológicos, una elipse es igual a una circunferencia, un cubo a una esfera y una taza de café a un toro.



Un ejemplo muy claro de topología es el plano de una red de metro en el que hay elementos puntuales (estaciones) y elementos lineales (líneas de metro) que unen las estaciones. Pero esa representación no es geoméricamente exacta, no conserva la forma real de las líneas con sus curvas, su longitud a escala ni tan siquiera la posición relativa de las estaciones. En cambio, el plano resulta muy útil y ofrece con exactitud la información que solicitada: la información topológica que mantiene intactos las estaciones y las líneas.

Este ejemplo topológico está relacionado con la teoría de grafos y los modelos en red, constituidos por nodos, bordes y caras.

- **Ficheros DIME (*Dual Independent Map Encoding*)**

La estructura vectorial de ficheros DIME recoge completamente la topología de los datos. Su utilidad principal es la de representar polígono. Se crea un listado de *coordenadas de vértices* con el nombre de cada vértice y sus coordenadas (X,Y) exactamente igual que en el Diccionario de Vértices (Tabla 2). Seguidamente se define la *descripción de los segmentos rectos* donde se codifican los segmentos lineales indicando el nombre del segmento, el vértice origen, el vértice final, el polígono que tiene a la izquierda y el polígono que tiene a la derecha (Tabla 4). Por último, se genera la *descripción de los polígonos* codificando los polígonos mediante su nombre y los segmentos rectos que los componen (Tabla 5).

Esta estructura recoge todos los elementos topológicos. Es posible averiguar con facilidad qué líneas conectadas, qué segmentos delimitan un polígono, qué polígonos son contiguos, etc.

DESCRIPCION DE LOS SEGMENTOS RECTOS				
Segmento	Poligono Dcha.	Poligono Izda	Vértice Origen	Vértice Final
1	A	B	1	2
2	A	B	2	3
3	A	B	3	18
4	A	B	18	4
5	A	D	4	5
6	A	D	5	6
7	A	Exterior	6	7
8	A	Exterior	7	16
9	A	Exterior	16	8
10	A	Exterior	8	1
11	B	Exterior	1	15
12	B	Exterior	15	14
13	B	C	14	13
14	B	C	13	20
15	B	C	20	9
16	B	D	9	4
17	C	Exterior	14	11
18	C	D	11	22
19	C	D	22	10
20	C	D	10	9
21	D	Exterior	11	12
22	D	Exterior	12	23
23	D	Exterior	23	6
24	A	A	16	17
25	A	A	17	18
26	B	B	18	19
27	B	B	19	20
28	C	C	20	21
29	C	C	21	22
30	D	D	22	23
31	A	A	24	24
32	C	C	25	25
33	D	D	26	26

Tabla 4. Fichero 2: Descripción de lo Segmentos Rectos

DESCRIPCION DE LOS ELEMENTOS	
Polígono A	1 2 3 4 5 6 7 8 9 10
Polígono B	1 2 3 4 11 12 13 14 15 16
Polígono C	13 14 15 17 18 19 20
Polígono D	5 6 16 18 19 20 21 22 23
Línea E	24 25 26 27 28 29 30
Punto F	31
Punto G	32
Punto H	33

Tabla 5. Fichero 3: Descripción de los elementos

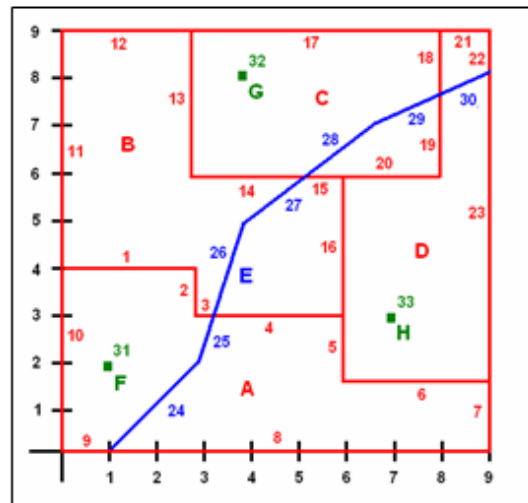


Figura 2. Segmentos Rectos

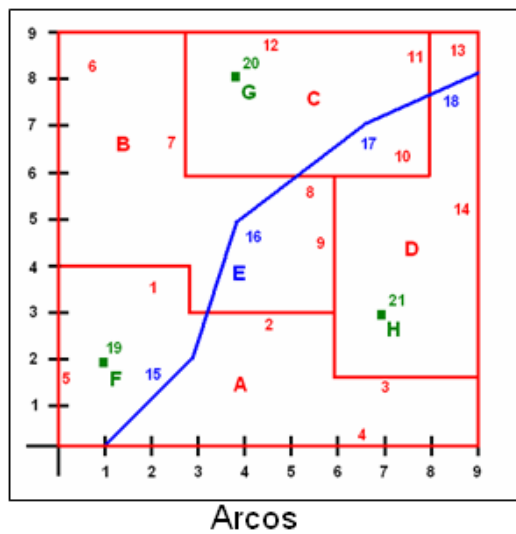
▪ **Modelo cadena/nodo**

La estructura de ficheros DIME se adapta perfectamente al registro de entidades espaciales geográficas de carácter artificial (límites administrativos, parcelas, etc.) cuya forma original se puede descomponer en un conjunto de segmentos rectos sin perder el aspecto del elemento geográfico.

La complejidad aumenta al codificar los fenómenos naturales (ríos, suelos, vegetación, etc.). El aspecto físico de estos fenómenos requiere una enorme cantidad de segmentos que dificultan la descripción topológica completa. Este problema se resuelve empleando la estructura cadena/nodo, en la cual el elemento base ya no es el segmento recto, si no la “cadena/nodo” o “arco/nodo” formado por la sucesión de todos aquellos segmentos rectos que tienen la misma

topología. Un segundo elemento fundamental es el “nodo”, vértice en el que se cruzan tres o más arcos o punto terminal de una línea o arco (Figura 3).

La estructura cadena/nodo está compuesta del *listado de coordenadas* de los puntos, tanto nodos como vértices de cada arco (Tabla 6). En este listado figura el nombre del arco, las coordenadas del nodo origen, de los vértices intermedios, si los hubiera, y del nodo final. A continuación tres ficheros definen la topología de los elementos geográficos. Un fichero de *topología de arcos* donde se codifican los arcos mediante su nombre, el nombre del nodo origen, el del nodo final, el polígono que se encuentra a su izquierda y el polígono que se encuentra a su derecha (Tabla 7). Otro fichero recoge la *topología de polígonos* que codifica los polígonos mediante su nombre y los arcos que lo componen (Tabla 8). Por último la *topología de los nodos* codifica los nodos mediante el nombre del nodo y los arcos que en él confluyen (Tabla 8).



Coordenadas de los nodos y de los vértices de cada arco			
Arco	Coordenadas Nodo origen	Coordenadas Vértices intermedios	Coordenadas Nodo final
1	(0,4)	(3,4)(3,3)	(3,5,3)
2	(3,5,3)		(6,3)
3	(6,3)	(6,2)	(9,2)
4	(9,2)	(9,0)	(3,6)
5	(3,6)	(0,0)	(0,4)
6	(0,4)	(0,3)	(3,3)
7	(3,3)	(3,6)	(6,6)
8	(6,6)		(6,6)
9	(6,6)		(6,3)
10	(6,6)	(8,6)	(8,7,5)
11	(8,7,5)		(8,3)
12	(3,3)		(8,3)
13	(8,3)	(9,3)	(9,3)
14	(9,3)		(9,2)
15	(3,6)	(3,2)	(3,5,3)
16	(3,5,3)	(4,5)	(6,6)
17	(6,6)	(7,7)	(8,7,5)
18	(8,7,5)		(9,3)
19	(1,2)		(1,2)
20	(4,3)		(4,3)
21	(7,3)		(7,3)

Figura 3. Arcos.

Tabla 6. Coordenadas de Nodos y Vértices.

Topología de arcos				
Arco	Nodo origen	Nodo final	Poligono Derecha	Poligono Izquierda
1	1	18	A	B
2	18	4	A	B
3	4	6	A	D
4	6	16	A	Exterior
5	16	1	A	Exterior
6	1	14	B	Exterior
7	14	20	B	C
8	20	9	B	C
9	9	4	B	D
10	9	22	D	C
11	22	11	C	D
12	11	14	C	Exterior
13	11	23	D	Exterior
14	23	6	D	Exterior
15	16	18	A	A
16	18	20	B	B
17	20	22	C	C
18	22	23	D	D
19	24	24	A	A
20	25	25	C	C
21	26	26	H	H

Tabla 7. Topología de Arcos.

Topología de elementos	
Elementos	Nodo final
A	1,2,3,4,5
B	1,2,6,7,8,9
C	7,8,10,11,12
D	3,9,10,11,13,14
15	15,16,17,18
19	19
20	20
21	21

Tabla 8. Ficheros 2 y 3. Topología de Elementos y de Nodos.

Topología de nodos	
Nodo	Arcos
1	1,6,5
2	1,2,15,16
3	2,3,9
4	3,4,14
5	4,5,15
6	6,7,12
7	7,8,16,17
8	8,9,10
9	10,11,17,18
11	13,14,18
12	19
13	20
14	21

▪ Modelos en red

La topología se encuentra íntimamente ligada a la teoría de grafos. La teoría de grafos estudia las propiedades de los grafos, que son colecciones de objetos llamados vértices (nodos) conectados por líneas (arcos) que pueden tener orientación (dirección asignada).

Típicamente, un grafo está diseñado por una serie de puntos (vértices) conectados por líneas (aristas).

Existen dos tipos de estructura de datos que describen un grafo:

- Estructura de lista:
 - De incidencia: Las líneas se representan con un vector de pares (ordenados, si el grafo es dirigido) de nodos que conecta esa línea.
 - De adyacencia: Cada vértice tiene una lista de vértices adyacentes a él.
- Estructura matricial:
 - De incidencia: Matriz de A (aristas) \times V (vértices) donde el elemento $[a \times n]$ vale 0 o 1 indicando si hay o no conexión.
 - De adyacencia: Matriz cuadrada $M \times M$ (número de vértices). Si entre el vértice x e y hubiera una arista, el valor del elemento $m_{x,y}$ sería 1, 0 en caso contrario.

En cuanto a su representación, existen grafos planares y no planares. Los grafos planares son aquellos que se pueden dibujar sin que dos segmentos se crucen. No es fácil establecer que grafo es plano, y tiene que ver con la topología.

Un problema de grafos y de topología es el problema de los puentes de Königsberg. Esta ciudad rusa tiene 7 puentes para salvar los dos brazos de un río que rodean a una isla (ver figura).



El problema consiste en localizar que trayectoria atraviesa los 7 puentes pasando solo una vez por cada uno de ellos.

Euler demostró el porqué de la imposibilidad de resolver este problema.

Figura 1: Los puentes de Königsberg

▪ Modelo espacial ISO

La normativa ISO que afecta a la estandarización de información geográfica vectorial y topología viene dada principalmente por las normas ISO19107 (esquemas espaciales), ISO19136 (GML) e ISO19137 (núcleos de perfiles de esquemas espaciales). La relativa a imágenes, coberturas y datos ráster viene dada por las normas ISO19123 (esquema para coberturas), ISO19129 (marco de trabajo), ISO19130 (sensores y modelos de datos).

Los datos vectoriales definidos en ISO19107 se caracterizan por ser tridimensionales, estar formados por primitivas simples y/o complejas y mantener relaciones topológicas.

Las primitivas geométricas fundamentales que define la norma son: el punto, la curva, la superficie y el sólido.

8.3. Modelo ráster de datos espaciales

8.3.1. Conceptos Generales

El modelo ráster, a diferencia del modelo vectorial, codifica el interior de los objetos geográficos registrando implícitamente la frontera del elemento. Esta codificación se consigue superponiendo al mapa analógico original una rejilla de unidades regulares, de igual forma y tamaño (Figura 5). En cada unidad de la rejilla se registra el valor que el mapa analógico o vectorial adopta (Figura 6). La rejilla puede estar formada por tres tipos de figuras geométricas elementales: cuadrados (rectángulos), triángulos regulares o hexágonos.

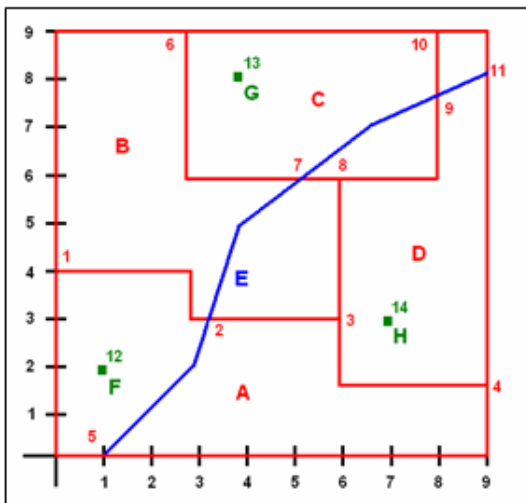


Figura 5. Nodos.

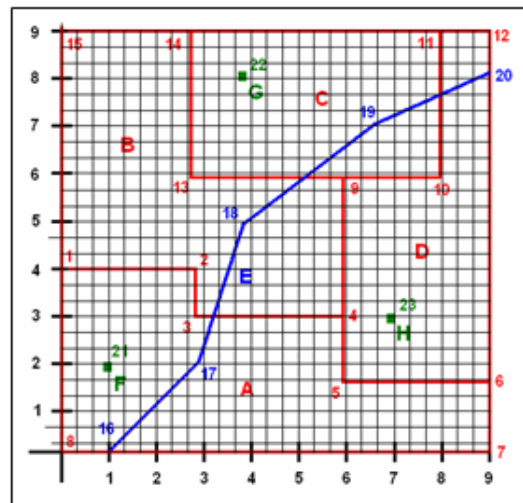


Figura 6. Superposición de malla rectangular.

Un parámetro fundamental en la rejilla cuadrada, la más utilizada de las tres, es la longitud de los lados de la cuadrícula base o “píxel”, cuanto más pequeña sea más precisa será la representación digital del mapa. La longitud del píxel en unidades del terreno nos proporciona la escala del mapa ráster generado.

El modelo ráster establece el origen de coordenadas (0,0) en la esquina superior izquierda del mapa. La razón de este origen se encuentra en la manera de trabajar de muchos equipos usados en la captura y tratamiento de datos ráster: scanner, sensores espaciales, impresoras, etc. El establecimiento del origen de coordenadas y la regularidad de la rejilla permiten recoger la topología del mapa ráster. De esta manera es posible conocer cuales son los vecinos de una celda de la rejilla. (Figura 7)

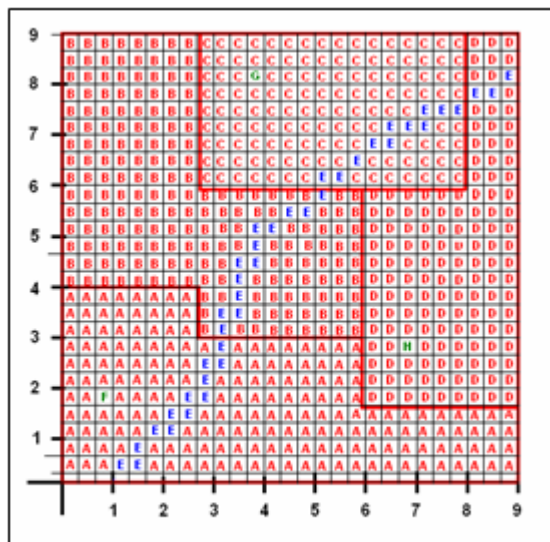


Figura 7. Estructura Ráster.

8.3.2 Estructuras de datos en el modelo ráster

Como sucede con el modelo vectorial, el modelo ráster permite establecer varias estructuras de datos. Las diferencias entre ellas estriban en la manera de almacenar la información registrada en este modelo de datos.

- **Estructura sin compresión. Enumeración Exhaustiva**

El valor de cada una de las unidades de la rejilla se registra individualmente. Esta estructura redonda información de manera innecesaria almacenando una gran cantidad de información. El mismo valor numérico aparece en píxeles contiguos (Figura 8).

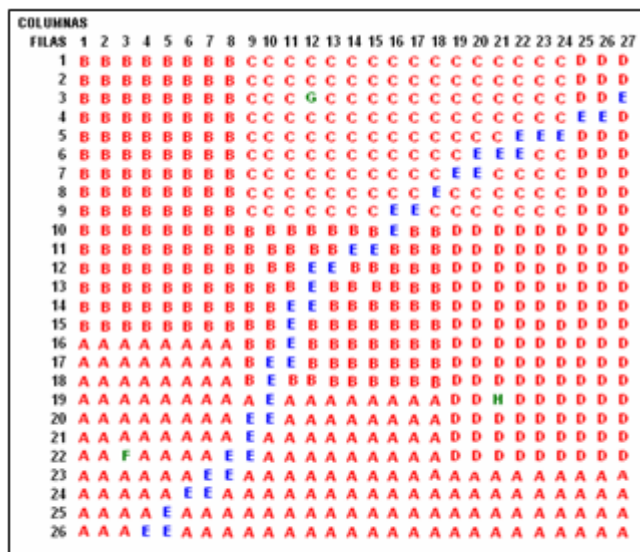


Figura 8. Fichero Enumeración Exhaustiva (702 caracteres).

▪ **Estructuras con compresión.**

En un mapa ráster es habitual que la información digital de un pixel se repita en pixeles contiguos. Esta situación permite compactar el almacenamiento de esa información reduciendo considerablemente el tamaño de fichero original. A continuación se exponen estructuras de almacenamientos que hacen uso de esta técnica de compactación.

▪ **Codificación Chain Code.**

Método ideado para la codificación de elementos superficiales y regiones compactas irregulares. Las fronteras de las diferentes regiones se codifican partiendo de un origen y una secuencia de vectores unitarios en las direcciones cardinales. Estas direcciones se numeran (Este=0, Norte=1, Oeste=2, Sur=3) y a continuación se recorre la región partiendo del punto origen en el sentido de las agujas del reloj. Cada dirección (0,1,2,3) va acompañada de un superíndice que indica el número de pixeles frontera en esa dirección. Continuando con el ejemplo de la figura 7, la región A quedaría codificada según el método chain code de la siguiente manera comenzando por la columna 1 fila 16:

$$\text{Región A-C1F16: } 0^8, 3^5, 0^8, 0^{10}, 0^8, 3^4, 0^8, 3^9, 0^8, 3^4, 0^8, 2^7, 0^8, 2^{11}$$

▪ **Codificación Run-Length.**

Este sistema recoge en cada fila el valor temático existente en el pixel y el número de columnas en las que se va repitiendo (Tabla 9). Se parte del primer pixel (columna 1) y se indica su valor y el número total de columnas en las que se repite recorriendo la fila hacia la derecha. A continuación se pasa al siguiente valor, y así sucesivamente hasta completar la fila.

La compactación del método *run-length* mejora considerablemente el almacenamiento de la información ráster, especialmente en aquellos casos en los que el volumen total de datos suponga una limitación del sistema. Por otro lado, se precisan realizar procesos que descompriman la información ralentizando el tratamiento de información cartográfica.

Es un método idóneo para trabajos con áreas extensas y uniformes. En cambio no resulta útil para la codificación de fenómenos con variaciones continuas en las que cada pixel contiene un valor único y diferente a sus contiguos.

B	8	C	24	D	27				
B	8	C	24	D	27				
B	8	C	11	G	12	C	24	D	26
B	8	C	24	E	26	D	17		
B	8	C	21	E	24	D	27		
B	8	C	19	E	22	C	24	D	27
B	8	C	18	E	20	C	24	D	27
B	8	C	17	E	18	C	24	D	27
B	8	c	15	E	17	C	24	D	27
B	15	E	16	B	18	D	27		
B	13	E	15	B	18	D	27		
B	11	E	13	B	18	D	27		
B	11	E	12	B	18	D	27		
B	10	E	12	B	18	D	27		
B	10	E	11	B	18	D	27		
A	8	B	10	E	11	B	18	D	27
A	8	B	9	E	11	B	18	D	27
A	8	B	9	E	10	B	18	D	27
A	9	E	10	A	18	D	20	H	21
A	8	E	10	A	18	D	27		
A	8	E	9	A	18	D	27		
A	2	F	1	A	7	E	9	A	18
A	6	E	8	A	27				
A	5	E	7	A	27				
A	4	E	5	A	27				
A	3	E	5	A	27				

Tabla 9. Fichero Run-Length (226 caracteres).

- **Estructura lineal comprimida de sentido único**

La codificación “Run-Length” requiere realizar un barrido de la imagen línea a línea para comprimir/descomprimir la información del fichero ráster. La modificación que ofrece la compresión lineal de sentido único permite empalmar los valores finales de una línea con la siguiente, todo ello representado en una única línea. Cuando se produce ese caso, se reduce considerablemente el tamaño del fichero.

En el ejemplo de la tabla 9 el fichero finalizaría comprimiendo el final de la fila 23 (A 27) con el inicio de la fila 24 (A 5) con un único valor (A 32). Lo mismo sucede con las filas 24-25, 25-26. El resultado de esta compresión con el ejemplo de la tabla 9 sería (en negrita la filas comprimidas):

B8 C24 D27 B8 C24 D27 B8 C11 G12 C24 D26 E27 ...A6 E8 **A32** E7 **A31** E5 A30 E5 **A27**

- **Estructura lineal comprimida de doble sentido o en greca**

Similar a la codificación “Run-Length” pero en este caso la imagen se barre enlazando el final de la primera línea con el final de la segunda y se avanza hacia el inicio de ésta. Posteriormente se enlaza con el inicio de la tercera y se avanza hacia su final para enlazar con el final de la cuarta, y así sucesivamente. Este método no provoca saltos temáticos al pasar de una línea a la otra ampliando la capacidad de compactación.

El ejemplo de la tabla 10 sería (en negrita la filas comprimidas):

B8 C24 **D54** C24 **B16** G11 C24 D26 E27 D17 C26 B24 **B16**...E8 **A54** E7 **A9** E5 **A54** E5 A3

- **Ordenación N de Morton.**

Se realiza un barrido de la imagen mediante la secuencia de la figura 8 aumentando la correlación espacial entre pequeñas áreas contiguas de la imagen y haciendo mínima el tamaño de almacenamiento. La secuencia de exploración minimiza el número de saltos y la dota de una estructura de ordenación jerárquica. Esta ordenación es aplicable únicamente en imágenes ráster cuadradas de tamaño N, potencia de 2.

- **Estructura Hilbert**

En este caso la imagen se barre mediante la curva de Hilbert, cuya numeración se muestra en la figura 9.

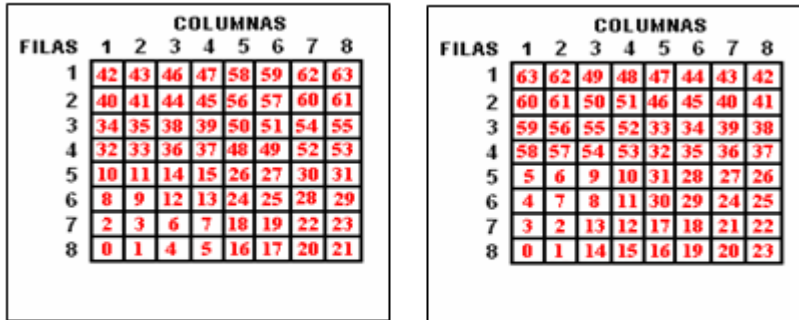


Figura 9. Estructuras N-Morton y Hilbert.

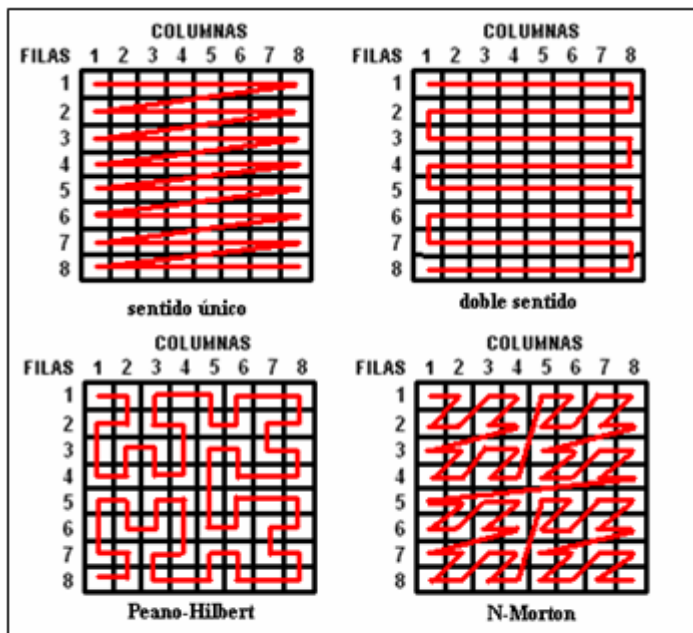


Figura 10. Barridos sentido único, doble sentido, Hilbert y Morton

- **Codificación por Bloques**

El método *run-length* codifica los datos en una única dimensión a nivel de fila. Esta idea se puede extender a las dos dimensiones del ráster, fila y columna, descomponiendo la matriz original en bloques cuadrados. La estructura de datos consta de 3 números, el origen del cuadrado expresado a través del par de coordenadas de uno de sus píxeles (el central o el inferior izquierdo) y el radio del bloque cuadrado.

La figura 11 muestran este procedimiento para el caso de la región A. Esta región se registra con 10 bloques simples, 8 bloques de 4 pixeles y 1 un bloque de 16 pixeles. Se necesitan 28 números para los orígenes de cada bloque (19 bloques por las 2 coordenadas fila y columna), más los 3 números del tamaño del bloque, en total 31 números.

La región A la componen:

$$1 [(10,2);(9,4);(10,3);(10,4); \dots; (15,10)] + 4 [(12,3); (10,5);(10,7); \dots; (14,9)] + 16 [(16,5)]$$

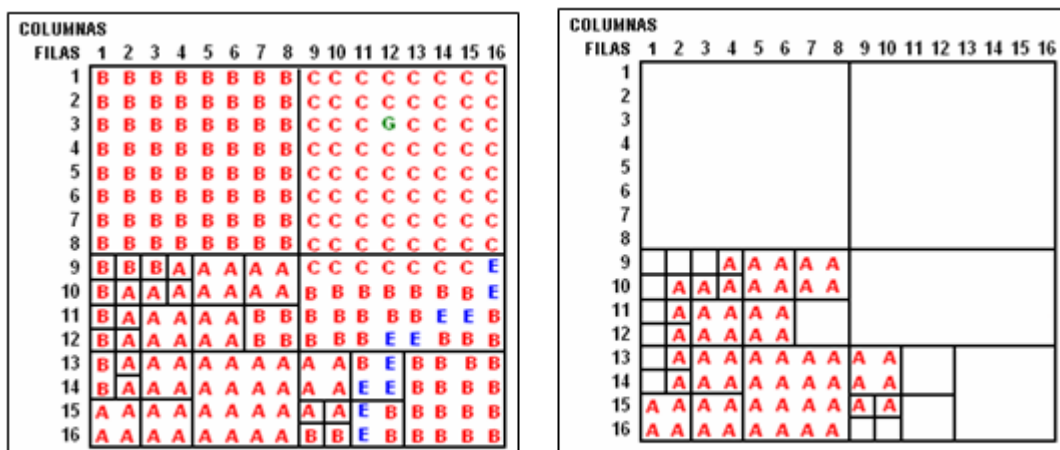


Figura 11. Codificación por bloques

Árboles binarios y Quadrees

Los métodos de compactación más efectivos para ficheros ráster con información geográfica están basados en la sucesiva división jerárquica de $2^n \times 2^n$. Si la división se realiza dividiendo el área por la mitad, se denomina árbol binario. Si la división se realiza en cuadrantes y se observa que cada cuadrante pertenece a la misma región, se denomina *quadtree*, siendo el método más utilizado.

La figura 12 muestra la división sucesiva de la zona A en cuadrantes. Esta estructura queda descrita por un árbol de grado cuatro conocido por *quadtree*, representado en la figura 13.

COLUMNAS		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
FILAS	1	1								2							
	2	1								2							
	3	1								2							
	4	1								2							
	5	1								2							
	6	1								2							
	7	1								2							
	8	1								2							
9	3	4	7	8	A	A	A	A	16	17	28						
10	5	6	9	10	A	A	A	A	16	17	28						
11	11	12	A	A	A	A	A	15	18	19	28						
12	13	14	A	A	A	A	A	15	18	19	28						
13	20	21	A	A	A	A	A	A	A	A	29	30	36				
14	22	23	A	A	A	A	A	A	A	A	29	30	36				
15	A	A	A	A	A	A	A	A	A	A	31	32	36				
16	A	A	A	A	A	A	A	A	A	A	31	32	36				

Figura 12. Codificación *quadrees* de la zona A.

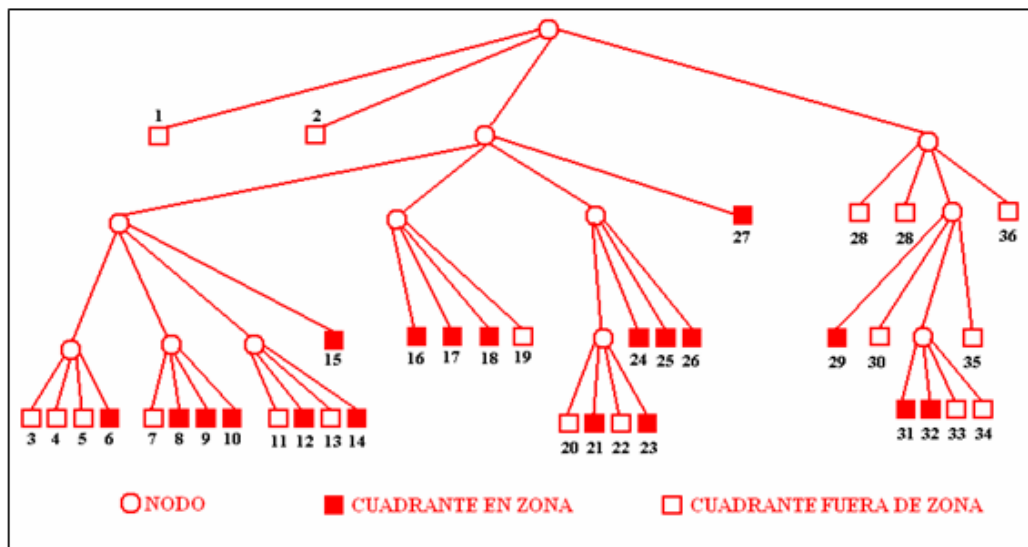


Figura 13. Jerarquía *quadrees* de la zona A.

La imagen ráster corresponde al nodo raíz del árbol y la altura del árbol es de 4 niveles. Cada nodo tiene 4 hijos. Los nodos hoja son aquellos nodos que no tienen subdivisión.

La estructura *quadtree* presenta ventajas respecto a otras estructuras, agiliza cálculos sobre las zonas, ofrecen una resolución variable representando únicamente aquellos detalles característicos de dicha resolución y permite separar zonas e incluir agujeros. Este método es variante en la traslación siendo un problema característico de estas estructuras, es decir, dos áreas con la misma forma y el mismo tamaño pueden tener árboles *quadtree* diferentes

haciendo. Este inconveniente provoca que análisis de forma y correlaciones de patrones sean procesos irreversibles, especialmente en áreas que presentan continuos cambios en el tiempo.

8.4. Ventajas e inconvenientes de los modelos vector/ráster.

Comparando ambos métodos de representación digital de datos geográficos se extraen una serie de ventajas y desventajas entre ellos.

El modelo vectorial necesita menos espacio de almacenamiento que la técnica ráster mediante altas tasas de compactación. La representación de las entidades geográficas es más precisa sobre todo en las entidades lineal muy pequeño o entidades puntuales. También permite medir distancias, superficies y volúmenes con alto grado de precisión. Soporta la descripción topológica de las entidades permitiendo el análisis de redes. Es posible aplicar transformaciones de coordenadas y cambio de datum con facilidad. Además soporta la actualización, la búsqueda y la generalización de entidades.

Las desventajas del modelo vectorial son el elevado consumo de recursos durante la visualización y la impresión de la información, especialmente en alta resolución, el análisis espacial es más lento, la captura vectorial de datos requiere más medios y más tiempo, además de la complejidad de la estructura de los datos.

El modelo ráster presenta estructura de datos simples, la localización y manipulación de valores temáticos concretos es más sencilla, soporta análisis espaciales y filtros específicos, todos los elementos tienen la misma forma elemental y regular (el píxel) facilitando el modelado matemático. Además la tecnología es más barata.

Su principal desventaja es el volumen de información que aumenta considerablemente si se quiere obtener una representación muy precisa. El uso de celdas de mayor tamaño reduce el tamaño de los ficheros suponiendo una pérdida de resolución espacial y de información. Tampoco reconoce explícitamente la existencia de objetos geográficos, limitando las operaciones de medición de distancias y superficies. Cualquier transformación de coordenadas consume muchos recursos cuyos resultados provocan distorsiones en los píxeles.

8.5. Orientación a Objetos

Existen una serie de definiciones de vital importancia para comprender la concepción de orientación a objeto (OO), término complejo aplicado a lenguajes de programación, bases de datos, metodologías de programación o algoritmos matemáticos.

Un objeto, es una entidad provista de una serie de propiedades (atributos) dotada de cierta funcionalidad o comportamiento (métodos) creada como instancia de una clase.

La clase, es la matriz en la que se definen las propiedades y el comportamiento de todos los objetos que a partir de ella se creen.

Los atributos, son los valores que describen el estado y las características de un objeto identificándolo de forma única y diferenciándolo del resto de objetos de la misma clase.

Los métodos, son las formas que tiene un objeto de interactuar con el resto del mundo modificando sus propios atributos, o bien, avisando al resto de objetos (incluso de diferentes clases) de ciertos mensajes.

Las características de la orientación a objetos son:

- Abstracción: cada objeto en el sistema sirve como modelo de un "agente" abstracto que puede realizar trabajo, informar y cambiar su estado, y "comunicarse" con otros objetos en el sistema sin revelar cómo se implementan estas características. Los procesos, las funciones o los métodos pueden también ser abstraídos y cuando lo están, una variedad de técnicas son requeridas para ampliar una abstracción.
- Polimorfismo: diferentes comportamientos de objetos distintos pueden tener el mismo nombre. En función del objeto que se esté utilizando se invocará un método u otro.
- Encapsulamiento: un objeto es una entidad indivisible y cuando se crea, encierra dentro de él todas las propiedades y métodos que ha heredado de su clase. Se maneja como un todo y en cualquier momento se puede analizar su estado a través de sus propiedades.
- Herencia: las clases se relacionan unas con otras formando una estructura jerarquizada. Los objetos heredan las propiedades y los métodos de las clases a la que pertenecen. Una clase hija puede heredar el comportamiento y la funcionalidad de otra clase padre, o de otras muchas clases padre (herencia múltiple). La herencia organiza y facilita el polimorfismo y el encapsulamiento.

Bibliografía

- [1] Bosque Sendra, J. (1992) *Sistemas de Información Geográfica*. Ediciones Rialp. ISBN 84-321-2922-4
- [2] Goodchild, M.F. (1984) *Geocoding and Geosampling. Spatial Statistics and models*.
- [3] Goodchild, M.F., Maguire, D. y otros (2001) *Geographic Information Systems and Science*. Editorial Wisley. ISBN 0-471-89275-0
- [4] Moldes Teo, F.J. (1995) *Tecnología de los Sistemas de Información Geográfica*. Editorial Ra-Ma. ISBN 84-7897-164-5
- [5] Burrough, P.A., McDonnell, R.A. (2000). *Principles of Geographical Information Systems*. Editorial Oxford. ISBN 0-19-823366-3
- [6] Kresse, W, Fadaie, Kian. (2004). *ISO Standards for Geographic Information*. Editorial Springer. ISBN 3-540-20130-0

Tema 9. Formatos de intercambio de datos geográficos. El estándar Geographic Markup Language (GML)

9.1. Introducción

La extensión en el nombre de un fichero nos da una idea normalmente del formato en el que se ha almacenado su contenido. Muchos de estos formatos pueden ser:

- no específicos para la información geográfica
- propietarios, se necesita un programa específico para poder tratarlo.
- no preparado para el transporte de la información a través de Internet
- no estructuran la información

Si se necesita intercambiar datos geográficos a través de la Web de una manera estandarizada que permita:

- Codificar la información geográfica de manera que puedan visualizarse con la resolución que se precise.
- Almacenar el contenido del mapa (posición de los fenómenos, su geometría, tipos y atributos) pero no como estos datos deben ser representados. El estilo de representación es personalizado.
- Que el formato sea editable, de manera que se puedan realizar anotaciones sobre la información geográfica
- Posibilidad de introducir enlaces asociados a los fenómenos, mediante URL u otros tipos.
- Posibilidad de interrogar un fenómeno para poder obtener la información asociada al fenómeno.
- Tener control sobre los contenidos, mediante la aplicación de operaciones de filtrado sobre los fenómenos, para poder diferenciar unos de otros.
- Permitir fenómenos animados
- Permitir el encadenamiento de servicios, como por ejemplo realizar una transformación de sistema de referencia y después poder visualizarlos.

El formato que cumple todas estas características es el formato *Geographic Markup Language*: GML.

GML es un lenguaje basado en XML que proporciona un vocabulario especializado para trabajar con datos geoespaciales. GML permite la codificación de fenómenos geográficos y permite la modelización, almacenamiento y transporte de los datos brutos de los fenómenos como puntos, líneas, atributos, objetos temporales y coberturas.

Un fenómeno geográfico es una abstracción de un fenómeno del mundo real; un fenómeno geográfico esta asociada a una localización de la Tierra, por ejemplo una carretera, un límite administrativo, un río. El estado de un fenómeno se define por un conjunto de propiedades, donde cada propiedad puede estar definida por su nombre, tipo y valor.

El número de propiedades que un fenómeno debe tener, junto con nombre y tipo, están determinadas por su tipo de definición. Los fenómenos geográficos espaciales, es decir con geometría, son aquellas cuyas propiedades son valores geométricos, por ejemplo: puntos, líneas, curvas o áreas. En GML también se admiten fenómenos sin geometrías, pero no serán fenómenos geográficos. Una colección de fenómenos es un conjunto de fenómenos el cual puede ser considerado como un fenómeno en sí y tiene sus propias propiedades, por ejemplo una ciudad.

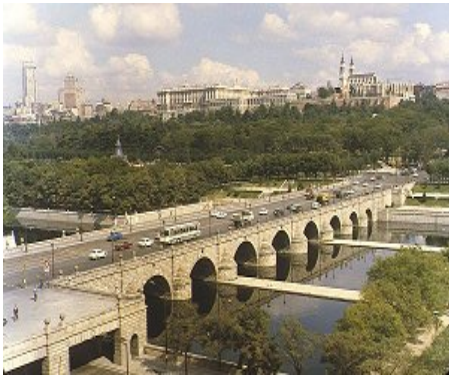
Utilizando GML, se puede codificar la descripción la localización o extensión espacial de un río, carretera o puente. Un registro GML incluye una geometría que puede representar un puente o una carretera, así como información de apoyo a la geometría como en número de carriles de la carretera.

Como cualquier codificación XML, GML representa la información geográfica en formato texto. Al ser texto, le da simplicidad y es visible por si mismo, fácil de revisar y de cambiar. Por tanto, GML representa la información geográfica por medio de texto, para representar geometrías y objetos geográficos. No representa los fenómenos geográficos de manera gráfica.

Con GML se puede transferir las descripciones del fenómeno geográfico a través de Internet, y controlar cómo los fenómenos son mostrados. GML separa el contenido de la presentación; en otras palabras, GML ofrece el dato geográfico real (información sobre las propiedades y geometría de objetos en un mundo físico) sin tener en cuenta y de modo independiente a la

visualización de ese dato (colores, ancho de las líneas, símbolos, etc.). Mientras que los mapas son una de las presentaciones más comunes de GML, GML no es dependiente de una especificación gráfica en particular.

Se puede decir que GML es una aplicación práctica para transferir información cartográfica por la Web. Además de transmitir un mapa en formato GIF o JPG, GML permite que los datos sean controlados en el navegador por el usuario, quien decide cómo serán mostrados los datos.



Contenido del fenómeno: puente

- Nombre: Puente de Segovia
- Posición: $X_1, Y_1, X_2, Y_2, \dots, X_n, Y_n$,
- Longitud: 240 m
- Altura: 21 m
-

Figura 1. Ejemplo de un fenómeno geográfico: Puente

GML fue desarrollado por OpenGIS Consortium, ahora, Open GeoSpatial Consortium (OGC) para definir la geometría y las propiedades de los objetos que comprenden la información geográfica.

9.2. Definición de GML

Geographic Markup Language (GML) es una codificación XML para el transporte, y almacenamiento de información geográfica que modela según el esquema de modelización conceptual utilizado en la serie de Normas Internacionales ISO19100, incluyendo tanto las propiedades espaciales como las no espaciales de los fenómenos geográficos.

El GML define la sintaxis del XML Schema, mecanismos y convenciones que:

- Proporcionan un marco de trabajo abierto y neutral de la definición de objetos y aplicaciones de esquemas geoespaciales.

- Proporcionan un conjunto de clases de objetos para describir elementos geográficos como fenómenos, sistemas de referencia espaciales, geometrías, topologías, tiempo, unidades de medida y valores generales.
- La definición de GML se realiza utilizando esquemas (*schemas*). Los *schemas* de la especificación pueden ser personalizados para un modelo de datos determinado, mediante extensión o especialización del mismo.
- Con GML la integración de datos espaciales (geométricos) y no espaciales es fácilmente realizable, especialmente en casos donde los datos no espaciales están codificados en XML y además permite separar el contenido de su representación.
- GML es independiente de plataformas y aplicaciones propietarias.
- Soporta el almacenamiento y el transporte de esquemas de aplicación y conjunto de datos.
- Incrementa la capacidad de organizaciones para compartir esquemas de aplicación geográfica y de información geográfica.

Se puede decir que es un sublenguaje de XML descrito como una gramática en XML Schema para el intercambio de la información geográfica. Su importancia radica en que a nivel informático se constituye como un lenguaje lengua franca para el manejo y trasvase de información entre los diferentes software que hacen uso de este tipo de datos, como los Sistemas de Información Geográfica.

GML se diseñó a partir de la especificación abstracta producida por el grupo *Open Geospatial Consortium* y de la serie de documentos ISO 19100. GML no contiene información específica sobre como se debe hacer la visualización de los datos representados. Para ello se utilizan estilos que se relacionan a GML.

9.3. GML Esquema.

GML proporciona objetos para describir geográficamente, incluyendo lo fenómenos, los sistemas de referencia de coordenadas, geometría, topología, tiempo, unidades de medida y los

valores de generalización. Un documento GML es un fichero de texto codificado XML que por si mismo no hace nada.

Existe un esquema GML para tres categorías de objetos:

1. Un **fenómeno** geográfico es una abstracción de un fenómeno del mundo real; es un fenómeno geográfico si esta asociado con un localización relativa a la Tierra. Una representación digital del mundo real puede entonces ser a través de un conjunto de fenómenos. El estado del fenómeno es definido mediante un conjunto de propiedades, donde cada propiedad tiene un nombre, tipo y valor.
2. Una **cobertura** puede representar un fenómeno o una colección de fenómenos para mostrar las relaciones espaciales entre fenómenos terrestres así como su distribución espacial.
3. Una **observación** modeliza el acto de observación, o reconocimiento. Una observación es considerada que es un fenómeno GML con un tiempo y con un valor.

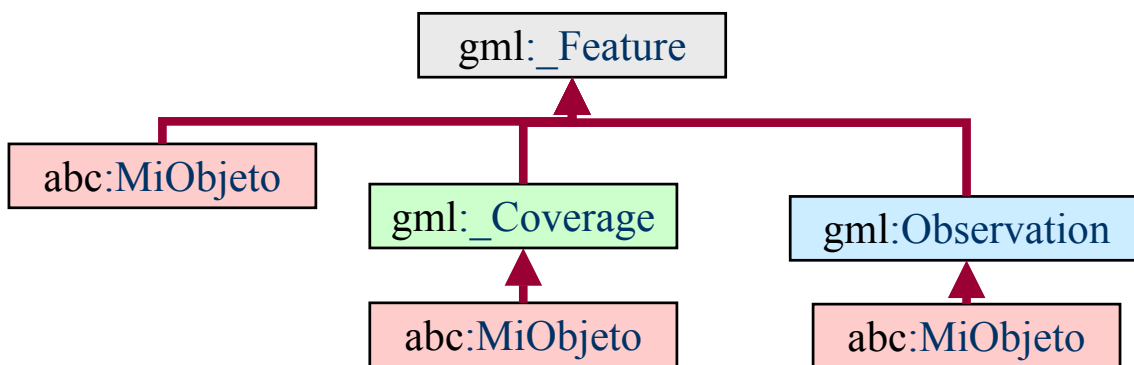


Figura 2: Tipos de objetos GML

El marco de trabajo de GML es muy extensible: se pueden crear y definir tipos de fenómenos relacionados específicamente para un área de especialización.

El GML esta normalizado a través de la norma internacional ISO19136 que especifica la codificación XML de clases conceptuales definidas en las normas internacionales de la serie ISO 19100 y en las especificaciones de OGC. Los modelos conceptuales están definidos en:

- ISO/TS 19103 – *Conceptual schema language* (unidades de medida)
- ISO 19107 – *Spatial schema* (geometría especial y topología)
- ISO 19108 – *Temporal schema* (geometría temporal y topológica y sistema de referencia temporal)

- ISO 19109 – *Rules for application schemas* (fenómenos)
- ISO 19111 – *Spatial referencing by coordinates* (sistema de referencia de coordenadas)
- ISO 19123 – *Schema for coverage geometry and functions* (coberturas y grids)

Además de estas normas, GML proporciona codificación XML para conceptos adicionales que no están modelados ni en la serie de normas internacionales ISO19100, ni en las Especificaciones de OGC.

9.4. Geografía, gráficos y mapas

Es importante aclarar las distinciones entre datos geográficos (que son codificados en GML) y la interpretación gráfica de esos datos que podría aparecer en un mapa o en otro formato de visualización. Los datos geográficos están relacionados con una representación del mundo en términos espaciales y es independiente de una determinada visualización de esos datos. Cuando se habla de datos geográficos se trata de capturar la información sobre las propiedades y geometría de los objetos del universo del problema. Cómo se simbolice esto en un mapa, el color o el grosor de las líneas que se usen para representarlos es independiente del GML.

Esta propiedad de separar contenido de representación proviene de su derivación de XML.

Por supuesto, puede usarse GML para crear mapas. Esto podría ser obtenido mediante una herramienta de dibujo que interprete datos GML, como por ejemplo la herramienta *TatukGIS Viewer*. Por tanto, para crear un mapa a partir de GML sólo se necesita aplicar un estilo a los elementos de GML dentro de un formato que pueda ser interpretado por un navegador Web. Entre los formatos gráficos potenciales están W3C *Scalable Vector Graphics* (SVG), *Microsoft Vector Markup Language* (VML) y X3D.

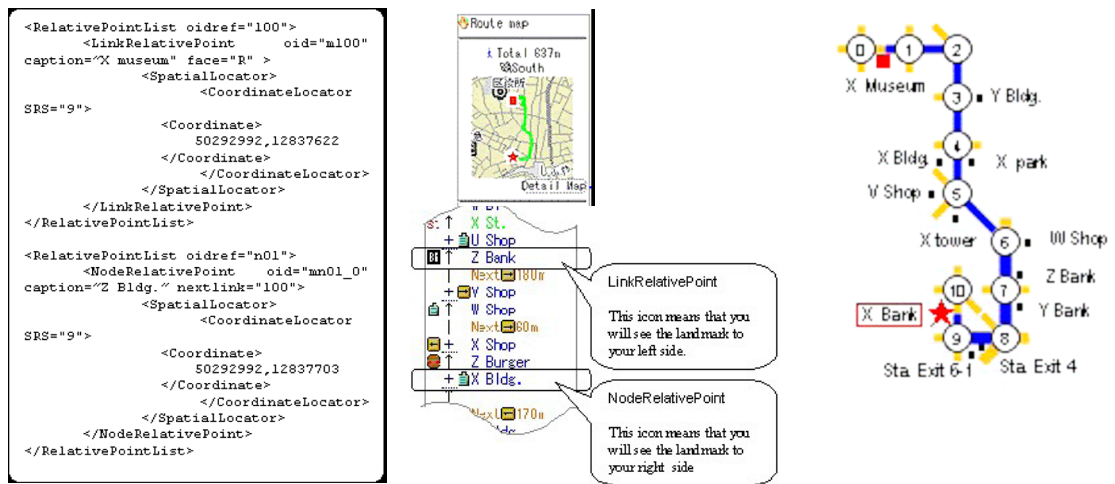


Figura 3: Diferentes formas de representación de la información geográfica

9.4.1. GML es texto

Como cualquier documento XML, GML representa la información en forma de texto. El texto aporta como ventaja su simplicidad y visibilidad. Es fácil de inspeccionar y cambiar. Y al ser XML facilita también que sea controlado.

Los formatos de texto para geometría y geografía han sido empleados anteriormente (SAIF format, por ejemplo). Otro ejemplo del uso de texto para conjuntos de datos geométricos complejos es VRML (*Vector Markup Language*). Se han construido modelos VRML grandes y complejos, que han sido accedidos a través de la Web, usando una codificación basada en documentos de texto.

9.4.2. GML codifica fenómenos geográficos y propiedades

GML se basa en el modelo geográfico abstracto desarrollado por el OGC. Este modelo describe el mundo en término de fenómenos geográficos. Esencialmente, un fenómeno geográfico no es más que una lista de propiedades y geometrías.

Las propiedades tienen normalmente un nombre, un tipo, un valor y una descripción. Las geometrías están compuestas de geometrías básicas tales como puntos, líneas, curvas, superficies y polígonos. Por simplicidad, la especificación GML inicial está restringida a geometría 2D. De todas formas, extensiones que gestionen geometrías 2 ½ y 3D pueden aparecer próximamente, así como las relaciones topológicas entre fenómenos.

GML ya permite la creación de fenómenos complejos. Un fenómeno geográfico puede, por ejemplo, estar compuesto de otros fenómenos. Un único fenómeno como un aeropuerto podría estar compuesto por otros fenómenos como carril-bus, autovía, hangares y terminales. La geometría de un fenómeno geográfico también puede estar compuesta de muchos elementos geométricos. Un fenómeno con geometría compleja puede estar constituido de la mezcla de tipos de geometría como puntos, líneas y polígonos,

Para codificar la geometría de un fenómeno como un edificio simplemente se escribe:

```
<Feature fid="142" featureType="colegio" Description="Un colegio">
  <Polygon name="extent" srsName="epsg:27354">
    <LineString name="extent" srsName="epsg:27354">
      <CDATA>
        491888, 5458045, 491904, 5458044. 491908, 5458064, 491924, 5458064, 491925,
        5458079, 491977, 5458120, 491953, 5458017
      </CDATA>
    </LineString>
  </Polygon>
</Feature>
```

Un fenómeno no tiene propiedades no geométricas. Pero podrían añadirse de la siguiente forma:

```

<Feature fid="142" featureType="colegio" >
  <Description>Pablo Ruiz Picasso</Description>>
  <Property Name="NumeroPlantas" type="Integer" value="3"/>
  <Property Name="NumeroEstudiantes" type="Integer" value="987"/>
  <Polygon name="extent" srsName="epsg:27354">
    <LineString name="extent" srsName="epsg:27354">
      <CDATA>
        491888, 5458045, 491904, 5458044, 491908,5458064, 491924,5 458064, 491925,
        5458079, 491977, 5458120, 491953,5458017
      </CDATA>
    </LineString>
  </Polygon>
</Feature>

```

A continuación se presenta el diagrama UML que muestra la jerarquía de clases GML:

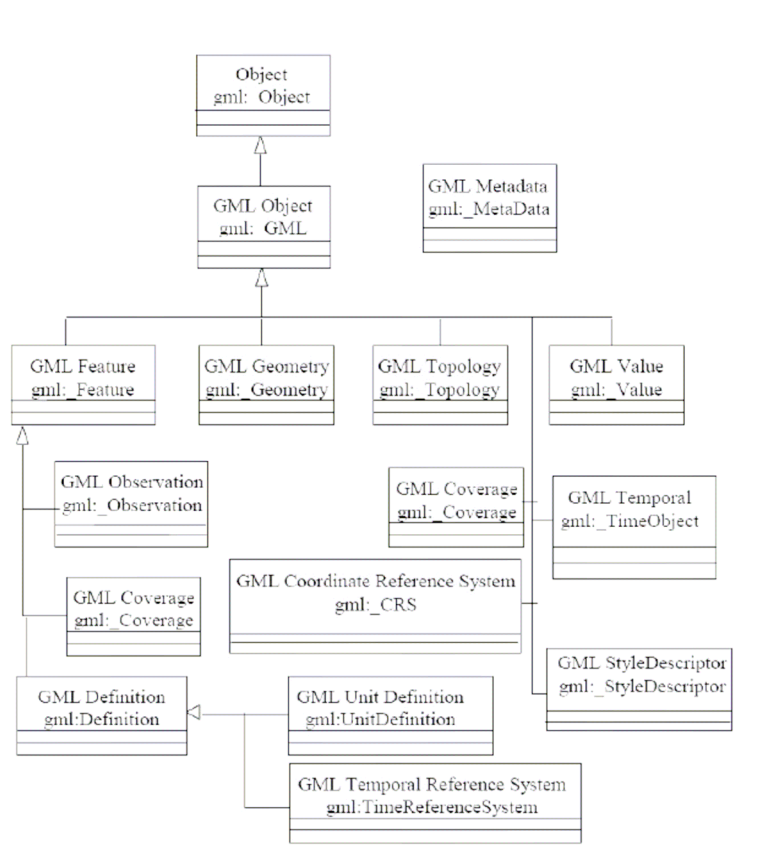


Figura 4: Diagrama UML que muestra la jerarquía de la clase GML

La clase GML_Feature, como ya se ha explicado, puede ser una cobertura, una observación o un fenómeno geográfico.

La clase GML_Geometry, es decir, el modelo de geometría de GML, sigue las especificaciones de OGC que son idénticas a ISO DIS 19107, que proporciona:

- Geometrías de 0, 1, 2 y 3 dimensiones.
- Curvas compuestas, superficies y sólidos.
- Primitivas de curvas (segmentos).
- Primitivas de superficies
- Coordenadas

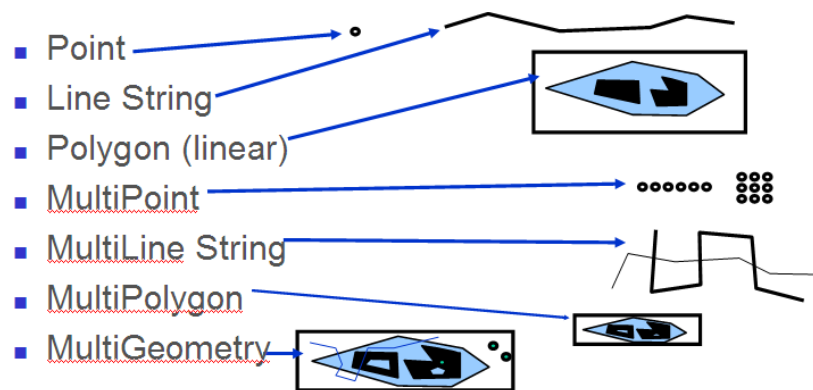


Figura 5: Tipos de geometría en GML

La clase GML_Topology, describe la correspondencia topológica y las relaciones geométricas hasta 3 dimensiones, proporcionando objetos de topología estándar:

- Nodos, aristas, caras y sólidos.
- Límites y relaciones de adyacencia con orientación.
- Topología compleja.
- Expresiones topológicas.

9.4.3. GML codifica sistemas de referencia espaciales

Un componente esencial de un sistema geográfico es un medio para referenciar los fenómenos geográficos a la superficie terrestre o a alguna estructura relativa a la misma. La versión actual

de GML incorpora un sistema de referencia espacial que es extensible y que incorpora los principales marcos de referencia geocéntricos y proyecciones de hoy en día.

Esto permite codificar todos los sistemas de referencia que pueden encontrarse en el portal Web del *European Petroleum Standards Group* (EPSG). Además los esquemas codificados permiten al usuario definir unidades y parámetros de los sistemas de referencia. Probablemente, versiones futuras de GML proveerán mayor flexibilidad para gestionar sistemas de coordenadas locales.

¿Por qué codificar un sistema de referencia espacial y no simplemente proporcionar un único nombre? En muchos casos tal aproximación es suficiente y GML no requiere que el proveedor de datos geográficos también proporcione una codificación de los sistemas de referencia para poder georreferenciar los datos. Pero hay casos, donde tal información es muy valiosa e incluye:

- Validación del cliente de un SRS del servidor. Un cliente puede solicitar una descripción del SRS (un documento XML) y compararlo con sus propias especificaciones o mostrárselas al usuario para su verificación.
- El cliente muestra el SRS específico del servidor.
- Usado por un Servicio de Transformación de Coordenadas para validar el SRS de una fuente de datos.
- Un Servicio de Transformación de Coordenadas puede comparar la descripción del SRS con sus propias especificaciones para ver si el SRS es consistente con la transformación seleccionada.
- Control automático de la transformación de coordenadas al añadir como entrada y salida los nombres de los SRS y los valores de los parámetros.

Con la codificación GML para sistemas de referencia, es posible crear un sitio Web en el que se almacene cualquier número de definiciones de sistemas de referencia espacial.

9.4.4. GML: Colecciones de Fenómenos Geográficos

La recomendación XML 1.0 del W3C se basa en la noción de documento. La versión actual de GML está basada en XML 1.0, y usa una *FeatureCollection* como base de sus documentos.

Una *FeatutreCollction* es una colección de *Features* (fenómenos) on un *Envelope* (entorno) (que limita el conjunto de *Features*), una colección de *Properties* (propiedades) que se aplican

al *FeatureCollection* y una lista opcional de definiciones de Sistemas de Referencia Espaciales. Una *FeatureCollection* también puede contener otras *FeatureCollections*, siempre que el *Envelope* de la *FeatureCollection* englobe los *Envelopes* de todos los *FeatureCollections* contenidos.

Cuando se realiza una petición de datos GML a un servidor GML, los datos siempre son devueltos en *FeatureCollections*. No hay límite en el GML RFC del número de fenómenos que pueden ser contenidos en un *FeatureCollection*. Debido a que un *FeatureCollection* puede contener otros *FeatureCollections* es un proceso relativamente simple unir los *FeatureCollections* recibidos desde un servidor en una gran colección.

9.4.5. GML: Más que un formato de transporte de datos.

Aunque GML es un medio efectivo para transportar información geográfica de un lugar a otro, se espera que también llegue a ser un importante medio de almacenamiento de información geográfica. El elemento clave aquí es *XLink* y *Xpointer*. Estas dos especificaciones mantienen grandes expectativas para crear construcciones complejas y conjuntos de datos distribuidos.

Los datos geográficos son, obviamente, eminentemente geográficos. Están distribuidos naturalmente sobre la superficie de la Tierra. Al mismo tiempo, hay aplicaciones que necesitan un mayor alcance y obtener datos de una base global para análisis de escalas grandes o por el interés en un dominio vertical más estrecho. Las aplicaciones de este último tipo también abundan en una colección diversa de campos desde protección medioambiental hasta minería, construcción de carreteras y gestión de desastres. Sería muy útil que los datos fueran desarrollados en la escala local e integrados a escala regional y local.

La mayoría de las veces, los datos geográficos son obtenidos por agencias particulares para un propósito específico. Las oficinas forestales recolectan información de la disposición de los árboles (diámetro, condiciones del hábitat, ratios de crecimiento) para la gestión efectiva del comercio forestal. Los departamentos de medioambiente recolectan información sobre la distribución de animales y su hábitat. El departamento de Desarrollo está interesado en mantener información sobre demografía y fenómenos existentes en el entorno de las construcciones. Los problemas del mundo real raramente se ciñen a los límites de los departamentos agencias y oficinas. Sería muy útil si los datos generados para un propósito pudieran ser integrados con datos generados con otro fin.

GML como formato de almacenamiento, combinado con *XLink* y *XPointer* proporcionarán algunas contribuciones muy útiles para tratar esos problemas.

9.5. Las tecnologías de las que depende GML

GML está basado en XML. XML, aunque algunas veces es mencionado como un sustituto de HTML, es mejor pensar en él como un lenguaje para la descripción de datos. Más correctamente, XML es un lenguaje para expresar lenguajes de descripción de datos. XML, de todas formas, no es un lenguaje de programación. No hay muchos mecanismos en XML para expresar comportamiento o operaciones de computación. Esto se deja a otros lenguajes como Java y C++.

9.5.1. XML Versión 1.0

XML 1.0 proporciona un medio de describir (etiquetado) datos usando etiquetas definidas por el usuario. Cada segmento de un documento XML está limitado por una etiqueta de inicio y una etiqueta de fin. Esto se expresa del siguiente modo:

```
<Feature>
    .... Más descripciones XML ...
</Feature>
```

La validez de los nombres de las etiquetas queda determinada por el esquema XML (archivos .xsd). Qué etiquetas puede aparecer dentro de un par de etiquetas de inicio y fin también se determina por el esquema XML.

Las etiquetas XML también tienen atributos asociados. También hay restricciones en el esquema XML sobre el nombre y en algunos casos el dominio de valores que los atributos pueden tomar.

Normalmente XML es leído por un intérprete o *parser*. Todos los intérpretes XML chequean que los datos están bien formados de modo que no pueda haber una corrupción de datos (por ejemplo, una etiqueta de fin olvidada) que no sea detectada. Muchos intérpretes de XML también validan, es decir, chequean que el documento es conforme a un esquema XML asociado.

Usar XML para generar y validar estructuras de datos jerárquicos complejas es relativamente fácil. Tales estructuras son comunes en aplicaciones geográficas.

9.5.2. SVG, VML y X3D – Gráficos vectoriales para la Web

XML es utilizada en muchos ámbitos, uno de los cuales es el campo de los gráficos vectoriales. Se han desarrollado varias especificaciones basadas en XML para describir gráficos vectoriales, incluyendo *Scalable Vector Graphics (SVG)*, *Microsoft's Vector Markup Language (VML)*, y X3D, la encarnación XML de la sintaxis y comportamiento de VRML (*Virtual Reality Markup Language*).

Estas especificaciones tienen muchas similitudes con GML, pero tienen un objetivo muy diferente. Cada una de ellas tiene un medio de describir la geometría. Las especificaciones gráficas, de todas formas, están enfocadas sobre la apariencia y a partir de ahí incluyen propiedades y elementos para colores, grosor de líneas y transparencia, por nombrar unos pocos aspectos. Para ver un fichero de datos SVG, VML o X3D es necesario tener un visor de datos gráficos compatible. En el caso de VML está construido dentro de IE 5.0 (y en ningún sitio más). En el caso de SVG, Adobe ha desarrollado una serie de plug-ins para Internet Explorer y Netscape Communicator así como Adobe Illustrator, mientras que IBM y otras compañías, están, o ya tienen desarrollado, visores SVG o soportal librerías gráficas.

Para dibujar un mapa desde datos GML se necesita transformar GML a uno de los formatos de datos gráficos (SVG, VML o VRML). Esto significa asociar un “estilo” gráfico (es decir, símbolo, color, textura) con cada tipo de fenómeno GML o instancia.

El siguiente dibujo ilustra el dibujo de un mapa usando el estilo XSLT en un cliente de mapas compatible, haciendo cartografía con XSLT y SVG:

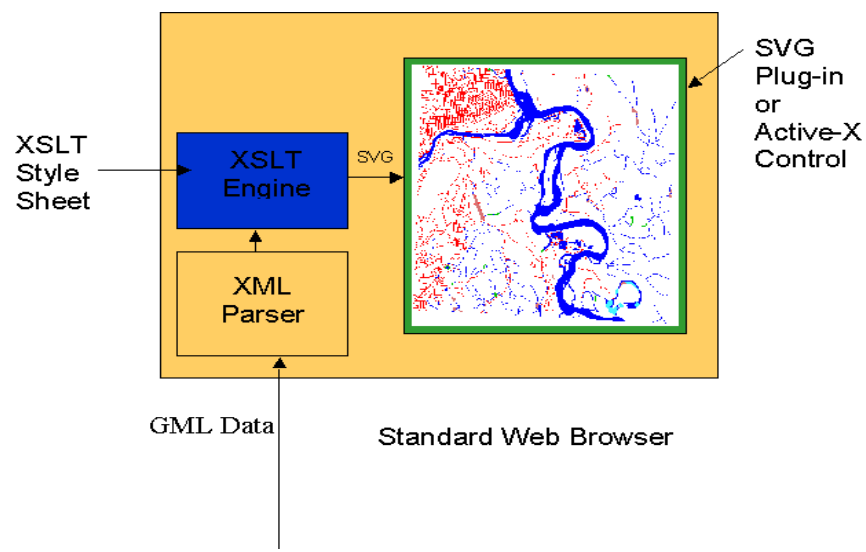


Figura 6. Ejemplo de implementación de XSLT a partir de XML

9.5.3. SVG, XLink y XPointer – Enlazando un sitio a otro

Con la tecnología HTML actual es posible construir conjuntos de datos geográficos enlazados. Uno puede construir mapas que son enlazadas a otros mapas. El mecanismo de enlazado HTML tiene, de todas formas, muchas limitaciones, y como resultado de esto, no es práctico construir conjuntos de datos distribuidos que sean grandes y complejos como ocurre en los sistemas del mundo real.

La mayor limitación significativa es que los enlaces HTML son codificados en el código tanto en la fuente () como en el documento objetivo, un hecho que hace a un sistema frágil e imposible de escalar. XLink evita estos problemas al permitir enlaces “fuera de línea”.

En los enlaces en línea y fuera de línea, las fuentes apuntan sólo a bases de datos enlazadas y es la base de datos enlazada la que proporciona el puntero al elemento específico del documento direccionado. El enlace no está codificado en ningún documento. Esto tiene una gran importancia para GML al hacer posible construir conjuntos de datos geográficos distribuidos y escalables. Incluso más importante es que XLink y XPointer permiten construir aplicaciones específicas indexadas por fechas. ¿Se necesita tener un grupo de edificios organizados por su dirección postal o se quiere crear un índice de parcelas de una granja basadas en el tipo de cosecha? Con XLink y XPointer, estos y otros muchos esquemas de índices pueden ser contruidos, y todos sin alterar la fuente de datos.

9.6. Ventajas y Desventajas de GML

9.6.1. Ventajas de GML

XML se ha convertido en la herramienta general de la descripción del dato y GML se ha convertido en el conductor tecnológico de la información geoespacial a través de la Web. GML ha estandarizado la forma de describir los fenómenos geográficos, por lo que no es necesario comprar un software SIG para visualizar los mapas: GML puede fácilmente transformarse a otros formatos como *Scalable Vector Graphics* (SVG) o *eXtensible 3D Graphics* (X3D).

A continuación una lista de razones explicativa de por qué es aconsejable utilizar GML:

- GML es texto y algunos de esos formatos también (ej. SAIF).
- GML está basado en un modelo común de datos geográficos, la Especificación Abstracta del OGC, acordada por gran número de fabricantes. Hay un consenso para convertirse en un estándar (ISO/DIS 19136).
- Lo más importante de GML: que está basado en XML y eso abre la puerta a todas las ventajas que ello conlleva.
- GML descansa sobre una amplia cantidad de estándares públicos que se han adoptado. Lo cual asegura que los datos GML pueden ser visualizados, editados y transformados por una gran variedad de herramientas comerciales y gratuitas.
- Hay una gran cantidad de profesionales de las Tecnologías de la Información que dominan XML y lo pueden aplicar al desarrollo de herramientas GML.
- Verificación automática de la integridad de los datos
- GML puede ser leído por herramientas públicas o genéricas
- GML puede ser fácilmente editado
- GML puede ser integrado fácilmente con datos No-Espaciales. Los formatos binarios dificultan este proceso porque necesitamos conocer las estructuras y ser capaces de modificarlos
- Mapas de mayor calidad
- Funcionan en los navegadores sin la necesidad de comprar software en el lado del cliente.
- Estilos de mapa personalizados. Separación de los datos y la presentación.
- Mapas editables basados en GML-SVG
- Mejores capacidades de consulta frente a los mapas raster
- Control sobre los contenidos. La selección y filtrado de la información
- Fenómenos animados

- No se tiene que pensar sólo en un navegador Web: dispositivos móviles, etc.
- Encadenamiento de servicios

9.6.2. Desventajas de GML

A continuación veamos una serie de desventajas del GML:

- El formato de GML es un fichero de texto, estos ficheros son de gran tamaño y por tanto la información no está optimizada desde el punto de vista del almacenamiento como si lo están los formatos binarios. Para ello hay dos soluciones:
 - Comprimir los archivos en formato zip para el transporte y almacenamiento, obligando a introducir en las aplicaciones mecanismos de compresión/descompresión. Por otro lado, son operaciones sencillas de implementar. Ratios de compresión 5:1 o superiores.
 - BinXML por ExpWay (expway.com). XML binario que es soportado tanto por API's nativas como abiertas (e.j. Sax, Dom). BinXML también da la posibilidad de streaming y expansión incremental del XML binario.
- Coste de la adaptación y formación a la tecnología GML.
- Dificultades para manejar ficheros raster de gran tamaño. Las fotografías aéreas y ortofotos se siguen almacenando en los formatos raster usuales (Tiff, etc). Aunque está previsto para las próximas actualizaciones poder trabajar con grandes ficheros raster, incluso vídeo.
- Es una tecnología de lenta implantación, que lleva desarrollada desde el año 2000 con su versión 1, que aún necesita popularizarse y convertirse en un estándar de facto. ¿Es éste el momento de comenzar a trabajar con GML?

9.7. Conclusiones

El GML está diseñado para la modelización, el transporte y almacenamiento de la información geográfica. Los esquemas predefinidos GML proporcionan un rico vocabulario que puede ser utilizado para crear esquemas de aplicación GML en dominios específicos. El GML sirve como una fundación para la Web geoespacial y para la interoperabilidad independientemente del desarrollo de las aplicaciones distribuidas incluyendo servicios basados en la localización.

GML es texto y esta basado en XML

Posee un modelo de datos basado en la Especificación abstracta de entidades del OGC y se revela como la tecnología idónea para la interoperabilidad y de las IDE porque:

- Está desarrollado sobre la tecnología que comienza a dominar la información en Internet.
- GML acumula ya una gran experiencia como modelo de información geográfica. No necesitamos reinventar nada. Podemos partir de GML y extender sus funcionalidades según nuestras necesidades o por contra, especializarlo utilizando sólo aquellas partes que vayamos a utilizar

El GML constituye una capa semántica sobre XML para expresar fenómenos geográficos y la norma ISO que normaliza el GML es la ISO 19136.

Bibliografía

- [1] ISO 19136 Geographic information - Geography Markup Language
- [2] OpenGIS® Geography Markup Language (GML) Implementation Specification
<http://www.opengeospatial.org/standards/gml>
- [3] Canadian Geospatial Data Infrastructure. CGDI Training Home
http://www.geoconnections.org/publications/training_manual/e/07/07_00/07_00_00.htm
- [4] http://mapas.topografia.upm.es/geoserviciosOGC/documentacion/GML/GML_en_la_IngGeografica_ver2005.pdf

Tema 10. Captura de datos geográficos. Métodos de captura: Digitalización. Escaneado. Fotogrametría. GPS. Trabajos de campo. Importación de datos. Captura de atributos.

10.1. Introducción

La Asociación para la Información Geográfica (www.agi.org.uk) define la Información Geográfica (IG) como “Información sobre objetos o fenómenos que están asociados a una posición sobre la superficie de la Tierra”.

Disponer de la IG adecuada ayuda en la toma de decisiones, pues de ésta dependen un sinnúmero de aplicaciones del ámbito social, económico y político. Algunos ejemplos son:

- Planificación de las prestaciones sanitarias.
- Gestión del litoral y defensa contra inundaciones.
- Órdenes de preservación de árboles.
- Observación por satélite del uso de la tierra.
- Previsiones geográficas para servicios de emergencia.
- Análisis de crímenes y disturbios.
- Orientación de vehículos y gestión de tráfico en general.
- Seguimiento en el reparto de paquetería.
- Planificación de la construcción.
- Prestación de agua, gas y electricidad a los usuarios.
- Uso de GPS para navegación y prestación de servicios.
- Transporte público: planificación y previsiones.
- Localización de puntos de venta al por menor.

- Supervisión del calentamiento global.
- Reciclaje del suelo contaminado.

Este amplio abanico de posibles aplicaciones de la IG pone de manifiesto la necesidad de obtención, integración, utilización y explotación de datos espaciales procedentes de fuentes muy dispares en el tiempo, en la distancia y en los formatos. Pero la IG ha sido realmente costosa, especialmente su producción y validación, cuya competencia siempre ha recaído en los gobiernos del correspondiente ámbito.

Internet y la revolución informática han cambiado notablemente, entre otros aspectos, el concepto y la disposición de la IG. La Red mundial, los SIG, el GPS, los satélites de alta resolución, etc. permiten disponer de IG de cualquier punto del planeta en cuestión de segundos, y generar cartografía a la carta, obtener imágenes captadas unos minutos antes, capturar datos geográficos de manera rápida y sencilla, etc. Capturar y recoger toda esta IG para su posterior aprovechamiento es una tarea hasta cierto punto compleja, debido a la heterogeneidad de los productores, las fuentes, los equipos utilizados, los medios de difusión, el formato de almacenamiento y la aplicación que los explota.

Se entiende por captura de IG aquellos procesos encaminados a convertir los fenómenos reales que suceden sobre el territorio a formato digital, sufriendo modificaciones y simplificaciones, para ser tratados posteriormente por un SIG.

Hohl describe la importancia del proceso de captura de datos mediante el siguiente triángulo:

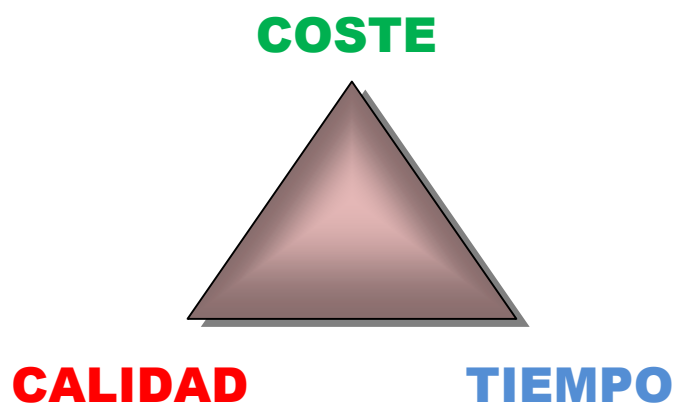


Figura 1. Triángulo de Hohl

La obtención de IG con calidad requiere una alta inversión económica y mucho tiempo de trabajo. Disminuyendo el tiempo de captura, disminuimos los costes y la calidad de los datos obtenidos. Con escaso presupuesto se tiene poco tiempo para capturar IG de la cual no se puede

verificar su calidad.

El análisis y la comparativa de los métodos de captura de la IG son el objeto de este tema.

10.2. Métodos de Captura

Existen diferentes técnicas de captura de dato atendiendo al formato de dato. Se pueden diferenciar entre técnicas de captura de información ráster y técnicas para la captura de información vectorial. A su vez, dependiendo del estado original de la información podemos diferenciar entre técnicas manuales, semiautomáticas o automáticas.

	Automático	Escaneado
RASTER	Automático	Teledetección / Satélites
	Manual - (Semi)Automático	Fotogrametría Digital
	Manual - (Semi)Automático	Digitalización
VECTORIAL	Automático	GPS - Topografía
	Manual - (Semi)Automático	Fotogrametría Analítica

Tabla 1. Métodos de captura

Si bien es cierto que están fronteras no están del todo claras como se verá a continuación.

10.2.1. Tabletas digitalizadoras. Digitalización manual

Una tableta digitalizadora es un dispositivo magnético o electromagnético que consta de un tablero sobre el que se sitúan los mapas o documentos analógicos. Las coordenadas de puntos en la superficie del digitalizador se envían al ordenador usando un lápiz magnético o un ratón.

Para que estas coordenadas sean válidas es necesario aplicar una transformación para convertirlas al sistema terreno y obtener el valor final de las mismas. Esta transformación se calcula previamente sobre puntos de coordenadas terreno perfectamente identificados en el

papel.

Existen dos formas de digitalización:

- Modo continuo: el operador va siguiendo la entidad a digitalizar y el aparato registra coordenadas a intervalos constantes de tiempo transcurrido o espacio recorrido.
- Modo punto a punto: el operador muestrea y selecciona los puntos que desea registrar pulsando el botón del lápiz o el ratón. Requiere cierta experiencia previa.

10.2.2. Escaneado y Seguidores de línea (Digitalización automática y semiautomática)

El escáner es un aparato electro-óptico utilizado en el ámbito de los SIG para la adquisición de datos geográficos desde los años 60. Los dispositivos del escáner generan una matriz de valores que representa la desigual reflectancia en pequeñas partes de la imagen original. Todos los escáneres tienen una resolución mínima de 8 bpp (bits por píxel ; 256 colores o niveles de gris) y pueden llegar hasta 24 y 32 bpp (millones de colores); la resolución espacial (número de píxeles por unidad de distancia) es también variable entre 72 ppp (píxeles por pulgada), 600 y 1200 ppp. Los escáneres de sobremesa (planos) suelen admitir documentos hasta un tamaño A3 (297x420 mm) y los de tambor hasta A0 (1188x840mm).

Los datos de entrada al escáner se encuentran en soporte analógico (papel) y una vez escaneado se pueden obtener:

- Datos geométricos: tras un proceso de vectorización.
- Atributos: el escáner es capaz de reconocer el color y en base a una clasificación por colores determinar el atributo de la entidad escaneada.
- Texto: mediante un sistema de reconocimiento óptico de caracteres (*ORC – pattern recognition*) extraer la información textual del documento escaneado.

Existen dos formas de trabajo:

- Barrido óptico teselar: generación de un fichero ráster mediante el barrido.
- Barrido óptico vectorial: generación de un fichero vectorial mediante un seguidor de líneas.

En los escáneres el operario sitúa el documento analógico en la superficie de escaneado y espera a que el proceso finalice, siendo un proceso totalmente automático (digitalización automática).

En los seguidores de línea el operador sitúa el dispositivo al principio de una línea y éste la sigue hasta que encuentra una intersección que debe ser resuelta por el operario. Se trata de un proceso semiautomático (digitalización semiautomática).

10.2.3. Fotogrametría

La Fotogrametría es la técnica que haciendo uso de una o varias fotografías tomadas sobre un objeto, es capaz de definir con precisión la forma, la posición y las dimensiones de dicho objeto.

Si los objetos fotografiados son objetos geográficos, la Fotogrametría se convierte en una fuente más de captura de IG. Esta técnica se caracteriza por la fuente de información: la fotografía. Una fotografía es un registro objetivo y completo del objeto que se adquiere de forma instantánea, es indirecto pues no hace falta tocar el objeto y además el punto de toma puede estar en movimiento, a diferencia de otras técnicas de captura de datos. A su vez, las fotografías son documentos de fácil manipulación y conservación, tanto en formato analógico, como el digital.

Un proceso fotogramétrico se resume en tres etapas:

1. Proyecto y ejecución del vuelo fotogramétrico.
2. Apoyo fotogramétrico o aerotriangulación.
3. Restitución de la información.

Para la obtención de coordenadas de fenómenos geográficos a partir de fotografías, es necesario que el objeto haya sido fotografiado desde dos puntos de vista que cumplan ciertas condiciones geométricas. Al par de fotografías obtenidas a partir de esta toma se le conoce como par estereoscópico.

Aún así, con una única fotografía también se pueden extraer datos de cierta relevancia geográfica cualitativa. A esta técnica se le conoce por Fotointerpretación.

El resultado de la restitución, último proceso de todo trabajo fotogramétrico, es la información tridimensional (coordenadas) de los objetos geográficos capturados. Este proceso se realiza en los restituidores. Estos equipos han evolucionado constantemente, pasando de los primeros modelos óptico-mecánicos, a los analógicos y posteriormente a equipos analíticos basados todos ellos en negativos de la fotogrametría. La última generación son los sistemas fotogramétricos digitales que trabajan directamente con imágenes digitales, captadas directamente por cámaras

aéreas digitales o bien por el escaneado de imágenes analógicas.

La ventaja de los sistemas digitales es la automatización de procesos laboriosos: orientaciones, restitución automática, extracción de la orografía, del MDT y la creación de ortofotos. En estos casos la información capturada se encuentra directamente en formato digital.

La Fotogrametría es una de las fuentes de IG más empleada por la excelente relación entre el coste, la rápida adquisición y la precisión conseguida.

- **Productos obtenidos.**

Los diferentes productos que se pueden obtener a partir de la Fotogrametría son:

- Imágenes, documentos gráficos de información continua.
 - Fotografías aéreas, obtenidas directamente de la cámara, ya sean convencionales o digitales.
 - Mosaicos, composiciones fotográficas partiendo de la fotografía área.
 - Imágenes rectificadas, en ellas se ha eliminado el efecto de la inclinación del eje principal con la vertical.
 - Ortofotos, en ellas se elimina el efecto de la inclinación del eje principal y el efecto del relieve (escala uniforme).
 - Ortofotomapas, en los que aparecen curvas de nivel y toponimia.
 - Ortomosaicos, composición de ortofotos.
 - Mosaico radar, composición de imágenes radar en una franja del terreno.
- Datos vectoriales o de línea.
 - Mapas planimétricos o de curvas de nivel o Topográficos.
 - Mapas temáticos.
 - Perfiles del terreno.
 - Modelos tridimensionales del objeto.
- Datos numéricos.
 - Coordenadas de los puntos de apoyo, datos obtenidos a partir de la aerotriangulación.
 - Otras magnitudes geométricas, superficies y volúmenes.
 - Modelos Digitales del Terreno, obtenidos a partir de puntos aleatorios o puntos obtenidos según una cuadrícula.

Otra de las fuentes vinculadas a la Fotogrametría es el LIDAR, acrónimo de *Light Detection And Ranging*. Se trata de una tecnología de visualización activa muy similar al RADAR (*Radio Detection And Ranging*). Este sistema está revolucionando la adquisición de datos digitales de elevación para las aplicaciones de mapas a gran escala. LIDAR presenta un coste efectivo, una fuente alternativa para los datos de Modelo Digital de Terreno.

Los sistemas LIDAR recogen datos posicionales (x,y) y de elevación (z) en intervalos predefinidos. Los datos resultantes LIDAR son una red de puntos muy densa, típicamente a intervalos de 1 a 3 metros. Los sistemas más sofisticados proporcionan datos de primero y segundo retorno los cuales proporcionan alturas del terreno y vegetación. Las alturas de la vegetación pueden proporcionar las bases para el análisis de aplicaciones de diferentes tipos de vegetación ó de separación de altura.

Una ventaja significativa de LIDAR es que los datos pueden ser adquiridos en tiempos en que la fotografía aérea convencional no puede hacerlo. Las misiones LIDAR pueden ser voladas durante la noche y en periodos nublados ó de bruma.

Los productos estándar fotogramétricos derivados de los datos LIDAR incluyen modelos de contorno y elevación para ortofotos.

10.2.4. GPS. Trabajos de campo

Se denomina levantamiento topográfico al conjunto de operaciones necesarias para representar topográficamente un terreno. El objetivo es obtener el levantamiento de todos los detalles que aparezcan sobre el terreno y que deban figurar en el plano y/o sean objeto de levantamiento, teniendo en cuenta la escala y objetivo del mismo.

La primera labor será establecer una Red Topográfica que se apoyará en la red o marco de referencia geodésico oficial. El objetivo de esta Red es la de proporcionar mayor cobertura topográfica a la zona de levantamiento, mediante la determinación de una serie de puntos materializados en el terreno, denominados vértices o puntos topográficos, con la mayor precisión posible. Estos puntos servirán de armazón a toda la zona objeto de levantamiento.

Cuando la Red Topográfica no es lo suficientemente densa en la zona objeto de levantamiento o en parte de ella, es habitual proyectar poligonales de relleno que parten y terminan en puntos de la Red Topográfica.

Posteriormente, se realiza el levantamiento topográfico, mediante la observación de ángulos y

distancias desde puntos de coordenadas conocidas (puntos pertenecientes a la Red Topográfica o a las poligonales de relleno). El método más sencillo y usualmente utilizado es el método de radiación.

Los aparatos capaces de medir ángulos y distancias simultáneamente se denominan estaciones totales. Existen diferentes tipos en función de las necesidades del trabajo, de la precisión requerida, del precio, etc.

Actualmente son de uso común los receptores GPS (Global Positioning System) tanto en la labor de observación de Redes Geodésicas y Topográficas, como en el levantamiento final de puntos del terreno. Su utilización mejora la eficiencia y los costes de producción significativamente, siendo una opción siempre a considerar.

A partir de las observaciones realizadas (ya sea por Topografía clásica o por GPS) se calculan las coordenadas de los puntos levantados. Con el listado final de coordenadas de los puntos (nube de puntos) se puede representar gráficamente la zona objeto del levantamiento (mapa o plano), así como crear un modelo digital del terreno (MDT), productos ambos de uso común en todos los Sistemas de Información Geográfica.

10.2.5. Teledetección

La Teledetección es la técnica que permite obtener información sobre un objeto, superficie o fenómeno a través del análisis de los datos adquiridos por un instrumento que no está en contacto con él. Se basa en que cada objeto, área o fenómeno emite un espectro electromagnético específico, en función de su propia naturaleza y de las radiaciones que recibe. La reflectancia de ese espectro electromagnético se denomina firma espectral, la cual hace distinguible a ese objeto, superficie o fenómeno de los demás y es la relación entre la energía que llega al objeto y la que sale de él.

Por lo general los datos son recogidos a través de sensores instalados en plataformas aerotransportadas o en satélites artificiales, los cuales captan la radiancia emitida o reflejada, obteniéndose una imagen, habitualmente en falso color con una banda para cada una de estas regiones del espectro. Los avances en tecnología han permitido contar con instrumentos cada vez más precisos basados en electrónica y experimentación con materiales que permiten obtener información cada vez más completa contenida en imágenes satelitales.

La técnica Teledetección tiene dos fases bien diferenciadas:

1. Adquisición de las imágenes.

2. Tratamiento de las imágenes adquiridas.

La adquisición de imágenes está afectada por la resolución del sistema. Existen cuatro tipos de resoluciones relacionadas con la imagen: la resolución espacial, distancia mínima que es capaz de captar el sistema (tamaño del pixel); la resolución temporal, intervalo de tiempo transcurrido entre dos pasos consecutivos del satélite por el mismo punto; la resolución radiométrica, capacidad del sensor para detectar variaciones de la radiación espectral que recibe; y la resolución espectral, número y anchura de las bandas del espectro que el sensor es capaz de distinguir.

Una vez adquiridas, las imágenes son pre-tratadas para ser corregidas de efectos provocados por las condiciones atmosféricas, los errores instrumentales, etc. Esos efectos afectan a la geometría de la imagen así como a su apariencia (histograma y contraste).

Posteriormente las imágenes están listas para ser tratada adecuadamente con el fin de obtener la información deseada. Existen dos tipos básicos de operaciones:

1. Operaciones de realce:

Conjunto de operaciones que facilitan la interpretación por parte del operador.

2. Operaciones de clasificación:

Cuya finalidad es separar la imagen en áreas homogéneas que compartan una misma propiedad del suelo.

10.2.6. Importación de datos

La importación de datos procedentes de otras fuentes implica que éstos ya se encuentran en formato digital aunque se trate de formatos diferentes. Para poder utilizarlos deberá hacerse un cambio de formato dependo del tipo de software utilizado.

Algunos de los formatos gráficos utilizados en aplicaciones SIG se muestran en la tabla 2.

FORMATOS VECTORIALES		FORMATOS RASTER	
Atlas	BNA	Arc Digitized Raster Graphics	ADRG
AutoCAD Drawing	DWG	ASCII (texto delimitado)	
AutoCAD Drawing Exchange Format	DXF	ASCII (texto)	XYZ
Automated Mapping System	AMS	AT&T Truevision Targa format	TGA
Computer Graphic Metafile	CGM	Band Interleaved by Line	BIL
Digital Feature Analysis Data	DFAD	Band Interleaved by Pixel	BIP
Digital Line Graph	DLG	Band SeQuential	BSQ
Dual Independent Map Encoding	DIME	Binario Genérico	RAW
Encapsuled Postscript	EPS	Compressed Arc Digitized Raster Graphics	CADRG
ESRI ArcView	SHP	Controlled Image Base	CIB
ESRI Coverage		Device Independet Bitmap	DIB
ESRI Geodatabase		Digital Terrain Elevation Data	DTED
Initial Graphics Exchange Standard	IGES	ERDAS 7,5	GIS
Interactive Graphic Desing Software	IGDS	ERDAS IMAGINE	IMG
Land Use and Land Cover Data	GIRAS	ERMMapper	ECW
Manifold	MFD/MDB	ESRI GRID file	GRID
Map Information Assembly Display System	MIADS	Graphics Interchagne Format	GIF
MapBase file	ETAK	IDRISI raster format	IMG ENVI IMG
MapInfo Interchange Format	MID/MIF	Image/surface format from ENVI	
MicroStation Drawing File Format	DGN	JPEG Interchange Format	JFIF
MOSS Export File	MOSS	MapInfo raster	TAB
Spatial Data Transfer Standard	SDTS/TVP	MFWorks Metafiel	MFM
Tele Atlas Impor Format	TAIF	Multi-resolution Seamless Image Database	
Topologically Intergrated Geographic	TIGER	PC Paintbrush format	PCX
UK National Transfer Format	NTF	Portable Network Graphics	PNG
Vector Product Format	VPF	Portable PixMap	PPM
		Silicon Graphics / Wavefront	PIX
		Silicon Graphics worstation format	SGI
		SPOT satellite images	SPOT
		Sun worstation graphics format	SUN TIFF GeoTIFF
		Tag Image File Format	
		USGS Digital Elevation Model	DEM
		Window BitMap	BMP

Tabla 2. Formatos de archivos SIG

La Cartografía ya existente como fuente de un SIG

Otra fuente de información muy utilizada en la actualidad es la reutilización de cartografía analógica oficial ya existente. Esta cartografía impresa se transforma a formato digital ráster, mediante un proceso de escaneado, o a formato digital vector, mediante la digitalización directa sobre la hoja o a partir de la imagen escaneada.

Una vez obtenida la IG en formato vectorial la carga a un SIG presenta una serie de aspectos a tener en cuenta:

- Creación de un modelo de datos. La información vectorial debe ser clasificada, ordenada y estructura en base a un modelo lógico de datos que constituya la base de datos geográficos (BDG). Debe existir una correspondencia entre los fenómenos geográficos recogidos en el mapa y los representados en el catálogo de fenómenos de la BDG. Éstos contienen una serie de atributos que diferencian unas entidades de otras. Es necesario crear una serie de herramientas que importen la información vectorial de un fichero grafico a la BDG.
- La cartografía impresa se encuentra separada por hojas. Para garantizar una continuidad en la BDG final es necesario, una vez cargada la información vectorial en el sistema, solucionar el problema conocido como CASE de hojas, esto es la falta de continuidad de elementos en los marcos de hojas entre hojas contiguas.

Solucionados estos aspectos, se consigue capturar información existente en soporte papel e introducirla en una BDG lista para ser explotada por SIG.

10.3. Captura de atributos

Los atributos son datos informativos no espaciales que detallan características propias de elementos geográficos espaciales permitiendo diferenciar unos de otros. El disponer de atributos con calidad y precisión que caractericen la IG es tan importante como la propia información espacial y deben ser adquiridos con el mismo rigor y exactitud que ésta.

Algunos de estos atributos son el nombre o topónimo del fenómeno geográfico; el tipo de fenómeno, si éstos están agrupados en torno a una misma característica (por ejemplo aeropuerto

público o aeropuerto privado); las medidas del fenómeno (ancho, largo, volumen, etc.); tipo de uso del suelo; tamaño de la población; número de incidencias, etc.

Las fuentes para la obtención de estos atributos son muchas y variadas, tanto en el acceso a dicha información como en la calidad del dato aportado. Las más empleadas son:

- Recolección directa en campo: consiste en acudir a la zona de trabajo y consultar a los habitantes del lugar sobre los nombres y topónimos.
- Fuentes oficiales públicas y privadas: gobiernos de los diferentes ámbitos, organismos públicos, empresas,...que publican datos sectoriales relativos a fenómenos geográficos, naturales, sociales, económicos, políticos, demográficos, etc...
- Datos libres: accesibles a través de Internet con la consecuente advertencia en cuanto a su fiabilidad y calidad.

Los errores más comunes en los atributos son por omisión, por incorrección o por transcripción errónea. La manera de localizarlos es mediante la impresión o ploteado de los datos en papel vegetal o transparente a la misma escala que el mapa base. Seguidamente los dos mapas se superponen sobre una mesa con foco de luz y se comparan visualmente de izquierda a derecha y de arriba abajo. Los errores detectados se marcan en la impresión. Otra manera no visual de chequearlos consiste en imprimirlos, ordenarlos y comprobar línea a línea cada uno de los datos.

Bibliografía

- [1] Bernhardsen, T. (XXXX) *GIS&Science XXXX*. Editorial XXXX. ISBN XXXX
- [2] Bosque Sendra, J. (1992) *Sistemas de Información Geográfica*. Ediciones Rialp. ISBN 84-321-2922-4
- [3] Burrough, P.A., McDonnell, R.A. (2000). *Principles of Geographical Information Systems*. Editorial Oxford. ISBN 0-19-823366-3
- [4] Chueca Pazos, Manuel; Herraiz Boquera, José (1996). *Métodos Topográficos*. Ediciones Paraninfo, Madrid.
- [5] Chuvieco, Emilio. (1999). *Fundamentos de Teledetección Espacial*. Ediciones Rialp, Madrid
- [6] Domingo Clavo, Luis (1985). *Apuntes de Fotogrametría*. E. U. I. T. Topográfica. E. T. S. de Ingeniería en Topografía, Geodesia y Cartografía. Madrid.
- [7] Goodchild, M.F., Maguire, D. y otros (2001) *Geographic Information Systems and Science*. Editorial Wisley. ISBN 0-471-89275-0
- [8] Lerma García, J.L. (2002). *Fotogrametría Moderna: Analítica y Digital*. Editorial de la Universidad Politécnica de Valencia. Valencia
- [9] Staller Vázquez, A (2005) *Captura de Información Geográfica*. UPM. E. T. S. de Ingeniería en Topografía, Geodesia y Cartografía. Madrid.

Tema 11. Edición de datos. Corrección de errores. Métodos de indexación espacial. Actualización de la información

11.1. Edición de datos

En el proceso de captura de datos geográficos, ya sean espaciales o no espaciales, es frecuente el que se obtenga información en bruto, no muy elaborada, siguiendo la política de acelerar los procesos de captura, que son habitualmente caros, largos y costosos.

En el caso de que los datos se importen, lo habitual es que el modelo original de los datos de entrada no coincida con nuestro modelo de datos, es decir que la información de entrada no sea directamente aprovechable, debido a posibles errores y deficiencias y al hecho de que necesitemos de dotarlos de características, aspectos y atributos que originalmente no tienen.

Por lo tanto, en todo proyecto SIG es necesario contemplar una fase de edición de los datos, posterior a la captura, para dotarlos de las características y estructura del modelo de datos deseado para el sistema final. Este proceso de tratamiento o refinamiento de los datos, es lo que llamamos edición de datos. Depende del método de captura empleado, de las diferencias entre los dos modelos de datos, el de origen y el de nuestro sistema, y de la calidad de los datos fuente.

11.1.1. Edición geométrica

La causa de los diferentes tipos de anomalías geométricas que se encuentran en los datos de un proyecto SIG, pueden tener su origen en diversos factores, tales como:

- Errores humanos en la interpretación de la realidad geográfica.
- Uso de métodos de trabajo inadecuados.
- Deficiencias en los métodos de captura.

Si a los errores así generados, se añaden los que se producen por la integración de conjuntos de datos heterogéneos, se presentarán una serie de anomalías o incoherencias espaciales, cuya corrección se ha de realizar en esta fase de edición geométrica. A continuación, se indican algunos de los errores que se presentan con más frecuencia.

- **Arcos colgantes**

También llamados extremos libres, se dan cuando se tiene una poligonal (secuencia encadenada de líneas) con al menos un extremo libre es decir, un punto final no coincidente con ningún otro extremo de otra poligonal o de sí misma. Los arcos colgantes pueden tener uno o dos extremos libres.

La detección de arcos colgantes se basa en la localización de puntos de otros elementos gráficos que tengan las mismas coordenadas que sus respectivos extremos libres. Si no existe ningún punto que cumpla este criterio de coincidencia, el elemento gráfico es un candidato a ser considerado como erróneo. En algunos casos es correcto que existan arcos colgantes (el nacimiento de un río siempre es un arco colgante), y en otro no lo es (una capa de límites municipales no debe tener ningún arco colgante), dependiendo del significado de los datos.

Los denominados “overshoot” (sobretazos) y “undershoot” (subtazos), representan un tipo particular de arcos colgantes.

- Sobretazo: Poligonal de longitud mínima con un extremo libre, detectado como erróneo por haber sobrepasado la línea con la que debería coincidir.
- Subtazo: Poligonal con un sólo extremo libre, detectado como erróneo por no ser coincidente con la línea –o con una de las líneas- existente en sus proximidades, y en la que podría finalizar.

Para localizar y corregir este tipo de errores, los programas de SIG disponen de un conjunto de funciones que suelen utilizar criterios basados en tolerancias gráficas de tipo lineal. Dado un determinado valor de tolerancia o distancia máxima de búsqueda D_{max} , cuando se localice un extremo libre sólo se detectará como error, si en el entorno circular de radio $R=D_{max}$ trazado a su alrededor, no se encuentra ningún otro elemento gráfico.

La **corrección** de errores se puede llevar a cabo de forma automática, semiautomática o manual.

Los segmentos de longitud mínima, mínima en función de la resolución del conjunto de datos al que pertenezcan, también pueden considerarse en ocasiones un tipo de arcos colgantes, cuya detección y corrección se realiza según los criterios expuestos.

- **Slivers**

Los *slivers*, astillas o cuñas, son falsos polígonos generados por la superposición de poligonales que debieran ser coincidentes y no lo son. Muchos de ellos se generan en los procesos de confluencia antes citados.

Estos polígonos se caracterizan por ser recintos cerrados, generalmente muy alargados, de gran perímetro y superficie mínima.

Los programas de SIG suelen incorporar funciones de localización y corrección, automática o semiautomática, de este tipo de anomalía geométrica. El criterio más utilizado en la detección de *slivers*, se basa en establecer un valor de tolerancia superficial (T_S). Si la relación entre la superficie y el perímetro al cuadrado de un polígono verifica esta condición, el polígono se considera un *sliver*

$$(1) \text{ (Superficie) / (Perímetro)}^2 \leq \text{T_S}$$

- **Bucles**

Son falsos polígonos que resultan de la autointersección errónea de una poligonal. Hay que tener en cuenta que en algunos casos se dan los bucles en la realidad, como en algunos enlaces de autopista.

- **Líneas duplicadas**

Son errores producidos por la duplicación innecesaria de elementos geométricos con coincidencia espacial exacta.

- **Intersecciones sin resolver**

En ocasiones dos líneas se cruzan, pero no está calculado el punto de intersección ni se ha incluido como un vértice más en cada una de las líneas. Es necesario hacerlo así para que un sistema informático localice y pueda utilizar esa intersección.

- **Otras anomalías**

Como vértices desalineados, densidad de puntos excesiva para la escala del modelo, etc.

11. 1.2 Edición semántica

La edición semántica consiste en la localización y corrección de los errores existentes en la identificación y caracterización de entidades, cuando éstos no se corresponden con la realidad. Las causas de este tipo de errores, pueden tener su origen en factores del tipo:

- Errores humanos en la interpretación de la realidad geográfica.
- Uso de criterios diferentes en la elaboración de cartografía. Por ejemplo, los diversos organismos oficiales que publican cartografía temática de un cierto tipo (edafología, geología, etc.), pueden utilizar diferentes clasificaciones estándar o incluso sus propias clasificaciones, lo que da lugar a cartografías no homogéneas.

Una de las tareas más difíciles de llevar a cabo sin cometer ningún error, es el chequeo y verificación del nombre de los fenómenos geográficos, labor harto complicada ya que es necesario recoger con finalidad acentos graves, agudos, apóstrofes, ce cedillas, eñes, mayúsculas acentuadas, minúsculas, guiones y toda suerte de caracteres especiales utilizados en el idioma o idiomas de trabajo.

Por añadidura, es habitual la situación en la que hay que gestionar nombres propios de ríos, montes etcétera en los distintos idiomas oficiales de un territorio e incluso en variedades dialectales en las que se conoce el nombre de un lugar o fenómeno.

El tratamiento temático también incluye la clasificación de los fenómenos geográficos que existen en el mundo real en las clases o tipo de fenómenos que sea necesario considerar en el modelo, por ejemplo: ríos, lagos, embalses, carreteras, caminos, autopistas, etc.

En esta labor, el problema es establecer criterios lo más claros, objetivos y homogéneos posible para clasificar los fenómenos, cosa a veces difícil de conseguir debido a la misma naturaleza de la información geográfica, inclasificable muchas veces de por sí. Porque ¿cómo se diferencia una pista forestal de un camino, o de un cortafuegos, o de una senda, o de una cañada?

Los programas de SIG suelen incorporar funciones de validación de atributos. Estas funciones permiten realizar algunas comprobaciones tales como:

- Verificar que el rango de valores numéricos de un atributo sea el correcto.
- Verificar que los valores (no numéricos) de un atributo pertenezcan a su dominio, es decir, al conjunto de todos sus posibles valores (lista de valores).
- Comprobar que existen unas determinadas relaciones entre los valores de un atributo ($>$, $<$,...).

11.2. Corrección de errores

En el proceso de captura o importación de los datos y en la fase posterior de edición de los datos pueden producirse todo tipo de errores y deficiencias, debido a múltiples causas, como son errores en el proceso de abstracción e interpretación del mundo real, errores técnicos en los procedimientos de captura de las coordenadas y de la información semántica, errores en la transmisión, errores en el procesamiento automático de los datos, errores interactivos, etc.

Independientemente de la fuente que los produzca, los errores pueden clasificarse siguiendo la misma taxonomía existente para describir los distintos aspectos de la calidad, y así se habla de errores:

- De exactitud posicional
- De exactitud temática
 - o Corrección de la clasificación
 - o Corrección de los atributos cualitativos
 - o Exactitud de los atributos cuantitativos
- De compleción, por omisión o por comisión
- De consistencia lógica, que pueden ser
 - o De consistencia conceptual
 - o De consistencia de dominio
 - o De consistencia de formato
 - o De consistencia topológica
- De exactitud temporal

11.3. Métodos de indexación espacial

11.3.1 Organización de ficheros

Una de las cuestiones básicas en la gestión de Información Geográfica consiste en cómo acceder a grandes volúmenes de datos en un tiempo razonable, entendiéndose, claro está que se trata de un acceso selectivo, es decir se trata de recuperar una parte de los datos con arreglo a un cierto criterio de selección. En este sentido, hay que tener en cuenta que el acceso físico a disco es una de las operaciones más lentas dentro de lo que es el proceso de datos. Esto es debido a que implica el que una cabeza lectora se posicione espacialmente en el lugar adecuado, un proceso de lectura y una transferencia de datos entre la unidad de almacenamiento y la CPU. El fondo de la cuestión es que el proceso mecánico de posicionamiento se hace a una velocidad de una magnitud muy inferior a las velocidades de los procesos digitales.

Como ejemplo, sirva la comparación entre las unidades en las que se mide la velocidad de obtención de datos en el acceso a disco y en el acceso a memoria real:

DISPOSITIVO	T. MEDIO	UNIDAD	CANTIDAD DE DATOS
Disco	milisegundos	Bloque	512 a 19.000 Bytes
Memoria	nanosegundos	Palabra	8, 16, 32 ó 64 Bytes

Tabla 1. Velocidades de acceso a datos en disco y en memoria. Datos del año 2000.

Entre los tiempos de acceso hay una relación aproximada de 1 a 300.000, y en el caso de bloques de 4.096 B y palabras de 16 B, valores muy frecuentes, la relación que resulta teniendo ya en cuenta el volumen de la unidad de recuperación, es de 1 a 1.200, favorable al acceso a memoria real.

Disco 0.13 MB/seg. Memoria 160 MB/seg.

En cualquier caso, también habría que tener en cuenta la velocidad de transmisión de los datos una vez recuperados que, al menos en el caso de la RAM, es significativo.

Vamos a ver las distintas formas en que puede organizarse un fichero atendiendo al mecanismo de acceso que se contempla.

- **Ficheros secuenciales**

En un fichero secuencial, los registros se colocan sencillamente uno a continuación de otro, y no se dispone de información previa alguna acerca de en qué lugar del fichero se encuentra un registro en particular. Por lo tanto, para localizar un registro concreto, es necesario leer secuencialmente el fichero hasta encontrarlo. Si el registro deseado está al principio del fichero, el tiempo de búsqueda puede ser bastante corto, pero si es el último el único modo de localizarlo es leer todo el fichero registro a registro.

Si tenemos un fichero de N registros, y el tiempo de lectura de un registro es t, el tiempo de búsqueda medio es igual a:

$$(2) \quad T = (N+1) * t / 2$$

Para un fichero de 100.000 registros y un tiempo de lectura individual de 0.1 milisegundos, el tiempo de búsqueda medio T para un sólo registro sería de 5 segundos.

Evidentemente, en las aplicaciones reales no es el caso mas frecuente el realizar una búsqueda de un único registro, lo más habitual es acceder al conjunto de registros que cumplen una determinada condición, y buscar alguna información complementaria en algún fichero auxiliar.

Por otro lado, los tiempos de respuesta dependen de multitud de factores, desde el volumen total de los datos, hasta las características físicas del equipo, pasando por el diseño de la base de datos. Sin embargo, las estimaciones que se van a proporcionar sirven para dar una idea de las eficacias relativas de los distintos métodos de acceso.

Volviendo al caso de los ficheros secuenciales, el añadir nuevos registros al fichero es trivial, basta con escribirlos a continuación del que ocupa la última posición.

- **Ficheros secuenciales ordenados**

El siguiente paso lógico para facilitar el acceso a un fichero secuencial, es ordenar sus registros por valores crecientes o decrecientes de un campo que se toma como clave principal de búsqueda, de la misma manera que se ordenan las definiciones contenidas en un diccionario por orden alfabético del término definido. Dicha clave principal sirve como identificador de cada registro y para buscar un registro en particular, se lee el registro central del fichero y se examina si su clave es mayor o menor que la buscada. De esta manera se deshecha ya la mitad del fichero, y se vuelve a proceder con la mitad restante del mismo modo, y así sucesivamente hasta encontrar el registro deseado. Es fácil demostrar que ésta es la manera óptima de proceder en una búsqueda automática y a este procedimiento se le llama búsqueda dicotómica.

El tiempo de búsqueda medio es ahora:

$$(3) \quad T = t * \log_2 (N+1)$$

Lo que supone en el caso del ejemplo anterior (N=100.000, t = 0.1 miliseg.) un tiempo de búsqueda medio notablemente menor, de 1.7 milisegundos.

A cambio de esta mejora en el tiempo de búsqueda, el añadir un nuevo registro conlleva el reordenar todo el fichero.

- **Ficheros indexados directos**

Otro método para hacer más efectiva la búsqueda secuencial, consiste en almacenar dentro de cada registro del mismo fichero, información acerca de durante cuantos registros un número determinado de caracteres de la clave de búsqueda permanecen constantes, es decir cuantos registros hay que saltar para llegar al siguiente valor distinto de dicha clave.

Por ejemplo, si se tiene un diccionario ordenado alfabéticamente y se tienen n_A registros que comienzan por la letra A, n_B que comienzan por la letra B, ..., n_Z que comienzan por la Z:

el bloque de registros que comienzan por A comenzará en el registro número 1

B	$n_A + 1$
C	$n_A + n_B + 1$
D	$n_A + n_B + n_C + 1$
...	...
Z	$n_A + n_B + n_C + \dots + n_Z + 1$

En este caso:

$$(4) \quad T = t * (M + 1) / 2 + (n + 1) / 2$$

Dónde M es el número de valores distintos de la clave de búsqueda, en el ejemplo anterior A, B, C...Z, $N=24$, y n el número medio de registros por valor de la clave, es decir el número total de registros N dividido por el número de valores distintos de la clave M.

Añadir un nuevo registro exige reordenar todo el fichero y actualizar el valor ni correspondiente. En las condiciones de nuestro ejemplo comparativo ($N = 100.000$ y $t = 0.1$ miliseg.) y optimizando los valores de M y n ($M = n = 316$), el tiempo medio de búsqueda es de 32 milisegundos, superior al correspondiente a la búsqueda dicotómica.

Hay distintas maneras de organizar ficheros indexados directos. Un ejemplo es el formato DGN, en el que cada primitiva geométrica tiene en su cabecera un parámetro WTF (*Words To Follow*), que indica cuántas palabras de 16 Bytes es necesario avanzar para llegar hasta la cabecera de la siguiente primitiva.

En los tres casos vistos hasta ahora el tiempo de búsqueda sin resultado, es decir el tiempo necesario para verificar que el registro buscado no existe, es el máximo posible, lo que supone además un consumo de recursos máximo.

- **Ficheros indexados inversos**

En este caso, se extrae un fichero externo de índices, un fichero auxiliar en el que se almacena en qué posición del fichero está el registro correspondiente a cada uno de los valores de la clave de búsqueda. De manera análoga a como se construye un índice analítico en un libro, en el que se recogen los términos mas significativos y la página o páginas en las que figura, se puede almacenar cada uno de los valores de la clave de búsqueda y la posición o posiciones en las que aparece. La clave de búsqueda no tiene por qué ser la clave principal del fichero, puede repetirse varias veces.

Para acceder a su vez al fichero de índices, puede repetirse el mismo esquema, generándose así in fichero de índices de los índices primarios, y así sucesivamente hasta tener varios niveles de índices. Esta estructura es lo que se llama un árbol de índices invertidos.

Añadir un nuevo registro supone reconstruir toda la lista de índices invertidos.

Este mecanismo de indexación tiene dos ventajas muy importantes, además de su alto rendimiento. La primera es que se pueden definir tantas claves de búsqueda como se deseen, creando un fichero de índices para cada una de ellas. Por lo tanto, no hace falta que un campo sea clave única para poder indexarlo.

La segunda, es que si se intenta buscar un registro con un valor de la clave de búsqueda que no existe, no es necesario acceder a los datos para verificarlo, como en los casos anteriores. Al no encontrar el valor solicitado en la lista de índices, el proceso se detiene con el consiguiente ahorro de tiempo y recursos.

Este es el modo en el que están organizados habitualmente los ficheros bajo un Gestor de Base de Datos.

El tiempo de búsqueda medio depende de gran variedad de factores, como el número de valores distintos que puede tomar la clave de búsqueda, su longitud en caracteres, el número de medio y máximo de registros con un mismo valor de la clave, etcétera. Pero en el ejemplo que estamos manejando, de un fichero de 100.000 registros, si suponemos una clave de longitud 12 Bytes, que puede tomar 1000 valores distintos y por cada valor de la clave hay 100 registros, podemos aproximar el tiempo de búsqueda medio a 0.4 milisegundos.

- **Árbol de índices invertidos**

Antes hemos dicho que se almacena la posición en la que se encuentra cada uno de los registros para cada valor de la clave. Lo que se almacena en realidad es el número de bloque en el que se encuentra, ya que los procesos de lectura y escritura en disco se realizan por bloques completos. Una vez localizado el bloque en el que está el registro deseado, se lee ya se localiza el registro dentro del bloque mediante una búsqueda secuencial en memoria real, proceso, como ya hemos dicho, de muy alta velocidad.

Supongamos que tenemos un fichero con datos acerca de una serie de personas y hemos definido el nombre propio como clave de búsqueda. Si buscamos los registros cuyos nombres empiezan por A (suponemos que sólo hay cuatro valores que cumplen esa condición: Alberto, Alfredo, Alfonso y Antonio), el camino a seguir sería: accediendo al primer nivel de índices invertidos, vemos que debemos acceder al bloque 15 para localizar los nombres que empiezan desde la A hasta la CH; en el segundo nivel, accediendo al bloque 15, leemos que los datos acerca de los nombres que empiezan por A están en el bloque 19; en tercer el nivel, accediendo al bloque 19, sabemos que los registros cuyo nombre empieza por A se encuentran en los bloques 123, 89, 142 y 76. (ver tabla 2)

BLOQUE 14	De A a CH BLOQUE 15	De D a L BLOQUE 16	De M a R BLOQUE 17	De S a Z BLOQUE 18
BLOQUE 15	A BLOQUE 19	B BLOQUE 20	C BLOQUE 21	CH BLOQUE 22
BLOQUE 19	ALBERTO BLOQUE 123	ALFREDO BLOQUE 89	ALFONSO BLOQUE 142	ANTONIO BLOQUE 76

Tabla 2. Estructura de un árbol de índices invertidos.

Nótese que no es necesario que el fichero de datos esté ordenado por valores crecientes de la clave de búsqueda, pero que en tal caso aumentará la eficiencia de la búsqueda.

Así, se dispone de los niveles que sean necesarios para que en el nivel superior haya siempre un sólo bloque. El número de lecturas para acceder a un registro es igual al número de niveles más uno, y en cada una de estas lecturas se accede a un sólo bloque.

La búsqueda es muy rápida y el precio que hay que pagar por ello es mantener un árbol de índices que puede llegar a ocupar mucho espacio dependiendo del número de claves de búsqueda de que se quiera disponer, incluso tanto como el propio fichero de datos, y que necesita ser reconstruido cuando hay nuevas incorporaciones.

- **Ficheros aleatorios** (en desuso)

En un fichero aleatorio, también llamado de acceso directo, se define un campo clave de búsqueda, se ordenan en orden creciente todos sus posibles valores. A continuación el sistema transforma cada uno de esos valores en un valor numérico distinto mediante un cálculo. Si la clave es numérica, es suficiente con su valor, y si es alfanumérica es sencillo disponer de un algoritmo que transforme una cadena de n caracteres en un número de manera unívoca. Una vez obtenido ese valor numérico, el sistema calcula la posición en la que escribir el registro correspondiente mediante una función matemática aleatoria, aleatoria en el sentido que para cada número de entrada distinto arroja un resultado diferente.

$$\text{Posición} = F(\text{Clave})$$

Buscar un registro por el valor de la clave, significa por lo tanto realizar el cálculo de la función aleatoria, lo que consume simplemente un cierto tiempo de CPU, y realizar una lectura en la posición indicada. El tiempo de búsqueda medio será prácticamente independiente del tamaño del fichero, y en nuestro ejemplo siempre menor que:

$$T = 0.2 \text{ miliseg.}$$

Para acceder a un único registro, el acceso aleatorio es más rápido que el acceso a través de listas invertidas. Pero para acceder a un conjunto de registros que tiene el mismo valor de la clave de búsqueda, el acceso aleatorio es más lento que el de los ficheros indexados inversos.

Otro inconveniente no desdeñable de estos ficheros, estriba en que un fallo en el funcionamiento de la función aleatoria o en la ubicación de un registro, hace el fichero irrecuperable al no servir

ya la función aleatoria para acceder a los registros. En caso de desorganización del fichero la información es irrecuperable.

La incorporación de nuevos registros hace necesaria la reorganización periódica del fichero.

11. 3.2 Indexación espacial

- **Mínimo Rectángulo Envolvente (MER)**

El Mínimo Rectángulo Envolvente (MER), también llamado *Bounding Box*, es el método más utilizado y extendido en programas y utilidades SIG que gestionan datos vectoriales. Consiste en almacenar las coordenadas máximas y mínimas (Máx. X, mín. X, Máx. Y, mín. Y, [Máx. Z, mín. Z]) del conjunto de todas las coordenadas de los vértices que componen una primitiva geográfica o un fenómeno geográfico:

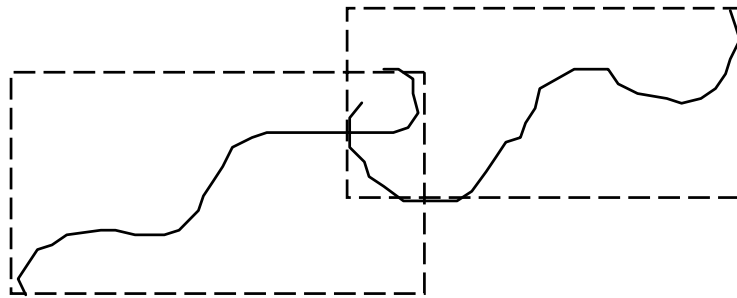


Figura 1. Mínimos rectángulos envolventes

Los datos que describen ese MER pueden almacenarse o bien como una cabecera del objeto, junto a la descripción de todos los vértices, o bien en un fichero de indización separado.

En cualquier caso, el tener el MER calculado y almacenado para cada primitiva geométrica (polígonos, superficies o curvas), permite realizar una primera criba espacial cuando se están realizando búsquedas espaciales orientadas a recuperar los objetos que se encuentran dentro o que tocan un ámbito geográfico determinado.

- **Quadtrees**

Quadtree es un nombre genérico para varios tipos de índices espaciales contruidos sobre una división recursiva del espacio en cuadrantes. Se utilizan tanto para comprimir datos, como para crear índices espaciales que permitan encontrar más rápidamente la información.

Para codificar en qué zona se encuentra un elemento geográfico del espacio total que cubre un conjunto de datos, se divide la extensión total en cuatro cuadrantes, y cada uno de ellos se vuelve a dividir recursivamente, o bien hasta una resolución límite o bien hasta que se encuentre un mínimo cuadrante envolvente. Por ejemplo si tenemos un objeto superficial:

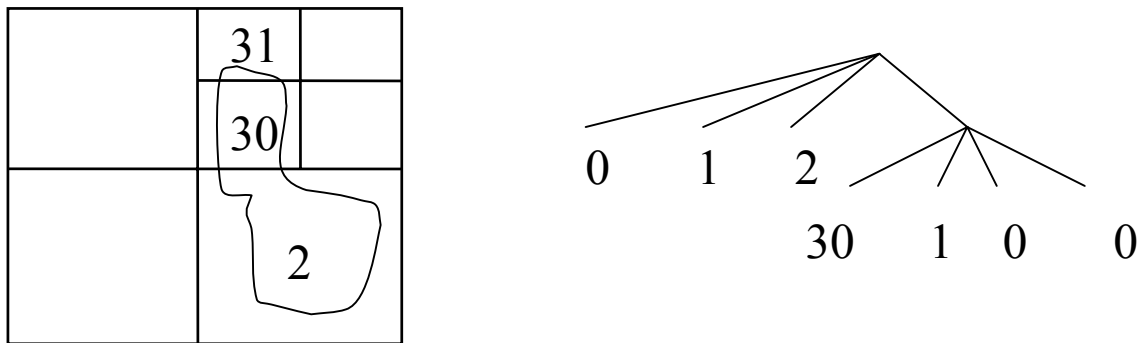


Figura 2. Quadtrees

La superficie sólo ocupa los cuadrantes SE y NE, y de éste último solo los subcuadrantes NO y SO. Para cada objeto, como la superficie de la figura, se almacenan secuencialmente sólo los índices de los cuadrantes en los que se encuentra. Si el orden es SO, NO, SE, NE, en el ejemplo de la figura los índices se almacenarían: 2, 30, 31.

Los *quadtrees* son ampliamente utilizados como método de indexación espacial en Bases de Datos espaciales por su efectividad, sencillez conceptual y por ser aplicables a muchos tipos de datos diferentes: puntos, líneas, polígonos, datos ráster, etc.

- **Índices R-tree**

Los índices R-tree utilizan el MER para aproximar la zona en la que se encuentra el objeto a indexar. En realidad es un método que se basa en indexar jerárquicamente utilizando sucesivos MER en lugar de cuadrantes, tal y como se ve en la figura.

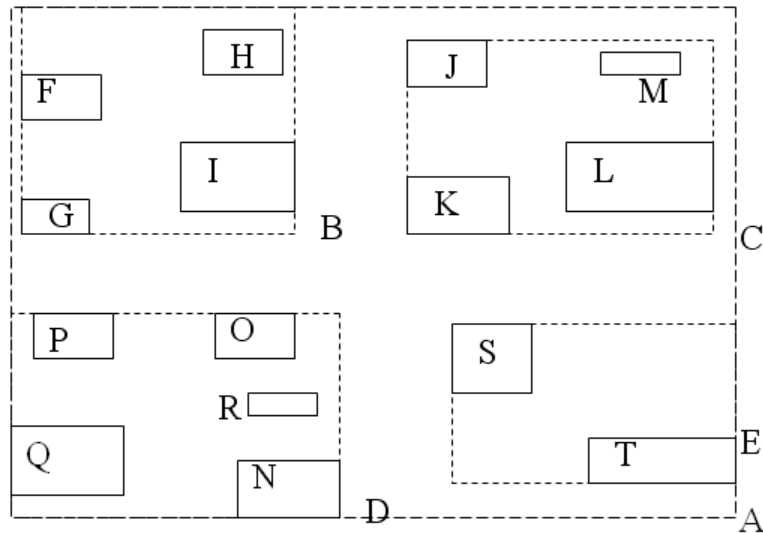


Figura 3. Indexación

El árbol de índices depende de los MER que haya dentro del MER de nivel superior. Cuando hay que introducir un nuevo objeto, un algoritmo calcula en qué rectángulo se puede incluir causando una modificación mínima de sus dimensiones.

Los R-tree son muy utilizados para crear índices espaciales por su gran flexibilidad y eficacia.

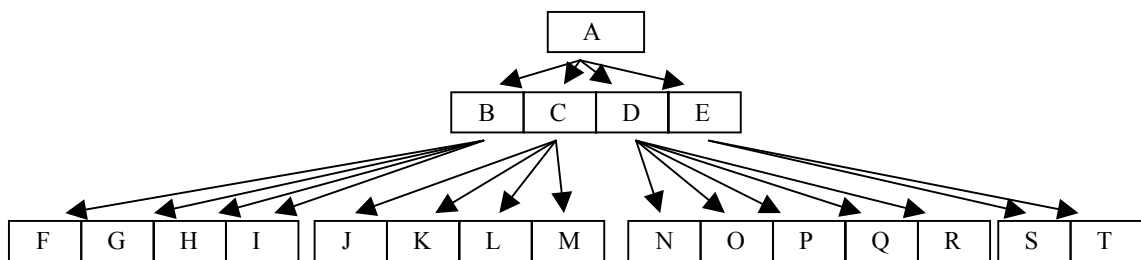


Figura 4. Índices R-Tree

11.4. Actualización de la información

El problema de la actualización de los datos cargados en un SIG es infravalorado muy a menudo y debe ser tenido en cuenta desde el principio del proyecto y planificado desde su inicio, por la gran complejidad técnica que supone y por los costes que implica.

La Información Geográfica es muy dinámica, debido por un lado a la constante variación de los objetos que se pueden llamar de obra humana (urbanismo, vías de comunicación) que pueden llegar a suponer actualizar el 10% anual de los objetos de un SIG, y por otro a la tasa residual de errores siempre presente en un SIG, que es necesario revisar y corregir en la medida de lo posible, de manera casi continua.

Por ese motivo resulta especialmente útil el disponer de varios puestos de trabajo desde los que se actualice el contenido de la Base de Datos

En la actualización de las Bases de Datos alfanuméricas, el concepto de transacción es fundamental. En el modelo relacional, la información relativa a un objeto, un río o una carretera, por ejemplo, se reparte en una serie de tablas enlazadas entre sí por claves externas. Al modificar un objeto, es importante asegurar la coherencia entre los registros relativos a tal objeto. Para ello, se manejan transacciones. Una transacción es el conjunto de actualizaciones de registros que completa una modificación en un objeto con todas sus consecuencias, es por tanto el proceso que pasa desde un estado consistente de la Base de Datos, a otro estado consistente de la Base de Datos, en la que el objeto que se deseaba actualizar, se ha modificado.

Los Gestores de Bases de Datos (GBD), se encargan garantizar que la Base de Datos no queda en un estado inconsistente, es decir que la transacción completa o se produce completamente hasta el final, o no se produce en absoluto. Para ello, ante de iniciarse la transacción, el GBD bloque todos los registros afectados para que otro usuario no pueda modificarlos al mismo tiempo, y no los libera hasta que se ha completado la transacción.

Al intentar extender esta filosofía de trabajo a las Bases de Datos geográficas, es decir a los SIG, aparece el problema de que el número de registros a bloquear para actualizar, por ejemplo una carretera, es enorme, a que si se modifica una carretera a menudo se ven afectados un puente, un río, varios cascos de población, etc. y al ser muy voluminosa la información a bloquear, se generan transacciones muy largas, que hacen el sistema poco eficiente y aumentan la probabilidad de que ocurran incidencias durante la transacción.

Por ese motivo, en entornos SIG en los que interesa que varios operadores actualicen los datos simultáneamente, se emplea la gestión de versiones, el versionado. Un usuario autorizado para actualizar, extrae una copia, o versión, de la Base de Datos, otro usuario extrae una segunda versión y ambos actualizan y trabajan sobre su versión. Una vez acabado su trabajo, cada uno salva su versión. Si no hay conflicto, ambas versiones se salvan sin problema, y si lo hay, el sistema avisa de la incoherencia y un administrador elige entre cada versión, con la posibilidad de recuperar cualquier versión anterior. Cada versión no incluye una copia completa de la Base de Datos, sólo almacena los cambios, por lo que no ocupa mucho espacio.

Por último, hay que decir que un productor de datos que difunde o comercializa IG debe estar preparado para proporcionar actualizaciones no sólo como nuevas versiones completas, sino

también como suministro de sólo cambios, proporcionando los datos que han sufrido modificaciones. De esta manera los usuarios que han enriquecido la información original con sus propios atributos y relaciones, no pierden toda la información añadida.

Para poder proporcionar actualización como sólo cambios, es imprescindible difundir la información con identificadores únicos de los objetos, y que los usuarios los almacenen, para poder así identificar unívocamente qué objetos han desaparecido, qué objetos han aparecido como nuevos y qué objetos se han modificado.

Bibliografía

- [1] Buirrough, P. A. “Principles of Geographical Information System for Land Resources Assesment”, 1986, OXFORD SCIENCE PUBLICATIONS (UK)
- [2] “Geographical Information Systems” Tör Bernhardsen, 1992, 318 pág. VIAK IT (Noruega).
- [3] “Innovations in GIS 1” Michael F. Worboys, 1994, Taylor & Francis.
- [4] “Innovations in GIS 2” Peter Fisher, 1995, Taylor & Francis.

Tema 12. Calidad de datos. Fuentes de error en un SIG. Propagación de errores. Descripción de la calidad: exactitud posicional, temática y temporal, compleción, consistencia lógica, propósito, linaje y uso. Métodos estadísticos de determinación de la calidad. Calidad de procesos.

12.1. Introducción

Antes de profundizar conviene que nos preguntemos ¿qué es calidad? Una contestación inmediata, intuitiva y muy válida nos dice que “una obra bien hecha es/tiene calidad”. Por su parte el Diccionario de la Lengua Española de la Real Academia ofrece la siguiente definición: “Propiedad o conjunto de propiedades inherentes a una cosa que permiten apreciarla como igual, mejor o peor que las restantes de su especie”. Esta definición viene a indicar que: a) la calidad puede expresarse como una o más propiedades, b) la calidad supone apreciación, y por ello existirán tantas apreciaciones como sujetos distintos, c) el considerar algo mejor, igual o peor, supone comparar. Estos tres aspectos son base de todo lo concerniente a la calidad, aquí brotan numerosas incógnitas (qué aspectos, qué propiedades, cuáles son más importantes, qué se aprecia, qué considera cada cuál, cómo se compara, etc.) que tratan de resolverse en todas las organizaciones que tienen preocupación sobre este asunto.

En este tema no nos va a preocupar la calidad teórica del producto (según su diseño), sino la calidad de su realización, de su producción o materialización. Se trata de una perspectiva utilitaria desde el punto de vista del productor y que consiste en que lo producido se ajuste a lo establecido en sus especificaciones.

En nuestro campo de trabajo la principal actividad es la producción de información geográfica (IG), de datos o bases de datos geográficas (BDG). Un dato es todo aquello que nos informa/contesta a una serie de preguntas como: dónde, cuándo, cómo, cuánto, etc. y por ello hemos de considerar que el dato geográfico posee varias dimensiones: espacio, tiempo, atributos, etc. Y lo importante para este tema es que cada una de ellas puede ser observada desde la perspectiva de su calidad. Así vista, la cosa podría parecer fácil, pero la IG es complicada, todos lo sabemos: es voluminosa, dependiente de la escala, borrosa, dinámica, encadena numerosos procesos, y muchos de esos procesos y algoritmos son realmente modelos.

De una forma abstracta, y muy general, podemos decir que el hacer cartográfico actual no es más que un proceso de transformación de datos e informaciones del mundo real en productos cuya explotación sirve para la toma de decisiones, las cuáles muchas veces tienen importantes consecuencias económicas, sociales, ambientales, etc. sobre ciudadanos y comunidades. Es justamente por este desempeño de base para la toma de decisiones que la calidad de la IG es muy importante, tanto más cuanto más importantes sean las consecuencias del uso de la IG. Por ello la calidad no debe ceñirse sólo a las cualidades del dato geográfico sino que también importa la manera en que se gestiona o procesa (por ejemplo al aplicar un modelo determinado).

En relación al propio dato geográfico, una de las principales características de su calidad, entendida también como posibilidades, es la resolución. Se puede hablar de tantas resoluciones como dimensiones o componentes del dato geográfico, entre otras:

- Resolución espacial: Se define como el tamaño del menor elemento que puede ser distinguido o separado de sus partes constituyentes o del conjunto al que pertenecen. En los sistemas vectoriales se refiere a la mínima unidad cartografiable, mientras que en los sistemas ráster se refiere a la dimensión de la celdilla.
- Resolución temática: Debe distinguirse entre datos cuantitativos y cualitativos. En el caso de los primeros es similar a la resolución espacial en tanto y cuanto se trata de una variable cuantitativa; mientras que en el caso de los datos cualitativos se refiere a la adecuación/precisión de la definición de las clases.
- Resolución temporal: Se refiere a la duración temporal (intervalo) del muestreo que siempre deberá ser inferior a la duración del fenómeno a recoger.

Advertir que en el caso de resolución espacial y temporal no conviene confundir éstas con las tasas de muestreo espacial o temporal.

12.2. Fuentes de error

Calidad y error son dos conceptos íntimamente ligados al trabajar con IG. La calidad es algo relativo y utilitario, ya que unos datos de mala “calidad” para un trabajo de “precisión” pueden ser muy adecuados para otro tipo de requerimiento no tan exigente (idoneidad al uso).

Error significa algo equivocado o falso, por lo que se relaciona con la verdad o con un patrón tomado como verdadero. El error se cuantifica como discrepancia entre el dato y el valor verdadero. Por tanto, y en cierta medida, el error es un término complementario al de calidad (mucho error → mala calidad, calidad → poco error), pero que debe ser matizado por la idoneidad

al uso, como ya se ha indicado. A pesar de su extendido uso el término error no es del todo correcto. Si se conoce el error bastaría con corregirlo para tener un dato exacto, coincidente con la realidad. Sin embargo, esto no es posible: no se trabaja con errores sino con incertidumbres. Éste es el término que debería utilizarse. La incertidumbre se puede estimar o medir, pero no se puede eliminar. De cualquier forma, en este tema se seguirá trabajando con la palabra error en muchas ocasiones.

Los errores que se pueden clasificar en: **groseros, sistemáticos y accidentales o aleatorios**. Los errores groseros son equivocaciones perfectamente evitables, y su control y detección debe ser un objetivo del método de trabajo seleccionado. Los errores sistemáticos (constantes o variables) ocurren con el mismo signo y, frecuentemente, con la misma magnitud, en un número consecutivo o relacionado de observaciones, y se generan por causas permanentes, de tal manera que serán eliminables si éstas quedan determinadas. Los errores accidentales son aquellos que se generan por procesos aleatorios; individualmente son pequeños, no presentan una ley fija y se compensan parcialmente al incrementar el número de observaciones o mediciones. En la mayoría de los casos se puede suponer que los errores aleatorios siguen un modelo gaussiano (distribución Normal).

Ligados a los errores aleatorios existen dos conceptos básicos: la **exactitud** y la **precisión**. La exactitud es la aproximación de los resultados de la observación (medida, cálculo o estimación), respecto a los valores verdaderos, o a aquellos que se consideran como verdaderos. El valor verdadero puede ser el obtenido de una media de observaciones o el procedente de un método considerado como más refinado. La precisión es una medida del agrupamiento en la repetición de una misma medida. La precisión sólo se ve afectada por la componente aleatoria, no tiene en cuenta los sistematismos y, por tanto, no guarda relación con el valor tomado como real. Unos datos pueden ser muy precisos, con poca dispersión, y poco exactos o alejados del valor real. En otros ámbitos la precisión se relaciona con la calidad del método de medida (por ejemplo, nivelación de alta precisión). Otra acepción común de precisión hace referencia al número de cifras significativas que soporta un tipo de dato (enteros y reales cortos o largos, etc.) desde el punto de vista informático. La limitación en los sistemas con soporte informático procede más de la falta de exactitud en los datos con los que se trabajó que en la falta de manejo de decimales por parte del sistema.

Dado que las tipologías básicas del dato son la cuantitativa y la cualitativa, los errores también adoptan estas casuísticas, y los índices y métodos para estimarlos y evaluarlos deben ser coherentes con las mismas, estos aspectos se tratarán más adelante dentro de este mismo tema.

Son varias las taxonomías o clasificaciones de las tipologías de error que se han propuesto: Burrough (1986), Veregin (1989), etc., pero todas ellas se pueden encajar en la siguiente clasificación más amplia de las posibles fuentes de error:

- Fuente de información: En el ámbito SIG es muy común utilizar fuentes de información procedentes de terceros. Estas fuentes son externas a nuestra organización y poseerán más o menos error en función de su antigüedad, cobertura espacial, escala, relevancia, técnica de levantamiento y compilación, soporte, etc.
- Operario: La persona que realiza los trabajos puede ser una fuente de error con independencia del método e instrumental utilizados.
- Método: El método de trabajo puede ser más o menos riguroso. Además en el trabajo con IG el orden de ciertas operaciones puede influir en la incertidumbre final de una BDG.
- Instrumental: Las herramientas físicas (por ejemplo nivel, GPS) poseen ciertas limitaciones y posibilidades de introducir errores en las medidas (por ejemplo graduación de los círculos, conversores analógico-digitales, etc.). Por su parte, las herramientas lógicas (software) a la hora de procesar los datos se basan en modelos (por ejemplo método de resolución de una ecuación, cálculo de la pendiente según el algoritmo de Horn) que también pueden introducir más o menos error.

La manera de reducir las posibilidades de introducción de error para cada una de estos casos es muy diversa y muchas de ellas pueden encontrar satisfacción dentro de un sistema de gestión de la calidad. Cuando se va a trabajar sobre fuentes de información ya existentes un aspecto básico es conocer lo que existe y, si es posible, elegir aquella que más se ajuste a nuestro propósito, realizando evaluaciones de su calidad antes de proceder a su uso. En el caso de lo operarios se deben cuidar los aspectos formativos respecto a las prácticas a desarrollar. También pueden influir aspectos relativos a la ergonomía y el ambiente. En algunos ámbitos se ha propuesto la certificación y calibración del personal para ciertas tareas (por ejemplo digitalización y restitución). Respecto al método lo más importante es que esté descrito convenientemente. De esta forma se podrá mejorar sucesivamente.

El método debe ser eficaz y eficiente respecto al propósito y desde el punto de vista del error en la IG debe haber sido diseñado teniéndolo en cuenta: reduciendo las posibilidades de equivocación, de incorporar incertidumbre, describiendo adecuadamente cómo se deben realizar las operaciones, etc. Esto ha sido un hacer clásico en campos como el de la topografía, basta recordar cómo se decide un método y equipo en función del error de cierre, y cómo los estadillos de campo incluían elementos de verificación (por ejemplo cierre de la vuelta de horizonte). Pero esta perspectiva no siempre está presente al trabajar con IG en entornos SIG. El instrumental es un elemento clave dentro de las fuentes. En la actualidad la rapidez con que evoluciona el mercado ocasiona que no siempre sean

conocidos a fondo. En el caso del software la situación es aún más crítica, la mayoría de ellos trabajan como cajas negras y los usuarios no conocen nada de los modelos y algoritmos que incorporan. En esta línea se echa de menos mayor transparencia por parte de los productores de software e incluso la necesidad de certificación.

El dato geográfico sigue un ciclo de vida en el cual hay abundantes oportunidades de introducción y/o generación de errores, y donde intervienen operarios, metodologías e instrumental. La Tabla 1 indica diversos procesos a los que se somete el dato geográfico y que ofrecen oportunidades de introducción de error.

Proceso	Motivo
Modelización conceptual	Errores en el modelo conceptual
Recogida de datos	Error en los trabajos de campo. Error en las fuentes de información utilizadas.
Captura de datos	Inexactitud en la digitalización. Inexactitud inherente a los elementos geográficos.
Almacenamiento	Insuficiente precisión numérica y/o espacial. Errores de procesamiento
Manipulación	Intervalos de clase inapropiados. Errores de superposición. Propagación de errores.
Salidas cartográficas	Errores en la transformación de coordenadas Inexactitud escala. Inexactitud del dispositivo de salida. Deformaciones en el soporte.
Uso de los resultados	Entendimiento incorrecto. Uso inapropiado.

Tabla 1.- Procesos comunes donde se generan errores en manejo de datos geográficos

Fuente: Tras Aronoff (1989)

Desde el punto de vista de los errores un buen proceso cartográfico es aquel en que:

- a) todas las operaciones están diseñadas para reducir al máximo la incertidumbre que se incorpora al ejecutarlas.
- b) todas las operaciones o subprocesos están debidamente caracterizados desde el punto de vista de la incertidumbre. La caracterización de los procesos o subprocesos puede realizarse según modelos físicos o estadísticos. Los primeros suponen un conocimiento muy preciso del propio proceso, de cada una de sus fases, y de los modelos analíticos que las soportan y de los parámetros que los precisan a la situación concreta. Los segundos son del tipo caja negra, es decir, se conoce que para unas entradas determinadas se obtienen unas salidas sin importar demasiado lo que ocurre dentro. Ambos procesos requieren de experimentaciones adecuadas, si bien los de base física o analítica son más costosos debido a su minuciosidad. La mayoría de los manuales de fotogrametría o topografía al tratar los diversos métodos de captura o

levantamiento establecen modelos físicos. En el caso de procesos u operaciones cartográficas o de índole SIG (por ejemplo digitalización, rasterización, superposición de capas, etc.) también se pueden plantear modelos físicos para determinar el error que se genera en cada una de ellas.

12.3. Propagación de la incertidumbre

El trabajo con IG (datos y modelos), es un proceso complejo en el que se concatenan transformaciones que incrementan el nivel final de incertidumbre. Este proceso de concatenación de operaciones es clásico en geodesia (redes de 1º, 2º, 3º y 4º orden). En este caso existe una clara jerarquía en las incertidumbres y procesos aplicados. Esta situación también se da en los trabajos temáticos y SIG.

El álgebra de mapas permite generar nuevas BDG a partir de ecuaciones o reglas que transforman los valores de la(s) BDG(s) original(es). Las incertidumbres presentes en las BDG también se someten a dichas transformaciones. Por ello, el modelado de la propagación de la incertidumbre debe hacerse mediante representaciones matemáticas que dependerán del proceso involucrado, de hipótesis de propagación del error, y de ciertos conocimientos previos sobre la calidad de los datos y el propio comportamiento de las funciones aplicadas (por ejemplo el cálculo de la pendiente u orientación sobre un modelo de elevaciones de tipo malla). En algunos casos los modelos matemáticos no resultan operativos por su complejidad (intercorrelaciones, componente espacial, etc.), en estos casos se recurre a la simulación.

12.3.1. Operaciones numéricas

Sea una variable aleatoria $x_i(\cdot) = b_i(\cdot) + V_i(\cdot)$ con $i \in [1, p]$, donde $b_i(\cdot)$ son los valores almacenados por el sistema (coordenadas o atributos), y $V_i(\cdot)$ es una variable aleatoria explicativa del error o incertidumbre presente. Una transformación es pues cualquier operación $G(\cdot)$ que origina una salida $H(\cdot)$ que también posee un comportamiento aleatorio:

$$(1) \quad H(\cdot) = G [x_1(\cdot), x_2(\cdot), \dots, x_p(\cdot)]$$

Los modelos pretenden determinar el valor de la media de $H(\cdot)$ (μ_H) y de su incertidumbre expresada mediante su varianza (σ_H^2). En relación a este último parámetro conviene recordar que cuando se dispone de dos o más variables se habla de covarianzas.

Los métodos que se van a presentar son el Gauss y el Taylor para las linearizaciones de primer orden. Otras alternativas como Taylor de segundo orden, de Rosenblueth, geoestadística, etc. pueden estudiarse en Heuvelink (1998). Conviene indicar que estos métodos tienen una importante base analítica y que resuelven los problemas de manera “puntual”. Si se desea desarrollar aplicaciones de evaluación de la incertidumbre en un entorno más geográfico la simulación es una herramienta mucho más adecuada.

Método de Gauss: La ley de propagación de los errores para el caso de ecuaciones de tipo lineal se conoce como modelo de Gauss, de tal forma que si dos variables aleatorias multidimensionales (X, Y) se relacionan mediante una matriz de constantes R $\{ \vec{y} = R \vec{x} \}$, la ley que relaciona las matrices de varianza-covarianza (C_x y C_y) de ambas variables queda expresada de la forma:

$$(2) \quad C_y = R C_x R^T$$

Método de primer orden de Taylor: Si las funciones no son lineales sólo se puede aplicar el modelo de Gauss si previamente se procede a linealizarlas. El proceso consiste en utilizar la aproximación mediante el desarrollo en serie de Taylor:

$$(3) \quad f(x) = f(a) + \frac{(x-a)}{1!} f'(a) + \frac{(x-a)^2}{2!} f''(a) + \dots + \frac{(x-a)^n}{n!} f^{(n)}(a)$$

El desarrollo se realiza hasta el primer término, considerándose que el resto es despreciable. Lo que supone linealizar la función en un entorno del punto de interés mediante su tangente, por lo que es condición indispensable que la función sea continua y derivable ($G'(\cdot)$ debe existir).

12.3.2. Operaciones lógicas

Las operaciones lógicas son de gran interés y uso dentro del modelado cartográfico. En este apartado sólo se analizarán las básicas: intersección o “Y” lógico (AND) y unión u “O” lógico (OR). Ambas pertenecen a lo que se denomina lógica booleana, la cuál se basa en las denominadas funciones de pertenencia discretas (booleanas). La función de pertenencia de tipo booleano $FP^B(z)$ no es más que un proceso de reclasificación discreto de los atributos cuyos rangos de clase están bien definidos, por ejemplo:

$$(4) \quad \begin{aligned} FP^B(A_x) &= 1 && \text{si } b_1 \leq A \leq b_2 \\ FP^B(A_x) &= 0 && \text{si } A < b_1 \text{ ó si } A > b_2 \end{aligned}$$

En el modelado cartográfico es usual el encadenamiento de hipótesis para la obtención de cartografía derivada: se definen funciones de pertenencia conjunta de tipo booleano FPC^B , que

no son más que una extensión de las FP^B ya comentadas, por ejemplo: *SI pendiente >0 10% Y textura = arenosa Y cubierta vegetación =< 25% ENTONCES peligro erosión ES severo.*

Utilizando como métrica de la calidad global, para toda una cobertura o tema, el índice $P[E]$ equivalente al porcentaje de acuerdo (Pa), para el Y-lógico, en el caso de trabajar con dos capas, el modelo se formula de cualquiera de las siguientes formas:

$$(5) \quad P[\bar{E}_g] = P[\bar{E}_1 \cap \bar{E}_2] \quad P[\bar{E}_g] = P[\bar{E}_1]P[\bar{E}_2 / \bar{E}_1]$$

Y de una manera más general (n capas de entrada):

$$(6) \quad P[\bar{E}_g] = P[\bar{E}_1 \cap \bar{E}_2 \cap \dots \cap \bar{E}_n]$$

Para el O-lógico, la formulación analítica es equivalente con la salvedad que los valores son complementarios a la unidad.

12.4. Descripción de la calidad

Se va a seguir la filosofía de ISO 19100, en concreto la norma ISO 19113, sobre la calidad de la IG. El objetivo de esta norma es establecer los principios para describir la calidad de un conjunto de datos, e informar sobre la misma. La calidad tiene aspectos cualitativos o descriptivos (información no cuantitativa según ISO 19113), y aspectos cuantitativos (elementos de la calidad). Aspectos cualitativos son el linaje historia de los datos, el uso o recopilación de casos de uso reales y el propósito o descripción de los usos teóricos para los que los datos han sido producidos. La idea básica de los elementos de la calidad es la “medida” o determinación de la calidad. La obtención de dichas medidas se realiza en un proceso de evaluación, objetivo de ISO 19114. Finalmente, cabe destacar otro aspecto importante de esta Norma: da directrices para realizar el informe sobre la calidad en forma de metadatos.

12.4.1. Elementos generales

Es información de carácter general y no cuantitativo, de gran interés para conocer el objetivo e historial de la IG, así como para conocer posibles usos en aplicaciones distintas a las consideradas comúnmente. Esto se describe mediante los denominados “elementos generales de la calidad”: Propósito, uso y linaje.

Propósito: El propósito son las razones de la creación de la BDG e información sobre el uso al que se pretende destinar. Por ejemplo:

“La Carta Digital del Mundo es una base de datos global, digital, y de propósito general diseñada para soportar aplicaciones en los SIG.”(En MIL-D-89009, apartado 3.4.1 descripción del producto).

Uso: Descripción de la(s) aplicación(es) distinta(s) al propósito para las cuales el conjunto de datos geográficos se ha usado eficazmente. Por ejemplo:

***Uso n° X:** Desarrollo de series de mapas electrónicos: “Se ha preparado una serie mapas digitales en formato WHEAT a partir de la Carta Digital del Mundo; un conjunto de mapas digitales E1000K basados en las Cartas de Navegación Operacional de la Defense Mapping Agency (EE.UU.). Estos fragmentos de la Carta Digital del Mundo fueron importados con varios propósitos: facilitar la explotación de recursos naturales en los países en vías de desarrollo para impulsar el desarrollo de sus necesidades básicas, para proporcionar cartografía adecuada para trabajos de ayuda en el Tercer Mundo, y para facilitar conjuntos de datos de ejemplo para emplearlos con WHEAT. Se confía que proporcionando cartografía topográfica regional, en un formato fácil de emplear, se facilitará la exploración de las aguas subterráneas, la planificación agronómica y la logística de los proyectos de ayuda“. (En User’s Manual for Digital Chart of the World 1 Quadrangles, Geohidrology Section, Kansas Geological Survey, este uso ha sido hallado en Internet.)*

Linaje: Historial del un conjunto de datos geográficos, atendiendo fundamentalmente a las fuentes y proceso de producción. Por ejemplo:

Fuente: El contenido de la Carta Digital del Mundo está basado principalmente en los contenidos de la serie de Navegación Operacional a E1000K (para todas las regiones exceptuando la región antártica) de la Defense Mapping Agency. Las Cartas de Navegación Operacional utilizadas para crear el producto fueron generadas por la DMA entre los años 1974 y 1991.

Pasos del proceso: se generaron positivos sobre una base estable a partir de los negativos de reproducción originales (hasta 35 negativos por hoja de la Carta de Navegación Operacional) y se digitalizaron mediante un escaneado y conversión raster-vector o por medio de digitalización manual en formato vectorial. Los datos vectoriales fueron etiquetados con la información de sus atributos utilizando el software ARC/INFO. La transformación a coordenadas geográficas fue realizada usando las retículas de la proyección de cada hoja. La información digital fue depurada y enlazada en los bordes de hoja para crear grandes conjuntos de datos de carácter regional. Estos conjuntos fueron luego subdivididos en celdas 5°x5°, y convertidos desde ARC/INFO al formato VPF. A

continuación los datos fueron premasterizados para su difusión en CD-ROM. El control de calidad fue realizado para cada paso del proceso por un grupo independiente. El proceso fue completado en enero de 1991.

12.4.2. Elementos cuantitativos

Es información cuantitativa de la calidad. Se describe mediante los “elementos de la calidad” (*Data Quality Elements*), que vienen a ser los denominados tradicionalmente como componentes de la calidad del dato geográfico. Los propuestos por ISO 19113 son: exactitud posicional, temática y temporal, compleción y consistencia lógica. Para cada elemento se establecen subelementos que permiten concretar el aspecto de interés. El conjunto de elementos y subelementos indicados en la norma es un conjunto inicial que puede extenderse tanto como se necesite si se cumplen ciertas reglas de coherencia. La descripción de la pareja elemento/subelemento se realiza por medio de un paquete de seis descriptores obligatorios (ámbito, medida, procedimiento de evaluación, tipo de valor, resultado, unidad y fecha) que informan sobre cada medida.

Las medidas están normalizadas en ISO 19138. El objetivo es utilizar un conjunto conocido que pueda ser bien y fácilmente interpretado por productores y usuarios. La normalización consiste en una selección indexada acompañada de una ficha de especificación para cada medida. Otro aspecto importante de esta norma es que define “medidas básicas” de la calidad. Éstas se introducen para evitar la repetición en la definición de conceptos y evitar así problemas y ambigüedades. Las medidas son:

- **Conteo de errores:** Se expresan mediante medidas de conteo del número de errores (defectos) o casos correctos. Son adecuadas para los aspectos de la calidad en los que la medida es el conteo de ocurrencias de una circunstancia (error).
- **Estimación de la incertidumbre:** Se expresan mediante medidas de incertidumbre con base en un modelo estadístico y son adecuadas para los aspectos medibles.

Exactitud posicional: En una BDG la posición de un fenómeno del mundo real se describe por medio de valores (coordenadas) en un sistema de referencia. La exactitud posicional significa la mayor o menor proximidad entre las posiciones de los fenómenos representados y las posiciones consideradas como buenas o verdaderas. La componente posicional queda recogida como un elemento de la calidad en la Norma ISO 19113. Esta norma considera los siguientes subelementos:

- **Exactitud externa o absoluta:** Es la proximidad entre los valores de coordenadas indicados y los valores verdaderos o aceptados como tales.

- Exactitud interna o relativa: Se refiere a las posiciones relativas de los objetos de un conjunto de datos y sus respectivas posiciones relativas verdaderas o aceptadas como verdaderas.
- Exactitud para datos de rejilla: Es la proximidad de los valores de posición de los datos en estructura de malla regular a los valores verdaderos o aceptados como verdaderos.

Algunas instituciones incluyen dentro de ésta la fidelidad geométrica (por ejemplo escuadrado, alineamiento, paralelismo, línea de visión, armonía, etc.), y establecen la siguiente jerarquía: fidelidad geométrica >> exactitud relativa >> exactitud absoluta.

Dado que es la componente más tradicional del dato geográfico existen numerosas metodologías para evaluar o cuantificar esta componente (por ejemplo NMAS, EMAS, NSSDA, STANAG, etc.). En esta línea, y a modo de ejemplo, la Tabla 2 presenta la metodología NSSDA. En Ariza y col. (2007) se puede encontrar una revisión amplia de las principales metodologías. La Norma ISO 19138 propone numerosas medidas de incertidumbre que le pueden ser de aplicación a esta componente.

1.	Seleccionar una muestra de n puntos, siendo $n \geq 20$.								
2.	Calcular el error de cada punto en cada componente:								
	$e_{x_i} = x_{t_i} - x_{m_i} \quad e_{y_i} = y_{t_i} - y_{m_i} \quad e_{z_i} = z_{t_i} - z_{m_i}$								
3.	Calcular el error medio cuadrático de cada componente:								
	$EMC_X = \sqrt{\frac{\sum e_{x_i}^2}{n}} \quad EMC_Y = \sqrt{\frac{\sum e_{y_i}^2}{n}} \quad EMC_Z = \sqrt{\frac{\sum e_{z_i}^2}{n}}$								
4.	Obtener el valor $NSSDA_H$ horizontal según una de las siguientes expresiones:								
	<table style="border: none; width: 100%;"> <tr> <td style="border-right: 1px solid black; padding-right: 10px;"> si $EMC_X = EMC_Y$ </td> <td> $NSSDA_H = 1,7307 \cdot EMC_r = 2,4477 \cdot EMC_X$ </td> </tr> <tr> <td style="border-right: 1px solid black; padding-right: 10px;"> donde: $EMC_r = \sqrt{EMC_X^2 + EMC_Y^2}$ </td> <td></td> </tr> <tr> <td style="border-right: 1px solid black; padding-right: 10px;"> si $EMC_X \neq EMC_Y$ </td> <td> $NSSDA_H = 2,4477 \cdot 0,5 \cdot (EMC_X + EMC_Y)$ </td> </tr> <tr> <td style="border-right: 1px solid black; padding-right: 10px;"> y $0,6 < (EMC_{min} / EMC_{max}) < 1,0$ </td> <td></td> </tr> </table>	si $EMC_X = EMC_Y$	$NSSDA_H = 1,7307 \cdot EMC_r = 2,4477 \cdot EMC_X$	donde: $EMC_r = \sqrt{EMC_X^2 + EMC_Y^2}$		si $EMC_X \neq EMC_Y$	$NSSDA_H = 2,4477 \cdot 0,5 \cdot (EMC_X + EMC_Y)$	y $0,6 < (EMC_{min} / EMC_{max}) < 1,0$	
si $EMC_X = EMC_Y$	$NSSDA_H = 1,7307 \cdot EMC_r = 2,4477 \cdot EMC_X$								
donde: $EMC_r = \sqrt{EMC_X^2 + EMC_Y^2}$									
si $EMC_X \neq EMC_Y$	$NSSDA_H = 2,4477 \cdot 0,5 \cdot (EMC_X + EMC_Y)$								
y $0,6 < (EMC_{min} / EMC_{max}) < 1,0$									
5.	Obtener el valor $NSSDA_Z$ vertical según la siguiente expresión:								
	$NSSDA_Z = 1,9600 \cdot EMC_Z$								
6.	Incluir junto a la leyenda del mapa:								
	“Se ha verificado una exactitud horizontal de _ metros al 95% de nivel de confianza”. “Se ha verificado una exactitud vertical de _ metros al 95% de nivel de confianza”.								

Tabla 2. Metodología NSSDA para la determinación de la exactitud posicional

Se desea advertir, que todos los métodos aplicables sobre la componente posicional (planimétrica, X e Y independientes o sobre la Z) son igualmente aplicables a cualquier atributo cuantitativo de la IG para el que se pueda asumir una hipótesis de comportamiento gaussiano.

Exactitud temática: La componente temática de una cartografía es fundamental. Esta componente no sólo está presente en las denominadas cartografías temáticas. En toda cartografía topográfica existen aspectos temáticos muy importantes para la calidad del producto.

ISO 19113 considera la componente temática como un elemento (cuantitativo) de la calidad, distinguiendo los siguientes subelementos: corrección de la clasificación, corrección de los atributos cualitativos y exactitud de los atributos cuantitativos. Por tanto, se observan dos niveles distintos: el de las clases y el de los atributos, con distinción de si éstos últimos son cualitativos o cuantitativos. La corrección de la clasificación y de los atributos se mide de manera similar. Los atributos de carácter cuantitativo se tratan como variables unidimensionales, por lo que lo presentado para la posición es de plena aplicación para los mismos cuando. La corrección de la clasificación puede caracterizarse con diversas medidas, pero la más aplicada y potente es la matriz de confusión y los índices derivados de la misma (el porcentaje de acuerdo Pa y el índice $Kappa$). La norma ISO 19138 contempla estos índices.

A diferencia de lo que ocurría para la componente posicional, y a pesar de la importancia de esta componente, no existe ningún estándar relativo a la evaluación de exactitud temática.

La evaluación suele hacerse en forma de matriz de error (tabla de contingencia o matriz de confusión). Su ordenamiento suele ser tal que la verdad del terreno aparece en columnas, mientras que las unidades cartografiadas (BDG) aparecen en las filas. La tabla así formada presenta una visión general de las asignaciones correctas (elementos de la diagonal) y de las migraciones o fugas (elementos fuera de la diagonal). La formación de la matriz a partir de una muestra requiere de ciertas condiciones:

- Las clases que se establezcan deben ser independientes, mutuamente excluyentes, exhaustivas y en número suficiente.
- Se utilizarán como unidades de muestreo elementos puntuales, superficiales, o agrupaciones de aquellos.
- Deben usarse métodos de muestreo estadísticos que eviten la autocorrelación.
- Conviene el uso de métodos estratificados para asegurar la presencia de clases extrañas o minoritarias.

Los parámetros Pa y $Kappa$ resumen la matriz de error dado una visión global de la corrección. Éstos siguen un modelo binomial por lo que se pueden plantear contrastes de hipótesis sobre valores distintos, sobre tolerancias, etc., permitiendo aceptar o rechazar trabajos con base estadística. Tienen la siguiente interpretación y forma de cálculo:

- *Pa* (*Porcentaje de Acuerdo*). Es considerado como la probabilidad global de estar bien clasificado. Sus valores se encuentran en el intervalo $[0, 1]$, siendo el valor 1 el caso de total acuerdo. Para M clases, se calcula como el cociente entre los valores presentes en la diagonal (n_{ii}) frente al total de casos (N).

$$(7) \quad Pa = \frac{1}{N} \sum_{i=1}^M n_{i,i} \quad \text{y con varianza: } \sigma_{Pa}^2 = \frac{Pa(1-Pa)}{N}$$

- *K* (*Coefficiente Kappa de ajuste*). Su valor da idea del porcentaje de acuerdo obtenido en una clasificación una vez se ha eliminado la parte que se debe al azar. Sus valores se encuentran en el intervalo $[0, 1]$, siendo el valor 1 el caso de total acuerdo. Por definición siempre muestra valores inferiores a *Pa*.

$$(8) \quad K = \frac{Pa - Ca_{ps}}{1 - Ca_{ps}} \quad \text{y con varianza: } \sigma_K^2 = \frac{Pa(1-Pa)}{N(1 - Ca_{ps})^2}$$

$$\text{Donde: } Ca_{ps} = \frac{1}{N^2} \sum_{i=1}^M n_{i+} n_{+i}$$

Exactitud temporal: La inclusión del tiempo en las BDG es uno de los campos de investigación y desarrollo actuales. En otras aplicaciones (por ejemplo gestión de banca, nóminas, stock, etc.) está superada pero su incorporación a los datos espaciales es más compleja. La forma más sencilla de incluir el tiempo es como un atributo, aunque lo deseable sería poder gestionarlo como una verdadera coordenada (XYZT) más de la IG. El tiempo es una característica fundamental para poder juzgar la bondad de muchos datos, pero la gestión de esta información lleva a plantearse a qué tiempo nos referimos:

- Al tiempo lógico o del evento, es decir cuando ocurre el cambio en el mundo real.
- Al tiempo de observación o evidencia, es decir cuando se observa.
- Al tiempo en el que se incluyen los cambios en la base de datos, denominado tiempo de transacción, tiempo de la base de datos o tiempo de captura.

Por supuesto, los valores anteriores no suelen ser idénticos. Lo común es registrar el tiempo de observación, dado que raras veces se puede estar al acecho de la ocurrencia de un evento. En este sentido, también hace falta definir una especie de modelo abstracto temporal. Por otra parte la propia base de datos suele guardar registro de las transacciones que se realizan, por lo que este tiempo también queda recogido en el sistema, con su importancia para la genealogía de los datos. Otro aspecto de interés es la necesidad de tomar una unidad de medida (segundos, minutos, horas), un origen de tiempos, etc., o decidimos por usar una simple escala ordinal (por ejemplo, la geológica).

Se considera que la estructura de los datos temporales se gestiona mejor asumiendo un modelo topológico. Esto permite hacer un estudio paralelo entre contigüidad temporal y contigüidad geográfica y aplicar a la exactitud temporal los mismos procedimientos de la exactitud topológica y, por tanto, de la coherencia lógica de la misma.

Desde un punto de vista mucho más aplicado tal vez lo que más importa al usuario es la actualidad, es decir, la proximidad temporal entre las fechas de captura y uso actual (exactitud temporal). La Tabla 3 presenta algunos índices y medidas tomados del proyecto de norma europea prEN 12656.

Indicador	Medida	Comentario
Error en tiempo	Error medio temporal	Error en las unidades establecidas
Actualidad	Fecha de última actualización	Relacionada con la compleción. Definible por clases.
Ratio de cambio	Número de cambio por unidad de tiempo	Estimación de la tasa de cambio por unidad establecida. Junto con la última fecha de actualización puede dar idea de los cambios.
Lapso temporal	Unidad de tiempo	Promedio de tiempo entre cambios en la verdad nominal y su representación en los datos.
Validez temporal	Validez	Validez respecto a su uso en función de la fecha: válido, fuera de uso, etc.

Tabla 3. Indicadores y medidas de la calidad para la exactitud temporal

Fuente: prEN 12656:1996.

Los subelementos considerados por ISO 19113 son: exactitud de la medida del tiempo, consistencia temporal, validez temporal. Éstos hacen referencia, respectivamente, a la corrección de: las referencias temporales asignadas a un elemento (informe del error en la medida del tiempo asignado), los eventos o secuencias ordenadas, si se indican, y a la validez de los datos respecto al tiempo. La norma ISO 19138 propone medidas de incertidumbre lineal para la exactitud de la medida del tiempo, y de conteo/ratio de errores para la validez temporal, para la consistencia temporal no ofrece ninguna medida.

Compleción: La compleción es la cualidad de completo. Para conocer si una cosa está/es completa se deberá comparar con un patrón. En el caso de las BDG la compleción debe entenderse frente a unas especificaciones de producto. La compleción no es evaluada por comparación directa con el mundo real, sino frente al terreno nominal, es decir el subconjunto de elementos del mundo real que cumplen con las especificaciones de producto (universo de discurso). Por lo general, esta componente se evalúa mediante trabajos muestrales con visita a campo.

La compleción es una cualidad necesaria de toda BDG. La dinámica del territorio afecta mucho a la compleción de una BDG en una fecha determinada. Por ello la compleción debe referirse siempre a

un momento temporal, y a la hora de utilizar la BDG en otro momento se ha de tener en cuenta el lapso temporal transcurrido y las tasas de cambio.

La norma ISO 19113 considera este elemento y define dos subelementos: omisión y comisión. Las omisiones son los elementos en falta y las comisiones los elementos en exceso. Una BDG de calidad será aquella en la que sus elementos sean correctos y coincidentes con los que existen en la realidad. La norma ISO 19138 propone para este elemento las medidas que se corresponden con contadores y tasas de error.

Existe una compleción de objetos, la cual indica en qué medida, de omisión o comisión, los elementos presentes en la BDG cumplen con las prescripciones establecidas. También existe una compleción de atributos, aquí puede ocurrir que no todos los atributos de un objeto hayan sido recogidos. Por otro lado la compleción del modelo cartográfico hace referencia a la inclusión de las clases, atributos y relaciones necesarias para plasmar adecuadamente la realidad en el modelo cartográfico.

En cuanto a la compleción de objetos, hay que comprobar que los datos representan todas las entidades del universo de discurso. Éste estará claramente definido si hay una referencia explícita (listado, inventario, definiciones, etc.). Si el universo de discurso viene definido de forma implícita, habrá que recurrir a la interpretación, el conocimiento de expertos y la opinión individual. Una vez el universo de discurso queda definido, la compleción de objetos se establece simplemente mediante una comparación de la BDG frente al terreno nominal. El informe de compleción debe incluir información exacta acerca de la forma de derivar el universo de discurso, indicando las listas, estándares, mapas, inventarios, informaciones de experto, etc., que sean menester.

Consistencia lógica: La coherencia lógica es un requisito fundamental en los productos digitales. Hace referencia al grado de conformidad respecto a la estructura interna descrita por unas especificaciones (estructura, atributos, relaciones de compatibilidad, etc.). La codificación digital está estrechamente ligada a lógico y formal por la matemática y los elementos lógicos que soportan tanto a esos datos, sus modelos y a las aplicaciones de software que los gestionan y explotan. No existe consistencia lógica en una base de datos cuando se encuentran contradicciones lógicas entre los elementos contenidos en la misma.

Es un elemento de la calidad que, normalmente, puede controlarse con inspecciones al 100%, automatizadas por medio de rutinas de software. Las rutinas se aplican a las BDG y chequean mediante reglas todas y cada una de las especificaciones impuestas. Este chequeo puede

permitir la corrección automática de algunos problemas, o ser la manera de aviso para un trabajo de corrección posterior por parte de algún operario.

ISO 19113 propone los siguientes subelementos:

- Consistencia conceptual: Tiene en cuenta la adherencia a las reglas del esquema conceptual.
- Consistencia de dominio: Tiene en cuenta la pertenencia de los valores a los dominios o recorridos establecidos (por ejemplo un intervalo). Los valores fuera de dominio o no contemplados dentro de una lista que enumera los casos posibles son el objeto de este subelemento.
- Consistencia de formato: Grado en que los datos se almacenan según la estructura física establecida. La adherencia al formato puede ser tanto a un formato analógico como digital. En el caso analógico deberá ser evaluada por un operario y seguramente se realizará por medio de muestreos. En el caso digital las pruebas de adherencia al formato pueden ser fácilmente implementadas de forma automática dado que las especificaciones estarán claramente indicadas.
- Consistencia topológica: Corrección de las características topológicas explícitamente codificadas de un conjunto de datos. Se deben cumplir las reglas de la topología, ejemplo de estas reglas son la ecuación de Euler o los chequeos sobre las cadenas de nodos (cadenas de celdas alternativas de 1 y 2 dimensiones alrededor de cada nodo), test de cadenas de polígonos (ciclos de nodos y celdas de una dimensión que forman el cerramiento de una celda de dos dimensiones), etc.

Por su parte, la norma ISO 19138 propone hasta 19 medidas en este ámbito, que pertenecen a las tipologías de contadores de error y tasas de error, principalmente.

12.5. Métodos estadísticos de determinación de la calidad.

Según ISO 19114, para la evaluación de la calidad se pueden aplicar métodos directos (aquellos que se basan en la comparación o medida). Los métodos directos internos utilizan datos de la propia BDG (por ejemplo consistencia lógica topológica). Los métodos directos externos necesitan fuentes externas a la BDG, como por ejemplo en la verificación de un topónimo una fuente más exacta o la consulta a los lugareños. Dentro de los métodos directos existen procesos plenamente automatizables, que permiten lo que se denomina una inspección o control al 100%. No obstante, muchos elementos de la calidad requieren procesos de control con un desarrollo manual, lo cual lleva al uso de técnicas de muestreo o inspección.

Este apartado presenta la teoría de muestreos como metodología estadística para la estimación de la calidad. La idea es dar una guía para evitar costosas inspecciones manuales sobre el 100%. Las bases estadísticas del muestreo deben ser conocidas tanto para entender la problemática (posibilidades y limitaciones) como para diseñar muestreos *ad hoc*. También se introducen las normas ISO 2859 y 3951 que dan la perspectiva del uso de estas técnicas estadísticas en los entornos de producción.

El objetivo de la teoría de muestras es la obtención de información sobre una característica A (por ejemplo error posicional, omisiones, comisiones, etc.) de una población “X” de tamaño “N”, a partir de un subconjunto de la misma “x”, de tamaño “n”, que se denomina muestra y que cumple que “ $n < N$ ”. Se llama fracción muestral ($f=n/N$, / $0 < f <= 1$) a la parte de la población investigada. El uso de muestras frente al análisis de toda la población tiene una clara ventaja económica dado que, al ser $n < N$, resulta un estudio más ventajoso.

Para que los resultados de una muestra sean de valor respecto a la población deben ser representativos de dicha población, esto se asegura con muestreos estadísticos bien diseñado. El diseño de muestreos tiene como fin recoger la máxima información sobre la población a partir de la muestra, con la máxima precisión y el mínimo coste posible.

12.5.1. Visión general del diseño de muestreos

Para que los resultados de la evaluación tengan una representatividad adecuada hace falta que el muestreo tenga una base estadística, por ello esta es una fase crítica. Para que la inferencia sea correcta la muestra debe conformarse de tal manera que sea representativa de la población. Por ello se plantean tres aspectos básicos e interrelacionados: a) establecer la población de interés, b) estimar el tamaño de la población, y c) determinar del tamaño de la muestra. El primer elemento es fundamental dado que condiciona de manera radical la orientación que se da a la investigación. El establecer la población, o marco de muestreo, puede parecer algo obvio, pero existen diversas alternativas que complican la decisión. Su establecimiento no consiste sólo en definir qué elementos conforman la población, hace falta tener un listado completo de todos ellos. Éste es un aspecto crítico de algunos muestreos, por lo que existen alternativas en las que no es necesario conocer, en el sentido anterior, el marco. Una vez determinada la población conocer su tamaño es fundamental desde el punto de vista estadístico. Hay casos en los que se desconoce este tamaño poblacional y se necesita un muestreo previo para determinarlo (por ejemplo una población de animales salvajes). Junto al tamaño poblacional, al plantear el tamaño

de una muestra se hace necesario previamente tomar alguna decisión sobre los siguientes aspectos:

- Característica a estimar.
- Tipo de variable a muestrear y características de su función de distribución.
- Nivel de confianza.
- Precisión requerida en los trabajos.
- Esquema de muestreo.

La característica “ A ” a estimar hace referencia a qué parámetro se quiere determinar y puede referirse a un valor medio, a una proporción, a una suma de valores, etc. En nuestro caso las características están ligadas a los subelementos, por lo que podrá tratarse de un porcentaje (por ejemplo número de violaciones de una regla), una variable cuantitativa (por ejemplo el error medio en posición), etc.

Estimación es la obtención de un valor relativo a la característica “ A ” de la población “ X ” a partir de un conjunto de observaciones-mediciones $\{a_1, a_2, a_3, \dots, a_n\}$ obtenidas de la muestra “ x ” de tamaño “ n ”. Para la obtener la estimación se utilizan estimadores. Un estimador $\hat{\theta}(A)$ es una fórmula o modelo que permite derivar el valor θ de la característica poblacional “ A ” a partir de las observaciones. El estimador debe ser insesgado, eficiente y consistente. La insesgadesz significa que la estimación no introduzca desviaciones sistemáticas y la eficiencia que, a igual tamaño de muestra, sea capaz de dar una mejor estimación que otro estimador. Si el estimador es consistente se cumple que al aumentar el tamaño de muestra n la estimación tiende al valor real, aumentando la confianza $1 - \alpha$ o disminuyendo el error de la estimación E .

Respecto al tipo de variable básicamente existen dos opciones: cualitativas y cuantitativas. Para las primeras el trabajo consiste en determinar si la categoría asignada a las mismas es o no correcta, aspecto se analiza utilizando una distribución binomial (correcto o incorrecto); o multinomial. Para las variables cuantitativas lo más correcto es determinar la función de distribución que siguen y, a partir de ella, derivar el tamaño adecuado. En la mayoría de las aplicaciones se asume la normalidad de los datos. Según lo anterior, la característica a estimar en el caso de variables cualitativas suele ser la proporción de cada una de ellas, mientras que en el caso de las variables cuantitativas suele ser el valor medio poblacional e incluso la variabilidad.

El nivel de confianza ($1-\alpha$) de una prueba estadística se refiere a la probabilidad del error de tipo I (α) que se asume. Esta probabilidad es que corresponde al rechazo de elementos que realmente cumplen la condición. Normalmente se asume un nivel de confianza del 95%, es decir, se limita el error de tipo I a un 5% de casos.

El error de la estimación es la diferencia entre el valor real de la característica en la población y el obtenido por medio del estimador. Este error se suele acotar $E \leq \varepsilon$, con una confianza de $1-\alpha$. La precisión del muestreo se refiere al error máximo admisible en el proceso y suele tomarse como la semiamplitud del intervalo de confianza. Este parámetro influye mucho en el tamaño de la muestra. Para este tipo de trabajos es usual que se tome en un 5%.

Los elementos anteriores están también muy condicionados por la propia variabilidad de la muestra, de la cual se necesita cierta información para poder proceder a plantear la argumentación matemática que permite derivar los tamaños. Este es un aspecto crítico de los diseños. Para afrontarlo hay varias opciones, entre ellas desarrollar un muestreo piloto para obtener esta información. Otra opción es tomar un valor correspondiente al peor de los casos posibles.

El esquema de muestreo se refiere a cómo se orienta la captura de las muestras. Las alternativas más comunes son: el muestreo aleatorio simple, el muestreo sistemático, el muestreo estratificado y el muestreo por conglomerados.

Muestreo aleatorio simple (MAS): Es el caso más simple de muestreo probabilístico. Su base es que todos los elementos de la población tienen la misma probabilidad de ser incluidos en la muestra, por lo que todas las muestras son equiprobables. Es el proceso base de muchos métodos de control. Suele ser la tipología más eficiente desde un punto de vista estadístico. La formulación matemática y los niveles de información necesarios para aplicarlos son los más reducidos.

Muestreo sistemático (MS): Consiste en elegir las distintas unidades que conforman la muestra espaciándolas regularmente “ k ” unidades en el marco. Si los “ n ” elementos de la muestra son los más heterogéneos posibles, el muestreo sistemático es más eficiente que el aleatorio, pero si están relacionadas las observaciones sucesivas puede ser ineficiente (caso variaciones periódicas). El MS tiene varias ventajas: a) el inicio del marco se establece al azar, por lo que cualquier elemento de la población tiene la misma probabilidad de ser elegido, b) es fácil de aplicar y no es necesario conocer el tamaño N de la población, y c) suele dar resultados representativos, y es poco costoso y menos sujeto a errores en el trabajo de campo. Como

inconvenientes se puede apuntar que no es posible obtener todas las muestras posibles. Además, si en la ordenación del marco existe alguna periodicidad la muestra puede ser ineficiente. Lo anterior se puede amortiguar utilizando varios muestreos con puntos de inicio distintos, o introduciendo cierta aleatoriedad en el paso.

Muestreo estratificado (ME): Es común que una población se pueda considerar dividida en grupos o estratos. Un estrato es un conjunto de elementos de la población que comparten alguna(s) propiedad(es). Los estratos deben ser lo más homogéneos internamente y lo más heterogéneos unos con otros. Deben ser excluyentes unos de otros y cubrir toda la población (exhaustividad). El ME permite optimizar los recursos destinados al muestreo. La característica poblacional que interesa se estima para cada uno de los estratos separadamente lo que permite obtener una información desagregada que puede ser de interés. El ME permite una mejora en la precisión (minimización de la varianza) que aumenta en función de la semejanza de los individuos considerados en cada estrato y de la disimilitud de los que están dentro de estratos diferentes. Por todo lo anterior, el trabajo de campo más fácil y menos costoso. Los resultados obtenidos por estratos pueden compararse entre sí dando información de valor. Al hacer grupos más homogéneos que la población las varianzas muestrales son menores. Se incrementa la eficiencia de las estimaciones siempre que el criterio de estratificación sea adecuado. Sin embargo también posee algunos inconvenientes, básicamente la definición de los estratos, que puede no ser trivial (se necesitan criterios, por ejemplo tamaño \times desviación = cte.) y la distribución de la muestra entre los estratos (afijación), que debe ser estudiada convenientemente (por ejemplo proporción al peso de cada estrato, según variabilidad, etc.).

Muestreo por conglomerados (MC): Al realizar un muestreo de una población grande o muy dispersa los costes de muestreo y la duración de los trabajos de campo pueden dispararse si se utilizan las alternativas MAS, ME, MS. El MC puede soslayar este problema si bien puede tener una mayor ineficiencia en las estimaciones. El MC consiste en dividir la población en " M " grupos o conglomerados disjuntos y exhaustivos que presenta cierto agrupamiento espacial con objeto de reducir los costes de muestreo. Los conglomerados deben ser lo más heterogéneos posible, lo ideal es que se parezcan, en la mayor medida de lo posible, a la población. De entre estos " M " conglomerados se elige una muestra aleatoria simple de " m " conglomerados dentro de los cuales se observan todas las muestras. Es preferible usar conglomerados pequeños y mayor número. La información que proporciona este método depende pues del número de conglomerados " m " que forman parte del muestreo, del tamaño relativo de éstos y del propio conjunto de elementos que se incluyen en los conglomerados. En general, a igual tamaño muestral, el MC es menos eficiente que el ME e incluso que el MAS. Sin embargo, aunque para un mismo tamaño muestral el MC es menos eficiente que el ME o el MAS, a igual coste total

del muestreo el diseño por conglomerados permite, habitualmente, obtener tamaños muestrales mayores que los otros diseños, lo que supone una mayor eficiencia si el presupuesto es fijo. Es apropiado para grandes proyectos.

12. 5.2. Determinación de tamaños de muestrales (bases estadísticas)

A continuación se presentan los aspectos básicos de determinación de tamaños de muestras para situaciones de variables cualitativas y cuantitativas, siempre bajo la perspectiva de muestreo aleatorio simple (MAS).

Tamaño de muestreo para estimar una proporción (binomial): Para el caso binomial sólo hay dos alternativas, el elemento es correcto o incorrecto, lo que significa su adecuación para casos de variables cualitativas dicotómicas que se consideran distribuidas como una binomial $B(n, p)$ de parámetros n (número de elementos), y p (probabilidad de asignación correcta, y parámetro complementario $q = 1 - p$, o probabilidad de asignación incorrecta). En cartografía esta opción es adecuada, entre otros, para los errores de omisión y comisión, corrección, etc. El tamaño de muestra se formula por la siguiente expresión:

$$(9) \quad n = \frac{N \cdot P \cdot Q}{P \cdot Q + \frac{(N-1) \cdot \epsilon^2}{z_{\alpha/2}^2}} \quad \text{donde:}$$

p	Probabilidad de asignación correcta.
$q = 1 - p$	Probabilidad de asignación incorrecta.
N	Tamaño de la población.
n	Tamaño de la muestra.
ϵ	Error permitido.
$Z_{\alpha/2}$	Estadístico correspondiente a la distribución normal con significación α .

Simplificando cuando N es grande, y tomando $p = q = 0.5$, que es el caso más desfavorable según la variabilidad de los datos, se tiene el mayor tamaño de muestra: $n = Z_{1-\alpha/2}^2 / 4 \epsilon^2$

Tamaño de muestreo para estimar una media: La normalidad suele ser la hipótesis más asumida en el control de variables cuantitativas. En algunos controles cartográficos usualmente se controla un valor medio correspondiente no tanto a una población de mensurables existente en la realidad (por ejemplo elevaciones) como de las discrepancias o errores entre los valores correspondientes a un conjunto de datos seleccionados del producto y los valores para esos mismos elementos en una fuente de mayor exactitud. La siguiente expresión permite calcular el tamaño muestral:

$$(10) \quad n = \frac{N\sigma^2}{\sigma^2 + \frac{Ne^2}{Z_{\alpha/2}^2}}$$

12. 5.3. Procedimientos de muestreo para la inspección por atributos y variables.

Las normas ISO 2859 e ISO 3159 tienen gran importancia en el control de calidad de procesos industriales (aceptación por muestreo). La norma ISO 19114 constata su potencial de aplicación a la IG pues permiten decidir la aceptación/rechazo de productos en un marco estadístico en el que se asumen ciertos riesgos (riesgos de productor y de usuario), bajo el beneficio de evitar la inspección al 100%. Estos métodos estadísticos comparten sus bases con los ya analizados.

Las normas ISO 2859 e ISO 3951 establecen planes de muestreo para la aceptación. Estos planes pueden ser de aplicación a: elementos finales, componentes y materias primas, actividades, materiales en proceso, existencias de almacén, operaciones de mantenimiento, datos o registros, procedimientos administrativos... De manera práctica las normas ISO 2859 e ISO 3159 presentan las instrucciones para la utilización de un conjunto de tablas y gráficos que determinan un plan de muestreo eficaz. Conociendo el tamaño del lote N y la calidad considerada, las normas facilitan el tamaño de muestra n y el número máximo de unidades defectuosas que se admiten en la inspección (número de aceptación, Ac) para la inspección por atributos o la constante de aceptabilidad (k , basada en estimaciones de la tendencia central y la variabilidad) para la inspección por variables. El Nivel de Calidad Aceptable (NCA) es el concepto más importante: el máximo número de defectuosos que admite el cliente como promedio de los porcentajes de defectuosos que aparecen en los lotes. Este valor se expresa en porcentaje.

ISO 2859 hace referencia a los procedimientos de muestro para la inspección por atributos, mientras que ISO 3951 lo hace para la inspección por variables. Son *variables* las características susceptibles de ser medidas sobre una escala continua (por ejemplo un error de posición). Se denominan *atributos* las características cuya consideración hace que una unidad pueda ser clasificada como “buena” o “defectuosa” (por ejemplo el cumplimiento de una relación topológica, una omisión/comisión).

Inspección por atributos: De la inspección por atributos, lote a lote, se encarga la norma ISO 2859-1. A los lotes se les impone: ser homogéneos y de tamaño adecuado, dado que las muestras a tomar en lotes grandes son, en %, más pequeñas que para lotes de menor tamaño. Cada lote estará formado por unidades de producto (por ejemplo dato individual, un conjunto de

datos, una hoja, etc.). La norma distingue tres severidades en la inspección: normal, rigurosa y reducida. La normal tiene un criterio de aceptación que asegura al fabricante una alta probabilidad de aceptación cuando la calidad media de su proceso es mejor que el *NCA*. La rigurosa, para el mismo tamaño de muestra, reduce *Ac* y se utiliza cuando se tienen indicios de que la media del proceso es peor que el *NCA*. La reducida propone un tamaño muestral más reducido para un criterio de aceptación comparable al de la normal y se utiliza cuando se tienen indicios de que la media del proceso es mejor que el *NCA*. Durante la inspección de una serie de lotes, la norma indica cuáles son los criterios de paso entre estas opciones. Además, se distinguen diferentes niveles de inspección que fijan la cantidad relativa de la inspección. El nivel II es el que debe utilizarse a menos que se indique otro expresamente. Mediante tablas, con el nivel de inspección (por ejemplo el normal II) y el tamaño del lote *N* la norma ofrece un código que permite determinar, para el *NCA* exigido, el tamaño de la muestra *n* y el valor de *Ac*.

Inspección por variables: La norma ISO 3951-1 establece los planes y reglas de muestreo para la inspección por variables en lotes, siendo complementaria de la norma ISO 2859, con la que mantiene una filosofía, reglas y vocabulario comunes, mostrando numerosas semejanzas. Esta inspección genera una información más completa respecto a la característica de calidad que se inspecciona. Esta característica debe expresarse mediante un valor numérico y, además, debe seguir una función de distribución Normal. ISO 3951 distingue varios métodos de cálculo de la variabilidad, tal que afectan al tamaño de muestra. Los métodos más adecuados son los denominados “ σ ” y “*s*”. El método “ σ ” requiere el menor tamaño de muestra, pero debe determinarse primero el valor de σ de la característica. La norma recomienda comenzar por el método “*s*”. No obstante, si la desviación típica es estable y conocida, se puede aplicar el método “ σ ”, aceptando que σ es igual a la media cuadrática de los valores de *s*. Mediante tablas, con el nivel de inspección en uso y el tamaño del lote *N*, la norma ofrece un código que permite determinar, para el *NCA* exigido, el tamaño de la muestra *n* y un valor *k* (constante de aceptabilidad) que se compara con una estimación obtenida de la muestra para decidir si se acepta o rechaza el lote.

12.6. Calidad de los procesos

El entorno influye notablemente en que un producto tenga mayor o menor calidad técnica. Pero es más, la propia filosofía de cada productor afecta a la definición del producto, a su calidad teórica. Por todo lo anterior, la perspectiva de la calidad del dato debe ser integrada dentro de la organización productora de una manera coherente. Esto se consigue con los sistemas de aseguramiento de la calidad, o en su versión más avanzada con los Sistemas de Gestión de la

Calidad (SGC). Con los primeros se pretende asegurar que las especificaciones del producto sean efectivamente alcanzadas en la producción. Los SGC tienen una perspectiva más amplia que desemboca en los sistemas de excelencia donde existe una clara orientación hacia la mejora continua y la satisfacción del cliente.

Un SGC es aquella parte del sistema de gestión que está enfocada al logro de los resultados en relación a los objetivos de la calidad, para satisfacer las necesidades, expectativas y requisitos de las partes interesadas (ISO 9000:2000). La gestión de la calidad debe incluir y coordinar las actividades de la organización que tienen como objetivo implantar y delimitar la política de la calidad, los objetivos y las responsabilidades mediante la planificación de la calidad, control de la calidad, aseguramiento de la calidad y mejora de la calidad. La gestión de la calidad es pues la función directiva que pretende obtener el grado de calidad establecido para la producción con la reducción de costes. Debe quedar recogida en la estructura organizativa de la empresa.

Desde un punto de vista de la gestión, la calidad consiste en hacer las cosas bien a la primera, pues cuando mejor pueden descubrirse y evitarse las causas de los defectos o errores es en el origen de los mismos, siendo las consecuencias tanto menos costosas cuanto antes se descubren. La idea importante de esta filosofía es pues que la calidad se hace, no se controla. No se trata tanto de rechazar los productos defectuosos como de no llegar a producirlos, es decir, detectar las posibles causas de defecto y desarrollar las actividades preventivas necesarias para que se tenga la confianza de que los productos que se generan cumplan con las especificaciones marcadas. La calidad la hacen las personas, por lo que el factor humano requiere una especial atención.

Los SGC quedan definidos, en una línea básica, según el modelo establecido como estándar a nivel mundial en las normas de la serie ISO 9000. El objetivo de las normas ISO 9000 es mejorar la organización empresarial, eliminar los costes superfluos, la no calidad, los costes de control, arreglo y rectificación, generando un sistema que, mediante los canales adecuados de comunicación, permita evolucionar mejorando como forma de supervivencia y desarrollo de la propia empresa gracias al aseguramiento de la calidad y el enfoque al cliente. Las normas ISO 9000 se conforman como una familia que establece un marco general y adecuado para su aplicación a cualquier organización, independiente de su tamaño y tipo, a la hora de implementar y operar un sistema de gestión de la calidad que sea eficaz.

Las normas son:

- ISO 9000:2000. Establece los fundamentos de los sistemas de gestión y la terminología a usar para evitar ambigüedades.
- ISO 9001:2000. Especifica los requisitos para los sistemas de gestión de la calidad. Esta propuesta es genérica, supone un marco conceptual aplicable a cualquier caso. Es la propuesta normativa de ISO para aquellas organizaciones con fines contractuales o de certificación, aunque también puede ser utilizada a nivel interno. Los requisitos de los productos quedan fuera del alcance de esta normativa.
- ISO 9004:2000. Forma un par consistente con la ISO 9001:2000 donde se incluyen recomendaciones sobre un rango más amplio de objetivos para un SGC, orientadas a mejorar las prestaciones globales de una organización en el camino de la “excelencia en los negocios” (mejora del desempeño de la organización, de la satisfacción al cliente y de las demás partes implicadas).

El diseño e implantación de un SGC depende de las necesidades concretas de una organización, de sus objetivos particulares, de los productos y los servicios suministrados, y de los procesos y prácticas específicas empleadas. El objeto de estas normas no es imponer la uniformidad de los SGC, sino establecer un modelo adaptable a cualquier tipo de empresa o institución. La aplicación de estas normas permite a las empresas obtener un certificado de calidad, para lo cual han de someterse a auditorias o inspecciones. No obstante, las organizaciones también pueden adoptarlas como un marco para la mejora, sin pretender la certificación.

Bibliografía

Fuentes de error:

- [1] Ariza, F.J. (2002). Control de Calidad en la Producción Cartográfica. Ra-Ma. Madrid.
- [2] Aronoff, S. (1989). Geographic Information Systems: A management perspective. WDL Publications. Ottawa, Canada.
- [3] Burrough P.A. (1986). Principles of Geographical Information Systems for Land Resources Assessment. Clarendon Press, Oxford.
- [4] Hunter, Gary J.,(1998) Managing Uncertainty in GIS. NCGIA Core Curriculum in GIScience, <http://www.ncgia.ucsb.edu/giscc/units/u187/u1871.html>
- [5] ISO (1995). Guide to the expression of uncertainty in measurements. International Organization for Standardization. Geneva.
- [6] Kenneth E., Donald, J. The Geographer's Craft Project. Department of Geography, The University of Colorado at Boulder.
http://www.colorado.edu/geography/gcraft/notes/error/error_f.html
- [7] Maling, D. (1989). Measurements from Maps. Pergamon Press. Oxford.
- [8] Veregin, H. (1989). A Taxonomy of Errors in Spatial Data Bases. NCGIA Technical Paper 89-12, NCGIS, Santa Bárbara, California.

Propagación de la incertidumbre:

- [9] Ariza, F.J. (2002). Control de Calidad en la Producción Cartográfica. Ra-Ma. Madrid.
- [10] Burrough, P.A. (1986). Principles of Geographical Information Systems for Land Resources Assessment. Clarendon Press, Oxford.
- [11] Giordano, A., Veregin, H. (1994). Il controllo di qualità nei sistemi informativi territoriali. Il Cardo, Venezia.
- [12] Heuvelink, G. (1998). Error propagation in Environmental Modelling. Taylor & Francis. Londres.

Descripción de la calidad

- [13] Guptill, S., Morrison, J. (Ed) (1995). *Elements of Spatial Data Quality*. ICA, Elsevier, Oxford.
- [14] ISO 19113:2002. *Geographic information - Quality principles*.
- [15] ISO 19138:2006. *Geographic Information - Quality Measures*.

Componente posicional

- [16] Ariza, F.J., Atkinson, A.D.; Nero, M. (2007). *Análisis de algunas metodologías de evaluación de la componente posicional*. Topografía y Cartografía, Vol XXIV, nº 140, pp 32- 45.
- [17] STANAG (1998). *STANAG 7074: Digital Geographic Information Exchange Standard (DIGEST)*. North Atlantic Treaty Organization Standardization Agreement. Brussels
- [18] ASCE (1983). *Map Uses, scales and accuracies for engineering and associated purposes*. American Society of Civil Engineers, Committee on Cartographic Surveying, Surveying and Mapping Division, New York.
- [19] ASPRS (1990). Accuracy standards for large scale maps. *PE&RS*, vol. 56, nº7, 1068-1070.
- [20] FGDC (1998). *FGDC-STD-007: Geospatial Positioning Accuracy Standards, Part 3*. National Standard for Spatial Data Accuracy. Federal Geographic Data Committee, Reston.
- [21] USBB (1947). *United States National Map Accuracy Standards*. U.S. Bureau of the Budget. Washington.
- [22] USGS (1997). *Standards for Digital Elevation Models. Part 3: Quality Control, Standards for Digital Elevation Models*. United States Geological Survey, Reston.

Componente temática

- [23] Congalton, R.; Green, K. (1998). *Assessing the accuracy of remotely sensed data: Principles and Practices*. Lewis Publishers. Boca Raton, Florida.

Componente temporal

- [24] Langran, G. (1992). *Time in Geographic Information Systems*. Taylor & Francis, London.
- [25] Ott, T.; Swiaczny, F. (2001). *Time-Integrative GIS*. Springer, Berlin.
- [26] Christakos, G., Bogaert, P., Serre, M. (2002), *Temporal GIS*. Springer, Berlin.

Métodos estadísticos para la determinación de la calidad

- [27] Ariza, F.J. García, J.L; Amor, R. (2004). *Casos prácticos de calidad en la Producción Cartográfica*. Universidad de Jaén. Jaén.
- [28] Ariza, F.J., García, J.L. (2007). *Curso “Determinación de la calidad de datos geográficos”*. IGN, Madrid.
- [29] Cochran, W.G. (1977). *Sampling Techniques*. John Wiley & Sons. Nueva York.

- [30] ISO 19114:2003. *Geographic information - Quality evaluation procedures.*
- [31] ISO 2859-1:1999. *Sampling procedures for inspection by attributes -- Part 1: Sampling schemes indexed by acceptance quality limit (AQL) for lot-by-lot inspection.*
- [32] ISO 3951-1:2005. *Sampling procedures for inspection by variables -- Part 1: Specification for single sampling plans indexed by acceptance quality limit (AQL) for lot-by-lot inspection for a single quality characteristic and a single AQL.*
- [33] Webster, R.; Oliver, M.A. (1990). *Statistical methods in soil and land resource survey.* Oxford University Press. Oxford.

Calidad de los procesos:

- [34] Ariza, F.J. (2002). *Control de Calidad en la Producción Cartográfica.* Ra-Ma, Madrid.
- [35] ISO 9000:2000. *Sistemas de Gestión de la calidad. Fundamentos y vocabulario.*
- [36] ISO 9001:2000. *Sistemas de Gestión de la calidad. Requisitos.*
- [37] ISO 9004:2000. *Sistemas de Gestión de la calidad. Directrices para la mejora del desempeño.*

Tema 13. Metadatos. Definición. Normas de metadatos y sus elementos: ISO19115, Núcleo Español de Metadatos y Dublín Core.

13.1. Metadatos. Definición

El término *metadato* ha venido utilizándose ampliamente en los últimos 10 años. Así, por ejemplo, las fichas existentes en los catálogos de las bibliotecas presentan infinidad de metadatos que durante años han servido para su gestión así como para facilitar las búsquedas de los libros que se encuentran catalogados en ellas. Pero ha sido con el nacimiento de las Infraestructuras de Datos Espaciales (IDE) cuando los metadatos empiezan a tomar protagonismo pasando de ser un actor de segunda a ocupar el papel protagonista en los proyectos relacionados con el tratamiento de Información Geográfica.

Una IDE es un sistema distribuido de información geográfica a través de Internet que nos va a permitir visualizar y consultar información de tipo geográfico (datos de referencia y temáticos) de diferentes instituciones u organismos y que puede estar ubicada en diferentes servidores localizados en distintos lugares. Uno de los pilares fundamentales en el que una IDE se sustenta es el Servicio de Catálogo. Este servicio permite a los usuarios la búsqueda, localización, y selección de los datos geográficos almacenados en los diferentes servidores, gracias a que gestiona los metadatos de cada uno de los datos objeto de las búsquedas. Para que los catálogos que proceden de diferentes instituciones puedan ser interoperables y admitir búsquedas distribuidas, principios básicos de una IDE, es necesario crear los metadatos de acuerdo a normas y criterios comunes.

Los metadatos se definen comúnmente como "datos acerca de los datos". Describen el contenido, la calidad, el formato y otras características que llevan asociadas un recurso de información geográfica, constituyendo un mecanismo para caracterizar datos y servicios de forma que usuarios (y aplicaciones) puedan localizar y acceder a ellos. Van a dar respuestas a preguntas del tipo:

- El qué: el título y descripción del recurso.
- El cuándo: la fecha de creación de los datos, periodos de actualización, etc.
- El quién: el creador de los datos.
- El dónde: la extensión geográfica.
- El cómo: el modo de obtención de la información, formato, etc.

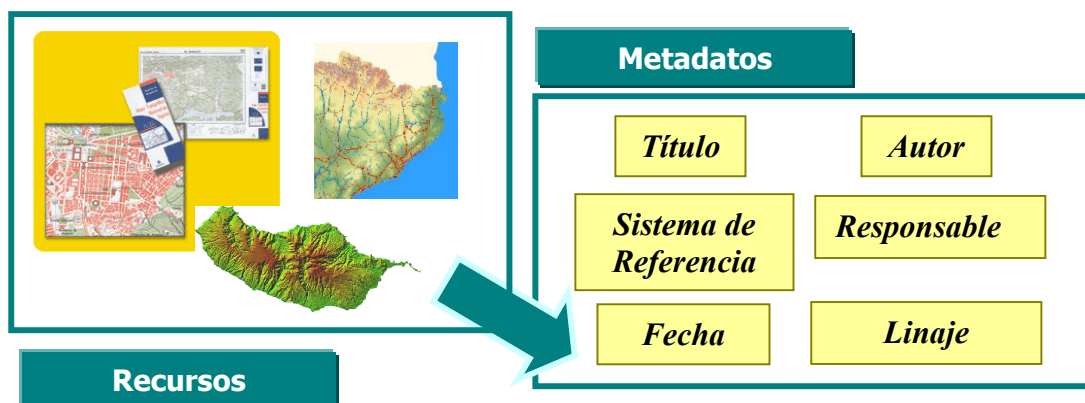


Figura 1. Recursos de Información Geográfica y ejemplos de metadatos

El concepto de metadato se está convirtiendo en una herramienta familiar para aquellas personas que trabajan con información espacial. Así la leyenda de un mapa es un ejemplo de aplicación de los metadatos que nos proporciona información sobre el autor, la fecha de publicación, la escala y otras características propias de un mapa.

Ahora bien, desde la primera concepción del término metadato como “ datos acerca de los datos” se ha producido una significativa evolución hasta dar lugar a la idea de metadatos como datos que describen no sólo datos geográficos sino también servicios (“datos acerca de datos y servicios”), como pueden ser los servicios de publicación de mapas en Internet (*Web Map Service*), servicios de transformación de coordenadas (*Web Coordinate Transformation Service*), servicios de fenómenos (*Web Feature Service*), etc. Todos ellos son servicios accesibles a través de Internet, que necesitan ser descritos mediante el uso de metadatos.

Con la creación de metadatos se persiguen diferentes objetivos que se clasifican de acuerdo a tres niveles:

- **Nivel de Descubrimiento:** mínimo de información que se necesita para transmitir la naturaleza y el contenido de los datos geoespaciales. Estos van a responder a las preguntas del tipo: "el qué, por qué, cuándo, quién, dónde y cómo". Su uso fundamental es permitir realizar búsquedas en los catálogos, para ello hay que saber que datos existen y cuales son sus principales características.
- **Nivel de Exploración:** dan la suficiente información para permitir al usuario, entre otras cosas: estar seguro de que existen datos apropiados para un propósito determinado, valorar sus propiedades y poder hacer referencia a algún punto de contacto para obtener más

información sobre los datos. Los metadatos de exploración incluyen aquellas propiedades necesarias para permitir al usuario final saber si los datos van a satisfacer los requisitos generales necesarios para poderlos utilizar para un determinado uso.

- **Nivel de explotación:** ayuda a los usuarios finales y a las organizaciones proveedoras a almacenar, volver a utilizar, mantener y archivar con efectividad sus colecciones de datos. Incluyen aquellas propiedades necesarias para el acceso, transferencia, carga, interpretación y uso de datos en la aplicación final, en donde van a ser explotados. Esta clase de metadatos con frecuencia incluye detalles de su diccionario de datos, la organización o esquema de los mismos, proyección, características geométricas y otros parámetros útiles para el uso apropiado de los datos.

Los metadatos buscan fomentar la reusabilidad de la información sin tener que recurrir al equipo humano que se encargó de su creación inicial. El valor de los datos es dependiente de la documentación que se posea de ellos, por ello, es conveniente que las organizaciones productoras de datos geográficos se encarguen de organizar y mantener los datos ayudándose para ello de metadatos.

Por otro lado las organizaciones, a través de los sistemas de catálogo, pueden publicar recursos de información geográfica, y así las organizaciones pueden encontrar los datos a utilizar y compartir con otras organizaciones y, por otro lado, los usuarios pueden conocer que información se produce en las organizaciones y cuales son sus características.

Los metadatos deberían acompañar siempre a los datos y de este modo servir tanto a usuarios como a las organizaciones productoras de la información, para el almacenamiento, procesamiento e interpretación de los propios datos.

13.2. Normas de metadatos y sus elementos: ISO19115, Núcleo Español de Metadatos y Dublín Core

Dentro del mundo de la información geográfica, que gira a unas velocidades cada vez más desconcertantes, se han ido definiendo recomendaciones para la creación de metadatos, cuya finalidad principal es proporcionar una estructura “jerárquica y concreta” que permita describir exhaustivamente cada uno de los datos a los que hacen referencia. Estas recomendaciones, creadas y aprobadas por organismos de normalización a partir de opiniones de expertos en esta materia:

- Suministran a los productores de datos criterios para caracterizar sus datos geográficos con propiedad.
- Facilitan la gestión de los metadatos y su organización.
- Permiten a los usuarios utilizar los datos de un modo más eficiente, determinando si serán de utilidad para ellos.
- Facilitan el acceso a los datos, su adquisición y una mejor utilización de los datos logrando una interoperabilidad de la información cuando esta procede de fuentes diversas.

La “Organización de Estandarización Internacional (ISO)” a través de su familia ISO 19100 define, entre otras temáticas, normas relacionadas con metadatos. Dicha familia se encuentra dividida en Comités. En el ámbito de los metadatos vamos a encontrar:

- El Comité Técnico 211, denominado “Geomática/Información Geográfica”, que ha definido las siguientes normas:
 - *ISO 19115:2003- Geographic Information Metadata*
 - *ISO/TS 19139-Geographic Information- Metadata -- XML schema implementation*
 - *ISO/CD 19115-2 Geographic information- Metadata-Part 2:Extensions for imagery and gridded data*
- El Comité Técnico 46, denominado “Información y documentación”, que ha definido la siguiente norma:
 - *ISO 15836:2003 “Information and Documentation- The Dublin Core Metadata Element Set”.*

13.2.1. ISO 19115:2003- Geographic Information Metadata

Los datos geográficos digitales pretenden modelar y describir el mundo real para su posterior análisis y visualización mediante medios muy diversos. Sus características principales así como sus limitaciones deben estar completamente documentadas mediante los metadatos. Con el fin de definir una estructura que sirva para describir los datos geográficos se creó la norma Internacional *ISO 19115:2003- Geographic Information Metadata*.

Para la elaboración de esta norma fue necesaria la colaboración de 33 países miembros de ISO/TC211 y un total de 16 países que aportaron expertos al Grupo de Trabajo (WG) encargados de su definición. En 1996 se disponía ya de un primer borrador, en el año 2003 se aprobó el texto definitivo como Norma Internacional de metadatos que fue adoptada como Norma Europea por CEN/TC 287 en 2005. AENOR (Asociación Española de Normalización) ha decidido también su adopción como Norma Española. (UNE-EN ISO19115).

Esta norma internacional proporciona un modelo y establece un conjunto común de terminología, definiciones y procedimientos de aplicación para los metadatos. Mediante la definición de elementos de metadatos se va a poder describir información sobre la identificación, la extensión, la calidad, el modelo espacial y temporal, la referencia espacial y la distribución de los datos geográficos. Se aplica a:

- La catalogación de conjuntos de datos, actividades de *clearinghouse*, y la descripción completa de conjuntos de datos;
- Diferentes niveles de información: conjuntos de datos geográficos, series de conjunto de datos, fenómenos geográficos individuales, propiedades de los fenómenos, etc.

Por lo tanto es aplicable a los distintos niveles de agregación de la información geográfica así, se pueden crear metadatos para:

- Nivel de organización: hay una serie de metadatos que describen el organismo, público o privado, productor de los datos.
- Nivel de serie o colección: existen gran variedad de productos geográficos que se agrupan en series que van a compartir ciertos metadatos.
- Nivel de unidad: es posible asociar metadatos a una hoja, una unidad, una imagen dentro de una serie.
- Nivel de clase de objeto: cada una de las agrupaciones dentro de un catálogo de objetos, por ejemplo las ciudades, los embalses, las carreteras en una base de datos.
- Nivel de instancia de objeto: se corresponden con las clases particulares dentro de una clase, como el río Ebro dentro de la clase Hidrografía.

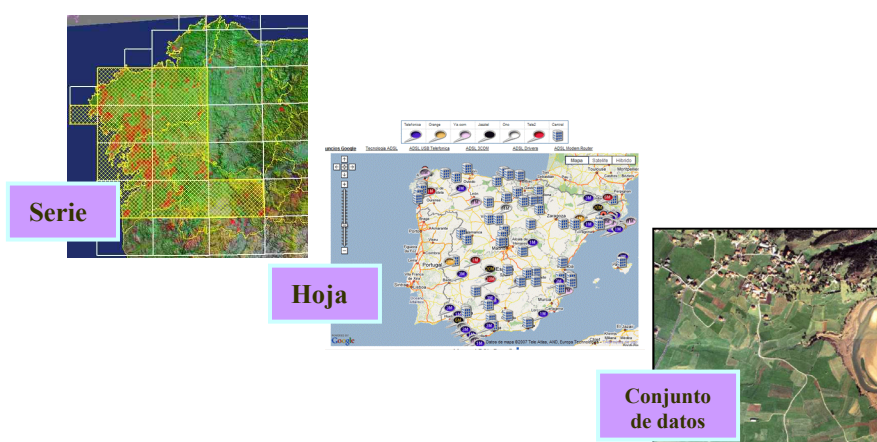


Figura 2. Ejemplos de diferentes niveles de información sobre los que se puede aplicar la norma ISO19115

Esta Norma de Metadatos es de una gran complejidad e incluye una larga serie de elementos de metadatos, unos obligatorios y otros opcionales. El documento consta de 140 páginas, incluye un total de 409 ítems y define 27 listas controladas, mediante las que se definen los posibles valores válidos de ciertos campos.

Aunque el número de elementos de metadatos descritos en esta norma es muy extenso, se define dentro de ella un “conjunto mínimo” (el núcleo o *Core*), que son obligatorios para todo el rango de aplicaciones de los metadatos (descubrimiento de datos, acceso a los datos, transferencia de datos y utilización de datos digitales, etc). Este núcleo está formado por elementos obligatorios y otros opcionales que usados todos ellos aumentan la interoperabilidad de los datos y permite a los usuarios entenderlos sin ambigüedades.

Todo perfil que se defina a partir de esta norma debe estar formado como mínimo por los elementos definidos en dicho núcleo.

Los elementos que se requieren como mínimo (*Core*) para que quede identificado un recurso asociado a información geográfico son los siguientes:

Título del Conjunto de Datos	Tipo de representación espacial
Fecha de Referencia	Sistema de Referencia
Parte responsable del Conjunto de Datos	Linaje
Localización geográfica de los Datos	Recurso en línea
Idioma del Conjunto de Datos	Identificador del Fichero de Metadatos
Conjunto de caracteres del Conj. de Datos	Norma de Metadatos
Categoría del tema	Versión de la Norma de Metadatos
Resolución espacial del conjunto de datos	Idioma de los Metadatos
Resumen descriptivo	Conjunto de caracteres de los Metadatos
Formato de Distribución	Punto de contacto para los Metadatos
Extensión vertical y temporal	Fecha de los Metadatos

Elementos obligatorios
 Elementos optativos
 Elementos condicionales

Tabla 1. Elementos que componen el Core de ISO 19115:2003

El Comité Técnico 211 de ISO revisó esta norma y debido a errores encontrados se redactó en Abril de 2006 el documento “**Technical Corrigendum 1**”. Este documento corrige erratas encontradas y realiza alguna modificación del documento ISO 19115:2003 como: variaciones de los atributos de alguna clase (inserción del atributo *locale* en la clase *MD_Metadata* debido a su presencia en ISO/TS 19139), cambios en algunos elementos en el Diccionario de datos y eliminación de otros, cambios en el apartado *CodeList* y Enumeraciones, etc.

La norma ISO 19115:2003 proporciona una estructura para describir información geográfica mediante elementos de metadatos y establece una terminología común para los mismos pero no

desarrolla como poder llevar a cabo su implementación, para ello se ha aprobado “ISO 19139”, especificación técnica que desarrolla una implementación en XML (*eXtensible Markup Language*) del modelo de metadatos descrito por ISO 19115:2003 y, que para ello, define un conjunto de esquemas en XML que van a describir los metadatos asociados a cada nivel de información, permitiendo así su descripción, asegurando su validación y su posterior intercambio en los catálogos a través de archivo de metadatos.

Aunque existen deficiencias e imprecisiones, ISO19115: 2003 constituye la referencia inexcusable para todo el que quiera trabajar en el campo de los metadatos referentes a datos y servicios geográficos. Se trata de la normativa a seguir, tanto por ser la Norma Internacional, como por estar ya implementándose a todos los niveles (serie, unidad, clase de objeto, instancia, etc) y en todos los ámbitos de la información geográfica (temáticos, de referencia, etc.

13.2.2. Núcleo Español de Metadatos

NEM, acrónimo de Núcleo Español de Metadatos, nace gracias a las opiniones, comentarios y aportaciones de un grupo abierto de expertos en metadatos pertenecientes a diferentes organizaciones e instituciones en el ámbito nacional, autonómico y local.

Es una **recomendación** de metadatos aprobada por el Consejo Superior Geográfico, a través de su Comisión de Geomática. NEM se define como un conjunto mínimo de metadatos entendidos como un perfil de ISO 19115:2003 de acuerdo con el concepto de perfil definido en la Norma ISO 19106 “*Geographic Information-Profiles*”, es decir, es un modo particular y concreto de aplicar y utilizar una Norma, seleccionando un conjunto de ítems y un conjunto de parámetros opcionales. Para ello este perfil va a tener en cuenta otras iniciativas y acciones relevantes que en la actualidad se están desarrollando en materia de metadatos.

Este perfil constituye por lo tanto un núcleo “Core”, **conjunto de metadatos “mínimo”** aconsejable por su utilidad y relevancia que va a permitir realizar (búsquedas, comparaciones,..) a partir de metadatos que proceden de diferentes fuentes, sobre distintos conjuntos de datos, de una manera rápida, práctica, fácil y fiable. Se define, para ser utilizado por todos los catálogos generados en las diferentes organizaciones relacionadas con la información geográfica de manera que se consiga la interoperabilidad de metadatos en toda España.

No es, por lo tanto, un **perfil normativo** o **restrictivo**, no se pretende que se implemente directamente sino que se aconseja su utilización, cada institución u organismo debe estudiar

cuales son los metadatos que considera adecuados para satisfacer sus necesidades, y una vez establecidos se recomienda incluir al menos los ítems que establece el perfil NEM, garantizando así la compatibilidad con el resto de iniciativas.

7 elementos obligatorios del Núcleo de ISO.	15 elementos opcionales y condicionales del Núcleo de ISO.
<ul style="list-style-type: none"> • Título • Fecha de referencia de los datos • Idioma de los datos • Categoría de tema • Resumen • Punto de contacto de los metadatos • Fecha de creación de los metadatos 	<ul style="list-style-type: none"> • Parte responsable de los datos • Formato de distribución • Tipo de representación espacial • Resolución espacial • Sistema de referencia • Recurso en línea • Información de extensión • Calidad: Linaje • Nombre del estándar de metadato • Versión del estándar de metadatos • Identificador del archivo de metadatos • Conjunto de caracteres de los datos • Idioma de los metadatos • Conjunto de caracteres de los metadatos • Localización geográfica
2 elementos pertenecientes a Dublin Core	
<ul style="list-style-type: none"> • Información de Agregación. • Créditos • Restricciones del recurso 	
2 elementos pertenecientes a la Norma ISO 19115 y propuestos por su utilización en la Directiva Europea Marco del Agua (WFD).	
<ul style="list-style-type: none"> • Propósito • Uso específico 	
Otros elementos adicionales pertenecientes a la Norma ISO 19115 y que se ocupan de profundizar en el tema de la calidad.	3 elementos adicionales, pertenecientes a la Norma ISO 19115, propuestos por expertos en metadatos y aprobados por el Subgrupo de Trabajo del NEM.
<ul style="list-style-type: none"> • Calidad: Información cuantitativa 	<ul style="list-style-type: none"> • Palabras claves descriptivas • Nivel jerárquico • Forma de representación

Tabla 2. Elementos que forman el Núcleo Español de Metadatos

La implementación de metadatos es una tarea difícil y complicada que requiere cierta especialización y considerable dedicación, pues, además de conocer bien las características técnicas y básicas del recurso a catalogar, hay que saber qué información hay que recoger en cada elemento de metadatos, cómo y con qué criterios.

Para facilitar esta tarea se ha elaborado el documento, “La Guía de Usuario de NEM” que describe para cada uno de los elementos que forman NEM, los criterios a seguir para rellenarlo. Por otro lado existen diversas herramientas que permiten crear metadatos conforme a este perfil, entre ellas tenemos la herramienta “CatMDEdit”, que está desarrollada como proyecto *Open Source* (código abierto), multilingüe, multiplataforma y compatible con otros estándares o normas de metadatos.

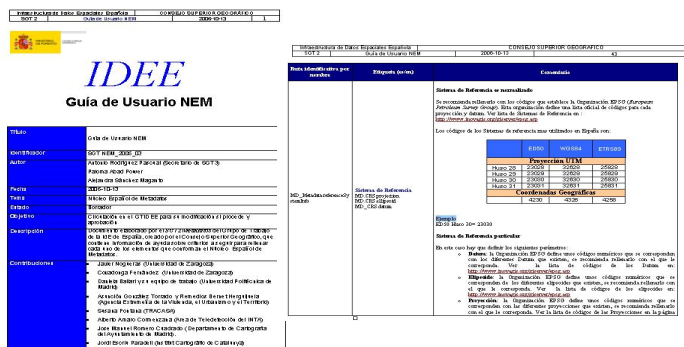


Figura 3. Guía de Usuario para crear metadatos conforme a NEM

13.2.3. Dublin Core

Un buen ejemplo de norma de metadatos de propósito general es la propuesta de la iniciativa "Dublin Core Metadata Initiative" (DCMI). Esta iniciativa, creada en 1995, promueve la difusión de estándares/normas de metadatos interoperables y el desarrollo de vocabularios de metadatos especializados que permitan la construcción de sistemas de búsqueda de información más inteligentes.

Dublin Core es una norma para la descripción de recursos de información en dominios cruzados, es decir, descripción de todo tipo de recursos independientemente de su formato, área de especialización u origen cultural.

Dublin Core ha atraído un apoyo internacional y multidisciplinario ya que muchas comunidades deseaban adoptar un núcleo común semántico para la descripción de recursos. La norma Dublin Core se ha traducido a más de 20 idiomas, y tiene un carácter oficial ya que se ha aprobado como norma americana (ANSI/NISO Z39.85), se ha adoptado dentro del comité técnico europeo CEN/ISSS (*European Committee for Standardization / Information Society Standardization System*), y desde Abril de 2003 también tiene carácter de norma ISO internacional (ISO 15836).

Esta norma consiste en quince descriptores básicos que son el resultado de un consenso internacional e interdisciplinario. Podemos clasificar estos elementos en tres grupos que indican la clase o el ámbito de la información que se guarda en ellos:

1. Elementos relacionados principalmente con el contenido del recurso
2. Elementos relacionados principalmente con el recurso cuando es visto como una propiedad intelectual
3. Elementos relacionados principalmente con la caracterización del recurso

Contenido	Propiedad Intelectual	Caracterización
Título	Creador	Fecha
Tema	Publicador	Tipo
Descripción	Contribuidor	Formato
Fuente	Derechos	Identificador
Idioma		
Relación		
Cobertura		

Tabla 3. Elementos de Dublin Core Iniciativa

La simplicidad de Dublin Core, sólo formado con quince elementos básicos, permite un fácil emparejamiento con otros esquemas de metadatos más específicos. Por tanto, muchas organizaciones en el dominio de la Información Geográfica consideran la adopción de Dublin Core en determinadas situaciones:

- Para servir de formato de intercambio entre sistemas que utilizan distintos estándares de metadatos (por ejemplo, exponer metadatos geográficos ISO19115 en otras comunidades interesadas en la utilización puntual de información geográfica);
- Para hacer recopilación (*harvesting*) de metadatos provenientes de distintas fuentes;
- Para facilitar la rápida creación de contenidos de catálogo. De hecho, la especificación de Servicios de Catálogo, propuesta por el *Open Geospatial Consortium*, propone utilizar Dublin Core como modelo básico de búsqueda y presentación de metadatos para la descripción de recursos geográficos.

Dublin Core ofrece una gran flexibilidad para la creación de perfiles de aplicación para dominios específicos. El concepto de perfil de aplicación nació dentro de Dublin Core como medio de declarar qué elementos se deberían usar en una aplicación, proyecto o dominio particular. Los perfiles de aplicación se adaptan al dominio específico mediante la posibilidad de definir cualificadores (refinamientos y esquemas de codificación) o la inclusión, si fuese necesario, de nuevos elementos definidos en el dominio específico

Existe, en la actualidad, un uso combinado de los elementos de Dublin Core con la tecnología RDF (*Resource Description Framework*), recomendación W3C para el modelado e intercambio de metadatos que se expresa en formato XML, que está adquiriendo gran importancia porque es una de las tecnologías básicas en la nueva concepción de la Web: “la Web Semántica”,

extensión de la Web actual dentro de la cual la información recibe un significado bien definido, permitiendo que computadores y personas puedan trabajar en cooperación.

Bibliografía

- [1] Documento Núcleo Español de Metadatos
(<http://www.ideo.es/resources/recomendacionesCSG/NEM.pdf>)
- [2] Documento Guía de Usuario de NEM
(<http://www.ideo.es/resources/recomendacionesCSG/GuiaUsuarioNEM.pdf>)
- [3] Rodríguez A, Abad P, Romero E, Sánchez Maganto A, 2004^a. La norma ISO 19115 de Metadatos: Características y aplicabilidad, VIII Congreso Nacional de Topografía y Cartografía TOPCART 2004.
(<http://www.cartesia.org/geodoc/topcart2004/conferencias/58.pdf>)
- [4] Sánchez Maganto A, Rodríguez A, Abad P, Romero .El Núcleo Español de Metadatos, perfil mínimo de metadatos recomendados para España. Jornadas Técnicas de las Infraestructuras de Datos Espaciales de España 2005.
(http://www.ideo.upm.es/jidee05/descargas/sesion_04_02.pdf)
- [5] Sánchez Maganto A, Rodríguez A, Abad P, Romero E,2005.El Núcleo Español de Metadatos, perfil mínimo de metadatos recomendados para España. Jornadas Técnicas de las Infraestructuras de Datos Espaciales de España 2005.
- [6] DCMI (2007). Homepage of the Dublin Core Metadata Initiative. Dublin Core Metadata Initiative (DCMI), <http://www.dublincore.org>.
- [7] ANSI (2001). The Dublin Core Metadata Element Set. ANSI/NISO Z39.85-2001, American National Standards Institute.
- [8] CEN (2003). Dublin Core Spatial Application Profile. CWA 14858, CEN/ISSS Workshop - Metadata for Multimedia Information - Dublin Core.

Tema 14. Análisis vectorial. Análisis de redes. Teoría de Grafos. Camino mínimo. Algoritmos. Problema del viajante. Análisis de polígonos.

14.1. Análisis vectorial

El modelo vectorial de datos asume un espacio geográfico continuo que cumple los postulados de la geometría euclidiana. Los objetos espaciales que se recogen con este modelo se representan codificando de forma explícita sus bordes, límites o fronteras. Un objeto puntual se representa por un par de coordenadas (X,Y) en el espacio de representación que dan la posición del objeto. Por ejemplo una fuente, un poste de telégrafos o un vértice geodésico.

Un objeto lineal se ajusta mediante segmentos lineales que tienen un vértice común donde se articulan y dan continuidad al objeto y se representan por el conjunto de parejas de coordenadas de tales puntos o vértices. Los objetos superficiales son representados por polígonos que no a su vez se representan por una serie de contornos, formados por un conjunto cerrado de segmentos.

Una representación vectorial del mundo usa un conjunto de líneas definidas por segmentos con un punto inicial y otro final conectados de alguna forma por un método de interpolación que suele ser lineal. Estos puntos inicial y final son como el origen y destino de un vector, y el conjunto de los mismos representa la forma de la entidad-objeto junto con el sentido de concatenación o “segmento siguiente” para indicar al ordenador como se conectan los segmentos.

Para el análisis de la información geográfica sobre datos vectoriales, éstos se han de recoger en estructuras complejas que permiten realizar el mismo de forma rápida y eficaz. Las bases de datos de información geográfica vectorial pueden incluso estar desarrolladas para tecnología Orientada a Objetos (OO) y permiten manejar realizar múltiples operaciones de análisis sobre los mismos gracias a que incorporan las relaciones implícitas entre los datos.

La estructura de datos vectorial se basa en la definición de puntos, líneas, y áreas.

Las entidades puntuales comprenden cualquier objeto cartográfico o geográfico que pueda ser georreferenciado mediante un par de coordenadas (X,Y) o en su caso de (X,Y,Z) caso de ser manejada la altitud del punto. Según la función cartográfica se pueden distinguir tres tipos de entidades puntuales:

- Objeto geográfico puntual, que identifica y representa una entidad geográfica que esta asociada con una posición espacial puntual (por ejemplo un vértice geodésico).
- Punto geométrico como objeto cartográfico complementario que especifica una posición geométrica relacionada con la representación del mapa (por ejemplo las coordenadas geográficas asignadas a puntos de control dentro del mapa, a las esquinas de hoja o a un centroide).
- Punto de situación de etiquetas y textos que permiten asociar un texto o una etiqueta a la representación de una entidad cartográfica o geográfica (por ejemplo la toponimia, rótulos, etc.)

Las entidades lineales están integradas por objetos lineales elementales o segmentos, de manera que pueden ser:

- Polilíneas de segmentos de recta, como sucesión de pares/tripletas de coordenadas de los vértices de la poligonal.
- Polilíneas de segmento de curva, como concatenación de segmentos de curvas definidas mediante los parámetros que permiten establecer la función matemática de la curva (por ejemplo: spline line, arco de circunferencia, etc.).

El segmento viene definido y delimitado por sus extremos, llamados a veces nodos, vértices, definidos como los puntos de discontinuidad en un objeto cartográfico lineal en el que se establece una relación de conectividad con otros objetos lineales.

El segmento puede ir acompañado de dirección, definida por sus nodos inicial y final, y un sentido dado mediante un signo + (mas) o un signo – (menos), un dígito 0 y 1, o bien mediante un valor lógico t (true) y f (false), que indicaría la dirección en la que debe considerarse dentro de un objeto lineal de mayor complejidad.

Las entidades superficiales resultan de la modelización bidimensional de las entidades geográficas y caracterizadas por contener una extensión o área. Quedan definidas por los segmentos lineales que forman los contornos de los polígonos que la forman.

Cuando se habla de polígono, se entiende que es un objeto superficial continuo, es decir que puede tener agujeros o descuentos, pero no anejos o adiciones; y cuando se habla de superficie se entiende que no es necesario que sea continua y admite descuentos y adiciones.

La identificación o etiquetado de los polígonos se puede hacer mediante la edición de un punto interior al área del objeto llamado centroide y que tiene coordenadas (X,Y) o (X,Y,Z). Este

punto es el que se utiliza para asignar los atributos, etiquetas, textos y demás características que definan al objeto “polígono” en la correspondiente base de datos.

Para la realización del análisis espacial es especialmente útil considerar la topología, que puede almacenarse de forma permanente o calcularse en el momento en el que se requiere realizar el análisis (topología *on the fly* o al vuelo).

La Topología es una rama de las matemáticas y en ella se estudian las interrelaciones espaciales cualitativas, que no varían cuando se deforma el espacio de forma continua, entre los elementos de un conjunto de objetos.

Para un SIG las primitivas topológicas son nodo, arco o borde y cara, y las interrelaciones las características comunes o de relación y orden que tiene entre sí estos objetos con otros de sus mismos modelos (intrínsecas) o con los demás (extrínsecas). Las relaciones entre los objetos son definidas y modelizadas en cada SIG y atendiendo a cada tipo de datos o sus objetos relacionados. Son típicas las relaciones de ordenación, concatenación, proximidad, inclusión, sentido de recorrido, vecindad, etc.

Dichas relaciones han de establecerse adecuadamente para lo cual es imprescindible una correcta y completa codificación de los objetos, así como una consistencia concreta en la plasmación de la geometría.

Relaciones topológicas. Son aquellas propiedades de los objetos relacionados que permanecen invariables a pesar de que el mapa sufra variaciones o distorsiones (cambios de escala o referenciación, proyección, alargamiento o reducción, etc.). En este último caso variarían propiedades como distancias y ángulos, pero no variarían las relativas a adyacencias, conexiones, inclusión, etc.). En general, dichas relaciones quedan recogidas en la correspondiente Base de Datos Espaciales lo que permite el análisis espacial en el mundo vectorial.

Estas relaciones topológicas se establecen entre los objetos cartográficos individualizados y pueden ser:

- De coincidencia. Cuando la situación de dos objetos cartográficos es la misma en todo o en parte, (por ejemplo un segmento de eje de calle y un segmento de límite de distrito municipal; un segmento de contorno de manzana y un segmento de contorno de parcela).

- De inclusión. Cuando un objeto cartográfico queda completamente dentro de otro, sin ser componente de aquél. Pueden distinguirse:
 - De objeto puntual en objeto lineal. El objeto puntual debe coincidir con un nodo.
 - De objeto puntual en objeto superficial. Identificado normalmente mediante algoritmos de determinación de punto interior.
 - De segmento de línea en objeto superficial. Caracterizado porque a derecha e izquierda del segmento se encuentra el mismo objeto superficial.
 - De objeto superficial en objeto superficial. Esta relación es paralela a la que se establece en la definición de objetos superficiales complejos.

- De conectividad. Cuando entre dos o más objetos cartográficos existe una conexión directa. Pueden distinguirse:
 - De objeto lineal con objeto lineal. A través de los puntos comunes o nodos, que definen segmentos de los objetos lineales.
 - De objeto lineal con objeto superficial. A través de los nodos o puntos de corte de la línea límite del objeto superficial.
 - De objeto superficial con objeto superficial. Mediante la existencia de un segmento lineal, del borde de los objetos, común. A esta relación podría denominársele también de Contigüidad.
 - De más de dos objetos superficiales. Mediante la existencia de un nodo común a varios segmentos lineales que separan los objetos superficiales.

- De superposición sin conexión. Circunstancia que se puede producir cuando se considera la tercera dimensión espacial de las entidades geográficas, ya que pueden existir objetos cartográficos puntuales, lineales y superficiales que coincidan en situación bidimensional, pero no tengan conexión directa por estar en distinto nivel.

- De influencia. Relación típicamente cartográfica que se puede establecer entre los objetos, por la cual se puede definir quien tiene prioridad en el momento de la presentación.

- De proximidad. Generalmente no explicitada, sino calculada analíticamente cuando es necesario.

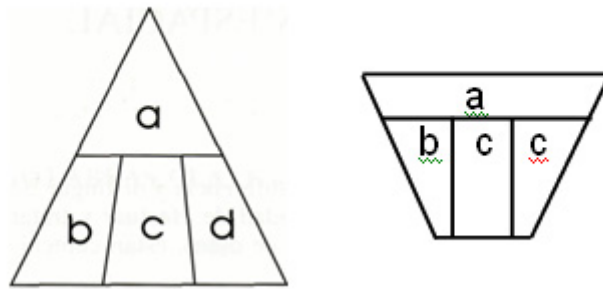


Figura 1. Los dos polígonos tienen áreas con distinta geometría pero con las mismas

14.1.1. Ejemplo de relaciones topológicas y su formulación práctica. Conectividad

El SIG reconoce qué segmentos conectan con los demás, chequeando en qué nodos empiezan y acaban cada uno de los segmentos. Así, el nudo 6 conecta los segmentos T9, T8, T6 y T10, por ser común a todos ellos. El ordenador reconocerá como posible ir del segmento T5 al segmento T6, pero no al segmento T9, pues no tienen ningún nodo común.

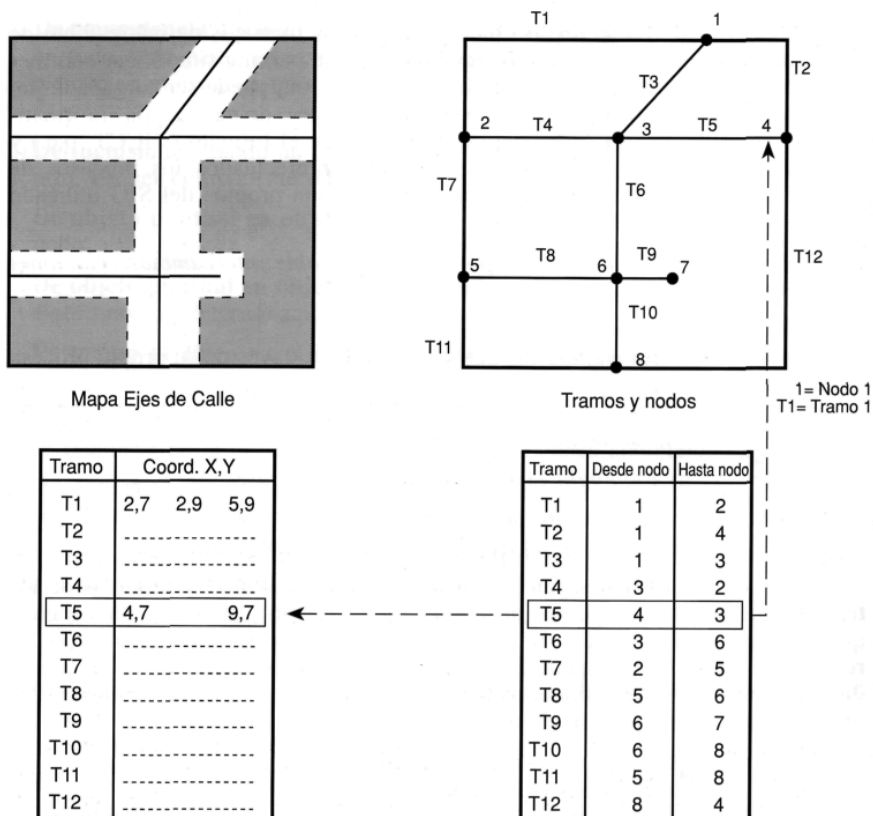


Figura 2. Ejemplos de relaciones topológicas: conectividad

14.1.2. Ejemplo de relaciones topológicas y su formulación práctica. Adyacencia

Como cada segmento tiene una dirección (“desde nodo”, “hasta nodo”), el SIG elabora una relación de polígonos a la izquierda y a la derecha de cada segmento. Cualesquiera polígonos que presentan un segmento común, el SIG los reconoce como adyacentes (p.e. polígonos 3 y 5). En los modelos de representación ráster, la topología va implícita en la propia estructura de datos, siendo mucho más simples los procesos de análisis topológicos celda a celda, píxel a píxel, y siendo muy poco apropiados para el análisis topológico.

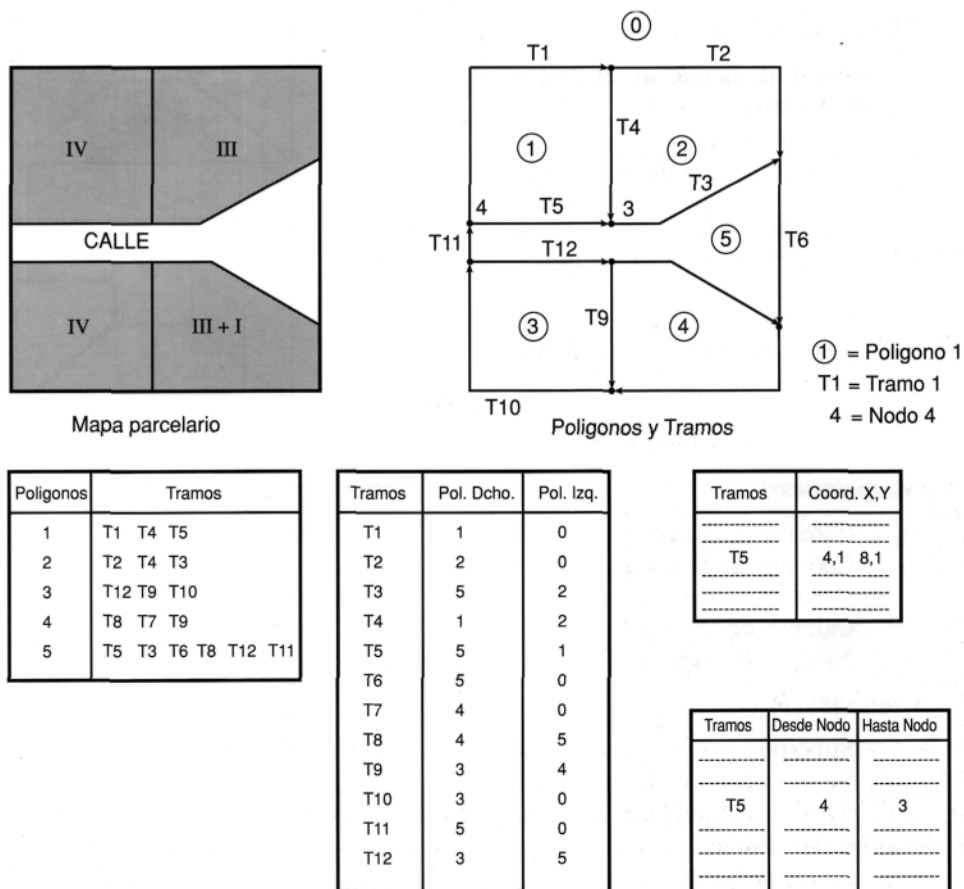


Figura 3. Ejemplos de relaciones topológicas: adyacencia

14.1.3. Funciones de Análisis Espacial

Ningún programa informático de SIG concreto contempla la totalidad de funciones de análisis espacial pues cada uno de los productos existentes enfatiza más unas u otras utilidades. De cualquier forma podemos enumerar las que son funciones habitualmente presentes en los sistemas SIG:

- Manipulación de datos
 - Reclasificación por atributos
 - Generalización
 - disolución y mezcla
 - suavizado de líneas
 - sombreado de superficies
 - Interpolación
 - ubicación de centroides
 - contornos
 - Cambios de escala
 - Transformación de coordenadas
 - Cambios de proyección
 - Cálculo de códigos geográficos y cartográficos
 - Operaciones lógicas de unión e intersección

- Mediciones
 - Cálculo de áreas
 - Cálculo de distancias
 - Dirección de ángulos
 - Cálculo de volúmenes
 - Número/frecuencia de observaciones

- Comparación
 - Superposición e intersección de polígonos
 - Intersección de polígono y líneas
 - Proximidad, vecindad, contigüidad, conectividad
 - Ruta óptima
 - Determinación localización óptima

- Interrogación
 - Búsqueda e interrogación
 - por atributo alfanumérico
 - de los atributos alfanuméricos
 - Búsqueda condicional

- Análisis de redes
- Representación tridimensional
- Estadísticas

14.1.4. Análisis espacial: Superposición de capas

Las distintas entidades geográficas que conforman un mapa son almacenadas en el ordenador, organizadas de una forma lógica mediante un conjunto de niveles o capas de información, contemplando cada una de ellas una información temática, tal como se ha visto en anterior apartado. Cada nivel o capa consiste, pues, en un conjunto de entidades geográficas relacionadas topológicamente y sus atributos descriptivos asociados.

La superposición de capas, aplicación muy usual en entornos SIG, permite visualizar y tratar la información agrupando varias capas simultáneamente como si fuera una nueva capa o nivel de información. Se utiliza esta posibilidad en casos como planeamiento territorial, interpolación de áreas, representación de redes y áreas de influencia. Es una de las operaciones más complejas de los SIG vectoriales. Matemáticamente se basa en el algebra de Boole y la superposición de conjuntos bajo distintas condiciones: AND, OR, XOR, y NOT.

La distribución de la información total en varias capas o niveles tiene importancia en los procesos de actualización y mantenimiento-modificación de la información.

Así, en los procesos de actualización, ésta puede afectar a varias CAPAS (por ejemplo, modificación de los límites de una parcela, segregación, nueva calle, etc.), por lo que dicho proceso de actualización y/o modificación de datos debe realizarse sobre una capa que sea la suma de todas ellas, para después restituir los datos modificados en todas las capas o niveles de información en que se encuentra dividida la información.

En el caso de construcción de una nueva cobertura o tema, por superposición de dos o más capas, se generan nuevos polígonos, intersecciones y otras circunstancias, que implican la generación y creación de nuevas relaciones entre los elementos, que deben ser calculadas y almacenadas, así como se producen nuevos ficheros de atributos, correspondientes a la nueva cobertura generada.

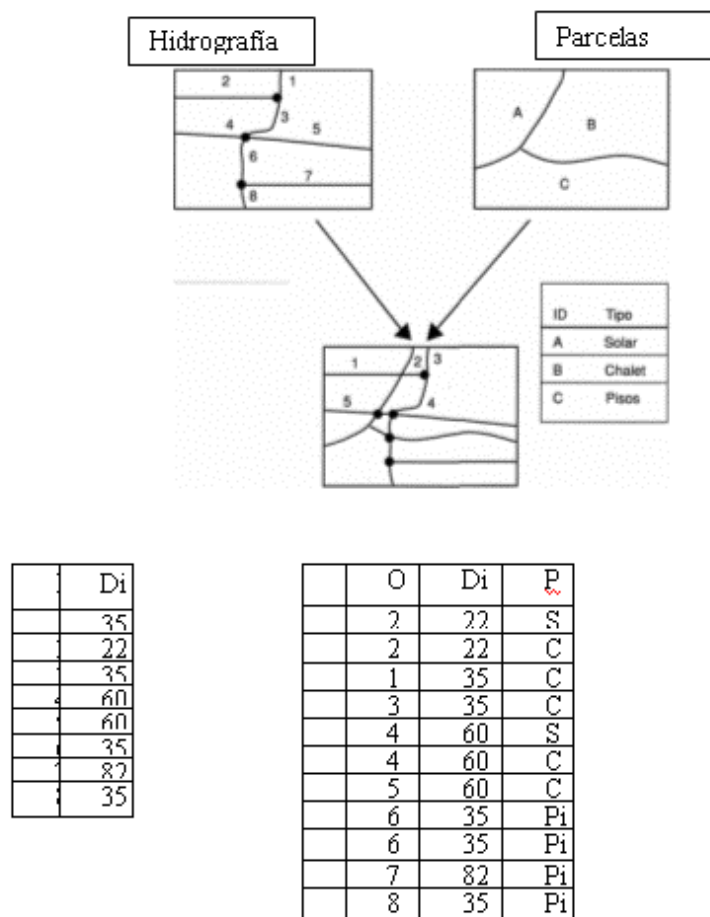


Figura 4. Superposición de Capas. Líneas sobre polígonos (Hidrografía & parcelas)

14.1.5. Análisis espacial: Áreas de influencia (buffering)

La generación de buffers o áreas de entorno, es una operación que permite identificar un área de tamaño predeterminado rodeando alguna entidad geográfica seleccionada. Se construye uno o más polígonos englobando los objetos cartográficos seleccionados.

Se utiliza fundamentalmente para determinar proximidad espacial, tomando como referencia cualquier tipo de objeto, punto, línea o polígono.

14.2. Análisis de redes

Una red es un conjunto de elementos lineales interconectados por los cuales hay un flujo de recursos (vehículos, personas, agua, energía, etc.) y que pueden tener una resistencia o costo al paso, denominada impedancia.

El análisis de redes tiene su principio en la topología de conectividad y conduce entre otros, a las siguientes posibilidades:

- **Ruta más corta:** Permite calcular cual es la ruta que menos distancia recorre entre una serie de paradas asignadas. El sistema realiza el análisis considerando **sólo la longitud** de los elementos.

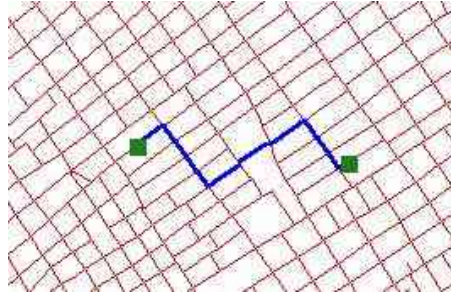


Figura 5. Análisis de redes. Ruta más corta.

- **Ruta más eficiente:** Permite calcular cual es la ruta que menos costo en tiempo u otra variable (combustible, dinero, etc) consume entre una serie de paradas asignadas. El sistema considera no sólo la longitud de los elementos sino también la **impedancia** asociada a cada uno.



Figura 6. Análisis de redes. Ruta más eficiente.

Así, por ejemplo una vía corta en longitud pero con alta impedancia (por ejemplo, baja velocidad promedio) es equivalente a una vía larga en términos de tiempo.

- **Utilidades más próximas.** El sistema permite encontrar cuales son las utilidades mas cercanas a un evento sobre la red y el recorrido más corto o más eficiente desde estas. Por ejemplo se puede calcular cuál es el recorrido más rápido que debe seguir una unidad de atención de desastres para atender una eventualidad desde la estación base.

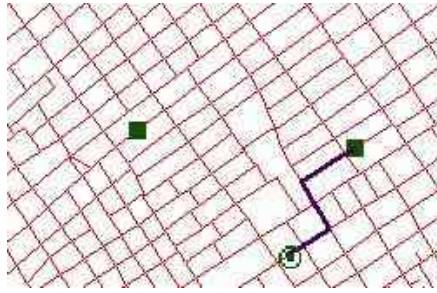


Figura 7. Análisis de redes. Utilidades más próximas.

- **Análisis de accesibilidad en distancia o tiempo.** Se puede obtener cual es el área que se puede cubrir en un determinado tiempo o distancia desde cierto punto. Es como un "Buffer" pero generado por desplazamiento a través de la red.

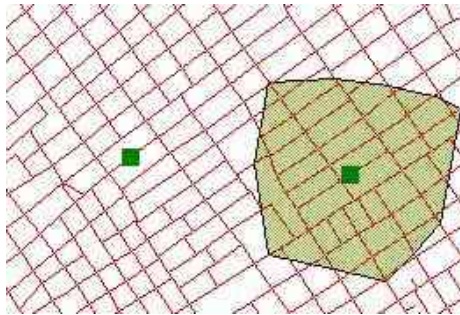


Figura 8. Análisis de redes. Análisis de accesibilidad en distancia o tiempo.

- **Análisis de redes específicos.** Hay aplicaciones SIG de redes desarrollados para análisis específicos como las redes hidrológicas, de servicios públicos, etc. Por ejemplo el sistema puede calcular cual es el caudal acumulado en un determinado punto de un cauce.



Figura 9. Análisis de redes. Análisis de accesibilidad en distancia o tiempo.

14.3. Teoría de grafos

En matemáticas y ciencias de la computación, la teoría de grafos estudia las propiedades de los grafos, que son colecciones de objetos llamados vértices (nodos, o nudos) conectados por líneas llamadas aristas (bordes, o arcos) que pueden tener orientación (dirección asignada).

Típicamente, un grafo está diseñado por una serie de puntos (los vértices) conectados por líneas (las aristas).

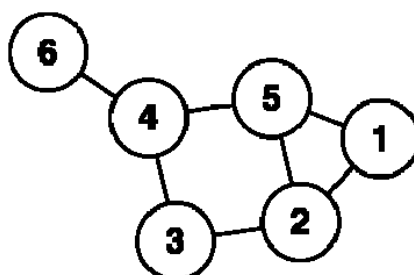


Figura 10. Grafo con 6 vértices y siete aristas

14.3.1. Introducción

El trabajo de Leonhard Euler, en 1736, sobre el problema de los puentes de Königsberg es considerado como uno de los primeros resultados de la teoría de grafos. También se considera uno de los primeros resultados topológicos en geometría (que no depende de ninguna medida). Este ejemplo ilustra la profunda relación entre la teoría de grafos y la topología.

El problema de los siete puentes de Königsberg (Prusia oriental en el siglo XVIII -ciudad natal de Kant- y actualmente, Kaliningrado, en la óblast rusa de Kaliningrado) es un célebre problema matemático que fue resuelto por Leonhard Euler en 1736 y dio origen a la Teoría de los grafos.

Consiste en lo siguiente:

Dos islas en el río Pregel que cruza Königsberg se unen entre ellas y con la tierra firme mediante siete puentes. ¿Es posible dar un paseo empezando por una cualquiera de las cuatro partes de tierra firme, cruzando cada puente una sola vez y volviendo al punto de partida?

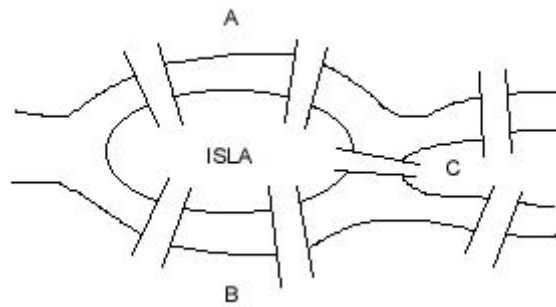


Figura 11.El problema de los siete puentes de Königsberg.

Euler enfocó el problema representando cada parte de tierra por un punto y cada puente, por una línea, uniendo los puntos que se corresponden. Entonces, el problema anterior se puede trasladar a la siguiente pregunta: ¿se puede recorrer el dibujo terminando en el punto de partida sin repetir las líneas?

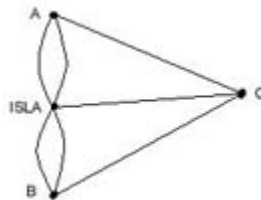


Figura 12.Enfoque de Euler al problema de los siete puentes.

Euler demostró que no era posible puesto que el número de líneas que inciden en cada punto no es par (condición necesaria para entrar y salir de cada punto, y para regresar al punto de partida, por caminos distintos en todo momento).

En 1845 Gustav Kirchhoff publicó sus leyes de los circuitos para calcular el voltaje y la corriente en los circuitos eléctricos.

En 1852 Francis Guthrie planteó el problema de los cuatro colores que plantea si es posible, utilizando solamente cuatro colores, colorear cualquier mapa de países de tal forma que dos países vecinos nunca tengan el mismo color. Este problema, que no fue resuelto hasta un siglo después por Kenneth Appel y Wolfgang Haken, puede ser considerado como el nacimiento de la topología. Al tratar de resolverlo, los matemáticos definieron términos y conceptos teóricos fundamentales.

14.3.2. Estructuras de datos en la representación de grafos

Existen diferentes formas de almacenar grafos en una computadora. La estructura de datos usada depende de las características del grafo y el algoritmo usado para manipularlo. Teóricamente se pueden distinguir las estructuras de listas y las de matrices, pero usualmente, lo mejor es una combinación de ambas. Las listas son preferidas en grafos dispersos porque tienen un eficiente uso de la memoria. Por otro lado, las matrices proveen acceso rápido, pero pueden consumir grandes cantidades de memoria.

- **Estructura de lista**
- **lista de incidencia** - Las aristas son representadas con un vector de pares (ordenados, si el grafo es dirigido) de vértices que conecta esa arista.
- **lista de adyacencia** - Cada vértice tiene una lista de vértices los cuales son adyacentes a él. Esto causa redundancia en un grafo no dirigido (ya que A existe en la lista de adyacencia de B y viceversa), pero las búsquedas son más rápidas, al costo de almacenamiento extra.
- **Estructuras matriciales**
- **Matriz de incidencia** - El grafo está representado por una matriz de A (aristas) por V (vértices), donde [arista, vértice] contiene la información de la arista (1 - conectado, 0 - no conectado)
- **Matriz de adyacencia** - El grafo está representado por una matriz cuadrada M de tamaño n, donde n es el número de vértices. Si hay una arista entre un vértice x y un vértice y, entonces el elemento $m_{x,y}$ es 1, de lo contrario, es 0.

14.3.3. Caracterización de Grafos

- **Grafos Simples.**

Un grafo es simple si a lo sumo sólo 1 arista une dos vértices cualesquiera. Esto es equivalente a decir que una arista cualquiera es la única que une dos vértices específicos. Un grafo que no es simple se denomina complejo.

- **Grafos Conexos.**

Un grafo es conexo si cada par de vértices está conectado por un camino; es decir, si para cualquier par de vértices (a, b), existe al menos un camino posible desde a hacia b.

Un grafo es fuertemente conexo si cada par de vértices está conectado por al menos dos caminos disjuntos; es decir, es conexo y no existe un vértice tal que al sacarlo el grafo resultante sea desconexo.

Es posible determinar si un grafo es conexo usando un algoritmo Búsqueda en anchura (BFS) o Búsqueda en profundidad (DFS).

En términos matemáticos la propiedad de un grafo de ser (fuertemente) conexo permite establecer en base a él una relación de equivalencia para sus vértices, la cual lleva a una partición de éstos en "componentes (fuertemente) conexas", es decir, porciones del grafo, que son (fuertemente) conexas cuando se consideran como grafos aislados. Esta propiedad es importante para muchas demostraciones en teoría de grafos.

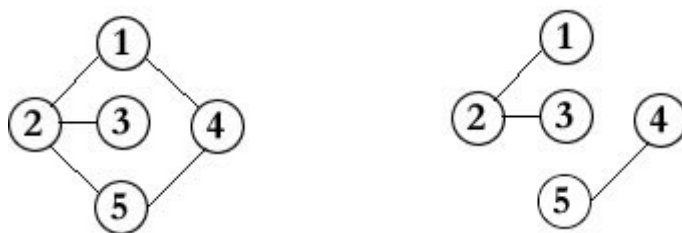


Figura 13. Ejemplos de Grafos. Grafo conexo y no conexo.

- **Grafos Completos**

Un grafo simple es completo si existen aristas uniendo todos los pares posibles de vértices. Es decir, todo par de vértices (a, b) debe tener una arista e que los une.

El conjunto de los grafos completos es denominado usualmente \mathbb{K} , siendo \mathbb{K}_n el grafo completo de n vértices.

Un \mathbb{K}_n , es decir, grafo completo de n vértices tiene exactamente $\frac{n(n-1)}{2}$ aristas.

La representación gráfica de los \mathbb{K}_n como los vértices de un polígono regular da cuenta de su peculiar estructura.

- **Grafos Bipartitos.**

Un grafo G es bipartito si puede expresarse como $G = \{V_1 \cup V_2, A\}$ (es decir, la unión de dos grupos de vértices), bajo las siguientes condiciones:

- V_1 y V_2 son disjuntos y no vacíos.
- Cada arista de A une un vértice de V_1 con uno de V_2 .
- No existen aristas uniendo dos elementos de V_1 ; análogamente para V_2 .

Bajo estas condiciones, el grafo se considera bipartito, y puede describirse informalmente como el grafo que une o relaciona dos conjuntos de elementos diferentes, como aquellos resultantes de

los ejercicios y puzzles en los que debe unirse un elemento de la columna A con un elemento de la columna B.

14.3.4. Operaciones con grafos

- **Subdivisión elemental de una arista**

$u \xrightarrow{e} v$ se convierte en $u \xrightarrow{e'} w \xrightarrow{e''} v$

Se reemplaza la arista $e = \{u, v\}$ por dos aristas $e' = \{u, w\}$ $e'' = \{w, v\}$ y un vértice w .

Después de realizar esta operación, el grafo queda con un vértice y una arista más.

- **Remoción débil de un vértice**

Si $v \in V_G$ y $g(v) = 2$ (Sea v un vértice del grafo y de grado dos) removerlo débilmente significa reemplazarlo por una arista que une los vértices adyacentes a v .

$u \xrightarrow{e'} w \xrightarrow{e''} v$ se convierte en $u \xrightarrow{e} v$

Entonces e' y e'' desaparecen y aparece $e = \{u, v\}$

14.3.5. Aplicaciones

Gracias a la teoría de Grafos se pueden resolver diversos problemas como por ejemplo la síntesis de circuitos secuenciales, contadores o sistemas de apertura.

Los grafos se utilizan también para modelar trayectos como el de una línea de autobús a través de las calles de una ciudad, en el que podemos obtener caminos óptimos para el trayecto aplicando diversos algoritmos como puede ser el algoritmo de Floyd.

Para la administración de proyectos, utilizamos técnicas como PERT en las que se modelan los mismos utilizando grafos y optimizando los tiempos para concretar los mismos.

La teoría de los grafos también ha servido de inspiración para las ciencias sociales, en especial para desarrollar un concepto no metafórico de "red social" que substituye los nodos por los actores sociales y verifica la posición, centralidad e importancia de cada actor dentro de la red. Esta medida permite cuantificar y abstraer relaciones complejas, de manera que la estructura social puede representarse gráficamente. Por ejemplo, la "red social" puede representar la estructura de poder dentro de una sociedad al identificar los vínculo (Aristas), su dirección e intensidad y da idea de la manera en que el poder se trasmite y a quienes.

14.3.6. Ejemplo: el algoritmo de Dijkstra

El algoritmo de Dijkstra, también llamado algoritmo de caminos mínimos, es un algoritmo para la determinación del camino más corto dado un vértice origen al resto de vértices en un grafo dirigido y con pesos en cada arista. Su nombre se refiere a Edsger Dijkstra, quien lo describió por primera vez en 1959.

La idea subyacente en este algoritmo consiste en ir explorando todos los caminos más cortos que parten del vértice origen y que llevan a todos los demás vértices; cuando se obtiene el camino más corto desde el vértice origen, al resto de vértices que componen el grafo, el algoritmo se detiene. El algoritmo es una especialización de la búsqueda de costo uniforme, y como tal, no funciona en grafos con aristas de costo negativo (al elegir siempre el nodo con distancia menor, pueden quedar excluidos de la búsqueda nodos que en próximas iteraciones bajarían el costo general del camino al pasar por una arista con costo negativo).

Este algoritmo está disponible en los programas de análisis de redes integrados en un SIG. En Daskin (1995) se puede encontrar una descripción de un algoritmo de este tipo.

14.4. Camino mínimo

En esta ocasión el problema reside, por ejemplo, en la posición "óptima" del trazado de carreteras, caminos, líneas férreas, canales, y en general, todo tipo de infraestructuras lineales. Como en el caso anterior aquí también se deben considerar criterios de diverso tipo: económicos (minimización de los costes de construcción) y ambientales (reducción de los impactos negativos que el trazado de la infraestructura produce). Existen diversos tipos de problemas sobre la localización óptima de elementos lineales:

- Localización de caminos/rutas óptimas sobre redes de transporte ya existentes. Ejemplo: establecer por donde debe viajar un repartidor de productos (leche, pan, etc.) de manera que el coste del viaje (o el tiempo de recorrido, etc.) sea mínimo. Clásico problema del viajante.
- Determinación del nuevo trazado de carreteras u otro tipo de infraestructuras lineales. También en este caso se pueden plantear dos modalidades: 1.º Se dan por encontrados varios trazados diferentes a una nueva ruta. Es necesario evaluar cual es el más adecuado. Cada ruta se evalúa por separado y se elige la que sea más corta o barata de construir. 2.º Se emplea un procedimiento de búsqueda ex-novo de un recorrido óptimo teniendo en cuenta los criterios necesarios que miden lo adecuado que es el paso por cada punto del territorio.

La ruta óptima encontrada debe maximizar la adecuación o, lo que es lo mismo, minimizar los inconvenientes totales del trazado. Es posible emplear diversos tipos de criterios para

determinar la localización óptima de una nueva infraestructura lineal: a) Económicos: establecen el coste de construir la ruta o las ventajas que su trazado supondrá para los usuarios o para las zonas servidas. b) Ambientales: se mide en cada punto del trazado el efecto negativo de la ruta para el medio ambiente.

14.4.1. Soluciones para el problema de determinar una ruta/trazado óptimo sobre una red existente.

Existen diversos algoritmos matemáticos que resuelven el problema. El más conocido es el de Dijkstra (ver punto 7 de este Tema 13). Por ejemplo, en el módulo Network de ARC/INFO se puede emplear la orden Route para llevar a cabo este proceso. En el caso de un SIG ráster (como IDRISI) la solución se puede esquematizar en lo siguiente:

1.º Establecer un mapa de fricciones unitarias, con valores desde 1 al máximo en las rutas ya existentes (fricción con valores en proporción inversa a la velocidad máxima en cada una de las rutas), en el exterior de las rutas valor -1 (imposibilidad de paso).

2.º Calcular el mapa de costos de recorrido acumulados hasta uno de los extremos de la ruta deseada, usando para ello el algoritmo COSTGROW (el cual no permite pasar por pixels con fricción de -1).

3.º Sobre este mapa de costes acumulados usar la orden PATHWAY con el otro extremo de la ruta como origen del camino. El resultado es una trayectoria que minimiza los costes acumulados de ir desde un extremo al otro de la ruta. El resultado difiere y es algo peor (más largo y más costoso) que el obtenido con ROUTE de ARC/INFO.

14.4.2. Soluciones para el problema de determinar una nueva ruta de trazado óptimo sin una red existente.

Para la variante del problema que consiste en evaluar trazados diferentes ya establecidos de una nueva ruta, el procedimiento sería el siguiente:

Crear estratos temáticos de cada uno de los criterios que valoran la adecuación de que la nueva ruta pase por un punto del territorio, en estos mapas cada punto del territorio se valora en cuanto a su adecuación para recibir la ruta.

Superponer cada trazado sobre los diversos criterios y extraer los valores de adecuación del paso de la ruta por cada punto.

Combinar los valores de adecuación extraídos en cada trazado para construir un único valor de medida de adecuación. Se pueden usar métodos de evaluación multicriterio para llevar a cabo esa combinación.

Elegir el trazado con el mayor valor de adecuación.

En la segunda variante del problema, la búsqueda de un recorrido óptimo para la nueva ruta, la solución se puede esquematizar así:

Se crean estratos temáticos con los criterios de adecuación.

Los criterios se convierten a valores de fricción unitaria o dificultad de paso de la ruta por cada punto.

Se combinan las fricciones unitarias consideradas en un solo mapa de fricción unitaria, en este mapa el valor existente en cada punto del territorio mide lo costoso que resulta que la nueva ruta pase por ese lugar.

Se obtiene una superficie de costes acumulados de ir hasta uno de los extremos de la ruta, el origen de ella, usando para ello la orden COSTPUSH de IDRISI.

El algoritmo de búsqueda determina, en la anterior superficie de costos, por donde pasa la ruta de manera que se minimice el costo acumulado, se emplea para esto la orden PATHWAY. El algoritmo utilizado por PATHWAY es el de la máxima pendiente sobre la superficie de costos, es decir a partir del otro extremo de la ruta, el destino, se va buscando una trayectoria que discurra bajando siempre por la máxima pendiente en la superficie de costos hasta el origen de la nueva ruta. Esa es la ruta mas adecuada. Este tipo de problema, buscar ex-novo una nueva localización para una ruta, es mas pertinente para su resolución con un SIG ráster, por el contrario, como ya se ha indicado, el problema de la trayectoria óptima en un red de carreteras preexistente es mas propicio para su estudio con un SIG vectorial.

14.5. Algoritmos

Por algoritmo se entiende una secuencia finita de operaciones (lógicas y/o aritméticas) a efectuar con las magnitudes de partida (no necesariamente numéricas) para obtener las magnitudes resultado (tampoco necesariamente numéricas), en una secuencia que se puede ejecutar de modo puramente automático por lo tanto entendemos por algoritmo un procedimiento que contiene de forma muy organizada y ordenada un conjunto de reglas que especifican una secuencia finita de operaciones que proveen de solución a un problema o a una clase específica de problemas. Sobre los algoritmos podemos mencionar las siguientes propiedades:

- **Finitud:** Un algoritmo debe estar constituido por un número finito de pasos ejecutados por un número finito de veces y se debe escribir mediante un número finito de símbolos.

- **Discreción:** Cada paso debe ser distinto de los demás.
- **Determinismo:** En correspondencia con cada paso deben estar bien definidas tanto las acciones a ejecutar como el paso sucesivo al que acceder.

Las magnitudes que se obtienen en un paso vienen determinadas por las magnitudes disponibles a partir de los pasos anteriores.

- **Masividad o generalidad:** Las magnitudes a las que se aplica el algoritmo se deben poder elegir de un conjunto tendencialmente extenso (dominio de aplicabilidad del algoritmo).

Por tanto, un algoritmo se entiende como un método de resolución de una clase de problemas análogos.

En programas informáticos SIG disponen de herramientas para resolver y dar solución a los problemas que se presentan en el manejo de información gráfica, tales como intersección de líneas, cálculo de áreas de polígonos, localización de centroides, etc.

Seguidamente analizaremos muy sucintamente el esquema de planteamiento del algoritmo de intersección de líneas. En los puntos 7 y 8 de este tema se encuentran descritos otros dos problemas algorítmicos.

14.5.1. Intersección de líneas

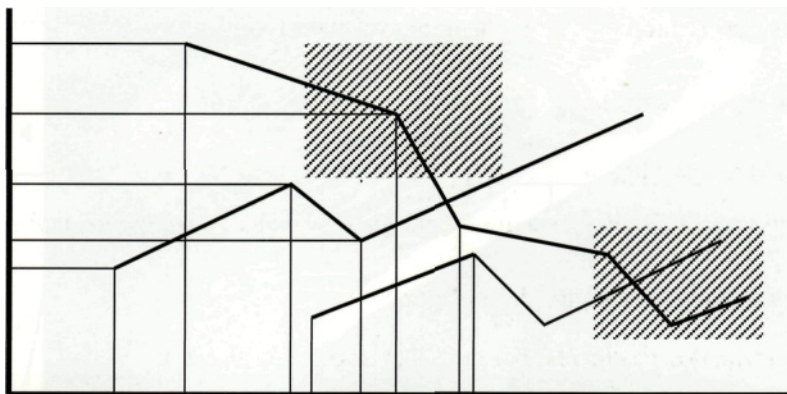


Figura 14. Intersección de líneas. El algoritmo actúa por zonas

La intersección de 2 líneas es una operación crítica en un SIG

Esta operación es usada en las tareas de superposición de polígonos, mezcla y disolución de polígonos y líneas, así como en el preproceso de generación de topología. Los procesos de intersección se apoyan en una previa ordenación de las coordenadas, para racionalizar las

búsquedas de posibles intersecciones y delimitar el ámbito de actuación. Las fases del proceso son, esquemáticamente:

- Ordenación de segmentos según eje X
- Análisis para cada segmento, de sus posibilidades de intersección con otros, dentro del ámbito considerado (selección de segmentos).
- Ordenación de los segmentos seleccionados según el eje Y.
- Seleccionar, según la ordenación efectuada, aquéllos susceptibles de intersectar.
- Aplicar algoritmo de intersección.

Es la base para todo lo relativo a las operaciones de puntos en polígonos.

Sean:

$$y_1 = a_1 + b_1x \quad \text{la ecuación de una recta } y_1$$

e

$$y_2 = a_2 + b_2x \quad \text{la ecuación de la recta } y_2$$

el punto de intersección se obtiene resolviendo el sistema de las dos ecuaciones lo que da:

$$x_i = -(a_1 - a_2) / (b_1 - b_2) \quad \text{coordenada } x \text{ del punto de intersección}$$

$$y_i = a_1 + b_1x_i \quad \text{coordenada } y \text{ del punto de intersección}$$

14.6. El problema del viajante

14.6.1. Base del problema

El problema del **Agente viajero** es un ejemplo que muestra y analiza la problemática que subyace tras algunos tipos de problemas matemáticos que a priori parecen tener una solución relativamente fácil, y en la práctica presentan un gran problema.

La respuesta al problema es conocida, es decir se conoce la forma de resolverlo, pero sólo en teoría, en la práctica la solución no es aplicable debido al tiempo que computacionalmente se precisa para obtener su resultado. El **problema del viajante** (también conocido como **problema del viajante de comercio** o por sus siglas en inglés: **TSP**) es uno de los problemas más famosos (y quizás el mejor estudiado) en el campo de la optimización combinatoria computacional. A pesar de la aparente sencillez de su planteamiento, el TSP es uno de los más complejos de resolver y existen demostraciones que equiparan la complejidad de su solución a la de otros

problemas aparentemente mucho más complejos que han retado a los matemáticos desde hace siglos.

14.6.2. Enunciado

Sean N ciudades de un territorio. El objetivo es encontrar una ruta que, comenzando y terminando en una ciudad concreta, pase una sola vez por cada una de las ciudades y minimice la distancia recorrida por el viajante. Es decir, encontrar una permutación $P = \{c_0, c_2, \dots, c_{n-1}\}$ tal

que $d_P = \sum_{i=0}^{N-1} d[c_i, c_{i+1 \bmod(N)}]$ sea mínimo. La distancia entre cada ciudad viene dada por la matriz $D: N \times N$, donde $d[x, y]$ representa la distancia que hay entre la ciudad X y la ciudad Y .

La solución más directa es la que aplica la fuerza bruta: evaluar todas las posibles combinaciones de recorridos y quedarse con aquella cuyo trazado utiliza la menor distancia. El problema reside en el número de posibles combinaciones que viene dado por el factorial del número de ciudades ($N!$) y esto hace que la solución por fuerza bruta sea impracticable para valores de N incluso moderados con los medios computacionales actualmente a nuestro alcance. Por ejemplo, si un ordenador fuese capaz de calcular la longitud de cada combinación en un microsegundo, tardaría algo más 3 segundos en resolver el problema para 10 ciudades, algo más de medio minuto en resolver el problema para 11 ciudades y... 77.146 años en resolver el problema para sólo 20 ciudades.

Por ejemplo las rutas posibles entre 12 ciudades son (479 millones) 479.001.600 combinaciones y los caminos individuales entre ciudades son el sumatorio de las 12-1 ciudades es decir 66.

Se puede demostrar que el requisito de volver a la ciudad de partida no cambia la complejidad computacional del problema.

14.6.3 Solución del problema

Desde el punto de vista práctico, el problema no está resuelto y desde el punto de vista teórico, las técnicas empleadas son sólo aproximaciones. No suponen una resolución real del TSP y sólo ofrecen soluciones aproximadas suficientemente aceptables.

Los algoritmos clásicos no son capaces de resolver el problema general, debido a la explosión combinatoria de las posibles soluciones. Por ello, a su solución se han aplicado distintas técnicas computacionales: heurísticas evolutivas, redes de Hopfield, etc.

14.6.4. Convergencia del problema

Una formulación equivalente en términos de la teoría de grafos es la de encontrar en un grafo completamente conexo y con arcos ponderados el ciclo hamiltoniano de menor coste. En esta formulación cada vértice del grafo representa una ciudad, cada arco representa una carretera y el peso asociado a cada arco representa la longitud de la carretera.

El TSP está entre los problemas denominados NP-completos, esto es, los problemas que no se pueden resolver en tiempo polinomial en función del tamaño de la entrada (en este caso el número N de ciudades que el viajante debe recorrer). Sin embargo, algunos casos concretos del problema sí han sido resueltos hasta su optimización, lo que le convierte en un excelente banco de pruebas para algoritmos de optimización que pertenezcan a la misma familia (lo que en jerga matemática se denominan problemas isomorfos).

14.7. Análisis de polígonos

14.7.1. Determinación del área de un polígono

El método más usual es un algoritmo que calcula el área de un número de trapezoides conformados por:

- una de las líneas-segmento del polígono
- el eje de abscisas X
- Líneas verticales desde los extremos del segmento al eje X

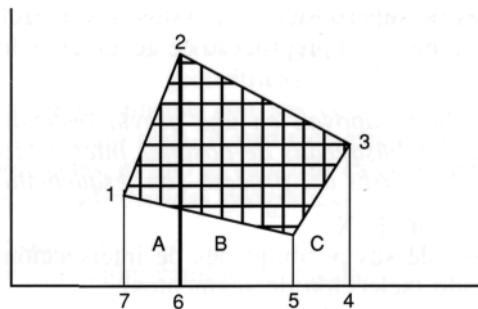


Figura 15. Determinación del área de un polígono.

Se efectúa a través del cálculo del área de cada uno de los trapezoides elementales así definidos y se suman después.

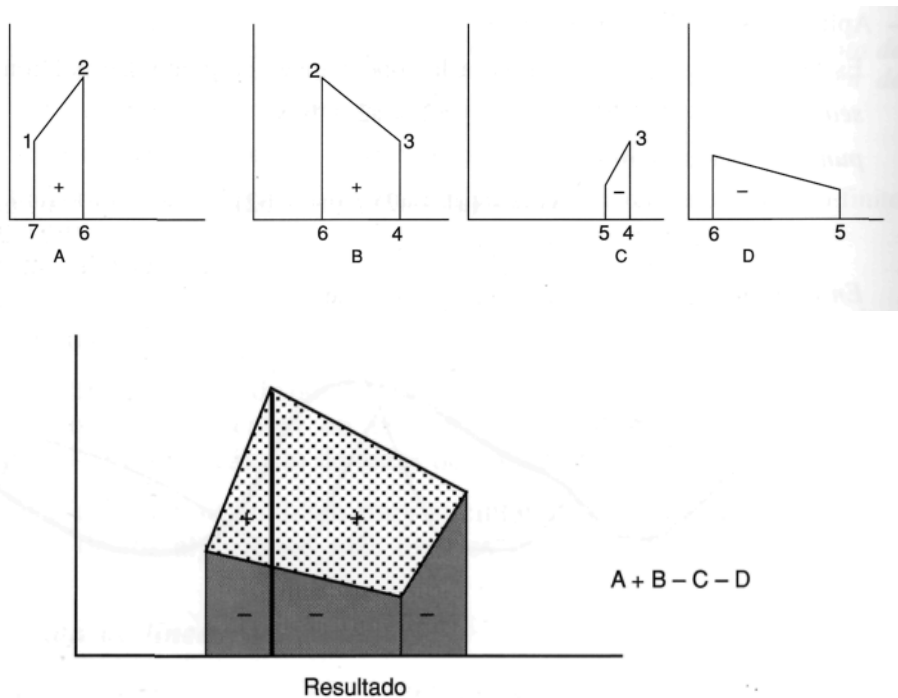


Figura 16. Cálculo del área de un polígono: Trapezoide

14.7.2. Punto en polígono

Su objetivo es el de determinar si un punto queda dentro o fuera de un polígono dado. Por ejemplo determinar qué parcelas contienen un objeto puntual tal como un pozo. Para ello se dibuja una línea vertical desde el punto al infinito, y se contabiliza el número de veces que la línea interseca con el perímetro del polígono; si este número es impar, el punto está contenido en el polígono, y si es par es externo al mismo.

Otro algoritmo muy utilizado es el que permite el posicionamiento automático de centroides, para lo cual debe determinarse el centro geométrico del polígono, donde se situará el centroide.

14.7.3. Superposición de polígonos

Es otro ejemplo típico de algoritmo de gran utilidad en las aplicaciones SIG de análisis espacial. Para realizar el proceso es preciso, en primer lugar, encontrar todas las intersecciones entre los perímetros de los polígonos que se superponen.

Ejemplos: supongamos los polígonos A y B.

Después de la intersección, se han formado 6 áreas nuevas y 4 polígonos que llamaremos A_0 , A_B y O_B

Segmento	Polígono derecho	Polígono izquierdo
1	A0	00
2	AB	0B
3	A0	00
4	00	0B
5	A0	AB
6	00	0B



Figura 17. Resultado obtenido

Los procesos informáticos de superposición son costosos en tiempo, manejando grandes cantidades de operaciones numéricas. En los equipos informáticos actuales, no obstante, pueden realizarse operaciones de esta índole que afecten a centenares de miles de polígonos en tiempos relativamente cortos.

Bibliografía

- [1] <http://es.wikipedia.org>
- [2] Network and Discrete Location Models. Algorithms and Applications. *Daskin, M. S.* Nueva York, J. Wiley. 1995.
- [3] “Fundamentals of Spatial Information System” R.LAURINI and D.THOMPSON, 1992, Academic Press.
- [4] “GIS. A Computing Perspective” M.F. Worboys, TAYLOR & FRANCIS, 1995.
- [5] “Sistemas de Información Geográfica” Joaquín Bosque Sendra, 1992, RIALP.

- [6] “Geographic Information Systems and Science” P.A. Longley, M.F. Goodchild, D.J. Maguire y D.W.Rhind, 2001, JOHN WILEY and SONS.

Tema 15. Análisis ráster. Autocorrelación espacial, reclasificación y superposición (*overlay*) de información geográfica.

15.1. Análisis ráster. Autocorrelación Espacial

La autocorrelación espacial es la relación entre los valores de una variable que se puede atribuir a la forma en la que los puntos de muestreo de esta variable están ordenados o dispuestos en un espacio unidimensional (transecto) o bidimensional (plano).

Decimos que existe autocorrelación positiva cuando los valores son similares entre puntos próximos o autocorrelación negativa cuando los valores son distintos. Se mide a través de las funciones de estructura como los correlogramas y los semivariogramas.

Existen varios tipos de autocorrelación espacial:

- Espúrea.
- Interpolada.
- Verdadera (relación causal entre vecinos).
- Inducida (por otra variable independiente con autocorrelación espacial).

Estas dos últimas provienen de procesos espaciotemporales como la dispersión, el establecimiento, la competencia, la migración, etc.).

La medida cuantitativa de la autocorrelación se hace a través de los: índices de correlación espacial. Los dos índices más importantes son:

- Índice de Moran (I). Es un índice que correlaciona parecidos:

$$(1) \quad I = \frac{N \sum_i \sum_j W_{ij} (x_i - \bar{x}) (x_j - \bar{x})}{\sum_i \sum_j W_{ij} \cdot \sum_{i=1}^N ((x_i - \bar{x})^2)}$$

- Índice de Geary (C). Es un índice basado en la covarianza que mide diferencias

$$(2) \quad C = \frac{n-1}{2 \sum_i \sum_j W_{ij}} \cdot \frac{\sum_i \sum_j W_{ij} (x_i - x_j)^2}{\sum_i (x_i - \bar{x})^2}$$

Estos índices tienen las siguientes características:

Índice de Moran:

- como un coeficiente de correlación, varía entre +1 y -1, aunque puede superar ambos límites.
- Autocorrelación positiva (valores similares entre puntos próximos, sean valores altos o sean valores bajos –no los distingue): I tiende a +1.
- Autocorrelación nula, $I = 0$ (más concretamente $I = -(N-1)^{-1}$, siendo N el número de puntos).
- Autocorrelación negativa: I tiende a -1.

Índice de Geary:

- Similar a la semivarianza, c suele variar entre 0 y 3
- Autocorrelación positiva: c tiende a 0 o es menor de 1 (sean valores altos o sean valores bajos –no los distingue).
- Autocorrelación nula: $c = 1$.
- Autocorrelación negativa: c tiende a 2 o es mayor de 1.

En la siguiente tabla aparece la interpretación del índice de autocorrelación espacial I

		Distancia	
		Corta (escala pequeña)	Larga (escala grande)
Signo de la autocorrelación (I)	Positivo	1) dispersión y agregación	1) superficies simétricas
		2) manchas favorables grandes	2) disposición en manchas cíclicas
		3) gradiente (tendencia)	
	Negativo	1) área heterogénea	1) gradiente (tendencia)
		2) manchas pequeñas	
		3) plantas regularmente espaciadas	
		3) escala de muestreo grande	

Tabla 1. Autocorrelación espacial I

Un estudio mas profundo de ambos índices nos lleva a las siguientes conclusiones:

- Índice I de Moran: parece estar menos afectado por la distribución de los datos que c.
- Pero, el índice I de Moran parece ser más sensible a los valores extremos.
- El índice I parece más intuitivo y fácil de interpretar.
- El índice c se acerca a la semivarianza γ , que es menos sensible que I a la falta de estacionariedad.

15.1.1. Correlogramas y semivariogramas

Un correlograma es la representación del índice de correlación en función de la distancia.

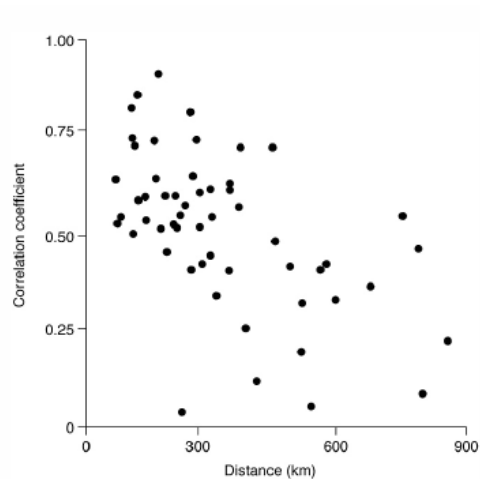


Figura 1. Correlograma

Para estudiar el correlograma se ha de obtener un valor de peso (α) de cada coeficiente de autocorrelación mediante un test de permutaciones. Estos coeficientes no son independientes ya que son varios tests simultáneos, por lo que para determinar si el correlograma es globalmente significativo se debe hacer una corrección a la media y calcular un nuevo nivel de significación (α'):

$$(3) \quad \alpha' = \alpha / k$$

donde k es el número de clases de distancia.

Para que el correlograma sea globalmente significativo al menos un coeficiente debe mostrar una significación menor o igual que α' . Por ejemplo: si $\alpha = 0,05$ y $k = 20$, nos queda que

$$(4) \quad \alpha' = 0,05 / 20 = 0,0025$$

15.1.2. Tipos de correlogramas

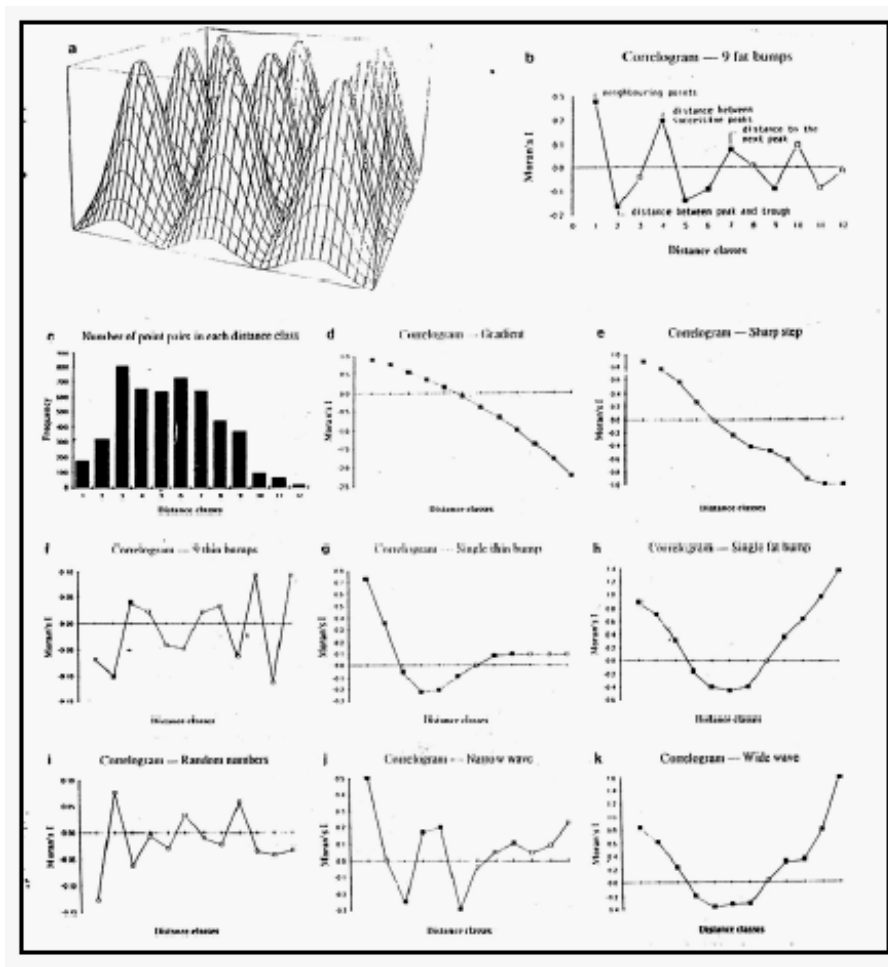


Figura 2. Ejemplos de correlogramas

15.1.3. Variograma

El variograma $\gamma(h)$ es una función matemática que mide la variabilidad, es decir el cambio una variable cuando sus valores se observan c-n puntos que están separados una distancia concreta (h). En concreto, el variograma se expresa mediante la ecuación:

$$(5) \quad \gamma(b) = \frac{\sum_i (x_i - \bar{X})(x_j - \bar{X})}{2N}$$

donde X_i e X_j son los valores de una variable en dos puntos del espacio (i, j) que están separados una distancia h, \bar{X} es la media aritmética de esa variable.

Para calcular el variograma se generan todas las parejas de puntos/objetos espaciales que estén separados 1, 2, 3, 4... unidades de distancia a lo largo de una dirección del plano (norte-sur, este-oeste o cualquier otra) y se calcula $\gamma(h)$ en cada uno de estos grupos de casos. Así se obtienen dos variables: h, la distancia de separación entre puntos/objetos espaciales y $\gamma(h)$. Con estas dos variables se puede construir un gráfico de dispersión que permite estudiar distintos aspectos de la autocorrelación espacial de X.

Se utiliza, como término de comparación, un modelo ideal de variograma, el denominado modelo esférico o de Matheron. Su ecuación es:

$$(6) \quad \begin{array}{l} \text{Para } b < r \\ \text{Para } b > r \end{array} \quad \gamma(b) = C \left[\frac{3b}{2r} - \frac{b^3}{2r^3} \right]$$

$$\gamma(b) = C$$

15.1.4. La correlación espacial en entornos de variable geográfica discreta

Ya se ha visto antes que la autocorrelación espacial tiene que ver tanto con la localización geográfica como con los valores hallados de la variable que se esté estudiando. Para determinar si el patrón de distribución espacial dista del meramente aleatorio, debe utilizarse un índice de comparación.

Todos los índices dedicados a la medida de la autocorrelación espacial (ver párrafo anterior) poseen una raíz común: la “matriz de producto cruzado”, o “estadístico general de producto cruzado”.

$$(7) \quad \Gamma = \sum_i \sum_j W_{ij} C_{ij}$$

Donde la matriz W_{ij} recibe el nombre de matriz de conexión, de contigüidad o de peso espacial. Sus valores representan una forma de medición de la contigüidad en los datos originales. La matriz C_{ij} , por su parte, es una medida de la proximidad de los valores i, j, en otra dimensión (por ejemplo, distancia euclídea, distancia esférica, distancia de Manhattan, etc.).

La matriz W_{ij} se compone de ceros y unos, según se considere la existencia de contigüidad, o no, entre localizaciones geográficas. Tras la presentación de los 3 criterios principalmente

utilizados por los geógrafos, incluidos en el software informático de análisis geográfico más extendido, introduciremos el que va a ser utilizado en este estudio debido a las diferencias en la obtención de datos respecto del análisis geográfico habitual.

En general, los datos geográficos se presentan en un mapa continuo, como en el siguiente ejemplo de tan sólo nueve localizaciones próximas:

a	b	c
d	e	f
g	h	i

Figura 3. Ejemplo para correlación espacial.

La contigüidad respecto de la localización central puede entonces definirse de tres maneras principales:

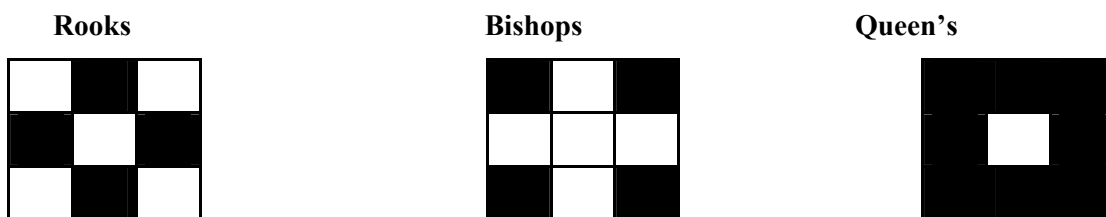


Figura 4. Tipos de modelos de contigüidad.

En el caso de Rooks, el más utilizado, se consideran adyacentes a la localización “e” las b, d, f, h; la contigüidad de Bishops, por el contrario, analiza las relaciones de proximidad en relación diagonal, y consideraría vecinas a la localización “e” las a, c, g, i; el criterio de Queen’s combina los dos anteriores. Las matrices W_{ij} de contigüidad quedarían de la siguiente manera, en función del criterio de vecindad elegido:

En la matriz de contigüidad-Rooks, por ejemplo, corresponde un “1” en las casillas de intersección de la celda “e” con las b, d, f, h, y un “0” en las casillas de intersección de la celda “e” con el resto. La celda “a” es considerada contigua a las celdas b y d, por lo que asignaríamos sendos “1” en los citados puntos de intersección, y ceros en las casillas correspondientes a la relación de la localización “a” con las restantes. Por convención, una localización no se considera adyacente a sí misma. Sólo se estudian sus posibles relaciones con

las localizaciones vecinas. Procediendo según lo descrito, pueden construirse las siguientes 3 matrices posibles de peso espacial:

	a	b	c	d	e	f	g	h	i
a	0	1	0	1	0	0	0	0	0
b	1	0	1	0	1	0	0	0	0
c	0	1	0	0	0	1	0	0	0
d	1	0	0	0	1	0	1	0	0
e	0	1	0	1	0	1	0	1	0
f	0	0	1	0	1	0	0	0	1
g	0	0	0	1	0	0	0	1	0
h	0	0	0	0	1	0	1	0	1
i	0	0	0	0	0	1	0	1	0

Figura 5. Matriz de contigüidad de Rooks

	a	b	c	d	e	f	g	h	i
a	0	0	0	0	1	0	0	0	0
b	0	0	0	1	0	1	0	0	0
c	0	0	0	0	1	0	0	0	0
d	0	1	0	0	0	0	0	1	0
e	1	0	1	0	0	0	1	0	1
f	0	1	0	0	0	0	0	1	0
g	0	0	0	0	1	0	0	0	0
h	0	0	0	1	0	1	0	0	0
i	0	0	0	0	1	0	0	0	0

Figura 6. Matriz de contigüidad de Bishops

	a	b	c	d	e	f	g	h	i
a	0	1	0	1	1	0	0	0	0
b	1	0	1	1	1	1	0	0	0
c	0	1	0	0	1	1	0	0	0
d	1	1	0	0	1	0	1	1	0
e	1	1	1	1	0	1	1	1	1
f	0	1	1	0	1	0	0	1	1
g	0	0	0	1	1	0	0	1	0
h	0	0	0	1	1	1	1	0	1
i	0	0	0	0	1	1	0	1	0

Figura 7. Matriz de contigüidad de Queen's

Este enfoque geográfico relativo a datos continuos presenta una limitación importante para el análisis de este trabajo: como ha sido explicado, el número de localizaciones vecinas queda limitado a 4 (criterios de contigüidad de Rooks y Bishops) o a 8 (criterio de Queen's). En el caso del presente análisis, va a otorgarse el valor 1 correspondiente a vecindad si existe una frontera común entre las regiones analizadas, y valor cero, si no existe ningún tramo fronterizo común. No se puede sistematizar a 4 o a 8 el número de regiones vecinas a cada región dada, por lo que no puede utilizarse directamente el mencionado software informático diseñado por los geógrafos.

El índice habitualmente utilizado es una adaptación del I de Moran, la primera medida de la autocorrelación espacial en el estudio de fenómenos estocásticos distribuidos en un espacio de dos o más dimensiones.

Este índice resulta análogo al coeficiente de correlación convencional, ya que su numerador se interpreta como la covarianza entre unidades contiguas, y sus valores oscilan entre +1 (significando fuerte correlación espacial positiva) y -1 (significando fuerte correlación espacial negativa).

La significatividad estadística del índice I de Moran puede ser obtenida a través de dos procedimientos: en primer lugar, mediante un proceso de contraste cuya hipótesis nula propone que los datos analizados no son sino una muestra aleatoriamente obtenida a partir de una de las $n!$ posibles distribuciones espaciales de las variables estudiadas entre las n localizaciones. Y, en segundo lugar, a partir de la utilización de estadísticos diseñados para la autocorrelación espacial basados en la aproximación normal.

Se puede demostrar que el valor esperado del I de Moran es

$$(8) \quad E_N(I) = \frac{-1}{(n-1)},$$

valor que tiende a cero conforme aumenta n, el número de conjuntos de datos analizados; la varianza del I de Moran viene dada, a su vez, por la siguiente expresión:

$$(9) \quad \text{Var}_N(I) = \left(\frac{1}{S_0^2(n^2-1)} (n^2 S_1 - n S_2 + 3 S_0^2) \right) - E_N(I)^2$$

donde:

n = número de observaciones,

$$S_0 = \sum_{i=1}^n \sum_{j=1}^n W_{ij}$$

suma de la matriz de peso espacial

$$(10) \quad S_1 = \frac{\sum_{i=1}^n \sum_{j=1}^n (W_{ij} + W_{ji})^2}{2}$$

si la matriz de peso espacial es simétrica, entonces

$$(11) \quad S_1 = 2 \sum_{i=1}^n \sum_{j=1}^n W_{ij}$$

$$S_2 = \sum_{i=1}^n (W_{i.} + W_{.i})^2 \quad \text{Es la suma de (la columna } i \text{ + la fila } i)^2 \text{ de la matriz de peso espacial.}$$

si la matriz de peso espacial es simétrica, entonces

$$(12) \quad S_2 = 4 \sum_{i=1}^n W_{i.}^2$$

La desviación típica y los valores z correspondientes a una distribución normal N(0,1) vienen dados por:

$$(13) \quad DT_N(I) = \sqrt{\text{Var}_N(I)}$$

$$(14) \quad z = \frac{(I - E_N(I))}{\sqrt{\text{Var}_N(I)}}$$

15.1.5. Correlación-autocorrelación de imágenes

Los métodos fotogramétricos de captura de datos digitales son ampliamente usados hoy en día para la obtención de coberturas de información ráster. Tanto por métodos de fotografías aéreas como obtención de imagen espacial se requiere un control terrestre.

Cuando un modelo fotogramétrico está orientado absolutamente en un instrumento, cualquier punto del modelo puede ser accedidos por la marca de medición del instrumento, por lo que sus coordenadas pueden registrarse en un archivo digital.

La técnica de correlación de imágenes (también llamada correlación cruzada), se fundamenta en la comparación de imágenes digitales a partir de pares estereoscópicos de fotografías aéreas o espaciales. Este es también un método automatizado de generación de MDT.

En este método se extraen subimágenes homólogas digitales de ambas imágenes, izquierda y derecha, que subsecuentemente se correlacionan por medio de su información radiométrica (los valores de la escala de grises), para que se determinen puntos de correspondencia a partir de píxeles homólogos de cada imagen.

Para calcular las coordenadas planimétricas se hace uso tanto de las coordenadas de "fotografía" obtenidas por transformación píxel a imagen para todos los pares de píxeles homólogos, así como de los elementos de orientación exterior de las imágenes.

En la figura se muestra el esquema en el que la primera matriz, o matriz de "referencia" (se supone la imagen izquierda), se toma como elemento para comparar los tonos de gris con la segunda matriz, o matriz de "búsqueda" (se supone la imagen derecha).

Se ha de considerar que tanto en el mundo fotogramétrico digital como en el mundo de la teledetección espacial el trabajo con imágenes parcialmente o totalmente coincidentes en el territorio, o desplazadas temporalmente se han de constituir como un conjunto de datos coherente en su dimensión planimétrica (X,Y). La superposición de imágenes georreferenciadas ha de ser perfecta a nivel de subpíxel para garantizar el ajuste geométrico multitemporal o el para el solape entre fotos digitales a fin de construir un MDT o una ortofoto. Existen distintas técnicas de ajuste de imágenes de forma que se lleva una de ellas alas referencia de la otra. Este ajuste siempre se ha de basar en que una de ellas ya esta rectificadas y georreferenciadas y es la segunda o mas imágenes las que se llevan a encajar con la primera. Esta tarea esta integrada en las herramientas de ajuste de imágenes en el mundo del tratamiento de imágenes digitales. Generalmente podemos destacar dos métodos:

- Basados en la geometría.
- Basados en signatura espectral (tonos de gris).

Los métodos basados en la geometría pueden a su vez dividirse en:

- Polinómicos
- Modelización geométrica paramétrica

15.1.6. Modelos Polinómicos

Si dos imágenes han de ser comparadas (por ejemplo para detección de cambios) es necesario registrarlas. En esta técnica una imagen se toma como referencia y la otra se adapta sobre la primera. Se ignora el problema de localización geográfica al elegir la imagen de referencia. La complejidad del registro depende de las diferencias entre las dos imágenes. Si las dos imágenes son de la misma escala y geometría, se pueden registrar mediante una rotación y una traslación de una imagen respecto a la otra. Una función de correlación cruzada es buena para calcular esta transformación. Para otros casos, como por ejemplo imágenes tomadas a diferente altitud, con diferentes sensores, etc, han de emplearse técnicas más sofisticadas.

La técnica más compleja es la representación (*mapping*) también llamada rectificación, y generalmente cuando hablamos de corrección geométrica siempre nos referimos a esta técnica. Las imágenes adquiridas por los satélites contienen diferentes distorsiones geométricas que les impiden ser utilizadas como mapas. El proceso de corrección es dar a la imagen esa cualidad de mapa, combinando la adopción de un sistema de coordenadas a la imagen con un sistema métrico de distancias sobre ella.

La complejidad de esta transformación la vemos en la Figura 9, donde se suponen coordenadas UTM. La imagen original está dada en un sistema de coordenadas (x,y) y debe ser transformada a un sistema de coordenadas rectangulares (u,v) de la proyección UTM. La salida son los correspondientes valores de los píxeles en la cuadrícula del sistema (u,v) .

El primer paso es la identificación de los mismos puntos en ambos sistemas (x,y) y (u,v) en sus correspondientes cuadrículas para determinar la transformación entre ambos espacios. Esta transformación permitirá después trasladar todos los puntos del sistema (x,y) al (u,v) mediante interpolación.

La corrección geométrica tiene por tanto dos fases. Primero determinar la transformación de imagen a mapa (corrección propiamente dicha o rectificación). Segunda asignar el

correspondiente valor de gris a la coordenada de salida en función del valor de entrada (interpolación).

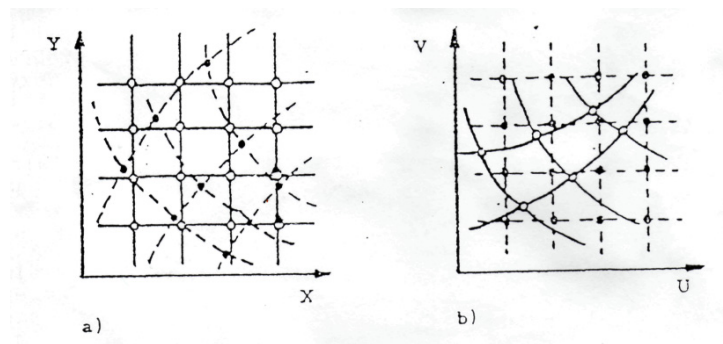


Figura 8. Transformación geométrica de imagen. La figura a) representa la imagen original en puntos negros en un sistema de malla regular en puntos blancos. La figura b) representa la imagen corregida geoméricamente en puntos negros con la malla original del sistema en puntos blancos, la cual ha sido distorsionada.

Los errores aleatorios que no se pueden corregir por desconocer su comportamiento exacto y se tienen en cuenta de forma única en el procedimiento de ajuste. Este método también corrige de los errores sistemáticos residuales. Hay dos formas de trabajar en este caso ya que se pretende obtener un camino de corrección de deformaciones y distorsiones entre la imagen y el correspondiente mapa. La diferencia entre ambos está en considerar el movimiento de la plataforma espacial o no. En el primer caso se usa la Representación Polinomial para la corrección, en el segundo caso se usa una modelación de la toma de vista del sensor. Vamos a ver el primer caso.

Las funciones que se usan para la transformación presuponen una distorsión residual suave, por lo que se utilizan polinomios de corrección bivariados de expresión general:

$$\begin{aligned}
 v &= \sum_{p=0}^N \sum_{q=0}^{N-p} a_{pq} x^p y^q \\
 u &= \sum_{p=0}^N \sum_{q=0}^{N-p} b_{pq} x^p y^q
 \end{aligned}
 \tag{15}$$

donde n es el orden del polinomio. Si $n=0$ los polinomios representan una traslación de coordenadas. Si $n=1$ se corrige de traslación, de rotación y de cambios de escala. Si $n=3$ se incluyen correcciones residuales aleatorias. La implementación de polinomios de orden superior a 2 se hace complicada siendo suficiente, dada la calidad de los datos de los sensores actuales, el orden $n=2$. Si trabajamos con imágenes aéreas (p.e.) habría que aplicar $n=3$ o hasta $n=5$.

A un píxel de la imagen generalmente no le corresponderá un punto en el sistema de salida (u,v) , ya que el polinomio dará valores reales (no enteros) para los valores enteros de salida. De aquí que sea necesario obtener un valor del píxel de salida mediante una interpolación.

También se pueden usar transformaciones más simples (con menor precisión), una de ellas es la transformación afín

$$(16) \quad \begin{aligned} X &= A u + B v + C \\ Y &= D u + E v + F \end{aligned}$$

donde los parámetros $A, B, C, D, E,$ y F se estiman por ajuste de mínimos cuadrados a las coordenadas de los puntos de control del terreno (GCP).

El número de GCP necesarios para el ajuste depende del orden del polinomio. Se suele tomar dos puntos por coeficiente del ajuste. Así para una deformación bilineal bastarían 6 puntos y para una cúbica 20 como mínimo. Para mejor ajuste por el método de Mínimos Cuadrados no importan que sean más puntos, aunque ello tiene asociado un mayor esfuerzo por parte del operador y por el ordenador (tiempo de CPU).

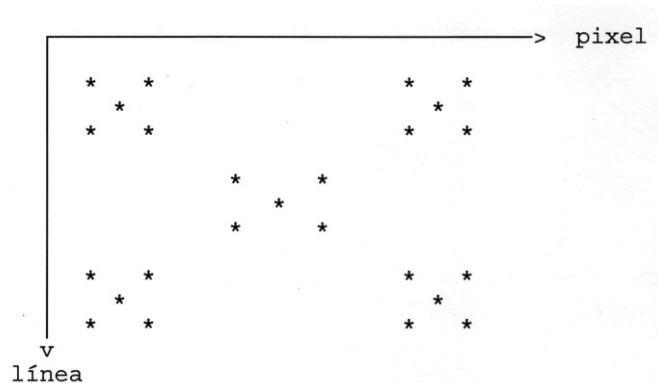


Figura 9. Distribución de los puntos de control.

Otro aspecto muy importante es la distribución sobre la imagen de estos GCP. La distribución debe ser más o menos uniforme, aunque los puntos de borde son más importantes que los del centro. La distribución adoptada en el IGN es de realizar en la imagen 5 aspas: cuatro de ellas en las esquinas y una en el centro, obteniéndose 25 puntos de control como mínimo.

La transformación inversa es necesaria ya que es a partir de ella como se interpolara el píxeles desde los valores de entrada conocidos para los puntos de salida calculados.

Cuando se tienen hallados los GCP de una imagen, no es necesario volverlos a tomar para otra imagen del mismo terreno, generalmente de otra poca. Se suele utilizar un procedimiento de autocorrelación que da la localización del GCP en la imagen en base a la máxima coincidencia entre ambas mediante la selección de ventanas alrededor de los puntos GCP de la imagen inicial. Los algoritmos son muy complicados y el tiempo de cálculo es muy elevado, depende del tamaño de la ventana y del método elegido. Da una idea que el número de operaciones para una imagen de 256*256 píxeles con una ventana de 64*64 oscila entre $7 \cdot 10^{**8}$ y $2.2 \cdot 10^{**6}$.

El problema del remuestreo es asignar los valores de gris a los píxeles obtenidos en la transformación de coordenadas (x,y) a (u,v) . Observando la Figura 18.y se ve que los círculos negros están ligados por píxeles de intensidad conocida. Como mínimo los círculos negros se obtendrán de los 4 píxeles más próximos, aunque la influencia de una vecindad mayor también puede tenerse en cuenta.

La más simple y más fácil de implementación de las funciones de interpolación es la del vecino más próximo, en la cual el valor de intensidad del píxel de salida se asigna como el del píxel más próximo en el espacio de entrada correspondiente. Este interpolador se utiliza en funciones de cambio de escala (magnificación, etc.) de imágenes.

Un segundo interpolador es el bilineal usando la siguiente función:

(17)

$$X(u,v) = a_1 X(x,y+1) + a_2 X(x,y) + a_3 X(x+1,y+1) + a_4 (x+1,y)$$

donde a_1 , a_2 , a_3 , y a_4 son las distancia euclídeas entre el punto de salida y los 4 de entrada mas próximos. Notar que este interpolador suaviza la imagen.

La teoría general de los interpoladores se encuadra dentro del dominio de las transformadas de Fourier. No entraremos aquí en ella. Solamente decir que el interpolador por el vecino más próximo pasa más las altas frecuencias lo cual provoca discontinuidades en la imagen y que en esta teoría todos los píxeles de la imagen original (a partir del espectro de potencias) influyen en el valor de gris del píxel de salida.

Muy utilizado es el tercer interpolador, llamado de convolución cúbica, en el que se usan polinomios bivariados cúbicos de tercer orden. En este esquema se tiene en cuenta 16 píxeles de

la vecindad de píxeles de remuestreo. Este método asegura que la interpolación pase por los valores conocidos en los centros de los píxeles y además que sus dos primeras derivadas sean suaves. Esto produce una gran conservación del aspecto visual de la imagen aunque introduce un suavizado muy poco considerable. Es el mejor de los tres, pero necesita mucho más cálculo. En el método de interpolación cúbica se usan los 16 píxeles circundantes a uno dado. Polinomios cúbicos son ajustados a lo largo de cuatro líneas de cuatro píxeles alrededor del punto de la imagen para formar cuatro interpolantes. Un quinto polinomio cúbico es ajustado con los cuatro valores anteriores con lo que se obtiene un valor síntesis de 16 para el valor de gris de salida en la corrección geométrica.

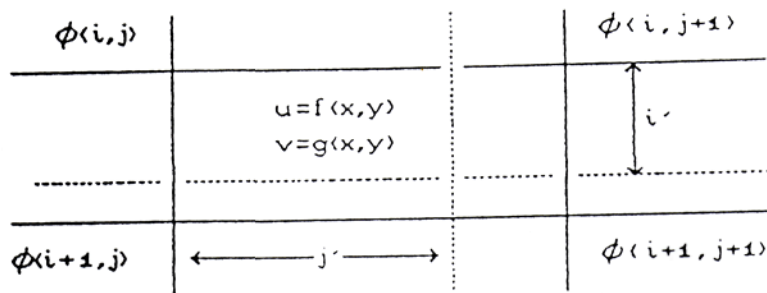


Figura 10. Interpolación por el vecino más próximo. ϕ representa el porcentaje de valores de gris.

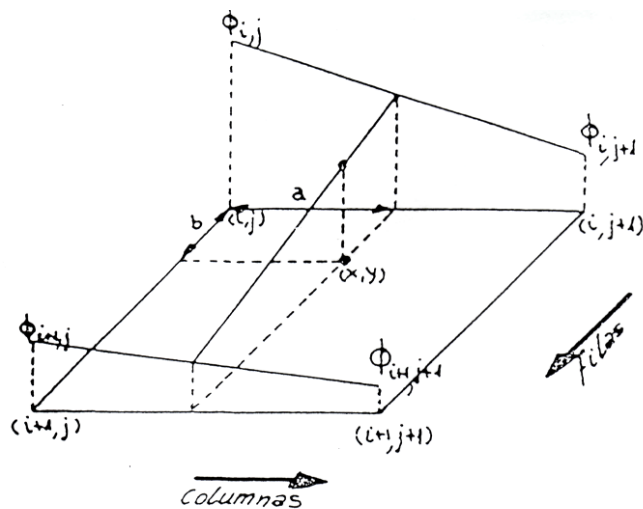


Figura 11. Interpolación bilineal.

La elección entre uno de los tres de métodos depende de la aplicación en particular. Si la imagen remuestreada es para posterior clasificación no se pueden cambiar los valores radiométricos iniciales y es mejor el método del vecino más próximo. Si la imagen corregida es para una fotointerpretación es mejor elegir una convolución o quizás la bilineal. De todas

formas depende bastante de la dinámica de la imagen, del ordenador disponible y del tiempo disponible en este ordenador.

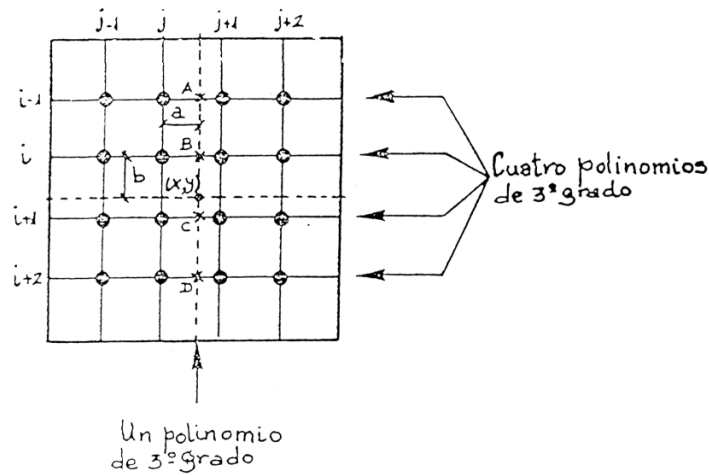


Figura 12. Esquema gráfico del Interpolador Cúbico.

15.1.7. Modelización geométrica paramétrica

Este método se aplica tanto en fotogrametría para restituir pares de fotogramas, o bandas o bloques, o conociendo los datos de la órbita del satélite con suficiente precisión se puede utilizar para corregir las imágenes de satélite. Es método es imprescindible cuando se quieren corregir pares estereoscópicos de imágenes SPOT, IKONOS, QuicBird, Terrasar, RADARSAT, etc., y para obtener modelos de relieve de la imagen común a comparar. También es imprescindible tener los valores de altitud de los puntos de apoyo del canevas de control necesario.

Un modelo geométrico de una imagen es la descripción de las relaciones que ligan la posición de un punto sobre el terreno, con su imagen captada por el sensor. La geometría de estas imágenes se suele decir que es cilindro-cónica.

En fotogrametría clásica, al ser tomada la imagen de forma instantánea, solo son necesarios 6 parámetros para modelizar la geometría, estando resuelto el problema desde hace tiempo con toda rigurosidad (Fotogrametría analítica). Sin embargo cuando la imagen tiene un tiempo finito de captura, caso de todos los sensores de barrido, hemos de relacionar cada punto de la imagen con la posición del rayo perspectivo a que corresponde. La ventaja añadida a esta modelización es la reducción de los puntos necesarios para la corrección de las imágenes, pasando de 20 de una escena entera a 10 para una banda completa de hasta 10 escenas consecutivas. Las principales aplicaciones de este método de modelización física tienen su principal aplicación en

el satélite SPOT, dadas sus características de estabilidad y corrección de parámetros orbitales, que no tienen con suficiente precisión el resto de plataformas actualmente en órbita.

Los métodos basados en signatura espectral establecen la correspondencia entre imagen e imagen mediante la comparación del tono de gris de cada una de ellas. Se supone ya presente un cierto ajuste de posición relativa entre las imágenes de forma que los desplazamientos que puedan existir entre ellas no han de ser superiores a 3 ó 4 píxeles en cualquiera de las direcciones (X, Y). Los ajustes se pueden realizar a tendiendo a dos técnicas principalmente: locales o de ventana, y Globales.

Los métodos locales establecen una correspondencia entre puntos de igual característica o identificación en ambas imágenes y se aplican filtros espaciales de convolución a las ventanas correspondientes centradas en los elementos comunes identificados a través de sus coordenadas de píxel (Columna, Línea) o (X, Y) geográficas si las imágenes se han georreferenciado independientemente y hay que ajustarlas. Se obtiene así una nube de puntos de correlación que alimenta un proceso de ajuste por mínimos cuadrados como en el caso anteriormente estudiado de rectificación-georreferenciación de la segunda imagen sobre la primera. Los errores han de ser siempre residuales del orden del subpíxel pero siempre se arrastrará la precisión de la corrección de la primera imagen sobre la segunda por la ley de composición de errores. Cuantos mas puntos-ventana se seleccione más exacto será el ajuste aunque también conlleva más carga de trabaja manual. Actualmente existen algoritmos de búsqueda de correlación de ventana muy eficientes que dan muy buenos resultados en la rectificación y ajuste entre imágenes. Habitualmente se usan aplicaciones de la transformada rápida de Fourier (*Fast Fourier Transformation-FFT*) sobre ventanas discretas de nos mas de 32x32 píxeles, o la transformada discreta de coseno (DCT).

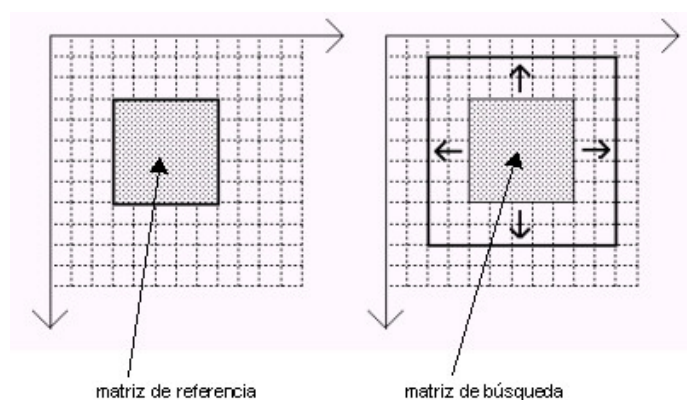


Figura 13. Ejemplos de matrices

Para cada posición se calcula el coeficiente de correlación de la matriz de búsqueda. El resultado de la aplicación de la función de correlación cruzada define la posición del mejor ajuste de la matriz de referencia y su correspondiente en la matriz de búsqueda.

(18)

$$\rho = \frac{\sum_{R=1}^R \sum_{C=1}^C (g_1(r,c) - \mu_1)(g_2(r,c) - \mu_2)}{\sqrt{\sum_{R=1}^R \sum_{C=1}^C (g_1(r,c) - \mu_1)^2 \sum_{R=1}^R \sum_{C=1}^C (g_2(r,c) - \mu_2)^2}}$$

donde:

- p= coeficiente de correlación.
- g1 (r,c) = valores individuales de gris de la matriz de referencia.
- μ_1 = valores promedio de gris de la matriz de referencia.
- g2 (r,c,) = valores individuales de la parte correspondiente de la matriz de búsqueda.
- μ_2 = valores promedio de gris de la parte correspondiente de la matriz de búsqueda.
- R,C= número de filas y columnas de la matriz de referencia.

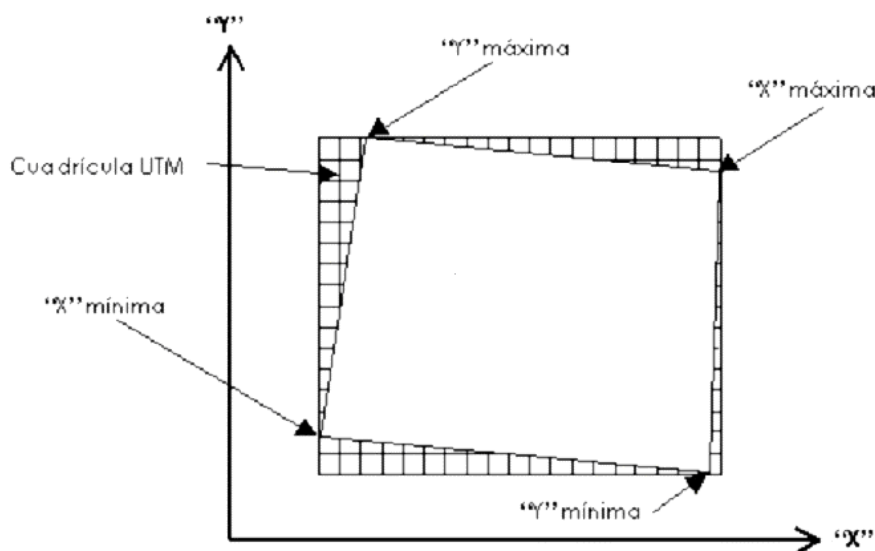


Figura 14. El esquema muestra con exageración la condición de lados verticales no paralelos de un mapa topográfico sobre la cuadrícula UTM.

Los **métodos globales** trabajan sobre ambas imágenes como un todo. Se emplea la técnica de la FFT. Para ello la imagen ha de tener un número par de filas y columnas. El ajuste se realiza a través de la selección de la ventana más idónea para el ajuste en el espectro de Fourier. La forma y tamaño de la ventana en el dominio de las frecuencias corresponde a un filtro espacial de convolución. Habitualmente se selecciona el disco central del espectro para la reconstrucción de la segunda imagen sobre la primera de forma que ambas tengan comunes los mismos componentes de las dos transformadas. Esta técnica se realiza cuando no se tiene ningún punto de control identificado y es habitual emplearla en datos radar, cartografía planetaria, correlación de Modelos del terrenos sin apoyo de cartografía, etc. El tiempo de cálculo es enorme dado la complejidad de obtener alguna correlación entre conjuntos de datos no homogéneos.

15.1.8. Consultas a un SIG ráster

El análisis de datos geográficos sobre estructuras ráster es la ventaja mas sobresaliente que tiene frente a los SIG vectoriales. Los análisis que se pueden realizar sobre un SIG se dividen en:

- Análisis de distancias:
 - Pasillos (*buffers*),
 - Contigüidad
- Análisis de superposición:
 - Intersección
 - Unión,
 - Ponderación
 - Matricial
- Filtros:
 - Mayoría
 - Máximo
 - Diversidad
 - Bordes

Esta clasificación de las operaciones de análisis sobre datos SIG proviene de Berry (1987), y a sido conceptualmente aceptada y operativamente puesta en varios paquetes de software de SIG. Berry establece tres formas de localizaciones espaciales: **punto** (representado por un píxel), **vecindad** (representado por los píxeles que rodean al punto), y **región** (representado por todos los píxeles que tiene el mismo valor de la característica temática representada).

Sobre estos tres tipos de localizaciones espaciales se puede establecer todo el conjunto de operaciones de análisis necesario para aprovechar la potencia de un SIG (Berry, 1987). Estas operaciones se pueden establecer como fundamentales dentro del análisis de mapas: reclasificación de mapas, superposición de mapas, medida de distancias y conectividad, y filtrado de mapas.

Las operaciones de análisis se pueden aplicar a un solo mapa (imagen o capa) o a un conjunto de ellos. En cualquiera de los casos la división operativa de Berry se mantiene como fundamental.

15.2. Reclasificación

La **reclasificación de mapas** es la más importante de las operaciones analíticas y actúa sobre las clases contempladas en la imagen. La entrada es una capa existente que sufre una reclasificación temática de sus valores produciendo otra capa. Los valores de salida pueden ser asignados como una función de los valores iniciales, su posición, contigüidad, tamaño o forma de la configuración espacial de las clases individuales.

15.3. Superposición (overlay) de la información geográfica.

La operación de **superposición** trabaja a partir de la coincidencia espacial, punto a punto, de la o las diferentes capas de datos que se quieren superponer. La operación de reclasificación trabajaba con una sola capa de datos de entrada, la operación de superposición debe trabajar al menos con dos capas de entrada. Aunque la operación trabaja con una comparación entre capas, esta se puede realizar bien a través del valor puntual de la 'localización específica' del píxel de coordenadas (cualesquiera que ellas sean) o bien a través de la composición de una 'región ancha' e las que se compararía clases enteras de diferentes capas que tuvieran un registro perfecto una sobre la otra.

Las operaciones clásicas que se suelen aplicar en este tipo de análisis son: suma, diferencia, multiplicación, división, raíz cuadrada, exponenciación, etc.). Otras operaciones que también se aplican siguiendo este esquema de superposición están relacionadas con parámetros estadísticos: máximo, mínimo, moda, mediana, desviación standard, etc.).

El operador de **medida de distancia** y conectividad refleja un conjunto de operaciones espaciales a través de las relaciones de un píxel con su entorno. Para poder establecer una medida cualquiera necesitamos un 'espacio de medidas' y un 'patrón de medida', en nuestro caso el espacio es la imagen y el patrón el píxel con su métrica asociada (metros, yardas, etc.).

A partir de aquí podemos establecer los criterios de distancia entre puntos del espacio de medida. La medida de distancias en SIG ráster es la normal euclidiana, excepto que interese por cualquier caso adoptar la definición en unidades de píxeles recorridos para ir de un punto a otro, o la distancia 'Manhattan'. Una vez definida una distancia podemos establecer otros operadores espaciales cuales son los de 'proximidad' y por extensión los 'pasillos'. Un extensión lógica de los operadores espaciales anteriores es la determinación del 'camino óptimo' entre dos píxeles, operador muy extendido en el análisis de los SIG vectoriales.

El conjunto de operaciones espaciales de **filtrado** se puede englobar en un conjunto mas amplio de tratamiento de imágenes digitales conocido como **filtrado espacial**. Si la ventana de filtro es de orden 3x3, interviene los 8 píxeles de alrededor de uno dado para contribuir al valor temático de salida de la operación resultante, esta es la vecindad de 'primer orden'. Cuando intervienen mas píxeles de la vecindad inmediata hablamos de vecindad de orden superior, y en caso de filtrado espacial digital, se expresa por el tamaño de la ventana de actuación del filtro, generalmente cuadrada, medido en unidades de píxel: 5x5, 9x9, ... Cuando definimos este tipo de ventanas de filtro, podemos hacerlas direccionales (filtros de Prewit), aunque su definición mas allá de uno orden de 3x3 tiene poca aplicación.

Bibliografía

- [1] Burrough, P.A. (1985): Principles of Geographical Information System for Land Resources Assesment.Oxford Science Publications
- [2] Antenucci, John C.; Brown, Kay; Croswell, Peter L.; Kevany, Michael J.; Archer, Hugh: Geographics Information Systems: A guide to the Technology. Van Nostrand Reinhold. New York.
- [3] Mulder, N.J.: Data Bases, Geo Information Systems. ITC. 1984
- [4] Zobrist, A.L.; Bryant, N.A.: Designing an Image Based Information System.
- [5] Ripple, W.J. (Editor): Fundamentals of Geographic Information Systems: A compedium. ASP&RS.1989
- [6] Vivas, P.: Sistemas de proceso de Imagen del IGN. Publicación Técnica nº24. IGN. Madrid. 1992.
- [7] Robinson, A.H. et al. Elementos de Cartografía. Editorial Omega.1987

Tema 16. Normas para la información geográfica. Normas ISO 19100. Contenido y campo de aplicación. Relación con otros organismos de normalización de información geográfica (CEN, AENOR, etc). El Open Geospatial Consortium. Especificaciones de interoperabilidad.

16.1. Normas para la Información Geográfica

Todos los sectores tecnológicos e industriales han pasado por diferentes etapas de desarrollo. En concreto, la normalización de procesos y productos supone alcanzar un grado de madurez cualitativamente esencial.

Nadie imagina tener que acudir al fabricante de un vehículo para comprarle una arandela o un tornillo que sólo se puede utilizar para un modelo específico o una marca determinada. Lo que espera cualquier usuario, es que el tornillo esté ajustado a norma y tenga unos valores estandarizados de tamaño y forma. De esta manera, la fabricación, distribución y manipulación de este tornillo no depende exclusivamente del fabricante del vehículo ni de ninguna solución tecnológica en particular.

De manera parecida, estamos acostumbrados a que todos los coches se conduzcan prácticamente igual, a que no tenga sentido aprender a manejar cada marca comercial de coche por separado. Estos dos aspectos, disponer de componentes universales e interfaces de uso normalizadas, son básicos para la madurez de cualquier sector industrial.

Pues bien, este tipo de situaciones, que parecen inadmisibles en el ámbito de la automoción, son las que se han sufrido y, en no pocos casos, se siguen sufriendo en el campo de la información geográfica.

Hasta hace pocos años, se disponía de información geográfica en diferentes formatos propietarios que seguían diferentes modelos conceptuales, y aplicaciones para trabajar con estos datos “tan particulares” que, por otra parte, no conseguían manipular ningún otro dato que tuviera diferente formato o que estuviera sobre otra plataforma.

Pensar en transferir información geográfica en esas condiciones a un cliente que trabaje con otro formato diferente al del suministrador, supone trabajos muy costosos, que no siempre se pueden automatizar y que, prácticamente siempre, implican pérdida de información en el trasvase. Bajo

este horizonte, las fronteras entre diferentes productores de datos suponían verdaderos muros no permeables que exigían gran cantidad de recursos para ser superados.

Por otra parte, los procesos no basados en normas, sino en soluciones específicas, están definidos para unas necesidades muy concretas y particulares, y no pueden solucionar por su propia concepción, el objetivo de compartir y reutilizar la información para un tercer cliente o para el público en general.

La verdadera solución consiste en establecer una normativa de ámbito lo más amplio posible que permita materializar mecanismos de intercambio, interoperabilidad y distribución de información geográfica digital.

16.1.1 Normas, estándares, recomendaciones y especificaciones

Para fijar el vocabulario a emplear en todo lo referente a normalización, es conveniente distinguir entre normas, estándares y recomendaciones.

Norma es todo documento que armoniza aspectos técnicos de un producto, servicio o componente, definido como tal por algún organismo de normalización, como son ISO, CEN o AENOR. En ocasiones se les llama normas de jure o normas de derecho.

Estándar es cualquier documento o práctica que, sin ser norma, está consagrado y aceptado por el uso y cumple una función similar a la de una norma. Incluye los documentos de tipo normativo que no han sido definidos por un organismo oficial de normalización. En ocasiones se les llama normas de facto o normas de hecho.

Ejemplos: las especificaciones de *Open Geospatial Consortium*, los formatos DGN, shape, etc.

Recomendación es una directriz que promueve un organismo que intenta armonizar prácticas y uso en una comunidad determinada. Su mayor o menor éxito depende de la influencia que es capaz de ejercer el organismo que la propone. Por ejemplo: EUROSTAT produce recomendaciones para armonizar las prácticas estadísticas en Europa; OSGEO recomienda la utilización de un servicio de mapas con parámetros específicos, el llamado WMS-C; el Consejo Superior Geográfico define recomendaciones basadas en el consenso, etc.

Especificación es una descripción técnica, detallada y exhaustiva de un producto o servicio, que contiene toda la información necesaria para su producción. Algunas especificaciones pueden ser adoptadas como normas o como estándares.

Existe cierta confusión ya que en inglés, una misma palabra (*standard*) se emplea para nombrar tanto a normas como a estándares.

16.1.2. Normas y estándares funcionales.

Han aparecido algunas iniciativas mediante las cuales organismos nacionales e internacionales de normalización han definido normas sectoriales o nacionales para el intercambio de Información Geográfica (IG).

Estas tareas de normalización en el campo de la información geográfica comenzaron a finales de los años ochenta, motivados por requisitos de normalización acuciantes en un sector determinado, y dieron lugar a lo que ISO/TC 211 ha denominado normas y estándares funcionales (*functional standards*):

- DIGEST (*Digital Geographic information Exchange Standard*) en el entorno de aplicaciones de defensa, definido por el grupo DGIWG (*Digital Geographic Information Working Group*).
- S-57, para el intercambio de cartas náuticas, definido por la (OHI) Oficina Hidrográfica Internacional.
- GDF (*Geographic Data File*) para tráfico por carretera, definida por CEN/TC278, y luego adoptada como norma ISO por ISO/TC 204

Es decir, normas o estándares prácticas, funcionales, bien definidas, útiles en un campo de aplicación concreto y con un problema de compatibilidad cuando se intenta definir una norma global que las acoja a todas ellas, en ocasiones definidas por un organismo de normalización, y en ocasiones sin ningún organismo oficial de normalización detrás.

• DIGEST

Dentro de las normas funcionales más conocidas actualmente tenemos el *Digital Geographic Information Exchange Standard* (DIGEST), desarrollada por Digital Geographic Información Working Group (DGIWG) en el que participan organismos cartográficos de países pertenecientes a la OTAN. Esta normativa tiene por objeto el permitir el compartir información geográfica en operaciones conjuntas de defensa de diferentes países que participan en DGIWG. Dentro de los documentos y especificaciones creados por este comité se puede reseñar el DIGEST 2.1: Part 4 que establece la definición de un catálogo de fenómenos que se ha utilizado en diferentes proyectos de ámbito internacional, como son EuroRegionalMap y

EuroGlobalMap, cartografía digital europea a escala 1:250.000 y 1:1.000.000 respectivamente, por ser un catálogo bien conocido, probado y de uso extendido.

DGIWG está realizando un gran esfuerzo para definir la normativa DIGEST como un conjunto de perfiles de la normativa ISO 19100, es decir como normas definidas por particularización de las normas definidas por ISO/TC211, mediante el procedimiento establecido que se denomina definición de perfiles.

- **S-57**

S - 57 es el estándar de transferencia electrónica de datos preparado por la Organización Hidrográfica Internacional (OHI) la comisión. La geodatabase modelo permite al usuario de importación Cartas de navegación electrónica (ENC) .000 archivos, que a su vez ofrece al usuario una herramienta de gran exactitud conjunto de datos vectoriales que es ideal como base de datos. Habilitar un valor añadido a la .000 ENC archivos dentro de un entorno de navegación no. Una hoja de estilo náutico genérico también está disponible para descargar. Tenga en cuenta que la simbología en el estilo náutico conjunto no corresponde a la S - 52, se utiliza la navegación orientada hacia otros símbolos.

- **GDF**

Geographic Data Files (GDF) es un formato de intercambio de datos definido para datos geográficos, que incluye reglas de captura detalladas, un catálogo de tipos de fenómeno, atributos y relaciones.

GDF se utiliza principalmente en sistemas de navegación para automóviles, pero es utilizable también para muchas otras aplicaciones de tráfico, como gestión de flotas, gestión de tráfico, etc.

Sin embargo al estar definido como un fichero plano de texto, es prácticamente inutilizable en aplicaciones a escalas grandes por su enorme volumen y normalmente es necesario convertirlo a otro formato más eficiente.

La mayoría de los proveedores de cartografía para navegación (NAVTEQ, TeleAtlas,...) soportan GDF.

GDF fue definido inicialmente por CEN/TC278 SWG3 (CEN GDF 3.0) y después ISO/TC 204 introdujo algunas modificaciones y se definió ISO FDG 4.0 como Norma Internacional.

16.2. Especificaciones OGC

También resulta interesante mencionar el trabajo de estandarización por consenso que ha realizado *Open Geospatial Consortium* (OGC), consorcio de más de 300 organizaciones industriales, agencias gubernamentales y universidades sin ánimo de lucro, cuyo objetivo es definir especificaciones de interoperabilidad por consenso, llevando la filosofía de los sistemas abiertos al mundo de los SIG. Por ese motivo, en un principio OGC respondía al nombre de *Open GIS Consortium*, pero al poner en práctica la interoperabilidad de los SIG mediante la definición de servicios web de interfaz estandarizada, apareció el concepto de IDE (Infraestructura de Datos Espaciales) como SIG distribuido y el mismo OGC modificó su nombre.

Las especificaciones de OGC se estructuran en dos grandes bloques:

1. Modelos Abstractos. Que proporcionan las bases conceptuales para el desarrollo de otras especificaciones OGC.

Topic 0 – Overview
Topic 1 - Feature Geometry
Topic 2 - Spatial Reference by Coordinates
Topic 3 - Locational Geometry Structures
Topic 4 - Stored Functions and Interpolation
Topic 5 – Features
Topic 6 - The Coverage Type
Topic 7 - Earth Imagery
Topic 8 – Relationships Between Features
Topic 10 - Feature Collections
Topic 11 - Metadata
Topic 12 - The OpenGIS® Service Architecture
Topic 13 – Catalog Services
Topic 14 - Semantics and Information Communities
Topic 15 - Image Exploitation Services
Topic 16 - Image Coordinate Transformation Services
Topic 17 – Location Based Mobile Services
Topic 18 – Geospatial Digital Rights Management Reference Model (GeoDRM RM)
Topic Domain 1 – Telecommunications Domain

Tabla 1. Modelos abstractos

2. Especificaciones para implementación. Están concebidas para una audiencia técnica con el nivel de detalle necesario para realizar una implementación.

OpenGIS Catalogue Service Implementation Specification 2.0.2.
OpenGIS Coordinate Transformation Service Implementation Specification 1.0
OpenGIS Filter Encoding Implementation Specification 1.1
OpenGIS Geographic Objects Implementation Specification 1.0.0
OpenGIS Geographic Markup Language (GML) Encoding Specification 3.1.1.
OpenGIS GML in JPEG 2000 for Geographical Imagery Encoding Specification 1.0.0
OpenGIS Grid Coverage Service Encoding Specification 1.0
OpenGIS Location Services Implementation Specification 1.1
OpenGIS Implementation Specification for G.I. Simple Feature Access 1 1.2.0
OpenGIS Implementation Specification for G.I. Simple Feature Access 2 1.2.0
OpenGIS Simple Features Implementation Specification for CORBA 1.0
OpenGIS Simple Features Implementation Specification for OLE/COM 1.0
OpenGIS Specification Styled Layer Descriptor Implementation Specification 1.0
OpenGIS Specification Symbology Encoding Implementation Specification 1.1.0
OpenGIS Specification Web Coverage service Implementation Specification 1.1.0
OpenGIS Specification Web Feature Service Implementation Specification 1.1
OpenGIS Specification Web Map Context Implementation Specification 1.1
OpenGIS Specification Web Map Service 1.3.0
OpenGIS Specification Web Service Common Implementation Specification 1.1.0

Tabla 2. Especificaciones de implementación.

16.3. Normativa Oficial. Organismos de Normalización en Información Geográfica

El verdadero protagonismo en este campo lo tienen los organismos nacionales, europeos e internacionales de normalización que han sabido aglutinar todo el esfuerzo realizado por las diferentes organizaciones en la realización de las llamadas normas funcionales, casi siempre sectoriales, y establecer una serie de normas más generales que las acojan bajo su paraguas.

16.3.1. ISO/TC 211. Geographic Information. Familia de Normas ISO 19100

Este comité internacional comenzó a trabajar en noviembre de 1994 con el objetivo de establecer normativa de referencia en el campo de la información geográfica digital, pensada tanto para la transferencia de datos y el mundo de los SIG aislados, como para los servicios interoperables y el universo de las IDE o SIG distribuidos.

Como resultado de este trabajo, apareció la familia ISO 19100, un conjunto de normas relacionadas con objetos o fenómenos que están directa o indirectamente asociados con una localización relativa a la Tierra. La normativa trata sobre los métodos, herramientas y servicios para la gestión de datos, adquisición, procesamiento, análisis, acceso, presentación y

transferencia de información geográfica en formato digital entre diferentes usuarios, sistemas y localizaciones. Este trabajo se realiza haciendo referencia, siempre que sea posible y oportuno, a la normativa ya existente en materia de Tecnologías de la Información y las Comunicaciones.

Como resultado se tiene una familia de más de 50 proyectos normativos en cuya elaboración han estado involucrados 29 países como miembro de pleno derecho, 30 países observadores, del orden de otras 30 organizaciones internacionales de máxima relevancia y otros 18 comités de normalización en ISO mediante los oportunos acuerdos de colaboración..

Actualmente, el TC211 tiene cinco grupos de trabajo, que dependen de la presidencia del comité:

- Grupo de Trabajo 4. Servicios Geoespaciales.
- Grupo de Trabajo 6. Imágenes.
- Grupo de Trabajo 7. Comunidades de la Información.
- Grupo de Trabajo 8. Servicios Básicos de Localización (Actualmente sin actividad).
- Grupo de Trabajo 9. Gestión de la Información.

También existen ocho grupos especiales, cuyo objetivo es dar soporte al trabajo de la presidencia del comité.

- Advisory Group on Strategy
- HMMG – Harmonized Model Maintenance Group
- Advisory Group on Outreach
- JAG – ISO/TC211 / OGC Joint Advisory Group
- TMG – Terminology Maintenance Group
- PMG – Programme Maintenance Group
- Task Force ISO/TC 204 and ISO/TC211 projects
- Task Force JTC 1/SC 24 and ISO/TC 211 projects.

Todos ellos son de suma importancia para mantener la coordinación y coherencia en el desarrollo de las actividades del comité al tratarse de un conjunto de 50 normas que se han elaborado esencialmente en paralelo.

Conviene destacar la actividad que desarrolla el Joint Advisory Group. Éste grupo está codirigido por miembros del TC211 y OGC con el objetivo coordinar los esfuerzos normativos de ambos organismos. Como resultado se consigue establecer una única normativa de referencia

en información geográfica digital, ISO 19100, recogiendo los fundamentos de las especificaciones OGC y asegurando la coordinación entre ambos ámbitos de estandarización.

Actualmente, existen treinta y tres documentos normativos publicados y casi otros 20 proyectos normativos en marcha. Continuamente se están realizando estudios sobre tareas a normalizar, que en el futuro supondrán nuevos proyectos normativos que terminarán dando, en muchos casos Normas Internacionales, y en otros casos Especificaciones Técnicas e Informes técnicos.

A fecha de Junio de 2007 el TC 211 ha desarrollado o está desarrollando los siguientes trabajos:

- CD: Committee draft (Borrador de Comité), es un borrador producido por un Grupo de Trabajo que se distribuye dentro del Comité Técnico (TC) a los países miembros del TC para que elaboran comentarios en sucesivas consultas hasta que el TC decide aprobarlo como DIS.
- DIS: Draft International Standard (Borrador de Norma Internacional), una vez que se ha alcanzado un consenso en el TC, se distribuye a todos los países miembros de ISO para votación y comentarios durante un período de 5 meses.
- FDIS: Final Draft International Standard (Borrador Final de Norma Internacional), una vez que se ha aprobado un DIS, se considera que es un documento estable, aceptado por la comunidad y utilizable para las primeras implementaciones. Sólo se espera que sufra cambios editoriales antes de ser aprobado como IS en un proceso que incluye una votación formal final y dura pocos meses.
- IS: International Standard (Norma Internacional), es la norma definitiva y para modificarla o actualizarla es necesario iniciar de nuevo todo el proceso con un nuevo documento de trabajo.

Hay otros dos tipos de documentos finales que no tienen carácter de norma, sino sólo informativo:

TS: Technical Specification

TR: Technical Report

Normas o especificaciones de carácter general	
DIS 6709 rev	Representación de latitud, longitud y altura.
IS 19101	Modelo de Referencia
PDS 19101-2	Modelo de Referencia para imágenes
IS19103	Lenguaje Conceptual
IS 19104	Terminología
IS 19105	Conformidad y Pruebas.
IS 19106	Perfiles

Normas o especificaciones de carácter general	
IS19107 Modelo espacial	IS 19108 Modelo temporal
IS19109 Reglas para aplicación	IS 19110 Met. de catalogación de Fenómenos.
IS19111 SR con coordenadas	IS 19112 SR con id. Geográficos
IS 19113 Principios de Calidad	IS 19114 Evaluación de la Calidad
IS 19115 Metadatos	IS 19118 Codificación
FDIS 19136 GML	IS 19137 Núcleo del Modelo Espacial.
TS 19138 Medidas de Calidad de los Datos.	DIS 19141 Schema for moving features

Normas de Servicios	
IS 19116	Serv. de posicionamiento
IS 19117	Representación
IS 19119	Servicios para coberturas
IS 19125-1	“Simple Feature Access” (SFA)
IS 19125-2	“SFA – SQL Option”
IS 19128	“Web Map Server Interface”
IS 19132	Posibles servicios de localización
IS 19133	Servicios de localización para seguimiento y navegación
IS 19134	“Multimodal location based services for routing and navigation”
CD 19142	“Web Feature Service”
CD 19143	“Filter encoding”
19147	“Location based services -- Transfer Nodes”
19148	“Location based services -- Linear Referencing System”

Normas de Formatos Ráster y Malla	
CD 19115-2	Metadatos. Part 2. Extensión para imágenes y malla.
TR 19121	Datos imagen y malla
IS 19123	Modelo para coberturas
19124	Componentes de datos imagen y malla
CD 19129	“Imagery and coverage data”

Normas Complementarias	
CD 19115-2	Metadatos. Part 2. Extensión para imágenes y malla.
TR 19120	Normas Funcionales
TR 19122	Cualif.y certif. del personal
CD 19126	Diccionario de datos FACC
TS 19127	Parámetros y códigos geodésicos
IS 19131	“Data Product Specifications”
IS 19135	Registration of GI items
IS 19139	Metadata - Implementation specifications
DIS 19141	“Schema for Moving Features!”
CD 19144-1	Sistemas de Clasificación – Part. 1: Estructura del sistema de clasificación.
CD 19144-2	Sistemas de Clasificación – Part. 2: “Land Cover Classification System” (LCCS)
19145	“Registry of representations of geographic point locations”
19146	“Cross-domain vocabularios”
19149	“Rights expression language for geographic information” - GeoREL
19150	Ontologías

Una de las principales ventajas de los trabajos del ISO/TC211, es que representan el resultado de las contribuciones de todos los actores y organismos que intervienen en el campo de la información geográfica: empresas productoras de *software*, productores de datos, agencias espaciales, organismos de normalización, asociaciones profesionales, universidades, sector público, etc. Y todos coinciden en la necesidad de tener una sola normativa de referencia en el sector, ISO 19100.

16.3.2. CEN/TC 287. Geographic Information

Este comité europeo de normalización comenzó su andadura antes que ISO/TC211, en 1991. En su primera etapa, hasta 1999, elaboró diez documentos normativos.

Una vez comenzados los trabajos de ISO/TC211, con el objeto de no existir duplicidad de responsabilidades en cuanto a normalización en el sector de la Información Geográfica, el CEN/TC 287 paralizó su actividad y, definió sus resultados como normativa experimental, no vinculante, con el fin de que no entrasen en colisión con la normativa que en el futuro produjera el ISO/TC211. De esa manera, además se abría la posibilidad de que ISO/TC211 tomará los documentos finales de CEN/TC287 como punto de partida para sus trabajos.

En el año 2004, el CEN/TC287 vuelve a activarse con el objetivo de aportar un marco normativo a la Directiva INSPIRE, cuyo objetivo es implantar una Infraestructura de Datos Espaciales (IDE) en Europa basada en las IDE nacionales, particularizando y aprovechando tanto la familia de normas ISO 19100 como las especificaciones OGC,

CEN/TC 287 en su nueva etapa, se basa en los siguientes principios:

1. Adopción de la serie ISO 19100 como estándares Europeos.
2. Desarrollo y la aceptación de nuevos estándares y perfiles en cooperación con ISO/TC 211, incluyendo unas reglas de implementación.
3. Facilitar la interoperabilidad con iniciativas de estándares relacionados mediante la armonización.
4. Promoción del uso de los estándares de información geográfica

En la actualidad, se han adoptado veinte normas de la familia ISO 19100 como normativa europea y existen otros ocho proyectos normativos para adopción de más normativa ISO.

También se ha producido el documento CEN/TR 15449:2006 “Geographic information” – “Standards, specifications, technical reports and guidelines, required to implement Spatial Data

Infraestructure”, que como su nombre indica es un informe que tiene como objetivo recoger información necesaria para la implementación de una Infraestructura de Datos Espacial.

16.3.3. Comité Técnico de Normalización 148 de AENOR, Información Geográfica. (AEN/CTN 148)

A nivel español, la colaboración con los organismos europeo (CEN/TC287) e internacional (ISO/TC211) de normalización se canaliza a través AEN/CTN 148.

Del trabajo de este comité, en 1997 vio la luz la normativa experimental MIGRA VI (Mecanismo de Intercambio de Intercambio de Información Geográfica Regional formado por Agregación) que suponía un primer intento normativo para facilitar el intercambio de información geográfica digital entre los diferentes agentes.

Al adoptarse ISO 19100 como normativa europea por el CEN/TC287, automáticamente se establece un proceso por el que dicha normativa se convierte en normativa española, teniendo que derogar la normativa nacional que en ese sentido estuviera vigente. Ese fue el caso de MIGRA VI.

Actualmente, el trabajo del comité se centra en tareas sus colaboraciones con los organismos europeos e internacionales de normalización, trabajos de traducción y adopción de normativa europea, difusión de la misma y, últimamente se ha comenzado a abordar un Proyecto Normativo de Control de Calidad Posicional, que proporcionará un documento de referencia dentro de nuestro sector en el ámbito nacional.

16.4. El Open Geospatial Consortium

El *Open Geospatial Consortium, Inc.* (OGC) es una organización de estandarización no lucrativa, internacional, basada en el consenso voluntario, que está conduciendo el desarrollo de estándares para los servicios geoespaciales y basados en la localización. A través de su programa de trabajo, dirigido a través de mecanismos de consenso y votación entre sus miembros, OGC trabaja con el gobierno, la industria privada, y la academia para crear los modelos, arquitecturas e interfaces de programación, para que los SIG sean abiertos e interoperen entre sí.

Los documentos consensuados se establecen como especificaciones OGC y están a pública disposición, de modo libre y sin coste alguno en <http://www.opengeospatial.org>.

1. Visión

La realización práctica completa de las ventajas sociales, económicas y científicas que resultan de la integración de recursos electrónicos de localización en todo tipo de procesos comerciales e institucionales por todo el mundo.

2. Misión

Servir como foro global para la colaboración de desarrolladores y usuarios de productos de IG y servicios de datos geográficos, y avanzar en el desarrollo de estándares internacionales para la interoperabilidad geoespacial de sistemas.

3 Objetivos estratégicos

- Proporcionar a la comunidad, de modo libre, abierto y gratuito los estándares disponibles, beneficios tangibles a sus miembros, y ventajas mensurables a los usuarios.
- Liderar a nivel mundial la creación y el establecimiento de estándares que permitan que los contenidos y los servicios geoespaciales se integren sin costuras, es decir se mimeticen, en los procesos de tratamiento de la información, ya sean públicos o privados, a través de la web espacial y la informática de empresa.
- Facilitar la adopción en todo el mundo empresarial de arquitecturas espaciales de referencia abiertas.
- Avanzar estándares para ayudar la formación de mercados nuevos e innovadores y nuevas aplicaciones de las tecnologías geoespaciales.
- Acelerar la asimilación en el mercado de las investigaciones sobre interoperabilidad a través de procesos colaborativos de participación en consorcios.

16.4.1. Especificaciones de interoperabilidad

Las siguientes especificaciones definidas por *Open Geospatial Consortium*, definen cómo deben ser servicios y componentes accesibles en La Red mediante un simple *browser*, para que sean interoperables, es decir para que puedan utilizar datos y otros servicios distribuidos en La Red y puedan ser encadenados entre sí. Los servicios permiten invocar dos operaciones básicas: GET CAPABILITIES, que devuelve las características esenciales del servicio.

- Web Map Service (WMS), o servicio de publicación de mapas, permite visualizar datos vectoriales y ráster en La Red con un simple *browser*, efectuar operaciones sencillas de visualización, como apagar y encender capas, hacer zoom, vuelo panorámico, etcétera y superponer datos independientemente del servidor físico en el que se encuentren, su formato y su sistema de referencia y proyección cartográfica.
- Styled Layer Descriptor (SLD), o descripción de estilo de capas, dedicada a permitir la definición temporal o permanente de la simbolización con la que se desean visualizar los objetos incluidos en un conjunto de datos determinado.
- Web Feature Service (WFS), o servicio de publicación de objetos, que ofrece la posibilidad de acceder en la Red a los objetos individuales contenidos en un conjunto de datos y recuperar todos sus atributos y características para poder efectuar análisis complejos.
- Web Coverage Service (WCS), o servicio de publicación de coberturas, que permite publicar en Internet coberturas ráster de imágenes de satélite, ortofotos, fotografías aéreas, ficheros ráster y, en general cualquier conjunto de datos que tenga esta estructura de la información, para su visualización, análisis y consulta.
- Web Coordinate Transformation Service (WCTS), o servicio de transformación de coordenadas, que oferte la posibilidad en la Red de transformar un fichero de datos geográficos de una proyección cartográfica a otra o de cambiar su sistema de referencia.
- Geographic Markup Language (GML), es una extensión de XML (*Extensible Markup Language*) para datos geográficos, que contempla las primitivas geométricas (puntos, polígonos, superficies, líneas,...), las topológicas (nodo, borde, cara,...) y su estructuración. Se está imponiendo como formato de intercambio de datos geográficos e, incluso, en algunos sistemas se utiliza como formato interno de almacenamiento.
- Catalog Service Web (CSW), o servicio de publicación de catálogos, especifica cómo hacer accesibles vía red catálogos de recursos, es decir catálogos de datos y servicios, para poder realizar búsquedas distribuidas que consulten un conjunto de catálogos en un conjunto de servidores. Para ello, es necesario que el catálogo que se consulta tenga catalogados el resto de catálogos en los que se va a desencadenar una búsqueda en cascada.
- Gazetteer (GAZ), o servicio de Nomenclátor, que permite realizar búsquedas de topónimos con varios criterios (nombre exacto, nombre incluido, empezando por, en una zona, sólo topónimos de una clase,...) y devuelve su situación geográfica.
- Geoparser, este servicio en Red consiste en procesar un documento textual en formato digital e identificar y marcar las palabras que aluden a una localización geográfica, en general nombres geográficos, por comparación con un nomenclátor. Posteriormente se pueden

realizar enlaces permanentes, crear elementos de hipertexto, hacer búsquedas bibliográficas y geográficas, etcétera.

- Geolinking Data Access Service (GDAS), o servicio de acceso a datos geolinkados, especificación en elaboración, pensada para ofrecer una interfaz normalizada de acceso a tablas de datos alfanuméricos (estadísticos, temáticos), que incluyan una columna con un identificador geográfico, una etiqueta de georreferenciación, que permita establecer una relación con un fichero de datos geográficos que haga posible su visualización. Por ejemplo una tabla de datos estadísticos relativos a municipios que incluya el código INE o los nombres de los municipios.
- GeoLinking Service (GLS), o servicio de geolinkaje, especificación todavía en preparación, concebida para poder acceder a una tabla de datos geolinkados, mediante un GDAS, seleccionar qué atributo se desea visualizar, acceder a un fichero de datos geográficos que contenga las áreas georreferenciadas la tabla en cuestión, acceder o definir interactivamente un fichero de leyenda dónde se especifique cómo se desea visualizar cada valor de un atributo, y finalmente visualizar en pantalla un mapa temático del atributo seleccionado.
- Sensor Web Enablement (SWE), o servicio de acceso a sensores, pensado para poder acceder en tiempo real a datos tomados por sensores tales como estaciones de aforo, estaciones pluviométricas, puntos de medida del tráfico por carretera, sensores meteorológicos, *webcam*, etc.
- Web Pricing and Ordering Service (WPOS), o servicio de comercialización y pedido electrónico, en elaboración, planteada para cubrir todos los aspectos de comercio electrónico (*e-commerce*) relacionados con datos espaciales digitales: precios, licencias de uso, pedidos, entregas, facturación, etcétera.
- Web Terrain Server (WTS), o servidor de vistas del terreno, cuya finalidad es producir vistas perspectivas de coberturas de datos geográficos tridimensionales.
- Web Processing Service (WPS), o servicio de procesamiento en la web, que permite definir utilidades de geoprocésamiento utilizando entradas de datos estándar OGC (GML, GDAS) y ofreciendo resultados también de forma estandarizada.
- Integrated Client for Multiple Services (IntClient), que permite almacenar de modo permanente un encadenamiento concreto de servicios con una serie de parámetros y selecciones específicas. Esta especificación supone disponer de un lenguaje de encadenamiento de servicios y supone el regreso del concepto de lenguaje de macros, en La Red, y del concepto de *Toolbox*.

Esta última especificación abre enormes posibilidades para integrar servicios disponibles en la red en la lógica de nuestras propias aplicaciones, que si están definidas como servicios públicos,

podrán ser a su vez integradas en otras aplicaciones hasta constituir una suerte de palimpsesto digital.

Las especificaciones subrayadas están ya disponibles en www.opengeospatial.org.

Del resto, también hay disponibles borradores de trabajo.

16.5. Conclusiones

La normalización supone la madurez de nuestro sector y aporta los siguientes beneficios:

- Aumenta la disponibilidad, accesibilidad, integración, facilita la distribución de información geográfica y posibilita la interoperabilidad de sistemas geográficos.
- Contribuye a aproximaciones unificadas para resolver problemas globales.
- Simplifica el establecimiento de Infraestructuras de Datos Espaciales.
- Facilitan la comprensión y el uso de la información geográfica.

En conclusión, aporta ventajas importantes para el sector.

Por otra parte, debido a la alta representación de todos los agentes que intervienen en los procesos relacionados con la información geográfica en el comité ISO/TC211, y a los acuerdos con la industria (*Joint Advisory Group* con OGC), ISO 19100 se ha convertido en la referencia normativa en el sector.

Bibliografía

- [1] López Romero, Emilio y otros, *Normas y estándares en el entorno de la IDEE*, Congreso Topcart 2004. Madrid
- [2] <http://www.cartesia.org/geodoc/topcart2004/conferencias/58.pdf>
- [3] Más Mayoral, Sebastián. *Experiencias y Proyectos del IGN/CNIG en el campo SIG. II Conferencia sobre Sistemas de Información Territorial*. Pamplona, 2002..
- [4] <http://www.cfnavarra.es/territorial2000/PONENCIAS/MASS.PDF>
- [5] ISO (International Standard Organization)
- [6] <http://www.iso.org>
- [7] ISO/TC 211
- [8] <http://www.isotc211.org>
- [9] Open GeoSpatial Consortium
- [10] <http://www.opengeospatial.org>
- [11] European Committee for Standardization (CEN)

[12] <http://www.cen.eu>

[13] AENOR

<http://www.aenor.es>

[14] Programa de Trabajo de ISO/TC 211

http://www.isotc211.org/pow_all.htm

[15] Procedimiento de elaboración de normas ISO

[16] http://www.iso.org/iso/standards_development/processes_and_procedures/stages_description.htm

Tema 17. Infraestructuras de Datos Espaciales. Definición y componentes. Geoportales. Arquitectura de una IDE. La Directiva INSPIRE. Iniciativas y proyectos IDE.

17.1. Infraestructuras de Datos Espaciales

Una Infraestructura de Datos Espaciales (IDE) es un marco formado por políticas, disposiciones institucionales, tecnologías, datos y personal que hacen posible el compartir y utilizar la Información Geográfica (IG) de modo eficaz, cómodo y sencillo.

Para compartir IG de manera óptima es imprescindible disponer de una red tan eficaz como Internet, por lo que la idea de tener varios SIG interoperando en la red está implícito en la definición.

Por otro lado, una IDE es más que un mero sistema informatizado integrado por hardware, software, datos y componentes de comunicaciones; es un concepto mucho más amplio, una estructura básica que incluye además personal y toda una organización: normas, acuerdos, políticas, estándares, planes, tecnologías,...

Una manera de explicarlo rápidamente es que una IDE es un SIG implementado sobre la Red, sobre Internet. Lo que implica un conjunto de estándares para que sea posible disponer de un SIG distribuido, una serie de políticas y disposiciones para coordinar a los actores que la integran, y una organización, claro.

La idea de una infraestructura de datos como un mecanismo para proporcionar acceso más efectivo a la IG apareció por primera vez en Canadá a principios de los años 80. En aquel momento había una gran presión sobre las instituciones cartográficas canadienses para que justificasen las enormes inversiones dedicadas a la generación de recursos geográficos y se produjese su explotación intensiva por la sociedad para amortizar el esfuerzo realizado. El concepto de infraestructura apareció como una idea atractiva que podría potenciar la aparición de un gran número de aplicaciones facilitando el acceso a los datos.

Sin embargo, el impulso definitivo se debió a la orden ejecutiva 1994/12906 del presidente Bill Clinton de USA para el establecimiento de la *National Spatial Data Infrastructure* de ese país.

Ese mismo año 1994 se fundó el *Open GIS Consortium*, con la intención de aplicar la filosofía de los sistemas abiertos a los SIG, pero pronto se vio que los SIG abiertos se habían convertido en algo radicalmente distinto, las IDE, y OGC cambió de su nombre por el de *Open GeoSpatial Consortium*.

Se trata de un consorcio de empresas privadas, organismos públicos y universidades que define por consenso y mediante votaciones electrónicas las especificaciones de interoperabilidad,

dónde se definen los modelos, protocolos e interfaces que hacen posible el que los sistemas geográficos interoperen entre sí.

Si nos restringimos al sistema informático que soporta una IDE, tendremos un conjunto de recursos integrados (catálogos, servidores, programas, datos, aplicaciones, páginas web, etc.) dedicados a gestionar Información Geográfica (mapas, ortofotos, imágenes de satélite, topónimos, etc), disponibles en Internet, que cumplen una serie de condiciones de interoperabilidad (normas, especificaciones, protocolos, interfaces,...) que permiten que un usuario, utilizando un simple navegador, pueda combinarlos según sus necesidades.

Un IDE ofrece la posibilidad de acceder a un conjunto de funcionalidades, llamadas servicios, desde cualquier ordenador con acceso a Internet y con un simple navegador:

- Buscar en un Catálogo qué datos geográficos de una zona determinada, a una escala concreta, y cumpliendo una serie de condiciones, hay disponibles, independientemente del servidor en el que estén almacenados.
- Visualizar los datos geográficos una vez encontrados, y efectuar operaciones sencillas de visualización (zoom, desplazamientos, ocultar capas, etc.).
- Superponer datos geográficos en pantalla, independientemente del servidor en el que se hallen, del formato en que se encuentren, de su escala y del sistema de coordenadas que tengan. La única condición es que se cumplan las condiciones de interoperabilidad oportunas.
- Hacer búsquedas por el nombre de un fenómeno geográfico en un catálogo de topónimos o nombres geográficos o Nomenclátor, que devuelve como resultado la ubicación geográfica del fenómeno buscado.
- Consultar qué atributos tiene un objeto que se está visualizando en pantalla.
- Descargar datos geográficos si el servidor que los almacena así lo permite.

Éstas son digamos las posibilidades básicas de una IDE, pero hay otras más complejas, que todavía es difícil encontrar implementadas, pero técnicamente realizables y esperamos que factibles en un futuro próximo como: la búsqueda no sólo de datos sino de todo tipo de recursos (como ya se ha dicho programas, aplicaciones, utilidades, datos de sensores, otros catálogos, libros, publicaciones, mapas en papel y un largo etcétera); el análisis de los datos que se están visualizando, por ejemplo para hallar cuál es el camino mínimo por carretera entre dos puntos o cómo se difunde en una red hidrográfica un contaminante (análisis de redes), o para averiguar qué área cumple una serie de condiciones, como pertenecer a un Municipio de menos de 50 000 habitantes y estar cubierto de coníferas (análisis de superficies); el seguimiento continuo de los datos recogidos por un conjunto de sensores (estaciones meteorológicas, estaciones de aforo en ríos, puntos de medida del tráfico, de contaminación, de ruido) distribuidos en un territorio.

Las posibilidades que ofrece el poder superponer, combinar, consultar y analizar libremente en pantalla datos geográficos, ortofotos, datos de catastro, modelos de digitales del terreno, imágenes de satélite, tablas de atributos alfanuméricos, etcétera son enormes (información al público, educación, mapas a la carta, gestión de redes de transporte, simulación de catástrofes, evaluación de impacto medioambiental, planificación) y muchas tareas y proyectos que hasta ahora sólo se podían llevar a cabo con un Sistema de Información Geográfica costoso y complejo, a partir de ahora y cada vez más, se podrán realizar explotando las capacidades de una IDE, de una manera mucho más sencilla, cómoda y barata, poniendo los datos y la información resultante al alcance de los ciudadanos.

Algunos de los principios y conceptos básicos de las IDE son:

- No centralizar ni los datos ni los metadatos, como se tendía a hacer en el SIG monolítico y clásico.
- El concepto central alrededor del que se concibe el sistema ya no son los datos, como en el SIG, sino los servicios.
- Un servicio es una funcionalidad que ofrece un software servidor en la red a un software cliente.
- Las IDE se basan en la arquitectura cliente/servidor, en la que se distribuye el proceso, pero no los datos. Es importante distinguir entre la parte cliente y la parte servidora de cada servicio.
- La interoperabilidad es clave, entendida como la capacidad para comunicar, ejecutar o transferir datos entre varias unidades funcionales de tal manera que el usuario no necesite tener ningún conocimiento o muy poco sobre las características particulares de tales unidades.
- Los recursos se comparten en la red para formar un repositorio común del que luego todos los participantes se benefician.

Por eso decimos que en el campo de la Información Geográfica está surgiendo un nuevo paradigma, llamado IDE, una nueva forma de trabajar que ha trascendido al mapa y al SIG, una nueva forma a la que es necesario adaptarse, que nos trae un mundo nuevo de posibilidades por explotar.

17.2. Definición y componentes

Como decíamos, una IDE no es nada más y nada menos que un SIG implementado sobre Internet, la Red con mayúscula y por antonomasia, con todo lo que ello conlleva y significa. Dicho de otro modo, una IDE es un SIG globalizado. Las distintas funcionalidades que integraban la caja de herramientas (*toolbox*) de un SIG adquieren entidad propia y se constituyen en servicios electrónicos, de interfaz estándar y descripción pública, y en consecuencia susceptibles de ser encadenados y combinados.

No se trata, por lo tanto, de que el usuario pueda realizar una mera conexión a un SIG a través de Internet para explotar en remoto el mismo sistema que puede tener disponible en una estación de trabajo. Más bien se trata de que el usuario pueda mediante un simple navegador, un cliente ligero, buscar qué datos geográficos hay disponibles, seleccionar cuáles son de su interés, verlos, superponerlos, consultarlos y trabajar con ellos gracias a una panoplia de utilidades creciente y asombrosa:

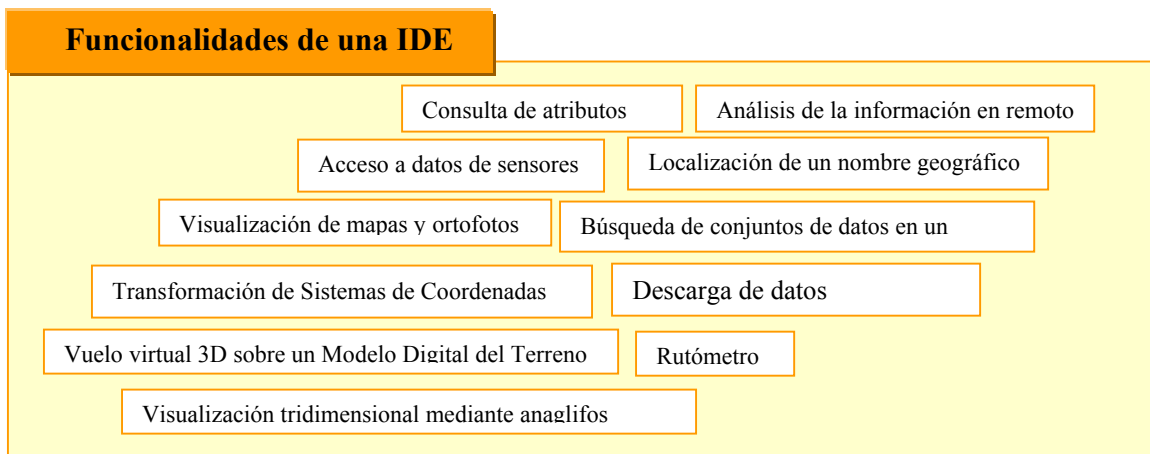


Figura 1. Ejemplos de utilidades en una IDE

Todo esto de manera transparente y sin preocuparse de en qué nodo reside cada componente ni de ninguna de sus características: formatos, sistemas de referencia de las coordenadas, etcétera. Parafraseando a Nicholas Negroponte, creador del MIT Media Laboratory, “no es que la Red interconecte Sistemas entre sí, es que la Red es el Sistema”.

Esta nueva forma de trabajar abre un considerable abanico de posibilidades, tanto técnicas como organizativas:

- Desde un punto de vista técnico, no sólo es posible utilizar los servicios de una IDE en remoto utilizando un navegador (cliente ligero), sino que desde una aplicación SIG clásica

(cliente pesado) que corre en el ordenador del usuario también se pueden invocar y utilizar servicios OGC, mezclando funcionalidad y comandos públicos y privados, y fusionando visualmente en pantalla también datos públicos, de una IDE, y datos privados, almacenados en el disco local. Incluso, si los servicios públicos disponibles no ofrecen toda la funcionalidad que se necesita, es posible en ocasiones descargarse los datos (cliente muy pesado) y trabajar a nivel local. Actualmente los principales productos clásicos en el sector SIG han incluido la posibilidad de actuar como clientes de servicios OGC. De esta manera, se alcanza la integración y sinergia del mundo IDE con el mundo SIG, abriendo la posibilidad de balancear la carga entre la parte cliente y la parte servidora en función de cada situación, para obtener la solución más eficaz.

- Desde un punto de vista organizativo, también han aparecido nuevas posibilidades de gran interés: establecimiento de *workflow* internos en una organización basados en servicios IDE, con el consiguiente abaratamiento de recursos; explotación mixta de recursos geográficos públicos y privados; utilización de la publicación de mapas como canal privilegiado de comunicación basado en datos geográficos hacia el público, los usuarios o los clientes, como soporte de un mensaje determinado; creación de *mashups* sobre un fondo cartográfico y basados en estándares; integración de componentes y servicios existentes, como nueva actividad profesional; aplicación de servicios OGC en sectores a los que la tecnología SIG no ha llegado por los costes que arrastra, como la enseñanza de geografía; etcétera.

Las especificaciones definidas por el mencionado *Open Geospatial Consortium* (www.opengeospatial.org), definen cómo deben ser los servicios y componentes accesibles en La Red mediante un simple *browser*, para que sean interoperables, es decir para que puedan encadenarse entre sí.

La mayoría de los servicios permiten al menos invocar dos operaciones básicas y generales:

- *GetCapabilities*, que devuelve una descripción del servicio y que nos sirve para saber si un servicio está funcionando.
- *Get <<Map, Feature, Coverage.>>*, que nos devuelve información (mapas, fenómenos, coberturas, etc.) del servicio

Los principales servicios, definidos mediante especificaciones OGC son:

- *Web Map Service (WMS)*, o servicio de publicación de mapas, permite visualizar datos vectoriales y ráster en La Red con un simple *browser*, efectuar operaciones sencillas de visualización, como apagar y encender capas, hacer zoom, vuelo panorámico, etcétera y

superponer datos independientemente del servidor físico en el que se encuentren, su formato y su sistema de referencia y proyección cartográfica.

- Catalog Service Web (CSW), o servicio de publicación de catálogos, especifica cómo hacer accesibles vía red catálogos de recursos, es decir catálogos de datos y servicios, para poder realizar búsquedas distribuidas que consulten un conjunto de catálogos en un conjunto de servidores. Para ello, es necesario que el catálogo que se consulta tenga catalogados el resto de catálogos en los que se va a desencadenar una búsqueda en cascada.
- Gazetteer (GAZ), o servicio de Nomenclátor, que permite realizar búsquedas de topónimos con varios criterios (nombre exacto, nombre incluido, empezando por, en una zona, sólo topónimos de una clase,...) y devuelve su situación geográfica. Es un caso particular de WFS (WFS-G o WFS-MNE).
- Web Feature Service (WFS), o servicio de publicación de objetos, que ofrece la posibilidad de acceder en la Red a los objetos individuales contenidos en un conjunto de datos y recuperar todos sus atributos y características para poder efectuar análisis complejos.
- Web Coverage Service (WCS), o servicio de publicación de coberturas, que permite publicar en Internet coberturas ráster de imágenes de satélite, ortofotos, fotografías aéreas, ficheros ráster y, en general cualquier conjunto de datos que tenga esta estructura de la información, para su visualización, análisis y consulta.
- Filter Encoding (Filter), o filtros, que define cómo expresar en XML condiciones de filtrado para seleccionar un subconjunto de los *features* accedidos. Está pensado para actuar como un SELECT sobre el resultado de un WFS.
- Web Map Context (WMC), o contexto, que permite almacenar todos los parámetros y características de una situación activa en un geoportal OGC para reproducirla más tarde.
- Styled Layer Descriptor (SLD), o descripción de estilo de capas, dedicada a permitir la definición temporal o permanente de la simbolización con la que se desean visualizar datos vía WMS.

- Web Coordinate Transformation Service (WCTS), o servicio de transformación de coordenadas, que oferte la posibilidad en la Red de transformar un fichero de datos geográficos de una proyección cartográfica a otra o de cambiar su sistema de referencia.
- Geographic Markup Language (GML), es una extensión de XML (*Extensible Markup Language*) para datos geográficos, que contempla las primitivas geométricas (puntos, polígonos, superficies, líneas,...), las topológicas (nodo, borde, cara,...) y su estructuración. Se está imponiendo como formato de intercambio de datos geográficos e, incluso, en algunos sistemas se utiliza como formato interno de almacenamiento.
- Web Processing Service (WPS), o servicio de procesamiento en la web, que permite definir utilidades de geoprocésamiento con interfaces bien descritas de entrada y de salida (ficheros GML, XML, o imágenes JPEG, PNG,...GDAS).
- Geolinking Data Access Service (GDAS), o servicio de acceso a datos geolinkados, especificación en elaboración, pensada para ofrecer una interfaz normalizada de acceso a tablas de datos alfanuméricos (estadísticos, temáticos), que incluyan una columna con un identificador geográfico, una etiqueta de georreferenciación, que permita establecer una relación con un fichero de datos geográficos que haga posible su visualización. Por ejemplo una tabla de datos estadísticos relativos a municipios que incluya el código INE o los nombres de los municipios.
- GeoLinking Service (GLS), o servicio de geolinkaje, especificación todavía en preparación, concebida para poder acceder a una tabla de datos geolinkados, mediante un GDAS, seleccionar qué atributo se desea visualizar, acceder a un fichero de datos geográficos que contenga las áreas georreferenciadas la tabla en cuestión, acceder o definir interactivamente un fichero de leyenda dónde se especifique cómo se desea visualizar cada valor de un atributo, y finalmente visualizar en pantalla un mapa temático del atributo seleccionado.
- Sensor Web Enablement (SWE), o servicio de acceso a sensores, pensado para poder acceder en tiempo real a datos tomados por sensores tales como estaciones de aforo, estaciones pluviométricas, puntos de medida del tráfico por carretera, sensores meteorológicos, *webcam*,...

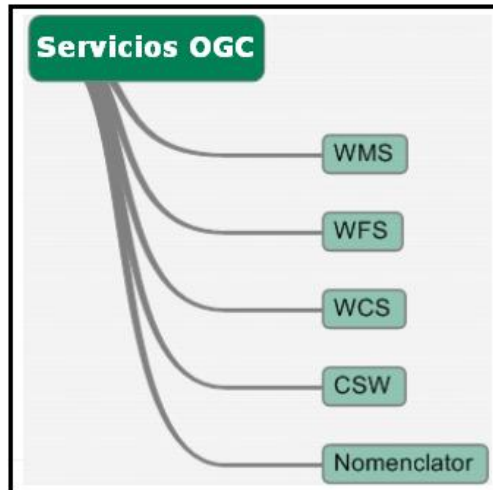


Figura 2. Algunos servicios OGC

17.3. Geoportales

Los geoportales son componentes fundamentales de las IDE, son la puerta de entrada y la cara visible del sistema distribuido, y es dónde se define la interfaz de usuario.

17.3.1. Definición

Un geoportal es un tipo de portal web utilizado para buscar, acceder y utilizar Información Geográfica a través de servicios en remoto, como pueden ser los mencionados servicios de visualización, edición, análisis, etcétera.

Todo ello a través de Internet y siguiendo la filosofía de los sistemas distribuidos.

Por lo tanto constituyen el punto de entrada a una IDE, y esencialmente están basados en un conjunto de aplicaciones cliente que invocan y encadenan los correspondientes servicios.

Los productores de IG, como organismos de la Administración y proveedores privados, utilizan los geoportales para publicar sus datos, sus servicios y sus metadatos (datos acerca de los datos) que describen los conjuntos de datos que producen, y están accesibles a través de Catálogos de metadatos que permiten la búsqueda y selección de datos.

Los usuarios de la IG, ya sean profesionales o ciudadanos sin especialización, hacen uso de los Geoportales para buscar y acceder a los datos geográficos que necesitan.

Por lo tanto los geoportales juegan un papel esencial y de importancia creciente en el compartir la IG y pueden evitar duplicaciones de esfuerzo, inconsistencias entre datos, retrasos, situaciones confusas y derroche de recursos.

17.3.2. Ejemplos

El portal de la *National Spatial Data Infrastructure* (NSDI) de USA, abierto en 1994, se considera la primera implementación del concepto de geoportal, dentro de una iniciativa coordinada por el FGDC (*Federal Geospatial Data Committee*). Utilizaba un gran número de catálogos independientes, que se podían consultar mediante un cliente y una interfaz comunes.

En 2003, apareció el *Geospatial One-Stop* (GOS) geoportal como parte de una iniciativa gubernamental. A diferencia del sistema anterior, se basó en un catálogo de metadatos centralizado, con una arquitectura que conectaba a cada usuario con cada proveedor de datos mediante servicios estándar. El usuario de GOS puede utilizar un simple navegador (cliente ligero) o acceder a los servicios desde un *software* SIG (cliente pesado).

Más recientemente, ha habido una proliferación de geoportales que permiten compartir Información geográfica, basados en proyectos nacionales, iniciativas que cubren una región o una temática específica. Por ejemplo:

El geoportal del proyecto INSPIRE (Infraestructura de Información Espacial en la Comunidad Europea), disponible en <http://www.inspire-geoportal.eu>.

El geoportal NatCarb, <http://www.natcarb.org>, que proporciona información sobre la captura y almacenamiento del CO en USA.

El geoportal GeoNetwork de la FAO, <http://www.fao.org/geonetwork> .

17.3.3. Características

Un buen geoportal debe cuidar todos aquellos aspectos que hacen más eficaz su utilización por todo tipo de usuarios, en particular:

- Universalidad, un geoportal debe ser multilingüe y tener, al menos, una versión inglés; debe ser además visible desde los sistemas operativos más usuales (Windows, Unix, Linux, Macintosh,...) y utilizando los navegadores más frecuentes (Microsoft Explorer, Netscape, Firefox,...).

- Usabilidad, un geoportal debe ser cómodo, fácil de utilizar, autoexplicativo y sencillo, es decir usable en el sentido que define ISO 9241: 1992, especialmente las partes 11 y posteriores.
- Accesibilidad, en el sentido que define ISO/TS 16071:2003, para facilitar el acceso a usuarios discapacitados, al menos a las páginas estáticas, dadas las dificultades técnicas de hacer accesibles los contenidos cartográficos.
- Diseño, de manera que sea atractivo, ordenado, seductor y transmita los valores que el proyecto considere más importantes.
- Disponibilidad, del geoportal y de las aplicaciones cliente en él integradas.
- Rendimiento, de las aplicaciones cliente, de manera que la respuesta sea rápida incluso en momentos de alta demanda y con gran número de usuarios concurrentes. El internauta típico abandona cualquier portal web que tarde en responder más de unos pocos segundos.

Clearinghouse

Un concepto muy cercano al de geoportal y al de IDE, es el de *clearinghouse*.

El término inglés *clearinghouse* alude a una institución que recolecta información y la distribuye, como una central de reservas de hotel, de reserva de billetes de avión o una cámara de compensación de cheques bancarios, cumpliendo una función de coordinación.

En el contexto de la IG, una *clearinghouse* es un sistema informático accesible en Internet que permite al usuario: buscar en un catálogo qué datos geográficos hay disponibles y cumplen sus requisitos y necesidades, de un conjunto de nodos; y, una vez que el conjunto de datos está identificado y seleccionado, realizar una descarga de esos datos para su posterior utilización.

Por lo tanto, al integrar un conjunto de nodos para proporcionar servicios de búsqueda y descarga, es algo similar a una IDE dedicada a la obtención de datos geográficos, en la que no se distribuye el procesamiento de los datos, sino que se gestiona el compartirlos.

Una *clearinghouse* suele basarse en un catálogo de metadatos centralizado, un servicio de descarga de datos distribuido que accede a varios nodos, y un geoportal que permite al usuario acceder al sistema y utilizar los servicios que ofrece.

17.4. Arquitectura de una IDE

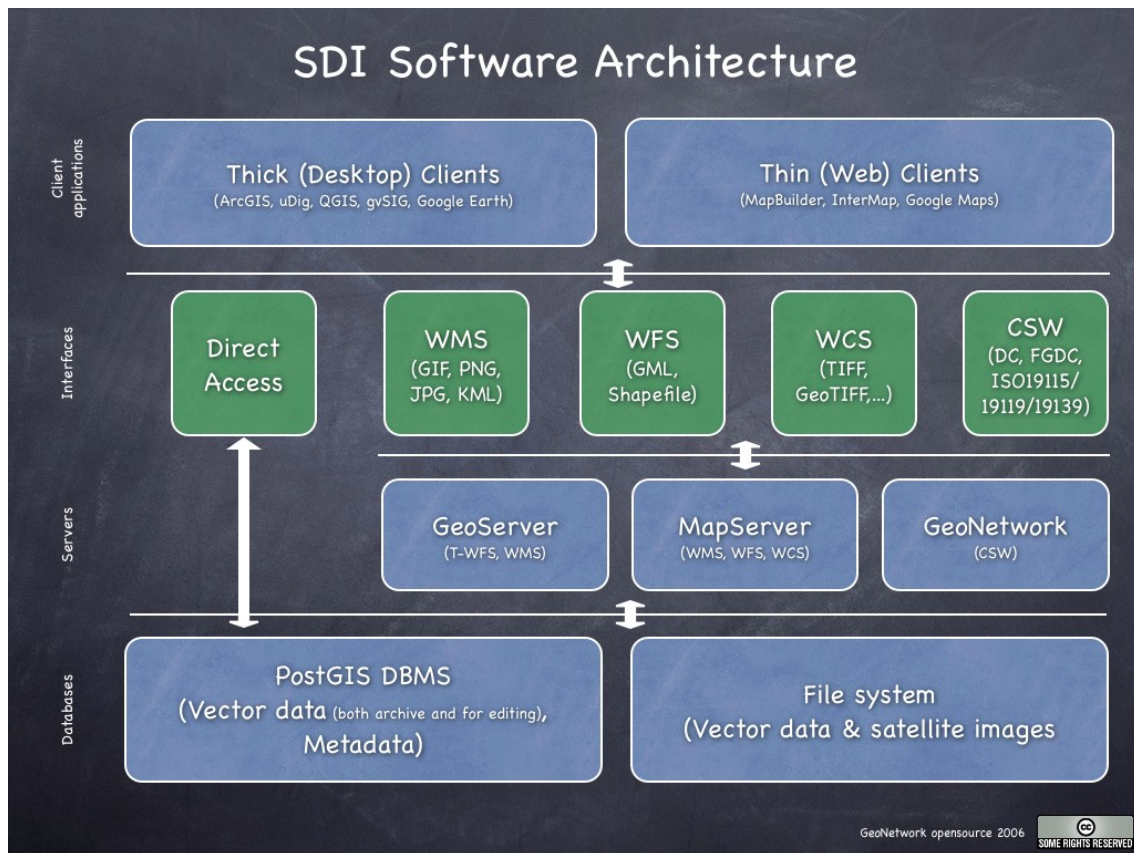


Figura 3. Imagen de Geonetwork http://geonetwork-opensource.org/img/general/sdiarch1/view?set_language=es

17.5. La Directiva INSPIRE

17.5.1. La iniciativa INSPIRE

En el año 2002 se lanzó la iniciativa europea INSPIRE (*IN*frastructure for *S*patial *IN*foRmation in Europe), cuyo objetivo era implantar una IDE Europea centrada en los problemas medioambientales, mediante una Directiva (2007/2/CE) que ya ha sido aprobada el 14 de marzo de 2007.

La motivación radicaba en los problemas de medioambientales europeos, como grandes inundaciones, lluvia ácida, contaminación atmosférica, mareas negras, que son evidentemente transfronterizos y exigen para su resolución y gestión el disponer de una IDE Europea.

La iniciativa INSPIRE nació en el año 2001, impulsada por varios organismos de la UE:

- EUROSTAT (Statistical Office of the UE)
- JRC (Joint Research Center)

- EEA (European Environmental Agency)
- IES (Institute for Environment and Sustainability)

con la formación de unos primeros grupos de expertos, dos por país, encargados de redactar los llamados *Position Papers*, cinco documentos publicados en 2002, que proporcionaban el bagaje técnico y organizativo para iniciar la redacción de la Directiva.

Durante el año 2003 se efectuó una consulta en Internet sobre el contenido de la entonces propuesta de Directiva y una consulta pública en una sesión abierta celebrada en Milán.

El Programa de Trabajo de INSPIRE ha seguido a continuación una serie de fases:

I. Fase Preparatoria (2005-2006)

- Se abre la participación electrónica de los países miembros a través de las LMO (*Legaly Mandated Organizations*), organizaciones responsables de la implementación una IDE en virtud de alguna norma legal, y de las SDIC (*Spatial Data Interest Communities*), asociaciones de grupos de interés.
- Ambos tipos de organización proponen expertos para iniciar la redacción de las Reglas de Implementación previstas en la Directiva para que la IDE Europea sea usable y sus componentes estén armonizados y sean compatibles.

II. Fase de Transposición (2007-2009)

- Las Reglas de Implementación se adoptan como Decisiones de la Comisión Europea.
- Se forman *Tematic Working Groups* para armonizar las especificaciones de los datos de cada uno de los temas definidos en ISPIRE.
- Se establece un punto de contacto con la Comisión único en cada país miembro, basado en el organismo de coordinación de la IDE nacional.
- Se transpone la Directiva Europea mediante las necesarias leyes, decretos, normas y otras disposiciones legales en los países miembros.

II. Fase de Implementación (2009-2019)

- La Directiva se pone en práctica.
- La Comisión Europea monitoriza la implantación de INSPIRE mediante informes periódicos remitidos por los países miembros.

La Directiva INSPIRE afecta a las Administraciones Públicas y establece la necesidad de implantar una IDE en Europa, basada en las IDE nacionales desarrolladas en los países

miembros, orientada a la aplicación de las políticas comunitarias de medio ambiente y a la gestión del territorio en general, conforme a cinco principios básicos.

17.5.2. Principios INSPIRE

- Debe ser posible combinar de modo continuo la Información Geográfica procedente de diferentes fuentes a lo largo de Europa y compartirla entre todo tipo de usuarios y aplicaciones.
- Debe de ser posible compartir a todos los niveles de administración, desde el más detallado al más general, la información capturada a un nivel determinado.
- La información geográfica necesaria para una buena gestión a todos los niveles debe ser abundante bajo unas condiciones que no restrinjan su uso. En particular los servicios de búsqueda y visualización deben ser públicos y gratuitos, salvo algunos casos especiales.
- Debe de ser fácil descubrir qué información geográfica está disponible, adaptarla para un uso particular y conocer bajo qué condiciones se puede adquirir y usar.
- Los datos geográficos deben ser fácilmente inteligibles e interpretables para ser visualizados dentro de un contexto apropiado y también fácilmente seleccionables por el usuario.

Se establece una diferencia clara entre dos tipos de datos: datos de referencia, descritos en los anexos I y II, que sirven como información básica y general descriptiva de la geografía de una zona que sirven para referenciar el otro tipo de datos, datos temáticos, descritos en el anexo III, que están orientados a un sector de aplicación específico. Y se estipulan plazos concretos de implementación de los servicios más elementales, dos años para los datos de referencia y cinco años para los datos temáticos, contados a partir de la aprobación de las Reglas de Implementación.

17.5.3. Filosofía y alcance

La filosofía de esta Directiva sigue una directriz genérica, la de facilitar el acceso y utilización de la IG a todo tipo de usuarios porque constituye un recurso estratégico para la protección del medio ambiente, en particular, y para la planificación de todo tipo de políticas, en general. En esa misma dirección, con diferentes matices, contribuyen el espíritu de la Convención de Aarhus, que reconoce el derecho de los ciudadanos a acceder a la información ambiental de las AA. PP., el de la Directiva de Reutilización de la Información Gestionada por el Sector Público (Directiva PSI) y el del recientemente aprobado Proyecto de Ley para el Acceso Electrónico de los Ciudadanos a las Administraciones Públicas.

La Directiva Europea 2007/2/CE INSPIRE, por la que se establece un infraestructura de información espacial en la Comunidad Europea, fue aprobada con fecha de 14 de Marzo de 2007.

Su objetivo es el establecimiento de una IDE Europea orientada a la aplicación de las políticas comunitarias de medio ambiente y de las políticas que puedan incidir en el medio ambiente.

Inspire se basará en las IDE nacionales, establecidas y gestionadas por los Estados miembros, y no prescribe la recopilación de nuevos datos.

La Directiva es aplicable a los conjuntos de datos geográficos que cumplan las siguientes condiciones:

Referirse a una zona sobre la que un Estado miembro tenga jurisdicción.

Estar en formato digital.

Tratar sobre alguno de los temas descritos en los anexos I, II y III.

Estar gestionados o bien por una Administración Pública, otra entidad que actúe en su nombre, u otra entidad que se integre voluntariamente en una IDE nacional y cumpla con las normas y estándares oportunos.

17.5.4. Metadatos

Los metadatos se generarán de acuerdo a las Reglas de Implementación de metadatos, que deberán ser elaboradas y aprobadas antes del 15 de Mayo de 2008.

Será obligatorio crear metadatos de los temas de los anexos I y II antes de dos años, y de los temas del anexo III antes de cinco años, contados siempre a partir de la aprobación de las citadas Reglas de Implementación de metadatos.

17.5.5. Interoperabilidad de datos y servicios

Las Reglas de Implementación que aseguren la interoperabilidad de datos y servicios espaciales tendrán en cuenta los requisitos de los usuarios, las iniciativas IDE ya en funcionamiento, las normas y estándares existentes, y su viabilidad y rentabilidad.

Para ello, la comisión Europea realizará los análisis de costes/beneficios y los estudios necesarios para garantizar la viabilidad y rentabilidad de las Reglas de Implementación relativas a la interoperabilidad.

Una vez aprobadas las Reglas de Implementación, que deberán estar disponibles antes del 15 de mayo de 2009 para el anexo I y antes del 15 de mayo del 2012 para los anexos II y III, los conjuntos de datos ya existentes deberán adaptarse en un plazo de 7 años, y los de nueva creación se adaptarán en un plazo de 2 años.

17.5.6. Servicios

Los estados miembros deberán crear en la web:

- Servicios de búsqueda de datos y servicios espaciales, basados en metadatos
- Servicios de visualización
- Servicios de descarga de datos
- Servicios de transformación de datos espaciales para que cumplan las Reglas de Implementación relativas a interoperabilidad.
- Servicios de acceso a servicios de datos espaciales

Los servicios de búsqueda permitirán al menos los siguientes criterios de búsqueda:

- Palabras clave
- Tema
- Calidad y validez
- Grado de conformidad con las Reglas de Implementación
- Ámbito geográfico
- Condiciones y restricciones de acceso y uso
- Entidad productora o distribuidora

Los servicios de búsqueda y de visualización serán libres, públicos y gratuitos y, excepcionalmente se podrán cobrar tasas por los servicios de visualización si así se garantiza el mantenimiento de los datos y la prestación del servicio, cuando se trata de grandes volúmenes de datos de actualización muy frecuente.

Los servicios de visualización pueden configurarse de manera que se impida su reutilización con fines comerciales.

En caso de que las AA. PP. cobren por alguno de los servicios mencionados, garantizarán la disponibilidad de servicios de comercio electrónico.

Por otro lado, se podrá limitar el acceso a los datos y servicios cuando dicho acceso pueda poner peligro:

- La confidencialidad de los procedimientos de las AA. PP.

- Las relaciones internacionales, la defensa nacional o la seguridad pública.
- El desarrollo de los procedimientos judiciales.
- La confidencialidad comercial o industrial.
- El secreto estadístico, el secreto fiscal y el derecho de las personas a su privacidad.
- Los derechos de propiedad intelectual.
- La protección del medio ambiente a que se refiere la información, por ejemplo localización de especies raras.
- El derecho a la intimidad y la privacidad, las relaciones internacionales, la seguridad pública o la defensa nacional.

Está prevista la implementación de un geoportal europeo, del que existe ya una primera versión, que sirva de punto de entrada principal a la IDE europea, y cada estado miembro deberá también implementar un geoportal nacional que sea el punto principal de acceso a cada IDE nacional.

17.5.7. Puesta en común de los datos

Cada Estado miembro adoptará medidas para la puesta en común de los conjuntos de datos espaciales y servicios relacionados con ellos entre sus autoridades públicas, mencionadas en el artículo 3, punto 9, letras a) y b). Estas medidas permitirán que dichas autoridades públicas tengan acceso a los conjuntos y servicios de datos espaciales y puedan intercambiar y utilizar tales conjuntos y servicios a efectos del desempeño de funciones públicas que puedan incidir en el medio ambiente.

Cada Estado miembro designará un punto de contacto, una autoridad pública que se encargue de los contactos con la Comisión en relación con la aplicación de la Directiva INSPIRE y de la coordinación de las distintas de competencias y responsabilidades en cada país.

Cada país llevará a cabo un seguimiento de la aplicación y utilización de su IDE, que se plasmará en informes anuales que estarán disponibles antes del año 2010 y que se harán públicos.

17.6. Iniciativas y proyectos IDE

Además de la ya mencionada iniciativa INSPIRE, para construir una IDE Europea orientada a la información medioambiental y a las Administraciones Públicas, otros proyectos IDE de relevancia, son:

17.6.1. GSDI

Global Spatial Data Infrastructure es una asociación de organizaciones, agencias, firmas comerciales y personas de todo el mundo, cuya finalidad es promover la colaboración y la cooperación internacional en apoyo del desarrollo de IDEs internacionales, nacionales y locales, que permitirán a las naciones afrontar mejor los problemas sociales, económicos y medioambientales.

Mantiene una web (www.gsdi.org) con documentos, información y enlaces a IDE de todo el mundo y organiza un congreso mundial anual.

17.6.2. UNSDI

La IDE de Naciones Unidas (<http://www.ungiwg.org/unsdi.htm>) es una iniciativa del UNGIWG (*UN Geographic Information Working Group*) que tiene como objetivo aprovechar los recursos y servicios existentes para coordinar y gestionar a lo largo del mundo las políticas de cooperación y ayuda de las organizaciones que componen las Naciones Unidas.

En estos momentos existen dos Oficinas de Coordinación de UNSDI: Una en Holanda y otra en Checoslovaquia. Está previsto abrir la Oficina Española de Coordinación de UNSDI en el año 2008.

17.6.3. IDEE

Como ejemplo de realización de una Infraestructura de Datos Espaciales a nivel nacional, podemos hablar de la IDE de España (IDEE), <http://www.idee.es>, un proyecto colectivo en el que se integran datos, servicios y recursos de:

- Organismos nacionales, como el Instituto Geográfico Nacional, la D.G. del Catastro, el Instituto Nacional de Estadística, el Fondo Europeo de Garantía Agraria, el Instituto Geológico y Minero de España, el Ministerio de Industria, el Ministerio de la Vivienda, y en breve el Ministerio de Medio Ambiente y varias entidades más.
- Prácticamente todas las Comunidades Autónomas, actualmente 14 de un total de 17, que con su empuje han contribuido de manera decisiva al éxito del proyecto. Las cinco restantes están desarrollando recursos interoperables y prevemos su puesta en marcha a lo largo del presente año 2007.
- La Diputación Provincial de A Coruña, los Cabildos insulares de Tenerife, La Palma y otros,...
- Un número creciente de Ayuntamientos, como Pamplona, Zaragoza, Getafe, Las Palmas de Gran Canaria y sendos proyectos colectivos que incluyen más de 60 Municipios en Cataluña,

más de 17 en la Rioja, más de 23 en Almería, 200 en Valencia,... hasta un total de más de 350 a finales de 2007.

- Una considerable variedad de entidades y organismos que publican datos temáticos, el Atlas Climatológico de la Península Ibérica desarrollado por la Universidad Autónoma de Barcelona; la IDE de Costas, integrada en la IDE de Cataluña; la IDE del Parque Nacional de Doñana; el Geoportal de IMEDEA (Instituto Mediterráneo de Estudios Avanzados) del Gobierno de las Islas Baleares; EUROPARC, una IDE de los espacios naturales protegidos de España; el proyecto Anthos del Real Jardín Botánico de Madrid, con un inventario de más de un millón de especies florales y su distribución geográfica; un Atlas Temático de Distribución de Aves Zonas Cinegéticas de España; y un largo etcétera que incluye un gran número de proyectos que surgen día a día en todos los sectores de aplicación y que están llevando la tecnología IDE prácticamente a todos los ámbitos de la actividad humana.

Estas entidades colaboran:

- Proporcionando servicios OGC, como los WMS del catastro, de las ortofotos del SIGPAC, o los de cartografía básica de las CC. AA.

- Con geoportales que sirven de punto de entrada a IDE de ámbito menor, como los de la mayoría de las CC.AA., la Diputación Provincial de A Coruña y los Cabildos de Tenerife, La Palma.

El Geoportal IDEE (www.idee.es), abierto en el verano del año 2004, con una interfaz en 7 idiomas (inglés, español, francés, portugués, gallego, vasco y catalán), ofrece nueve tipos de servicios OGC:

- Servicio Web de Mapas (WMS)
- Servicio Web de Catálogo (CSW)
- Servicio de Nomenclátor (Gaz)
- Servicio Web de Fenómenos (WFS)
- Servicio Web de Coberturas (WCS)
- Servicio Web de Contexto (WMC)
- Descriptor de Estilo de Capas (SLD)
- Servicio Web de Transformación de Coordenadas (WCTS)
- Servicios Web de Procesamiento (WPS).

Además dispone de algunas utilidades de análisis combinado en remoto y por municipios del Modelo Digital del Terreno 1:25.000, un mapa de pendientes derivado, un mapa de orientaciones y del Mapa de Coberturas del Suelo Corine-Land Cover 1:100.000.

Integra un total de más de 90 servicios, y permite el acceso a más de 1000 capas de información disponible.

Permite fácilmente visualizar y superponer entre sí el catastro de toda España, la cartografía 1:25.000 del IGN, las ortofotos del FEGA de 1 metro de resolución y una gran cantidad de datos geográficos de CC. AA., algunos Municipios y entidades de todo tipo.

En un próximo futuro está previsto el incluir en la IDEE los datos procedentes del Proyecto Nacional de Ortofotografía Aérea (PNOA), que contempla la obtención de ortofotos de toda España con resoluciones de 25 y 50 cm cada dos años, y la información del proyecto CARTOCIUDAD, o callejero oficial de la Administración Pública, que incluye la georreferenciación de direcciones postales, basado en los datos de Catastro, del INE., de Correos y del IGN y que ya está comenzando a dar sus primeros frutos.

Por otro lado, aunque *Google Earth* (GE) es un cliente pesado de servicios de visualización no estándar, hay en la IDEE una pasarela que permite utilizar el visualizador de GE como cliente de un servicio WMS estándar OGC que cumpla ciertas condiciones.

Creemos que una la línea de desarrollo futuro más interesante que se ha iniciado dentro del proyecto IDEE es la de desarrollo de Servicios de Procesamiento en la Red (WPS), que en la práctica supone la implementación de servicios electrónicos, públicos y normalizados de cálculo de caminos mínimos, zonas de influencia, análisis superficial, etc.

La Infraestructura de Datos Espaciales de España (IDEE), es un proyecto coordinado por el Consejo Superior Geográfico (CSG), órgano colegiado en el que están representados los productores de datos geográficos digitales de ámbito nacional, autonómico y local (a través de la FEMP) , cuya presidencia ejecutiva y secretaría desempeña el Instituto Geográfico Nacional. El CSG se organiza en comisiones y la Comisión de Geomática definió en Noviembre de 2002 un Grupo de Trabajo para la IDEE.

El GT IDEE está abierto a todos los actores que estén y trabajando realmente en el campo de las IDE: empresas privadas, AA. PP. y Universidades. Cuenta actualmente con más de 200 miembros personales que provienen de más de 80 entidades.

En este Grupo de Trabajo, sus miembros intercambian experiencias y llegan a los consensos necesarios para la implementación de una IDE en España, abierta y eficaz, de acuerdo con las

directrices marcadas por INSPIRE y siguiendo las especificaciones de interoperabilidad de Open Geospatial Consortium (OGC), en forma de Recomendaciones:

- Procedimiento para la elaboración de Recomendaciones.
- Núcleo Español de Metadatos (NEM)
- Modelo de Nomenclátor de España (MNE)
- Recomendaciones para la implementación de WMS

GT IDEE celebra tres reuniones anuales, unas Jornadas de la IDEE y se organiza en SubGrupos de Trabajo concentrados en aspectos técnicos.

17.6.4 Iniciativas regionales

PC IDEA (Permanent Committee on SDI for The Americas)

<http://www.cpidea.org>

PCGIAP (Permanent Committee on GIS Infrastructure for Asia and the Pacific)

<http://www.pcgiap.org>

17.6.5. Otros proyectos

- NSDI (USA), National Spatial Data Infrastructure

<http://gos2.geodata.gov/wps/portal/gos>

- CGDI (Canadá) Canadian Geographic Data Infrastructure

<http://cgdi.gc.ca>

- NSDI INDIA (India), National Spatial Data Infrastructure of India

<http://www.nsdindia.org>

- GDI-DE (Alemania), Infraestructura de Datos Geográficos de Alemania, que agrupa a las IDE de los 16 estados federados

http://www.gdi-de.org/de/f_start.html

- SNIG (Portugal), Sistema Nacional de Información Geográfica

<http://snig.igeo.pt>

- PROSIGA (Argentina), Proyecto SIG de Argentina

<http://www.sig.gov.ar>

- IDERC (Cuba), IDE de la República de Cuba

<http://www.iderc.co.cu>

- ICDE (Colombia), Infraestructura Colombiana de Datos Espaciales
<http://www.geoportal.gov.co>

- CNDG (Uruguay) Clearinghouse Nacional de Datos Geográficos
<http://cndg.clearinghouse.gub.uy>

Bibliografía

- [1] Groot, Richard y McLaughlin, John “Geospatial Data Infrastructure. Concept, Cases and Good Practice”, 2000, Oxford University Press.
- [2] GSIDI, “El recetario IDE v1.1”
http://www.gsdi.org/pubs/cookbook/recetario_es0515.pdf
- [3] Directiva Inspire
<http://www.idee.es/resources/leyes/INSPIRE.pdf>
- [4] VV. AA. “La Infraestructura de Datos Espaciales de España (IDEE): Un proyecto colectivo y globalizado” en las Jornadas de la IDEE 2006.
http://jidee06.uji.es/down/s11_rodriguez.pdf

Tema 18. Aplicaciones cliente y servidor en una IDE. Clientes ligeros y pesados. Tipos de servicios: Servicios de Mapas, de Catálogo, de Nomenclátor y otros. Encadenamiento de servicios. Recomendación WMS-C. Peticiones teseladas y gestión de teselas. Mashups geográficos.

18.1. Aplicaciones cliente y servidor en una IDE

18.1.1. Definición arquitectura cliente-servidor

La **arquitectura cliente-servidor**, en un sistema distribuido, está formada por clientes que solicitan servicios y servidores que responden a las peticiones.

La arquitectura cliente-servidor es un modelo para el desarrollo de sistemas de información en el que las transacciones se dividen en procesos independientes que cooperan entre sí para intercambiar información, servicios o recursos.

Se denomina **cliente** al proceso u ordenador que realiza peticiones de servicios a otros ordenadores o procesos (servidor). Es el que inicia el diálogo o solicita los recursos.

El **servidor** procesa la petición y envía el resultado de vuelta al cliente. Es el que procesa y responde a las solicitudes.

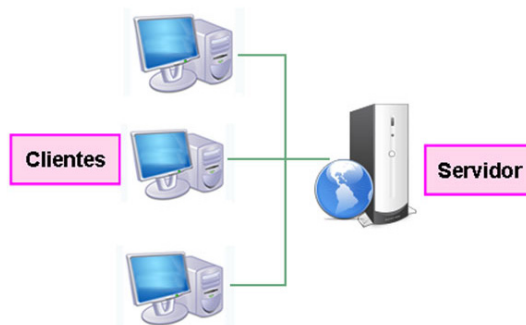


Figura 1. Arquitectura cliente-servidor

En este modelo las aplicaciones se dividen de forma que el servidor contiene lo que debe ser compartido por varios usuarios, y en el cliente permanece sólo lo particular de cada usuario. El cliente se comunica con el servidor a través de una interfaz de software que es la parte del servicio que permite el flujo de información entre el cliente y el servidor ocultando las características del servicio.

Entre las **principales características** de la arquitectura cliente-servidor se pueden destacar las siguientes:

- El servidor brinda servicio a múltiples clientes.
- El servidor presenta a todos sus clientes una interfaz única y bien definida.
- El cliente no necesita conocer la lógica del servidor, sólo su interfaz externa.
- El cliente no depende de la ubicación física del servidor, ni del tipo de equipo físico en el que se encuentra, ni de su sistema operativo.
- Es posible repartir la carga de procesamiento entre el cliente y el servidor.
- Los cambios en el servidor implican pocos o ningún cambio en el cliente.
- La comunicación con el servidor es transparente para el usuario.

1. Cliente

Es el proceso u ordenador que **inicia un requerimiento** de servicio. El requerimiento inicial puede convertirse en múltiples requerimientos de trabajo a través de redes de comunicaciones. La ubicación de los datos o de las aplicaciones es totalmente transparente para el cliente.

El cliente se encuentra conectado a una red, que le permite acceder y gestionar una serie de recursos.

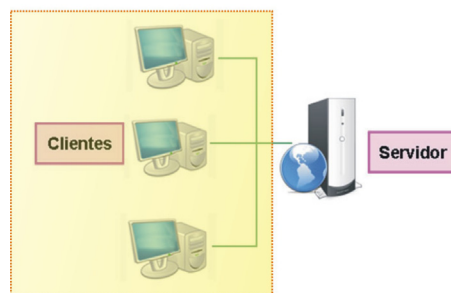


Figura 2. Clientes en arquitectura cliente-servidor

Existen dos tipos de clientes:

Cliente ligero (<i>Thin client</i>)	Cliente pesado (<i>Thick client</i>)
<ul style="list-style-type: none">• Tiene muy poca o ninguna lógica del programa.• Los procesos son realizados principalmente por el ordenador central.• La palabra ligero se refiere a lo pequeña que es la imagen de arranque, quizás no más grande que la requerida para conectar a la red y arrancar un navegador Web.• Ejemplo: navegador Web, Google Maps.	<ul style="list-style-type: none">• Realiza la mayor parte del procesamiento de datos.• Los datos son almacenados en el servidor.• Ejemplos: Aplicación Java o C++ que se instala y ejecuta, Google Earth.

Tabla 1. Cliente ligero y pesado

2. Servidor

Es cualquier aplicación dedicada a **responder a los requerimientos** del cliente.

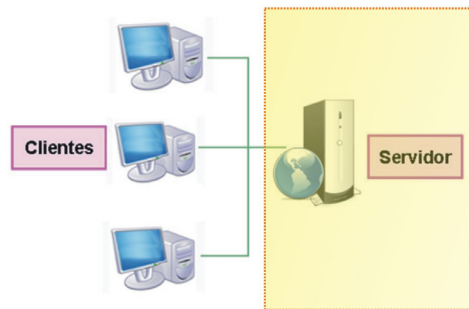


Figura 3. Servidor en arquitectura cliente-servidor

Existen muchos tipos de servidores entre los que se destacan:

- **Servidor de archivos:** Servidor donde se almacena archivos y aplicaciones de productividad como por ejemplo procesadores de texto, hojas de cálculo, etc.
- **Servidor de bases de datos:** Servidor donde se almacenan las bases de datos, tablas, índices.
- **Servidor Web:** Se utilizan para acceder a páginas Web. Este servidor permite transacciones con el acondicionamiento de un navegador específico.
- **Servidor de correo:** Es una aplicación que permite enviar mensajes (correos) de unos usuarios a otros, con independencia de la red que dichos usuarios estén utilizando.

- **Servidor de aplicaciones:** Servidor que ejecuta una serie de aplicaciones. Suele ser el encargado de realizar la mayor parte de los cálculos y procesamiento del sistema de información (la lógica del sistema).
- **Servidor FTP:** Servidor que facilita la transferencia de ficheros entre sistemas conectados a una red.

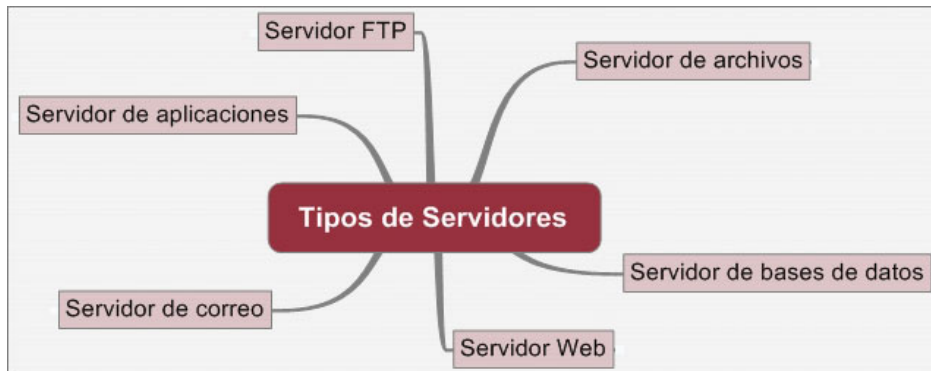


Figura 4. Tipos de servidores

18.1.2. Arquitectura Web

En Internet, la arquitectura por excelencia es la denominada arquitectura cliente-servidor, en la que una serie de clientes (navegadores Web) solicitan una serie de servicios (servidores Web). Estos últimos procesan las peticiones de los navegadores (realizadas según el protocolo HTTP) y devuelven páginas HTML. De este modo, cuando navegas en una página Web, en realidad estás utilizando un cliente Web, que se conecta al servidor de páginas Web.

1. Servidor Web

Un servidor Web es un programa que implementa el **protocolo HTTP** (*hypertext transfer protocol*).

De acuerdo al tipo de la petición recibida, el servidor Web buscará una página Web o ejecutará un programa en el servidor. Una vez procesada la petición, devolverá al cliente o navegador Web la respuesta.

El **navegador Web** realiza una petición al servidor, en este caso la petición consiste en solicitar una página Web: <http://www.google.es>. El servidor responde al cliente enviando el código HTML de la página; el cliente, una vez recibido el código, lo interpreta y lo muestra en pantalla.



Figura 5. Petición del cliente web al servidor

El cliente es el encargado de interpretar el código HTML, es decir, de mostrar las fuentes, los colores y la disposición de los textos y objetos de la página; el servidor tan sólo se limita a transferir el código de la página sin llevar a cabo ninguna interpretación de la misma.

Sobre el servidor Web clásico se puede disponer de **aplicaciones Web**. Éstas son fragmentos de código que se ejecutan cuando se realizan ciertas peticiones o respuestas HTTP. Puede distinguirse entre:

Aplicaciones en el lado del cliente	Aplicaciones en el lado del servidor
<p>Son las aplicaciones tipo Java o Javascript que se ejecutan en la máquina del usuario: el servidor proporciona el código de las aplicaciones al cliente y éste, mediante el navegador, las ejecuta. Es necesario, por tanto, que el cliente disponga de un navegador con capacidad para ejecutar aplicaciones (scripts).</p> <p>Además de aplicaciones Javascript y Java, puede ejecutarse otros lenguajes mediante el uso de plugins.</p>	<p>El servidor Web ejecuta la aplicación; ésta, una vez ejecutada, genera cierto código HTML.</p> <p>El servidor toma el código creado y lo envía al cliente por medio del protocolo HTTP.</p>

Tabla 2. Aplicaciones en cliente y servidor

2. Navegador Web

Para poder acceder a la World Wide Web es necesario utilizar un programa cliente. Este se llama **navegador Web o Web browser** y es una aplicación software que permite al usuario recuperar y visualizar páginas Web, almacenadas en los servidores Web.

Los navegadores Web ofrecen una interfaz gráfica que permite navegar por Internet y su **funcionalidad básica** es permitir la **visualización** de páginas Web, documentos de texto, recursos multimedia, etc. Las páginas Web poseen hipervínculos que enlazan un texto o imagen a otro documento, tejiendo de esta forma la telaraña de la WWW.

Ejemplos de navegadores Web:

Internet Explorer

Mozilla Firefox

Opera

Safari



Figura 6. Navegadores Web

18.1.3. Cliente-servidor en una IDE

Las Infraestructuras de Datos Espaciales se basan en el modelo cliente-servidor, en el que se distribuye el proceso pero no los datos.

El usuario que entra en el Geoportal de una IDE, es el cliente y toda su interacción con la IDE se entiende en términos de un cliente que realiza peticiones estándar a una aplicación servidora.

El usuario dispone de un ordenador cliente y de una aplicación cliente, normalmente un simple navegador, que es el que lanza las peticiones a la parte servidora. Las aplicaciones cliente, como el visualizador, residen en un servidor de la IDE, habitualmente el mismo en el que está el Geoportal de la IDE, y cuando el usuario las arranca, se cargan en la memoria de su ordenador.

El nodo IDE que proporciona los servicios tienen unas máquinas servidoras y unas aplicaciones que son los que dan el servicio, las aplicaciones servidoras, que se encargan de recibir las peticiones estándar del aparte cliente, analizar su contenido, construir la imagen que se desea visualizar, o la respuesta que se solicita, y de enviarla al usuario.

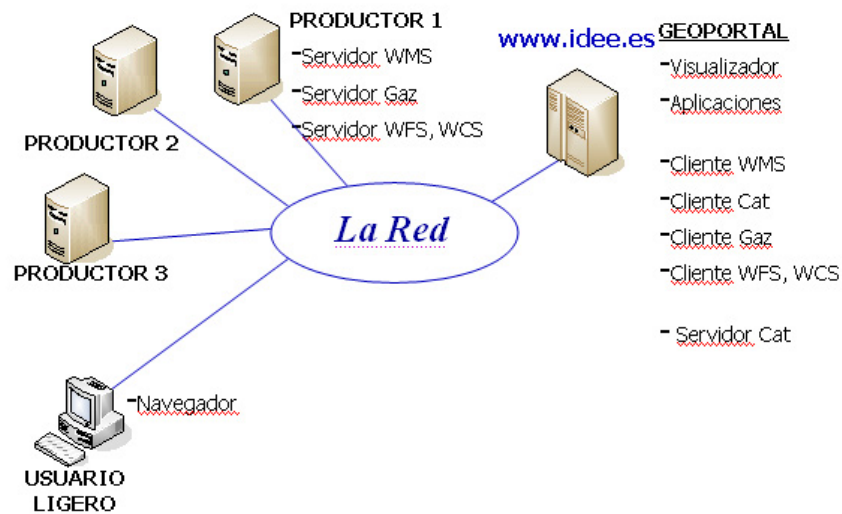


Figura 7. Cliente-servidor en una IDE

18.2. Clientes ligeros y pesados

Existen dos grandes tipos de aplicaciones cliente dentro de una IDE:

- Los clientes ligeros, que son simples navegadores o *browsers*, como Microsoft Explorer, Firefox Mozilla, Opera, Netscape, etc. desde los que se pueden enviar peticiones estándar a los servidores OGC, directamente como peticiones http o utilizando clientes elaborados (como son los visualizadores) que están en un servidor externo.
- Los clientes pesados, que en realidad son aplicaciones SIG, como gvSIG, Geomedia, ARC/GIS, u otros. Son *plug-ins* que es necesario instalar, aplicaciones que corren en el ordenador del usuario y que tienen la posibilidad de realizar peticiones estándar a servidores OGC y permiten mezclar en pantalla datos privados que gestiona la propia estación de trabajo, con datos y servicios públicos que están en la red.

18.3. Tipos de servicios

18.3.1. Servicio de Mapas

Web Map Service (WMS), o servicio de publicación de mapas, permite visualizar datos vectoriales y ráster en La Red con un simple *browser*, efectuar operaciones sencillas de visualización, como apagar y encender capas, hacer zoom, vuelo panorámico, etc. y superponer datos independientemente del servidor físico en el que se encuentren, su formato y su sistema de referencia y proyección cartográfica.

18.3.2. Servicio de Catálogo

Catalog Service Web (CSW), o servicio de publicación de catálogos, especifica cómo hacer accesibles vía red catálogos de recursos, es decir catálogos de datos y servicios, para poder realizar búsquedas distribuidas que consulten un conjunto de catálogos en un conjunto de servidores. Para ello, es necesario que el catálogo que se consulta tenga catalogados el resto de catálogos en los que se va a desencadenar una búsqueda en cascada.

18.3.3. Servicio de Nomenclátor

Gazetteer (GAZ), o servicio de Nomenclátor, que permite realizar búsquedas de topónimos con varios criterios (nombre exacto, nombre incluido, empezando por, en una zona, sólo topónimos de una clase, etc.) y devuelve su situación geográfica. Es un caso particular de WFS (WFS-G o WFS-MNE).

18.3.4. Otros servicios

- Styled Layer Descriptor (SLD), o descripción de estilo de capas, dedicada a permitir la definición temporal o permanente de la simbolización con la que se desean visualizar datos vía WMS.
- Web Feature Service (WFS), o servicio de publicación de objetos, que ofrece la posibilidad de acceder en la Red a los objetos individuales contenidos en un conjunto de datos y recuperar todos sus atributos y características para poder efectuar análisis complejos.
- Web Map Context (WMC), o contexto, que permite almacenar todos los parámetros y características de una situación activa en un geoportal OGC para reproducirla más tarde.
- Web Coverage Service (WCS), o servicio de publicación de coberturas, que permite publicar en Internet coberturas ráster de imágenes de satélite, ortofotos, fotografías

aéreas, ficheros ráster y, en general cualquier conjunto de datos que tenga esta estructura de la información, para su visualización, análisis y consulta.

- *Web Coordinate Transformation Service (WCTS)*, o servicio de transformación de coordenadas, que oferte la posibilidad en la Red de transformar un fichero de datos geográficos de una proyección cartográfica a otra o de cambiar su sistema de referencia.
- *Geolinking Data Access Service (GDAS)*, o servicio de acceso a datos geolinkados, especificación en elaboración, pensada para ofrecer una interfaz normalizada de acceso a tablas de datos alfanuméricos (estadísticos, temáticos), que incluyan una columna con un identificador geográfico, una etiqueta de georreferenciación, que permita establecer una relación con un fichero de datos geográficos que haga posible su visualización. Por ejemplo una tabla de datos estadísticos relativos a municipios que incluya el código INE o los nombres de los municipios.
- *GeoLinking Service (GLS)*, o servicio de geolinkaje, especificación todavía en preparación, concebida para poder acceder a una tabla de datos geolinkados, mediante un GDAS, seleccionar qué atributo se desea visualizar, acceder a un fichero de datos geográficos que contenga las áreas georreferenciadas la tabla en cuestión, acceder o definir interactivamente un fichero de leyenda dónde se especifique cómo se desea visualizar cada valor de un atributo, y finalmente visualizar en pantalla un mapa temático del atributo seleccionado.
- *Sensor Web Enablement (SWE)*, o servicio de acceso a sensores, pensado para poder acceder en tiempo real a datos tomados por sensores tales como estaciones de aforo, estaciones pluviométricas, puntos de medida del tráfico por carretera, sensores meteorológicos, *webcam*, etc.
- *Web Processing Service (WPS)*, o servicio de procesamiento en la web, que permite definir utilidades de geoprocésamiento con interfaces bien descritas de entrada y de salida (ficheros GML, XML, o imágenes JPEG, PNG, ... GDAS).

18.4. Encadenamiento de servicios

Los servicios OGC que ofrece una IDE se pueden encadenar, para formar servicios más complejos y sofisticados, ya que al disponer de interfaces estándar y bien definidas es posible construir una aplicación cliente que lo invoque de manera consecutiva y encadenada.

Ejemplos de encadenamiento son: el que cuando se invoca el servicio de Nomenclátor sea posible ver dónde se encuentra el resultado automáticamente mediante un servicio de visualización WMS; el que al encontrar un conjunto de datos determinado en el Catálogo de Metadatos sea posible que el usuario, desde el mismo entorno, pueda visualizar el WMS que publica ese conjunto de datos.

18.4.1. El modelo SOA

El modelo SOA (*Services Oriented Architecture*, Arquitectura Orientada a Servicios), ha sido definido en el año 2005 por el consorcio OASIS (<http://www.oasis-open.org>), un consorcio dedicado a producir estándares para e-comercio, y se adapta muy bien a la filosofía de las IDE, en las que el concepto fundamental alrededor del que gira toda la arquitectura del sistema son los servicios, a diferencia de los SIG, en los que el concepto fundamental son los datos.

Se basa en tres acciones

- Publicar servicios en la Red
 - WSDL: Web Service Definition Language
- Buscar servicios
 - UDDI: Universal Discovery, Description and Integration
- Solicitar servicios
 - SOAP: Simple Object Access Protocol

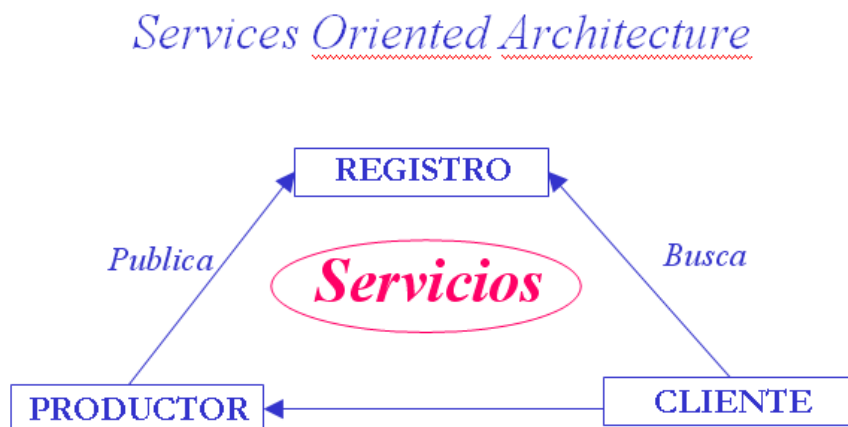


Figura 8. Arquitectura SOA

El consorcio OASIS ha definido un estándar de servicios, los servicios SOAP, a los que no se han ajustado los servicios OGC por ahora, pero se está estudiando la convergencia de estos dos estándares de facto.

18.4.2. Servicios y aplicaciones de valor añadido

La principal utilidad de una IDE no es que los usuarios puedan conectarse a su Geoportal y consultar y visualizar la información geográfica disponible. Con ser ésta una finalidad importante, existe otra que los es más.

18.5. Recomendación WMS-C

18.5.1 Ventajas e inconvenientes del WMS

El estándar WMS (Web Map Service) es la especificación OGC de la que existen más implementaciones y también una de las más estables y eficaces, por la sencillez de su concepción y por lo probada que está en estos momentos.

Sin embargo, produce en muchas ocasiones respuestas demasiado lentas, por el hecho de que si se trata de visualizar un fichero vectorial, (1) la parte servidora debe localizar los vectores que deben ir en la ventana que se pide, muchas veces realizando una consulta a una Base de Datos (proceso lento), (2) luego debe renderizar, es decir formar la imagen ráster en formato PNG, JPG u otros, que se va a mostrar en pantalla (proceso también lento), (3) a continuación transferir esa imagen hasta la aplicación cliente (proceso más lento todavía, que depende del ancho de banda del usuario y de la rapidez de su sistema), y por último (4) el sistema del usuario debe mostrarlo en pantalla (proceso en el que a veces una tarjeta gráfica poco potente arruina todo el rendimiento). Cuando hablamos de operaciones lentas, queremos decir lentas para ser tiempos de espera del internauta medio conectado a la Red, para el que esperar 5 segundos es una eternidad.

Como una posibilidad para acelerar al menos las dos primeras partes, ha aparecido recientemente la propuesta llamada WMS-C.

18.5.2 La propuesta de OSGEO

La recomendación WMS-C (abreviatura de *WMS Tile Caching*, o WMS para cachear teselas) ha sido definida en el 2006 por OSGeo, la fundación *Open Source Geospatial*, una organización sin ánimo de lucro financiada mediante donaciones, cuya finalidad es apoyar y promover el desarrollo colaborativo de tecnologías geoespaciales abiertas y libres.

El objetivo de la recomendación WMS-C es encontrar un método para optimizar la publicación de cartografía mediante servicios WMS a través de Internet, acelerar la publicación de cartografía, mediante el uso de una memoria de almacenamiento intermedio o caché.

La idea es definir una teselación estándar que permita componer cualquier imagen seleccionada al realizar una petición estándar a un servicio WMS para visualizar una ventana de visualización determinada mediante un conjunto de teselas, definidas mediante una cuadrícula independiente de la ventana, y en el nivel de escala fijo más cercano al solicitado. De esta manera, al margen de la ventana y de la escala solicitada, la imagen siempre se compondrá de las mismas teselas en formato imagen (JPG, PNG o similares) de una escala predeterminada (la más cercana a la solicitada de un conjunto de escalas fijas), y las teselas se podrán cachear, con lo que su recuperación será mucho más rápida. El tiempo de respuesta se puede reducir hasta un factor 30. Por un lado, el servidor de mapas debe utilizar en realidad un perfil de un servicio WMS, que sólo servirá teselas estándar. La propuesta WMS-C necesita que el servidor WMS declare de alguna manera que sólo admite peticiones de teselas estándar y para ello se define un parámetro `tiles = true`.

Por otro lado, la aplicación cliente debe encargarse de calcular qué teselas son las necesarias para componer la imagen solicitada en función de la zona deseada y de la escala de visualización. Para ello es necesario establecer una teselación única y bien conocida y unos niveles de escala fijos.

A la hora de preparar el almacenamiento necesario para cachear las imágenes a servir, hay dos posibilidades:

- Preparar un almacenamiento de las áreas más visitadas, en un país serán típicamente las zonas de costa y los centros de población, mediante un programa de gestión de memoria caché que se encargue de almacenar las teselas pedidas más frecuentemente.
- También es posible almacenar no la parte del conjunto de datos más accedida, sino todos los datos renderizados y preparados en forma de teselas. En tal caso ya no se puede hablar de caché, sino de una versión pasterizada de los datos para acelerar su publicación. Esta opción sólo es posible si los datos no se actualizan muy frecuentemente.

18.6. Peticiones teseladas y gestión de teselas

Existen varias teselaciones propuestas en diferentes proyectos, sin embargo la teselación estándar de hecho más frecuentemente adoptada se basa en teselas de 256 por 256 píxeles, y los saltos de escala se toman de manera que cada escala sea doble de la anterior. Se definen dos perfiles dentro de WMS-C:

1) El Perfil no Proyectado que utiliza:

Sistema de Referencia EPSG: 4326 es decir WGS84 (latitud, longitud)
Bounding Box inicial -180 -90, 180 90
Resoluciones 0,703125, 0351625,...

2) El Perfil Mercator que utiliza:

Sistema de Referencia OSGEO: 41001 WGS84 (Proyección simple de Mercator)
Bounding Box inicial -20037508.34 -20037508.34 20037508.34 20037508.34 Resoluciones
156543.03390625 78271.516953125 ...

En el caso 1, se comienza con una teselación de sólo dos teselas, para que puedan ser cuadradas y cubrir todo el mundo:

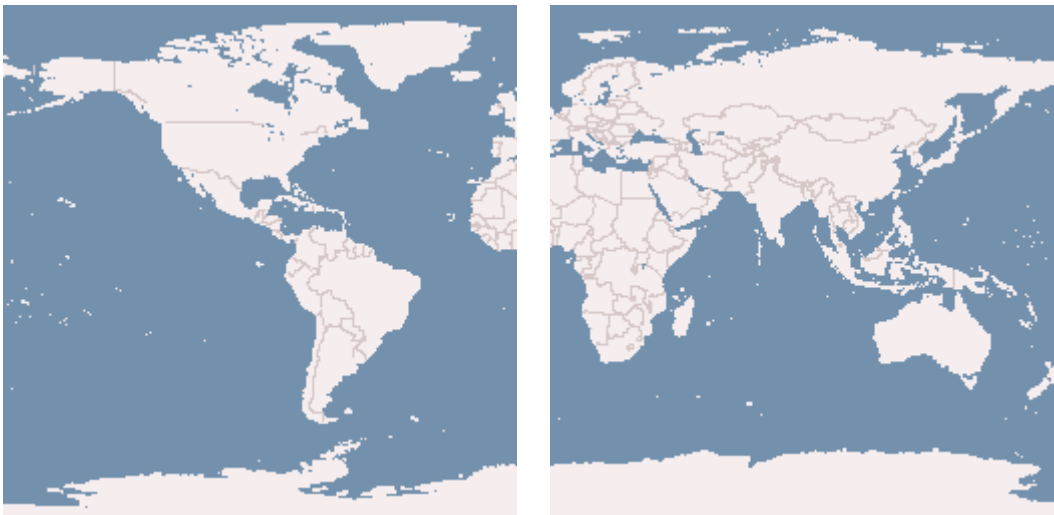


Figura 9. Teselación de 2.

una tesela desde (-180, -90) hasta (0, 90) y la otra desde (0, -90) hasta (180, 90).

Por lo tanto tanto en longitud como en latitud las teselas tienen de lado 180 grados sexagesimales, y la resolución en grados por píxel es de:

$$180^\circ / 256 \text{ píxeles} = 0,703125$$

Si la primera teselación es la de orden 0 ($n = 0$), la resolución de un orden n será:

$$0,703125 / 2^n \text{ grados por píxel}$$

El tamaño aproximado de cada tesela, en una pantalla de 15 pulgadas, es de 7 por 7 cm para una resolución de 1024 por 768, lo que equivale a una escala aproximada en la primera teselación, la de orden 0 formada por sólo 2 teselas, de 1:285 M, y de 9 por 9 cm para una resolución de 800 por 600, lo que equivale a una escala aproximada de 1:225 M.

Al realizar una operación de zoom + con un cliente que cumple WMS-C, la escala aumenta al doble y se tienen cuatro teselas cubriendo el área que antes ocupaba una:



Figura 10. Teselación de 4

Si en cada salto de escala, la escala se multiplica por dos, con sólo nueve niveles de zoom, se puede pasar de una escala 1:2M ½ hasta una escala 1:5.000, y con 10 niveles hasta una escala 1:2.500. La escala de un nivel de orden n es aproximadamente

$$E^n = 1: 250.000.000 \cdot 2^n$$

Cada tesela de información vectorial en formato JPG, bastante optimizado para datos vectoriales, tiene un tamaño aproximado de unas 10 Kb; cada tesela de datos ráster en formato PNG, más apropiado para datos ráster, ocupa unas 16 Kb.

El tamaño de la caché debe estar preparado para almacenar la parte más accedida, por ejemplo, un 25% de los datos totales, aunque esto varía mucho con el grado de concentración que tengan las peticiones de los usuarios en una superficie más o menos reducida. O alternativamente puede prepararse para almacenar el 100% de la información.

Si se parte de una escala 1:2M $\frac{1}{2}$, para representar 1.200.000 km² (la superficie del mínimo rectángulo envolvente de la península), será necesario cubrir de teselas una superficie de

$$1.200.000.000.000 \text{ m}^2 / 6.250.000.000.000 = 0,192 \text{ m}^2 = 1920 \text{ cm}^2$$

Si una tesela tiene una superficie aproximada de 7 por 7 cm, 49 cm², serán necesarias unas 40 teselas. En el siguiente nivel de zoom harán falta cuatro más y así sucesivamente. Si queremos llegar hasta una escala de visualización de 1:1.000, tendremos 11 niveles de zoom, y lo que tenemos es una progresión geométrica de razón 4 de la que queremos calcular la suma de los 12 primeros términos. La suma total será, aplicando la fórmula que da la suma de los n términos de una progresión geométrica de razón r cuyo primer término es a_n:

$$S_n = a_1 (r^n - 1) / (r-1) = 40 * (4^{12} - 1) / (4 - 1) = 223.696.200$$

es decir unas 225.000.000 teselas aproximadamente, si cada una ocupa unas 16 Kb aproximadamente 3.600 Gb, prácticamente 4 Tb.

Si se quiere tener almacenado el 25% de los datos en la memoria intermedia, y se cuenta con cierto margen de seguridad, hay que disponer de un almacenamiento para la memoria caché de 1 Tb.

Si se manejan una memoria caché de ese tamaño puede ser conveniente definir dos niveles de caché: un primer nivel de, digamos el 5%, unos 250 Gb, muy rápida, que contenga las zonas de muy alta demanda; y un segundo nivel, con unos 750 Gb, que no importa que sea un poco más lenta con las áreas de demanda un poco más baja.

La caché se puede colocar en la parte servidora, en la parte cliente, o en la red mediante un conjunto de servidores distribuidos que almacenan los ficheros necesarios.

18.7. Mashups geográficos

18.7.1. Los clientes pesados 3D

Desde hace pocos años, el sector de la Información Geográfica ha sufrido de *Google Earth*, y otros clientes pesados no estándar de servicios de cartografía.

Se trata de un cliente pesado, un *plug-in* que es necesario instalar en el ordenador del usuario, que permite acceder a las imágenes y cartografía proporcionados por servicios web de mapas no estándar, que publican datos adquiridos por *Google Earth*, normalmente en exclusiva, en forma de cartografía tridimensional navegable y con relieve.

El servicio de mapas se basa en teselas, lo que permite que las imágenes se “cacheen” en un gran número de servidores repartidos por todo el mundo, integrados en un sistema distribuido en el que colaboran para servir la respuesta.

Las últimas versiones incluyen la posibilidad de invocar servicios WMS estándar, como una capa adicional sobre los datos que sirve *Google Earth*.

El gran éxito que ha tenido *Google Earth* se ha basado en proporcionar un servicio de alta calidad, con un rendimiento, una usabilidad y una disponibilidad excepcionales. Han llevado a la práctica la idea fundamental de las IDE, que se basan en el servicio, como concepto central, en lugar de los datos, como ocurría en los SIG.

Existen otras alternativas a *Google Earth* con funcionalidades similares, y en ocasiones superiores, y la misma filosofía:

- Skyline

http://www.skylinesoft.com/corporate/corporate_home.asp

- NASA World Wind

<http://worldwind.arc.nasa.gov>

- Microsoft Virtual Earth

<http://www.microsoft.com/virtualearth>

- WW2D (S.L.)

<http://ww2d.softonic.com>

- Géoportail

<http://www.geoportail.fr>

18.7.2. Contribuciones y APIs

Sobre *Google Earth* es posible para el usuario colocar contribuciones o post, que enriquecen la cartografía con datos puntuales georreferenciados, como: fotos de Panoramio, panoramas, notas de la Wikipedia, edificios 3D realizados con Google Sketch UP, etc.

Por otro lado, *Google Earth* ha puesto a disposición de los usuarios, una API muy cómoda que permite crear un portal utilizando como fondo cartográfico la información por él servida, e integrarla así en un visualizador al efecto.

Una API (del inglés *Application Programming Interface* - Interfaz de Programación de Aplicaciones) es el conjunto de funciones y procedimientos (o métodos si se refiere a programación orientada a objetos) que ofrece cierta librería para ser utilizado por otro software como una capa de abstracción.

Una API representa un interfaz de comunicación entre componentes software. Se trata del conjunto de llamadas a ciertas bibliotecas que ofrecen acceso a ciertos servicios desde los procesos y representa un método para conseguir abstracción en la programación, generalmente (aunque no necesariamente) entre los niveles o capas inferiores y los superiores del software. Uno de los principales propósitos de una API consiste en proporcionar un conjunto de funciones de uso general, por ejemplo, para dibujar ventanas o iconos en la pantalla. De esta forma, los programadores se benefician de las ventajas de la API haciendo uso de su funcionalidad, evitándose el trabajo de programar todo desde el principio

18.7.3. Neocartografía

Se llama Neocartografía a una nueva manera de producir información geográfica en la red, basada en la colaboración de los usuarios, que mediante aportaciones, o post, contribuyen al enriquecimiento de una información cartográfica de base, que puede ser una imagen, una foto o un mapa, añadiendo información semántica georreferenciada, como fotos, puntos de interés, etiquetas, notas, nombres, panoramas, topónimos, etc.

Se puede decir que la Neocartografía no es en realidad más que la consecuencia de la Web 2.0 en el campo de la publicación de cartografía en Internet. Web 2.0 es un término acuñado en 2004 por Dale Dougherty (O'Reilly Media) para referirse a una Web caracterizada por:

- Evolución de las aplicaciones tradicionales hacia aplicaciones web enfocadas al usuario.
- Aparición de aplicaciones que generan colaboración, con el florecimiento de los *mashups*, o web híbridas.
- Utilización de la red como auténtica plataforma.
- Emergencia de soluciones innovadoras a partir de aportaciones características distribuidas generadas por desarrolladores independientes.
- Sindicación de contenidos.
- Cooperación, explosión del fenómeno blog (bitácoras), extensión del *tagging* (etiquetado), construcción de folksonomías.
- Es más una actitud, una forma de trabajar, que una tecnología.

18.7.4 Mashups geográficos

Una aplicación web híbrida (*mashup* o remezcla), es un sitio web o aplicación web que usa contenido de otras aplicaciones Web para crear un nuevo contenido completo, consumiendo servicios directamente siempre a través de protocolo http.

El contenido usado en un *mashup* es típicamente usado de terceros a través de una interfaz pública o usando un API. Otros métodos que constituyen el origen de sus datos incluyen: sindicadores web (RSS o Atom), Screen scraping, etc.

Así como los weblogs han revolucionado la publicación en línea, los *mashups* están revolucionando el desarrollo web, permitiendo que cualquiera combine, de forma innovadora, datos que existen en eBay, Amazon.com, Google, Windows Live y Yahoo. Las grandes facilidades brindadas por simples y ligeras API's han hecho que los *mashups* sean relativamente fáciles de diseñar. El que se requiera de mínimos conocimientos técnicos ha hecho que los *mashups* sean creados en su mayoría de casos por innovadores, quienes combinan en formas nuevas y creativas datos disponibles públicamente. Así como hay *mashups* muy útiles, existen otros que no pasan de sólo ser novedosos o publicitarios, con mínima utilidad práctica.

Desde la aparición de *Google Earth*, *Google Maps* y sus correspondientes APIs, que han disparado la aparición de soluciones de ese tipo, se han abierto un bien número de *mashups* de carácter geográfico o cartográfico:

- GEONAMES (CC: by)

<http://www.geonames.org>

Contiene más de 6 millones de topónimos georreferenciados (nombres con coordenadas) de todo el mundo, con gran número de exónimos obtenidos de la wikipedia, procedentes de más de

35 fuentes de datos. Esta base de datos está abierta a modificaciones y correcciones mediante una wiki que tiene en cuenta la fiabilidad de la fuente si es un organismo oficial.

- OPEN STREET MAP (CC: by-nc-sa)

<http://www.openstreetmap.org>

Proyecto pensado para que voluntarios capturen con trazadores GPS trazas de calles y carreteras y las suban a este repositorio común.

- VEOGEO

<http://www.veogeo.com>

Permite subir recorridos descritos con coordenadas y un vídeo, capturados con un móvil que tenga cámara de fotos y GPS, o equipamiento parecido, y consultarlos para ver de forma sincronizada el video y cómo un cursor recorre el itinerario sobre un croquis.

- SHAPEWIKI (CC: by)

<http://shapewiki.com>

Diseñado para dibujar *shapes*, o rectángulos, sobre la cartografía de *Google Maps*, cargarlos de etiquetas y subirlos, para luego compartirlos y consultarlos, aun repositorio.

- WIKIMAPIA

<http://wikimapia.org>

Similar al anterior, permite marcar zonas mediante rectángulos y describirlos mediante comentarios. Las contribuciones admiten votos negativos (con más peso) y votos positivos, que sirven para evitar cibervandalismo y consolidar la información más interesante.

- WIKIMAP

<http://www.wikimap.es>

Mapa colaborativo en el que los usuarios añaden puntos de interés sobre un callejero.

- WIKILOC (CC: by-nc)

<http://www.wikiloc.com>

Proyecto abierto a los usuarios registrados para que compartan rutas y localizaciones capturados mediante GPS para practicar 30 tipos diferentes de actividades al aire libre. Utiliza como cartografía de fondo Google Earth, el Catastro y el IDEE-Base, entre otros.

- TAGZANIA

<http://www.tagzania.com>

Proyecto para etiquetar el mundo. Permite colocar puntos georreferenciados, descritos por etiquetas; los usuarios ven en cada momento las etiquetas más utilizadas, con lo que se crea una folksonomía.

- GOOLZOOM (google Maps + Catastro + SIGPAC)

<http://www.goolzoom.com>

- CARTESIA (Google Maps + Geonames + Wikipedia)

<http://gmaps.cartesia.org>

- CHICAGO CRIME

<http://www.chicagocrime.org>

Estos mashups tienen en común varias características que los hacen especialmente interesantes:

- Son colaborativos, es decir están pensados para que los usuarios cargen el sistema con sus aportaciones y de esta manera compartan sus gustos, conocimientos, consejos, etc. con la comunidad.

- La información se puede utilizar de manera bastante libre, en la mayoría de los casos bajo una de las licencias Creative Commons (CC), que están definidas en la red (véase <http://es.creativecommons.org>) y permiten hacer todo tipo de usos y operaciones con los datos, siempre que se respeten algunas condiciones:

by: Atribución. Si los datos se redistribuyen o difunden, se ha de mencionar la fuente original.

nc: No comercial. No se permiten los usos comerciales de los datos.

sa: Compartir igual. Si se redistribuyen los datos o se difunden, ha de hacerse bajo el mismo tipo de licencia.

Bibliografía

- [1] Groot, Richard y McLaughlin, John “Geospatial Data Infrastructure. Concept, Cases and Good Practice”, 2000, Oxford University Press.
- [2] GSDI, “El recetario IDE v1.1”
http://www.gsdi.org/pubs/cookbook/recetario_es0515.pdf
- [3] Recomendación WMS-C de OSGEO
http://wiki.osgeo.org/index.php/WMS_Tile_Caching
- [4] Teselación estándar
http://wiki.osgeo.org/index.php/WMS_Tiling_Client_Recommendation

Tema 19. Modelos Digitales del Terreno: Análisis, formatos y aplicaciones (análisis de pendientes y orientaciones, mapas de visibilidad, perfiles). Algoritmos y métodos.

19.1. Modelos Digitales del Terreno

La representación del relieve del territorio ha sido siempre una de las claves de interpretación de los mapas desde su origen. En épocas anteriores a la representación por técnicas digitales se utilizaba una simbología específica que iba desde roquedo, terraplén, cortado, farallón, acantilado, etc. hasta las isolíneas de altitud constante conocidas generalmente por curvas de nivel.

Las nuevas tecnologías de la información han querido recoger todo lo bueno de la representación del relieve clásica e introducir el tratamiento automatizado de las mismas en vistas a tener una representación digital y por lo tanto susceptible de generación, manejo y explotación de esta representación del territorio. El tener los datos altimétricos de un territorio en forma digital les hace susceptibles de un tratamiento impensable hace algunas décadas obteniéndose mediante el análisis de los mismos productos derivados de gran importancia e interés.

La implementación de estos datos digitales en un SIG ha sido cuestión de tiempo [Miller y Laflamme, 1958 y Guevara, 1990]. Desde un principio se sabía que este trabajo requería elevadas capacidades de proceso para un computador [Miller y Laflamme, 1955-60] así como un almacenamiento masivo, en cualquier caso prohibitivo en las primeras generaciones de la informática. Actualmente la revolución tecnológica en materia de informática ha permitido ponernos sobre la mesa los datos y los procedimientos para generar y tratar lo que se conoce con el acrónimo de MDT, o sea los Modelos Digitales del Terreno.

19.2. Introducción: La guerra de las siglas

Conceptualmente la representación digital del terreno puede establecerse desde distintos enfoques, lo que ha dado lugar a una guerra de siglas que en el fondo hacen referencia todas a lo mismo, difiriendo solo en la forma de abordar la solución para obtener el modelo de la orografía de la superficie terrestre. Así tenemos:

Modelos Altimétricos Digitales [Christensen, 1992]- MAD

Modelos de Elevación Digital - MED (DEM son sus siglas en inglés)

Modelos Digitales de Alturas - MDA, o de elevaciones, o de cotas.

Modelos Topográficos Digitales [Cebrián y Mark, 1985] - MTD

Modelos Digitales del Terreno - MDT

Otras nomenclaturas serían Matrices de Altitudes [Evans, 1980 o Matrices de Superficies].

En sentido estricto un MDT (DTM en inglés) es una matriz regular uniforme y ortogonal de puntos definido usualmente como “una matriz regular de datos bi-dimensionales” [Kratky, 1978]. Collins [Collins, 1978] incluye en su definición las características de esa matriz regular: “Un MDT es una matriz de números que representa la distribución espacial de un conjunto de propiedades del terreno”, no solo las altitudes.

Frente al MDT, el MED (DEM en inglés) se usa solo cuando se trata de una única propiedad: cotas [Collins, 1978], aunque también es verdad que las restantes propiedades del terreno, como pueden ser cobertura vegetal, uso del suelo, etc. no se incorporan de forma rutinaria [Doyle, 1978]. En general en un MDT deben estar implícitas una serie de propiedades representadas a través de unas variables de superficie [Nagy, 1979]: $F=f(x,y)$, siendo F una función uniforme de valor único para cada par de valores (x,y) , necesariamente discretos.

Para llegar a obtener esta función podemos de partir de cierta información del terreno que podemos enumerar en:

- Isolíneas (curvas de nivel).
- Puntos aislados.
- Líneas geomorfológicas, o líneas de cambio de pendiente.
- Red Irregular de Triángulos (TIN).
- Secciones verticales.
- Perfiles o líneas de barrido (scanlines), ‘drop lines’, paralelos o semiparalelos.
- Matriz uniforme o variable de muestreo de valores de altitud.

Notar que la función $F=f(x,y)$ no se ha definido en función de la tercera coordenada, la z, ya que lo que se genera al construir un MDT es un único valor para esta coordenada y sería más riguroso hablar de una función $z=f(x,y)$ con valores discretos para las tres coordenadas. En general una superficie sería una función $f(x,y,z)$ con valores para todo el rango de z, lo que ocurre es que los MDT solo obtienen un valor de esa coordenada z por lo que no es estrictamente tridimensional (3D) y se les suele llamar modelos 2.5D.

Los modelos 3D son muy útiles en otras ramas de la ciencia: meteorología, geofísica, etc. Existiendo software SIG que puede manejar estos datos 3D, pero en nuestro interés, dentro de los MDT, solo queremos un valor de esa z que podíamos suponer como una condición de contorno en un modelo tridimensional estricto.

En los SIG se suele trabajar con estos datos 2.5D ya que se muestran suficientes para los objetivos de análisis y simulación de los datos altimétricos gestionados en el SIG.

Este tema se ha estructurado en las distintas fases que debe abarcar el estudio de los Modelos Digitales del Terreno, a saber:

- Características.
- Generación.
- Almacenamiento.
- Explotación: análisis y representación de resultados.
- Estudio de errores.

19.3. Características

Un MDT o Modelo Digital del Terreno en su acepción más amplia es cualquier representación en formato digital (soporte magnético) de una superficie topográfica, apropiado para ser gestionado por tecnologías de la información. Dentro de los SIG esta acepción se amplía incluyendo otras propiedades del terreno: tipos del suelo, vegetación, hidrografía, red vial, etc.

La necesidad de tener un MDT se hace patente al considerar el elevado número de aplicaciones que los utilizan, que sin interés de ser exhaustivas, podemos enumerar en distintas áreas:

- a) Producción de mapas: curvas de nivel, sombreado del relieve.
- b) Ciencias de la Tierra: Hidrología, Geomorfología, Geología, Geodesia.
- c) Ingeniería de Forestal y Montes.
- d) Planificación Urbana y Regional
- e) Ingeniería Civil, Obras Públicas y de Minas
- f) Teledetección: simulación de los efectos radiométricos en la topografía, eliminación de las distorsiones debidas al relieve en imágenes de satélite (corrección geométrica).
- g) Aplicaciones militares.
- h) Telecomunicaciones: visibilidad de radio-enlaces, recubrimiento radioeléctrico.
- i) Cartografía: ortofotos, cartografía temática, manejo de datos radar.

- j) Ciencias espaciales: relieve de planetas o satélites del Sistema Solar
- k) Cartografía planetaria.

Muy a menudo se dice que el MDT es modelo ráster y hay que aclarar que aunque puede serlo, no tienen por que serlo forzosamente. Como veremos mas adelante la estructura que puede tener un MDT puede ser matricial, puramente ráster o estar basada en una red irregular de triángulos; algunos programas de ordenador generan el MDT necesario directamente de los datos de muestreo según la resolución requerida, no guardando este MDT construido ya que su fin es algún producto derivado como la ortofoto digital [Bernhardsen, 1992].

El modelo conceptual que se tiene de un MDT esta muy ligado a la estructura subyacente digital que contiene la información guardada en los datos. Una superficie continua teselada como una malla regular ortogonal de datos tal como una estructura ráster nos da la idea del terreno a través de prismas de base cuadrada y distinta altura (Figura, 9.x) sobre una superficie, en general un elipsoide de revolución en cartografía, aunque esta superficie puede variar con la escala. A muy grandes escalas puede tomarse como una superficie plana.

Cuando utilizamos una red de triángulos irregulares la descripción del terreno se realiza mediante una concatenación de triángulos (tan próximos a triángulos equiláteros como sea posible). Esta descripción permite tener de una forma mucho más compacta la información que en una ráster, pero en contra tiene que es mucho más complicada generarla y tratarla en algunas aplicaciones, además, se han de guardado mas parámetros que las simples cotas de los vértices pues la topología no está implícita como en el modelo ráster sino que hay que obtenerla.

Un estudio comparativo de las estructuras de los modelos digitales del terreno nos llevaría a establecer una relación entre solo dos de ellas que están en extremos opuestos de las que actualmente se encuentran mas en boga considerando además que el tema de las estructuras no es un tema cerrado y no es de solución única ya que depende esencialmente de las necesidades del usuario.

Por una parte hemos de considerar las estructuras de mallas ortogonales regulares y por otra parte las mallas de triángulos irregulares. Las primeras son artificiosas y no dependen del fenómeno concreto representado, lo que lleva a que trata de igual forma los terrenos llanos y los rugosos provocando una redundancia de datos en el primer caso o una clara insuficiencia en el segundo.

En cuanto a las mallas de triángulos irregulares son estructuras muy flexibles con una gran adaptación a la naturaleza de cada configuración topográfica concreta. Además en algunas cuestiones de análisis son claramente superiores a las estructuras ráster (mapas de pendientes). En contra, la derivación de las curvas de nivel es claramente laboriosa y la malla es muy costosa de obtener a partir de los puntos de muestra; también es inadecuado para visualizar sin dibujar y viceversa [Bernhardsen, 1992].

Entre ambos tipos de estructuras se encuentran otros con los que comparten ventajas e inconvenientes siendo este campo muy prolijo en investigaciones y publicaciones intentándose obtener resultados cada vez mejores tanto en la fidelidad al terreno, puntos de muestreo, velocidad de computo de los modelos o estructuras de archivo (optimización del volumen de datos y de acceso a los mismos). Entre estas estructuras intermedias merece la pena destacar las estructuras adaptables basadas en una descomposición de tipo regular como los *quadtree* o los *octree* [Lauzon, 1985], o los que definen el muestreo progresivo [Makarovic, 1973, 1978].

Se han mencionado las ventajas e inconvenientes de las dos grandes estructuras de mayor uso en los modelos digitales del terreno. La elección de utilizar una u otra se debe basar en considerar factores tales como los datos de partida para generar el modelo (como se han capturado, como se han guardado,...), las condiciones físicas y lógicas del sistema de información geográfica donde va a ser integrado (no todos los SIG en el ámbito de componentes físicos o lógicos están preparados para manejar datos ráster, aunque la mayoría pueden trabajar con datos RIT) y, sobre todo, las aplicaciones previstas a desarrollar en función de los objetivos (la explotación del MDT) [García, L., 1995].

Por todo lo anterior parece que la estructura RIT es más eficiente y se adapta mejor a la naturaleza de los SIG. También parece que es mucho más complejo trabajar con ellas en aplicaciones tales como perspectivas, sombreado, representación gráfica (visualización,...), etc. Sin embargo no hay que olvidar que la estructura adoptada para los MDT en la mayoría de los SIG presentes en el mercado utiliza la malla regular debido a su diseño simple y facilidad de explotación, tanto es así que se ha convertido *de facto* en una forma de intercambio y comercialización. En cualquier caso la conversión de estructuras RIT a malla regular y viceversa siempre es posible.

Algunas formas primitivas de almacenar un MDT eran a través de tablas secuenciales de coordenadas (x, y, z) cuando se seguía un determinado paso de malla obteníamos un modelo matricial que podía ser ráster si solo se guardaban las cotas. Posteriormente la sistematización en la entrada de datos afianzó la construcción de mallas regulares que de algún modo compacto

y optimizado al análisis y explotación se guardaban en el ordenador. Sin embargo la malla variable no es desdeñable ya que se adapta mucho mejor a las características del terreno que la malla regular, ya que considera muy baja densidad de puntos en terrenos llanos (pocos cambios de altitud) y alto densidad de puntos en terrenos rugosos (gran cantidad de cambios en altitud). Para guardar mallas variables se suelen utilizar estructuras de árbol cuaternario (*quadtree*) por su capacidad de compactación de datos preservando fielmente la descripción original del terreno.

Hay que reafirmar que un MDT es, como su nombre indica, un modelo o maqueta de la realidad y que por lo tanto solo es una aproximación a la misma ya que la realidad del terreno se hace especialmente inadaptable a una descripción o comportamiento matemático. En general una descripción analítica del terreno es imposible por diferentes razones: datos insuficientes y afectados de error, superficies complejas a interpolar, ámbito local o global de estas superficies, falta de generalidad de los métodos matemáticos considerados, tiempo de cálculo, errores e imprecisiones locales elevados en los MDT generados, etc. La modelización de datos es una abstracción de la realidad a través de un número de objetos o características de la misma, entonces, definiendo estos objetos, sus interrelaciones y su comportamiento preciso nos permite generar el modelo de la realidad lo mas ajustado posible a nuestros requerimientos [Morehouse, 1990].

19.4. Métodos de generación

Los métodos de generación de un MDT están condicionados por las características de los datos de partida, por la forma en que estos son capturados o convertidos a un formato digital susceptible de tratamiento con ordenador, y por el resultado o aplicación que se quiere hacer de los mismos.

Trataremos por tanto las distintas formas de entrada de datos para la obtención de los MDT y analizaremos las técnicas de generación de los modelos digitales del terreno que se adaptan mejor a las características de los datos de partida.

Las superficies del terreno son superficies del relieve de este que varía continuamente. Las características prominentes del terreno se describen selváticamente a través de muchos accidentes geográficos incluyendo pendientes, depresiones, altos, líneas características, farallones, cortados, picos, roquedos, etc. La geometría sin embargo solo tiene tres términos: punto, línea y área. En principio no se puede describir un terreno que varía continuamente (en

sentido matemático) usando solo tres variables discretas, por lo tanto toda descripción será necesariamente una aproximación a la realidad.

Esencialmente los modelos digitales del terreno comprenden la colocación de puntos individuales de los que conocemos sus coordenadas (x, y, z). En la mayoría de los casos el propósito es calcular nuevos puntos de altitud a partir de los originales, lo que se conoce como interpolación. “La definición de un conjunto de *funciones-modelo* convenientes para codificar las relaciones de vecindad entre las proyecciones de los puntos de entrada en el plano XY es el primer paso necesario para construir una representación efectiva de la superficie” [Floriani, 1987].

La estructura de los modelos digitales del terreno que se pretenden generar con los datos muestrales que contamos puede adoptar principalmente tres formas:

- mallas regulares
- mallas variables
- red irregular de triángulos (RIT)

En el párrafo 19.4.1. trataremos de la generación de los modelos digitales del terreno en estructura de malla continua y en el punto 19.4.2. de los RIT.

19.4.1. Muestreo de los datos

Hay dos formas de construir un MDT partiendo de datos discontinuos, una es por muestreo regular y la otra es partiendo de un muestreo irregular.

En un principio EE.UU. ha trabajado sobre muestreo regular, principalmente el USGS, en fotogramas estereoscópicos; actualmente el USGS esta muestreando las curvas de nivel (muestreo irregular) para obtener MDT de gran calidad. En Europa se trabajaba sobre todo con muestreo irregular hasta hace unos años. Actualmente las herramientas informáticas disponibles han obligado a cambiar esta tendencia prefiriendo ahora usar datos malla o ráster.

El muestreo regular consiste en aplicar una malla regular de coordenadas (x, y) sobre un mapa o documento fotogramétrico alternativo (ortofoto, minuta de restitución, etc.) y anotar los valores z de los puntos intersección de la malla.

El muestreo irregular se genera mediante un grafo lineal digital (DLG-Digital Line Graph) en el que el operador elige una información representativa del terreno (por ejemplo una curva de nivel de cota determinada).

El muestreo regular esta sesgado por la distancia entre líneas adyacentes de la malla o resolución de la malla:

$$\Delta x = x_{i+1} - x_i \quad \Delta y = y_{i+1} - y_i$$

En general se acepta que $\Delta x = \Delta y$ por simplicidad del modelo de datos.

El muestreo irregular esta sesgado por la elección por la elección del fenómeno orográfico a elegir: líneas de cambio de pendiente (*break-line*), curvas de nivel, etc.

En un principio la restitución era analógica, luego surgieron los restituidores analíticos: seleccionada una z, automáticamente el ordenador busca la línea de igual paralaje.

Sin embargo aparecen “artefactos” en los MDT construidos como son los “oleajes” y las “estrellas de mar” (Figura 9.x) en zonas llanas (superficies de agua, etc.), es lo que se llama “ruido del modelo” provocado por la distancia muestral o por el tipo de muestreo. Para construir los MDT se buscan divergencias o convergencias de la pendiente (primera y segunda derivadas de la superficie) (ver Figura 9.x).

Cuando no tenemos posibilidades de un muestreo regular debemos utilizar un muestreo irregular, pero la topología superficial es complicada de generar con un muestreo irregular ya que no existe una relación topológica clara entre los elementos de la superficie como son las curvas de nivel. Estas curvas son vectores dentro de un SIG. Hay que complementar la información de las curvas con puntos de deformación súbita que se quieran señalar (*break-line*, farallones, acantilados, etc.). Para esto se toma una muestra irregular de puntos de los que hay que averiguar la z.

19.4.2. Captura de datos

Ya Miller y Laflamme [1958] adelantaron los métodos operativos de captura de datos topográficos: la medida del mapa y la medida del modelo estereoscópico. Estos mismos autores sugieren también la automatización “sustituyendo al operador humano por un sensor de barrido automático”.

Torlegar [1981] inventarió los diferentes modos de muestreo de altitudes en:

1. Muestreo aleatorio (puntos).
2. Registro de elementos especialmente característicos del terreno en cuestión (puntos y líneas).
3. Perfiles.
4. Mallas rectangulares.
5. Curvas de nivel.

Es muestreo aleatorio del terreno para la obtención de datos no se considera un método apropiado para describir la topografía por su misma característica aleatoria.

En cuanto al segundo modo, la toma de muestras de elementos característicos del terreno, ya sean puntuales (cimas collados, depresiones, etc.) o lineales (hidrografía, divisorias de cuencas, etc.) exige su captura de forma no automatizada o como mucho semiautomatizada para el caso de la hidrografía por lo que el esfuerzo y el costo es importante pero son de un gran valor para la obtención de MDT de precisión (con menos errores de los habituales) y pueden introducirse en procesos de interpolación como puntos críticos, lo que habitualmente elimina distorsiones importantes en los MDT producidos.

Anteriormente se ha mencionado el modo de muestreo por mallas rectangulares homogéneas o su variante irregular por perfiles, añadir a lo ya comentado que este método de muestreo no se adecua a las características de la topografía del terreno que pretenden describir. El caso es que los valores de altitud del terreno en puntos o líneas no muestreados ha de obtenerse por interpolación, con los pros y contras que eso conlleva.

El método de muestreo se destaca como el más adecuado por diversas razones [Clarke *et al.*, 1982]:

1. Es selectivo: las características del terreno son explícitamente o implícitamente registradas tales como cumbres, líneas de ruptura de pendiente, divisorias, cauces,...
2. Es homogéneo: se consideran todas las partes del estereomodelo.
3. Es progresivo: la densidad de datos se incrementa a medida que el terreno se hace más articulado.

4. Considera la geomorfología local a través de la estructura de las curvas de nivel, así como que la distancia entre líneas nos indica el límite de las elevaciones que pudieran quedar sin registrar.

19.4.3. Digitalización de curvas de nivel

Según Doyle [1978] existen en la actualidad cuatro formas genéricas de digitalizar curvas de nivel, aunque solo tres de ellas son operativas. La primera forma es hacerlo manualmente. Un operador digitaliza la curva en una mesa de digitalización siguiendo la curva dibujada en el documento de trabajo. Esta forma tradicional de entrada de datos se aplica solo cuando el trabajo a realizar es sobre una zona muy reducida o no se tienen medios para aplicar alguno de los tres métodos restantes. Se sabe que los errores que se producen, así como la duplicidad de elementos de captura y de omisión, son realmente elevados por lo que en realidad solo se debe aplicar para la depuración interactiva de ficheros provenientes de otro origen. Por otra parte hay que mencionar que el operador tiene delante de sí un trabajo pesado y tedioso lo que lleva al cansancio prematuro y a no prestar la suficiente atención a la labor de digitalización. Por lo tanto no podemos sino destacar como poco operativo este primer método.

La digitalización automática de líneas vino a resolver un problema dentro del mundo de la digitalización de curvas de nivel. El operador solo debe posicionar el cursor sobre una curva de nivel, darle la cota adecuada y lanzar el trabajo. El sistema es capaz de seguirle la pista a la curva de nivel hasta terminar en algún lugar donde hay que tomar una decisión de continuidad o de cambio de curva a digitalizar.

El tercer modo es registrar digitalmente los movimientos de la marca de referencia calibrada de un estereoploter, o cualquier restituidor, analógico durante el trazado de cada curva de nivel. Los puntos que configuran las curvas de nivel se van guardando bien por muestreo temporal, bien por muestreo longitudinal, en un fichero digital al ir la marca de referencia trazando la curva de nivel en el restituidor.

La última posibilidad nos permite una gran capacidad de trabajo y consiste en partir de un fichero digital tipo ráster que contenga las curvas de nivel [Sircar *et al.*, 1990]. Se procede por tanto a rasterizar o escanear el documento analógico donde están recogidas las curvas de nivel (por ejemplo minuta de restitución de altimetría), también se puede usar una cámara de video. Obtenido el fichero digital se procede a aplicar algoritmos de reconocimiento de formas y de seguimiento de características lineales que van extrayendo del fichero digital las coordenadas que configuran las distintas curvas de nivel. Los problemas que se presentan no son simples pero

tienen solución. Por una parte la limpieza del fichero digital generado por la rasterización se debe aplicar para eliminar las imágenes del polvo, suciedad o rasgos indeseables presentes en el documento original mediante técnicas de proceso de imagen (por ejemplo filtrado de moda) para eliminar puntos aislados. Por la misma técnica podemos hacer resaltar más las líneas sobre el fondo aplicando filtros de realce de bordes que permiten definir y contrastar mejor las curvas de nivel frente al fondo del documento. Otra cuestión a considerar es el tamaño de ventana de rasterización o muestreo. Para una perfecta definición de las curvas de nivel se debe trabajar con no menos de 200 dpi (puntos por pulgada) ya que si no los algoritmos pueden tener graves problemas de seguimiento lo que llevaría a la interrupción de la rasterización automática.

En general los ficheros obtenidos son imágenes bitmap o mapas de bit que indican en formato ráster la presencia (bit a valor 1) o ausencia (bit a valor 0) de la curva de nivel en el fichero digital, por lo que el tamaño de este fichero, a pesar de la elevada resolución muestral necesaria, no es muy grande (además se guardan comprimidos - formato RLE de Intergraph) y el tiempo de ordenador más operador se hace totalmente asumible frente a otras técnicas de rasterización de curvas de nivel.

19.5. Generación de malla continúa

Como se ha mencionado anteriormente distinguimos entre mallas regulares y mallas variables. En las primeras el paso de malla se mantiene constante en las dos direcciones (aunque este paso no sea el mismo en una dirección que en otra - mallas regulares rectangulares), en las segundas el paso de malla puede variar a conveniencia atendiendo a distintas características del territorio (relieve llano, relieve montañoso, etc.)

Las mallas regulares constituyen hoy en día la estructura más utilizada en el establecimiento de un MDT. En ella los puntos que se obtienen y que constituirán el MDT se encuentran equiespaciados en los dos ejes del sistema de coordenadas adoptado, formando una matriz de altitudes que puede utilizarse muy fácilmente. Si bien esta estructura presenta evidentes similitudes a la estructura de datos ráster, no olvidemos su gran diferencia: puntos frente a superficies (las teselas regulares o píxeles) [García, L. 1995].

Las mallas suelen ser cuadradas, es decir igualmente equiespaciadas en ambos ejes aunque también se puede considerar una malla rectangular aunque esto complica un poco su explotación posterior: sombras, perspectivas, pendientes, etc. En cualquier caso se mantiene una propiedad muy importante de esta forma de estructura del modelo ya que basta explicitar las

coordenadas de uno sólo de sus puntos para referenciar geográficamente toda la malla facilitando enormemente su uso.

Las mallas variables surgen para resolver el gran inconveniente del uso de la malla regular que es su deficiente adaptación a las características del terreno. Para que una malla regular pueda describir con suficiente precisión la morfología del terreno se debe hacer muy densa (muy fina) con lo que se complica su uso al generarse un elevado volumen de datos. Una forma de conseguir que se reduzca este volumen de datos es utilizar una malla adaptativa que permita seleccionar en cada caso un tamaño de malla más adecuado al terreno en el que se trabaja en cada momento.

No se nos escapa que para el uso de estas mallas variables se debe categorizar el terreno en áreas de morfología similar y posteriormente asociar a cada área un paso de malla óptimo. Cualquiera de las dos tareas lleva a complicar el trabajo de producción automática de los modelos digitales del terreno, además de que los posibles tipos de terreno y su paso de malla asociados nos llevarían a darle una complejidad al modelo que es precisamente lo que queremos evitar a la hora de su explotación.

La generación de los MDT en estructura de malla continua atiende a distintas técnicas según el tipo de datos de partida: puntuales (cotas), lineales (curvas de nivel) o continuos (imágenes).

En los dos primeros casos son necesarias aplicar técnicas de interpolación para generar una malla densa de datos continuos a partir de los datos muestrales. En el tercer método se aplican técnicas de estéreo-correlación automática de imágenes digitales ya sean estas de origen aéreo o espacial.

Un método de interpolación consiste en establecer, para cada punto de la malla, la media ponderada de las altitudes de los puntos próximos. Los factores de ponderación de las altitudes (pesos) se calculan normalmente como el inverso del cuadrado de las distancias al punto de la malla. En todo caso, los pesos suelen ser función del inverso de la distancia pero también pueden utilizarse otras ponderaciones que, por ejemplo, consideren la mayor o menor influencia que debe tener cada tipo de datos.

La generación de una malla regular puede hacerse considerando todos los datos de partida cualesquiera que sean estos: nube de puntos acotados de distribución irregular, puntos aislados relevantes (cimas, vértices geodésicos, cerros, etc.), curvas de nivel, líneas estructurales del terreno (divisorias, vaguadas, etc.), ortoperfiles fotogramétricos, secciones verticales, etc., o

cualquier combinación de ellos. En principio se puede manejar esta masa de puntos como si tuviera una distribución irregular.

19.5.1. Generación partiendo de curvas de nivel

Si tenemos una descripción digital (por series de pares de coordenadas bidimensionales) de las curvas de nivel de un mapa topográfico, queremos deducir de estos datos las altitudes correspondientes a los puntos de una malla ortogonal que cubra el mismo espacio, construyendo un modelo digital del terreno lo mas adaptado posible al relieve de la zona. Cualquier algoritmo genérico de interpolación puede ser aplicado a la serie de puntos que describen las isolíneas (líneas conectando puntos de igual altitud) como se ha planteado en el punto 3.4 considerándolos aleatoriamente distribuidos. Sin embargo los resultados de estos algoritmos son menos precisos que aquellos que se han desarrollado explícitamente para interpolar puntos generando una superficie de la que se tiene solo la descripción de sus isolíneas. (Kriging, etc.).

Podemos clasificar los métodos de generación de modelos digitales del terreno a partir de curvas de nivel en cuatro grupos [Clarke *et al.*, 1982; Cebrián *et al.*, 1985] basados en la forma en que se interpolan los puntos para generar la malla ortogonal de datos y el número de ellos que intervienen.

1. - Interpolación lineal entre cada dos puntos pertenecientes a curvas de nivel intersecadas en cada una de las direcciones previamente definidas. Se suelen tomar una, dos (bilineal) o cuatro direcciones [Figura 9.x], lo que nos da dos, cuatro u ocho puntos incluidos en las curvas de nivel correspondientes que intervendrán en la interpolación del nuevo valor. Cuando intervienen más de dos puntos se ha de realizar algún tipo de promedio entre los valores interpolados linealmente a partir de cada par de puntos situados en la misma dirección.
2. - Interpolación lineal entre dos puntos pertenecientes a curvas de nivel localizados en la dirección de mayor pendiente. Las direcciones para el cálculo de la pendiente suelen estar limitadas a cuatro (los ejes de coordenadas o Norte, Sur, Este y Oeste) o a ocho (los cuatro anteriores y las bisectrices de los cuadrantes).
3. Interpolación no lineal entre más de dos puntos pertenecientes a las curvas de nivel intersecadas en una dirección definida. Esta dirección puede ser alguna previamente determinada o la de mayor pendiente. En este caso se suele usar un polinomio de segundo o tercer grado, siendo necesarios 6 puntos en el primer caso y 10 en el segundo.

4. Interpolación por ajuste en un plano de regresión a varios puntos próximos, pertenecientes a curvas de nivel, encontrados en una serie de direcciones determinadas previamente.

Otro método de interpolación usa únicamente los fragmentos de curvas de nivel que se encuentran en el semiplano a la derecha de la columna del punto de la malla que se está calculando. Los puntos a la izquierda de dicha columna, necesarios para la interpolación, se toman de puntos ya interpolados o puntos auxiliares interpolados linealmente entre ellos. Este método se aplica en proceso paralelo [Clarke; Lee *et al.*, 1984] ya que los segmentos que constituyen las curvas de nivel pueden ser clasificadas en regiones y aplicar el algoritmo a cada una de las regiones simultáneamente.

En estudios de precisión de los modelos digitales generados por distintos métodos [Clarke; Sendra, 1992] se destaca como mejor el algoritmo de interpolación usando la dirección de mayor pendiente, siendo este el método que mejor modeliza el relieve de la superficie terrestre y que mejor interpreta un mapa de curvas de nivel. En cuanto al resto de los métodos no se encontraron diferencias significativas entre ellos, por lo que el método de interpolación lineal nos es más recomendable.

Como ya se ha mencionado se han de considerar tres factores para evaluar la precisión de los modelos digitales del terreno generados: la distribución de los datos, el interpolador utilizado y el tamaño de malla de salida en el modelo generado. Es obvio que cuanto mayor sea la densidad de la malla (paso más fino), más fielmente representará esta la superficie topográfica a la que se refiere. También está claro que a menor tamaño de malla, mayor cantidad de puntos hay que calcular, luego mayor tiempo de cómputo, y también mayor volumen de datos a guardar. Por lo tanto lo que tenemos que hacer es evaluar la relación coste/beneficio para establecer si es oportuno densificar la malla del modelo y que grado de mejora en la precisión alcanzamos con ello.

Analíticamente podemos pensar en obtener una precisión máxima si contamos con un número infinito de puntos, en rigor sería así; pero no debemos olvidar que estamos generando un modelo de una realidad que es muy compleja y que por lo tanto estamos admitiendo simplificaciones de ella. ¿Cuál es el grado de simplificación admisible? La respuesta depende de varios factores, pero el más importante es el conjunto de aplicaciones previsto a desarrollar o utilizar con el modelo generado. Por tanto debemos tener siempre en cuenta el grado de fidelidad que pretendemos obtener en función de las aplicaciones que va a tener el MDT generado. No es lo mismo generar un modelo para diseñar un trazado de una autopista que para obtener un mapa de distribución de señales de TV por repetidores o radio-enlaces.

En experiencias realizadas con un mismo tipo de datos [Doytsher y Shmutter, 1982] se ha demostrado que la introducción de elementos característicos de la topografía (break-line, divisorias, hidrografía, etc.) permite conseguir mayor precisión que el disminuir el tamaño de malla, aunque esta disminución sea de hasta un 35 por ciento. Actualmente algunos organismos oficiales responsables de la generación masiva de modelos digitales están ignorando este resultado y aplican tamaños de malla muy reducidos para conseguir elevadas precisiones. El motivo podríamos buscarlo en que esto conlleva la introducción exógena de características topográficas que no se encuentran en formato digital compatible con el esquema de producción de los modelos o que incluso no se encuentran en formato digital y su digitalización provocaría el retraso en la producción o la pérdida del automatismo en la cadena de producción al ser necesaria la inclusión de un operador.

19.5.2. Generación por estéreo correlación automática de imágenes digitales

En este caso los datos muestrales son continuos y presentan una estructura ráster. El concepto de interpolación se aplica al igual que en el caso anterior pero de distinta forma y con sentido diferente. Hay que distinguir entre imágenes obtenidas desde vuelos fotogramétricos e imágenes de origen espacial. En el primer caso el formato original no es digital, si no que hay que proceder a la conversión correspondiente mediante la rasterización de los fotogramas correspondientes que en su zona de solape van a permitir generar el modelo digital mediante la reconstrucción de las respectivas tomas de vista y la intersección de los haces para el mismo punto del terreno mediante la determinación de la paralaje.

Cuando trabajamos con imágenes de satélite, esta ya han sido capturados por detectores digitales a bordo de las plataformas espaciales. Estos sensores dan la información de muestreo sobre la superficie terrestre en formato digital y en estructura ráster. En este caso las medidas del suelo son radiométricas correspondiente a una área del suelo y no a un punto como en los casos anteriores y su correspondiente información geométrica ha de ser implementada a posteriori mediante una rectificación y una georreferenciación.

Un caso especial de geometría de toma de vista, por lo tanto de captura de datos, es cuando el sensor trabaja en el rango de las microondas (radar) y no en el rango óptico del espectro electromagnético, que es el caso que vamos a estudiar en este punto. En este caso de imágenes radar la generación del modelo digital del terreno se produce por interferometría y se estudia en el punto 3.3.3.3.

Se ha de distinguir en este estudio entre imágenes de origen aéreo e imágenes de origen espacial dado que la geometría y las formas de captura de datos son muy distintas. En este tema solo abordaremos el tratamiento de las imágenes aéreas.

Imágenes aéreas

La generación de modelos digitales del terreno generados a partir de pares estereoscópicos se basa en la fotogrametría. La fotogrametría es fundamentalmente un método de medida espacial. Son objeto de la fotogrametría las medidas de localización de un cuerpo, su longitud, dirección, tamaño, forma, área, volumen y magnitudes semejantes.

Desde un punto de vista geométrico, la fotografía es una proyección de una perspectiva del objeto fotografiado. En una perspectiva todos los rayos procedentes del objeto convergen en un punto, siendo la proyección de la perspectiva la traza de la intersección de los rayos convergentes con un plano. El plano de la película o placa fotográfica es el plano de intersección. La métrica de la fotografía permite la recuperación del punto de convergencia de los rayos procedentes del objeto, tanto física como matemáticamente, y una vez recuperado el foco (punto de convergencia), la reconstrucción de los rayos mismos.

Para una mejor comprensión de la generación de los modelos digitales del terreno partiendo de imágenes debemos estudiar la reconstrucción física del objeto tridimensional fotografiado y el instrumento que la permite. Para un estudio más detallado y profundo se puede consultar cualquier referencia de la bibliografía citada al final del capítulo.

Cuando se tomó la fotografía, los rayos procedentes del objeto incidieron en la película o placa fotográfica. Si colocamos en la misma posición de la cámara del proyector (que puede ser considerado desde el punto de vista geométrico similar a la cámara, ya que la única diferencia es que invierte el sentido de los rayos luminosos), los rayos volverán a su punto de origen en el objeto. De manera semejante, si disponemos de un segundo proyector situado en otra posición de la cámara proyectando la fotografía correspondiente que forma el par fotogramétrico que queremos reconstruir, los rayos procedentes de un punto imagen que aparezca en ambas fotografías (puntos homólogos) se intersecarán en el espacio en el lugar de su punto (en el objeto) de origen común. En estas condiciones estamos reconstruyendo los haces de la toma de vista y ya que estamos proyectando un modelo tridimensional que coincide en el espacio con el objeto original. Démonos cuenta que el modelo espacial formado no puede ser percibido, a no ser que a cada ojo se le presente únicamente una de las dos imágenes proyectadas. Cuando se hace así el cerebro se encarga de fundir las dos imágenes en un único modelo tridimensional. Este es el fundamento de la visión estereoscópica. La intersección de los rayos procedentes de

los puntos homólogos se encontrará en una posición sobre una superficie de referencia, la coordenada z de esta posición será la cota o la altitud de ese punto. Si se reconstruye una imagen digital píxel a píxel podemos obtener una malla de coordenadas z para cada pareja de puntos homólogos y si esto se hace de forma digital habremos generado una malla continua de altitudes o modelo digital del terreno.

Para poder realizar medidas sobre el modelo estereoscópico que sean válidas como medidas del objeto original se necesita la escala del modelo y su localización en un sistema de referencia establecido. La escala y la orientación del modelo pueden determinarse si se conoce la posición horizontal de dos puntos en el modelo. Para establecer el plano horizontal de referencia, de manera que planos horizontales en el objeto lo sean también en el modelo, necesitamos las diferencias de altitud de tres puntos (ya que tres puntos determinan un plano) en el modelo. Normalmente las altitudes están referidas a un datum seleccionado acorde con el resto de información geográfica con la que se trabaja en el sistema de información geográfica; en este caso se requiere además conocer la altitud respecto del datum seleccionado de uno de los tres puntos utilizados en la definición del plano horizontal de referencia.

La captura de medidas en el modelo estereoscópico se realiza a través de un punto visible de referencia que tiene movimiento libre en las tres dimensiones, como un punto flotante, en el espacio imagen y que además puede calibrarse de forma que siempre conocemos sus coordenadas tridimensionales. La intersección de los rayos homólogos se pone de manifiesto por la perfecta visibilidad de este punto sobre el modelo que se está construyendo de forma virtual cuando nos encontramos ‘posados’ sobre el terreno. Moviendo el punto de referencia sobre el modelo se pueden seguir las características de este, fundamentalmente planimetría y altimetría. En el caso de altimetría se busca el terreno que mantiene una cota seleccionada con lo que se van siguiendo las curvas de nivel de correspondientes a esa cota.

Según se van siguiendo las características del terreno a través del punto flotante es indispensable que los valores que van tomando las coordenadas (x, y, z) de los puntos de la traza del mismo se vayan recogiendo en algún medio digital para su uso posterior e incorporación a un sistema de información geográfica (planimetría y altimetría).

El restituidor fotogramétrico es el aparato que permite la formación del modelo estereoscópico, su medida y el registro de la misma. Las máquinas que realizan restitución del relieve utilizando las técnicas fotogramétricas son instrumentos fotogramétricos equipados con módulos de correlación automática de imágenes.

Para la generación de los modelos digitales del terreno se debe proceder a la rasterización de los fotogramas aéreos que constituyen el par estereoscópico que sirven de base al modelo fotogramétrico los cuales son digitalizados utilizando un escáner. Una vez obtenidos los ficheros digitales correspondientes solo se trabaja ya con ellos. Tenemos que destacar que para proceder a dar los pasos siguientes se han debido realizar dos tareas fundamentales con los fotogramas, es decir, el apoyo fotogramétrico y la aerotriangulación. La primera consiste en dar coordenadas a una serie de puntos fácilmente distinguibles en la foto, y por lo tanto en la imagen digital, labor que se realiza mediante trabajo de campo por arrastre de coordenadas desde los vértices geodésicos próximos a los puntos seleccionados en la foto. La aerotriangulación nos va a permitir referir los fotogramas a un sistema de referencia común, relacionándolos entre sí. Esta referencia común nos permite obtener una serie de transformaciones que aplicadas a cada uno de los fotogramas consideran las distintas posiciones que tuvieron estos en el momento de hacerse las fotografías aéreas. En principio se refieren las fotos a un sistema de referencia local (orientación relativa) y posteriormente a un sistema de referencia geodésico (orientación absoluta). Para realizar todo esto se ha de proceder a la obtención de los coeficientes de la transformación para lo que necesitamos las coordenadas en el terreno y en la imagen digital del conjunto de puntos de control (al menos seis).

El último paso lo realiza el correlador de imágenes digitales, el cual identifica a partir de las dos fotografías digitalizadas toda una serie de pares de puntos homólogos, calculando en función de sus coordenadas en las imágenes digitales y de los coeficientes de transformación calculados previamente, las coordenadas terreno (x, y, z) de la posición que representan. Los algoritmos de correlación automática suelen restituir puntos espaciados regularmente en ambos ejes de coordenadas horizontales, por lo que se genera una malla regular ortogonal por lo que tendremos al final una matriz de elevaciones que se corresponde con los vértices de esta malla.

La resolución, sensibilidad, ajuste de la curva de respuesta, etc. es una materia muy específica de cada equipo de rasterización en particular aunque podemos establecer algunos criterios para realizar esta labor. En primer lugar los documentos a digitalizar han de estar limpios de marcas, señales, polvo y deben ser de buena calidad fotográfica sin saturaciones ni zonas oscuras. Cuando tenemos fotos a digitalizar de baja calidad fotográfica debemos ajustar las curvas de respuesta (curva γ) de los sensores electrónicos (CCD, etc.) del escáner para intentar paliar los defectos de la fotografía modificando la curva γ para que tome la forma mas parecida a una rampa de 45° , aunque esta labor se debe realizar también para la calibración del propio escáner y en cualquier caso podemos realizar las pruebas que consideremos necesarias hasta obtener una imagen que mejore al original. Para el tamaño de malla se ha de pensar en la utilidad que va a tener el modelo digital generado: ortofoto digital, generación de curvas de nivel, sombreados,

etc. Y por supuesto la escala gráfica que se le van a dar a esos productos. En cualquier caso si el objetivo e la detección de objetos (fallas sísmicas, fracturas, etc.) se habrá de rásterizar con un tamaño de malla que corresponda a la mitad de la dimensión más pequeña del objeto buscado.

Los procedimientos actuales de generación de modelos digitales del terreno por estéreo correlación digital de imágenes utilizan la llamada “geometría epipolar” a aplican el proceso de correlación por relajación jerárquica (HRC) [Helava, U.V., Chapelle, W.E.; 1972]. La geometría epipolar es un concepto fotogramétrico desarrollado en esta disciplina cartográfica que permite simplificar la búsqueda de puntos homólogos (correlación) en un proceso unidimensional.

Según podemos apreciar en la figura 9.x todo plano que pase por los centros proyectivos de un par fotográfico (plano epipolar o plano nuclear) corta a los dos planos focales en dos rectas (rectas epipolares) que tienen la siguiente característica: cualquier punto de uno de ellas tiene su homólogo en la otra. Por tanto el homólogo de un punto cualquiera de la foto izquierda habrá que buscarlo exclusivamente en la recta de intersección del plano epipolar que pasa por ese punto, con el plano focal en la foto derecha. Dada esta característica geométrica se procede a la simplificación del proceso de autocorrelación preparando ambas imágenes mediante una transformación geométrica de ambos fotogramas (proceso de epipolarización) de forma que las nuevas imágenes digitales sean rectas epipolares. Con esto conseguimos simplificar el proceso de cálculo del modelo digital del terreno y además eliminar las paralajes verticales posibilitando por tanto una visualización perfecta del relieve en pantalla estereoscópica, por otra parte utilísimos para una posterior depuración del modelo generado.

El proceso de correlación ya se ha descrito anteriormente. Una estrategia para abordar este proceso consiste en usar un algoritmo eficiente como es el de relajación jerárquica que pretende reducir la intervención humana al máximo en la tarea de correlación consiguiendo una autentica auto-correlación. El proceso se basa en atacar el problema por aproximaciones sucesivas, para ello se preparan imágenes de los fotogramas con distinta resolución espacial a modo de distintas capas (ver figura 9.x) en una pirámide de base rectangular siguiendo un procedimiento de remuestreo en las que se va reduciendo el tamaño del fichero digital de forma que cada píxel representa cada vez mas píxeles (2x2, 4x4, 8x8, 16x16, 32x32,...) de la imagen de partida según vamos ascendiendo por las secciones de la pirámide. Dichas imágenes remuestreadas se utilizan sucesivamente en el proceso de correlación en sentido inverso a como fueron generadas. Empezando por la imagen 1:32 se van calculando mallas cada vez más densas, hasta llegar a la malla de densidad deseada, que se calcula con la imagen original 1:1. Por este procedimiento cada paso intermedio va generando un modelo digital del terreno cada vez más exacto y más fiel a los detalles del relieve de la zona.

Una ventaja de este método es que, basándose en la cota aproximada de cada punto obtenida con el nivel de resolución anterior según descendemos por la pirámide, se restringe al máximo el intervalo de paralajes posible y por tanto la longitud del segmento en que hay que buscar el punto homólogo cuando se va a refinar dicha cota. Esto conlleva una reducción del tiempo de cálculo y además conseguimos reducir los errores de correlación en falsos puntos homólogos ya que estos se producirán en un cierto nivel, si bien serán posteriormente arrastrados hasta la base de la pirámide, pero mantenemos bien correlado el nivel anterior que podemos recuperar para relanzar el proceso modificando la estrategia (ver mas adelante).

Dependiendo de las características geométricas (rugosidad) y radiométricas (contraste) del terreno, el algoritmo se comportará mejor o peor en la fase de autocorrelación. Por eso se han de indicar una serie de características del terreno que se aplican a cada una de las etapas de la relajación jerárquica. Para cada nivel de resolución se definen unos parámetros que cuyos valores condicionan la estrategia para cada etapa de autocorrelación. Estos parámetros son esencialmente la longitud y anchura de la ventana de correlación, longitud de la zona de búsqueda de posibles puntos homólogos (mayor cuanto más elevado sea el terreno) y otros relativos a la función de correlación (criterios para la elección de los picos, umbrales de rechazo de una correlación, etc.).

La elección acertada de los parámetros anteriores condicionan fuertemente el comportamiento del algoritmo y la calidad del modelo digital construido [Theede, J. *et al.*; 1983]. Los problemas más comunes que presenta este algoritmo se deben a la presencia de distintos tipos extremos de relieve sobre la misma imagen, en cuyo caso el correlador se comporta de forma equívoca. Así mismo el correlador no tiene un buen comportamiento al abordar desniveles muy grandes (crestas, farallones, acantilados, etc.), grandes zonas sin detalles de poco o nulo contraste en la imagen (bosques cerrados, aguas embalsadas, sombras, etc.) y zonas muy montañosas.

Una estrategia que permite disminuir la influencia de lo accidentado del terreno es partir de un modelo digital aproximado que construiríamos con el mismo procedimiento pero con un paso de malla muy superior (por ejemplo ocho veces) al definitivo, corregirlo interactivamente, lo que llevará poco trabajo por ser un fichero muy pequeño, y remuestrearlo al paso de malla definitivo. Este modelo nos serviría como etapa de lanzamiento dentro del algoritmo de relajación jerárquica [Vivas, P, *et al.*; 1991].

19.6. Generación de RIT

Las redes irregulares de triángulos, como su nombre indica [García, L., 1995], representan una triangulación espacial del terreno establecida usando como vértices los puntos que constituyen los datos de partida. Como triangulación, su adaptación al terreno es óptima (esto lo saben muy bien los geodestas y los topógrafos) [García, L, 1995]. Como, además, no contempla un solo punto más de los que se usan para establecerla, se trata de una estructura también óptima en eficiencia.

Ya se han mencionado a lo largo de este capítulo algunos inconvenientes de esta estructura de modelo digital del terreno en particular la ausencia de topología *a priori*, lo que implica tener que calcularla y guardarla en la propia estructura RIT. En efecto es necesario establecer la topología de los triángulos, es decir, establecer una estructura que relacione cada triángulo con los vértices que lo delimitan y con los triángulos contiguos (o también, cada vértice con todos los que forman con él un lado de un triángulo). Otros inconvenientes también señalados son su complejidad a la hora de su explotación y su elevado volumen de datos.

En la generación de un RIT no interviene la interpolación de altitudes como en el caso de la generación de estructuras de malla, en este caso el problema es establecer la topología de una triangulación con puntos de altitud conocida previamente seleccionados.

Cuando queremos genera un RIT tendemos a conseguir el máximo número posible de triángulos equiláteros (o próximos a equiláteros) puesto que son estos los más adecuados para describir la topografía. Un método que se aplica comúnmente es la triangulación de Delauney.

La triangulación de Delauney se realiza en dos pasos: primero se construyen los polígonos de Thiessen y segundo se genera la triangulación propiamente dicha. Los polígonos de Thiessen se construyen alrededor de cada punto muestral para ello se conecta cada punto en cuestión con los puntos muestrales de su vecindad más inmediata, luego se construyen las mediatrices de esta conexión, por último se selecciona el polígono más próximo resultante de las intersecciones de estas mediatrices, este es el “polígono Thiessen” y sus vértices son llamados “vecinos Thiessen” Una vez que el polígono Thiessen de un punto se ha determinado, el procedimiento continúa de forma sistemática para construir los polígonos Thiessen de todos los puntos de entrada. Después de esto, la triangulación de Delauney se obtiene directamente a partir de que cada triángulo está formado por un punto muestral y dos “vecinos Thiessen”; ver gráfico.

La triangulación de Delauney verifica los requisitos establecidos por Stefanovic [1988] para la generación de estructuras RIT, a saber:

- Todos los puntos muestrales deben quedar incluidos en la RIT.
- los triángulos no pueden solaparse.
- Los lados de los triángulos deben tener longitud mínima.
- Los triángulos deben ser lo más equiláteros posible (ángulos próximos a 60 grados).
- El RIT producido debe ser independiente del algoritmo utilizado.

Además de la triangulación Thiessen, existe otro algoritmo que cumple los requisitos anteriores y que también se utiliza de forma sistemática para la generación del RIT, este es el algoritmo de “barrido radial” [Mirante y Weingarten, 1982].

El algoritmo de barrido radial incluye las siguientes etapas (ver gráfico 9.x):

1. Se selecciona como punto inicial de la triangulación el punto más cercano al centroide. Desde este punto se calcula la distancia y el acimut a todos los puntos del conjunto de puntos muestral. A continuación se ordenan respecto al azimut.
2. Se conecta el punto central con todos los demás, lo que produce una red de triángulos finos y alargados (la mayoría son isósceles, un ángulo de mas de 90 grados).
3. Las concavidades que aparecen por la triangulación del barrido inicial son cerradas creándose nuevos triángulos.
4. Para optimizar sus formas, siguiendo a Stefanovic, cada triangulo es comparado con sus vecinos. Este proceso se repite iterativamente hasta que no se producen cambios.

19.7. Detección y corrección de errores

La calidad de un modelo digital del terreno debe darse tanto en forma cualitativa como en forma cuantitativa. Si bien la derivación de curvas de nivel puede darnos una forma cualitativa de la calidad del modelo construido debemos de paramétrica esta calidad con un valor cuantitativo. Este valor cuantitativo se expresa en la precisión del modelo a través de medidas estadísticas sobre el mismo y más concretamente con el cuadrático medio. La influencia que sobre este valor tiene la calidad de los datos de entrada no es desdeñable, ya que sabemos que los datos muestrales están afectados de error y no solo en coordenadas (x, y) sino en la coordenada z que

es la que más nos interesa en este estudio. Un MDT se da por bueno si el error cuadrático medio es tres veces menos que el intervalo entre líneas de nivel que se pretenda utilizar en la generación de estas a partir del modelo digital del terreno [Allder *et al.*, 1982].

La expresión matemática usada es:

$$\sqrt{\frac{\sum_{i=1}^n (\sum T_i - \sum MDT_i)^2}{n}}$$

Donde $\sum T$ es el sumatorio de las altitudes medidas sobre el terreno de una serie de puntos de control; $\sum MDT$ es el sumatorio de los mismos puntos el modelo digital del terreno construido; y n es en número de puntos de control.

Otra forma de estimar la calidad de una modelo digital construido es la generación de curvas de nivel e investigar aquellas que en superficie presentan una diferencia mayor del 10% respecto a curvas de nivel que tuvieran otro origen (fotogrametría) [Clarke *et al.*, 1982]. También es fácil contrastar los accidentes morfológicos del MDT construido con el mapa de la zona ya que existen herramientas de proceso de imagen que pueden extraer estos accidentes del MDT ya que tiene una estructura ráster, por lo que las divisorias, vaguadas, acantilados, etc. Se pueden verificar visualmente de forma muy rápida.

Dado el alto grado de automatismo que se está alcanzando en la generación de los modelos digitales del terreno, actualmente se está trabajando mucho sobre la depuración de los mismos de tal forma que cuando lleguen a incorporarse en un SIG se encuentren libres de errores groseros y mantengan una calidad uniforme por toda la malla regular de altitudes. En este sentido los programas generadores de MDT incorporan herramientas de depuración cada vez más automáticas y cada vez más sofisticadas. No solo se utiliza una pantalla gráfica de alta resolución para la representación visualizar del relieve ya sea en coloración hipsométrica, relieve sombreado, o perspectivas, actualmente se incorpora tecnología para ver el relieve estereoscópicamente bien por pantallas de luz polarizada o bien por emisores de rayos infrarrojos que presentan a cada ojo de forma intermitente una imagen ligeramente distinta del MDT siendo el cerebro el que consolida un relieve de forma virtual y permite localizar las anomalías buscadas ya sea por propio conocimiento del terreno por parte del operador o bien contrastándolo con ortofotomapas, fotos aéreas originales o cualquier otro documento gráfico o mapa de la zona.

Los errores más frecuentes en la generación del MDT son de escasa importancia como pueden ser las discontinuidades en los bordes de los distintos modelos estereoscópicos que han generado el MDT o el retocado de las líneas de borde de los elementos hidrológicos. Errores importantes son la omisión de puntos en determinadas zonas (calvas) por incorrección de las imágenes digitales originales (lagos, zonas brillantes o muy oscuras de las imágenes originales), omisión de puntos en un determinado perfil que se deben corregir utilizando los puntos del entorno y de forma interjectiva, la alternancia de pendiente positiva/negativa en elementos hidrográficos (detección que se puede realizar de forma automática (proceso de imagen) o semiautomática (hidrografía previamente digitalizada) y obtener los perfiles de ellos).

Dado que los errores pueden afectar a un punto aislado o a un grupo de puntos, las herramientas de corrección de estos errores pueden actuar modificando solo el punto en cuestión a través de la información de los puntos que le rodean en una ventana cuyo tamaño es elegido por el operador en cada caso, o bien si el error afecta una zona seleccionando toda la zona en el terminal gráfico y decidir que tipo de corrección hay que aplicar: “subir” o “bajar” el modelo para que se ajuste al terreno que se está apreciando de forma interactiva en el terminal.

Las correcciones sistemáticas se suelen aplicar cuando el error está distribuido por todo el modelo y le afecta por igual. Se aplican técnicas de proceso de imagen del tipo filtrado espacial. Un caso muy común al extraer las curvas de nivel de un MDT es que aparezcan pequeñas curvas de nivel cerradas y aisladas que dan una apariencia de ‘viruela’ a todo el mapa. La eliminación de estas curvas de nivel se realiza aplicando un filtro de suavizado al MDT, lo que no resta rigurosidad al modelo ya que este suavizado elimina elementos poco significativos en un MDT y además suaviza los contornos de las curvas de nivel significativas haciéndolas más agradables a la vista. Otros procesos automáticos también se han descrito [Wong, 1982] para la corrección del modelo con vistas a eliminar las incoherencias entre este y la red hidrográfica.

19.8. Explotación y aplicaciones

Los datos de un modelo digital del terreno se encuentran recogidos bien en forma matricial o ráster o bien en un RIT. Según el caso la extracción de datos derivados del MDT será más fácil o complicada. La explotación, en cualquier caso, se puede plantear como productos derivados directos del MDT y aplicaciones de estos productos.

En un simple análisis de la geometría del MDT (ráster o RIT) podemos deducir cuáles de los productos derivados lo serán directamente o indirectamente. Por otra parte en un primer bloque vamos a tratar la explotación general del MDT y en un segundo bloque citaremos la explotación aplicada de los mismos.

19.8.1 Explotación general

Vamos a tratar en este apartado la derivación de algunos de los principales subproductos que podemos obtener desde un modelo digital del terreno:

Como podemos apreciar de la lista anterior todos estos productos se obtienen a partir de un estudio geométrico del modelo digital del terreno.

Cálculo de pendientes y orientaciones

Un excelente tratado del cálculo de pendientes lo encontramos en Cebrián [Cebrián, A, Mark, D.M.; 1985]. Siguiendo a estos autores encontramos que existen varios métodos simples para el cálculo de las pendientes, todos ellos asumen que las altitudes correspondientes a los vértices de una malla son puntuales y que los cuatro puntos que definen cada elemento unitario de la malla son coplanares en el espacio tridimensional (ver gráfica 9.x).

Definimos un plano que contiene a los cuatro vértices de la malla mediante dos vectores \mathbf{u} y \mathbf{v} cuyas expresiones vienen dadas en función de las altitudes de los cuatro vértices del elemento de la malla tomando como origen local de coordenadas el vértice inferior izquierdo y ordenando las cotas en sentido contrario a las agujas del reloj; nos queda entonces que:

$$\mathbf{u} = [1, 0, ((z_2 - z_1) + (z_3 - z_4))/2]$$

$$\mathbf{v} = [0, 1, ((z_4 - z_1) + (z_3 - z_2))/2]$$

Otra forma de establecer los valores de \mathbf{u} y \mathbf{v} [Kiruchi, *et al.*, 1982] es considerar solo tres vértices del elemento de malla, quedando:

$$\mathbf{u} = [1, 0, (z_2 - z_1)]$$

$$\mathbf{v} = [0, 1, (z_4 - z_1)]$$

Evidentemente esta segunda expresión es más simple de calcular que la primera y no exige que los cuatro vértices del elemento de malla sean coplanares.

En cualquier caso, la ecuación genérica de un plano puede expresarse como $Ax + BY + Cz + D = 0$. El vector de coordenadas (A, B, C) es perpendicular al plano por lo que el plano definido sobre el elemento de malla puede ser representado por su vector unidad perpendicular. Para obtener un vector perpendicular al plano basta obtener el producto vectorial de los vectores \mathbf{u} y

v contenidos en el plano. Si expresamos u y v en forma abreviada (sin perder generalidad) como:

$$\mathbf{u} = [1, 0, z_u]$$

$$\mathbf{v} = [0, 1, z_v]$$

entonces el producto vectorial $\mathbf{u} \times \mathbf{v}$ es:

$$\mathbf{u} \times \mathbf{v} = \begin{pmatrix} \mathbf{u}_x & \mathbf{u}_y & \mathbf{u}_z \\ 1 & 0 & z_u \\ 0 & 1 & z_v \end{pmatrix} = \begin{pmatrix} 0 & z_u \\ 1 & z_v \end{pmatrix} \mathbf{u}_x - \begin{pmatrix} 1 & z_u \\ 0 & z_v \end{pmatrix} \mathbf{u}_y + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \mathbf{u}_z = A\mathbf{u}_x + B\mathbf{u}_y + C\mathbf{u}_z$$

donde: $A = -z_u$; $B = -z_v$; $C = 1$

El vector unidad en esa dirección lo obtendremos dividiendo los valores de A, B y C por el módulo del vector perpendicular a la superficie (A, B, C), con lo que obtendremos, además, los cosenos directores de ese vector α , β y γ , en las respectivas direcciones de los ejes OX, OY, OZ del sistema de referencia.

$$\cos \alpha = A/(A^2 + B^2 + C^2)^{1/2} ; \cos \beta = B/(A^2 + B^2 + C^2)^{1/2} ; \cos \gamma = C/(A^2 + B^2 + C^2)^{1/2}$$

La pendiente es el ángulo entre el plano de referencia (perpendicular al eje vertical OZ) y el plano definido por los vértices del elemento de malla considerados. Este ángulo es el mismo que forman el vector perpendicular unidad y la dirección vertical OZ.

La orientación de esa pendiente también define el aspecto de ese mismo plano y este caso es equivalente al ángulo que forman el vector perpendicular unidad y la dirección del eje OY.

Por tanto la pendiente del plano es el valor del ángulo dado γ , y el aspecto del plano esta dado por β .

$$\text{pendiente del plano} = \arccos \gamma$$

$$\text{aspecto del plano} = \arccos \beta,$$

que también puede expresarse en función de la tangente de la forma

$$\text{aspecto del plano} = \arctg (\alpha/\beta)$$

Un estudio más profundo de la geometría del modelo digital del terreno se debe a Monmonier [1982] el cual propone una solución diferente basada en el cálculo de las derivadas parciales de la función lineal $z=f(x, y)$ a la que se supone pertenecen los cuatro vértices del elemento del plano. Este análisis es bastante más riguroso que el anterior y obtiene los mismos resultados

para el análisis de pendientes y orientaciones, permitiendo además el estudio de convexidad y concavidad de superficies compuestas por varios elementos de malla.

La pendiente y curvatura de un punto están asociadas con la primera y segunda coordenadas de la función del terreno $z=f(x, y)$. Usando el formalismo de la geometría diferencial podemos evaluar las componentes del vector gradiente o su magnitud y el ángulo de dirección del mismo. La curvatura en un punto se puede definir por el radio principal de curvatura de la superficie o por su función. Las derivadas de primer orden $\partial z/\partial x$ y $\partial z/\partial y$ se utilizan para dar directamente la pendiente a partir de la expresión:

$$p = \sqrt{(\partial z / \partial x)^2 + (\partial z / \partial y)^2}$$

que nos da el valor de la pendiente p , y

$$d = \arctg\left(\frac{\partial z / \partial y}{\partial z / \partial x}\right)$$

Que nos el valor de la dirección del ángulo de pendiente (orientación).

La curvatura en un punto se evalúa a partir de expresiones mas complicadas, en las que intervienen, en general, las derivadas de segundo orden $\partial^2 z/\partial x^2$, $\partial^2 z/\partial x\partial y$ y $\partial^2 z/\partial y^2$. Si C_M y C_T son las curvaturas medias y total de la superficie, podemos obtener estas a partir de las expresiones [Thomas, 1952]:

$$C_M = \frac{1}{R_1} + \frac{1}{R_2} = \frac{E \cdot D'' + G \cdot D - 2F \cdot D'}{E \cdot G - F^2}$$

$$C_T = \frac{1}{R_1 \cdot R_2} = \frac{D \cdot D'' - D'^2}{E \cdot G - F^2}$$

Donde R_1 y R_2 son los radios principales de curvatura; y E , F , G y D , D' , D'' son respectivamente las cantidades fundamentales de primer y segundo orden de la superficie.

Otra medida de la curvatura de la superficie que es mucho más fácil de evaluar es el Laplaciano:

$$C_L = \partial^2 z/\partial x^2 + \partial^2 z/\partial y^2$$

La diferencia entre C_L y C_M es una función de la derivada de primer orden, y para terrenos con pendientes de hasta el 40% ($p \leq 0.4$) la diferencia es menor del 4%.

Entonces para evaluar p , d y C_L en un punto, necesitamos calcular la primera y segunda derivada del terreno en ese punto. Para ello hemos de seleccionar una “superficie móvil”, a

modo de ventana, de puntos de la malla, sobre la cual ajustaremos una función matemática, en general un polinomio de segundo o tercer grado $z=z(x, y)$. La ventana puede ser de 3x3 (9 puntos), o 5x5 (25 puntos) centrada en el punto a calcular. La forma general de tal polinomio es:

$$z_u = z_u(x, y) = \sum_{q=0}^u \sum_{p=0}^{u-q} a_{pq} \cdot x^p \cdot y^q$$

Donde u es el grado del polinomio y x, y son las coordenadas cartesianas del punto, en el subconjunto de puntos de malla que están en la ventana elegida, respecto al punto central. La determinación de los coeficientes a_{pq} se realiza ajustando por mínimos cuadrados los polinomios dado que hay mas puntos en el subconjunto que coeficientes son necesarios, aunque esto depende del orden del polinomio, pero como el lógico a mayor orden del polinomio la superficie $z=z(x,y)$ a construir es mas complicada y se necesitaran mas puntos, luego una ventana de tamaño superior.

Usando una terminología matemática convencional, podemos decir que el MDT que estamos manejando hasta ahora esta representado en el dominio espacial, llamémosle MDT-Z, dado que esta construido en la variable z . Sin embargo existe otro dominio de representación de los fenómenos en el que se hacen con gran facilidad operaciones que en el dominio espacial resultan muy complejas y laboriosas. Este dominio se le llama dominio de las frecuencias espaciales y en él z esta definido como la suma de un conjunto de superficies armónicas; llamémosle MDT-G, dado que en él las características geométricas se representan con gran claridad.

Existen transformaciones matemáticas directas e inversas entre uno y otro dominio o representación, el paso entre ellas se realiza a través de la transformada discreta de Fourier, y en general se implementan a través del algoritmo de la Transformada Rápida de Fourier (FFT - Fast Fourier Transform). En este dominio la representación del modelo digital del relieve no se hace con los valores de las cotas, z , de los puntos de la malla, sino se obtienen con la FFT los valores de los coeficientes de las funciones armónicas que lo definen.

En estos coeficientes se hallan recogidas las características geométricas del MDT y basta seleccionar un conjunto de valores de estos coeficientes para obtener las líneas de nivel, o a través de un intervalo de coeficientes seleccionar unos determinados accidentes geomorfológicos (fallas, quebradas, alineaciones, etc.); o por eliminación de algunos valores, filtrar el MDT-G, eliminando la rugosidad del terreno (filtro de paso bajo que suprime las altas frecuencias), lo que permite que el posterior curvado del MDT nos dé curvas de nivel más suaves y agradable a la vista.

Cálculo de sombras

Según la figura 9.x, podemos apreciar que para que un elemento del modelo digital del terreno se encuentre en sombra al ser iluminado por una fuente de radiación, por ejemplo el Sol, este debe tal que el ángulo α formado por el vector fuente-elemento, p , y el vector normal al elemento, n (obtenido del mapa de pendientes) ha de ser mayor de 90° . Si el ángulo es de 90° la incidencia del rayo es rasante y si es de menor de 90° el elemento se encuentra iluminado.

Normalmente los elementos del modelo digital del terreno que se encuentran iluminados se suelen representar con una gama de valores (generalmente codificados en niveles de gris) para destacar los que se encuentran mas iluminados por la fuente de radiación de los que se encuentran menos, así se daría un nivel de reflectancia máximo a aquellos elementos que tienen en ángulo α con valor 0° y mínimo a los que tienen α de valor 90° (incidencia rasante). Entre ambos extremos se distribuye el rango de los 90° en 256 niveles de gris que es lo que permite el codificar el sombreado en un byte (8 bits). Los elementos en sombra se representarían en negro con el valor mínimo (cero).

Cuando la fuente de iluminación es del tipo radar (sensor activo) el aspecto de la imagen al ser visualizada es como si ya tuviera incluido el sombreado. Esto es debido a la naturaleza de la imagen radar que nos da la reflectancia del suelo en el rango de las microondas y es muy sensible al relieve del mismo, presentando zonas oscuras cuando la incidencia sobre el elemento de superficie (píxel) es normal ($\alpha=0^\circ$) y zonas brillantes cuando es perpendicular ($\alpha=90^\circ$), de forma inversa a cuando la fuente de iluminación esta en el rango óptico.

Bibliografía

- [1] Bourrough, P. A. "Principles of Geographical Information System for Land Resources Assesment" 1986, OXFORD SCIENCE PUBLICATIONS (UK) 1ª edición.

Tema 20. Comunicaciones, redes e Internet. Medios de transmisión y codificación de datos. Modelo de referencia OSI. Familia de protocolos TCP/IP. Redes de área personal, PAN. Redes de área local, LAN. Redes de área extensa, WAN. Internet.

20.1. Comunicaciones, redes e Internet

Una red de computadora (también llamada red de ordenadores o red informática) es un conjunto de computadoras y/o dispositivos conectados por enlaces, a través de medios físicos (medios guiados) o inalámbricos (medios no guiados) y que comparten información (archivos), recursos (CD-ROM, impresoras, etc.) y servicios (e-mail, antivirus, ftp), etc.

La comunicación por medio de una red se lleva a cabo en dos diferentes categorías:

- Capa física
- Capa lógica.

La capa física incluye todos los elementos de los que hace uso un equipo para comunicarse con otros equipos dentro de la red, como, por ejemplo, las tarjetas de red, los cables, las antenas, etc.

La comunicación a través de la capa física se rige por normas muy rudimentarias que por sí mismas resultan de escasa utilidad. Sin embargo, haciendo uso de dichas normas es posible construir los denominados protocolos, que son normas de comunicación más complejas (mejor conocidas como de alto nivel), capaces de proporcionar servicios.

Los protocolos son un concepto muy similar al de los idiomas de las personas. Si dos personas hablan el mismo idioma, es posible comunicarse y transmitir ideas.

20.1.1. Clasificación de redes

Existen muchas formas de clasificar las redes. A continuación se detallan tres clasificaciones que serán útiles a lo largo de este tema.

Por extensión geográfica:

- Red de área personal (PAN)

- Red de área local (LAN)
- Red de área metropolitana (MAN)
- Red de área amplia (WAN)

Por estructura:

- Red OSI o Modelo OSI
- Red TCP/IP

Por topología de la red

- En bus
- En Estrella
- En Anillo
- Malla

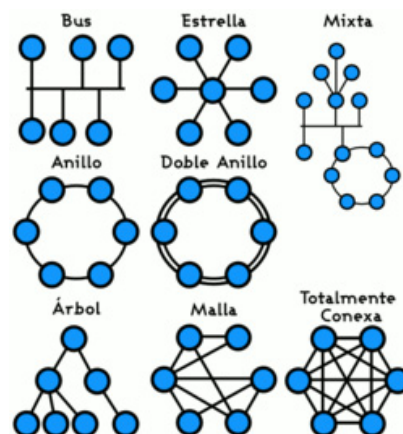


Figura 1. Topologías de redes. Fuente: Wikipedia

Por su alcance

- Intranet: Red interna que se limita en alcance a una sola organización o entidad y que utiliza la familia de protocolos TCP/IP y software de red de uso general.
- Extranet: Red interna que se limita en alcance a una sola organización o entidad pero que también permiten conexiones limitadas a determinados contenidos, no necesariamente a organizaciones o entidades de confianza (por ejemplo, los clientes de una compañía pueden tener permitido el acceso a una cierta parte de su Intranet).

- Internet: Interconexión mundial de las redes gubernamentales, académicas, públicas, y privadas basadas en el *Advanced Research Projects Agency Network* (ARPANET) desarrollado por ARPA del departamento de defensa de los EE.UU.

20.1.2. Componentes básicos de una red

En este epígrafe se presentan los elementos fundamentales de una red de ordenadores.

- Estación de trabajo: Los ordenadores que toman el papel de estaciones de trabajo aprovechan o tienen a su disposición los recursos que ofrece la red así como los servicios que proporcionan los Servidores a los cuales pueden acceder.
- Servidor: El servidor es aquel o aquellos ordenadores que van a compartir sus recursos hardware y software con los demás equipos de la red.
- Tarjeta de red: También se denominan NIC (Network Interface Card). Básicamente realiza la función de intermediario entre el ordenador y la red de comunicación. En ella se encuentran grabados los protocolos de comunicación de la red.
- Bridges, Switches y Routers: Los servidores y las estaciones de trabajo son considerados como los extremos de la comunicación. En una red existen otros elementos intermedios que hacen fluir la información entre estos extremos. A estos elementos les diferencia la “inteligencia” que proporcionan a la comunicación.
- Gateways o pasarelas: Es un hardware y software que proporciona la interconexión entre diferentes redes, como por ejemplo la conexión de una red LAN interna con la red de Internet. Su función es la de adaptar los distintos protocolos de las redes.
- El medio: Constituido por el cableado y los conectores que enlazan los componentes de la red. Los medios físicos más utilizados son el cable de par trenzado, par de cable, cable coaxial y la fibra óptica (cada vez en más uso esta última).

20.2. Medios de transmisión y codificación de datos

En el campo de las telecomunicaciones, el medio de transmisión constituye el soporte físico a través del cuál emisor y receptor pueden comunicarse en un sistema de transmisión.

Los medios de transmisión pueden ser guiados y no guiados. En ambos casos la transmisión se realiza por medio de ondas electromagnéticas.

En un medio guiado las ondas son conducidas (guiadas) a través de un camino físico, mientras que en uno no guiado el medio solo proporciona un soporte para que las ondas se transmitan, pero no las guía.

Como ejemplo de medios guiados pueden ser mencionados el cable coaxial, la fibra óptica y los cables de pares como el del teléfono doméstico.

Entre los no guiados existen el aire y el vacío.

Dependiendo de la naturaleza del medio, las características y la calidad de transmisión se verán limitadas de forma distinta. Así en un medio guiado será de éste del que dependerán, principalmente, la velocidad de transmisión, el ancho de banda y el espaciado entre repetidores. Sin embargo, en el caso de un medio no guiado resulta más determinante el espectro de frecuencias de la señal transmitida que el propio medio de transmisión en sí mismo.

20.2.1. Cable coaxial

El cable coaxial es un cable formado por dos conductores concéntricos:

- Un conductor central o núcleo, formado por un hilo sólido o trenzado de cobre.
- Un conductor exterior en forma de tubo o vaina, y formado por una malla trenzada de cobre o aluminio o bien por un tubo, en caso de cables semirrígidos. Este conductor exterior produce un efecto de blindaje y además sirve como retorno de las corrientes.
- El primero está separado del segundo por una capa aislante llamada dieléctrico. De la calidad del dieléctrico dependerá principalmente la calidad del cable.

El cable coaxial no es habitualmente afectado por interferencias externas, y es capaz de lograr altas velocidades de transmisión en largas distancias. Por esa razón, se utiliza en redes de comunicación de banda ancha (cable de televisión) y cables de banda base (Ethernet).

El cable coaxial se reemplaza por la fibra óptica en distancias superiores a varios kilómetros, porque el ancho de banda de esta última es muy superior, lo que justifica su mayor costo y su instalación más delicada.

20.2.2. Fibra óptica

La fibra óptica es un conductor de ondas en forma de filamento, generalmente de vidrio, aunque también puede ser de materiales plásticos. La fibra óptica es capaz de dirigir la luz a lo largo de su longitud usando la reflexión total interna. Normalmente la luz es emitida por un láser o un LED.

Las fibras son ampliamente utilizadas en telecomunicaciones, ya que permiten enviar gran cantidad de datos a gran velocidad, mayor que las comunicaciones de radio y cable. También se utilizan para redes locales. Son el medio de transmisión inmune a las interferencias por excelencia. Tienen un costo elevado.

20.2.3. Cable de pares

Un cable de pares es el formado por grupos de 2 hilos (par) de material conductor recubiertos de plástico protector.

El cable multipar es aquel formado por un elevado número de pares de cobre, generalmente múltiplo de 25. Existen cables multipares normalizados con capacidad de 25, 50, 125, 250 y hasta 3600 pares en un único cable físico.

Los cables de pares son usados para la conexión física de equipos de telefonía, en redes de datos, como por ejemplo en redes LAN. En estas redes de datos se utilizan pares de cobre trenzados (UTP), donde los conductores se “trenzan” entre sí, y apantallados, es decir cubiertos de una pantalla o malla de material conductor. Estas mejoras permiten la transmisión de datos a capacidades altas y minimizan interferencias hacia/desde otros sistemas.

20.2.4 Codificación de datos

Se entiende por codificación de datos las técnicas que se emplean para mandar un tipo de datos utilizando señales analógicas o digitales a través de un medio de transmisión.

La información puede transmitirse tal como es (banda base) o puede modificarse (modulación) para obtener una serie de ventajas a la hora de enviarla por el medio de transmisión. En este

último caso, sería necesario volver a dejar la señal en su estado original para interpretarla (demodulación).

En los sistemas de transmisión, la banda base (*señal original*) es generalmente utilizada para modular otra señal (*señal portadora*).

Las frecuencias de banda base se caracterizan por ser generalmente mucho más bajas que las resultantes cuando éstas se utilizan para modular una portadora. Por ejemplo, es señal de banda base la obtenida de la salida de video compuesto de dispositivos como grabadores/reproductores de video y consolas de juego, a diferencia de las señales de televisión que deben ser moduladas para poder transportarlas vía aérea (por señal libre o satélite) o por cable.

Las técnicas más habituales de codificación, en función de la naturaleza de los datos que se desean transmitir son las siguientes:

a) Datos analógicos:

- AM. Amplitud modulada, o modulación de amplitud es un tipo de modulación que consiste en hacer variar la amplitud de la onda portadora de forma que esta cambie de acuerdo con las variaciones de nivel de la señal moduladora, que es la información que se va a transmitir.
- PM. Fase modulada. Es el caso de modulación donde tanto las señales de transmisión como las señales de datos son analógicas. Es un tipo de modulación exponencial al igual que la modulación de frecuencia. Se caracteriza porque la fase de la onda portadora varía directamente de acuerdo con la señal modulante, resultando una señal de modulación en fase.

b) Datos digitales

- ASK. Modulación por desplazamiento de amplitud, en inglés Amplitude-shift keying, es una forma de modulación en la cual se representan los datos digitales como variaciones de amplitud de la onda portadora tomando estas variaciones un conjunto discreto de valores. El nivel de amplitud puede ser usado para representar los valores binarios 0s y 1s. Podemos pensar en la señal portadora como un interruptor ON/OFF. Tanto los procesos de modulación ASK como los procesos de demodulación son relativamente baratos. La técnica ASK también es usada comúnmente para transmitir datos digitales sobre la fibra óptica.

- PSK. La modulación por desplazamiento de fase o Phase Shift Keying es una forma de modulación angular consistente en hacer variar la fase de la portadora entre un número de valores discretos.
- FSK. La Modulación por desplazamiento de frecuencia o Frequency Shift Keying es la más simple de las modulaciones digitales. Es similar a la modulación de frecuencia (FM), pero más sencillo, dado que la señal moduladora es un tren de pulsos binarios que solo varía entre dos valores de tensión discretos. Así, por ejemplo, un 0 binario se puede representar con una frecuencia f_1 , y el 1 binario se representa con una frecuencia distinta f_2 .

Una técnica de codificación especial es la denominada Modulación por Impulsos Codificados (MIC) en inglés Pulse Code Modulation (PCM), que es un procedimiento de modulación utilizado para transformar una señal analógica en una secuencia de bits, formando así una señal digital.

20.3. Modelo de referencia OSI

20.3.1. Introducción

Por mucho tiempo se consideró al diseño de redes un proceso muy complicado de llevar a cabo, esto es debido a que los fabricantes de ordenadores tenían su propia arquitectura de red, y ésta era muy distinta al resto, y en ningún caso existía compatibilidad entre marcas. De la misma forma en que las personas que no hablan un mismo idioma tienen dificultades para comunicarse, las redes que utilizaban diferentes especificaciones e implementaciones tenían dificultades para intercambiar información.

Para tal fin, la ISO (Organización Internacional de Normalización) en 1977 desarrolló una estructura de normas comunes dentro de las redes, que se conocen como el Modelo de Referencia OSI (interconexión de sistemas abiertos), modelo bajo el cual empezaron a fabricar computadoras con capacidad de comunicarse con otras marcas.

20.3.2. Principales características

Este modelo se basa en el principio de "divide y vencerás", de forma que las redes se diseñan como una secuencia de capas, cada una construida sobre la anterior.

Las características del modelo OSI son las siguientes:

1. Estructura multinivel: Se diseña una estructura multinivel con la idea de que cada nivel resuelva solo una parte del problema de la comunicación, con funciones específicas.
2. El nivel superior utiliza los servicios de los niveles inferiores: Cada nivel se comunica con su homólogo en las otras máquinas, usando un mensaje a través de los niveles inferiores de la misma. La comunicación entre niveles se define de manera que un nivel N utilice los servicios del nivel N-1 y proporcione servicios al nivel N+1.
3. Dependencia de Niveles: Cada nivel es dependiente exclusivamente del nivel inferior y del nivel superior.
4. Cabeceras: En cada nivel, se incorpora al mensaje una cabecera con la información necesaria para realizar sus funciones. Por esta razón se considera que un mensaje está constituido de dos partes, el encabezado y la información. Sin embargo, como la computadora receptora retira los encabezados en orden inverso a como se enviaron desde la computadora emisora, el mensaje original no se afecta.

Las capas se pueden dividir en dos grupos:

1. Servicios de transporte (niveles 1, 2, 3 y 4).
2. Servicios de soporte al usuario (niveles 5, 6 y 7).

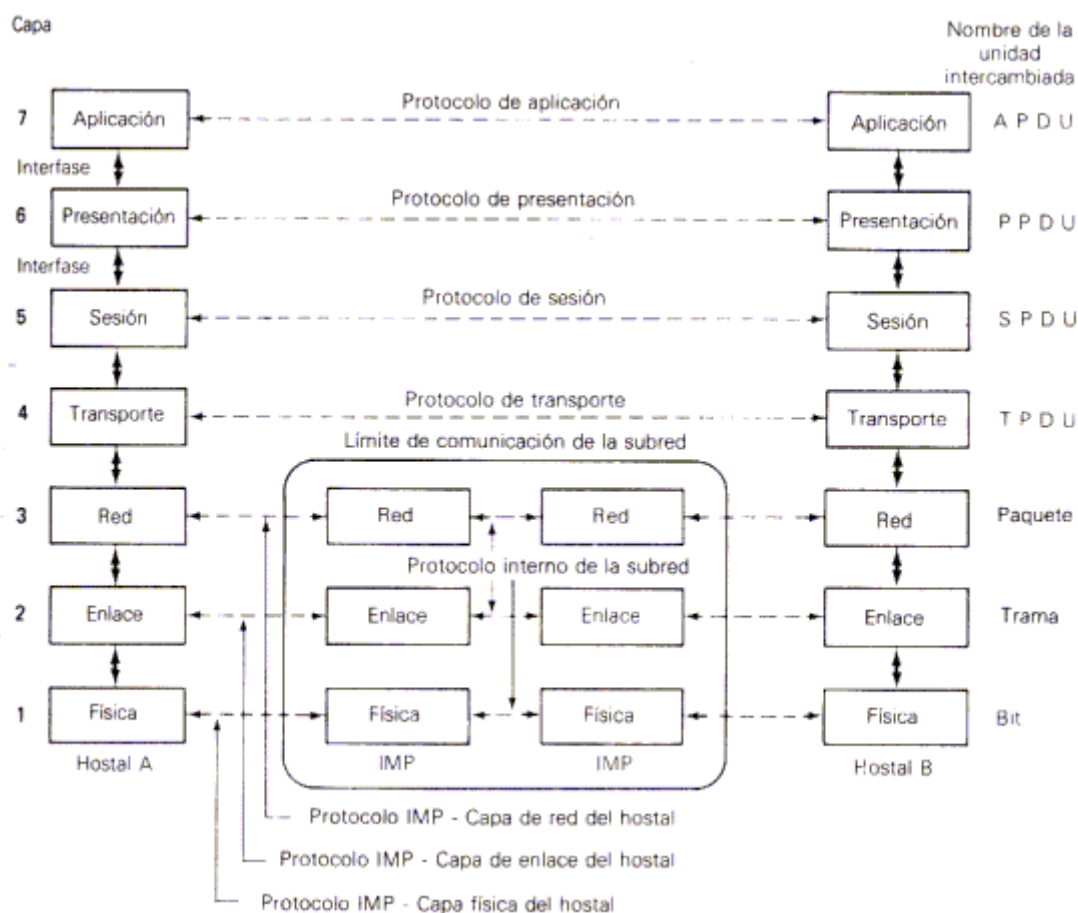


Figura 2. Arquitectura de red del modelo OSI de ISO.

En esta figura se representan, a modo de esquema explicativo, cuatro nodos de la comunicación: los dos extremos correspondientes a los cuadros de los laterales y dos nodos intermedios correspondientes a los cuadros del medio. Estos nodos intermedios son los que conforman la red que interconecta a los dos extremos.

Como puede apreciarse, sólo los nodos extremos tienen todos los niveles de la arquitectura, mientras que los nodos intermedios sólo proporcionan los tres primeros niveles, que permiten encaminar la información entre los dos extremos.

Cada una de las capas libera a la posterior del conocimiento de las funciones subyacentes. Esto hace necesario establecer interfaces de comunicación entre capas que definen los servicios y operaciones. Una capa-*i* de una máquina A establece comunicación con la capa-*i* de una máquina B, a través de unas determinadas reglas y convenciones para llevarla a cabo, lo cual se denomina protocolo de la capa-*i*.

A la configuración de capas y protocolo se le llama *arquitectura de red*.

20.3.3. Capas del modelo OSI

1. Capa Física: Aquí se encuentran los medios materiales para la comunicación como las placas, cables, conectores, es decir los medios mecánicos y eléctricos.
2. Capa de Enlace: Elimina los errores de la transmisión.
3. Capa de Red: Encamina cada paquete a su destinatario a través de encaminadores (routers) intermedios.
4. Capa de Transporte: La función principal es la de aceptar los datos de la capa superior y dividirlos en unidades más pequeñas, para pasarlos a la capa de red, asegurando que todos los segmentos lleguen correctamente, esto debe ser independiente del hardware en el que se encuentre.
5. Capa de Sesión: Permite a los usuarios coordinarse entre sí permitiendo acceder a un sistema de tiempo compartido a distancia, o transferir un archivo entre dos máquinas.
6. Capa de Presentación: Se ocupa de los aspectos de sintaxis y semántica de la información que se transmite, por ejemplo la codificación de datos según un acuerdo.
7. Capa de Aplicación: Por nivel de aplicación se entiende el programa o conjunto de programas que generan una información para que ésta viaje por la red (Por ejemplo, el correo electrónico).

El modelo, junto con las definiciones de servicio y protocolos asociados, es muy complejo. Cuando se apilan normas impresas, ocupan una fracción significativa de un metro de papel. También son difíciles de realizar e ineficientes en su aspecto operativo. Pero a pesar de estos inconvenientes, tiene gran utilidad como modelo de referencia para el análisis de otros modelos.

20.4. Familia de protocolos TCP/IP

La familia de protocolos de Internet es un conjunto de protocolos de red en los que se basa Internet y que permiten la transmisión de datos entre redes de computadoras. En ocasiones se la denomina conjunto de protocolos TCP/IP, en referencia a los dos protocolos más importantes que la componen: Protocolo de Control de Transmisión (TCP) y Protocolo de Internet (IP), que fueron los dos primeros en definirse, y que son los más utilizados de la familia.

Existen tantos protocolos en este conjunto que llegan a ser más de 100 diferentes, entre ellos se encuentra el popular HTTP (*HyperText Transfer Protocol*), que es el que se utiliza para acceder a las páginas web, además de otros como el ARP (*Address Resolution Protocol*) para la resolución de direcciones, el FTP (*File Transfer Protocol*) para transferencia de archivos, y el SMTP (*Simple Mail Transfer Protocol*) y el POP (*Post Office Protocol*) para correo electrónico, TELNET para acceder a equipos remotos, entre otros.

El TCP/IP es la base de Internet, y sirve para enlazar computadoras que utilizan diferentes sistemas operativos, incluyendo PCs, minicomputadoras y computadoras centrales sobre redes de área local (LAN) y área extensa (WAN). TCP/IP fue desarrollado y demostrado por primera vez en 1972 por el departamento de defensa de los Estados Unidos, ejecutándolo en ARPANET, una red de área extensa del departamento de defensa.

La familia de protocolos de Internet puede describirse por analogía con el modelo OSI, que describe los niveles o capas de la pila de protocolos, aunque en la práctica no corresponde exactamente con el modelo en Internet.

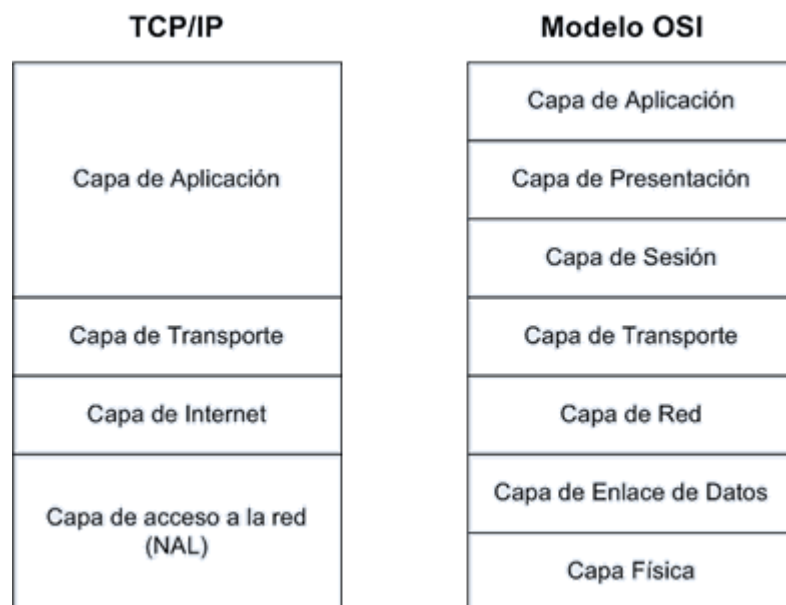


Figura 3. Comparación del modelo TCP/IP con el modelo de referencia OSI de ISO.

La diferencia más notable entre los modelos de TCP/IP y OSI es el nivel de Aplicación, en TCP/IP se integran algunos niveles del modelo OSI en su nivel de Aplicación.

Una relación simplificada de ejemplos de protocolos de la familia TCP/IP se muestra a continuación:

5	Aplicación	Ej. HTTP, FTP, DNS (protocolos de enrutamiento como BGP y RIP, que por varias razones funcionen sobre TCP y UDP respectivamente, son considerados parte del nivel de red)
4	Transporte	Ej. TCP, UDP (protocolos de enrutamiento como OSPF, que funcionen sobre IP, son considerados parte del nivel de red)
3	Internet	Para TCP/IP este es el Protocolo de Internet (IP) (protocolos requeridos como ICMP e IGMP funcionan sobre IP, pero todavía se pueden considerar parte del nivel de red; ARP no funciona sobre IP)
2	Enlace	Ej. Ethernet, PPP, Frame Relay, RDSI, ATM, IEEE 802.11, FDDI
1	Físico	Ej. medio físico, y técnicas de codificación, T1, E1

Tabla 1. Ejemplos de protocolos TCP/IP

20.5. Red de área personal (PAN)

Una red de área personal (PAN) es una red de ordenadores usada para la comunicación entre los dispositivos de la computadora (teléfonos incluyendo las PDAs o Asistentes Digitales Personales) en un alcance reducido, típicamente algunos metros.

Las PAN se pueden utilizar para la comunicación entre los dispositivos personales, o para conectar con una red de alto nivel e Internet.

Se pueden conectar con cables a los buses de la computadora tales como USB y FireWire. Una red de área personal inalámbrica (WPAN) se puede realizar con tecnologías de red tales como IrDA (infrarrojos) y radiofrecuencia, como la solución más conocida en el mercado *Bluetooth*.

20.6. Redes de área local LAN

Una red de área local, o red local, es la interconexión de varios ordenadores y periféricos. LAN es la abreviatura inglesa de *Local Area Network* (red de área local). Su extensión está limitada físicamente a un edificio o a un entorno de pocos kilómetros. Su aplicación más extendida es la interconexión de ordenadores personales y estaciones de trabajo en oficinas, fábricas, etc., para compartir recursos e intercambiar datos y aplicaciones. En definitiva, permite que dos o más máquinas se comuniquen.

Las características fundamentales de una LAN son las siguientes:

- Medio de transmisión compartido.
- Cableado específico instalado normalmente a propósito.
- Capacidad de transmisión comprendida entre 1 Mbps y 10 Gbps.
- Extensión máxima no superior a 3 km
- Uso de un medio de comunicación privado.
- La facilidad con que se pueden efectuar cambios en el hardware y el software.
- Gran variedad y número de dispositivos conectados.
- Posibilidad de conexión con otras redes.

Algunos de estos parámetros son generales y varían con el tiempo de forma vertiginosa al aparecer nuevas tecnologías que elevan los umbrales antes mencionados. Este es el caso de la fibra óptica, que permite transmisiones a grandes distancias con anchos de banda de varios Gigabits por segundo.

La oferta de redes de área local es muy amplia, existiendo soluciones casi para cualquier circunstancia. Podemos seleccionar el tipo de cable, la topología e incluso el tipo de transmisión que más se adapte a nuestras necesidades. Sin embargo, de toda esta oferta la solución más extendida es *Ethernet*.

El otro tipo de red local, muy utilizado actualmente, es la tecnología inalámbrica o Wireless LAN (WLAN), que permite establecer una red local sin cables. Existe un estándar muy

extendido en el mercado, el 802.11 o WiFi (*Wireless Fidelity*), de la organización IEEE (*The Institute of Electrical and Electronics Engineers*)

Esta tecnología es muy útil a la hora de instalar la red pero conlleva un inconveniente inherente a esta tecnología, la seguridad. Esto se debe a que cualquier equipo con un receptor inalámbrico puede recibir la información de los equipos que se encuentran en su radio de cobertura. Para ofrecer seguridad, se emplean técnicas de encriptación de la información.

Además las redes inalámbricas ofrecen velocidades de acceso inferiores a las tecnologías por cable, siendo la más común de 54 Mbps (802.11b).

20.7. Red de área extensa WAN

Una Red de Área Extensa (Wide Area Network o WAN, del inglés), es un tipo de red de computadoras capaz de cubrir distancias desde unos 100 hasta unos 1000 km, dando el servicio a un país o un continente. Un ejemplo de este tipo de redes sería RedIRIS, Internet o cualquier red en la cual no estén en un mismo edificio todos sus miembros (sobre la distancia hay discusión posible). Muchas WAN son construidas por y para una organización o empresa particular y son de uso privado, otras son construidas por los proveedores de Internet (ISP) para proveer de conexión a sus clientes.

Hoy en día Internet proporciona WAN de alta velocidad, y la necesidad de redes privadas WAN se ha reducido drásticamente mientras que las VPN (Redes Privadas Virtuales) que utilizan cifrado y otras técnicas para hacer esa red (Internet) privada, aumentan continuamente.

Normalmente la WAN es una red punto a punto, es decir, red de paquetes conmutados como GPRS (Servicio general de paquetes por radio) y UMTS (Sistema Universal de Telecomunicaciones Móviles). Las redes WAN pueden usar sistemas de comunicación vía satélite o de radio.

Cuentan con una infraestructura basada en poderosos nodos de conmutación que llevan a cabo la interconexión de los elementos, por los que además fluyen un volumen apreciable de información de manera continua. Por esta razón también se dice que las redes WAN tienen carácter público, pues el tráfico de información que por ellas circula proviene de diferentes lugares, siendo usada por numerosos usuarios de diferentes países del mundo para transmitir información de un lugar a otro. A diferencia de las redes LAN, la velocidad a la que circulan los datos por las redes WAN suele ser menor que la que se puede alcanzar en las redes LAN.

Además, las redes LAN tienen carácter privado, pues su uso está restringido normalmente a los usuarios miembros de una empresa, o institución, para los cuales se diseñó la red.

La infraestructura de redes WAN la componen, además de los nodos de conmutación, líneas de transmisión de grandes prestaciones, caracterizadas por sus grandes velocidades y ancho de banda en la mayoría de los casos. Las líneas de transmisión (también llamadas "circuitos", "canales" o "troncales") mueven información entre los diferentes nodos que componen la red.

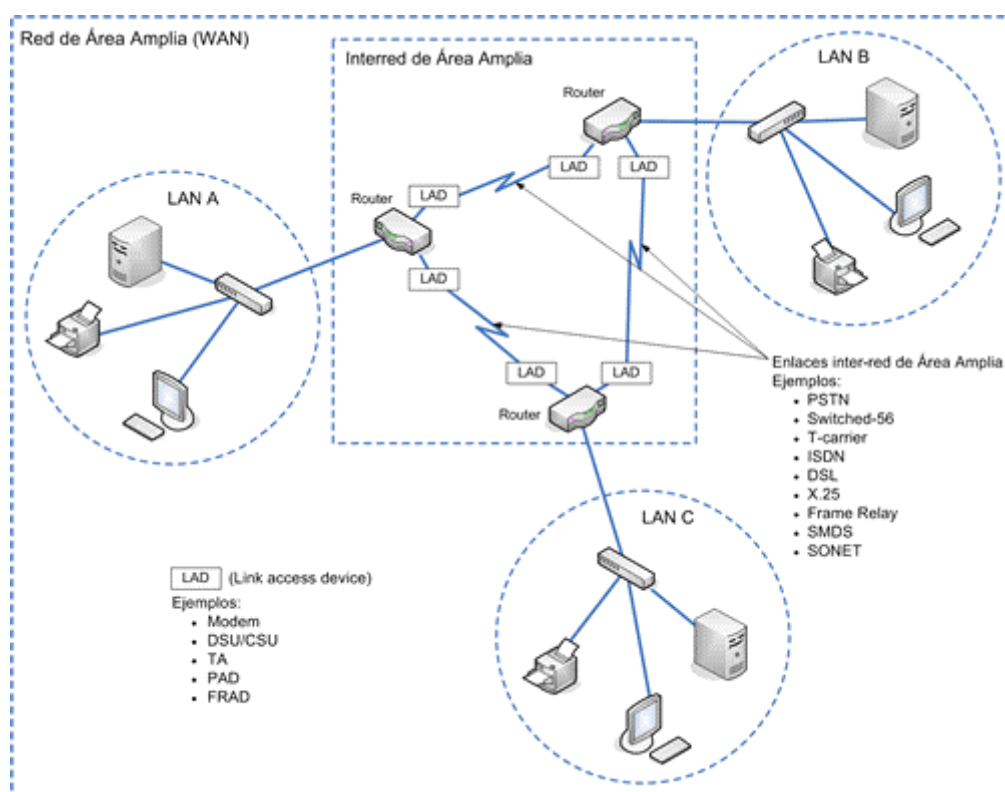


Figura 4. Arquitectura de una red WAN. Fuente: www.textoscientificos.com.

Las características fundamentales de las redes WAN son:

- Cubren una región, país o continente.
- Generalmente se dividen en subredes intercomunicadas.
- Conectan múltiples LAN.
- División entre líneas de transmisión y elementos de conmutación (enrutadores).
- Usualmente los routers son computadoras de las subredes que componen la WAN.

Las soluciones tecnológicas que existen en el mercado son muy abundantes y variadas. A continuación se exponen las más utilizadas.

Las que usan cable como medio físico:

- ADSL: Consiste en una línea digital de alta velocidad, apoyada en el par simétrico de cobre que lleva la línea telefónica convencional o línea de abonado. Es una tecnología de acceso a Internet de banda ancha, que parte de 8 Mbps teóricos máximos, aunque actualmente están apareciendo evoluciones que aumentan este umbral.
- RDSI: Red Digital de Servicios Integrados, es una red que facilita conexiones digitales extremo a extremo para proporcionar una amplia gama de servicios, tanto de voz como de otros tipos, y a la que los usuarios acceden a través de un conjunto de interfaces normalizados. Es una red que está siendo relegada por otras tecnologías como ADSL.
- Frame Relay: La técnica Frame Relay se utiliza para un servicio de transmisión de voz y datos a alta velocidad que permite la interconexión de redes de área local separadas geográficamente. Esta tecnología está siendo relegada por otras como ADSL de menor coste.
- ATM: El Modo de Transferencia Asíncrona es una tecnología de telecomunicación desarrollada para hacer frente a la gran demanda de capacidad de transmisión para servicios y aplicaciones. Se suele emplear para dar soporte a tecnologías como ADSL. Actualmente está siendo desplazada por otras tecnologías como Ethernet.
- Fibre Channel: El Canal de fibra, es una tecnología de red a velocidad de gigabit principalmente utilizada para redes de almacenamiento. A pesar de su nombre, la señalización del Canal de Fibra puede funcionar tanto sobre pares de cobre, como sobre cables de fibra óptica.

Las tecnologías inalámbricas:

- WiMax: Es un estándar de transmisión inalámbrica de datos (IEEE 802.16) que proporciona accesos concurrentes en áreas de hasta 48 km de radio y a velocidades de hasta 70 Mbps, utilizando tecnología que no requiere visión directa con las estaciones base. WiMax es un concepto parecido a Wi-Fi (Wireless Fidelity), pero con mayor cobertura y ancho de banda.

- GPRS/EDGE//UMTS: Tecnologías de telefonía móvil celular. Proporcionan anchos de banda entre 110 Kbps (GPRS) y 2 Mbps (UMTS).
- Otras tecnologías como comunicaciones vía satélite.

Bibliografía

- [1] Boehm, Barry W. IEEE Software. Defense Advanced Research Projects Agency. “Software Risk Management: Principles and Practices”.
- [2] Groth, David; Toby Skandier (2005). “Guía del estudio de redes”. 4ª Edición”.
- [3] Temario de la Asociación Profesional del cuerpo superior de Sistemas y Tecnologías de la Administración del Estado. Edición 2007.

Wikipedia