



Spatial Information Management

Location is an inherent part of business data: organizations maintain customer address lists, own property, ship goods from and to warehouses, manage transport flows among their workforce, and perform many other activities. A majority of these activities entail managing locations of different types of entities, including customers, property, goods, and employees. Those locations need not be static—in fact, they may continually change over time. For instance, goods are manufactured, packaged, and channeled to warehouses and retail/customer destinations. They may have different locations at various stages of the distribution network.

Let's consider an example of parcel services to illustrate how location is used. We have become increasingly accustomed to monitoring the status of parcel deliveries on the Web by locating our shipment within the distribution channel of our chosen service supplier. The simplicity and usefulness of this service is the result of a very complex underlying information system. The system relies on the ability to locate the parcel as it moves across different stages of the distribution network. Many information systems share location information in this process, which can be used to estimate, for instance, transit or delivery times. Systems such as RFID¹ are used to automatically record the movements of parcels along the distribution chain. Aircraft, trains, container ships, or trucks that move goods between distribution hubs use systems such as Global Positioning System (GPS) to locate their positions in real time. Even the “last mile”—that is, the delivery of an individual parcel to the end customer—is based on the geographical optimization of the delivery schedule as well as on the ability to locate the truck drivers in real time, to guide them to their destinations, and to estimate delivery times.

All of this location information is stored, analyzed, and exchanged between multiple systems and is the basis for making the entire operation cheaper, faster, and more reliable. Most of these systems are connected to each other through the Internet. The end user also uses the Internet to access the system and to query the current status of his parcel. By analyzing the system in its entirety, you can recognize that the added value is the result of the integration of various systems, of their interoperability, and of the pervasive role of spatial information across the entire process. Spatial information plays a crucial role in enabling the systems and processes to run smoothly and efficiently.

This example illustrates the pervasiveness of location or spatial information in day-to-day business. In fact, market research estimates that the majority of the data handled by organizations—perhaps as much as 80 percent of all data—has a spatial dimension.² The ability to properly manage the “where,” or the spatial information, is key to the efficiency of organizations and could translate to substantial costs savings and commercial competitiveness. For instance, healthcare, telecommunications, and local government organizations depend on spatial information to run their daily business. Other

-
1. RFID stands for Radio Frequency IDentification, a technology to exchange data between tags and readers over a short range. See *RFID Essentials* (O'Reilly Media, 2006).
 2. See Daratech Inc.'s analysis titled “Geographical Information Systems: Markets and Opportunities” (www.daratech.com/research/index.php).

organizations in the fields of retail, distribution, and marketing use spatial information for strategic decision making—for example, choosing store locations, making investment decisions, examining market segmentation, and supporting clients.

At one point in time, the Internet seemed to have made location irrelevant. The Web emerged as a locationless cloud, where we could contact anybody around the world instantly and shop anywhere without the usual constraints of geography. It seemed that the worlds of transport, logistics, and location received a critical blow. Of course, that thinking was naive. The Internet has made geography even more relevant and has bound digital and physical worlds closer than ever. It is now possible to do business over much farther distances, and tracking the locations of different components of a business and analyzing them have become all the more important.

The emergence of wireless and location services promises to add location to every information item that we use or process. Technologies such as RFID have the potential to radically alter the retail and distribution worlds, making it possible to cheaply locate and track individual items, however small they are. With these new developments, the relevance of location has grown, and this is why it has become increasingly important to master the tools that handle spatial information.

Software tools for spatial information management have been traditionally known under the name of Geographical Information Systems (GIS). These systems are specialized applications for storing, processing, analyzing, and displaying spatial data. They have been used in a variety of applications, such as land-use planning, geomarketing, logistics, distribution, network and utility management, and transportation.³ However, until recently GIS have employed specific spatial data models and proprietary development languages, which held them separate from the main corporate databases. This has represented a barrier for the full deployment of the added value of spatial data in organizations.

As the use of GIS in enterprises and in the public sector has grown in popularity, some of the limitations of GIS have become apparent. Organizations often have to deal with multiple and incompatible standards for storing spatial data, and they have to use different languages and interfaces to analyze the data. Furthermore, systems such as Customer Relationship Management (CRM) and Enterprise Resource Planning (ERP) or the systems used in logistics increasingly rely on the integration of spatial information with all other types of information. This has often been an operational and technical challenge that in some cases was solved by manually extracting information from one system and loading it into another to perform the necessary spatial analysis.

Oracle Spatial has an important role in changing this situation. Once the spatial data is stored in an Oracle database, it can be processed, retrieved, and related to all the other data stored in the database: spatial information, or location, is just another attribute of a business object. This eliminates both the need for coordinating multiple data sources because of an application's dependence on special data structures and using different languages to query the data. Relevant features of Oracle Spatial are the ability to access spatial data through SQL statements, just like any other database content, and support for industry standards for spatial information (SQL and Open Geospatial⁴). Above all, Oracle Spatial facilitates leveraging the full added value of spatial information, which becomes an integral part of the information assets of organizations.

Given this overview of what location information is and how it can be used, in this chapter we will elaborate on the following topics:

-
3. For an introduction to GIS and its applications, see *An Introduction to Geographical Information Systems, Third Edition* (Prentice Hall, 2006).
 4. See www.opengeospatial.org.

- First, we describe how location information is used in different industry segments. Chances are that this will relate to your application and give you a head start putting spatial information to good use.
- Next, we describe different sources for spatial data. The data could be location information from different applications, or it could be geographical data representing, for instance, political boundaries.
- We then describe typical functionality required for managing spatial/location information. This functionality involves storing and analyzing the spatial data. We look at a specific example to illustrate the different components of such spatial processing.
- Finally, we discuss the systems that enable spatial information management, such as GIS, and their evolution. We consider an out-of-the-box approach to spatial information and the Oracle Spatial approach that integrates spatial data with other data in an Oracle database. We elaborate on this comparison and highlight the benefits of using Oracle Spatial.

Using Spatial Information in Various Industries

Let's now consider a simple business application example. The database for this application contains data about available products (a `Products` table), customers (a `Customers` table), suppliers (a `Suppliers` table), delivery sites (a `Delivery` table), and competitors (a `Competitors` table). The tables for customers, suppliers, delivery sites, and competitors contain information on the location of each item in the table. For instance, the `Customers` table contains the address of each customer and also the x,y coordinates of the address.

Notice that only the address is usually known, but for many spatial analyses, such as the calculation of the distance between a customer's location and delivery sites, you need to know the x,y coordinates of this address. The conversion of address fields to x,y coordinates is one of the most fundamental spatial operations described in this book, called *geocoding*. It serves to translate a text string such as "Abbey Road, 3, London NW8" into something like "longitude = -0.1784; latitude = 51.5320," which is the information used to relate spatial information items to each other.

With this information available, we might want to conduct valuable business analyses that can help determine new marketing campaigns, opening of new stores, and discontinuation of poorly located stores, as well as identify more efficient home-delivery schedules, changes in the stores' product portfolios, and so on. Consider the following options:

- Identify customers that are close to a competitor store (say less than 5 kilometers). To prevent them from switching stores, you could design a specific marketing campaign proposing special discounts for these customers.
- Optimize the distribution network. By counting the number of customers who are located within a certain distance from a distribution center, you could see whether some centers are overloaded or underutilized. This may lead to a redesign of the distribution network.
- Identify routes from delivery sites to customer locations, and cluster goods in such a way that the same delivery can serve multiple customers and save money. Note that this analysis requires additional data, such as the road network.
- Superimpose the location of stores on a population map, and check whether the store locations are appropriate. If some areas are underserved, this would alert you to opportunities for new outlets. Note that additional demographic data is often useful for this analysis.
- Visualize table data and analysis results as maps (such as customer maps, delivery site maps, and so on) and produce rich visual material better suited for communication and decision making.
- Integrate these maps with existing applications, such as a CRM system, so that location information and analysis can promote effective customer relations.

To perform these types of analyses, you need to store location information for customers, delivery sites, and competitors. In practice, this will mean augmenting the corresponding tables with additional columns for storing location information. You also need to store additional information, such as street networks, rivers, city and state boundaries, and so on, to use in visualization and analysis.

The preceding analyses are representative of a vast class of uses for spatial information. The following list summarizes some of the main uses of spatial data, analysis, and visualization in various industries:

- *Banking and finance:* These industries use location data for analysis of retail networks and for market intelligence. The customer database combined with demographics and wealth information helps banks define an optimal retail network and define the best product mix to offer at each branch.
- *Telecommunications:* Location analysis helps telecom operators and carriers improve their competitive position. Spatial data is used for network planning, site location, maintenance organization, call-center and customer support, marketing, and engineering.
- *Local and central government:* Spatial information is heavily used by all government agencies, since they manage a multitude of assets distributed over large territories. Uses include natural resource management or land-use planning, road maintenance, housing stock maintenance, emergency management, and social services.
- *Law Enforcement:* Spatial information helps officers in operational duties, as well as in crime analysis and prevention. Location information is used by field officers to locate places and other resources in the field in real time. Investigators use spatial data for crime analysis. Spatial patterns of crime are used to better locate police resources and improve prevention.
- *Real estate and property management:* Geographic data and demographics are used to identify and assess locations for outlets, housing, or facilities. Land-use, transport, and utility networks are used to site industrial and production facilities.
- *Retail:* Location data serves as a basis for operational and strategic decisions. It can be used to identify the profile of the best customers and help reach similar prospects. Spatial data can increase the relevance and focus of marketing campaigns and find the best layout of a distribution network for maximum profit.
- *Utilities:* Many different utility systems can be found under almost every street. Utility companies use spatial information to design these underground systems, plan and monitor groundwork, and maintain their cable and pipe networks.
- *Communications, media, and advertising:* Location data are frequently used for increasing the return of communications campaigns. Segmentation and location-based targeting help companies finesse the timing and appropriateness of marketing campaigns, thereby increasing their expectation of success.
- *Wireless data services:* Wireless data services increasingly use location data to enrich the user experience and provide valuable services. Uses include personal navigation systems, friend finders, roadside emergency, location-based yellow page searches, and the like. Wireless location services are necessary for fast returns on investments made on third-generation telecom networks.

Sources of Spatial Data

In the previous section we described the uses of spatial information in applications and in various industries, and we introduced the distinction between application data and spatial data. The simplest example is that of address lists collected as text items and subsequently enhanced by associating geographical (longitude, latitude) coordinates to each address. This association makes it possible to analyze the address information from the spatial perspective, an otherwise impossible operation based on the original address list.

In general, the association between nonspatial objects and their corresponding geometry makes it possible to relate the objects based on spatial concepts (close, far, overlap, joined, and so on). Very often the tables derive their spatial dimension from some primarily spatial data sources. In the case of address geocoding, for instance, postal data provides the locations of individual addresses in the form of a reference address list with the associated coordinates.

This is only one of the multitudes of spatial datasets and sources used in practice. Some datasets, such as cadastral data, land-use data, road network data, administrative boundary data, rivers and lakes data, and so on, are almost always present in spatial analysis and visualization. This data is collected, updated, and distributed by public bodies or by companies (the latter is the case, for instance, for the road networks for car navigation). All these datasets are first of all spatial, because the geographic component of the data content defines the usefulness and relevance of the entire dataset, and they are often used as reference layers.

The vast majority of these datasets are dynamic, at least to some extent. However, there are several cases in which the reason for using spatial data is specifically because of their dynamics. For example, use of *real-time location* is increasingly common, thanks to the widespread use of GPS and the growing use of location systems such as Wi-Fi location or RFID tagging, to locate people or objects.

GPS receivers can be located with high accuracy and can feed a database with the real-time location of a moving person/object (a field engineer, a car, a truck, a container, and so on). Note that there are also many commercial GPS applications, such as car navigators, that use real-time location within closed applications that support a specific purpose (such as door-to-door navigation) without connection to corporate data infrastructures. However, in most cases, it is the ability to feed the enterprise databases with the location of the mobile users or assets of an organization that allows planning, scheduling, and logistics improvements.

This is increasingly becoming the case in the retail and distribution industries, where the use of RFID, instead of bar codes, makes it possible to track vast amounts of goods automatically while they travel through the distribution chain from supplier to end user. RFID tagging can be implemented at the level of single items, products, or even documents. With RFID, goods can be followed precisely—for instance, within a warehouse—and this information can be used to minimize inventory, optimize supply schedules, and create a unique opportunity to link logistics with administrative, CRM, and ERP systems. It is likely that these areas, often referred to as *location-based* or *sensor-based* systems and services, will stimulate a rapid increase in the use of spatial information in the near future.

Managing and Analyzing Spatial Data

In this section, we will examine how to manage spatial data and what the typical analysis functions on spatial data are. Note that a variety of spatial processing systems such as GIS and spatial-enabled databases can provide this functionality using their own types and functions. We first describe spatial processing using generic terminology without referring to any specific solution (such as Oracle Spatial).

Spatial operations typically include, but are not limited to, the following:

- *Storage of spatial data:* In most cases, this involves the following:
 - Storing the data in an appropriate form in the database. For instance, the database system could have a *geometry* type to store spatial information as points, lines, polygons, and other types of vector representations. The system may also have a *network* type for modeling road networks.⁵
 - Inserting, deleting, and updating these types of spatial data in the database.
- *Analysis of vector spatial data:* This typically includes the following analysis functionality:
 - *Within-distance:* This operation identifies all spatial data within a specified distance of a query location.
 - *Contains:* This operation identifies all spatial data that contain a specified query location (geometry). Functions to detect other types of relationships may also be defined.
 - *Nearest-neighbor:* This operation identifies all spatial data closest to a query location.
 - *Distance:* This operation computes the distance between two spatial objects.
 - *Buffer:* This operation constructs buffer zones around spatial data.
 - *Overlay:* This operation overlays different layers of spatial data.
 - *Visualization:* This operation presents spatial data using maps.
- *Analysis of network data:* Typically, most spatial data, such as road networks, can also be represented as network data (in addition to vector data). We can perform the preceding analysis on such data using network proximity rather than spatial proximity.

The subjects of spatial analysis and management have filled dozens of books and hundreds of university courses. Our goal here is not to repeat all this—the references at the end of this chapter will provide you with a good background on these topics. Here, we will illustrate spatial analysis and management by describing how you can apply them to solve a common problem in the retail industry: site selection.

The consideration of location in Figure 1-1 streamlines the selection of candidate sites for a shopping mall. The process involves limiting the choice to those locations that are the following:

- Included in areas where construction is allowed
- On sale and of a suitable size
- Not exposed to natural risks, such as floods
- Close to main roads to ensure good accessibility

5. Oracle Spatial includes an additional data type called raster, which is used for images and grid data. We cover raster data and the raster data model in Appendix D.

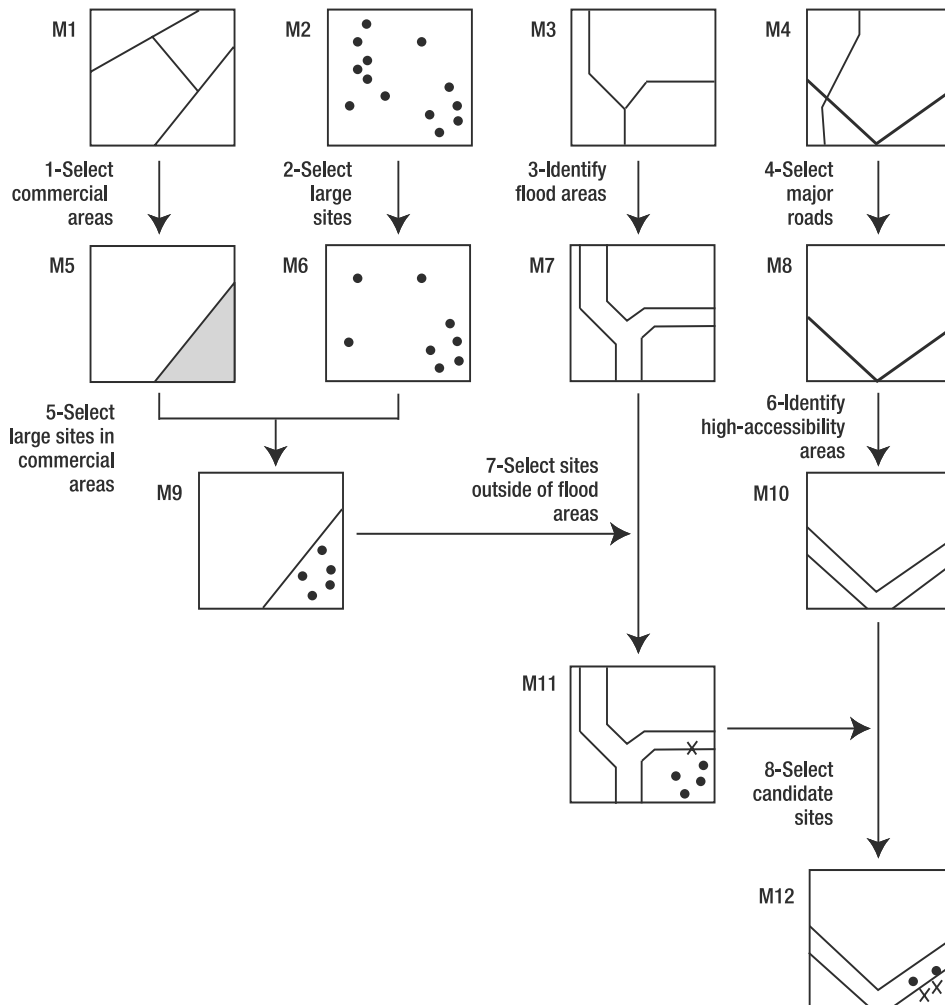


Figure 1-1. *Spatial data and spatial analysis for choosing a site for a shopping mall*

For the selection of suitable sites, we use spatial information and spatial analysis. To keep the example simple, however, we ignore demographic issues.

The main steps of the analysis are as follows:

1. From the land-use map (provided by a public organization), we first select areas for which we can obtain permits to build commercial sites. These areas are labeled as “commercial” and denote sites where new commercial activities can be located.
2. From a map that contains sites for sale (provided by a large real-estate agency), we restrict the choice to sites that are sufficiently large for a shopping mall.
3. On the basis of a risk map, which indicates safety buffer areas around rivers, we eliminate those sites that may be subject to floods.
4. Finally, of the remaining sites, only those close to main roads are deemed suitable for accessibility reasons.

Figure 1-1 shows the sequence of steps, the data used, and the spatial operations involved in this process. Note that the maps are numbered M1–M12 and the steps are numbered 1–8.

The combination of the first two steps leads to five candidate sites. One of them is excluded because of high flood risk, and two additional ones are excluded because they are located too far away from the main roads. This narrows the results to two suitable candidate sites.

Table 1-1 details the steps in this process. Note that the usual way of representing the data used in this example is through maps, as in Figure 1-1. Note also that the description can be easily translated into database and SQL terms. The various “maps” correspond to one or more database tables. The data objects (points, polygons, lines, grids, and so on) and their attributes are table records, while the analysis is performed with SQL statements. It is clear that some SQL extensions are needed to handle spatial and nonspatial objects simultaneously. The rest of this book will essentially deal with the models and tools available in Oracle Spatial for storing and processing this type of data for types of analysis like this one.

Table 1-1. *Steps, Data, and Analysis for Choosing a Site for a Shopping Mall*

Step	Data	Analysis	Result
1. Select commercial areas.	M1: Land use map. Collection of polygons, described by an attribute “land-use type.”	Select polygons where the attribute is “commercial.”	M5: Commercial areas. A set of polygons with the “commercial” attribute.
2. Select large sites.	M2: Sites for sale. Locations described by price, plot size, etc.	Select points where the size is larger than a certain value.	M6: Large sites. A selection of points corresponding to large sites for sale.
3. Identify flood areas.	M3: River map.	Create a buffer around the riverbed (e.g., 1 km) that is at risk of floods.	M7: Flood risk areas.
4. Select major roads.	M4: Road network map. Road segment attributes are “road type,” “max speed,” etc.	Select road segments where the attribute is “major roads.”	M8: Major roads.
5. Select large sites in commercial areas.	M5 and M6.	Overlay M5 and M6. Select “large” points within “commercial” polygons.	M9: A selection of points corresponding to large sites within commercial areas.
6. Identify high-accessibility zones.	M8.	Create a buffer of 500 meters on each side of a major road.	M10: High accessibility zones.
7. Select sites outside of flood areas.	M9 and M7.	Overlay M9 and M7, and eliminate sites in the flood areas.	M11: Points corresponding to large sites within commercial areas not subject to flood risks.
8. Select candidate sites.	M10 and M11.	Select safe sites within high-accessibility zones.	M12: Large sites in commercial areas that are not subject to floods and are highly accessible.

For simplicity, in the example we have assumed that a new map is created at the end of every step. This is certainly a possibility, but it is not necessarily the best option. Later in this book, we will discuss data modeling and how to optimize the sequence of operations. In particular, Chapters 8 and 9 cover spatial operators and functions that make it possible to cluster some of the steps in the example into single queries, making the process much simpler and more efficient.

Storing Spatial Data in a Database

Looking at vector data, we usually distinguish between the following:

- *Points* (for example, the plots for sale in Figure 1-1), whose spatial description requires only x,y coordinates (or x,y,z if 3D is considered)
- *Lines* (for example, roads), whose spatial description requires a start coordinate, an end coordinate, and a certain number of intermediate coordinates
- *Polygons* (for example, a residential area), which are described by closed lines

Figure 1-2 shows an example containing point, line, and polygon data. The figure corresponds to a small portion of the area used in the previous site selection example. The vector representation, here simplified for convenience, shows a point (the stadium), three lines (the roads), and four polygons (the built-up areas, clipped at the picture borders, and the sports complex).

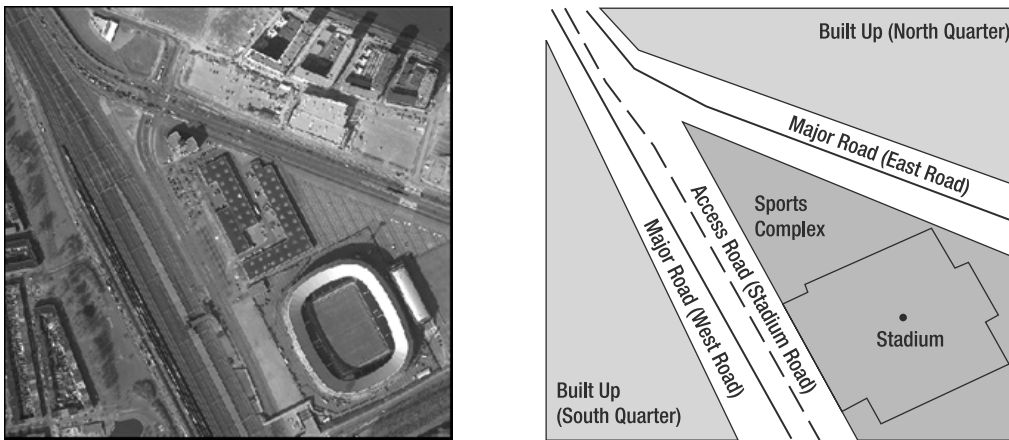


Figure 1-2. Vector representation of the spatial objects in the picture

The vector data in Figure 1-2 could be stored in one or multiple tables. The most natural way of looking at this data is to identify *data layers*—sets of data that share common attributes—that become data tables. Most spatial databases use a special data type to store spatial data in a database. Let's refer to this type as the *geometry*. Users can add columns of type *geometry* to a database table in order to store spatial objects.

In this case, the choice of tables would be rather simple with three main data layers present: “Road infrastructures,” “Land use,” and “Points of interest.” These three layers contain objects that share common attributes, as shown in the three tables later in this section. The same objects could have been aggregated into different data layers, if desired. For instance, we could have stored major and minor roads in different tables, or we could have put roads and land use in the same table. The latter would make sense if the only attributes of relevance for roads and land-use areas were the same, for instance, the province name and the city name. It is also worth stressing that every *geometry* column can contain any mix of valid spatial object (points, lines, polygons) and also that every table can contain one or more *geometry* columns.

Structuring spatial data into tables and defining the right table structure are the first logical activities of any spatial analysis. Fortunately, in most cases there is an intuitive correspondence between the data and the table structure used to store them. However, in several cases you may find that the spatial database design can be a complex activity. Proper designs may facilitate analysis enormously, while poor data structures may make the analysis complex and slow. These issues are addressed in various places in the book but in particular in Chapter 3.

Table 1-2 shows the road infrastructure table of Figure 1-2. This table contains three records corresponding to the east road, the west road, and the stadium road. All of them are represented as lines using the *geometry* type. Each road is described by three types of attributes: the road type (one column containing either “major,” “local,” or “access” road), the road name (a column containing the name of the road as used in postal addresses), and the area attributes (two columns containing the name of the province and city where the road is located).

Table 1-2. Road Infrastructure Table

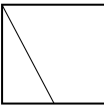
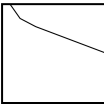
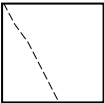
ID	Province	City	Road Name	Road Type	Road Geometry
1	Province name	City name	West road	Major road	
2	Province name	City name	East road	Major road	
3	Province name	City name	Stadium road	Access road	

Table 1-3 shows the land-use table. It contains three records corresponding to the north quarter, the south quarter, and the sports complex. In this case, all spatial objects are polygons. Each object has three types of attributes: the surface of the area (in square meters), the name of the area, and the area location (province and city names).

Table 1-3. *Land Use Table*

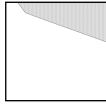
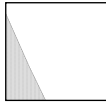
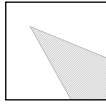

ID	Province	City	Area Name	Surface (Square Meters)	Area Geometry
1	Province name	City name	North quarter	10,000	
2	Province name	City name	South quarter	24,000	
3	Province name	City name	Sports complex	4,000	


Table 1-4 shows the points of interest (POI) in the area. It contains two records: a point (in this case, the center of the stadium complex) and a polygon (in this case, the contour of the stadium complex). Attributes include the type of POI from a classification list, the POI name, and the province and city where they are located.

Table 1-4. *Points of Interest Table*

ID	Province	City	POI Name	Type of POI	POI Geometry
1	Province name	City name	Olympic stadium	Sports location	●
2	Province name	City name	Olympic stadium	Sports infrastructure	

In the Table 1-4, we use two records to describe the same object with two different geometries. Another option for storing the same information is presented in Table 1-5, where we use two columns of type *geometry* to store two different spatial features of the same object. The first *geometry* column stores the POI location, while the second stores the outline of the complex. Under the assumption that all other nonspatial attributes are the same, Table 1-5 is a more efficient use of table space than Table 1-4.

Table 1-5. *Points of Interest Table: Two Geometry Columns*

ID	Province	City	POI Name	Location (POI) Geometry	Infrastructure Geometry
1	Province name	City name	Olympic stadium	●	

The objects in the preceding tables are represented with different line styles and fill patterns. This information is added for clarity, but in practice it is not stored in the *geometry* object. In Oracle Spatial, the *geometry* data are physically stored in a specific way (which we will describe in Chapters 3 and 4) that does not have a direct relationship to the visual representation of the data. Chapter 12, which describes the Oracle Application Server MapViewer, shows how symbology and styling rules are used for rendering *geometry* instances in Oracle.

Geometry models in the SQL/MM and Open Geospatial (OGC) specifications describe in detail the technical features of the *geometry* type and how points, lines, and polygons are modeled using this type.

Spatial Analysis

Once data is stored in the appropriate form in a database, spatial analysis makes it possible to derive meaningful information from it. Let's return to the site selection example and look again at the three types of spatial operations that we used:

- *Select*, used in the following:
 - Step 1 (to select areas where the attribute was a certain value)
 - Step 2 (to select large sites from the sites for sale)
 - Step 4 (to select major roads from the road network)
- *Overlay*, used in the following:
 - Step 5 (large sites in commercial areas)
 - Step 7 (sites away from risk areas)
 - Step 8 (sites within highly accessible areas)
- *Buffer*, used in the following:
 - Step 3 (areas subject to flood risk)
 - Step 6 (high accessibility areas)

Returning to our example, assuming we have the data stored in a database, we can use the following eight pseudo-SQL statements to perform the eight operations listed previously. Please note that for the sake of the example, we have assumed certain table structures and column names. For instance, we have assumed that M1 contains the columns LAND_USE_TYPE, AREA_NAME, and AREA_GEOMETRY.

1. Use

```
SELECT AREA_NAME, AREA_GEOMETRY
FROM M1
WHERE LAND_USE_TYPE = 'COMMERCIAL'
```

to identify available plots of land for which a construction permit can be obtained for a shopping mall.

2. Use

```
SELECT SITE_NAME, SITE_GEOMETRY
FROM M2
WHERE SITE_PLOT_AREA > <some value>
```

to identify available sites whose size is larger than a certain value.

3. Use

```
SELECT BUFFER(RIVER_GEOMETRY, 1, 'unit=km')
FROM M3
WHERE RIVER_NAME= <river_in_question>
```

to create a buffer of 1 kilometer around the named river.

4. Use

```
SELECT ROAD_NAME, ROAD_GEOMETRY
FROM M4
WHERE ROAD_TYPE='MAJOR ROAD'
```

to identify major roads.

5. Use the *contains* operator to identify the sites selected in step 2 that are within areas selected in step 1. You could also achieve this in one step starting directly from M1 and M2:

```
SELECT SITE_NAME, SITE_GEOMETRY
FROM M2 S, M1 L
WHERE CONTAINS(L.AREA_GEOMETRY, S.SITE_GEOMETRY)='TRUE'
AND L.LAND_USE_TYPE= 'COMMERCIAL'
AND S.SITE_AREA > <some value>;.
```

6. As in step 3, use the *buffer* function to create a buffer of a certain size around the major roads.
7. Use *contains* to identify sites selected in step 5 that are outside the flood-prone areas identified in step 3.
8. Use *contains* to identify safe sites selected in step 7 that are within the zones of easy accessibility created in step 6.

Oracle Spatial contains a much wider spectrum of SQL operators and functions (see Chapters 8 and 9). As you might also suspect, the preceding list of steps and choice of operators is not optimal. By redesigning the query structures, changing operators, and nesting queries, it is possible to drastically reduce the number of intermediate tables and the queries. M11, for instance, could be created starting from M9 and M3 directly by using the *nearest-neighbor* and *distance* operations together. They would select the nearest neighbor and verify whether the distance is larger than a certain value.

Benefits of Oracle Spatial

The functionality described in the previous section has been the main bread and butter for GIS for decades. In the past five to ten years, database vendors such as Oracle have also moved into this space. Specifically, Oracle introduced the *Oracle Spatial* suite of technology to support spatial processing inside an Oracle database.

Since GIS have been around for several years, you may wonder why Oracle has introduced yet another tool for carrying out the same operations. After all, we can already do spatial analysis with existing tools.

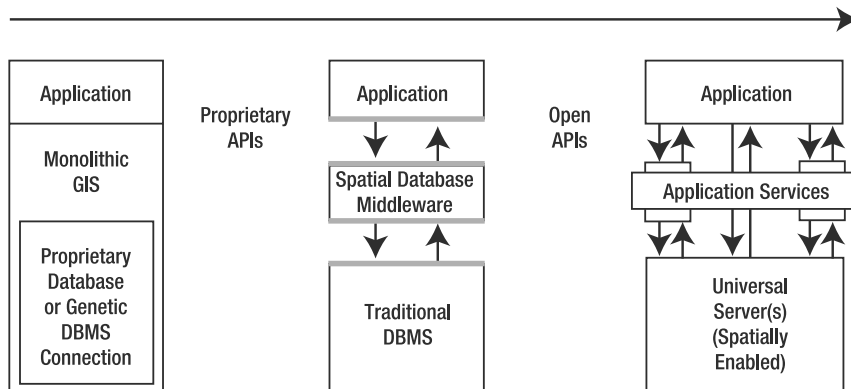
The answer lies in the evolution of spatial information technology and in the role of spatial data in mainstream IT solutions. GIS have extensive capabilities for spatial analysis, but they have historically developed as stand-alone information systems. Most systems still employ some form of dual architecture, with some data storage dedicated to spatial objects (usually based on proprietary formats) and some for their attributes (usually a database). This choice was legitimate when mainstream databases were unable to efficiently handle the spatial data. However, it has resulted in the proliferation of proprietary data formats that are so common in the spatial information industry.

Undesired consequences were the isolation of GIS from mainstream IT and the creation of automation islands dedicated to spatial processing, frequently disconnected from the central IT function of organizations. Although the capabilities of GIS are now very impressive, spatial data may still be underutilized, inaccessible, or not shared.

Two main developments have changed this situation: the introduction of open standards for spatial data and the availability of Oracle Spatial. Two of the most relevant open standards are the Open Geospatial Simple Feature Specification⁶ and SQL/MM Part 3.⁷ The purpose of these specifications is to define a standard SQL schema that supports the storage, retrieval, query, and update of spatial data via an application programming interface (API). Through these mechanisms, any other Open Geospatial-compliant or SQL/MM-compliant system can retrieve data from a database based on the specifications. Oracle Spatial provides an implementation for these standards⁸ and offers a simple and effective way of storing and analyzing spatial data from within the same database used for any other data type.

The combination of these two developments means that spatial data can be processed, retrieved, and related to all other data stored in corporate databases and across multiple sources. This removed the isolation of spatial data from the mainstream information processes of an organization. It is now easy to add location intelligence to applications, to relate location with other information, and to manage all information assets in the same way. Figures 1-3 and 1-4 summarize this paradigm shift.

Figure 1-3 illustrates the industrywide shift from monolithic/proprietary GIS to open, universal, spatially enabled architectures.



Source: UNIGIS-UNIPHORM project

Figure 1-3. From monolithic/proprietary GIS to universal, spatially enabled servers (Source: UNIGIS-UNIPHORM project. See www.unigis.org)

Figure 1-4 emphasizes the shift from geo-centric processing to information-centric processing, where the added value is not in the sophistication of the spatial analysis but in the benefits it delivers. Traditional geoinformation management tools emphasize geodata processing while separating geodata storage from attribute data storage (see the emphasis on *Geography* in “Gis” in the figure).

6. See www.opengeospatial.org for information on approved standards, for an overview of ongoing standardization initiatives for spatial information data and systems, and for an up-to-date list of compliant products.
7. See ISO/IEC 13249-3:2003, “Information technology - Database languages - SQL multimedia and application packages - Part 3: Spatial” (www.iso.org/iso/en/CatalogueDetailPage.CatalogueDetail?CSNUMBER=31369).
8. The ST_Geometry of Oracle Spatial is fully compliant with the OGC Simple Feature specification for the object model.

Oracle Spatial makes it possible to process geodata within the same information platform used for all other data types (see the emphasis on *Information Systems* in “gIS” in the figure).

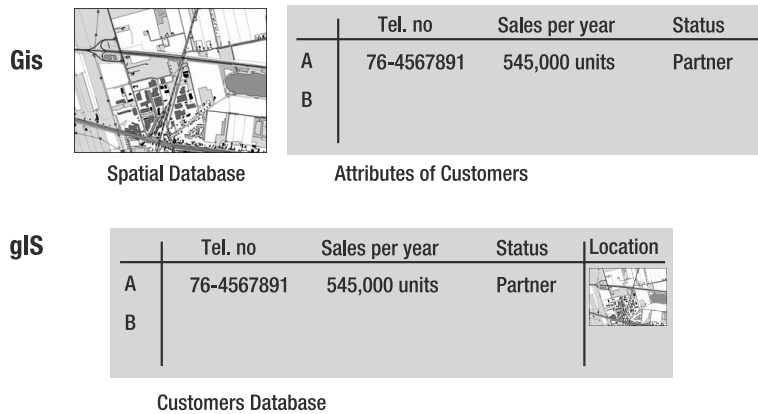


Figure 1-4. From *Gis* to *gIS*

The benefits of using Oracle Spatial can be summarized as follows:

- It eliminates the need for dual architectures, because all data can be stored in the same way. A unified data storage means that all types of data (text, maps, and multimedia) are stored together, instead of each type being stored separately.
- It uses SQL, a standard language for accessing relational databases, thus removing the need for specific languages to handle spatial data.
- It defines the `SDO_GEOMETRY` data type, which is essentially equivalent to the spatial types in the OGC and SQL/MM standards.
- It implements SQL/MM “well-known” formats for specifying spatial data. This implies that any solution that adheres to the SQL/MM specifications can easily store the data in Oracle Spatial, and vice versa, without the need for third-party converters.
- It is the de facto standard for storing and accessing data in Oracle and is fully supported by the world’s leading geospatial data, tools, and applications vendors, including NAVTEQ, Tele Atlas, Digital Globe, 1Spatial, Autodesk, Bentley, eSpatial, ESRI, GE Energy/Smallworld, Intergraph, Leica Geosystems, Manifold, PCI Geomatics, Pitney/Bowes/MapInfo, Safe Software, Skyline, and many others.⁹
- It provides scalability, integrity, security, recoverability, and advanced user management features for handling spatial data that are the norm in Oracle databases but are not necessarily so in other spatial management tools.
- It removes the need for separate organizations to maintain a spatial data infrastructure (hardware, software, support, and so on), and it eliminates the need for specific tools and skills for operating spatial data.
- Through the application server, it allows almost any application to benefit from the availability of spatial information and intelligence, reducing the costs and complexity of spatial applications.

9. For a list of partners, visit <http://otn.oracle.com/products/spatial/index.html>, and click the Partners link (in the Oracle Spatial and Locator Resources section).

- It introduces the benefits of grid computing to spatial databases. For large organizations that manage very large data assets, such as clearinghouses, cadastres, or utilities, the flexibility and scalability of the grid can mean substantial cost savings and easier maintenance of the database structures.
- It introduces powerful visualization of spatial data, eliminating the need to rely on separate visualization tools for many applications.

Summary

This first chapter provided an introduction to spatial information management, its importance in business applications, and how it can be implemented in practice. The example of situating a shopping mall illustrated the relationship between the logical operations necessary to make a proper choice and the spatial data and analysis tools that can be used to support it.

After describing the example, we discussed how database vendors such as Oracle enable spatial functionality. We enumerated the benefits of a database approach, specifically that of Oracle Spatial. We observed that the most basic and essential feature of Oracle Spatial is that of eliminating the separation between spatial and nonspatial information in a database. This separation was mainly the result of technology choices and technology limitations, but it does not have any conceptual ground or practical justification. On the contrary, all evidence points toward the need to integrate spatial and nonspatial information to be able to use the spatial dimension in business operations and decision making.

We have also made the explicit choice of emphasizing the relevance of adding the spatial dimension to mainstream database technology, thereby introducing spatial information starting from the database. A GIS specialist, a geographer, or an urban planner would have probably described the same examples with a different emphasis—for instance, highlighting the features and specific nature of spatial data and analysis. This would have been a perfectly legitimate standpoint and is one very common in literature and well served by the selected titles in the “References” section.

In the next chapter, we will give a brief overview of the functionality of Oracle Spatial. The subsequent chapters in the book present an in-depth tour of the different features and functionality of Oracle Spatial and how you can implement them in applications.

References

- Glover and Bhatt, *RFID Essentials*, Cambridge: O'Reilly Media, 2006.
- Grimshaw, David J. *Bringing Geographical Information Systems into Business, Second Edition*. New York: John Wiley & Sons, 1999.
- Haining, Robert. *Spatial Data Analysis: Theory and Practice*. Cambridge: Cambridge University Press, 2003.
- Heywood, Ian, Sarah Cornelius, and Steve Carver. *An Introduction to Geographical Information Systems*. New Jersey: Prentice Hall, 2006.
- Korte, George B. *The GIS Book, 5th Edition*. Clifton Park, NY: OnWord Press, 2000.
- Longley, Paul A., Michael F. Goodchild, David J. Maguire, and David W. Rhind, eds. *Geographical Information Systems and Science*. New York: John Wiley & Sons, 2005.