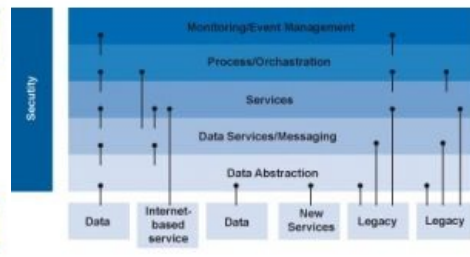
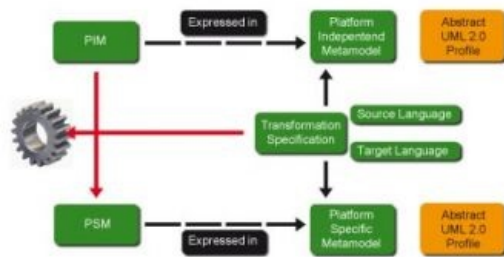


ADVANCES IN MANAGING LBS AND FRAMEWORK DATA

Spatial Database Challenges



When providing local information, a location-based service (LBS) exploits the geographical position of the mobile device used to access it. The management of very large spatial data databases requires a focus on production and quality control of data, both framework and for LBSs. Progress has lately been made at two levels: system architecture and knowledge management. Standards are key.

Mass-market LBSs have had a number of promising beginnings. However, there is a gap between concept and reality due to shortfalls in several areas. Issues concerning price of suitable handsets and availability of positioning capabilities are currently being resolved. The suitability and quality of location-based data, previously masked by hardware issues, are now becoming evident. Increasingly important in data synergy is quality framework data, provided by national mapping and cadastral agencies (NMCAs). Large data aggregators, value-adders and NMCAs share problems in managing very large, complex and dynamic data holdings and in delivering timely, quality data products and services. The challenges are being met.

Mainstream IT

A key change in recent times is the realisation that the geospatial industry is not independent; industry-specific standards organisations (ISO TC211) and the open geospatial consortium (OGC) build on standards implemented by mainstream IT industries. The benefits of this are particularly evident in the increasing adoption of enterprise architectures and formal data modelling techniques. Mutually reinforcing technology advances are emerging at system level and similar benefits are being realised at knowledge management level too.

Formal Data Modelling

Formal data modelling using Universal Modelling Language (UML) is well established in the geospatial domain. A recent EuroSDR (spatial data research) workshop on feature/object data models documented the emergence of a new generation of data models and their characteristics. UML is widely used for defining and documenting; it results in better communication and definition at conceptual level, the emergence of tools for automating generation of implementation-level from conceptual-level UML schema, and the ability to manage evolution of the data model in a controlled manner. This underpins the adoption of a model-driven architecture (MDA) approach steered by the conceptual or platform-independent model (PIM). The storage, or platform-specific, physical model (PSM) is derived from the PIM in an automated or near-automated fashion, as is mapping from the storage- to transfer-model (usually a geography markup language (GML) application schema). Changes in the conceptual model can thus be propagated across the whole system in a co-ordinated and error-free manner (see Figure 1).

Schema-Aware Software

Another key development has been that of schema-aware software components. The extensible markup language XML and XMLSchema, together with domain-specific extensions such as GML, have gained near universal acceptance and are supported by an impressive array of software tools. The role of XMLSchema as a machine-readable manual for the data model underpins powerful schema translation techniques. These in turn enable key functionality, including elaboration and data harvesting, and consolidation of diverse data sources as required in spatial data infrastructures (SDIs). Deployment of schema-aware software components in MDA allows the whole system to adapt automatically to data model change without new programming.

Workflow Technologies

These advances are complemented by developments in orchestration and workflow technologies, standardised in the Web Services Business Process Execution Language OASIS. Standard WS-BPEL 2.0 is an XML grammar for programming, and provides central orchestration of distributed components. Business process and workflow is implemented as a two-tier system. Large-scale process implementation is governed using BPEL to control high-level transitions in process; small-scale business logic implementation uses rules-processing engines, as described below.

Overarching Framework

OASIS defines service-oriented architectures (SOA) as: '...a paradigm for organising and utilising distributed capabilities that may be under the control of different ownership domains. It provides a uniform means to offer, discover, interact with and use capabilities to produce desired effects consistent with measurable preconditions and expectations'. The key elements of an SOA are illustrated in Figure 2. It provides an overarching framework within which integration may occur of the above advances. Key resulting characteristics include loose coupling with formal contracts, scalability and reusability, i.e. future-proof system architecture.

Unearthing Knowledge

Rapid advances in IT knowledge management are beginning to yield benefits in the geospatial realm. Knowledge has historically been buried inside data, by the juxtaposition, for example, of data components or in implicit relationships. It may also be hidden in point applications, or even inside people's heads. This results in serious problems in keeping such knowledge up to date. Progress towards rigorous semantics contributes to the removal of ambiguities. The accessible storage of knowledge/expertise, portable and independent of specific datasets and systems is equally important. A crucial component here is a rules language that enables specification of logical constraints. Such a language needs to be unambiguous, logical and portable, compact, intuitive, quantitative, web-compatible, declarative and susceptible to refining. Rapid progress is being made in this area through initiatives such as the semantic web community.

Defining Rules

Rules-based processing follows the *fact-pattern-action* dynamic: if the given facts (known data sources) meet any of the patterns (business rules obeyed by the data source), then they perform the defined action. 'Radius Studio' from 1Spatial is an example of a modern rules-based processing environment, implemented both as middleware and as a service. The rules base is a set of conditions that objects from the data store should satisfy. A rule is a tree of predicates against which objects can be tested. Rules are expressed in a form independent of the schema of any particular data store, so they can easily be used with different data sources. An intuitive web-based interface is provided for defining rules and building up a rules-base. It allows definition of potentially complex rules with an easy-to-use, tree-structured browser interface. The syntax used is very similar to English, so the user need not have programming experience. Processing is carried out through actions, procedures that are applied to objects. Actions may either be applied to all the objects from a data store or in a more targeted manner, using action maps.

Measuring Quality

The quality of data, in the sense of measuring the degree of its conformance to a rules base, can be measured using 'Radius Studio'. A document or data store can be processed and a report generated giving results of the conformance test at both summary and individual feature level. The former reports the proportion of compliant objects; the latter lists non-compliant objects and rules infringed. Quality can be improved by automatically applying 'fixes' to cases within allowable tolerances or parameters as defined by an action map. Alternatively, quality can be improved by referring the list of non-compliant objects to an interactive client. When an acceptable level of quality is reached, all workspace objects that have been changed are committed back to the main data store.

Impressive Progress

Although never fully exposed, the challenging demands on data management technologies driven, by the LBS phenomenon around the turn of the last century were recognised by system architects. Steady and, in some cases, impressive progress has been made in resolving these issues at both system architecture and knowledge management levels. These advances are firmly based on mainstream and geospatial industry standards.