GIS INTEROPERABILITY: THE "SAFE" WAY

S.Raghavendran[1]

[1]GIS Analyst
PIXEL INFOTEK PVT LTD
#194, Sankey Rd., Sadashiva Nagar
Bangalore-560 080, India
Tel: +91-80-3619351/2/3, Cell: +91-9845673851
Fax: + 91-80-3619354
Email: srg_gis@yahoo.com

Geography or GIS for that matter has never been so closer to our lives. "GIS Interoperability" rightly said, as the holy grail of GIS has never been so much discussed among the geospatial community. Interoperability has been in our daily lives, whether we liked it and realized it or not. If interoperability was not an issue in the lives of the Stone Age man but for sure it is in this 21st century hi-tech electronic, tech perceptive day today life. Be it the travel experiences of someone amongst us especially to foreign lands with languages different from what we are accustomed to and facing situations of inadequate communication have been definitely part of our memories. Imagine the interoperability if there was one global language for communication. Imagine, what it would be if music files of our heartthrob celebrities written on Samsung CD's could be played only in Samsung CD players. But for the interoperability, music CDs has transcended such barriers. "How nice it would be, had I taken one of those power adapters for my electrical appliances" – A common hearsay from the first timers to the United States from Asian countries.

Interoperability has been in our daily lives, whether we liked it and realized it or not and GIS is no exception to that. Unfortunately the kind of interoperability that exists in the other industries does not exist in the geospatial industry for several reasons. Geospatial interoperability though seen as a moving target has in the recent years definitely gained momentum. Thanks to the buzzword "Geospatial Democracy", dissemination of geospatial information being seen as a public right, the initiative efforts of various organizations in developing open standards and the increasing usage of SDI (Spatial Data Infrastructure) interoperability tools like FME (Feature Manipulation Engine): GIS Interoperability may soon be a reality!

What is GIS interoperability?

In the world of GIS, interoperability in simple terms would mean the ability to integrate or exchange information between different components of a geospatial solution, even though different organizations or departments on different GIS platforms may have developed the components. The components here would typically mean CAD/GIS files, Spatial/ Attribute Databases and other related electronic documents of different types and from multifarious organizations. However this could be as simple as a plain GIS format "A" to GIS format "B" translation or could be a complicated one requiring data transformation into the structure the user or the application needs. In other words true GIS interoperability would mean the right tools and technologies to exchange or transform geospatial data into the structure, wherein the end user can take full advantage of specific GIS software capabilities.

Why GIS interoperability?

The concept of Geospatial Data Infrastructure was first conceived in Canada in the early 1980s. There after other countries like USA, UK, Holland, Australia and Malaysia joined the race. Of late with the setting up of National Geospatial Data Infrastructure (NGDI) in India, the importance of geospatial data and the need to have a geospatial data infrastructure similar to any other infrastructure has increased manifold. Up to the early 1990's, GIS implementations were typically stand-alone systems not integrated with other business processes. The data used was stored in file-based formats optimized for fast access. The ability to share data within an organization was limited to Network File System (NFS) capability and typically required that all application software be from the same software vendor. Sharing of data with other organizations using another vendor's software typically involved the use of GIS data translators either provided with the GIS software or built for this purpose using popular programming languages and utilities. Contrastingly, in this cyber savvy era of rapid transition, government agencies and other organizations are required to provide quick response to natural disaster, environmental crises and other such issues based on critical information available to them. When lives are at risk, every second counts and quick access to critical information could mean life and death situation. But in reality such critical information is either unavailable or available in disjoint legacy systems and multitude of GIS file formats. The decision makers are often put up with a devil and the deep blue sea situation: lack of information needed to make accurate decisions on one hand and the issue of interoperability between the multifarious formats on the other. This not only underlines the critical need to have a Geospatial Data Infrastructure at various levels but also appropriate SDI tools and technologies for interoperability, to make available the *right information,* at the *right time,* in the *right form.*

GIS interoperability technologies

GIS interoperability is easier to be said than realized. Yes, this is no less to a Herculean task. There are several hurdles in the way even before one can even see an iota of this being accomplished. Starting from the basic problems like the maps being not up to date, non availability of maps to the complicated ones like missing projection system, scale discrepancies etc. But this is not the end of all troubles. As it is rightly said "one size never fits all". Blame it on the technology or blame it on the file formats, even if the geospatial data were readily available, they are in multitude file formats. But many a time this is a very serious bottleneck in the process of building a geospatial data infrastructure as evident from the experiences across the globe.

Data interoperability has long been a burning issue in the GIS community. For years GIS users and managers of organizations across the globe continue to maintain data locked up in some proprietary format until the gloomy cloud of interoperability issue suddenly snuck out from nowhere and its dark shadow descends over. A nightmare scenario, but for this digital alchemy: FME (Feature Manipulation Engine), a Spatial ETL (Extract, Transform, and Load) technology from Safe Software, Inc., founded in 1993 in Canada.

Every GIS file format has its own advantages and limitations. Even with the OGC (Open GIS Consortium) gaining momentum in coming up with the standardization of GIS file formats and interfaces, there is still a long way to go on this route to interoperability.

FME (Feature Manipulation Engine): The Spatial ETL tool

Simply put, the objective of **ETL (Extract, Transform and Load)** tools is to transfer data from one data store to another. The "Extract" function reads data from a specified source data store, extracting the desired data. Next, the "Transform" function processes the acquired data, transforming it and even perhaps combining it with other data to package it into the correct structure for the destination data store. Finally, the "Load" function writes the resulting data to a target data store.

Spatial ETL (a term coined by Safe Software) takes ETL to the next level by adding the spatial component. A Spatial ETL tool such as Safe's Feature Manipulation Engine (FME) is designed to provide the same bottom-line benefit as traditional ETL systems, but with one distinct difference. They also include the spatial information within an organization, either in helping it create/mine spatial data or in capitalizing on the spatial data it already has. Many a time, by virtue of the way in which data is represented varies greatly from system to system, wherein a simple format A to format B translation is often not enough. When a simple translation or migration from one format to the other alone does not solve the GIS interoperability issue, a Spatial ETL like FME comes handy.

FME as a Spatial ETL tool extracts data from a desired data store, and then transforms it to the requested projection, format, and view. The data can then be presented to a requesting user application or loaded into another data store. Again, the intended data can be obtained for temporary use by the end user or application or as part of a permanent data migration/translation project. In summary, the goal of Spatial ETL is to give users the *right view*, of the *right data*, at the *right time!*

What truly distinguishes FME from all interoperability approaches is its ability to perform data model transformations at the same time as performing the translation. FME provides a "thick pipe" connection between systems that allows for significant data model restructuring to be performed during the translation.

FME: Why GIS file formats do not matter?

In traditional GIS file format translations, which are sometimes compared to a thin pipe connecting one data source to another, attributes and spatial data are forced through a limited data model. The underlying data model is very weak resulting in only basic concepts moved across. This typically adds little or no value to the data but provides a quick and easy way to get data from one system to the other.

FME employs a "semantic translation" (thick pipe translation), which focuses on changing the view of the data to something that matches the desires of the end user or end system. Instead of viewing the problem as a means of moving data from one format to another, FME works entirely on generic features and concentrates on providing building blocks that enable users to manipulate the data into the desired representation. Changing the format during data translation is a small part of the translation process. Semantic translation provides an engine that enables the redefinition of the data, either on input or output. Underlying the engine is a rich data model, much richer than what is supported in the proprietary systems, allowing for a high degree of redefinition, which is internally consistent and inherently extensible.

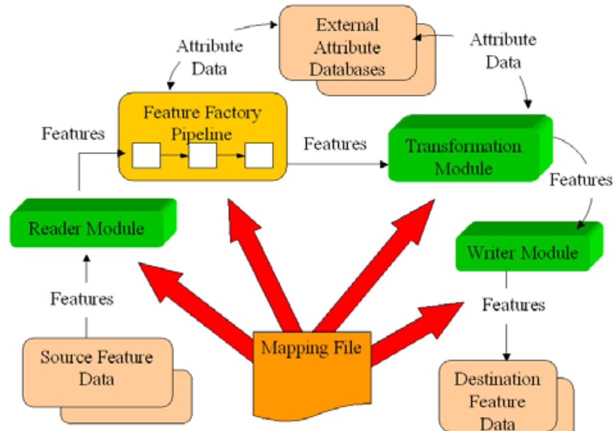
With an underlying rich data model and a powerful format neutral spatial data translation and transformation engine supporting over 100 file formats, it is needless to say that in FME, "CAD and GIS file formats don't matter". In addition to its format translation capabilities, FME also provides a wide range of data processing functions through its inbuilt *Functions* and *Factories*. These functions can be used to perform geometric manipulations (such as forming polygons or connecting lines) on the spatial data as it is being translated, as well as generate new attribute values (such as length or area). There are more than 65 processing functions (which operate on one spatial object at a time), and more than 47 processing factories (that operate on more than one spatial object at a time).

GIS interoperability: The "SAFE" way

FME has eight different suites, each suite catering to the multifarious needs of its users. Each FME Suite consists of three core components: FME Workbench, FME Universal Translator and FME Universal Viewer. **FME Universal Translator:** powerful data translation software with a drag-and-drop interface. Users can easily transfer and save data into as many as 123 different formats (including a variety of OGC ones) with accuracy. When users need to customize the translation of their data, they rely upon **FME Workbench**, which features a graphical data flow programming paradigm that makes performing very sophisticated processing tasks a breeze. The **FME Universal Viewer** allows quick viewing of data formats, including all the details of each graphical element they contain. Users can preview translation data (attributes and geometry), before performing translations using Workbench or Universal Translator. It is ideal for previewing data. FME supports over 4400 coordinate systems based on a variety of projections, ellipsoids and datum.

Functions and factories are the basic building blocks of FME provide users with sophisticated feature processing capabilities. Functions provide a flexible means of applying specific algorithmic operations to features as they are being transformed, thereby supplementing the transformation capabilities of the FME. FME factories enable the FME to perform operations on collections of features.

To transport features from one format to another, the FME considers features to be collections of attribute names and values associated with two-dimensional (2D) or three-dimensional (3D) geometry. The FME places no restrictions on the values or types of attributes. Attribute names consist of one or more ASCII characters. A typical FME translation is controlled by a series of rules specifying the guidelines and transformations of a data translation. The "mapping file", the currency of FME can be generated in multiple ways and can be edited using a simple application like Notepad. The mapping file drives the operation of all FME modules during a data translation. The translation of features from a source system to a destination system is completely controlled by rules specified in the semantic mapping file.



Feature Manipulation Engine (Architecture)

Since FME is a GIS format neutral Spatial ETL tool, it handles the features being read and transformed in a unique way, unique than most common GIS translators the market. Every time FME reads any dataset for a typical Spatial ETL operation, every feature read is tagged with a "Feature Type", "Feature Attribute" and "Feature Geometry". The Feature Type reflects the "grouping" of features in a dataset and is format dependent. The Feature Attributes hold the user data, format specific data and FME internal specifications. The Feature Geometry tells FME the physical geometry and the intent geometry of the feature being read and transformed. Functions and Factories enhance and add sophisticated feature processing capability thereby allowing data set to be transformed into the structure the user or the application needs.

A case of CAD to GIS interoperability: The "SAFE" way

Let us discuss a case of CAD to GIS interoperability with a simple example of Bentley's MicroStation DGN (version 7) as the CAD format and MapInfo's TAB as the GIS file format. Let us take the case of a MicroStation DGN file containing parcel data and other base map features of a City. The attribute information for these graphical features in Microsoft Access is linked through unique "MSLINK" values attached to the DGN elements. The intention is to migrate this data into MapInfo's TAB file format. Any common GIS translators can do this job. But when it comes to adding value, no doubt, FME is the right tool!

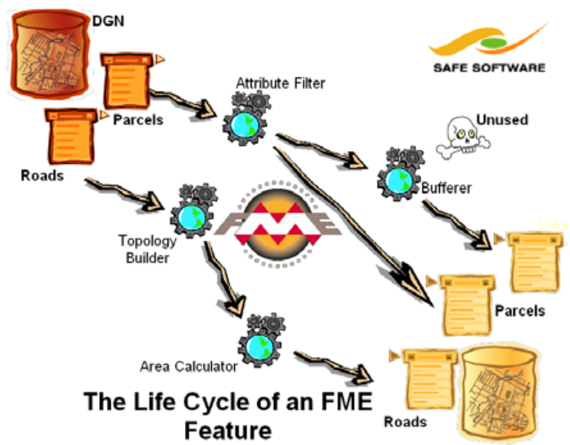
The idea here is to demonstrate the sophisticated feature processing capabilities in FME, when a user transforms data between these two systems by specific references to the Functions and Factories in FME. They can be categorized as:

Calculators: Calculate a value and supply it to a new attribute on a feature.

Example: The Calculators can be used to perform calculations like parcel acreage for the parcel polygons in the DGN file and then add this as an attribute to the TAB file in MapInfo.

Collectors: Operate on collections of features at a time. The collection of features may be replaced by new features based upon them, have their attributes or geometries merged, or have their orders altered.

Example: If the end user requires that the parcel polygons in the MapInfo TAB file be aggregated into multi-part features based on the area code in the parcel table, an "Aggregator" in FME can group the parcel before writing to MapInfo's TAB format from DGN.



Database: Allow interaction with external databases such as ODBC, Access MDB, and direct Oracle SQLNet to extract data and merge into the feature stream, or merge onto features.

Example: The attribute data for the parcel polygons in the DGN file can be extracted from the corresponding MS Access table by reading the MSLINK values and the required attributes attached to the data written in MapInfo TAB files.

Filters: Allows the feature to be routed to different destinations depending on the outcome of the tests on feature geometry or/and attributes.

Example: The end user needs the parcel polygons to be filtered based on the "tax category" value in table in MS Access and only write parcels of particular tax category into MapInfo TAB files. Using an "Attribute Filter" or a "Tester", the parcel polygons can be filtered for the particular "tax category" and written to MapInfo TAB file.

Geometric Operators: Operate on the geometry of individual features, or groups of features. A wide variety of operations are available, including overlays, snapping, line labeling, clipping, and intersection.

Example: The parcel data in DGN files do not have the topology built. Using the "Topology Builder" the topology can be automatically built for the features transformed to TAB format.

Manipulators: Modify (manipulate) the geometry or attributes of individual features in isolation from other features.

Example: It is required to buffer the street network data in the DGN file by 100 units to visually analyze in MapInfo, the parcels falling within the 100 units distance from the street centerline. Using the "Bufferer" the street buffer layer can be written into MapInfo TAB format by processing the street centerlines in DGN files, so that the data an be readily analyzed in MapInfo.

Strings: Operate on character strings held in FME attributes for searching, replacing, changing case, and extracting character encoding from strings held in FME attributes.

Example: It has come to attention that there has been a data entry error in the parcel table for the landuse code. For the residential landuse, instead of "RL" the code "RC" has been entered which needs to be rectified while transforming data from DGN to MapInfo. Using the "StringPairReplacer" in FME this can be easily rectified without the need for editing in MapInfo.

Conclusion

The examples discussed herein are a drop in the ocean of what FME can contribute to GIS interoperability. There are much more sophisticated functionality in FME that provide an edge to the users of FME worldwide by taking data transformation one step ahead of others not only by virtue of its ability to break the format barrier, but also to add value during the transformation and provide data in the structure the user or the application needs. FME may not be the paragon of all virtues but it could be important as one of several measures for addressing the geospatial interoperability issue confronting the GIS community of today. Constructive and creative suggestions are the need of the hour. Let us work towards finding ways and means of applying GIS interoperability tools such as FME to handle this burning issue, a major hurdle in building a geospatial data infrastructure.

GIS interoperability has been a moving target inspite of the laudable efforts of notable organizations such as Open GIS Consortium (OGC), Global Spatial Data Infrastructure (GSDI), Infrastructure for Spatial

Information in Europe (INSPIRE), National Spatial Data Infrastructure (NSDI).

There is still a long way to go on this route to interoperability and until then in this "SAFE" route to GIS interoperability, ***"GIS file formats don't matter!"***

Acknowledgment

The author records his deep sense of gratitude to Dale Lutz, VP of Product Development and Don Murray, President of Safe Software Inc, who have been a constant source of encouragement in providing an opportunity to the author to enhance his knowledge in this subject matter and to express his views on their Spatial ETL tool, viz. FME.

The author expresses his special thanks to the support team at Safe Software for their excellent support and for sharing the technical expertise of the FME.

The author is thankful to Mr. P. V. Rai, Managing Director, Pixel Infotek Pvt. Ltd., Bangalore, India, for his encouragement and moral support.

Last but not the least the author wishes to remember his tutors at the Division of Urban Systems Development, Anna University, Chennai, for being responsible for his first step into the world of GIS.

References

National Spatial Data Infrastructure
http://www.nsdiindia.org/about/back.htm

Joe Francica (2003), "Product Overview: Safe Software's Feature Manipulation Engine (FME)", Directions Magazine
http://www.directionsmag.com/features.php?feature_id=82

Directions Staff (1999), "Data Formats Don't Matter", Directions Magazine
http://www.directionsmag.com/article.php?article_id=33

Brig V. N. Nerikar, "Development of a National Geospatial Data Infrastructure – Theories and Technologies", GISdevelopment
http://www.gisdevelopment.net/policy/india/technology/intech037.htm

White papers on the theories behind semantic data translation and the processes, Safe Software Inc
http://www.safe.com/solutions/whitepapers/index.htm

Preetha Pulusani (2003), "Interoperability – Trend or Reality?", GIS development
http://www.gisdevelopment.net/magazine/gisdev/2003/september/itr.shtml