Geographic Information System (GIS) for Managing Survey Data To the Development of GIS-Ready Information

Zainal A. MAJEED, Malaysia and David PARKER, United Kingdom

Keywords: Geographic Information System (GIS), survey data management, ArcGIS, spatial modelling, geospatial development

SUMMARY

This paper describes the use of current geospatial handling technology to manage captured and processed survey data and other information in the Department of Surveying and Mapping Malaysia, abbreviated as JUPEM. The data empowerment consequently advances to the stage of producing and delivering of 'GIS-ready' information using current GIS technology. Basically, two management tools based on GIS fundamentals are used to carry out the objectives. Taking a practical, rather than conceptual approach, particular focus is placed on the need for the identification and application of existing geospatial standards and new available GIS technology in addressing some of these issues. A range of solutions have been developed for the management of survey and mapping information covering cadastral data, topographic and mapping data, aerial photographic images and geodetic information within Kuala Lumpur, the capital city of Malaysia. A modus operandi flow line is designed and tried to particularly fabricate a number of GIS-ready information out of the managed data and an on-line map visualisation and query of these geospatial data. The end product of the GIS is expected to resolve the issues of data integration and delivery from multiple heterogeneous data sources. The research employs current GIS software, ArcGIS and other spatial data handling application to carry out the task of data management, transformation and delivery. As widely as possible existing standards, current geospatial data handlings, protocols and GIS technologies have been used to develop solutions to the issues.

Geographic Information System (GIS) for Managing Survey Data To the Development of GIS-Ready Information

Zainal A. MAJEED, Malaysia and David PARKER, United Kingdom

1. BACKGROUND AND MOTIVATION

According to records, JUPEM commenced its survey operation under the Torrens System in 1885 when the first Johor Survey Department was established [2]. However, land surveying was carried out earlier than 1885 after the British arrival in Malaysia. Survey records existed only in paper and films thus were not well-managed. Government projects, national developments and researches using and investigating historical survey data were costly, time-consuming, and staff dependent. To better manage and allow more efficient access to survey-related data for both staff and customers, it is time that JUPEM should totally accept the task of compiling historic and notable survey-related information into digital format. The formatted data can then be easily available across the organisation by linking them to a new database or digital file storage. With these interfaces, time to explore historic data can be reduced relative to manually searching the paper records.

Currently, the department have wide range of 'geomatics' data captured and processed for land administration and various range of usages. Many of the datasets are scattered over the organisation and across the country in the state department. These data sources are often not easily available to either the public or the staff. A research task should be anticipated to bring together what is believed to be a fair sampling of all survey works that cover varied aspects of basic geography and GIS applications. It is trusted that with the new design of data assembling and computing management, the principles and properties of these datasets would have greater longevity and advantages to the geomatics and GIS world.

JUPEM is divided into sections that deal with survey-related information covering Geomatic fields of geodesy, aerial photography, cadastral, topographic mapping. Even though a digital data management system does exist in some sections of the organisation, it is still problematic. Each database has its own discrete interfaces and staff to be skill and proficient in each of them. Great expenditure is needed to train staff and customers. These stand-alone entities and interfaces allow searching of all relevant data for a given area of interest in separate database, which is time consuming. At this stage there is no management system that gathers all survey-related information into one source. Figure 1 shows the vision of the organisation to have all data accessed in a single view.

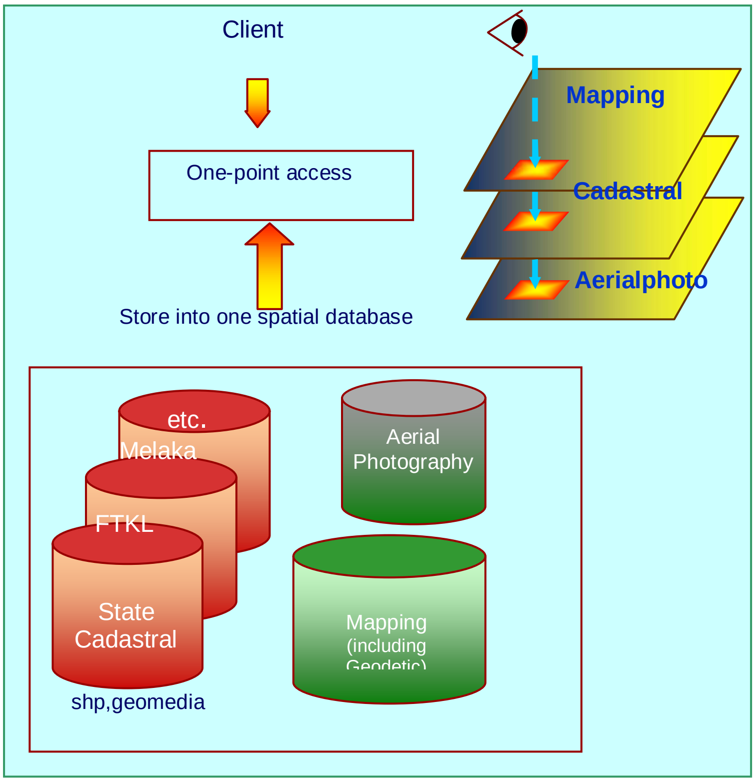

Figure 1: Context diagram of the single portal access of the various data

A goal for JUPEM should be to formulate a more efficient system of accessing data for data users and providers leading to a reduction in time for emergency procedure, decision making and national infrastructure development. GIS technology is one way foreseen to resolve this issue. Generally, positive participation by JUPEM in the field of GIS is expected. To have GIS functionality into the domain, it is necessary for the organisation to transfer the spatial description of existing historical paper survey information into spatially accurate GIS data files. There will be a probability of entering survey data to the point level rather that the polygonal level so that data related to a given historical survey could be readily examined on the GIS map-based software.

Because most sections in the organisation conduct data in tabular form, spatial feature-based queries were impossible or difficult. Spatial feature-based searches can enormously increase productivity and accuracy in survey-related enquiries. This is especially significant to staff or customers who are interested in locating data relevant to a given land parcel, highway or within administrative boundaries. GIS capability supports user-friendly map-based interface that are valuable for users wishing to visually scan available data over a given area of interest.

With the survey-related data management system on the shelf, a prospect of the development of GIS can be realised. The existing huge digital topographic data in the department presents an immensely useful database for GIS [4]. However, the ongoing development would go a long way to meet the requirement of GIS because it has limited capability for processing non-graphic attributes. Truly, GIS system combine graphic capabilities with strong non-graphic attribute linkages allowing complex queries, map overlay, polygon processing and geographic modelling operations. The department relies on the Computer-Aided Design (CAD) technologies and published hardcopy maps to meet public geospatial need. The demand for GIS data increases as time goes by, and JUPEM foresees taking the leading role in providing this spatial data. The department realises that it needs to develop an enterprise GIS and mapping system that produces sophisticated, integrated spatial information for the national's decision makers (in various government departments) and for the public as well. It is important to have a system facilitating seamless access to a variety of applications and formats. It is expected that thousands of existing CAD files can be automatically translated into GIS datasets, and other JUPEM's geomatics datasets can be managed using a range of GIS formats.

The development of a National Spatial Data Infrastructure (NSDI) enumerates JUPEM involvement in playing a leading role in the implementation and operations (6). The commitment is shown by the devotion of significant amounts of senior management and expertise in its main operation. The Director-General of Surveying, two senior Directors and three Directors in the JUPEM Headquarters were appointed to assist and evaluate feasibility study in the Coordinating Committee. Five senior Land Surveyors, six fresh Land Surveyors and supporting staff were seconded to the Secretariat, and all states directors are member of the Coordinating Committee. By introducing the concept of databases with full implementation of modern computer technology and their pivotal role in the development of NSDI, JUPEM should initiate a reform which has wide ranging implications for the department and the country as a whole.

With the deployment of digital field survey and office equipment, the department is now poised an important role in the establishment of land related information systems in the country in support of the government's effort in establishing an "Electronic Government" and a knowledge-based K-economy. The organisation is therefore carrying a weighty burden to fulfil its mission of providing digital spatial data as well as the geodetic reference framework in support of the geographic and land information systems in the country. It is hoped, at least for a start that the data from the topographic mapping database be used by other agencies as the basic geospatial data that can be utilised in establishing their respective database for GIS implementation. Therefore by the end of the day JUPEM should anticipate that a "master plan" of a contiguous coverage of the whole nation can be achieved and be the pride of the organisation.

In line with the government's effort to push Malaysia to achieve developed nation status by the year 2020, there are various initiatives that have been drawn up to bring the country closer to that objective. One of these is the e-government initiative of using technology for the improvement of many services and connectivity of various products rendered by JUPEM.

JUPEM has about 1.9 Terabytes of digital spatial data which are available for government, business, and public and individual's consumption (2). These data are stored in various seamless databases, data files and image files. These datasets should be able to serve communities who use geospatial data and information for businesses and developments.

There are a number of applications using GIS technology used in organisations to manage, maintain and distribute their data. Organisations maintain and provide specialised data according to their functions; however in many cases users or even organisations themselves need other datasets for a particular application (1). A major dilemma that is still persisting in the GIS community nowadays in Malaysia and some developing countries is the missing of geospatial data and information as well as the geospatial data services; it is how and where to find and access it. The growth of Internet access and use coupled with advancements in web based technologies over the past decade has provided new possibilities for the access, delivery and use of geospatial information (GI) (5). In recent years the GI sector has begun to recognise the importance and role of the web for the dissemination of spatial information, with many GI technology vendors now offering extended systems of Internet Map Server (IMS) to their desktop products e.g. ArcIMS, Geomedia, GE Smallworld IMS. The development of such systems has introduced and highlighted issues pertinent to the use of GI via the web.

Traditionally, a gap has existed between surveying standards and practices and those in GIS (3). GIS specialists had expressed that surveying community are slow in achieving map and data product, not to mention expensive and unfinished artefacts. They instead produce their own data which do not compromise to the standard of formality in which surveyors respond that clients are confused and given with incorrect norm of data as well as map. This creates misunderstanding and surveyors esteem may be challenged in term of their public mission and work. This mental disagreement at least can be tackled by some available software technology which surveyors can easily incorporate their measurements and calculations into GIS databases that serve all sections and applications in an organisation. Survey data should therefore be stored in a GIS environment.

2. AIM AND OBJECTIVES

The aim of this research is to manage captured and processed survey data within the associated organisation through the development of approaches to achieve information that is "GIS-ready" and eventually to disseminate the end product geospatial data over the Internet using current geospatial handling technology via all appropriate standards. Throughout the research the main objectives can be given as follows:

- To review the type and format of raw and processed survey and mapping data which were held in the Department of Survey and Mapping in Malaysia and suggest the vision of the end product of these data.
- To construct a survey data management system to facilitate the combination of data and information from raw and processed survey data from different seamless databases and sources within the JUPEM organisation using GIS technique.
- To design a flow line and modus operandi to remodel and transform the managed survey data into GIS-ready information using contemporary geographic information application and technology.

- To develop an on-line geographic information system to facilitate the delivery of geospatial data via the Web to meet the needs of corporate Intranet and demands of worldwide Internet access.

3. TEST SURVEY DATASETS

The Department of Survey and Mapping Malaysia (JUPEM) is among the main government organizations providing high quality spatial data, survey and mapping products and services to the government, business, public and individuals for the purpose of national development, security and defence. Three types of datasets were acquired for the implementation of the research task. They are:

- Cadastral dataset, in Environmental System Research Institute, Inc. or ESRI's format shapefile (SHP), a data set type that stores geometrical shape consisting of a set of vector coordinates and attribute information related one-to-one with the associated shape. This dataset has spatial reference in local coordinate system.

- Topographical mapping data which are produced and stored in vector file format namely Data Exchange File (DXF), a two-dimensional graphics file format supported by virtually all CAD products.. The Rectified Skew Orthomorphic (RSO) coordinate system is assigned for its spatial reference.

- Digital aerial photographs which are scanned and stored in raster form, Joint Photographic Experts Group (JPEG), non-rectified and not spatially-referenced.

These datasets are collected, stored, and maintained in several sections as well as in different systems. They are chosen for use in the early part of the research due to their heterogeneous and proprietary nature in terms of format, resolution, and source, amongst others. They are thought to highlight issues and provide a basis for exemplifying possible solutions to the problems of GI accessibility within the enterprise and distribution via the web. Various vector and raster representation of survey-related information covering Global Positioning System (GPS) stations, Levelling Network, Malaysian Active GPS Station (MASS), Gravimetric Reading Station, Triangulation Station and Control Traverses will be included in the later stage of the research.

4. MODELLING RAW AND PROCESSED SURVEY DATA

Geographic feature is defined by International Standards Organization (ISO) as an "abstraction of a real world phenomena", and a feature attribute as a "characteristic of a feature". Thus, survey data in points and lines are geographic feature which is a real world phenomena with a feature attribute that defines its dimension and location in space. We use a Structured Query Language (SQL) database to manage these data where features are managed in tables, an instance of a feature corresponds to a row, and an attribute of a feature to a column.

The spatial attribute of a geographic feature is defined by its dimension and location, and is referred to as the feature's geometry. The dimension of a geometry describes its form in space. The location of a geometry in space is defined by its coordinate system. The coordinate system contains information about the number of coordinate values, the mathematical rules for projecting the geometry coordinates, and how the coordinate system is related to a datum on the earth's surface.

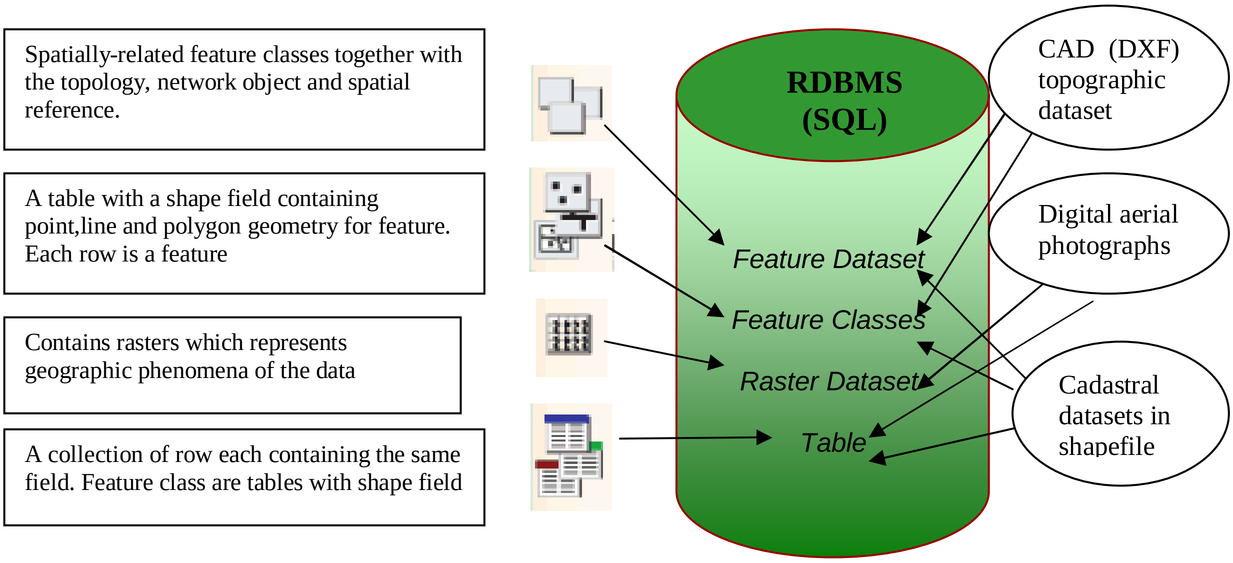

Figure 2: The structural GIS elements that are used to develop the geographic data model of the test dataset. The geographic geodatabase represents a generic data model for the test data.

Using standard relational database management system (RDBMS) technology, a data model for representing the test dataset as spatial information is created and called geographic database, in shorthand, geodatabase. This data model acts as a storage of geographic data implemented with the relational database, in this case, SQL database. All database elements are managed in standard RDBMS tables using standard SQL data types. In one case, a table is used to store topographic feature classes where each row in the table represents the feature. Each row in the table has a shape column used to hold the geometry or shape of the feature. This fundamental relational storage model is adhered to the Open GIS Consortium and the ISO simple features specification. All the topographic, cadastral and raster images are managed and stored in the relational tables. The table configuration for assigning the features in the geodatabase model allows the opportunity to manage all the spatial data in one database system. The geodatabase model is illustrated as in Figure 2.

Geospatial data handling tool ArcGIS Desktop is used to create and work with the geographic data in the relational geodatabase. ArcMap provides a complete set of tools for working with the data in the database using the interface application for mapping and editing the geographic feature in the model. Applications for managing and geoprocessing geographic datasets are ArcToolboc and ArcCatalog that help to create and manage the datasets between the relational geodatabase and the application protocol. They are a powerful suite of applications and capable of interfacing each other and working together to perform all GIS tasks. These GIS applications are utilised to transform and manipulate the raw surveyed data through the production of GIS ready information which then is stored and managed for intelligent GIS usage in a relational database.

Data access uses the standard model defined in the SQL. The access is enabled on the traditional model supporting spatial and raster constraint in a query. A database connection is a connection to the source geographic data. It represents the server which can respond to request for geographic data. Spatial Data Engine (SDE) technology is used as the interface gateway to the relational database. ESRI's SDE called ArcSDE is employed to manage the test data in the SQL server which can then serve the data to the GIS applications and Internet Map Service as well.

5. TRIAL IMPLEMENTATION

A few digital aerial photographs were stored and loaded using ArcSDE as the gateway into RDBMS SQL server. ArcSDE provides fast and scalable capabilities to access and manage raster data through ArcCatalog and ArcToolbox interfaces. The images in JPEG format were structured in the database using the raster catalog tool in a list of raster catalog. ArcSDE database supports various raster format and provides ease of interchange geospatial data between files storage and geodatabase or enterprise database.

High resolution aerial photographs acquired from Aerial Photography Section were rectified via ground control point (GCP) from digital topographic features using ERDAS IMAGINE software. After given a spatial reference, the photographs were mosaiced producing a larger image of the test area. It was then gridded into small bounding rectangles covering the relevant area for the implementation of the application. During the process it has to be transformed from raster to 'tagged image file format' (TIFF) file to Imagine (IMG) format and finally to JPEG format file to achieve a lower resolution and compact file size image in order to distribute on the network. These data were then stored back in ArcSDE database as GIS-ready information for overlay operation or as a background for vector visualisation.

The topographic vector data, in six map sheets, obtained from digital mapping section are captured in DXF file format. It contains features of point, line and polygon geometry describing various dataset categories namely boundary, water, building, relief, transport, utility, vegetation and landuse. Spatial reference was assigned in the original datasets with real world coordinate which is then defined in ArcMap as RSO type. ArcSDE capabilities was used to manage the CAD data which can be read as a single background layer (CAD Drawing Layer) or as a collection of point, line, polygon, and annotation feature classes (CAD Feature Layer). CAD layers can be viewed and displayed individually in ArcMap thus enable the ease of selecting which feature to convert or manipulate. Figure 3 shows the CAD layer overlay the rectified and mosaiced raster image. They can be accessed and displayed directly from ArcSDE interface database.

Feature Manipulation Engine (FME) software is used to convert the CAD file into shapefile before it can be stored and manipulated in ArcSDE database. In order to edit the polyline CAD files and form object-oriented features the layers are translated using FME Workbench interface. The translated data is assigned a group layer of feature datasets for GIS display covering contiguous map coverage. The source datasets were accessed from ArcSDE database and transferred into the same database directly during the translation. Editing is carried out using the extension tool of ArcInfo and ArcEditor. Figure 4 shows the diagrammatic flow of the translation in FME.

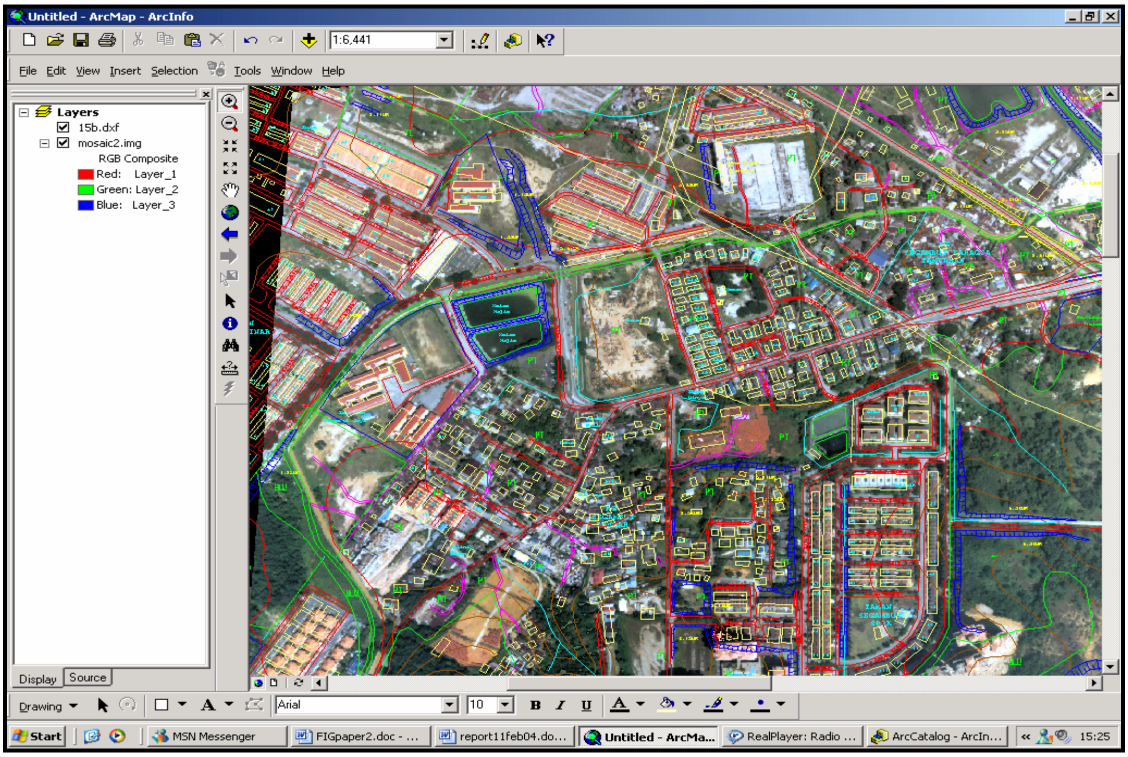

Figure 3: Original AutoCAD topographic data overlayed on raster data for managing the data manipulation to produce contiguous map objects for GIS usage.

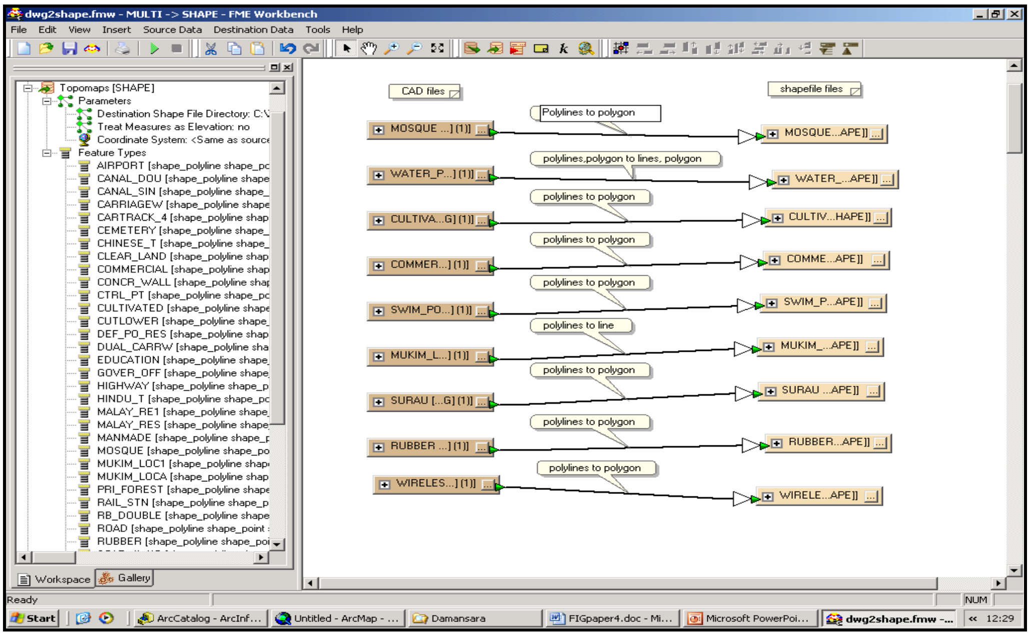

Figure 4: The operation model of FME Workbench in translation of geospatial data

Cadastral data of part of the Kuala Lumpur, capital city of Malaysia was utilised as a land administration and survey data coverage. The spatial information in the land parcel data includes land parcels boundaries, boundary markers and coordinates of boundary markers. The attribute information includes unique parcel identifier, lot number, parcel area, surveyed bearings, and surveyed distances, type of boundary marker, date surveyed and date approved. Since the data was in their local Cassini Soldner coordinate system, transformation of coordinates to RSO coordinate system was carried out. The coordinate transformation was undertaken in ArcGIS. Figure 5 is the result of the transformed cadastral data overlay on the raster aerial photograph. The overall flow line process of the datasets is depicted in Figure 6.

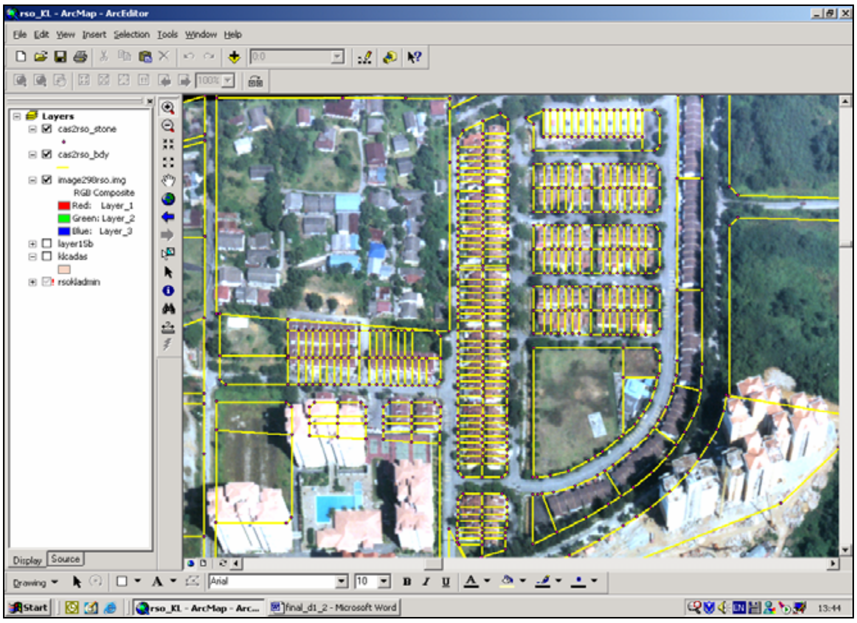

Figure 5: A spatially-referenced cadastral data as managed on raster data

6. OUTCOME OF THE TRIAL IMPLEMENTATION

The sample dataset achieved after trial implementation shows that the GIS-ready information can be produced using the RDBMS geodatabase and ArcGIS suite of application. ArcSDE gave a useful interface to fast access and effective management. Figure 7 shows the appearance of the end product when all the survey data have been populated into the relational database and current geospatial data handling functionality was used to explore the dataset within SDE Application Protocol Interface (API). The CAD data were transformed into contiguous coverage to enable spatial visualization and search within the application environment. Resultant cadastral land parcel polygon enables spatial and attribute search when seeking land property information using specific query.

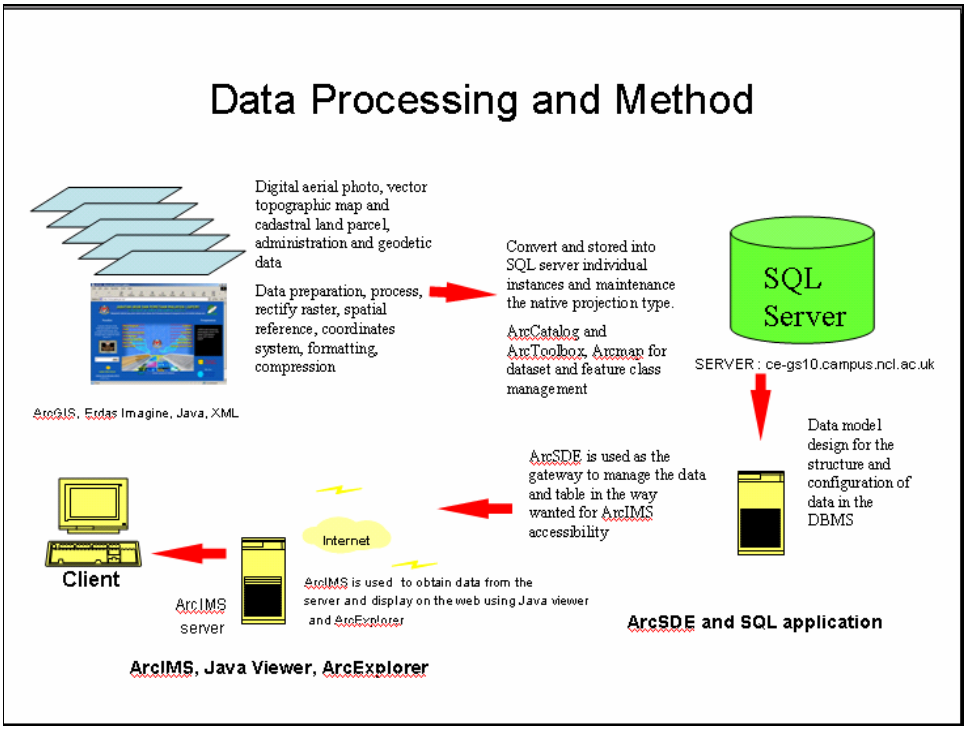

Figure 6: The flowline of the data processing, data populating, conversion and dissemination

Using ArcSDE connection, geographic datasets can be visualized across the Internet using the ArcIMS application interface. This was achieved as in Figure 8, except that the large raster images would need be compressed in order to be displayed in ArcIMS application viewer. Cadastral and topographic data in their raw or GIS-ready data can be accessed and query for spatial display. Different features can be displayed and explored for specific analysis. The implementation of map services is only at the early stage and a lot more needs to be done especially for application and web page customization for interactive usage.

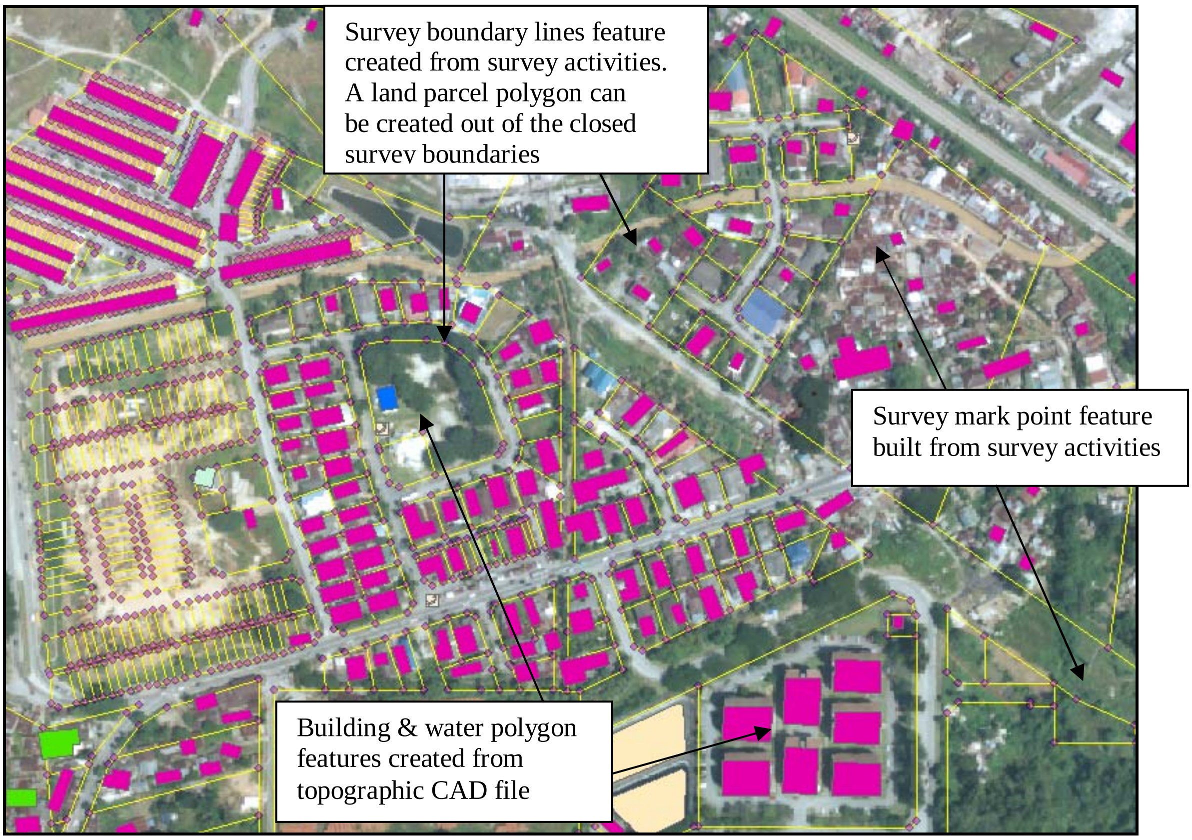

Figure 7: The sample datasets in the tested application. The project contains three GIS-ready datasets.

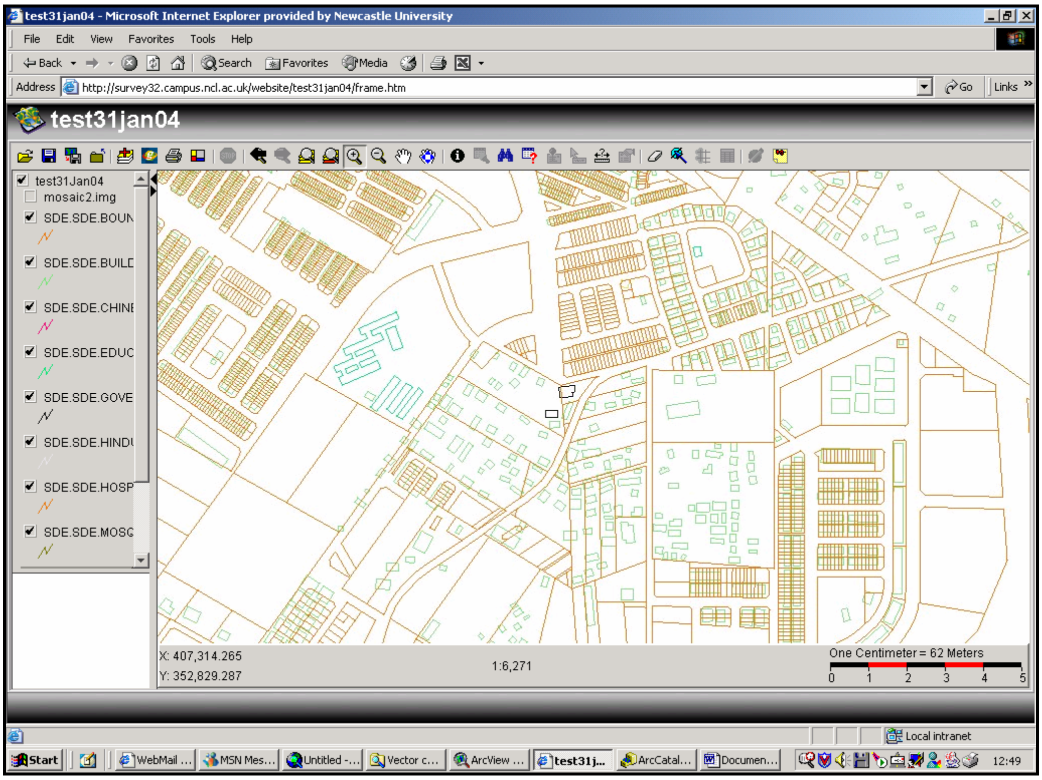

Figure 8: A test site shows on-line representation of the datasets which are stored in different servers.

7. CONCLUSION

This research was intended to introduce a design system to manage survey datasets through the production of GIS-ready information using appropriate standard and computing application. The trial implementation does instigate sufficient results at present stage whereby the test datasets consisting of raster image and feature classes were being managed carefully through the platform of producing and delivering GIS-ready information. However there is still a need for an improved flow line of the process as more dataset type and volume covering other survey datasets held in a survey organisation would be used. The testing of the design and flow line has clearly shown the possibility to disseminate, retrieve and combine those data for visualisation and query over the web from multiple different data sources. ArcSDE tool enables the operating department to keep using their existing proprietary system without physically deposits all required data into a single system. ArcGIS functionality is proved offering capabilities for geospatial data interchange, manipulation and management as well. The ArcGIS application has clearly shown the successes of the concept of data integration on-the-fly from multiple heterogeneous geospatial data servers. The ArcSDE interface is discovered to be well-off and powerful to overcome obstacles in a timely fashion, effectively and manageable.

ACKNOWLEDGEMENT

The authors would like to thank the Director-General of the Department of Surveying and Mapping Malaysia (JUPEM) in Jalan Semarak, Kuala Lumpur and the respective officers for preparing and supplying the useful datasets.

REFERENCES

Chunithipaisan, S., Majeed, Z.A., James, P., and Parker, D., Abele, S. (2003). Geospatial Interoperability via the Web: Supporting Land Administration in Kuala Lumpur. *Proceedings of MapAsia 2003 Conference*, Kuala Lumpur, Malaysia, 13-15 October 2003. http://www.gisdevelopment.net/proceedings/mapasia/2003/lis/index.htm

Department of Surveying and Mapping Malaysia Website.
http://www.jupem.gov.my.

Environmental Systems Research Institute (ESRI) 2003. Brochures on ArcGIS Survey Analyst Extension Software, 2003.
http://www.esri.com/library/brochures/pdfs/surveyanalyst.pdf

Kassim, M.M., Kadir, R.A. (1989). The development of national topographic and cartographic data bases for geographical information system (GIS) implementation in Malaysia. *Proceeding of Asian Conference on Remote Sensing*, Kuala Lumpur, Malaysia, 1989.
http://www.gisdevelopment.net/aars/acrs/1989/S/pls001.shtml

Open GIS Consortium. The Open Implementation Specifications. http://www.opengis.org/techno/implementation.htm

Tong, C.W. (2001). Department of Survey and Mapping Malaysia: The Major Building Block for NaLIS. *Open Seminar on Spatial Data Infrastructure in Asia and Pacific,* 7th

Permanent Committee on GIS Infrastructure for Asia and Pacific (PCGIAP), Tsukuba, Japan, April 2001.
http://www.gsi.go.jp/PCGIAP/tsukuba/seminar/seminar.htm

BIOGRAPHICAL NOTES

Zainal A Majeed
The author is currently studying research PhD in geospatial data management and Internet spatial data delivery to gain experience towards the programme of establishing NSDI in Malaysia. He has working experience of more than 20 years as a land surveyor in the Department of Surveying and Mapping Malaysia, the National Land Information System (NaLIS now called MaCGDI) and the National Institute of Land and Survey Malaysia. He has been a member of the Institution Surveyor Malaysia (ISM) and is a registered land surveyor to practice under the Licensed Surveyor Board, Malaysia.

Professor David Parker
He is a Professor in Geomatics at the University of Newcastle upon Tyne, currently holding the post of the Head of School of Civil Engineering and Geosciences.

CONTACTS

Zainal A Majeed and David Parker
School of Civil Engineering and Geosciences
Cassie Building, University of Newcastle upon Tyne
Newcastle upon Tyne, NE1 7RU
UNITED KINGDOM
Tel. + 44 191 2226 544
Fax + 44 191 2226 502
Email: Zainal.abdul.majeed@ncl.ac.uk, david.parker@ncl.ac.uk
Web site: http://www.ceg.ncl.ac.uk, http://www.jupem.gov.my