

Provincial Spatial Database Policy

GLOSSARY OF TERMS	3-4
INTRODUCTION	5
BACKGROUND.....	5
PSDB STRATEGY.....	6
SDE STRUCTURE AND DESIGN.....	6-7
STANDARDS.....	7-8
DATA MANAGEMENT.....	8
SPATIAL DATA MANAGEMENT SYSTEM.....	8
DATABASE MAINTENANCE.....	9
SECURITY	9
DATA SHARING METHODS.....	9-10
SDE ADVANTAGES	10
SDE DISADVANTAGES.....	10
TRAINING.....	10
CONCLUSION	11

Glossary of Terms

Co-ordinate System – it is the system whose units and characteristics are defined by a map projection. A common coordinate system is used to spatially register geographic data for the same area. See map projection.

Data Custodian - is an organ of state as recognized by the Committee for spatial Information, having the responsibility to ensure that a base data set is collected and maintained according to specifications and priorities determined in consultation with the Committee for Spatial Information (CSI) and the user community, and being made available to the community in a format that conforms to standards and policies established by the CSI.

Data set – is an identifiable collection of related spatial information.

Maintenance - is the updating or modification of spatial information to ensure that it remains usable and reliable.

Map Projection – is a mathematical model for converting locations on the earth's surface from spherical to planar coordinates, allowing flat maps to depict three dimensional features.

Metadata – is a description of the content, quality, condition and other characteristics of spatial information.

Provincial Spatial Database (PSDB) – is a database which stores all relevant provincial spatial data. It is stored in a Microsoft SQL database and uses ESRI's ArcSDE technologies.

Point – is a single x,y coordinate that represents a geographic feature too small to be displayed as a line or area e.g. a town on a small scale map.

Polygon – is a vector representation of an enclosed region e.g. Municipality boundaries.

Quality - is the degree to which spatial information which has been captured or collected satisfies stated or implied needs and includes geographic information about lineage, completeness, currency, logical consistency and accuracy of the spatial information.

Raster Data – is an abstraction of the real world where spatial data is expressed as a matrix of cells or pixels, with spatial position implicit in the ordering of the pixels.

Relational Database Management System (RDBMS) – is a database management system with the ability to access data organised in tabular files that may be related together by common field.

Spatial Information Management Unit (SIMU) – is the unit within the Office of the Premier (Eastern Cape) which deals with spatial information within the province.

Spatially related Data - is information about spatial objects or features and their attributes.

Vector Data – is an abstraction of the real world where positional data is represented in the form of co-ordinates. In the vector data the basic units of spatial information are points, lines and polygons.

Introduction

Data integration, sharing, quality control and duplication are challenging issues for the departments within the Eastern Cape. It is therefore crucial to promote the importance and benefits of sharing spatial data as stated by Spatial Data Infrastructure Act (SDI).

A central repository for storing provincial spatial data has been identified as a basic concept for the effective coordination of spatial data integration, management, maintenance and sharing. This central repository will reside in the Spatial Information Management Unit (SIMU) within the Office of the Premier. The chosen/preferred software is ESRI's ArcSDE with Microsoft SQL as the Database Management System (DBMS).

Definition: Spatial Database Engine (ArcSDE) is a tool that provides a gateway between the Geographic Information System (GIS) and the DBMS to share and manage spatial data as tables.

Background

Government departments in the Eastern Cape Province have been engaged in numerous projects, collecting and storing various spatial datasets in different file formats. This has led to the duplication of data, and ultimately wasted efforts. In general there has been no formal structure used to store specific departmental and project specific spatial data. Due to lack of organization and non existence of spatially related data standards, users experience difficulties when using a specific dataset.

The Spatial Information Management Unit (SIMU) of the Office of the Premier introduced the concept of developing an Enterprise Spatial Database solution for the province, which will enable dissemination and storage of spatial data. SIMU is responsible for developing strategies for organizational spatial data standardisation, integration and ensuring consistency between Government departments and private sectors. All the existing datasets in the Eastern Cape Province will be stored in ARCSDE, software that enables storing and management of spatial data. This spatial database will be maintained by all departments as regards to their individual information and managed by SIMU through the GITO Council. SIMU will be responsible for the awareness of existing datasets through the GITO Council. The datasets stored within the Provincial Spatial Database (PSDB) will be viewable via the Internet Mapping Service (IMS) and will be accessible in various media formats. In providing the PSDB, it is envisaged that efforts in GIS initiatives will be more effectively coordinated, which will allow for more timely and accurate strategic planning.

PSDB STRATEGY

The office of the Premier co-ordinated a workshop for all stakeholders in order to identify problems, initiate actions and generate strategies and policies which will facilitate the establishment of a standardised provincial spatial database for storage and management of spatially related data in the Eastern Cape Province. The following risks were identified during this workshop:

- There is no clear distinction between what should be classified as base data or sector specific data.
- Some stakeholders do not know what datasets are available and where to acquire them.
- There is insufficient buy-in from certain stakeholders and decision makers to establish a standardised spatial database structure.
- Stakeholders are not well trained to use or implement spatial database structures.
- Poor communication between stakeholders regarding existing or planned datasets.
- No or poorly defined standards for data formats and data accuracy between the various stakeholders.
- Metadata is poorly defined and not well documented.
- Most datasets are poorly maintained.

The SIMU has “unpacked” these risks and developed a strategy to address them.

Provincial Spatial Database Structure and Design

The PSDB makes use of Microsoft’s SQL Server technologies for an effective integration of spatial data, as the data is loaded into SQL Server through ArcSDE (Spatial Database Engine). The Provincial Spatial Database will be hosted on two servers: a development server and a hosting server. The development server will be used for all the processes which will be carried out whilst it’s still in the development stage. The second server will act as the final storage /hosting space for the different datasets which have been cleaned and approved (checked for quality) by the PSDB Administrator. Data Custodians/Owners will be responsible for the quality of their datasets, should any errors be detected they will be requested to clean them.

The PSDB Administrator will create users for all the Provincial Departments which are on the Provincial Common Core Network (PCCN) and assign them read/write access to the databases which will be created within SQL server.

The spatial data will be stored on different databases; each Department will have its own database e.g. one called Dept_Agric for Department of Agriculture and this database will be containing all the datasets own by the Department. See appendix I for the structure and datasets stored on PSDB.

Data custodians/owners must classify their data into **public/confidential or restricted** for the PSDB storage. All the datasets that have been classified by the custodian/owner as public data are accessible to the public with copyrights in line with the Promotion of Access to Information Act. PSDB Administrator will liaise with the copyright holder i.e. custodian/owner of the data for granting authority, e.g. Hydrology dataset whereby Department of Water Affairs and Forestry will be the data custodian of such data. Datasets that are classified as confidential and restricted wont be of access to the public, will only be accessed by those specified by the data custodian/owner.

Larger Departments with existing SDE's are advised to use their SDE's as development SDE's within their department which will allow them to manipulate their data prior to publishing it to the PSDB and making it available to all.

Standards

Standards task team was established by the Provincial Interdepartmental GIS committee to investigate and establish data standards.

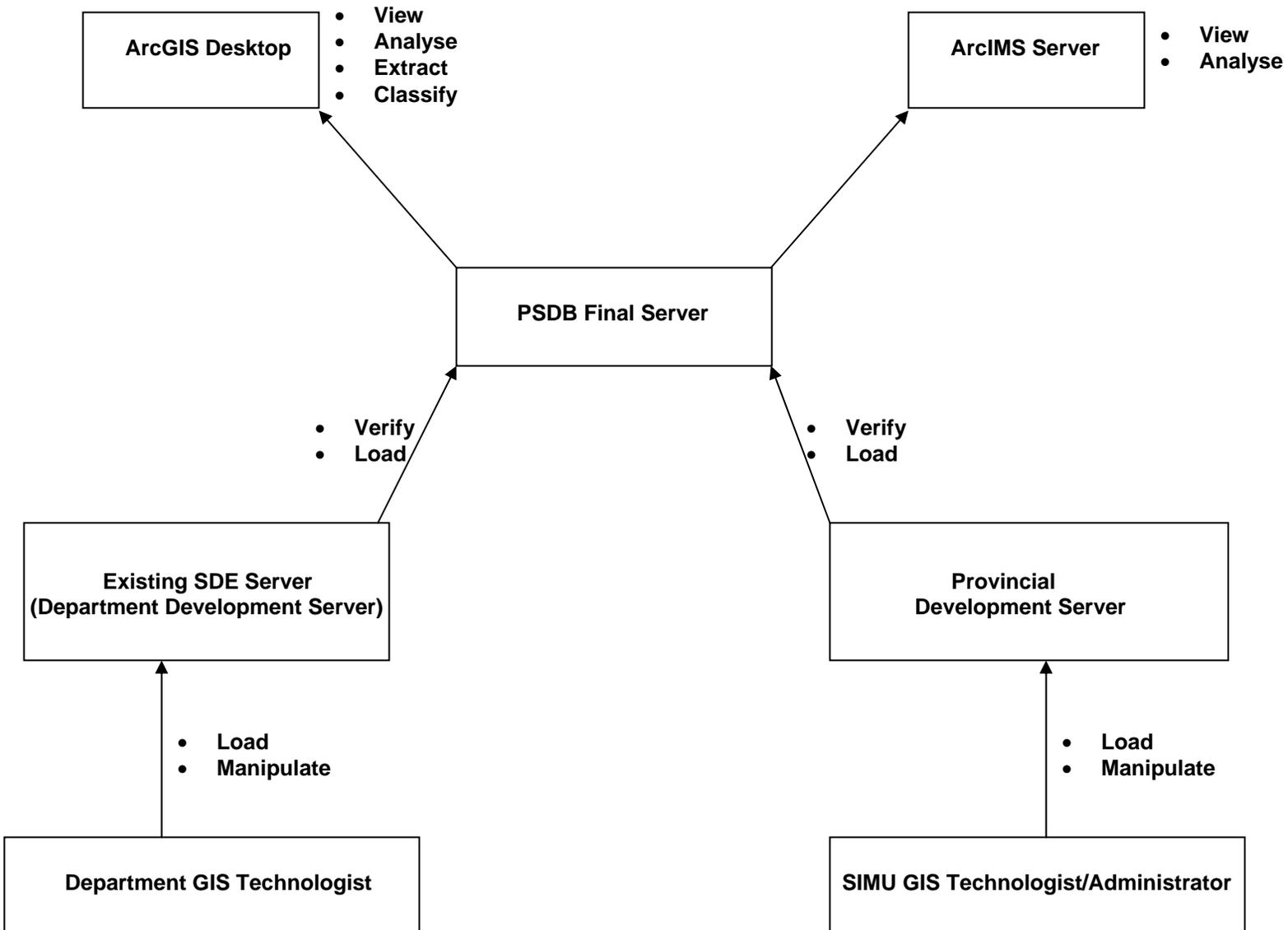
Standardisation includes:

- **A standard coordinate System:** This is very important for the efficient use and distribution of data across public and private organizations. The recommended spheroid for the Eastern Cape Province is WGS 1984 and Transverse Mercator as the map projection.
- **Metadata** is very crucial for an integrated GIS system as it describes each datasets in the database with regards to accuracy, source, currency etc. Metadata will be maintained on-line in a format which meets the requirements of the National Spatial Infrastructure Framework (NSIF).
- **Raster Catalogs:** were identified as the best way for storing the Provincial raster data as some of the images are overlapping. Raster catalogs preserve the overlap and yet still group the raster datasets together. Images on a raster catalog must have the same spatial reference, therefore any images that are not in on the standard co-ordinate system mentioned above will be reprojected to meet the requirements.
- **Standard Naming** and data conversion for attributes and feature classes
- **Symbology** standards for Buffalo City Municipality will be adopted as they have already developed their own in the Province (avoiding duplication) and are inline with Surveys and Mapping standards.

Data Management

It is crucial to establish formal responsibilities for the creation and maintenance of datasets that will be stored and disseminated for use. To avoid redundancy, a data custodian should be identified for each dataset. Each Department will have its own data custodian for its respective datasets. The data custodian will be responsible for capturing and maintenance of metadata for any spatial information held by it in accordance to the Spatial Data Infrastructure Act (SDI), 2003.

Spatial Data Management System



Database Maintenance

PSDB Administrator under the guidance of the GITO Council (Eastern Cape) will effectively maintain the database and grant privileges to authorised users. This access will be authorised by the Department GITO and implemented by the PSDB administrator. This means that only users with appropriate read-write access can edit the database.

Every data custodian/owner will have edit rights to their own dataset which will allow them to directly update their datasets on the PSDB. Data custodians/owners will be advised to maintain standard compatible software in order to be able to make edits to their datasets. They must also ensure that the spatial data they are publishing on the PSDB is compatible with the provincial data standards.

Security

SQL Server has a built in security protocol that ensures data stored in the Relational Database Management System (RDBMS) is protected from unauthorized access. This involves creating users, roles and access to specific database. ArcSDE also has a built-in functionality that allows PSDB administrator to grant roles to authorized users.

Developing and implementing a good security schema will enable Government departments to confidently store sensitive data into PSDB. This will be inline with the Provincial ISS policy.

For security and safety reasons the database will require a full backup every night. The database will also be backed up on an off-site storage as the Provincial backup policy states that all user data must be protected by means of full backup. Off-site storage will be secured and available only to SIMU staff.

Spatial Data Sharing Methods

One of the mechanisms for sharing data is through web based technologies. When the PSDB implementation is done, ArcIMS (Internet Mapping Server) serves as a method of obtaining data and also a mapping server. ArcIMS serves spatial data on the intranet/internet; with the potential for retrieving results for queries. The paths to the data for the ArcIMS maps will be redefined to view the data in the PSDB. This will ensure the data viewed on ArcIMS will be the most current and accurate.

The advantage of ArcIMS is that a user can view the Provincial spatial data using their existing web browser without being a GIS expert or purchasing specialised GIS software. ArcIMS does not replace GIS software but provide a way of sharing spatial data. ArcIMS will give spatial data widespread, easy access and sharing as it is distributed on the internet which connects people from all over the world.

Written requests will also be accepted for those who do not have access to the internet, and will be processed by SIMU. Any non-governmental organisation requesting spatial information will be requested to provide proof for use (letter of appointment from a government department) of any government owned data.

The other method that can be used is for all the stakeholders connected into the Government network to retrieve the data directly from PSDB final server through the network. The PSDB Administrator will create usernames and passwords to connect to the PSDB final server and assign privileges to different departments. This will also ensure that stakeholders are having the latest and current datasets available on the PSDB.

SDE Advantages

- Security and user access control
- Data Recovery
- Custom Topological Relationships
- Extremely large data volumes
- Reduced cost of ownership
- Ensures spatial data Integrity
- Improved coordination and management of datasets
- Minimise duplication

SDE Disadvantages

- Although ArcSDE is expensive, it is a return on investment because all government departments will benefit from it as it going to serve as a centralised storage for spatial data in the province and will be shared by all government departments. There will be one copy of data to maintain and backup.

Training

The training of users will be undertaken by SIMU. Users will be trained on how to access, display, query and analyse the data on the PSDB using ArcGIS Desktop. Users who have been assigned editing rights by the PSDB Administrator will also be trained on editing and versioning of the datasets.

Conclusion

The establishment of the PSDB will streamline data management and maintenance process for the Eastern Cape Province. It is going to serve as a spatially related data management, manipulation, analysis, modelling and decision making tool for all spheres of government.

The PSDB will serve as the backbone for the intranet and internet technologies for Eastern Cape Province through spatial data services via the use of ArcIMS.

The PSDB will assist managers in better decision making for planning and budgeting, as the large percentage of data will be spatially referenced. Managers will be able to access the latest integrated spatial data available on the PSDB from their desktop without physically going to SIMU and make data requests; as it will be accessible through the network.

Appendix 1: Eastern Cape Province Data Model