Geospatial data management within SA's national mapping organisation

by Patrick Vorster and Raoul Duesimi, CDNGI

South Africa's National Mapping Organisation, the Chief Directorate Surveys and Mapping (CDSM), was renamed the Chief Directorate National Geo-spatial Information (CDNGI) in 2009 (www.cdsm.gov.za). The CDNGI forms part of the National Department of Rural Development and Land Reform (the former Department of Land Affairs).

The CDNGI provides geodetic and topographic surveying and geospatial information services in terms the Land Survey Act (Act 8 of 1997). Products available are:

- National topographic line map series' at scales of 1:50 000, 1:250 000 and 1:500 000.
- An orthophoto map series, at 1:10 000 scale, covering metropolitan and national development areas.
- National horizontal and vertical control survey networks.
- An active Global Navigation Satellite System (GNSS) base station network named TrigNet.
- A national imagery database comprising aerial photography dating back to 1936.
- A National Digital Elevation Model (NDEM).
- International Civil Aviation Authority (ICAO) maps of Southern Africa, on behalf of the Civil Aviation Authority, at scales of 1:500 000 and 1:1 000 000.
- A topographic names database from which the nine provincial gazetteers have been produced for the South African Geographical Names Council (SAGNC).

The Directorate National Spatial Information Framework (NSIF) (www.nsif.dla.gov.za) has now been incorporated within the CDNGI structure. The NSIF's mandate is to co-ordinate South Africa's Spatial Data Infrastructure in terms of the Spatial Data Infrastructure Act (Act 54 of 2003). The Act establishes the South African Spatial Data Infrastructure (SASDI), the Committee for Spatial Information (CSI) and an electronic metadata catalogue. The CSI is required to administer and co-ordinate the capture and sharing of spatial information by identifying data custodians, specifying standards and other prescriptions to which spatial data must comply, capturing and publishing metadata of spatial data and avoiding duplication of spatial data capture.

Fig. 1: The NGI Offices in Mowbray, Cape Town.

Committee for Spatial Information

The CSI consists of elected members representing the National Department of Rural Development and Land Reform, Statistics South Africa, the National Department of Provincial and Local Government, each of the provincial governments, local municipalities (rural and urban), the Council of Government Information Technology Officers, geographic information science professional associations, geographic information science teaching/research institutions, each of the identified data custodians and other appropriate public entities referred to in Schedules 2 and 3 to the Public Finance Management Act, 1999 (Act 1 of 1999).

Fig. 2: 1:50 000 map extract.

The CSI is required to:

- Identify data custodians of spatial data in the public domain.
- Specify standards and prescriptions that will promote the integration and sharing of such spatial data.
- Ensure that data custodians capture, publish and update metadata of such spatial data.

- Ensure access to such spatial data in terms of the Promotion of Access to Information Act (Act 2 of 2000).
- Ensure that data custodians are responsible and accountable for the integrity of such spatial data.
- Ensure that data custodians honour usage and collaborative maintenance agreements that they have entered into.
- Improve data quality where data custodians are required to rectify identified spatial data deficiencies within a specified period.
- Ensure that data custodians provide adequate security to mitigate spatial data loss and unauthorised/unlawful access to, modification of, or disclosure of spatial data for which they are the custodians.
- Ensure that data custodians protect the copyright of the state, and other interested parties, of spatial data for which they are the custodians.

The CSI will prescribe spatial data standards to which data custodians must comply. The primary source of such standards is Standards' South Africa (STANSA).

Standards

Standards' South Africa (STANSA) is a statutory organisation, governed by the Standards Act (Act 29 of 1993), and is South Africa's official body for the preparation and publication of standards. STANSA's Technical Committee 71 is responsible for national standards for information technology, and its sub-committee, SC71E, for national standards for geographic information.

SC71E is South Africa's representative to the International Organisation for Standardisation's Technical Committee for Geographic Information/Geomatics (ISO/TC 211) (www.isotc211.org). SC71E holds permanent member status within ISO/TC211.

ISO/TC211 carries out research and generates international standards, technical specifications and guidelines by means of focused working groups (WG). There are five working groups that are currently active within ISO/TC211. They are Geospatial Services (WG 4), Imagery (WG 6), Information Communities (WG 7), Information Management (WG 9) and Ubiquitous Public Access (WG 10).

SC71E either adopts ISO/TC211 standards and technical specifications or generates South African standards that are aligned to these standards and specifications. The following South African Standards have been developed or are in the process of being developed:

- *SANS 1876 Feature instance identification standard:* Provides each data custodian, identified by the CSI, with a unique two-digit alpha identifier. The data custodian will use this alpha identifier as prefix in order to generate a unique fourteen-digit alphanumeric identifier for each feature instance. The feature instance will retain that unique alphanumeric identifier through its lifetime, facilitating temporal analysis (STANSA, 2004a).
- *SANS 1877 A standard land-cover classification scheme for remote sensing applications in South Africa:* Provides a land-cover classification scheme that has been used for land cover mapping in South Africa. The standard is to be revised once ISO/TC211 has published ISO 19144-2: Land Cover Classification System (STANSA, 2003).
- *SANS 1878 South African spatial metadata standard:* Provides a South African metadata standard that consists of selected metadata elements contained in the international metadata standard, ISO 19115 (STANSA, 2004b).
- *SANS 1880 South African Geospatial Data Dictionary (SAGDaD) and its application:* Provides a list of feature types, and definitions, to which a unique code has been allocated. Each feature type has a list of aliases to which a unique value has been assigned. The standard is referenced to ISO 19110: Methodology for feature cataloguing (STANSA, 2009b).
- *SANS 1883 South African Address Standard:* The standard consists of three parts catering for data formats of addresses, guidelines for addresses in databases, data transfer, exchange and interoperability and guidelines for address allocation and updates (STANSA, 2009a).

iTIS Development

The CDNGI commenced with the development of a geographic information system (GIS) in 1986 where the digitisation of the 1:500 000 national map series was carried out as a pilot project (Vorster, 2003). The pilot project was completed in 1992 and maintenance of this database was discontinued in 1996.

The development of the National Topographic Information System (NTIS) commenced in 1988 with the digitisation of the major features of the 1:50 000 national map series. GIS structuring would commence in 1992 and the database would be completed in 1999. The NTIS would prove its effectiveness when it was used by Statistics South Africa and the Independent Electoral Commission for the delineation of enumerator areas and the establishment of voting districts for the successful 1999 national elections.

The NTIS, however, had positional and content deficiencies due to having been digitised from cartographic maps and that only the major features had been captured with little or no attribute and metadata content. Complete topographic feature capture of the NTIS commenced in 1996 and is expected to be completed by March 2011.

Other areas of concern were identified that required resolution. These were:

- The topographic data capture processes produced valuable and relevant data that were housed in data silos and not integrated with the NTIS.
- Other CDNGI products were housed in data silos that were not integrated with the NTIS.
- The NTIS data content consisted of features contained in the 1:50 000 map series only and required enrichment to cater for the broader GIS community.
- The NTIS database is housed in an Oracle 8.05 database, a relational database, and required to be upgraded to an Oracle 10g database, an object-relational database.
- The NTIS did not comply with the South African standards developed through SC71E.

The CDNGI sourced expertise from the State Information Technology Agency (SITA) and initiated an investigation in order to enrich the NTIS content and integrate the various data silos within the CDNGI on an Oracle 10g platform. The project commenced in August 2004 and was completed in August 2006.

The project objectives were (CDSM, 2004a):

- Integrate CDNGI's diverse systems onto one common platform providing the ability to share data within the organisation and with other clients via local and wide areas networks and the internet.
- Obtain, maintain, manage and disseminate quality topographical data, in terms of best practices (lineage, metadata, object-based version management, long transaction management and topology).
- Maintain and manage the iTIS in terms of best practices.
- Conform to existing standards (i.e. ISO 19100 group of standards and related South African Standards and Open GIS Consortium (OGC) standards and specifications).
- Establish appropriate security access to the iTIS to provide for direct data/product access to the stakeholders.
- Comply with the requirements of Intergraph's Interoperability Grant, to which the CDNGI had subscribed, where GML will be used for data dissemination (including web services).

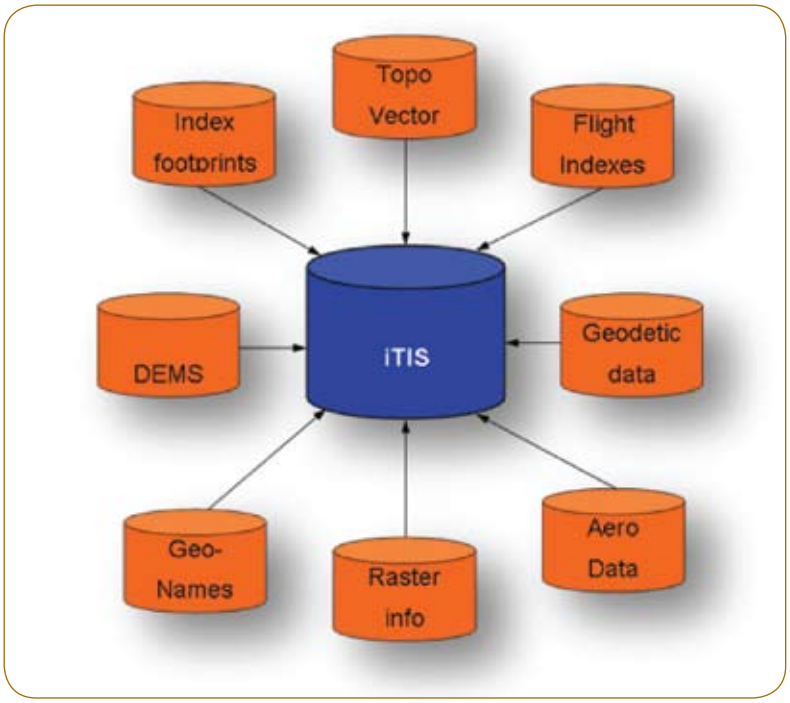

Fig. 3: The iTIS integrates all NGI data.

The project consisted of four phases, the phases and outcomes were:

- *Project Planning and Initiation (PPI):* The refinement of the project investigation and project deliverables (CDSM, 2004a);
- *System Requirements Analysis (SRA)*: Analyse and document current products and processes

and determine the functional requirements to be supported by the iTIS (CDSM, 2004b);

- *System Delivery Specification (SDS)*: Analyse and document user requirements and outline the design to produce a detailed specification of the system to be delivered. The current (AS IS) and future (TO BE) logical models were determined here (CDSM, 2005);
- *Technical System Design (TSD):* How the iTIS was to be built, based on the TO BE logical model, include implementation plans for system development, data capture and migration (CDSM, 2006).

iTIS Implementation

A summary of the four-phase implementation plan, documented within the TSD, is as follows (CDSM, 2006):

- *Phase 1:* Prototype data model: Construct the physical schema and test the model with sample data.
- *Phase 2:* Pilot phase: Acquire identified software, develop customisations and interfaces, build complete database infrastructure and migrate current data silos.
- *Phase 3:* Initial production phase: Test functionality, communication, performance, volumes, recovery, operations, environment and security and carry out acceptance test and accept system.
- *Phase 4:* Final implementation phase: Prepare and implement cut out of current system and switch over to new system. Review implementation, identify and implement interventions.

A detailed implementation plan was prepared and the actual implementation of the iTIS commenced in August 2007 and should be completed by March 2010. The implementation deliverables are:

- *Hardware and software infrastructure installation:* Install migration and iTIS database servers to migrate vector from Oracle 8.05 via Oracle 9i to 10g. Install migration and production imaging servers for raster data migration.
- *Vector data migration:* Convert to object-relational form in iTIS structure where the roads and rivers are networked. Quality checking and data cleaning carried out.
- *Raster data migration:* Quality check and structure imagery and port to image server with metadata housed in Oracle 10g database.
- *iTIS Production workflow specification:* the integration of various workflow processes to ensure version management within the iTIS.
- *DEM data migration*: Structure and populate the new National Digital Elevation Model (NDEM) which has an interval of 25 m (previously 50 m, 200 m and 400 m).
- *Divisional development:* Develop interfaces in order to populate various data, previously housed in various data silos, within the new iTIS.

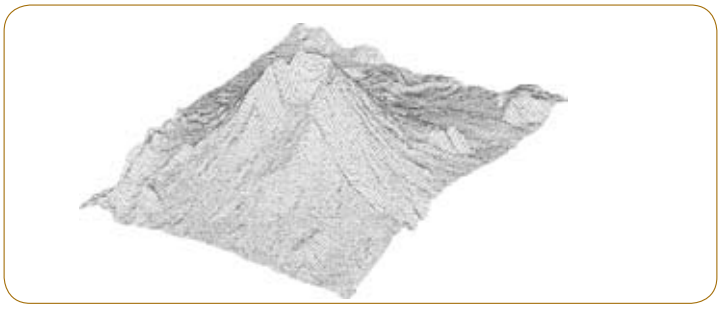

Fig. 4: A graphic example of a DEM.

Subject	NTIS	iTIS
Database	Oracle 8.05	Oracle 10g
Database implementation	*Relational:* Tables have identical data structure/ design, no relationships defined.	*Object/relational:* Tables have different data structure/ design and many tables have relationships.
Standards	None	Compliant SABS/ISO
Data content, integrated	No	All silo databases integrated into a single oracle database schema, access through oracle authentication.
Data management	Long Transaction and Version Management (LTVM) for vector data.	Geomedia Transaction Management (GTM) for vector. Terrashare for raster data and a standard logical file structure and file naming convention to facilitate bulk ingesting of image metadata.
Topographic data segmentation	All features segmented along logical map sheet boundaries, e.g. 1:50 000 map sheet footprints. A tiled database.	Seamless (un-tiled database), segmentation of road and river data at intersecting geometries forming a networked topological structure.

Table 1: Comparison between NTIS and iTIS architecture.

Subject	NTIS	iTIS
Feature groupings	0	11
Feature classes	23	50
Feature types	222	429

Table 2: iTIS content enrichment.

Feature groupings	Feature classes	
Hydrography	HYDR_AREAS	HYDR_ COASTAL_AREAS
	HYDR_COASTAL_LINES	HYDR_LINES
	HYDR_POINTS	

Table 3: Example: Hydrography feature classes in the iTIS.

iTIS class ID	iTIS feature type ID	Feature type name	SAGDaD feature code	SAGDaD attribute code	Custodian ID
18	195	Bog	15	1	NG015001nnnncc
18	196	Dam	14	1	NG014001nnnncc
18	197	Dry pan	15	5	NG015005nnnncc
NG: Custodian ID allocated by CSI. nnnn: Unique number allocated to feature instance. cc: Two check digits to verify custodian ID integrity.					

Table 4: Example: Three of the Hydro Area feature class types where SAGDaD codes are incorporated in Custodian ID.

- *Technical system design testing:* Operational testing of the iTIS, software packages and customised applications prior to rollout and complete system documentation.
- *Satellite user software installation:* Software rollout of off-the-shelf applications and developed interfaces to users, providing training and support.
- *Transition to the new Oracle 10g database:* Decommission NTIS and commission new iTIS in parallel, where production workflow changes from old to new schema over a period of approximately twelve months.
- *iTIS Portal – development and integration:* Develop interfaces and functionality to permit internet/ intranet access to iTIS.

The changes in architecture between the NTIS and the iTIS are outlined in Table 1.

The comparative data content enrichment of the iTIS, compared to the NTIS, is shown in Table 2.

The eleven logical feature groupings are Administrative, Aeronautical, Cultural, Geodetic, Hydrography, Hypsography, Land Cover Land Use, Physiographic, Reference Index, Transportation and DEM. Each of these feature groupings has a selection of feature classes.

The 50 features classes are each assigned to a particular feature grouping. For example, the five feature classes assigned to the Hydrography feature grouping within the iTIS are presented in Table 3.

The 429 feature types are each assigned to a particular feature class and are referenced to the SANS 1880 (SAGDaD) (STANSA, 2009b) feature codes, attribute values and unique identifiers contained in SANS 1876 (STANSA, 2004a). For example, three of the feature types within the feature class HYDR_AREAS can be seen in Table 4.

In summary

The iTIS implementation will provide for a multi-server environment in which off-the-shelf and customised software applications provide seamless access to all CDNGI's spatial and related data, within an Oracle 10g database environment. Spatial data is indexed and cross-referenced so that users can find data easily using map sheet references, flight plan references, photo references or by identifying a geographical point or area.

Imagery, in particular, was previously difficult to access as various manual searches had to be carried out in order to identify the particular photography job, strip and photograph numbers. Within the iTIS the flight index is geo-referenced, providing a footprint of the job coverage and where each individual photograph is located. The user then can use the imaging server to access the image, use a topographic map as backdrop, view the trigonometrical beacons in the area, and/or any other spatial and related data that is available within the iTIS.

The web portal will be so designed that intranet/internet users have access to the iTIS using a Web Feature data server (WFS), Geography Markup Language (GML) data server, Web Map Service (WMS) and access to imagery. The WFS and GML data servers will permit database connection to the iTIS. GML is an open format developed by the OGC. The WMS will permit thin-client access to the iTIS in order to create customised maps which can then be downloaded by the user. The limitation, however, will be available bandwidth. Initial tests have not been encouraging and it may be some time yet before this functionality can be fully utilised by external users.

Data is only valuable and relevant if it is extensively used and is found to be useful. Access to such data by internal and external users is vital. The iTIS has been designed to provide such access.

References

[1] Chief Directorate: Surveys and Mapping (2004a) PPI – *Project Plan and Initiation (Project Charter) for an Enterprise Business Plan for an Integrated Topographical Information System data and databases for DLA*, Chief Directorate: Surveys and Mapping. Internal Report.

[2] Chief Directorate: Surveys and Mapping (2004b) *System Prospectus: User Requirements Specifications (URS), System Requirements Analysis (SRA 1-4) for an Integrated Topographical Information System for the Chief Directorate: Surveys and Mapping.* Internal Report.

[3] Chief Directorate: *Surveys and Mapping (2005) System Delivery Specification for an integrated Topographical Information System for the Chief Directorate: Surveys and Mapping.* Internal Report.

[4] Chief Directorate: *Surveys and Mapping (2006) Technical System Design for an integrated Topographical Information System for the Chief Directorate: Surveys and Mapping.* Internal Report

[5] Standards South Africa (2003) SANS 1877: 2003: *A standard land cover classification scheme for remote sensing applications in South Africa*

[6] Standards South Africa (2004a) SANS 1876: 2004: *Feature instance identification standard (draft).*

[7] Standards South Africa (2004b) SANS 1878: 2004: *South African spatial metadata standard Part 1 – Core metadata profile (draft).*

[8] Standards South Africa (2009a) *SANS 1883-1,2,3: 2009: Address Standard Part 1, Part 2 (draft) and Part 3 (draft)*

[9] Standards South Africa (2009b) SANS 1880: 2009: *South African Geospatial Data Dictionary (SAGDaD) and Its Application (draft).*

[10] PJ Vorster: (2003) *The Development Of The South African National Topographic Information System. Proceedings of the 21st International Cartographic Conference, Durban, South Africa*, August 2003.

[11] PJ Vorster: (2005) The Development of an Integrated Topographic Information System for South Africa's National Mapping Agency. *Proceedings of the 7th Africa GIS Conference, Tshwane (Pretoria), South Africa, November 2005.*

Contact Patrick Vorster,
Chief Directorate: National Geo-spatial Information, Tel 021 658-4312,
pvorster@ruraldevelopment.gov.za