

An inside look at processes and data content at the NGI since the inception of the datamodel project.

Raoul Duesimi

Chief Directorate: National Geo-spatial Information. Department of Rural Development and Land Reform, Cape Town, South Africa, rduesimi@gmail.com

Abstract

It has been six years since the NGI (National Geo-spatial Information) undertook a project to restructure and integrate all its data into a single data repository. Data was stored in silos throughout the organisation in different formats and data structures. Public access to the data is through the sales team, the silo mentality inhibited the sales team from accessing data within its own organisation. With the inception of the “Promotion of Access to Information Act” and the “Spatial Data Infrastructure Act (SDI)”, an increase in the demand for data was expected and compliance to standards prescribed by the SDI Act required for the integration and restructuring of the NGI data. Processes, technology and data content at the NGI will be presented in this paper.

1. Introduction

National geo-spatial Information (NGI) provides many geo-spatial products to the geo-spatial industry, the most important product being the Aerial photography which is the base product from which all mapping products are produced. Products derived directly or indirectly from the photography include, Aerial triangulation data, Digital Elevation Model (DEM) data, Orthorectified imagery, Topographical vector data and a variety of mapping products at different scales. In order for the NGI to meet the demands of the dynamic changing environment due to the occurrences of aggressive social and economic activities, products have to be produced through seamless processes throughout the enterprise with optimal access to the data repository, hence the development of the Integrated Topographical Information System (ITIS) at the NGI. The increase in the temporal acquisition of digital aerial photography has lead to an increase in storage demands which is further compounded by an increase in the number of product revision cycles requiring more storage. Processes, technology and data content and future opportunities will be addressed in this paper. A typical daily life cycle at the NGI is illustrated in Figure 1.

The heart of all the processes and product activity is the data centre which is the repository for all metadata, imagery and products from which the itisportal and geospatial portal will consume and derive information for presentation. The datacentre consists mainly of an oracle spatial cartridge database, tape robot archive system and storage area network consisting of 240 TB of disc space. The SAN is made up of a collection of Enterprise Virtual Array (EVA) storage systems provided by Hewlett Packard (HP) and is managed by the StorNext Storage Manager application. The archive

system is supported by a tape system that stores two copies of each data element stored in the SAN. The databases are also allocated SAN storage and are backed up onto a tape robot system using the Legato application software. The establishment and management of the datacentre is a very expensive exercise, EVA enclosures and servers have to be purchased and maintained, discs and tapes need purchasing to meet the increasing data acquisition and product production cycles, Software licensing and maintenance is required, uninterrupted power supply (UPS) and alternate power source generators are required, air conditioners are required to keep all the electronic equipment in the datacentre cool and finally a team of IT specialist need to be contracted in order to manage the system. Additional costs will later come from the replacement of servers and EVA enclosures due to the advancement in technology and the purchasing of additional storage management licenses, which is charged proportionally to the amount of storage that needs to be managed on the SAN and tape robot system which is ever increasing on revision cycles of data.

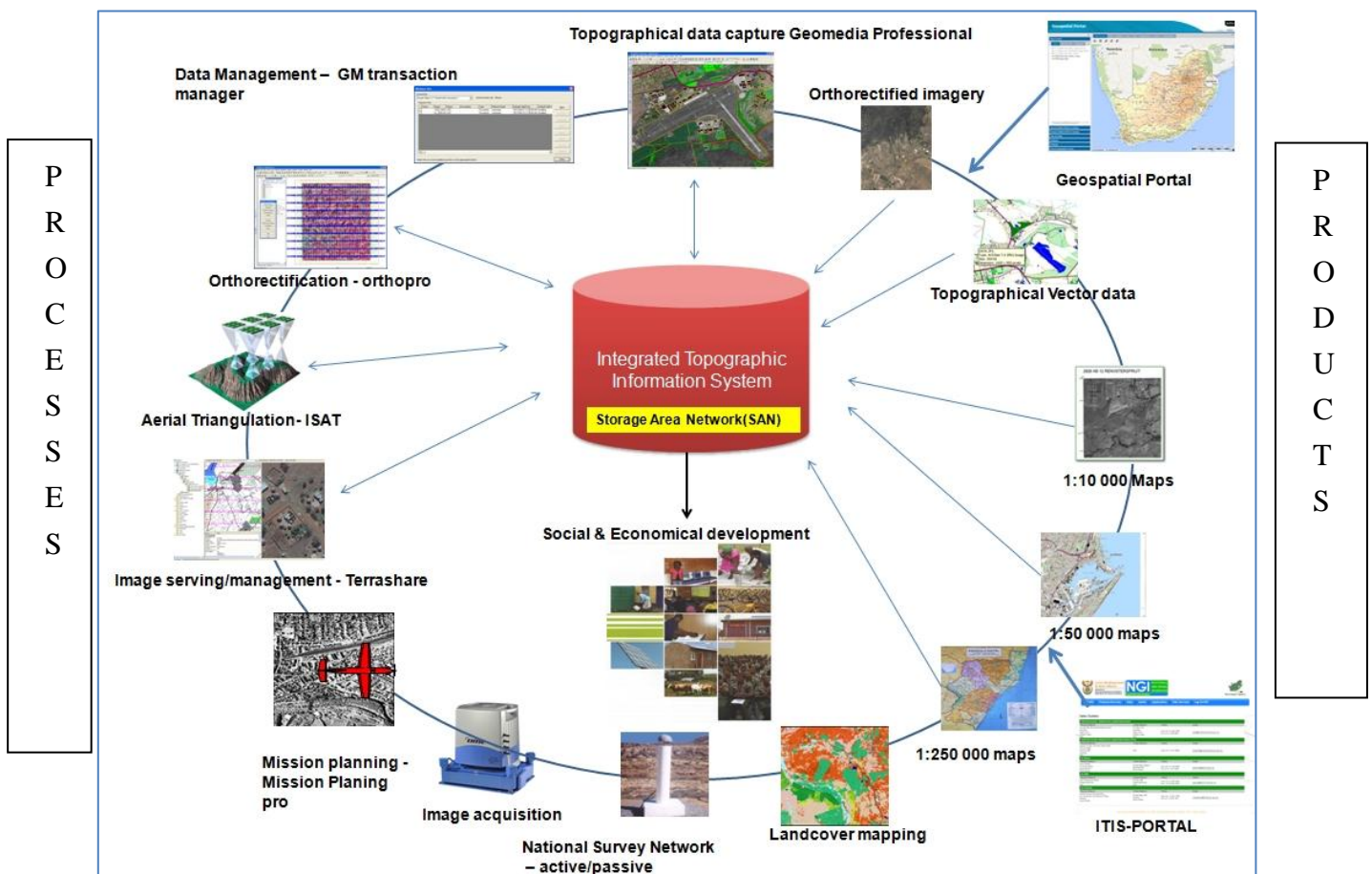


Figure 1. Typical daily life cycle at the NGI

2. Processes and technology used at the NGI:

2.1. National Survey Network

The national network consists of active and passive beacons. Strategic passive beacons are identified for maintenance annually. The maintenance is performed by the NGI beacon building team. The active network is monitored daily and data continuously streamed to the control centre at the NGI. GNNS data is packaged into products and is available online.

2.2. Image acquisition:

The NGI has acquired its own Digital Mapping Camera (DMC) for digital aerial photography acquisition. National imagery acquisition is undertaken within a three year cycle. A partnership is established with a service provider, through bid process, to acquire two thirds of the country utilising the NGI DMC and the other one third is acquired from the industry via another bid. All photography is acquired strictly according to the Digital Aerial Imagery Standard and flown according to mission/flight planning prepared by the NGI.

2.3. Mission Planning

Flight Planning is prepared by the NGI and is accepted after consultation with the service provider before the image acquisition cycle is started. The NGI will scrutinise the supplied imagery for conformance with the flight plan details. The software utilised for the flight planning is Leica Mission Pro from the Hexagon group.

2.4. Image Management and Serving:

A formal naming convention has been set up for all image products that contain the job number, strip number and photo number within the image identifier. The naming convention forms part of the validation process during database population. Ancillary data files also have formal naming conventions and the data within is specifically formatted/structured so that it can be consumed seamlessly into process activities down the line without out any formatting directly from the SAN and without any local disc copying thereof. On acquisition of a job the images are immediately archived into the system and information related thereto is ingested automatically into the ITIS database. Validation is carried out on ingestion and any image data not conforming to the flight plan index is logged for inspection and correction. The ingestion is carried out via the Terrashare image server client application supplied by the Intergraph/Hexagon group. On successful ingestion the image client creates a logical structure corresponding to the physical storage structure from which images and stereo photogrammetric 3D image models can be viewed by navigating the implied flight plan created by the ancillary support flight data. Stereo capability is very useful for the topographical vector data capture component in carry out photo interpretations during their capture process. The imagery is served throughout the enterprise immediately after ingestion via the Terrashare Client application. The ingestion process also updates the central database with metadata so that the availability of the images can be derived by the portal applications for presentation and quoting thereof by the sales team.

2.5. Aerial triangulation

The Aerial triangulation process is seamlessly performed directly from the SAN without copying any images to local drive or reformatting thereof. The ancillary flight data support files enables the consumption of the data directly into the aerial triangulation software for immediate processing. Aerial triangulation is performed using Image Station Aerial Triangulation (ISAT) and PATB

software applications. The output is stored on the SAN in specific formats so that it can be seamlessly consumed into The DEM and orthorectifications application.

2.6. Ortho Rectification

All the components for orthorectification, e.g. DEMS, calibrated images and final adjusted image orientation parameter files, are stored within the SAN as outputs from previous processes. The rectification process can seamlessly consume these components directly into the orthorectification application software. Areas consisting of hundreds of images can be consumed for processing during a single session. Rectification is performed on a single job area at a time and the final output is 100 mosaic images that corresponded to the 1:10 000 map reference index. The orthorectified images are stored on the SAN and ingested through the Terrashare image client for serving to the enterprise and population and validating of metadata into the database for publishing in the ITIS portal application. The most important component for orthorectification is DEM, which is performed after the aerial triangulation process using the Usmart and Image Station Automatic Elevation applications supplied by SmartTech and Intergraph/Hexagon respectively. The orthorectification process is performed using the Image station Orthopro Application supplied by Intergraph/Hexagon.

2.7. Data management

The topographical vector data is stored in a 10g Oracle Spatial Cartridge database and is a fully versioned implementation. The version management implementation is 'just in time', i.e. rows in the database are locked at the time of data manipulation and remain locked to particular users for exclusive use until it is returned/committed to the database, unlike conflict management which allows for multiple changes on a single row and then later resolving the conflicts for final committal. All modified, deleted and inserted data is recorded on the database as belonging to a revision set and these changes to the database will only become visible in the database as soon as the supervisor have accepted the revision set for committal, i.e. the authorised owner of the revision set will be able to observe all the changes in the revision set while the online live database user will only have access to the current state of the database and not the revision set. Versions of all edited data are stored within the database as a temporal entity, hence the database state is continuously changing and it consists of both current and temporal data. The more edits you have on the database will result in a larger storage demands so we need to make sure that the changes that are occurring on the database is for temporal updates and not correcting bad digitising else we will have both versions of edited and temporal data stored. The NGI has implemented topological data structuring rules that are checked against the data by a supervisor before final committal to the ITIS database in order to minimise the edits due to bad digitising. The management of versioned data is performed via Oracle Workspace Manger from Oracle and GeoMedia Transaction Manager from Intergraph/Hexagon.

2.8. Topographical vector data capture

Topographical vector data is primarily performed using NGI orthorectified imagery and ancillary data sourced by the NGI as an overlay for aiding the capture process. Data is interpreted and mono-digitised from the imagery into a revision set before final committal to the database. Geomedia Professional and GeoMedia Transaction Manager is utilised for performing this procedure. Since the inception of the data model project, all data is captured according to the new a classification system. The new data is now available from the NGI sales team.

2.9. Cartographic map production

The cartographic products are by-products of the topographical vector data and ancillary mapping data. Medium and small scale, 1:50 000, 1:250 000, 1:500 000 , 1:1000 000 and Provincial maps are produced using the Gothic Developer Digital Cartographic mapping system from 1Spatial solution and the large scale, 1:10 000 orthophoto, maps are produced using the ESRI ArcGIS product. Images and vector data are consumed seamlessly into this environment from the SAN and Oracle spatial database. Feature Manipulation Engine (FME) is used as a data translator for ingesting the topographical vector data from the oracle spatial database into the Gothic spatial Database for map productions. All mapping products are archived in the SAN according to formalised file structures and naming conventions. The map products are also ingested into the terrashare image server environment for enterprise viewing, metadata ingestions and validation for presentation through the portal application.

3. Portal application processes

Portal processes are performed by the sales team for searching for client product requests, creating quotes thereof and stock management. All imagery and map products that have not been ingested through the terrashare client application will not be available to the web portal application. The portal application can derive information about products via geographical name searches, coordinate searches (DMS and DD) and map reference index searches. The derived data can be viewed and added to a shopping cart facility for final quoting. The data is derived from the ITIS database. Figure 2 - Figure 5 below illustrate the various types of searches that can be performed via the ITIS portal web application. Figure 2 is the initial search for maps containing the geographical name “Newlands” and Figure 3 is the extended list of the desired map selected, the extended list will list the latest copy that is available on the shelf and all back copies of the map including the date, edition, map ellipsoid, geo-referenced ellipsoid and the processing method. The desired map may be previewed or selected for purchasing and added to the shopping cart for invoicing. Figure 4 and Figure 5 below illustrates a coordinate search, DMS, on the aerial photography library. The list of related jobs are served up and presented in Figure 4 and Figure 5 displays the specific photos that relate to the coordinate entered and job derived for selection and adding to the shopping cart.

Maps & Charts

Printed Map

50 Index:	3318CD
Name/Area:	KAAPSTAD/CAPE TOWN
LO Central Meridian:	19 EAST
Extended Sheet?:	No
Comments:	WGS84 Photo J1033
Distribution Caveats:	
Quality Rating:	Not Rated
Hardcopy Edition:	9
Hardcopy Edition Year:	2000
Withdrawn?:	
Withdrawn Date:	
When Next Published:	
Production Status:	
Printed Media:	<input type="text" value="Paper"/> <input type="button" value="Buy"/>

Digital Raster Images

50 Index	Previous Index	Media	Name/Area	Ed Nr	Ed Year	Processing	Source	Map Ellipsoid	GeoRef Ellipsoid	Hardcopy Spec	Preview	Buy
3318CD		TIF	KAAPSTAD/CAPE TOWN	3	1961	Scanned	Cartography Clarke 1880	WGS 84			Preview	Buy
3318CD		TIF	KAAPSTAD/CAPE TOWN	5	1980	Scanned	Cartography Clarke 1880	WGS 84			Preview	Buy
3318CD		TIF	KAAPSTAD/CAPE TOWN	6	1983	Scanned	Cartography Clarke 1880	WGS 84			Preview	Buy
3318CD		TIF	KAAPSTAD/CAPE TOWN	6	1988	Scanned	Cartography Clarke 1880	WGS 84			Preview	Buy
3318CD		TIF	KAAPSTAD/CAPE TOWN	7	1993	Scanned	Cartography Clarke 1880	WGS 84			Preview	Buy
3318CD		TIF	KAAPSTAD/CAPE TOWN	8	1998	Scanned	Cartography Clarke 1880	WGS 84			Preview	Buy
3318CD		TIF	KAAPSTAD/CAPE TOWN	9	2000	Digital	Cartography WGS 84	WGS 84			Preview	Buy

Figure 2. A geographical name search on 50K maps

Maps & Charts

1:50 000 Topographical Maps

Media

Geographical Names Feature Types:

Province:

Coordinates: Latitude(S): Longitude(E):

Search

Enter Index or Name (Full or partial) and [Search]

	50 Index	Name/Area
<input type="text"/>	<input type="text"/>	<input type="text"/>

Results

Select Index - or Name-heading to sort. Select Index for more information

Buy	LO Central Meridian	50 Index	Edition	Year	Name/Area	Extended Sheet?
<input type="checkbox"/>	23 EAST	3222BC	4	2005	BEAUFORT WEST	No
<input type="checkbox"/>	19 EAST	3318CD	9	2000	KAAPSTAD/CAPE TOWN	No
<input type="checkbox"/>	23 EAST	3323CD	3	1998	THE CRAGS	No

Figure 3. An extended list derived from results in Figure 2.

Digital Aerial Photography

Aerial Photography Job Listing

Media: Colour 900mmX1000mm

Geographical Names: Feature Types: All

Province: All

Coordinates: Degrees, minutes, seconds Latitude(S): 33 ° 18 ' 18 " Longitude(E): 18 ° 18 ' 18 "

Search

Enter Series or Name and (Search)

Type	Spectral	Job Nr	Name/Area	Year	Scale	Index ie. 2228CD
All	All			All		

Results

Buy	Acquisition Source	Spectral	Job Nr	Name/Area	Year	Scale	List Photos
<input type="checkbox"/>	Analogue	BlackWhite	1033	Cape Town	2000	1:50 000	List Photos
<input type="checkbox"/>	Analogue	BlackWhite	126	Cape Town / Piketberg1938	1:25 000		List Photos
<input type="checkbox"/>	Analogue	BlackWhite	454	Hopefield	1900	1:35 000	List Photos
<input type="checkbox"/>	Analogue	BlackWhite	498_355	Malmesbury	1999	1:30 000	List Photos
<input type="checkbox"/>	Analogue	BlackWhite	498_49	Ysterfontein	1973	1:30 000	List Photos
<input type="checkbox"/>	Analogue	BlackWhite	786	Western Cape	1977	1:30 000	List Photos
<input type="checkbox"/>	Analogue	BlackWhite	919	Cape Town	1989	1:50 000	List Photos
<input type="checkbox"/>	Analogue	Colour	498_685	Langebaan	2007	1:30 000	List Photos
<input type="checkbox"/>	Digital	Colour	3318A_2010_305	Langebaan	2010	Use GSD 0.50	List Photos
<input type="checkbox"/>	Digital	Colour	3318A_2013_697	Langebaan	2013	Use GSD 0.50	List Photos

First Previous 1 Next Last



Figure 4. A coordinate search on the Aerial photography library

Digital Aerial Photography

Aerial Photography Job Listing

Media: Colour 900mmX1000mm

Geographical Names: Feature Types: All

Province: All

Coordinates: Degrees, minutes, seconds Latitude(S): 33 ° 18 ' 18 " Longitude(E): 18 ° 18 ' 18 "

Search

Enter Series or Name and (Search)

Type	Spectral	Job Nr	Name/Area	Year	Scale	Index ie. 2228CD
All	All			All		

Results

Buy	Acquisition Source	Spectral	Job Nr	Name/Area	Year	Scale	List Photos
<input type="checkbox"/>	Analogue	BlackWhite	1033	Cape Town	2000	1:50 000	List Photos
<input type="checkbox"/>	Analogue	BlackWhite	126	Cape Town / Piketberg1938	1:25 000		List Photos
<input type="checkbox"/>	Analogue	BlackWhite	454	Hopefield	1900	1:35 000	List Photos
<input type="checkbox"/>	Analogue	BlackWhite	498_355	Malmesbury	1999	1:30 000	List Photos
<input type="checkbox"/>	Analogue	BlackWhite	498_49	Ysterfontein	1973	1:30 000	List Photos
<input type="checkbox"/>	Analogue	BlackWhite	786	Western Cape	1977	1:30 000	List Photos
<input type="checkbox"/>	Analogue	BlackWhite	919	Cape Town	1989	1:50 000	List Photos
<input type="checkbox"/>	Analogue	Colour	498_685	Langebaan	2007	1:30 000	List Photos
<input type="checkbox"/>	Digital	Colour	3318A_2010_305	Langebaan	2010	Use GSD 0.50	List Photos
<input type="checkbox"/>	Digital	Colour	3318A_2013_697	Langebaan	2013	Use GSD 0.50	List Photos

First Previous 1 Next Last

Figure 5. A coordinate search on the Aerial photography library, listing specifies job photos.

4. Geospatial Portal (Current prototype created)

The Geospatial Portal is in its infancy. A prototype implementation has been concluded and is being tested for comment and improvements before final implementation, illustrated in figure 6–figure 12.

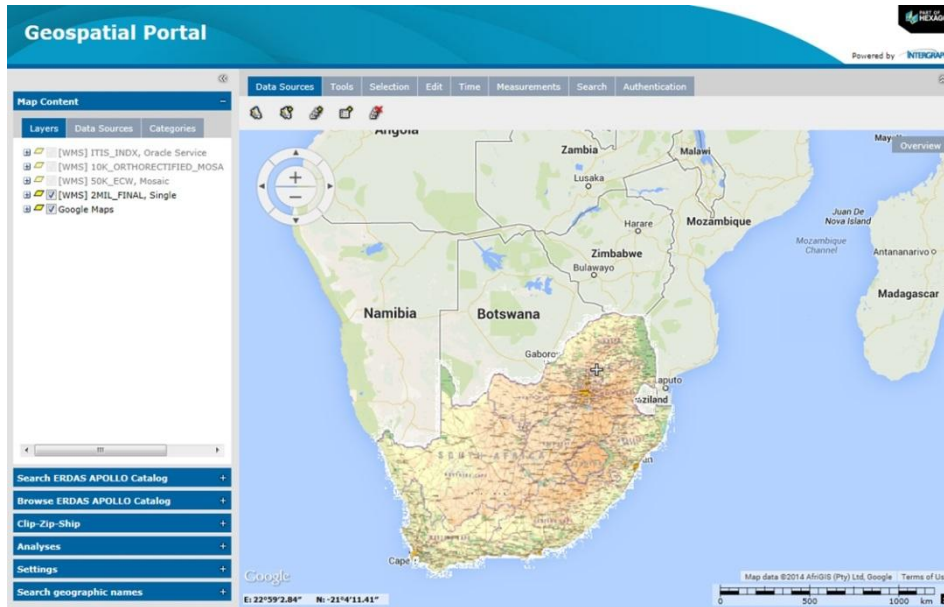


Figure 6. Default landing page with base datasets present via OGC services and Google map backdrop

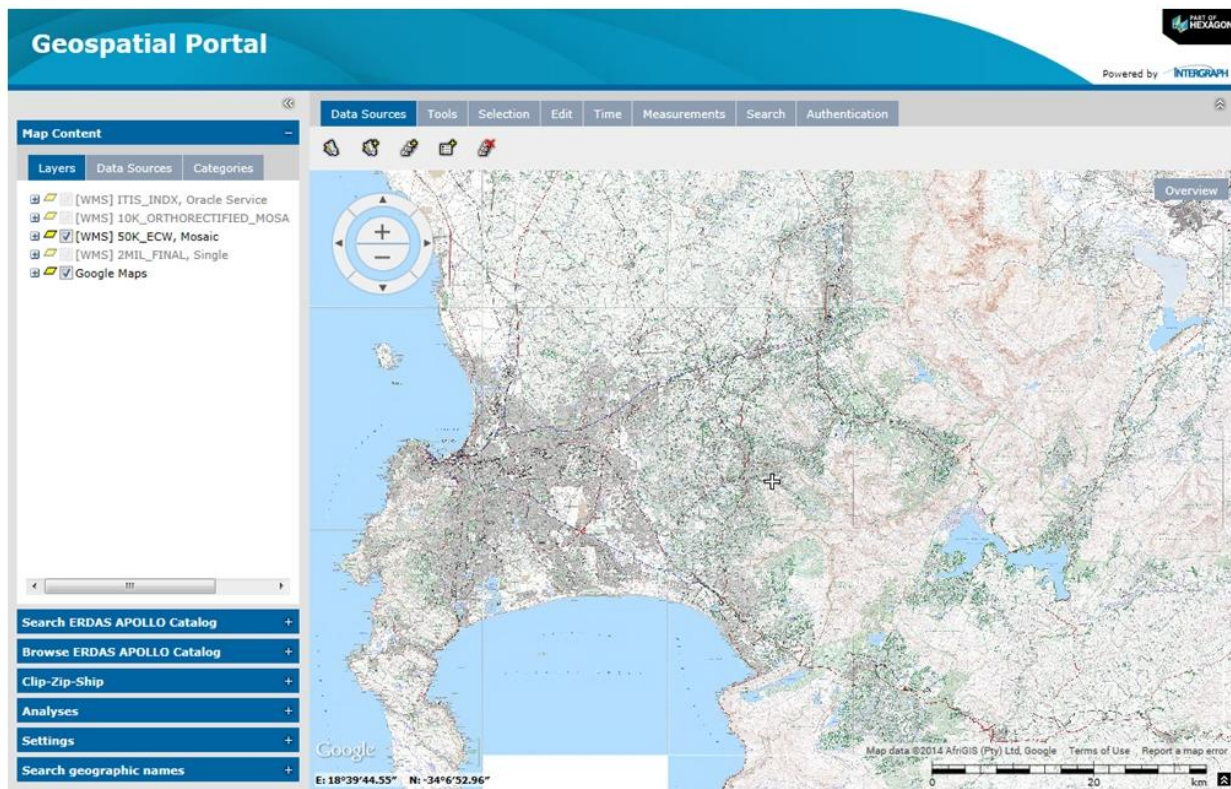


Figure 7. WMS implementation of the clipped 1:50 000 maps

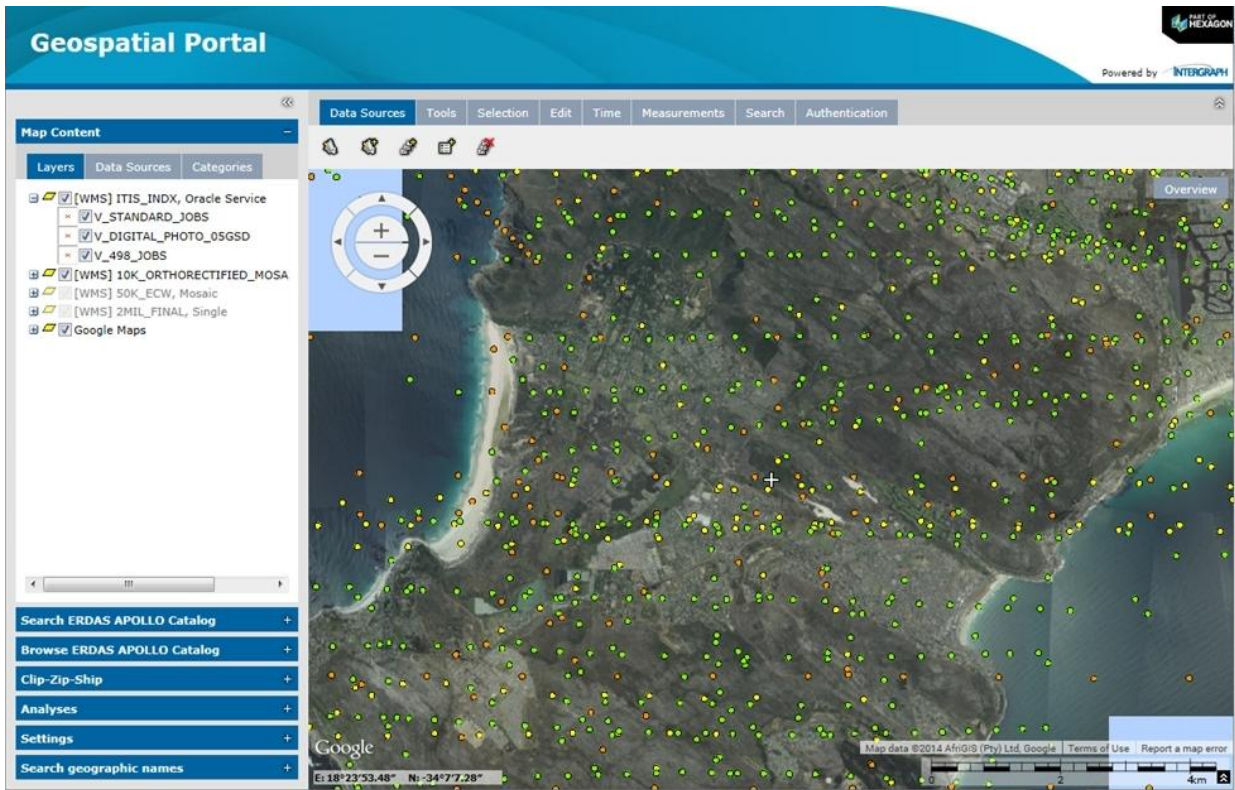


Figure 8. WMS of orthorectified imagery and image photo centres of the photography library

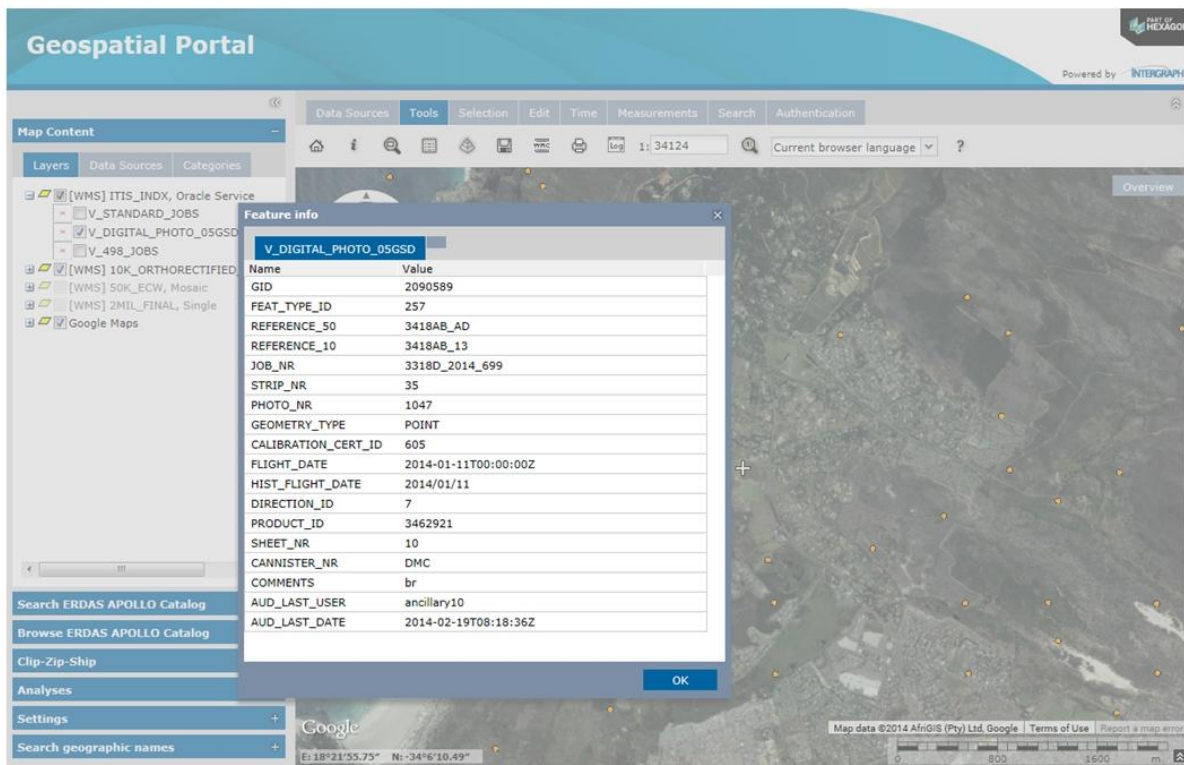


Figure 9. WMS query result invoked by selecting an image photo centre.

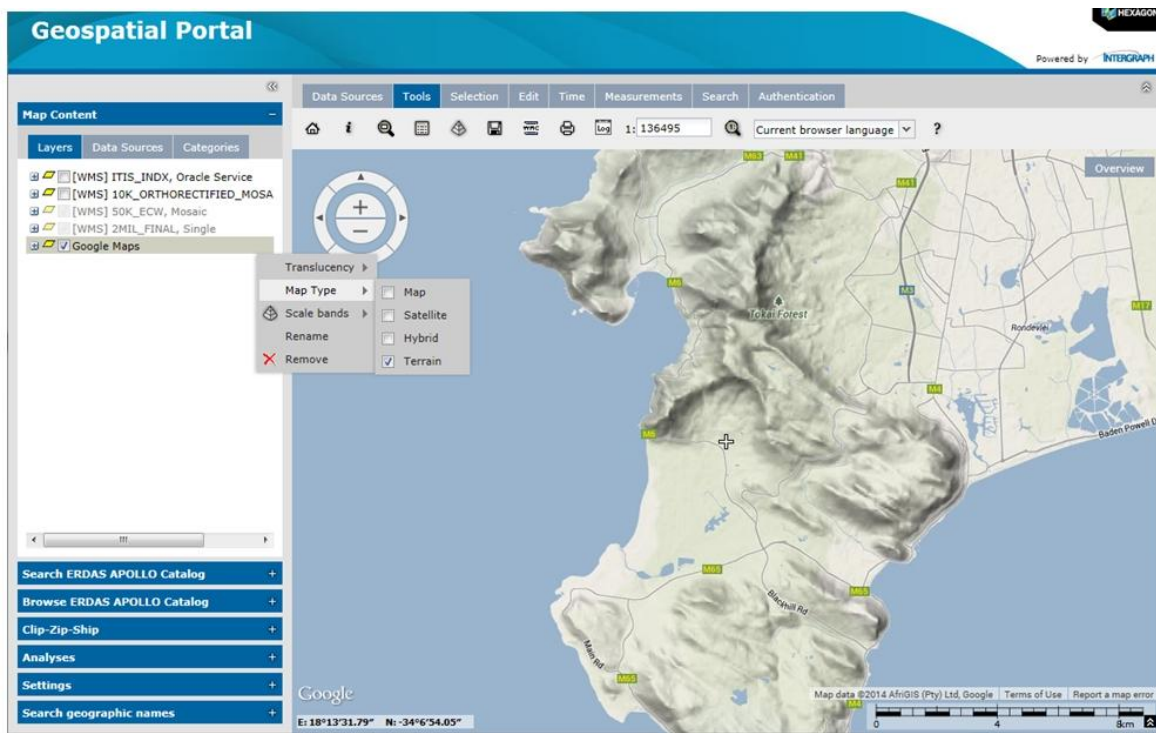


Figure 10. Google terrain map implementation within the portal

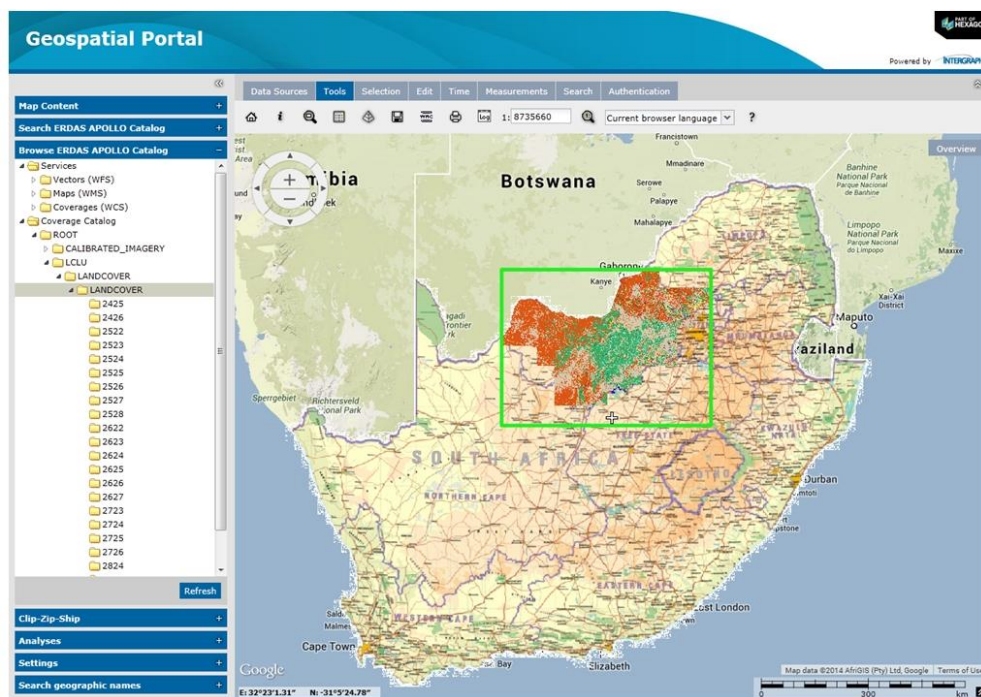


Figure.11. Landcover data presented directly from the SAN into the portal environment using a Catalog service.

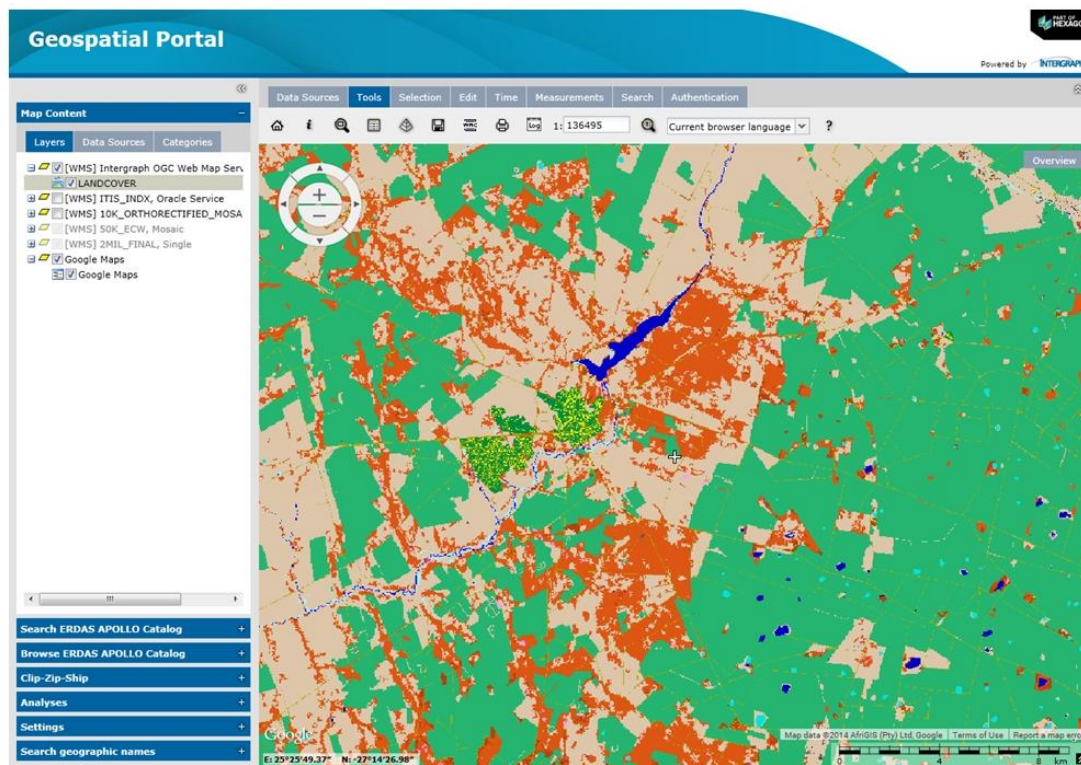


Figure12. Zoom in on the Landcover data crawled from the catalog service.

5. Opportunities and strategies to consider

The NGI has a large variety of Geo-spatial information content which includes the following.

- Topographical vector data:
 - Cultural
 - Hydrography
 - Hypsography
 - LULC
 - Physiographic
 - Transportation
 - Geodetic
- Dem Data
- Aerial Photography, current and historical.
- Ortho Rectified imagery
- Geographical names
- Flight plan indexes
- Landcover thematic raster
- Various Small scale map series
- Medium scale map series
- Orthophoto maps – large scale

It is important that the geo-spatial community has uninterrupted access to the data. The data has been accumulated at huge cost to the tax payer. Access to the data can be perceived as being restrictive. Data is via pre-packed product DVD's or hard drives that are sent to the sales team for downloading onto external hard drive. A trend is to request all the data, national coverage, even though only portion of the country is required. This could be as a result of the client not wanting to incur repeat costs related for courier services or business entrepreneurs wanting to setup their own portal facilities which may incur further costs to the tax payer in acquiring access to that data. A typical order scenario may be a request for all the orthorectified imagery of the country from 5 clients at the same time, this result in the downloading/copying of $5 * 4TB = 200TB$ of data from the SAN over the LAN to the external hard drive at a single instant. On average 36TB of data is transferred to external hard drives monthly by the NGI. Increased network traffic has an adverse effect on the daily production activities within the enterprise. NGI data is copied and scattered throughout the country, silos of the same data. The NGI foots the bill for the entire storage requirement as mentions in the introduction, servers, tape robot system, storage tapes, storage discs, EVA enclosures, software maintenance, hardware maintenance, storage licenses for the SAN and tape system which is charged per megabyte and is ever expanding, UPS, air conditioners, backup power generators and the IT specialists that are responsible for management and sound daily operation of the systems. These are the challenges the NGI face on a daily basis.

Data is very well organised and structured at the NGI, it can be easily consumed and presented through a Geospatial Portal. The internal access and processing of the data within the LAN is good. Web software applications are available, pilot implementation completed, but internet band width is restricting the publishing on WWW. OGC services have been available for a while and cannot be deployed due to the restrictive nature of the WAN (wide area network) and the band width that is available to the NGI. As a result data is duplicated throughout the country and on private geospatial portals. The cost of storing and managing the data is increasing annually and a new strategy should be considered.

The NGI could consider going it alone or look at other solutions such as cloud technology. Cloud technology is an opportunity and not a threat and should be considered for deployment. The creeping nature in the cost of maintaining and equipping an internal datacentre must be weighed up against a cloud solution. A cloud can be experienced as a virtual datacentre and all the hardware and operational elements become implicit to the user and matters related to data can become the main focus of the enterprise, i.e. create the data and allow the IT experts to optimise the physical environment. Technological trends and infrastructure/data security challenges are changing rapidly and a service level agreement with a supplier of a cloud facility may insure that the datacentre is always secure and equipped with the latest IT technology.

Government could consider procurement of a cloud service for all government departments participating in the geo-spatial community, together is better. There will inevitably be policy challenges.

6. References

Software application: Intergraph, viewed 10 APRIL 2014, <<http://www.intergraph.com/>>

Software application: ESRI, viewed 10 APRIL 2014, <<http://www.esri.com/>>

GeoSpatial Portal: INSPIRE, viewed 10 APRIL 2014, <<http://inspire.ec.europa.eu/>>

Geospatial and location standards: OGC (Open Geospatial Consortium), viewed 10 APRIL 2014, <<http://www.opengeospatial.org/>>