International Seminar on Land Administration Trends and Issues in Asia and the Pacific Region

WG3 Data Integration Project

Integration of Built and Natural Environmental Datasets within National SDI Initiatives

20th August 2008 – Kuala Lumpur, Malaysia
Data Integration Project

Introduction

Abbas Rajabifard
Project Coordinator
Vice Chair – PCGIAP-Working Group 3
Integration of Built (cadastral) and Natural (topographic) environmental datasets within National SDIs

One of the major initiatives of WG3 over the last three years is work on the project of the “Integration of Built (cadastral) and Natural (topographic) environmental datasets within National SDI initiatives.”
The aim of the project was to better understand and describe the technical, jurisdictional, institutional, legal and land policy perspective surrounding the two foundation datasets (cadastral and topographic) in a National SDI.
Project Aim

The project has investigated the justification for integrating built and natural environmental datasets in support of sustainable development and developing an integration model, framework and associated tools capable of being used in diverse jurisdictions.
Project Objectives

• Investigate data integration of natural (topographic) and built (cadastral) environmental datasets in the context of establishing and maintaining national SDIs within countries in Asia and the Pacific region.

• Identify problems, issues, similarities and differences through case studies in:
  – institutional arrangements;
  – related SDI policies;
  – data integration methods and technologies; and
  – human resource and capacity building
Project Objectives

• Identify the technical and non-technical tools and requirements of an effective data integration (Integration toolbox):
  – Standards
  – Technical tools
    • Integration data model
    • Spatial data integration validation tool
    • Metadata content
  – Institutional, policy, legal and social guidelines
• Design and implement the tools
Project Support

- Support of PCGIAP
- Research Funding and support of Australian ARC-Linkage scheme
- Support from /Case Study countries
  - Australia (three Australian government organisations)
  - Thailand
  - Japan
  - Malaysia
  - New Zealand
  - Brunei Darussalam
  - Indonesia
Project Case Studies

- Adopted case study approach to undertake research which enabled best practice solutions and improved capacity to evaluate the success of countries systems.
- Case studies were facilitated through PCGIAP-WG3 and 7 countries:
  - Australia
  - Thailand
  - Japan
  - Malaysia
  - New Zealand
  - Brunei Darussalam
  - Indonesia
Case Studies

Spatial data integration process

- Data Exploration
- Data Assessment
- Communication
- Access/Acquisition
- Geo-Processing
- Data Collation

Technical
- Framework
- Data Quality
- Capacity Building
- Metadata
- Effective Communication
- Spatial Data Dictionary
- (Single) Point of Access

Non-technical
- Collaboration
- Privacy Policy
- Data Models
- Restrictions on Use
- Specifications and Guidelines
- Pricing
- Standards
- Custodianship
- Interoperability
Case studies

Stage 1
Current situation of multi-source spatial data integration
Existing documents and resources
- Bhutan
- Canada
- India
- Botswana
- China
- Greece
- WA, Australia

Stage 2
The investigation of technical and non-technical issues
Questionnaires and country reports
- Australia
- Japan
- Singapore
- Malaysia
- Philippines
- Indonesia
- Brunei Darussalam

Stage 3
Detailed investigation of technical and non-technical issues
Agency visits, meetings and data audit
- Victoria, AUSTRALIA
- NSW, AUSTRALIA
- GeoScience, AUSTRALIA

Potential Integration Issues and Obstacles

Tools and Requirements of effective spatial data integration
Overall Plan and Project Methodology

![Flowchart diagram showing the project timeline and steps]

1. **Project Start**
   - June 2005

2. **Literature review**
   - Review of current concepts and situation

3. **Case study development**
   - Oct 05 – Apr 06
   - Local-State State-National in Australia
   - Integration template
   - Distributed to PCGIAP members June 06

4. **Pilot study**
   - June 2007

5. **Compare & contrast case study results**
   - Mar-June 07

6. **Develop integration framework, data model & associated tools**
   - March 2007
   - Sept 2006

7. **Develop implementation Strategy**
   - June 2007
   - Aug 2008

8. **LA Seminar – 14TH PCGIAP**
   - Dec 2008

9. **Project End**
   - June 2005

10. **International workshop – UNRCC-AP**
    - Sept 2006
Project Activities and Outcomes

• **Major Activities**
  – International Case Studies
  – International Workshop on Data Integration as part of the 17th UNRCC-AP, 2006
  – Pilot study of Geo-web Service Integration
  – Development of a data integration framework
  – Development of a web-based system using the data integration framework to facilitate data interoperability test and data integration.
  – Publication Strategy

• **Spatial data integration validation tool**
• **Spatial data integration tool**
• **Spatial data integration guideline**
• **Technical and policy reports**
• **Publications**
  – 5 international papers
  – 2 book chapters
  – Project website
PCGIAP Working Group 3
Research Project
Integration of Built and Natural Environmental Datasets
within National SDI Initiatives

Spatial Data Integration Validation Tool
Specification and User Guide

August 2008

Centre for Spatial Data Infrastructures and Land Administration
Department of Geomatics, University of Melbourne
UNRCC-AP & PCGIAP International Workshop on Data Integration

ARC Linkage Project - the Integration of Built and Natural Environmental Datasets in National Spatial Data Infrastructure Initiatives

This project is an Australian Research Council Linkage Project involving researchers from the Centre for Spatial Data Infrastructure and Land Administration at the University of Melbourne and three industry partners, as listed below. The project also involves collaboration with Working Group 3 (Cadastral) of the UN supported Permanent Committee for GIS Infrastructure for Asia Pacific (PCGIAP), which helps to involve seven different countries in the project.

Project Partners
Project Duration
Personnel
Project Overview
Project Background

Project Partners

Victoria
DSE, Victoria
Department of Lands
Department of Lands, NSW
Geoscience Australia
Data Integration Project

Technical Details

Hossein Mohammadi
PhD Candidate
Centre for SDIs and Land Administration
Department of Geomatics, The University of Melbourne
Why data integration?

- **Fundamental Data**: Cadastre, Roads, Imagery, Topography, Census, Administrative Boundaries, Street Directory, Waterways, Vegetation.
- **Infrastructure Data**: Electricity, Sub-stations, Gas, Water, Hydrants, Sewerage, Storm-water, Telecoms, Ferry Routes, Bush Fire Prone Zones, Flood Planes, 3D Buildings.

Rob Colles (EICU), 2005
Why data integration?

- Hazard management, Tsunami
- Gaps of Data, No Vertical Topology, Currency, Collaboration, Data Exchange Policy,
  - Restricted Privacy
  - Different Custodians
  - Not Completely Fits to Utility and Access Network
  - Privacy
  - Poor Attribution
- Data Barely Exists
- Not Created and Owned by Users
- No Nation-wide Data
- Poor Metadata
- Poor Metadata
- Data Barely Exists
- Cadastre
- Bathymetry
- Building Data
- Utilities

- Restricted Privacy
- Different Custodians
- Bathymetry
Why data integration?

• PSMA’s Experience

  • G-NAF
  • Inconsistencies in Data Model, Metadata, Standards, Specifications, Collaboration Model, and IP issues in Victoria, NSW, GA, …

  • CadLine
  • Transport and Topography
  • Admin Boundaries
  • Points of Interest

IDM
Example

- Data duplication
Example

- Completeness
## Data model

<table>
<thead>
<tr>
<th>Theme</th>
<th>Victoria</th>
<th>NSW</th>
<th>GA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Admin</td>
<td>1</td>
<td>4</td>
<td>11</td>
</tr>
<tr>
<td>2. Admin</td>
<td>23</td>
<td></td>
<td>12</td>
</tr>
<tr>
<td>3. Elevation</td>
<td>8</td>
<td>4</td>
<td>13</td>
</tr>
<tr>
<td>4. Hydrography</td>
<td>8</td>
<td></td>
<td>14</td>
</tr>
<tr>
<td>5. Planning</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Property</td>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Transport</td>
<td>8</td>
<td>4</td>
<td>15</td>
</tr>
<tr>
<td>8. Vegetation</td>
<td>3</td>
<td>4</td>
<td>16</td>
</tr>
<tr>
<td>9. Imagery</td>
<td></td>
<td>4</td>
<td>17</td>
</tr>
<tr>
<td>10. …</td>
<td></td>
<td></td>
<td>18</td>
</tr>
</tbody>
</table>

Theme

NSW

1. Admin 4
2. Aviation 1
3. Cartography 4
4. Culture 3
5. Drainage 7
6. Errors 1
7. Framework 4
8. Habitation 5
9. Industry 3
10. Marine 4
11. Physiography 1
12. Production 1
13. RailTransport 5
14. RoadTransport 3
15. SeriesIndex 4
16. Relief 2
17. Surveymarks 1
18. Utility 2
19. Vegetation 4
20. Waterbodies 6
21. …
Example

- Data model

<table>
<thead>
<tr>
<th>Data</th>
<th>Victoria</th>
<th>NSW</th>
<th>GA</th>
</tr>
</thead>
<tbody>
<tr>
<td>File Name</td>
<td>No particular category for Tank points (HY-WATER_STRUCT_POINT) Beside watering places and swimming pools</td>
<td>TankPoint</td>
<td>Point_watertank</td>
</tr>
<tr>
<td>Category</td>
<td>Hydro</td>
<td>TopoFD</td>
<td>Drainage</td>
</tr>
</tbody>
</table>
Example

- Feature Level and Theme Level Metadata

### Feature Level

**PlaceArea**

- Description: A polygon feature class defining a named place.
- GeometryType: GeometryPolygon

<table>
<thead>
<tr>
<th>Attribute Name</th>
<th>Data Type</th>
<th>Allow Nulls</th>
</tr>
</thead>
<tbody>
<tr>
<td>PlaceName</td>
<td>SmallInteger (9)</td>
<td>F</td>
</tr>
<tr>
<td>Population</td>
<td>Integer (7)</td>
<td>T</td>
</tr>
<tr>
<td>PopulationDate</td>
<td>DateTime</td>
<td>T</td>
</tr>
<tr>
<td>PlaceSignificance</td>
<td>SmallInteger (9)</td>
<td>F</td>
</tr>
<tr>
<td>GeneralNameON</td>
<td>Integer (9)</td>
<td>T</td>
</tr>
<tr>
<td>GeneralName</td>
<td>Char (50)</td>
<td>T</td>
</tr>
<tr>
<td>AlternativeLabel</td>
<td>Char (50)</td>
<td>T</td>
</tr>
<tr>
<td>OperationalStatus</td>
<td>SmallInteger (9)</td>
<td>F</td>
</tr>
<tr>
<td>TopoID</td>
<td>Integer (9)</td>
<td>T</td>
</tr>
<tr>
<td>CreateDate</td>
<td>DateTime</td>
<td>T</td>
</tr>
<tr>
<td>UpdateDate</td>
<td>DateTime</td>
<td>T</td>
</tr>
<tr>
<td>FeatureModDate</td>
<td>DateTime</td>
<td>T</td>
</tr>
<tr>
<td>ClassType</td>
<td>Integer (10)</td>
<td>T</td>
</tr>
<tr>
<td>FeatureRelativityType</td>
<td>DateTime</td>
<td>T</td>
</tr>
<tr>
<td>AttributeRelativityDate</td>
<td>DateTime</td>
<td>T</td>
</tr>
<tr>
<td>CaptureSourceCode</td>
<td>SmallInteger (9)</td>
<td>T</td>
</tr>
<tr>
<td>CaptureSource</td>
<td>SmallInteger (9)</td>
<td>T</td>
</tr>
<tr>
<td>PlanimeterAccuracy</td>
<td>Single (6.2)</td>
<td>T</td>
</tr>
<tr>
<td>VerticalAccuracy</td>
<td>Single (6.2)</td>
<td>T</td>
</tr>
</tbody>
</table>

### Theme Level

**Full Description Report**

- Name: VICMAP_TRANSPORT
- Spatial Reference: WGS84
- Data Source: VICMAP
- Data Accuracy: Good
- Data Version: 2009
- Data License: Open Access
- Data Distribution: Via the Victorian Government's Spatial Data Library
- Data Coverage: Victoria, Australia
- Data Availability: Online at [VICMAP Website](http://www.vicmap.vic.gov.au)

**Layer Design Summary**

- Layer Design Considerations: Compatibility with other layers, ease of use, data quality, and relevance to end users.
- Data Currency Information: Regular updates and maintenance.
- Data Status: Current
- Data Format: Geodatabase
- Data Content: Includes topographic, transportation, land use, and environmental data.
Stage 1: Current situation of multi-source spatial data integration
Existing documents and resources
- Bhutan
- Canada
- India
- Botswana
- China
- Greece
- WA, Australia

Stage 2: The investigation of technical and non-technical issues
Questionnaires and country reports
- Australia
- Japan
- Singapore
- Malaysia
- Philippines
- Indonesia
- Brunei Darussalam

Stage 3: Detailed investigation of technical and non-technical issues
Agency visits, meetings and data audit
- Victoria, Australia
- NSW, Australia
- GeoScience, Australia

Potential Integration Issues and Obstacles
Tools and Requirements of effective spatial data integration
Potential integration issues and challenges

- Funding Model Differences
- Utilizing Inconsistent Collaboration Models
- Lack of Awareness of Data
- Definition of Rights, Restrictions and Responsibilities
- Inconsistency in Copyright and Intellectual Property Rights (IPR) Approaches
- Different Data Access and Privacy Policies
- Weakness of Capacity Building Activities
- Cultural Issues
- Different Background of Stakeholder

- Lack of Supporting Legislations
- Inconsistency in Policy Drivers and Priorities
- Pricing
Tools and requirements for effective spatial data integration

- Integration data model
- Data validation tool
- Guideline
  - Custodianship program
  - Metadata requirements
  - Policy arrangements
  - Etc.
Integration data model

- **Physical Universe**: Parcel as environmental feature
  - resources, habitat, etc
  - aspatial + spatial area, buffer, point in polygon, topology etc.
  - Parcel as coordinates

- **Legal**: Parcel as legal entity
  - rights, restrictions, legal boundaries, interference of interests

- **Environment Expert**: Parcel as environmental feature

- **Surveyor**: Parcel as coordinates
  - (x, y)

- **GIS Expert**: Parcel as aspatial + spatial area
Integration data model: example

• Victoria
  – 12 road classes
  – Stored in a single layer and singled out by attributes

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Freeway</td>
</tr>
<tr>
<td>1</td>
<td>Highway</td>
</tr>
<tr>
<td>2</td>
<td>Arterial</td>
</tr>
<tr>
<td>3</td>
<td>Sub-Arterial</td>
</tr>
<tr>
<td>4</td>
<td>Collector</td>
</tr>
<tr>
<td>5</td>
<td>Local</td>
</tr>
<tr>
<td>6</td>
<td>2wd</td>
</tr>
<tr>
<td>7</td>
<td>4wd</td>
</tr>
<tr>
<td>8</td>
<td>Unknown</td>
</tr>
<tr>
<td>9</td>
<td>Proposed</td>
</tr>
<tr>
<td>11</td>
<td>Walking Track</td>
</tr>
<tr>
<td>12</td>
<td>Bicycle Track</td>
</tr>
</tbody>
</table>
Integration data model: example

- New South Wales
  - 7 road classes
  - Stored in different layers
Integration data model: example

- PSMA (Public Sector Mapping Agency)
Freeway (adopted from VicMap)

Hard surface formation, high volume, high speed roads declared as “Freeway”; comprising dual carriageway and full access control and grade separated intersections; ie no direct access from adjoining properties or side roads and all crossings are by means of overpass or underpass bridges with traffic entering or leaving carriageways by means of ramps. Single carriageway sections forming part of declared freeways may be included within this category.
Integration data model

- Freeway (adopted from VicMap)
  - hard surface
  - high volume
  - high speed
  - dual carriageway
  - full access control
  - grade separated intersections
  - no direct access from adjoining properties or side roads
  - all crossings are by means of overpass or underpass bridges
  - traffic enters and leaves the carriageway by means of ramps

Can be defined in attributes and extracted by SQL queries

Can be extracted by topological relations with other spatial datasets and extracted by spatial queries
Guideline: Metadata requirements

• Measurable content
  – Attribute and spatial accuracy
  – Restrictions/rules on attributes
  – Scale
  – Logical consistency
  – Etc.
Guideline: Custodianship arrangement

• Agreements on
  – Accuracy, completeness, currency and logical consistency
  – The provision of effective metadata
  – Effective collaboration
  – Adoption of a nationally consistent data model
  – Adoption of nationally consistent data content (theme-based)
  – etc
Data validation tool

Real World Applications
(land administration, emergency management, resource management)

User Data Validation

FITNESS-FOR-PURPOSE
Data manipulation

Spatial Data Integration
Data validation tool

- 
- Standards

- Guidelines

- Datum

- Format

- Spatial Extent

- Attributes

- Quality

- Restrictions etc.

- Integratability

- WebService

- Web-Servers (WFS, WMS)

- Local dBases

- Remote dBases

- Modify

- Revision Instruction

- Definition of Inconsistency

- Superimpose

- Comprehensive Report

- Restriction of use
Data validation tool
Data validation tool

- **Users**
  - Define queries to retrieve datasets
- **System coordinator**
  - Define measures
  - Make configurations
- **Data supplier**
  - Define new session for data register
  - Resume previously commenced sessions
  - Evaluate datasets
  - Amend data based on instructions
Structure of the tool

Database
Manages authority access level, data register sessions, data records, measures and rules, jurisdiction specifications

GIS environment
Provide display, analysis and amendment services

Development environment
Manages queries on database, link with GIS, data validation and manipulation

System architecture
System architecture: Relational Database

- Entities
  - System user
  - Register sessions
  - Issue record
  - Data record
  - Jurisdiction
  - Jurisdiction entities: Custodian, rules, datum, metadata (schema), bounding box, restrictions, scale and etc
System architecture: Development

- ESRI ArcGIS toolbar: developed by macro programming
System architecture: Development

User

Database coordinator

Data provider

User/data retriever
System architecture: Development

- Jurisdiction (context) selection
System architecture: Development

- Data and metadata entry, additional rule definition
### Data measures

<table>
<thead>
<tr>
<th>Measure</th>
<th>Accepted</th>
<th>Reason</th>
</tr>
</thead>
<tbody>
<tr>
<td>Policy agreement has been accepted.</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Data's coordinate system is the same as Jurisdiction-defined one (EPSG, GDA, 1994)</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Data format is compliant with the format list defined by Jurisdiction.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Metadata measures

<table>
<thead>
<tr>
<th>Measure</th>
<th>Accepted</th>
<th>Reason</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metadata Standard is the same as jurisdiction-defined one (AKNZ,1K) and has not been utilized in assessments.</td>
<td>✗</td>
<td></td>
</tr>
<tr>
<td>The currency of data is not accepted by jurisdiction.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The scale of data is accepted by jurisdiction.</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>The completeness of data is not accepted by jurisdiction.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The attribute accuracy of data is not accepted by jurisdiction.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The spatial accuracy of data is not accepted by jurisdiction.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The logical consistency of data is not accepted by jurisdiction.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The custody of data is accepted by jurisdiction.</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>User complied with the restrictions defined by jurisdiction.</td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>

### Rules

<table>
<thead>
<tr>
<th>Rule</th>
<th>Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rule1:</td>
<td>ELEVATION &lt;= 100</td>
</tr>
<tr>
<td>Rule2:</td>
<td></td>
</tr>
<tr>
<td>Rule3:</td>
<td></td>
</tr>
<tr>
<td>Rule4:</td>
<td></td>
</tr>
</tbody>
</table>
System architecture: Development

- Amendment services: attribute manipulation, format conversion, datum conversion
System architecture: Development

- Record and Display
Scenario: Setting

- Accepted Formats
  - ESRI shapefile, ESRI coverage, WFS/WMS, ERMapper ECW
- Metadata Standard and DTD/Schema
  - ANZLIC’s profile, anzmeta1.3.dtd
- Policy agreement
  - Privacy policy
  - Restrictions
  - Licensing
- Geographical extent
- Australia’s boundaries
- Datum
  - GDA_1994
- Currency
  - Not older than one month
- Minimum and Maximum scale
  - 1:1,000 to 1:1,000,000
- Completeness
  - complete
- Attribute and positional accuracies
- And user-defined rules:
  - No “null” is accepted for attributes
Dataset: contour lines

- Format
  - MapInfo TAB
- Metadata and DTD
  - ANZLIC’s profile, anzmeta1.3.dtd
- Geographical envelope
  - Australia, Victoria
- Coordinate system
  - WGS86
- Currency
  - Two weeks old
- Scale
  - 1:25000
- Complete
- Some “null” attributes
**Evaluation process**

- Information is extracted from data and metadata
  - Bounding box, restrictions, accuracies, scale, attribute content, format, datum, completeness
- This information is compared against the information defined by the system/jurisdiction
- If any incompliancy occurs, it will be reported and guidelines and if applicable tools to overcome will be provided
- Upon amendment, comparison module evaluates the data and if there is no more issues a record of data is added to system
Amendment tools

- Amendment services: attribute manipulation, format conversion, datum conversion
Record the dataset and display
System capabilities

• Capability to adopt new jurisdiction-specific rules and standards
• Extendable to accept new data formats and metadata standards
• Capability to be utilized as a data validator
• Capability to be a part of bigger spatial services
• Capability to develop new metadata and metadata schema
Data Integration Validation Tool is available online

- Software together with specification, manual and installation guide is available through the project website at:

  www.geom.unimelb.edu.au/research/SDI_research/Integrated

Or

- if you require further information or want to report a feedback, please contact PCGIAP-Working Group 3
Specification and Guideline

• Tools
  – Data validation
  – Datum conversion
  – Interoperability tool/data conversion
  – Attribute manipulation
  – Metadata entry tool
Tool Functions

- System configuration
Tool Functions

- System configuration
- Data validation
**Tool Functions**

- System configuration
- Data validation
- Data manipulation
Tool Functions

- System configuration
- Data validation
- Data manipulation
- Data registry
Tool Functions

- System configuration
- Data validation
- Data manipulation
- Data registry
- Data display
Hardware/Software requirements

- ArcGIS 9.x
- Internet connection (in case of using data or metadata/metadata schema from internet)
- Integration.zip (from project website)
Thank you for your attention