Integrating Geospatial Elements into the ABS Information Model

Workshop on Integrating Geospatial and Statistical Standards – Stockholm, November 2017

Lead Author: Tom Walter – Assistant Director
Presented by: Martin Brady – Director
Geospatial Solutions, Australian Bureau of Statistics
Transforming for the future

**ENVIRONMENT**
We collaborate with stakeholders to understand and respond better to the current and future external environment

**STRATEGY**
Our strategies enable rigorous statistics, strong partnerships, and effective use of resources

**GOVERNANCE**
Our governance supports responsive decision making, prioritisation and management of enterprise risk

**TRANSFORMATION GOALS**

**INFRASTRUCTURE**
Our infrastructure is effective, efficient and adaptable

**CULTURE**
We are high performing, aligned, engaged, innovative and accountable

**PEOPLE**
We have a diverse, expert, motivated and agile workforce
Government Investment Objectives

- reduce statistical risk
- reduce time to market
- reduce cost
- grow the business
- reduce red tape
The ABS is undergoing a major infrastructure transformation

- statistical modernisation project to upgrade and update aging siloed information technology infrastructure.
- adapting Generic Statistical Business Process Model (GSBPM), called Statistical Production Activity Model (SPAM).
- developing a new ABS Information Model (AIM) based on the Generic Statistical Information Model (GSIM).
- SPAM and AIM will help to reduce workloads through the reuse of data, metadata and processes - meaning less time is spent on repetitive work, and more time is available for analysis and innovation.

Problem: GSIM and GSBPM do not effectively incorporate concept of "Location“ or geospatial
The Generic Statistical Business Process Model

Geographic coverage and disaggregation requirements

Consider alternate data sources - geo

Geographic classifications

Geospatially enabled registers

Acquire geospatial/location data

Geospatial intelligence and tools for field operations

Geospatial intelligence and tools for coding, validation and editing

Link datasets based on location

Enhance data with geostatistics

Map data visualisations

Geo web services

Geographic correspondences

Geospatial disaggregation methods
Conclusions – Statistical Modernisation and Transformation

For this reason, any program to modernise statistical systems, processes, services and tools must consider and implement appropriate processes, concepts, value domains and metadata structures associated with location/geospatial information.

- Geospatial processes need to be recognised within GSBPM.
- Application and integration of existing geospatial metadata standards and terminology to this process is essential.
An updated statistical model

• Early development of the ABS Information Model (AIM) did not include geospatial elements
• Geospatial specialists identified this gap and work began with statistical information architects to address issues
• The first efforts tried to re-use existing statistical structures to hold geospatial classifications and geospatial data with little success

➢ To overcome these limitations, new AIM structures are being developed specifically for geospatial data and metadata
New geospatial elements

The current solution is undergoing testing and adds the following new elements to the AIM:

1. **Geometry Value Domain** – a new Value Domain to hold geospatial data formats

2. **Geospatial Feature Statistical Unit** – a new Statistical Unit type that data is captured from or about (e.g. a property or grid cell)

   [*Statistical Unit – different to INSPIRE/core data*]

3. **Geospatial Family Structure** – a new structure to hold the ABS’ own geospatial boundary classification for Australia, the Australian Statistical Geography Standard (ASGS)
Geometry value domain
Variable description - Sex

**Concept**
Sex

**Unit Type**
Person

**Variable**
Sex of Person

**Represented Variable**
Sex of person: sex code list, code items 1-3

**Enumerated Value Domain**

**Statistical Standard (Code List)**
Sex

**Code Item**
1. Male
2. Female
3. Other
Variable description – Location Address

- **Concept**: Location
- **Unit Type**: Person
  - **Variable**: Location of Person
    - **Represented Variable**: Location of person by: Address
      - **Described Value Domain**
      - **Text**
        - **Address Standard**
Variable description – Location Geometry

Concept
Location

Unit Type
Person

Variable
Location of Person

Represented Variable
Location of person by: Geometry

Geometry Value
Domain

Geometry Format

Others formats as required

Oracle SDO
Geometry
Describing a variable in AIM

- **Meaning, key idea measured**
  - **Concept**
  - **Unit Type**

- **Variable** (Concept + Unit Type)
  - **Represented Variable** (Variable + Value Domain)
    - **Instance Variable**
    - **Value Domain**
      - **Enumerated Value Domain**
        - **Statistical Standard**
        - **Geospatial Standard**
      - **Described Value Domain**
        - **Text**
        - **Numeric**
        - **Date / Time**
      - **Geometry Value Domain**
        - **Geometry Format**
Geospatial Feature Statistical Unit
Concept: Land Use

Unit Type: Land Parcel

Variable: Land Use of Land Parcel

Represented Variable: Land Use of Land Parcel: Land Use code list, code items 1, 2, 3, etc...

Enumerated Value Domain

Statistical Standard (Code List) Land Use

Code Item 1. Residential

Code Item 2. Commercial

Code Item 3. Etc…
Geospatial Family Structure
DIAGRAM 1: ASGS ABS STRUCTURES

- Remoteness
  - UCL/SOS
  - Indigenous
    - Main
    - GCCSA
    - SUA
  - Section of State (SOS)
  - Section of State Ranges (SOSR)
  - Urban Centres and Localities (UCL)
  - Indigenous Locations (ILOC)

- Indigenous
  - Indigenous Regions (IREG)
  - Indigenous Areas (IARE)

- Statistical Area
  - Level 4 (SA4)
  - Level 3 (SA3)
  - Level 2 (SA2)
  - Level 1 (SA1)

- Greater Capital City Statistical Areas (GCCSA)
  - Significant Urban Areas (SUA)

- Mechs Block (MB)
New Geospatial Structures in AIM

- **Geospatial Family**: Overarching geospatial structure
  - «Set of»
  - «References»
  - «Set of»

- **Geospatial Standard**: e.g. Australian Statistical Geography Standard (ASGS) 2016
  - «References»
  - «Set of»

- **Geospatial Path**: e.g. Statistical Area 2 to Significant Urban Area
  - «Set of»
  - «References»

- **Geospatial Feature**: e.g. North Sydney
  - «Tree Link»
  - «IsOfType» Meta

- **Geometry Value Domain**

- **Geospatial Feature Type**: e.g. Statistical Area 2
  - +Type
# Geospatial and Statistical Classification Structures

## Geospatial Family Structures and Statistical Classification Structures in the ABS Information Model

<table>
<thead>
<tr>
<th>Geospatial Family</th>
<th>Classification Family</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Geospatial Standard</strong></td>
<td><strong>Statistical Classification</strong></td>
</tr>
<tr>
<td><em>ASGS – geographic regions classification</em></td>
<td><em>International Standard Classification of Occupations</em></td>
</tr>
<tr>
<td><strong>Geospatial Feature Type</strong></td>
<td><strong>Classification Level</strong></td>
</tr>
<tr>
<td><em>Local Government Areas</em></td>
<td><em>Major occupation group</em></td>
</tr>
<tr>
<td><strong>Geospatial Feature</strong></td>
<td><strong>Classification Item</strong></td>
</tr>
<tr>
<td><em>North Sydney Council</em></td>
<td><em>Professionals</em></td>
</tr>
<tr>
<td><strong>Geospatial Path</strong></td>
<td></td>
</tr>
<tr>
<td><em>Mesh Blocks aggregate to Local Gov’t Areas, SA1s, Suburbs, etc</em></td>
<td></td>
</tr>
</tbody>
</table>
Challenges, lessons and recommendations - concepts and language across communities

**Challenge** = achieving common understanding across the different concepts and language used in the statistical and geospatial domains

**Recommendation** = The statistical and geospatial communities work together to clarify, standardise or correspond the concepts and language that are applied in both domains, and that this is applied in International Statistical Modernisation documentation and relevant standards.
Challenges, lessons and recommendations - GSBPM and GSIM incorporate geospatial

Challenge = include geospatial processes in Business Process Maps (BPMs) to ensure they are considered in future design work in the organisation.

Recommendation = The GSBPM and GSIM are updated to incorporate geospatial concepts, systems, processes, services and tools in the statistical business processes.
Challenges, lessons and recommendations
- engage with semantic web community

**Challenge** = inclusion of geospatial data within semantic web for creating and maintaining linkages between the geospatial and statistical domains

**Recommendation** = The statistical standards community should engage with the community developing semantic web standards and best practice documents to ensure the statistical community benefits from and stays relevant to the users of semantic web technologies.
Questions?