UN Expert Group on the Integration of Statistical and Geospatial Information (EG-ISGI) 5th meeting
➤ GSGF Principle 2 Definition

Martin Brady
Co-Chair UN Expert Group on the
  Integration of Statistical and Geospatial Information
Director – Geospatial Solutions
Australian Bureau of Statistics

Principle 2 – Progress

- Reviewed content based on feedback

- Expanded descriptions and addressed presumed knowledge
    - more context and description
    - more examples to provide context
    - linkage and description of benefits

- Made content more broadly applicable to different country situations and pathways

- Recommendation that statistical unit records have a location or geographic reference!

Principle 2 – Outline

Principle 2 of the Global Framework recommends that the linkage of a geocode for each statistical unit record occur within a data management environment. This enables the following to be achieved:

- Allow the statistics generated from these datasets to be produced for a wide range of geographic contexts

- Enable production of value added data from geospatial analysis

- Enable data linkage processes using a range of sources.

- Ensure data can be managed securely

Principle 2 – fit with other principles

Pr.3 Common geographic boundaries for dissemination of statistics

Definition of geographic regions and aggregation/ disaggregation of data to regions.

Pr.2 Geocoded unit record data in a data management environment

Apply Pr.1 elements to statistical and administrative data within statistical infrastructure

Pr.1 Use of fundamental geospatial infrastructure and geocoding

Primarily geospatial community data, tools and standards

Principle 1 & 2 - clarification

Principle 1 – focusses on infrastructure

- Provision of addressing standards and infrastructure, such as address registers

➢Ensures geocoding infrastructure (systems and tools) is as standardised as possible

Principle 2 – geocoding of unit records

- Application of geocoding infrastructure to unit records

➢Ensures statistical infrastructure can use geospatial infrastructure and standards for geocoding

Principle 2 – Objectives

- **All statistical unit records should include or be linked to a geocode.**

- Ensure the effective implementation of fundamental or national geospatial and geocoding infrastructure and demonstrate its broader value.

- Implement effective data management of statistical and geospatial data.

- Ensure appropriate protection of privacy and secrecy of unit record or microdata level datasets.

- Storage of consistent and interpretable geocodes, preferably linked from a "point-of–truth".

- Establish tools and methods to enable simplified geographic aggregation of data.

- Ensure that data is stored in a way that will facilitate flexible use of geocoded unit records in future aggregations, analysis and visualisation.

Principle 2 – Relationship to others

Principle 1:

– Draw on fundamental or national geospatial data and infrastructure and geocoding capabilities

Principle 3:

– Definition of common geographic regions for the dissemination of data and associated metadata and data.

– Methods for aggregation and disaggregation of data to regions.

Principle 2 – Inputs

Standards, frameworks, infrastructure, and best practice.

1. Agreed statistical and geospatial data management frameworks.

2. Addressing and/or location reporting standards and infrastructure.

3. Geocoding tools and metadata standards.

4. Promotion of point-of-entry address validation and geocoding.

Stat and Geo Communities

Pr.1

Pr.1

Pr.1

Principle 2 – Inputs

Standards, frameworks, infrastructure, and best practice (cont.)

5. National privacy laws and/or agreed privacy standards (UNFPOS).

6. Agreed geographic regions and infrastructure.

7. Global or national/regional Geodetic Reference Frames.

Stats Community

Pr.3

Geo Community

UN-GGIM | United Nations Commitee of Experts on Global Geospatial Information Management

BETTER DATA
BETTER LIVES

unstats.un.org
@UNStats #UNStatCom

Principle 2 – External dependencies

Statistical – GSBPM/GSIM managed by the UNECE HLG-MOS

These models are being examined so that they better incorporate geospatial tools, methods and processes into their definitions and resources. There are limited examples of current country applications

GEOSTAT projects are ESSnet projects to foster a better integration of geospatial information and statistics

The main objective of GEOSTAT 3 is to develop recommendations for a harmonised implementation of the GSGF for the ESS.

Provide countries and regions with an example of GSGF implementation plus more detailed resources and practical implementation guidance.

Geospatial – OGC/ISO/IHO Standards

These standards will evolve and change through OGC/ISO/IHO processes - OGC are currently developing geocoding API standards

Principle 2 – Community Roles

Geospatial community

– Provision of fundamental geospatial data and infrastructure, and geocoding capabilities

– Global or national/regional Geodetic Reference Frames and implementations

– Geospatial data management frameworks

– Geospatial data standards, particularly geocoding metadata specifications

– Supporting common geographic boundaries

Principle 2 – Community Roles

Statistical Community

– National and international privacy protocols (e.g. UN Fundamental Principles of Official Statistics)

– Statistical data management frameworks

– Supporting common geographic boundaries

– Implementation of principles to statistical and administrative unit record data and their storage and management

Administrative Data Community

– Implementation of principles to administrative unit record data and their storage and management

Note: differences between community roles may occur at the national level

1. Geocoding guidance material

2. Best practice data and metadata management

3. Guidance on protecting unit record privacy and secrecy

Principle 2 – Materials

Geocoding guidance material

A. Application of addressing/location reference standards.

B. Geocoding methods.
    - Direct coordinate capture
    - Address and location coding (incl. POE).
    - Geographic correspondences and allocations.

C. Data and metadata management.

Statistical Spatial Framework
Australian Bureau of Statistics

SSF Guidance Material –
Geocoding Unit Record Data Using Address and Location

UN-GGIM | United Nations Commitee of Experts on Global Geospatial Information Management

Statistical Commission

BETTER DATA
BETTER LIVES

unstats.un.org
@UNStats   #UNStatCom

Location and regional information pathways

External Environment

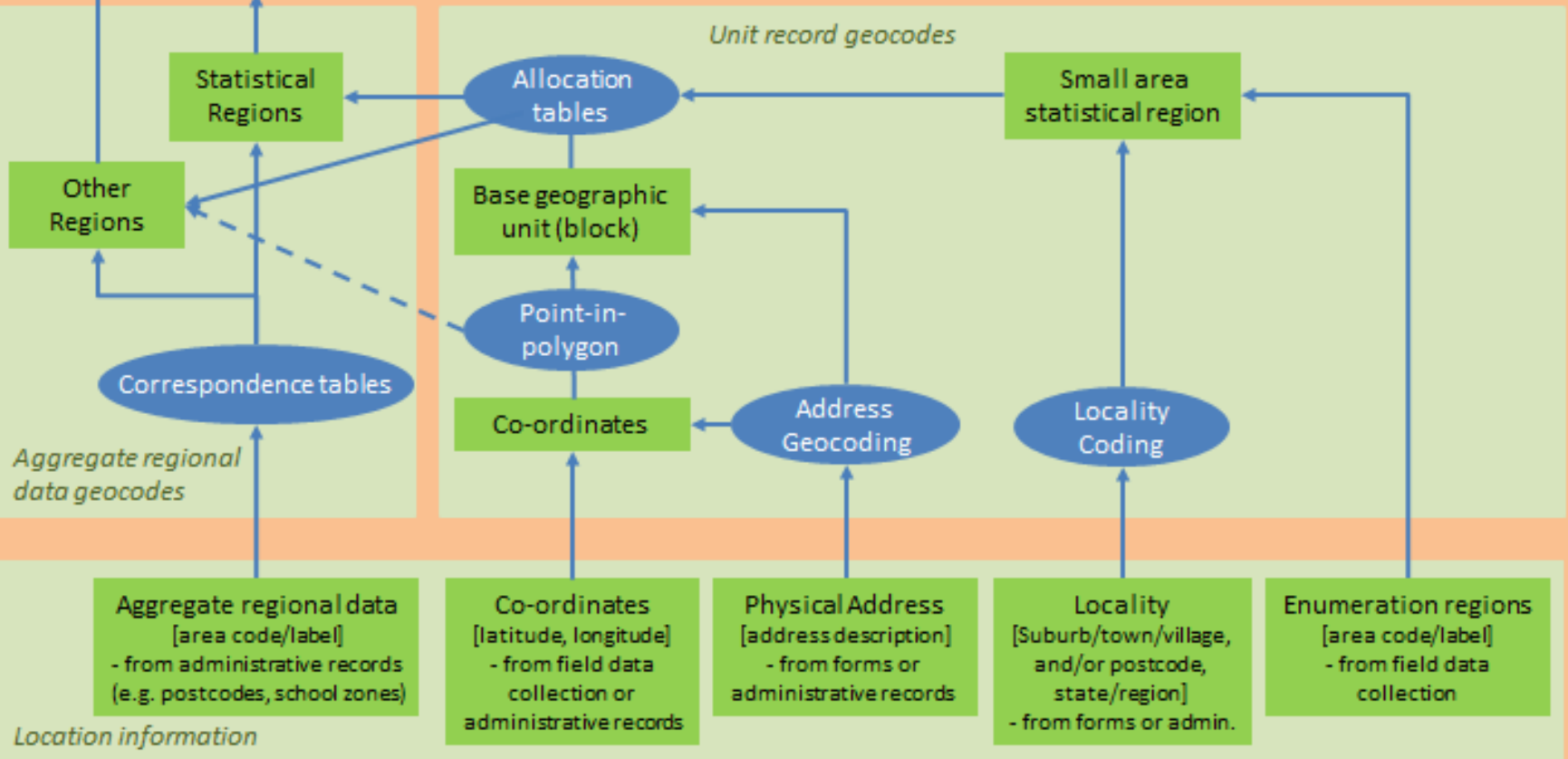Regional data / Map visualisations

Diagram legend

Data type

Data states

Preferred transitions
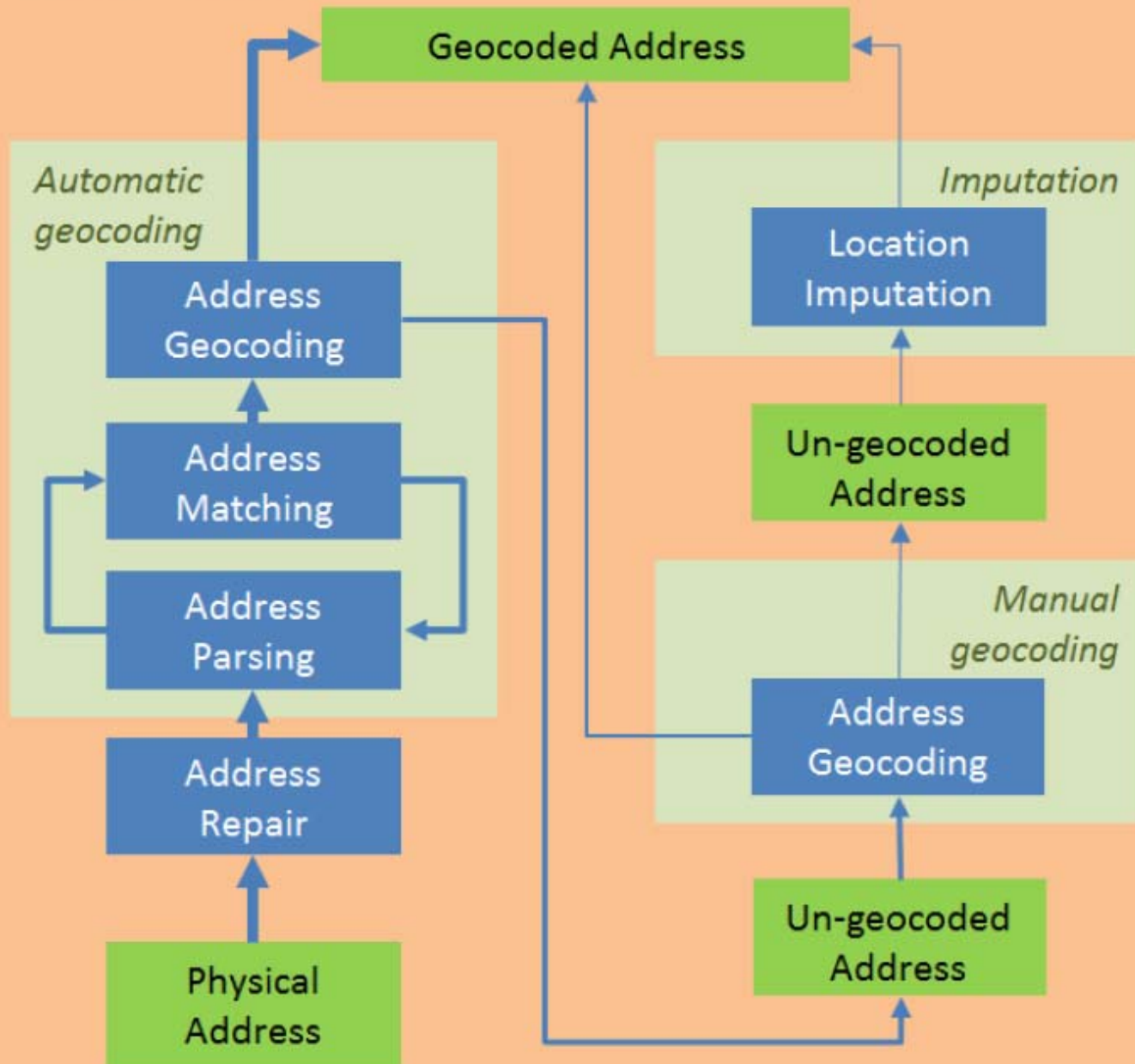
Other transitions

Data Processes

Internal Environment

Statistical/ GIS Tools

Unit record geocodes

Statistical Regions

Allocation tables

Small area statistical region

Other Regions

Base geographic unit (block)

Point-in-polygon

Correspondence tables

Co-ordinates

Address Geocoding

Locality Coding

Aggregate regional data geocodes

Aggregate regional data
[area code/label]
- from administrative records
(e.g. postcodes, school zones)

Co-ordinates
[latitude, longitude]
- from field data collection or administrative records

Physical Address
[address description]
- from forms or administrative records

Locality
[Suburb/town/village, and/or postcode, state/region]
- from forms or admin.

Enumeration regions
[area code/label]
- from field data collection

Location information

Address geocoding process

Geocoded Address

Automatic geocoding

Address Geocoding

Address Matching

Address Parsing

Address Repair

Physical Address

Imputation

Location Imputation

Un-geocoded Address

Manual geocoding

Address Geocoding

Un-geocoded Address

1. Geocoding guidance material

2. Best practice data and metadata management

   A. Statistical – GSBPM/GSIM & Case Studies

      ➢ GEOSTAT3 examining aspects

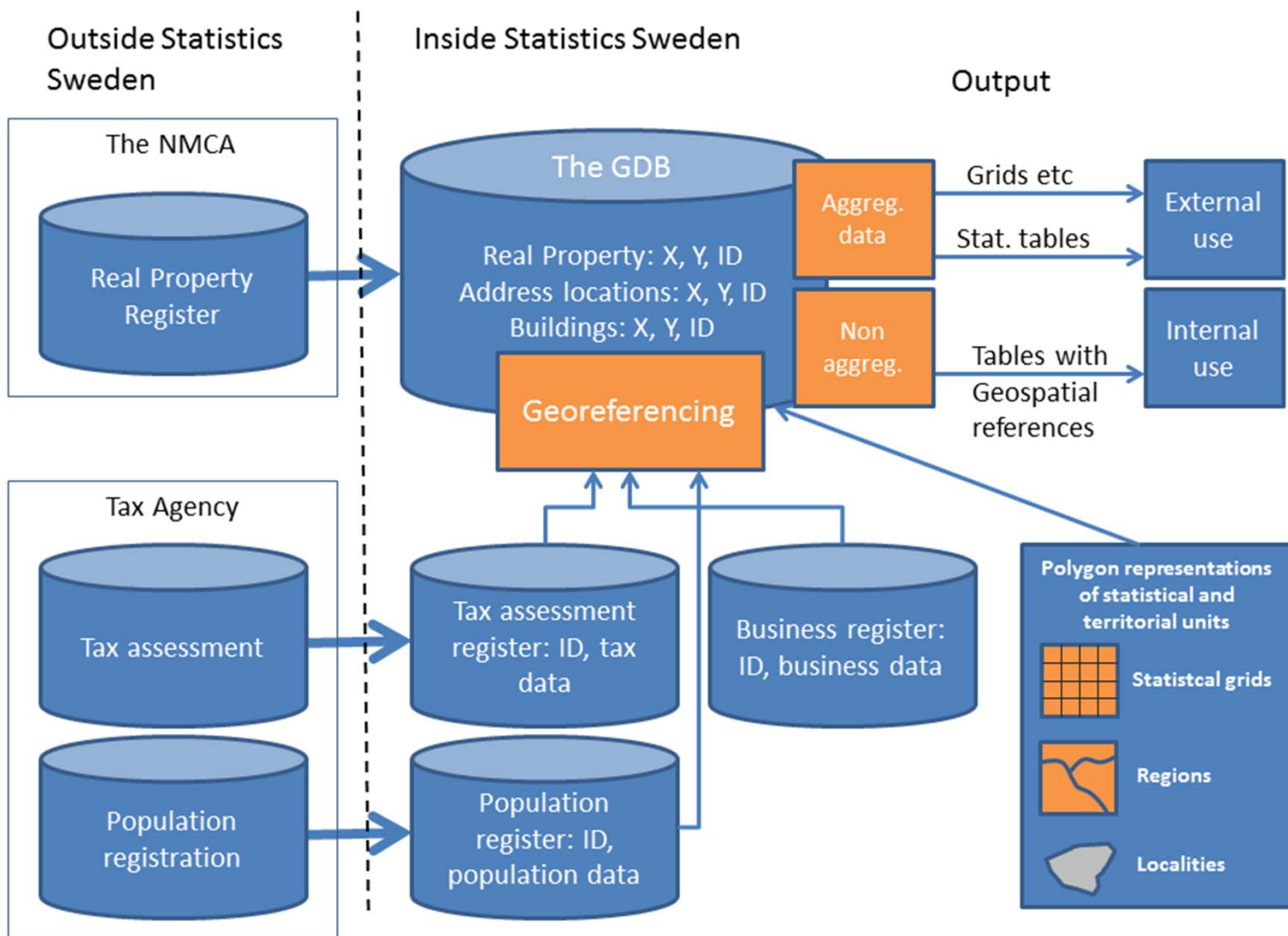   B. Geospatial – OGC/ISO Standards

      ➢ OGC Geocoding standards

UN-GGIM | United Nations Commitee of Experts on Global Geospatial Information Management

BETTER DATA
BETTER LIVES

unstats.un.org
@UNStats   #UNStatCom

Statistics Sweden geo data environment

Principle 2 – Concepts

Need to develop list of concepts:

– Geocoding

– Geocoding services

– Point-of-entry address validation

– Georeferencing

– Standards/Framework/Models

– Geographies *(related terms - statistical and administrative geographies, statistical regions, statistical units, regional divisions, statistical areas, functional areas)*

Principle 2 – Future plans

1. Collect country practices for geocoding, data management and privacy protection.

2. Geocoding guidance material
    – develop document from best practice

3. Data and metadata management
    – monitor and contribute to international developments: UNECE-HLG, OGC, GEOSTAT3

4. Guidance on protecting unit record privacy and secrecy
    – identify documentation of existing best practice