The Global Statistical Geospatial Framework and the Aggregation of Geocoded Unit Level Data

Tim Trainor

Former Co-Chair UN-GGIM

Former Chief Geospatial Scientist
U.S. Census Bureau

*Slides from members of the EG-ISGI

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

BETTER DATA BETTER LIVES
unstats.un.org
@UNStats #UNStatCom

Outline

1. Background to the Global Statistical Geospatial Framework

2. Bridge between statistical and geospatial communities

3. Geocoding and geocoded units

4. Geocoding process

International Mandate

UN Economic and Social Council (ECOSOC)

UN Statistical Commission (UNSC)

- Global Statistical-Geospatial program review – proposed a global framework

UN Committee of Experts on Global Geospatial Information Management (UN-GGIM)

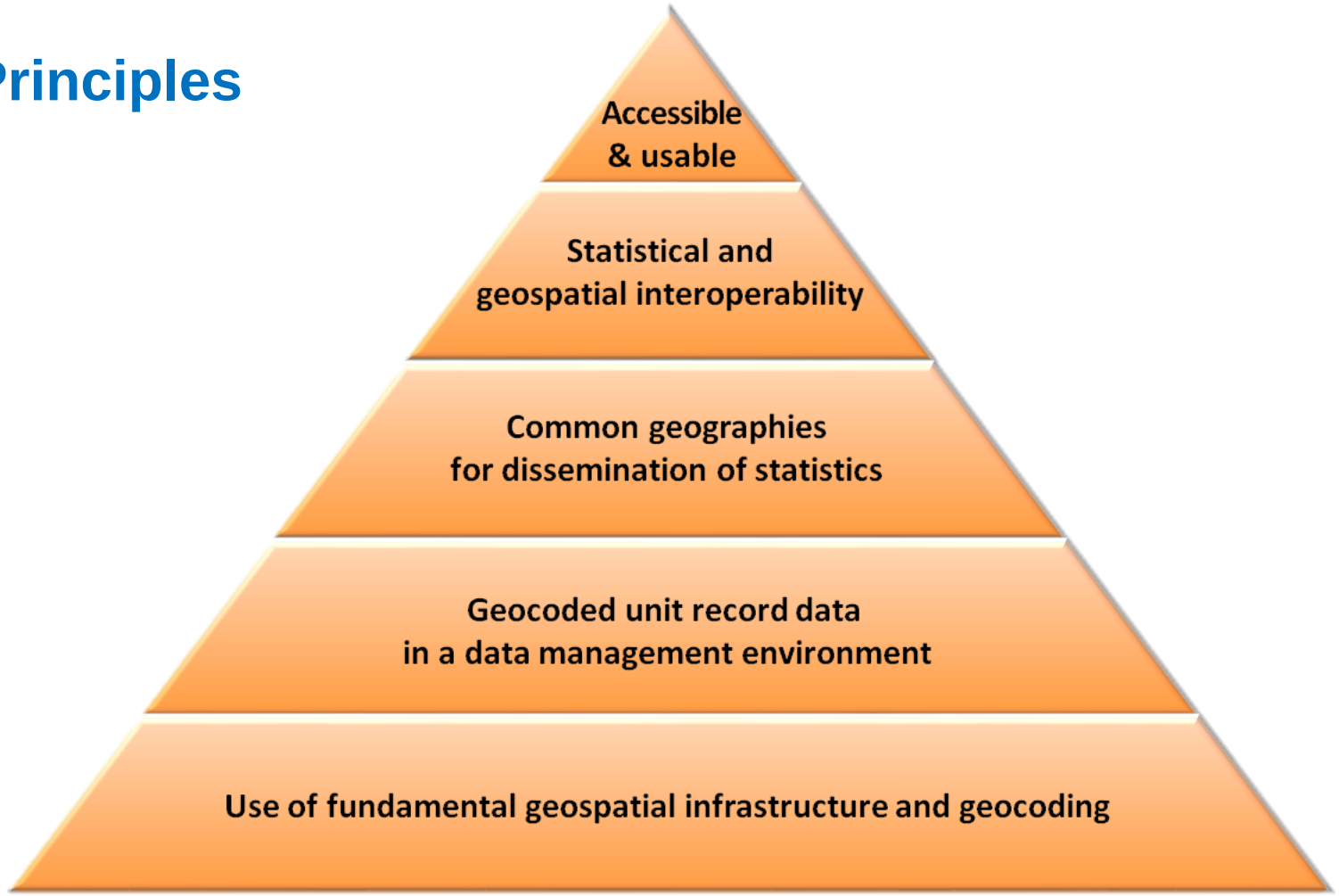
- List of nine issues included 'linking of spatial to statistics'

UN Expert Group – Integration of Statistical Geospatial Information

…. both communities

Global Statistical Geospatial Framework

5 Principles



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

Statistical Commission

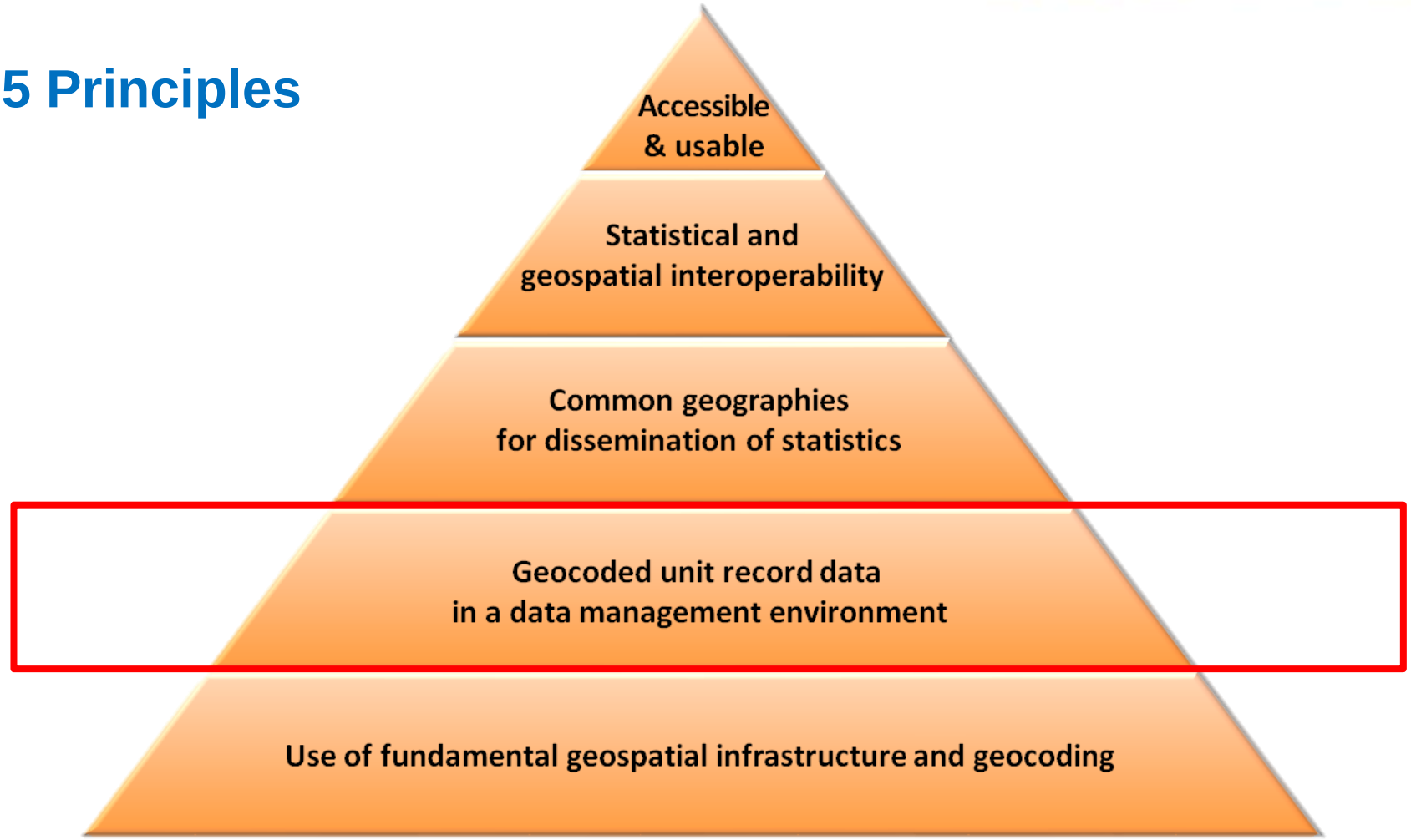
BETTER DATA
BETTER LIVES

unstats.un.org

@UNStats #UNStatCom

Global Statistical Geospatial Framework

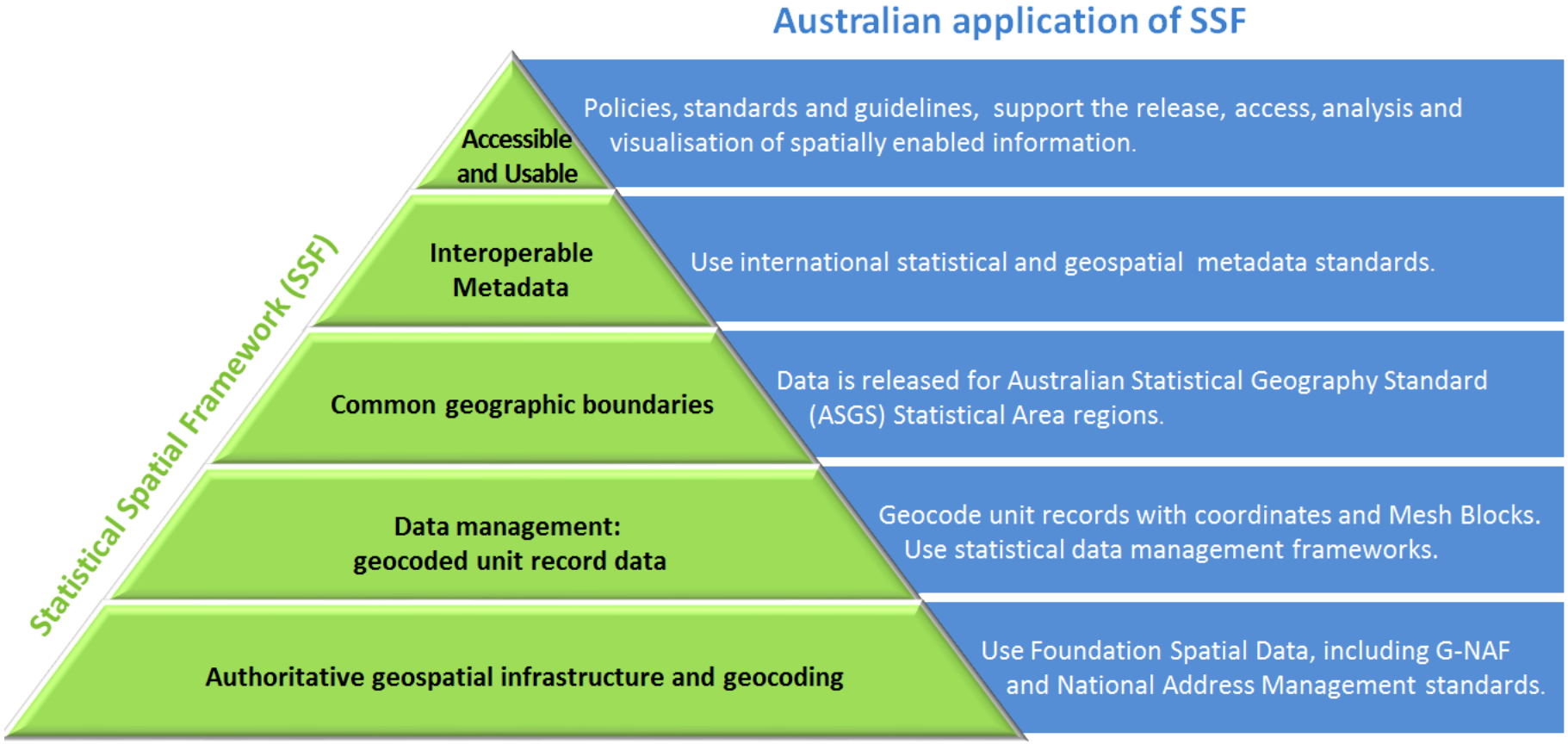
5 Principles



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

BETTER DATA BETTER LIVES
unstats.un.org
@UNStats #UNStatCom

Australian SSF



GSGF Purpose

"The Global Statistical Geospatial Framework will provide:

- a common method for geospatially enabling statistical and administrative data,
- ensure that this data can be integrated with geospatial information."

Proposal for a Global Statistical Geospatial Framework, UN-GGIM 6, 2016 New York

Bridging between two communities

Statistical Community

Socio-Economic Datasets

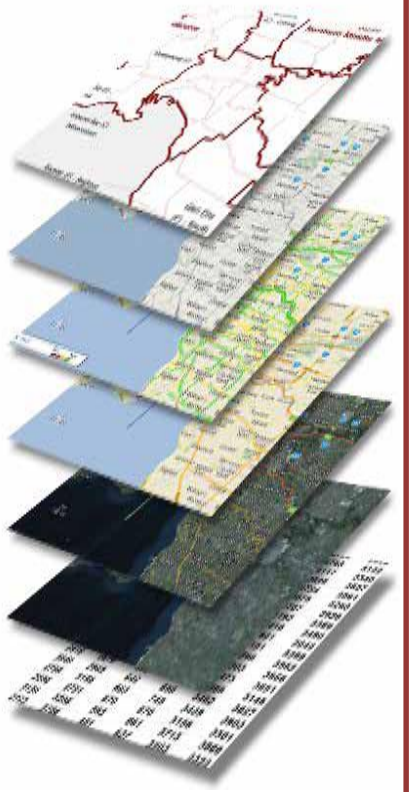
- **Core Statistical** — Census, Demographics, Agriculture, Building, Labour Force, etc.
- **Tax** — Income and Business Tax
- **Immigration**
- **Health** — Medicare, Pharmaceuticals, Workforce
- **Land** — Valuation and Use
- **Social Welfare** — Unemployment, Disability, Family Support
- **Others …**

GSGF bridge

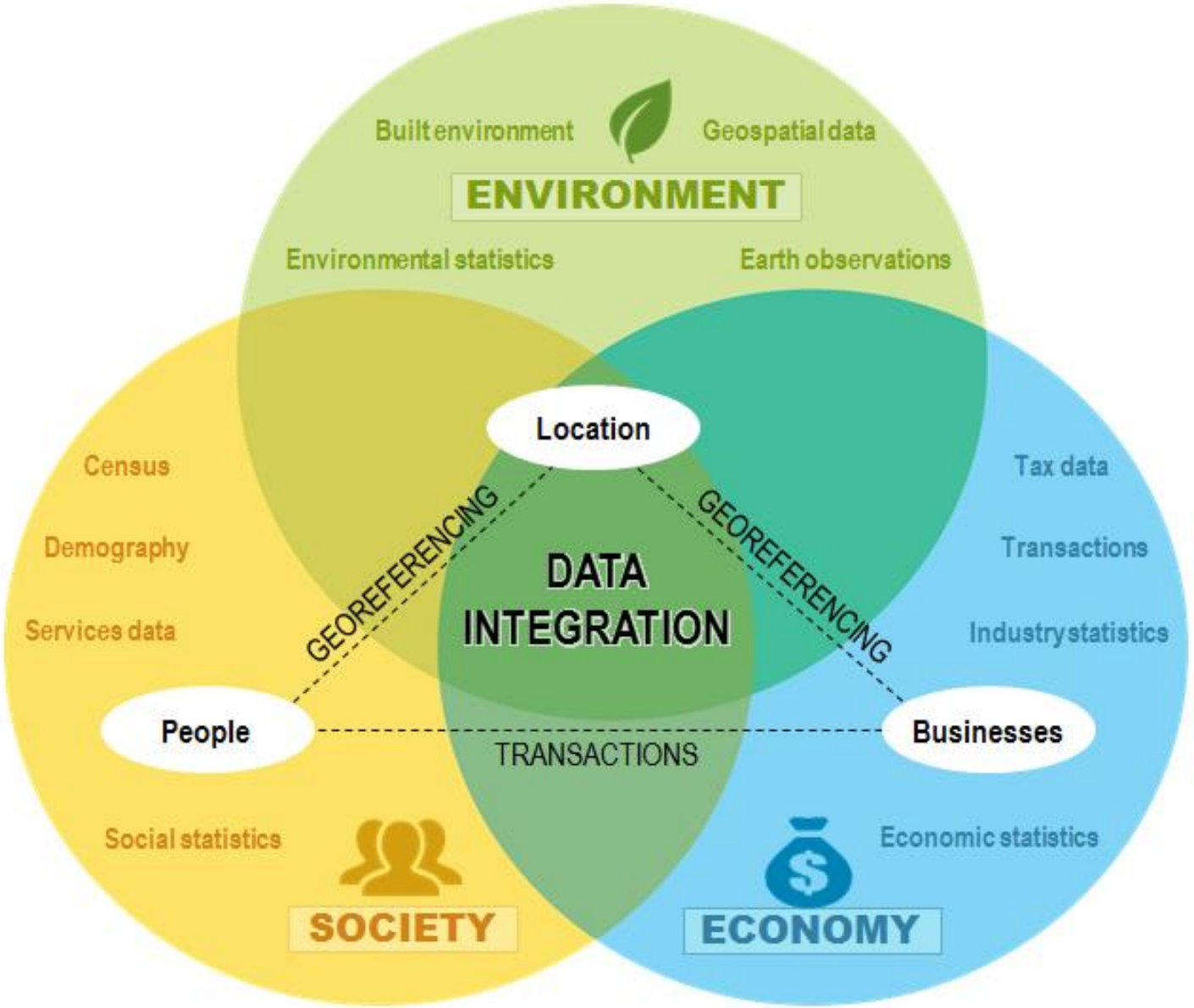
Spatial Community

Fundamental Geospatial Datasets

- Admin. & statistical boundaries
- Addressing, Place Names
- Transport, Water
- Land and Property
- Elevation and Depth
- Imagery
- Positioning



Location – bridging the 3 domains



GSGF – Status paper

Global Statistical Geospatial Framework: Linking Statistics and Place

Current status and plans for development July 2018

United Nations Expert Group on the

Integration of Statistical and Geospatial Information

EG-ISGI webpage

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

BETTER DATA BETTER LIVES

unstats.un.org

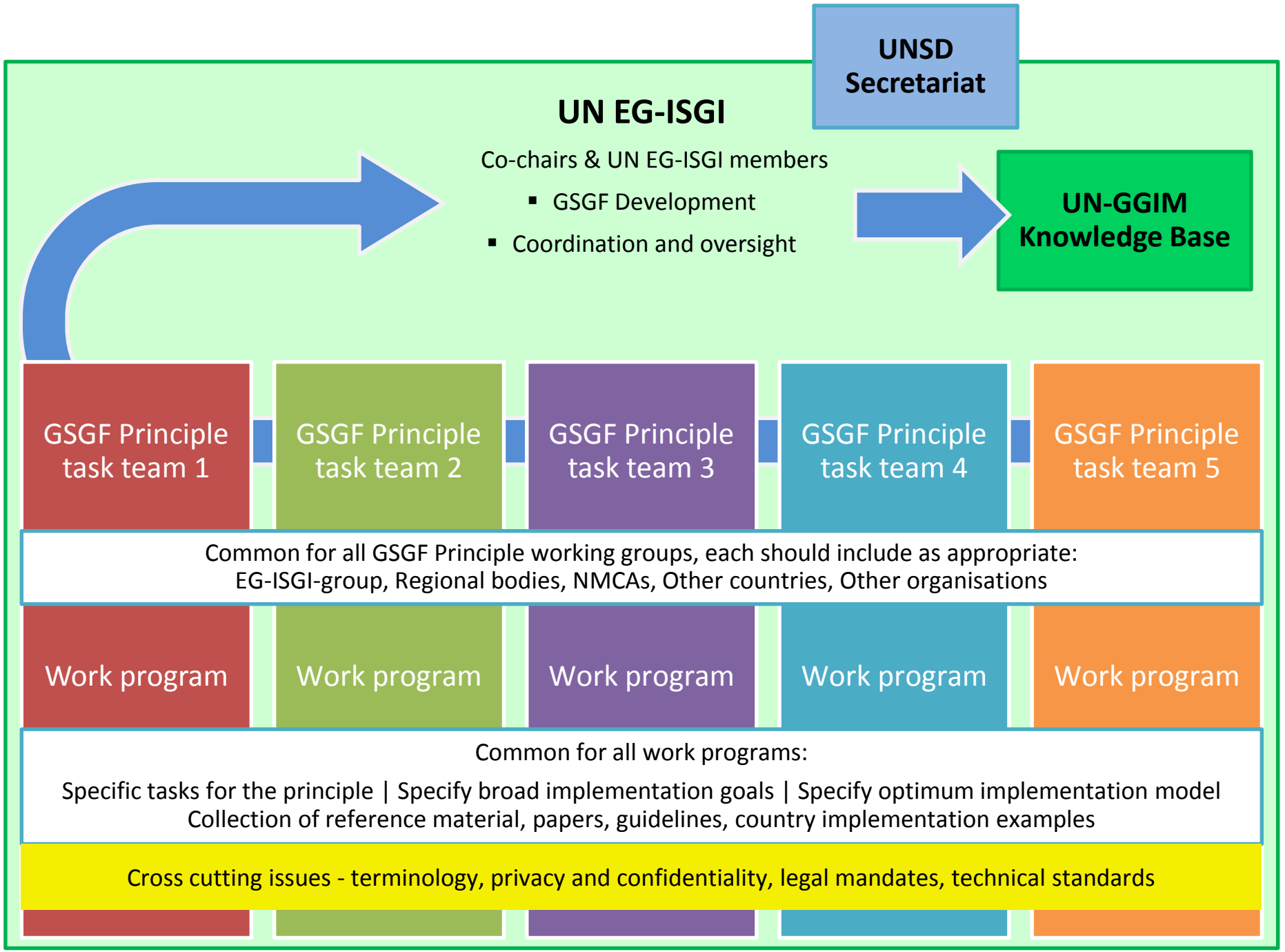
@UNStats #UNStatCom

GSGF Principle Working Groups

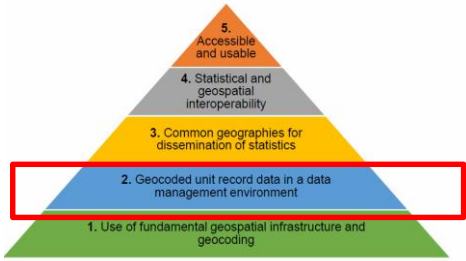
GSGF Principle Task Teams

Lead countries - GSGF principles

GSGF Principle	Lead countries/orgs
Principle 1: Use of fundamental geospatial infrastructure and geocoding	Mexico / Germany
Principle 2: Geocoded unit record data in a data management environment	Australia / Sweden
Principle 3: Common geographies for dissemination of statistics	Poland / Candada
Principle 4: Statistical and geospatial interoperability	Eurostat / UNECE
Principle 5: Accessible and usable geospatially enabled statistics	UK / USA


UNSD Secretariat
UN EG-ISGI
Co-chairs & UN EG-ISGI members
GSGF Development
Coordination and oversight
UN-GGIM Knowledge Base
GSGF Principle task team 1
GSGF Principle task team 2
GSGF Principle task team 3
GSGF Principle task team 4
GSGF Principle task team 5
Common for all GSGF Principle working groups, each should include as appropriate:
EG-ISGI-group, Regional bodies, NMCAs, Other countries, Other organisations
Work program
Work program
Work program
Work program
Work program
Common for all work programs:
Specific tasks for the principle | Specify broad implementation goals | Specify optimum implementation model
Collection of reference material, papers, guidelines, country implementation examples
Cross cutting issues - terminology, privacy and confidentiality, legal mandates, technical standards

Principle 2 – Geocoding in a data management environment



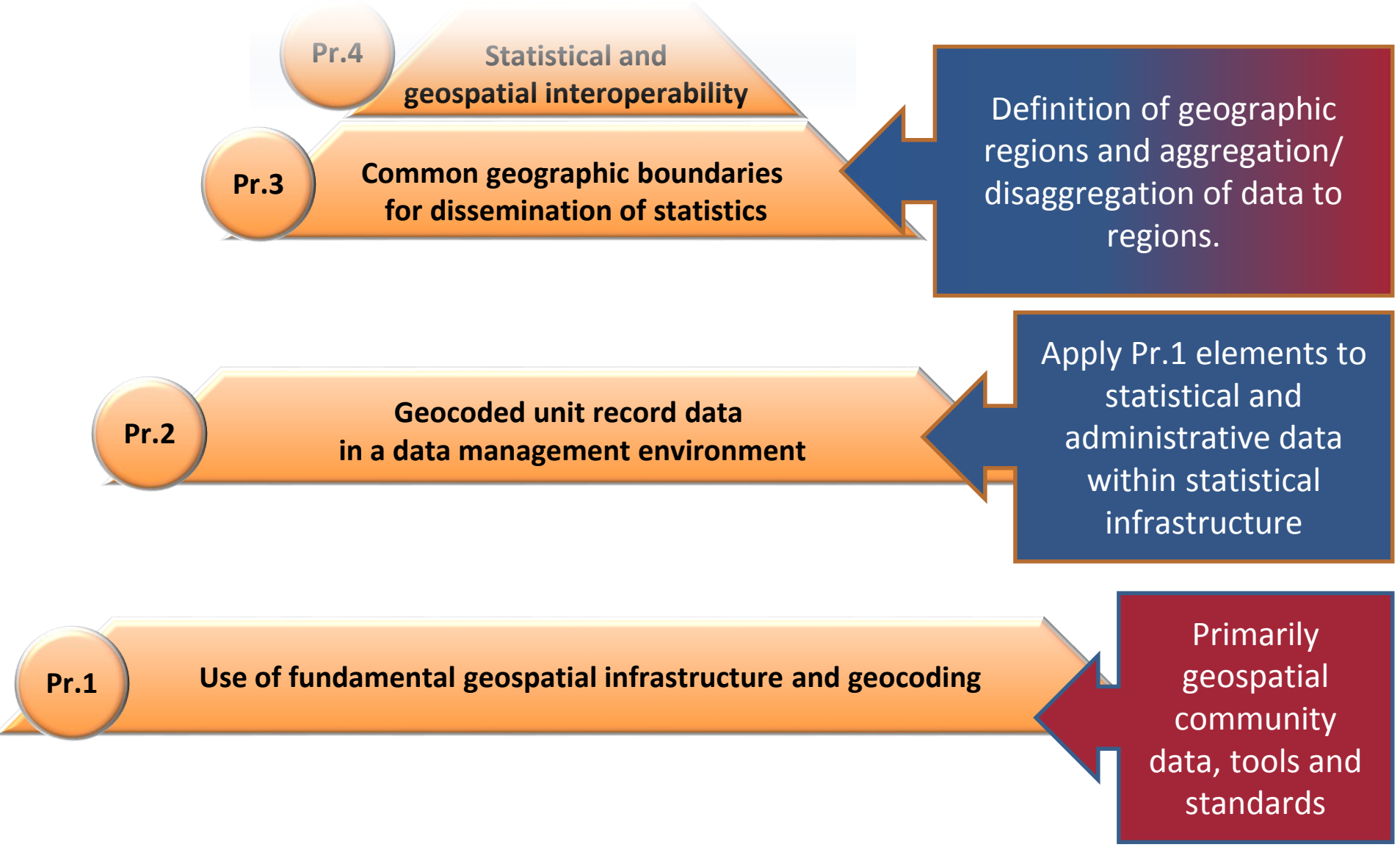
- Build an effective and secure data management environment
- Store location only once
- Ensure consistency and quality of geocoding results
- Develop consistent approaches to manage non-matching data
- Use point-of-entry validation in collection of administrative or statistical data

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



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

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

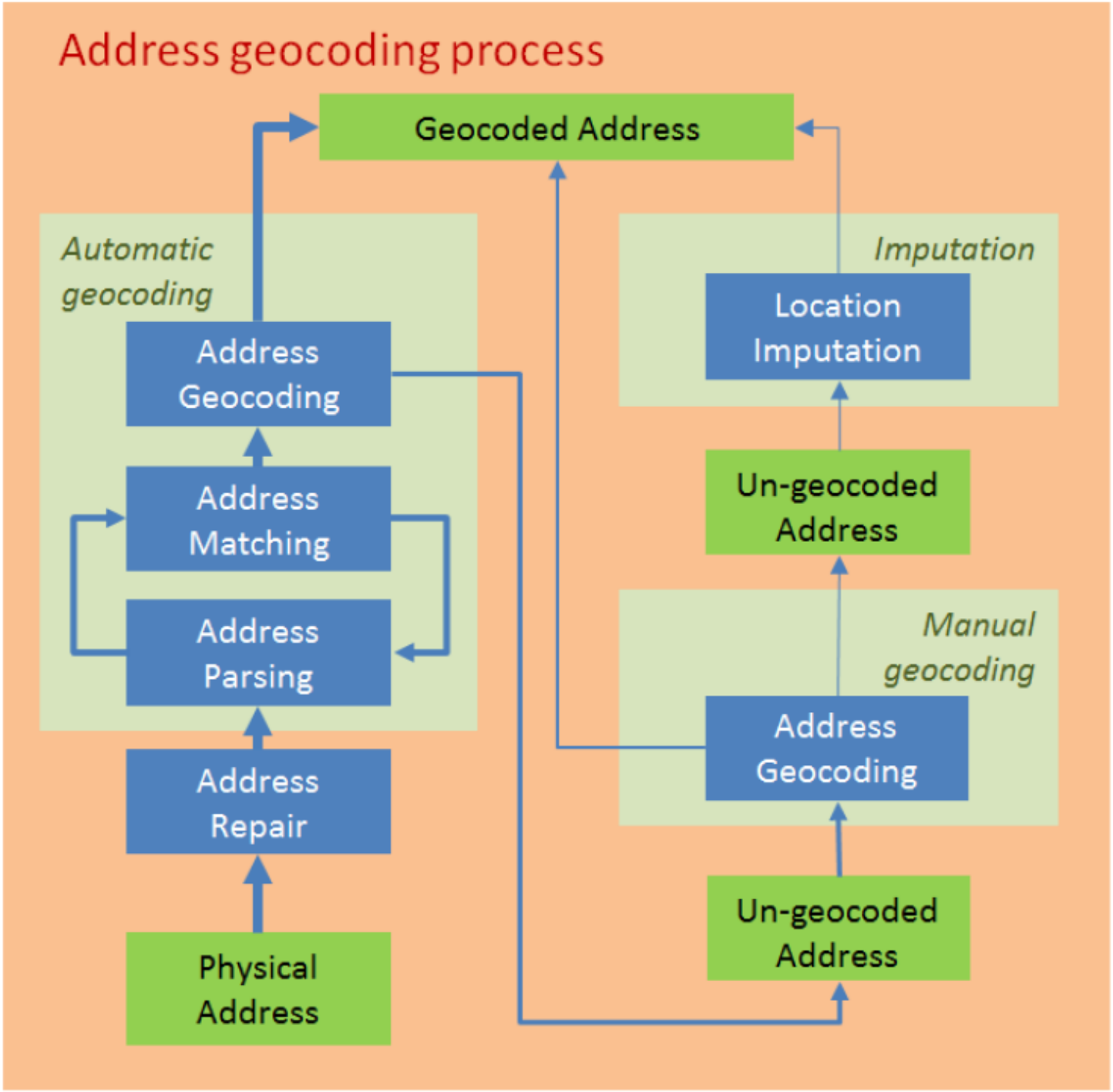
Statistical Community

- 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


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