

# UNECE Statistical Division – INGEST Project on Integration of Statistical and Geospatial data & **GeoGSBPM**

*Taeke Gjaltema (inter-regional adviser Statistics, UNECE)  
Sara Stewart (UNECE Consultant)*

UN-GGIM: Europe Webinar on “United Nations Geospatial  
Network Data Hub: One UN Geospatial Situation Room’  
*15 November 2023 online*



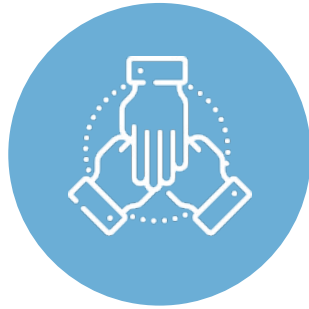
Funded by  
the European Union

# PROJECT OVERVIEW

UNECE is leading an EU-funded project to develop statistical capacity in geospatial and statistical data integration



Foster stronger links between statistical and geospatial communities across the UNECE region



Facilitate greater collaboration amongst statistical and geospatial organisations



Encourage greater integration of geospatial and statistical information



Promote stronger institutional partnerships and the use of common standards

# UNECE SURVEY INTEGRATION OF STATISTICAL AND GEOSPATIAL INFORMATION

- Obtain insight from the statistical and geospatial communities.
- Ensure the views of key stakeholders were included in the project.
- Acquire user feedback to inform the direction of project activities.

The survey was designed around 4 key areas:

1. Use of data and technology.
2. Involvement in wider activities at national and international levels.
3. Issues and obstacles to greater data integration.
4. Opportunity to register interest in future project activities.

Some recommendations:

- Identify and promote sustainable funding resources and models to support data integration activities at national levels.
- Enhance communication and engagement strategies to grow awareness of the benefits of data integration and better support the sharing of best practice and new technologies.
- Promote greater data standardisation and interoperability through harmonised standards, operating models, production processes and services.



## UNECE Survey on the Integration of Statistical and Geospatial Information

Spring 2023

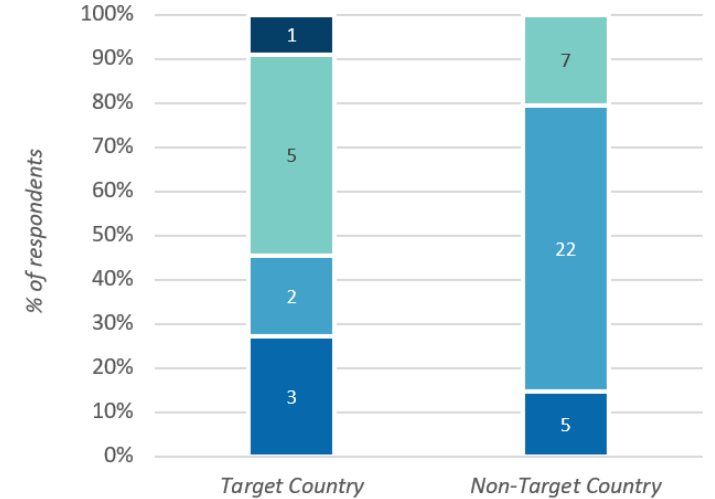
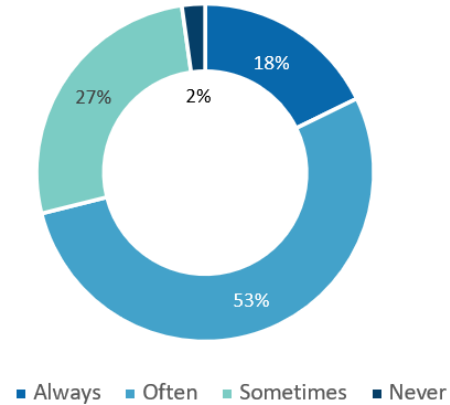
This survey will inform a European Union funded project which seeks to strengthen the integration of statistical and geospatial information across the UNECE region. Your feedback will help to set the direction of future activities relating to statistical and geospatial data integration.

# DATA AND TECHNOLOGY

## Common NSI uses of geospatial data:

- Census operations
- Geocoding
- Spatial analysis
- Dissemination (e.g. grid statistics)

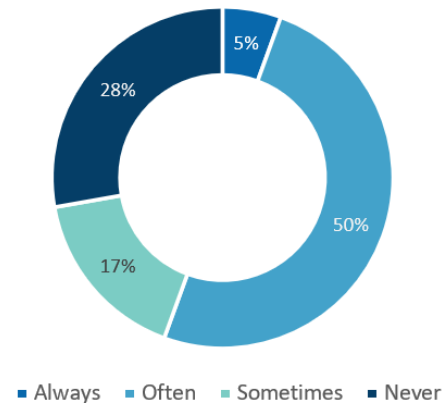
If you are an NSI, how often do you use geospatial data within your workflows?



## Common NMCA uses of statistical data:

- Data production and management
- Thematic map production
- Spatial analysis using demographic data

If you are a NMCA, how often do you use statistical data within your workflows?



# COLLABORATION AND PARTNERSHIPS

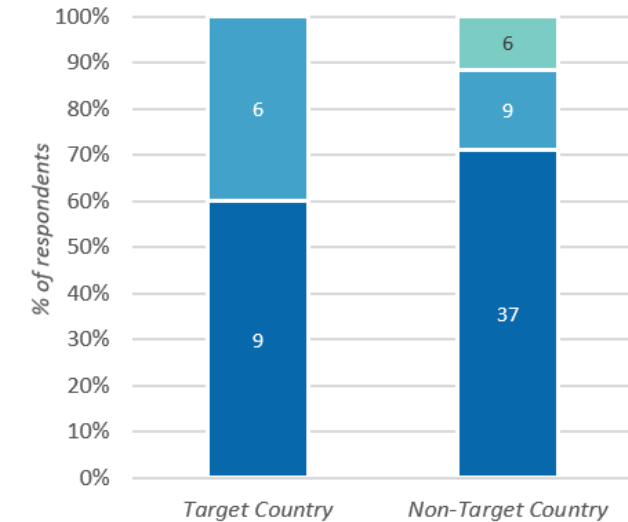
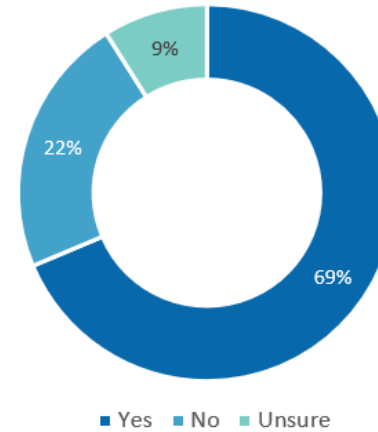
## National working group activities:

- INSPIRE Directive
- Land use classification
- Data validation
- Spearheading use of geospatial data

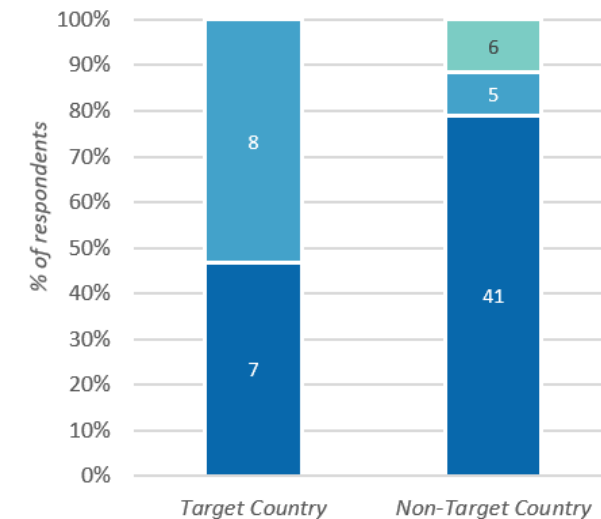
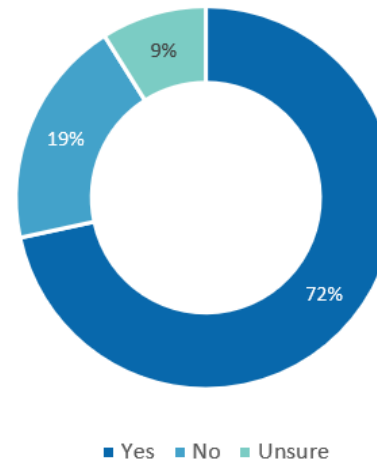
## Regional/International working groups:

- GISCO Working Group
- UN EG-ISGI
- UN-GGIM: Europe LoW on Data Integration
- EFGS

Do you participate in any national working groups with your statistical or geospatial counterpart?



Do you participate in any regional/international working groups relating to statistical or geospatial data?



# JOINT UNECE / EUROSTAT / UN-GGIM: EUROPE WORKSHOP ON INTEGRATING STATISTICAL AND GEOSPATIAL DATA



- UNECE in collaboration with Eurostat, UN-GGIM: Europe and our hosts, the Republic Geodetic Authority of the Republic of Serbia (4-5 October 2023, Belgrade, Serbia).
- Feedback from the UNECE Survey has guided the design and structure of the workshop.
- Innovative and best practices, as well as lessons learned, on the integration of statistical and geospatial data.
  - Effective Governance
  - Data and Technology
  - People and Partnerships
- Discussions: how to optimise current governance and policy frameworks, identify the skills needed to progress data integration activities, determine the best ways to communicate and engage with key stakeholders in the field, and identify the key building blocks needed to form successful and sustainable partnerships both now and into the future

Workshop wiki pages: <https://statswiki.unece.org/x/7AHnFw>



# WORKSHOP'S KEY TAKEAWAYS

- Sharing experiences, best practices and connecting statistical and geospatial communities provided inspiration for future work and further collaboration.
- Frameworks are an important means of bringing different datasets together to solve common challenges.
- Building trust in data is key and AI can help to harness the benefits of data and its integration.
- Stronger integrative framework needs to be developed.



- Shared understanding of common challenges at every level of organisation
- Benefits of integration data need to be more effectively communicated with decision-makers.
- Participation of EECCA countries in data integration activities will ensure that the benefits are equally achieved across the region.

Workshop wiki pages: <https://statswiki.unece.org/x/7AHnFw>



# TASK FORCE ON STANDARDS ISSUES

- Aim to support the greater harmonization and interoperability of statistical and geospatial information through the use of common standards.
  - Discuss the use of standards
  - Explore issues and constraints
  - Identify priorities and future actions

## **Activities (sub-groups on technical and on wider governance issues):**

- a) Understand the current use of standards and share use cases and best practice.
- b) Determine the suitability of those standards to data integration activities and identify domains where data integration is hampered by a lack of common standards.
- c) Identify priorities for standards harmonization work and recommend related actions.

## **Output:**

- A report containing recommended actions and methodological guidelines to support their implementation at a national level.





# WIKI SPACE



Integration of  
Geospatial and  
Statistical Data

Pages

Blog

PAGE TREE

- INGEST Blog Series
- ▼ Joint UNECE, Eurostat, U
  - 9th Joint UN-GGIM: Eu
  - GeoStat 2023 - Timeta
- Presentations



<https://statswiki.unece.org>

## Integration of Geospatial and Statistical Data Home

Welcome to our space on the integration of geospatial and statistical data

Data is a critical resource for evidence-based decision-making and the 2030 Agenda for Sustainable Development has brought the need for harmonised data of increasing quality, accuracy, currency and granularity to the forefront of global, regional and national agendas to support the measurement and monitoring of its SDGs. The integration of geospatial and statistical data is one of the most promising ways to maximise the value of data-driven decision-making and global efforts to progress data integration have been ongoing for a decade, centred upon the work of the United Nations, its various bodies and expert groups, and increasingly extended and adapted to regional and national contexts through the work of UNECE, Eurostat, UN-GGIM: Europe, EFGS and others. The benefits of integrating geospatial and statistical data are great: enhancing the value of the geospatial and statistical data itself through improved quality, improving the interoperability of datasets, providing new possibilities for analysis and presentation, and resulting in greater insights which can inform decision-making and policy development across multiple levels.

This space is designed to bring together information on upcoming workshops, links to key groups and resources relating to data integration, as well as the home of a new blog series on integrating geospatial and statistical data, INGEST, which has recently been launched. Happy exploring!

### Workshops

Workshops on integrating geospatial and statistical data	When and Where
<a href="#">Joint UNECE, Eurostat, UN-GGIM Europe Workshop on Integrating Statistical and Geospatial Data</a>	4-5 October 2023, Belgrade, Serbia
<a href="#">9th Joint UN-GGIM: Europe - ESS - UNECE Meeting, 3 October</a>	3 October 2023, Belgrade, Serbia
<a href="#">EFGS 2022 Conference - Towards standardised geospatial statistics</a>	27-28 October 2022, online
<a href="#">8th Joint UN-GGIM: Europe – UNECE Meeting on the Integration of Statistical and Geospatial Information</a>	24 March 2022, online
<a href="#">EFGS 2021 Conference - Unlocking value from data</a>	7-8 September 2021, online
<a href="#">Joint Workshop on the Integration between Geospatial and Statistical Information in International Organisations and Member States</a>	28 April - 12 May 2021, online
<a href="#">7th Joint UN-GGIM: Europe – ESS – UNECE Meeting on the Integration of Statistical and Geospatial Information</a>	24 March 2021, online
<a href="#">EFGS 2020 Webinar</a>	20-21 October 2020, online

<https://statswiki.unece.org/x/3wDnFw>



UNECE

eurostat



UN-GGIM  
UNITED NATIONS  
COMMITTEE OF EXPERTS ON  
GLOBAL GEOSPATIAL  
INFORMATION MANAGEMENT

### Quick Links

[INGEST Blog Series](#)

#### Key Organisations:

[UNECE Statistics Division](#)

[Eurostat](#)

[UN-GGIM: Europe](#)

[UN-GGIM \(Global\)](#)

[European Forum for Geography and Statistics](#)

#### Key Resources:

[The Global Statistical Geospatial Framework \(GSGF\)](#)

[The United Nations Integrated Geospatial Information Framework \(UN-IGIF\)](#)

# BLOG SERIES

6 blog posts published so far including:

1. Welcome to Ingest, our new blog series on integrating geospatial and statistical data
2. Data Integration: Key players and recent developments
3. UNECE Survey - Part 1: Background and Context
4. UNECE Survey - Part 2: Data and Technology
5. UNECE Survey - Part 3: Collaboration and Partnerships
6. UNECE Survey - Part 4: Issues and Obstacles
7. Low-down on the recent Joint UNECE / Eurostat / UN-GGIM: Europe Workshop

Blog / 2023 / July / 03

## Welcome to INGEST, our new blog series on integrating geospatial and statistical data

Created by Sara Stewart, last modified on 24 Jul, 2023



Do you want to learn more about integrating geospatial and statistical data but aren't sure where to start? Are you short on time and need bite-sized information on important developments and trends in the field of data integration? Well, you have come to the right place!

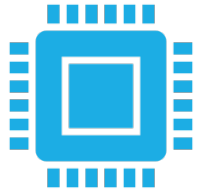
Welcome to our exciting new blog series, INGEST (which, in case you haven't noticed, is short for INtegrating GEospatial and SStatistical data) where we will be exploring a range of topics relating to data integration as well as sharing guidance, resources and best practice to help you on your journey towards greater data integration. Whether you are just beginning this journey of discovery, or you are a seasoned pro who is hoping to broaden your knowledge and keep up to date with developments in the field of data integration, we hope there will be something for everyone. We'd also love to hear your thoughts on the topics that we explore so please feel free to comment on the posts and we can connect you to the wider community and stimulate discussion.

In this introductory post to the INGEST Blog Series, we will start right at the beginning and explore what integrating geospatial and statistical data is all about and discuss why it is important. So are you ready to dive in? Let's go!

<https://statswiki.unece.org/x/OwLnFw>



# FUTURE ACTIVITIES



Task Force on Standards Issues relating to the Integration of Statistical and Geospatial Information



Report containing recommended actions and guidelines produced by the Task Force.



Further developing the wiki space with supportive material and resources.



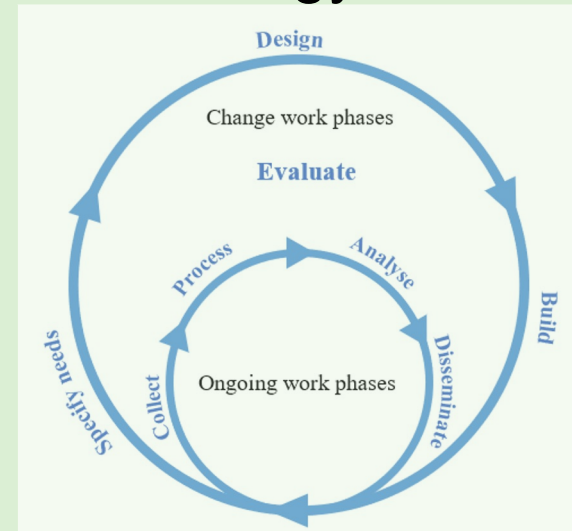
Producing regular blog posts for INGEST



Providing online consultations, training workshops or similar events.

# Generic Statistical Business Model: GSBPM

- International Standard or Model describing statistical business processes in a general and process-oriented way (exploration, new and ongoing)
- Independent of data source: survey, census, administrative registers, business portals, Big Data etc.. & All Statistical domains
- Usages:
  - Common reference framework & language/standardized process terminology
  - Process based organization structures
  - Benchmarking within and between statistical organizations
  - Efficiency (identifying synergies between processes)
  - Cost estimation and evaluation tool
  - Process Quality management/frameworks
  - Methodology framework
  - Software sharing
  - Develop/align systems and business/enterprise architectures (SOA)



Overarching Processes

Level 1: 8 Phases

Level 2:  
44 Sub-processes

Level 3:  
GSBPM  
Tasks

Overarching Processes							
Specify needs	Design	Build	Collect	Process	Analyse	Disseminate	Evaluate
1.1 Identify needs	2.1 Design outputs	3.1 Reuse or build collection instruments	4.1 Create frame and select sample	5.1 Integrate data	6.1 Prepare draft outputs	7.1 Update output systems	8.1 Gather evaluation inputs
1.2 Consult and confirm needs	2.2 Design variable descriptions	3.2 Reuse or build processing and analysis components	4.2 Set up collection	5.2 Classify and code	6.2 Validate outputs	7.2 Produce dissemination products	8.2 Conduct evaluation
1.3 Establish output objectives	2.3 Design collection	3.3 Reuse or build dissemination components	4.3 Run collection	5.3 Review and validate	6.3 Interpret and explain outputs	7.3 Manage release of dissemination products	8.3 Agree an action plan
1.4 Identify concepts	2.4 Design frame and sample	3.4 Configure workflows	4.4 Finalise collection	5.4 Edit and impute	6.4 Apply disclosure control	7.4 Promote dissemination products	
1.5 Check data availability	2.5 Design processing and analysis	3.5 Test production systems		5.5 Derive new variables and units	6.5 Finalise outputs	7.5 Manage user support	
1.6 Prepare and submit business case	2.6 Design production systems and workflow	3.6 Test statistical business process		5.6 Calculate weights			
		3.7 Finalise production systems		5.7 Calculate aggregates			
				5.8 Finalise data files			

National Adaptations/Extensions

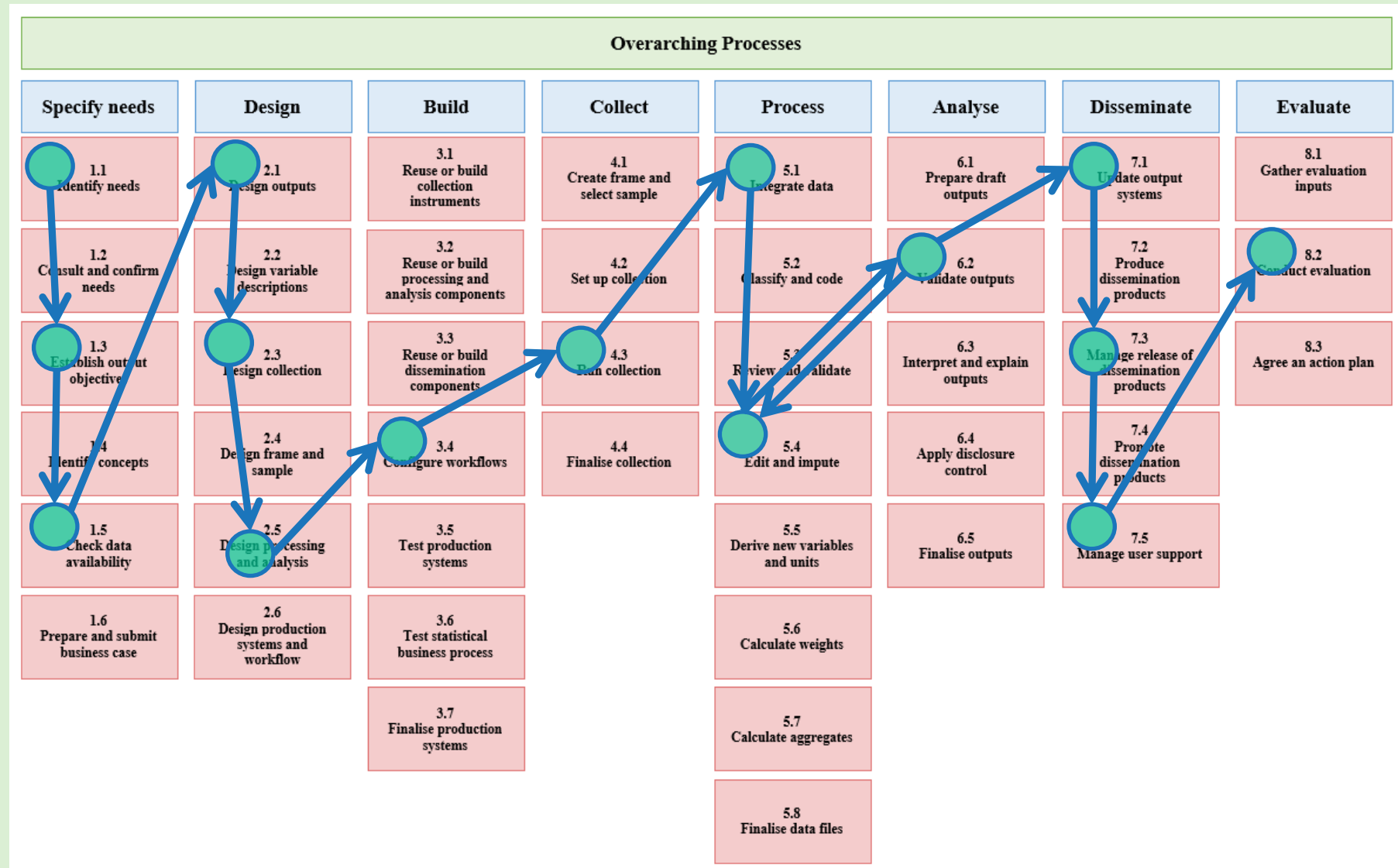


Overarching Processes

Level 1: 8 Phases

Level 2:  
44 Sub-processes

Level 3:  
GSBPM  
Tasks



National Adaptations/Extensions



# GSBPM and GSGF

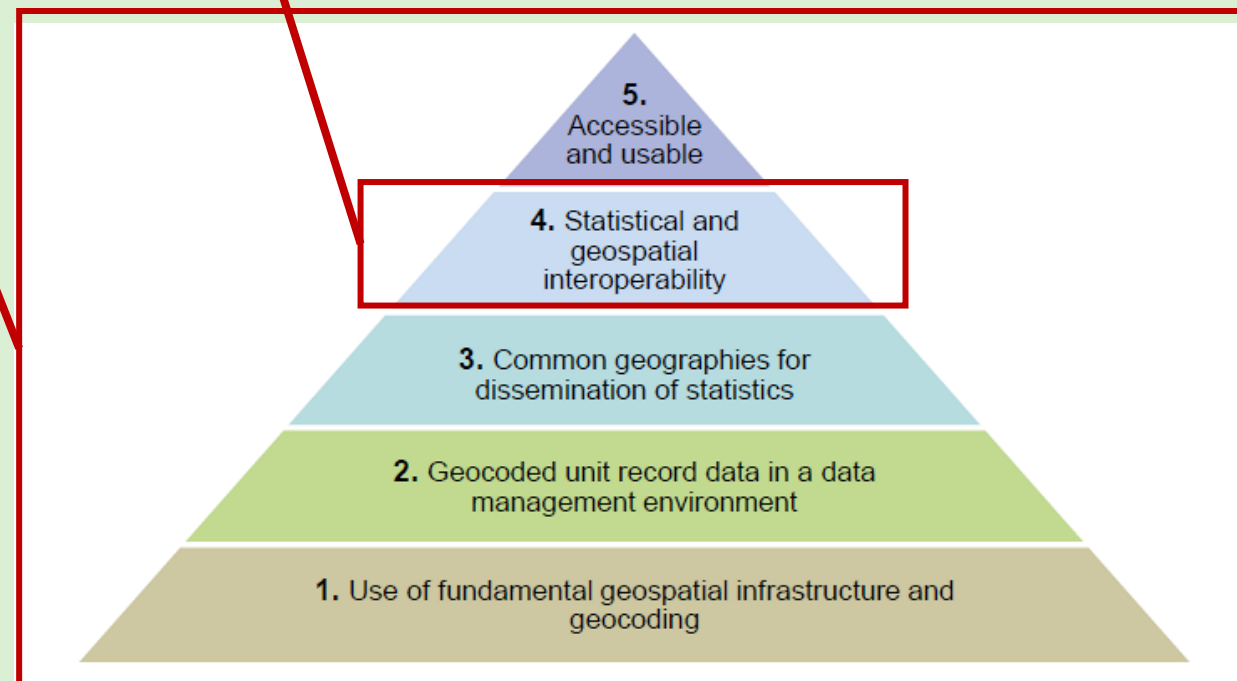
GSBPM as tool to ensure GSGF principles to be followed

## GSBPM

Overarching Processes							
Specify needs	Design	Build	Collect	Process	Analyse	Disseminate	Evaluate
1.1 Identify needs	2.1 Design outputs	3.1 Reuse or build collection instruments	4.1 Create frame and select sample	5.1 Integrate data	6.1 Prepare draft outputs	7.1 Update output systems	8.1 Gather evaluation inputs
1.2 Consult and confirm needs	2.2 Design variable descriptions	3.2 Reuse or build processing and analysis components	4.2 Set up collection	5.2 Classify and code	6.2 Validate outputs	7.2 Produce dissemination products	8.2 Conduct evaluation
1.3 Establish output objectives	2.3 Design collection	3.3 Reuse or build dissemination components	4.3 Run collection	5.3 Review and validate	6.3 Interpret and explain outputs	7.3 Manage release of dissemination products	8.3 Agree an action plan
1.4 Identify concepts	2.4 Design frame and sample	3.4 Configure workflows	4.4 Finalise collection	5.4 Edit and impute	6.4 Apply disclosure control	7.4 Promote dissemination products	
1.5 Check data availability	2.5 Design processing and analysis	3.5 Test production systems		5.5 Derive new variables and units	6.5 Finalise outputs	7.5 Manage user support	
1.6 Prepare and submit business case	2.6 Design production systems and workflow	3.6 Test statistical business process		5.6 Calculate weights			
		3.7 Finalise production systems		5.7 Calculate aggregates			
				5.8 Finalise data files			

Immediate connection

## GSGF

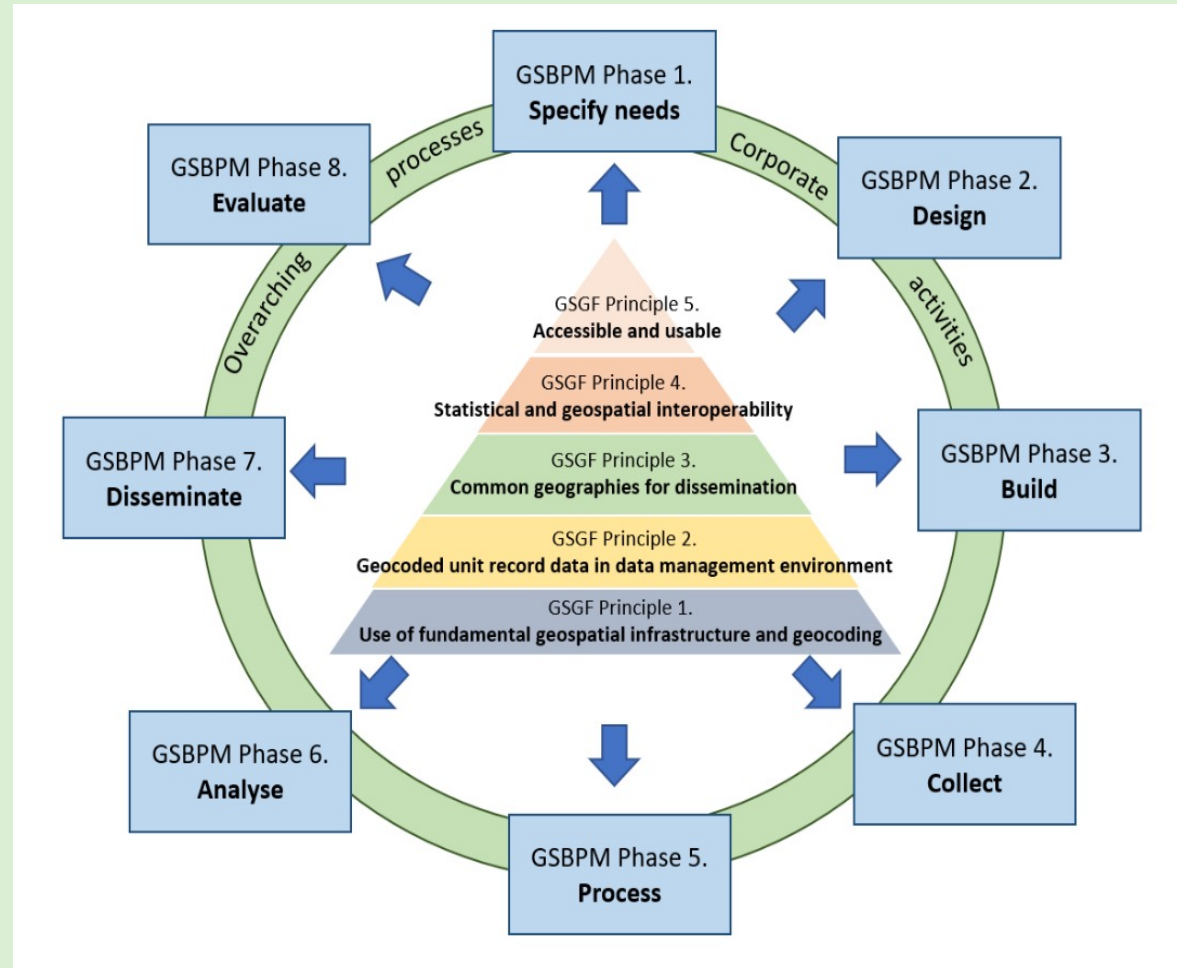


SOURCE: Australian Bureau of Statistics (ABS) / UN-GGIM, illustration by Statistics Sweden

Figure 2: The Global Statistical Geospatial Framework (GSGF)

# Geospatial view of GSBPM (GeoGSBPM)

GeoGSBPM describes geospatial-related activities and considerations using the framework of the GSBPM



# Geospatial view of GSBPM (GeoGSBPM)

- Developed by Geospatial task team of HLG-MOS Supporting Standards Group
- GeoGSBPM describes geospatial-related activities and considerations using the framework of the GSBPM

(Example of GSBPM sub-process 2.2 Design variable description)

## **2.2 Design variable description**

28. This sub-process defines the variables to be collected via the collection instrument, as well as any other variables that will be derived from them in sub-process 5.5 (Derive new variables and units), and any statistical or geospatial-classifications that will be used. It is expected that existing national and international standards will be followed wherever possible.
29. This sub-process may need to run in parallel with sub-process 2.3 (Design collection), as the definition of the variables to be collected, and the choice of collection instruments may be inter-dependent to some degree. Preparation of metadata descriptions of collected and derived variables, statistical and geospatial classification is a necessary precondition for subsequent phases.

# Geospatial view of GSBPM (GeoGSBPM)

- Developed by Geospatial task team of HLG-MOS Supporting Standards Group
- GeoGSBPM describes geospatial-related activities and considerations using the framework of the GSBPM

(Example of GSBPM sub-process 2.2 Design variable description)

**2.2 Design variable description**  
28. This sub-process defines the variables to be collected via the collection instrument, as well as any other variables that will be derived from them in sub-process 5.5 (Derive new variables and units), and any statistical or geospatial-classifications that will be used. It is expected that existing national and international standards will be followed wherever possible

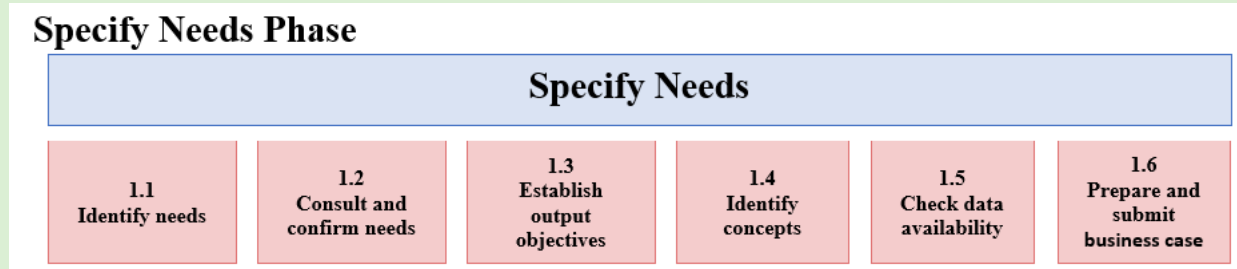
29. Geospatial variables (geographies) that are used while collecting data at a statistical unit level are not usually the same as those that are used for dissemination. Hence, they should be designed at the statistical unit level using point-based location<sup>8</sup> as the base geospatial variable, as it will provide a considerable adaptability to changes over time and flexibility to aggregate up to various dissemination-level geographies. For gridded geographies, it is important to use a grid system that is comparable with the existing regional or global grid system (e.g. Discrete Global Grid System (DGGS)<sup>9</sup>) as it will greatly increase usability of the output. Different types of grid (e.g. hexagon, rectangular) and their advantages and disadvantages can be assessed when designing gridded geographies

30. This sub-process may need to run in parallel with sub-process 2.3 (Design collection), as the definition of the variables to be collected, and the choice of collection instruments may be inter-dependent to some degree. Preparation of metadata descriptions of collected and derived variables, statistical and geospatial classification is a necessary precondition for subsequent phases.

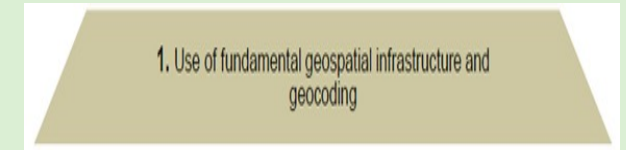
**GSBPM original text**

**Geospatial-related consideration and activities**

# GeoGSBPM – Phase 1. Specify Needs

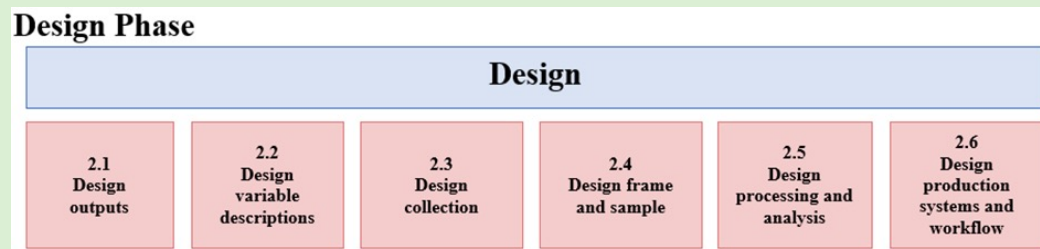


GSGF 1.



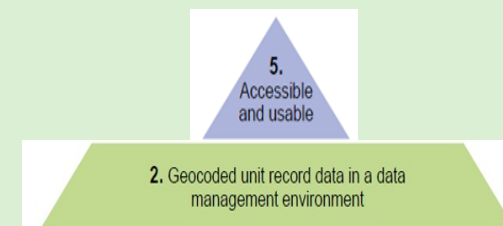
- **Type of geography (e.g. grid, administrative boundary) and size of geography unit can have a significant implication on the cost of production, reliability and the risk of privacy breaches**, they have to be carefully examined from the very beginning of the production process
- **When assessing data availability, the existence and availability of suitable geospatial information should be first identified from authoritative sources with the National Spatial Data Infrastructure (NSDI)**
- Geospatial information may exist at the point-based level (e.g. x-y coordinates) or at the coarse area-based level (e.g. administrative boundary) and this should be compared with geospatial requirements that users specified (e.g. geography type, size, date of reference of the data, availability of time series).

# GeoGSBPM – Phase 2. Design



GSGF 5.

GSGF 2.



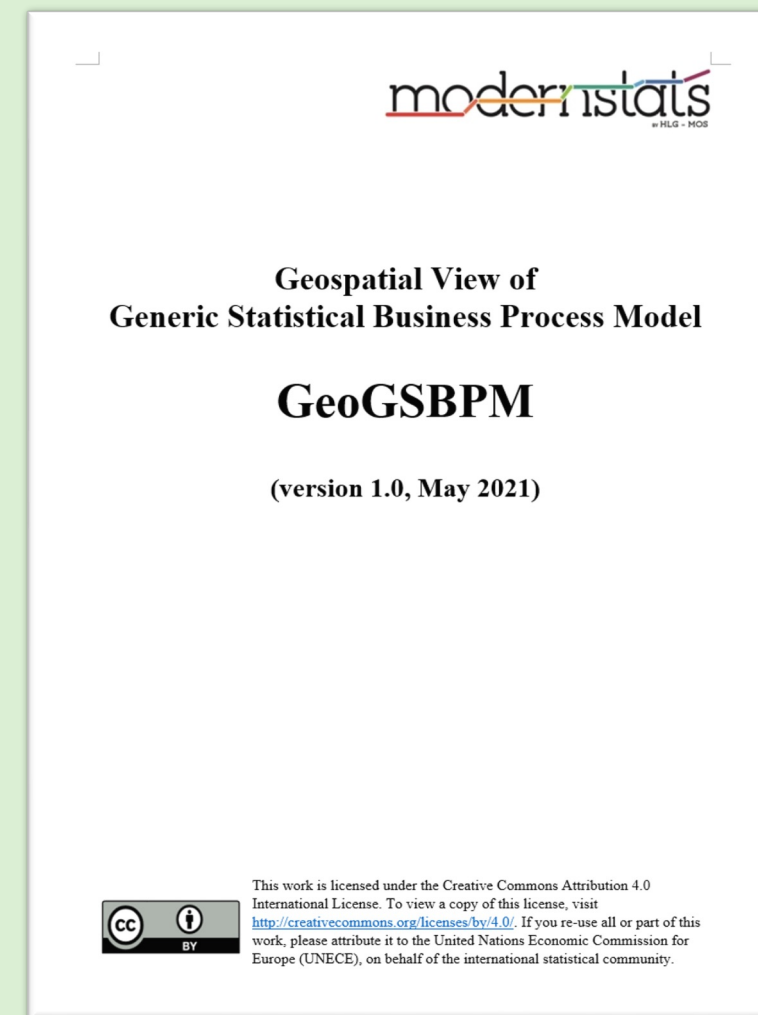
- Geographic viewing tools are a powerful way to help users to understand the geographic context of the issues that they are trying to solve with the data. Therefore, **it is recommended to include spatial visualisation and GIS services components in the output ... Accessibility and usability of geospatially enabled statistics and services can greatly increase by the use of standards and open data formats (e.g. XML, GeoJSON).**
- Geospatial variables should be designed at the statistical unit level using **point-based location as the base geospatial variable, as it will provide a considerable adaptability to changes over time and flexibility to aggregate up to various dissemination-level geographies.** For gridded geographies, it is important to use a grid system that is comparable with the existing regional or global grid system (e.g. DGGS) as it will greatly increase usability of the output.
- When geospatial information is collected along with data, the efficient and sustainable way to ensure quality is to make sure the information is accurate from the source. Therefore, design of the collection instrument should include a **point-of-entry validation tool.**





# GeoGSBPM – potential benefits

- Help production of geospatially enabled statistics to be conducted **in a systematic and consistent way**
- Provide a common framework to **manage quality and metadata** of statistical and geospatial information and services.
- Facilitate **sharing of geospatial services, methods and tools** that can be applied regardless of data types, domains and output formats
- [Wiki link GeoGSBPM](#)



# Questions



# Tools and Resources

- [Clickable GAMS0 v1.2](#) & [Clickable GSBPM v5.1](#)
- [Modernisation Maturity Model, Roadmap, and tools](#)
- [ModernStats World Workshops](#)
- [New: ModernStats Community of Practice](#)
- [GAMS0 Resources Repository](#) & [GSBPM Resources Repository](#)
- [GAMS0 Discussion Forum](#) & [GSBPM Discussion Forum](#)
- [GSBPM Training Materials](#)
- [Uses of GSBPM](#)
- [GeoGSBPM](#)
- More coming every day!

# Links













- GAMSO

- [Main space](#)
- [Clickable](#)
- [Resource repository](#)

- GSBPM

- [Main space](#)
- [Clickable](#)
- [Resource repository](#)
- [GeoGSBPM](#)

Welcome to GSBPM wiki space!

If you want to know	Relevant GSBPM resources
the latest version of GSBPM	 <b>GSBPM</b> Current version of GSBPM v5.1 available here
	 <b>Clickable GSBPM</b> Clickable GSBPM v5.1
introductory materials on GSBPM	 <b>Learning GSBPM</b> Introductory presentations from past training and workshops
how to implement the GSBPM in your organisations	 <b>Roadmap for Implementing ModernStats Standards</b> Modernisation Maturity Model (MMM) and MMM Roadmap Tools
how to model information flow within GSBPM using GSIM	 <b>Information flow within GSBPM using GSIM</b> Information flow within GSBPM described using Generic Statistical Information Model (GSIM)
how to use GSBPM for the production of geospatially-enabled statistics	 <b>Geospatial view of GSBPM</b> Geospatial-related activities and considerations needed for the production process
finer-level activities in GSBPM	 <b>GSBPM Tasks</b> <b>New</b> Finer-level of activities under the eight phases of the process and sub-processes
past versions of GSBPM	 <b>Previous versions of GSBPM</b> Previous versions of GSBPM (v1.0, v2.0, v3.1, v4.0 and v5.0)
how other statistical organisations implemented / used GSBPM	 <b>Uses of GSBPM</b> Use cases of GSBPM (mapping specific processes to GSBPM, using GSBPM for managing statistical programmes and etc.)
quality indicators across GSBPM	 <b>Quality Indicators for GSBPM</b> Quality indicators for each GSBPM subprocess
all resources available on GSBPM!	 <b>GSBPM resources repository</b> Click here to browse all resources in the GSBPM wiki space and search by
Or, just to ask questions about GSBPM	 <b>GSBPM Discussion Forum</b> Questions about GSBPM? Post them here!