GSBPM and other standards

Carlo Vaccari

UNECE High-Level Group for the Modernisation of Official Statistics (HLG-MOS)

UNECE
Conference of European Statisticians

modernistats
by HLG - MOS

High-Level Group for the Modernisation of Official Statistics

To promote common standards, models, tools and methods to support the modernisation of official statistics;

To drive new developments in the production, organisation and products of official statistics

To ensure that there is a maximum of convergence and coordination within the statistical "industry".

Modernisation of official statistics

Challenges for Official Statistics

New competitors in producing statistics

New data sources available (e.g. Big data)

Increased costs and quality problems to collect data

Reduced budgets for official statistics

Emerging information needs

More with less

Modernisation

Join efforts

Increase cooperation

Interactive and low-cost cooperation

Sharing of skilled resources

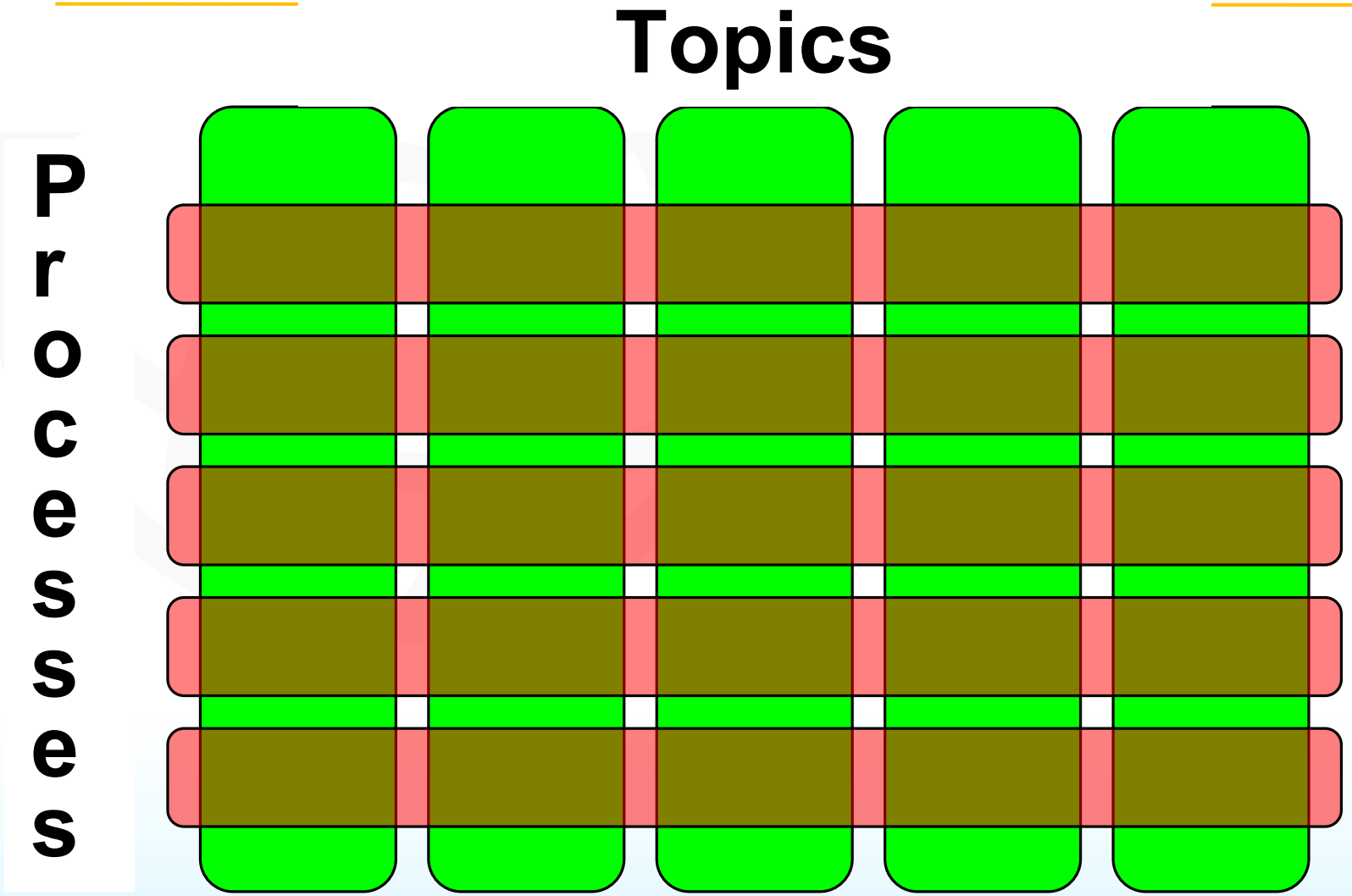
Sharing tools, methods, standards

Adopting common solutions

Why Modernisation?

- Statistical production has traditionally been organized by topic, e.g. transport, trade, …
- Financial pressures are encouraging new ways of thinking
- Some statistical organizations are moving towards a process-based approach
- Others are considering a matrix approach

Why Modernisation?



Official statistics challenges (by UNECE)

NSOs are facing many challenges

- Change how we work

Continuously changing environment

- Become agile and adaptive

We have to be more efficient

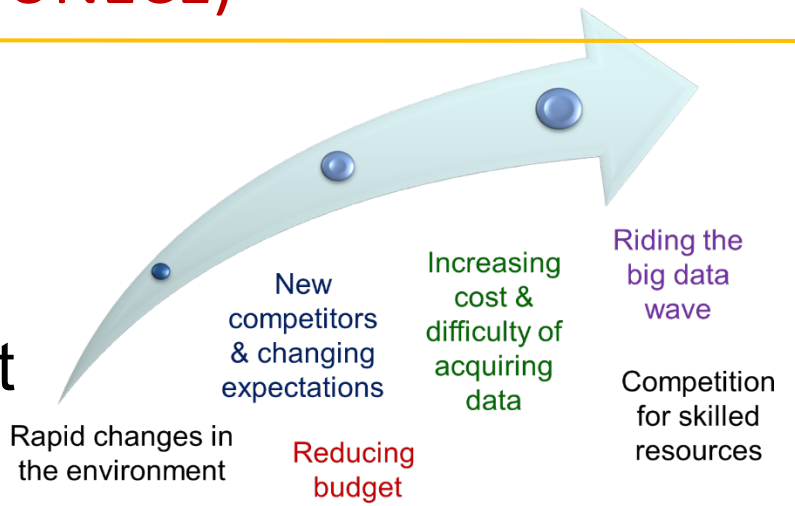
- **Share and Standardize**

Challenges too big to tackle on their own

- **Collaborate and Coordinate**

Official Statistics needs leadership

- Vision, Mission and Strategy



HLG-MOS Organization

HLG-MOS Organized in permanent groups:

- Supporting Standards
- Sharing Tools
- Organizational
- Blue Skies Thinking

Every year two projects, in 2019:

- Strategic Communication 2 (SC in 2018)
- Machine Learning (ML in 2018)

2012-2018: GSIM, CSPA I&II, Big Data I&II, Linked Metadata, Data Integration I&II, Data Architecture I&II, Strategic Communication Framework I

ModernStats World Structure 2019

UNECE

CES Bureau

HLG-MOS

HLG-MOS Seminars & Sprints

Modernisation WS

Executive Board

Modernisation updates

Innovation Topics

Blue Skies Thinking Network

Data Collection WS

Confidentiality WS

Data Editing WS

Dissemination Communication

Sharing Tools

Supporting Standards

Organisational Capability

Machine Learning Project

Strategic Communication Project II

CSPA Catalogue

CSPA Services

CSPA Promotion

ModernStats World WS

Sprints

GSBPM GSIM

GAMSO GSBPM

Ontology

Culture Change WS

Visions Mission

Skills & Capability

HRMT Forum

Sprint

Sprint

Sprint

modernstats by HLG - MOS

Modernisation Workshops

- Annual HLG-MOS Workshop in November (November 2019, Geneva)
- Annual ModernStats World Workshop (April 2019)
- HRMT:
  - Workshop on Culture Evolution (September 2019, Geneva, Switzerland)
  - Workshop on Human Resources Management and Training (September 2018, Oslo, Norway)
  - Workshop on Implementing Efficiencies and Quality of Output (September 2017 Geneva, Switzerland)
  - Workshop on Human Resources Management and Training: Developing Capabilities for the Future (September 2016, Krakow, Poland)
  - Workshop on Risk Management practices in Statistical Organisations (April 2016, Geneva, Switzerland)
- CSPA:
  - 2017 CSPA Workshop: *CSPA for Digital Transformation* (July 2017, Wiesbaden, Germany)
  - 2016 CSPA Workshop: *Implementing CSPA* (October 2016, Geneva, Switzerland)
- Other
  - Workshop on Data Integration: Realising the Potential of Statistical and Geospatial Data (May 2019, Belgrade, Serbia)
  - HLG Seminar on the Process Oriented Apprach (April 2017, Republic of Korea)
  - Workshop on Integrating Geospatial and Statistical Standards (November 2017)
  - HLG-MOS Sprint Session: Strategies to stay relevant (November 2017, Geneva, Switzerland)
  -
  - Also thematic workshops on Data Editing, Communication, Confidentiality and Data Collection

Supporting Standards Group

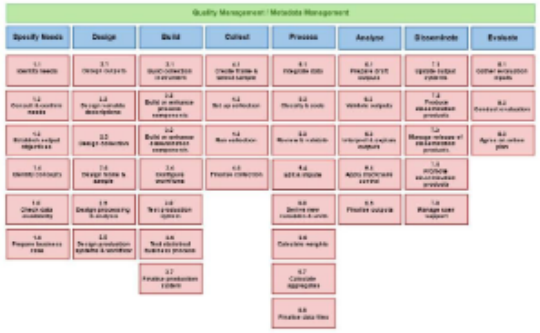
https://statswiki.unece.org/display/hlgbas/Modernisation+Groups

The goal of the group is to find ways how to develop, enhance, integrate, promote, support and facilitate implementation of the range of standards needed for statistical modernisation

Maintenance and development of the
Generic Activity Model for Statistical Organizations (GAMSO),
Generic Statistical Business Process Model (GSBPM),
Generic Statistical Information Model (GSIM) and the documentation of the
Common Statistical Production Architecture (CSPA)

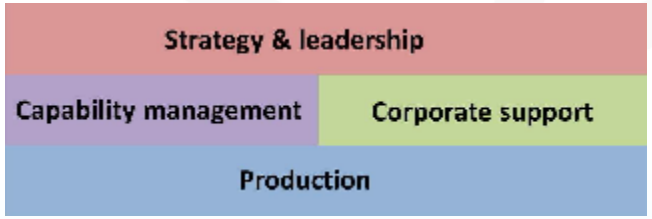
GSBPM and GAMSO
When have they been developed?

GSBPM



- ✓ Current version v5.1 – January 2019
- ✓ First release v1.0 – March 2008

GAMSO



- ✓ Current version v1.2 - January 2019
- ✓ First release v1.0 - March 2015

GSIM and CSPA
When have they been developed?

✓ Current version 1.2 - August 2019

✓ Current version v1.5 - December 2015

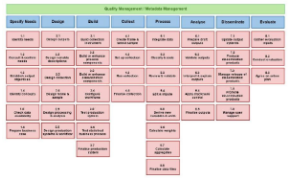
Common purposes of the standards

- ✓ Improve communication by introducing a common language
- ✓ Gain efficiency (rationalise processes, information, flow, assign responsibilities,..)
- ✓ Support industrialisation process (re-use of methodologies, tools, software, sharing of solutions, …)
- ✓ Build staff competencies around the standards (enhance capabilities)

Management of the standards

- ✓ Owner = High-Level Group
- ✓ Maintenance is delegated to the Modernisation Committee on Standards
- ✓ Discussion forums to gather feedback
- ✓ Importance of stability over time
- ✓ Reviews every 5 years (GSBPM in revision)
- ✓ Revisions only if really needed

The GSBPM



https://statswiki.unece.org/display/GSBPM/Generic+Statistical+Business+Process+Model

The Generic Statistical Business Process Model (GSBPM) is an international standard that can be used to document any kind of statistical business process, from the more traditional survey to the administrative data acquisition or to the statistical compilation

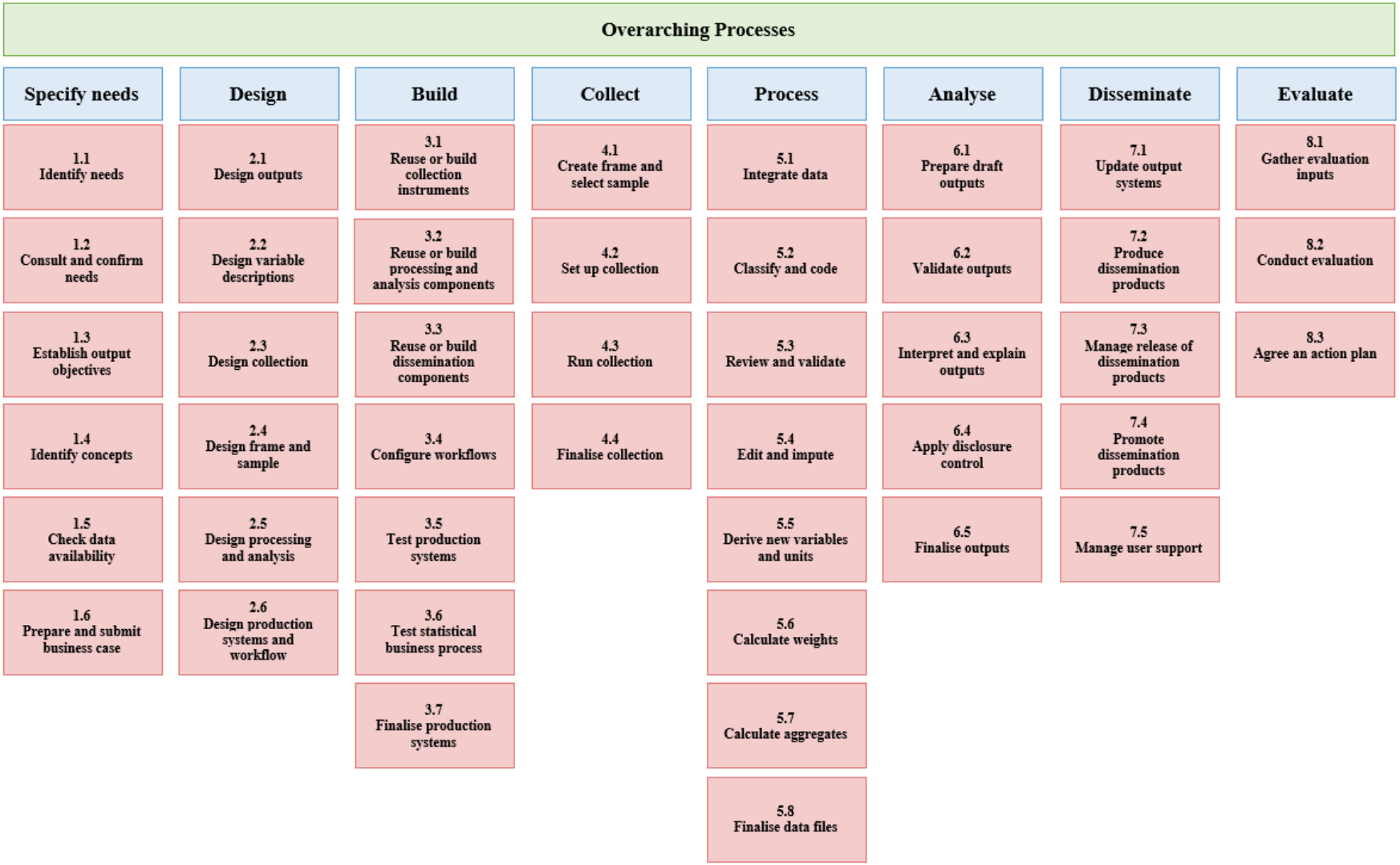
- Big data and new data sources included
- Mixed-sources

The GSBPM

Why do we need the GSBPM?

- ✓ To define and describe statistical processes in a coherent way
- ✓ To compare and benchmark processes within and between organisations
- ✓ To make better decisions on production systems and organisation of resources

The GSBPM


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

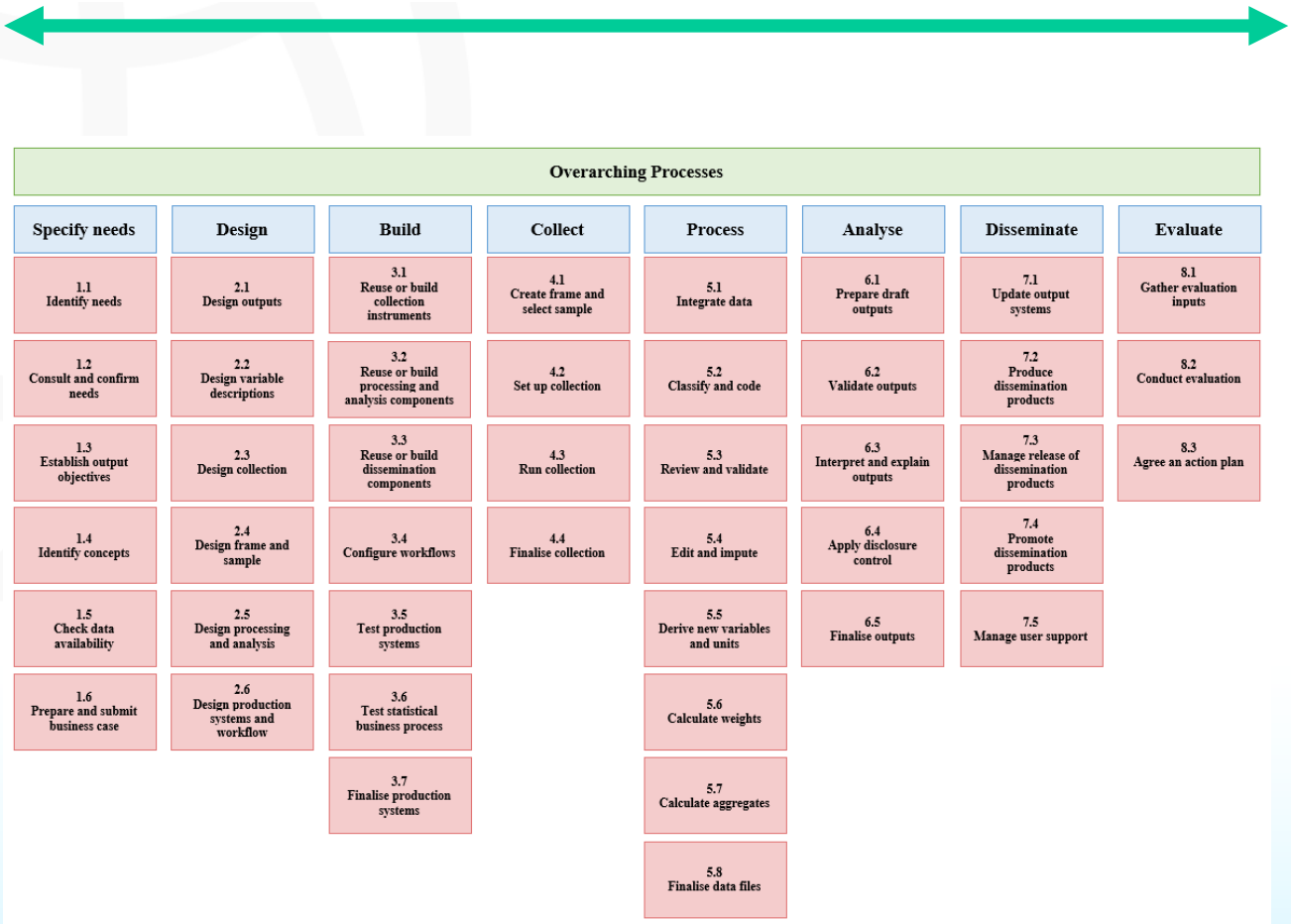
The GSBPM

Process

Phases

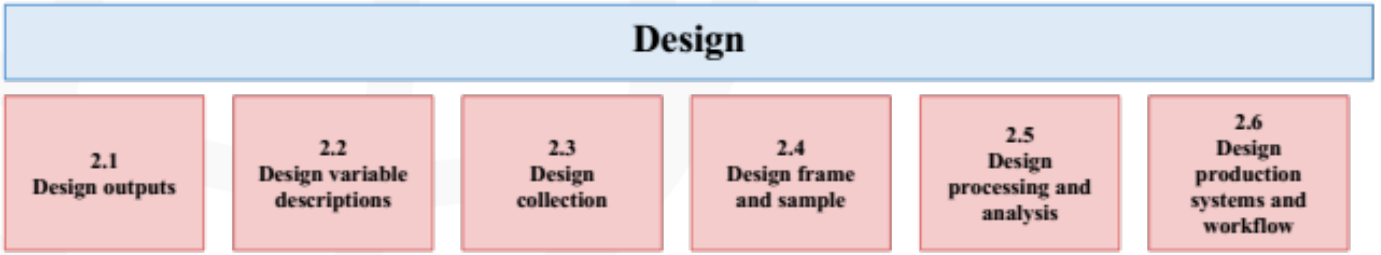
Sub-processes

(Descriptions)

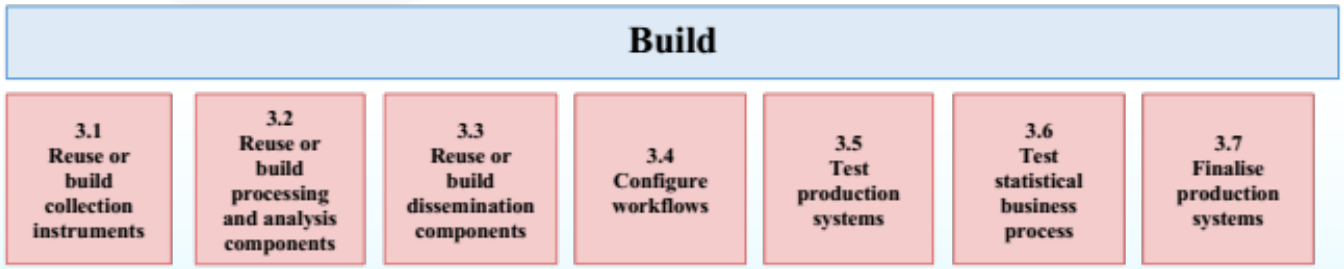


The GSBPM

Dialog with users, Identification of needs (new or additional), Definition of high level solution



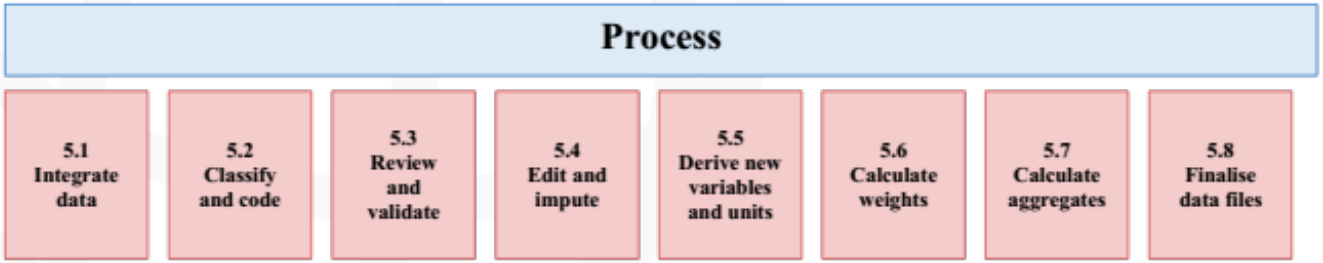
Definition of all methods and tools that will be used in the realisation of the statistical process



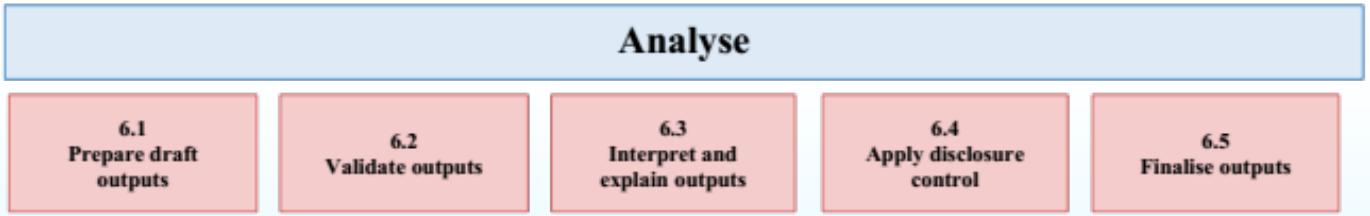
Set up and test of all methods and tools defined in the design stage

The GSBPM

The actual data acquisition, whatever the source or the method used, including data entry



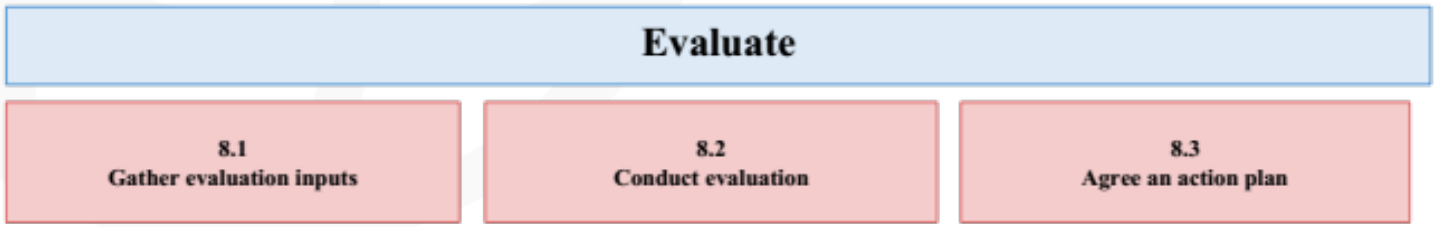
The traditional phases of data treatment till the macrodata estimates are produced



It includes the production of complex statistics (e.g. indices), macrodata validation, confidentiality treatment

The GSBPM

The release of statistical outputs to users



The quality evaluation done at the end of a specific edition of a statistical business process

The GSBPM

- National implementations may need additional levels
- Over-arching processes
  - Quality management
  - Metadata management
  - Statistical framework management
  - Statistical programme management
  - …….. (others – see page)

The GSBPM Applicability

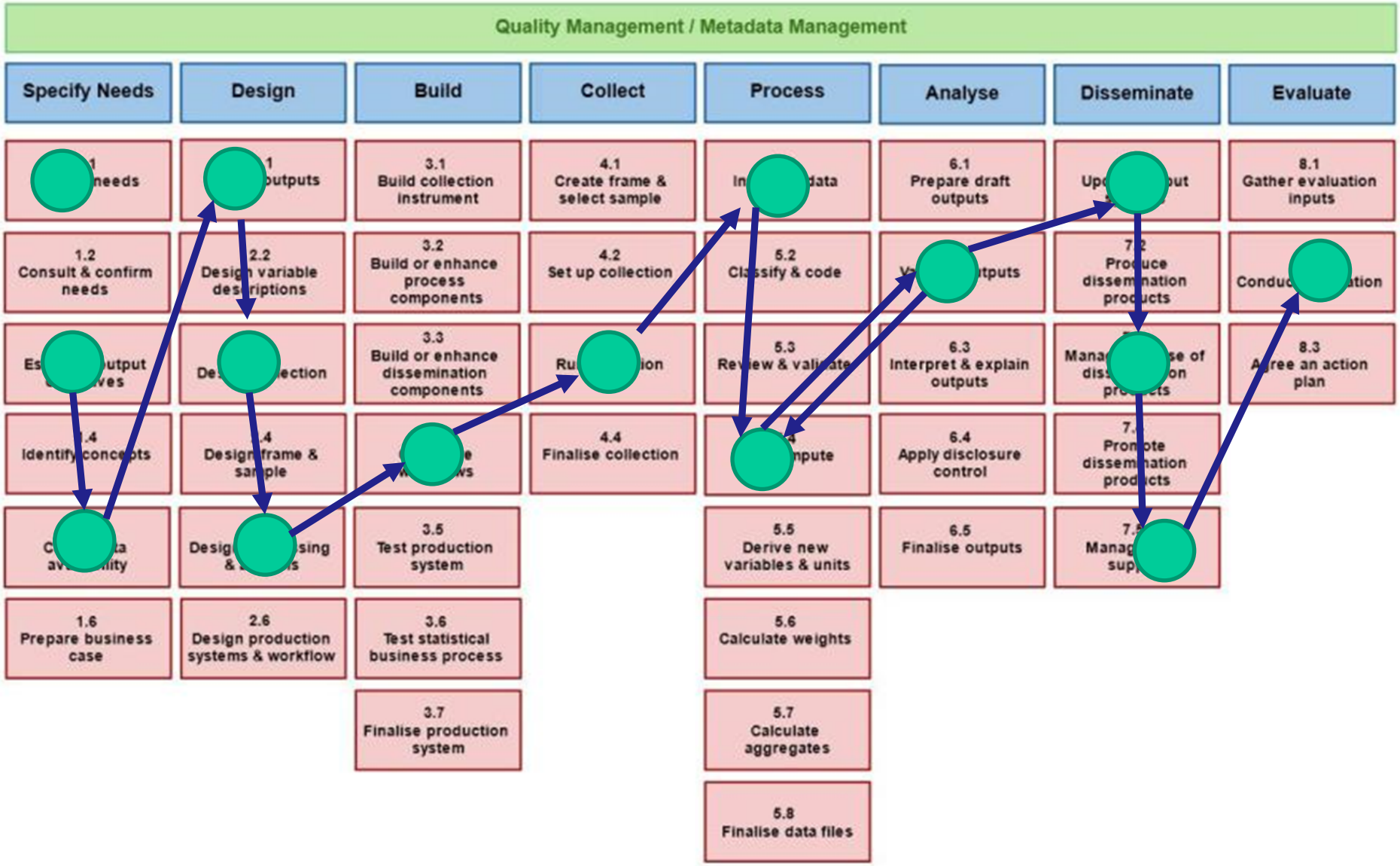
- ✓ All activities undertaken by producers of official statistics which result in data outputs
- ✓ All statistical domains
- ✓ National and international statistical organisations
- ✓ All types of data source:
  - Surveys / censuses
  - Administrative sources / register-based statistics
  - Mixed sources
  - “Big Data”

The GSBPM key features

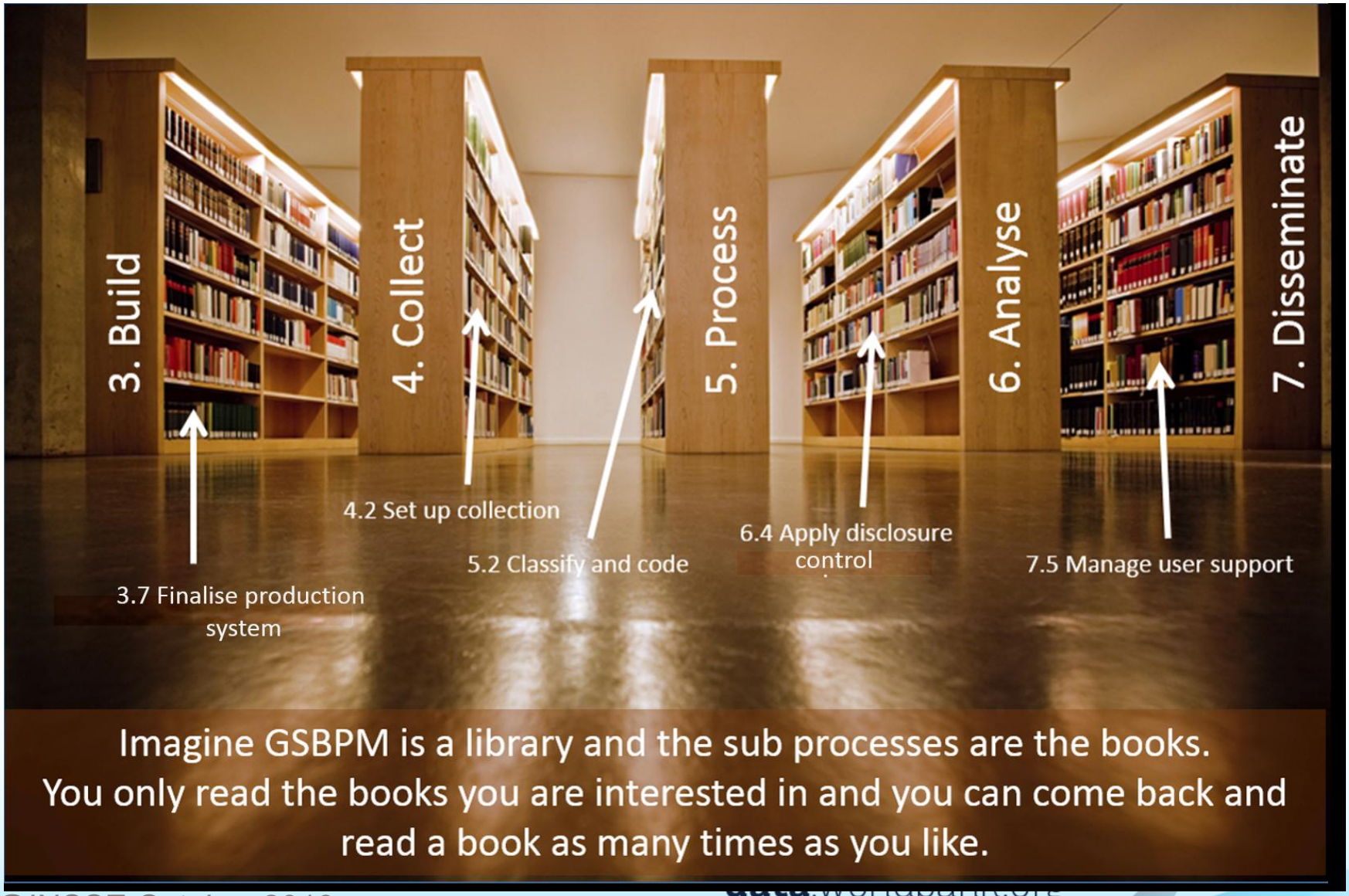
Not a linear model

- Sub-processes are not followed in a strict order
- It is a matrix, through which there are many possible paths undertaken by producers of official statistics which result in data outputs

The GSBPM key features



The GSBPM key features



GSBPM other utilizations

- ✓ Managing statistical programmes
- ✓ Cost / resource allocation
- ✓ Documenting statistical processes
- ✓ Framework for quality assessment
- ✓ Sharing statistical software

National implementations

See in the link experiences from Armenia, Australia, Canada, Denmark, Eurostat, Ireland, Korea, Spain, Sweden, Turkey

Other uses:

- Process-related metadata
- Managing Statistical programs
- Tool for Cooperation: Software Inventory

GSBPM revision: why?

Users needs

- Recent experiences from users in implementing GSBPM in their organisations and a number of developments in business landscape where statistical organisations operate are requiring changes in GSBPM.

CES

- This model was endorsed by the CES in 2017 on the understanding that it should be updated every 5 years.
  - GSBPM v.1 -2008, v.4 -2009 and v.5 -2013 https://statswiki.unece.org/display/GSBPM/GSBPM+v5.0
  - Revisions only if really needed

Other Models - GAMSO

- The recent developments of GAMSO requires some updating to GSBPM in order to ensure alignment between the two reference models.

GSBPM revision purposes

- Periodic revision every 5 years to ensure the model remains relevant and reflects the changes in the production environment
- Increase clarity
- Improve consistency with other modernisation models
- Recognise the growing importance of integrating statistical data with geospatial data
- Make the model less survey-centric, account for new data sources including big data and relationships with data providers
- Ensure continuity with previous version. Strong case study before making any changes

GSBPM revision

The revision focused on three major tasks:

- Compilation of user feedback (July 2017 - October 2017)
- Review of feedback and revision process (October 2017 - June 2018)
- Public consultation and revision (July 2018 - November 2018)

Released (January 2019) – Approved by CES June 2019

The GSBPM revision working group met virtually every three weeks

The new versions have to be approved by the HLG-MOS and by CES

GSBPM revision principles

- To the largest extent possible, keep the existing structure unchanged
- Revision needs to encompass other exchange channels and data sources, such as administrative data, big data and geospatial data
- Ensuring consistency with other standards (GSIM, GAMSO, CSPA)

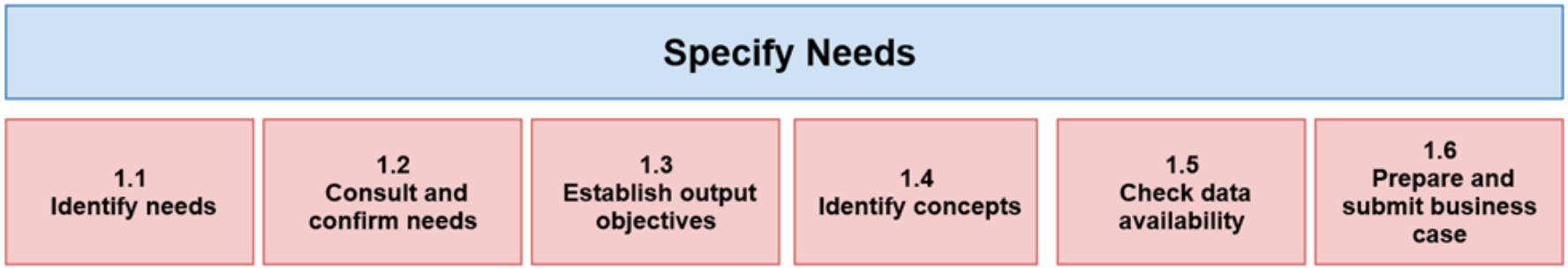
GSBPM revision challenges

- Achieving clarity about the model without reducing the simplicity of the model
- Determining the boundaries of the subprocesses more precisely
- Allowing for the business processes for all data sources
- Review and expand the definition of the overarching processes
- Difficulty to reflect some users' specific needs in the generic model
- Make sure to keep in GSBPM what has been a success (i.e., a general framework of the data lifecycle for official statistics)

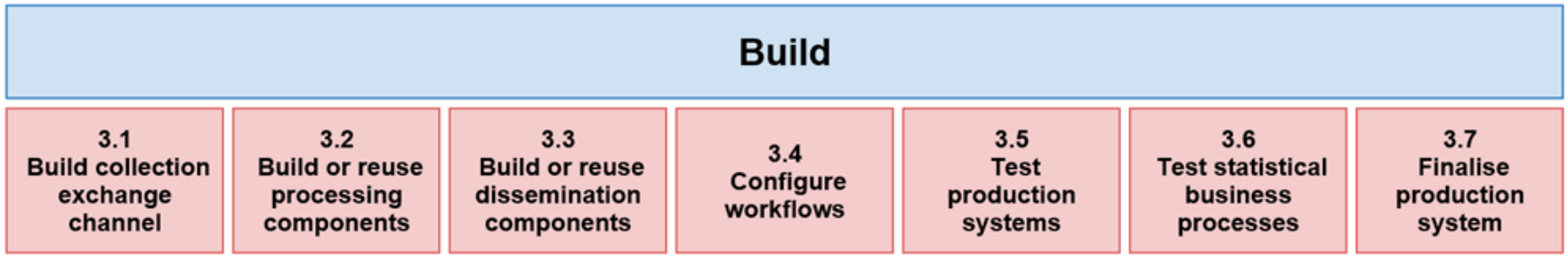
GSBPM revision: main changes

- Few sub-processes have been re-named to improve clarity
- The duplication between the overarching processes in the GSBPM version 5.0 and the GAMSO has been resolved
- Descriptions of the phases and the sub-processes have been updated to be less survey-centric and activities related to working with non-statistical data providers have been added
- In recognition of the growing importance of integrating statistical data with geospatial data, descriptions have been expanded
- Terminology has been modified where necessary to improve consistency with the GAMSO and the GSIM
- Examples and descriptions have been updated and expanded

GSBPM revision: main changes



1.6 **Prepare and submit** - (instead of prepare)



3.1 **Collection exchange channel** - (instead of collection instrument) changed in final review

3.2 **Build or reuse** (instead of build or enhance)

GSBPM revision: example of changes

Give a look to GSBPM v5_1 (track change) document

GAMSO

GAMSO stands for Generic Activity Model for Statistical Organizations

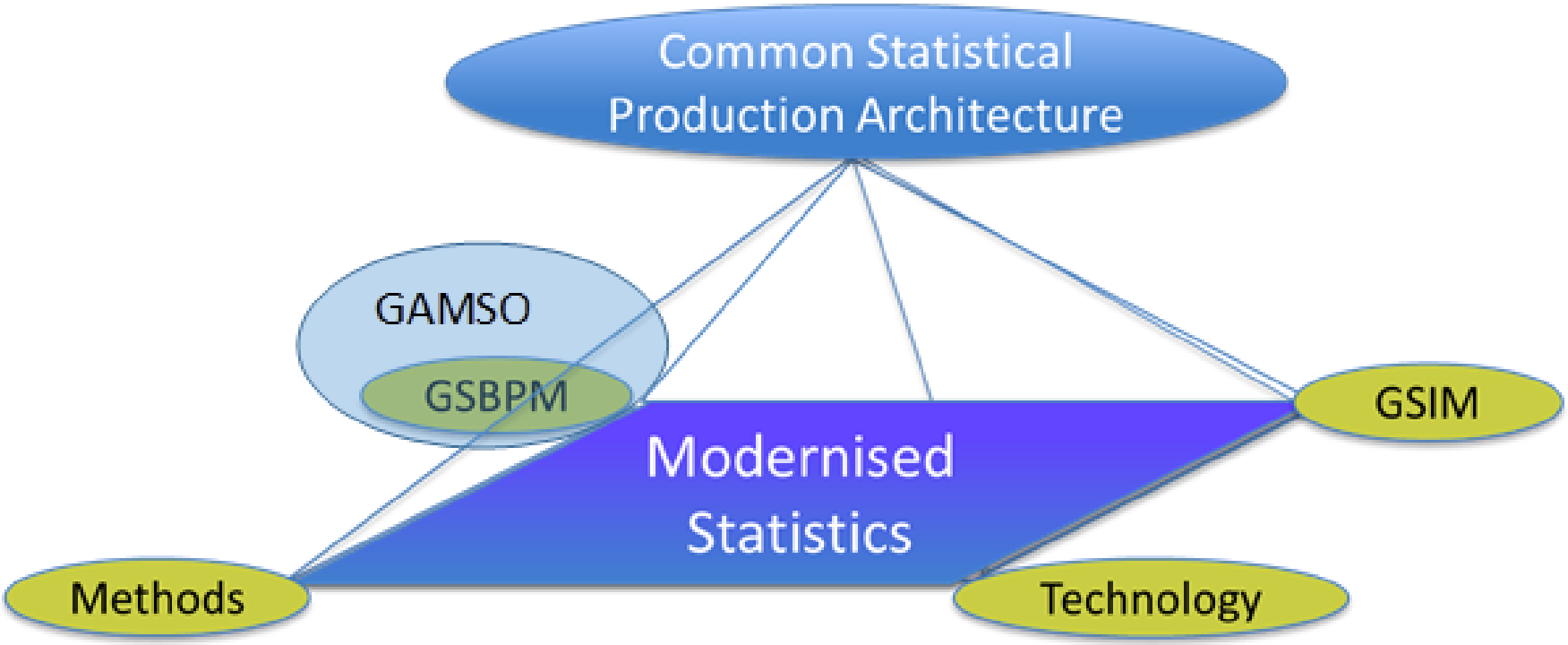
GAMSO extends and complements the Generic Statistical Business Process Model (GSBPM) by adding additional activities needed to support statistical production.

When the GSBPM was developed, such activities were referred to as over-arching processes but not elaborated in any great detail.

The GAMSO is based on two existing models:

- The GSBPM v5.0, which provides the contents of the Production activity area
- The Statistical Network business activity model, which provides the basis for the Strategy and leadership, Capability management and Corporate support activity areas.

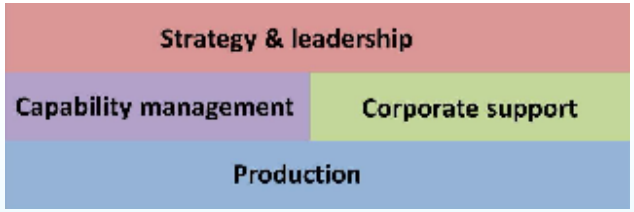
GAMSO


Common Statistical Production Architecture
GAMSO
GSBPM
GSIM
Modernised Statistics
Methods
Technology

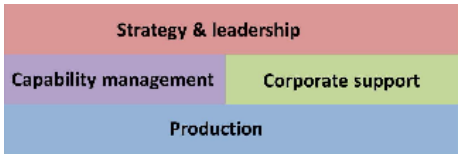
Generic Activity Model for Statistical Organisations - GAMSO

What is GAMSO?

It describes and defines the activities that take place within a typical statistical organization. It extends and complements the Generic Statistical Business Process Model (GSBPM) by adding additional activities needed to support statistical production



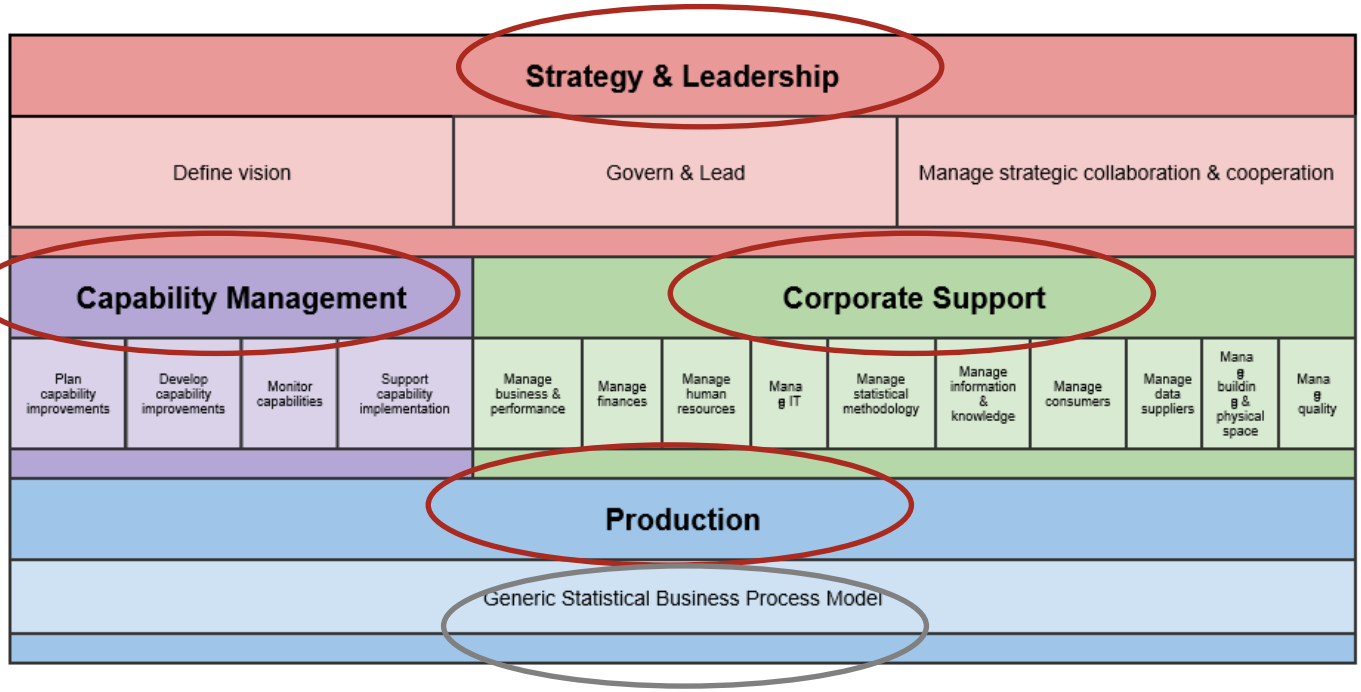
GAMSO



Common vocabulary and framework

- basis for resource planning within a statistical organisation
- basis for the measurement of costs of producing official statistics in a comparable way
- tool to help assess the readiness of organisations to implement different aspects of modernisation, in the context of a proposed “Modernisation Maturity Model”
- support risk management systems

Activity areas

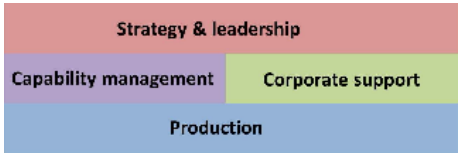


The GAMSO comprises three hierarchical levels.

The top level comprises four broad activity areas

The production area is described by the GSBPM

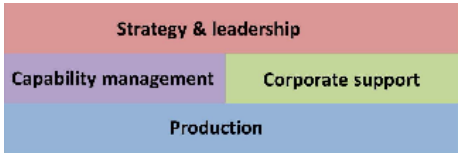
Strategy and leadership



Strategy and leadership		
1.1	**1.2**	**1.3**
Define vision	**Govern & lead**	**Manage strategic collaboration & cooperation**

- ✓ high-level strategic activities that enable statistical organisations to deliver the products and services needed by governments and communities nationally and internationally

Capability management

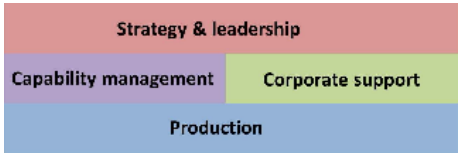


Capability management			
2.1	**2.2**	**2.3**	**2.4**
Plan capability improvements	**Develop capability improvements**	**Monitor capabilities**	**Support capability implementation**

support the successful development and monitoring of the capabilities that underpin an organisation's ability to conduct its business.

promoting the re-use and sharing of infrastructure (statistical and technical), both inside the organisation and across organisations, thus facilitating harmonisation and coherence of statistical outputs.

Corporate support



Corporate Support									
3.1	3.2	3.3	3.4	3.5	3.6	3.7	3.8	3.9	3.10
Manage business & performance	Manage finances	Manage human resources	Manage IT	Manage statistical methodology	Manage information & knowledge	Manage consumers	Manage data suppliers	Manage buildings & physical space	Manage quality

- ✓ These activities cover the cross-cutting, functions required by the organisation to deliver its work programme efficiently and effectively.

GAMSO

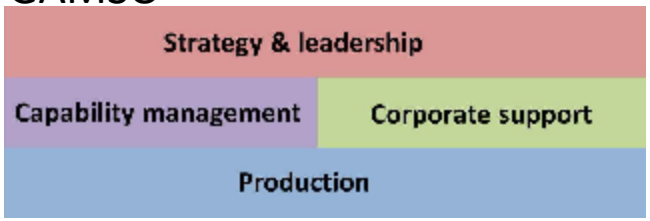
Like the GSBPM, the GAMSO aims to provide a common vocabulary and framework to support international collaboration activities, particularly in the field of modernisation.

Some expected uses of the GAMSO are listed below:

- As a basis for resource planning within a statistical organisation
- As a basis for the measurement of costs of producing statistics
- As a tool to help assess the readiness of organisations to implement different aspects of modernisation
- To support risk management systems
- To support the implementation of enterprise architectures
- To help to measure and communicate the value of statistical modernisation activities across an organisation

Interrelations among the standards

GAMSO



The information flow

GSBPM