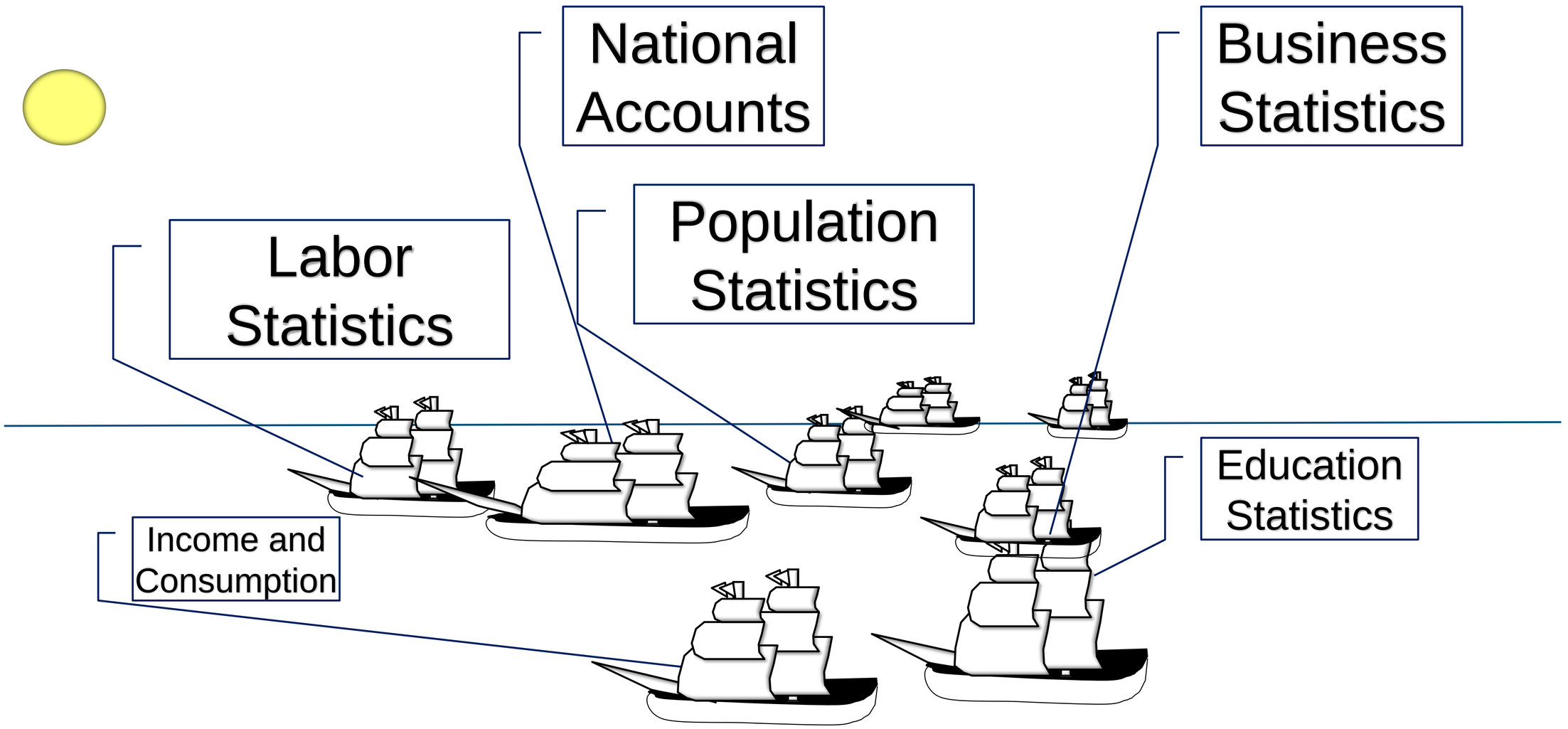
# Driving Official Statistics to its Modernization

**Panel on Transparency and Reproducibility of Federal Statistics for NCSES, Meeting 2**
**The National Academies of Sciences, Engineering, and Medicine**
**Washington, D. C.. September 9th 2019**

Juan Muñoz López

- Some reasons to modernize official statistics
- HLG-MOS and the Statistical Community
- Models proposed by the HLG-MOS
- Standards outside the HLG-MOS coordination
- How these models fit together
- Examples of how these standards and models are being used in the Official Statistics Offices to modernize the production of statistics
- Importance for Transparency and Reproducibility of Official Statistics
- The way forward
- Questions to the audience

# Some reasons to modernize official statistics




# Illusion of the “Status Quo”

Reduced Budgets

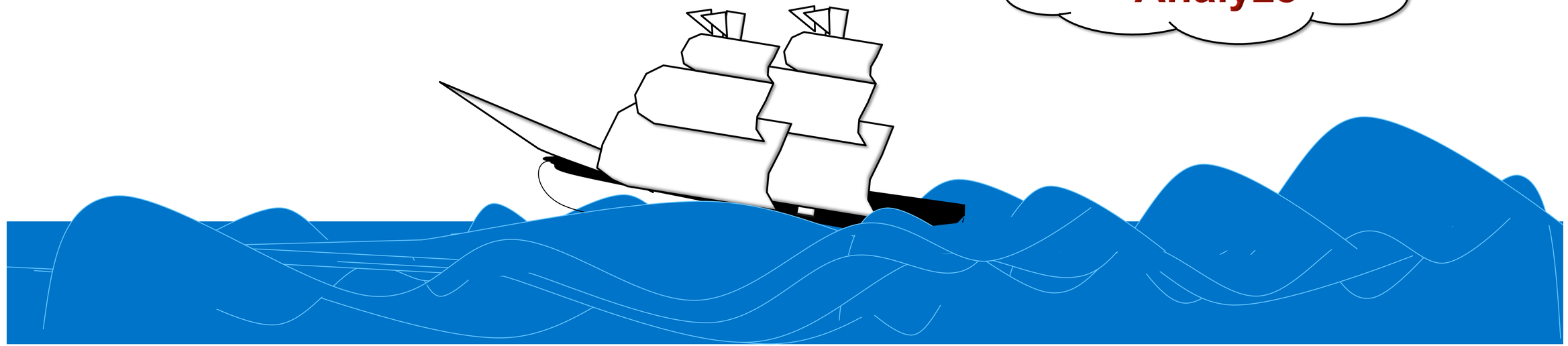
New Connected Technologies

Decay of User's Confidence – Fake News -

New User Needs

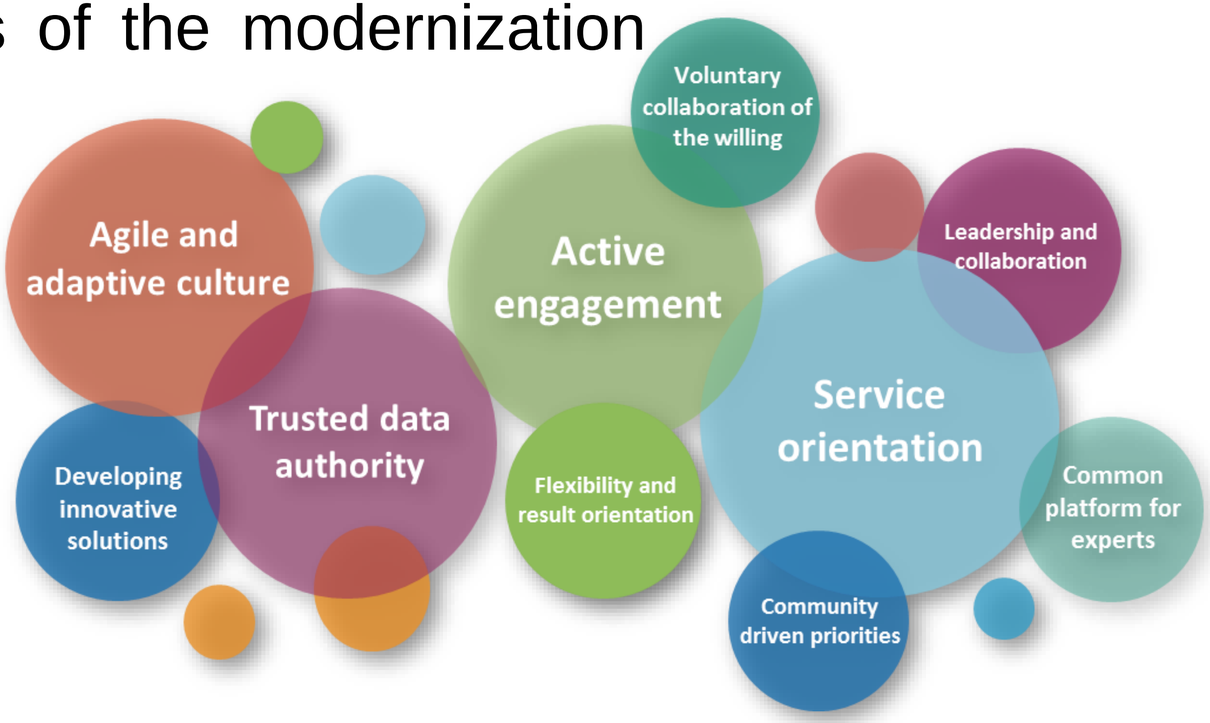
Shorter times to deliver results

More Facts to Analyze



# HLG-MOS and the Statistical Community

The High Level Group for the Modernization of Official Statistics (HLG-MOS) is responsible for deciding on the annual flagship international collaboration projects undertaken within the United Nations Economic Commission for Europe (UNECE) statistical modernization program, as well as overseeing and providing strategic direction to the work programs of the modernization groups.

- Australia
- Canada - Chair
- Ireland
- Italy
- Mexico
- Netherlands
- New Zealand
- Republic of Korea
- Slovenia
- United Kingdom
- Eurostat
- OECD
- UNECE

- **High-Level Group for the Modernisation of Official Statistics setup in 2010* under the Conference of European Statisticians (CES)**
- **Statistical Modernisation Community – Statement of Intent (2015)**
- **Since 2016 the community is branded as:**

modernistats BY HLG - MOS

*originally as HLG for the Modernisation of Statistical Production and Services & HLG Group for strategic Developments in Business Architecture in Statistics (HLG-BAS)

# HLG-MOS members and history

- Australian Bureau of Statistics
- Central Bureau of Statistics (Netherlands)
- Central Statistics Office (Ireland)
- Eurostat
- Hungarian Central Statistical Office
- INEGI (Mexico)
- Istat (Italy)
- OECD Statistics Directorate
- Puerto Rico Institute of Statistics
- Statistical Office of the Republic of Serbia
- Statistical Office of the Republic of Slovenia
- Statistics Canada
- Statistics Korea
- Statistics New Zealand
- Statistics Poland
- Statistics Sweden
- Statistical Office of Montenegro
- UNECE Statistical Division

# Statement of Intent Endorsed

CES and Bureau

HLG-MOS

Executive Board


Blue Skies Thinking Network


Sharing Tools

- CSPA Catalogue
- CSPA Services
- CSPA Promotion


Supporting Standards

- GSBPM GSIM
- GAMSO GSBPM
- Ontology

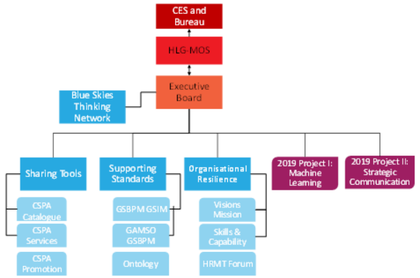
Organisational Resilience

- Visions Mission
- Skills & Capability
- HRMT Forum


2019 Project I: Machine Learning

2019 Project II: Strategic Communication

- The HLG-MOS reports directly to the Conference of European Statisticians
- Executive board: Oversees the work of the HLG-MOS
- Blue-Skies Thinking Network: Innovations, proposals of projects
- Supporting Standards Group: Evolution, maintenance and support of HLG-MOS standards and models
- Capabilities and Communication Group: Organizational evolution
- Sharing Tools Group: Sharing of ITC solutions, implementing of CSPA
- Each year two priority projects are selected
- It addresses some specialized topics too: Dissemination and Communication, Data Collection, Statistical Data Editing and Statistical Confidentiality




# HLG-MOS Organization 2019

# Models proposed by the HLG-MOS

# Generic Activity Model for Statistical Organizations GAMSO 1.2 January 2019

## Strategy and Leadership

| Define Vision | Govern and Lead | Manage Strategic Collaboration and Cooperation |
|---|---|---|
| • Understand national and international directions and factors<br>• Determine vision, mission and strategic goals<br>• Determine organisational value proposition<br>• Determine and communicate values and expectations<br>• Create interest and awareness | • Develop strategies for achieving organisational goals<br>• Prioritise capability portfolio<br>• Prioritise statistical product and service portfolio<br>• Define and manage statistical programme<br>• Allocate project and programme portfolio budgets<br>• Build and maintain internal statistical and professional excellence<br>• Ensure general coordination and alignment<br>• Define general organisational policies<br>• Publish policies, guidelines and normative documents | • Build and maintain strategic relations, nationally a nd internationally<br>• Build and maintain external statistical excellence<br>• Advance inter -agency and international collaborations<br>• Secure support for statistical product and service and capability portfolio<br>• Coordinate the national statistical system |

## Capability Development

## rt

| Plan Capability Improvements | Develop Capability Improvements | Monito Capabil Improvem | | | | | | | | Manage Finances | Manage Human Resources | Manage IT | Manage Buildings and Physical Space |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| • Identify disruptions and capability improvements<br>• Propose capability improvement projects<br>• Manage capability improvement programmes | • Undertake background research<br>• Define detailed capability requirements<br>• Design capability solution<br>• Build/procure and deploy capability solution | • Maintain capab improvements<br>• Promote capab improvements<br>• Evaluate capab improvements | | | avoidance<br>• Data linkage | | standards and access rights<br>• Manage metadata and data | | | • Maintain accounts (including assets and liabilities)<br>• Manage procurement and contracts<br>• Manage suppliers of equipment, office supplies and services | • Manage employee performance<br>• Manage and develop skills<br>• Manage talent<br>• Manage recruitment<br>• Ensure succession planning | • Manage IT assets and services<br>• Manage IT security<br>• Manage technological change<br>• | • Manage environmental, mechanical, and electrical needs<br>• Manage arrangement of office space<br>• Manage distribution of offices within space |

## Production

## Generic Statistical Business Process Model

INEGI | GAMSO: Organizational Context

| 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 | 4.2 | 5.2 | 6.2 | 7.2 Produce dissemination products | 8.2 Conduct evaluation |
| 1.3 Establish output objectives | 2.3 Design collect | | | | | 7.3 Manage release of dissemination products | 8.3 Agree an action plan |
| 1.4 Identify concepts | 2.4 Design frame a sample | | | | | 7.4 Promote dissemination products | |
| 1.5 Check data availability | 2.5 Design process and analysis | | | | 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 | | | |

# Generic Statistical Business Process Model GSBPM 5.1 January 2019

Quality Management

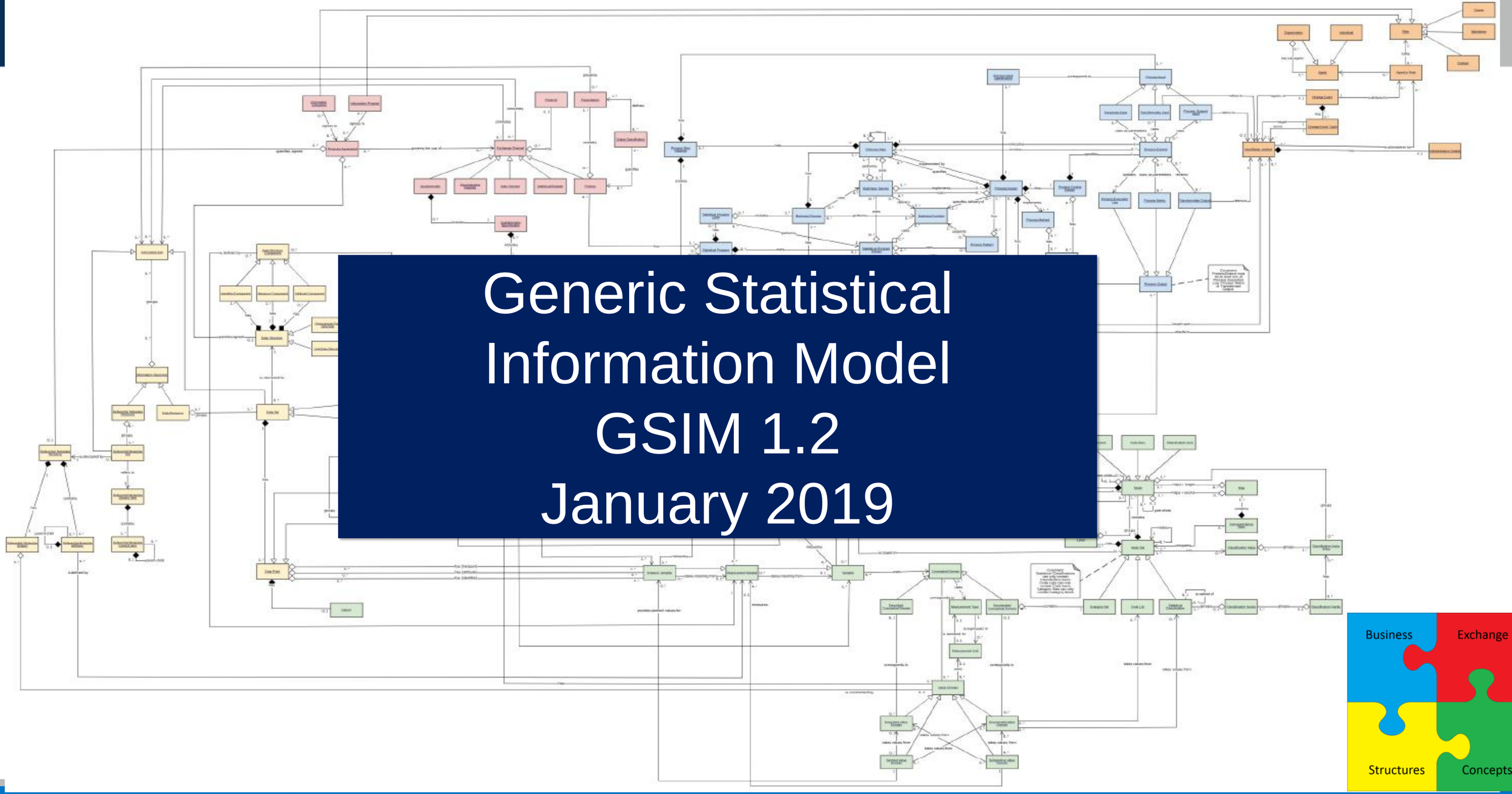
Metadata Management

Data Management

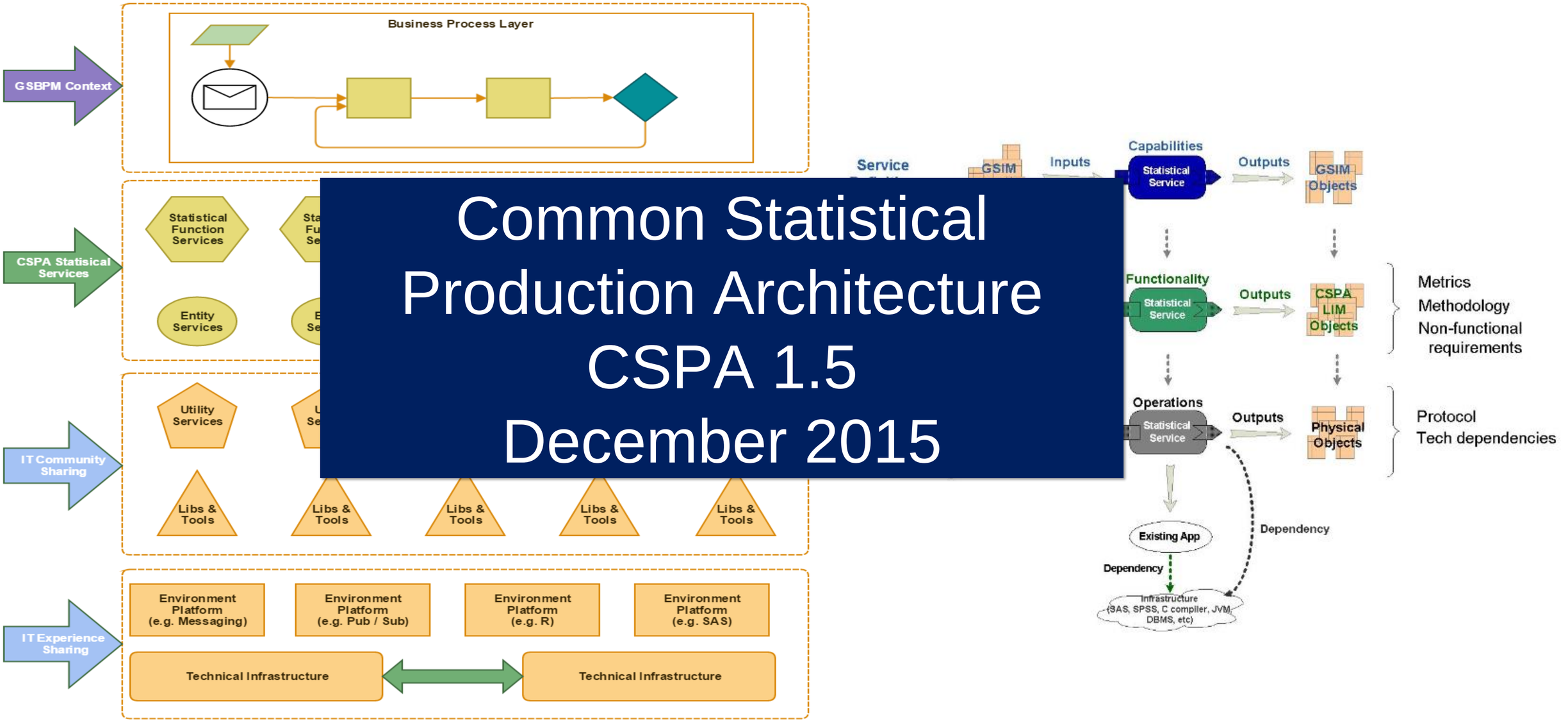
Process Data Management

Knowledge Management

Provider Management

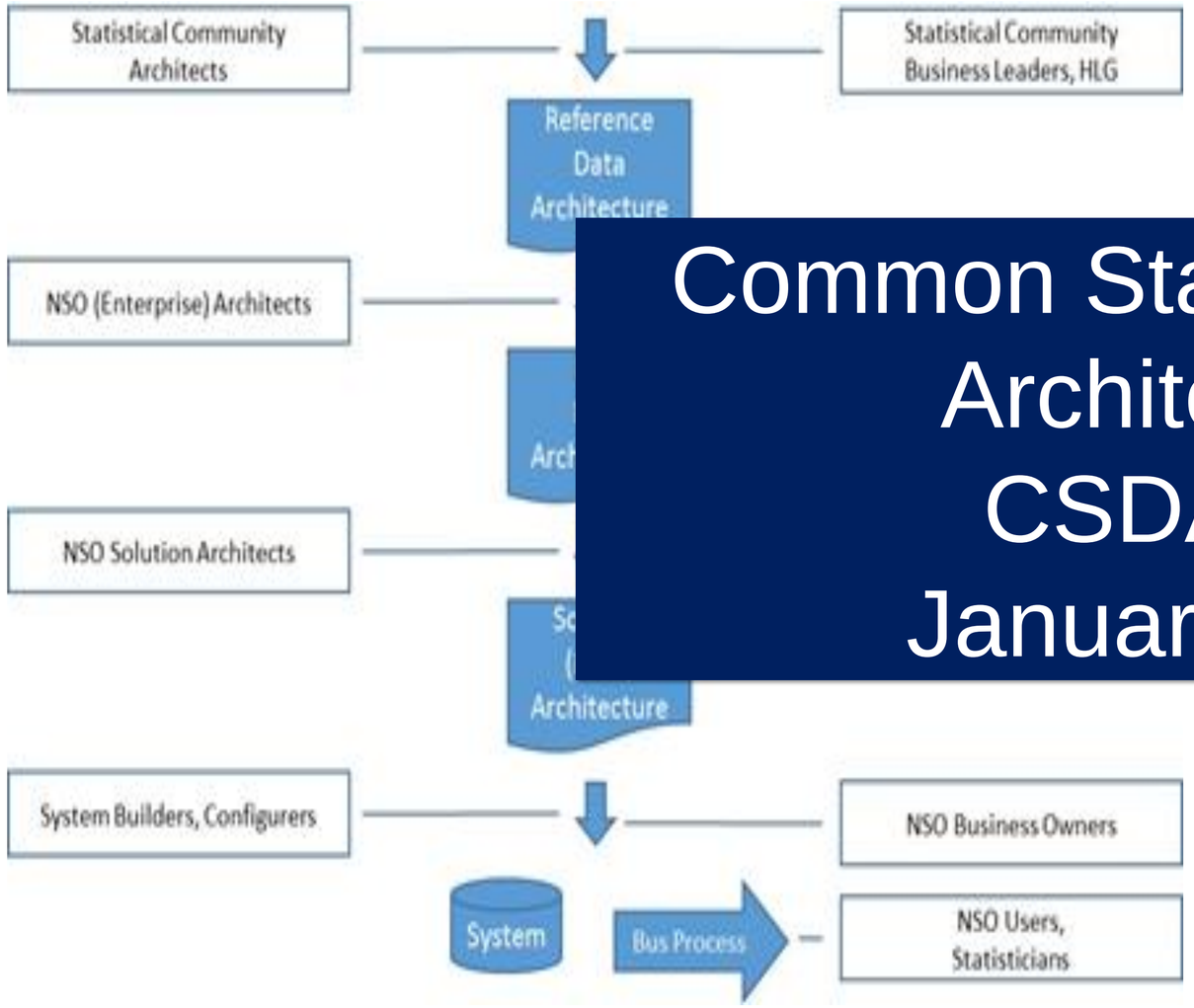


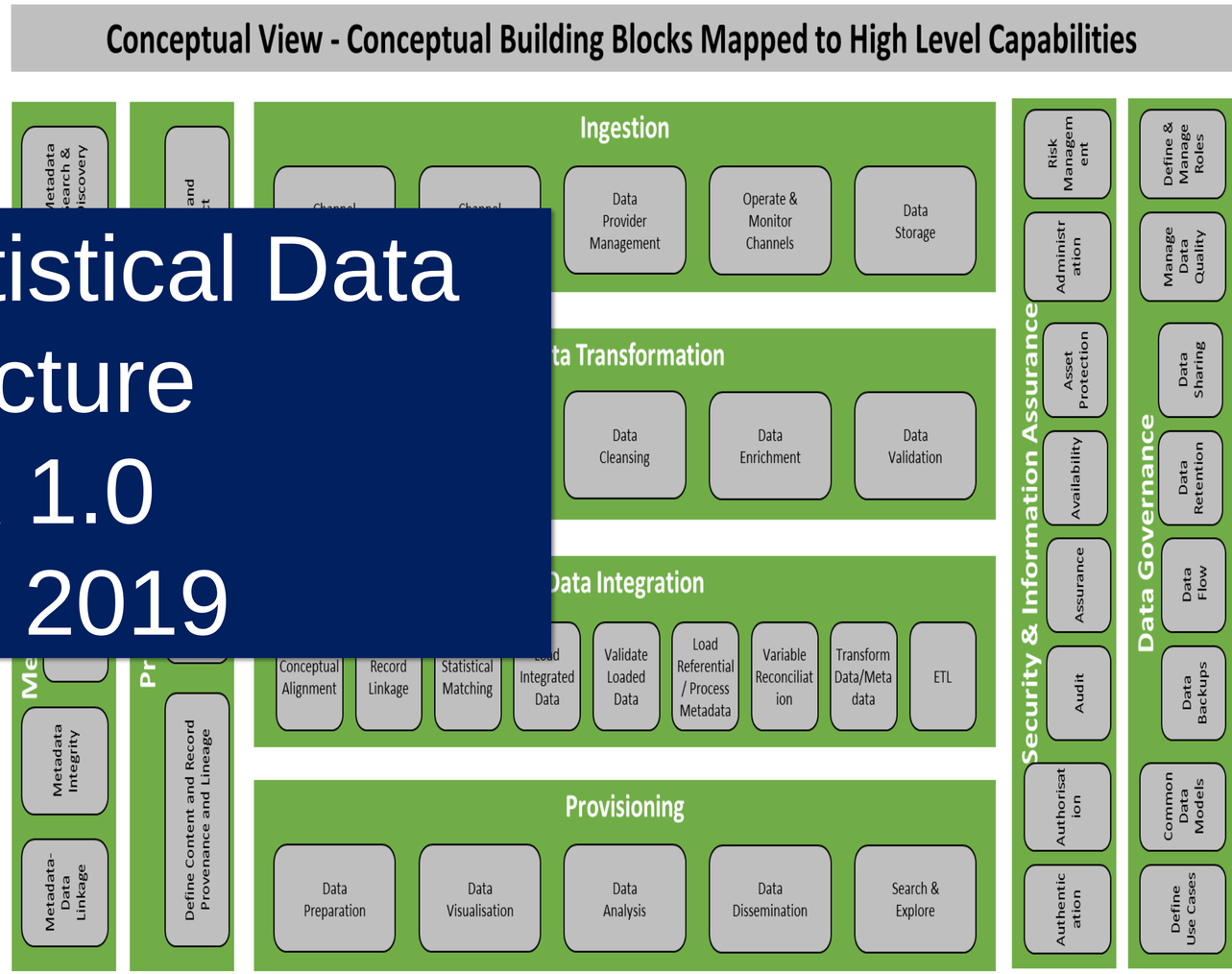
# Generic Statistical Information Model
# GSIM 1.2
# January 2019




# Common Statistical Production Architecture CSPA 1.5 December 2015

# Common Statistical Data Architecture
# CSDA 1.0
# January 2019



# Standards outside the HLG-MOS coordination

- Sponsored by 7 international organizations: BIS, ECB, Eurostat, IMF, OECD, UNSD, and the World Bank.
- Promotes the us
  technologies to e
  metadata efficien
- Harmonizes con
  and structures.
- It provides tools to better understand and interpret statistical information.

# Statistical Data and Metadata Exchange SDMX 2.1 (ISO IS 17369) April 2013

- Standardize metadata concepts to describe surveys, questionnaires, statistical data files, and social level information
- Used to Document different stages in the research data lifecycle, such as conceptualization, collection, processing, distribution, discovery, and archiving.


Data Documentation Initiative
DDI lifecycle 3.2 February 2014,
DDI codebook 2.5 January 2015

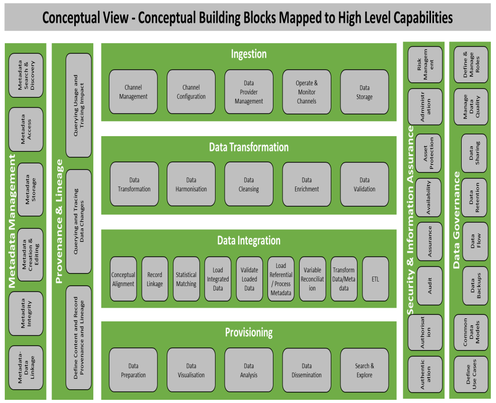
# How These Models Fit Together

**GAMSO**: Corporative Business Support

Organizational Context

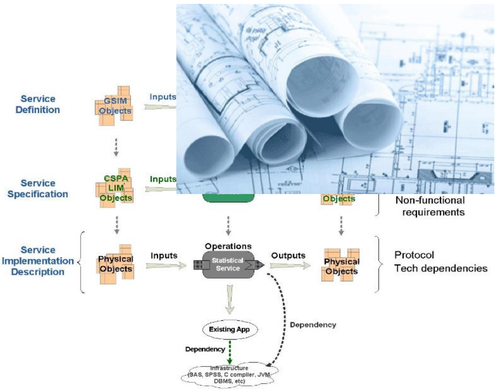

**GSBPM**: Organizing the Process

Statistical Process Life-cycle



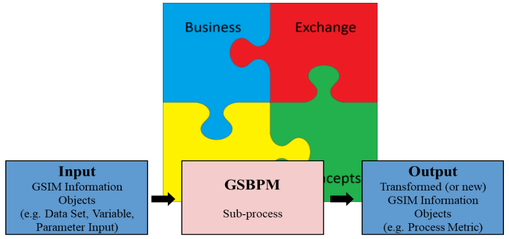

**CSDA**: Managing Data

Needed Practices



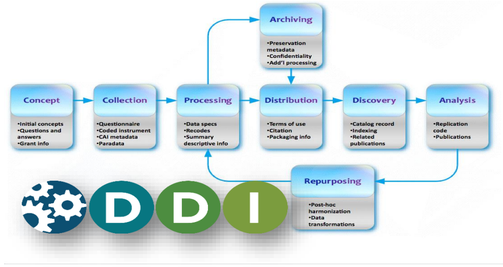

**CSPA**: Developing IT Systems

Technological Support



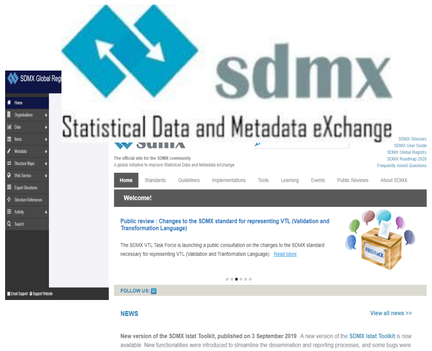

**GSIM**: Naming Conventions

Common Semantic/Language




**DDI**: Metadata Structures

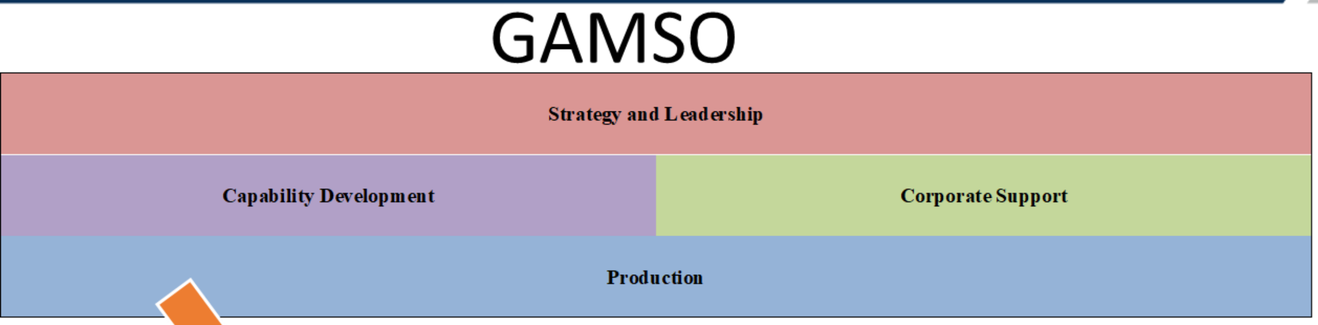
Process Documenting

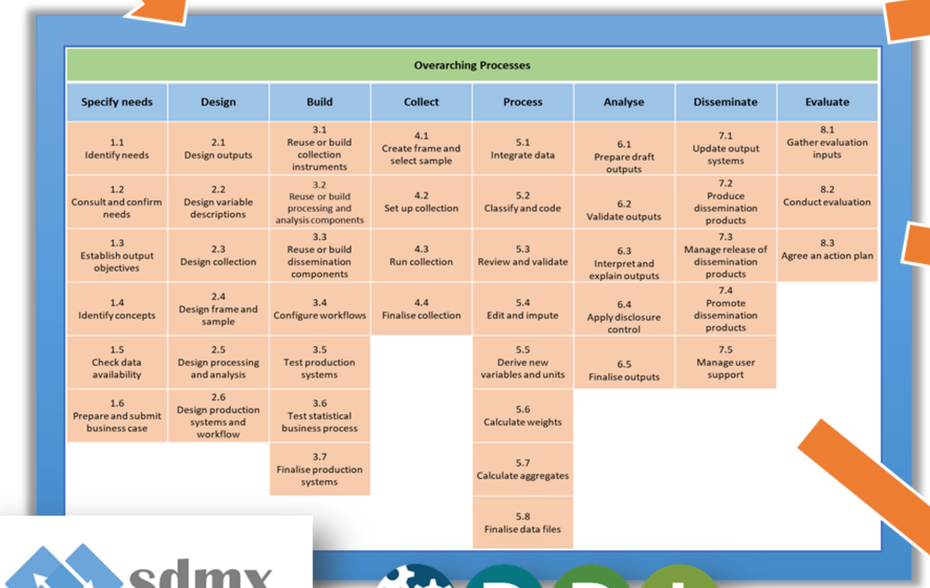



**SDMX:** Analyzable and Comparable Data

Exchange and Dissemination

INEGI | Creating an Ecosystem of Standards



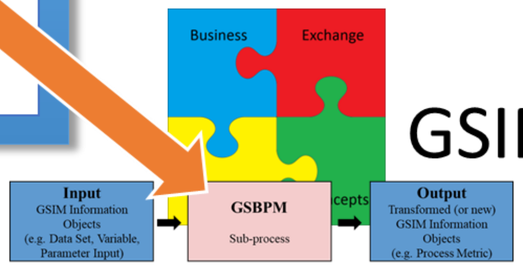






CSPA


CSDA

GSIM




- Modernization of Statistics is a broad concept.
- It must be addressed attending different aspects and focused on specific goals.
- One standard is not enough to deal with all the aspects.
- A coherent set of them must be established.

# High Level View of the Models Ecosystem

# Examples of how these standards and models are being used in the Official Statistics Offices to modernize the production of statistics

- Simplification/streamlining of procedures

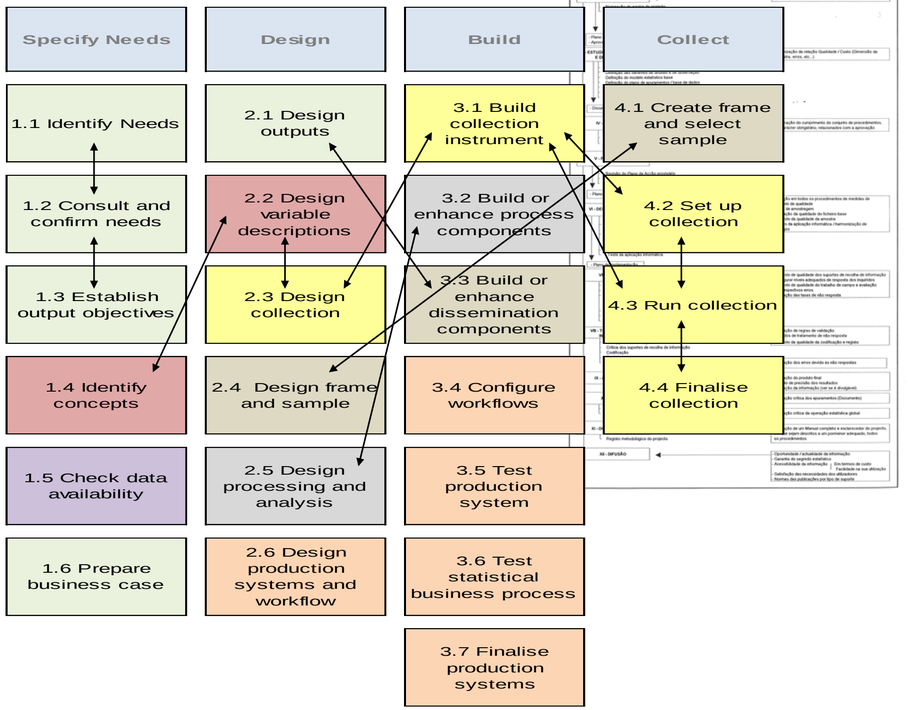
- Production and standardization of core reference documentation for the business process (work instructions, guidelines, templates)

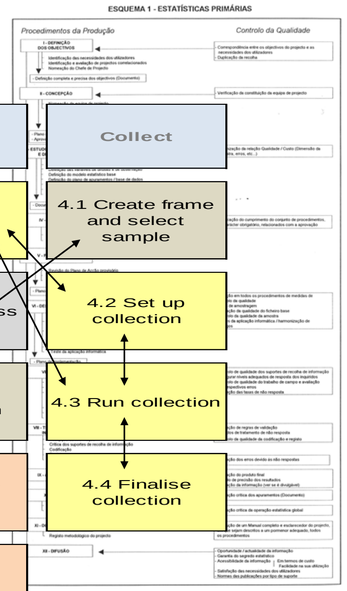
INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

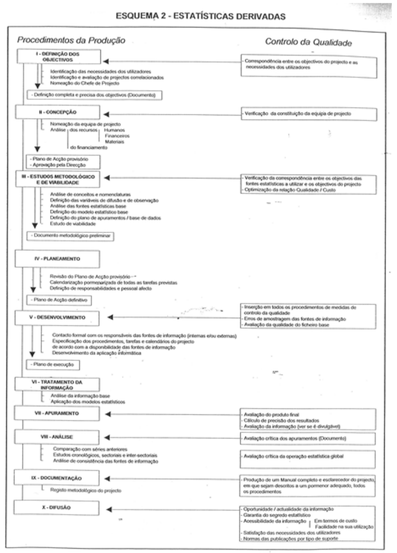
**Description of the Business Model**

Primary statistics

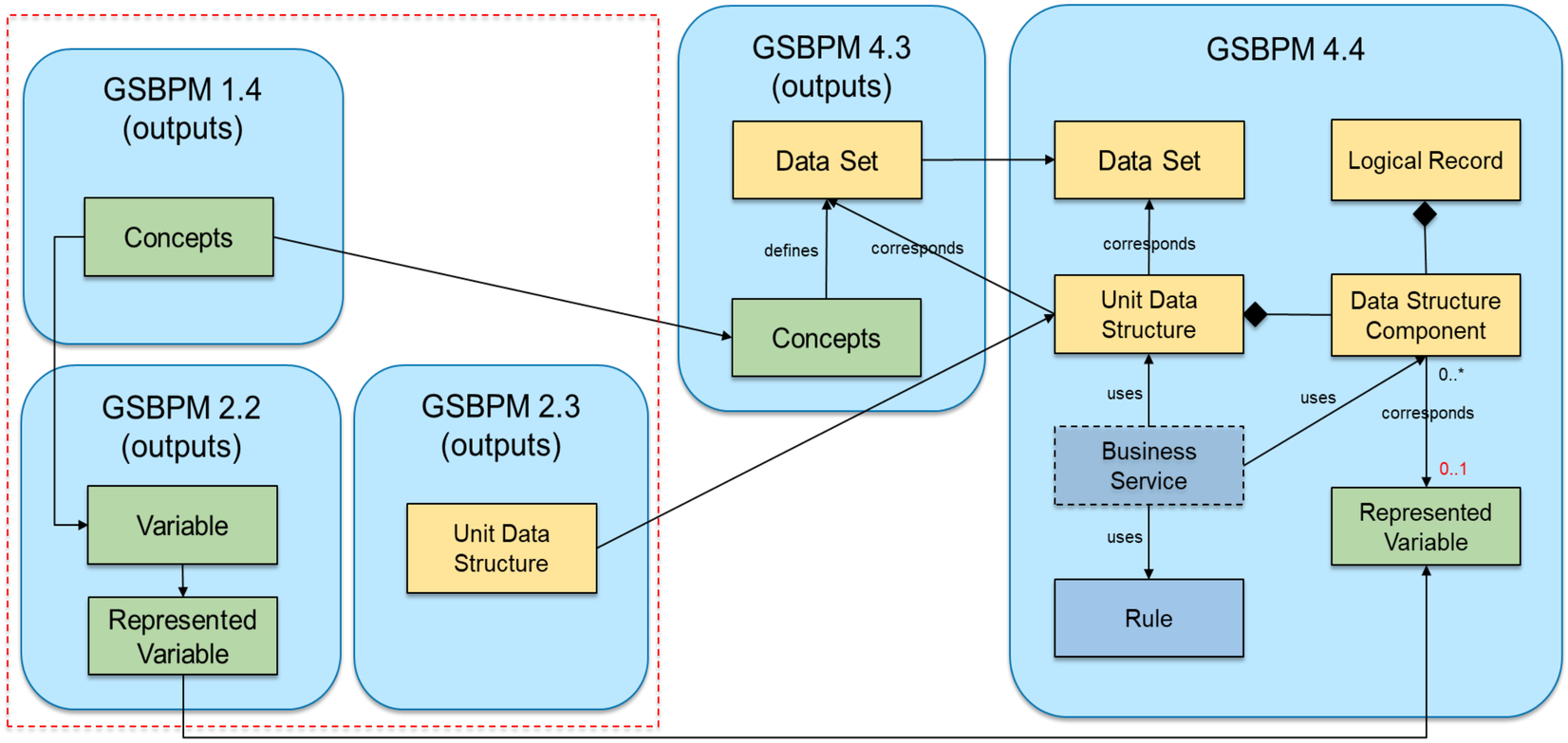
Derived statistics



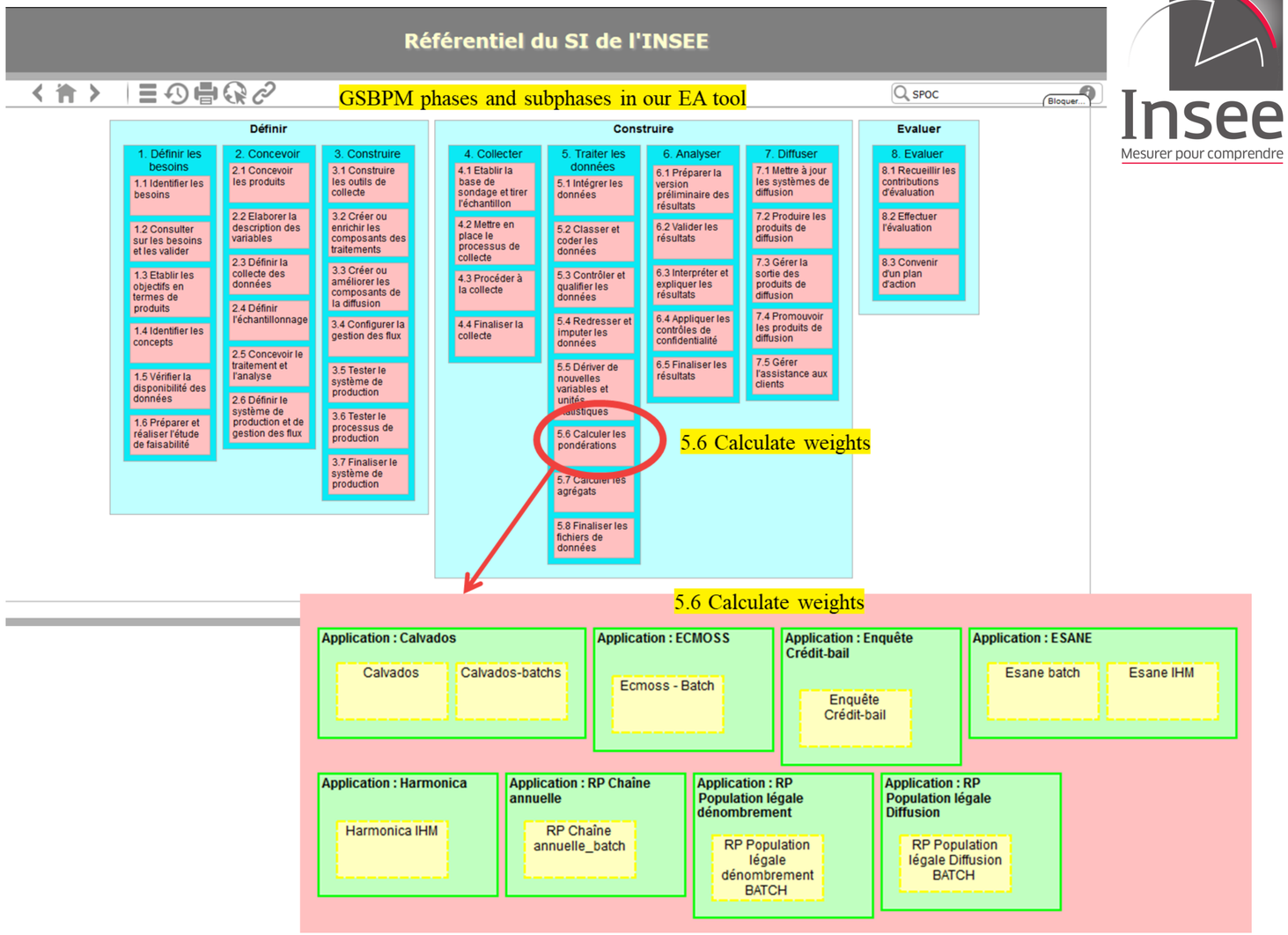



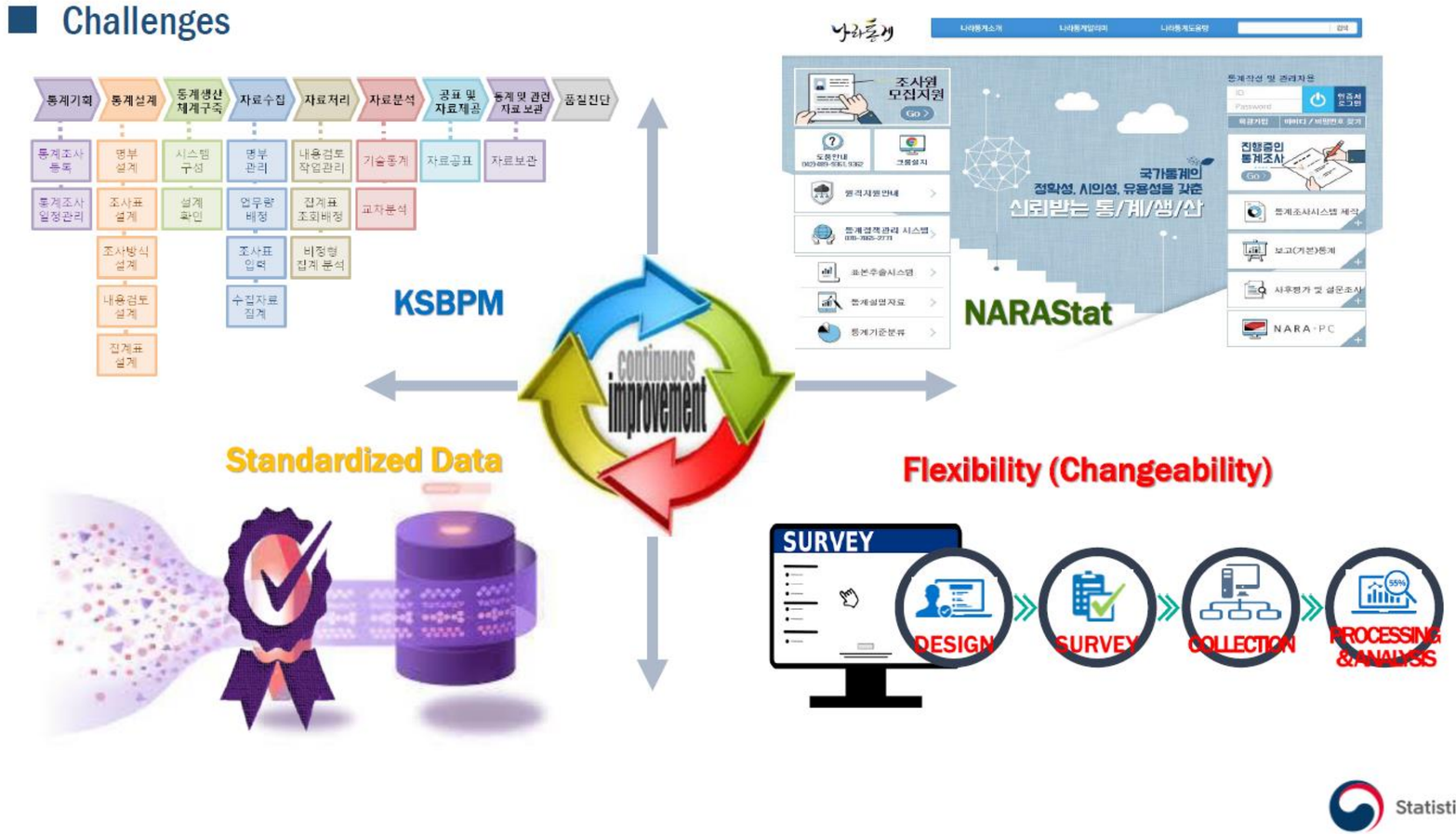
- Metadata Driven
- Fully Automated Process
- Service Oriented. Data processing is based on micro services (CSPA aspect)
- Covers different data collection, methods. Registers, Web data collections, surveys
- Based on Data Lake Zones
- Transparent Process. All tasks gather process metric (in respect to CSPA LIM)



- GSBPM is used for process review
- IT applications are mapped to GSBPM
- A metadata repository is designed to describe metadata related to each phase of statistical operations.
- IT applications are optimized to avoid duplications and to increase re-using.



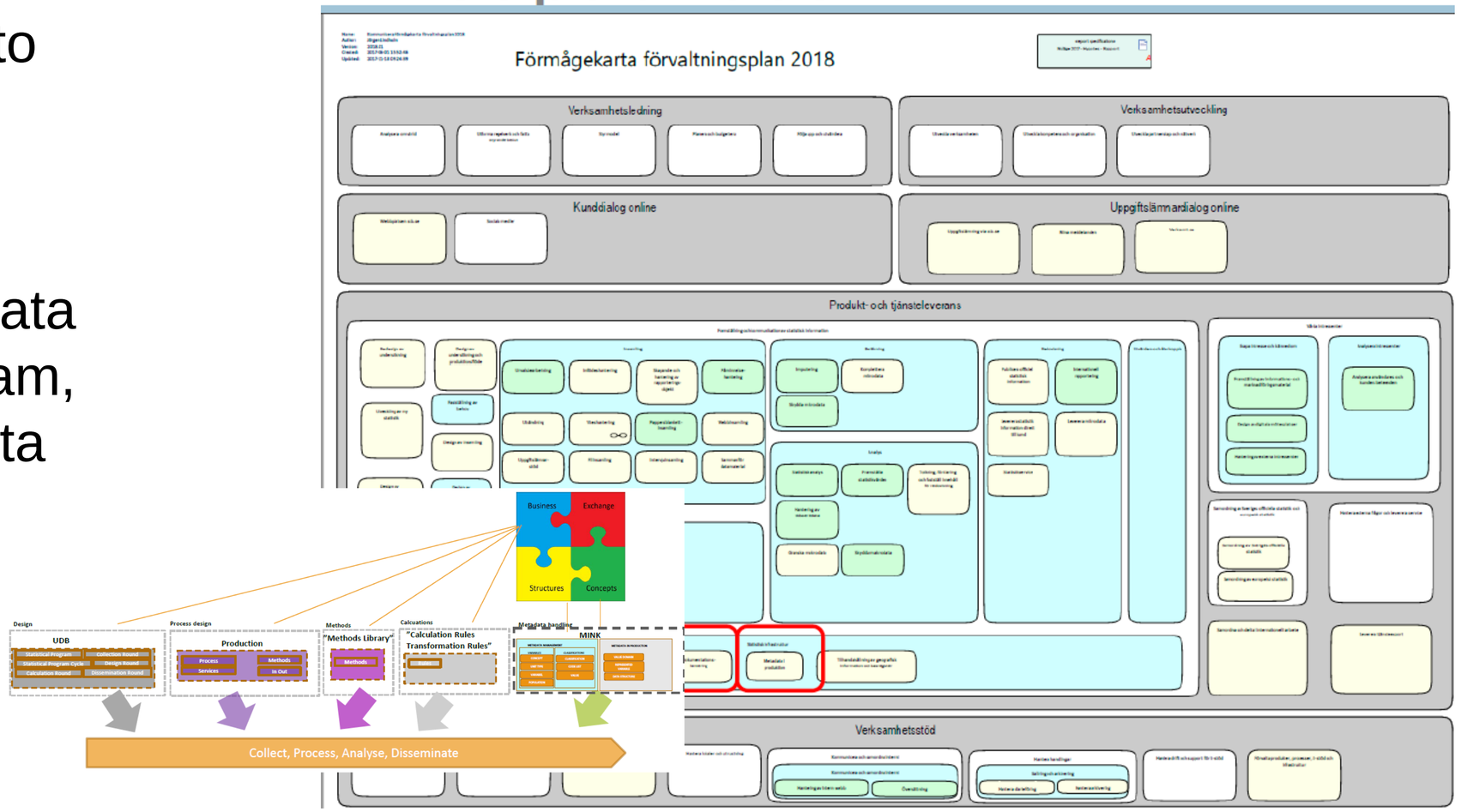
- South Korea uses GSBPM as the basis to design the statistical production in its national statistical system
- One system is used by many institutions to produce statistics
- Improves timeliness and quality




# KOREA: GSBPM to drive survey process

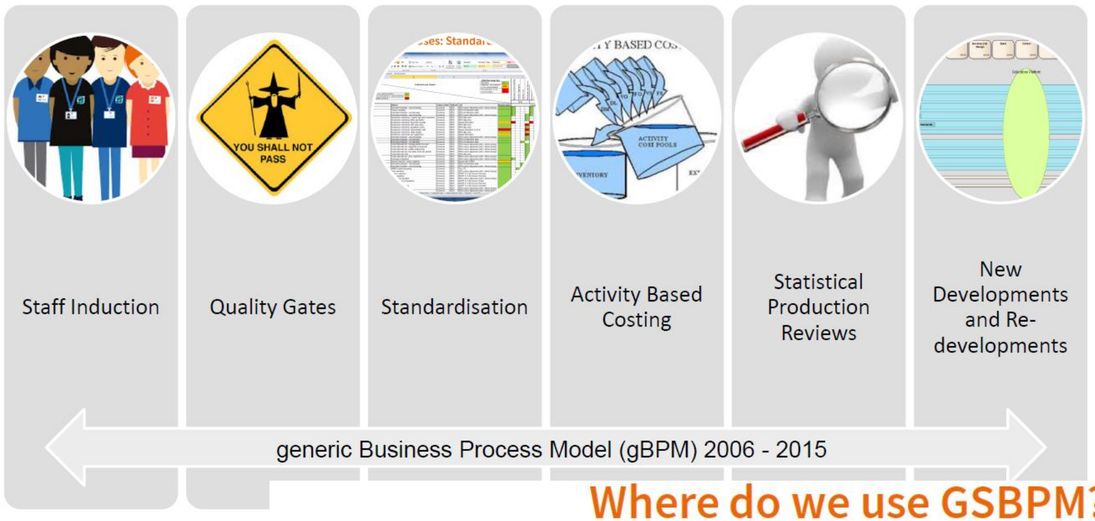
- Use of GSBPM and GSIM
- Metadata Management to improve  comparability
- Use of Metadata in Production to link metadata with the statistical program, in a specific time to a data structure.

# Capabilities



- The roots of GSBPM are in the New Zealand's generic Business Process Model.
- Organizational flexibility and agile production.
- Free up statistical analysts and methodologists for in-depth complex analysis
- The focus is on adding value to data.

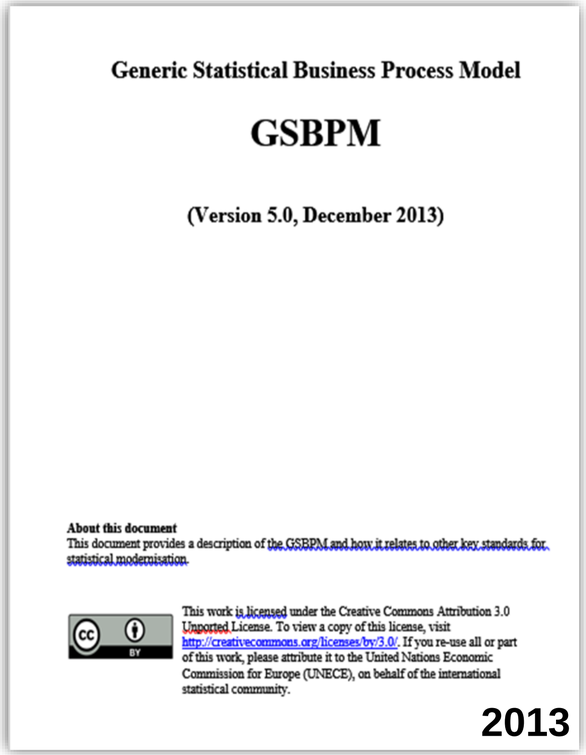
## Where did/do we use it?




## Where do we use GSBPM?


# NEW ZEALAND: Extensive Use of GSBPM

## GSBPM 5.0



Generic Statistical Business Process Model

GSBPM

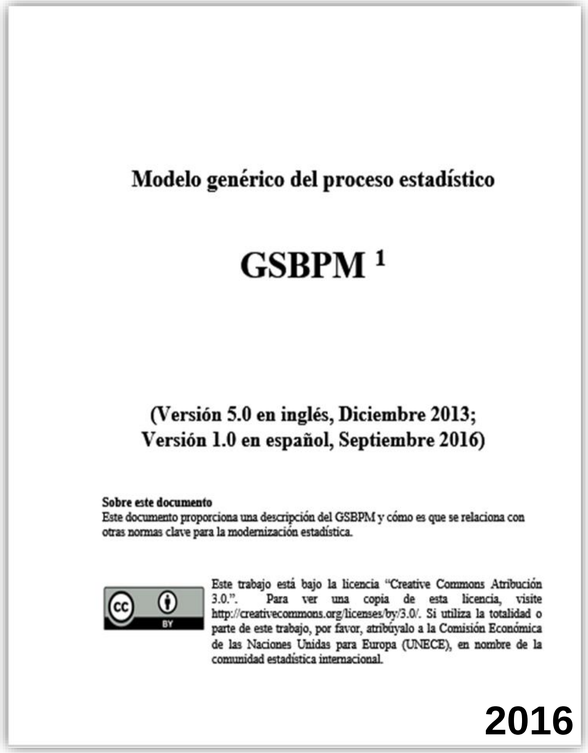
(Version 5.0, December 2013)

About this document
This document provides a description of the GSBPM and how it relates to other key standards for statistical modernisation.

This work is licensed under the Creative Commons Attribution 3.0 Unported License. To view a copy of this license, visit http://creativecommons.org/licenses/by/3.0/. If you re-use all or part of this work, please attribute it to the United Nations Economic Commission for Europe (UNECE), on behalf of the international statistical community.

2013

## GSBPM 5.0 Spanish



Modelo genérico del proceso estadístico

GSBPM [1]

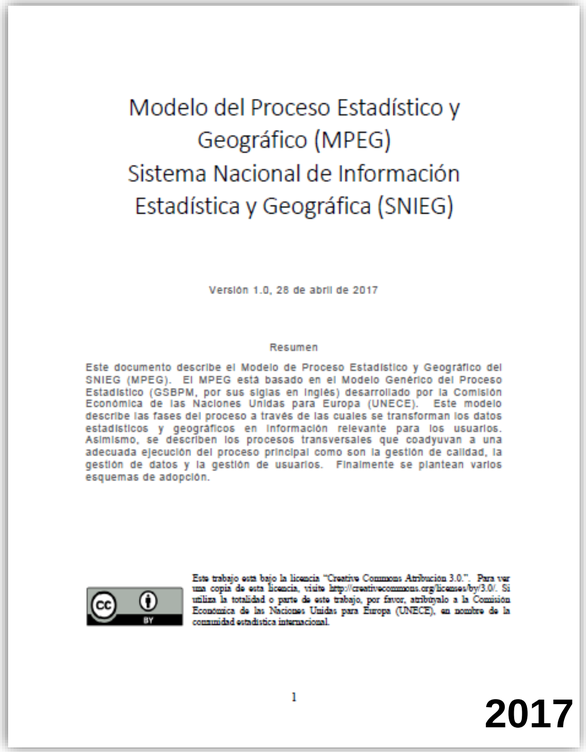
(Versión 5.0 en inglés, Diciembre 2013;
Versión 1.0 en español, Septiembre 2016)

Sobre este documento
Este documento proporciona una descripción del GSBPM y cómo es que se relaciona con otras normas clave para la modernización estadística.

Este trabajo está bajo la licencia "Creative Commons Atribución 3.0.". Para ver una copia de esta licencia, visite http://creativecommons.org/licenses/by/3.0/. Si utiliza la totalidad o parte de este trabajo, por favor, atribúyalo a la Comisión Económica de las Naciones Unidas para Europa (UNECE), en nombre de la comunidad estadística internacional.

2016

## MPEG



Modelo del Proceso Estadístico y Geográfico (MPEG)
Sistema Nacional de Información Estadística y Geográfica (SNIEG)

Versión 1.0, 28 de abril de 2017

Resumen

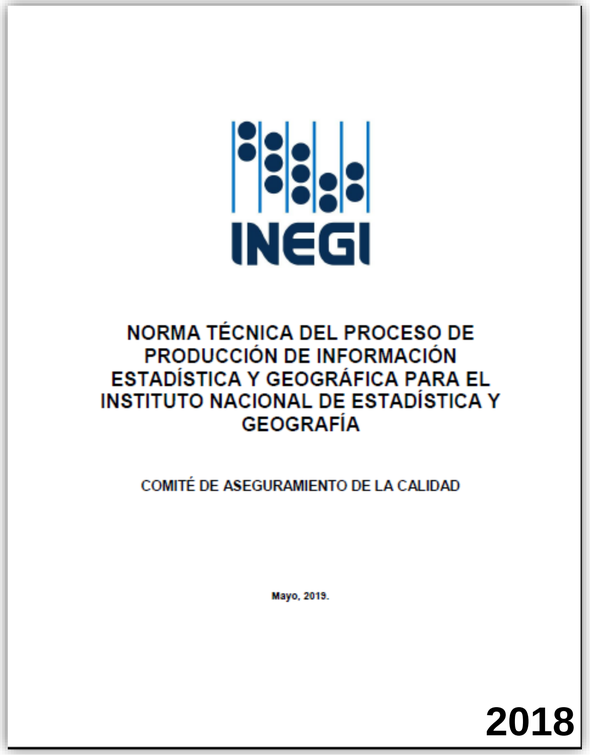
Este documento describe el Modelo de Proceso Estadístico y Geográfico del SNIEG (MPEG). El MPEG está basado en el Modelo Genérico del Proceso Estadístico (GSBPM, por sus siglas en inglés) desarrollado por la Comisión Económica de las Naciones Unidas para Europa (UNECE). Este modelo describe las fases del proceso a través de las cuales se transforman los datos estadísticos y geográficos en información relevante para los usuarios. Asimismo, se describen los procesos transversales que coadyuvan a una adecuada ejecución del proceso principal como son la gestión de calidad, la gestión de datos y la gestión de usuarios. Finalmente se plantean varios esquemas de adopción.

Este trabajo está bajo la licencia "Creative Commons Atribución 3.0.". Para ver una copia de esta licencia, visite http://creativecommons.org/licenses/by/3.0/. Si utiliza la totalidad o parte de este trabajo, por favor, atribúyalo a la Comisión Económica de las Naciones Unidas para Europa (UNECE), en nombre de la comunidad estadística internacional.

1

2017

## Technical Standard for the Process of Production of Statistical and Geographical Information for INEGI



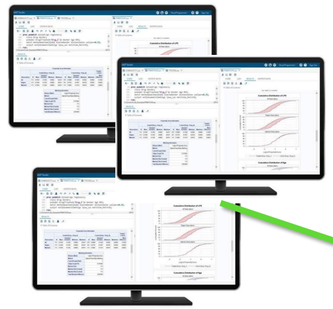
INEGI

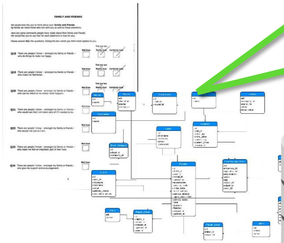
NORMA TÉCNICA DEL PROCESO DE PRODUCCIÓN DE INFORMACIÓN ESTADÍSTICA Y GEOGRÁFICA PARA EL INSTITUTO NACIONAL DE ESTADÍSTICA Y GEOGRAFÍA

COMITÉ DE ASEGURAMIENTO DE LA CALIDAD

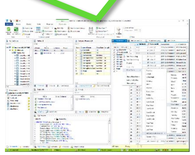
Mayo, 2019.

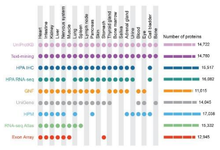
2018

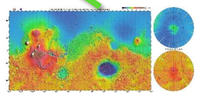



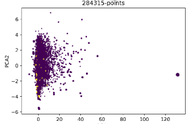


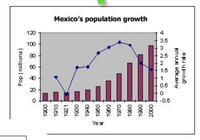


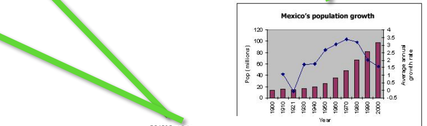






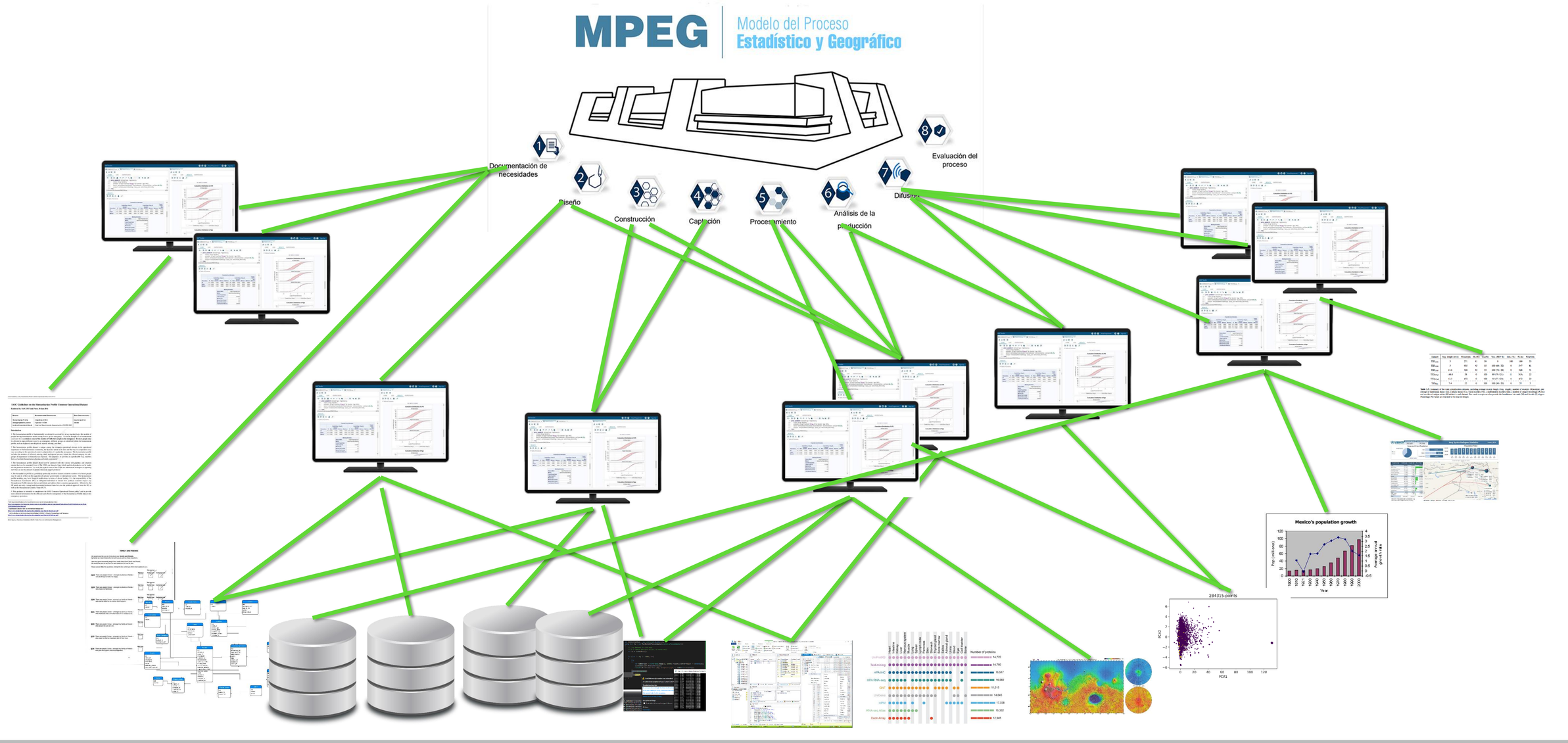




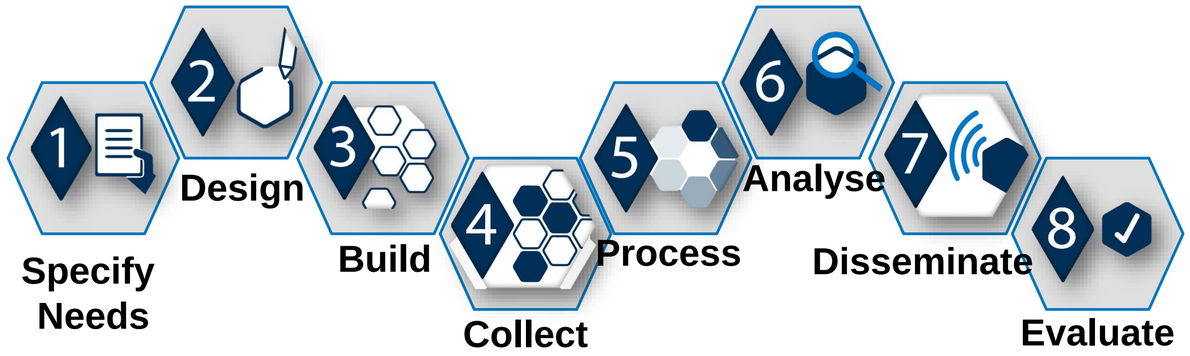




INEGI

# Multiple Information Systems

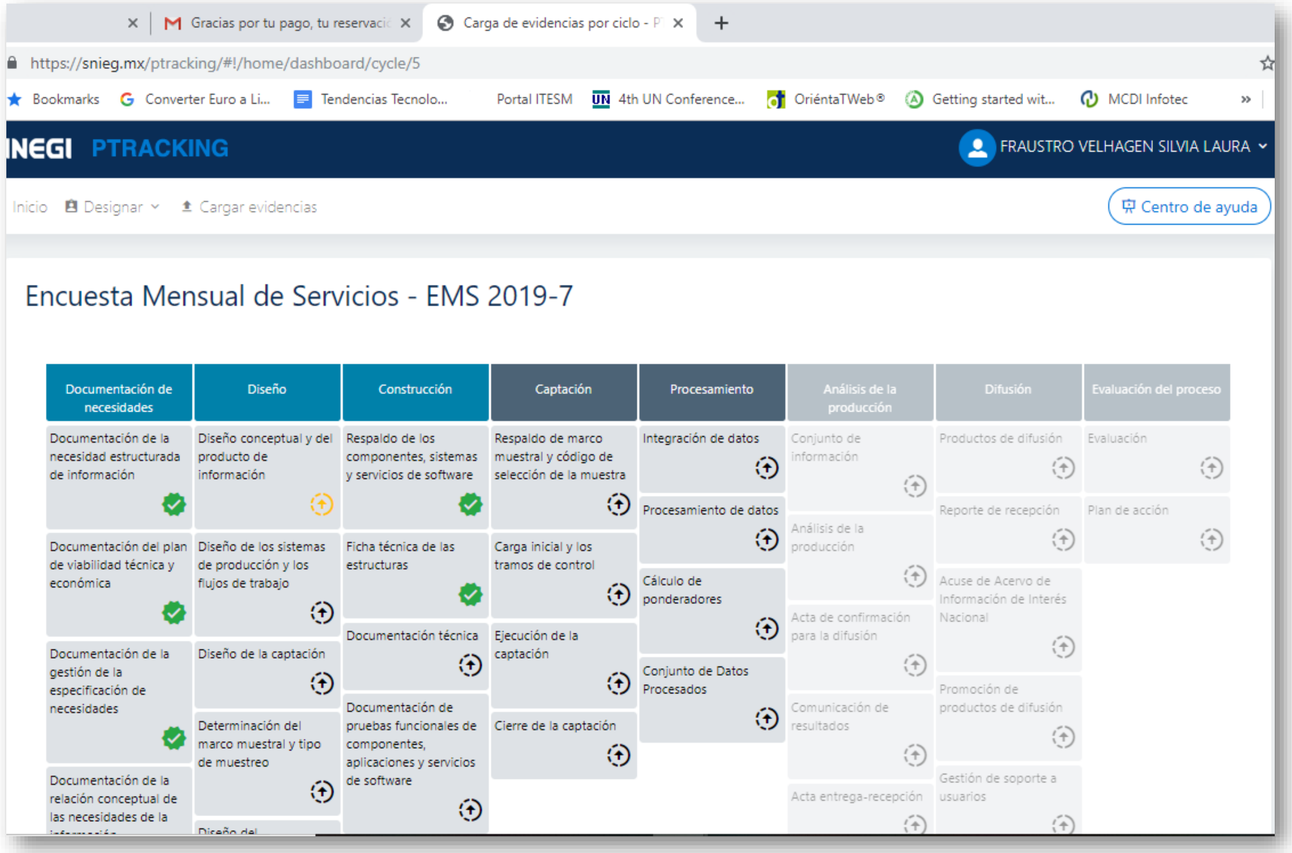



# Association of Evidences to the Process Activities

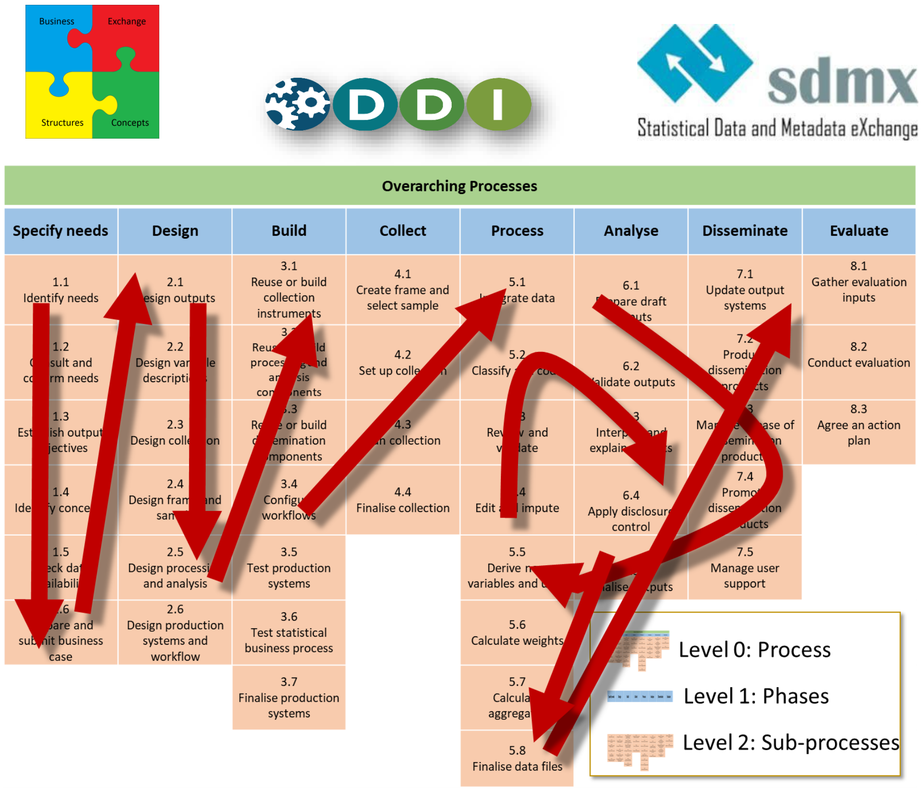



GSIM has been used to map the process to the actual practices, like:

- Statistical Program
- Cycle
- Phase



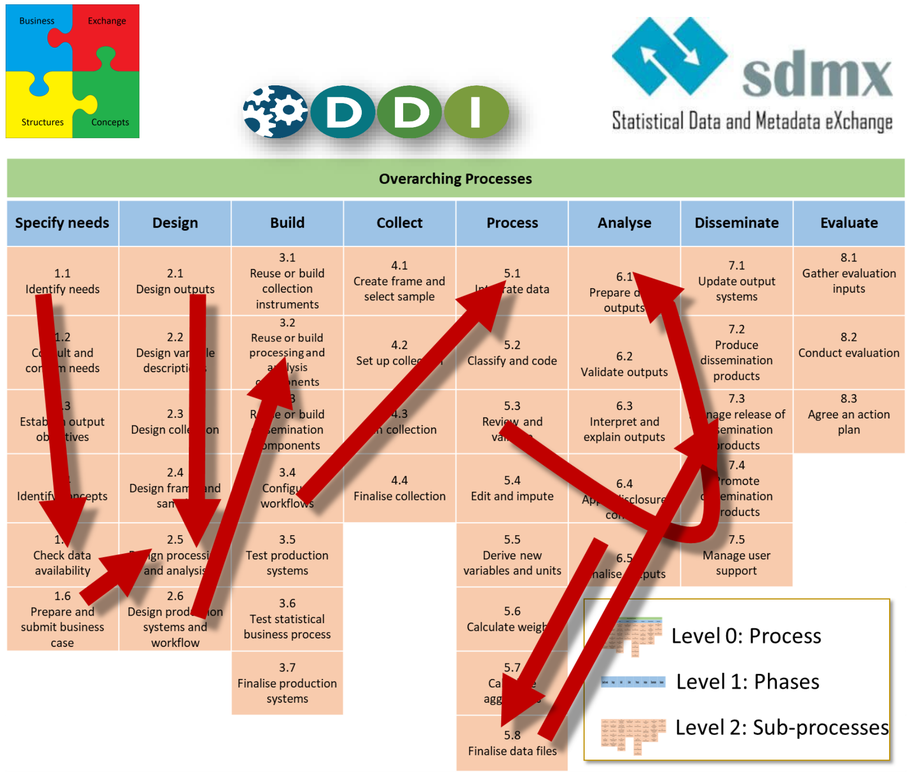

# Importance for Transparency and Reproducibility of Official Statistics




**Transparency**: We can show what is made on each sub-process.

**Reproducibility**: We can provide precise documentation of the processes and methodologies.

Outcomes may be compared too under the same terms




**Country A Survey X**

**Country A Survey Z**

**Country B Survey X**

# The way forward

- Maturation in the use of the standards
- Wider adoption
- Integration of the environment
- Creation of supporting tools
- Geo-referencing statistical information
- Use of new technologies and new data sources (p. e. AI and Big Data)
- Everywhere Statistics

# Questions to the audience

# General Discussion

Conociendo
# México


01 800 111 46 34
www.inegi.org.mx
atencion.usuarios@inegi.org.mx


INEGI Informa

- All the logos are property of their respective owners/institutions/organizations/companies
- Some examples and illustrations took from presentations made by the respective institutions
- Illustration and Photos took from the Internet, may be subject to copyright, contact the source for more information about reproduction rights
- Site of HLG-MOS: https://statswiki.unece.org/display/hlgbas
- Site of SDMX: https://sdmx.org/
- Site of DDI: https://ddialliance.org/
- This presentation is not for profit purposes, is not for being sell

# Disclousures