

SDMX in the Standards Landscape

Zoltán Vereczkei

UNECE Supporting Standards Group

10th Expert Group Meeting on Statistical Data and Metadata eXchange

January 25-28, 2021

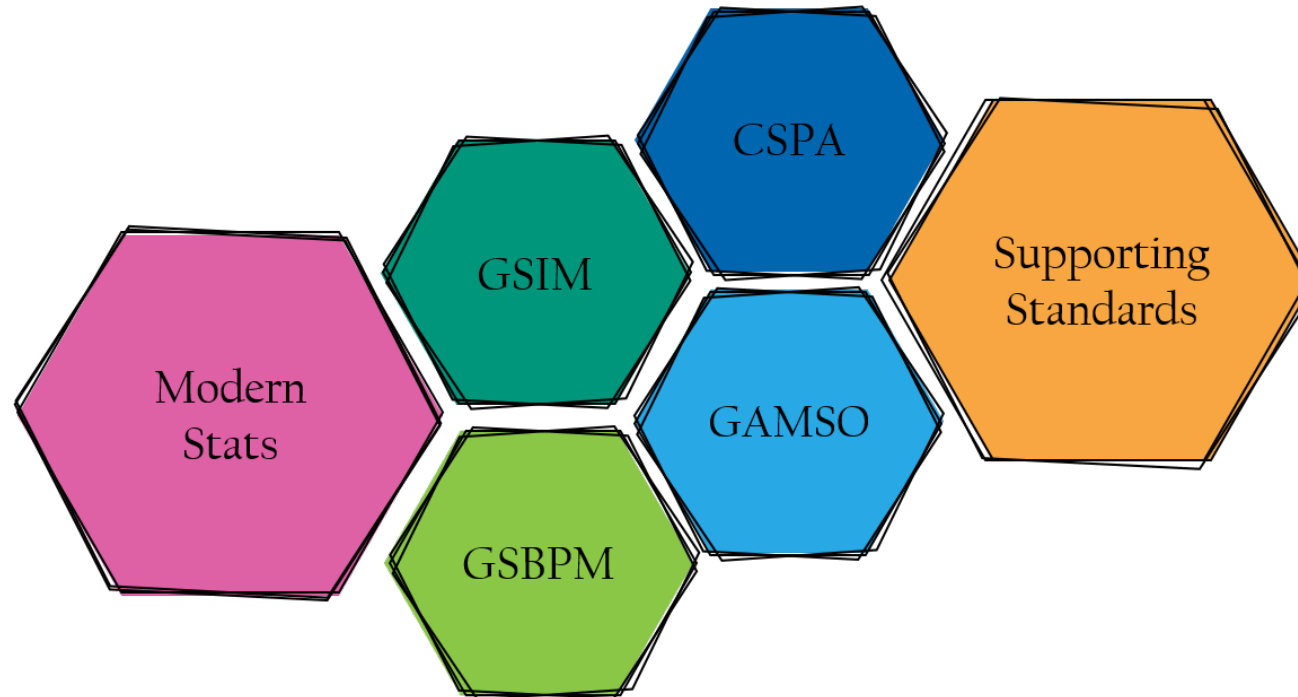
Outline...

- Standards in general, statistical potential
- UNECE Supporting Standards Group & ModernStats models
- SDMX as an important statistical standard
- Architecture layers & SDMX
- Standards working together: integration
- Relevant ModernStats work
- Integrating SDMX with other standards: potential
- Questions for you

Standards and their statistical potential

- What is a standard?
 - ISO definition: „A standard is a document, established by consensus and approved by a recognized body, that provides for common and repeated use, rules, guidelines or characteristics for activities or their results, aimed at the achievement of the optimum degree of order in a given context.”
- Why are standards so important?
 - „Perfect fits” for standardisation goals
 - Providing common language: starting point for every framework
 - Practical solutions: methodologies, tools to support actual implementation
 - Success stories, good practices, experience sharing; neverending modernisation
- Why are standards relevant to official statistics?
 - Developing, producing and disseminating statistics is a final goal for every statistical system all around the world
 - Common mission? Common frameworks? Common standards?

UNECE Supporting Standards Group



Source: <https://statswiki.unece.org>

- develop,
- enhance,
- integrate,
- promote,
- support and facilitate implementation of

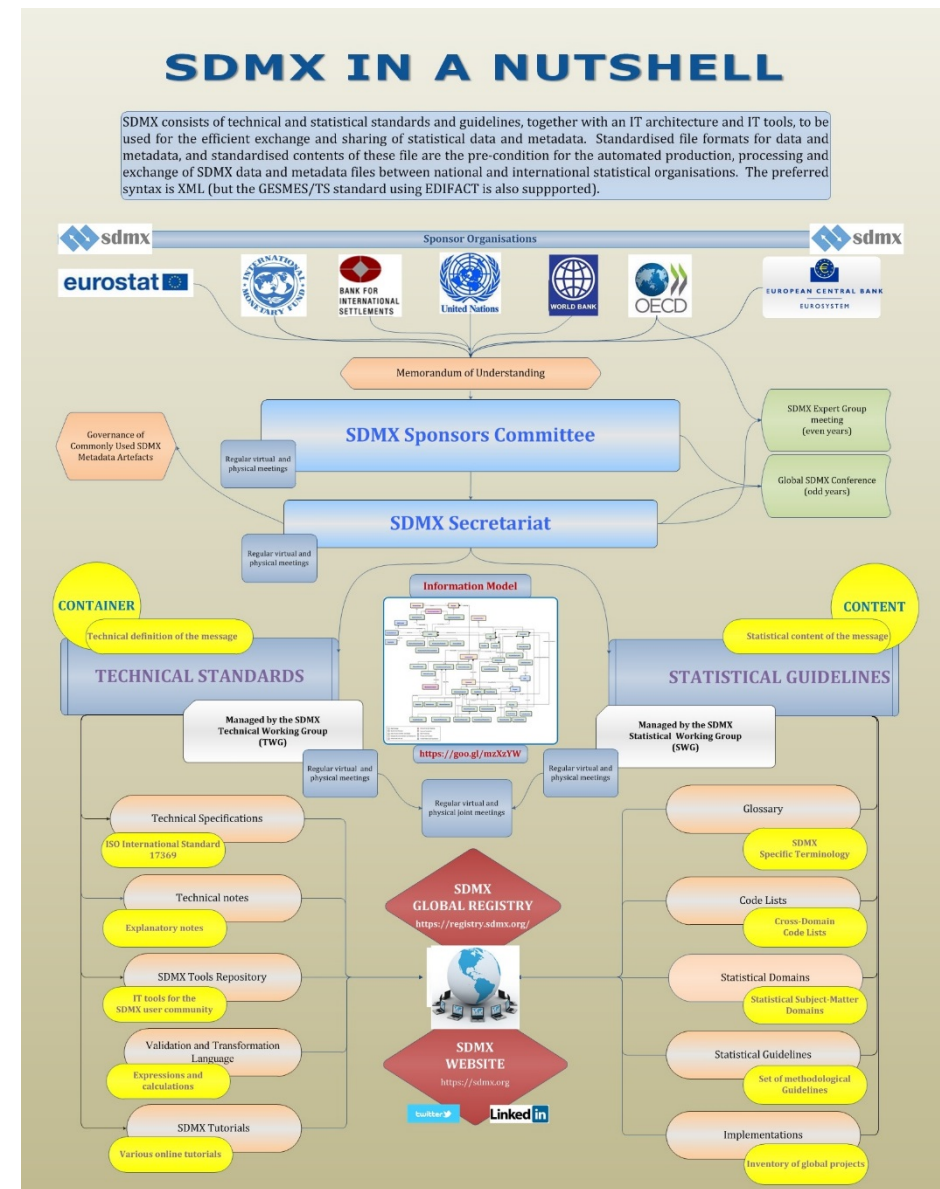
the range of standards needed for statistical modernisation

In focus: SDMX (reminder)

SDMX is a standard ([ISO standard, 17369:2013](#)) to describe statistical data and metadata; to normalise their exchange, and to enable them to be shared more efficiently among organisations.

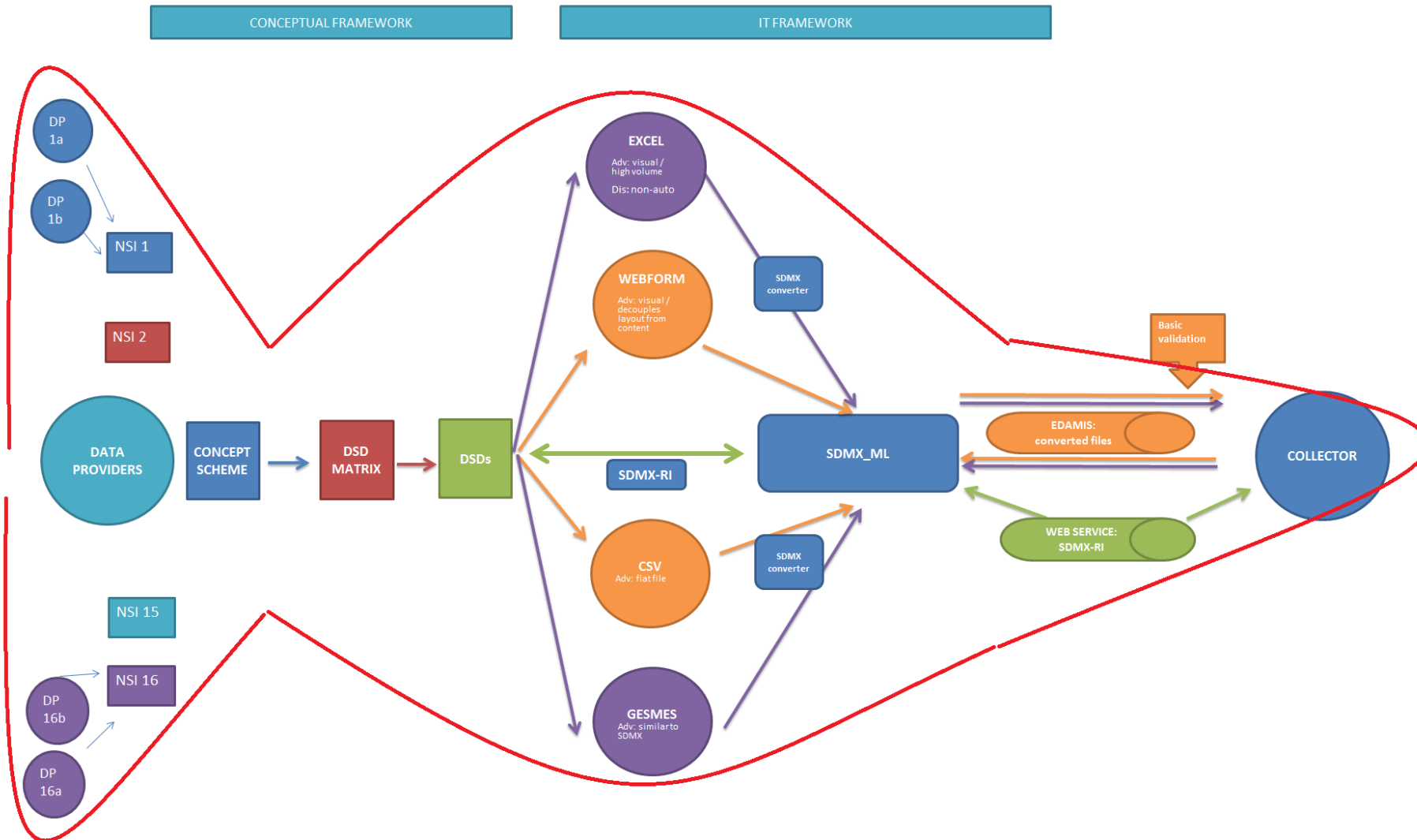
To meet these three requirements, SDMX has three key components:

- a set of standards, including guidelines for
 - Coding of statistical datasets
 - Best practices and guidelines for modelling, exchange
 - Technical implementation notes
- Technical standards (metadata registry, information model, etc.)
- Reusable tools

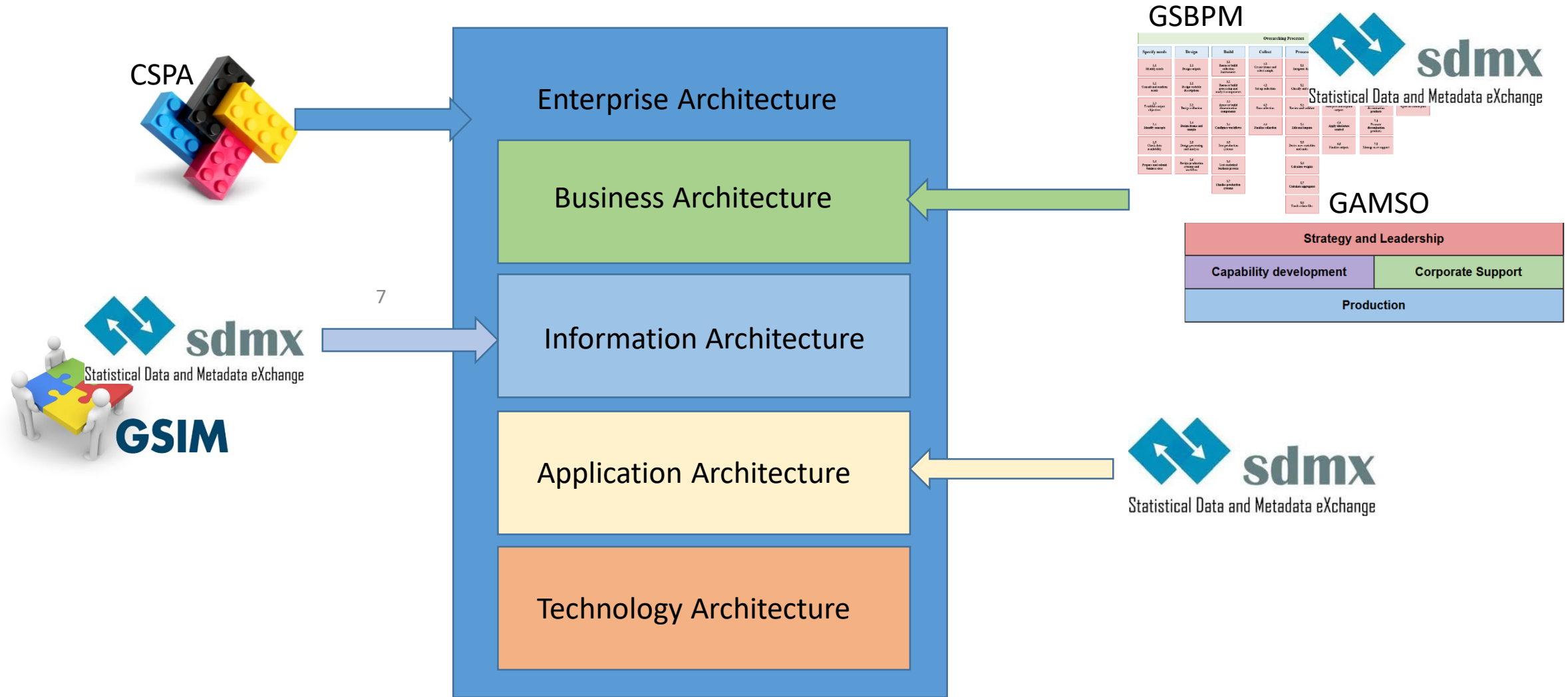


Source: https://sdmx.org/wp-content/uploads/SDMX_map_3_0.jpg

SDMX as a statistical standard



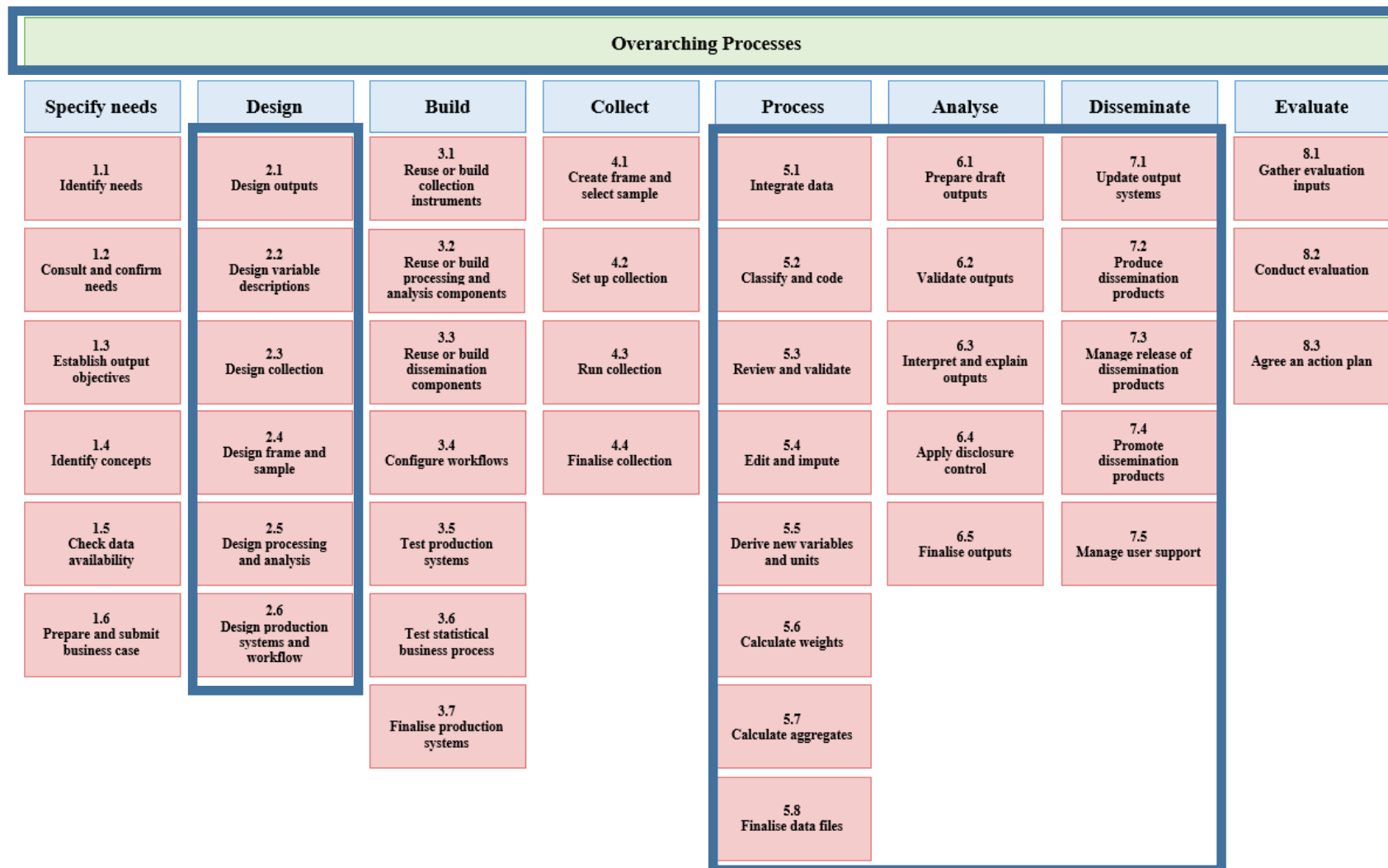
SDMX & ModernStats standards



Enterprise Architecture & SDMX

- CSPA is a reference architecture for our statistical community.
- An important concept in architecture is the "separation of concerns". For that reason, the architecture is separated into a number of "perspectives". These "perspectives" are:
 - **Business Architecture** which defines what the industry does and how it is done (statistics in our case),
 - **Information Architecture** which describes the information, its flows and uses across the industry, and how that information is managed,
 - **Application Architecture** which describes the set of practices used to select, define or design software components and their relationships, and
 - **Technology Architecture** which describes the infrastructure technology underlying (supporting) the other architecture perspectives.

Business Architecture & SDMX



Information Architecture & SDMX

- GSIM: It provides a **set of standardised, consistently described information objects**, which can be used as inputs and outputs in the design and the production of statistics.
- SDMX Content-Oriented Guidelines (Statistical standards)
 - **SDMX Glossary Version 2.0**: The Glossary provides definition of terms found in the SDMX Information Model, Data Structure Definitions (DSDs), and Metadata Structure Definitions (MSDs).
 - **SDMX cross-domain code lists**: Code lists are predefined sets of terms from which some statistical coded concepts take their values. The use of common code lists help users to work even more efficiently, easing the maintenance of and reducing the need for mapping systems and interfaces delivering data and metadata to them.
 - **Classification of Statistical Subject-Matter Domains**: A statistical activity that has common characteristics with respect to variables, concepts and methodologies for data collection and the whole statistical data compilation process.

Application Architecture & SDMX

- SDMX IT tools – for data and metadata sharing

<https://ec.europa.eu/eurostat/web/sdmx-infospace/sdmx-it-tools>

https://sdmx.org/?page_id=4500

- SDMX Technical Specification – applications

https://sdmx.org/?page_id=5008

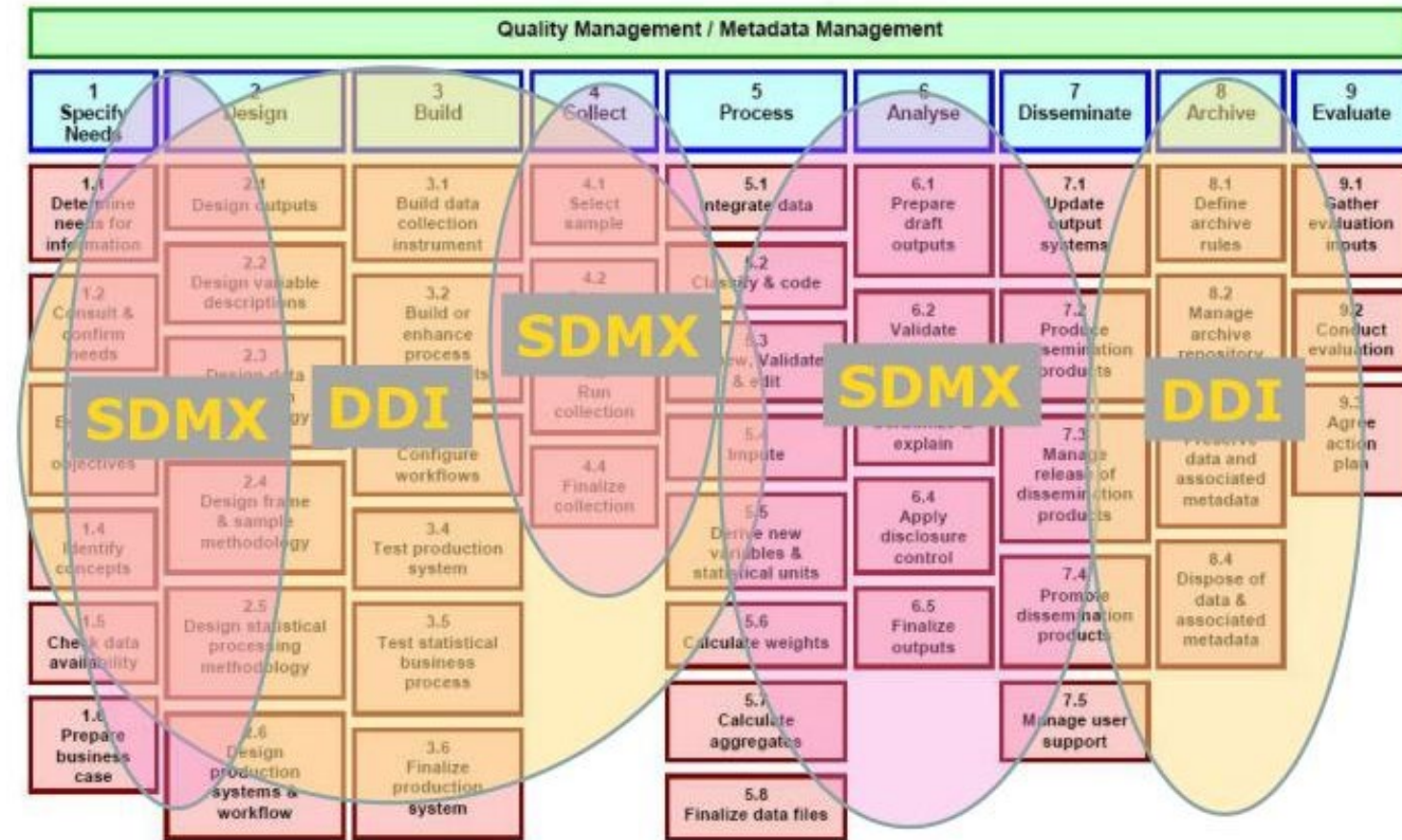
Technology Architecture

- No real specific standard at the moment?
- Technology is very specific.

SDMX & standards outside of ModernStats

DDI

a cooperative standard

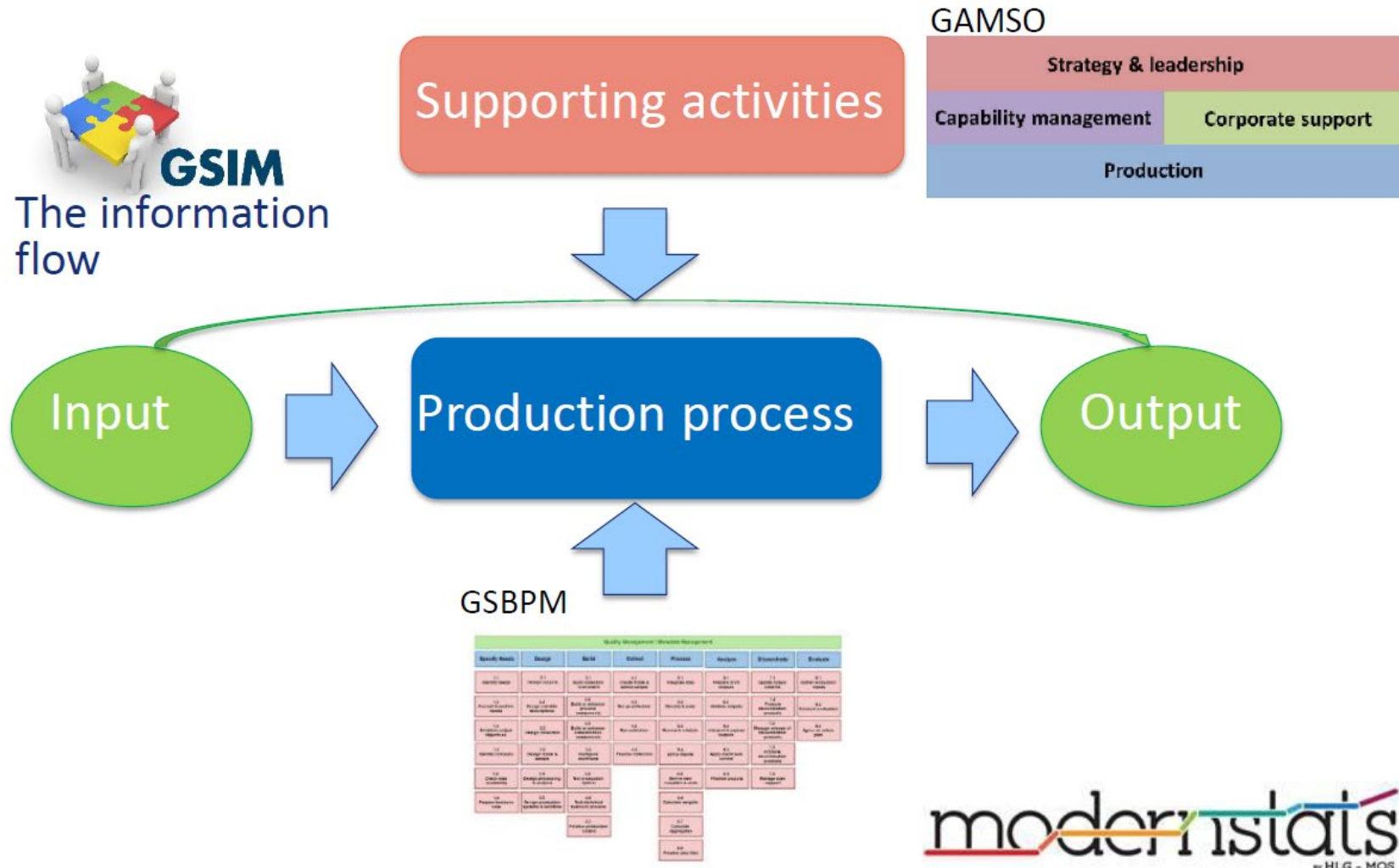


Based on source: UNECE: Exploring the relationship between DDI, SDMX and the Generic Statistical Business Process Model

SDMX & standards outside of ModernStats

- Geospatial standards
 - The geospatial community uses the General Feature Model (GFM) and developed the ISO19115 metadata standard, plus a number of application-specific standards.
 - Geospatial and statistical information are getting closer: integrating geospatial information is important for ModernStats as well: Task Team ongoing for the application of GSBPM for geospatial information
- Technical standards (technical formats)
 - CSV, JSON, XML, RDF
 - GESMES
 - XBRL
 - etc.

Standards working together: integration



Integrating standards: improving architectures

- Standards are not independent of each other, they are connected; most benefit is realised when this integration potential is recognised.
- No ultimate „best way” on what standards to adopt first. But!
 - A very clear understanding of our core business (producing statistics) is needed, no matter where we go with standardisation (GSBPM)
 - A generic process model („world” around the statistical production) is usually a next step (as this will provide the full business architectural framework) (GAMSO)
 - Further specification of the core business usually results in information objects thus the implementation of an information standard (GSIM / SDMX / CSPA LIM)
 - Organising the whole enterprise architecture in one clear framework might be the ultimate goal (CSPA)

Integrating standards: improving architectures

- Different national statistical systems have different maturity to make the next step towards better architectures
- Maturity models and definition of roadmaps help to improve the enterprise architecture or different layers of it

Integrating standards: assessment of maturity

- Modernisation Maturity Model (MMM) as a self-assessment tool to assess the maturity of the organisation against different criteria and dimensions.
- Could this be useful for SDMX? Help to define a roadmap for implementation?

Matrix for GSBPM/GSIM/GAMSO/CSPA

	Initial awareness	Pre-implementation	Early implementation	Corporate implementation	Mature implementation
Business					
Methods					
Information					
Applications					
Technology					

Relevant ModernStats work



Especially interesting for the SDMX community:

- **Core Ontology for Official Statistics:** establishing common foundations for semantic standards; using OWL as a tool for the integration of models into a coherent and interoperable framework. Task Team to launch now.
- **Linking GSBPM and GSIM:** good example on how to put the business and information architecture models together. Task Team to finish work by the end of 2021.
- **GSIM Task Team:** GSIM soft update is currently ongoing. Task Team to finish work by the end of 2021.

Other Task Teams:

- **Defining and detailing GSBPM tasks:** Task Team to launch around June 2021
- **Application of GSBPM for Geospatial Information:** Task Team to conclude the work around April 2021.

SDMX: from standardisation perspective

- SDMX fits perfectly with the business / information / application architectures. This makes it linked to ModernStats, especially GSBPM, GSIM (cooperating standards):
 - SDMX & GSBPM: already mapped and positioning of SDMX in statistical production is usually represented using the GSBPM
 - SDMX & GSIM: it is a „natural” next step to find out how SDMX and GSIM can be better integrated (detailed mapping possible)

Questions

- What is your motivation to implement different standards in your organisation?
- In what cases (at what level) do you think it is important to integrate different standards for statistics?
- What would help you the most when you move to the actual implementation of standards (use cases, guidelines, etc.)?

Thank you for the attention!

 zoltan.vereczkei@ksh.hu