Generic Statistical Business Process Model

Carlo Vaccari

BiH June 2012

Contents

Background

GSBPM key features

Relevance to SDMX

Next steps

Linked initiatives:

- CORA/CORE ESSnet

- SAB (Sharing Advisory Board)

- GSIM

Thanks to Steve Vale (UNECE) for GSBPM presentation

Istat

❖ Defining and mapping business processes in statistical organisations started at least 10 years ago

- "Statistical value chain"
- "Survey life-cycle"
- "Statistical process cycle"
- "Business process model"
- …

Istat

❖ Defining and mapping business processes in statistical organisations started at least 10 years ago

- ~~"Statistical value chain"~~ **X**
- ~~"Survey life-cycle"~~ **X**
- ~~"Statistical process cycle"~~ **X**
- ~~"Business process model"~~ **X**

Generic Statistical Business Process Model

Istat

❖ Reached a stage of maturity where a generic international standard is the logical next step

❖ Many drivers for a common model:

- Metadata systems development
- Harmonization of terminology
- Software sharing
- Process-based organization structures
- Quality management requirements
- ...

Istat

❖ To define, describe and map statistical processes in a coherent way

❖ To standardize process terminology

❖ To compare / benchmark processes within and between organisations

❖ To identify synergies between processes

❖ To inform decisions on systems architectures and organisation of resources

Istat

❖ Based on the business process model developed by Statistics New Zealand

❖ Many modifications in three rounds of comments

❖ Terminology and descriptions made more generic

❖ Wider applicability?

Istat

❖All activities undertaken by producers of official statistics which result in data outputs

❖National and international statistical organisations

❖Independent of data source, can be used for:

- Surveys / censuses
- Administrative sources / register-based statistics
- Mixed sources

Istat

❖Producing statistics from raw data (micro or macro-data)

❖Revision of existing data / re-calculation of time-series
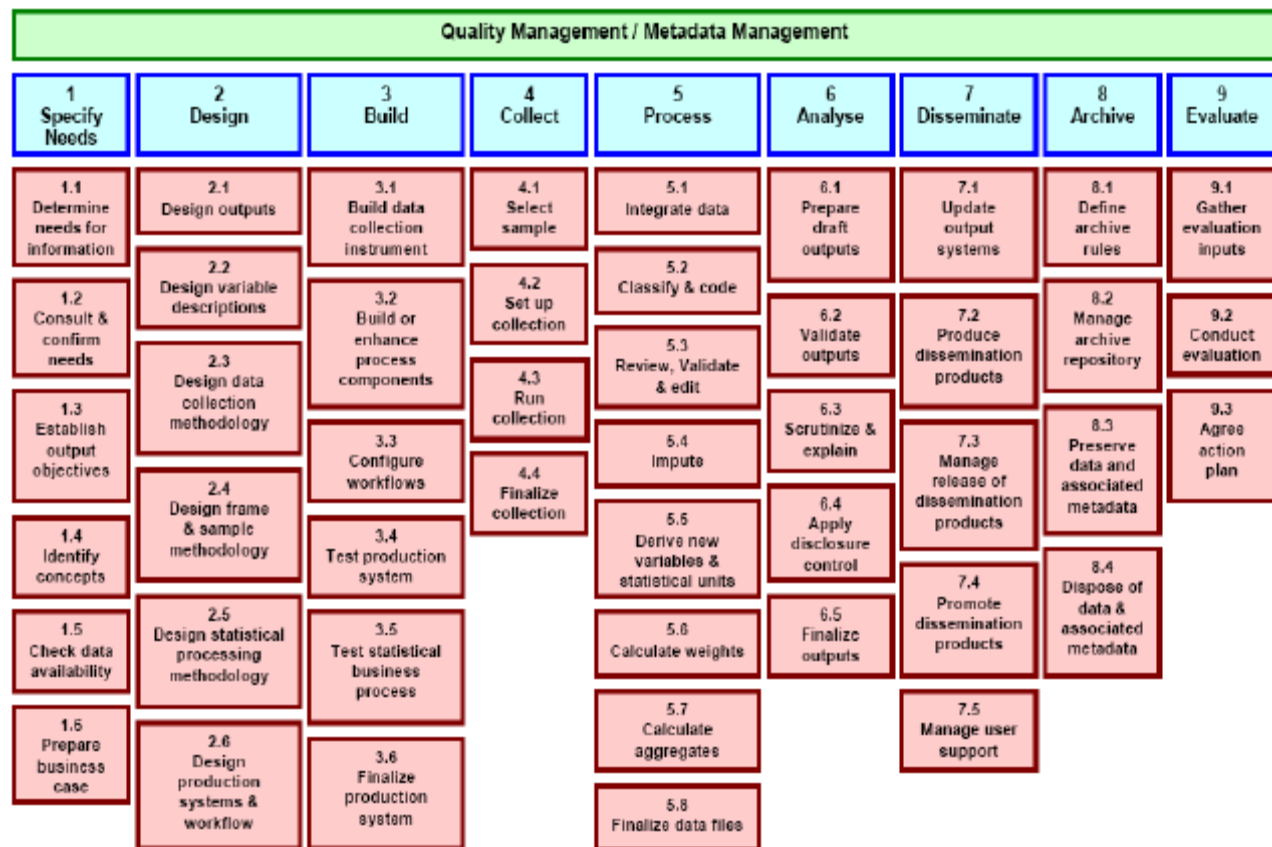
❖Development and maintenance of statistical registers

Process

Phases

Sub-processes

(Descriptions)

Quality Management / Metadata Management								
1 Specify Needs	2 Design	3 Build	4 Collect	5 Process	6 Analyse	7 Disseminate	8 Archive	9 Evaluate
1.1 Determine needs for information	2.1 Design outputs	3.1 Build data collection instrument	4.1 Select sample	5.1 Integrate data	6.1 Prepare draft outputs	7.1 Update output systems	8.1 Define archive rules	9.1 Gather evaluation inputs
1.2 Consult & confirm needs	2.2 Design variable descriptions	3.2 Build or enhance process components	4.2 Set up collection	5.2 Classify & code	6.2 Validate outputs	7.2 Produce dissemination products	8.2 Manage archive repository	9.2 Conduct evaluation
1.3 Establish output objectives	2.3 Design data collection methodology	3.3 Configure workflows	4.3 Run collection	5.3 Review, Validate & edit	6.3 Scrutinize & explain	7.3 Manage release of dissemination products	8.3 Preserve data and associated metadata	9.3 Agree action plan
1.4 Identify concepts	2.4 Design frame & sample methodology	3.4 Test production system	4.4 Finalize collection	5.4 Impute	6.4 Apply disclosure control	7.4 Promote dissemination products	8.4 Dispose of data & associated metadata	
1.5 Check data availability	2.5 Design statistical processing methodology	3.5 Test statistical business process		5.5 Derive new variables & statistical units	6.5 Finalize outputs	7.5 Manage user support		
1.6 Prepare business case	2.6 Design production systems & workflow	3.6 Finalize production system		5.6 Calculate weights				
				5.7 Calculate aggregates				
				5.8 Finalize data files				

Istat

Zoom!

Quality Management / Metadata Management

1 Specify Needs	2 Design	3 Build	4 Collect	5 Process	6 Analyse	7 Disseminate	8 Archive	9 Evaluate
1.1 Determine needs for information	2.1 Design outputs	3.1 Build data collection instrument	4.1 Select sample	5.1 Integrate data	6.1 Prepare draft outputs	7.1 Update output systems	8.1 Define archive rules	9.1 Gather evaluation inputs
1.2 Consult & confirm needs	2.2 Design variable descriptions	3.2 Build or enhance process components	4.2 Set up collection	5.2 Classify & code	6.2 Validate outputs	7.2 Produce dissemination products	8.2 Manage archive repository	9.2 Conduct evaluation
1.3 Establish output objectives	2.3 Design data collection methodology		4.3 Run collection	5.3 Review, Validate & edit	6.3 Scrutinize & explain	7.3 Manage release of dissemination products	8.3 Preserve data and associated metadata	9.3 Agree action plan
1.4 Identify concepts	2.4 Design frame & sample methodology	3.3 Configure workflows	4.4 Finalize collection	5.4 Impute	6.4 Apply disclosure control		8.4 Dispose of data & associated metadata	
1.5 Check data availability	2.5 Design statistical processing methodology	3.4 Test production system		5.5 Derive new variables & statistical units	6.5 Finalize outputs	7.4 Promote dissemination products		
1.6 Prepare business case	2.6 Design production systems & workflow	3.5 Test statistical business process		5.6 Calculate weights		7.5 Manage user support		
		3.6 Finalize production system		5.7 Calculate aggregates				
				5.8 Finalize data files				

Istat
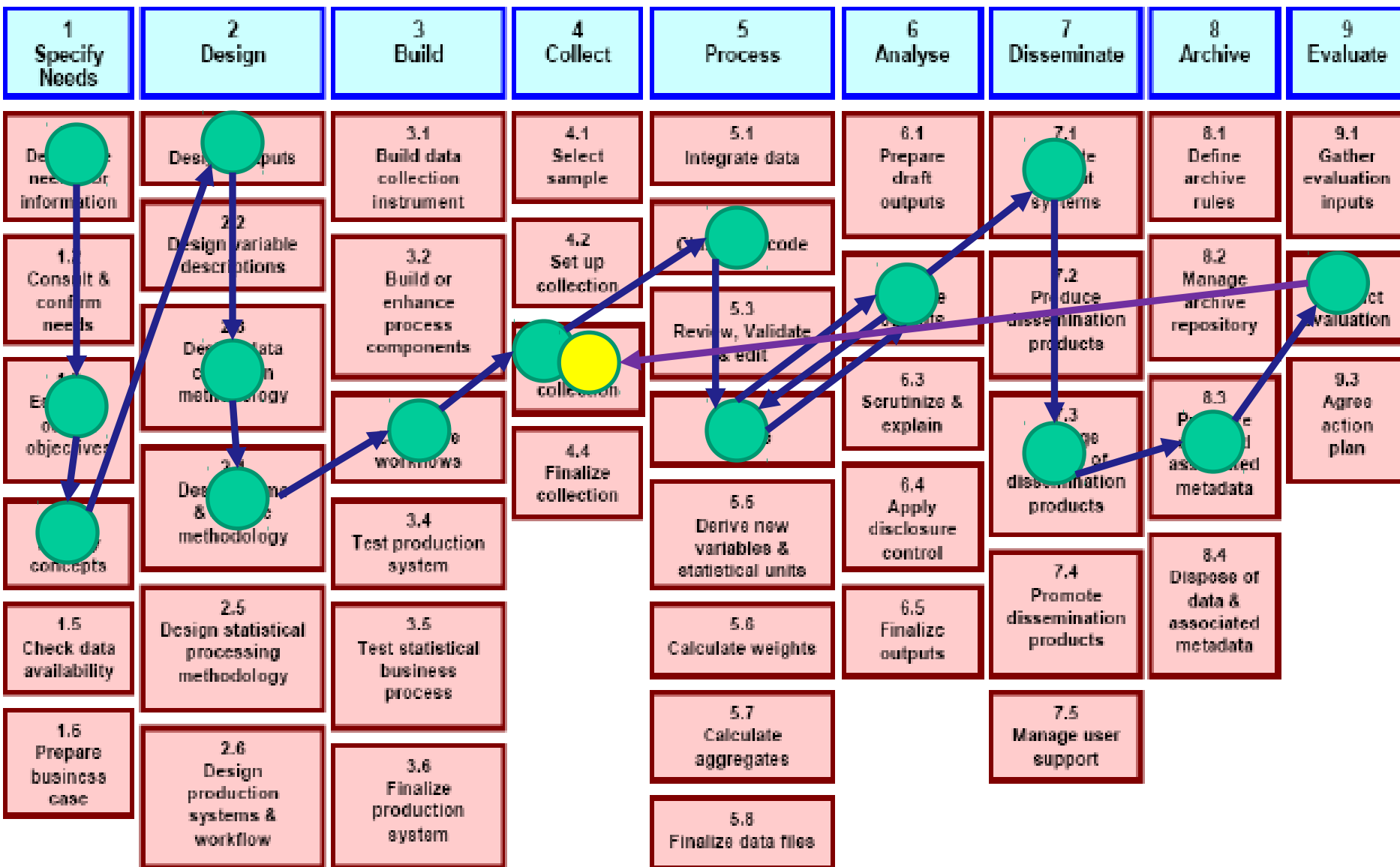
❖National implementations may need additional levels

❖Over-arching processes

- Quality management

- Metadata management

- Harmonizing statistical computing architectures

- Facilitating the sharing of statistical software

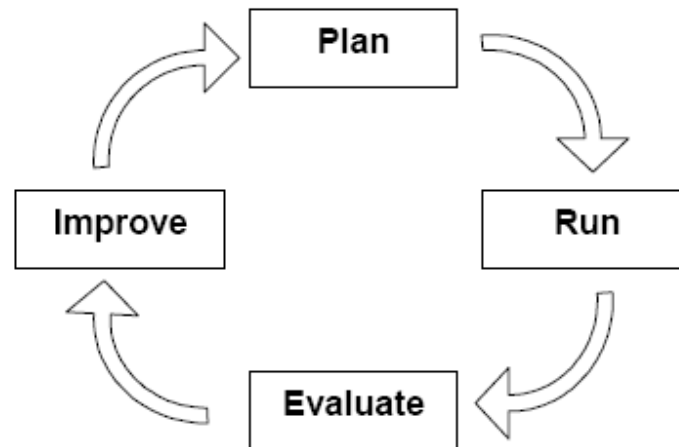- .... (8 more – see paper)

Istat

❖ **Not a linear model**

❖Sub-processes do not have to be followed in a strict order

❖It is a matrix, through which there are many possible paths, including iterative loops within and between phases

❖Some iterations of a regular process may skip certain sub-processes

Quality Management / Metadata Management

1 Specify Needs	2 Design	3 Build	4 Collect	5 Process	6 Analyse	7 Disseminate	8 Archive	9 Evaluate

1 Specify Needs
- Determine needs for information
- 1.2 Consult & confirm needs
- Establish output objectives
- concepts
- 1.5 Check data availability
- 1.6 Prepare business case

2 Design
- Design outputs
- 2.2 Design variable descriptions
- Design data collection methodology
- Design frame & sample methodology
- 2.5 Design statistical processing methodology
- 2.6 Design production systems & workflow

3 Build
- 3.1 Build data collection instrument
- 3.2 Build or enhance process components
- workflows
- 3.4 Test production system
- 3.5 Test statistical business process
- 3.6 Finalize production system

4 Collect
- 4.1 Select sample
- 4.2 Set up collection
- collection
- 4.4 Finalize collection

5 Process
- 5.1 Integrate data
- Classify & code
- 5.3 Review, Validate & edit
- Derive new variables & statistical units
- 5.6 Calculate weights
- 5.7 Calculate aggregates
- 5.8 Finalize data files

6 Analyse
- 6.1 Prepare draft outputs
- Validate outputs
- 6.3 Scrutinize & explain
- 6.4 Apply disclosure control
- 6.5 Finalize outputs

7 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

8 Archive
- 8.1 Define archive rules
- 8.2 Manage archive repository
- 8.3 Preserve data and associated metadata
- 8.4 Dispose of data & associated metadata

9 Evaluate
- 9.1 Gather evaluation inputs
- Conduct evaluation
- 9.3 Agree action plan

Istat

❖ In theory the model is circular:

- Evaluation can lead to modified needs and design

❖ In practice it is more like a multiple helix:

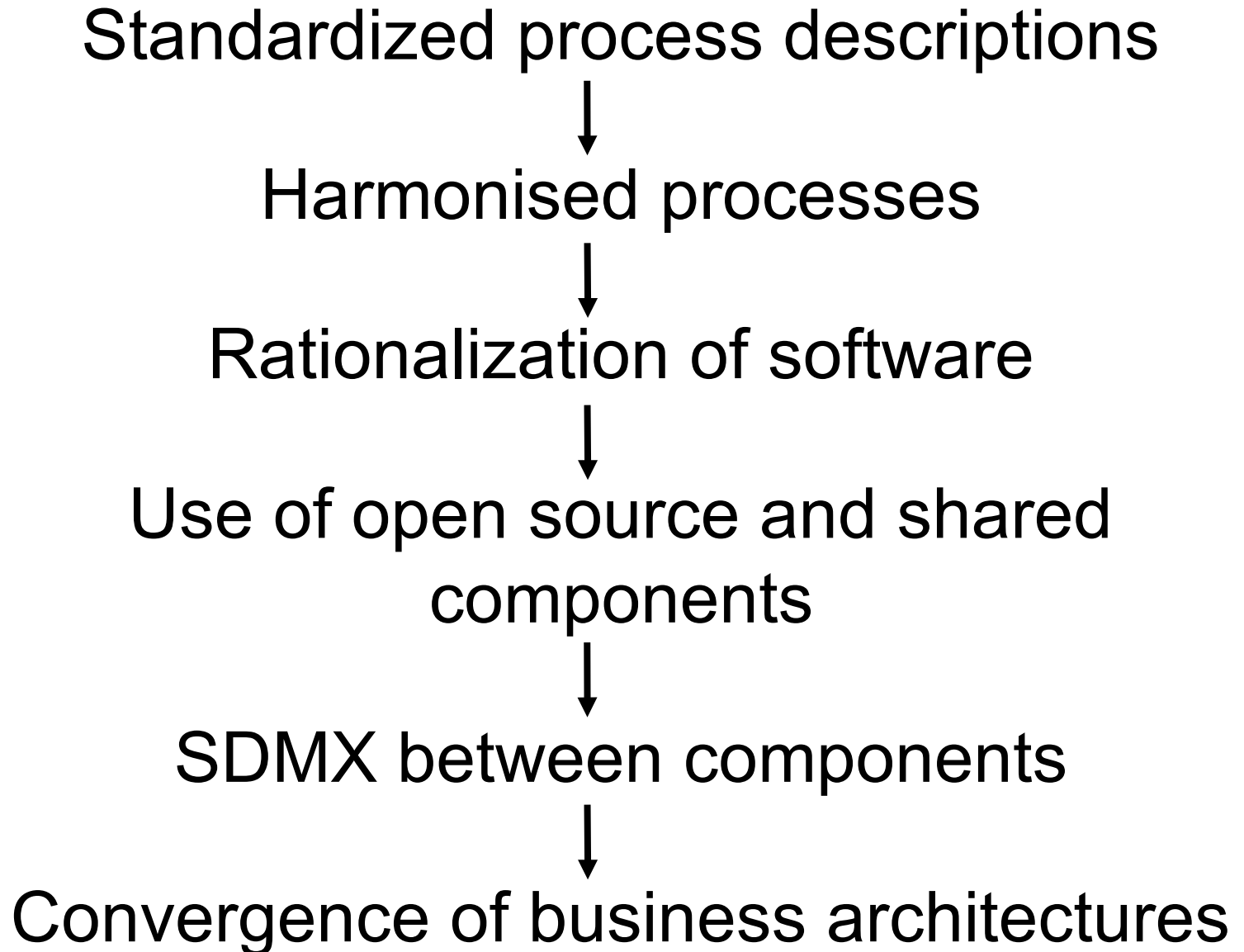- There may be several iterations of a process underway at any point in time

Key features (3)

Mapping to Other Models

Generic Statistical Business Process Model	Information Systems Architecture Model	DDI 3.0 Combined Life Cycle Model
1 Specify Needs	Planning - Specify survey contents - Establish survey procedures	Study Concept Repurposing (part)
2 Design		
3 Build		
4 Collect	Operation (part) - Frame creation - Sampling - Measurement	Data Collection
5 Process	Operation (part) - Data preparation - Observation register creation	Data Processing (mostly) Repurposing (part)
6 Analyse	Operation (part) - Estimation and analysis Evaluation (part) - Check survey outputs	Data Discovery Data Analysis Data Processing (part)
7 Disseminate	Operation (part) - Presentation and dissemination	Data Distribution
8 Archive		Data Archiving
9 Evaluate	Evaluation (part) - Evaluate feedback metadata	
Quality Management		
Metadata Management		

❖ Process modeling already mentioned in:

- SDMX User Guide
- V2 Technical Standards
- Euro SDMX Metadata Structure

❖ Common terminology

❖ If inputs and outputs use SDMX formats, why not the intermediate processes?

Istat

Standardized process descriptions

↓

Harmonised processes

↓

Rationalization of software

↓

Use of open source and shared components

↓

SDMX between components

↓

Convergence of business architectures

Istat

- ❖ Several organisations are implementing this model or similar ones
- ❖ Gather implementation experiences and other comments as input for Part C of the " Common Metadata Framework "
- ❖ Adopted by the Bureau of the Conference of European Statisticians
- ❖ Role in SDMX, DDI

Istat

❖ CORA network:

  ❖ ESSNet: European Statistical System network

  ❖ Eurostat funded

  ❖ 7 NSIs: ch, dk, it (coord.), lv, nl, no, se

  ❖ final meeting september 2010

  ❖ coordination with UNECE Sharing Advisory Board

Istat

❖CORA tasks

- ❖shared software repositorycreation (starting from GSBPM phases)
- ❖definition of a common architecture based on GSBPM using a **layered model**
- ❖list business models for software sharing
- ❖face licensing questions (EUPL ok, but not only)

Istat

CORE network:

- ❖ Eurostat funded
- ❖ 6 NSIs: fr, it (coord.), nl, no, pt, se
- ❖ final meeting January 2012
- ❖ coordination with UNECE Sharing Advisory Board, MSIS, GSIM

Istat

CORE tasks

❖ definition of a common environment in which run statistical applications

❖ selection of work-flow tools for statistical institutions

❖ Scenario with practical execution of different statistical applications both in .NET and in Java

❖ Cooperation also with METIS, GSIM, HLG-BAS, Statistical Network …

❖ MSIS SAB tasks

  ❖ coordinate and encourage collaboration between statistical organizations

  ❖ emerging economies are entitled to benefit from the work of others

  ❖ creating tools such as a web repository

  ❖ support the creation of a common reference architecture to ensure that the different components we are sharing will fit together efficiently

Istat

❖GSBPM use as:
  ❖Integrating work on statistical metadata and quality
  ❖Harmonizing statistical computing architectures between Directorates
  ❖Common framework for activities and/or organizational units and/or software
  ❖Sharing statistical software (other NSIs)
  ❖Measuring operational costs and/or system performance
❖Integration with CORA/CORE, SAB, GSIM, SDMX, DDI, ...