



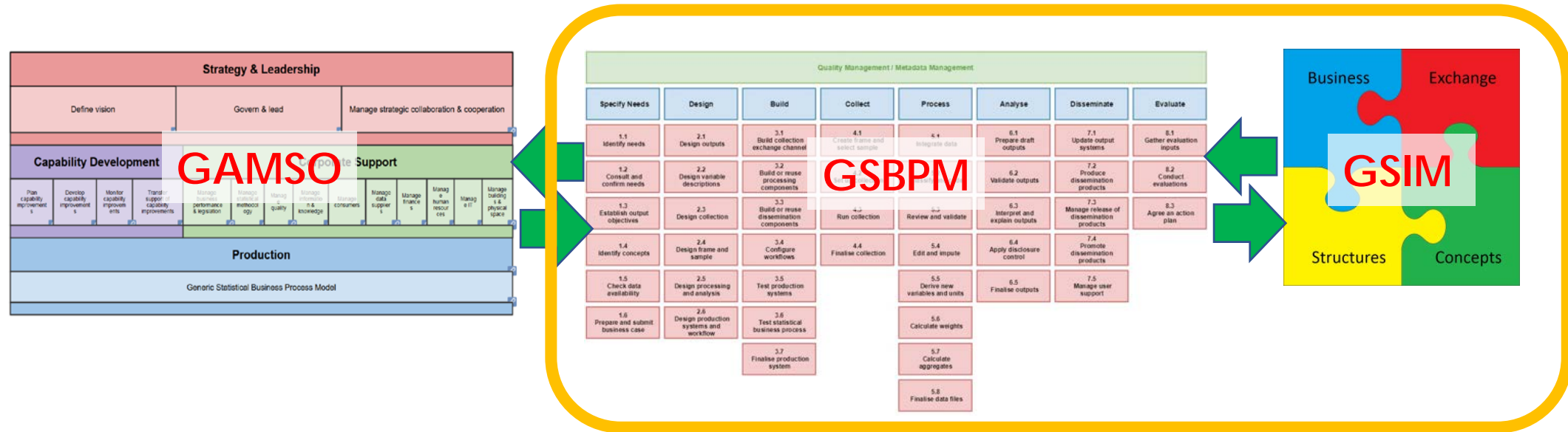
Linking GSBPM and GSIM: Status Update

Flavio Rizzolo
Statistics Canada



Delivering insight through data for a better Canada

ModernStats model integration



Business and information models need to align to be used effectively

Linking GSBPM and GSIM – Task Team

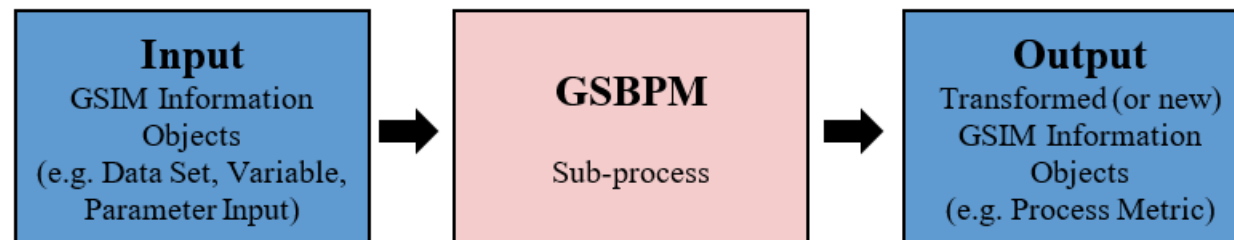
- 18 members from 10 participating institutions
- Main goal → help NSOs understand the models better to implement metadata-driven solutions, e.g. CSPA services, and enable semi-automated business process

Country	Member
Italy	Marina Signore - Chair Giorgia Simeoni Mauro Scanu Carlo Vaccari
Canada	Flavio Rizzolo - Chair Francine Kalonji
Mexico	Juan Munoz Ildeliza Ramos Jesús Togno
Sweden	Patrik Wahlgren

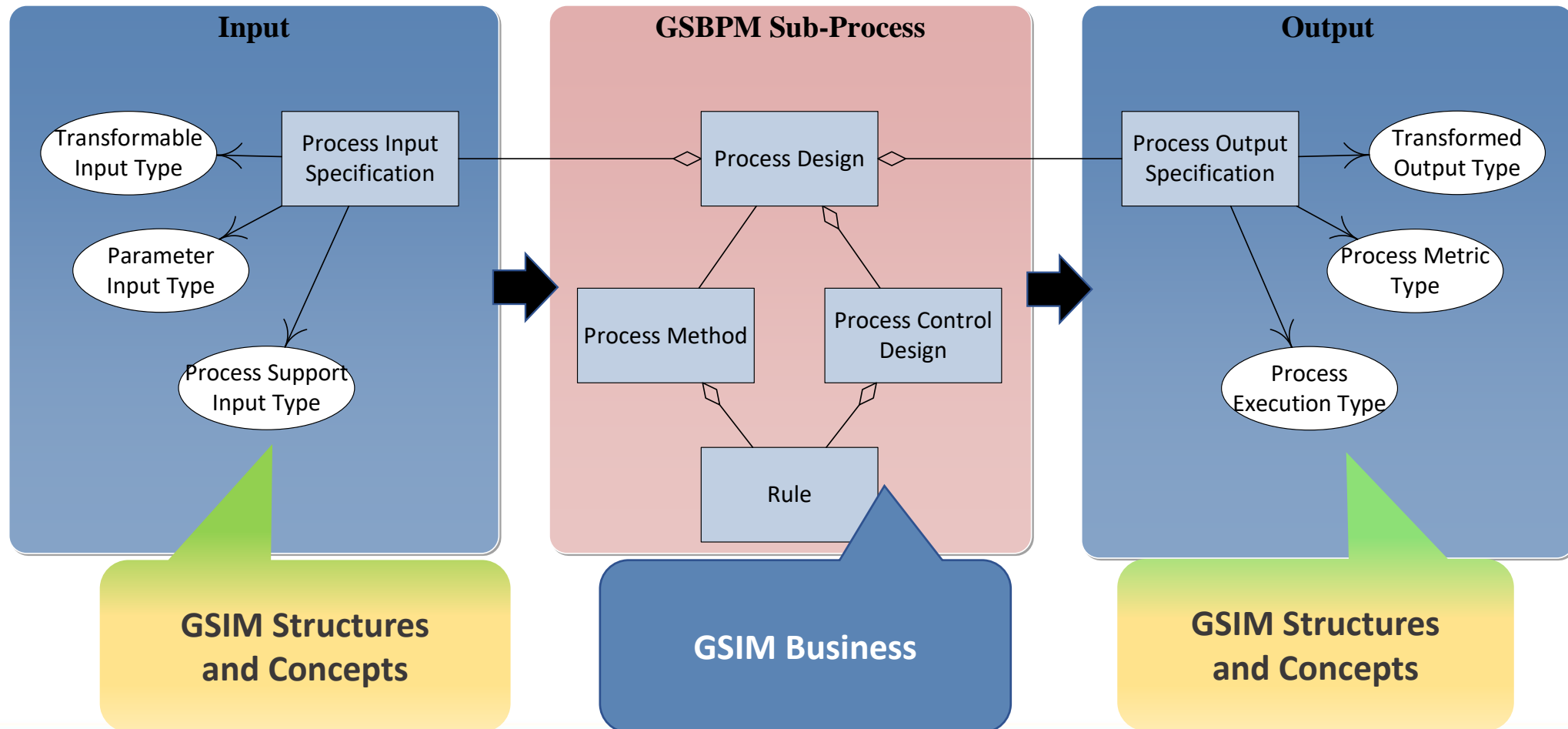
Country	Member
Hungary	Zoltán Vereczkei Csaba Ábry Eva Hajosn Ender
Poland	Anna Dlugosz
South Korea	Yulla Choi
Australia	Annette Tyler
ILO	Edgardo Greising
DDI	Jay Greenfield

Linking GSBPM and GSIM - Mapping

- The **Generic Statistical Business Process Model (GSBPM)** provides a framework to describe the building blocks of statistical production processes
- The **Generic Statistical Information Model (GSIM)** provides a catalogue of information objects to describe statistical data and metadata
- Mapping inspired by paper **Metadata Flows in the GSBPM** (UNECE Work Session on Statistical Metadata, 2013)



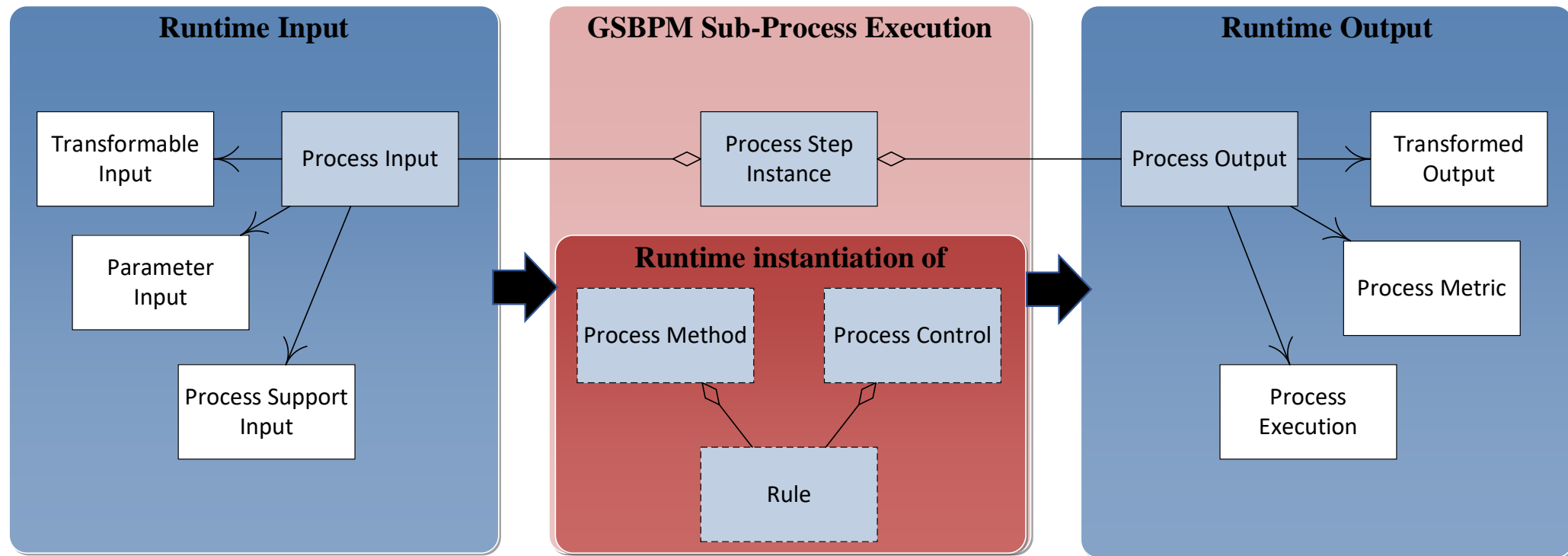
Describing GSBPM sub-processes with GSIM - Specification



Specification template example: 2.1 Design outputs (Hungary)

<i>Process Input Specification</i>	<i>Process Step Design</i>	<i>Process Output Specification</i>
<p>Transformable Input type:</p> <p><i>Statistical Need:</i></p> <ul style="list-style-type: none"> User needs concerning outputs (products and services) to be produced; that are not exact parameters for the <i>Process Step Design</i>. <p><i>Statistical Program Design</i></p> <ul style="list-style-type: none"> If there is previous design available (that is refined in the sub-process), especially <i>Data Structure</i> <p>Parameter Input Type</p> <p><i>Statistical Need:</i></p> <ul style="list-style-type: none"> User needs concerning outputs (products and services) to be produced; that can used as parameters for output design. Dissemination rules and policies (including data protection provisions, quality policies, etc.). <p>Process Support Input type:</p> <p><i>Referential Metadata:</i></p> <ul style="list-style-type: none"> If there are referential metadata (descriptions of the methodology, quality) available from previous iterations, then this can be used as <i>Process Support Input</i> to design outputs. 	<p>Process Method:</p> <p>Apply processes to design the output elements of the <i>Statistical Program Design</i>.</p>	<p>Transformed Outputs type:</p> <p><i>Statistical Program Design</i></p> <ul style="list-style-type: none"> <i>Product</i> specification elements, especially <i>Data structures</i> and quality criteria for the outputs. <i>Process Method</i> specifications for creating the draft outputs and the dissemination products <p>Process Metric type:</p> <ul style="list-style-type: none"> Quality measures (e.g. check if all user needs are incorporated into the designed outputs). <p>Process Execution Log type:</p> <p><i>Process Execution Log</i></p> <ul style="list-style-type: none"> Performance measures

Describing GSBPM sub-processes with GSIM - Runtime



100



GSBPM coverage (so far)

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			

GSIM definitions update (in progress)

- Main Inputs
 - Transformable Inputs → needn't be transformed
- Auxiliary Inputs
 - Process Support Inputs → influences the execution
 - Parameter Inputs → run-time configuration
- Main Outputs
 - Transformed Outputs → information flows to other steps
- Secondary Outputs
 - Process Metrics → execution summary
 - Process Execution Log → timestamped events
- Good synergy with GSIM task team

Outcomes and Future Work

- Two templates, i.e. a high-level specification and illustrative runtime examples, capture all necessary information
 - Elicitation of common types of inputs and outputs across examples provided by NSOs
- New mapping diagrams showing metadata flows
- GSIM definitions update
- Detailed report
- Finalized mapping of remaining GSBPM sub-process (2021, if proposal approved)

Challenges

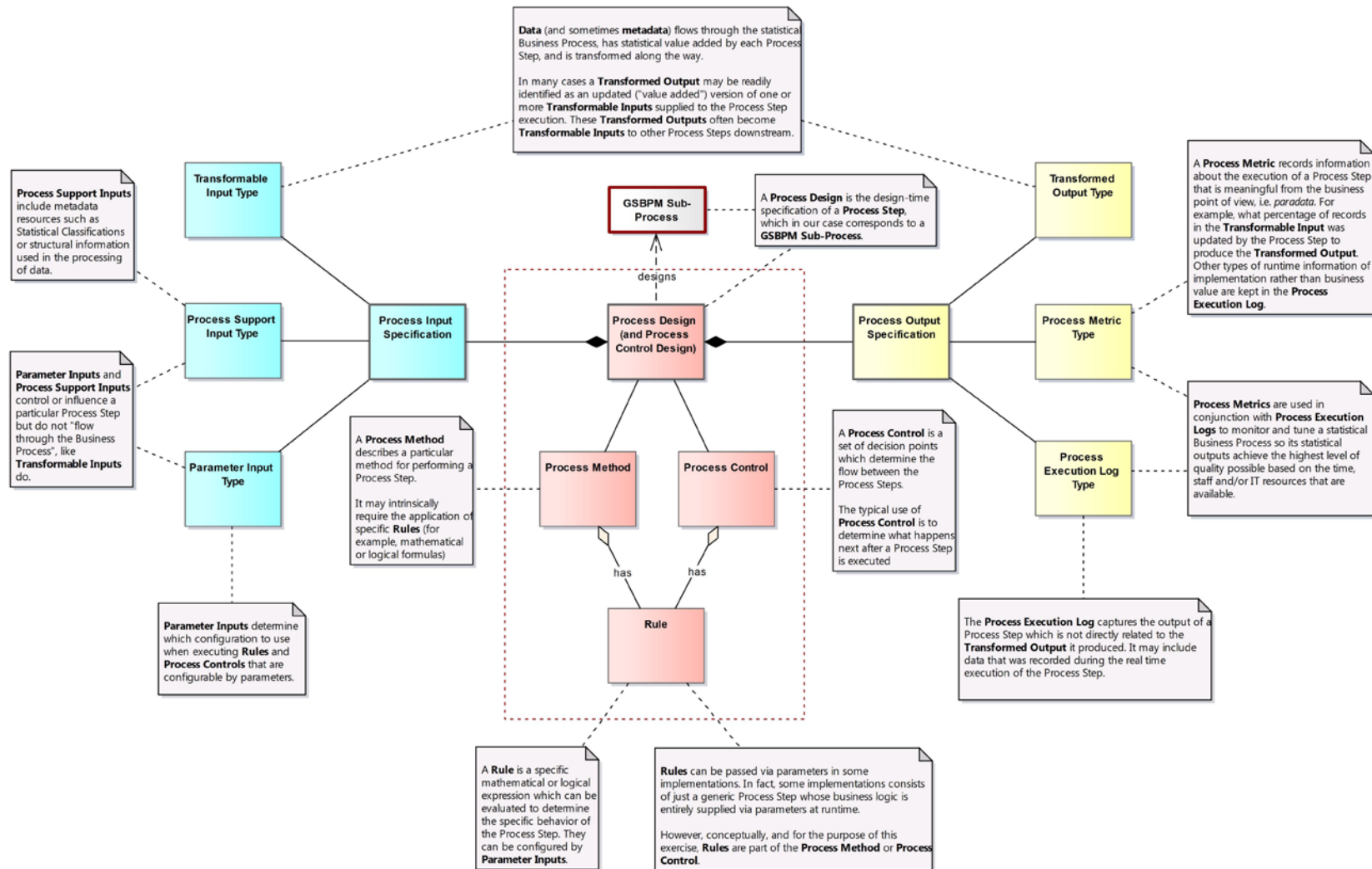
- Mapping exercise proved to be more challenging for Phases 1 (Specify Needs) and 2 (Design)
 - Processes around requirements and design are less clear
 - Terminological issues between the two models became more problematic
- Access to internal subject matter experts by the NSOs teams to fill out the templates
- Diversity of approaches and interpretations of models



Annex

class GSIM-GSBPM

GSIM objects used in the template



class GSIM-GSBPM Process Step Instance

GSIM instance objects used in the template

