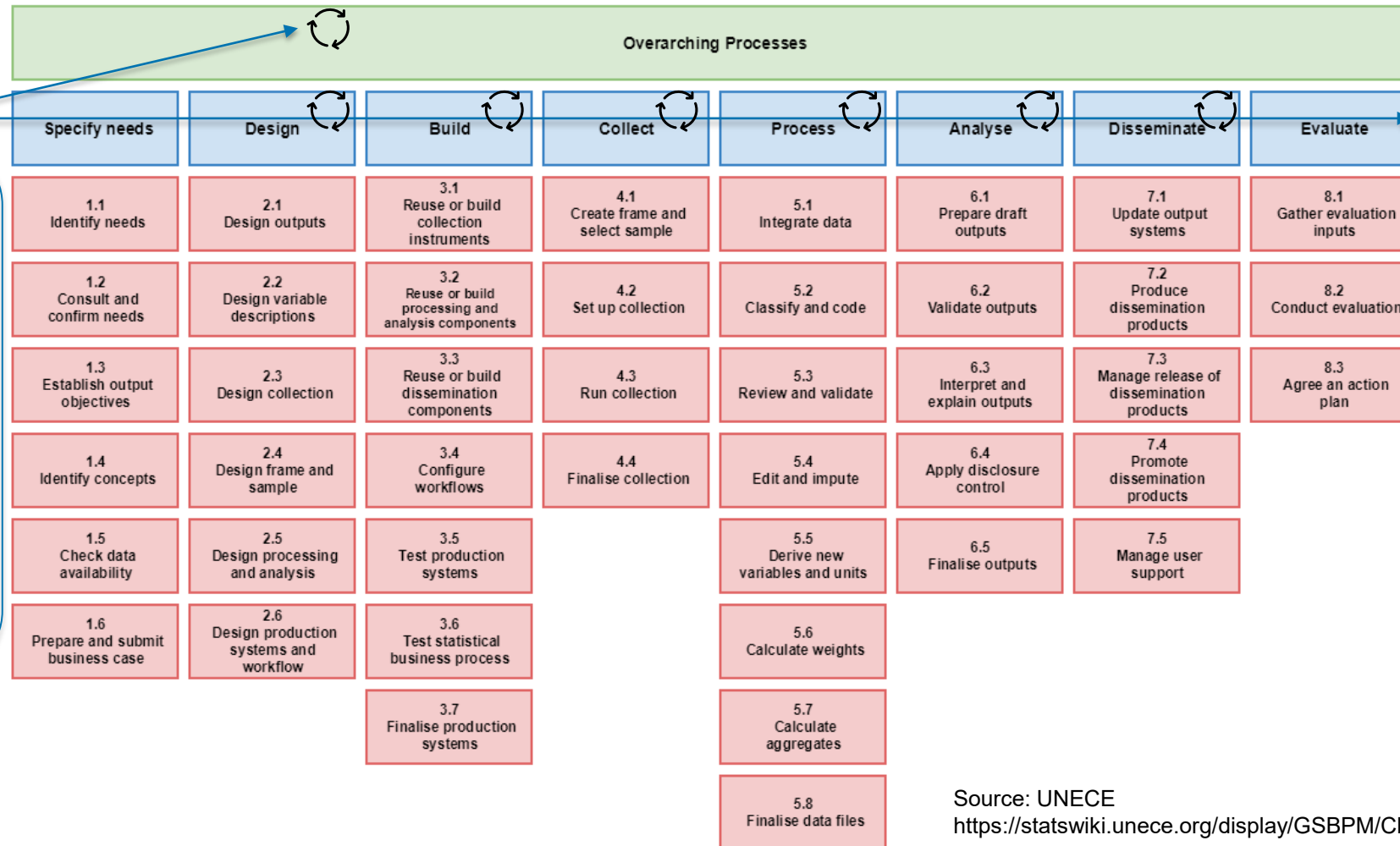




GSIM supporting the changing business case in Statistics Finland: Case Data Acquisition Process

Mikko Saloila (with co-authors)
ModernStats World Workshop, Geneve 26-28.6.2019


Changing business case is having a substantive impact on information management throughout GSBPM



Source: UNECE
<https://statswiki.unece.org/display/GSBPM/Clickable+GSBPM+v5.1>

Background: Transformed administrative data sources



- One of the drivers in Statistics Finland are new and transformed administrative data sources
 - For example the Income register from The Tax administration or right to study and achievement register from The Finnish National Agency for Education
 - Typically Data is provided frequently, even on a daily basis, and usually consists of events.
 - Not tailored for NSO needs, instead a standard API for all agencies in the public sector
- These data sources are seen partly as an answer to the need to get closer to real time data
 - ...But it also creates a need to change and automate production processes 

Case Data Acquisition 2.0:

Need to develop the collection process

Design

Build

Collect

- Automation of Data Acquisition processes requires a more detailed description of incoming data
- Desire to move further into Single Point of Access for administrative data
- Effective handling of GDPR* responsibilities requires new services to handle personal information

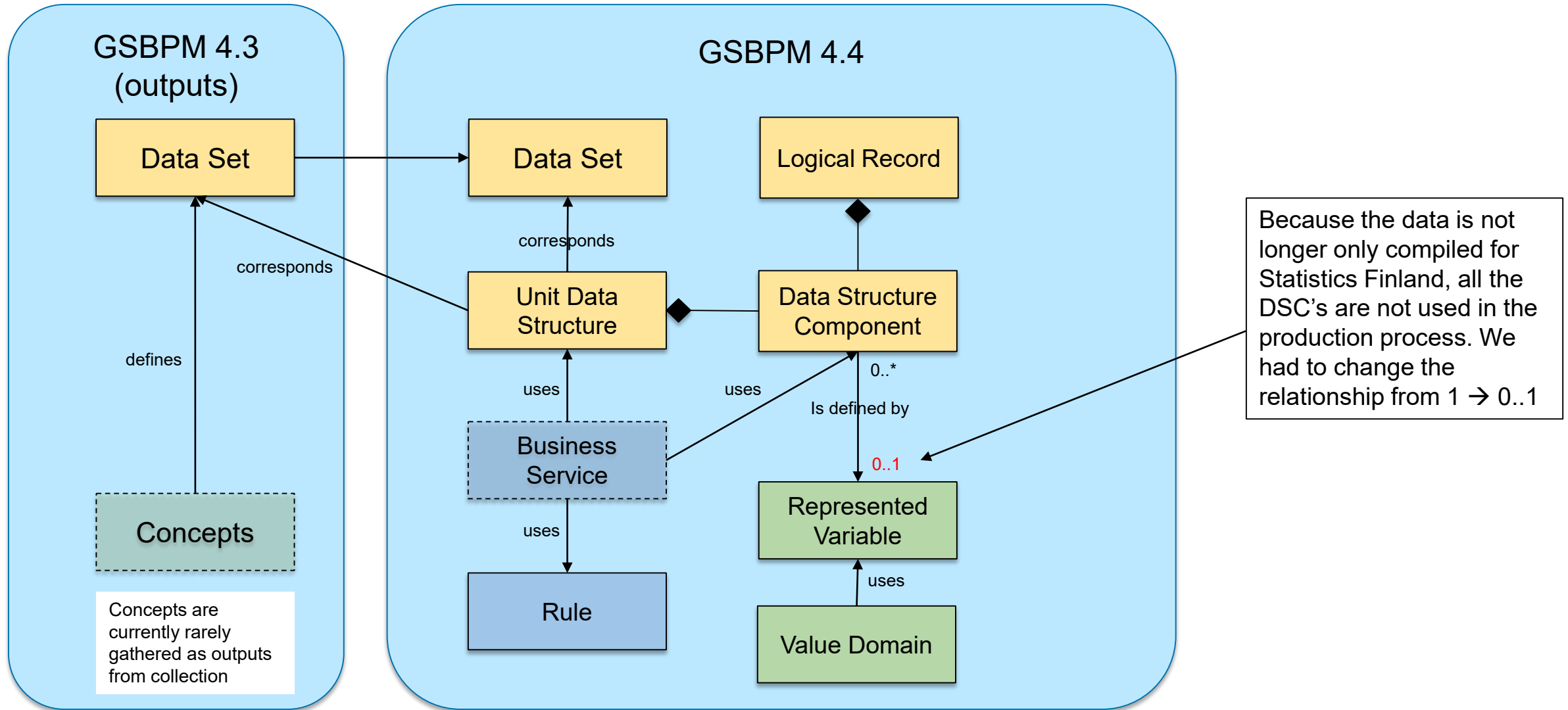
*General data protection regulation

https://ec.europa.eu/info/law/law-topic/data-protection/reform/what-does-general-data-protection-regulation-gdpr-govern_en

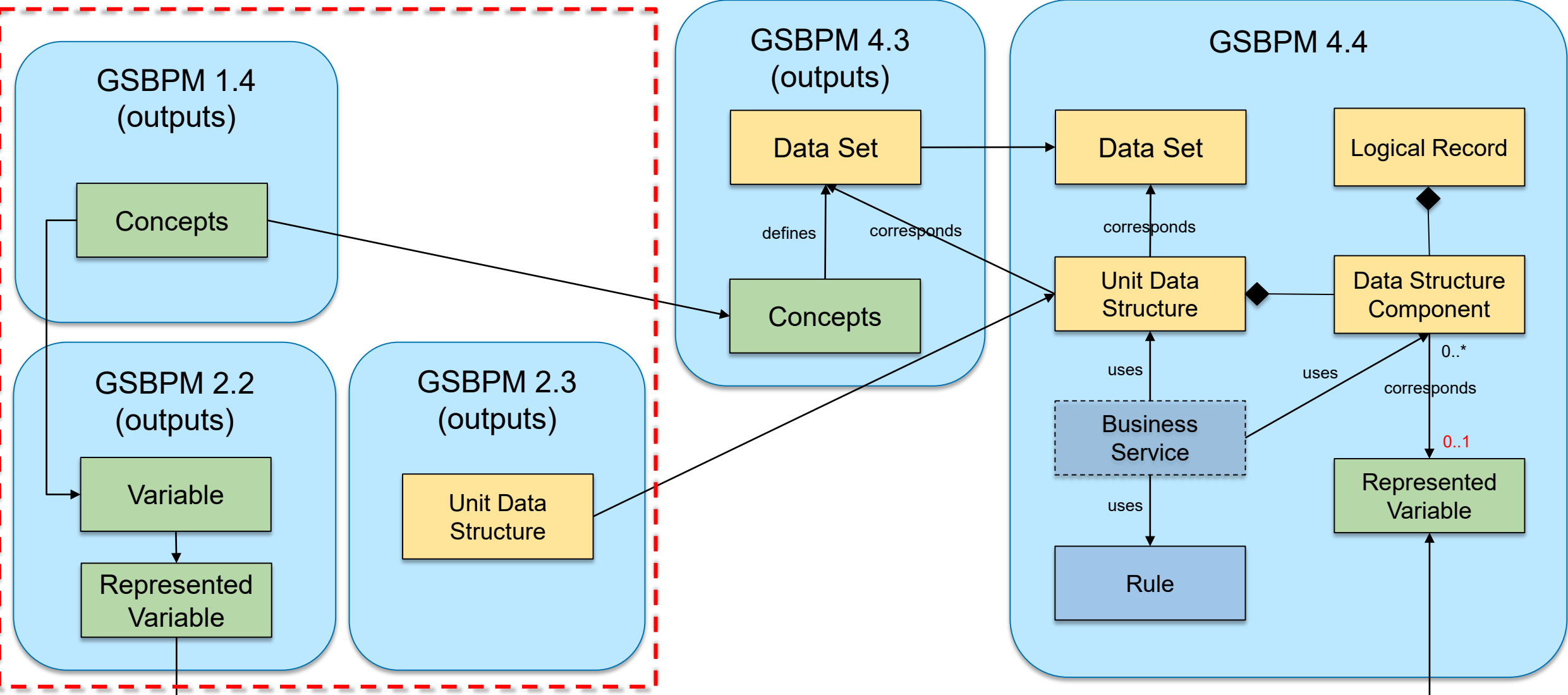
Data Acquisition 2.0 Project targets: Re-design of data acquisition in GSPBM 2-4

- ❑ Metadata Driven
 - The processing of the acquired data is based on GSIM based metadata objects and relevant processing instructions
- ❑ Fully Automated Process
 - The data acquisition process is executed using an in-house process management system
 - Human interaction is required only if problems arise within the 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
 - Data is refined step by step in different zones
- ❑ Transparent Process
 - Visual process monitoring
 - All tasks gather process metric (in respect to CSPA LIM)

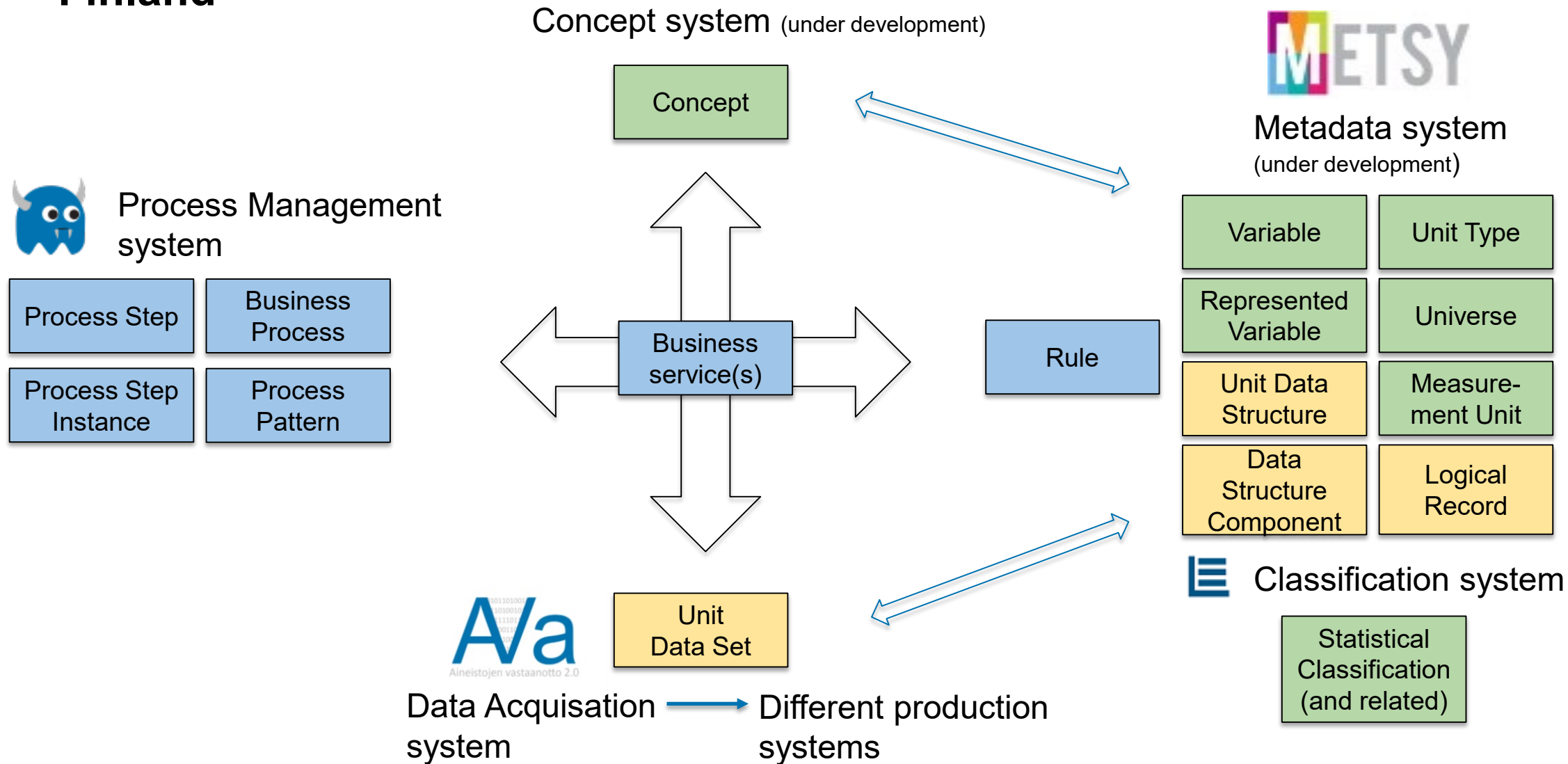
So how does Data Acquisition look with GSIM objects in GSBM framework



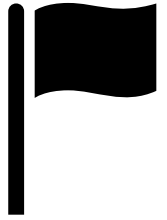
Automation of processes requires more focus towards the design phase



How GSIM-objects are currently being implemented in Statistics Finland



Challenges related to GSIM and service development



- Managing Lineage is difficult when there are manual Process Steps in between automated processes.
- Actually changing the work processes towards planning and design is challenging
 - It is not always so straight forward to establish short term benefits, especially when automated processes are not ready
- Integrating commercial of the self products into all this requires work, especially when they are related to GSBPM 5-6 phases (e.g. BI-tools for macro statistics)



Thank you!

Mikko Saloila mikko.saloila@stat.fi

Toni Räikkönen toni.raikkonen@stat.fi

Aura Pasila aura.pasila@stat.fi

ModernStats World Workshop, Geneve 26-28.6.2019

Statistics Finland 