

Creating a Model for a Central Metadata Repository

Klas Blomqvist

EDDI 14, London 2014-12-02



Background

- Statistics Sweden - NSI
- > 200 “statistical products”
- Traditional stove-pipes production system
 - “Product” oriented
- Decentralised organisation
 - Relatively independent departments
 - 50% financed by commissions – “solution for each customer”
- Strategies (Decision by the Director General)
 - Register system must be coordinated
 - Data storage changed from accessing individual products to accessing the entire data warehouse



Content

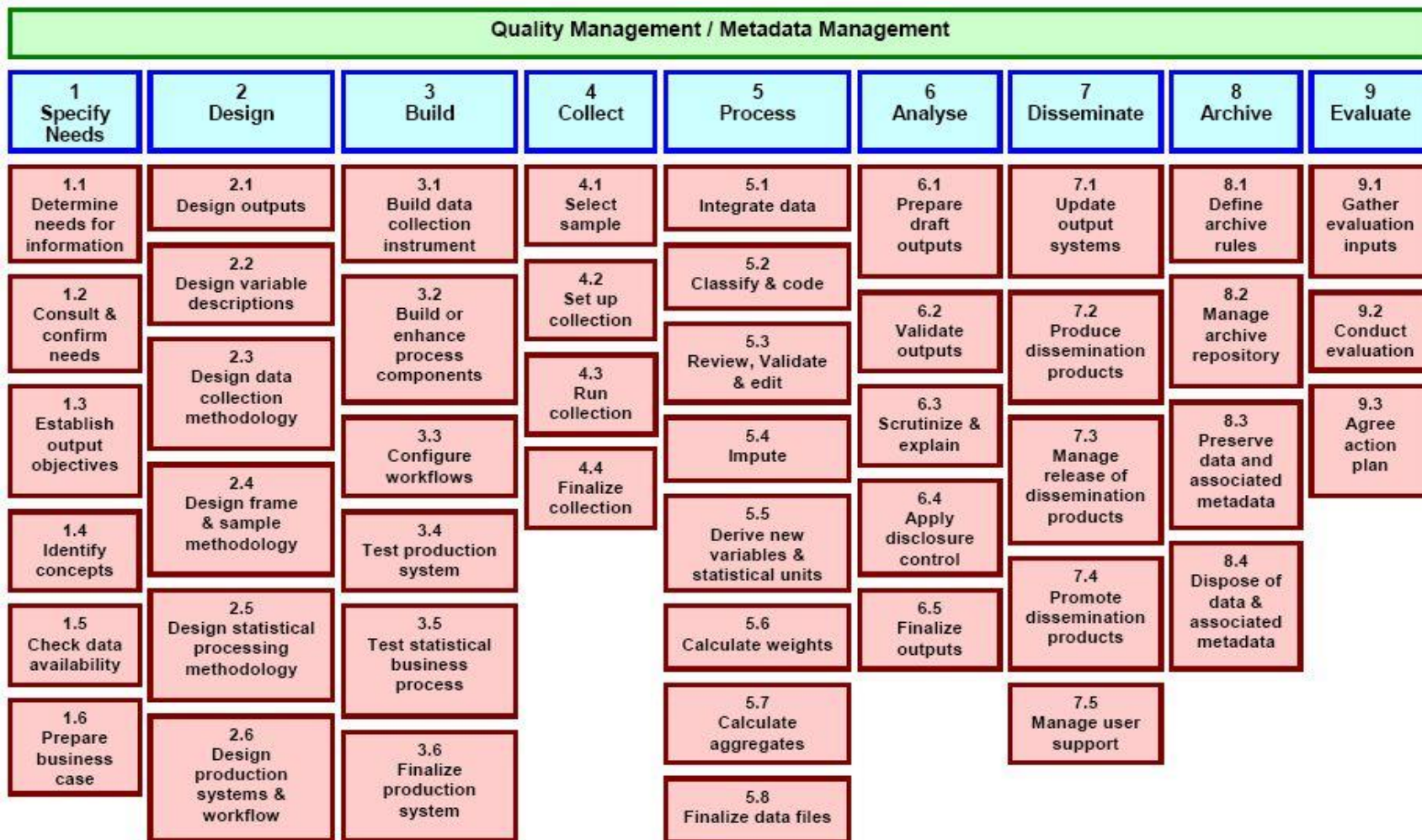
- Register and data warehousing vision and strategy
- A central metadata repository
- Standards based
 - GSBPM
 - GSIM
- Testing in production systems



GSBPM - Generic Statistical Business Process Model

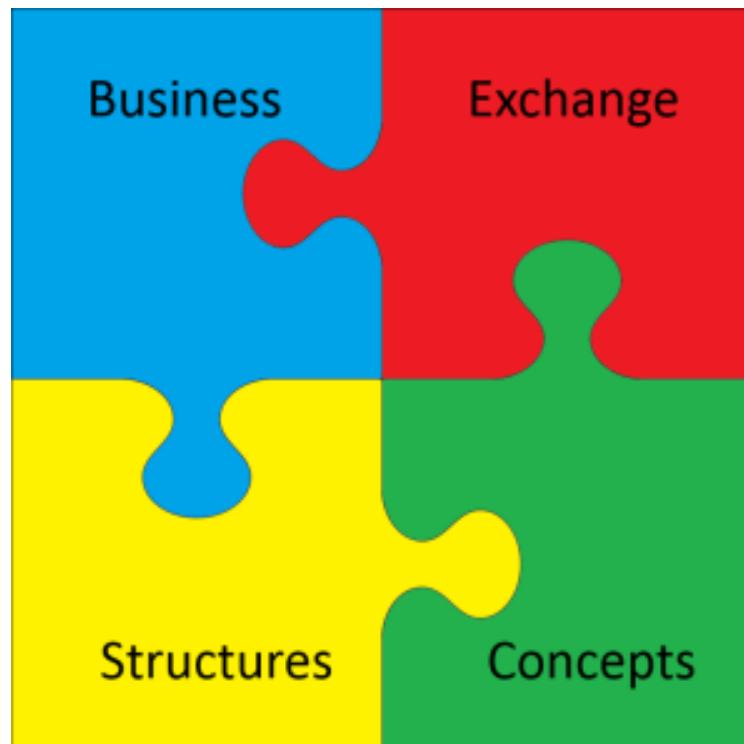


Statistiska centralbyrån Statistics Sweden





GSIM

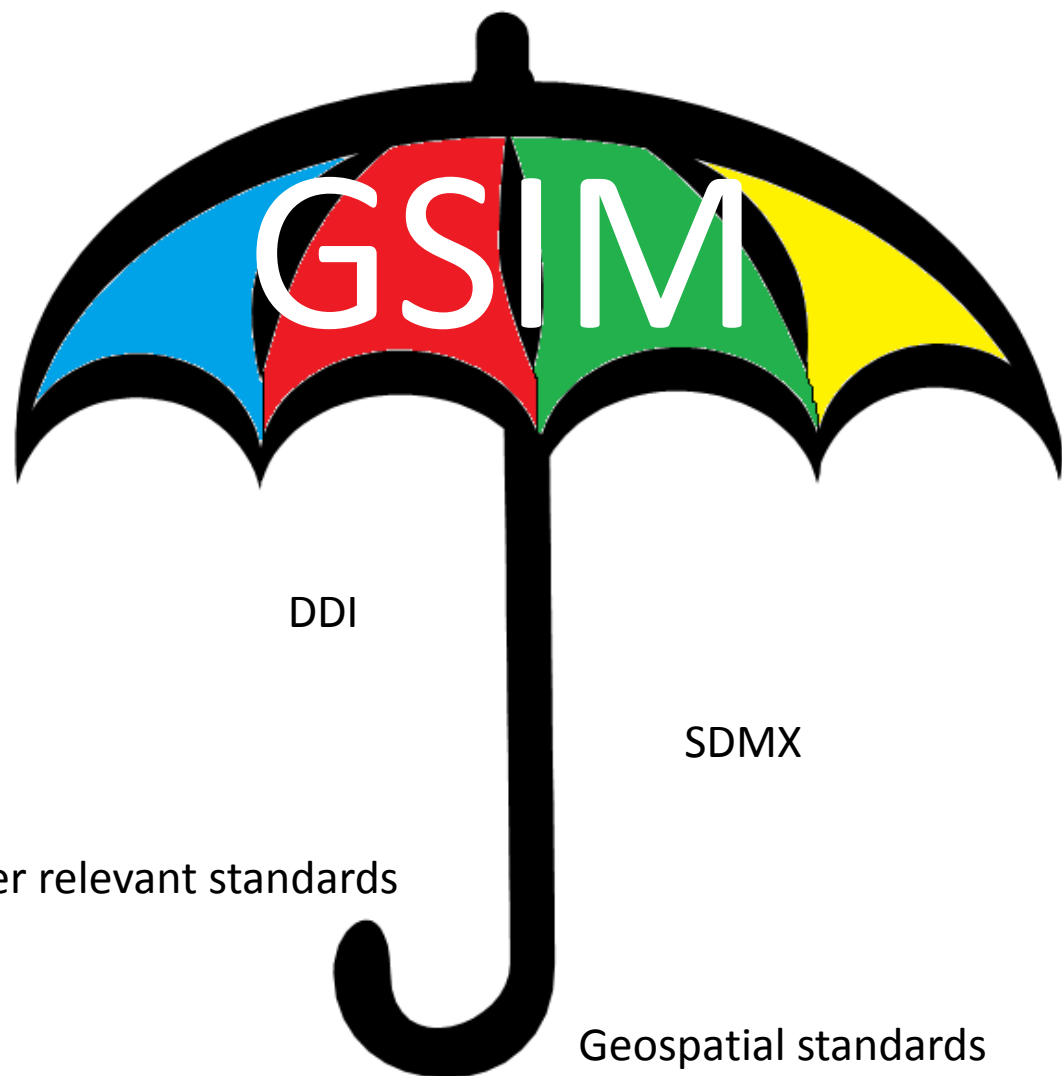




Conceptual model

Implementation standards

Other relevant standards

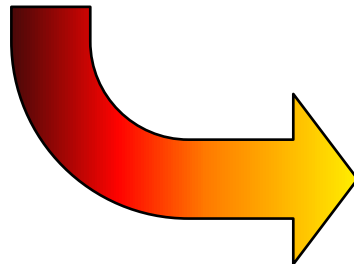
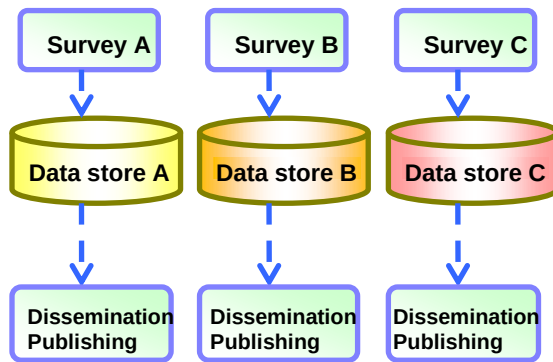


DDI

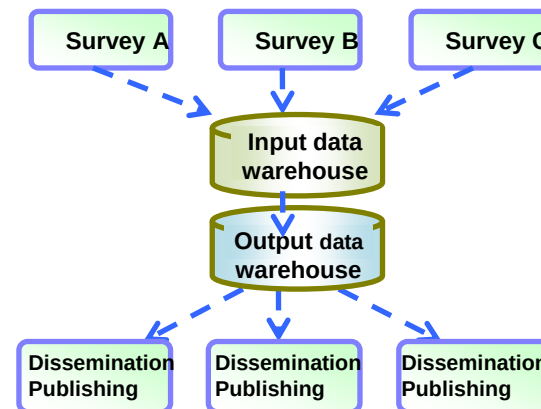
SDMX

Geospatial standards

Strategy

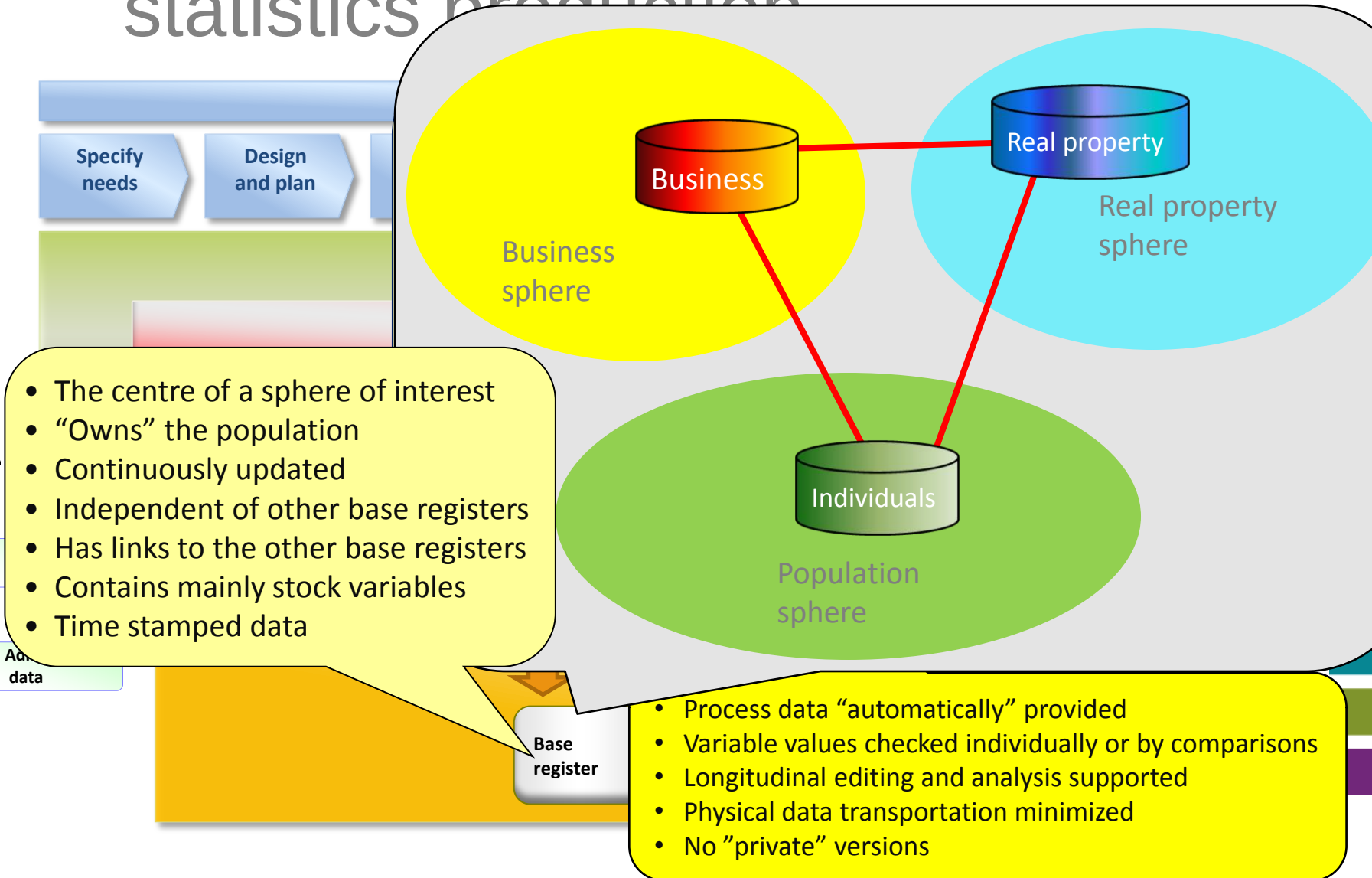


- Support process oriented production
- Active Communication with suppliers
- Use of registers when possible
- Standardization



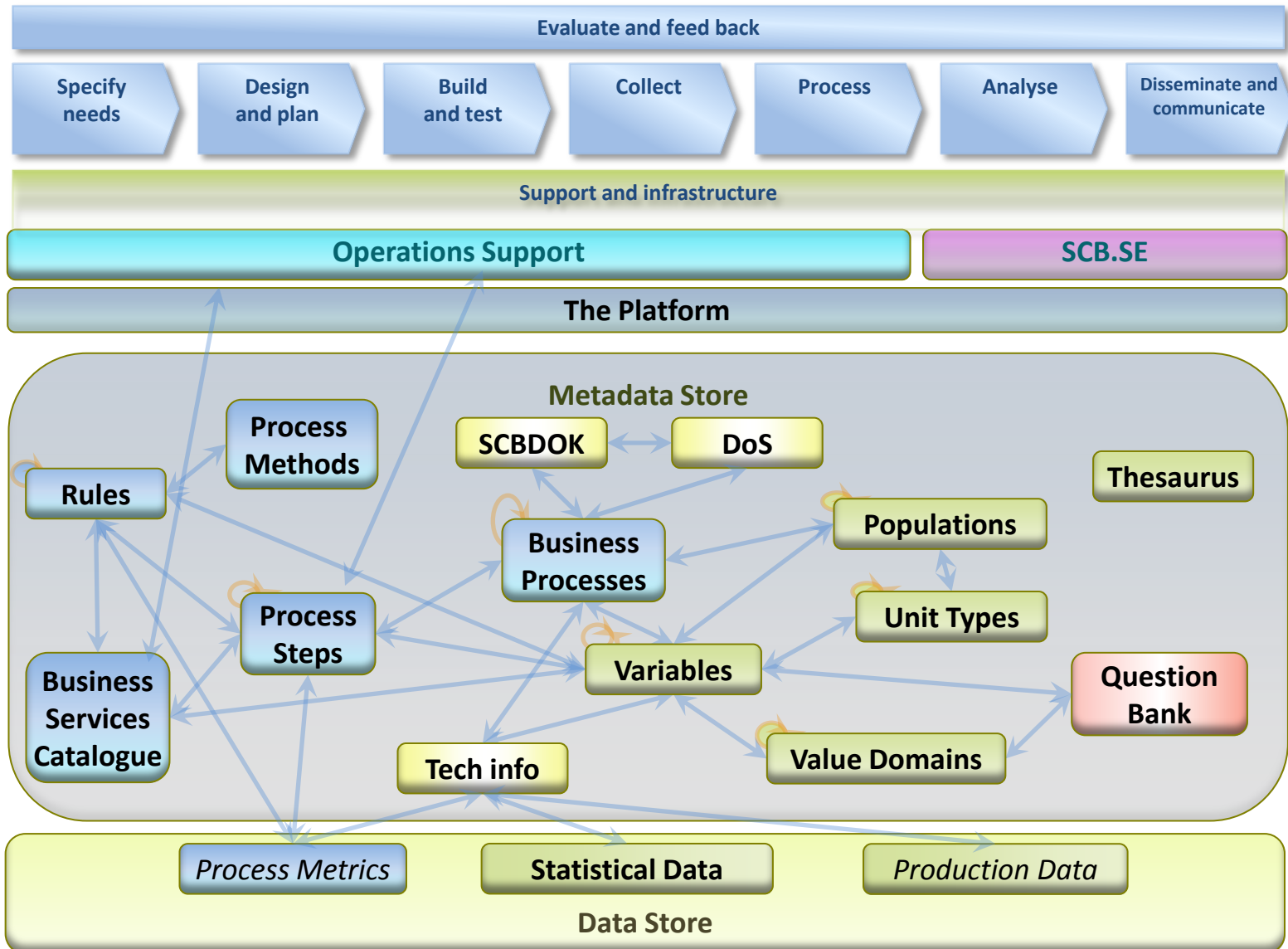
SCB's strategy on coordinated statistics production

Statistiska centralbyrån Statistics Sweden

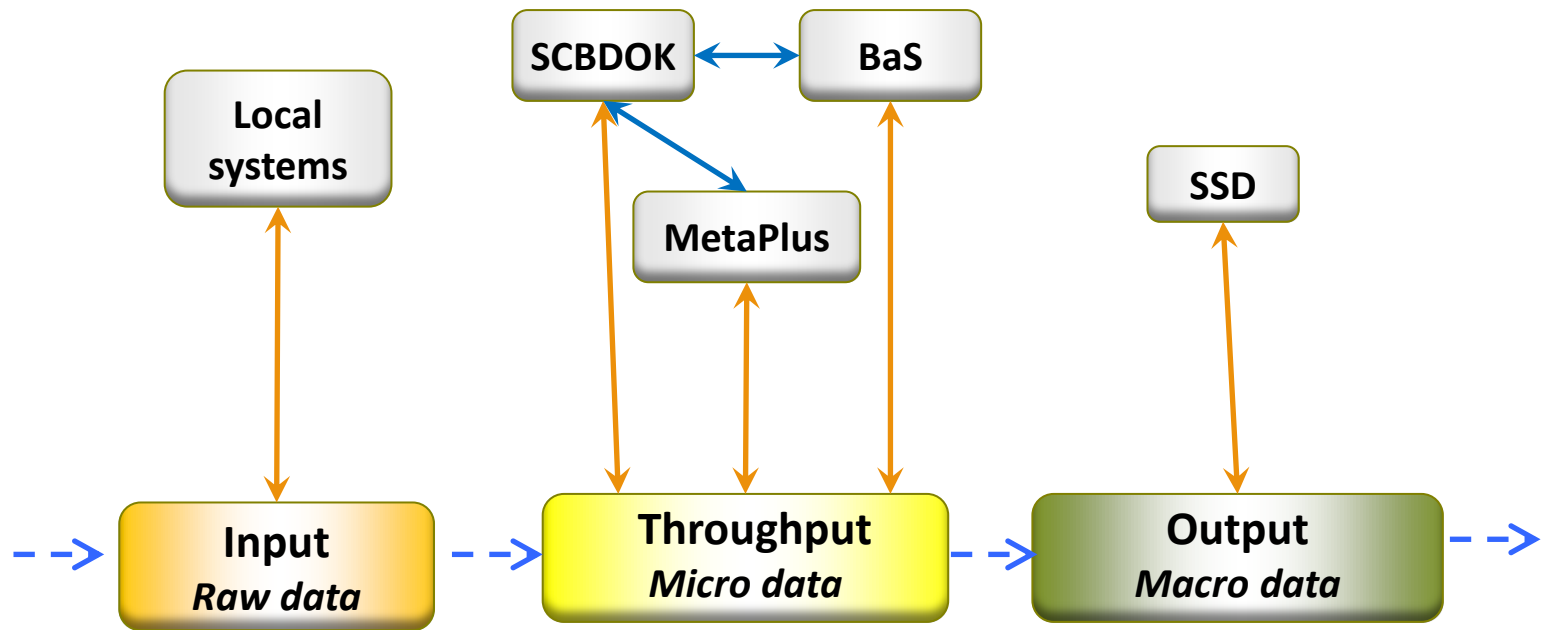


A Central Metadata Repository

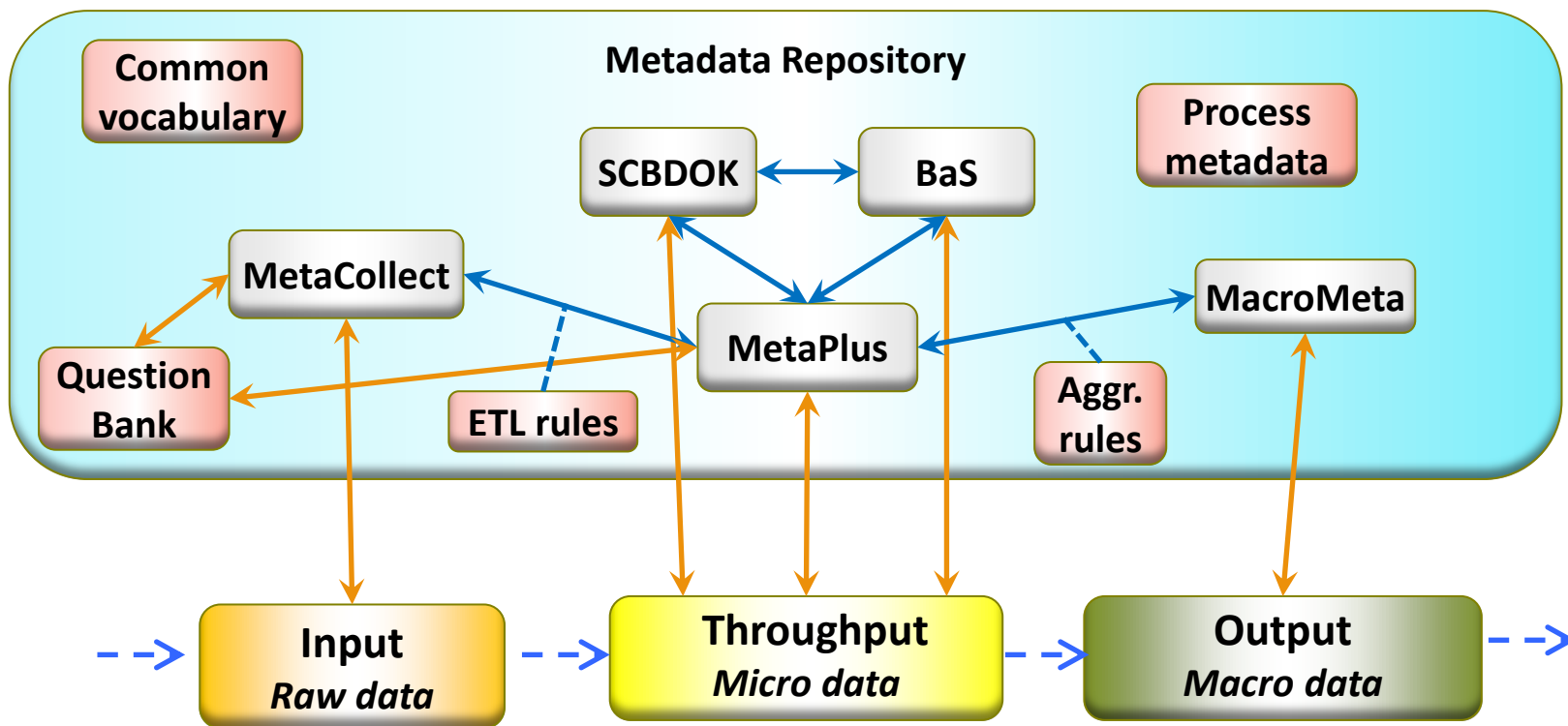
Statistiska centralbyrån Statistics Sweden



Current Metadata Situation at SCB



Desired Metadata Situation

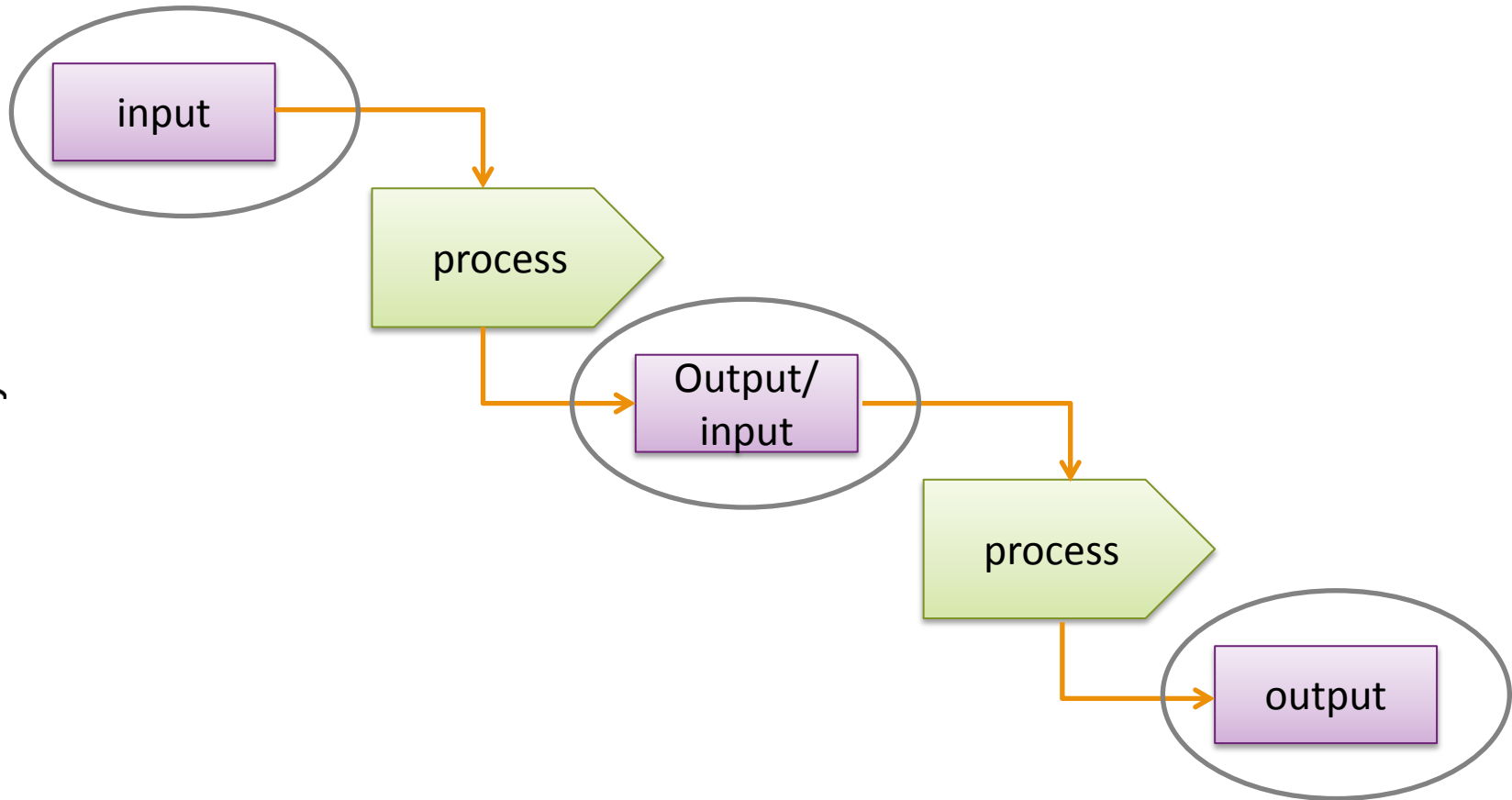


Metadata Expectations

- More active metadata
 - Documentation during the design phase
 - Metadata driven production
- Enable tracing
 - What was the question that led to this published figure?
 - Ideally a two-stage tracing for NA
- More and better process data
 - Who did what when and why?
- More comprehensive metadata
 - Improved search mechanisms
- Unified terminology
 - Thesaurus



Runs and feeds the processes

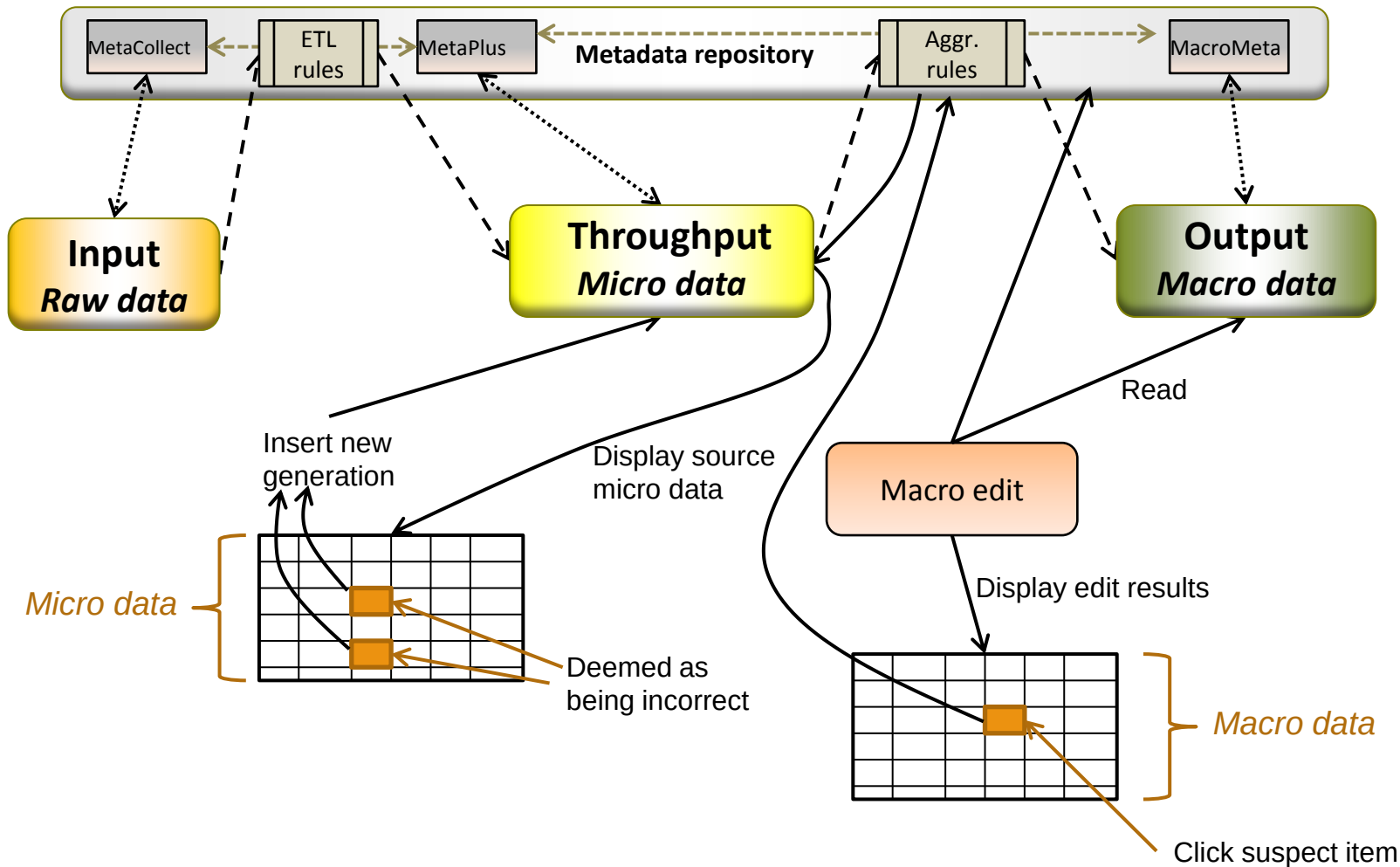


Macro editing with drill-down

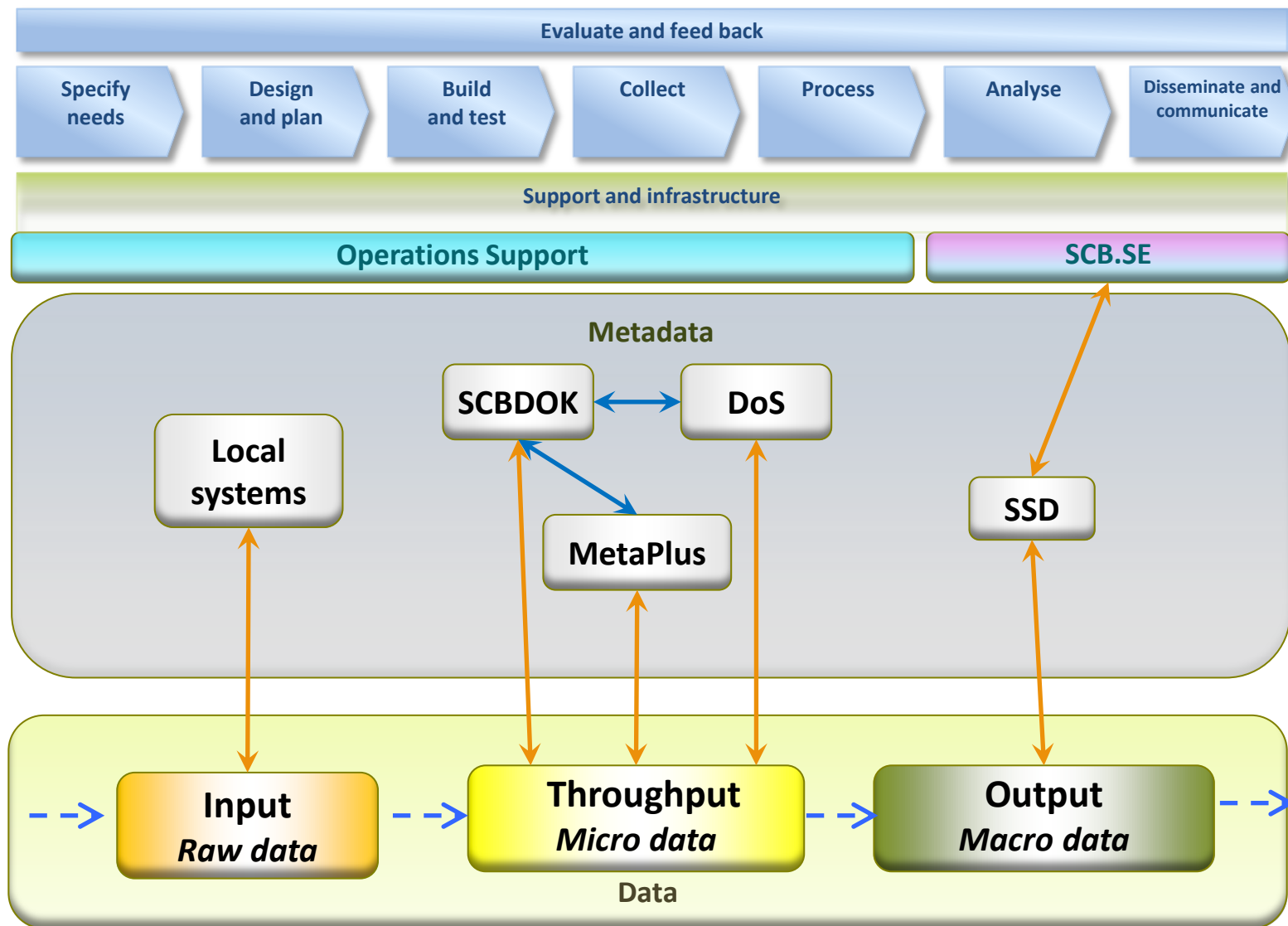


Statistics Sweden

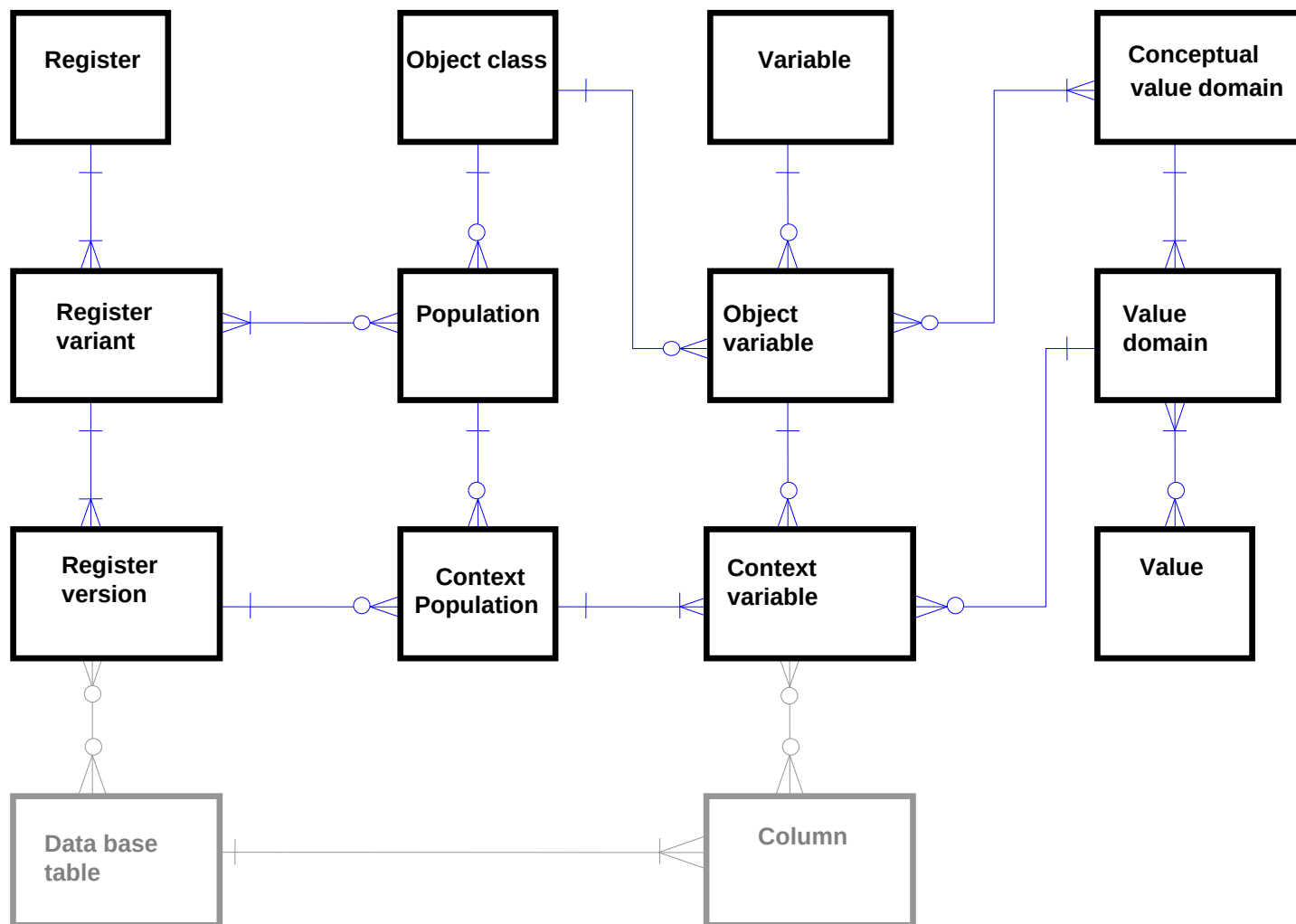
Statistiska centralbyrån



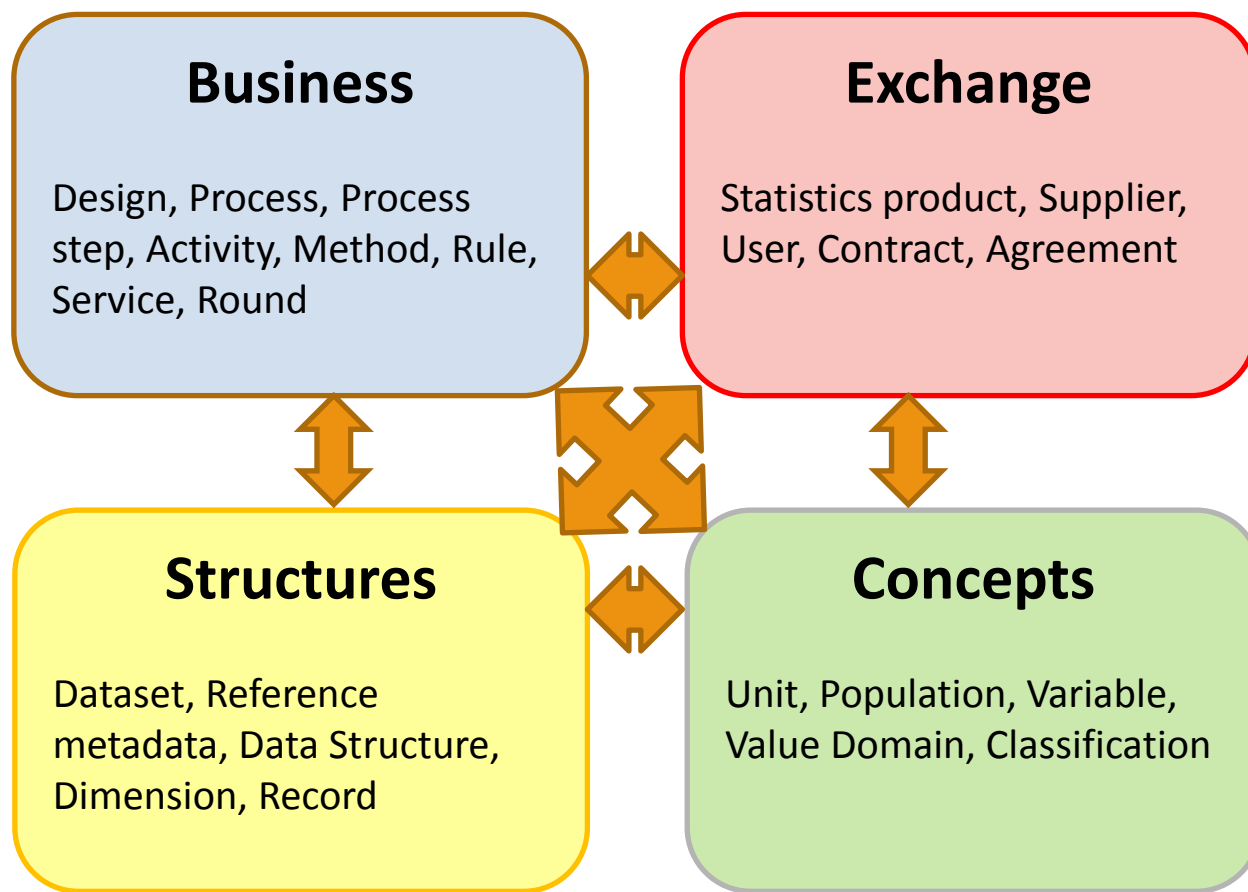
The Current Metadata Situation



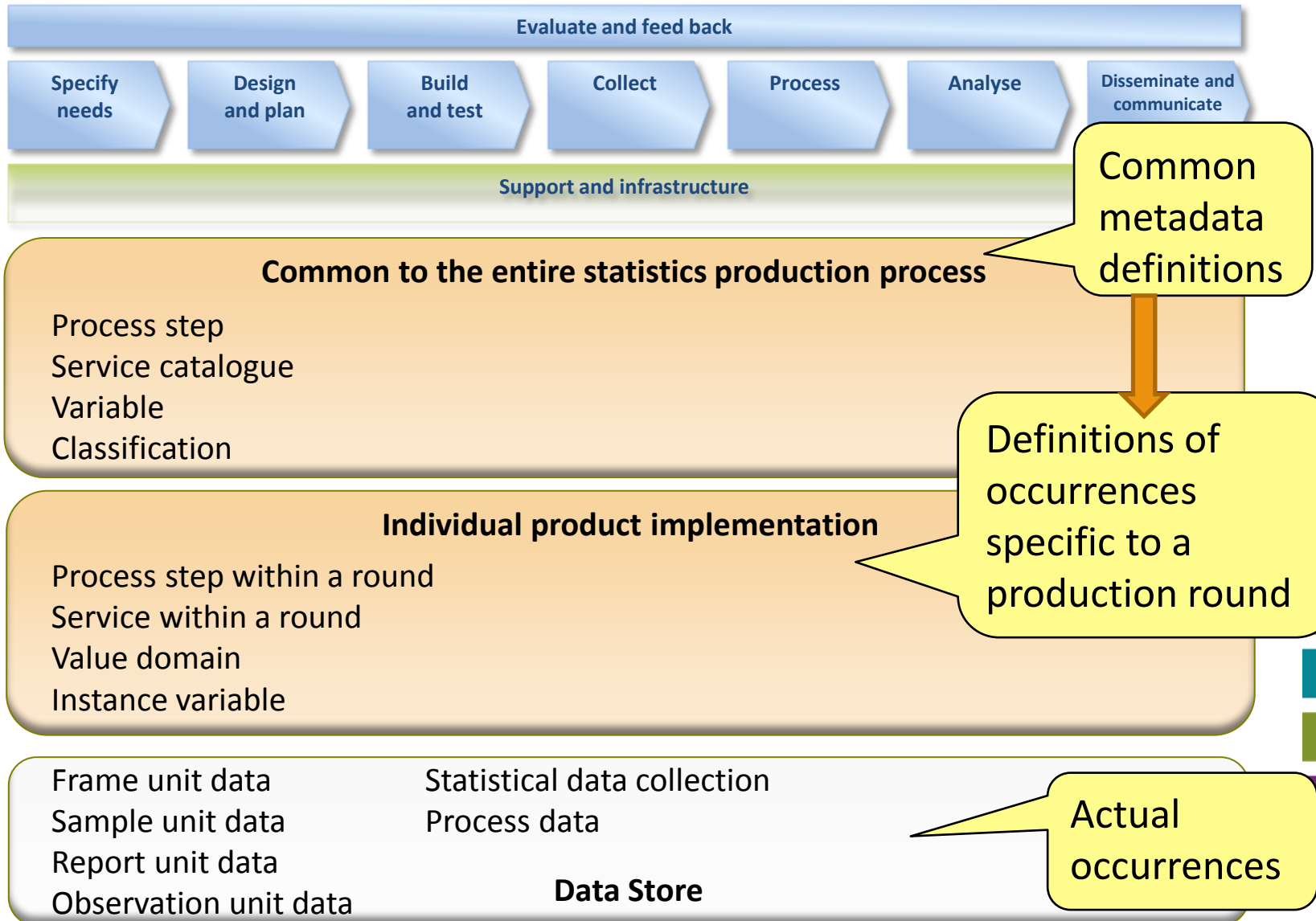
MetaPlus Structure

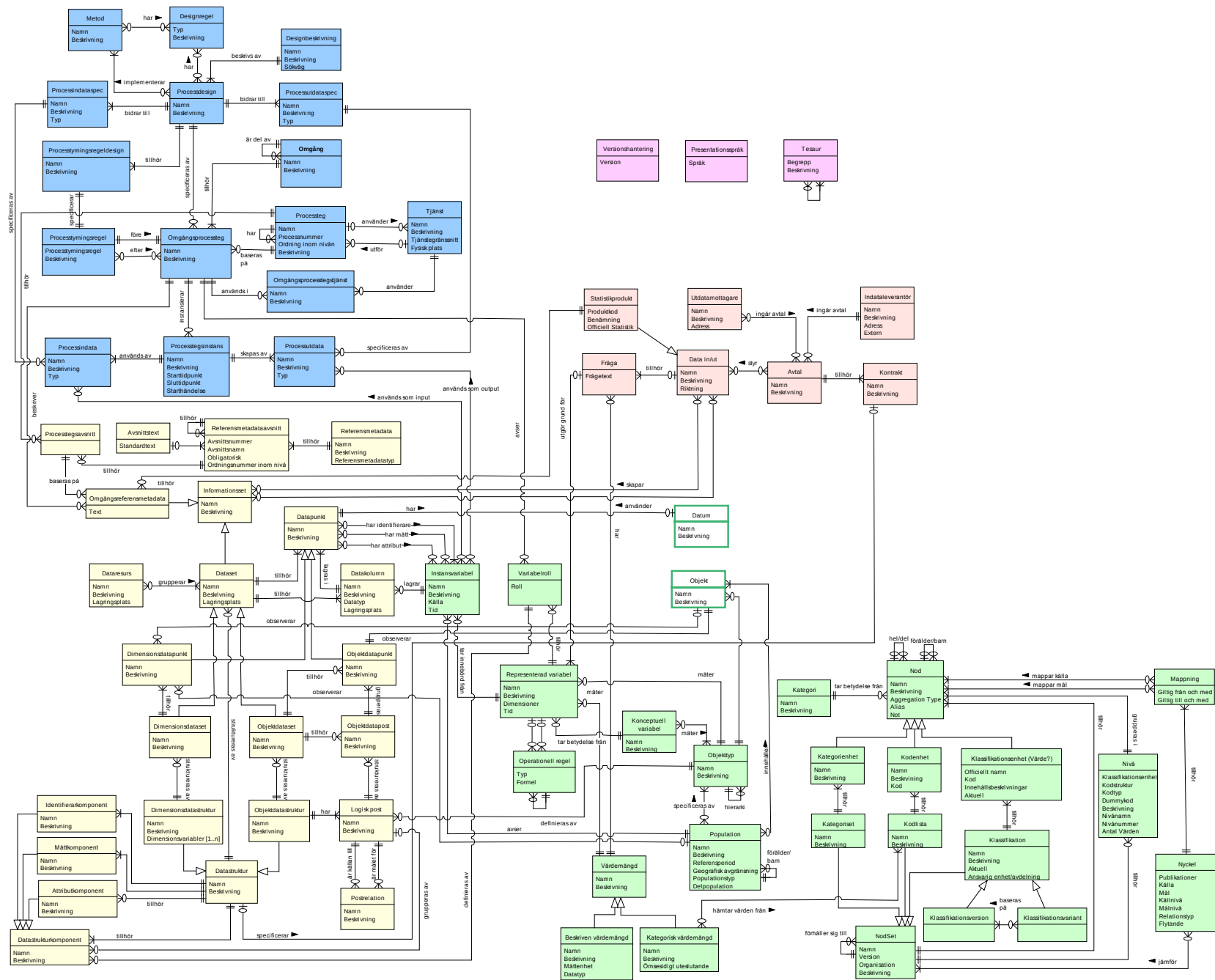


Central Metadata Store for efficient statistics production



Common and specific metadata

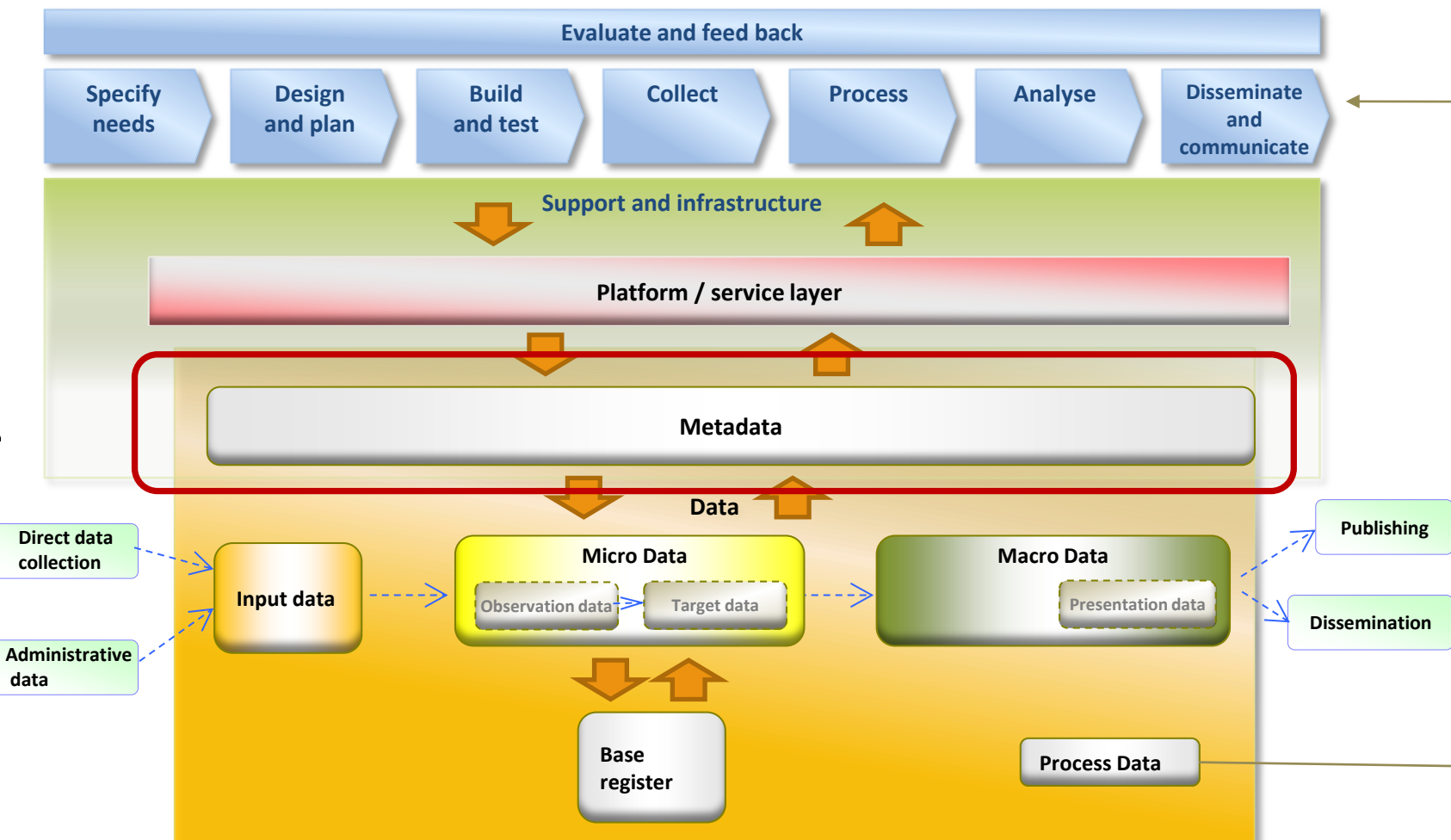




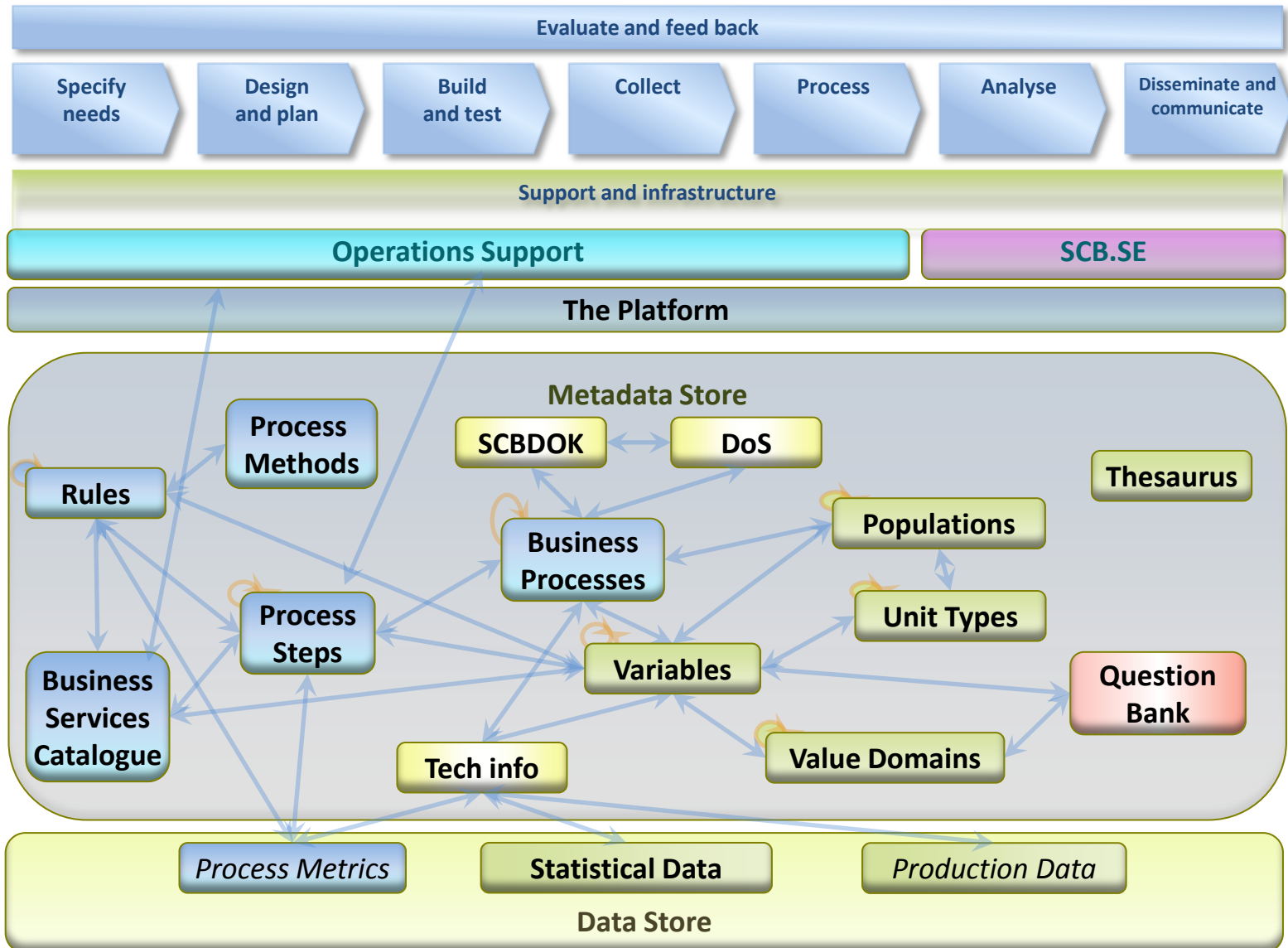
SCB's strategy on coordinated statistics production

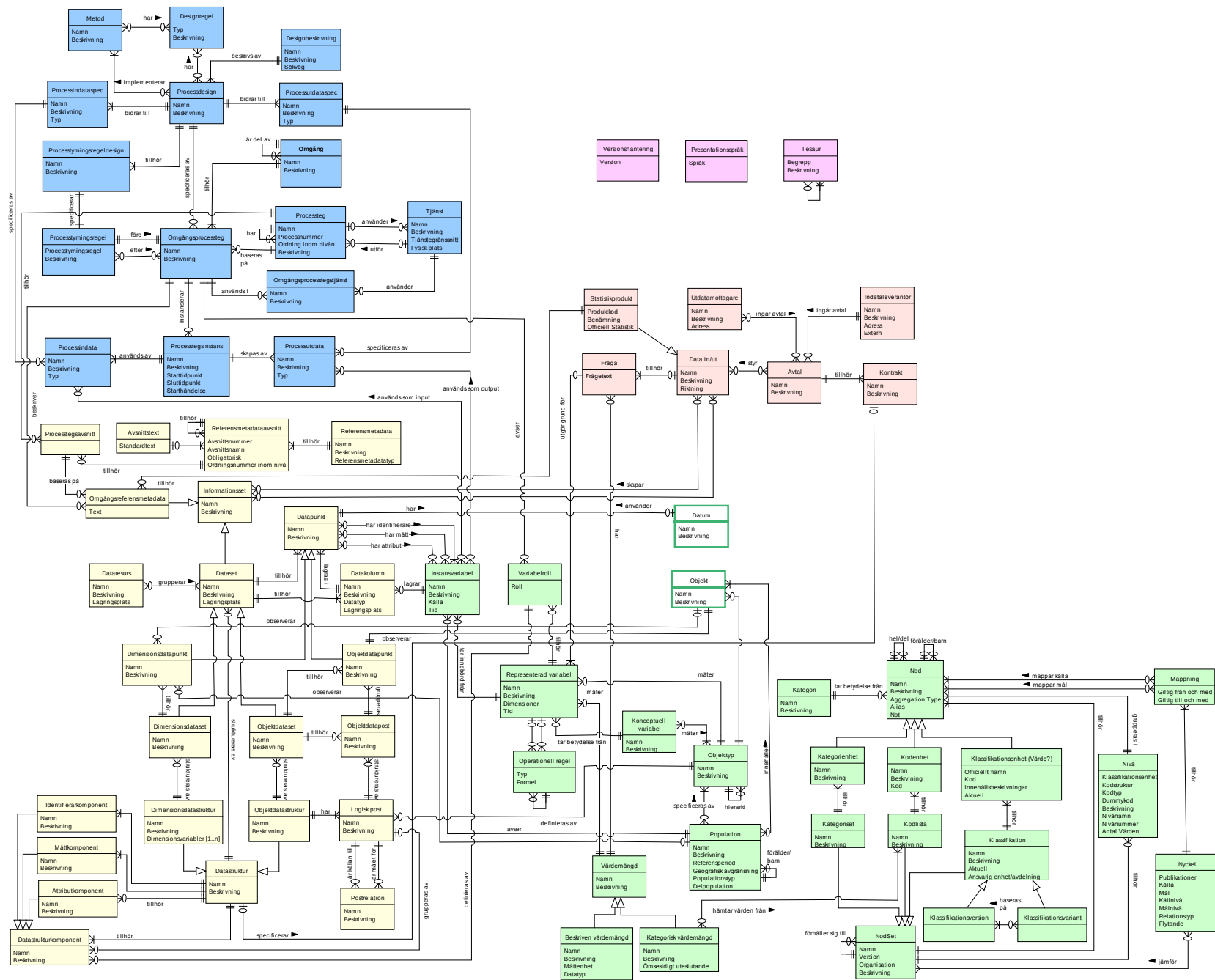
Statistics Sweden

Statistiska centralbyrån



A Central Metadata Repository

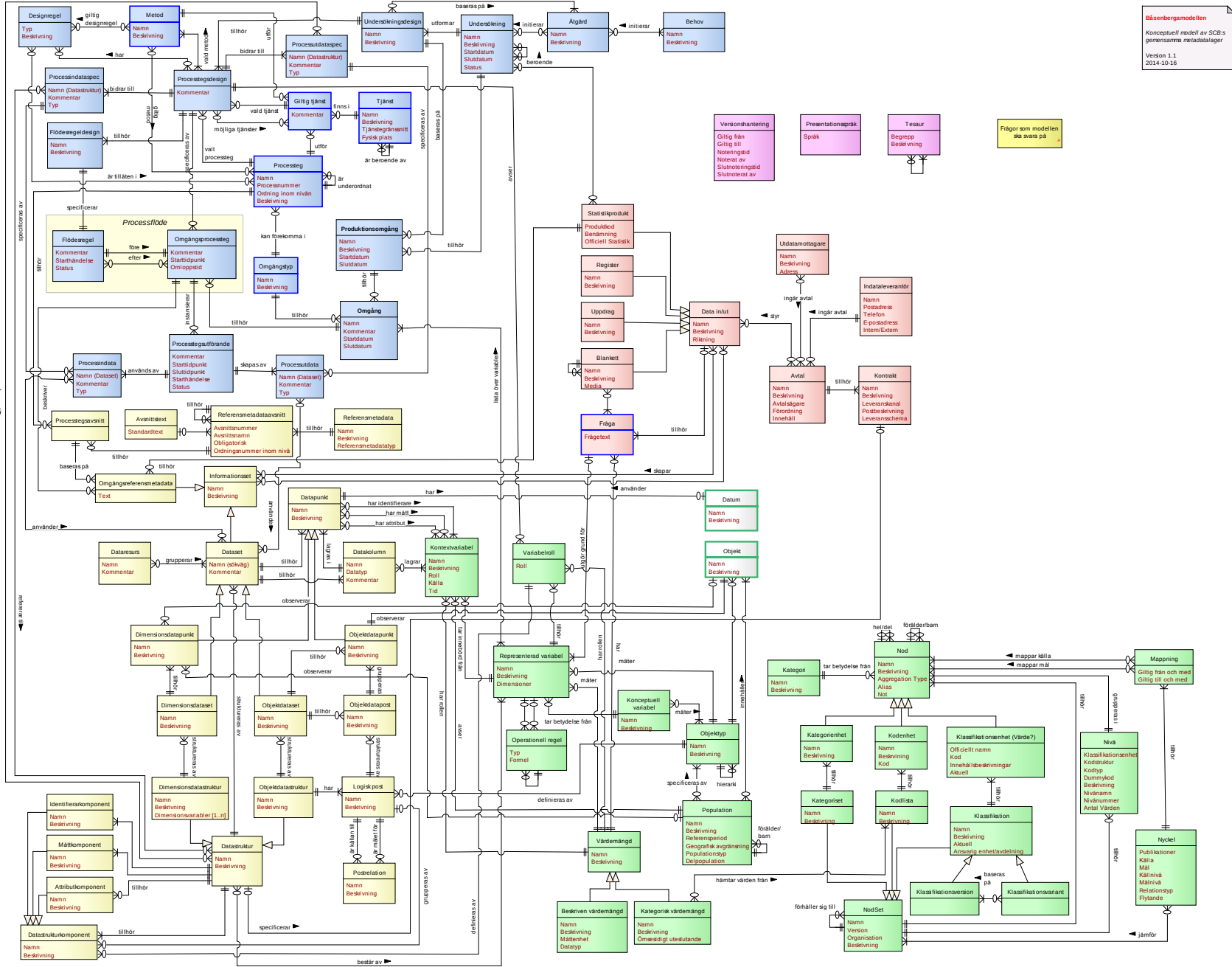




Ongoing

- Prototype - Access database
- Coordinating economic short term statistics (5 products)
- Testing in Structural business statistics and National accounts (Excessive Deficit Procedure - EDP)
- Case on sampling filling the blue part
- Version 1.1
- Next step - Incorporating in/with the platform





The end...

