Fiber, a new production system based on GSIM.

Statistics Sweden
Patrik Wahlgren
Agenda

• Statistical Program Design and Process Design
• Dataset catalog
• Visualization of the dataset catalog with a graph database
Statistical Program design and Process Design
Processes

(Statistical Program)
(Statistical Program Design)
(Business Process)
(Process Step)
(Process Design)
“At some level it will be appropriate to consider a Process Step to be a discrete task (unit of work) without warranting further subdivision.”
“At some level it will be appropriate to consider a Process Step to be a discrete task (unit of work) without warranting further subdivision.”
“At some level it will be appropriate to consider a Process Step to be a discrete task (unit of work) without warranting further subdivision.”
“At some level it will be appropriate to consider a Process Step to be a discrete task (unit of work) without warranting further subdivision.”
“At some level it will be appropriate to consider a Process Step to be a discrete task (unit of work) without warranting further subdivision.”
“At some level it will be appropriate to consider a Process Step to be a discrete task (unit of work) without warranting further subdivision.”
"At some level it will be appropriate to consider a Process Step to be a discrete task (unit of work) without warranting further subdivision."
“At some level it will be appropriate to consider a *Process Step* to be a discrete task (unit of work) without warranting further subdivision.”
New production system based on GSIM, “Fiber”.

- Import (Provision agreement)
- Calculation
- Validate
- Transcoding
- Manual editing
- Analysis and reports
- Export of datasets for reporting
Dataset catalog
Workflow execution

GSIM objects:
- Statistical Program
- Statistical Program Cycle (T)

GSIM objects:
- Concepts
- Codelist/Classification
- Rep. Variable/Variable
- Unittype
- Population
- Valuedomain
- Etc..

GSIM: ProcessDesign (User Interface: ProcessStepDesignArea)
SAS, VTL, etc..
(Mapping, Edit, Validering etc.)

Business Service (SAS, VTL, Edit Mapping, import…)

VTL
Manual edit
Validation
Mappning
Reports

GSIM: ProcessMethod
GSIM: Rule

GSIM objects:
- Statistical Program
- Statistical Program Cycle (T)

GSIM objects:
- Concepts
- Codelist/Classification
- Rep. Variable/Variable
- Unittype
- Population
- Valuedomain
- Etc.
Services (Dataset, Datapoint and Datastructure)

GSIM objects:
- Statistical Program
- Statistical Program Cycle (T)

GSIM objects:
- Concepts
- Codelist/Classification
- Rep. Variable/Variable
- Unittype
- Population
- Valuedomain
- Etc..

GSIM: Process Step Instance

SAS, VTL, etc. (Mapping, Edit, Validation, etc.)

Datapoint Service

Datastructure Service

GSIM: Business Services

Dataset Service

GSIM: ProcessMethod

GSIM: Rule

GSIM objects:
- Statistical Program
- Statistical Program Cycle (T)

GSIM objects:
- Concepts
- Codelist/Classification
- Rep. Variable/Variable
- Unittype
- Population
- Valuedomain
- Etc..
Handoverpoints (Checkpoints, Qualitygates or Steady states)
Relative references with “T” for access to datasets

All datasets has an URI or a path: \\
\\[StatisticalProgram\\Statistical Program Cycle\T\\Dataset Name

Example:
“NAPRSQ\2022Q2\A calculation result”

Datasets can be accessed relative “T-x”, “T” is set on the “Statistical Program Cycle”.
Visualization of the dataset catalog with a graph database
This is a test using a graph database (NEO4J) to display all the dataset in the dataset catalog.
The picture shows a couple of the datasets created in production cycle for quarter 2, 2020 and for the statistical program called NRPSRQ.
The datasets specifications in design are now dataset's. (TransformedOutput)
<table>
<thead>
<tr>
<th>Sektor</th>
<th>Cofog</th>
<th>Area</th>
<th>Transaction</th>
<th>Coicop</th>
<th>Unit</th>
<th>Ref.Tid</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>1002</td>
<td>46</td>
<td>S1</td>
<td>701</td>
<td>1</td>
<td>2021</td>
<td>1213</td>
</tr>
<tr>
<td>7</td>
<td>1002</td>
<td>46</td>
<td>S2</td>
<td>701</td>
<td>1</td>
<td>2021</td>
<td>332</td>
</tr>
<tr>
<td>7</td>
<td>1003</td>
<td>46</td>
<td>S1</td>
<td>702</td>
<td>1</td>
<td>2021</td>
<td>2355</td>
</tr>
<tr>
<td>7</td>
<td>1003</td>
<td>46</td>
<td>S2</td>
<td>702</td>
<td>1</td>
<td>2021</td>
<td>789</td>
</tr>
</tbody>
</table>

This is not correct content!!
You could see the rule that created this dataset.
The edges (ProcessInput) correspond to the lines (ProcessInputSpecification’s) drawn in the process design.
Patrik Wahlgren

Business architect
Methodology and architecture governance and Enterprise architecture
SCB, Statistics Sweden

Mail: patrik.wahlgren@scb.se