Fiber, a new productionsystem based on GSIM.

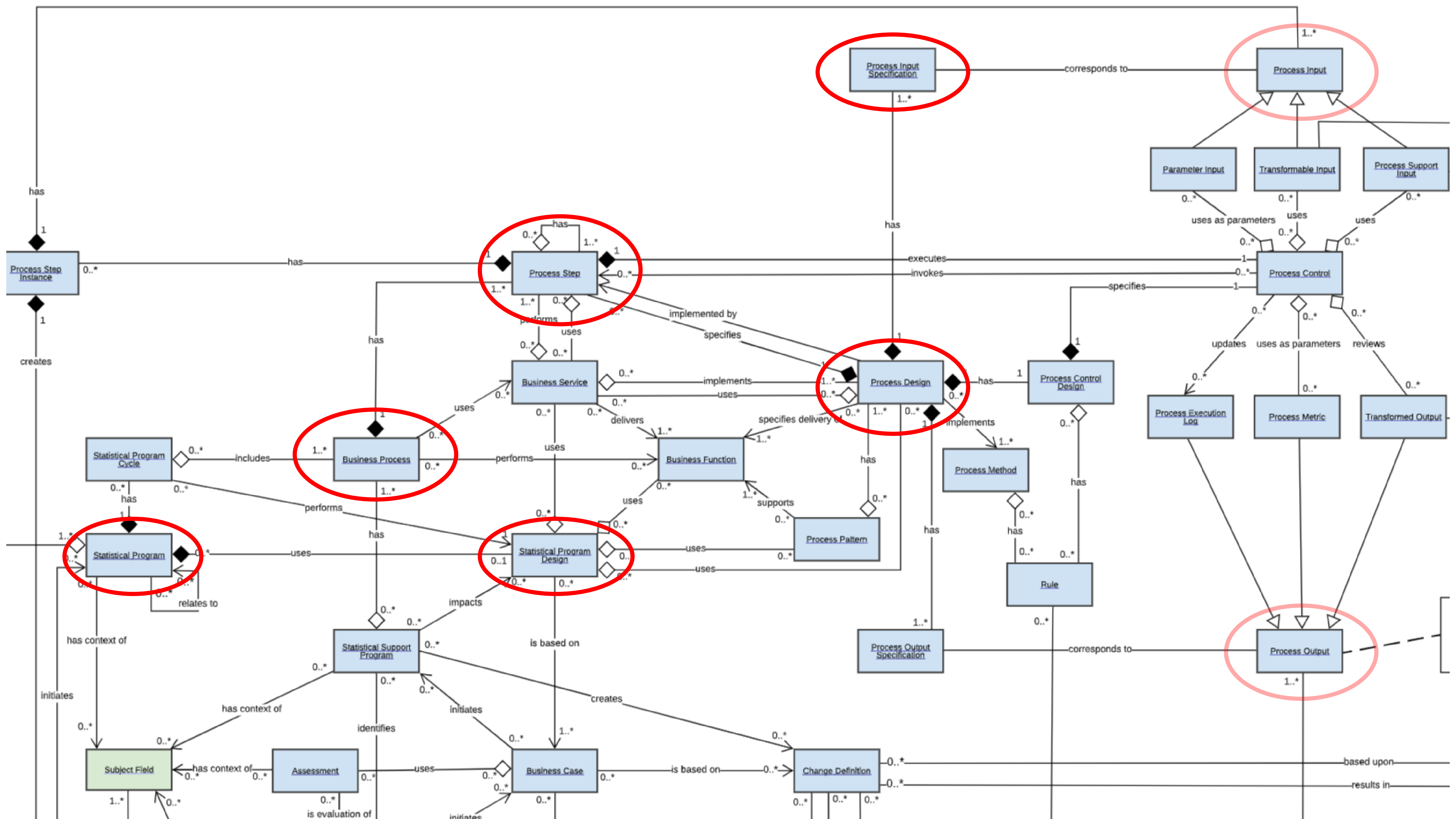
Statistics Sweden

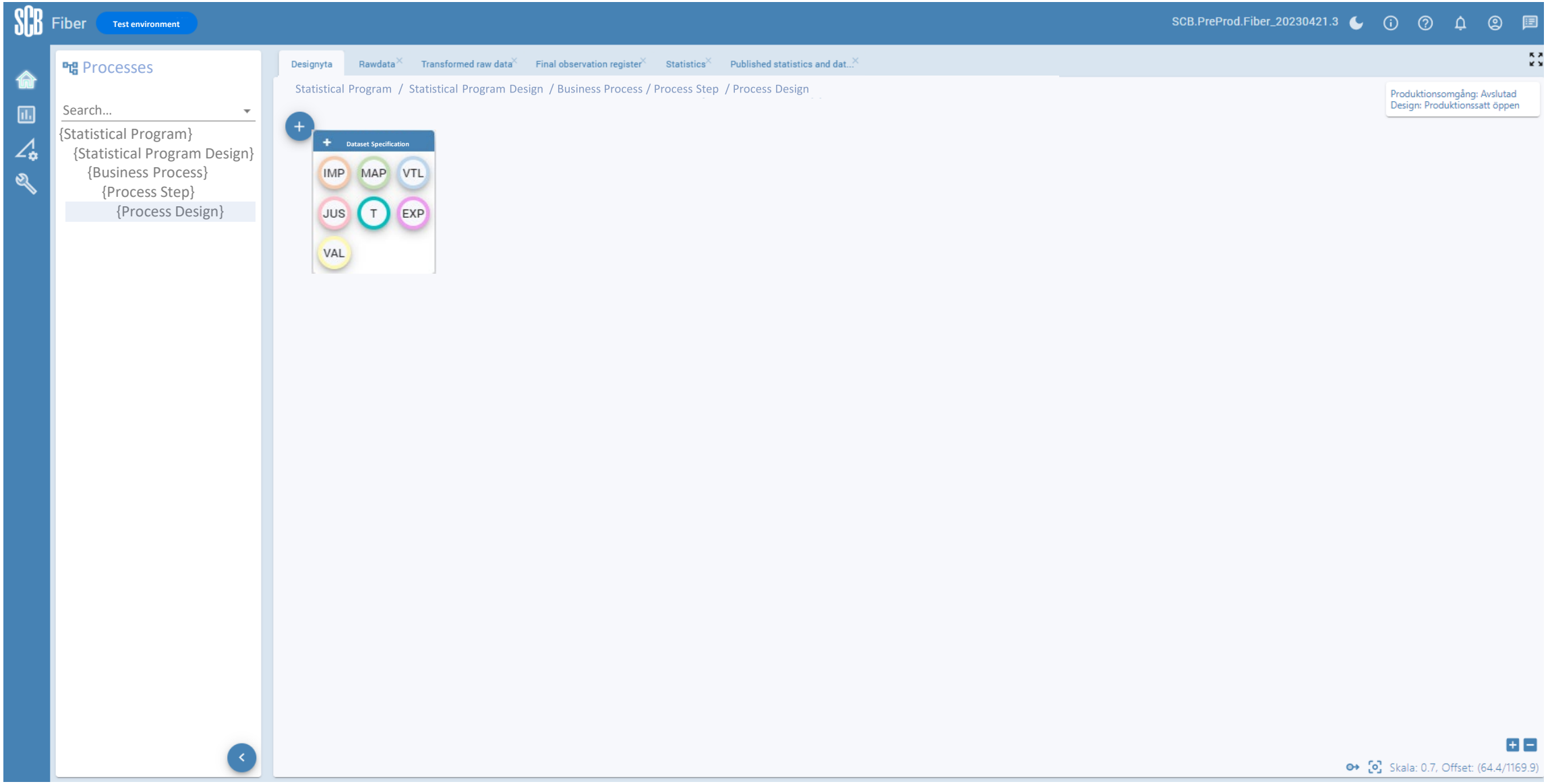
Patrik Wahlgren

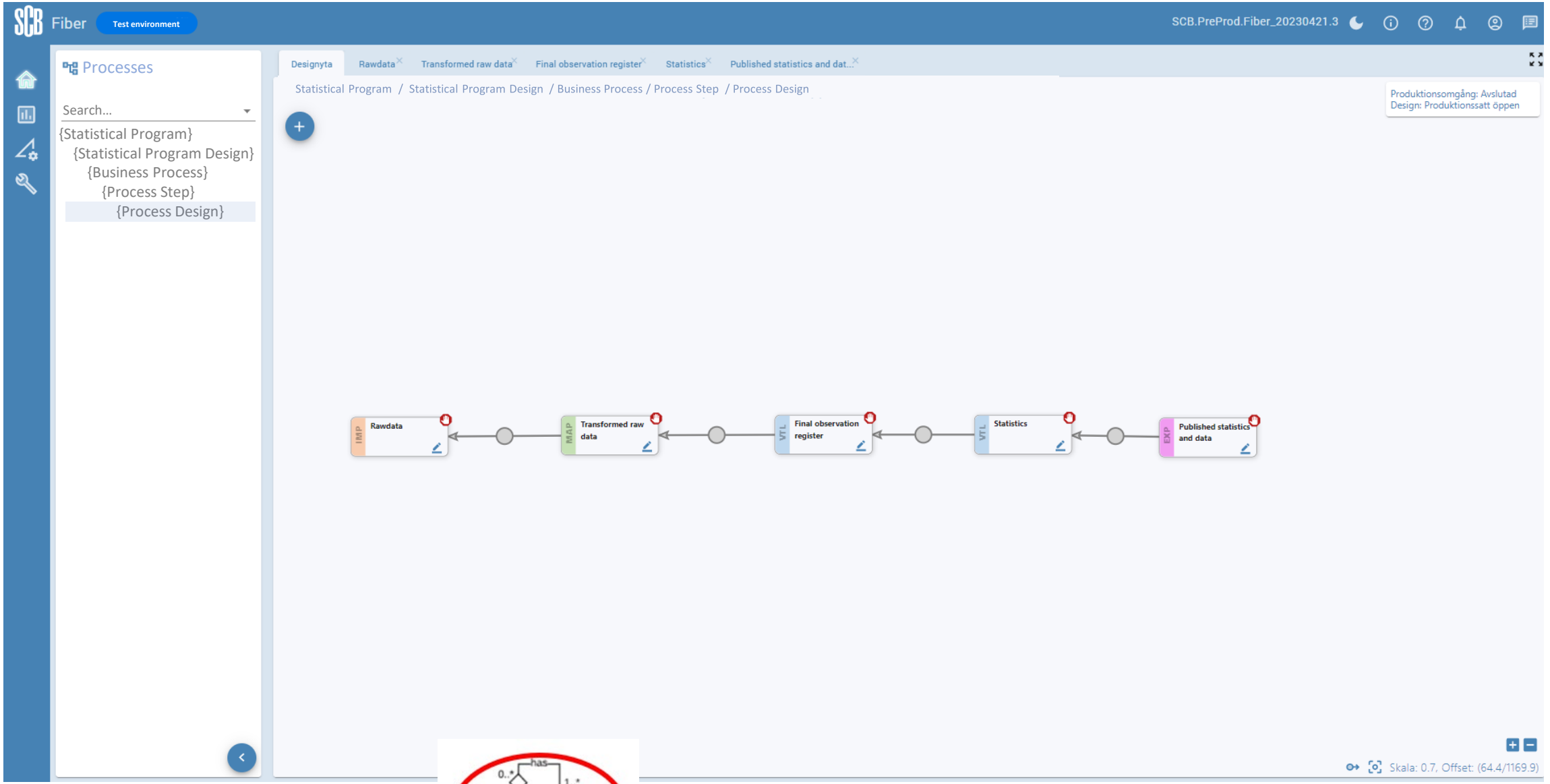
Agenda

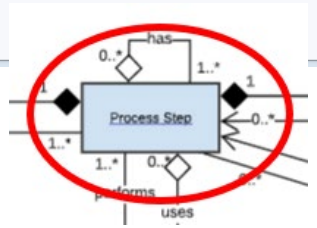
- Statistical Program Design and Process Design
- Dataset catalog
- Visualization of the dataset catalog with a graph database

Statistical Program design and Process Design

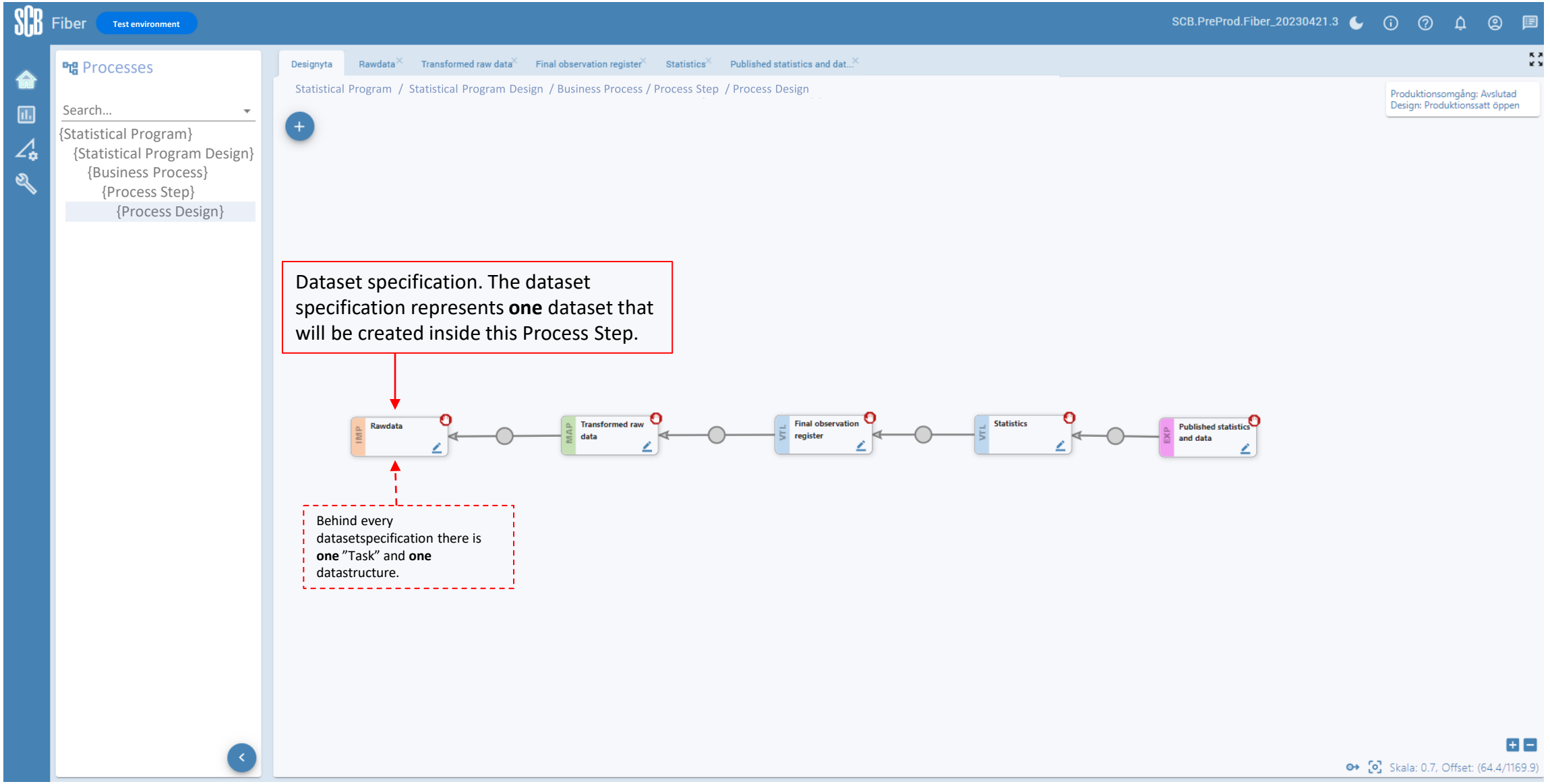

Process Input Specification
corresponds to
Process Input
Parameter Input
Transformable Input
Process Support Input
uses as parameters
uses
uses
has
Process Step Instance
has
has
creates
Process Step
has
executes
invokes
Process Control
specifies
implemented by
specifies
performs
uses
updates
uses as parameters
reviews
Business Service
implements
uses
Process Design
has
Process Control Design
Process Execution Log
Process Metric
Transformed Output
uses
delivers
specifies delivery of
implements
Statistical Program Cycle
includes
Business Process
performs
uses
Business Function
Process Method
has
has
has
has
performs
supports
Process Pattern
has
Statistical Program
uses
Statistical Program Design
uses
uses
has
relates to
impacts
Rule
has context of
initiates
Statistical Support Program
is based on
Process Output Specification
corresponds to
Process Output
has context of
identifies
initiates
creates
Subject Field
has context of
Assessment
uses
Business Case
is based on
Change Definition
based upon
results in
is evaluation of
initiates


SCB
Fiber
Test environment
SCB.PreProd.Fiber_20230421.3
Processes
Search...
{Statistical Program}
{Statistical Program Design}
{Business Process}
{Process Step}
{Process Design}
Designyta
Rawdata
Transformed raw data
Final observation register
Statistics
Published statistics and dat...
Statistical Program / Statistical Program Design / Business Process / Process Step / Process Design
Produktionsomgång: Avslutad
Design: Produktionssatt öppen
Dataset Specification
IMP
MAP
VTL
JUS
T
EXP
VAL
Skala: 0.7, Offset: (64.4/1169.9)

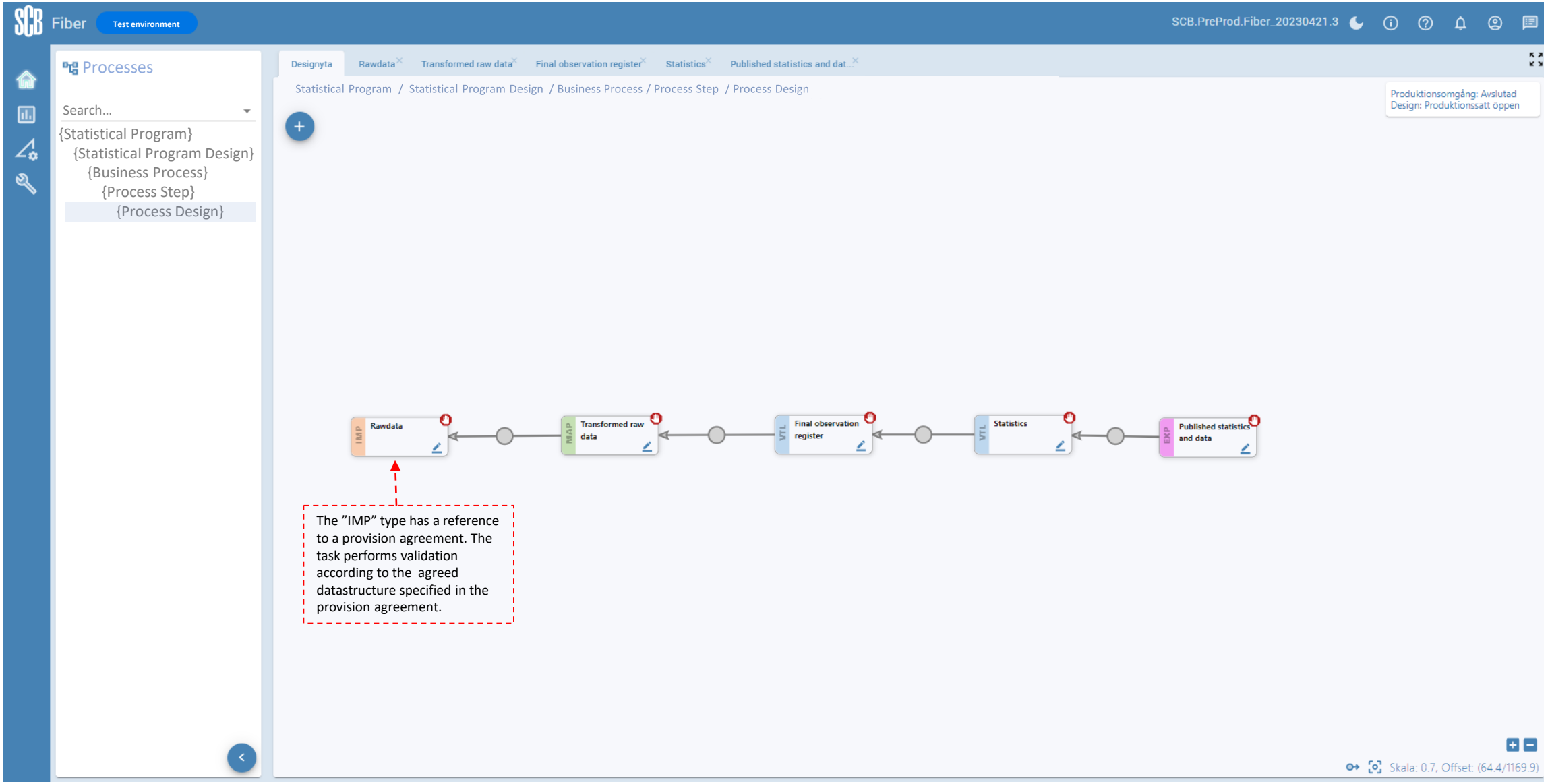




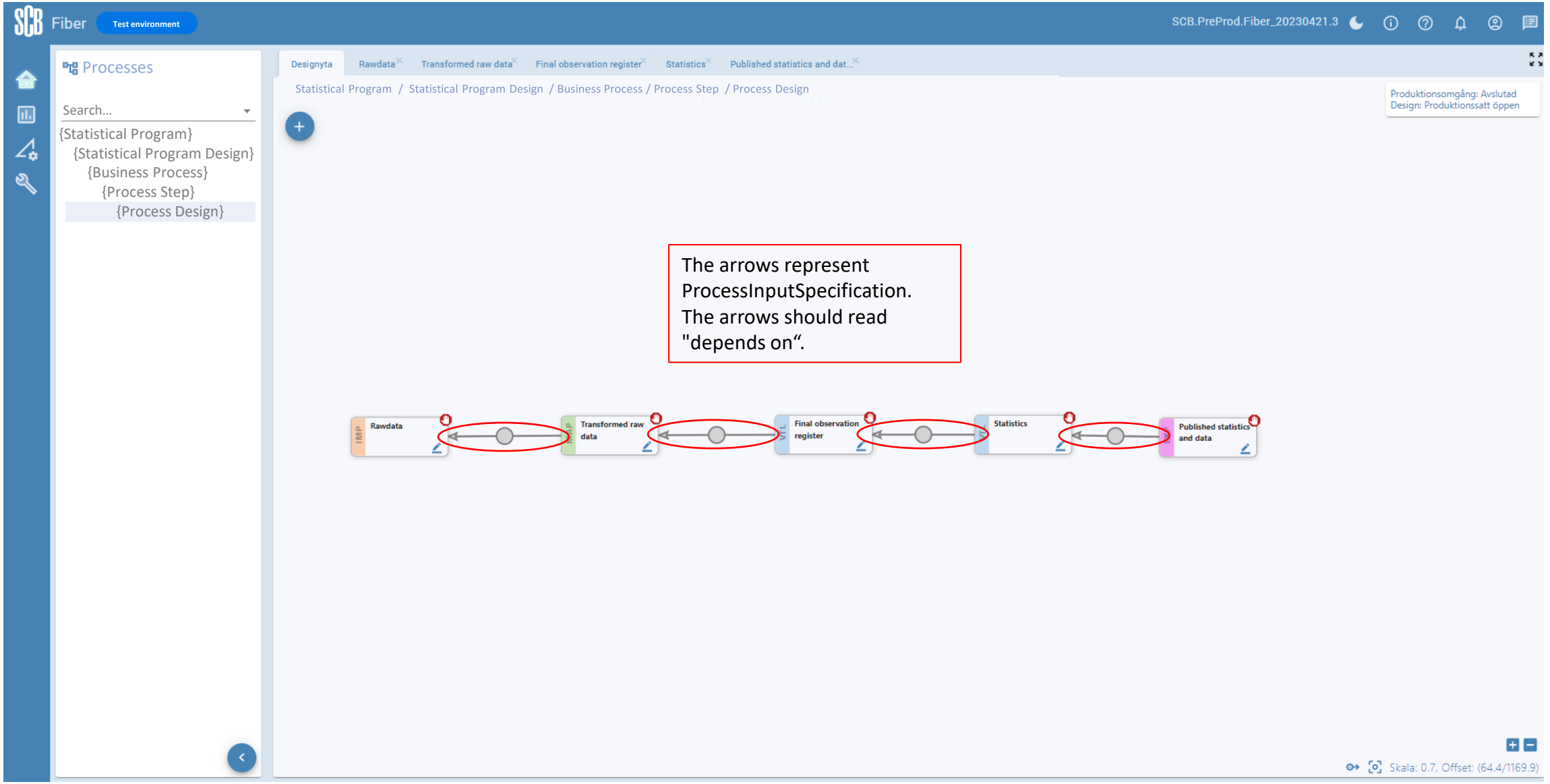
"At some level it will be appropriate to consider a *Process Step* to be a discrete task (unit of work) without warranting further subdivision."



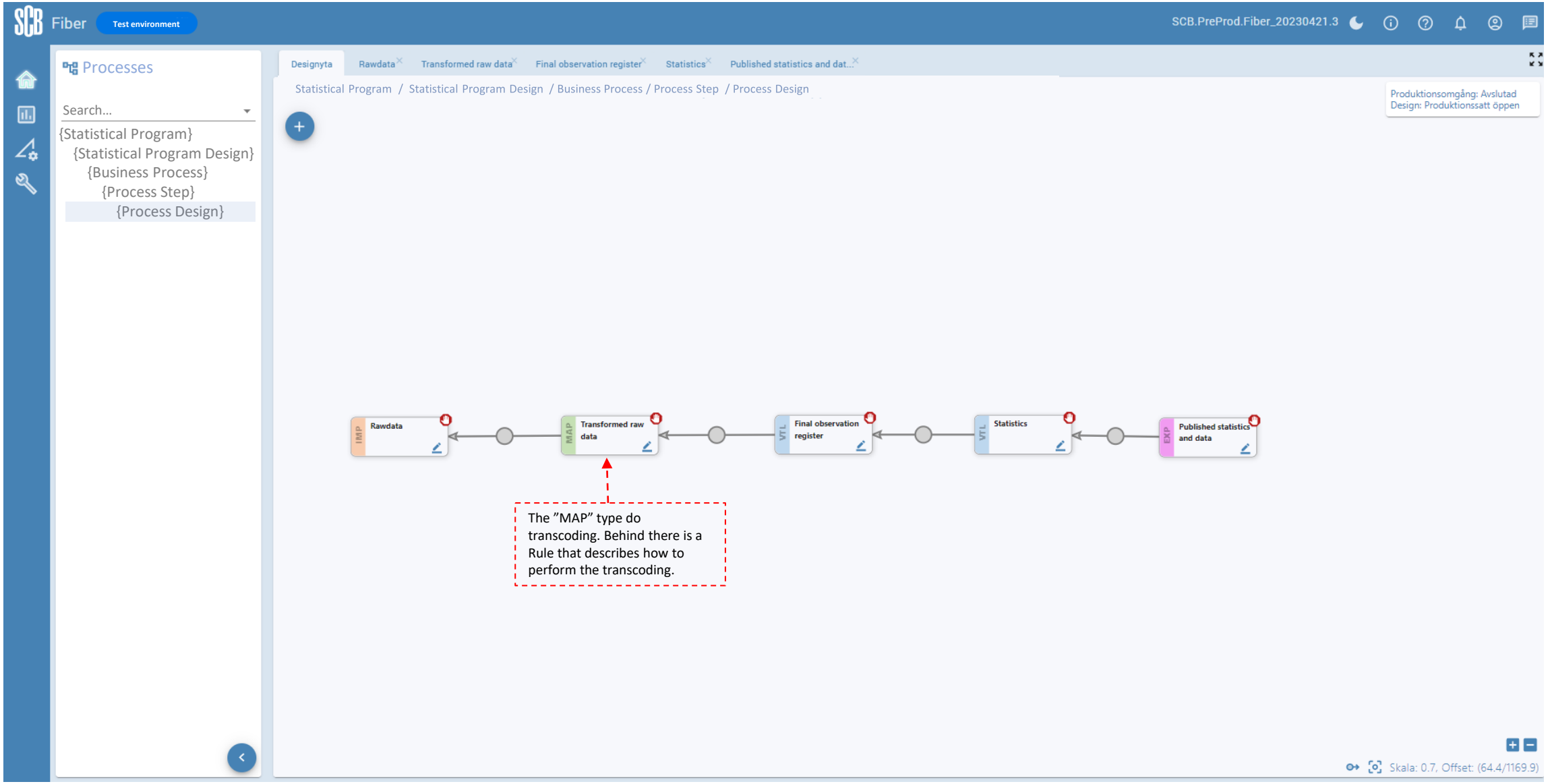
"At some level it will be appropriate to consider a *Process Step* to be a discrete task (unit of work) without warranting further subdivision."



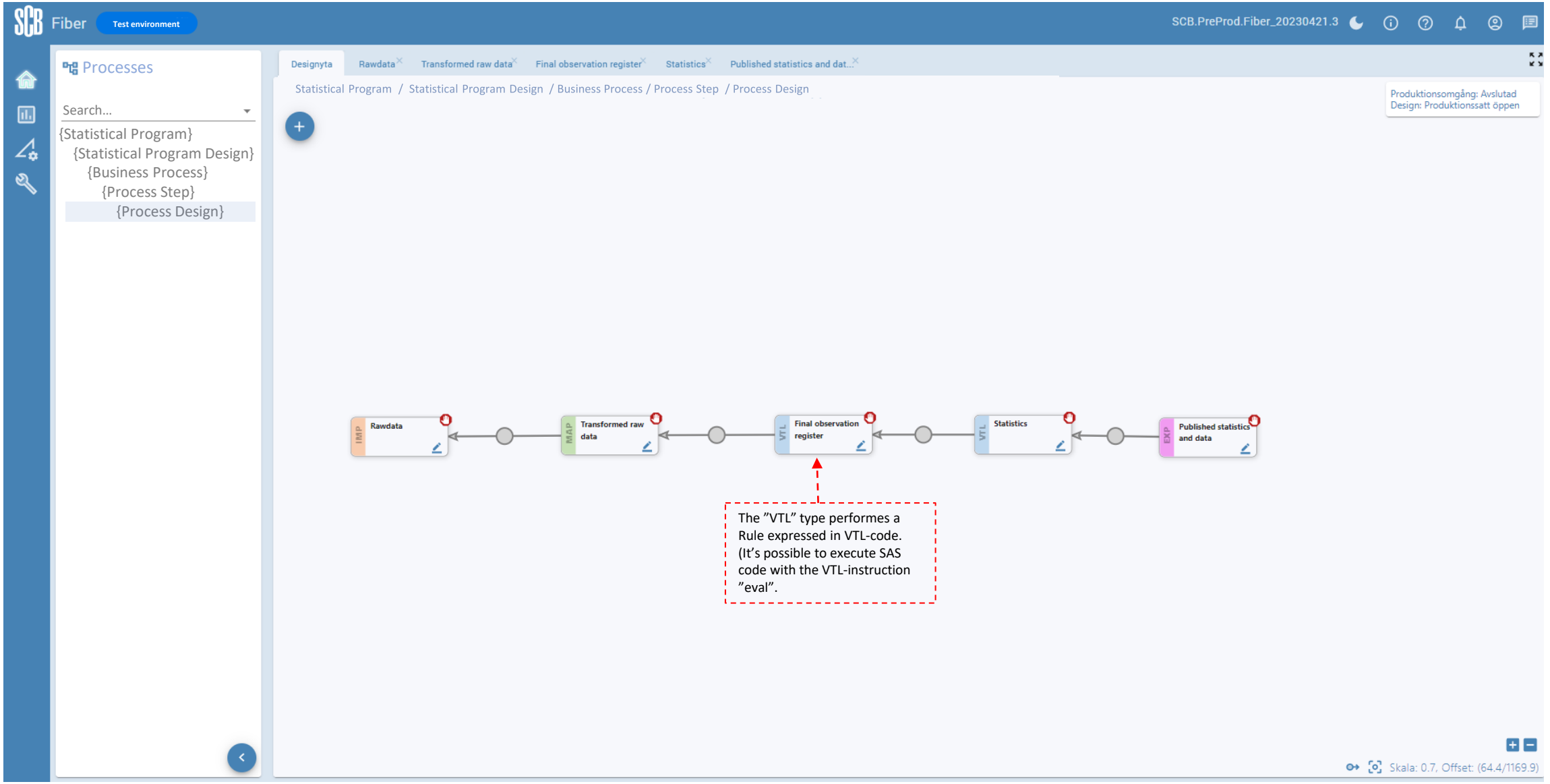
“At some level it will be appropriate to consider a *Process Step* to be a discrete task (unit of work) without warranting further subdivision.”



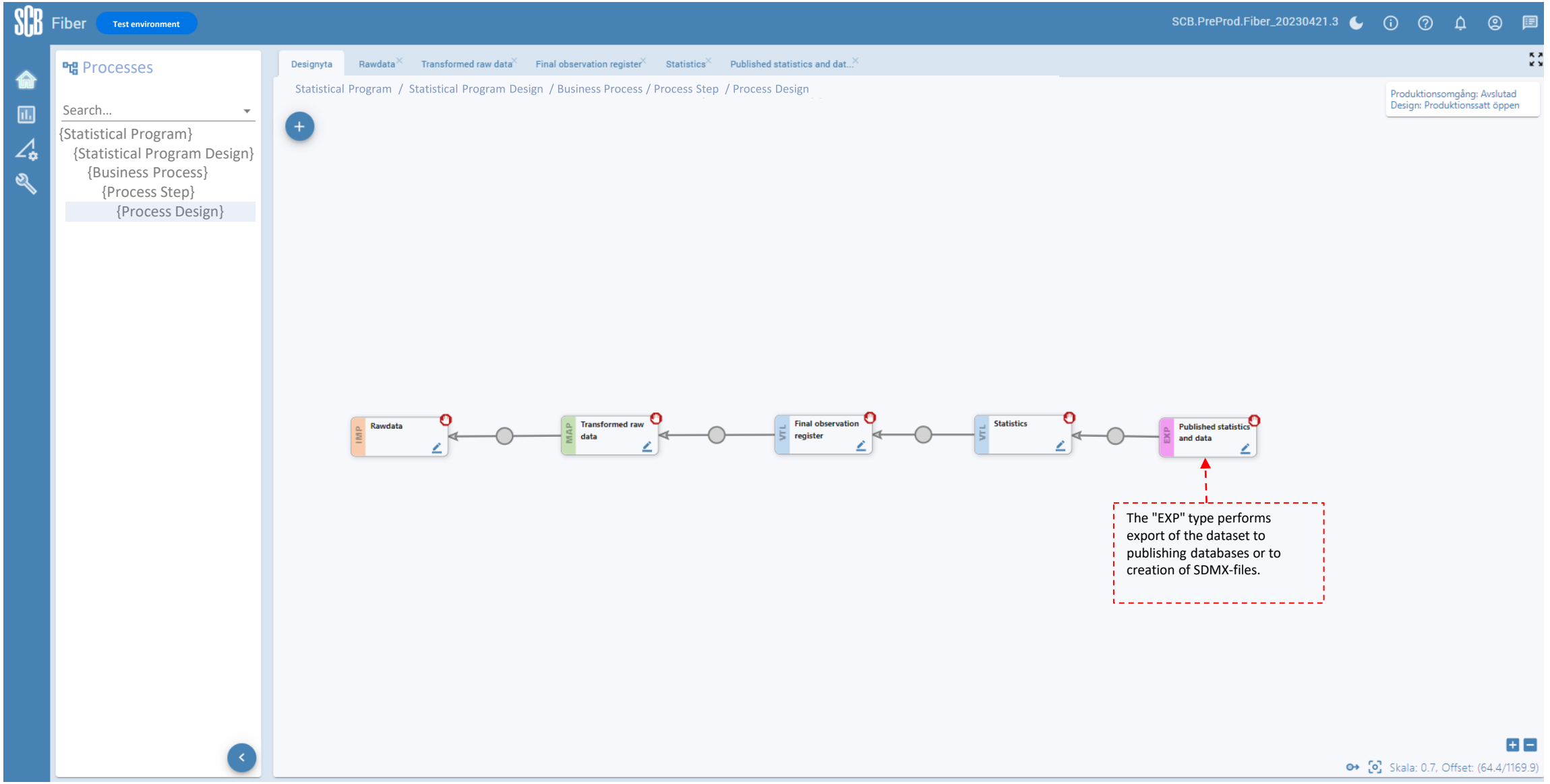
"At some level it will be appropriate to consider a *Process Step* to be a discrete task (unit of work) without warranting further subdivision."



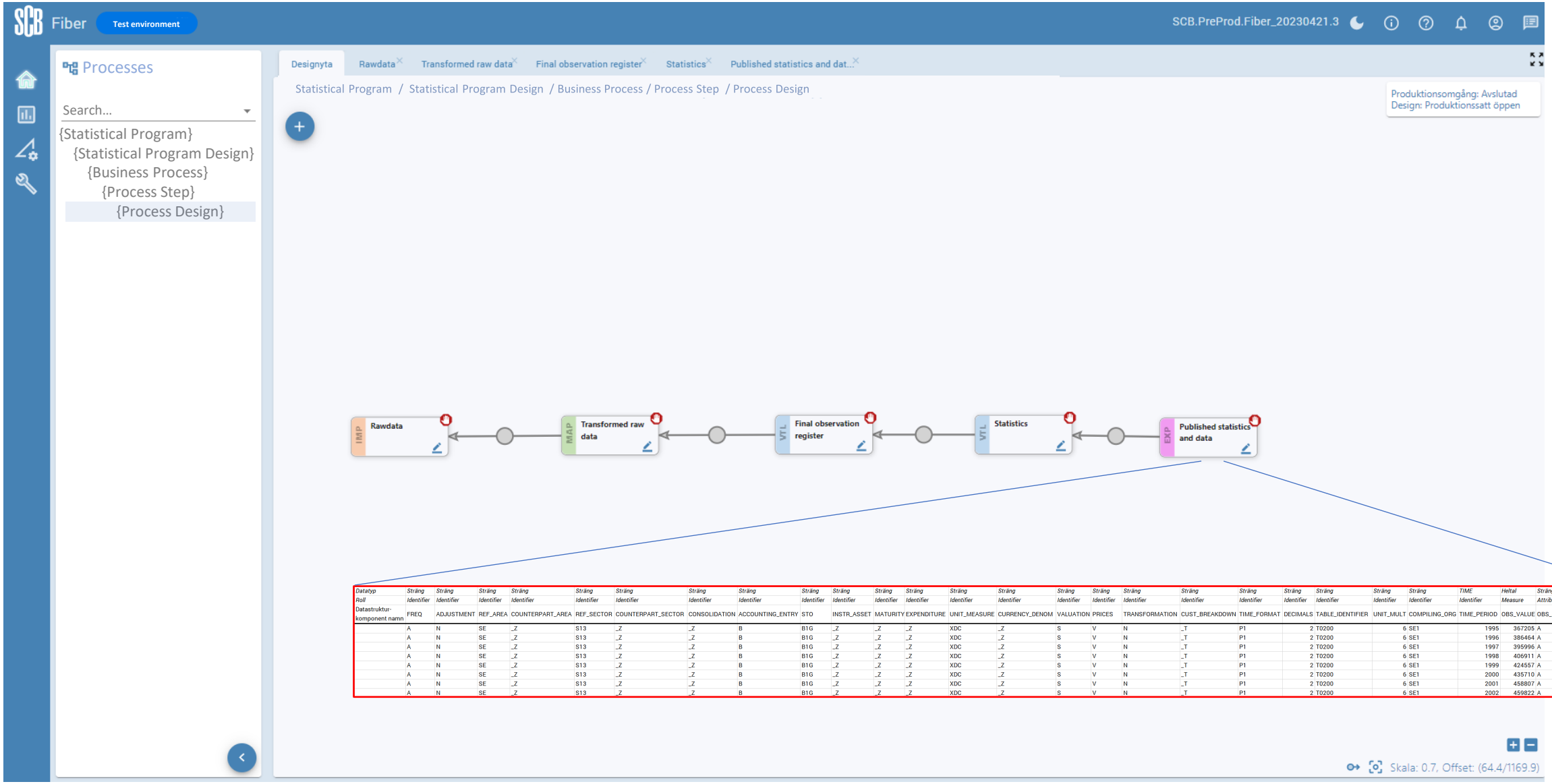
“At some level it will be appropriate to consider a *Process Step* to be a discrete task (unit of work) without warranting further subdivision.”



“At some level it will be appropriate to consider a *Process Step* to be a discrete task (unit of work) without warranting further subdivision.”

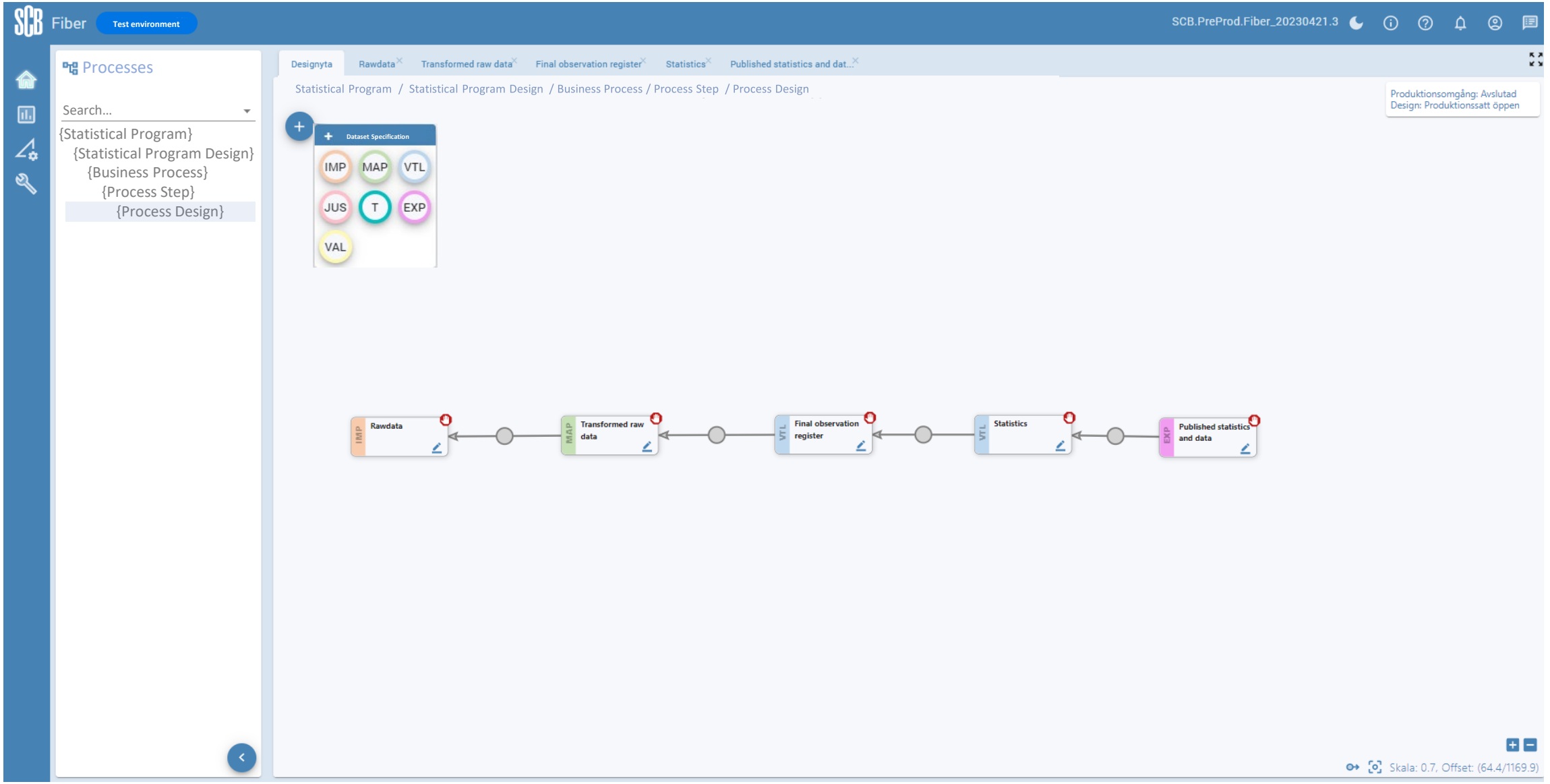


"At some level it will be appropriate to consider a *Process Step* to be a discrete task (unit of work) without warranting further subdivision."



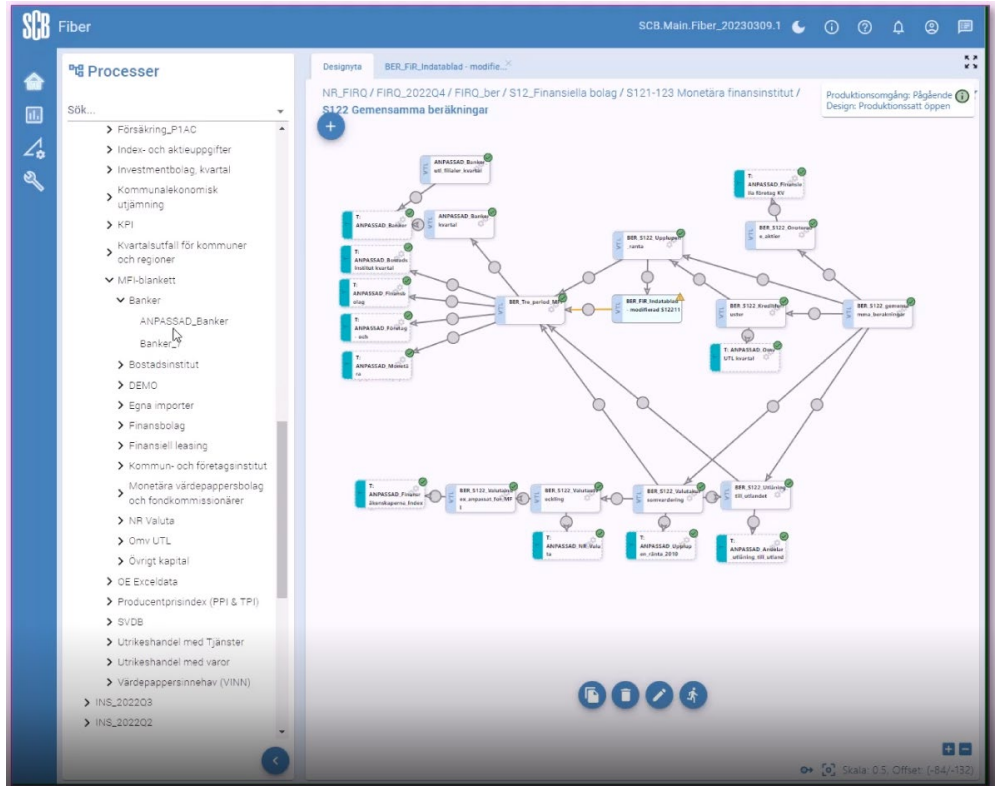
Datatyp	Sträng	Sträng	Sträng	Sträng	Sträng	Sträng	Sträng	Sträng	Sträng	Sträng	Sträng	Sträng	Sträng	Sträng	Sträng	Sträng	Sträng	Sträng	Sträng	Sträng	Sträng	Sträng	TIME	Heltal	Sträng	Sträng
Roll	Identifier	Identifier	Identifier	Identifier	Identifier	Identifier	Identifier	Identifier	Identifier	Identifier	Identifier	Identifier	Identifier	Identifier	Identifier	Identifier	Identifier	Identifier	Identifier	Identifier	Identifier	Identifier	Identifier	Measure	Attribute	Attribute
Datastruktur-komponent namn	FREQ	ADJUSTMENT	REF_AREA	COUNTERPART_AREA	REF_SECTOR	COUNTERPART_SECTOR	CONSOLIDATION	ACCOUNTING_ENTRY	STO	INSTR_ASSET	MATURITY	EXPENDITURE	UNIT_MEASURE	CURRENCY_DENOM	VALUATION	PRICES	TRANSFORMATION	CUST_BREAKDOWN	TIME_FORMAT	DECIMALS	TABLE_IDENTIFIER	UNIT_MULT	COMPILING_ORG	TIME_PERIOD	OBS_VALUE	OBS_STATUS	CONF_STATUS
	A	N	SE	_Z	S13	_Z	_Z	B	B1G	_Z	_Z	_Z	XDC	_Z	S	V	N	_T	P1	2	T0200	6	SE1	1995	367205	A	F
	A	N	SE	_Z	S13	_Z	_Z	B	B1G	_Z	_Z	_Z	XDC	_Z	S	V	N	_T	P1	2	T0200	6	SE1	1996	386464	A	F
	A	N	SE	_Z	S13	_Z	_Z	B	B1G	_Z	_Z	_Z	XDC	_Z	S	V	N	_T	P1	2	T0200	6	SE1	1997	395996	A	F
	A	N	SE	_Z	S13	_Z	_Z	B	B1G	_Z	_Z	_Z	XDC	_Z	S	V	N	_T	P1	2	T0200	6	SE1	1998	406911	A	F
	A	N	SE	_Z	S13	_Z	_Z	B	B1G	_Z	_Z	_Z	XDC	_Z	S	V	N	_T	P1	2	T0200	6	SE1	1999	424557	A	F
	A	N	SE	_Z	S13	_Z	_Z	B	B1G	_Z	_Z	_Z	XDC	_Z	S	V	N	_T	P1	2	T0200	6	SE1	2000	435710	A	F
	A	N	SE	_Z	S13	_Z	_Z	B	B1G	_Z	_Z	_Z	XDC	_Z	S	V	N	_T	P1	2	T0200	6	SE1	2001	458807	A	F
	A	N	SE	_Z	S13	_Z	_Z	B	B1G	_Z	_Z	_Z	XDC	_Z	S	V	N	_T	P1	2	T0200	6	SE1	2002	459822	A	F

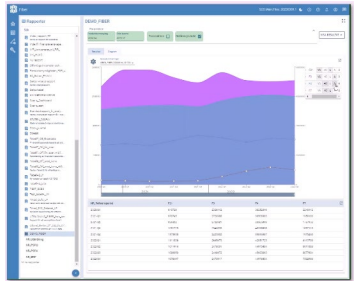
"At some level it will be appropriate to consider a *Process Step* to be a discrete task (unit of work) without warranting further subdivision."

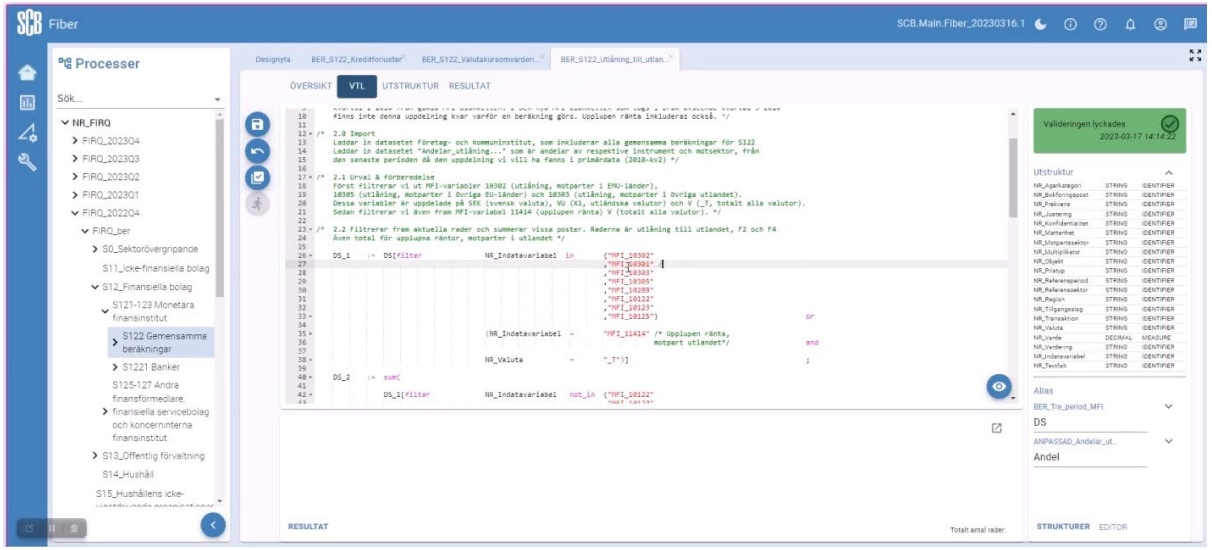

Fiber
Test environment
SCB.PreProd.Fiber_20230421.3
Processes
Search...
{Statistical Program}
{Statistical Program Design}
{Business Process}
{Process Step}
{Process Design}
Designyta
Rawdata
Transformed raw data
Final observation register
Statistics
Published statistics and dat...
Statistical Program / Statistical Program Design / Business Process / Process Step / Process Design
Produktionsomgång: Avslutad
Design: Produktionssatt öppen
Dataset Specification
IMP
MAP
VTL
JUS
T
EXP
VAL
IMP
Rawdata
MAP
Transformed raw data
VTL
Final observation register
VTL
Statistics
EXP
Published statistics and data
Skala: 0.7, Offset: (64.4/1169.9)

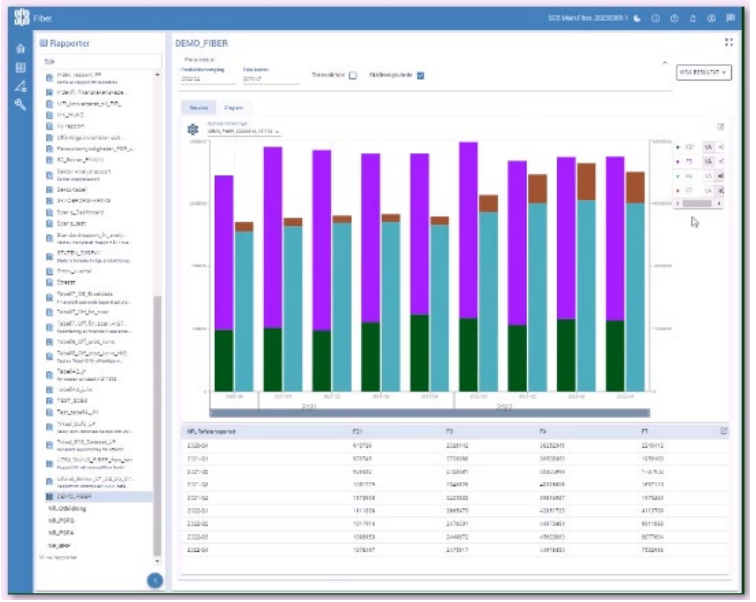
New production system based on GSIM, “Fiber”.

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



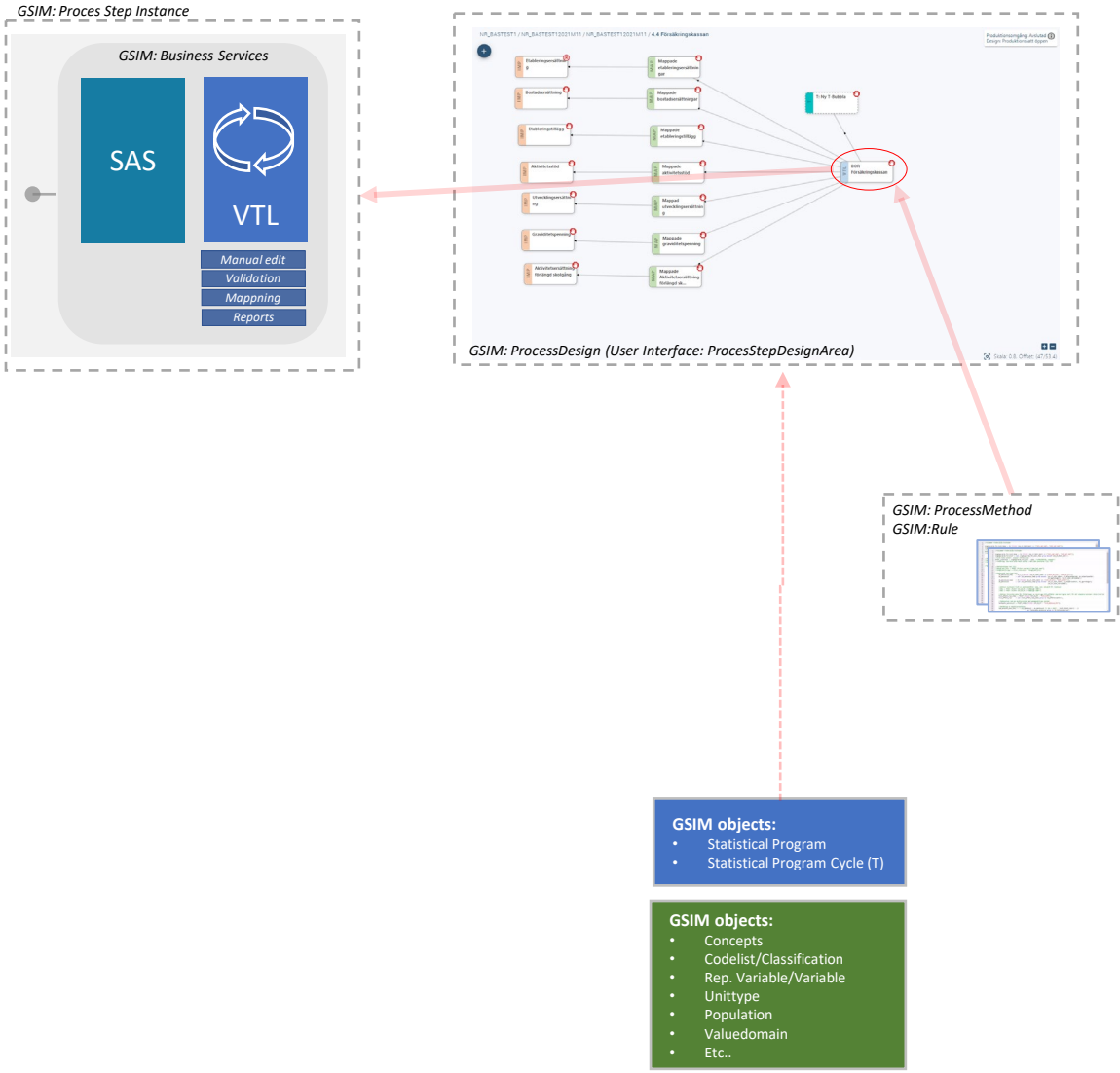




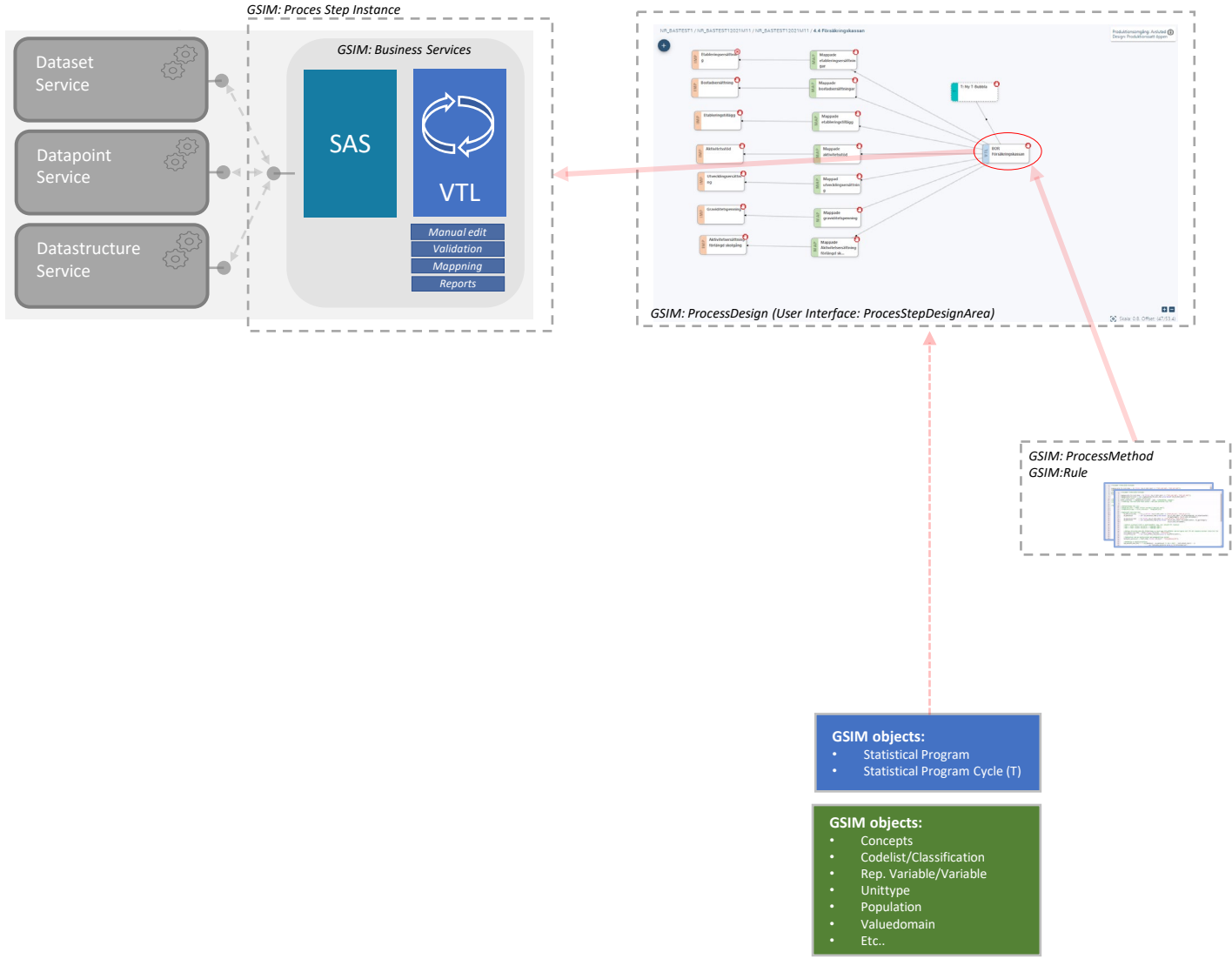


Dataset catalog

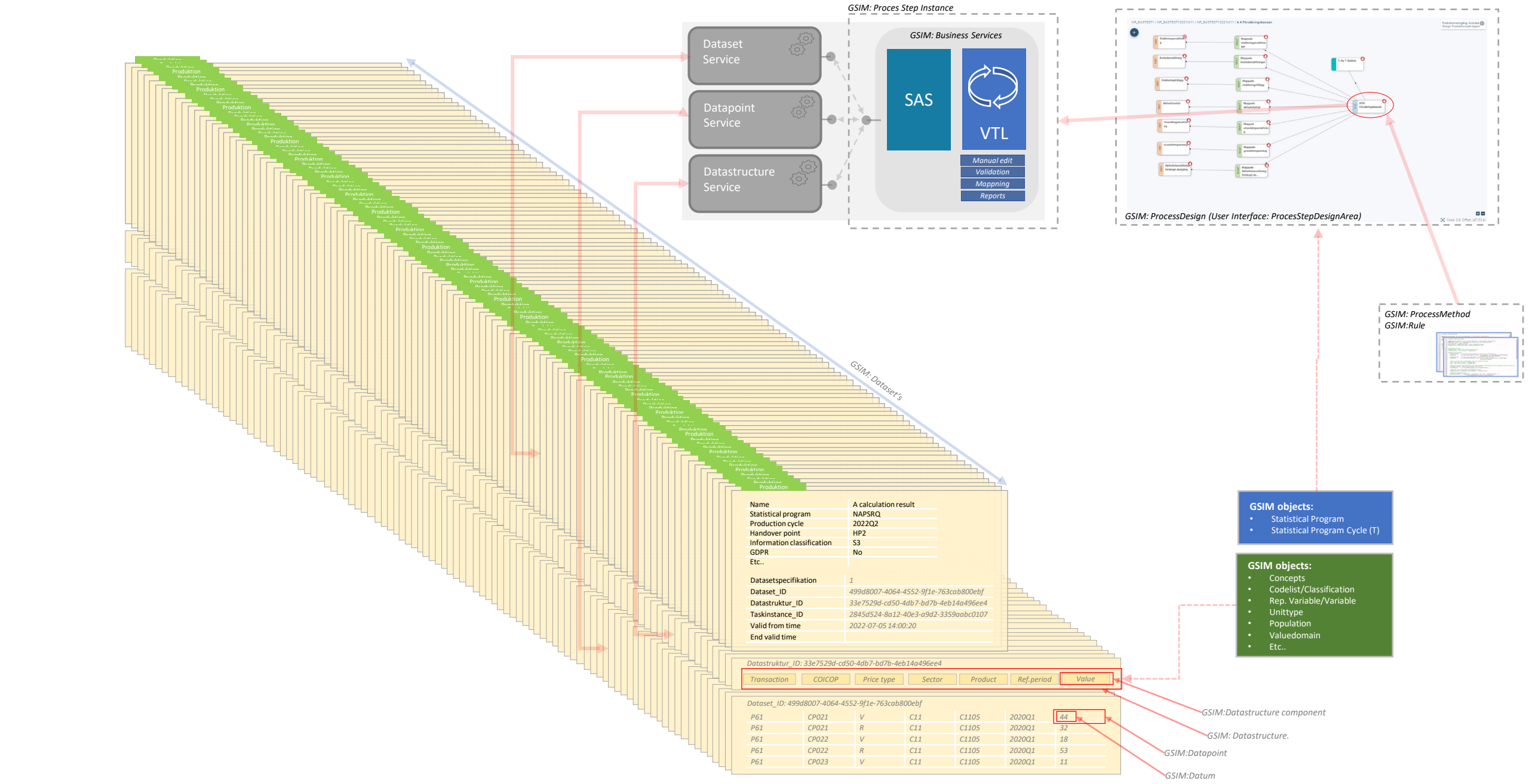
Workflow execution



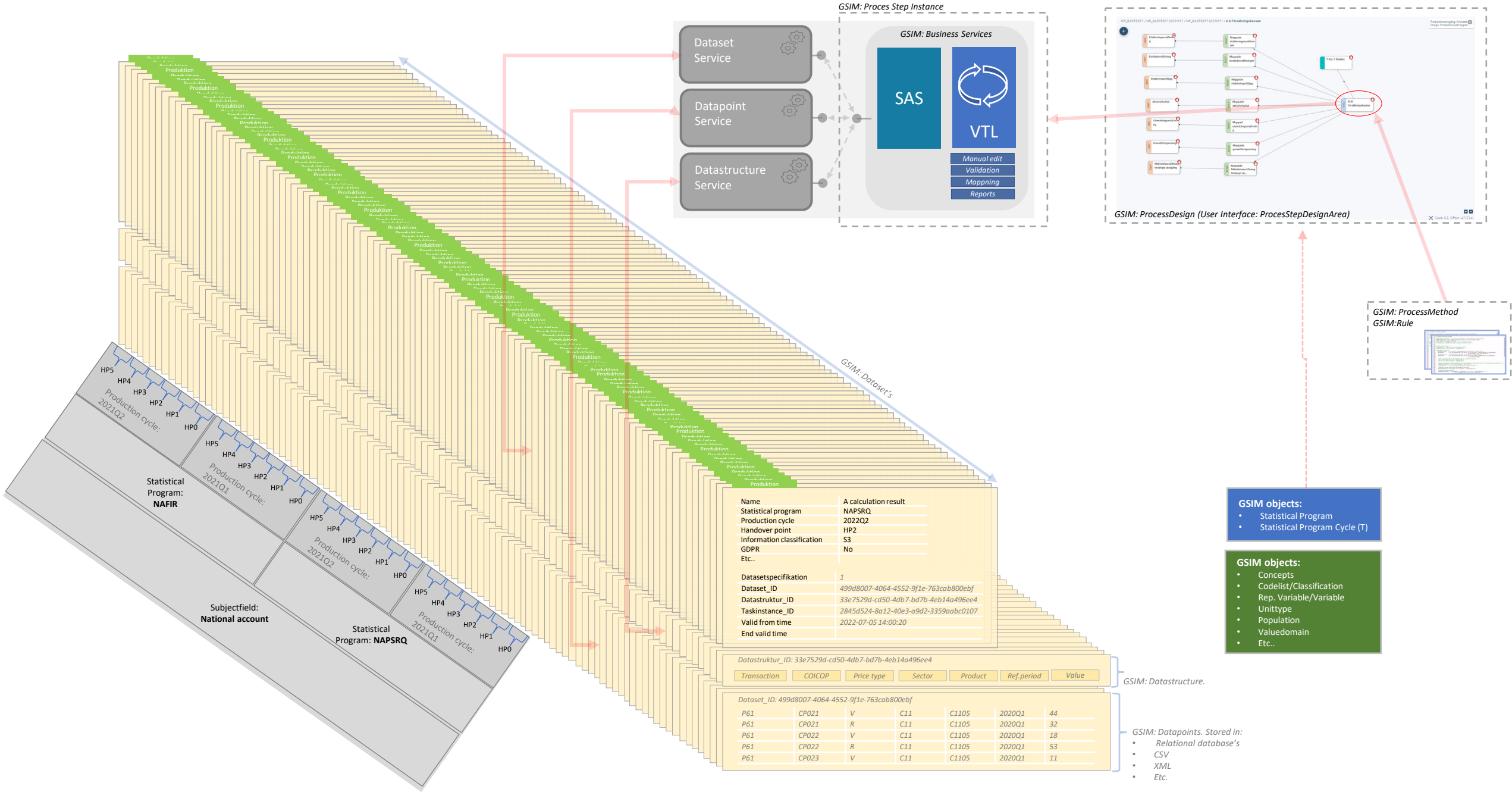
Services (Dataset, Datapoint and Datastructure)



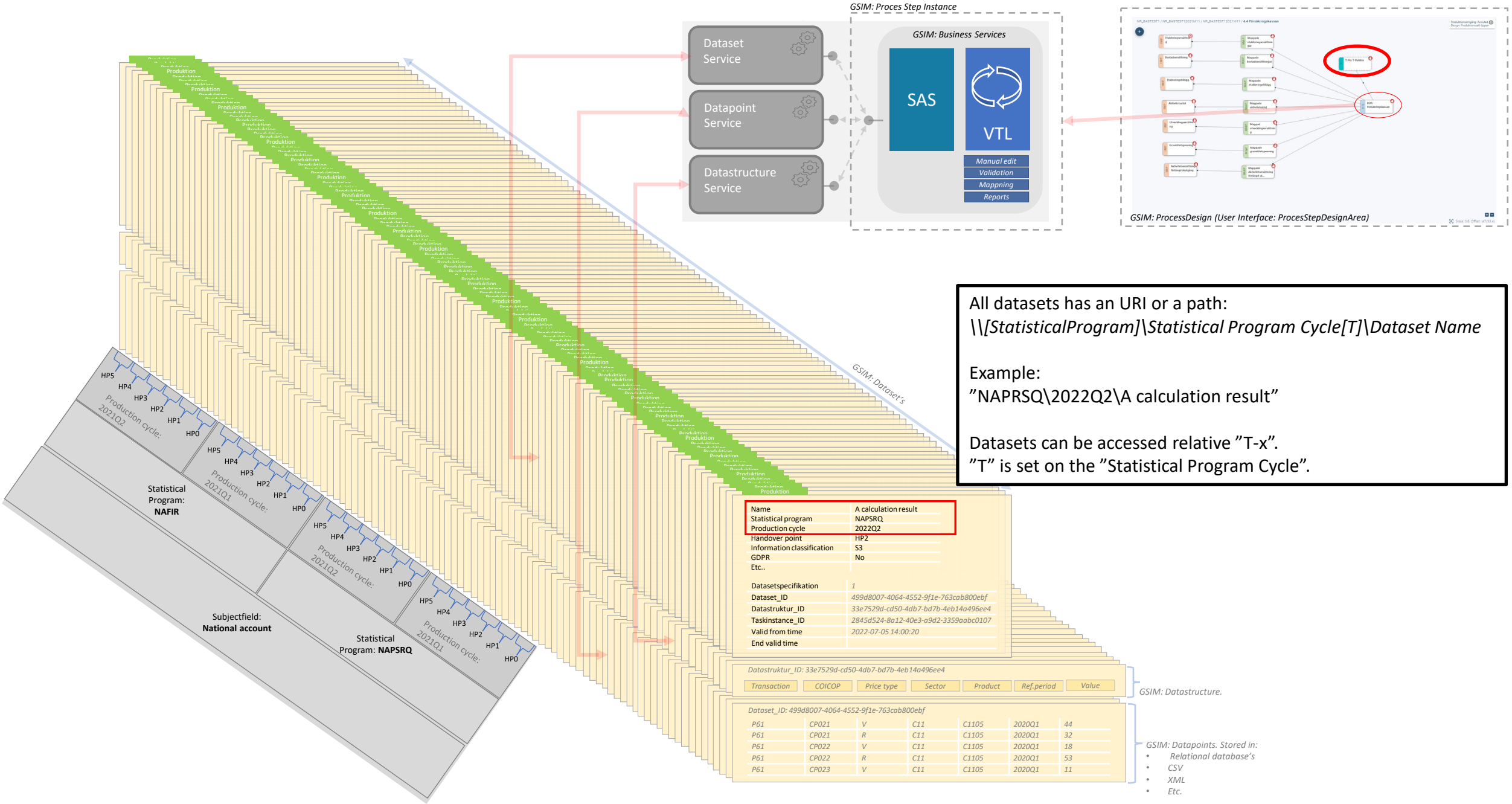
Dataset catalog



Handoverpoints (Checkpoints, Qualitygates or Steady states)



Relative references with ”T” for access to datasets

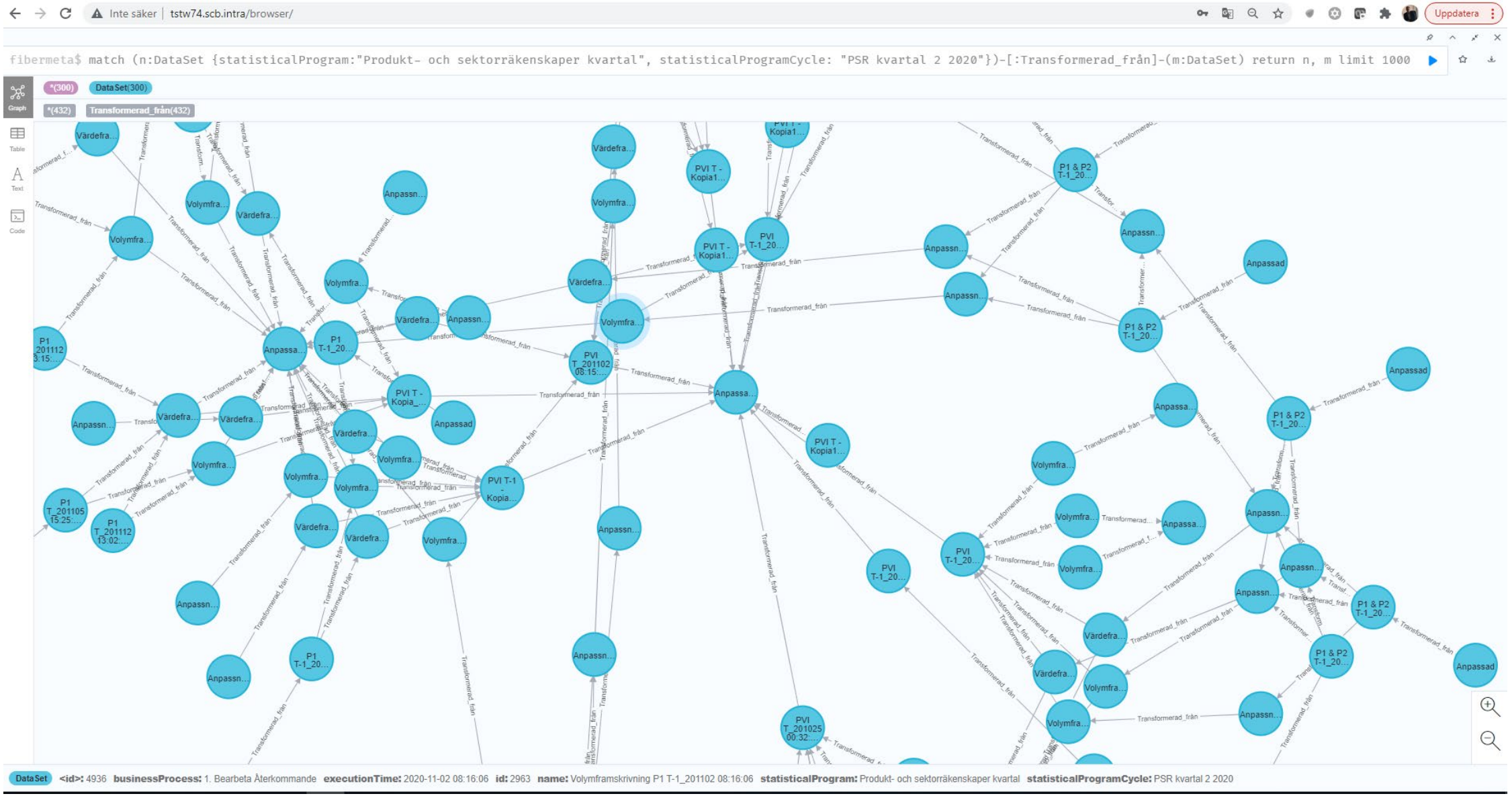


All datasets has an URI or a path:
\\[StatisticalProgram]\Statistical Program Cycle[T]\Dataset Name

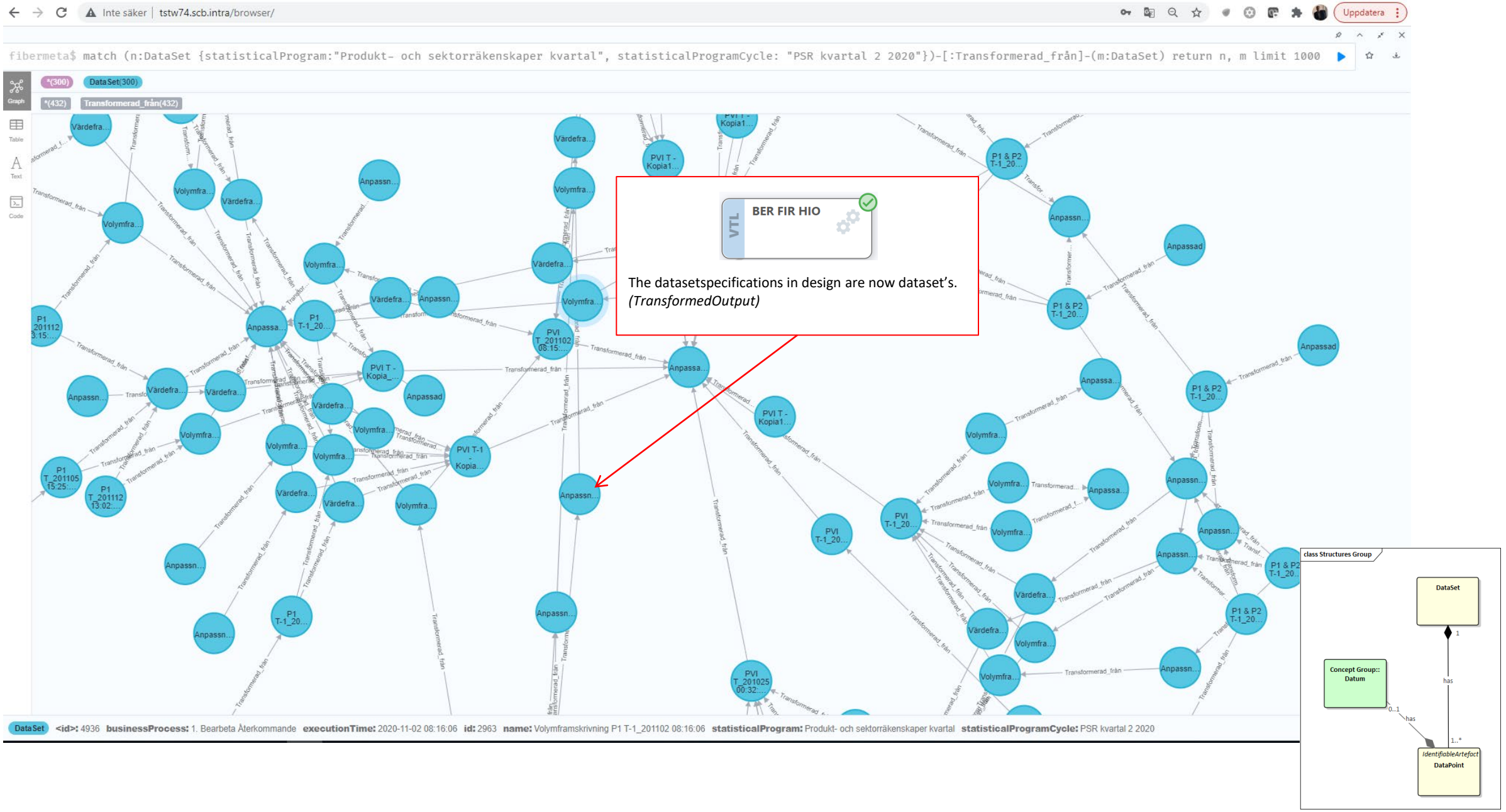
Example:
”NAPRSQ\2022Q2\A calculation result”

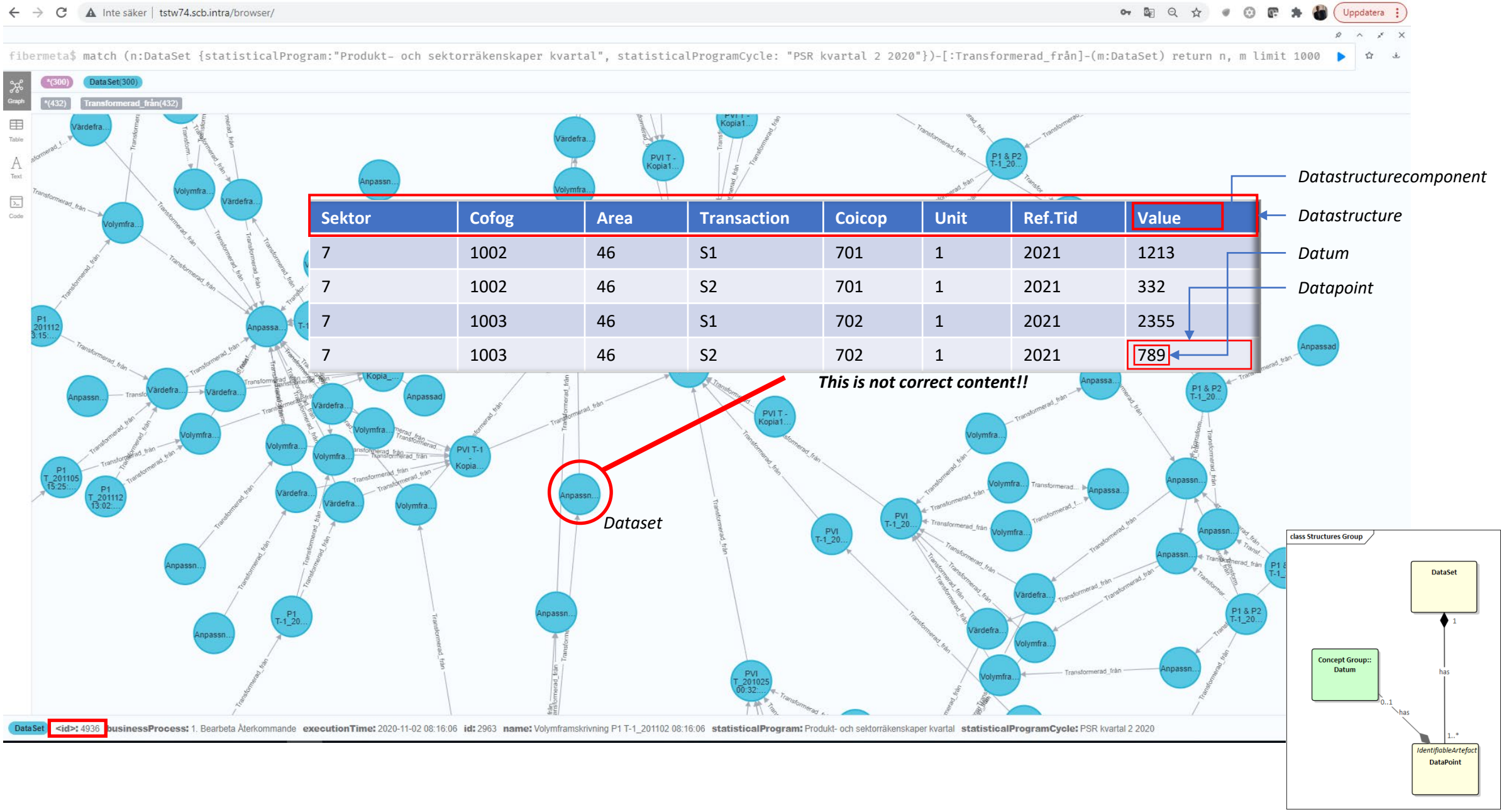
Datasets can be accessed relative ”T-x”.
”T” is set on the ”Statistical Program Cycle”.

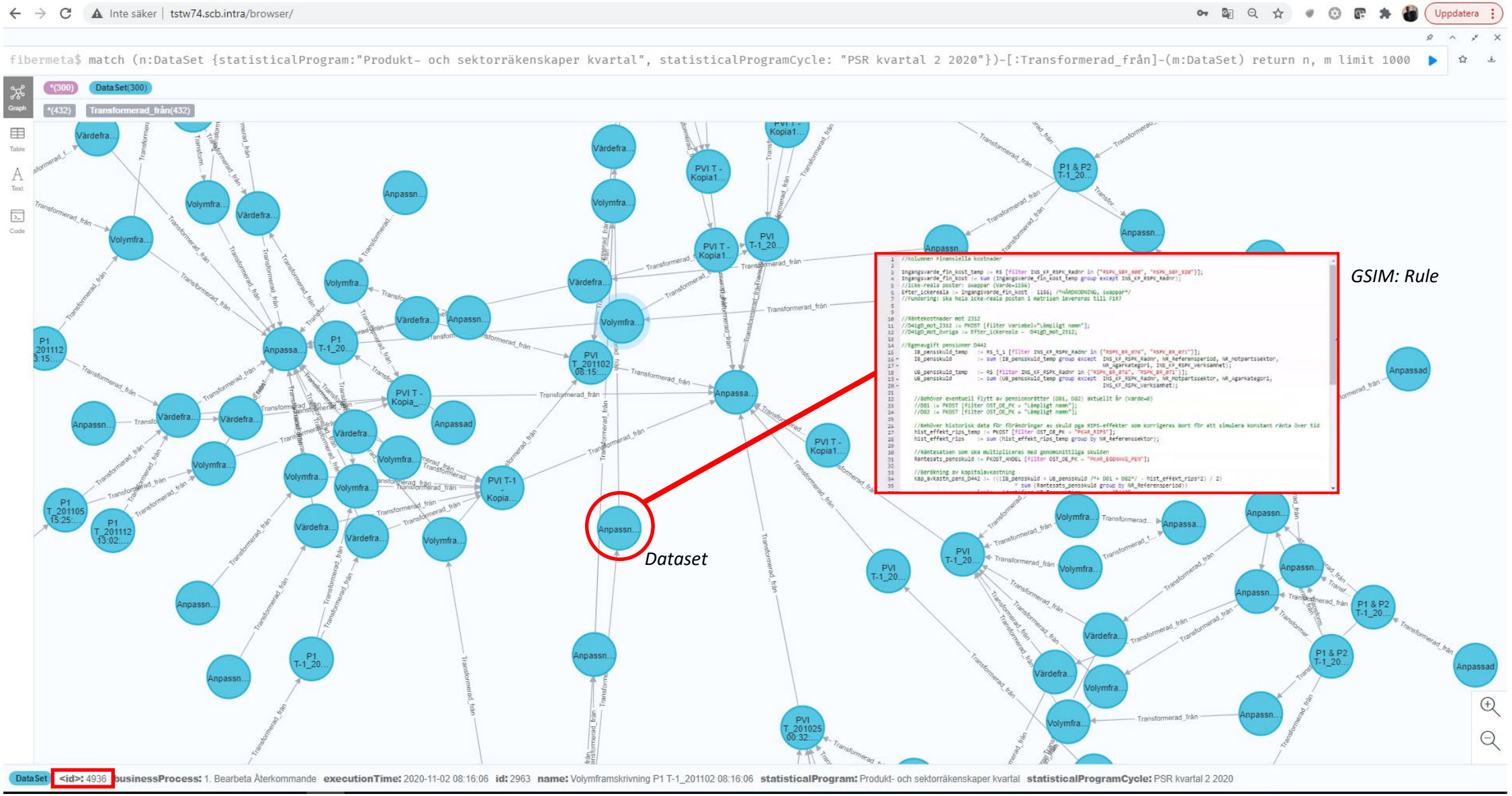
Visualization of the dataset catalog with a graph database



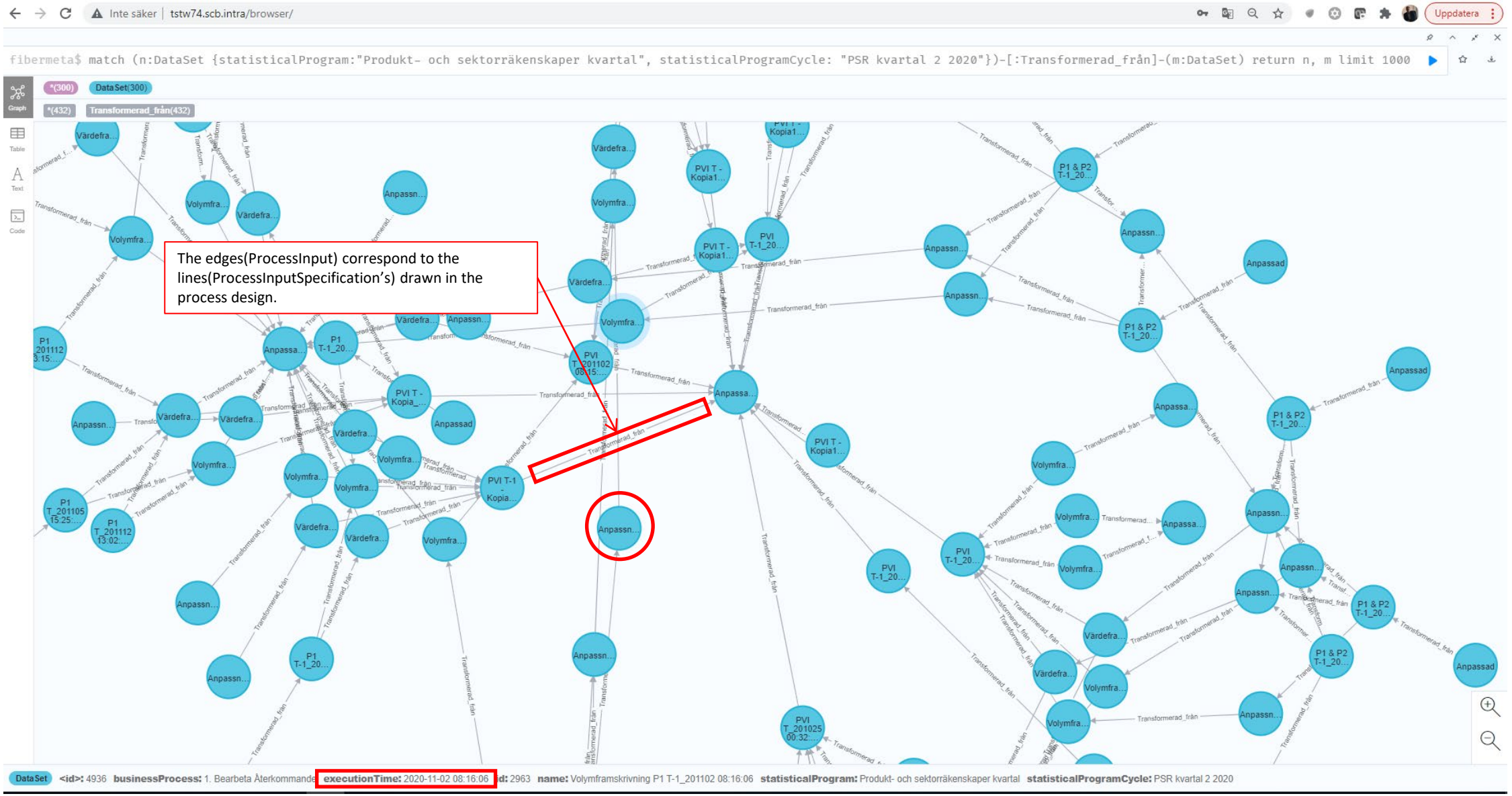
This is a test using a grapph database (NEO4J) to display all the dataset in the datasetcatalog
The picture shows a couple of the datasets created in production cycle for quarter 2, 2020 and for the statistical program called NRPSRQ


Inte säker | tstw74.scb.intra/browser/
Uppdatera
fibermeta$ match (n:DataSet {statisticalProgram:"Produkt- och sektorräkenskaper kvartal", statisticalProgramCycle: "PSR kvartal 2 2020"})-[:Transformerad_från]-(m:DataSet) return n, m limit 1000
*(300)
DataSet(300)
*(432)
Transformerad_från(432)
Graph
Table
Text
Code
VTL
BER FIR HIO
The datasetspecifications in design are now dataset's.
(TransformedOutput)
DataSet <id>: 4936 businessProcess: 1. Bearbeta Återkommande executionTime: 2020-11-02 08:16:06 id: 2963 name: Volymframskrivning P1 T-1_201102 08:16:06 statisticalProgram: Produkt- och sektorräkenskaper kvartal statisticalProgramCycle: PSR kvartal 2 2020
class Structures Group
DataSet
1
has
Concept Group::
Datum
0..1
has
1..*
IdentifiableArtefact
DataPoint


Inte säker | tstw74.scb.intra/browser/
Uppdatera
fibermeta$ match (n:DataSet {statisticalProgram:"Produkt- och sektorräkenskaper kvartal", statisticalProgramCycle: "PSR kvartal 2 2020"})-[:Transformerad_från]-(m:DataSet) return n, m limit 1000
*(300)
DataSet(300)
*(432)
Transformerad_från(432)
Graph
Table
Text
Code
Sektor | Cofog | Area | Transaction | Coicop | Unit | Ref.Tid | Value
7 | 1002 | 46 | S1 | 701 | 1 | 2021 | 1213
7 | 1002 | 46 | S2 | 701 | 1 | 2021 | 332
7 | 1003 | 46 | S1 | 702 | 1 | 2021 | 2355
7 | 1003 | 46 | S2 | 702 | 1 | 2021 | 789
Datastructurecomponent
Datastructure
Datum
Datapoint
This is not correct content!!
Dataset
DataSet <id>: 4936 businessProcess: 1. Bearbeta Återkommande executionTime: 2020-11-02 08:16:06 id: 2963 name: Volymframskrivning P1 T-1_201102 08:16:06 statisticalProgram: Produkt- och sektorräkenskaper kvartal statisticalProgramCycle: PSR kvartal 2 2020
class Structures Group
DataSet
1
has
Concept Group::
Datum
0..1
has
1..*
IdentifiableArtefact
DataPoint



You could see the rule that created this dataset.


Inte säker | tstw74.scb.intra/browser/
Uppdatera
fibermeta$ match (n:DataSet {statisticalProgram:"Produkt- och sektorräkenskaper kvartal", statisticalProgramCycle: "PSR kvartal 2 2020"})-[:Transformerad_från]-(m:DataSet) return n, m limit 1000
*(300)
DataSet(300)
*(432)
Transformerad_från(432)
Graph
Table
Text
Code
The edges(ProcessInput) correspond to the lines(ProcessInputSpecification's) drawn in the process design.
DataSet <id>: 4936 businessProcess: 1. Bearbeta Återkommande executionTime: 2020-11-02 08:16:06 id: 2963 name: Volymframskrivning P1 T-1_201102 08:16:06 statisticalProgram: Produkt- och sektorräkenskaper kvartal statisticalProgramCycle: PSR kvartal 2 2020

Patrik Wahlgren

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

Mail: patrik.wahlgren@scb.se