



# GSIM Implementation at Statistics Finland

**Session 1: ModernStats World - Where to begin with standards based  
modernisation?**

UNECE ModernStats World Workshop on  
sharing and supporting standards 2018

Essi Kaukonen and Mikko Saloila / Statistics Finland

# GSIM Implementation at Statistics Finland

## 4) GSIM implementation at Statistics Finland: challenges and solutions today



2020??

The GSIM (or the national version of it) is fully adopted to describe information objects in its business processes and development of methods and services

There is widespread awareness of the GSIM (or the national version of it) and a consistent approach to its use across the organisation

2017-18

Use of the GSIM (or the national version of it) is spreading, but limited to individual projects in an isolated manner

3) The role of GSIM in the emerging Information Architecture as well as in a couple of other internal projects (2017-2018)

2016

Use of the GSIM (or the national version of it) is basic and limited to a few individuals

2) GSIM as a common language with other organisations (2016→)

2015

A few individuals are becoming interested in the potential value of using the GSIM

1) First actions: GSIM exploration and renewing the Classifications System based on the GSIM Statistical Classification Model (2015)

→2014

The organisation is not using the GSIM (or the national version of it) and does not have intention to use it

## How it all began

- First attempts to utilize GSIM were made in Eurostat ESRBs -project
- GSIM was explored in-house for the first time around 2015 in a couple of workshops

## Implementing the GSIM Statistical Classification Model

- Around 2015 GSIM Statistical Classification Model was adopted in the new Classification System, launched in 2016
- Studying the model, benchmarking, comparing with other models, as well as testing the model with several classifications preceded the decision
- With few extensions to the model, implementation was a success leading to easier communication

# Cooperation with other organisations 2016 onwards and the role of GSIM

2016

Use of the GSIM (or the national version of it) is basic and limited to a few individuals

- **Nordic cooperation project** NordMan developing the service on microdata with other statistical Institutes 2016 to 2017
  - **GSIM enabled conversations on a lot more detailed level** as all participants already knew GSIM
  - This work also produced feedback for GSIM revision work
- **National cooperation project** for developing a national metadata description model with the National Institute for Health and Welfare (2016 to 2018)
  - Good experiences of the approach: **first GSBPM, then GSIM**
  - GSIM and GSPBM are applicable no matter what the domain (e.g. health data)

# Experience gained in the beginning of GSIM implementation:

It is important to make a difference between GSIM objects and Domain concepts

**GSIM objects (Concept group) give role to Domain concepts**

## GSIM: Unit Type

- Person
- Household
- Dwelling
- Building
- Location (Municipality)

## GSIM: Variable

- Size of household
- Floor area (m<sup>2</sup>)
- Person's municipality
- Person's gender

## GSIM Instance Variable

(e.g. 2016 published table)

- *Sp1* (technical var. name)
- *akoko* (technical var. name)

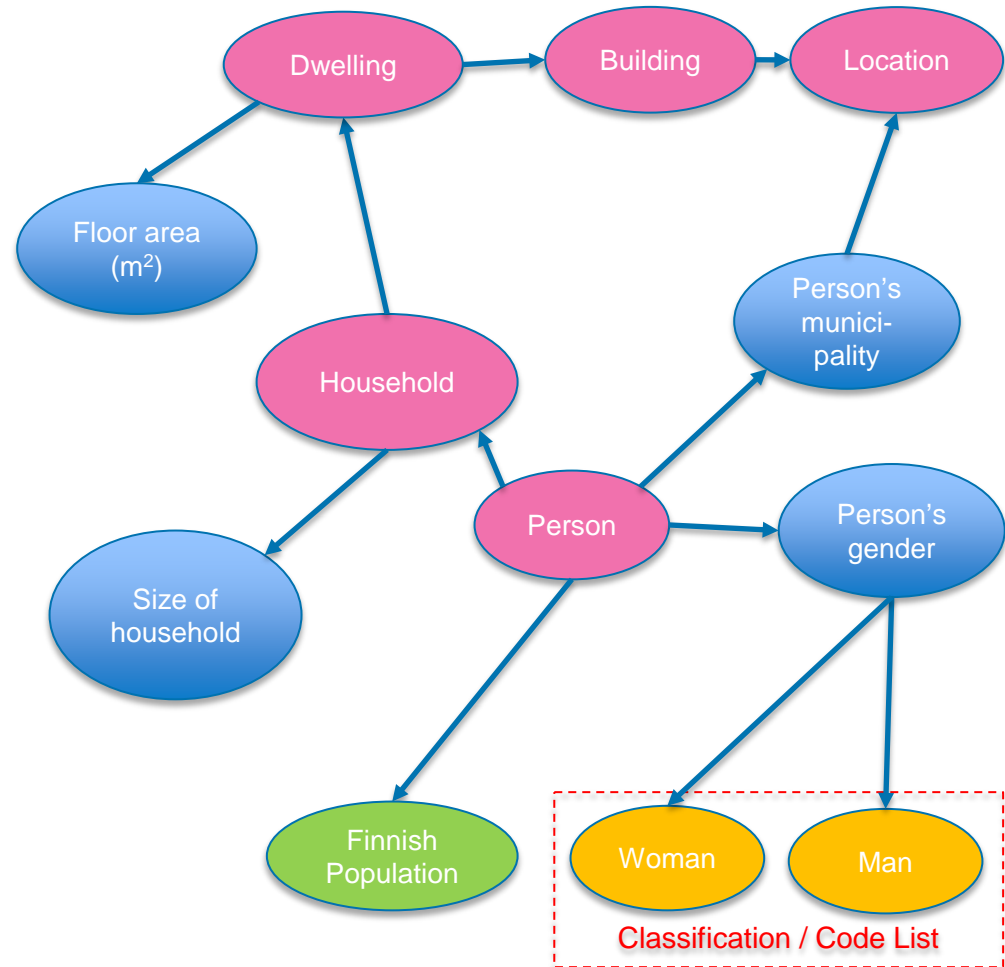
## GSIM Population

- *Finnish Population 2017*

## GSIM Classification / Code List

- *Gender classification*

Examples of **Domain concepts** defined in STIINA projects in our NSO



# Challenges and solutions in the pre-implementation phase

- ❖ It takes time to understand the model
  - *Studying the examples of other NSO's, cooperation, "step by step" approach starting with the Concept part*
- ❖ Modelling and information expertise are required in GSIM implementation
  - *Reserve time and resources, integrate with IA efforts (possibly nationally guided)*
- ❖ Language and communication: GSIM is not as easy to communicate as GSBPM
  - *Focus on communicating different issues to different stakeholders*

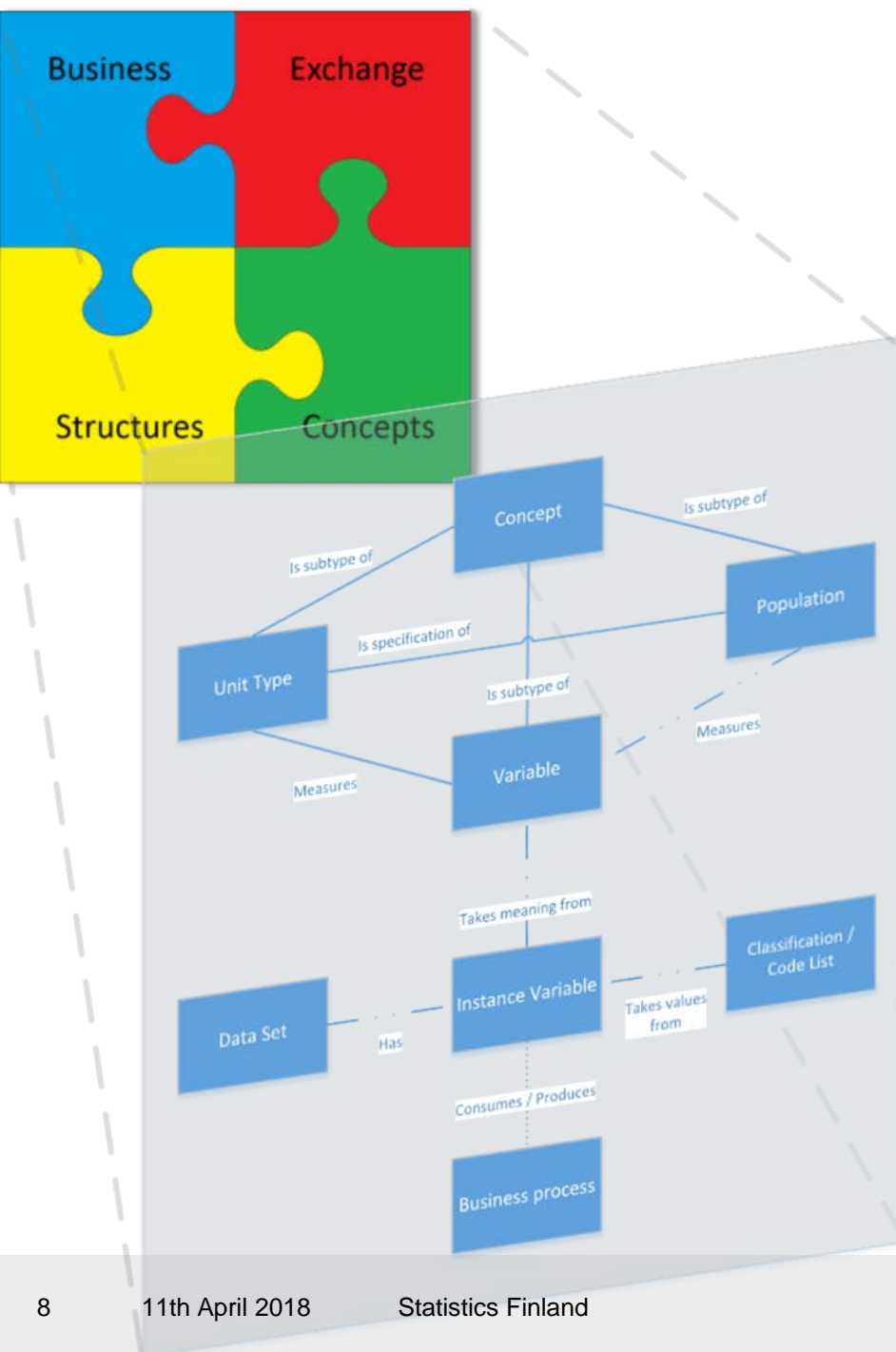


# The role of GSIM in the emerging Information Architecture (IA) and in a couple of other internal projects (2017 to 2018)

- **GSIM Core Elements** were approved in the first IA project in 2017
- At the same time, **several other development projects started utilising GSIM**
  - Social Statistics Integrated Information System: GSIM-based Metadata “sandbox”
  - Metadata related to producing statistical tables
  - Plans to develop new common Identifier Service
  - Interest is also rising i.e. in connection with developing the process governance system further and modernising the data collection, as well as structuring disseminated data and metadata

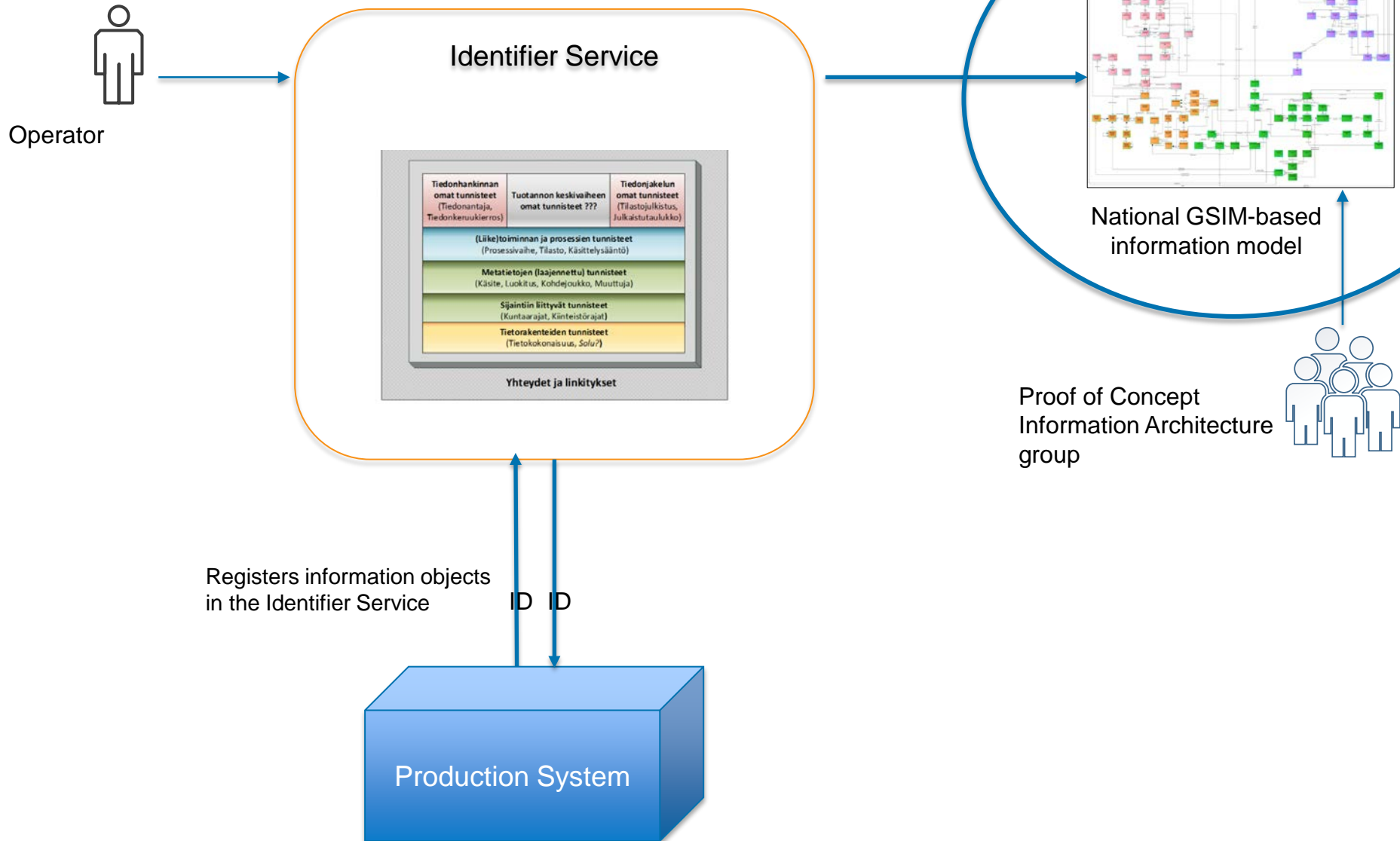
# GSIM Core Elements – approved in the first IA project in autumn 2017

- These Core Elements are easier to **communicate to a general audience** than the whole GSIM
- **The Core Model will be extended and further specified** in future according to the needs of selected development projects





# Future plans: Identifier Service and its governance model



# Challenges and solutions today

target

2020??

There is widespread awareness of the GSIM (or the national version of it) and a consistent approach to its use across the organisation

- ❖ Ongoing development projects might go in slightly different directions in terms of Information Architecture. This might lead to partially inconsistent and multiple data and metadata pools and is also resource consuming.
  - **Well governed Information Model based on GSIM**
- ❖ Governance of a common information model is difficult in a large, heterogeneous organisation
  - **Responsible party, building IA capability, benchmarking, something else?**
- ❖ It is not easy to show where resources are saved as the savings are realised in hundreds of other projects and activities in the longer run.
  - **Communication with the most important stakeholders, something else?**

# Conclusions:

## What can we achieve by using GSIM?

- GSIM is a common data-centric language in the complex world of statistics and their production → We can talk about the same information objects with the same names → this harmonises data and metadata contents, enhances data flows and simplifies data governance
  - Master data thinking: reuse information by saving the same information objects only once → this supports data exchange and programming of standardised interfaces leading to more efficient solutions
- *For our customer, this means better quality, as well as better findability and usability of the data and metadata we provide*

