Fulfilling user-needs, improving quality and efficiency using GSIM and other standards

UNECE Workshop on Implementing Standards for Statistical Modernisation,
21 – 23 September 2016

Mogens Grosen Nielsen
Statistics Denmark
(mgn@dst.dk)

STATISTICS DENMARK

The vision

Statistical information must help users in the "turbulent information-sea"

Metadata about content and quality must
a) help users in their knowledge processes
b) give users precise information about our products

International standards and standard software must enable
a) Cost efficient solution with few ressources
b) Sustainable long term solutions
c) Common terminology

Vision: Integrated and reusable metadata

STATISTICS DENMARK
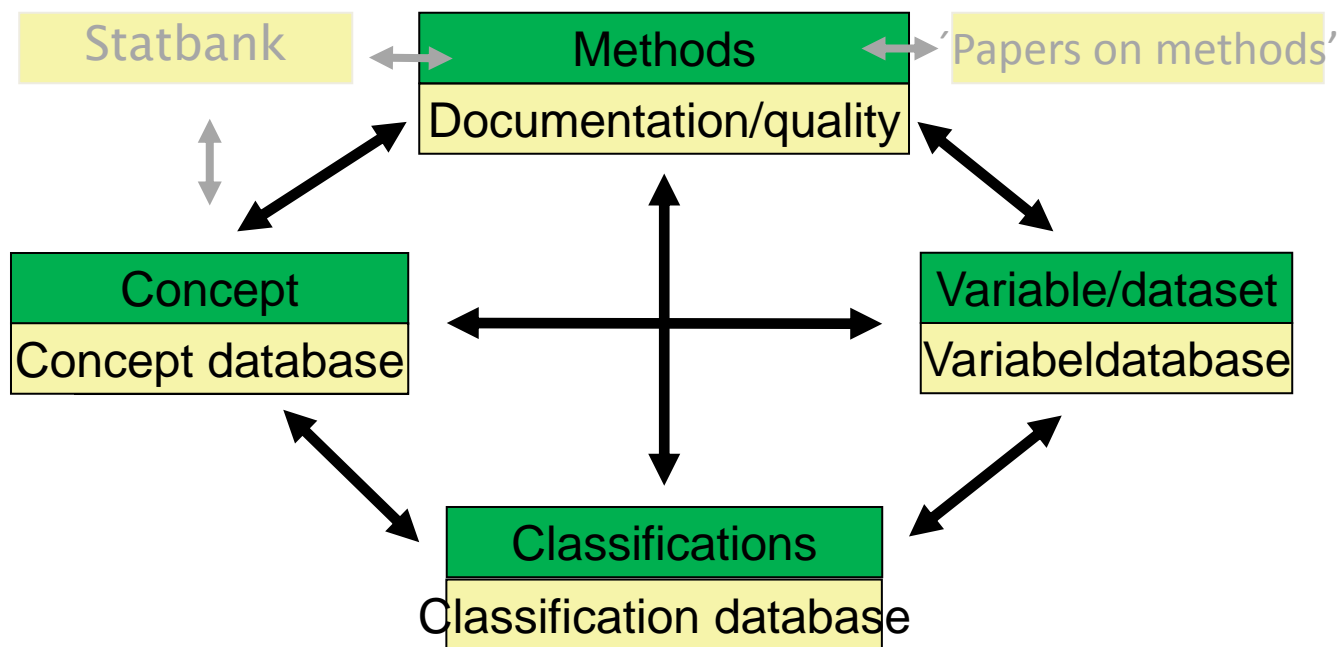
Claims in the paper

Common reusable metadata require improved understanding of

1. metadata-terminology
2. the use of in production processes in GSBPM
3. the role of metadata in relation to users
4. systematic use of Business Process Management and Enterprise Architecture

STATISTICS
DENMARK

Metadata-terminology and frames of reference

	Frames of reference of producers	Frames of reference of users
General terminology for statistical metadata	Complex metadata terminology. E.g. instance variable vs represented variable, logical record	Simplified metadata terminology. E.g. classification, variable, concept, population
Domain specific metadata	Domain specific metadata E.g. detailed definition of income of person	Domain specific metadata tailored for various users. E.g. short and long definition of income for a person

STATISTICS
DENMARK

Metadata portal

Unit Types

Arbejdssted

BBR enhed

Bolig

Bygning

Ejendom

Firma

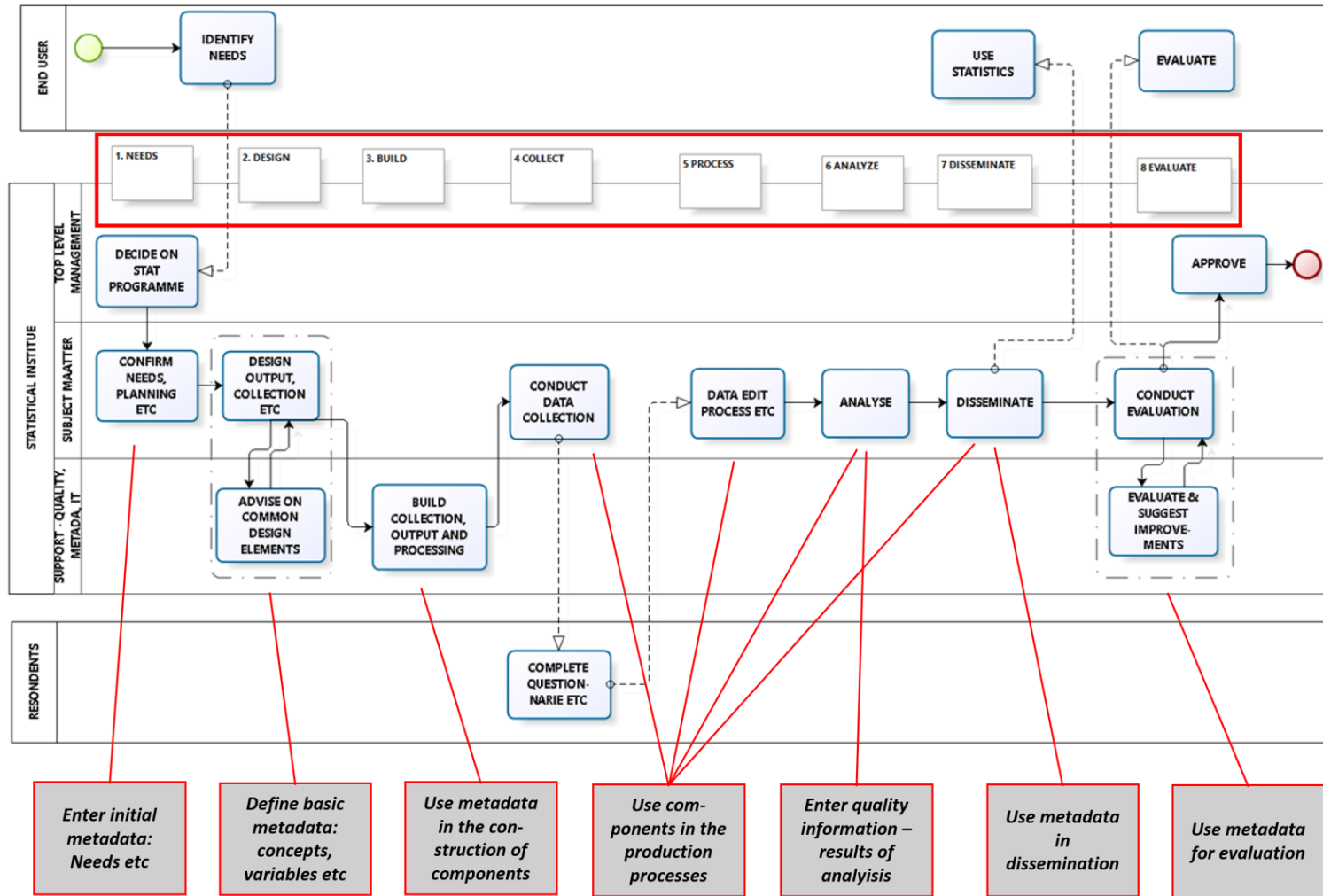Company - no of employees

Husstand

Job

Person

Person gender conceptual variable

Person's age at onset of disability conceptual variabnle

Person's address

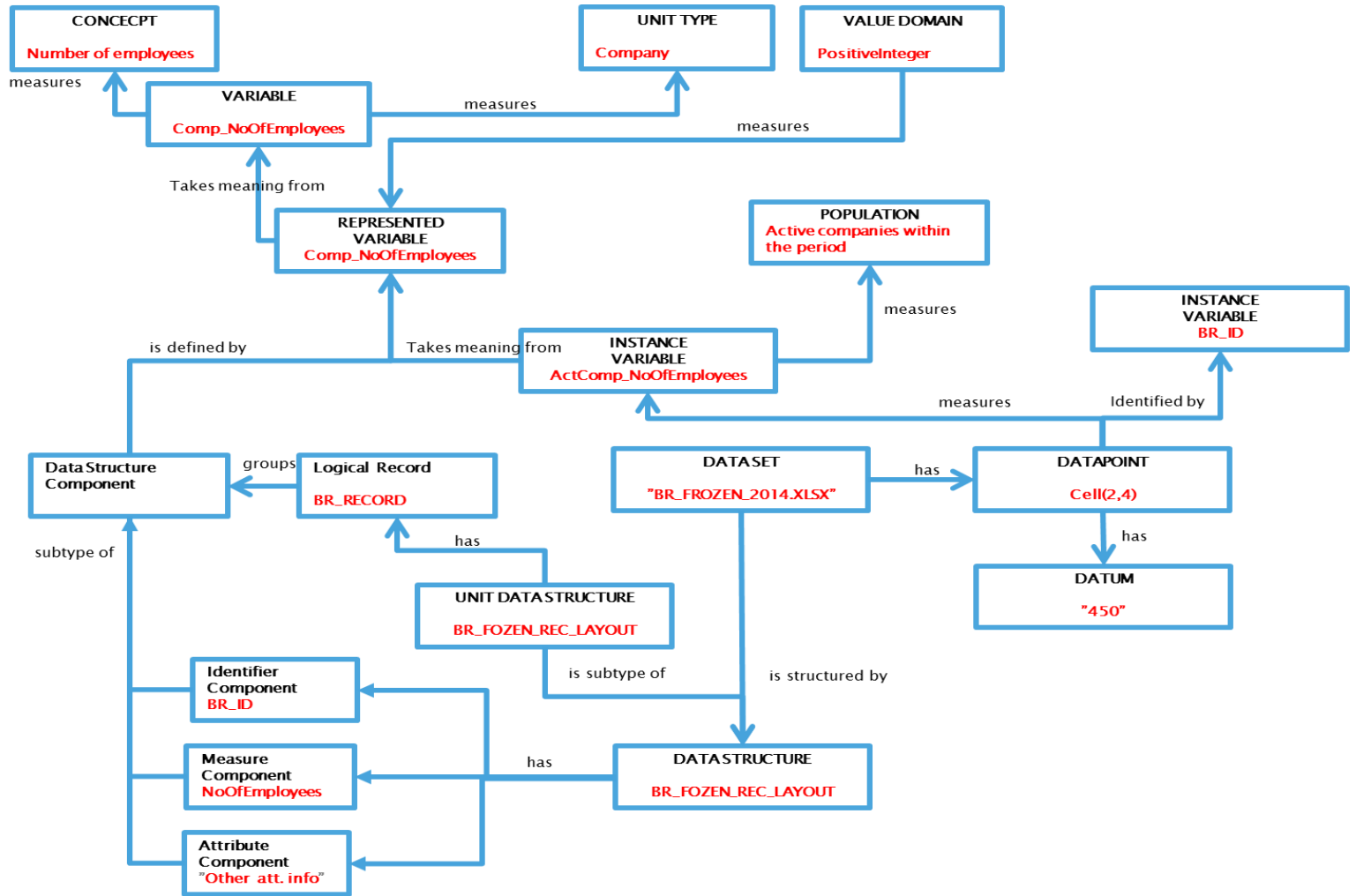Civil status conceptual variable

Meta-data and business proces-ses

From GSIM to DDI and implementation in Colectica

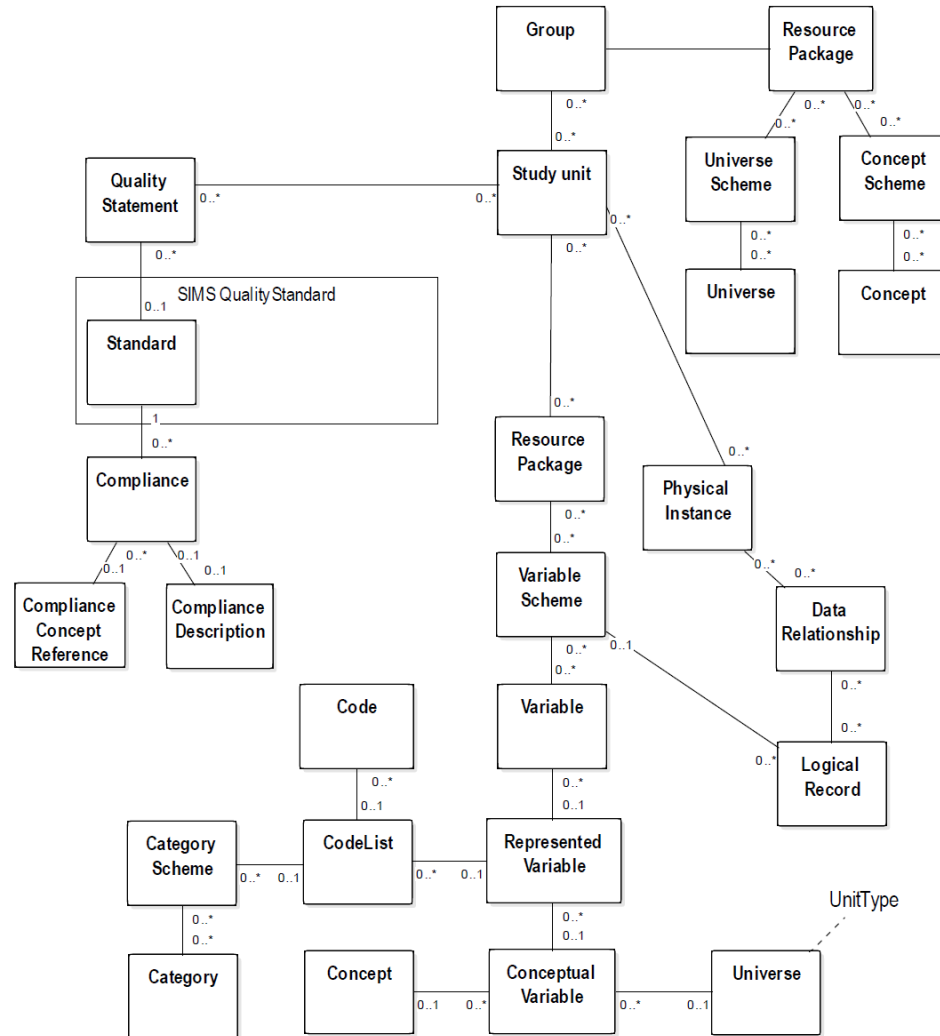| Level	Scope of model and standards used
Conceptual 1	Selected terms from GSIM: concept, variable, dataset etc.
Conceptual 2	Selected terms from DDI 3.2 complying with GSIM terms
Logical	Selected elements from DDI 3.2 used for implementation
Physical	Logical model extended with Colectica implementation details

Conceptual model Selected elements from GSIM

CONCECPT
Number of employees

UNIT TYPE
Company

VALUE DOMAIN
PositiveInteger

VARIABLE
Comp_NoOfEmployees

measures

measures

measures

Takes meaning from

REPRESENTED VARIABLE
Comp_NoOfEmployees

POPULATION
Active companies within the period

INSTANCE VARIABLE
BR_ID

is defined by

Takes meaning from

INSTANCE VARIABLE
ActComp_NoOfEmployees

measures

measures

Identified by

Data Structure Component

groups

Logical Record
BR_RECORD

DATASET
"BR_FROZEN_2014.XLSX"

has

DATAPOINT
Cell(2,4)

subtype of

has

has

DATUM
"450"

UNIT DATA STRUCTURE
BR_FOZEN_REC_LAYOUT

is subtype of

is structured by

Identifier Component
BR_ID

Measure Component
NoOfEmployees

has

DATA STRUCTURE
BR_FOZEN_REC_LAYOUT

Attribute Component
"Other att. info"

Logical model using DDI 3.2 implemented in Colectica.
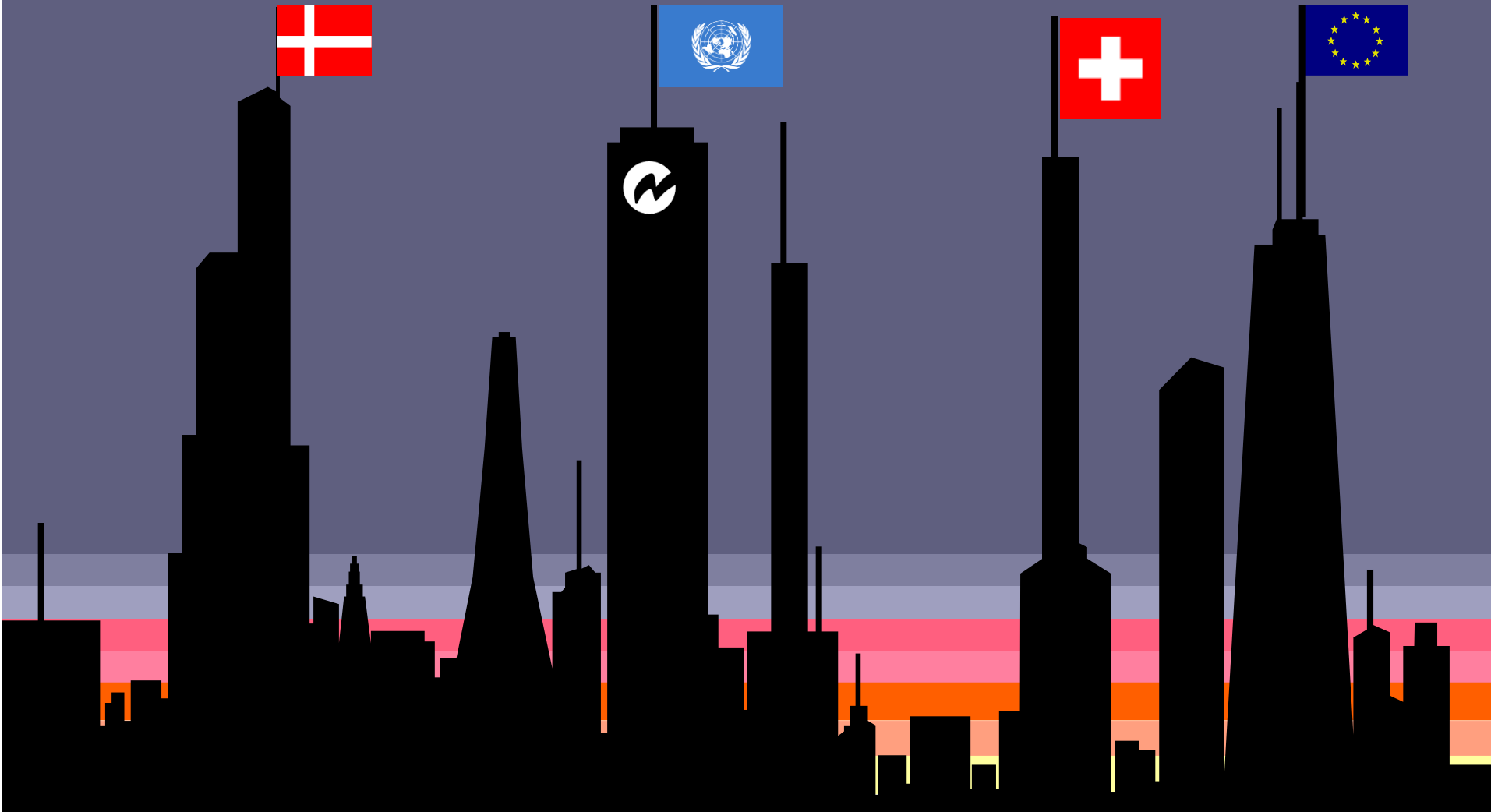
Basic parts comply with GSIM

Conclusions

- No simple road towards the vision for metadata

- Important to have a precise understanding of metadata terminology, the use of metadata in processes and metadata in relation to users.

- Systematic use of Business Process Management and Enterprise Architecture is important

- Need for international standards and cooperation to succeed with the implementation (e.g. this workshop, cooperation between Nordic countries)

Thanks for your attention

Remember:

- DDI conference in Cologne 6-7 December 2016

The End!