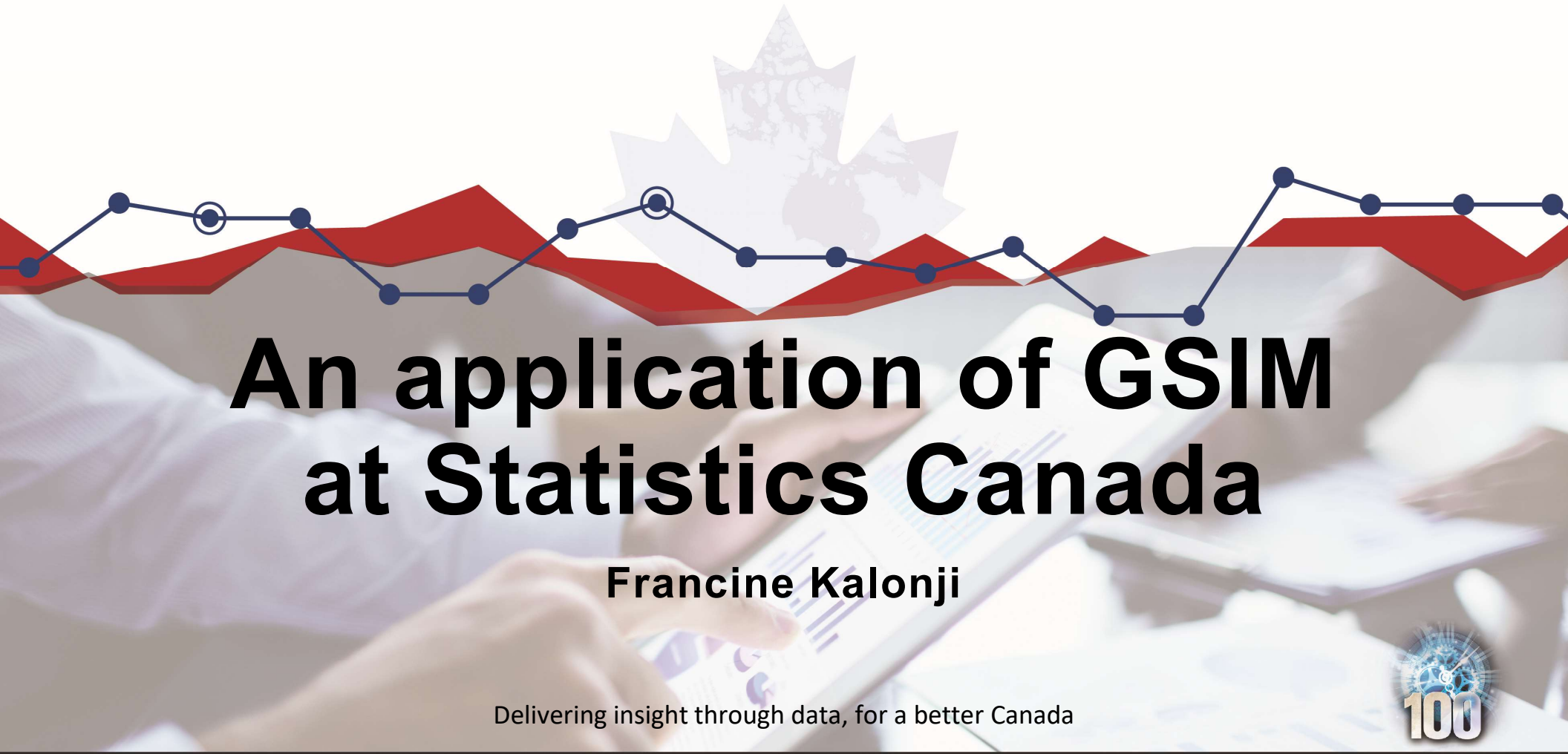


An application of GSIM at Statistics Canada

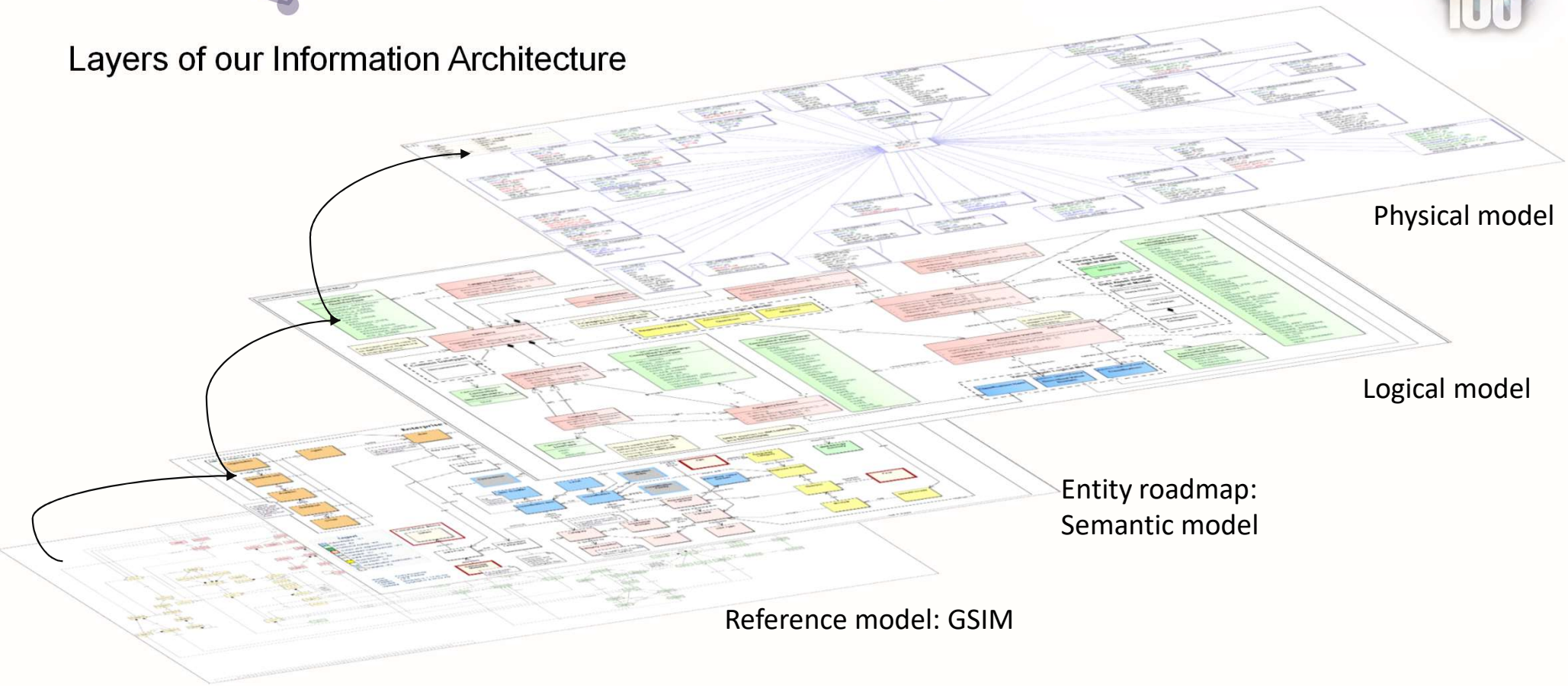
Francine Kalonji

Delivering insight through data, for a better Canada

Outline

- GSIM as a framework for our reference models
- An synopsis of our implementation process
- Snapshots of our Picasso application
- Benefits of using the GSIM model

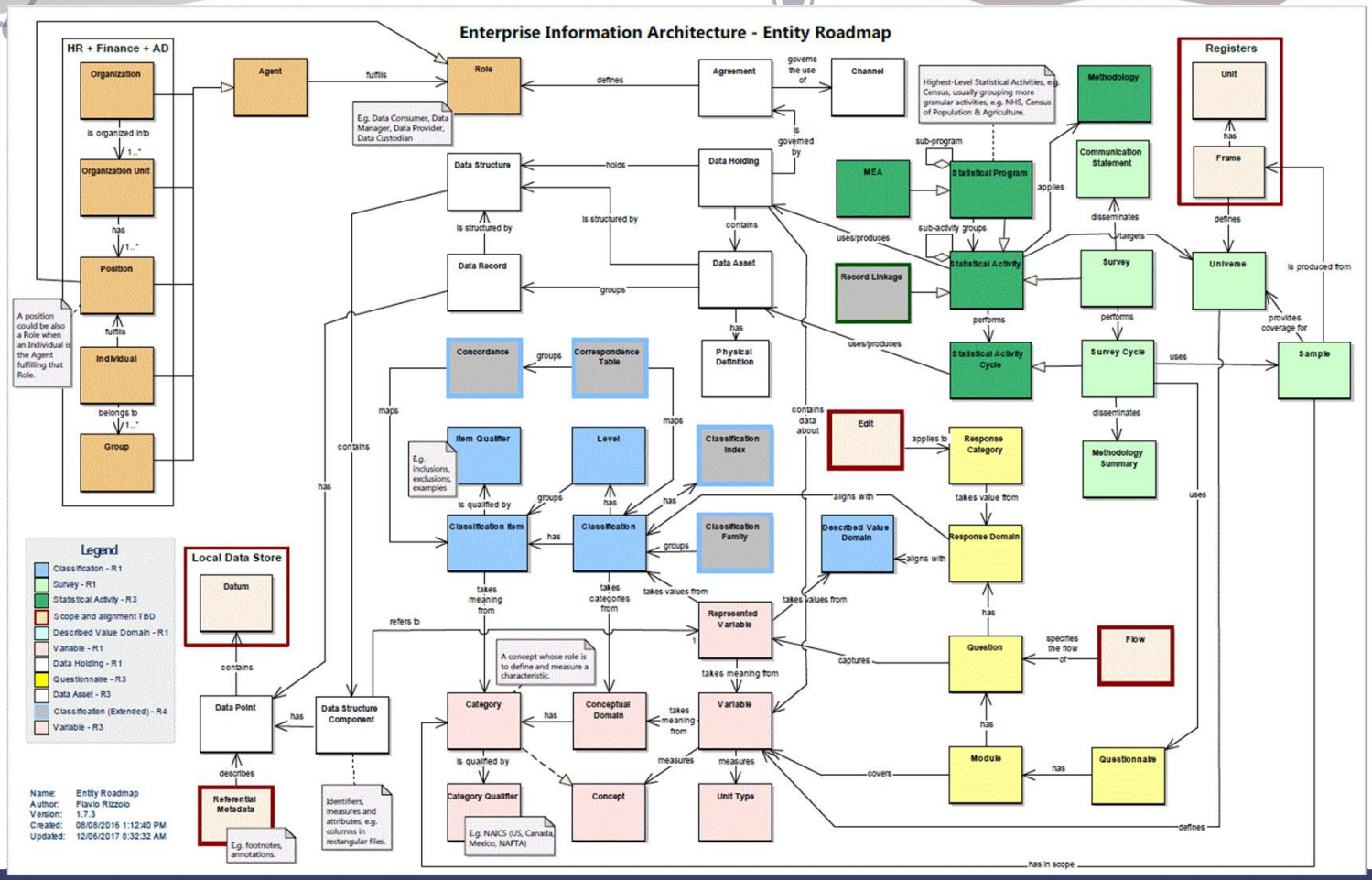
Layers of our Information Architecture



Entity roadmap

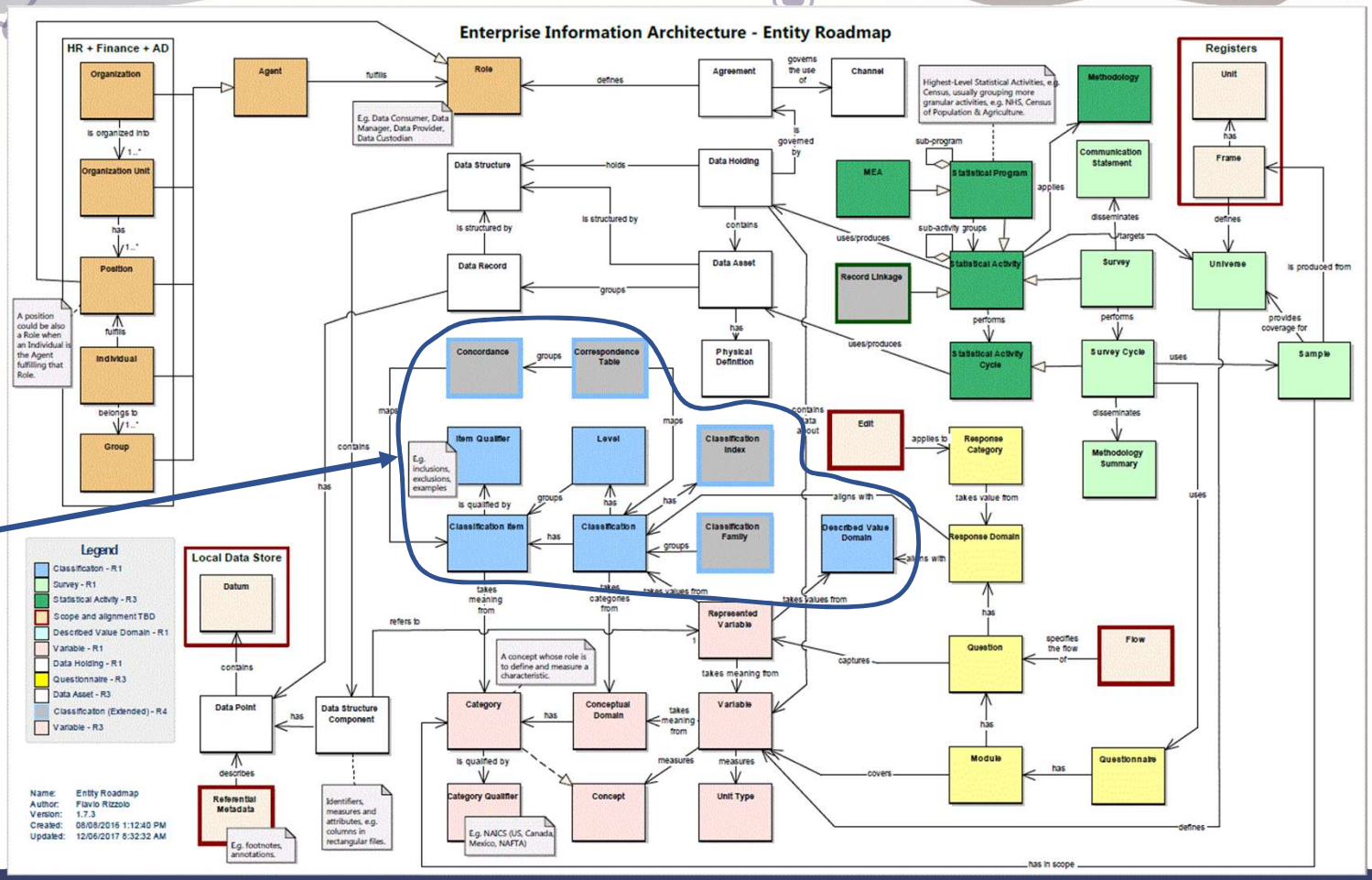
- Models Statistics Canada's core statistical entities
- Maps main concepts needed to conduct statistical activities
- Is designed to align with GSIM objects, sometimes with our own internal names
  - Data Asset for GSIM Dataset, Survey for GSIM Statistical Program
- Broken down into domains at the logical level
  - Variable domain, Data Asset domain, Survey/Statistical Activity domain, etc.

Entity Roadmap



Entity Roadmap

Elements that are part of GSIM Concept group
↓↓
Value domain logical model

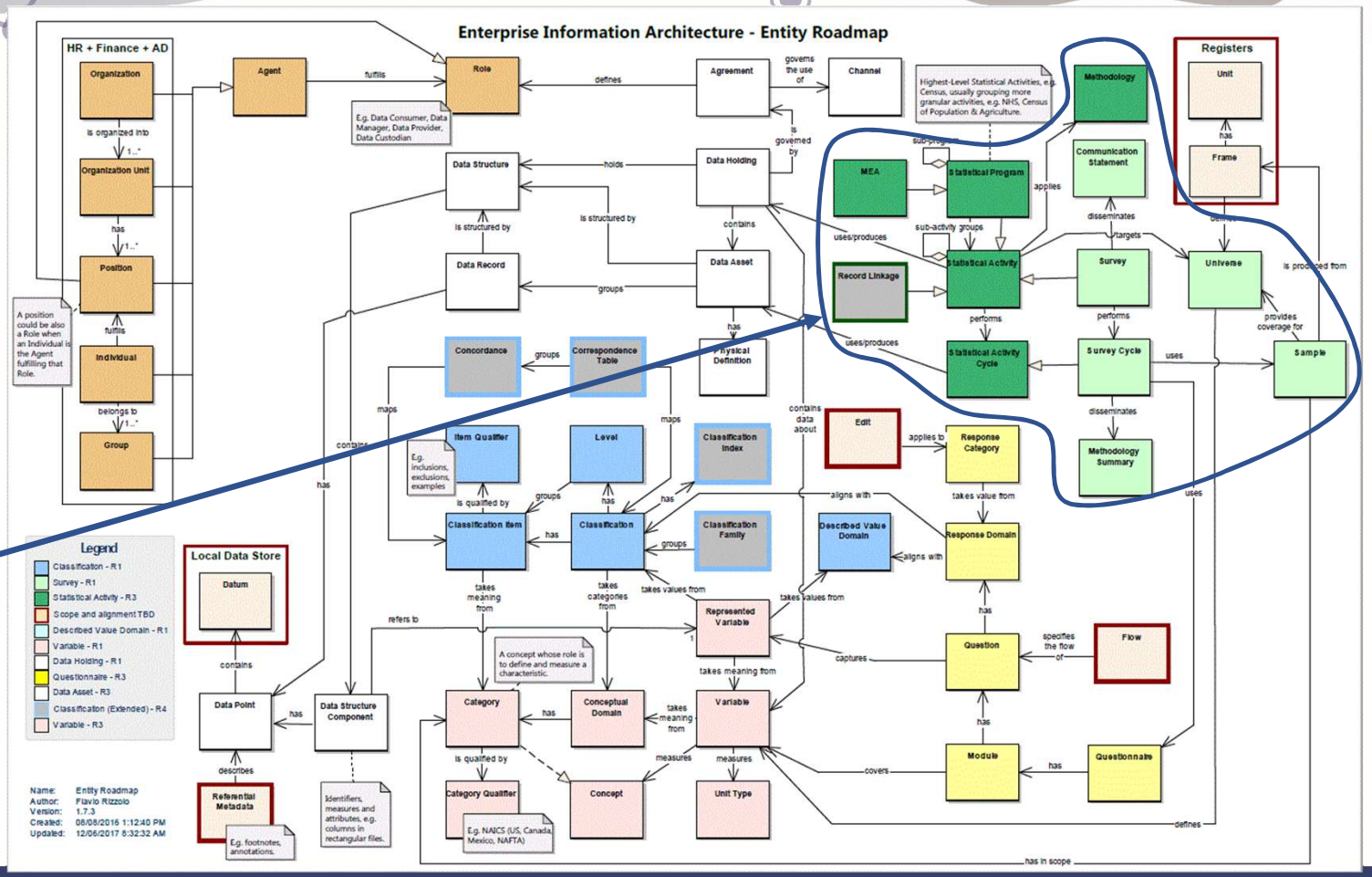


Entity Roadmap

Elements that are part of GSIM Business group

↓↓

Statistical activity domain logical model

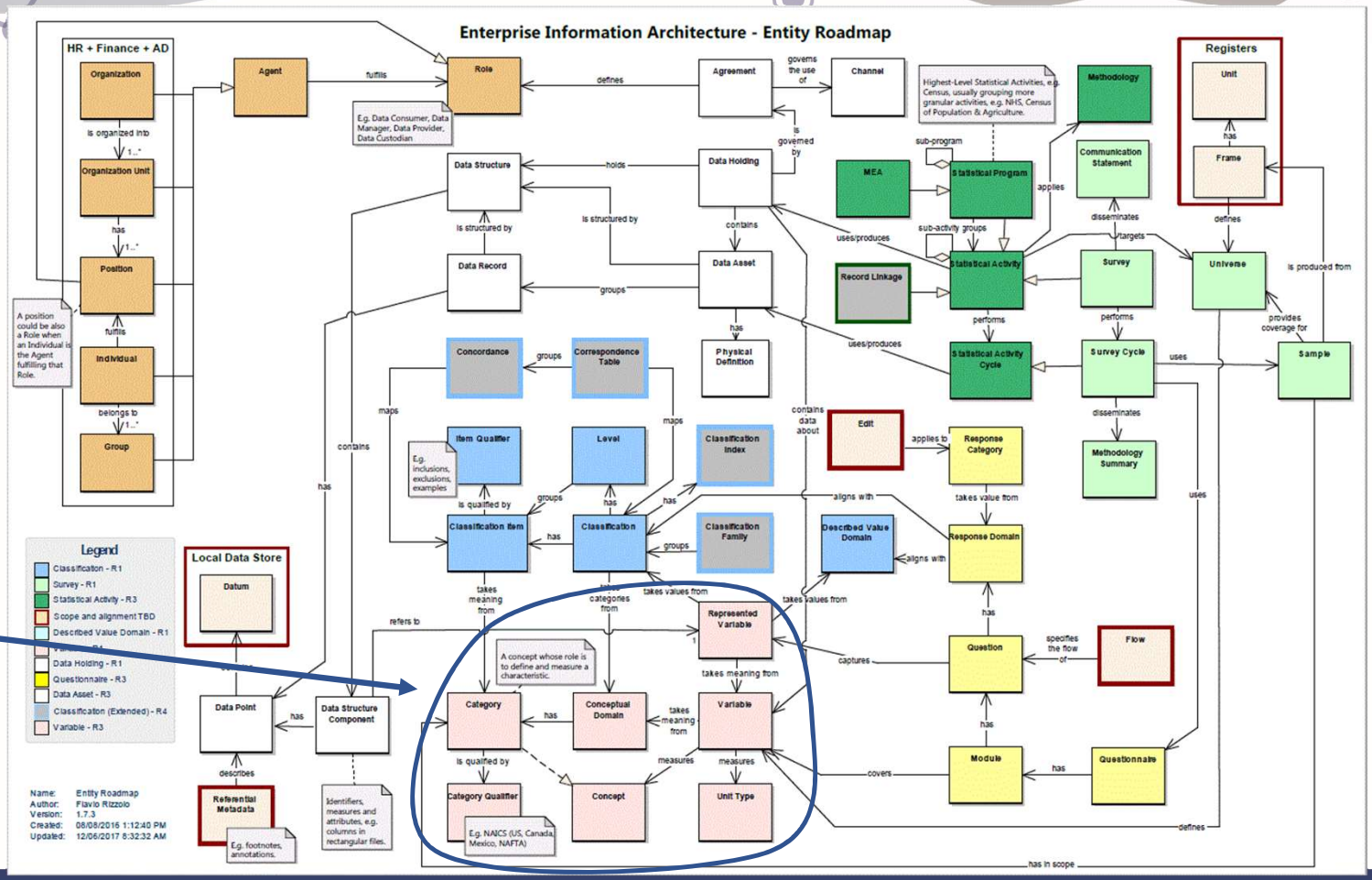


Entity Roadmap

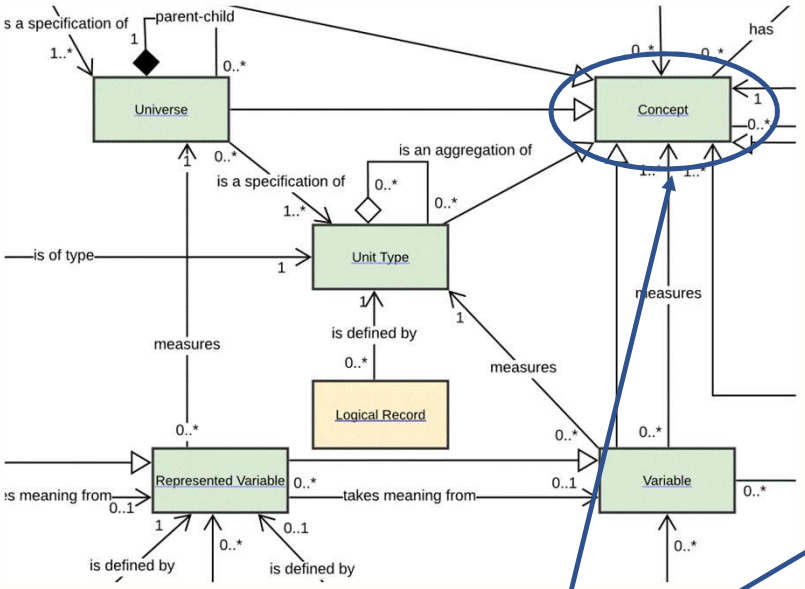
Elements that are part of GSIM Concept group

↓↓

Variable domain logical model

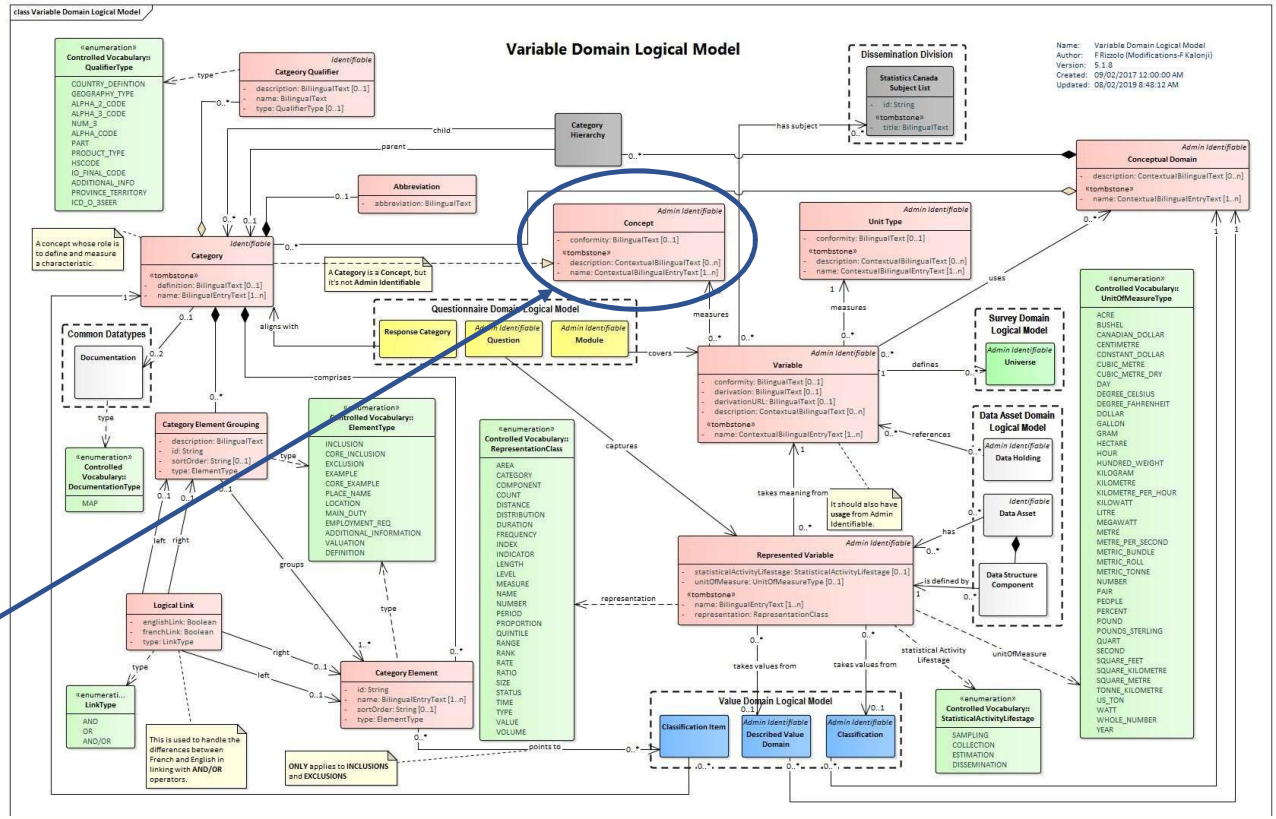


GSIM Concept Group

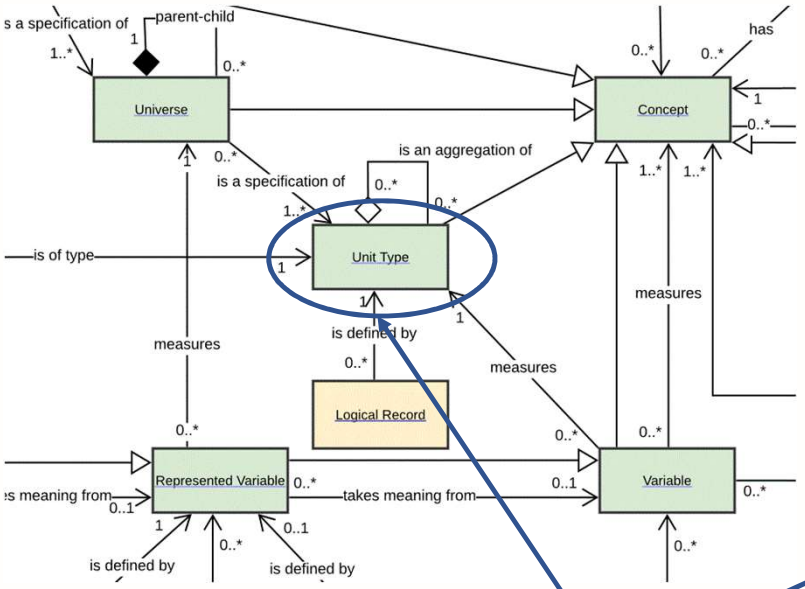


Concept

Variable Domain Logical Model

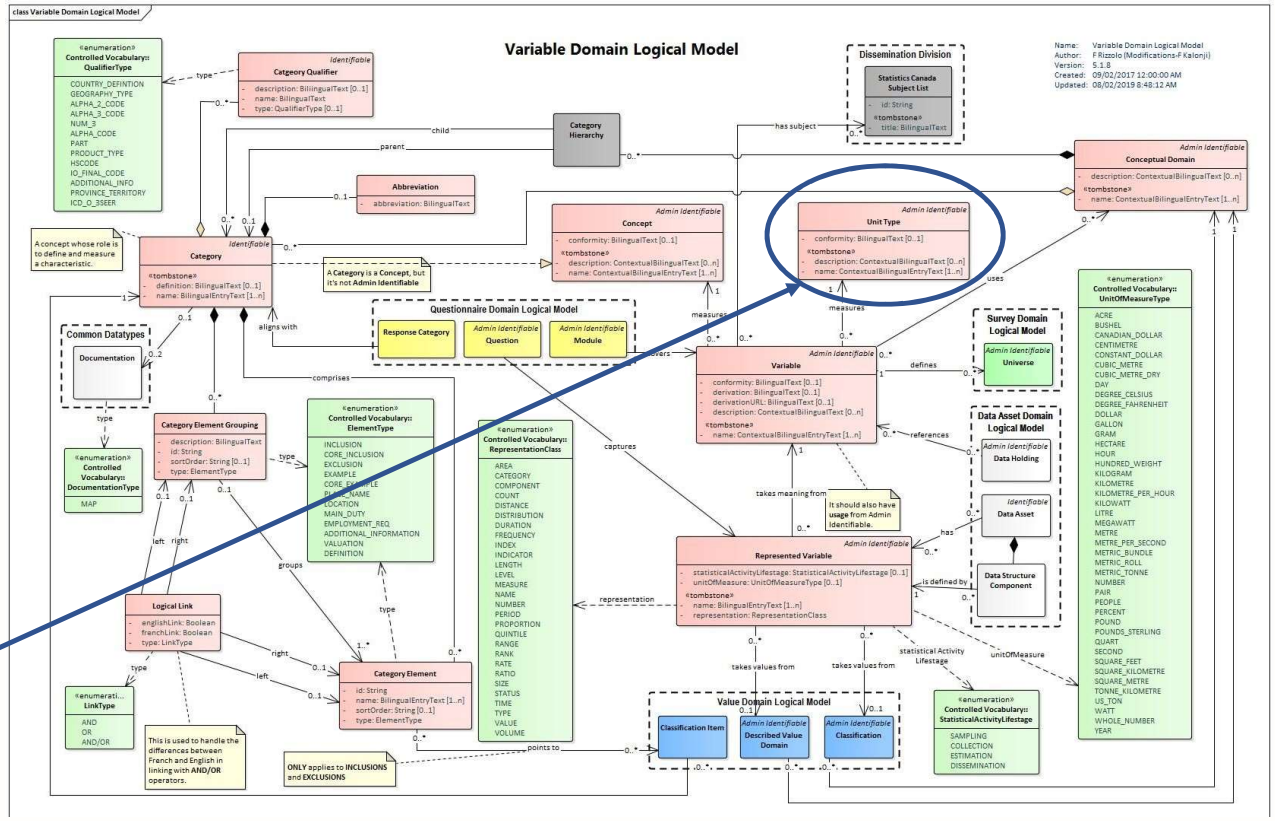


GSIM Concept Group

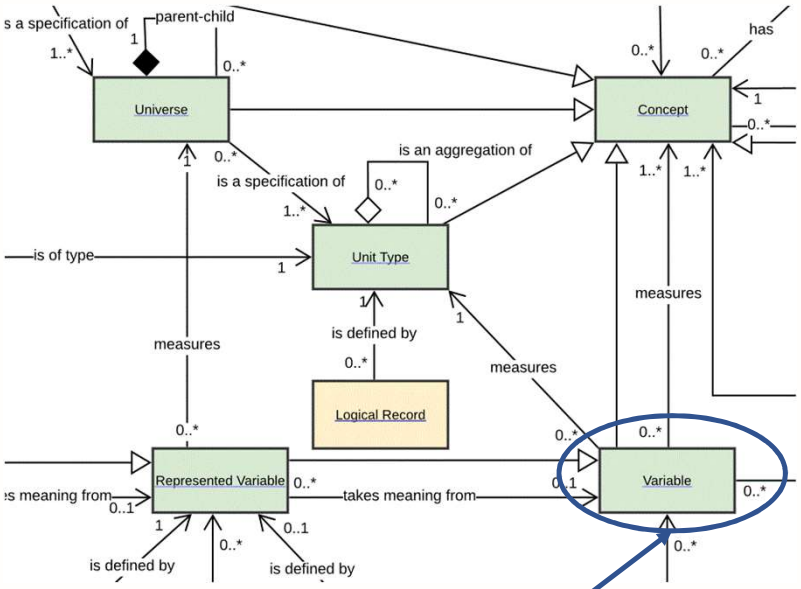


Unit Type

Variable Domain Logical Model

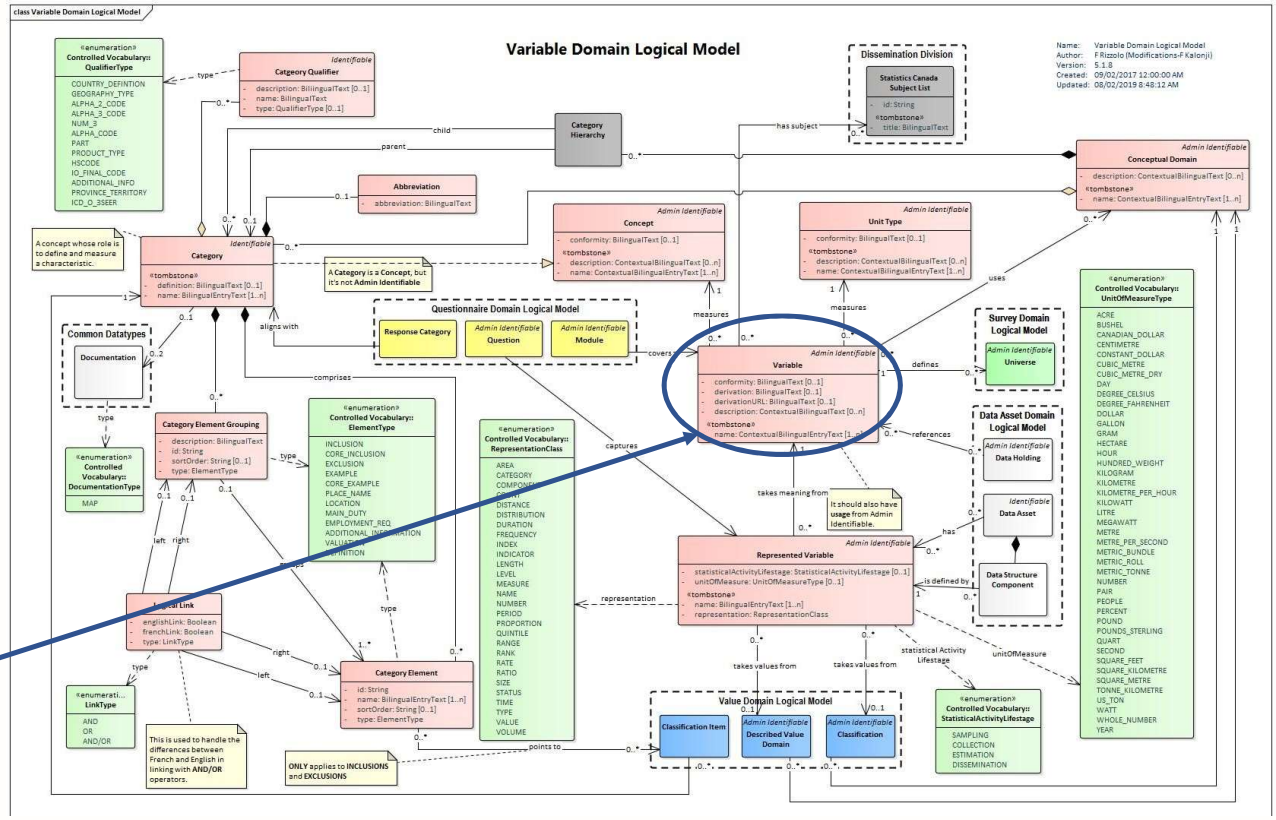


GSIM Concept Group

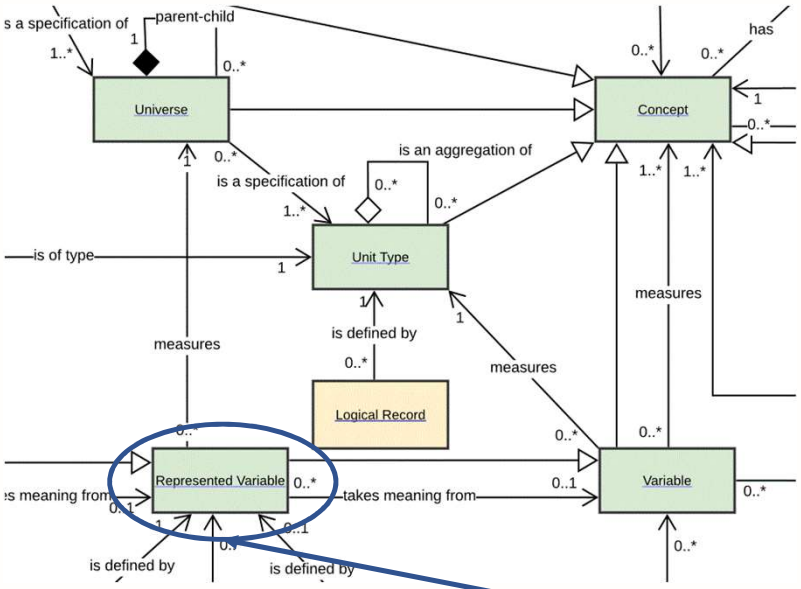


Variable

Variable Domain Logical Model

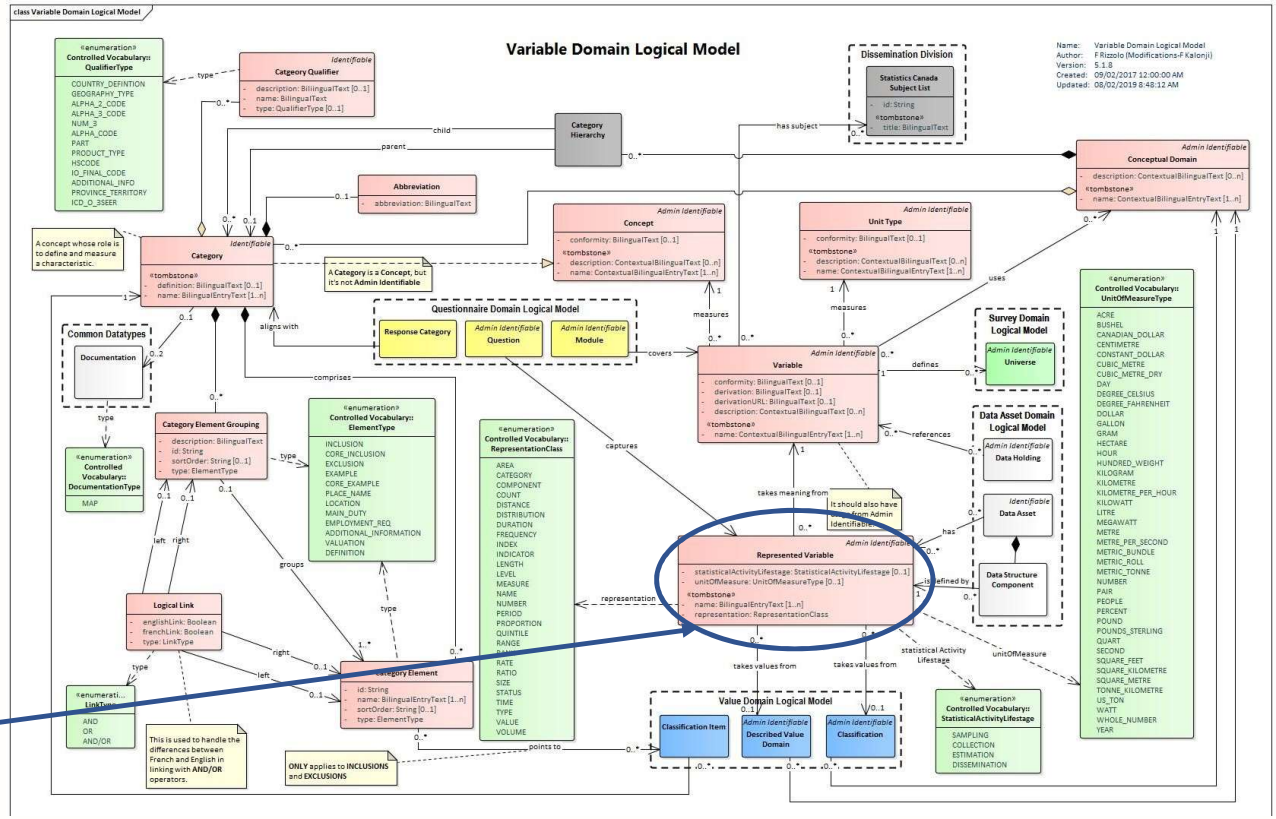


GSIM Concept Group



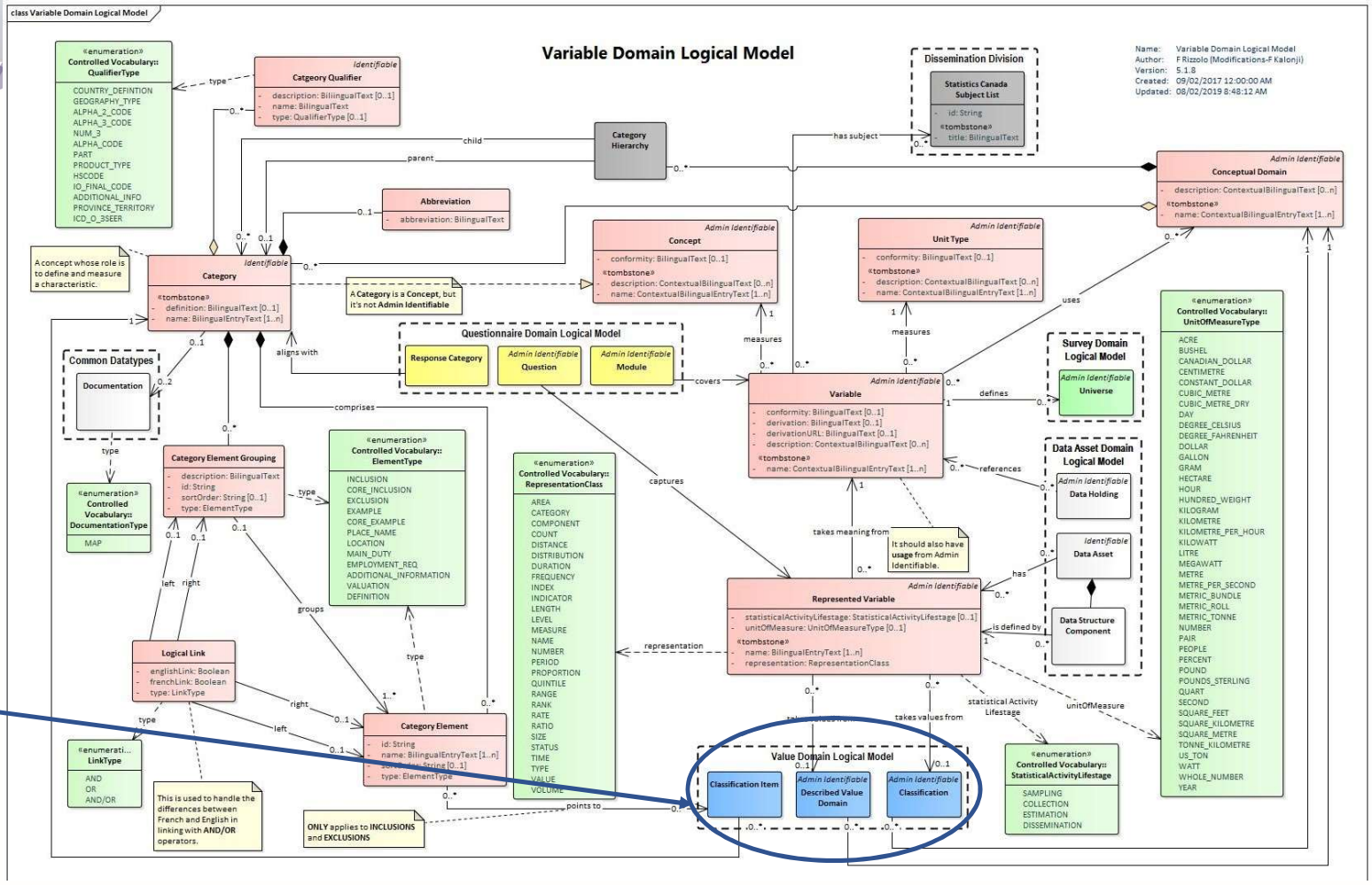
Represented Variable

Variable Domain Logical Model

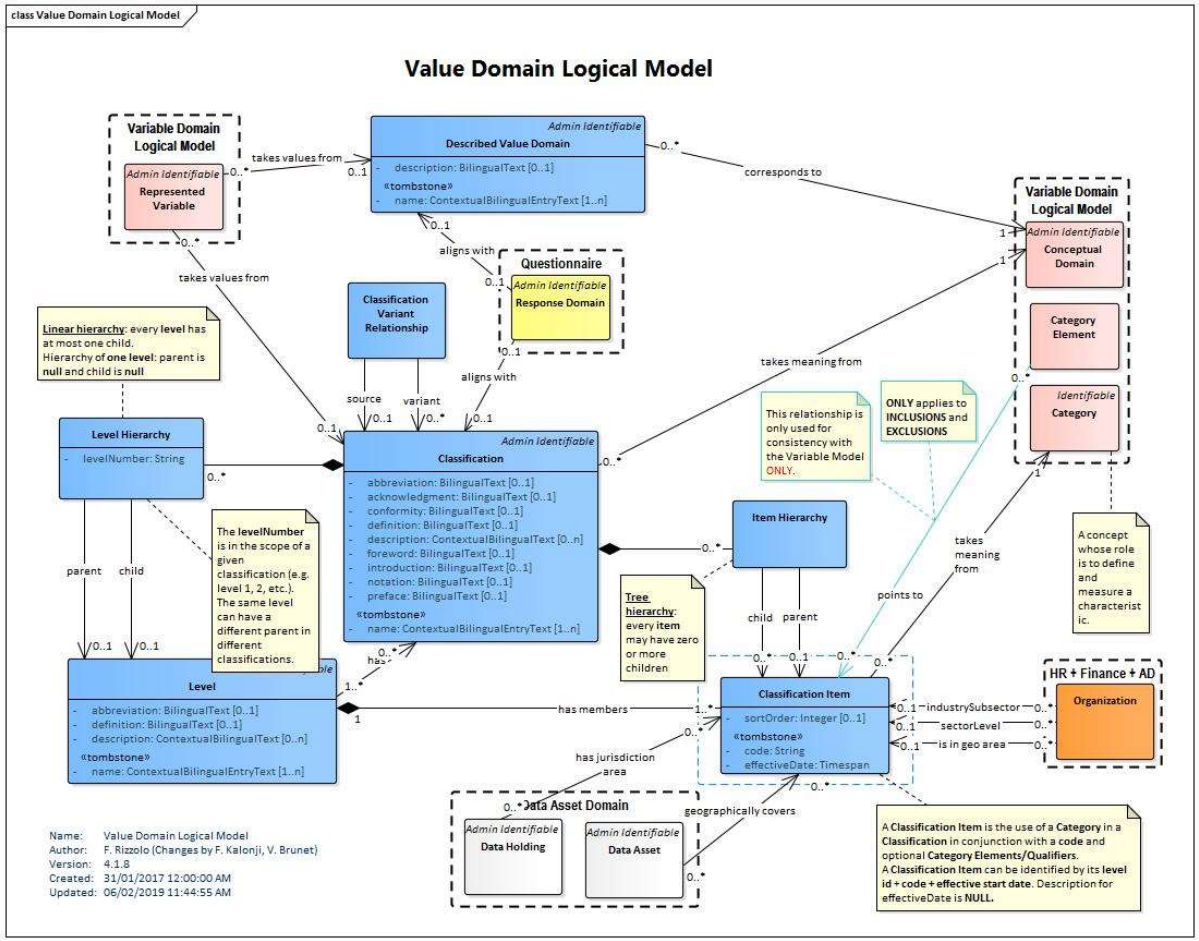


Variable domain logical model

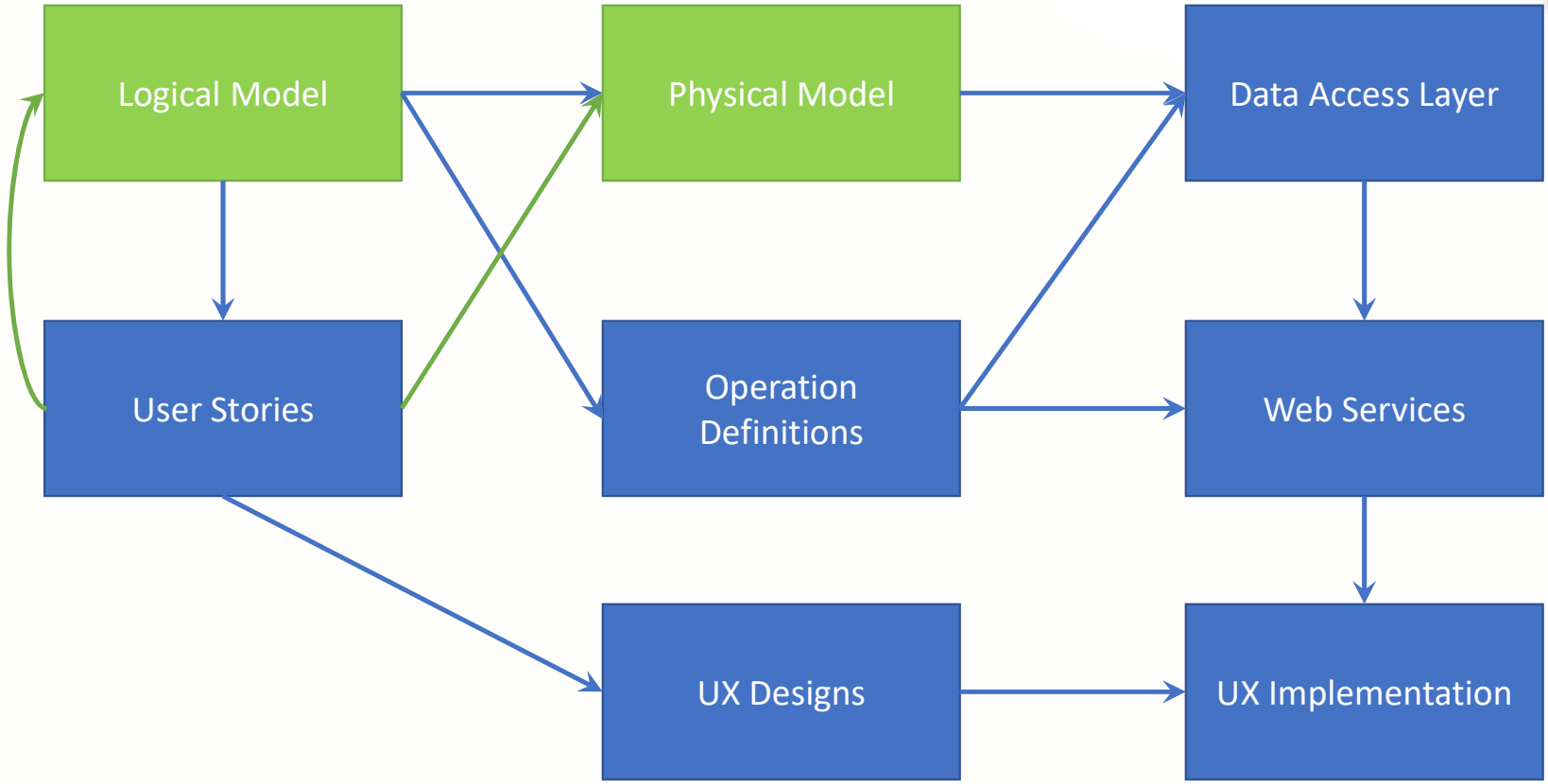
Links with other domains, Value Domain in this case



Value domain logical model



Overview of the process



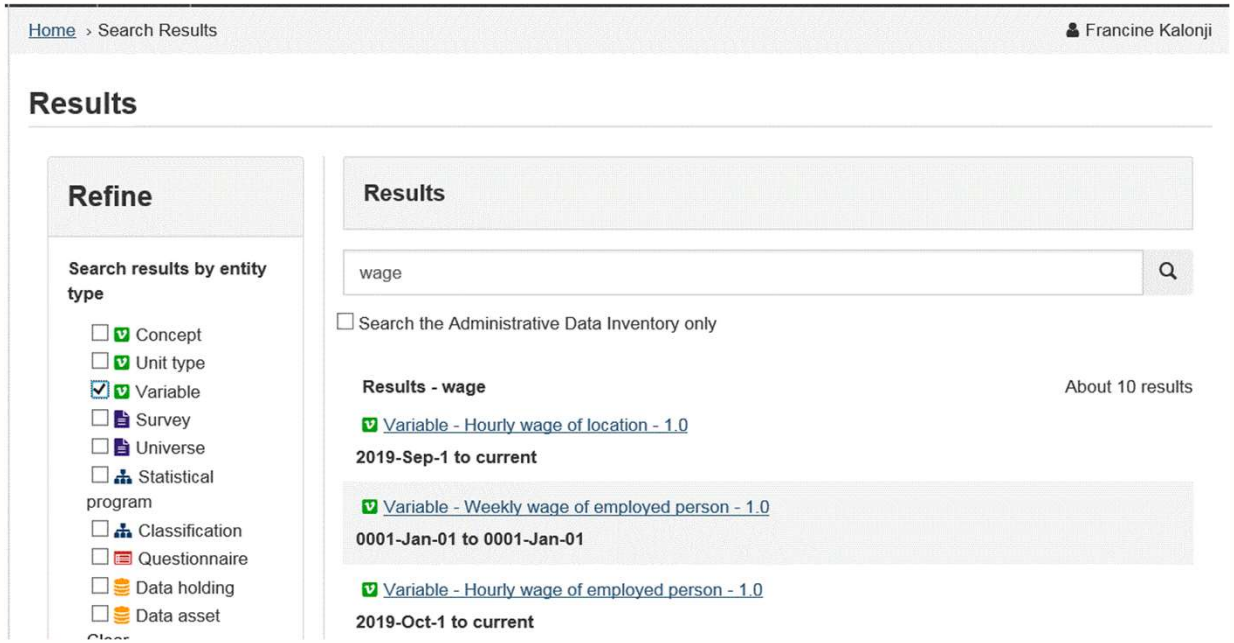
Operations Definition: Variable

Operation name	Return type	Parameters
getVariable	Variable[0..1]	String id Boolean[0..1] includeAssociatedTombstones ParamList[0..1] adminFilters ParamList[0..1] otherFilters String [0..1] contextId Boolean[0..1] includeAllContexts
getRepresentedVariable	RepresentedVariable[0..1]	
getVariableSummaries	VariableSummary[0..n]	Date[0..1] updatedAfter String[0..1] nameSubstring ParamList[0..1] adminFilters ParamList[0..1] otherFilters String [0..1] contextId Boolean[0..1] includeAllContexts
getRepresentedVariableSummaries	RepresentedVariableSummary[0..n]	
getVariableSeries	VariableSeriesTombstone[0..n]	String seriesId ParamList[0..1] adminFilters ParamList[0..1] otherFilters String [0..1] contextId Boolean[0..1] includeAllContexts
getRepresentedVariableSeries	RepresentedVariableSeriesTombstone[0..n]	

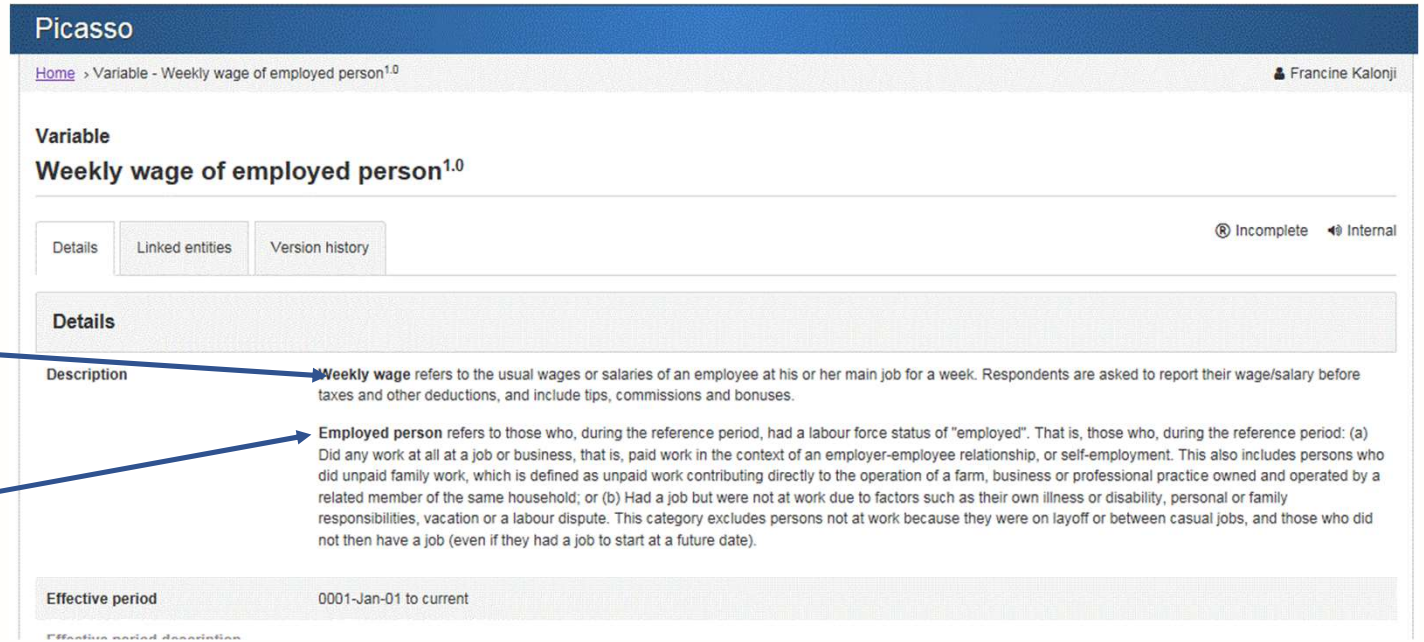
Picasso Application

- Picasso is Statistics Canada's one stop portal
- Enables management of data and statistical metadata
- A tool for search and discovery of data assets and statistical metadata
- Standards/frameworks
  - Is built on GSIM for data and statistical metadata
  - Uses ISO-11179 standard for the registration process
  - Uses Prov-O terms for tracking provenance information

Picasso: Search results for Wage



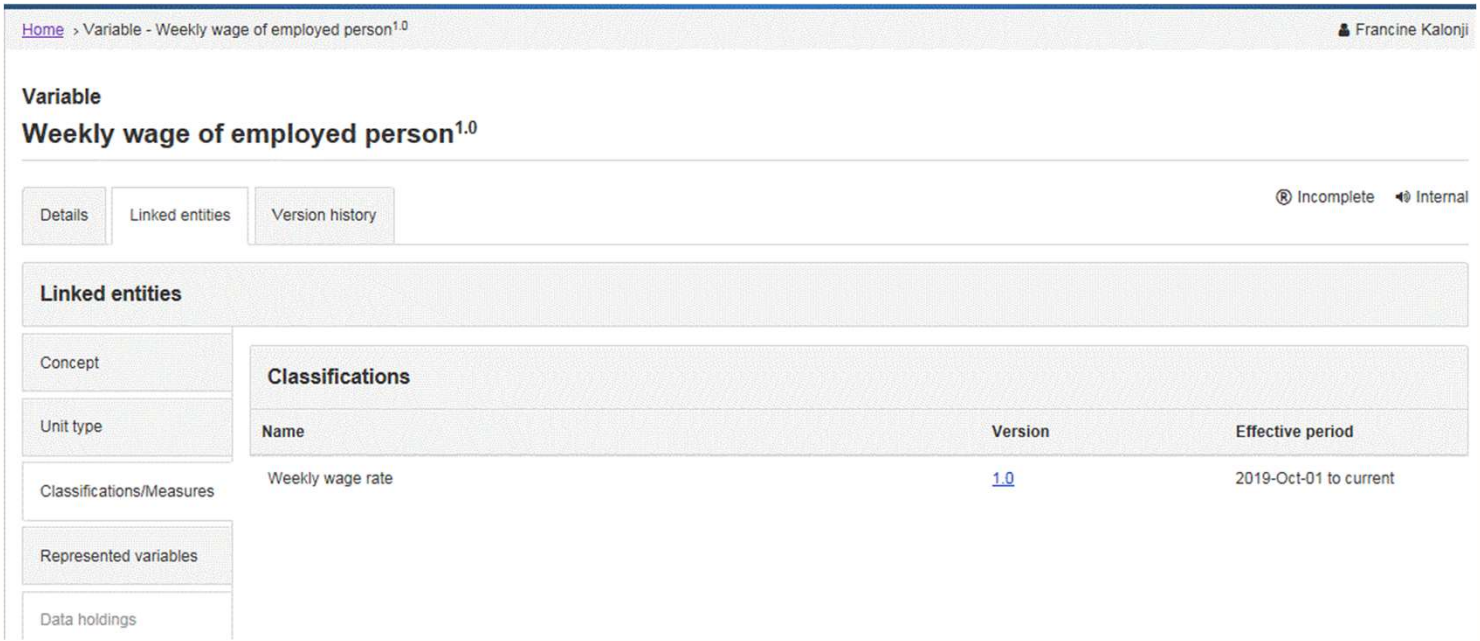
Picasso: Weekly wage of employed person



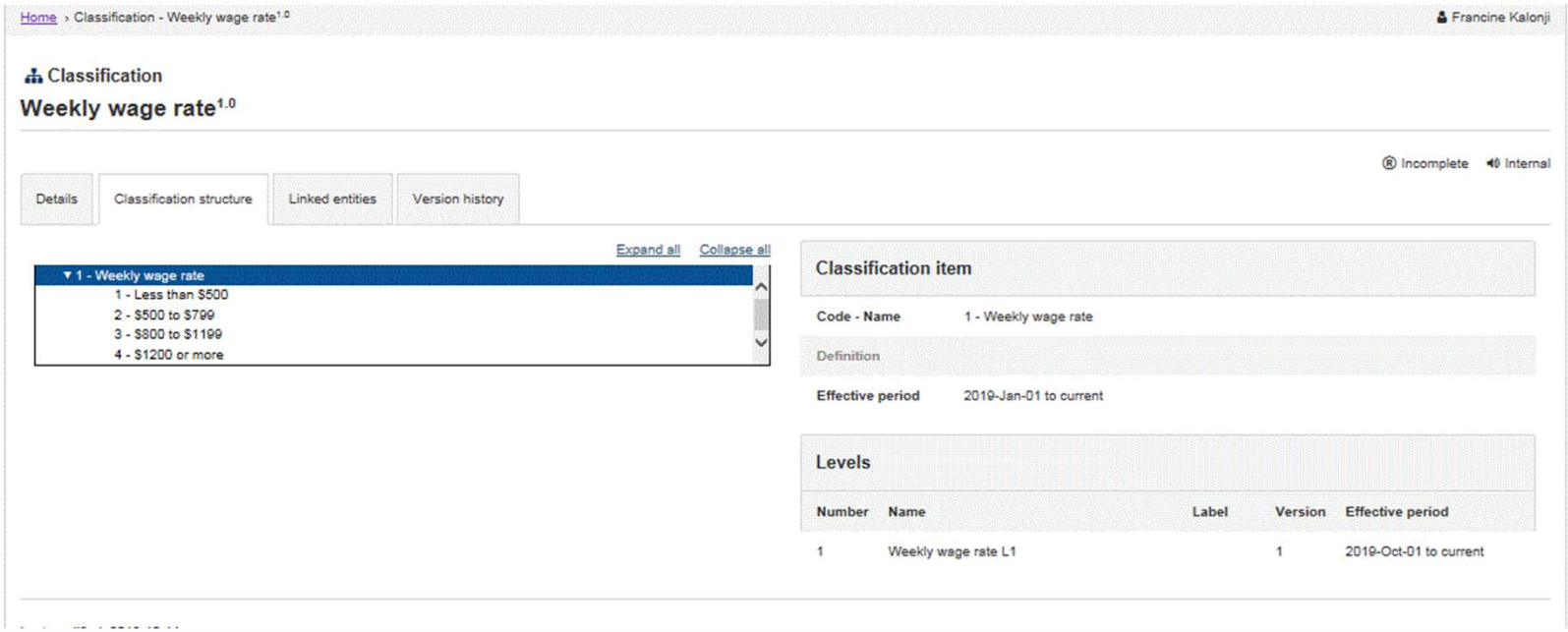
Concept

Unit Type

Picasso: Classification of Weekly wage rate



Picasso: Classification Structure of Weekly wage rate



Using the GSIM model

- Harmonization of our statistical metadata
- Interpretability and coherence
  - Common language leads to common solution
- Interoperability with the development of web services
- Increase of the visibility of our administrative data
  - Solid framework for representing our administrative data
- Links to standards like DDI and SDMX, used for our micro and aggregate data
- Mapping objects can be a challenge
  - Definitions, descriptive texts, examples, the community are helpful

Thank you!
Questions?