XKOS and how to use it

What is XKOS?

An extension of SKOS
for statistical classifications

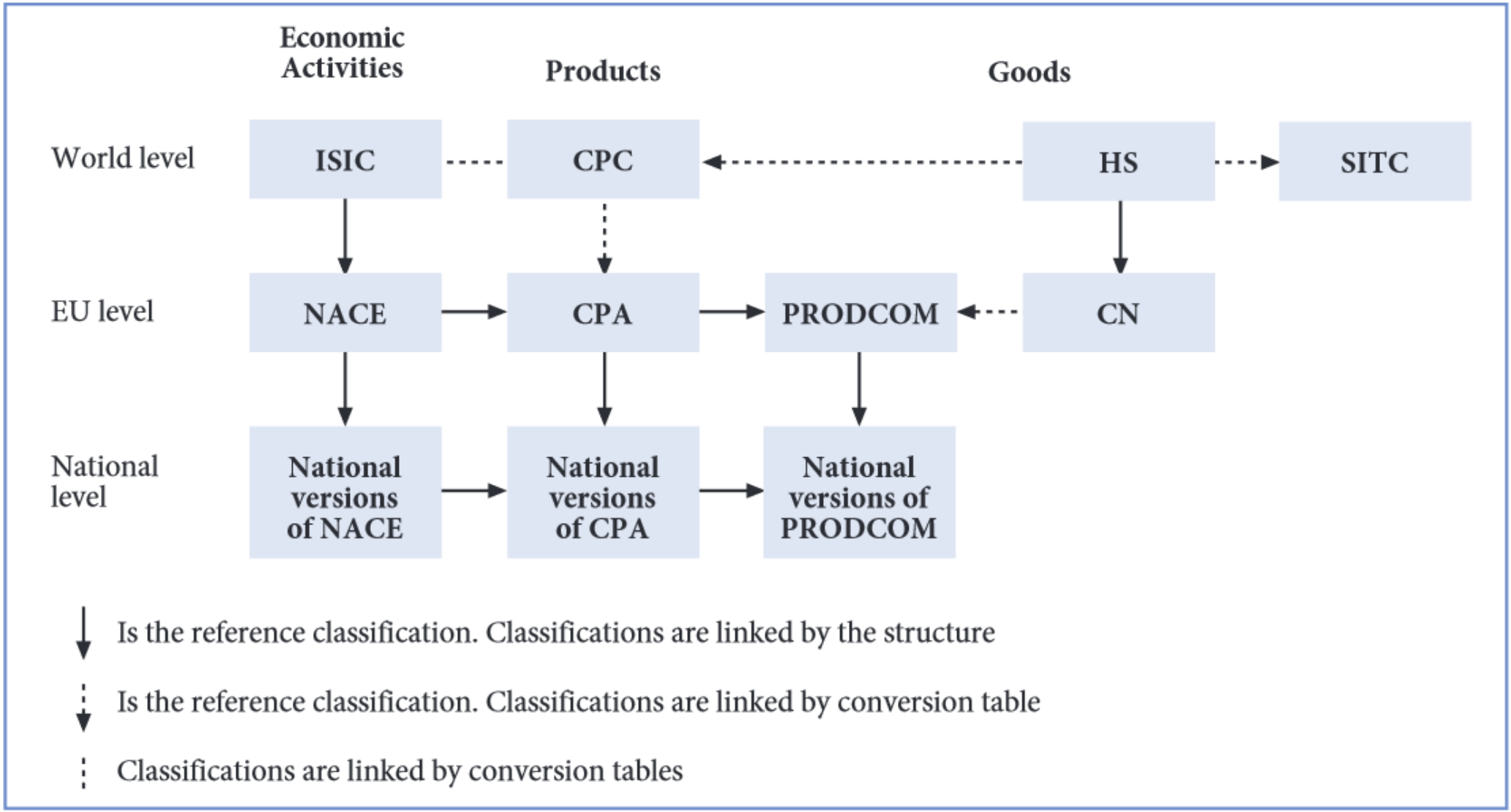
What are statistical classifications?

- Hierarchical structures of concepts
  - One or several levels of detail
  - Usually balanced hierarchies
- Covering a specific field
  - Economic activity, occupations, level of education, legal form...
  - Usually exhaustively and mutually exclusively

What are statistical classifications?

- Living objects
  - Published in major versions (classification schemes)
  - Minor modifications on a regular basis (notes)
- Linked by correspondence tables
  - Between one major version and the next
  - Between classifications on the same field
  - Between classifications on related fields

What are statistical classifications?



What is SKOS?

An RDF vocabulary
for Knowledge Organization Systems

What is an RDF vocabulary?

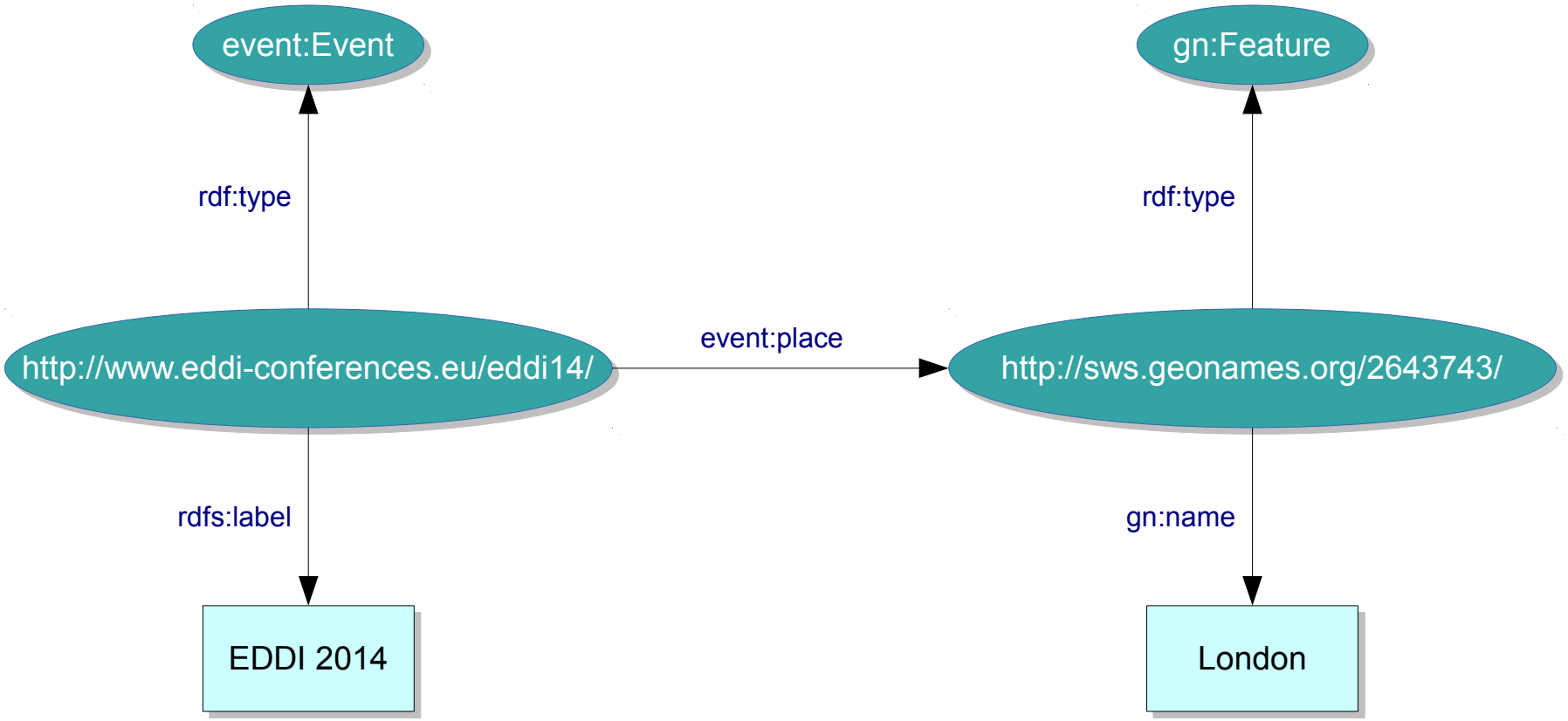
- RDF is a framework for expressing information about resources (RDF 1.1 Primer)
  - Information is modeled as statements (or triples)

    'EDDI 2014' 'takes place in' 'London'

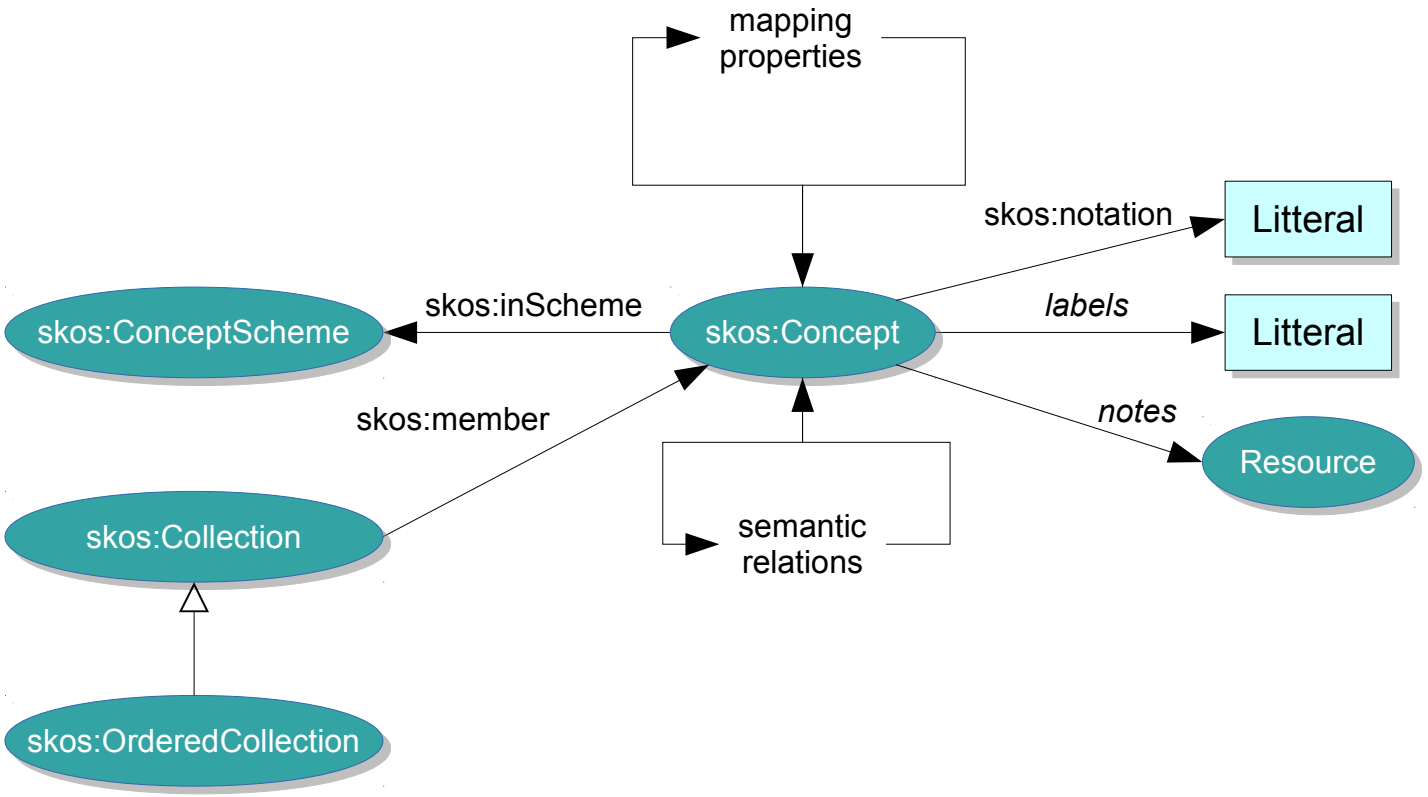
    `<subject>` `<predicate>` `<object>`

    `<subject>` and `<predicate>` are IRIs, `<object>` is IRI or literal
- RDF vocabularies add semantics to RDF data
  - EDDI 2014 is an event, London is a city
  - Vocabularies are defined in schema languages (RDFS, OWL)

An RDF example



The SKOS data model

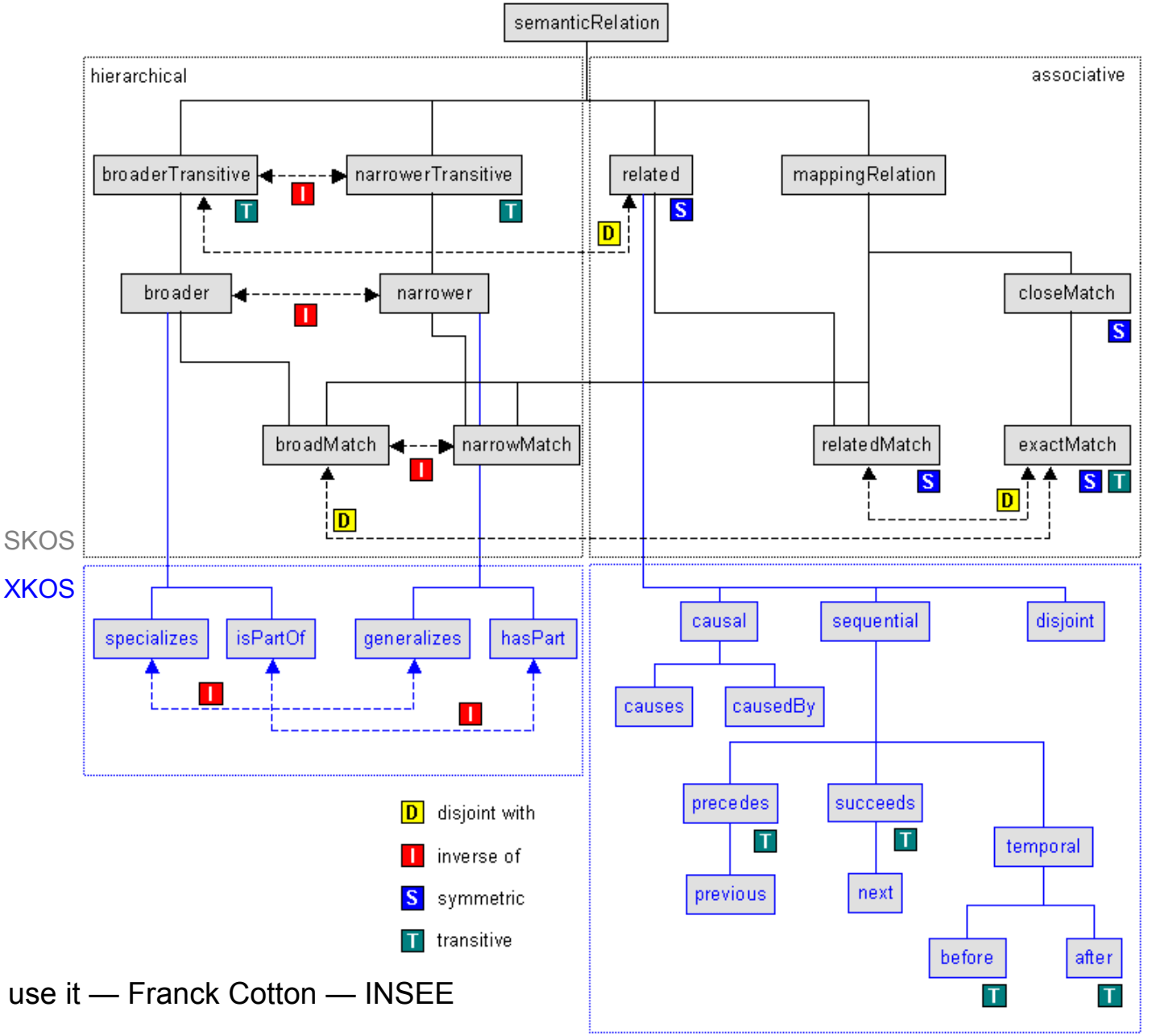


How does XKOS extend SKOS?

- Specific classes and properties
  - 3 additional classes
    - Classification level (a kind of SKOS Collection)
    - Correspondence and Concept Association
  - 36 additional properties
    - Business properties of classification management
    - More semantic relations

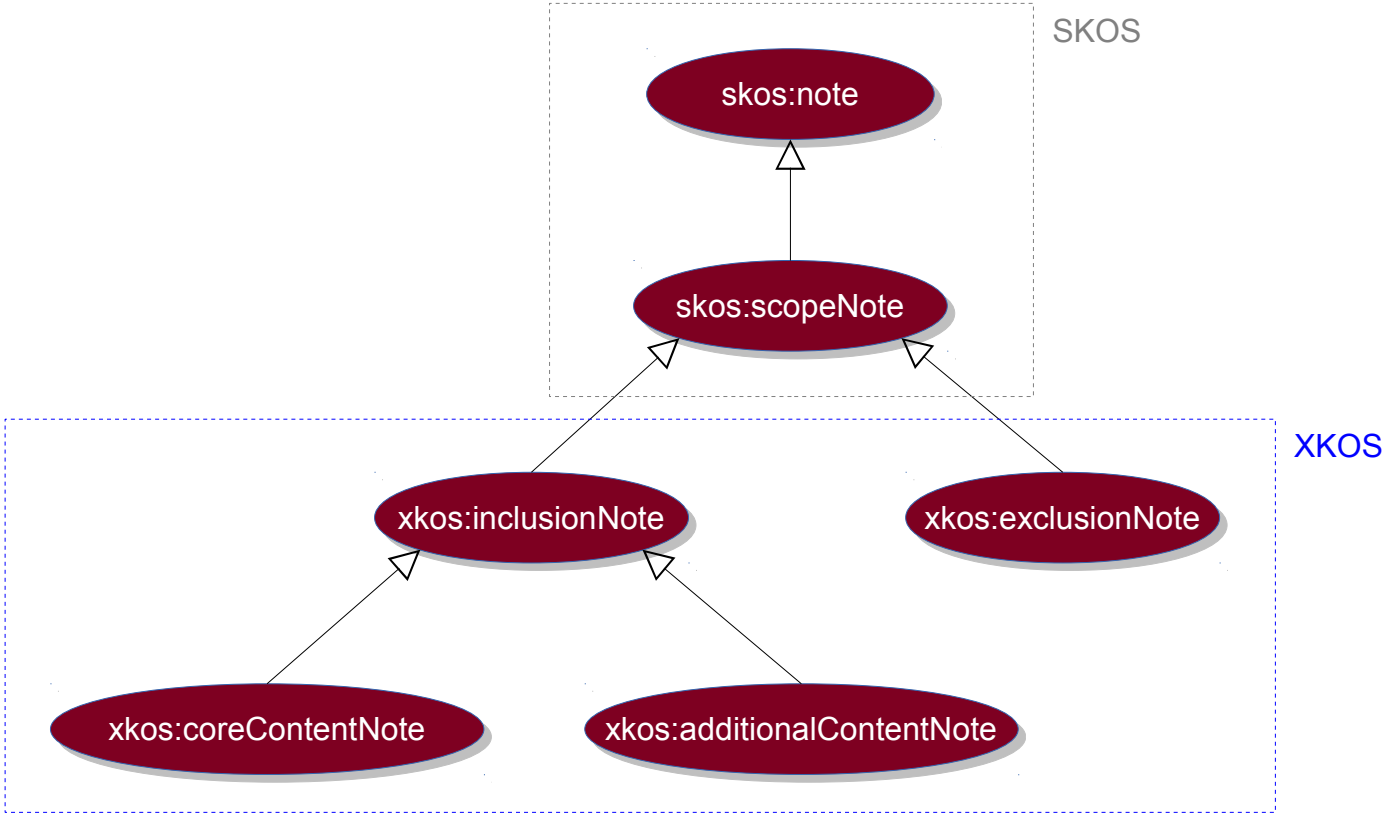
How does XKOS extend SKOS?

- Additional semantic relations



How does XKOS extend SKOS?

- Additional types of notes

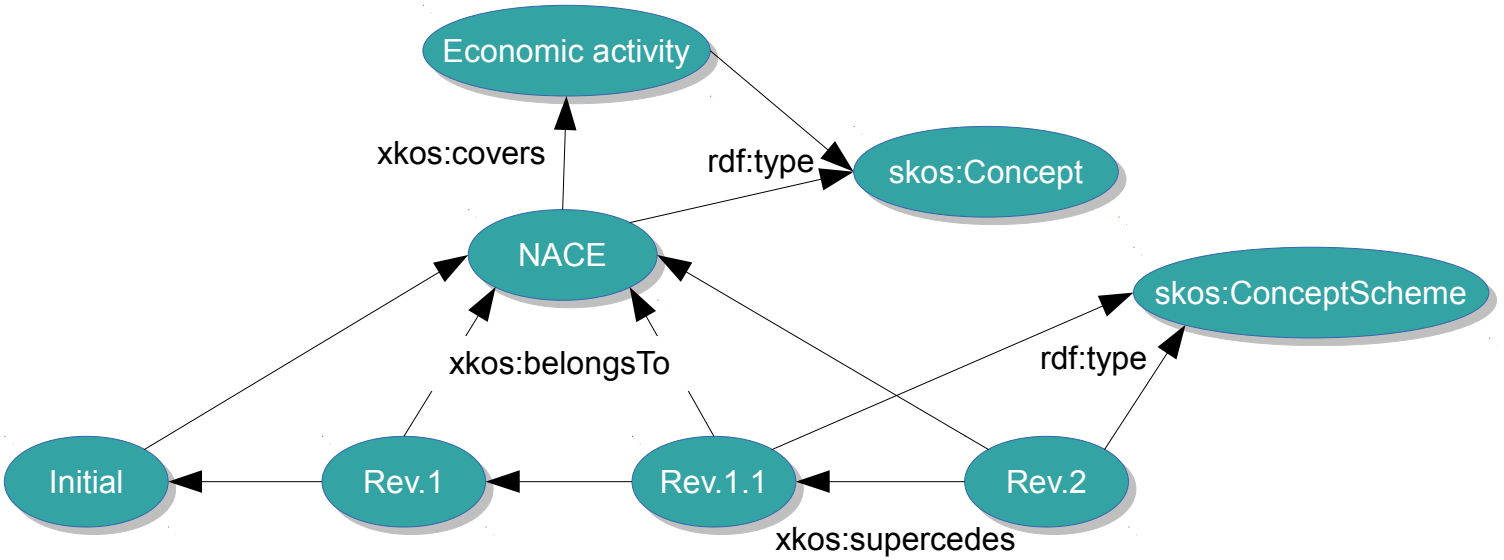


How to use XKOS?

- For classifications (obviously)
  - Some dedicated features
  - Important to publish classifications as Linked Data
    - Fundamental in the structuration of statistical data
    - Coordinated at the international level
- But not only
  - Semantic relations, for example:
    - Generic / specific and partitive relations
    - Sequential / causal relations for representing processes
  - Notes (inclusion, exclusion)

XKOS for statistical classifications

- Better representation
  - Linking a classification with the field it covers
  - Linking a classification and its major versions
  - Formalizing the classification levels



XKOS for statistical classifications

- Better representation
  - Better modelization of the correspondences
  - More precise typology of textual material
  - Richer semantics for relations between items
  - Possibility to represent how the actual objects are classified

XKOS for statistical classifications

- XKOS is SKOS
  - Use of SKOS mapping properties
  - Use of classifications and levels in Data Cube
- XKOS is RDF
  - Use of standard properties
    (for example rdfs:seeAlso, owl:sameAs)
  - Use of other vocabularies (PROV-O, DC, ADMS)
  - Inference
  - IRIs

XKOS for statistical classifications

- Enhances their value as Linked Metadata
  - Richer browsing
    - Classification structure is clearer
    - Links to other classifications or versions
  - Smarter queries
    - Avoid to match on exclusion notes
    - Queries on concept associations
  - Easier extractions, in multiple formats
    - Levels
    - Correspondences

XKOS current state

- Most of the work done during two Dagstuhl workshops (September 2011 and October 2012)
  - Other specification authors: Richard Cyganiak, Daniel Gillman, R.T.A.M. Grim, Yves Jaques, Wendy Thomas... and many more contributors
- 2012-2013: real implementations (Istat, INSEE)
- Public review completed in 2014
- Final publication at the beginning of 2015 by the DDI Alliance

XKOS and how to use it

Thank you