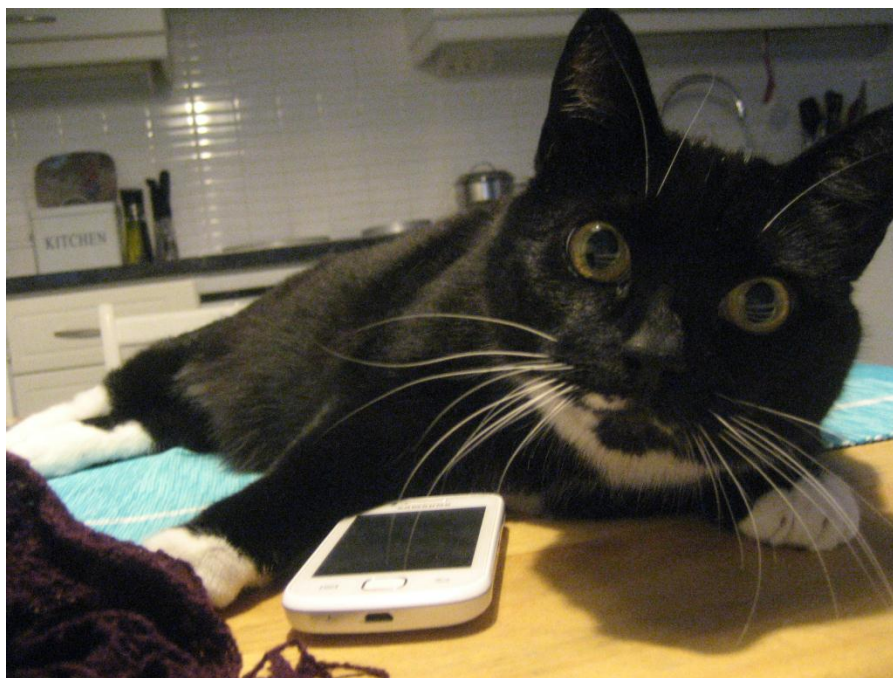


# Building a National Ontology Infrastructure

Matias Frosterus, Mirja Anttila,  
Mikko Lappalainen, Susanna Nykyri,  
Tuomas Palonen, Sini Pessala  
SWIB 2013

# This presentation

- Overview of the ONKI project
- Linked ontology approach
- Trilingual ontology



# ONKI project

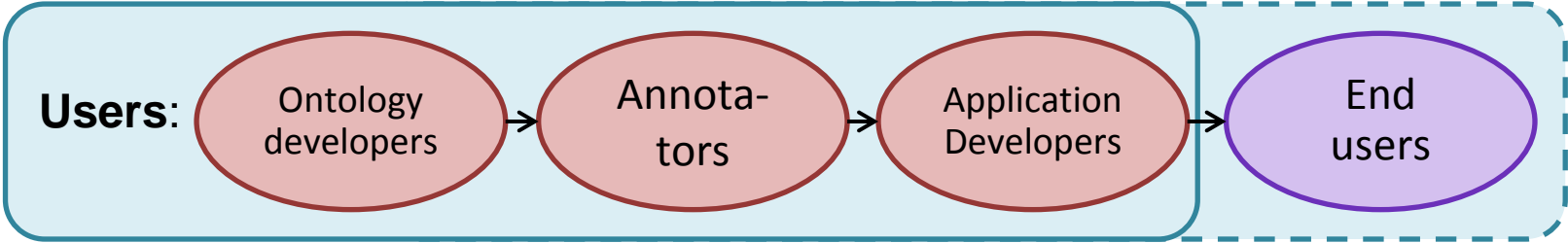
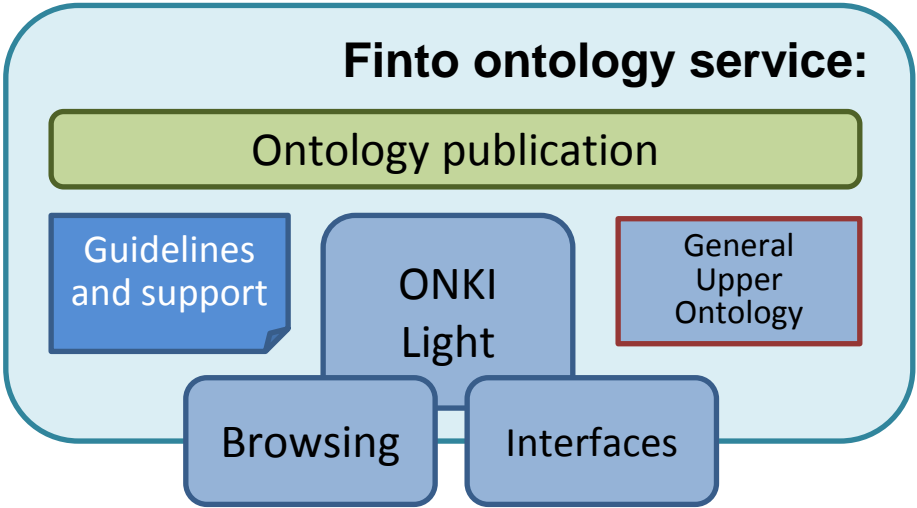
- A joint project of the National Library of Finland, the Ministry of Finance and the Ministry of Education and Culture
- The aim is to build a reliable, centralized, national ontology service named Finto

# ONKI project

- What does the ONKI project offer?
  - Publication of ontologies
  - Using ontologies in applications through various interfaces
  - The development of the General Finnish Upper Ontology YSO
  - Coordination of ontology work on national scale
- Improving interoperability across the spectrum by harmonizing annotations

# ONKI project

- Based on the FinnONTO research project, which ran in Aalto University and the University of Helsinki 2003-2012
- Focus on light-weight SKOS ontologies intended for annotations
- Powered by ONKI Light
  - Open source
  - <https://code.google.com/p/onki-light/>



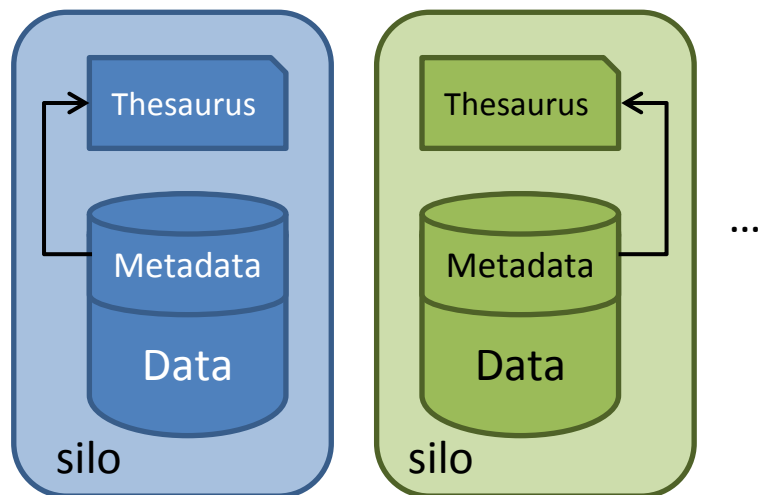
# The second part

- Linked ontology approach



# Linked ontology approach

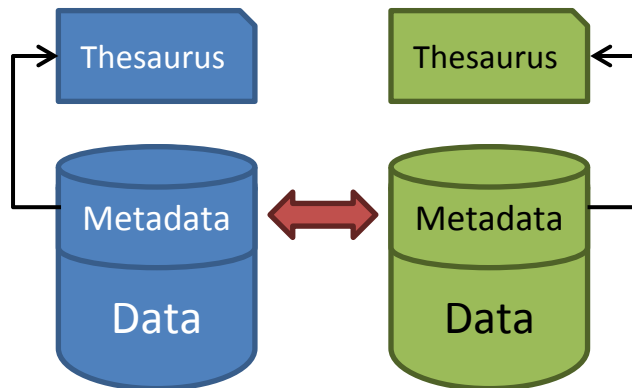
- What we have:
  - Silos
  - Expert-made thesauri
  - A large amount of data and annotations





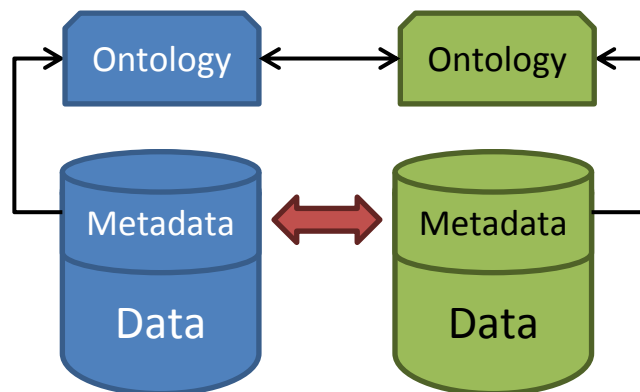
# Linked ontology approach

- What we want:
  - Eliminate the silos
  - Harmonize the annotations



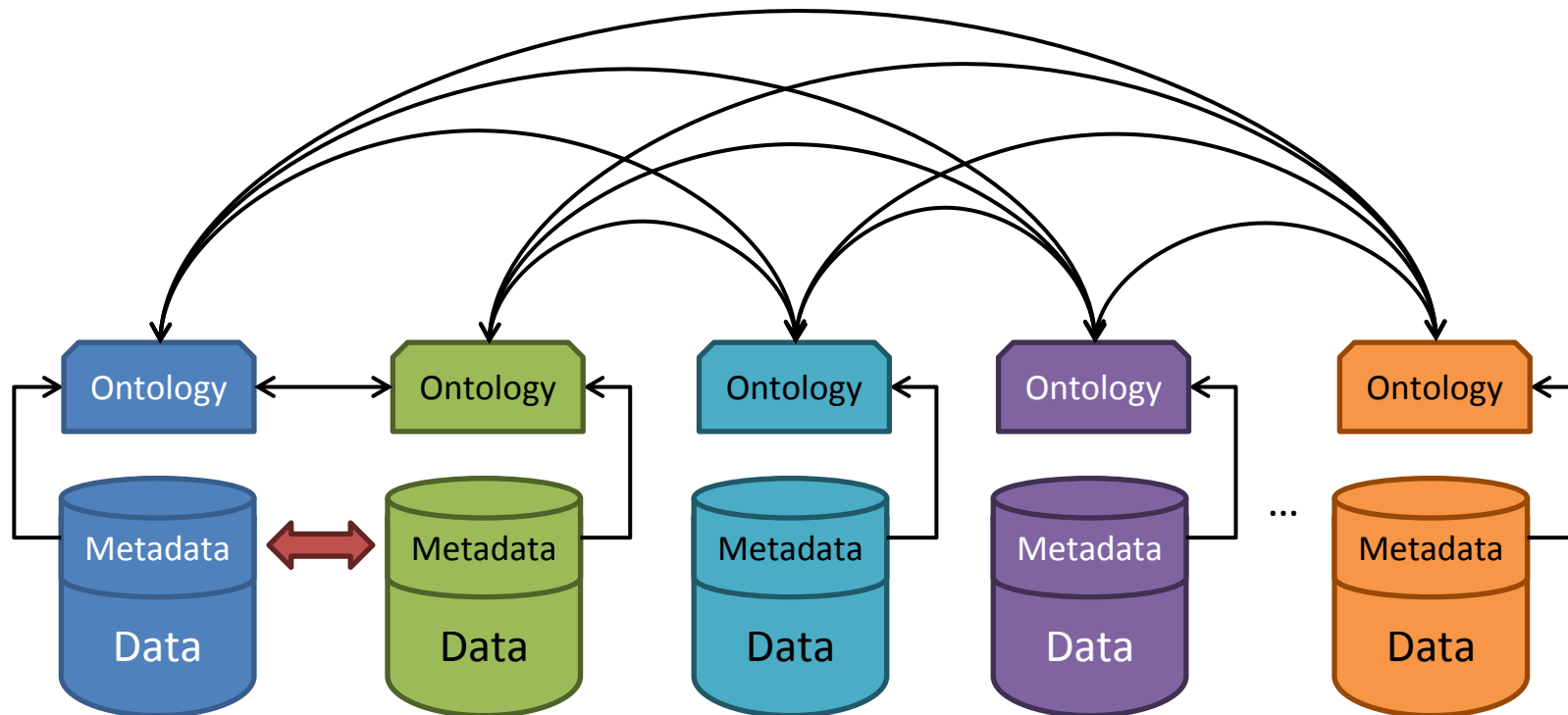
# Linked ontology approach

- How?
  - Ontologies are much easier to link together than thesauri
    - Concepts as opposed to terms
    - Explicit relations



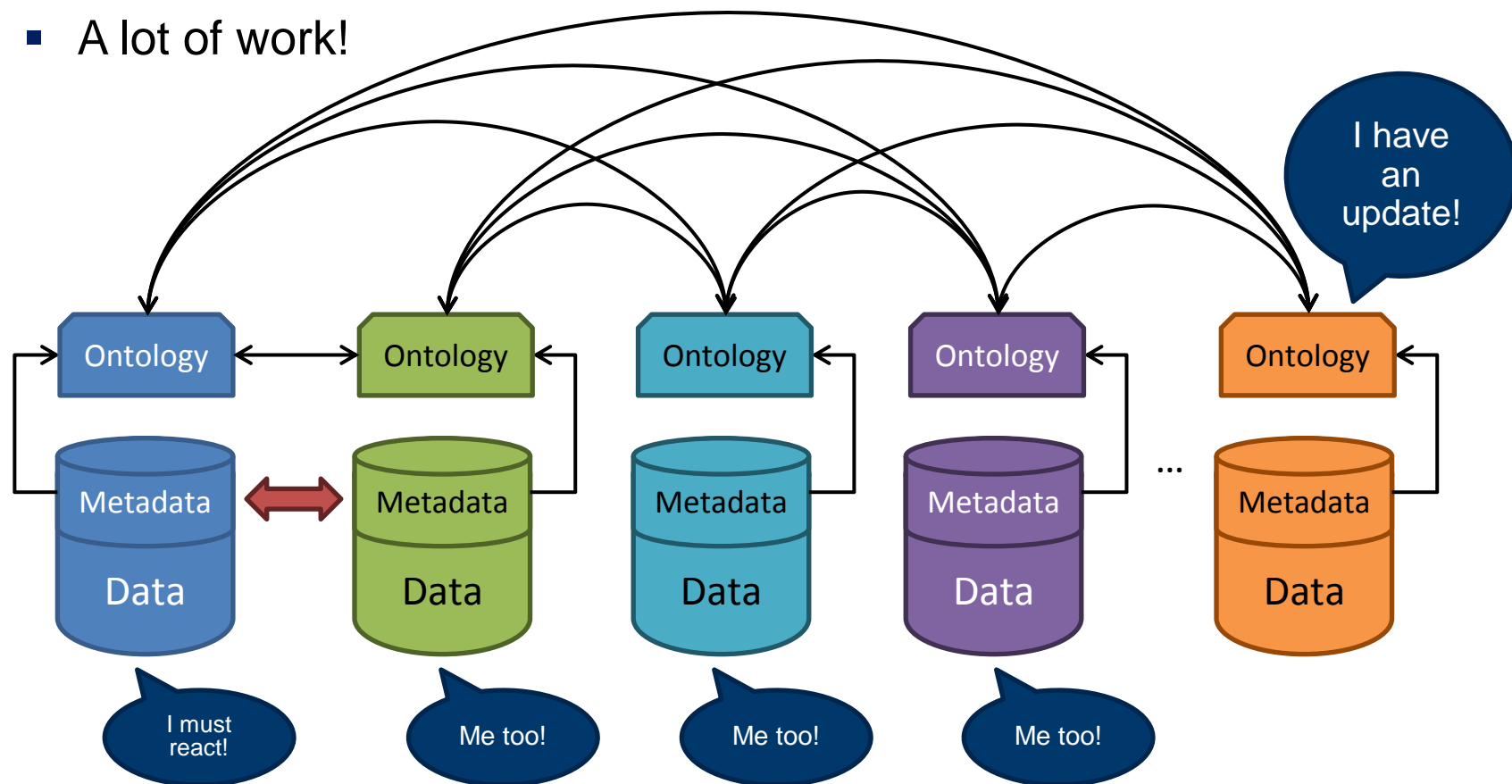
# Linked ontology approach

- The problem:



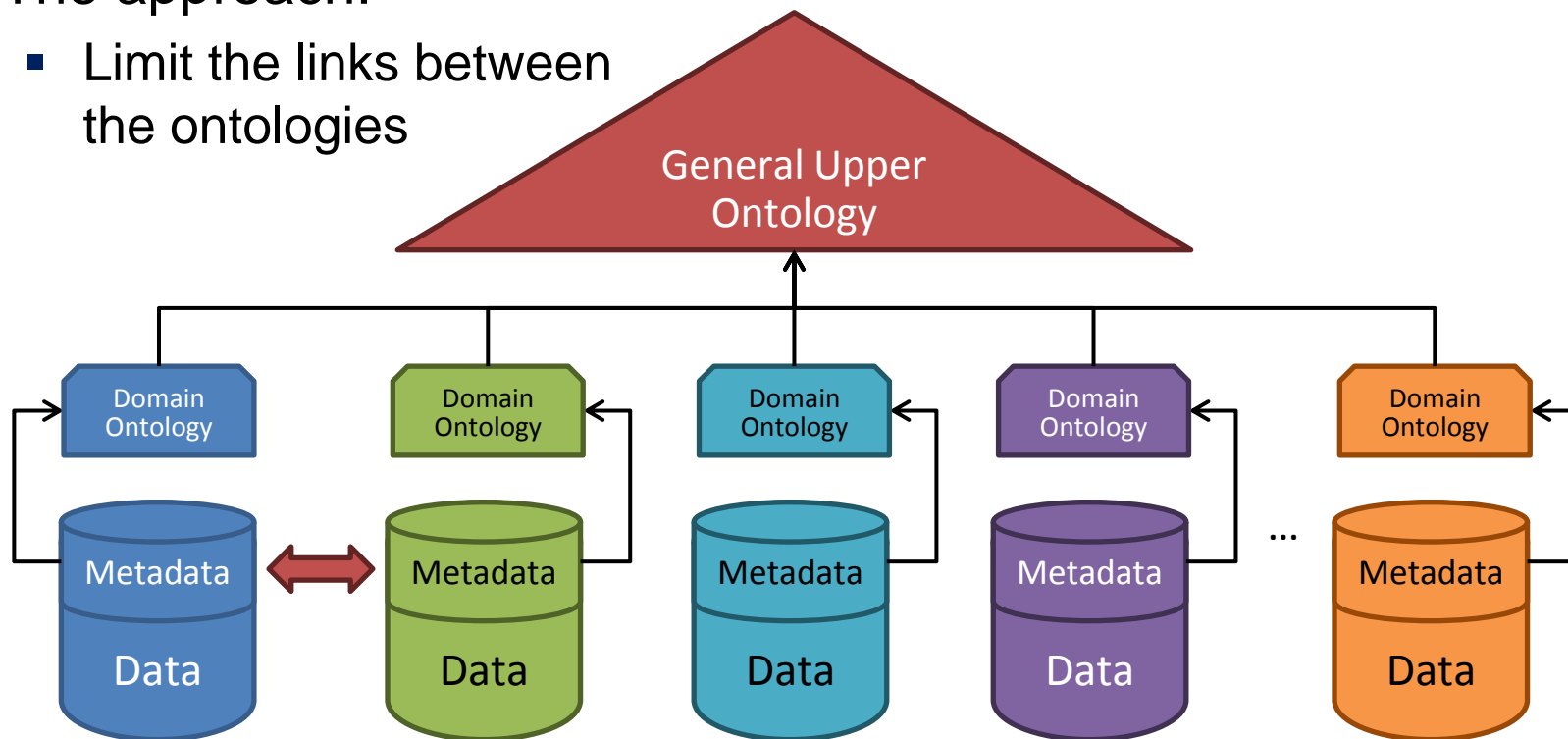
# Linked ontology approach

- The problem:
  - A lot of work!



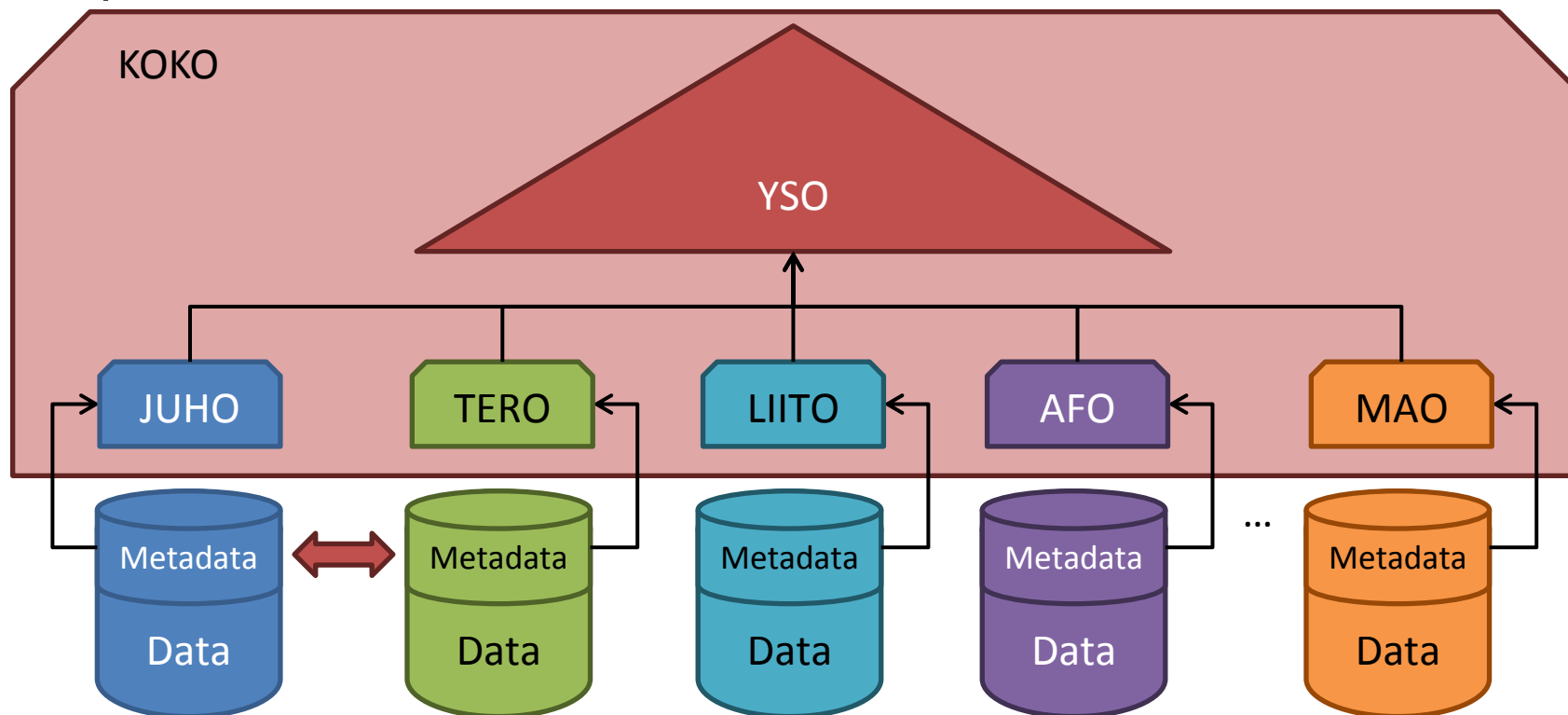
# Linked ontology approach

- The approach:
  - Limit the links between the ontologies



# Linked ontology approach

- In practice:



# Linked ontology approach: KOKO

Ontology	Domain	Concepts
YSO	General upper ontology	24 800
MAO	Museum artifacts	6 800
MUSO	Music	1 000
TAO	Design	3 000
TERO	Health	6 500
VALO	Photography	2 000
AFO	Agriculture	7 000
JUHO	Government	6 300
KAUNO	Literature	5 000
KTO	Linguistics	900
KITO	Literary research	850
KULO	Cultural research	1 500
LIITO	Economics	3 000
MERO	Seafaring	1 300
PUHO	Military	2 000

# Challenges to be tackled

- Propagating the changes in the upper general ontology to the domain ontologies
- Locating the overlapping concepts between the domain ontologies
  - Not always simple
    - Labels might be misleading
    - Ontological structure can help
- Coordinating the use and development of ontologies on a national level



# The third part

- Trilingual ontology



# Ontology design

- The relations between concepts can be designed in several ways
- What affects these choices?
  - Corpora
  - Language
  - Culture

# Trilingual ontology

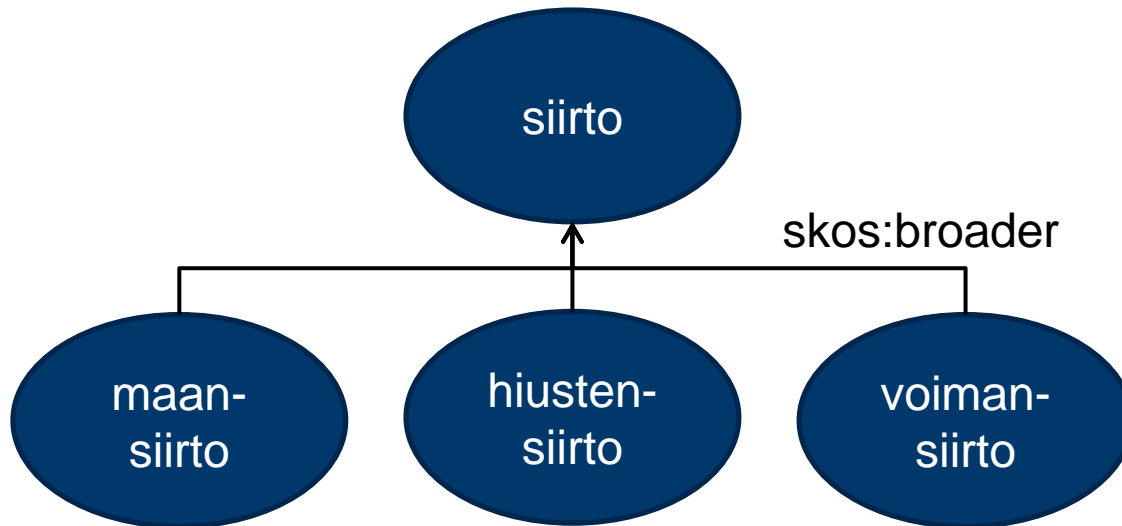
- In practice
- YSO: General Finnish Upper Ontology
  - "Finnish" as a culture
  - Finland has two official languages: Finnish and Swedish
    - Very different from one another
  - Lingua franca: English

# YSO

- Topmost hierarchy is inspired by DOLCE
- Offers the general concepts needed for annotation in many domains
- Complemented with a number of domain ontologies for specific use cases
- Based on the General Finnish Thesaurus YSA
  - Used and developed for decades in the annotation of all Finnish non-fiction literature

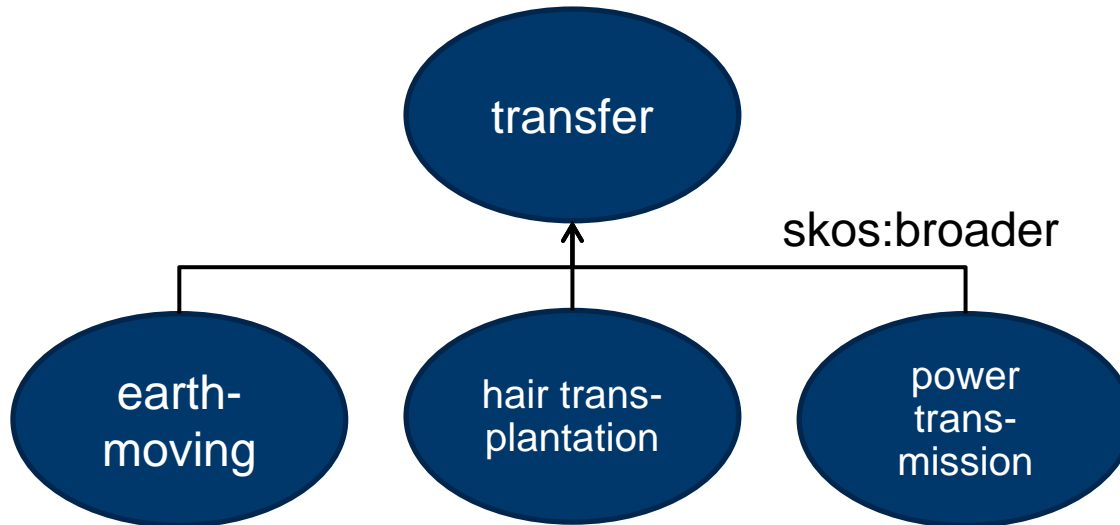
# Language affects the hierarchy

- Finnish word 'siirto' means transfer



# Language affects the hierarchy

- Finnish word 'siirto' means transfer



# Language affects the hierarchy

- Finnish has a single concept for rivers
- Swedish has three
  - Älv = Scandinavian river situated north of Göta älv (a specific river)
  - Å = Scandinavian river situated south of Göta älv
  - Flod = non-Scandinavian river
- A distinction not used in Finland

# Culture before language

- Looking beyond the language
- Realizing that language does affect the way we perceive the world
- Building an ontology for a specific cultural sphere
- Key to the harmonization of different annotations in different domains



# The development of YSO

- Mapping to other ontologies
  - Mapping to LCSH is underway
- Building the guidelines for the development
  - How to choose the correct approach when language clash leads to concept clash?

# Thank you!

[matias.frosterus@helsinki.fi](mailto:matias.frosterus@helsinki.fi)

[onki-posti@helsinki.fi](mailto:onki-posti@helsinki.fi)

