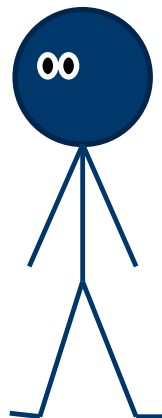# From Librarian to Datarian: How to Combine Technical Expertise, Librarian Skills and Build a Competence Network

Matias Frosterus
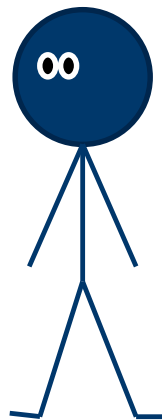
13th International Bibliotheca Baltica Symposium
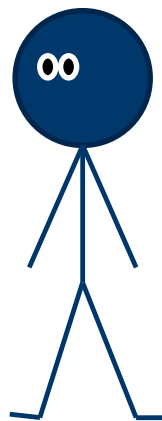
Riga 20.10.2015

A long, long time ago
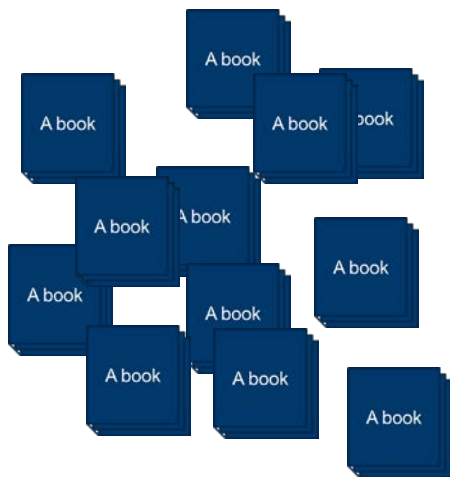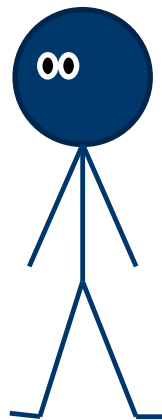

A book

A long, long time ago

A book

A book storage guy
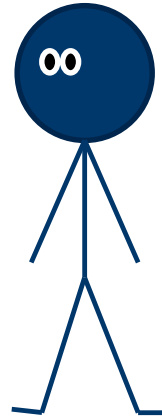
A long, long time ago

A book storage guy

A long, long time ago

A book storage guy
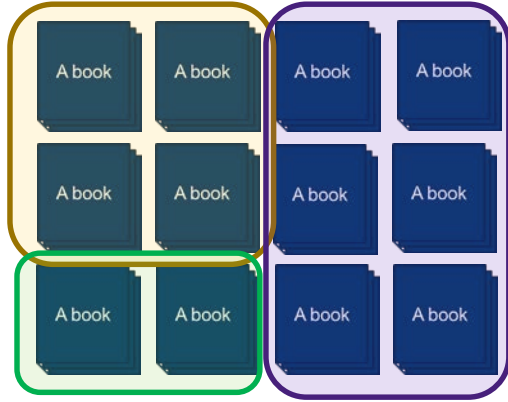
A long, long time ago

A book storage guy

A long, long time ago
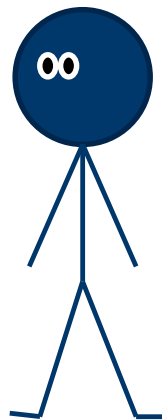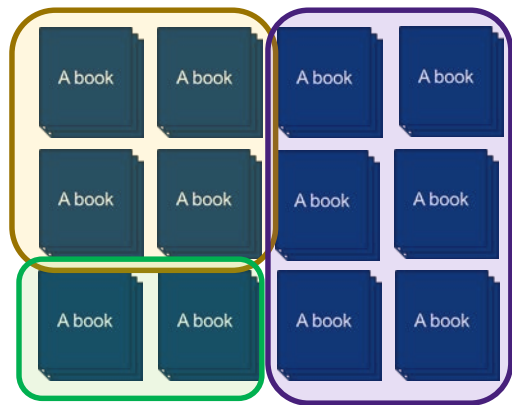
A librarian

The need for librarianship

- Every middle- to large-sized organization these days is producing and consuming a library's worth of information
    - Yet they do not generally employ librarian know-how
- Text search is great!
    - But we all know the limitations of it

- Granting access to information is one of the key aspects of librarianhood
    - Metadata know-how

New openings: FinnONTO Research project

2003-2012

FinnONTO Research Project
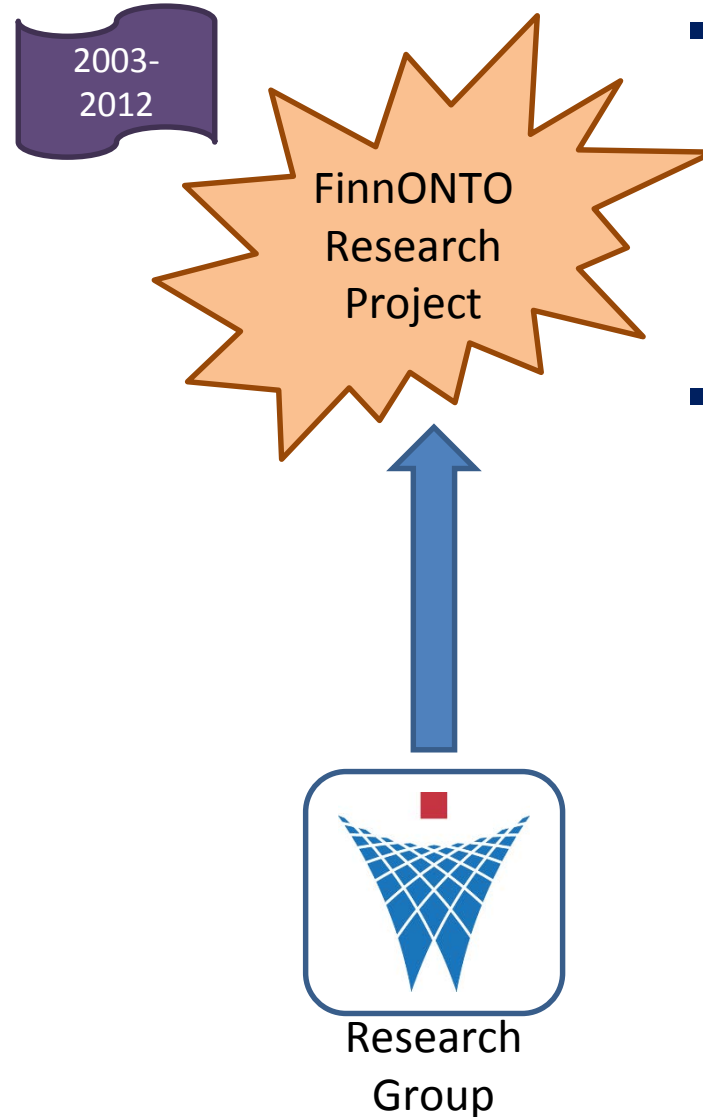
- A research project in the Aalto University and the University of Helsinki
- Aim of building a semantic web infrastructure for Finland
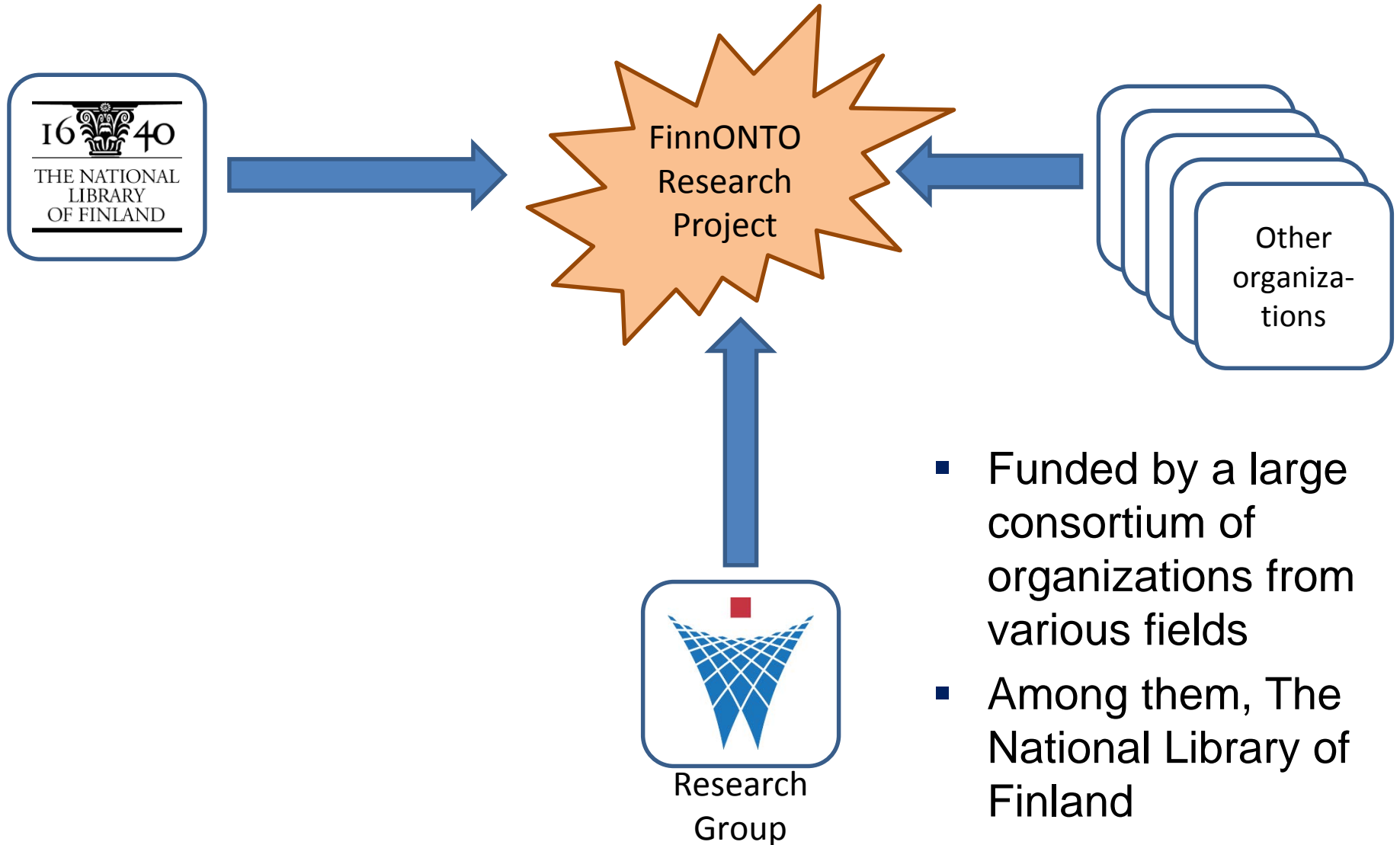
Research Group

New openings: FinnONTO Research project

New openings: FinnONTO Research project

FinnONTO
Research
Project

Other organiza-tions

Research, prototypes

Research Group

- Researchers, publications, prototypes
  - Testing various approaches
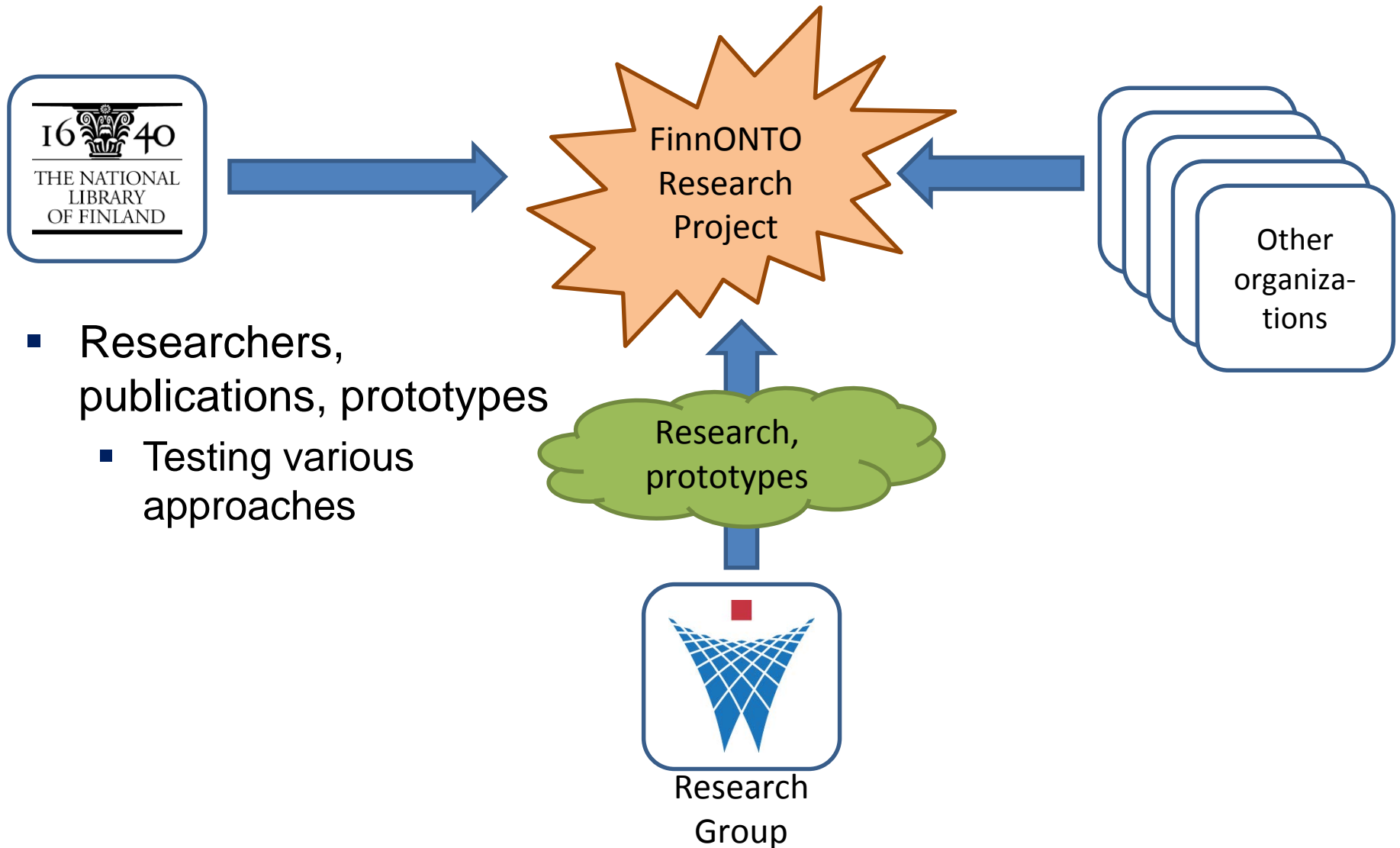
New openings: FinnONTO Research project
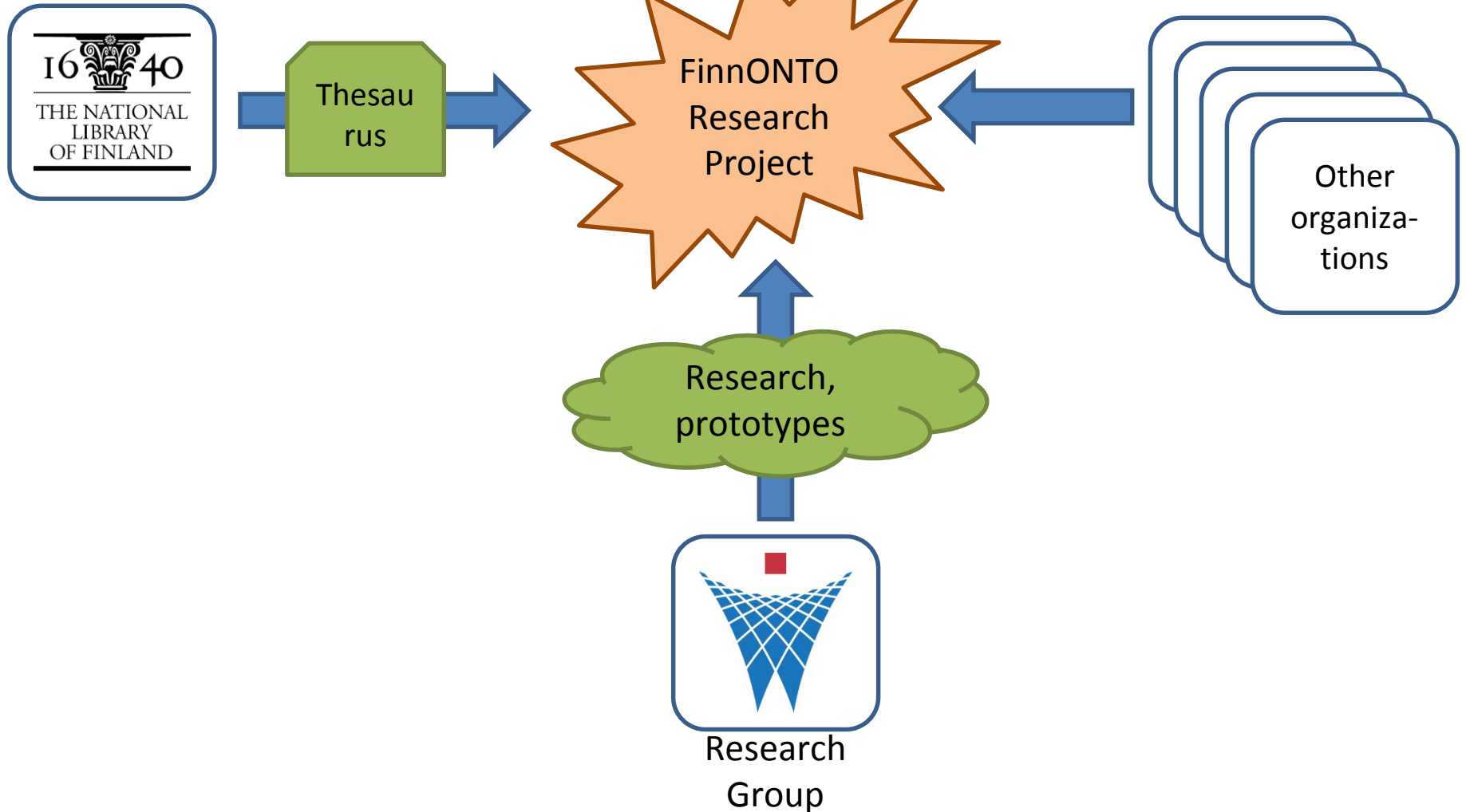
New openings: FinnONTO Research project
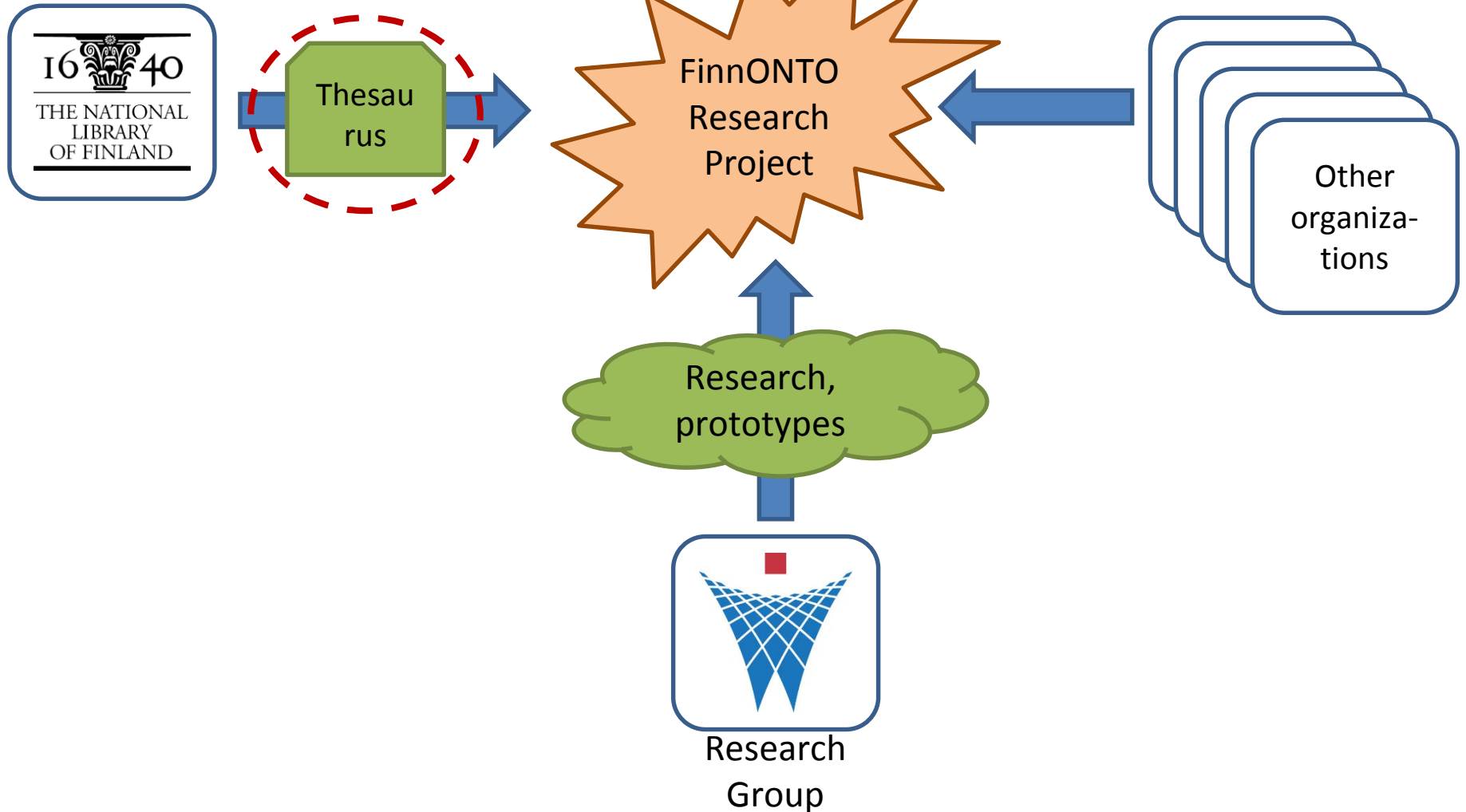
From thesauri to light-weight ontologies

- National Library's General Finnish thesaurus YSA was/is the most used thesaurus in Finland
    - Developed originally in the 1980's mainly for book indexing
        - Used to annotate, among other things, all non-fictional literature published in Finland
    - Monolingual, has a Swedish equivalent Allärs
    - Structure tied to MARC21-format, for example use of MARC-subdivisions

From thesauri to light-weight ontologies

- Semantic web and linked data
- Enabling use outside of the library domain
    - multiliguality
    - machine processability
    - open linked data

From thesauri to light-weight ontologies

- From YSA to YSO, the General Finnish Ontology
  - From terms to concepts
    - Multilinguality
      - Finland has two official languages so very important
    - Linkability
  - Use dereferenceable URIs
  - Machine-processable relations between concepts
    - Complete hierarchy
  - Depicted using RDF

YSO
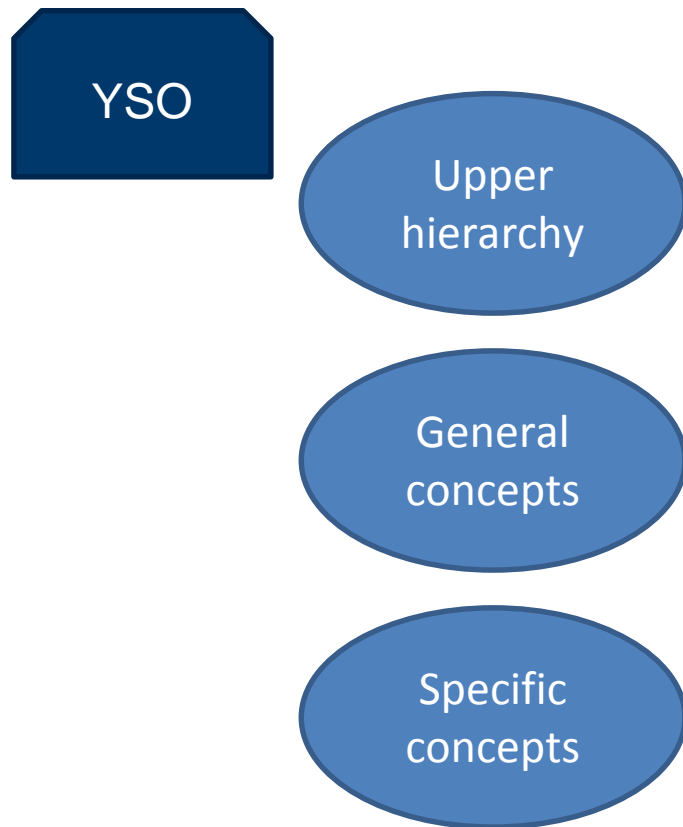
YSO
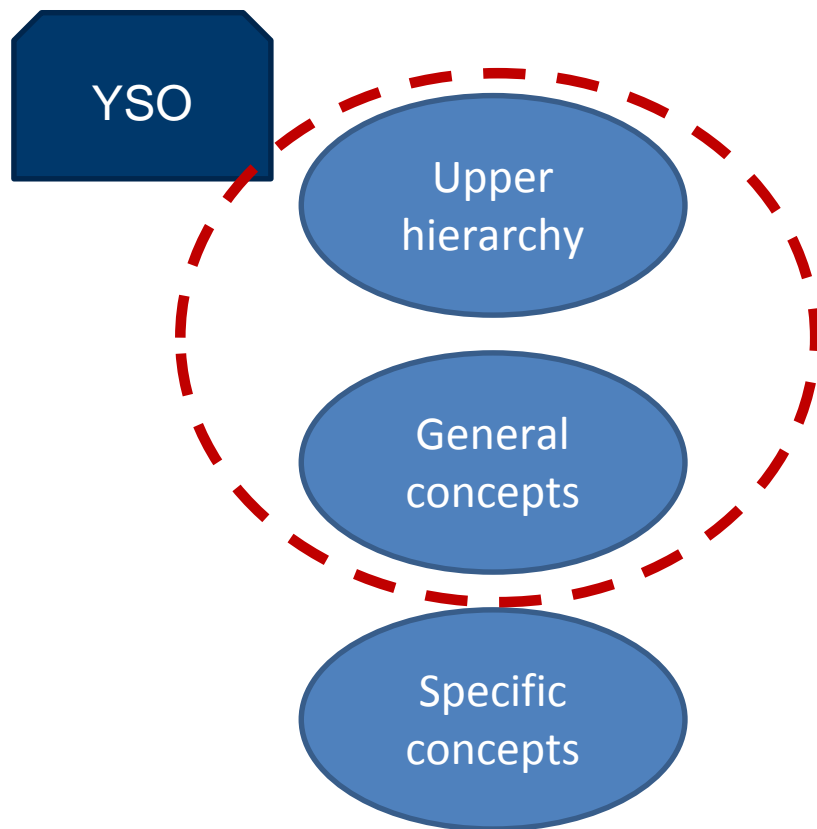
Upper hierarchy

General concepts

Specific concepts

YSO

YSO

Upper hierarchy

General concepts

Specific concepts

Adapted into use outside the library domain

- Extended with domain ontologies
  - Using the core provided by YSO
    - Helps interoperability!
  - Developed by the domain experts in various organizations
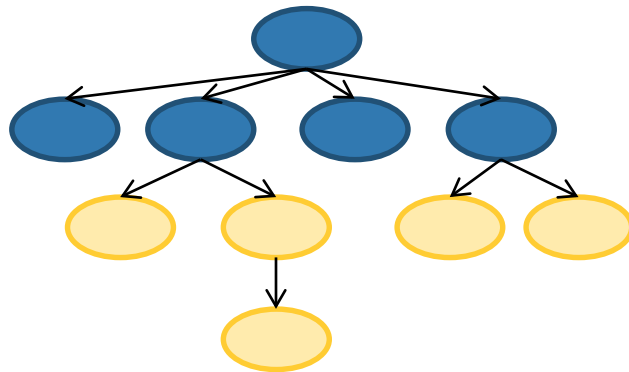
Adapted into use outside the library domain

- Extended with domain ontologies
  - Using the core provided by YSO
    - Helps interoperability!
  - Developed by the domain experts in various organizations
  - Over a dozen domain ontologies such as:
    - AFO - Agriculture - 7 000 concepts
    - JUHO - Government - 6 300
    - KAUNO - Literature - 5 000
    - KULO - Cultural research - 1 500
    - LIITO - Economics - 3 000
    - PUHO - Military - 2 000
    - TERO - Health - 6 500
    - And others

Lessons learned

- A valuable resource
  - Offered freely
- Cultivated for new applications and audiences
  - Extensible design allows for interoperability

New openings: FinnONTO Research project
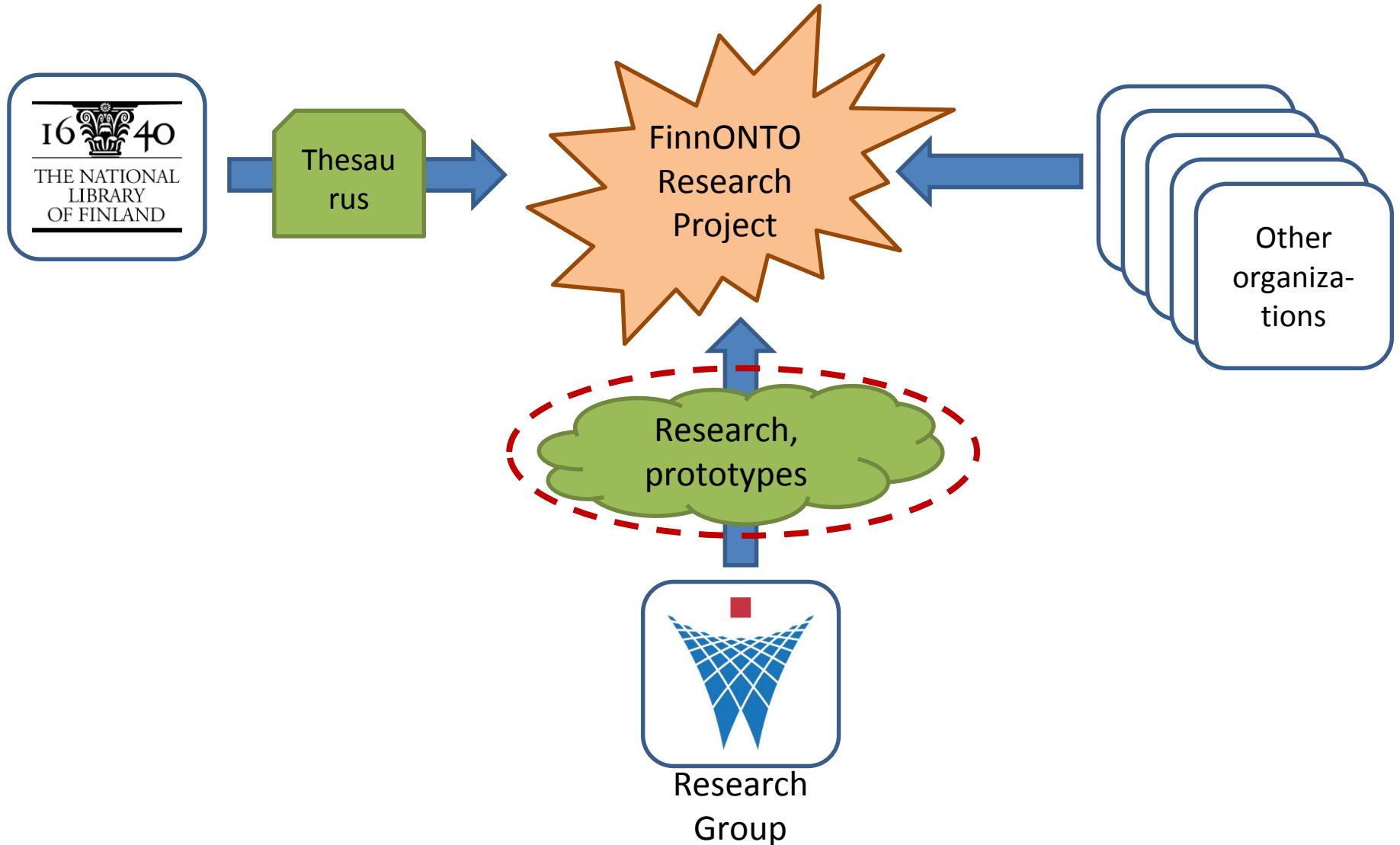
Thesaurus and ontology service

- FinnONTO created ONKI ontology service
  - Several iterations
  - Research prototype
    - New features added when it resulted in interesting research
    - Usability not a high concern
    - Stability not a high concern
  - Proof of concept

Thesaurus and ontology service

- Production-level national service
- National Library jumped at the chance to build it
  - Vocabulary expertise
  - Old thesaurus service
  - Annotation expertise
    - Translates into usability expertise

- Thus, the National thesaurus and ontology service Finto

Finto

- The aim of Finto is to provide a dependable, centralized, national thesaurus and ontology service

- http://finto.fi

- Developed in the National Library of Finland for use in the whole public sector

  - Recruited some of the technical people from the research project and added these to the people with library expertise

- Funded by the Ministry of Finance and the Ministry of Culture and Education

Finto

- The aim of Finto is to provide a dependable, centralized, national thesaurus and ontology service.
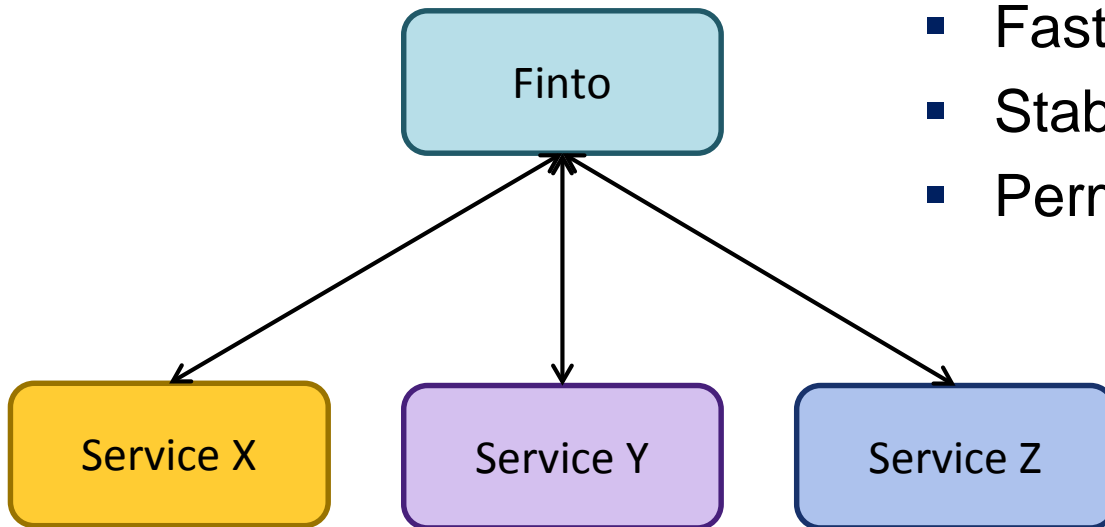
- Fast
- Stable and reliable
- Permanent

Finto

- The aim of Finto is to provide a dependable, centralized, national thesaurus and ontology service.

Finto

O1   O2   O3
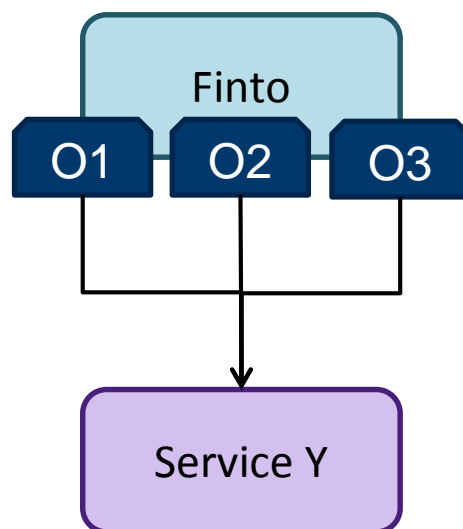
Service Y

- One place to publish
- One place to find
- Common APIs
- Easy to switch between resources

Finto

- The aim of Finto is to provide a dependable, centralized, national thesaurus and ontology service.
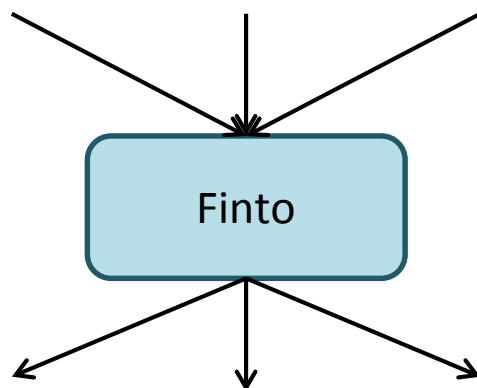
Finto

- Working with all sorts of organizations with differing needs and capabilities
- Standards not shackled to the library world

from all ▾ English ▾ [   ] ✕ **Search**

Welcome to the Finto service!

Finto is a Finnish thesaurus and ontology service, which enables both the publication and browsing of vocabularies. The service also offers interfaces for integrating the thesauri and ontologies into other applications and systems.

Available vocabularies and ontologies

GENERAL
Allärs - Allmän tesaurus på svenska
Finnish Corporate Names
KOKO Ontology
Metatietosanasto
Pondus categories
YSA - General Finnish thesaurus
YSAn ja YSOn käsite-ehdotukset
YSO - General Finnish ontology

SOCIETY
JUHO - Julkishallinnon ontologia
Julkisten palvelujen luokitus
JUPO - Finnish Ontology for Public Administration Services
KEKO - Ontology for Education for Sustainable Development
Korkeakoulujen tutkimustiedonkeruussa käytettävä tieteenalaluokitus
LIITO - Liiketoimintaontologia
MERO - Merenkulkualan ontologia
PUHO - Puolustushallinnon ontologia
TSR ontology

GEOGRAPHY AND GEOINFORMATION

Tweets by @Fintopalvelu

Fintopalvelu @Fintopalvelu

Suomalainen esityskokoonpanosanasto SEKO on julkaistu Fintossa: finto.fi/seko/fi/ #musiikki #soittimet #kotilotorvi #tuulikannel

♡ ↪ 06 Oct

Fintopalvelu @Fintopalvelu

#Skosmos 1.8 on nvt

YSO – General Finnish ontology

Content language **English ▾**

[   ] ✕   **Search**

A-Z	Hierarchy	Groups	New

- events and action
- objects
  - abstract objects
  - physical objects
  - physical whole
  - place
    - areas and regions
    - other places
    - places created by nature
      - **coral reefs**
      - craters
      - natural sites
      - potholes
    - places defined by humans
  - systems
- properties

objects > place > places created by nature > coral reefs

PREFERRED TERM   **coral reefs**

TYPE	General concept
BROADER CONCEPT	places created by nature
RELATED CONCEPTS	lagoons
BELONGS TO GROUP	11 Geography. Cartography. Geodesy. Geology. Palaeontology
	13 Hydrology
IN OTHER LANGUAGES	koralliriutat Finnish
	atollit
	korallrev Swedish
	atoller
URI	http://www.yso.fi/onto/yso/p14886
Download this concept:	RDF/XML TURTLE JSON-LD Last modified 10/7/16

CLOSELY MATCHING CONCEPT	Coral reefs and islands	Library of Congress Subject Headings
	koralliriutat (fi)	YSA - General Finnish thesaurus
	korallrev (sv)	Allärs - Allmän tesaurus på svenska

Images indexed with the term in Finna 6

▲ Image ▲

«

»

Postikorttipainanteinen lähivalokuva värikkään keltasinisest ...

Postikorttipainanteinen ilmakuva Cairnsin kaupungista Austra ...

Postikorttipainanteinen valokuva Australian suuren koralliri ...

Postikorttipainanteinen valokuva Australian Green Islandista ...

Postikorttipainanteinen valokuva Australian Green Islandin r ...

Lessons learned

- Using domain expertise as a basis for service development
    - Recruiting experts when needed
    - Building on a national level: library know-how but for the whole of the public sector

New openings: FinnONTO Research project
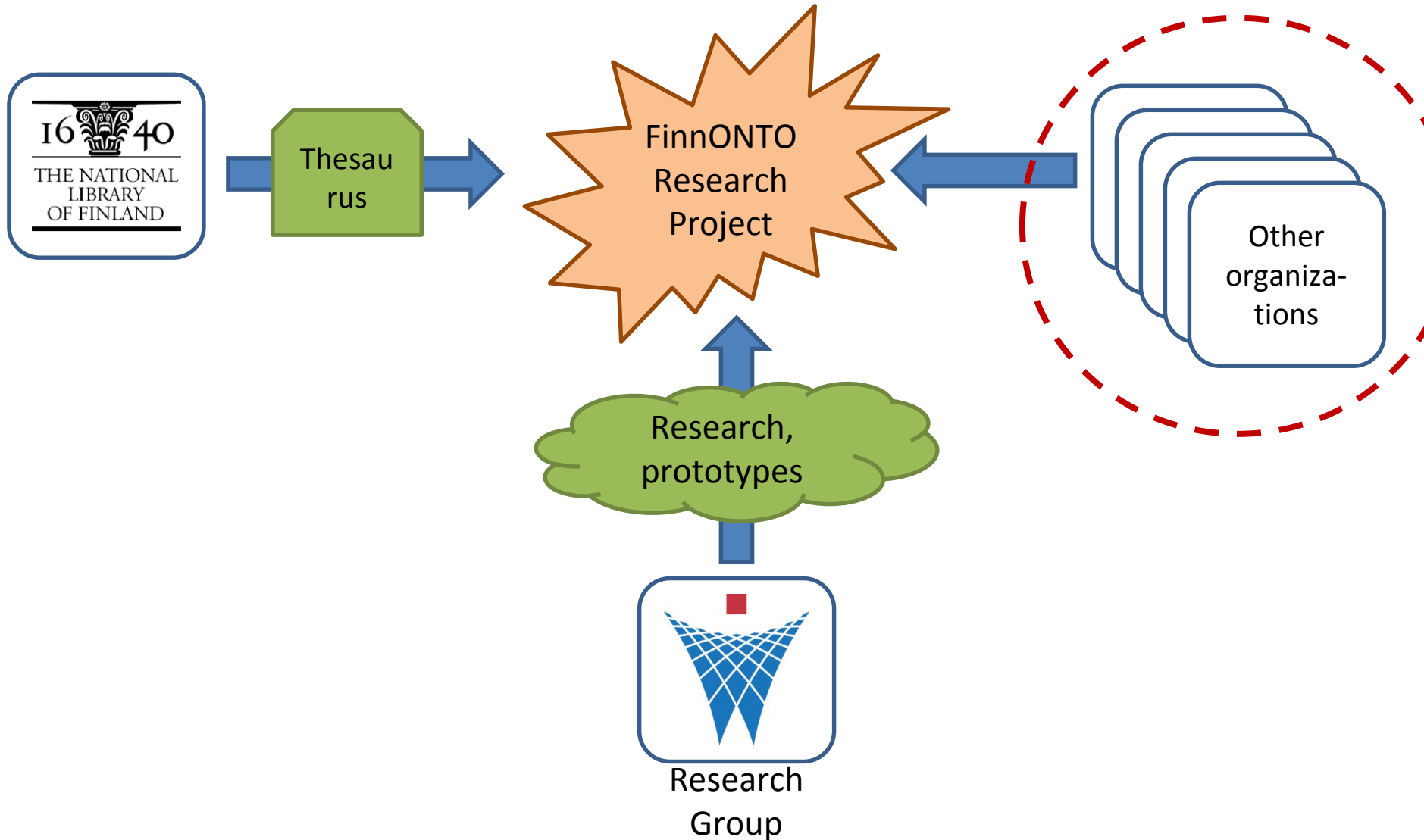
Contacts from the research project

- Formed a co-operation network
    - Users for Finto
    - Publishing vocabularies
        - A number of domain ontologies Extending YSO
    - Using APIs
    - Taking part in usability tests
    - A source of feature requests
    - A varied bunch:
        - Museums, archives, the national broadcasting company, national institute for health and welfare, various ministries and other public sector organizations, etc.

The Law on Interoperability between Information Systems in the Public Sector

- In the summer of 2011

- Everyone should follow the definitions and descriptions of interoperability

  - Through the project, controlled vocabularies and ontologies were included

  - A national vision of interoperability with semantic interoperability through descriptive metadata playing one key role
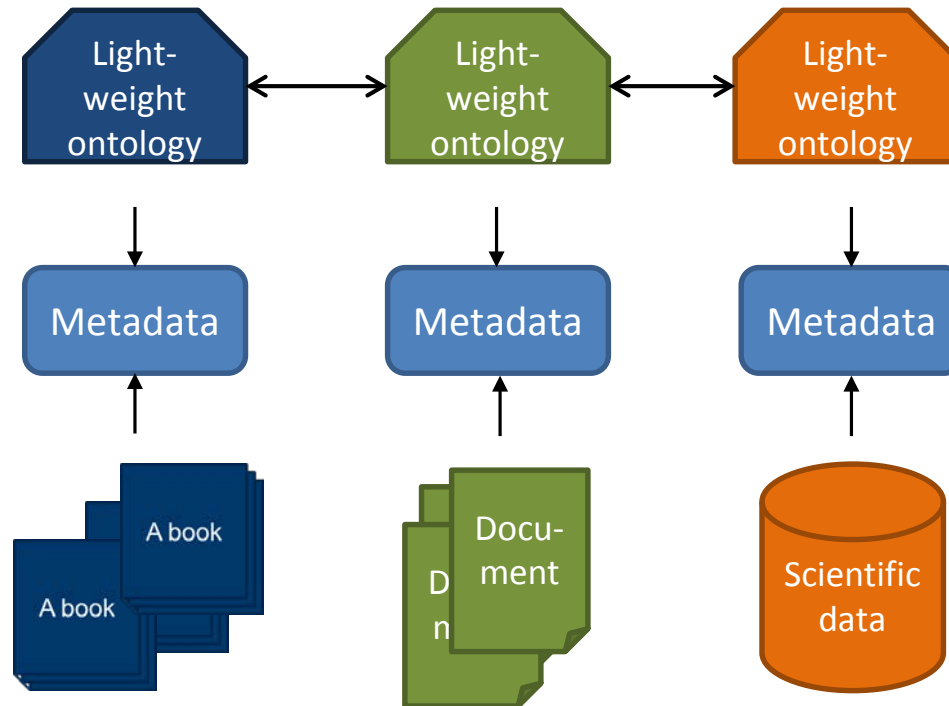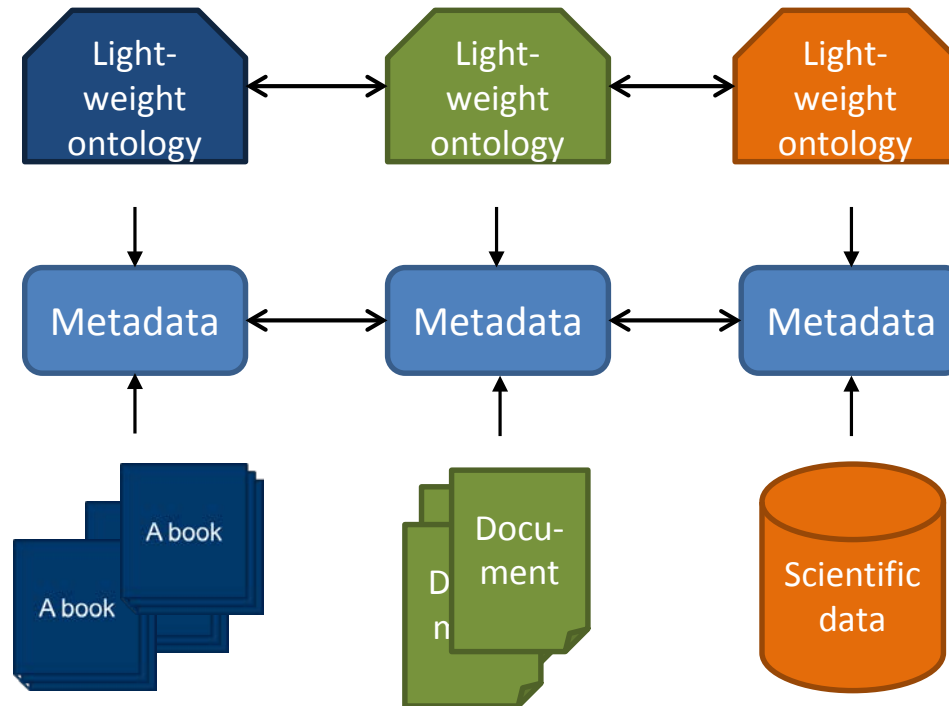
Semantic interoperability through linked ontologies
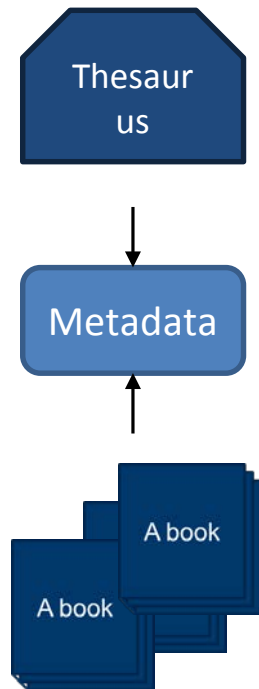
Semantic interoperability through linked ontologies
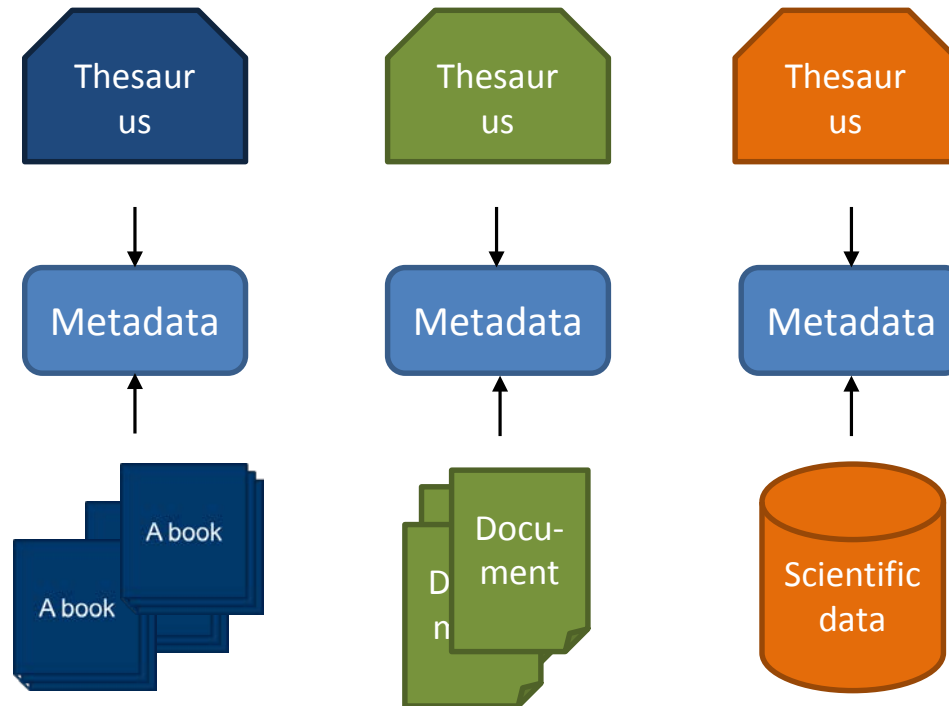
Semantic interoperability through linked ontologies

Semantic interoperability through linked ontologies

Lessons learned

- Networking!
- Several National Library personnel are serving as experts in various groups in different ministries
  - Unique expertise
  - A chance to affect things

Skosmos

Skosmos

- The heart beating inside Finto
- Open source SKOS vocabulary browser
  - http://skosmos.org
- Publication and use of light-weight ontologies, thesauri and classifications
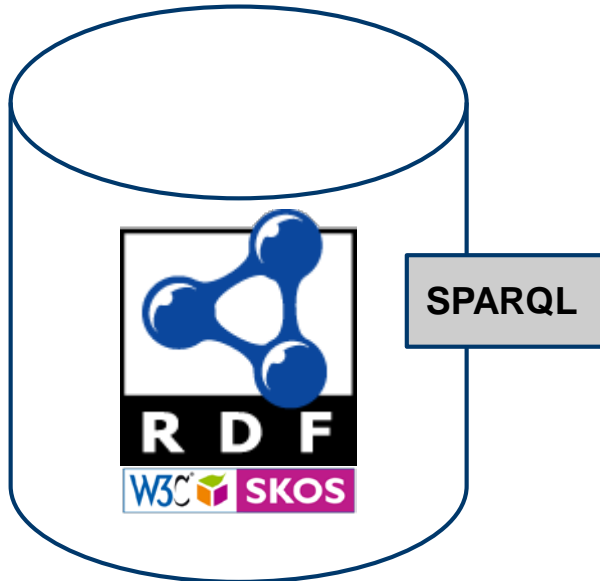 - Web interface
  - REST API
  - SPARQL endpoint
- Community
  - https://groups.google.com/forum/#!forum/skosmos-users

How does it work?
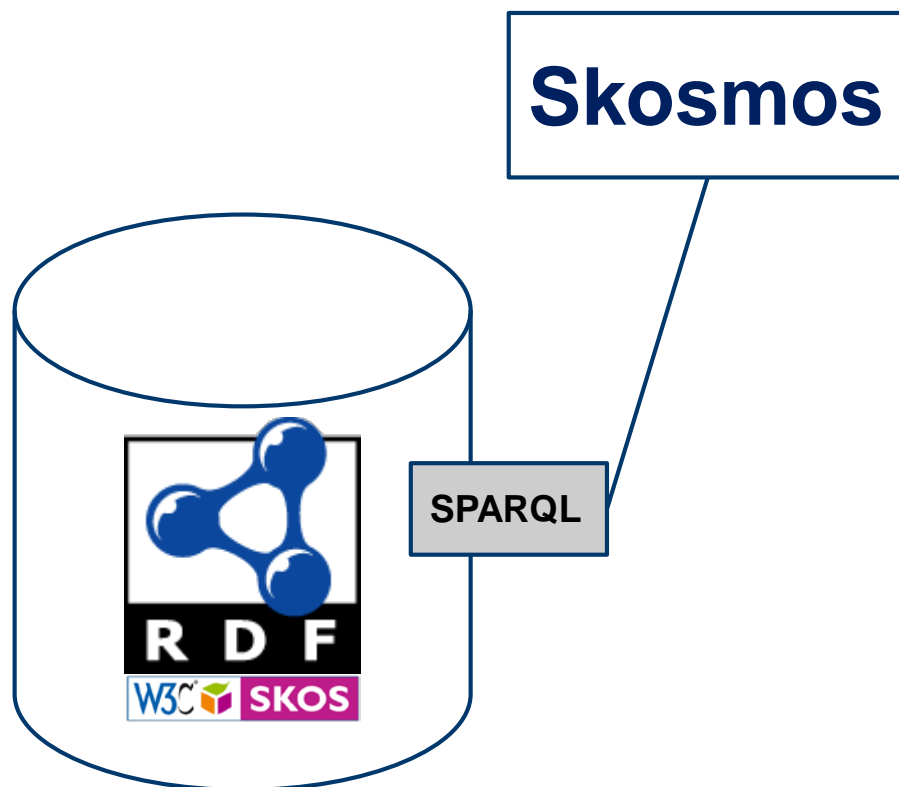
- Make your thesaurus into SKOS

How does it work?

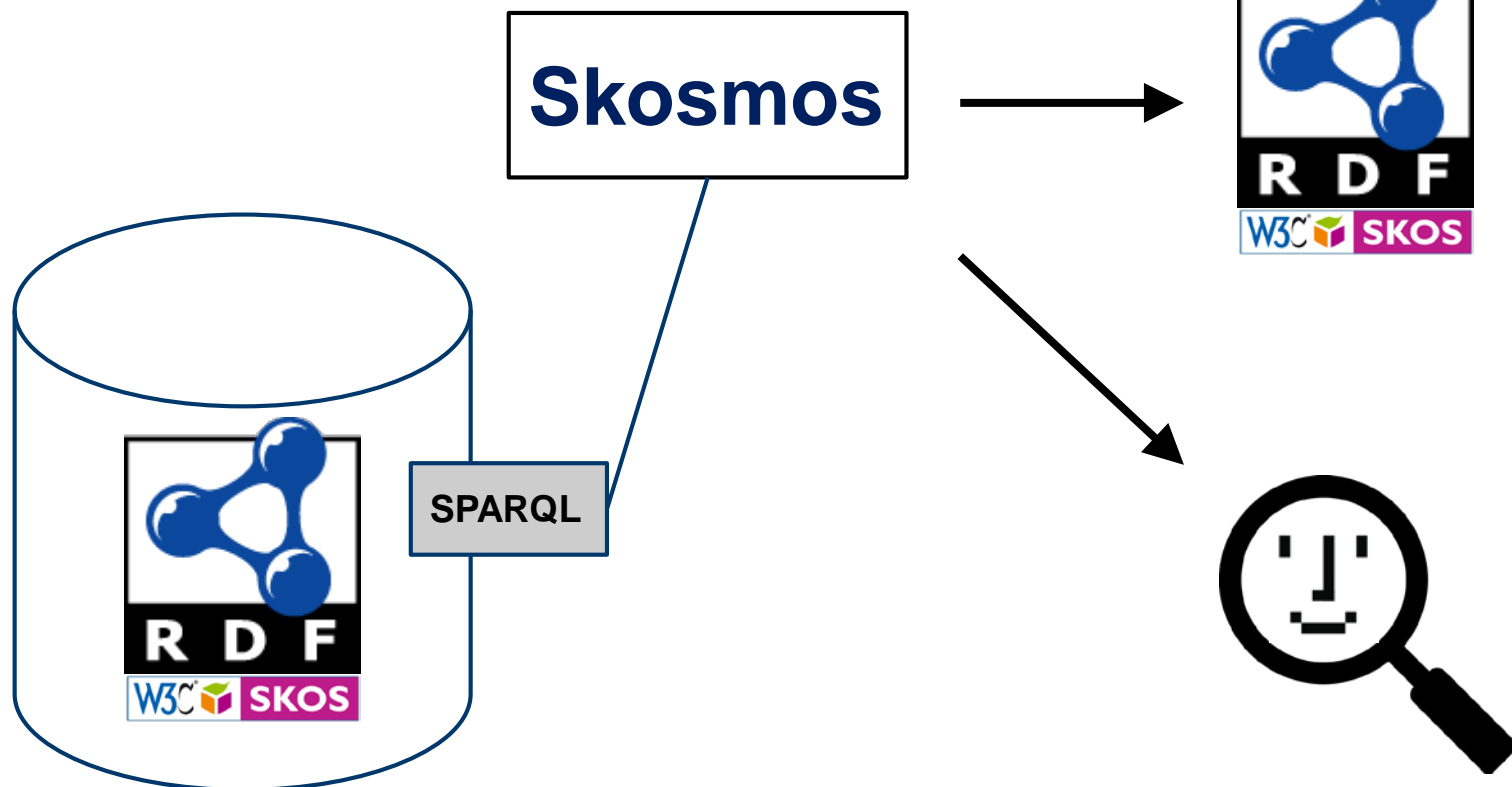- Put in in a SPARQL triple store

How does it work?

- Point Skosmos at your SPARQL endpoint

How does it work?

- And serve your thesaurus for humans, Linked Data agents, and REST API access

Skosmos

SPARQL

Key features

- Multilingual browser interface (10 languages)
  - Autocomplete search
  - Alphabetical index
  - Concept hierarchy display
  - Concept groups (thematic index)
  - New concepts
- REST API for enabling use of vocabularies in other applications
  - responses usually JSON-LD

Skosmos around the world

UN	FAO: Skosmos used for publishing AGROVOC
	GACS project (FAO, CABI (UK), NAL (USA)) is using Skosmos
	UNESCO is using Skosmos for publishing its thesaurus
Germany	Rheinland-Pfalz spatial data infrastructure: publishing classifications
	German National Library of Economics (ZBW)
	North Rhine-Westphalia university library centre (HBZ-NRW): publishing classifications
France	Institute for Demographic Studies INED: publishing Demovoc thesaurus
	Sparna (helping UNESCO and INED in deploying Skosmos)
Italy	University of Tor Vergata (developers of VocBench thesaurus editor)
Austria	University of Vienna, Phaidra project
Norway	Oslo University Library Norwegian General Thesaurus project
Finland	Terminology Centre TSK
	Lingsoft Oy
	SeCo research group light.onki.fi
	Mikkeli University of Applied Sciences

AGROVOC browser

http://aims.fao.org/standards/agrovoc/functionalities/search

Lessons learned

- General parts of the software developed as open source
  - Garners interest
  - Global interoperability!

Summary

- As in ice hockey: go not where the puck is but where it is going to be

- Boldly market your expertise
    - Do not be afraid to apply that expertise into new domains
    - From a librarian to datarian

- Open-mindedly network widely

Thank you! Questions?

matias.frosterus@helsinki.fi

finto-posti@helsinki.fi

http://finto.fi

http://skosmos.org