

Integrating Semi-formal Knowledge Organization Structures

5th ECDL 2006 NKOS workshop
Alicante, Spain, 21SEP2006

Gilles Falquet
Luka Nerima
Claire-Lise Mottaz Jiang
Jean-Claude Ziswiler

Centre universitaire d'informatique
University of Geneva – Switzerland

Outline

Introduction: Building ontology-based Digital Libraries

Context

Hyperbook Alignment Process

Hyperbook Alignment Process (findings)

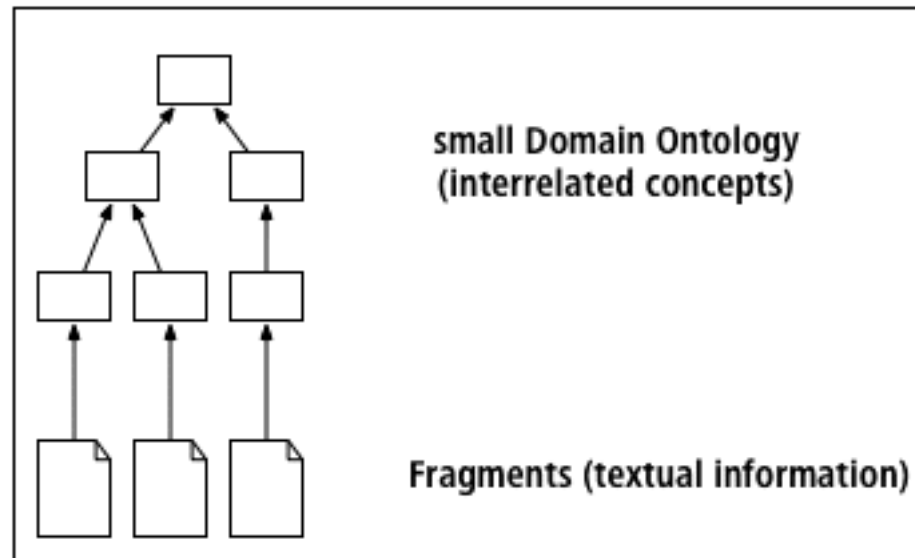
Selecting / validating relations through social networks

Conclusion

Introduction: Building ontology-based Digital Libraries

Hyperbooks

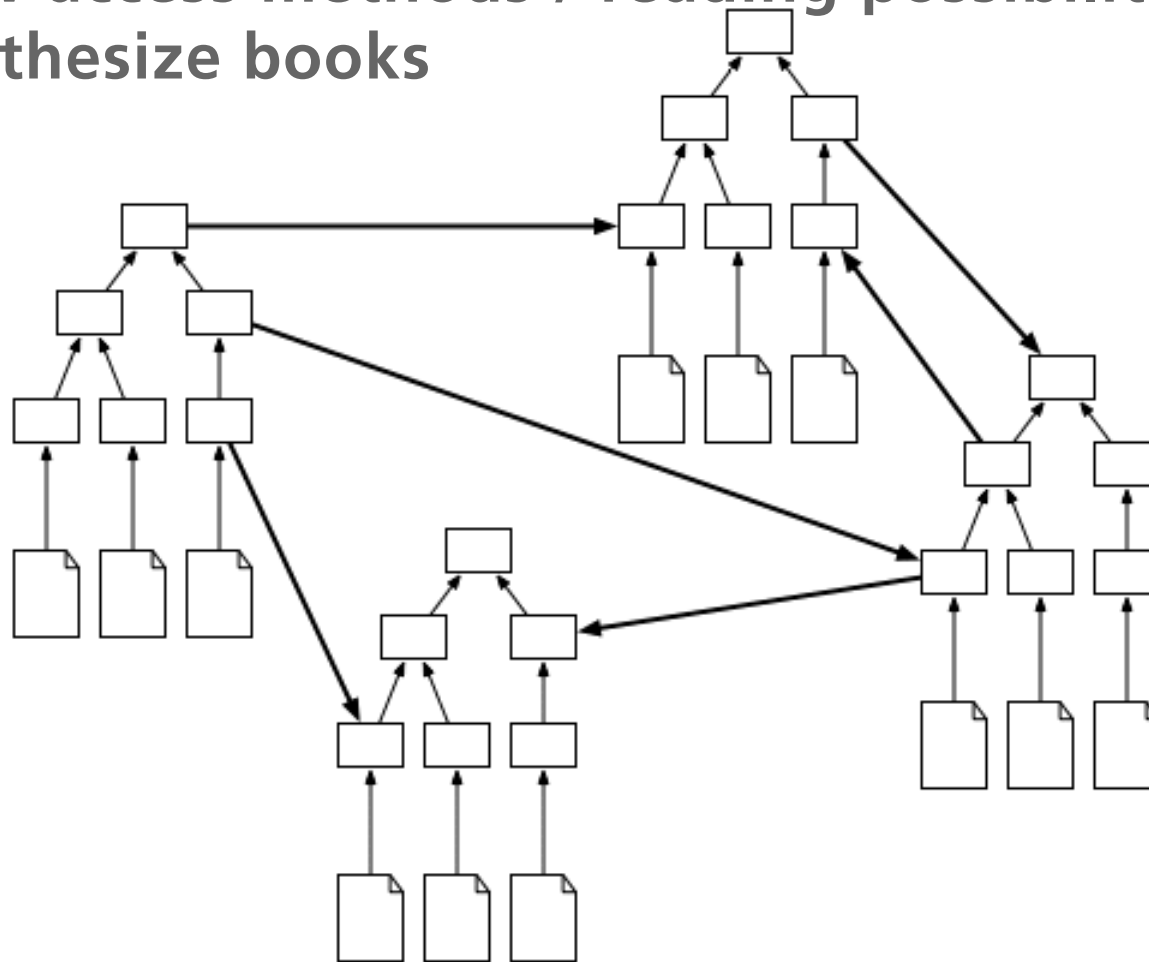
- Break the monolithic aspect of the documents



Introduction: Building ontology-based Digital Libraries

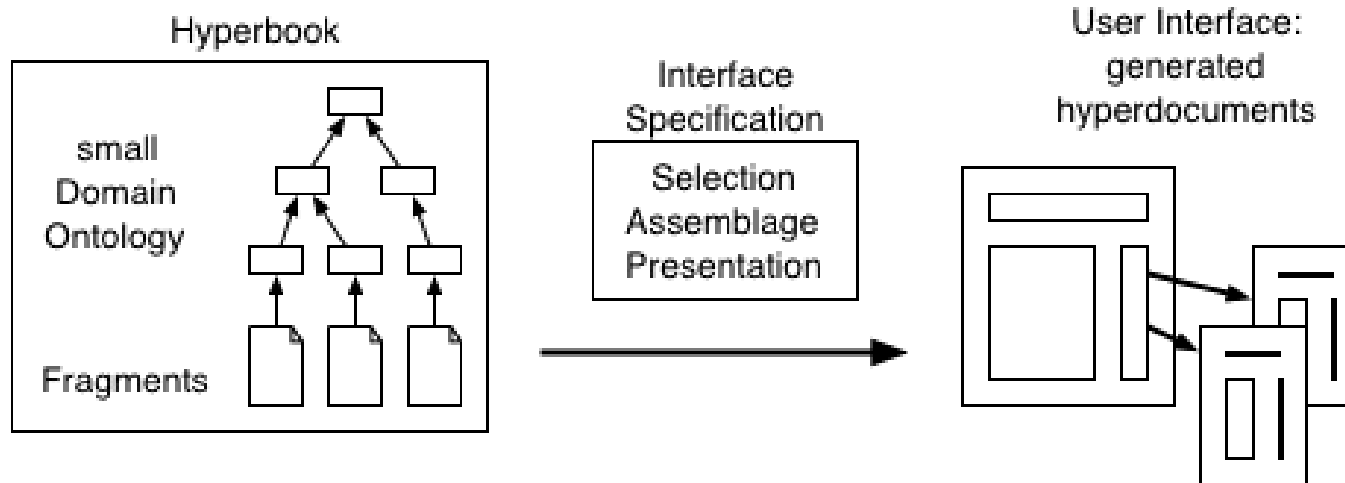
Digital Library built of hyperbooks

- Compare books
- Extend a book with the contents of others
- => New access methods / reading possibilities
- => Synthesize books



Introduction: Building ontology-based Digital Libraries

Generating virtual documents out of the hyperbook structure through an interface specification



[Crampes], [Garlatti], [DeBra], [Brusilowsky], [Falquet]

Introduction: Building ontology-based Digital Libraries

Example: Virtual document of a Digital Library

Node: hb5.linkedfragments[1871,1884]
http://cuisun1.unige.ch:8080/lazy/ns?eip=ZY

Multifunctional Agriculture [concepts] [fragments]

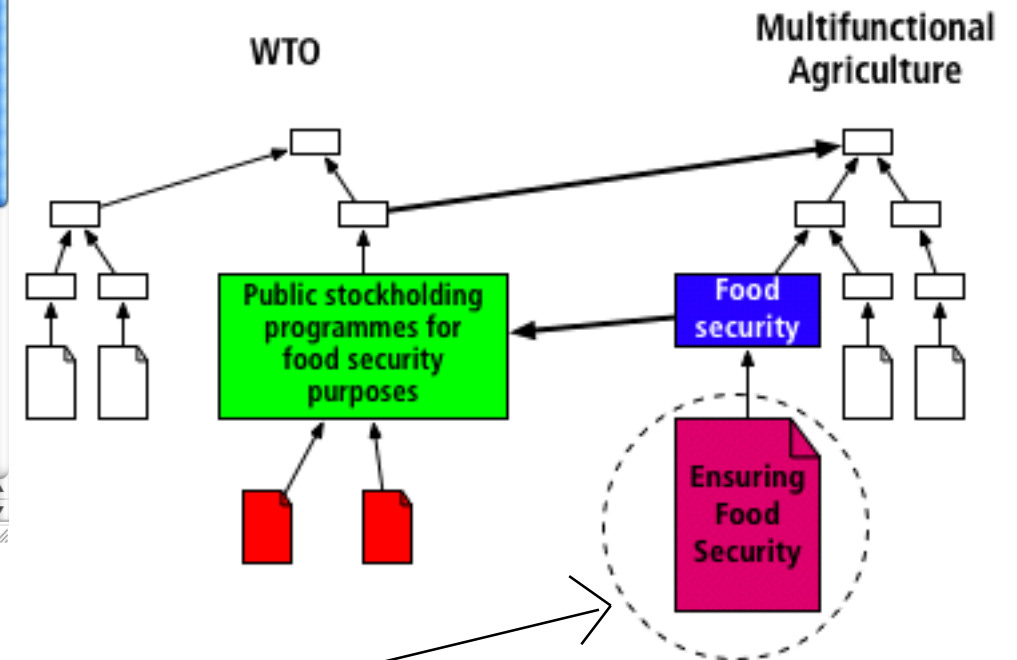
Ensuring Food Security

All countries have to ensure food security for their people, through domestic production, stockholding and imports. Several background elements contribute to such a conclusion. First, market mechanisms alone may not be sufficient to ensure food security. Food security therefore has the characteristics of a public good the provision of which may require government intervention. Second, the WTO agricultural reform process must be consistent with other relevant multilateral commitments such as those relating to the right to food. Article 11 of the International Covenant on Economic, Social and Cultural Rights recognises the fundamental right of everyone to be free from hunger, emphasises the responsibility of the state in this respect and underlines the necessity of taking into account the problems of both food-importing and food-exporting countries, to ensure an equitable distribution of world food supplies in relation to need. Third, we need to recognise that agriculture is a complex and longterm activity, which only slowly responds to changes in demand. This has important implications for our national food security strategies. One, since agriculture is a biological production characterised by seasonality and constrained by climatic conditions and the limits that exist for...

Related concepts: food security

Similar concepts in other hyperbooks

- Public stockholding programmes for food security purposes (in WTO)
- Stockholding programmes in Europe
An assessment of public stockholding policies
- Domestic food aid (in WTO) [view fragments]



Context

Wikipedia

encyclopaedia

architecture:

- articles (sub-articles)
- categories

collaboratively created (contains discussions)

content in plain text, difficult to process it automatically

Wikibooks (Wikiversity)

-> reuse of existing content?

Context

Semantic Wikipedia

[Völkel et al., WWW 2006]

encyclopaedia

architecture:

- articles (typed links between articles)
- ontology (RDF, RDFS, OWL)

RDF export

=> Where is the wiki in Semantic Wikipedia?

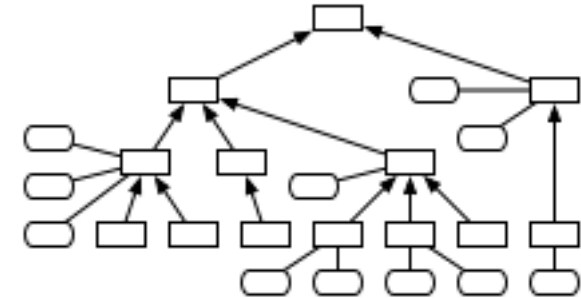
Context



Electronic versions of documents
Book chapters
Journal articles
Web pages



Weakly structured KOS
Glossaries
Directories
Indexes, Folksonomy
Metadata annotated models (Learning Objects)



Highly structured KOS
Ontologies
Thesauri
Concept maps
Taxonomies

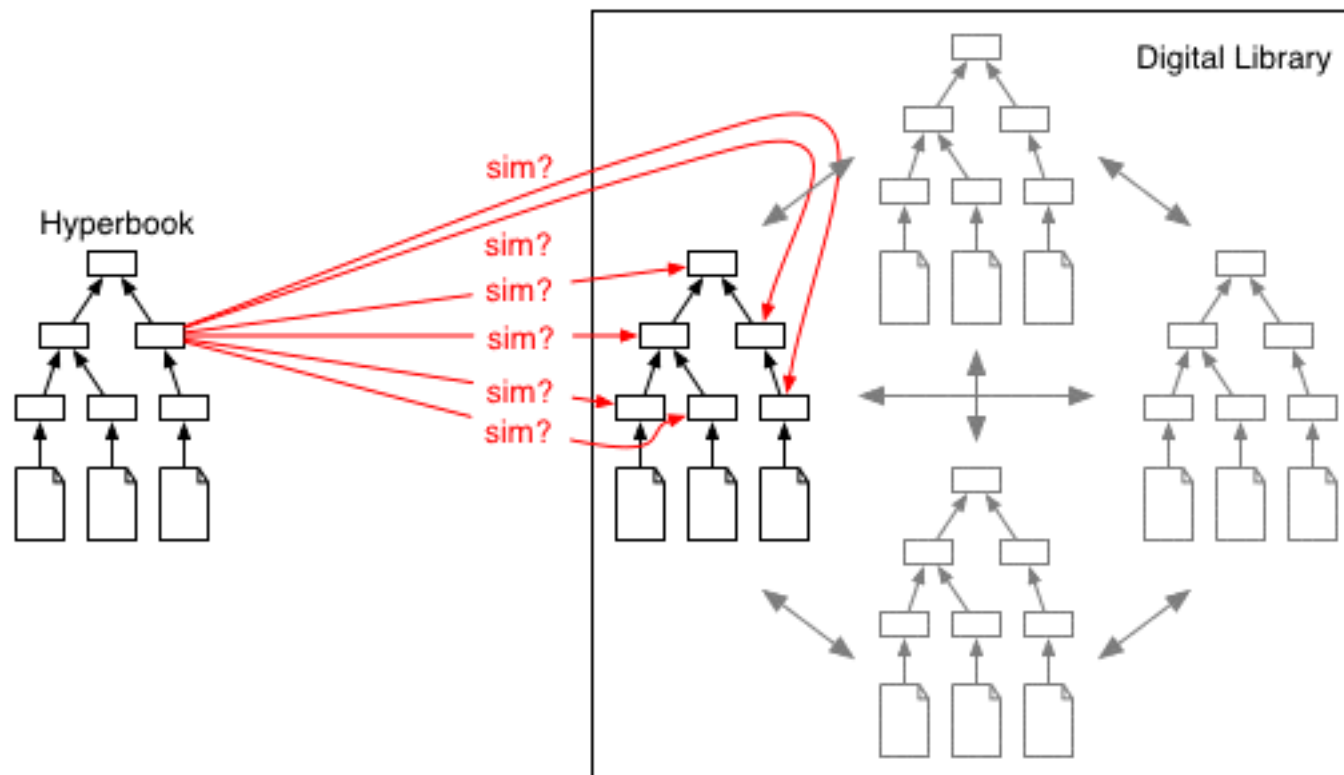
- + availability
- no or weak semantic structure
- **not or hardly machine-processable**

- + strong semantic structure
- not available
- **difficult to construct**
- + **machine-processable**

Hyperbook alignment process

Integrating the ontologies of the hyperbooks

Compute **semantic similarity** between concepts of the hyperbook ontologies



Hyperbook alignment process

Alignment method

Alignment method that works with less formalized (incomplete) ontologies, adapted from similarity function of [Rodríguez&Egenhofer03] by **involving fragments**:

$$a * WM + b * FragM + c * NeighM(r) > t$$

- **Word Matching** (terms of the concepts, often 0.00)
- **Fragment Matching** (terms of the fragments, often > 0)
- **Semantic Neighborhood Matching** (terms of the concepts and fragments in the neighborhood)

Experiment with students that wrote hyperbooks about the course topics (within the hyperbook system that allows creating fragments / concepts / links)

Hyperbook alignment process (findings)

Relations that indicate **equal** concepts

-> depends on the setting (variables a,b,c,r / threshold t)

Setting S1: 19 relations were found

-> 12 of 12 equal relations (precision 63%, recall 100%)

Setting S2: 11 relations were found

-> 9 of 12 equal relations (precision 82%, recall 75%)

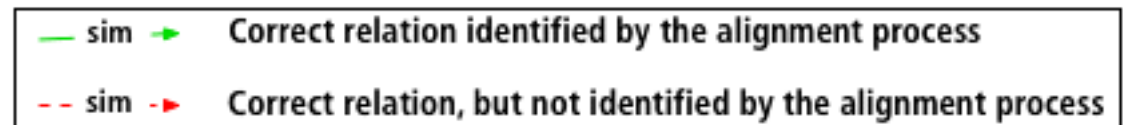
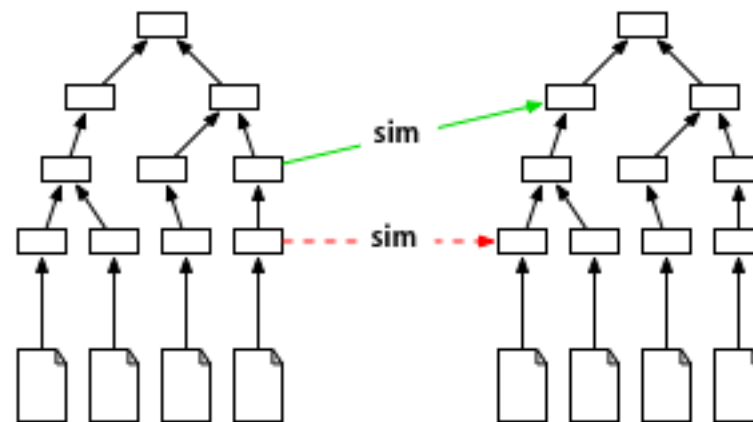
Relations that indicate **similar** concepts

-> with S1 and S2, precision around 50%, very low recall

Hyperbook alignment process (findings)

Strategy

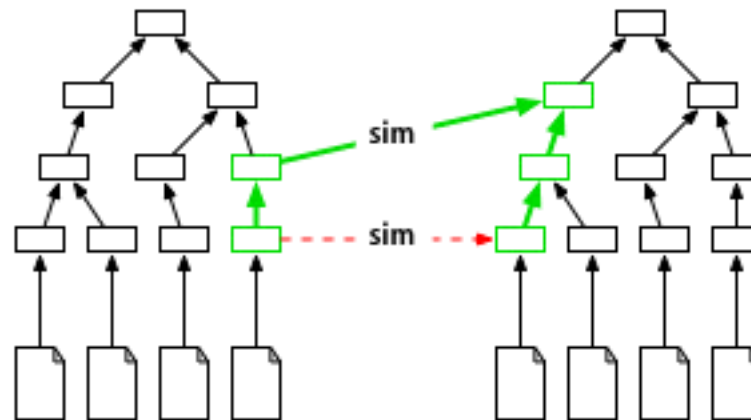
- > focus on high precision concerning relations indicates equal concepts?
- => few, but good quality relations



Hyperbook alignment process (findings)

Strategy

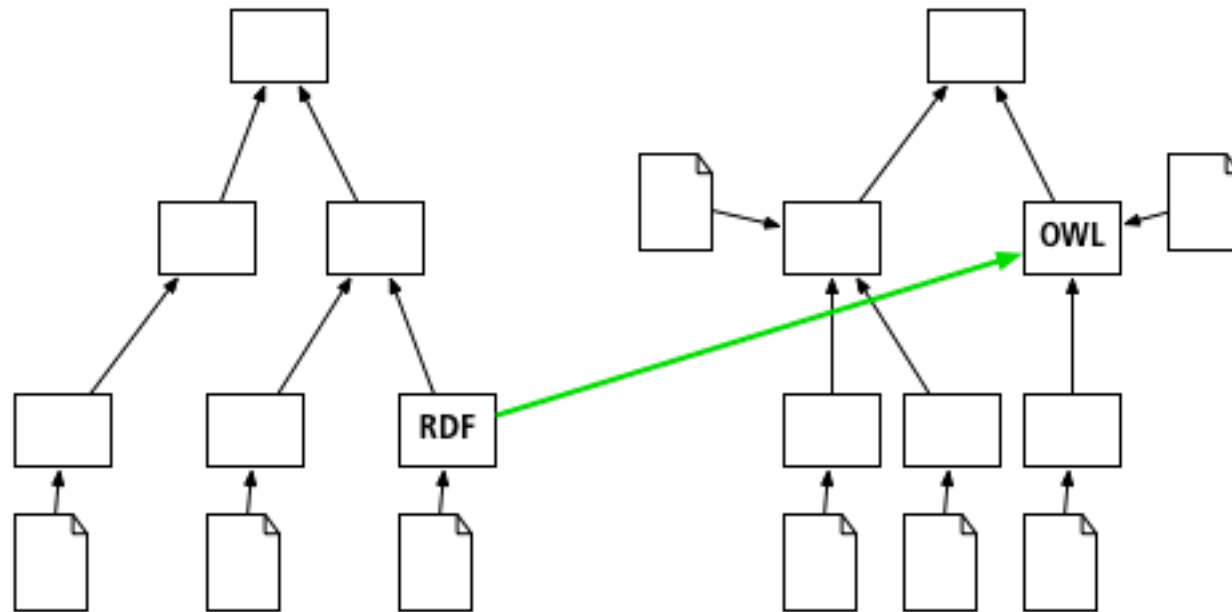
- > focus on high precision concerning relations indicates equal concepts?
- => few, but good quality relations



- => alternative access path (different from IR)

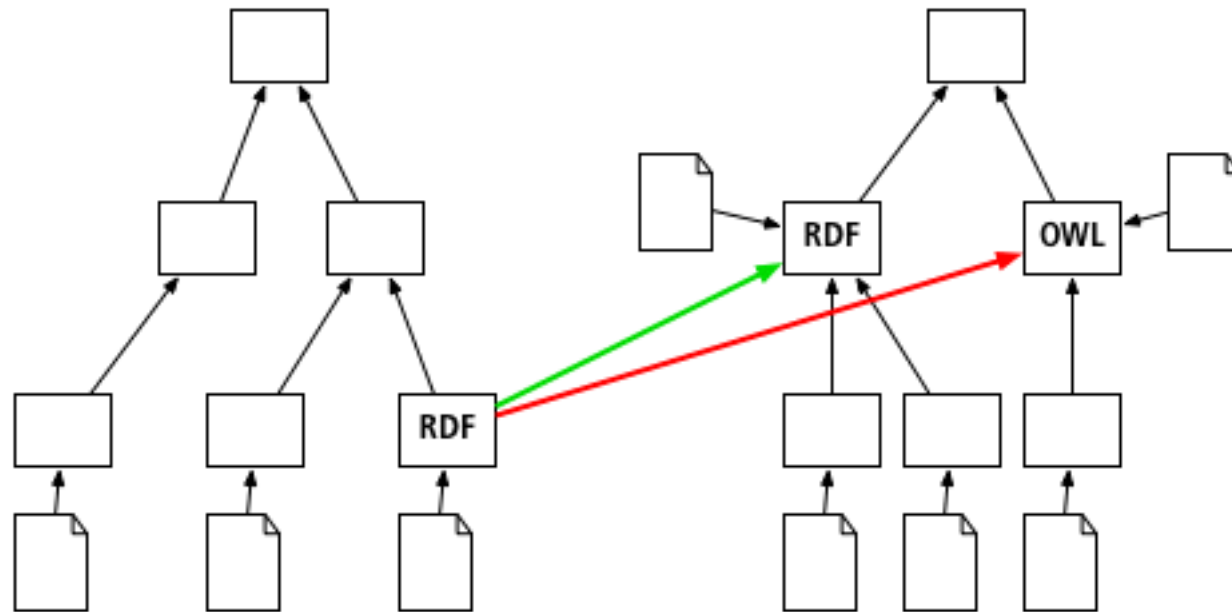
Hyperbook alignment process (findings)

Strategy



Hyperbook alignment process (findings)

Strategy



Selecting / validating relations through social networks

Starting point

- > All links automatically generated through the alignment process are presented

3 possible interactions of the user:

- > Follow these links
- > Definitely accept a link in the user's view
- > Definitely reject a link in the user's view

=> not only a static validation of links (by statistical data)

Selecting / validating relations through social networks

Processes

- => Following or definitely accepting a link **reinforces the quality** of it (the quality of a link is measured by the similarity value calculated in the alignment process).
 - Local effect (one user): validation bottom disappears
 - Global effect (all users): rises the link quality
- => An explicit acceptance of a link by a user is stronger than the simple click on a link

- => Rejecting a link:
 - Local effect (one user): link disappears
 - Global effect (all users): decreases the link quality

- > not limited to generated similarity links, applicable also to manually created links

Conclusion

- > **Hyperbook: small domain ontology and fragments**
- > **Building Digital Libraries by aligning hyperbook ontologies**
- => **Alignment process to find equal concepts works also with small ontologies if fragments and the semantic neighborhood are included**
- => **Setting of variables / thresholds (= [Villa et al., 2004])**
- => **Instead of finding the right strategy for the alignment process settings, we use Social Networks to select / validate relations (!= [Villa et al., 2004])**

Conclusion

Explicit vs. implicit representation of semantics

- > Explicit representations are costly to produce
- > Re-use of existing sources difficult, often not appropriated

- => Explicit representations in hyperbooks are useful
 - > weakly structured KOS (not expensive to create)
 - > allows (automatic) interlinkage of hyperbooks
 - > pedagogical skills
 - > Internal re-use of resources is possible (Wikipedia -> Wikibooks)

- => Next step: Integration of the hyperbook system into a (semantic) Mediawiki

Integrating Semi-formal Knowledge Organization Structures

Thank you!

**Gilles Falquet
Luka Nerima
Claire-Lise Mottaz Jiang
Jean-Claude Ziswiler**

**Centre universitaire d'informatique
University of Geneva – Switzerland**