



Interoperability for Repository Federations

November 21st, 2007, Graz, Austria


Fridolin Wild
Vienna University of Economics and Business Administration



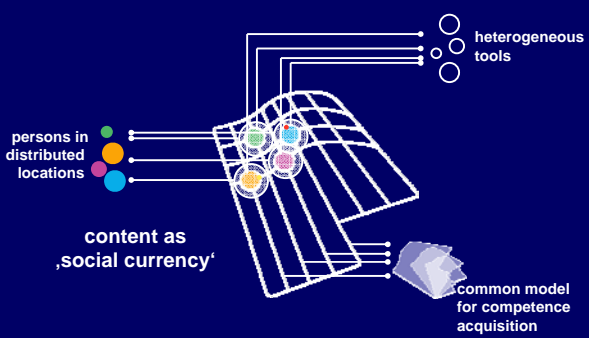


Agenda

- The iCamp Project
- Interoperability
- Retrieval in Repository Federations (ObjectSpot)
- (Preliminary) Retrieval Evaluation
- Future Work
- Outlook: Networking Feed-Based Microcontent



The iCamp Project



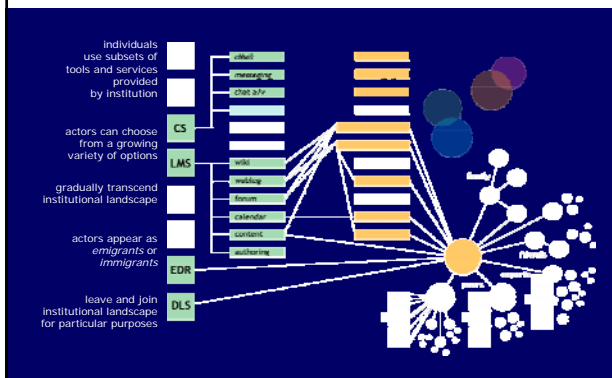
persons in distributed locations

heterogeneous tools

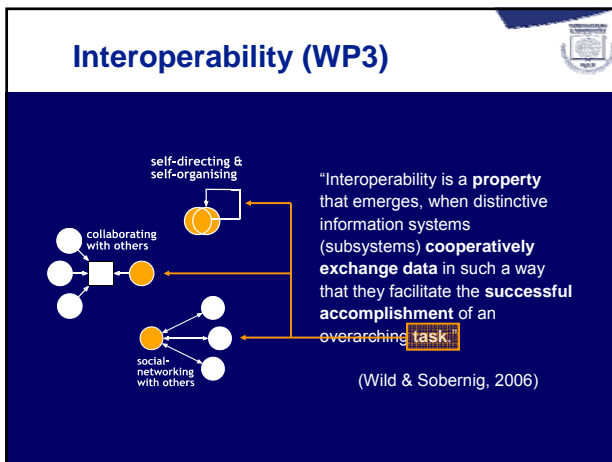
content as ,social currency'

common model for competence acquisition

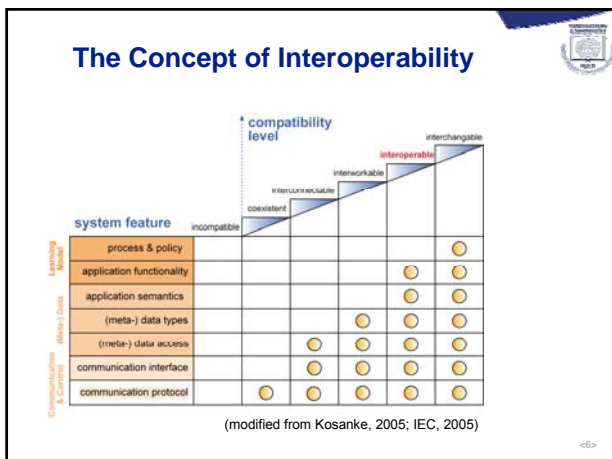
Augmented Landscapes



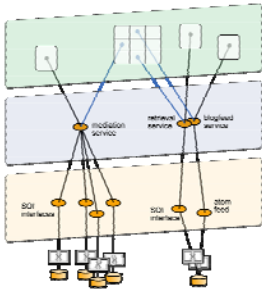
Interoperability (WP3)



The Concept of Interoperability



Potential Levels of Integration



- Presentation Integration**
 - portlets
 - webapp mash-ups
- Remoting: Service Orientation**
 - Information Retrieval
 - Information Filtering
- Information Integration & Dissemination**
 - Data Integration
 - Data Exchange

(Wild, 2007)

<->

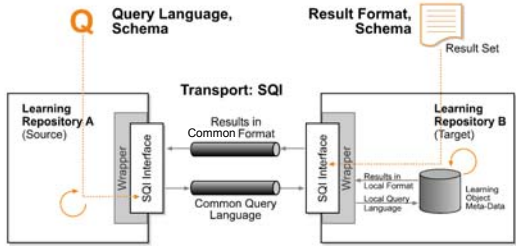
Repository Interoperability

- SQI
- Mediator Service
- The Net: SQI Targets
- Typical Integration Patterns
- Portlet & Portlet Configurator
- Federated Ranking



<->

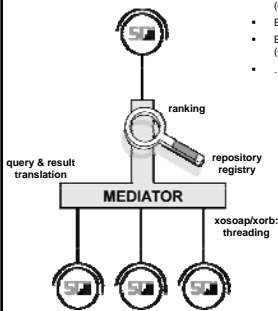
Retrieval via SQL



<10>

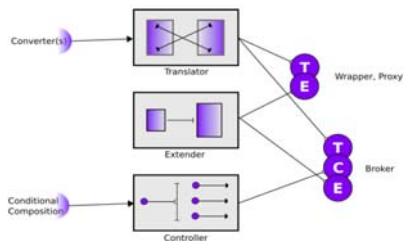
Mediator

- Middleware
- Encapsulates complexity
- Eases consumer access
- Enables performance optimisation (e.g. caching, aggregation, ...)
- Enables query routing
- Basis for an overlay structure (if necessary)
- ...



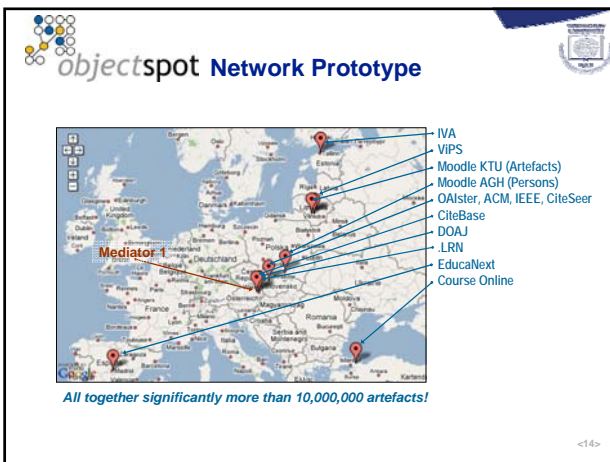
<11>

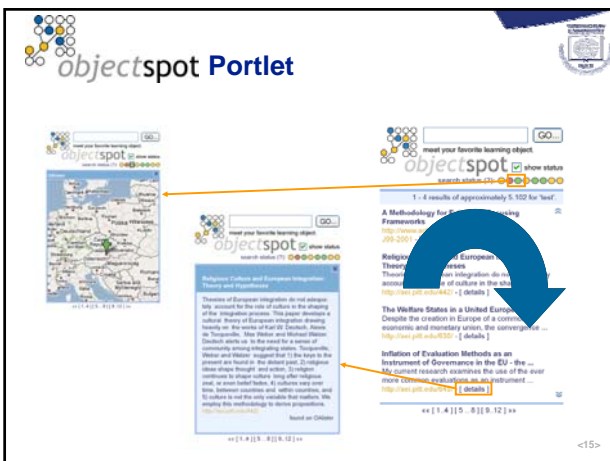
Typical Integration Patterns



<12>







objectspot Configurator

iCamp

Interoperability

The Concept

Framework

Methodology

State of the Field

Theoretical Approaches

Information Integration and Dissemination

Naming

Provenance Integration

Repository Address

Collaborative APIs

Connectors

SQ Consumer Profile

Publications

Interfaces

Configurable search portlet

Enter a search term and click search to see your results!

Search: iCamp

GO

Results: 12/14 of items go

1. iCamp

2. iCamp

3. iCamp

4. iCamp

5. iCamp

6. iCamp

7. iCamp

8. iCamp

9. iCamp

10. iCamp

11. iCamp

12. iCamp

13. iCamp

14. iCamp

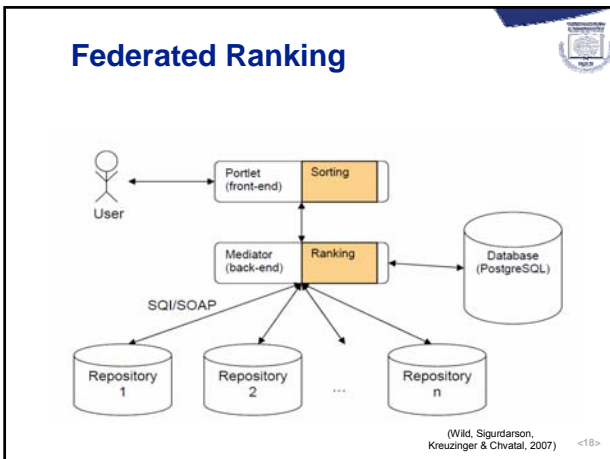
To use the code above, click on the code box, go to the 'Edit' menu of your browser, and choose the 'Select all'.

<16>

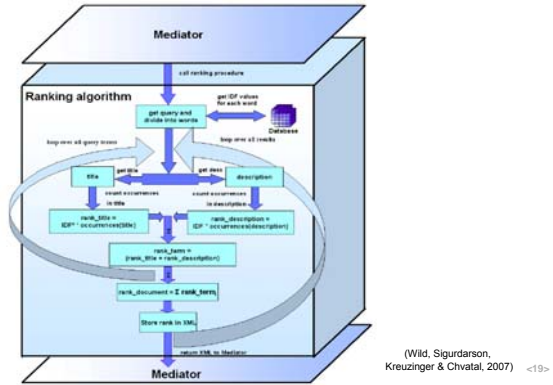
Trend: Even more Scattered Collections

- Survey in (mainly) Central Europe (100 universities) showed:
 - Universities like sharing with selected partners (42%) and publicly for free (29%)
 - But: they don't manage their assets
 - Only 18% have a repository accessible to outside
 - Only 16% have a cross-organisational repository
- Institutional Repositories emerge (LCMS, Video-Lectures, Fides, ...)
- Microformats are popping up (enabling for embedded, even individual repositories)
- Trading platforms emerge and grow (e.g. slideshare, EducaNext, preprint servers, ...)

<17>



Federated Ranking Algorithm



(Preliminary) Evaluation Results

RHL vs. Precision & Recall
Preliminary Results

Ranked Half-Life (Borlund 1998)

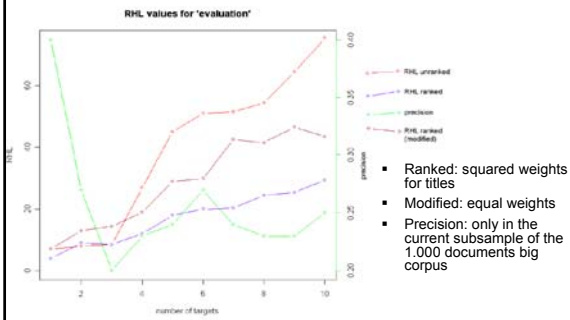
- Median of grouped continuous data
- Result position where half of the relevant documents are found before (lower value = better)
- e. g. RHL value 5: Half of all relevant documents in the answer set within top 5 results.

$$M_g = L_m + \left(\frac{n/2 - \sum f_2}{F(\text{med})} \times CI \right)$$

- L_m = lower real limit of the median class
- n = number of observations
- $\sum f_2$ = cumulative relevance values up to and including the class preceding the median class
- $F(\text{med})$ = relevance value of the median class;
- CI = class interval, commonly in IR = 1.

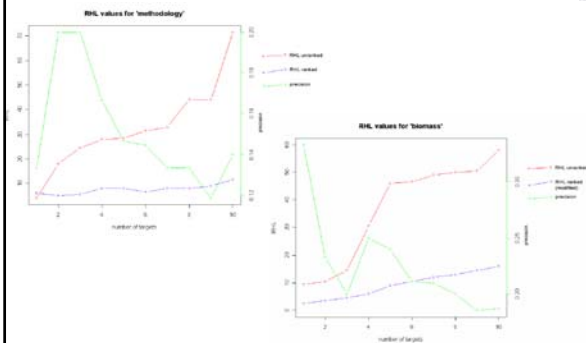
<21>

Results



<22>

Results (2)



<23>

Future Work: Retrieval

Implicit Feedback
 Collaborative Foraging Strategies
 Thorough Evaluation

Future Work

- Investigate implicit feedback means: harvest behaviour traces
 - e.g. local feedback from detail views
 - e.g. track query modifications to adjust weights for disambiguation
- Investigate Collaborative Foraging Strategies
 - e.g. through investigating tagging
 - e.g. through investigating digest sharing
- Repeat Evaluation with Reuters Corpus
 - Simulation of specific situations (bad fast repositories vs. good slow reps; varying 'noise' levels to simulate varying topicality of a repository federation)

<25>

Thank You



www.icamp.eu

<26>

Outlook: Distributed Feed Networks

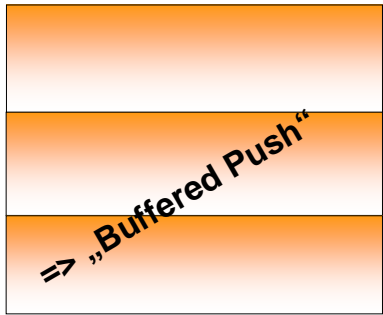
FeedBack Spec
Wordpress Plug-In

Interaction Standards

Feature / API specification	Publish Posting (put, edit, del)	Publish Media Objects	Get Categories	List Postings	Get Postings	Publish Comment	Retrieve Comment	Instant Linkbacks	List 'friends' on same server	Get Blogrolls	Advertise Feeds	Request Update Pings	Instant Indexing (Update Ping)	Extensive Content Model	Simple Content Model
LiveJournal	x	x	x	x											x
Blogger API 1.0	x														x
Blogger GData API	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
MetaWeblog API	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Atom Publishing Protocol	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Trackback Specification															x
Pingback Specification															x
	publishing					referring		passive netwkg.		active networking			content model		

collaboration

The Missing Link: FeedBack



- 1 OFFER
- 2 REQUEST update notifications
- 3 NOTIFY

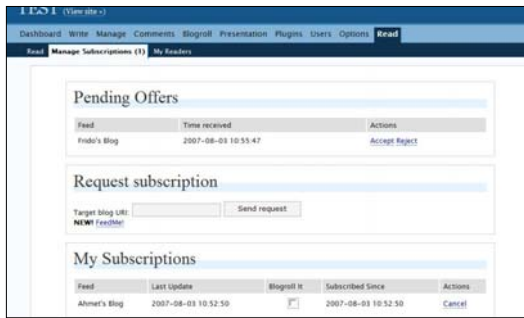
<29>

Example: WordPress (1)

Description	Subscribed Since	# Success	# Failure	Last Success Time	Action
Stein	2007-08-03 10:58:50	1	0	2007-08-03 10:51:20	Cancel

<30>

Example: WordPress (2)

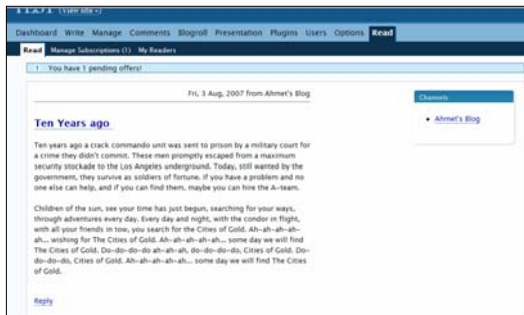


The screenshot shows the 'My Readers' section of the WordPress Admin Dashboard. It features three main components:

- Pending Offers:** A table with columns 'Feed', 'Time received', and 'Actions'. One entry is visible: 'Fidd's Blog' received on '2007-08-01 10:55:47' with an 'Accept/Reject' action.
- Request subscription:** A form with a 'Target blog URL' field (containing 'NEW! Feeder') and a 'Send request' button.
- My Subscriptions:** A table with columns 'Feed', 'Last Update', 'Blogroll ID', 'Subscribed Since', and 'Actions'. One entry is visible: 'Ahmer's Blog' with a last update of '2007-08-03 10:52:50' and a 'Cancel' action.

<31>

Example: WordPress (3)

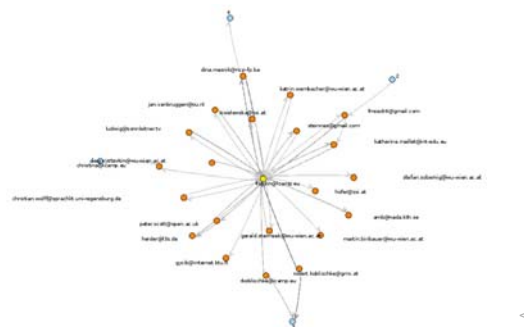


The screenshot shows a 'You have 1 pending offer!' notification. The offer is from 'Ahmer's Blog' dated 'Fri, 3 Aug, 2007'. The offer text reads: 'Ten Years ago. Ten years ago a crack commando unit was sent to prison by a military court for a crime they didn't commit. These men promptly escaped from a maximum security shokkade to the Los Angeles underground. Today, still wanted by the government, they survive as soldiers of fortune. If you have a problem and no one else can help, and if you can find them, maybe you can hire the A-team. Children of the sun, see your time has just begun, searching for your ways, through adventures every day. Every day and night, with the condor in flight, with all your friends in tow, you search for the Cities of Gold. Ah-ah-ah-ah-ah... wishing for The Cities of Gold. Ah-ah-ah-ah-ah... some day we will find The Cities of Gold. Do-do-do-do-ah-ah-ah, do-do-do-do, Cities of Gold. Do-do-do-do, Cities of Gold. Ah-ah-ah-ah-ah... some day we will find The Cities of Gold.'


<32>

Future Work

- Study interaction behaviour with a trial



<33>



#eof.
BEWARE, THE END IS NEAR.
