

# A System for the Evaluation of Interactive Information Retrieval

Norbert Fuhr University of Duisburg-Essen, Germany



#### Talk structure

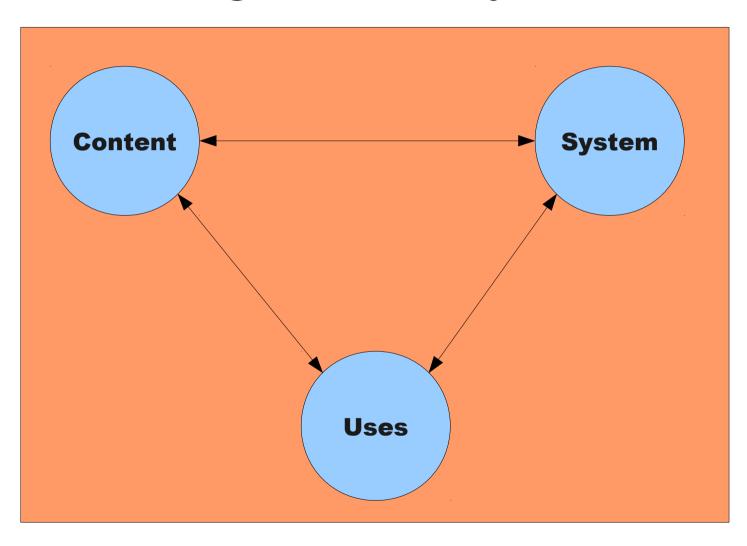
- Digital library systems
- Architecture levels of information retrieval / digital library systems
- Levels of experimentation
- The ezDL system
- ezDL experimentation
- Conclusion



## Digital library systems

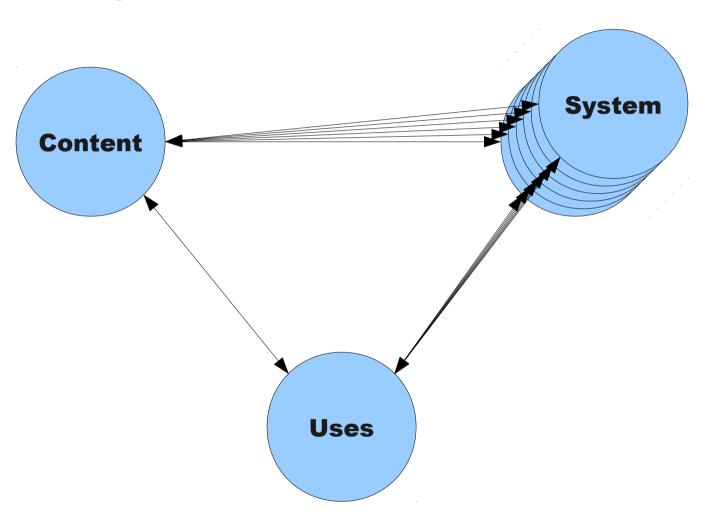


## Digital Library Model





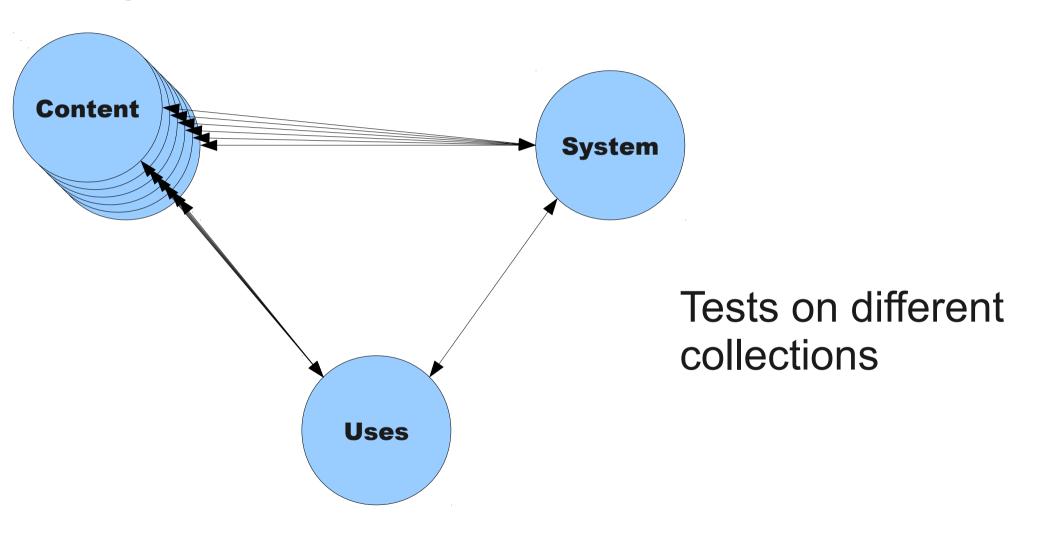
## Experimentation: Variation of Systems



Classic Evaluation campaign

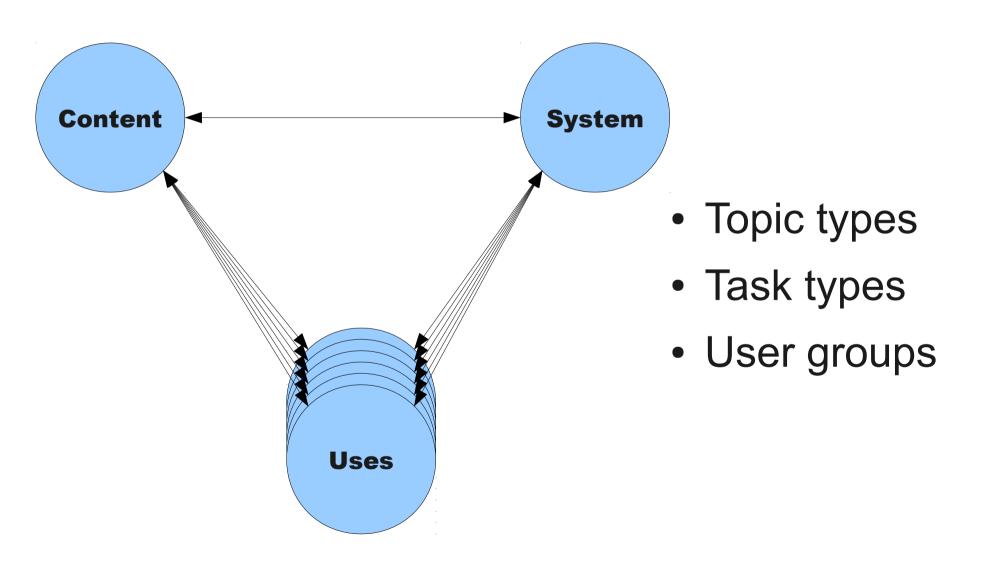


## Experimentation: Variation of Content





## Experimentation: Variation of Uses

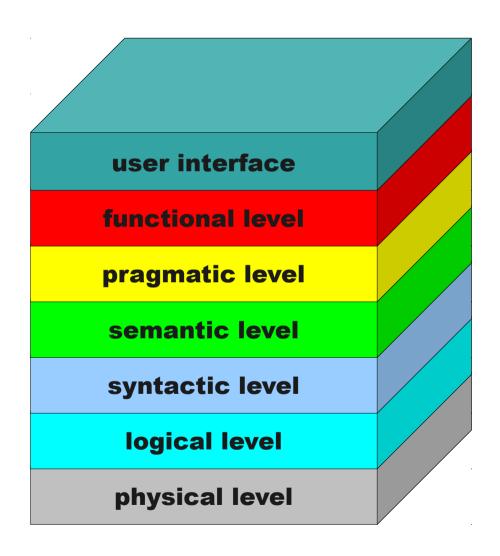




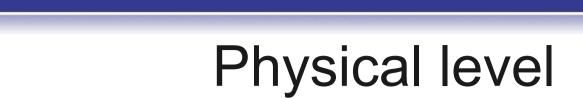
## Architecture levels of information retrieval / digital library systems

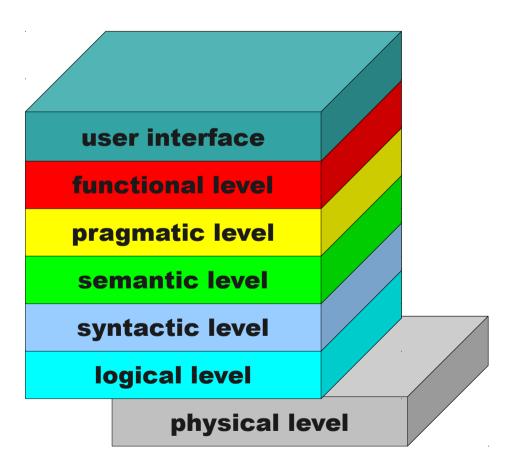


## Architecture Levels of IR Systems

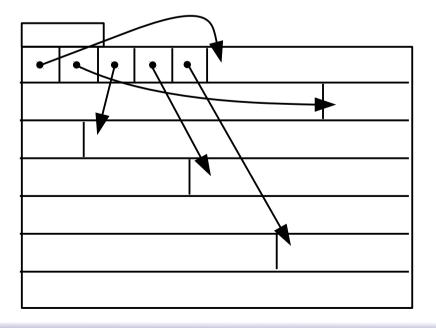








$t_1$	$d_2$	$d_{15}$	$d_{23}$	$d_{89}$	
$t_2$	$d_5$	$d_{15}$	$d_{89}$		







user interface

functional level

pragmatic level

semantic level

syntactic level

**logical level** 

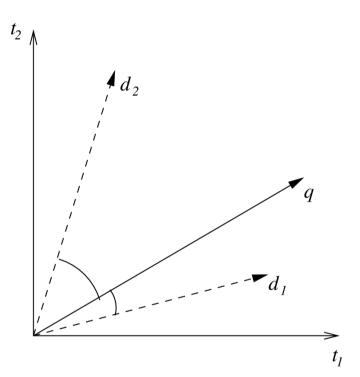
physical level

Vector space

Boolean

Fuzzy

Probabilistic



Language model

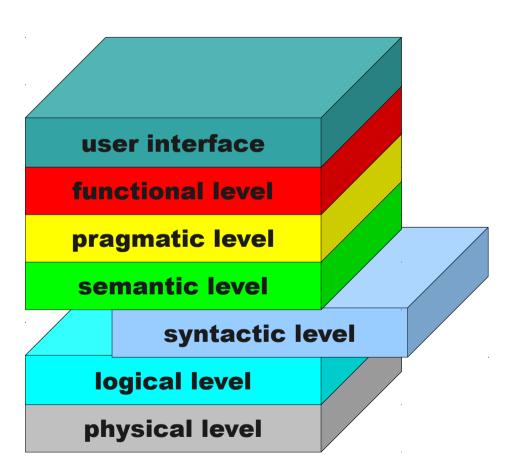
Learning to rank

$$P(R|q,d) \propto \prod \lambda * P(w|d) + (1-\lambda) * P(w|C)$$





## Syntactic level

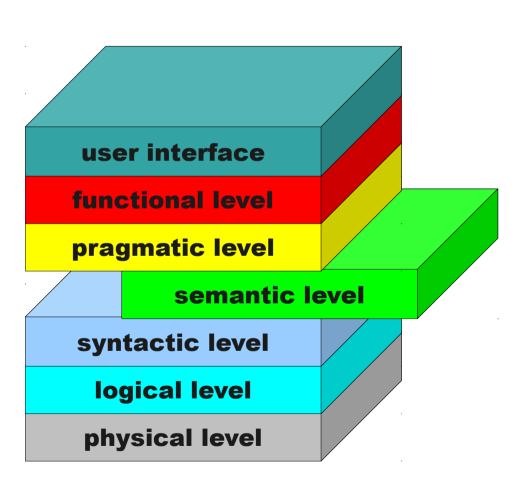


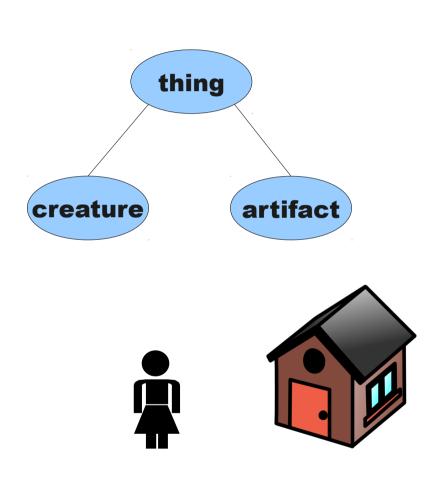
informat.\* retrie[f|v].\*





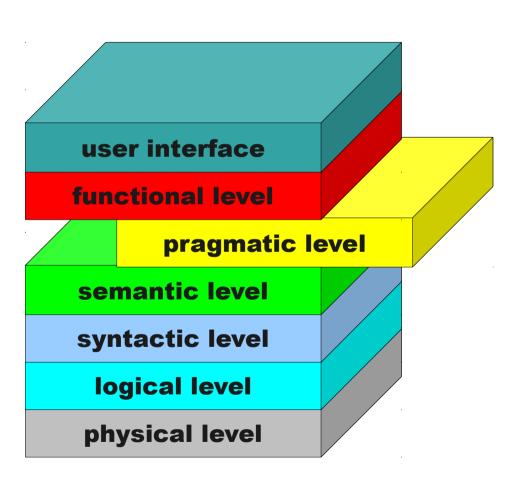














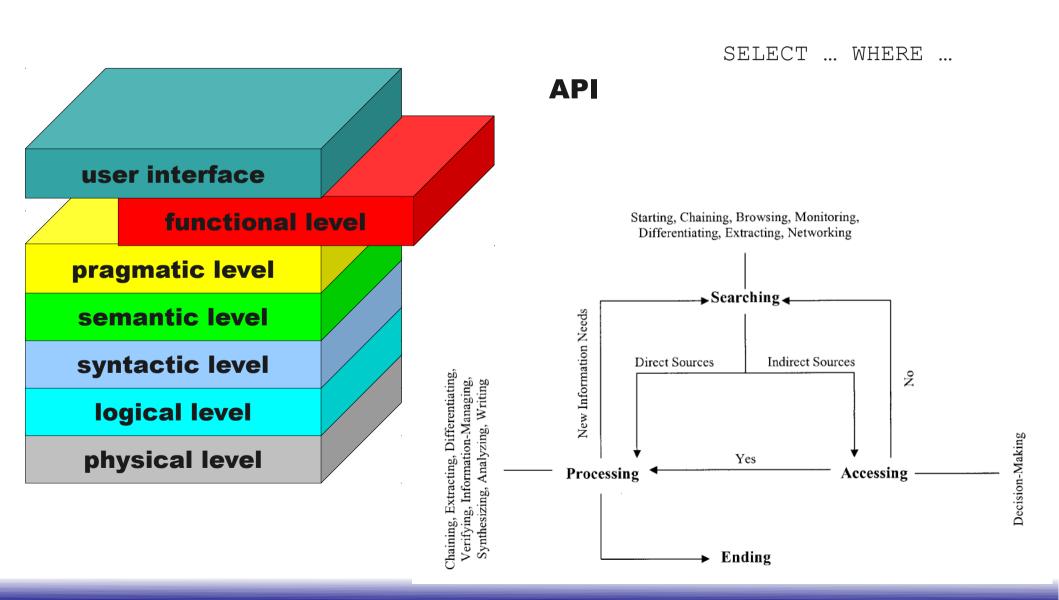
Problem



Work task

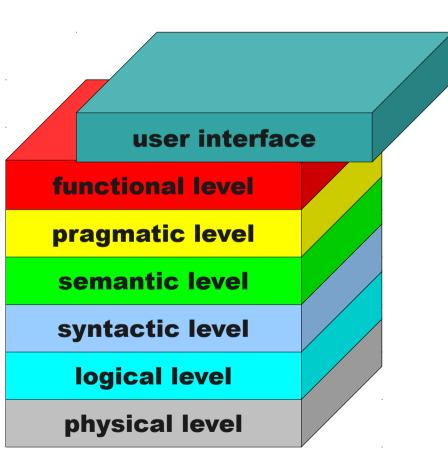


#### **Functional level**





#### User Interface



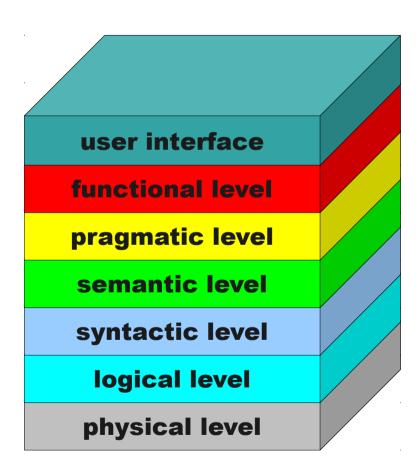




## Levels of experimentation

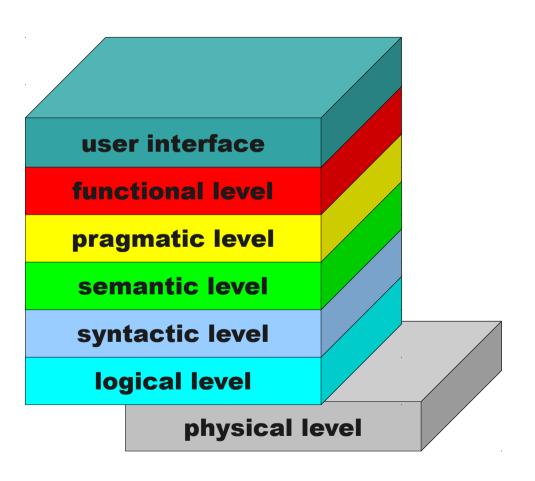








## Levels of Experimentation: Physical

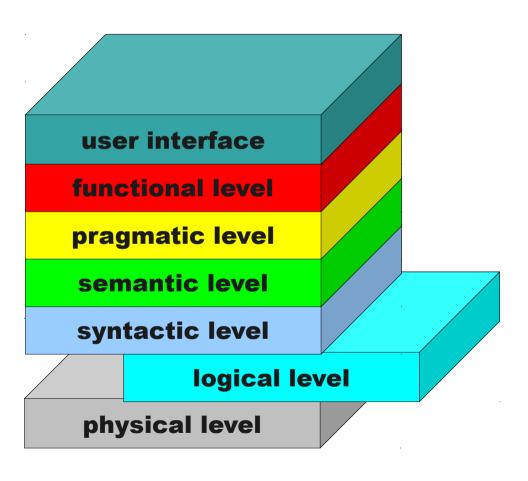


#### Performance experiments:

- Efficiency of data structures
   & algorithms
- Efficiency vs. effectiveness



## Levels of Experimentation: Logical

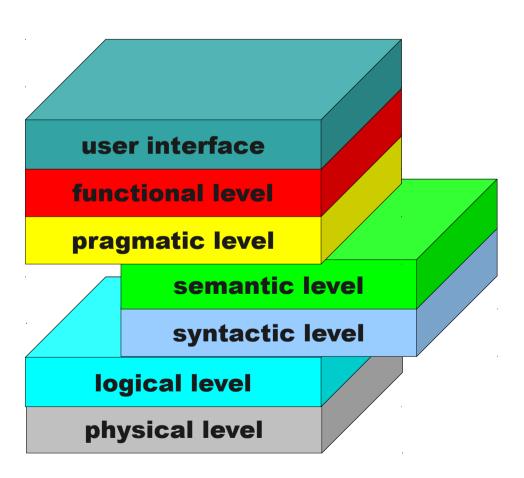


#### Test of IR models

Focus on effectiveness



### Levels of Experimentation: Syntactic/Semantic

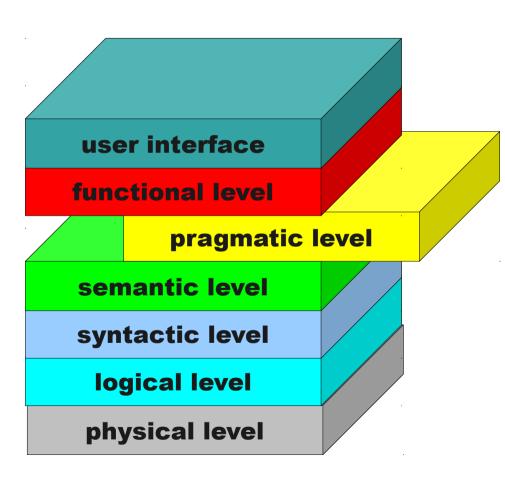


#### Test of representations

- Focus on effectiveness
- Efficiency vs. effectiveness



## Levels of Experimentation: Pragmatic

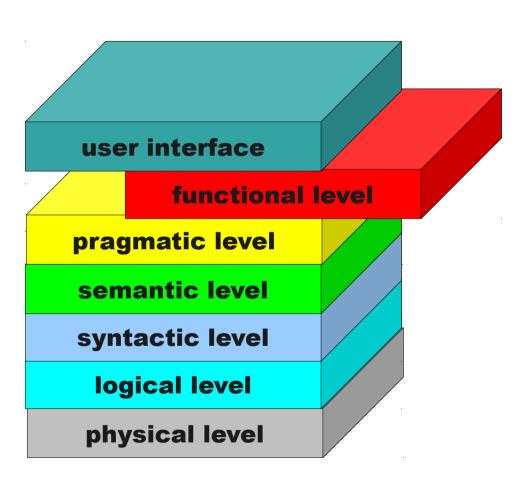


#### Field studies:

- System support for
  - problem solving
  - work tasks



## Levels of Experimentation: Functional

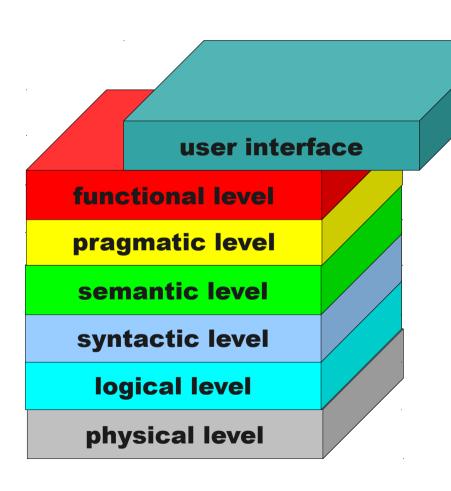


## Test of search/cognitive functions

- user experiments
  - lab
  - living labs



## Levels of Experimentation: User Interface

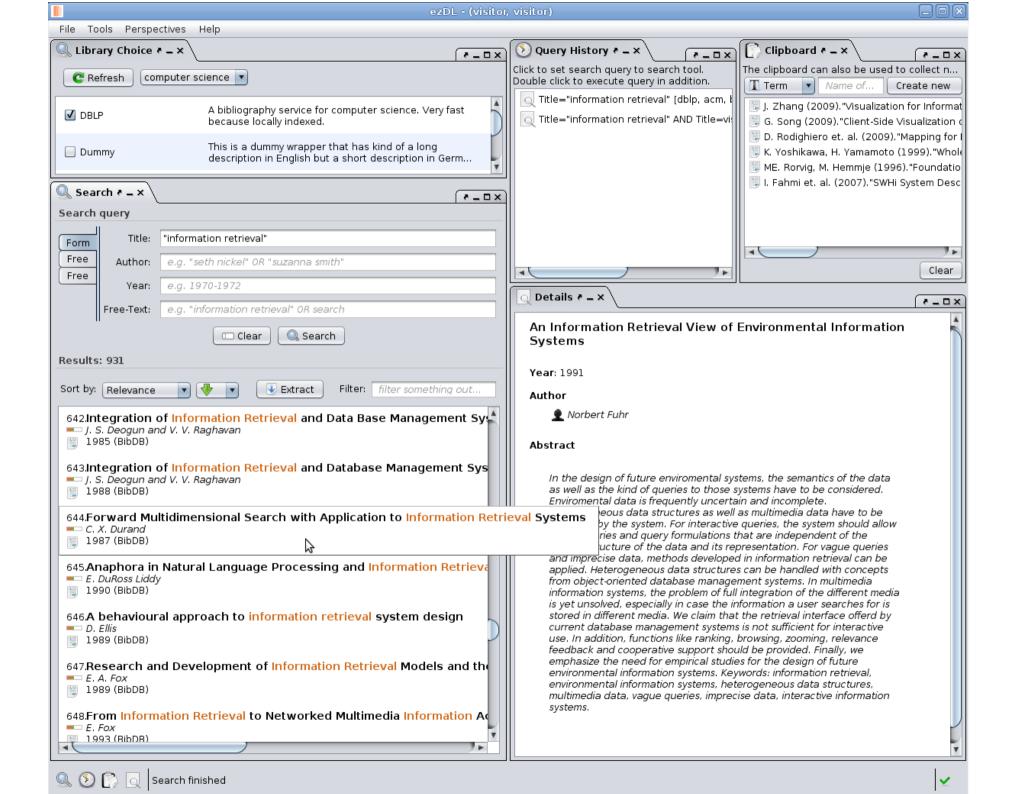


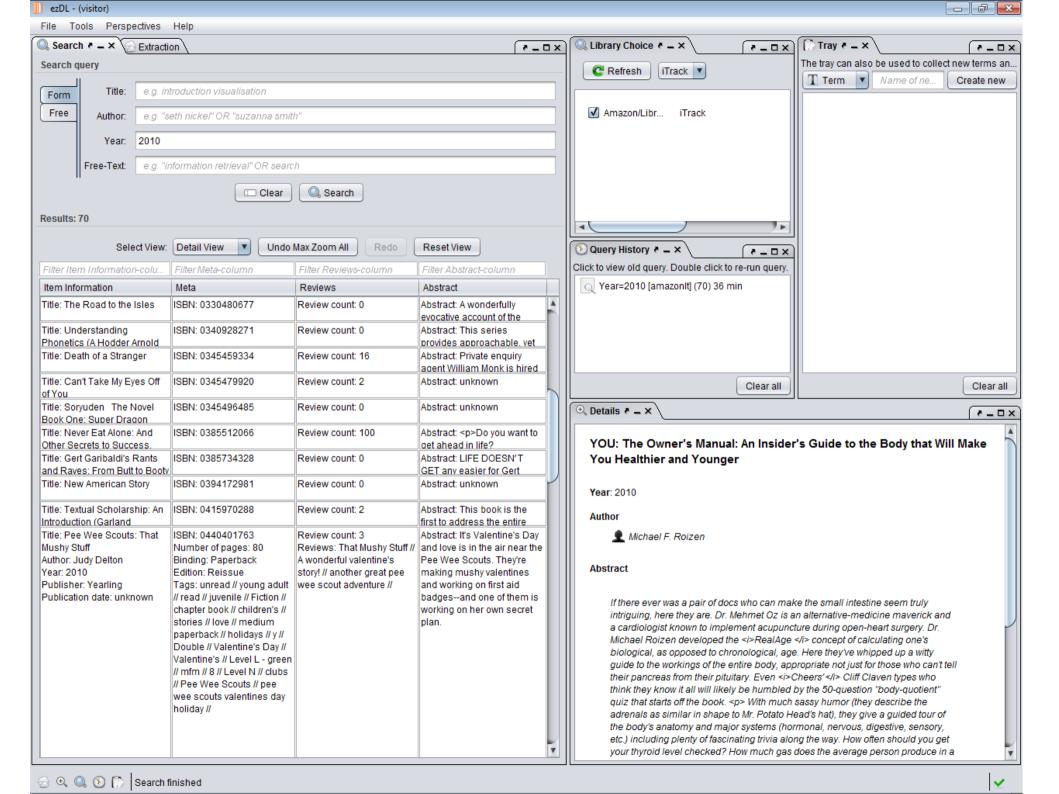
#### User interface tests

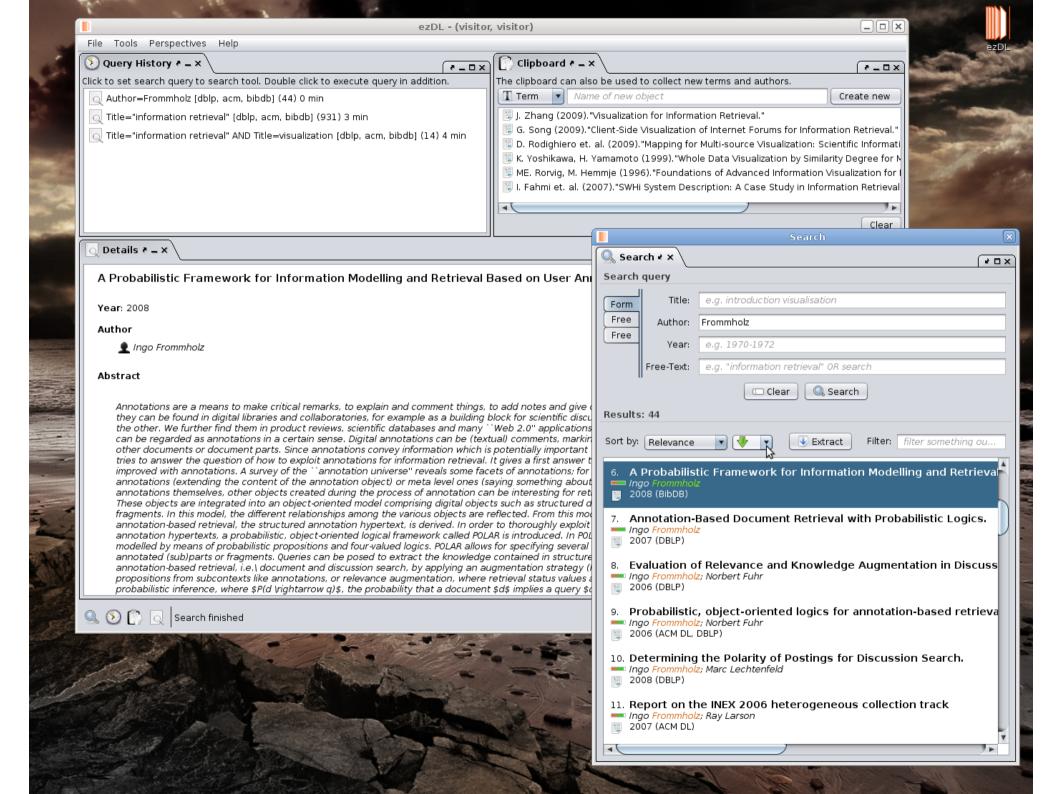
- user experiments
  - lab
  - living labs



## The ezDL system

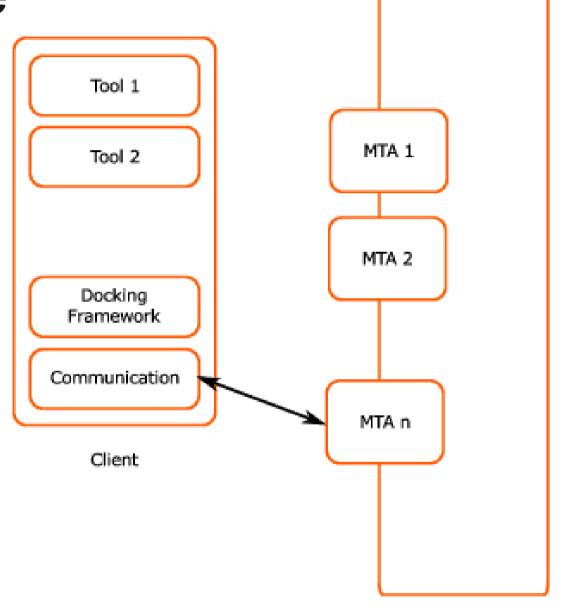






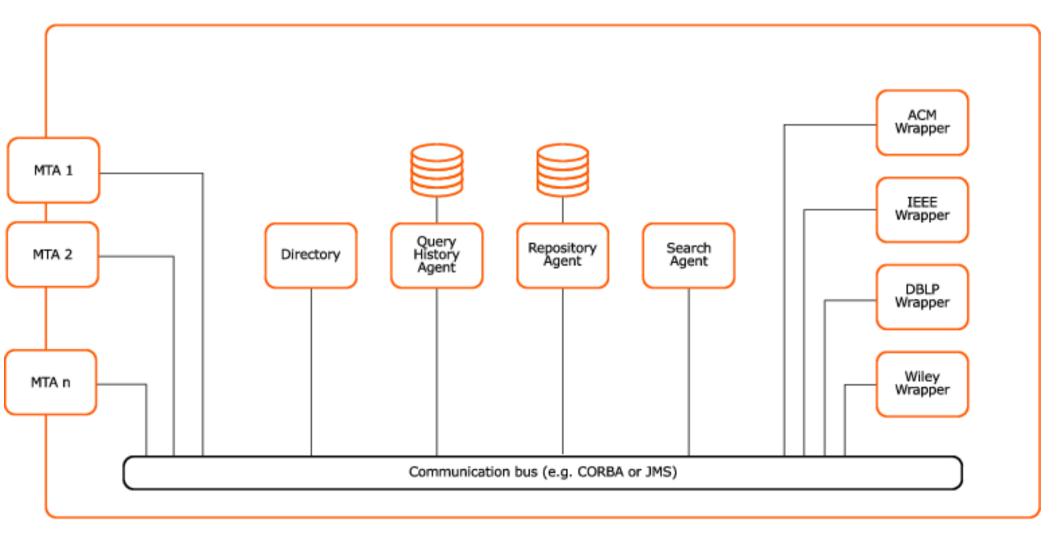


### ezDL Architecture





#### ezDL Backend Architecture







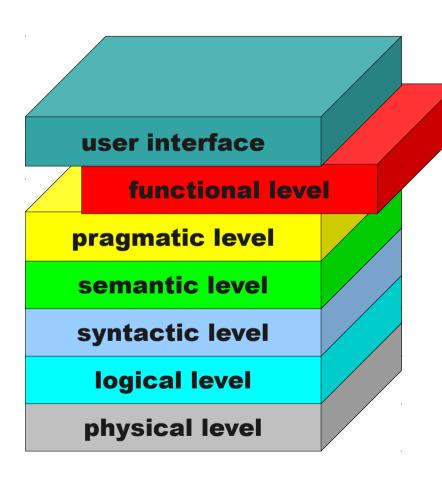
## user interface functional level pragmatic level semantic level syntactic level logical level physical level

#### Wrappers to IR Backends

- Amazon books (SOLR)
- Computer Science DLs: ACM, LEABIB, DBLP, Wiley, SpringerLink, IEEE Xplore, CiteseerX, ...
- Medicine: PubMed,...



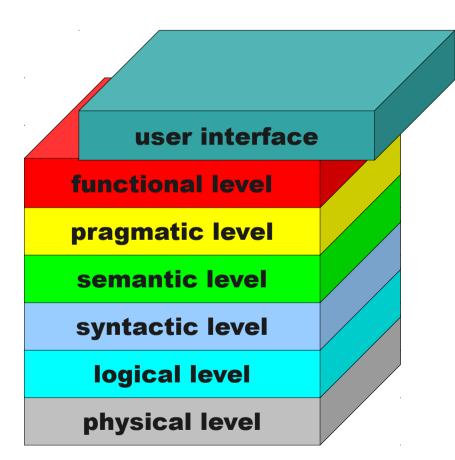


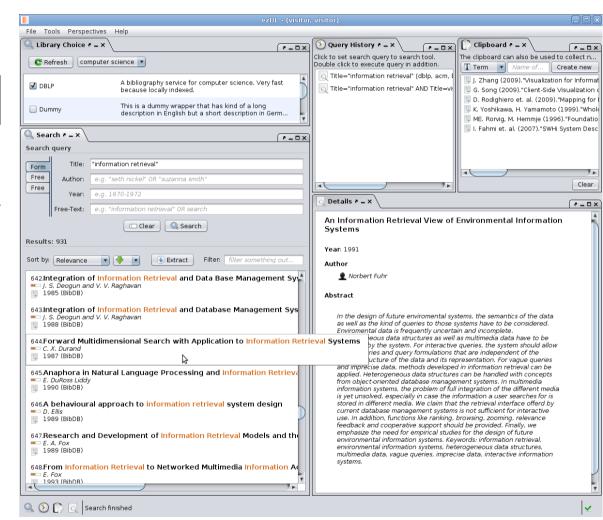


- higher level search functions: tactics, stratagems
- proactive support
- session support
- support of the digital library lifecycle



#### User Interface



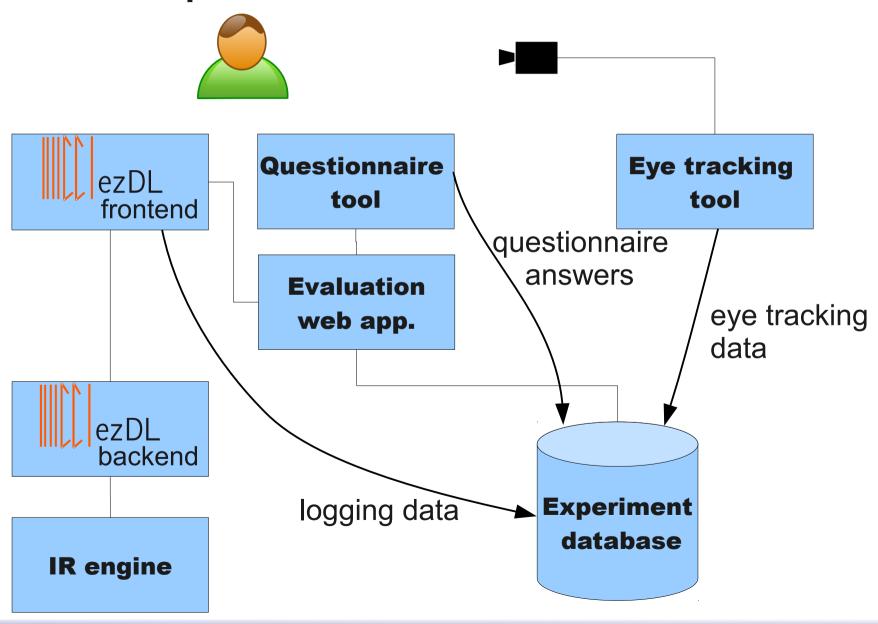


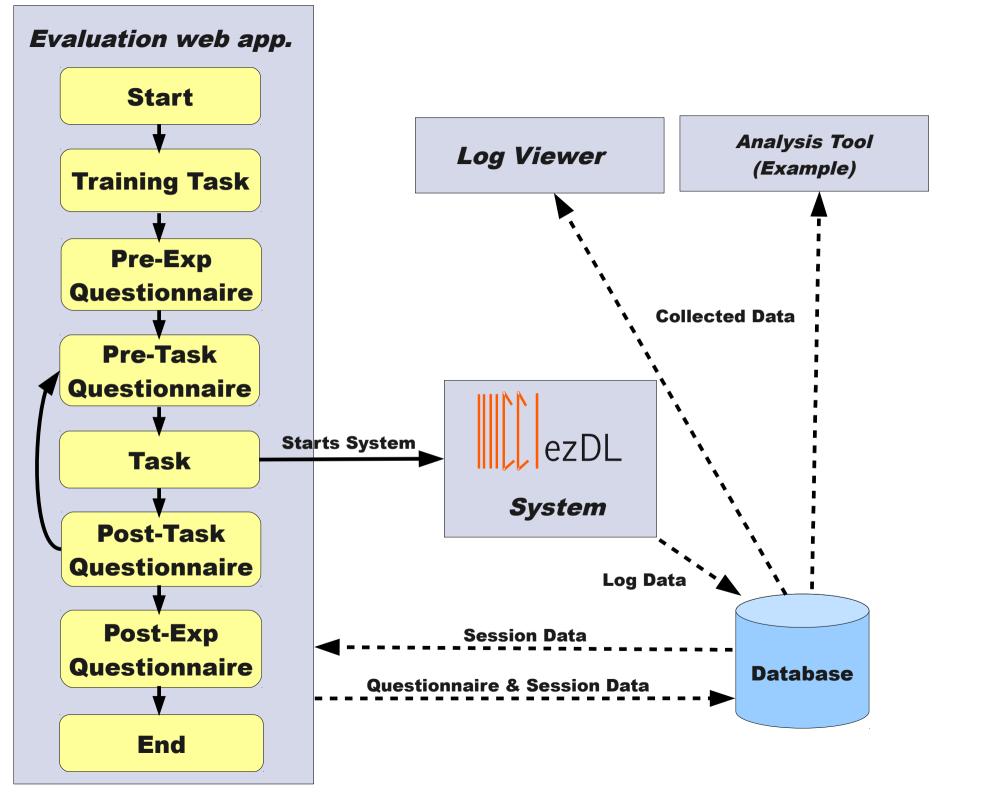


## ezDL experimentation



## Experimental architecture







## Experiment database

- Data controlling the experiments
  - task scheduling (latin square design)
  - scheduling of questionnaires
- Results:
  - logging data: actions, result items
  - questionnaire answers
  - eye tracking data

User:	visitor
Password:	•••••
First Task:	task1  ▼
Second Task:	task4   ▼
Third Task:	task7   ▼
Start	

#### Task 1 (task1)

You are at an early stage of working on an assignment, and have decided to start exploring the literature in order to get an overview of your topic. Your initial idea has led to the following research need:

Find trustworthy books discussing the conspiracy theories which developed after the 9/11 terrorist attacks in New York.

#### Task 2 (task2)

You are at an early stage of working on an assignment, and have decided to start exploring the literature in order to get an overview of your topic. Your initial idea has led to the following research need:

Find controversial books discussing the climate change and whether it is man-made or not.

### Task 3 (task3)

You are at an early stage of working on an assignment, and have decided to start exploring the literature in order to get an overview of your topic. Your initial idea has led to the following research need:

Find highly acclaimed novels that treats issues related to racial discrimination.

### Task 4 (task4)

You are in the data gathering stage of an assignment and need to collect a series of books that should serve as data for further analysis. This has led to the following research need:

Age:

### TPDL 2011 Workshop: User-Oriented Evaluation of Digital Library Interfaces

### **Pre-Experiment Questionnaire**

Dlasca	fill	out questionnaire	

Gerider.	Tibuse Offer		
First language:			
Current Occupation:			
Which diplomas/degrees have you b	een awarded?		
High School		area/field:	
Bachelor or equivalent		area/field:	
Master or equivalent		area/field:	₽
PhD, Doctoral equivalent		area/field:	N
Professional degree (medicine, la	aw etc.)	area/field:	
If you are a student, which universit  Bachelor or equivalent  Master or equivalent  PhD, Doctoral equivalent  Professional degree (medicine, la		area/field: area/field: area/field: area/field: area/field:	
Overall, how many years have you u	ised the Web or other comput	erized resources to search for information?	
How often do you search? (1: Never, 2: Once/twice a year, 3: Once/twice a month, 4: Once/twice a week, 5: once/twice a day)Digital libraries of scholarly articles (e.g. institutional repositories, ACM Digital Library)Web search engines (e.g. Google)Digital bookstore, such as e.g. Amazon		day)  0 1 0 2 0 3 0 4 0 5  0 1 0 2 0 3 0 4 0 5  0 1 0 2 0 3 0 4 0 5	
If you already have used a digital be to find a known book for purchas to explore bookstore for leisure; to explore bookstore for study/w	e purposes	ch purpose?	

# Logging

### **Log Viewer**

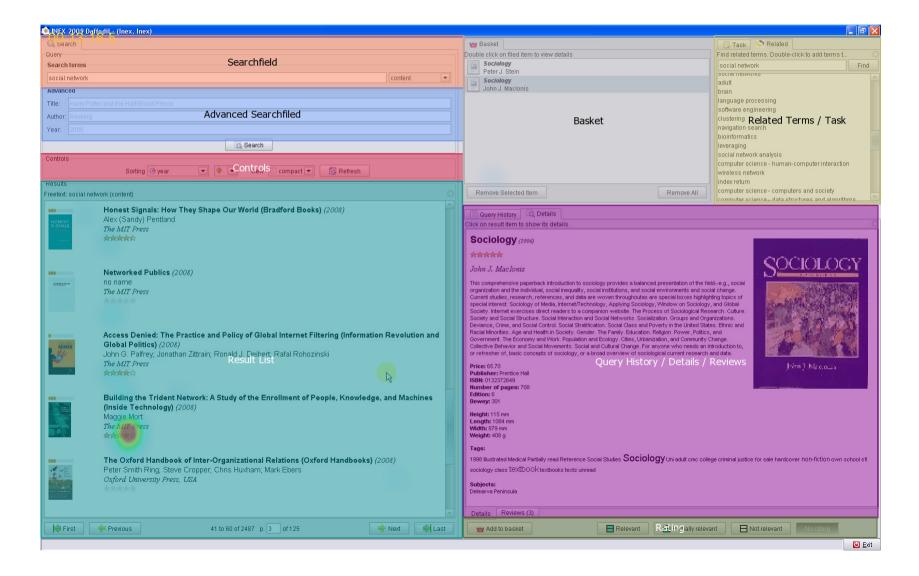
### **Session overview**

Consider Island	Lineau Tel	Chart	l a a Data
Session Id	User Id	Start End	Log Data
3b658a78-8761-429c-8641-33e4d9ab8bed	oslo-13	Wed, 31 Aug 2011 16:02:48 +0200 Wed, 31 Aug 2011 16:14:28 +0200	XML
cfea03fb-6dd8-4384-be4c-c8c623907aa7	oslo-12	Wed, 31 Aug 2011 15:54:34 +0200 Wed, 31 Aug 2011 16:04:41 +0200	XML
bf611c9f-ba71-4c86-91eb-1eba6dd42ae7	oslo-11	Wed, 31 Aug 2011 15:53:30 +0200 Wed, 31 Aug 2011 16:05:14 +0200	XML
959e0013-8c31-4805-9af4-b17341e8f299	oslo-10	Wed, 31 Aug 2011 15:49:09 +0200 Wed, 31 Aug 2011 15:59:02 +0200	XML
d52d8559-a6b7-4558-80e9-321c07378a5d	oslo-13	Wed, 31 Aug 2011 15:47:52 +0200 Wed, 31 Aug 2011 15:59:45 +0200	XML
31aa11c1-d95b-43b7-909b-ae5cd8bf510a	oslo-12	Wed, 31 Aug 2011 15:45:13 +0200 Wed, 31 Aug 2011 15:51:20 +0200	XML
ed845846-b198-485b-a364-f95eba10fe81	oslo-12	Wed, 31 Aug 2011 15:35:43 +0200 Wed, 31 Aug 2011 15:38:17 +0200	XML
58ae76a3-51e0-4661-b2f5-c5390476f89d	oslo-11	Wed, 31 Aug 2011 15:35:40 +0200 Wed, 31 Aug 2011 15:50:47 +0200	XML
f3a7801d-002e-46c2-897f-75fabb38513e	oslo-12	Wed, 31 Aug 2011 15:35:32 +0200 Wed, 31 Aug 2011 15:37:27 +0200	XML
31fdc70c-dddc-46a6-878c-e759b2339c88	oslo-10	Wed, 31 Aug 2011 15:31:56 +0200 Wed, 31 Aug 2011 15:47:04 +0200	XML
263c5b48-0195-49a8-ae31-fa915ff60deb	oslo-12	Wed, 31 Aug 2011 15:25:25 +0200 Wed, 31 Aug 2011 15:34:56 +0200	XML
e20454af-70b8-4f19-95f4-4a6c35eba044	oslo-12	Wed, 31 Aug 2011 15:25:12 +0200 Wed, 31 Aug 2011 15:35:15 +0200	XML
7b6d5032-aaa7-412b-834f-0101c41dc400	oslo-13	Wed, 31 Aug 2011 15:24:38 +0200 Wed, 31 Aug 2011 15:39:20 +0200	XML
481069c6-4234-426d-8797-3925a72311ab	oslo-11	Wed, 31 Aug 2011 15:23:42 +0200 Wed, 31 Aug 2011 15:31:55 +0200	XML
95c57258-a7bc-4993-a141-f198d01d87b8	oslo-10	Wed, 31 Aug 2011 15:20:58 +0200 Wed, 31 Aug 2011 15:29:42 +0200	XML
75ce5d7f-dc05-456f-a7a0-def3c40f621a	oslo-13	Wed, 31 Aug 2011 15:15:21 +0200 Wed, 31 Aug 2011 15:20:59 +0200	XML
144949f7-3b32-4ecc-9f1e-52b46f89b283	oslo-13	Wed, 31 Aug 2011 15:13:48 +0200 Wed, 31 Aug 2011 15:15:19 +0200	XML
a443b142-7196-4054-93ab-caaf9bb868d9	oslo-12	Wed, 31 Aug 2011 15:13:45 +0200 Wed, 31 Aug 2011 15:19:25 +0200	XML
f7c96b09-2481-4233-b455-566c61d7ce11	oslo-11	Wed, 31 Aug 2011 15:12:39 +0200 Wed, 31 Aug 2011 15:18:37 +0200	XML
58362f1c-9d74-456e-99c3-6b75460c837b	oslo-10	Wed, 31 Aug 2011 15:12:36 +0200 Wed, 31 Aug 2011 15:17:38 +0200	XML
5ebf05da-6117-4ee8-ab23-86382bfc5f9e	oslo-7	Wed, 31 Aug 2011 14:20:26 +0200 Wed, 31 Aug 2011 14:33:36 +0200	XML
9c399787-b92d-48b8-84bd-9e6ebd7b5ced	oslo-9	Wed, 31 Aug 2011 14:13:15 +0200 Wed, 31 Aug 2011 14:28:35 +0200	XML
f9af0877-0d05-4d73-a0ac-66bce49513b4	oslo-6	Wed. 31 Aug 2011 14:07:31 +0200 Wed. 31 Aug 2011 14:22:39 +0200	) XMI

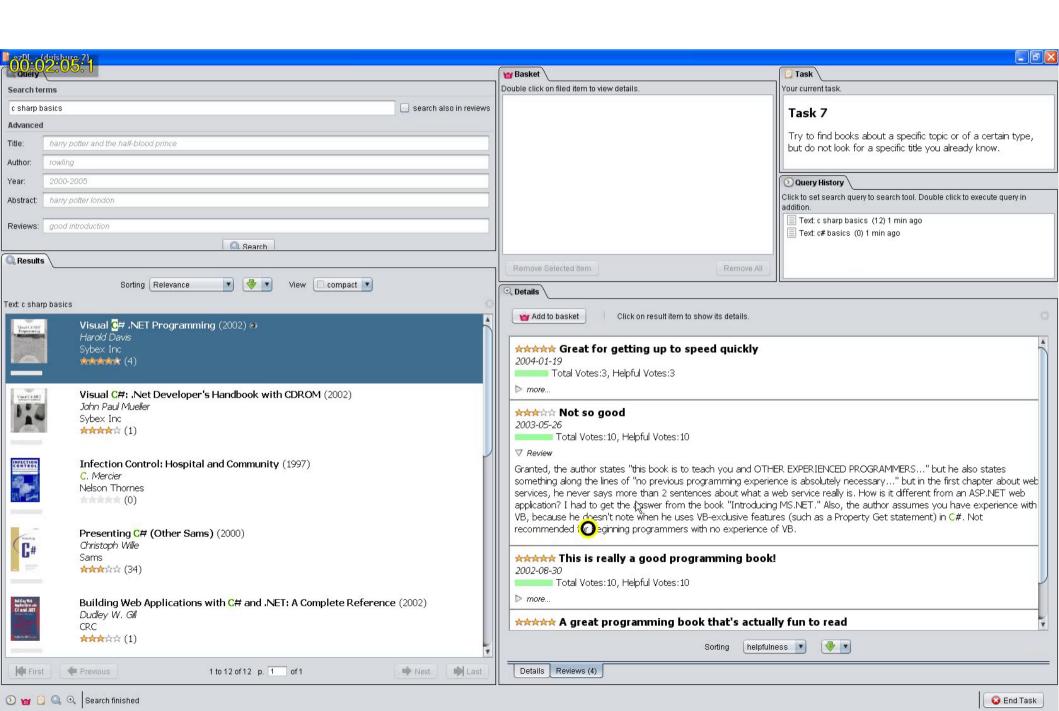
```
-<log>
  - <event name="start" timestamp="1314799370195" timestampClient="1314799409473">
      <user>oslo-13</user>
      <version>a</version>
    </event>
  - <event name="taskselected" timestamp="1314799370263" timestampClient="1314799409473">
      <task>task3</task>
    </event>
    <event name="clientstart" timestamp="1314799370519" timestampClient="1314799410020"/>
  - <event name="desktopinfo" timestamp="1314799370551" timestampClient="1314799410020">
      <screenX>1680</screenX>
      <screenY>1050</screenY>
      <os>Windows 7</os>
      <java>1.6.0 26</java>
    </event>
  - <event name="search" timestamp="1314799395637" timestampClient="1314799435519">
      <solrOuery>4:racial AND 4:discrimination</solrOuery>
      <textQuery/>
      <titleQuery>racial discrimination</titleQuery>
      <authorQuery/>
      <yearQuery/>
      <abstractQuery/>
      <reviewQuery/>
      <searchType>content</searchType>
      <startDocNum>0</startDocNum>
      <endDocNum>19</endDocNum>
      <sorting order="descending">9998</sorting>
    - <fields>
        <field>7</field>
        <field>1003</field>
        <field>4</field>
        <field>31</field>
        <field>8029</field>
        <field>1018</field>
        <field>8030</field>
      </fields>
      ction>compact
    </event>
  - <event name="results" timestamp="1314799396417" timestampClient="1314799436174">
```



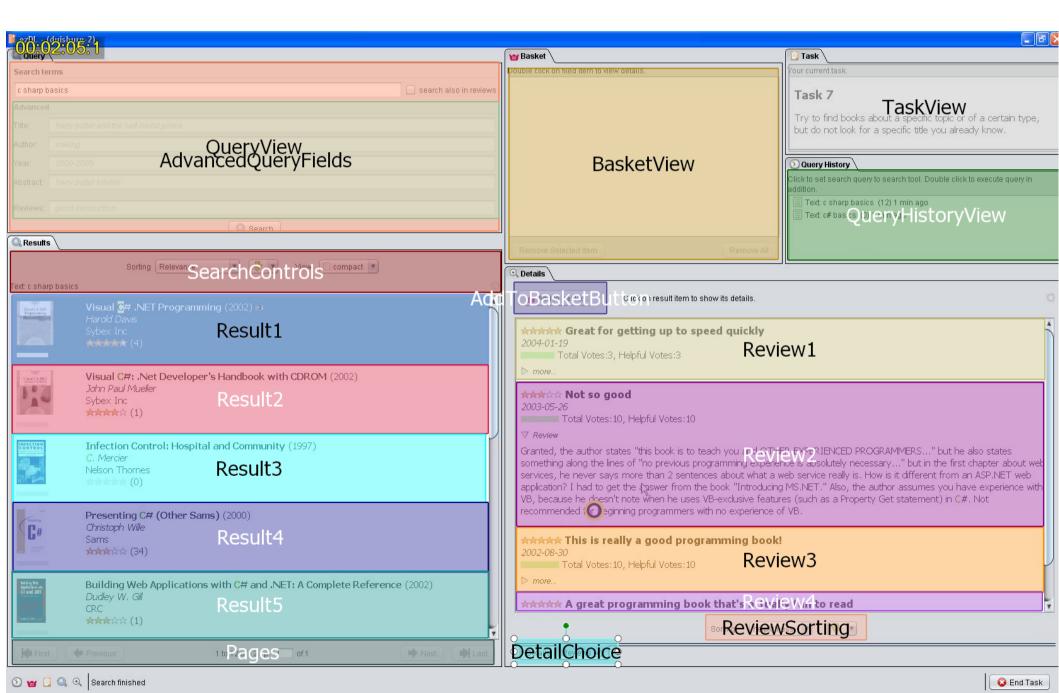
## Eye Tracking: Areas of Interest (AOI)



# Eye tracking: Dynamic AOIs

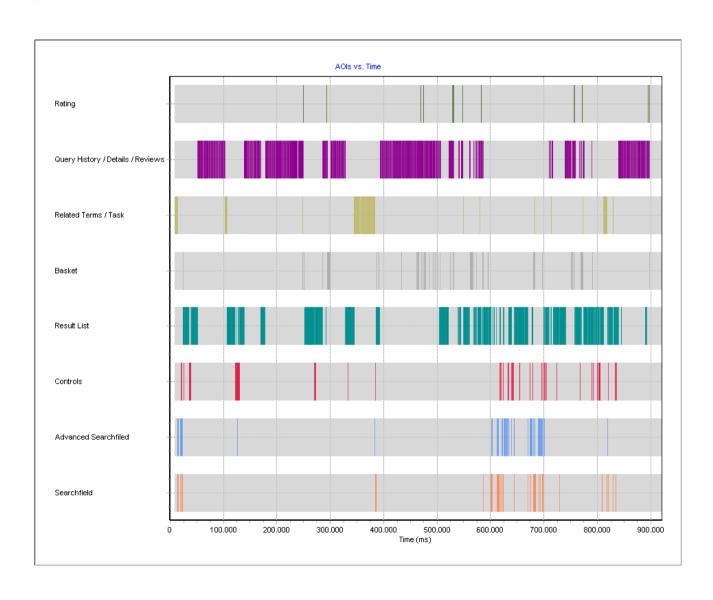


# Eye Tracking: Dynamic AOIs



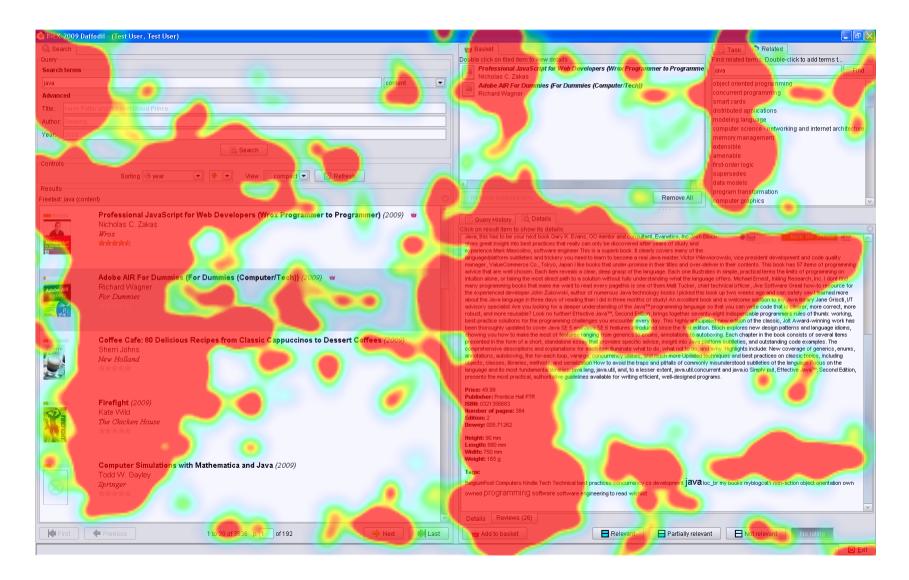


## Eye Tracking: AOIs vs. Time





## Eye Tracking: Heatmap









### Conclusion

- Easy performance of IIR experiments
- Flexible configuration
- Extensible system
- Distributed system
- Centralized collection of experimental data
- Flexible analysis of data