gesis

Search term recommendation and non-textual ranking evaluated

Philipp Mayr, Peter Mutschke, Philipp Schaer, York Sure
GESIS Leibniz Institute for the Social Sciences

9th European NKOS Workshop at the 14th ECDL Conference
Glasgow, Scotland

9th September 2010

Leibniz Gemeinschaft

Topic of the talk

1. Value-added services to **improve the search** in Digital Libraries (DL)
   - Short introduction of the services
2. Results from a **retrieval evaluation** exercise
   - Evaluation of a KOS-based Search Term Recommender (STR)
   - Evaluation of re-ranking approaches: Bradfordizing and author centrality

Situation

Major difficulties in most text-oriented DL

1. **Vagueness between search and indexing terms** (language problem)
2. **Too much documents and unstructured result sets** (information overload)
3. **Pure text-statistical rankings (e.g. tf-idf) have serious weaknesses** (short document problem)

Consequence: Users are unsatisfied with a retrieval system and **search is far beyond its possibilities**

IRM project

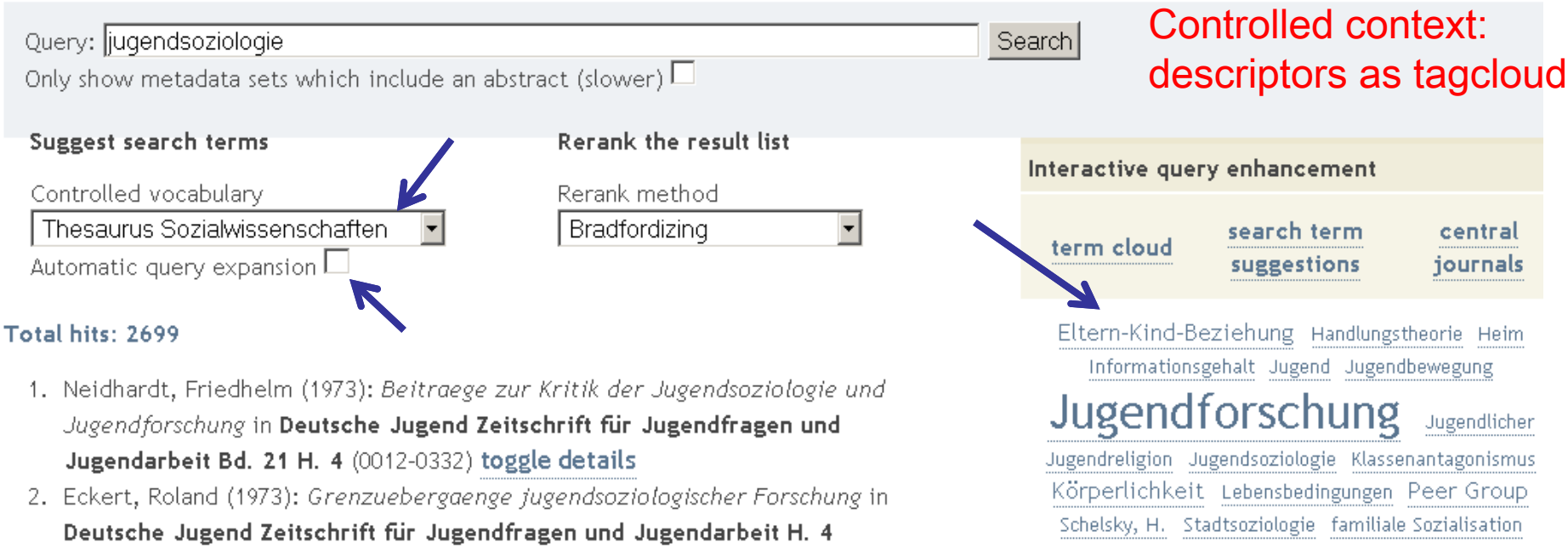
Funded by DFG (grant number INST 658/6-1)

Motivation:

- Implementation and evaluation of value-added services for retrieval in DL
- Using science models (**specialties in languages**, **networks of authors**, **centrality of journals**) (Bates, 1990)
- Improvement of search results/experience:
  - Search term recommendation (bridge between user and controlled specialized vocabulary/ies)
  - Non-textual ranking techniques (providing computational approaches to rank documents, provide context)

Search Term Recommender (STR)

Recommendation of highly associated controlled terms
Expansion of the top n ranked controlled terms



Controlled context: descriptors as tagcloud

Statistical relations:
- Co-word analysis
- Complete database and KOS as training set
- Automatic mapping of terms

Make use of knowledge/specialties in a KOS and the statistical relations computed by the co-word analysis

http://www.gesis.org/beta/prototypen/irm/

Example

- **Starting with a unspecific term "Unemployment"**
- **Pure text-based standard ranking produces many and relatively poor results**
- **STR recommends controlled terms from a pre-selected KOS (thesaurus)**

Query: Unemployment Search

Total hits: 9902

1. by Herbert Leder (1963): *Unemployment benefit and unemployment assistance* toggle details
2. (1939): *Black coffins and the unemployed* toggle details
3. (1983): *Unemployment: cause and cure* (0-85520-622-5) toggle details
4. (1971): *Employment and unemployment in Ceylon* toggle details

Search Term Recommendation (STR)

Interaction: User **refines search** with a recommended controlled term e.g. „Job Search"

Total hits: 412

1. Kahn, Lawrence M.; Low, Stuart A. (1982): *The relative effects of employed and unemployed job search* in **The review of economics and statistics Vol. 64 Nr. 2** (0034-6535) toggle details
2. Bolinder, Margareta (2006): *THE IMPORTANCE OF SPACE FOR NEGOTIATING FOR UNEMPLOYED PERSONS' EXPERIENCES, STRATEGIES FOR ACTION AND CHANCES FOR OBTAINING A JOB* (1104-2508) toggle details
3. Thomas, Jonathan P.; McFadyen, Ruth G. (1997): *Economic and Psychological Models of Job Search Behavior of the Unemployed* in **Human Relations 1997, 50, 12, Dec, 1461-1484.** (0018-7267) toggle details

Service: Reranking (Bradfordizing)

Sorting central papers in core journals/publishers on top

- **A build-in functionality of the SOLR system aggregates articles**
- **The frequency counts of the journals are used to boost documents**

Query: jugendsoziologie and collection:iz* Search

Only show metadata sets which include an abstract (slower)

Suggest search terms

Controlled vocabulary: Thesaurus Sozialwissenschaften

Automatic query expansion

Rerank the result list

Rerank method: Bradfordizing

Journal context

Interactive query enhancement

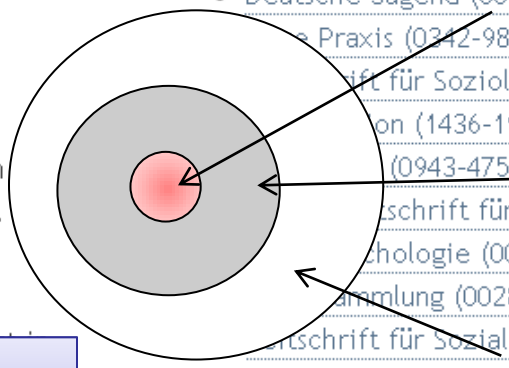
term cloud | search term suggestions | central journals

Interaction: User selects a certain journal

Total hits: 2539

1. Bendit, Rene; Gaiser, Wolfgang (1992): *Einführung in das Thema: 'Europa und sein Nachwuchs'* in **DISKURS Studien zu Kindheit, Jugend, Familie und Gesellschaft Jg. 2 H. 2** (0937-9614) toggle details
2. Mauger, Gerard (1992): *Jugend in Frankreich : Reproduktion sozialer Ungleichheit im gesellschaftlichen Wandel* in **DISKURS Studien zu Kindheit, Jugend, Familie und Gesellschaft Jg. 2 H. 2** (0937-9614) toggle details
3. (1992): *Die Horizontalisierung der Lebenswelt italienischer Jugendlicher* in **DISKURS Studien zu Kindheit, Jugend, Familie und Gesellschaft Jg. 2 H. 2** (0937-9614) toggle details
4. Bendit, Rene; Gaiser, Wolfgang; Nissen, Ursula (1992): *Aufwachsen in der*

- DISKURS (0937-9614) [150]
- Aus Politik und Zeitgeschichte (0479-611X) [150]
- Deutsche Jugend (0012-0332) [146]
- Praxis (0342-9857) [147]
- ift für Soziologie der Erziehung und
- on (1436-1957) [89]
- (0943-4755) [86]
- schrift für Soziologie und
- hologie (0023-2653) [74]
- mmlung (0028-3355) [69]
- tschrift für Sozialisationsforschung und Erziehungssoziologie (0724-4967) [58]
- Medien und Erziehung (0723-399X) [52]



Make use of publication regularities on a topic (coreness of journals)

Service: Reranking (Author centrality)

- **A background procedure computes a co-authorship network for the whole result**
- **The centrality values of the computed authors are used to boost documents**

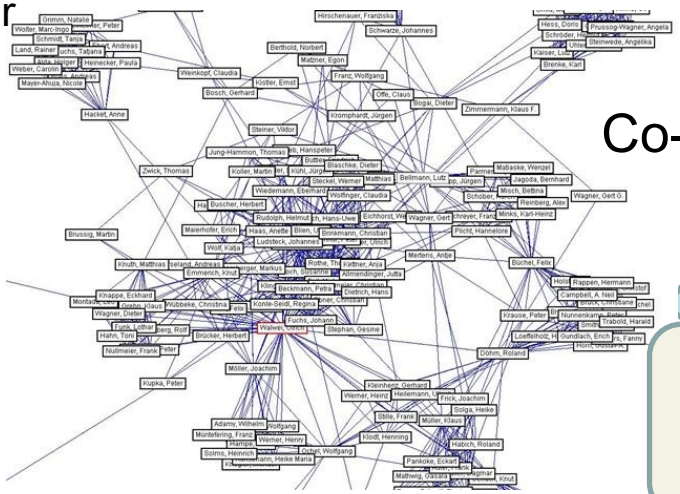
Query: Unemployment Search

Total hits: 9902

1. by Herbert Leder (1963): *Unemployment benefit and unemployment assistance* toggle details
2. (1939): *Black coffins and the unemployed* toggle details
3. (1983): *Unemployment: cause and cure* (0-85520-622-5) toggle details
4. (1971): *Employment and unemployment in Ceylon* toggle details

Interaction: User selects a certain author

Re-Ranking



Co-authorship network

- Bellmann, Lutz [0.017583124]
- Bender, Stefan [0.011249411]
- Lehmann, Hartmut [0.008765155]
- Walwei, Ulrich [0.007932616]
- Layard, Richard [0.007899008]
- Hujer, Reinhard [0.007864751]

Author networks (AUTH)

Make use of knowledge about the interaction and cooperation of authors

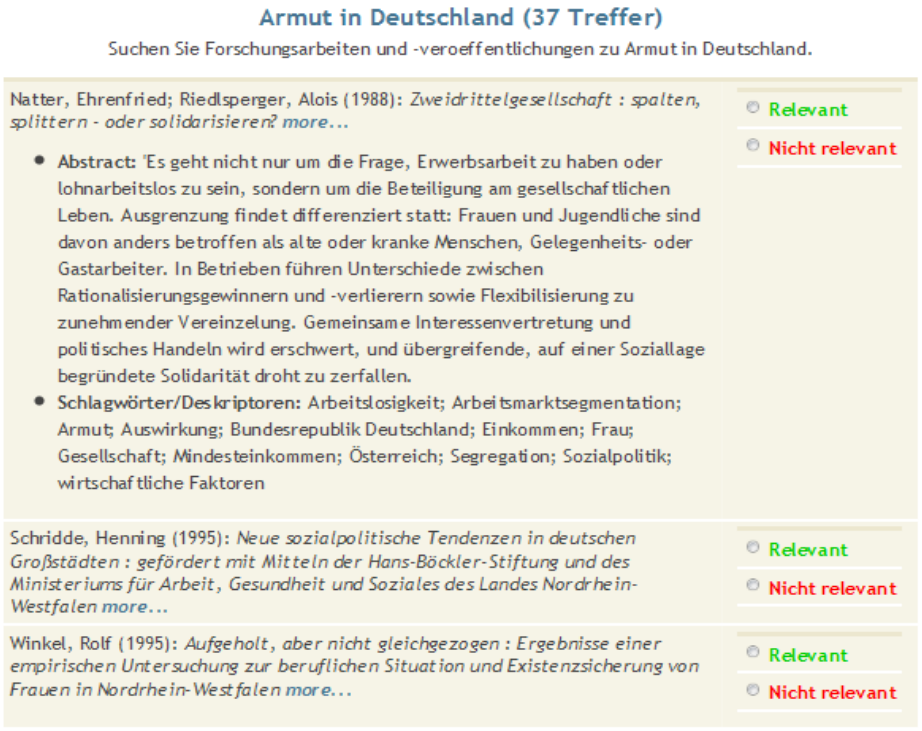
Main questions

- What is examined?
  - the quality of different services to improve search
  - or the quality of the associated search of the services
- Can we improve search in a DL with different KOS-enhanced databases?
  - In one discipline
  - Between at least two disciplines
- The services are helpful to whom?
  - Experts/scientists
  - Students

Evaluation

10 different topics from the CLEF conferences (title and descriptions), 73 LIS students, German SOLIS database as test collection

Pooled document set in the assessment tool



1. Selection of a topic
2. Assessment of documents: top 10 ranked documents for each service
   - **STR**
   - **BRAD**
   - **AUTH**
   - **SOLR (tf-idf as baseline)**
3. Evaluation of the relevance assessments (MAP, P@10)

Operationalization

All terms and topics are translated from German

Original query for the service SOLR:

povert* AND german*

STR: Expanded STR-enhanced query:

(povert* AND german*) **OR "poverty" OR "Federal Republic of Germany" OR "social assistance" OR "immiseration"** (highest confidence value)

BRAD: top 10 documents after BRAD-reranking (most frequent core journal)

AUTH: top 10 documents after AUTH-reranking (highest betweeness value)

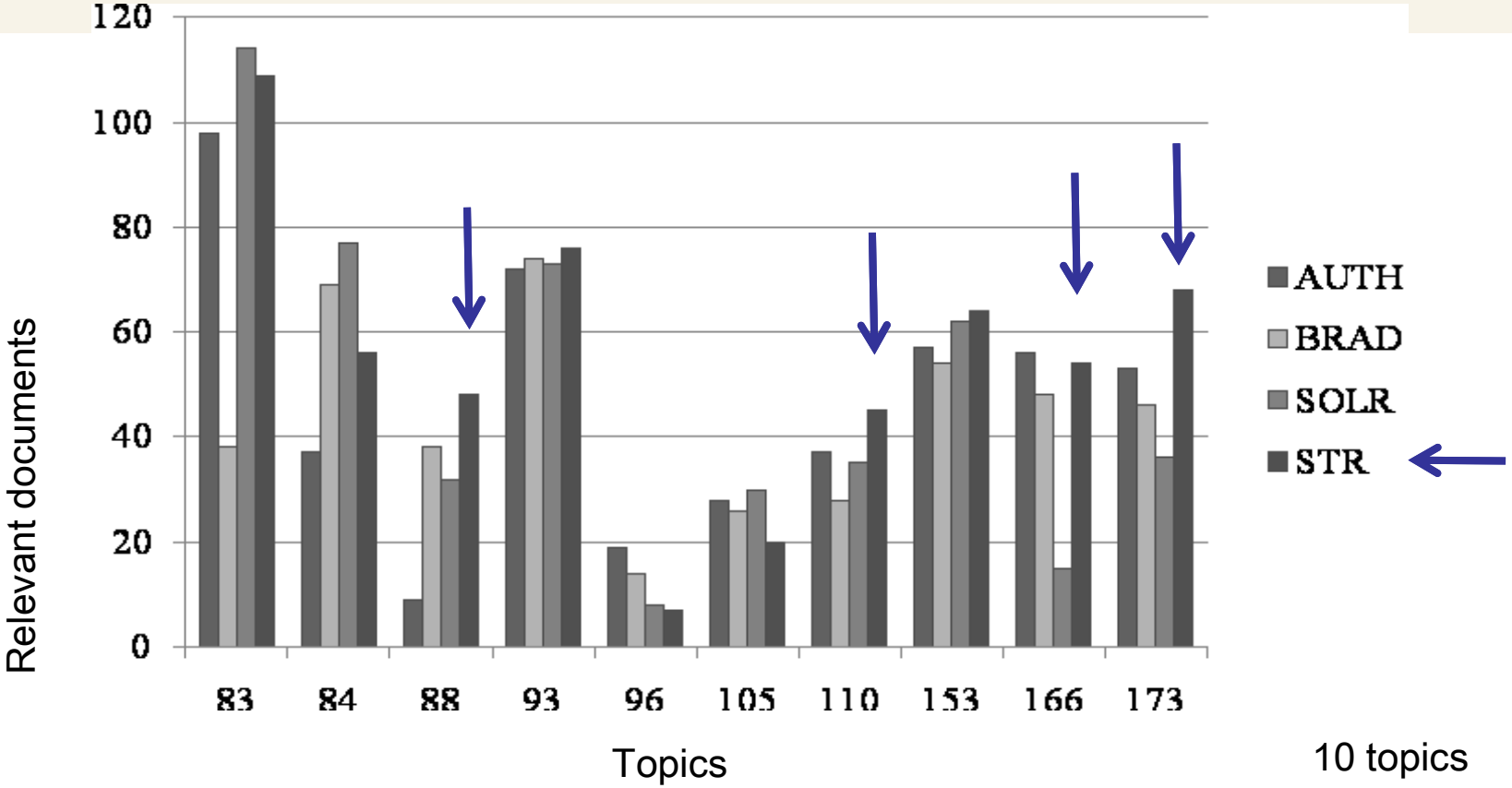
Topics

- 73 assessors did 3,196 single relevance judgments
- topic 83 was picked 15 times – topic 96 twice
- fair and moderate inter-rater agreement

topic	title	description	*n*	*N*	*k*	K	>=0.8	=1
83	Media and War	Find documents on the commentatorship of the press and other media from war regions.	15	40	2	0.522	0.727	0.225
84	New Media in Education	Find documents reporting on benefits and risks of using new technology such as computers or the Internet in schools.	11	40	2	0.304	0.5	0.15
88	Sports in Nazi Germany	Find documents about the role of sports in the German Third Reich.	6	40	2	0.528	0.75	0.425
93	Burnout Syndrome	Find documents reporting on the burnout syndrome.	10	40	2	0.411	0.8	0.35
96	Costs of Vocational Education	Find documents reporting on the costs and benefits of vocational education.	2	40	2	0.488	0.775	0.775
105	Graduates and Labour Market	Find documents reporting on the job market for university graduates.	5	40	2	0.466	0.675	0.525
110	Suicide of Young People	Find documents investigating suicides in teenagers and young adults.	5	40	2	0.222	0.625	0.425
153	Childlessness in Germany	Information on the factors for childlessness in Germany	10	40	2	0.202	0.325	0.175
166	Poverty in Germany	Research papers and publications on poverty and homelessness in Germany.	9	40	2	0.438	0.5	0.25
173	Propensity towards violence among youths	Find reports, cases, empirical studies and analyses on the capacity of adolescents for violence.	10	40	2	0.411	0.55	0.2
				avg.		*0.4*	*0.622*	*0.35*

Schaer et al.

Results



Number of documents judged relevant by the assessors per topic

- STR by adding the 4 descriptors with the highest confidence
- STR was best service (topics 88, 110, 166 and 173 with an improvement of 20% at least compared to the baseline)

10 topics
1 database
top 10 docs
4 services
73 students

Mayr et al., submitted

Results

Precision for each topic and service

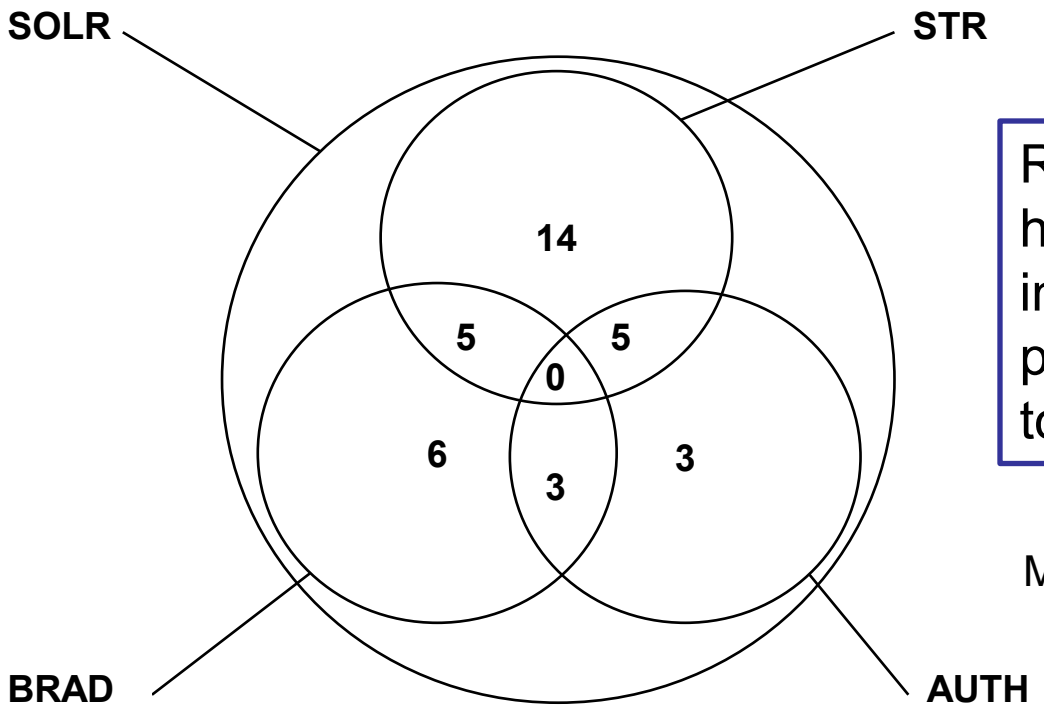
-Weaknesses of each service on a concrete topic can be observed

-Different thresholds applied make a more diverse result

	non relevant				relevant				precision			
topic	AUTH	BRAD	SOLR	STR	AUTH	BRAD	SOLR	STR	AUTH	BRAD	SOLR	STR
83	42	104	36	25	98	38	114	109	0.70	0.27	0.76	0.81
84	71	38	27	26	37	69	77	56	0.34	0.64	0.74	0.68
88	51	22	28	12	9	38	32	48	0.15	0.63	0.53	0.80
93	28	26	25	26	72	74	73	76	0.72	0.74	0.74	0.75
96	3	8	12	13	19	14	8	7	0.86	0.64	0.40	0.35
105	15	18	15	24	28	26	30	20	0.65	0.59	0.67	0.45
110	13	22	15	5	37	28	35	45	0.74	0.56	0.70	0.90
153	42	40	36	32	57	54	62	64	0.58	0.57	0.63	0.67
166	30	39	72	33	56	48	15	54	0.65	0.55	0.17	0.62
173	41	48	57	26	53	46	36	68	0.56	0.49	0.39	0.72
								avg.	*0.60*	*0.57*	*0.57*	*0.68*
								avg.(kappa >= 0.4)	*0.61*	*0.56*	*0.52*	*0.64*
								avg. (overlap >= 0.35)	*0.63*	*0.63*	*0.61*	*0.65*

Schaer et al.

Results



Read: SOLR and STR have 14 document intersections from 100 possible (for all 10 topics)

Mayr et al., submitted

- 36 intersections of suggested top n=10 documents over all topics and services
- Total of 400 documents: (4 services * 10 per service * 10 topics)
- Result sets are close to disjoint, services provide quite different views onto the document space

Implications

1. Precision values of the retrieval services are the **same or better than standard text-statistical ranking** (tf-idf)
2. Each service produces its **own document space** (measured in intersections)
3. Each service has weaknesses which can have implication on the relevance ratio
   - STR: misleading search terms (statistical artefacts) but high descriptive power of controlled terms
   - BRAD: high frequent but some irrelevant journals (outdated)
   - AUTH: central but misleading authors

Next steps

Searching and evaluating in a **multi-thesauri-scenario** (connecting STR services)

Interactive retrieval with services support

Evaluation of scientists on their specific research topics

Evaluation of **combination of services**

Reimplementation of the commercial categorizer

Short conclusions

1. Valuable alternative views on document spaces -> **especially in interactive retrieval (?)**
2. Improved relevance ratio and low overlap of the services results
3. Contexts matter!
   - Providing context to stimulate and support users
   - Support interactive systems for exploration

gesis

References

1. Philipp Schaer, Philipp Mayr, Peter Mutschke (accepted for ASIST 2010): **Demonstrating a Service-Enhanced Retrieval System**

2. Philipp Schaer, Philipp Mayr, Peter Mutschke (accepted for LWA 2010): **Implications of Inter-Rater Agreement on a Student Information Retrieval Evaluation**

3. Mayr, Philipp; Mutschke, Peter; Schaer, Philipp; Sure, York (Abstract accepted for Scientometrics Special Issue on Science Models): **Science Models as Value-Added Services for Scholarly Information Systems**.

Thank you!

Dr. Philipp Mayr
GESIS
Lennéstr. 30
53113 Bonn

mailto:philipp.mayr@gesis.org

IRM project
http://www.gesis.org/index.php?id=2479

IRM prototype
http://www.gesis.org/beta/prototypen/irm/