What Information Behaviour Mean for Use of Metadata and KOS in business environments

Marianne Lykke

Human-centered Science and Digital Technologi

Department of Communication and Psychology
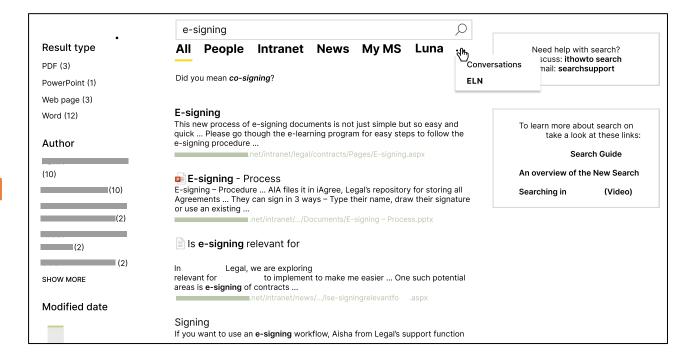
Aalborg University

Today's program

- **Context of talk:** Enterprise search and workplace search tasks

- **What do we know:** Information behavior and practices at the workplace

- **Findings from case study:** International biotech company

- **Lessons learnt:** Role and use of metadata and KOS in enterprise search

Enterprise search

Enterprise search is managed search environment that enables employees to find information they can rely on in making decisions that will achieve organizational and personal objectives (White, 2015)

Search of digital textual materials owned by an organization, including search of their external website, company intranet and other electronic text that they hold such as email, database records and shared documents (Hawking, 2004, White 2015)

Enterprise search system based on Microsoft Sharepoint

Workplace information retrieval and behavior

- Workplace information retrieval is often targeted a single "right answer" (Fagin et als, 2003; Freund, Toms & Waterhouse, 2005)

- Workplace searching is simple, with few search terms and little use of facets (Stenmark, 2005, 2007, 2010; Teevan et als., 2004)

- Corporate information is stored in hetegenous documents (Hawking, 2004)

- Several professional languages and perspectives in play -> the Vocabulary problem (Furnas et als. 1987; Lykke Nielsen, 2005, Cleverley & Burnett, 2015)

- Workplace searchers prefer to browse as opposed to searching (Stenmark, 2010; Teevan et als., 2004; Stocker et al., 2015)

- People search and expertise search is important (Hertzum & Pejtersen, 2000; Guy, 2012; Hertzum, 2014; Freund, 2015)

- Metadata are useful, but often limited and simple use in workplace retrieval systems (Schymik et al., 2015; Stocker et al., 2015)

Information behavior in international biotech company

7500 employees, organised into 7 organizational units:
Top management, Administration, R&D, Production and 3 business application areas

Search log analysis – 4-month period

- 5,854 active users
- Data: user ID, data and time, source used, search queries, organizational affiliation and job category
- Categorization of search queries into subject categories

Questionnaire survey

- Sent to 226 most active users over prior three days, response rate 43.4% (98 users)
- Data: Search frequency, time spent on searching, satisfaction, sources, information types, reformulation strategies

Interviews

- 8 frequent users selected from administration, R&D, production, and the three business areas. General management was not available for the study.
- Demonstrations of 19 search tasks

Subject categories – per organisational areas

SUBJECT CATEGORIES	ADMINISTRATION	BUSINESS AREAS	PRODUCTION	R&D	TOP MANAGEMENT	GRAND TOTAL
People Search	81.20%	69.69%	71.03%	68.18%	84.36%	72.24%
Quality	1.57%	4.63%	8.98%	8.22%	0.00%	6.02%
IT	4.63%	4.73%	5.60%	4.74%	3.34%	4.88%
HR	3.11%	3.04%	3.66%	3.59%	3.34%	3.38%
Product	0.27%	4.24%	1.46%	2.48%	0.70%	2.10%
Finance	1.92%	1.96%	1.97%	1.68%	2.28%	1.85%
Facility Services	1.12%	1.19%	1.06%	1.57%	1.93%	1.29%
Sourcing	0.91%	1.30%	0.70%	1.86%	0.00%	1.28%
Zymernet	1.32%	1.27%	0.83%	1.44%	1.76%	1.26%
R&T	0.18%	1.55%	0.61%	1.49%	0.00%	1.01%
Location	0.89%	0.72%	0.99%	0.85%	0.53%	0.86%
Communications	0.98%	1.10%	0.54%	0.79%	0.70%	0.85%
Sales	0.37%	1.92%	0.16%	0.14%	0.53%	0.56%

Information source used – per organisational area

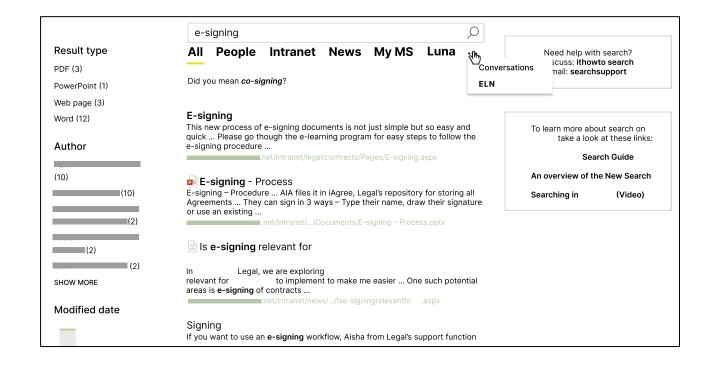| Source used	Administration	Business Areas	Production	R&D	Top management	Grand Total
All	74.03%	74.40%	71.02%	73.64%	83.88%	73.37%
People	22.34%	18.72%	21.86%	18.13%	13.02%	19.97%
Intranet	1.75%	2.71%	3.20%	2.56%	2.33%	2.53%
Luna	0.33%	2.24%	1.06%	2.26%	0.00%	1.56%
MyMS	0.60%	1.16%	2.55%	2.75%	0.16%	1.88%
ELN	0.04%	0.14%	0.00%	0.12%	0.00%	0.08%
Conversations	0.22%	0.12%	0.03%	0.14%	0.31%	0.13%
News	0.57%	0.30%	0.24%	0.38%	0.31%	0.38%
Total	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

Searcher groups

	Number of searches (% total number)	Number of active employees (% all active)	Lowest to highest searches per employee	Average searches per employee	Standard deviation
Highly frequent	115,112 (39.9)	479 (8.2)	148-878	240.3	108.2
Frequent	92,624 (32.1)	967 (16.5)	64-147	95.8	22.7
Infrequent	80,627 (28.0)	4408 (75.3)	1-63	18.3	16.6
All active employees	288,363 (100.0)	5,854 (100.0)	1-878	49.3	72.9

Search tasks – identified during 8 interviews

Work tasks	Search tasks	Expected information	Search strategy
1 Applying products	Find new application areas for products	Non-specific	Exploratory
	Find out how a potential application area is technically carried out, and what problems characterize the area	Non-specific	Exploratory
2 Researching products and technologies	Find info about specific technology	Non-specific	Exploratory
	Find background lab data for a specific sample	Non-specific	Exploratory
	Find information about previous research results	Non-specific	Exploratory
3 Supporting operations around products	Find documentation for patent disputes	Non-specific	Exploratory
	Find master list of quality standards	Specific	Exploratory
	Find sales report for product	Specific	Lookup
4 Administrating e.g. staff, customers, economy	Find rules for	Specific	Exploratory
	Find payment information for customer	Specific	Lookup

Search behavior and practices

- Exploratory search strategy

- Tracing search strategy – searching and browsing across sources and information types

- Historical, contextual information – need of historical, contextual information about relationships between enzymes, products, projects, business areas, people

- Metadata used when possible

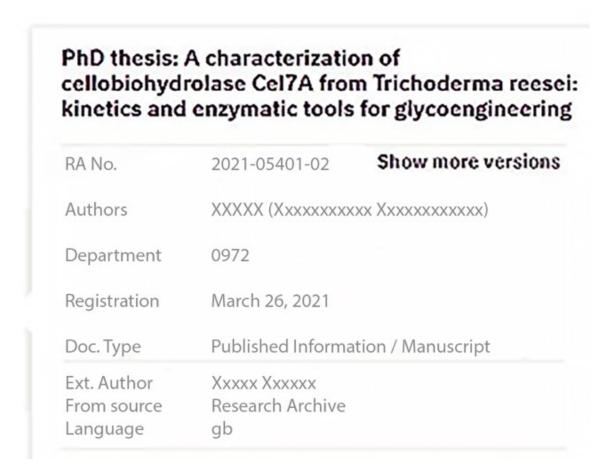- People search – 72 % of all searches

People search

- Contact information: to obtain contact information for a person.
- Documents: to find documents written by a person.
- Expertise: to get expertise knowledge
  - Professional knowledge about e.g. enzymes, technologies, products and application area
  - Organizational knowledge about previous projects, lab analyses, and research groups

- As part of tracing search strategy
  - Shortcut: to get help to identify a person that know about a certain topic
  - Steppingstone: to get help to decide how to move forward in the tracing search strategy

Role of metadata – provide context and relations

- Contextual metadata:
  - Person – as author, project leader, project member, developer, researcher
  - Enzyme, organism, and products
  - Project
  - Technology
  - Business area
  - Lab test
  - Research group

PhD thesis: A characterization of cellobiohydrolase Cel7A from Trichoderma reesei: kinetics and enzymatic tools for glycoengineering

RA No.	2021-05401-02	**Show more versions**
Authors	XXXXX (Xxxxxxxxxxx Xxxxxxxxxxx)	
Department	0972	
Registration	March 26, 2021	
Doc. Type	Published Information / Manuscript	
Ext. Author	Xxxxx Xxxxxx	
From source	Research Archive	
Language	gb	

ML metadata by SciBite

Lessons learnt: ML metadata and KOS

- Evaluation of ML metadata
  - Provides important contextual and historical information
  - Relevance depends on search task, most relevant for R&D
  - Increased visitibity of information sources
  - Short cut to relevant background information

  - Explanability – nature and use of ML metadata was not obvious
  - Use of ML metadata demands background knowledge
  - Mistrust to ML metadata

- Assignment of ML metadata by taxonomy
  - Combi of enriched commercial and company specific taxonomies
  - Categorization problems:
    - Person metadata difficult to assign
    - Difficulty in distinguishing between organism variants
    - Difficulty in distinguishing between company as collaborator or supplier

References

Lykke, M. (2022). Supporting exploratory enterprise search with an AI/ML-based semantic tool. I *Search Insights 2022* (s. 40-42)

Lykke, M., Bygholm, A. K. M., Søndergaard, L. B., & Byström, K. (2022). The role of historical and contextual knowledge in enterprise search. *Journal of Documentation*, *78*(5). Advance online publication. https://doi.org/10.1108/JD-08-2021-0170

Svarre, T., Lykke, M. & Bygholm A. K. M. (2024). Searching for people in the workplace: Aims, behavior, and challenges (Accepted, revision submitted)