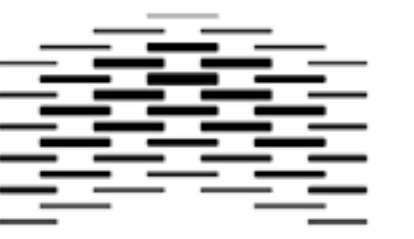




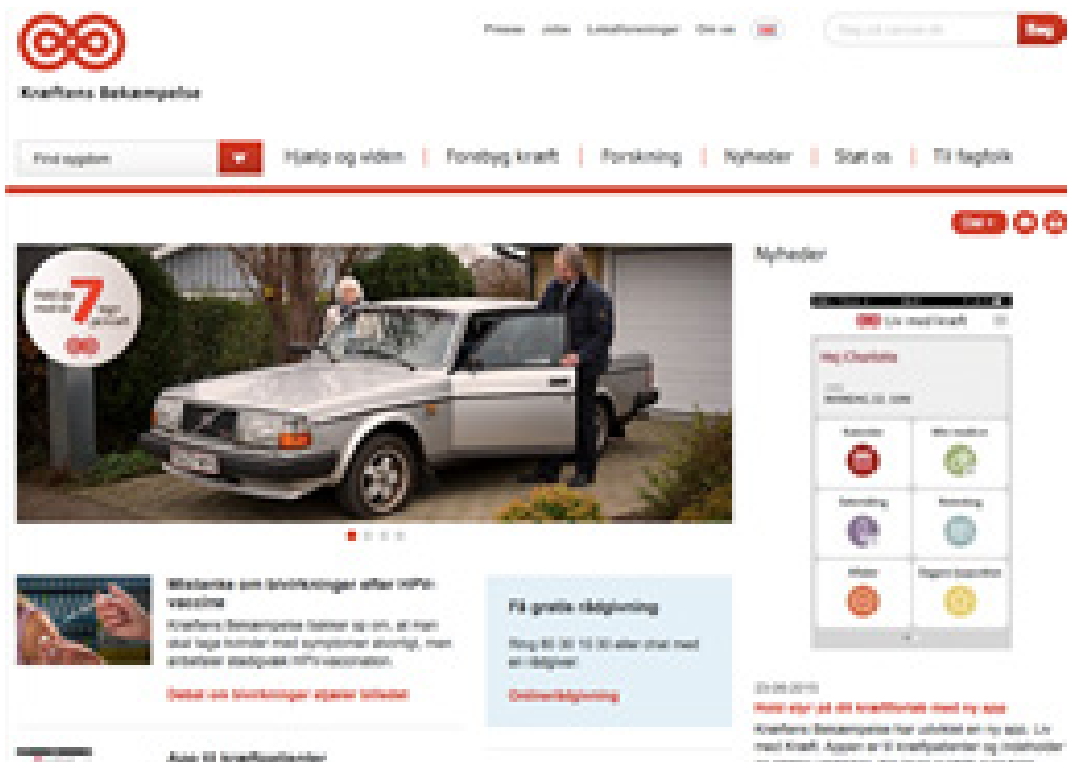
Web analytics as tool for improvement of website taxonomies



Taxonomy

- Hierarchical vocabularies for classification, browsing, and searching of content in networked environments.
- Support navigation by hierarchical organisation
- Operate as lead-in by synonym rings
- Guide the user by “see also” references to related categories in the hierarchy
- Allow the user to filter and narrow search results
- Guide the user by scope notes specifying the meaning of the vocabulary

(ISO/TC 46/SC 9, 2013).



Subject area and facets

- Events: “Denmark is biking”
- Recipes: “Celery beef”
- Social rights: “Care leave”
- Alternative treatment: “Traditional Chinese medicine”

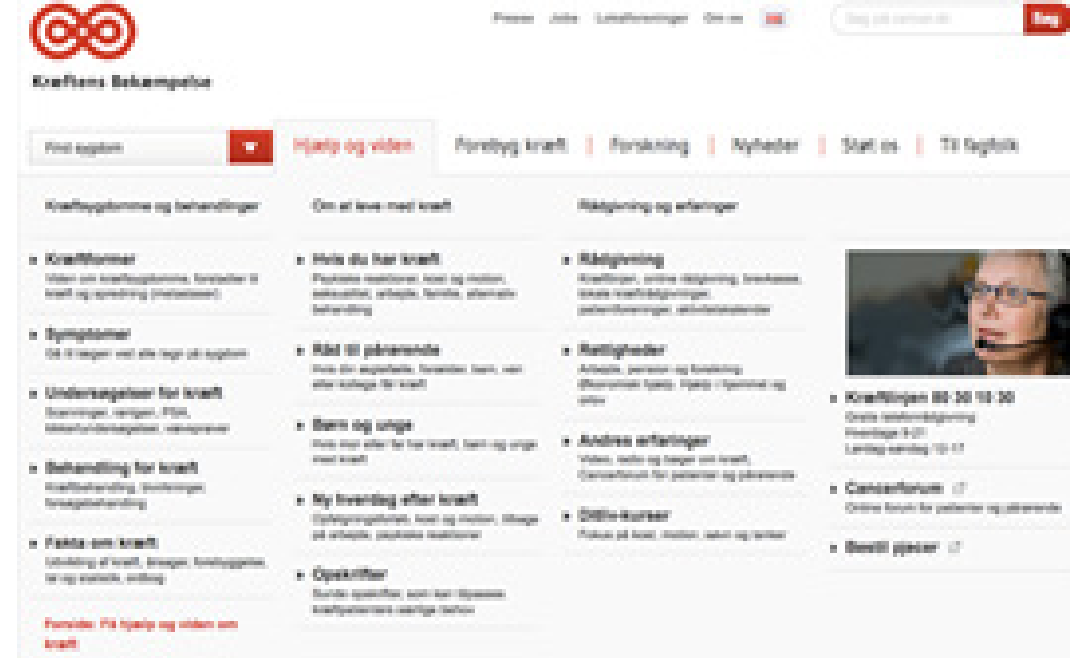
Vocabulary use

- Layman for symptom: “Too many white blood cells”
- Layman for disease: “Cancer in the tongue”
- Professional for symptom: “Helicobacter pylori”
- Professional for disease: “Non-Hodgkin”

Taxonomy design – challenges

- Viewpoint – facets and perspective
- Semantics – meaning and use
- Generics – specificity and multi hierarchy

(Holm & Rasmussen, 1961; Lykke Nielsen, 2008)



Google Analytics - Search terms

Search terms	Sessions	% new session!	New users	Pages per session
(not provided)	261437	51,69%	135138	2,89
kræftens bekæmpelse	168587	40,49%	68259	4,77
cancer.dk	23529	26,51%	6237	4,63
hudkræft	18546	54,56%	10118	2,46
lymfekræft	14447	50,85%	7347	2,61
hpv vaccine	10518	60,71%	6385	1,32
www.lotteri.dk	10485	51,09%	5357	2,67
modermærkekræft	9797	49,89%	4888	2,37
hele danmark cykler	7837	70,54%	5528	1,28
lungekræft	7068	42,88%	3031	3,21
testikelkræft	5960	62,30%	3713	1,85
leverkræft	5929	45,32%	2687	3,16
prostatakræft	5830	47,34%	2760	3,10
kræft	5720	41,92%	2398	4,31
tyktarmskræft	5701	45,03%	2567	3,02
brystkræft	5639	41,35%	2332	3,37
cancer	5598	27,46%	1537	4,33
www.lotteri.dk/vind	5577	45,74%	2551	3,42
rødbedesalat	5325	56,30%	2998	1,40
pet scanning	5253	46,53%	2444	1,96
celleforandringer	5231	59,64%	3120	3,28
dallund	5015	20,16%	1011	4,76
www.cancer.dk	5011	32,31%	1619	4,34
selleribof	4867	54,51%	2653	1,45
cancerforum	4798	12,98%	623	2,00
rødkålssalat	4788	60,19%	2882	1,25
kemoterapi	4717	39,81%	1878	3,35
endetarmskræft	4385	50,03%	2194	2,61
lotteri.dk	4377	52,36%	2292	2,45
blomkålssuppe	4310	57,31%	2470	1,29
kræftens bekæmpelse jo	4219	35,89%	1514	3,28
psa	4171	49,70%	2073	2,88
rosenkålssalat	3989	52,19%	2082	1,36
gullash	3827	60,47%	2314	1,28
bugspytkirtelkræft	3819	45,27%	1729	2,66
modermærkekræft billec	3733	38,79%	1448	2,06
livmoderhalskræft	3585	45,75%	1640	2,28
kræft i bugspytkirtlen	3569	43,43%	1550	3,04
plejeorlov	3523	46,24%	1629	4,12
syptomer på kræft	3513	54,48%	1914	3,59

Home-grown transaction log - User tags

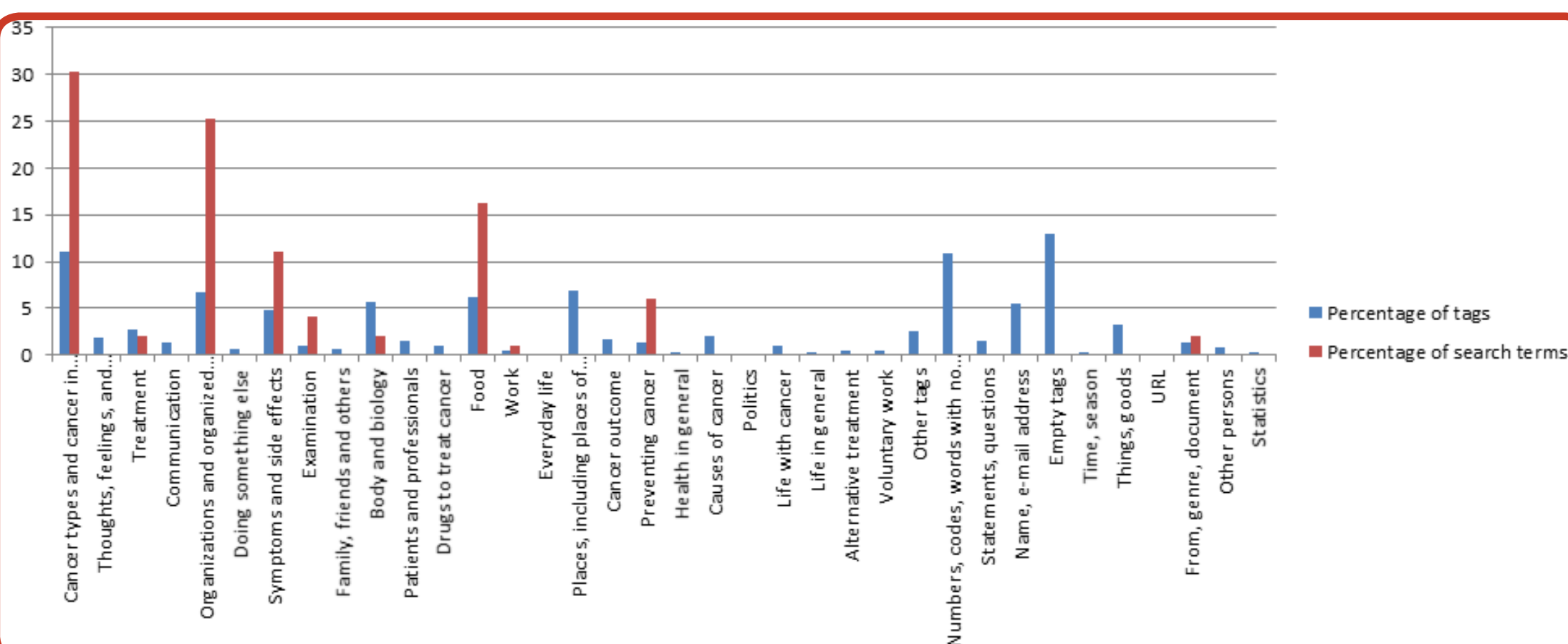
id	TimeStamp	UserIP	SessionId	Url	Tag	Internal/ekstern	Aboutness	Ab. Oppsum	Facet	Layman	Professional
1	08-11-2011 13:04	127.0.0.0.knspjhm0nt5r	/Nyheder/m	ProActiveTester		i	9	51	26	4	
2	08-11-2011 13:55	94.101.1.kmpif0ubmda	/forebyg/sol	sol		i	5	51	20	2	
3	08-11-2011 14:12	94.101.1.stpuus45tid5c	/Hjaelp+vide	celler		i	4	51	10	1	
4	08-11-2011 14:13	94.101.1.stpuus45tid5c	/Hjaelp+vide	forebyggelse		i	4	51	18	2	
5	08-11-2011 14:13	94.101.1.stpuus45tid5c	/Hjaelp+vide	livsstil		i	5	51	20	2	
6	08-11-2011 14:14	94.101.1.stpuus45tid5c	/Hjaelp+vide	årsager		i	18	52	20	2	
7	08-11-2011 14:15	94.101.1.stpuus45tid5c	/Hjaelp+vide	statistik		i	5	51	36	2	
8	08-11-2011 14:15	94.101.1.stpuus45tid5c	/Hjaelp+vide	kræftforskning		i	7	52	14	1	
9	08-11-2011 14:16	94.101.1.stpuus45tid5c	/Hjaelp+vide	tal		i	7	52	36	2	
10	08-11-2011 14:17	94.101.1.stpuus45tid5c	/Hjaelp+vide	kræftceller		i	5	51	1	1	
11	08-11-2011 14:17	94.101.1.stpuus45tid5c	/Hjaelp+vide	arvmateriale		i	5	51	10	2	
12	08-11-2011 14:17	94.101.1.stpuus45tid5c	/Hjaelp+vide	gener		i	5	51	10	2	
13	08-11-2011 14:18	94.101.1.stpuus45tid5c	/Hjaelp+vide	godartede knudi		i	5	51	1	1	
14	08-11-2011 14:19	94.101.1.stpuus45tid5c	/Hjaelp+vide	ondartede knudi		i	5	51	1	1	
15	08-11-2011 14:19	94.101.1.stpuus45tid5c	/Hjaelp+vide	spredning		i	5	51	1	1	
16	08-11-2011 14:19	94.101.1.stpuus45tid5c	/Hjaelp+vide	metastaser		i	5	51	1	3	
17	08-11-2011 14:20	94.101.1.stpuus45tid5c	/Hjaelp+vide	årsager		i	5	51	20	2	
18	08-11-2011 14:21	94.101.1.stpuus45tid5c	/Hjaelp+vide	forebyggelse		i	3	51	18	2	
19	08-11-2011 14:21	94.101.1.stpuus45tid5c	/Hjaelp+vide	råd		i	4	51	18	2	
20	08-11-2011 14:21	94.101.1.jkfc0j45qcllbd	/Hjaelp+vide	metastase		i	3	51	1	3	
21	08-11-2011 14:21	94.101.1.stpuus45tid5c	/Hjaelp+vide	europæiske kræ		i	3	51	5	1	
23	08-11-2011 14:22	94.101.1.stpuus45tid5c	/Hjaelp+vide	Den Europæiske		i	3	51	18	1	
24	08-11-2011 14:22	94.101.1.stpuus45tid5c	/Hjaelp+vide	samfund		i	5	51	26	2	
25	08-11-2011 14:22	94.101.1.355qmgqkm5	/Hjaelp+vide	nydiagnosticere		i	7	52	11	1	
26	08-11-2011 14:23	94.101.1.stpuus45tid5c	/Hjaelp+vide	risiko		i	5	51	18	2	
27	08-11-2011 14:23	94.101.1.jkfc0j45qcllbd	/Hjaelp+vide	behandlingssted		i	4	51	16	2	
28	08-11-2011 14:23	94.101.1.stpuus45tid5c	/Hjaelp+vide	undersøgelse		i	5	51	8	2	
29	08-11-2011 14:23	94.101.1.jkfc0j45qcllbd	/Hjaelp+vide	behandlingssted		i	4	51	16	2	
30	08-11-2011 14:23	94.101.1.jkfc0j45qcllbd	/Hjaelp+vide	afdeling		i	4	51	16	2	
31	08-11-2011 14:25	94.101.1.stpuus45tid5c	/Hjaelp+vide	konkrete råd		i	5	51	2	2	
32	08-11-2011 14:26	94.101.1.stpuus45tid5c	/Hjaelp+vide	nedsatte risiko		i	4	51	18	2	
33	08-11-2011 14:26	94.101.1.stpuus45tid5c	/Hjaelp+vide	forebygge		i	3	51	18	2	
34	08-11-2011 14:26	94.101.1.stpuus45tid5c	/Hjaelp+vide	røgfri		i	5	51	18	2	
35	08-11-2011 14:26	94.101.1.stpuus45tid5c	/Hjaelp+vide	overvægt		i	5	51	20	2	
36	08-11-2011 14:27	94.101.1.stpuus45tid5c	/Hjaelp+vide	bevægelse		i	5	51	18	2	
37	08-11-2011 14:27	94.101.1.stpuus45tid5c	/Hjaelp+vide	motion		i	5	51	18	2	
38	08-11-2011 14:28	94.101.1.stpuus45tid5c	/Hjaelp+vide	spis sundt		i	5	51	18	2	
39	08-11-2011 14:28	94.101.1.stpuus45tid5c	/Hjaelp+vide	fuldkorn		i	5	51	13	2	
40	08-11-2011 14:28	94.101.1.355qmgqkm5	/Hjaelp+vide	nydiagnosticere		i	18	52	11	1	

Synonym variations

- “Symptom”, “sign for cancer”
- “Pancreas cancer”, “kræft i bugspytkirtlen” (Danish synonym)
- “Basocellular carcinoma”, “Skin cancer”
- “Hope”, “Optimism”, “Positive thoughts”

Level of specificity

- “Metastases”, “Bone metastasis in breast cancer”
- “Treatment”, “Life-sustaining treatment”
- “Mouth”, “Tongue”
- “Prevention”, “Sunscreen”



Take aways

– transaction logs provide knowledge about:

- User’s interests and subjects
- Importance of subjects (facets)
- User’s special languages
- User level of specificity
- User tags provide a broader picture and more nuances compared to search terms

Literature:

Holm, B.E. & Rasmussen, L.E. (1961). Development of a technical thesaurus. American Documentation, July 1961.

ISO Working group 8. International Organization for Standardization: ISO/TC 46/SC 9 (2013). Information and documentation- Thesauri and interoperability with other vocabularies: Part 2: Thesauri for information retrieval. International Organization for Standardization: ISO/TC 46/SC 9.

Lykke Nielsen, M. (2008). Special libraries and specialized vocabularies in the digital age. Bates, M. J. & Maack, M. D. (eds.). Encyclopedia of Library and Information Science, vol 1/1.