

Searching in a Public Library

Some Experiences with the Search Behaviour of Patrons of a Public Library

At the public library in Waalre (NL) there is a dedicated catalogue system in operation now for 2 years. All search actions are being logged and it is interesting to analyse these in order to learn how anonymous users of the system behave.

The results of this analysis have been used to extend the functionality of the system with the aim to eventually being developed to a one stop information resource for patrons (and other users) of the catalogue system.

Many of the concepts thus developed can be "translated" to Wikipedia (and possibly other wiki's as well). The added value of this work is that it is based on the behaviour of all kind of users, from kids to seniors, from (computer)dummies to experienced information seekers.

Topics

- Introduction / Background
- SmartSearch
 - A tool assisting the user to find what he meant
 - Thesaurus
- Feedback from logs
 - To improve composition of holdings
- Helping the user
 - Integrating various sources of information
 - In finding what he intuitively is searching
- Conclusions
 - Transfer concepts to Wikipedia

Public Library Functions

- Holdings
 - Books / Magazines (and other items: CD's etc.)
 - On loan
 - To be consulted on site
- Inter Library Loans
- Source for Information
 - Printed (books / magazines)
 - Electronic (dedicated databases, internet)
 - Librarian as intermediate

The Public Library in Waalre (NL)

Some Data

- Community population: 17.000
- Patrons: 5.000
- Holdings: 40.000
- Loans: 120.000 per annum
- Financing:

Municipality	€ 260.000 (70%)
Patrons	€ 100.000 (30%)

Similarities and Differences between a Library Catalogue and an Encyclopaedia

- Similarities
 - End-users (not editors) are very similar
 - End-users are looking for information
- Differences
 - Single index - Multiple Indexes
 - title, author, keyword, etc.

Normally an encyclopaedia is searched on the title of the lemma.

In a library catalogue there are multiple indexes.

This has only consequences for the technical implementation, not for the concepts as presented here.

Design / History of the (new) Library's Catalogue

- Front-end to existing commercial library system
 - Developed from 1-1-2003 onwards
 - Commissioning first version 1-7-2003
- Prerequisites
 - Simplicity
 - Intuitive use
 - Performance
 - Log everything in order to learn
(about use, future needs, etc.)
- Ultimate goal
 - Be the premier search tool for the patrons
 - Weekly searches have grown from 400 to >1.600

The simplicity of Google's home page has inspired the design. There are numerous examples of commercial catalogue systems, which are too complicated for an ordinary user.

The development has been done on an old server. That forces the development process to continuously search for smart solutions, in order to provide good performance. Thus a built in warranty for scalability is realised.

Analysis of Search Logs

DT	IP	QS	Q	Qnew	RsltNu	SrchT	Smart	Bro
1-8-2005 16:06:27	192.168.0.2	cattype=cat&xct=y&pglen=25&srt=	film		33	312	343	IE5
1-8-2005 16:06:37	192.168.0.2	cattype=cat&xct=y&pglen=25&srt=	script		4	281	312	IE5
1-8-2005 16:06:49	192.168.0.2	cattype=cat&xct=y&pglen=25&srt=	schrijven		225	953	984	IE5
1-8-2005 16:06:58	192.168.0.2	cattype=cat&xct=y&pglen=25&srt=	scriptwriter		0	250	1140	IE5
1-8-2005 16:07:04	192.168.0.2	cattype=cat&xct=y&pglen=25&srt=	scriptschrijven		0	265	952	IE5
1-8-2005 16:07:10	192.168.0.2	cattype=cat&xct=y&pglen=25&srt=	script schrijven	script schrijven	18	265	996	IE5
1-8-2005 16:07:54	192.168.0.2	cattype=cat&xct=y&pglen=25&srt=	script schrijven	script schrijven	18	265	1307	IE5
1-8-2005 16:08:00	192.168.0.2	cattype=cat&xct=y&pglen=25&srt=	script schrijven	script schrijven	18	265	1618	IE5
1-8-2005 16:08:18	192.168.0.2	cattype=cat&xct=y&pglen=25&srt=	script schrijven	script schrijven	18	265	1961	IE5
1-8-2005 16:08:23	192.168.0.2	cattype=cat&xct=y&pglen=25&srt=	script schrijven	script schrijven	18	265	2288	IE5
1-8-2005 16:08:29	192.168.0.2	cattype=cat&xct=y&pglen=25&srt=	script schrijven	script schrijven	18	265	2615	IE5
1-8-2005 16:08:40	192.168.0.2	cattype=cat&xct=y&pglen=25&srt=	script schrijven	script schrijven	18	265	2942	IE5
1-8-2005 16:09:49	192.168.0.2	cattype=cat&pglen=25&srt=rlv&sub	wereld magie		3	218	249	IE5
1-8-2005 16:09:57	192.168.0.2	cattype=cat&pglen=25&srt=rlv&sub	magie		37	31	46	IE5
1-8-2005 16:10:21	192.168.0.2	cattype=cat&pglen=25&srt=rlv&sub	wet magie		9	140	155	IE5
1-8-2005 16:10:43	192.168.0.2	cattype=cat&pglen=25&srt=rlv&sub	wet magie		9	140	155	IE5
1-8-2005 16:10:52	192.168.0.2	cattype=cat&pglen=25&srt=rlv&sub	wet magie		9	140	155	IE5
1-8-2005 16:10:54	192.168.0.11	cattype=cat&xct=y&pglen=25&srt=	scriptschrijven		0	265	843	IE5
1-8-2005 16:10:55	192.168.0.2	cattype=cat&xct=y&pglen=25&srt=	scriptwriter		0	250	1000	IE5
1-8-2005 16:10:55	192.168.0.2	cattype=cat&xct=y&pglen=25&srt=	schrijven		225	953	984	IE5
1-8-2005 16:10:55	192.168.0.2	cattype=cat&xct=y&pglen=25&srt=	script		4	281	312	IE5
1-8-2005 16:10:56	192.168.0.2	cattype=cat&xct=y&pglen=25&srt=	film		33	312	343	IE5
1-8-2005 16:11:01	192.168.0.2	cattype=cat&pglen=25&srt=rlv&sub	wereld magie		3	218	233	IE5
1-8-2005 16:11:37	192.168.0.2	cattype=cat&pglen=25&srt=rlv&sub	filmen	film	87	250	530	IE5
1-8-2005 16:11:47	192.168.0.2	cattype=cat&pglen=25&srt=rlv&sub	filmen	film	--	--	--	--
1-8-2005 16:11:58	192.168.0.2	cattype=cat&pglen=25&srt=rlv&sub	zelf filmen	zelf film				
1-8-2005 16:12:25	192.168.0.2	cattype=cat&pglen=25&srt=rlv&sub	homevideo	reuze video				
1-8-2005 16:12:35	192.168.0.2	cattype=cat&pglen=25&srt=rlv&sub	video					
1-8-2005 16:12:43	192.168.0.2	cattype=cat&pglen=25&srt=rlv&sub	film maken	film maken				
1-8-2005 16:12:51	192.168.0.2	cattype=cat&pglen=25&srt=rlv&sub	film maken	film maken				

Thousands of these records, in particular the query phrase, have been browsed to understand the behaviour of the users.

Mistakes by Users

- Too many words (each word needs to match one of the indexes)
 - Add more, rather than delete
- Wrong index being searched
- Ordinary misspellings
- Plural / Single nouns
- Combine / Split words

Solution: SmartSearch

- Attempts to correct (or improve) the search string typed in by the user
1. Proactive in case of null results
 - Possible improvements are automatic
 2. Passive in case of results returned
 - Possible extended search is suggested

SmartSearch: Result

Search the Type your search word(s) and press Enter

Library Catalog mulisch

Syner-IM B.V. Visually impaired Help Search Tips Go Advanced

Indexes to search

Title Author(s) Keyword - list

Searched for: mulisch in Title, Author Results 1 - 22 of 22

SmartSearch® adaptations: Original query: Search for mualis in Title, Author mualis (does not exist) replaced by mulisch

Auto Modifications Click for detailed information

Title	Author(s)	Availability	Loc	Rating
1. De aanslag : roman	Mulisch, Harry	✓	WAA V	■■■■■
2. De aanslag : roman	Mulisch, Harry	✗	AAL V	■■■■■
3. De elementen	Mulisch, Harry	✗	AAL V	■■■■■
4. De ontdekking van de hemel : roman	Mulisch, Harry	✗	AAL V	■■■■■
5. De ontdekking van de hemel : roman	Mulisch, Harry	✗	AAL V	■■■■■
6. De procedure : roman	Mulisch, Harry	✗	AAL V	■■■■■
7. De procedure : roman	Mulisch, Harry	✗	AAL V	■■■■■
8. De verhalen	Mulisch, Harry	✗	AAL V	■■■■■
9. Het stenen bruidsbed : roman	Mulisch, Harry	✗	AAL V	■■■■■
10. Het theater, de brief en de waarheid : een tegenspraak	Mulisch, Harry	✗	AAL V	■■■■■
11. Het theater, de brief en de waarheid : een tegenspraak	Mulisch, Harry	✗	AAL V	■■■■■
12. Het zwarte licht : kleine roman	Mulisch, Harry	✗	AAL V	■■■■■

Search Tip

Here again thinking about performance is important.

The Auto Modifications are executed immediately, because there would be a null result otherwise.

The Search Tip appears some time after the page is displayed, as this is processed in the background.

SmartSearch: How it works

- Plural / Single nouns
 - Default
 - Search on beginning of words
 - **house** will also find **houses**, but **houses** will not find **house**
 - SmartSearch
 - If **houses** typed in AND single form (remove last s) is an existing word, than replace houses by **house**
 - Special case: language dependent exceptions, e.g.
 - huis – huizen (Dutch)
 - haus – häuser (German)
 - duif – duiven (Dutch)

SmartSearch: How it works (2)

- Ordinary misspellings (example muullis)
 - One character
 - m*uullis
 - mu*llis, etc
 - Two characters
 - m*llis
 - mu*iss, etc
- Join/split words
 - Example: science fiction - sciencefiction

An asterisk means that this could be any number of characters.

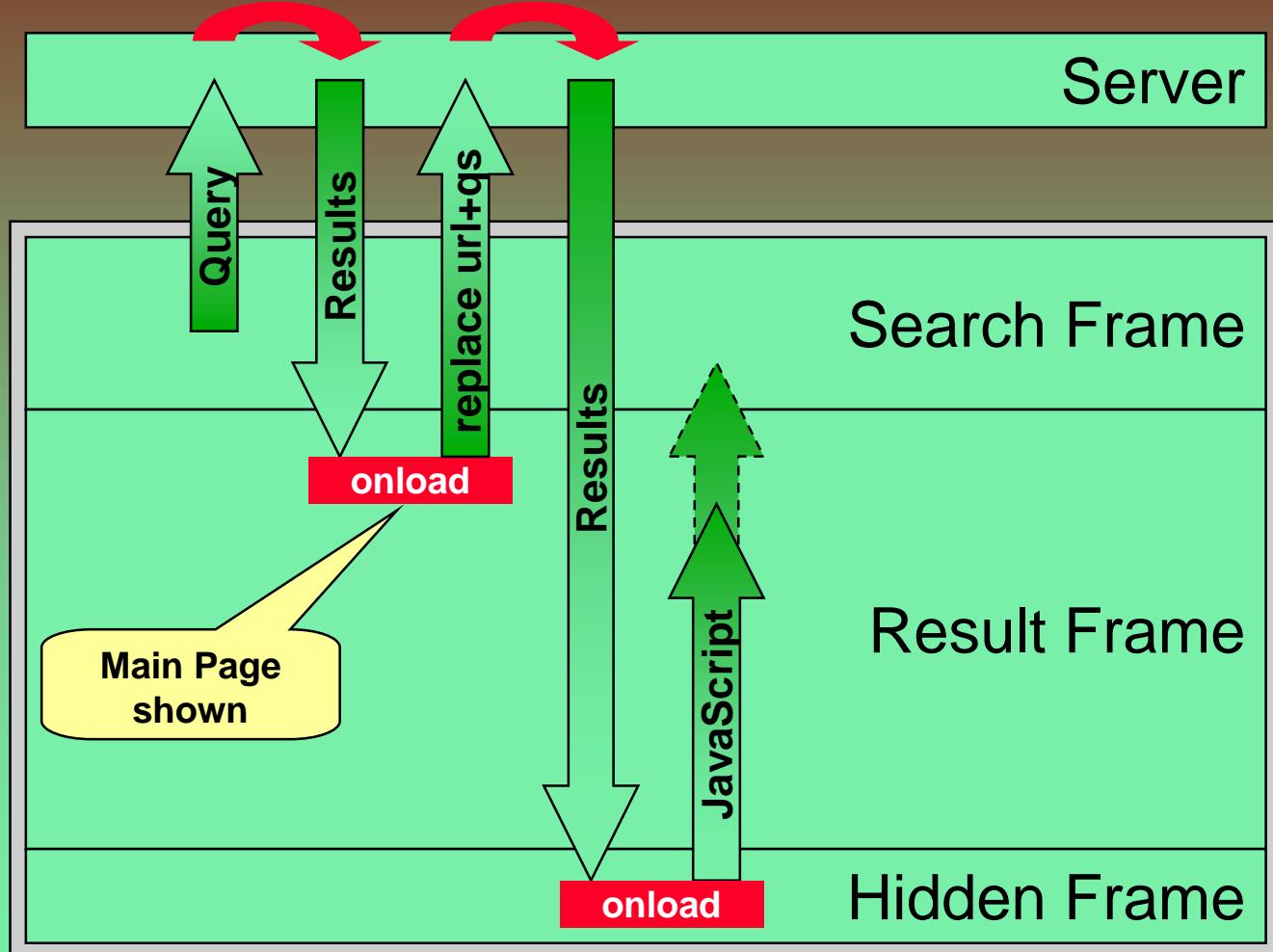
SmartSearch: Performance

- Display raw results (if any) immediately
- Restrict number of SQL queries on Server
 - Escape from SmartSearch procedure if result/improvement found
 - By using UNION queries where possible

SmartSearch follows a sequence to analyse the query phrase. As soon as an improvement is found that sequence is stopped and results become visible.

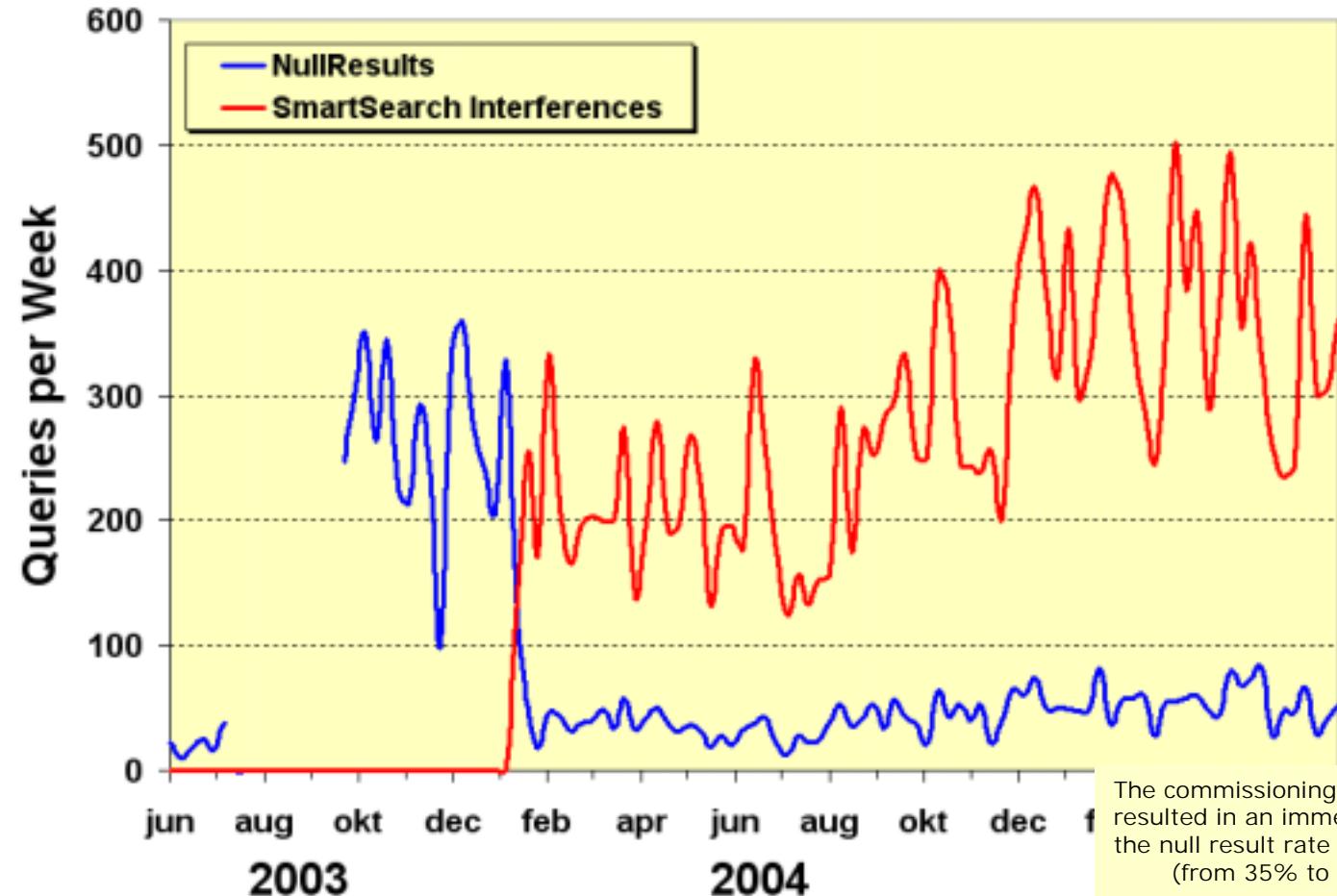
In particular the SQL query, investigating the occurrence of words similar to the (possibly misspelled) word(s) as typed in, should be constructed as an UNION SQL statement, in order to keep performance within reasonable limits.

Background Processing



The Result

Percentage of null results dropped from 30-35% to well below 5%



The commissioning of SmartSearch resulted in an immediate drop of the null result rate (from 35% to < 5%).

Both graphs follow the trend of increasing use in terms of searches per week.

Searching on Keywords

- Keywords to "Title Descriptions" are not consistent
- Solution:
 - Dedicated thesaurus
 - Derived from standard MS-Word thesaurus
 - Only words and synonyms that are present in "all words" table

Title Descriptions (or Bibliographic Records) are produced centrally for the Netherlands.

If one analyses all keywords (attached to the bibliographic record, if it concerns an informative book), these appear to be non consistent. This is the main reason why a thesaurus has been added.

Dedicated Thesaurus

Word	Meaning	Alias	PoS	Syn	StrType
Acteren	toneelspelen	toneelspelen	3	<input checked="" type="checkbox"/>	15
Acteren	doen alsof	spelen	3	<input checked="" type="checkbox"/>	15
Acteren	doen alsof	spelen	3	<input checked="" type="checkbox"/>	11
acteren	acteren	Spelen	3	<input checked="" type="checkbox"/>	11
acteren	acteren	Spelen	3	<input checked="" type="checkbox"/>	15
acteren	acteren	Toneelspelen	3	<input checked="" type="checkbox"/>	15
acteren	spelen	Toneelspelen	3	<input checked="" type="checkbox"/>	15
acteur	figurant	figurant	1	<input checked="" type="checkbox"/>	11
acteur	figurant	speler	1	<input checked="" type="checkbox"/>	11
acteur	figurant	toneelspeler	1	<input checked="" type="checkbox"/>	11
acteur	speler	actrice	1	<input checked="" type="checkbox"/>	11
acteur	acteur	figurant	1	<input checked="" type="checkbox"/>	11
acteur	acteur	speler	1	<input checked="" type="checkbox"/>	11
acteur	acteur	toneelspeler	1	<input checked="" type="checkbox"/>	11
acteur	speler	toneelspeler	1	<input checked="" type="checkbox"/>	11
Actie	aandeel	aandeel	1	<input checked="" type="checkbox"/>	11
Actie	aandeel	effect	1	<input checked="" type="checkbox"/>	11
Actie	campagne	campagne	1	<input checked="" type="checkbox"/>	11
Actie	campagne	strijd	1	<input checked="" type="checkbox"/>	11
Actie	daad	daad	1	<input checked="" type="checkbox"/>	11
Actie	expeditie	expeditie	1	<input checked="" type="checkbox"/>	11
Actie	handeling	daad	1	<input checked="" type="checkbox"/>	11
Actie	handeling	optreden	1	<input checked="" type="checkbox"/>	1
Actie	stap	stap	1	<input checked="" type="checkbox"/>	1
Actie	stap	stap	1	<input checked="" type="checkbox"/>	1
actie	actie	Stap	1	<input checked="" type="checkbox"/>	1
actie	actie	voortgang	1	<input checked="" type="checkbox"/>	1

This automatically generated thesaurus, consists of > 50.000 (valid/meaningful) records (all words in the column 'Alias' do have a reference to an existing keyword in one of the (40.000) bibliographic records.

Feedback to Librarians

- List of null results
 - Seamless link to Google to understand meaning of entered word(s)
- List of SmartSearch interferences
 - Derive obvious misspellings
 - Can be used for dedicated entries in (automatically) working thesaurus

Null-Results after SmartSearch

Catalogus - Analyse Zoekresultaten				
Nulresultaten				
Toon:		Nul-Resultaten	vanaf:	1-4-2005
#	Datum Zoekopdracht	Oorspronkelijk	Zoekterm	Aangepast
	#	Laatste	Eerste	
13	9-4-2005	2-4-2004	bint	
11	4-7-2005	2-4-2005	davinci code	
10	24-5-2005	28-5-2004	php	
6	29-7-2005	8-3-2004	kertesz	
6	8-6-2005	13-12-2004	stravaganza	
5	1-7-2005	8-1-2005	preston&child	
5	28-6-2005	23-3-2005	korenwolf	
5	28-6-2005	12-1-2005	tsunami	
4	18-5-2005	2-6-2004	fototoestel	
4	14-5-2005	9-3-2005	kaapverdie	
4	6-5-2005	24-4-2004	\	
3	29-7-2005	23-5-2005	dalh	
3	25-7-2005	14-6-2004	zeezeiler	
3	12-7-2005	24-1-2005	spiegelologie	
3	8-7-2005	23-2-2005	wondergympen	
3	4-7-2005	1-6-2005	code	
3	31-5-2005	25-8-2004	indridason	
3	28-5-2005	15-1-2005	arendsoog	

This is a list of null results.
The librarian can click on the original search phrase, resulting in a Google search. In this way the librarian can find out the possible meaning.
PHP is a good example, a librarian will most likely not know the meaning of this acronym, but may find out very quickly.

Interferences by SmartSearch

Catalogus - Analyse Zoekresultaten

SmartSearch Aanpassingen

Toon: vanaf: sorteer op:

#	Datum Zoekopdracht		Zoekterm	
	Laatste	Eerste	Oorspronkelijk	Aangepast
16	29-7-2005	29-3-2004	rooyen	royen
14	27-7-2005	5-11-2004	hoe over leef ik	hoe overleef ik
11	15-7-2005	19-4-2004	mathilda	mathilde
11	15-7-2005	10-5-2004	turksfruit	turks fruit
10	29-6-2005	25-5-2004	onpersoonlijkheid	persoonlijkheid
9	7-6-2005	19-4-2004	gossipgirl	gossip girl
9	1-6-2005	17-5-2004	romulus mijn vader	mijn vader
8	23-6-2005	30-6-2004	elfquest	question
8	14-5-2005	22-4-2004	starwars	starters
7	18-7-2005	9-7-2004	dolfjeweerwolfje	dolfje weerwolfje
6	27-7-2005	29-3-2004	heleen rooyen	heleen royen
6	23-7-2005	22-5-2005	fernhoot	turnhout
6	22-7-2005	15-7-2004	mcgraw	macgraw
6	27-6-2005	24-5-2004	mc donald	mcdonald
6	8-6-2005	13-7-2004	charmed	charme
6	3-6-2005	21-9-2004	trimbaan	trimbach
6	2-6-2005	25-10-2004	capitool	capitals
6	28-5-2005	27-2-2004	verotte leven floortje bloem	verrotte leven floortje bloem

Provide User with alternative Information Sources

Search the Type your search word(s) and press Enter

Library Catalog

Indexes to search

Title

Author(s)

Keyword - list

© Syner-IM B.V. Visually impaired Help Search Tips Go

Searched for: **amsterdam** in Title, Author Results 1 - 10 of 136

Title	Click for detailed information	Author(s)	Availability	Loc	Rating
1. Amsterdam		Baeten, Marja	✓	AAL J	
2. Amsterdam		Weidemann, Siggi	✓	AAL V	
3. Amsterdam		Baeten, Marja	✓	WAA J	
4. Amsterdam		MacEwan, Ian	✗	AAL V	
5. Amsterdam in de Gouden Eeuw		Emeis, M.G.	✓	AAL J	
6. Amsterdam in de Gouden Eeuw		Emeis, M.G.	✓	WAA J	
7. Amsterdam Zuidoost		Osch, Yvonne van	✓	AAL J	
8. Amste	Zoek in	Osch, Yvonne van	✓	WAA J	
9. Maleis	Wikipedia met: unei	Asterisk*, Amsterdam	✓	AAL V	
10. Boeda	Amsterdam Amsterdam-Centrum Amsterdam-eiland Amsterdam-Noord Amsterdam-Rijnkanaal Amsterdam-Zuidoost ... meer ...	Limburg, Inger Asterisk*, Amsterdam	✗	AAL V	

1 2 3 4 5 6 7 8 9 10 11 12 13 14 Next

Zoek met **amsterdam** ook in

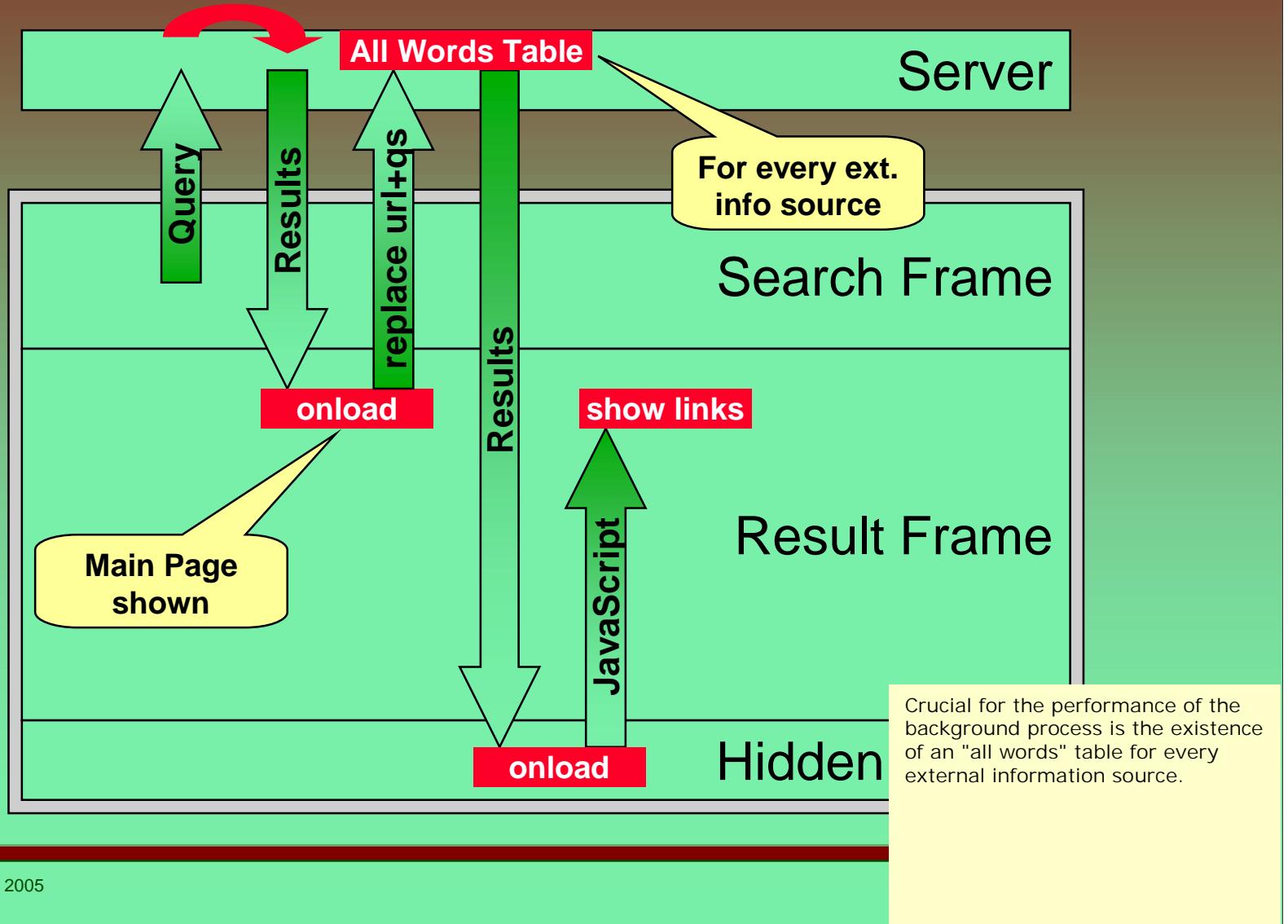
obwaalre.nl bibliothek.nl dmoz.org

startpagina.nl

Lower bar with buttons appears after background processing. It shows only buttons of selected sites and/or Wikipedia, if that information resource has indeed an entry with in this case Amsterdam in either title or text.

In the case of Wikipedia all articles with Amsterdam somewhere in the title are shown.

How it works



"All Words" Table

- Where clauses in SQL statements
 - Like* clauses are "expensive"

	cur_id	cur_title	cycle
109	90	Algemene_Centrale_Der_Liberale_Vakbonden_Van_België	0
110	90	Centrale_Der_Liberale_Vakbonden_Van_België	1
111	90	Der_Liberale_Vakbonden_Van_België	2
112	90	Liberale_Vakbonden_Van_België	3
113	90	Vakbonden_Van_België	4
114	92	Aboe_Nidal	0
115	92	Nidal	1
116	93	Arabische_Liga	0
117	93	Liga	1
118	94	Afrikaans_Nationaal_Congres	0
119	94	Nationaal_Congres	1
120	94	Congres	2
121	95	Anti-Revolutionaire_Partij	0
122	95	Partij	1
123	96	Afro-Aziatische_talen	0
124	96	talen	
125	97	Anatolische_talen	
126	97	talen	
127	98	Archosauria	

A (Wikipedia) table as shown here, significantly contributes to the performance of the background process.

Between ... And ...z SQL clauses are much more efficient than ordinary Like clauses in the SQL statement.

If the User is not really knowing what to look for

- The patron's dilemma
 - I've read a book by, did like it (or dislike)
 - I'm looking for something similar
- What do you suggest?

Simple Cross-references

similar to e.g. Amazon buying suggestions

Search the Type your search word(s) and press Enter

Library Catalog Advanced

Syner-IM B.V. Visually impaired Help Search Tips Go Default values

Indexes to search Title Author(s) Keyword - list

Results 1 - 10 of 22

Crossed for mulisch in Title, Author

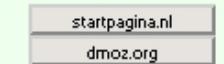
Alternatieve Auteurs Help X

Leners van boeken van **Mulisch, Harry** hebben ook boeken van de volgende auteurs geleend.

Auteur	Aantallen			Relevante
	Leners	Bezit	Uitde-	
Dorresteijn, Renate	18	21	888	100%
Steel, Danielle	28	92	3.889	80%
Andrews, Virginia	11	83	2.586	79%
Winter, Leon de	12	16	465	59%
Rendell, Ruth	15	35	1.223	47%
Enquist, Anna	10	9	408	46%
Wolkers, Jan	9	14	421	39%
Gemmell, David	14	11	224	38%
Eduardo, Mariano	5	21	1.101	29%

Next

Zoek met **mulisch** ook in

 obwaalre.nl
 obeindhoven.nl
 startpagina.nl
 dmoz.org

The pop-up screen is shown when the user clicks on an yellow square.

Complex Cross-references

Uw Leesvoorker(en) **Suggesties voor** **Update Methode**

Ontspannend Auteurs Automatisch
 Spannend Titels Handmatig
 Romans

LeenHistorie (118 titels)			
	Titel	Auteur	Geleend op
<input checked="" type="checkbox"/>	Het Venetiaanse huis	Nickson, Mary	26-7-2005
<input checked="" type="checkbox"/>	Onderstromen	MacKinley, Tamara	12-7-2005
<input checked="" type="checkbox"/>	Dagboek van een bezeten bruid	Wolf, Laura	6-5-2005
<input checked="" type="checkbox"/>	Dagboek van een maffe moeder-in-spe	Wolf, Laura	2-5-2005
<input checked="" type="checkbox"/>	Samen ben je minder alleen	Gavalda, Anna	22-4-2005
<input checked="" type="checkbox"/>	Leren zwemmen : roman	Chambers, Clare	4-4-2005
<input checked="" type="checkbox"/>	Vang me als ik val	French, Nicci	15-3-2005
<input checked="" type="checkbox"/>	De vogelkooi	Willett, Marcia	4-2-2005
<input checked="" type="checkbox"/>	Het verwende nest	Sambeek, Liza van	21-1-2005

LeenAdvies (71 titels)				
	Titel	Beweeg de muis naar een titel voor een samenvatting		
		Auteur	Relevante	Aanwezig
<input type="checkbox"/>	Het kersthuis	Wiggin, Helene	100%	<input checked="" type="checkbox"/>
<input type="checkbox"/>	Erfenis in november	Maljers, Margreet	90%	<input checked="" type="checkbox"/>
<input type="checkbox"/>	Bevrijd	Ster-Bouwmeester		
<input type="checkbox"/>	Kostbaar kleinood	Leni van der Wageningen, Gerd van		
<input type="checkbox"/>	Vertrouw op mij	Woolley, Lucy		

The upper pane shows the loan history of an user.

The lower pane shows the suggestions, taking into account the loan history of all patrons, but the sorting is heavily influenced by the user's own loan history.

The ultimate Goal of a free Source of Knowledge

- Current wiki projects are predominantly focusing on **active users** (contributors)
- Next phase:
 - Wikipedia to be a natural info source for **passive users**
 - Passive users will behave very similar as library patrons

Will make all kind of mistakes

Have limited skills to define proper search strings

Conclusions and Recommendations (1)

- SmartSearch
 - Concepts can equally be used in e.g. **Wikipedia**
 - Algorithms for Dutch need to be converted to other languages
 - Requires another mechanism to start a new article
- Feedback to Librarians
 - Can be used as input to "**Requested Articles**"
 - Can be used to build meaningful **Redirects**

Conclusions and Recommendations (2)

- Integration of other Information Resources with Wikipedia
 - Could be made 'on the fly' if an 'all words' (in titles) table in the on-line database could be **queried**
- Cross-references (far future?)
 - Can be used to provide suggestions for "**further reading**"

Catalogue on the Internet

www.obwaalre.nl

www.obwaalre.nl/cat/default.asp?lang=en

Home page and catalogue.
Most of the pages can be displayed
in English.

Contact author on:
w: nl:gebruiker:RonaldB