

uogin

eric sieverts

LexisNexis®

Google

bing

YAHOO!

OvidSP

Onder de motorkap:
technieken van zoeksystemen

het zoekproces

1. zoeksystemen

- hoe ze werken
- hun basis-functionaliteit

2. zoekmethode

- opbouw van eenvoudige zoekacties

het zoekproces

vrijwel alle zoeksystemen zijn gebaseerd op dezelfde basistechniek en bieden in principe gelijksoortige basisfuncties

- web-zoekmachines
- betaalde zoeksystemen (*dialog, stn, lexisnexis, ebSCO, ovid, csa, ...*)
- web-databases
- bibliotheekcatalogi

waarop de nadruk ligt, hoe je ze "bedient" en hoe ze de resultaten presenteren, kan soms sterk verschillen

soorten acties



	<i>silverplatter zoekstelsel</i>	<i>webzoekmachines</i>
zoeken	perception <input type="button" value="Search"/>	perception <input type="button" value="Search"/>
in index bladeren	percep <input type="button" value="Go To Term"/>	<i>niet</i>
Booleaans combineren van termen of van eerdere resultaten	perception and time perception or perceived perception not visual (#1 or #2) and #3 not #4	perception AND time perception OR perceived perception -visual <i>niet</i>
getrunceerd zoeken	percept * organi?ation	percept * <i>(alleen Exalead)</i> <i>niet</i>
op specifiek veld zoeken	perception in ti peereboom * in au	intitle: perception <i>(bijv. google, yahoo)</i> <i>niet</i>
nabijheidszoeken (maximum woordafstand, in zelfde zin, ...)	time adj perception time with perception time near perception	"time perception" <i>niet</i> time NEAR perception <i>(yahoo, exalead)</i>
zoekresultaat tonen in ander format	<i>(automatisch)</i> <input type="button" value="Change Display"/>	<i>(automatisch)</i> <i>niet</i>
zoekgeschiedenis	<input type="button" value="Search History"/>	<i>niet</i> <i>(google history wel een beetje)</i>

waarin zoekt zoekstelsysteem?

in (bijna) alle digitaal beschikbare tekst

- bij webzoekmachine in alle zichtbare
(en soms in deel van onzichtbare)
tekst van webpagina's
structuur? <title> / link tekst / url / ...
- bij bibliografisch bestand in alle
onderdelen
(met veldstructuur) >> voorbeeld

01064258 INSPEC Abstract Number: A77049149

Title: Summer weather in south-east England, 54 BC

Author(s): Meaden, G.T.

Author Affiliation: Tornado & Storm Res. Organisation, Trowbridge, UK

Journal: Weather vol.32, no.1 p.33-5

Publication Date: Jan. 1977

CODEN: WTHRAL ISSN: 0043-1656

Language: English

Abstract: The meteorological aspects of Julius Caesar's expedition to Britain in 54 BC are discussed. It is shown that the invasion was delayed by a persistent north-west wind, and did not take place until about July 6. The ships reached land near Deal, where they were badly damaged by a storm on the night of 8-9 July. The evacuation, in mid-September, was also hampered by the continuing northwest winds. The occurrence of these northwest winds, and of a drought in Gaul, suggests a prolonged northward extension of the Azores high, with the main anticyclone lying west of the British Isles. (1 Refs)

Descriptors: atmospheric pressure and density; history; meteorology; rain; storms; wind

Identifiers: 54 BC; storm; drought; Gaul; Azores high; SE. England summer weather; July to September 54 BC; Julius Caesar; NW. winds; N. Atlantic anticyclone

Class Codes: A9260G (Winds and their effects); A9260Q (Storms); A9330G (Europe)

titel
auteurs
werkadres
bron
datum
tijdschr
taal
abstract

trefwoorden
(gecontr.)
trefwoorden
(vrije)

classificatie

index raadplegen

zoeken gebeurt via alfabetische indexen op de inhoud (alle woorden) van in het systeem aanwezige tekst

- bij web-zoekmachines is die index niet zichtbaar voor de gebruiker
- bij andere zoeksystemen is die index meestal wel te raadplegen, bijv. om schrijfwijze, woordvorm of frequentie van voorkomen van zoektermen te bekijken

index raadplegen

wat zijn de “termen” die in de index worden gezet?

- woord-index *losse woorden uit de tekst*
- "phrase"-index" *bevat "zinnetjes" uit de tekst
bijv. samengestelde trefwoorden*

wat wordt met “velden” gedaan?

- afzonderlijke indexen per veld
- één totale index *(met daarin ook veld-info)*

expand weather

Ref	Items	Index-term
E1	2	WEAST
E2	2	WEATE
E3	1542	WEATHER — los woord
E4	1	WEATHER BROADCASTS
E5	25	WEATHER FORECASTING — phrase
E6	2	WEATHER MODIFICATION
E7	4	WEATHER REPORTS
E8	8	WEATHER SATELLITES
E9	2	WEATHER VANES
E10	2	WEATHERALD
E11	5	WEATHERALL
E12	1	WEATHERBEE
E13	1	WEATHERBY
E14	1	WEATHERCASTS
E15	1	WEATHERDISC
E16	12	WEATHERED
E17	27	WEATHERFORD
E18	1	WEATHERFORD HAMMOND MOUNTAIN COLLECTION
E19	1	WEATHERFORD PUBLIC SCHOOL DISTRICT OK
E20	1	WEATHERGLASS
E21	3	WEATHERHEAD
E22	90	WEATHERING
E23	27	WEATHERIZATION
E24	1	WEATHERIZATION

zoeken

- zoeken:
 - soms alleen maar een woord intikken,
 - soms zoekknop om zoeken te starten,
 - soms zoekopdracht nodig
- zoeken op een woord geeft als resultaat een verzameling (set) van alle documenten (records) waarin dat woord voorkomt
- systeem meldt meestal hoeveel is gevonden

Set	Term Searched	Items
S1	RETRIEVAL	57349
S2	DATABASE	94522

zoeken

- zoeken op een woord:

weather

- specifiek zoeken: specifiek veld

weather/ti
AU=jansen, a.

- ruimer zoeken: trunkeren

*informat**

trunkeren (en maskeren)

l i b r a r *

l i b r a r i a n
l i b r a r i a n s
l i b r a r i a n s h i p
l i b r a r i e s
l i b r a r y

c a t *

c a t
c a t a c o m b
c a t a l o g
c a t a l y t i c
c a t a r a c t
c a t a s t r o p h e
c a t e c h i s m
c a t e g o r y
c a t e r i n g
c a t e r p i l l a r
c a t h a r s i s
c a t h o d e
c a t h o l i c
.....

trunkeren (en maskeren)

cat+	cat cats catz
cat+++	cat cater cats cattle
wom?n	woman women
pol ari ?ati on	pol ari sati on pol ari zati on

- NB:*
- elk zoekstelsel gebruikt zijn eigen trunkatiesymbolen
 - webzoekmachines kunnen helemaal NIET trunkeren
(m.u.v. Exalead)

combineren

gebruik van booleaanse operatoren om woorden (of sets) te koppelen

- AND beide woorden gelijktijdig aanwezig, aan beide criteria voldaan

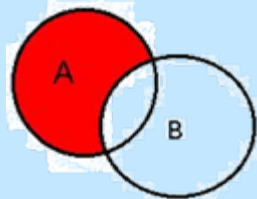
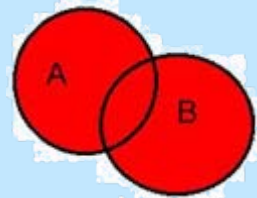
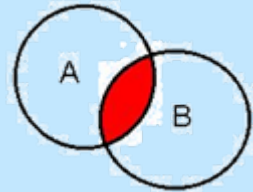
computer AND bibliotheek

- OR minstens één van de woorden aanwezig, aan minstens één criterium voldaan

fiets OR rijwiel

- NOT woorden/resultaten uitsluiten

greenhouse NOT climatic change



combineren

- combineren van woorden (*ook bij webzoekmachines*)
library AND (web OR internet)
- combineren van sets (*niet bij webzoekmachines*)
(#1 OR #2 OR #3) AND (#5 OR #6) NOT #4
- specifiekere AND-combinatie:
 - beide woorden binnen bepaalde afstand
electronic ADJ4 publishing
electronic NEAR publishing (*zelfs bij Yahoo, Bing, Exalead*)
 - beide woorden in zelfde zin
library SAME computer

best match

"best match" & "relevance ranking"

- gebruiker tikt reeks zoektermen in
- computer vindt documenten die daar het "best" bij aansluiten en presenteert die in volgorde van berekende "relevantie"

maar hoe doet ie dat ?

- *standaard methode bij webzoekmachines*
- *ook mogelijk bij sommige andere zoeksystemen*

OID

Basic Search (Natural Language)	Natural language search: Works best when entering a phrase or question about the topic of your research. Geared towards novice or occasional researchers but also a powerful tool for professionals that quickly need to find relevant results. e.g.: <i>use of Ritalin for adults with ADHD</i>
Ovid Advanced Search	Search on keyword, title author or journal. Will try to map your search term directly to the correct tree or thesaurus term when possible. Supports Boolean search terms and command line syntax searching, e.g. <i>cats.ti,ab. not dogs.ti,ab.</i>
Multi Field Search	Easily create Boolean searches. Use the dropdown to select specific fields and choose from the Boolean operators "Not", "Or", "And" to build your precision search.

EBSCO



Searching: **SocINDEX** | [Choose Databases >](#)

therapy alcoholism alcohol abuse adolescents youth

Search **Clear** ?

▼ [Search Options](#) | [Basic Search](#) | [Advanced Search](#) | [Visual Search](#) | [Search History/Alerts](#) | [Preferences >](#)

Search Options **Reset**

- Search modes ?
- Boolean/Phrase
 - Find all my search terms
 - Find any of my search terms
 - SmartText Searching** [Hint](#)

SmartText Searching ?

Enter as much text for your search as you want - a phrase, a sentence, paragraph, or even whole pages. This mode searches EBSCO Databases only.

[Close](#)

Limit your results

Linked Full Text

(Peer

best match

"best match" & "relevance ranking"

wat "matcht" het best?

- hoe meer termen aanwezig hoe beter
*(alleen nog gebruikt bij "kleinere" zoeksystemen;
bij webzoekmachines nu standaard: "all of the terms";
keuze "any of the terms" rankt hier niet meer op)*
- document met zoekwoorden in titel of heading belangrijker
- document met herhaalde zoekwoorden belangrijker
- document met zoekwoorden dicht bij elkaar belangrijker
- zeldzame zoekwoorden belangrijker
- webpagina's waarheen veel gelinkt wordt belangrijker

resultaten tonen

opvragen van bepaalde gevonden gegevens,
uit bepaald zoekresultaat, in bepaalde presentatievorm

- bepaald zoekresultaat *set-nummer*
- bepaalde gevonden gegevens *item-nummers*
- bepaalde presentatie *format / velden*

*veel zoeksystemen laten na elke actie automatisch al
meteen de eerste 10 gevonden titels zien
(bij zoekmachines kan dat zelfs niet anders)*

... en dan ook zo downloaden of printen

zoekgeschiedenis

overzicht van gebruikte zoekopdrachten en daarmee verkregen resultaten (aantallen)

geheugenopfrisser bij maken van nieuwe combinaties of bij opvragen van resultaten voor printen of downloaden

Set	Term Searched	Items
S1	RETRIEVAL	57349
S2	DATABASE?	121053
S3	DATABANK?	1335
S4	ADVANCES OR TRENDS	88199
S5	S1 AND (S2 OR S3) AND S4	487

*bij sommige web-zoekmachines kunnen gebruikte zoekvragen worden bewaard en weer opgevraagd, maar **geen** te combineren sets*

zoekmethode

"oriënterende" aanpak zoekvraag: "best-match methode"

1. maak inhoudelijke analyse van je vraagonderwerp, en bepaal daarvoor essentiële "vraagconcepten"
2. bedenk de meest specifieke zoekterm voor elk concept
3. tik die 3 of 4 woorden in (*bijv. in webzoekmachine*)
4. beoordeel gevonden resultaten op basis van titels / trefwoorden / samenvattingen / tekst
5. probeer zo nodig nader te specificeren met extra term(en)
6. probeer nieuwe combinaties met alternatieve woorden voor sommige van de concepten

zoekmethode

gestructureerde aanpak zoekvraag: "bouwsteenmethode"

1. maak inhoudelijke analyse van je vraagonderwerp en bepaal daarvan de "bestanddelen" of "vraagconcepten"

voorbeeld: *literatuur over moderne bouwkunst in Nederland*

bestanddelen: **bouwkunst** **modern** **nederland**

2. bedenk zoektermen voor elk van die bestanddelen

zoekmethode

"bouwsteenmethode"

2. bedenk zoektermen voor elk van die bestanddelen

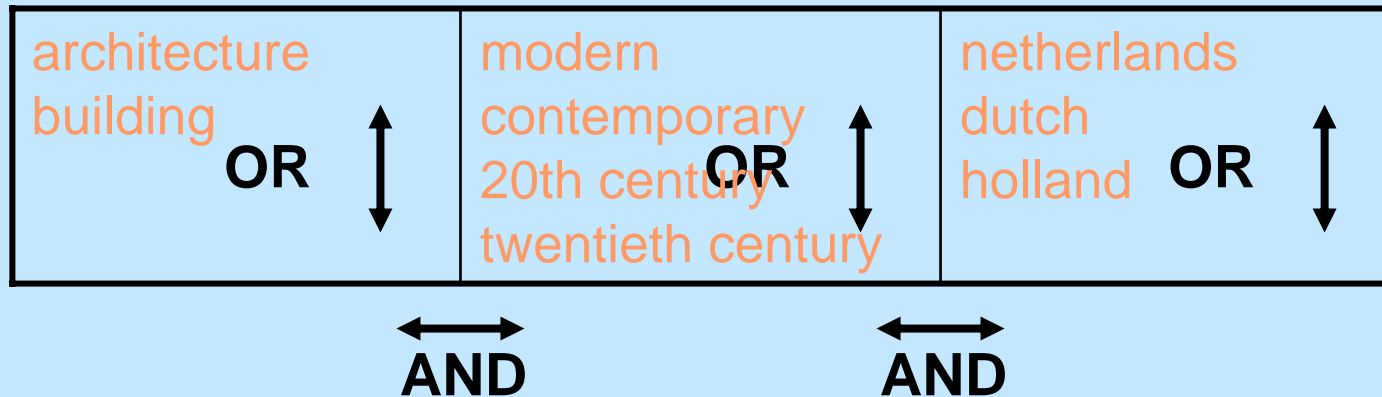
- (quasi)synoniemen *[modern >> hedendaags]*
- spellingvarianten *[20ste eeuws >> twintigste eeuws]*
- samenstellingen *[bouwen >> hoogbouw]*
- antoniemen *[werkloosheid >> werkgelegenheid]*
- specifiekere deelbegrippen *[nederland >> rotterdam]*
- vervoegingen, verbuigingen, *[bouwen >> gebouwd]*
- vertalingen *[bouwen >> building]*
-

architecture building	modern contemporary 20th century twentieth century	netherlands dutch holland
--------------------------	---	---------------------------------

zoekmethode

"bouwsteenmethode"

3. zoek en combineer zoektermen per bestanddeel met **OR**
4. combineer resultaten van de bestanddelen met **AND**



5. beoordeel gevonden resultaten op basis van titels / trefwoorden / samenvattingen / tekst
6. stel zoekactie zonnodig bij

zoekproces - praktijk

- beginnen *te gebruiken zoekstelsel kiezen
(eventueel) bestand kiezen*
- zoekactie *in index kijken
zoeken op termen
combineren
tonen van resultaten
beoordelen van resultaten
bijstellen van zoekactie
zoeken op termen
combineren
tonen van resultaten
.....
inperken op secundaire criteria
afdrukken of downloaden van resultaten*

