

Zoeken is Saai

Informatica van de universiteit naar school

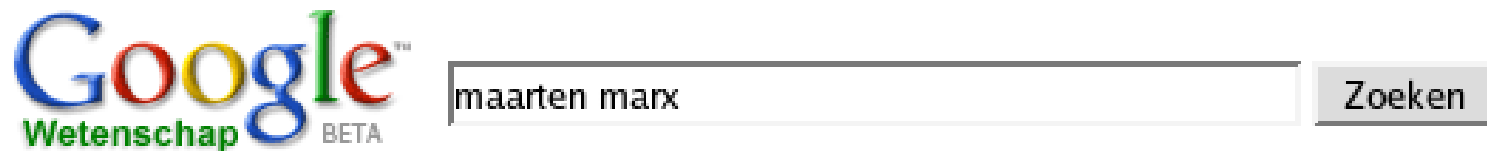
Maarten Marx

Universiteit van Amsterdam

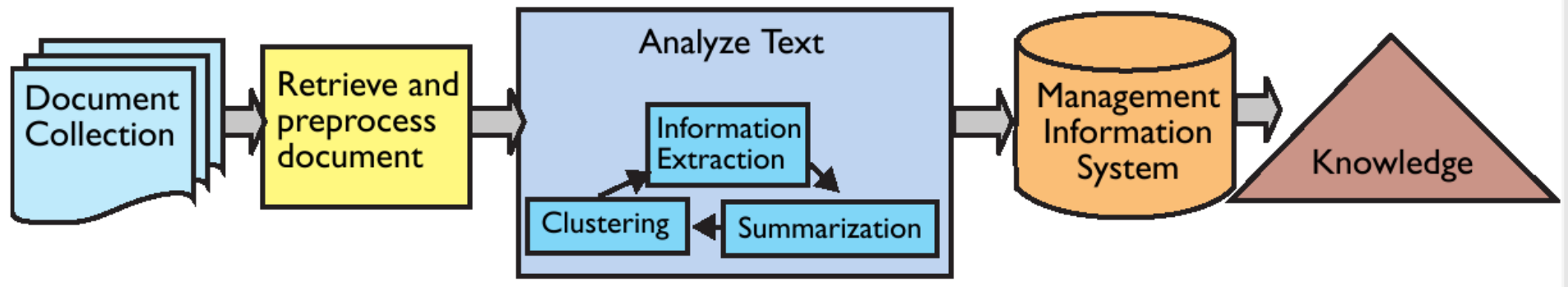
November 2008

Wie ben ik?

- Dat vraag je toch aan Google.



Wat doen we?



Intelligente Informatie Ontsluiting

Uitgangspunt van ons werk

- Zoeken is géén activiteit, maar
- een **verstoring** van een activiteit.

Zoeken is saai en vaak vervelend.

Toch moet je best vaak iets (op)zoeken....

Speld in de hooiberg: hier is Google goed geschikt voor.

Zoeken en turven • Wie was het populairst deze week?
• Wat werd er allemaal over Obama gezegd en hoe vaak?

Zoeken naar mensen Ik wil een Xbox aansluiten op een MacBook.
Wie op school kan me hierbij het best helpen?

Zoeken in heel veel, hele lange, vaak saaie documenten
• iets terugvinden in de video opnamen van de les

Zoeken naar mensen

- Veel makkelijker dan het lijkt ...
- (1) Verzamel van ieder mens zoveel mogelijk van wat ie zegt en schrijft
- (2) Maak hiervan een **Persoonlijk Taal Model**
- (3) Doe dit voor een hele hoop mensen
- (4) En klaar ben je.

Demo: Welke tien politici passen het beste bij dit bericht?

De Telegraaf **Binnenland** Zoek in deze site powered by

VIDEO BINNENLAND | STELLING VAN DE DAG | PRINSJESDAG 2008

HOME > NIEUWS > BINNENLAND

SNELNIEUWS vr 26 sep 2008, 19:17
Maandag 29 september

Non-actieve militairen blijven in Uruzgan

Binnen- en Buitenland RSS DEN HAAG - Er is vooralsnog geen reden om de 24 militairen die eerder deze week tijdelijk op non-actief zijn gesteld, terug te halen uit Uruzgan. Dat heeft staatssecretaris Jack de Vries (Defensie) vrijdag gezegd.

17:57 Schilderij Breughel ontdekt

17:55 Minder uitgaven geboortebepanking

17:54 Rusland levert Zuid-Korea gas

16:53 Politie verliest inbrekers

 16:46 Geen vervolging om 'Schipholgate'

 16:39 Kipcampagne niet 100 procent oké

16:27 Rel over gekaapte tanks

Donderdagavond laat werd bekend dat een heel peloton op non-actief is gesteld omdat zij volgens hun eigen commandant de werkzaamheden dusdanig belemmeren dat ze niet kunnen worden ingezet. Er loopt momenteel een oriënterend feitenonderzoek dat naar verwachting enkele dagen zal duren. Zoals gebruikelijk wordt



De 24 militairen die eerder deze week tijdelijk op non-actief zijn gesteld, blijven

- http://www.telegraaf.nl/binnenland/2039440/___Groep_militairen_Uruzgan_.html
- Acht van de tien hits zitten in de commissie defensie.

Zoeken en Turven

- Wetenschap begint heel vaak gewoon met **tellen en turven** ...
- woordjes tellen
- mensen tellen
- hoe vaak iemand genoemd wordt (in de krant/op het web/op TV) tellen
- een stuk lastiger dan je denkt!

Voorbeeld

- Vraag:

Met wie wordt Wilders het meest geassocieerd in de krant en reacties

- Antwoord: Neem alle (web)kranten uit een bepaalde periode en ga turven ...

- <http://www.politicalmashup.nl/2008/04/reacties-op-geert-wilders/>

Wie is wie?

Paul de L, P.de L, P de Leeuw, Paul de leeuw,
Paul, Leeuw, paul de leeuw, paul, de leeuw,
leeuw, De Leeuw

Anneke Grönloh, Anneke Gr, Anneke Gronloh,
Mevrouw G, Anneke, AG, Gronloh, Gr ?nloh,
anneke, gronloh, Anneke Kreunlo, Anneke
Grohnloh, Anneke **Kreuntzo**

Dit staat echt zo in de krant

- http://www.telegraaf.nl/prive/74586221/Anneke_Gr%C3%B6nloh_verwittigt_advocaat_over_persiflage_Paul_de_Leeuw.html?p=1,1
- Wij hebben de reacties ook nog bewaard:
- http://staff.science.uva.nl/~mahboob/dutchnews/output/news_377487.htm

Zoeken in heel veel, hele lange, vaak saaie documenten

- Zoeken in video is lastiger dan zoeken in tekst of een boek
- Zoeken in heel veel videos is nog lastiger!
- Hoe maak je dat eenvoudiger?
- Probleem: zoeken gaat met tekst, maar video is beeld
- Hoe verbindt je die?

Hoe verbindt je tekst met beeld?

1. Mensen er wat bij laten schrijven . . .
2. Veels te veel werk! ■

Ondertitels,
spraakherkenning,
kopjes/tekst in beeld,
gebruik de dias in een lezing,
gebruik aantekeningen/verslagen

3 Voorbeelden

- <http://zookma.science.uva.nl/islatv/>
- <http://openkamer.unwind.nl/>
- Bob Dylan's multimedia clip
http://zookma.science.uva.nl/islatv/demo_dylan.html

Samenvatting

- Zoeken blijft een **onderbreking** van een nuttige taak,
- maar informatica maakt het **leuker, makkelijker** en **sneller** om iets te vinden.
- Universiteit samen met Google, Yahoo, Beeld en Geluid, Bibliotheken, . . .

Einde

PoliticalMashup



PoliticalMashup

