



UPPSALA  
UNIVERSITET

# Informationsåtkomst

Eva Forsbom

`evafo@stp.lingfil.uu.se`

Metoder och tillämpningar inom språkteknologi  
Språkteknologiprogrammet

4 december 2007



UPPSALA  
UNIVERSITET

Översikt

Dokumenthantering

Sortering/Filtrering

Yahoo! Pipes

Utvinning av  
information

Informationssökning

Informationsextraktion

Textsammanfattning

# Översikt



# Dokumenthantering

- Skapa
  - Skrivhjälpmedel
- Analysera
  - Indexering
- Lagra
  - Komprimeringsalgoritmer
- Söka
  - Sökalgoritmer

Översikt

Dokumenthantering

Sortering/Filtrering

Yahoo! Pipes

Utvinning av  
information

Informationssökning

Informationsextraktion

Textsammanfattning



# Sortering/Filtrering

- Klassificering (fastställda kategorier)
  - Språk
  - Ämnen
  - Genrer
  - Känslor
- Klustring (fastställt antal kategorier)
- Rensning
- Områdes-/Händelsebevakning
- Rekommendationssystem

Översikt

Dokumenthantering

Sortering/Filtrering

Yahoo! Pipes

Utvinning av  
information

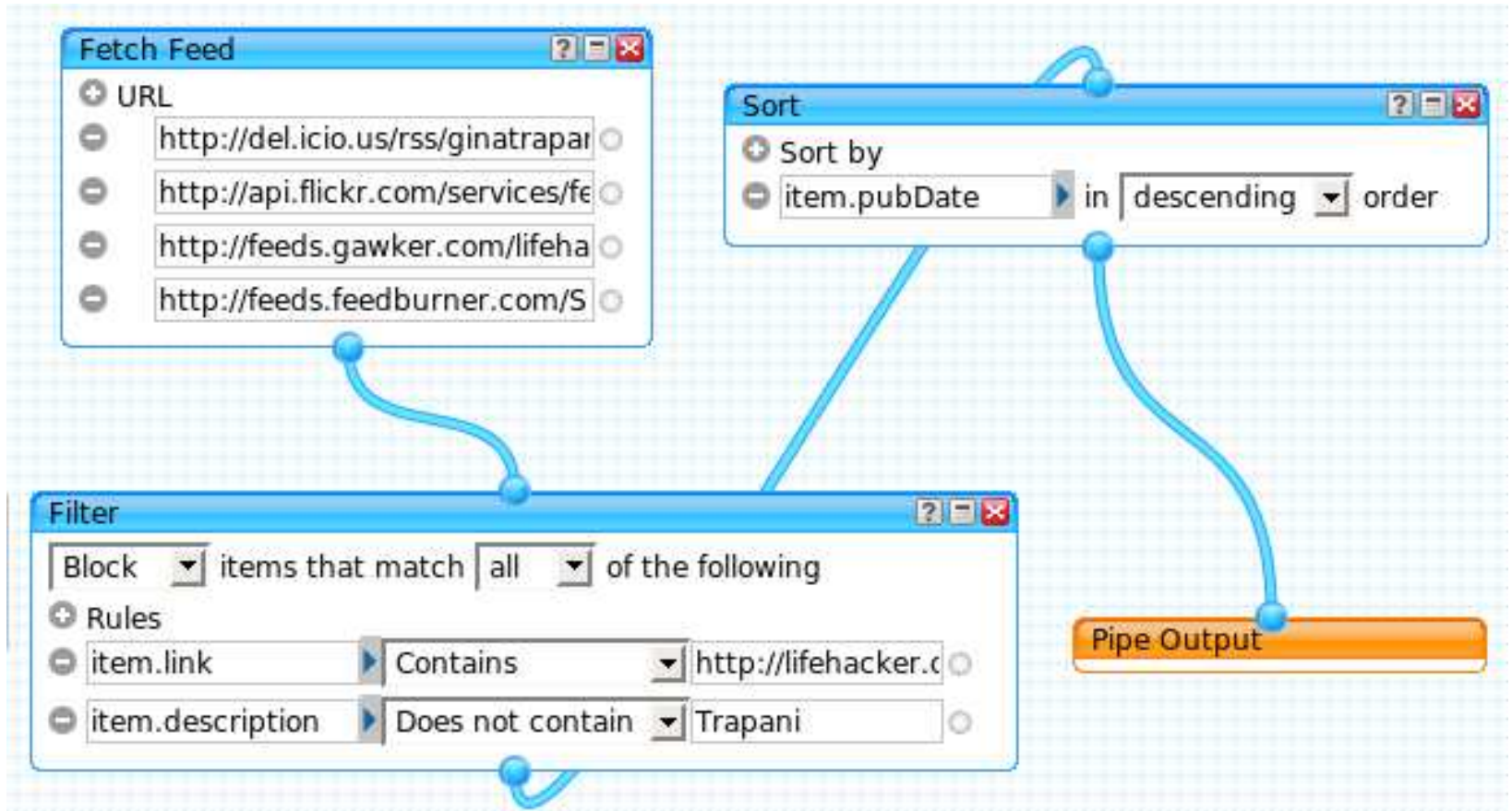
Informationssökning

Informationsextraktion

Textsammanfattning



# Yahoo! Pipes



Översikt

Dokumenthantering

Sortering/Filtrering

Yahoo! Pipes

Utvinning av  
information

Informationssökning

Informationsextraktion

Textsammanfattning



# Utvinning av information

- Informationsökning
  - Dokumentsökning
  - Passagesökning
  - Faktasökning
- Informationsextraktion
  - Mönstermatchning (*x killed y, y was killed by y*)
  - Variabelsökning (`killed(x, y)`)
- Informationsutvinning ur text
  - Samförekomstanalys (*migrän...magnesiumbrist*)
  - Relationssökning (`z(migrän, y)`)

Översikt

Dokumenthantering

Sortering/Filtrering

Yahoo! Pipes

Utvinning av  
information

Informationssökning

Informationsextraktion

Textsammanfattning



UPPSALA  
UNIVERSITET

Översikt

Informationssökning

Grundmetoder

Systemarkitektur

Indexering

Inverterat index

Frågebearbetning/Sökning

Rankning

Utvärdering

Informationsextraktion

Textsammanfattning

# Informationssökning





# Grundmetoder

- Boolesk modell (sanningsvärdeslogik)
- Vektorrumsmodell (oskarp logik)
- Statistisk modell ( $P(R = 1|q, d)$ )
  
- Språkmodell ( $P(q|M_d)$ )

Översikt

Informationssökning

Grundmetoder

Systemarkitektur

Indexering

Inverterat index

Frågebearbetning/Sökning

Rankning

Utvärdering

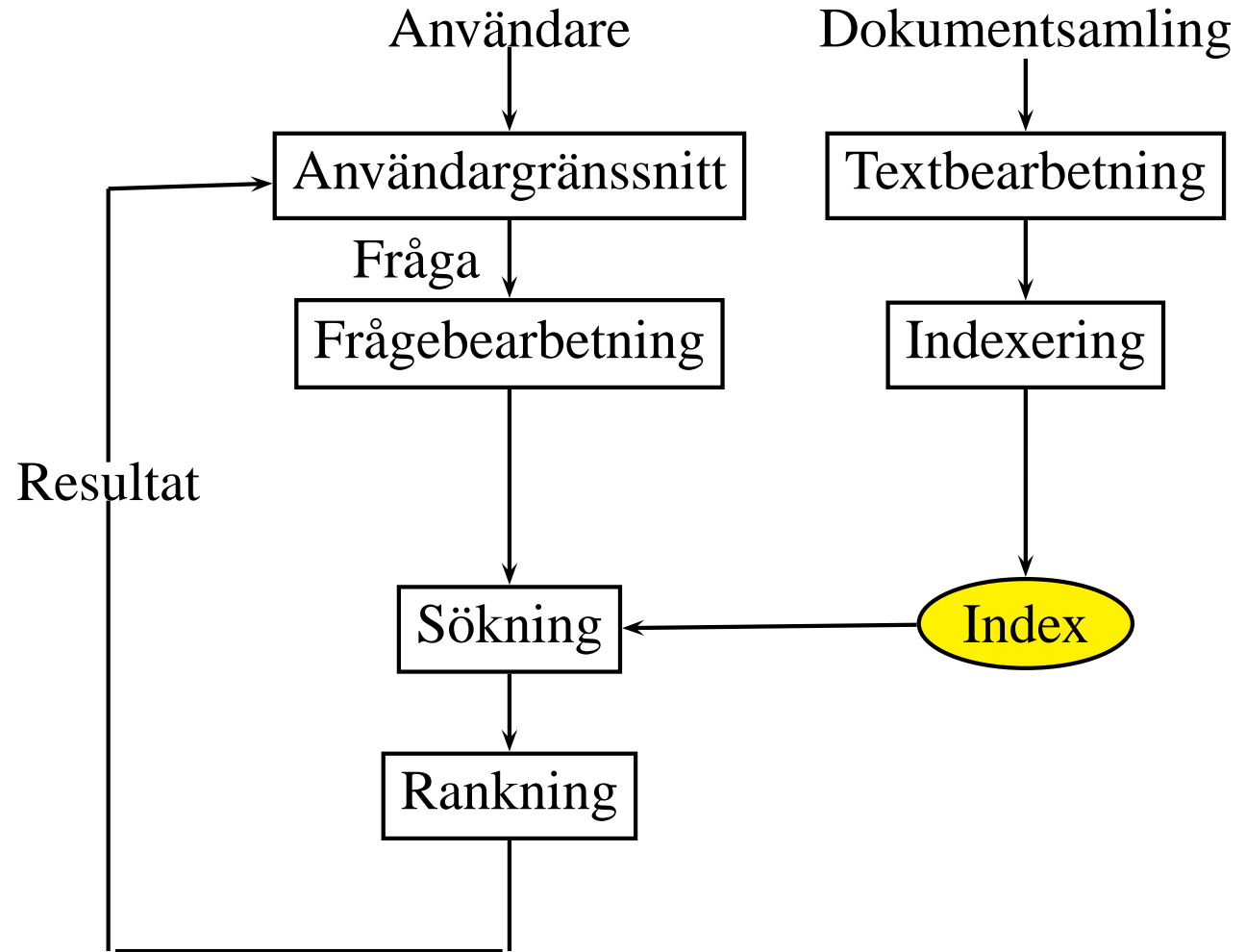
Informationsextraktion

Textsammanfattning





# Systemarkitektur



Översikt

Informationssökning

Grundmetoder

Systemarkitektur

Indexering

Inverterat index

Frågebearbetning/Sökning

Rankning

Utvärdering

Informationsextraktion

Textsammanfattning



# Indexering

- Innehållsord beskriver innehållet (filtrera funktionsord).
- De mest frekventa orden är inte särskiljande (filtrera).
- De minst frekventa orden är inte karakteristiska för innehållet (filtrera).
- Ord i vissa positioner är viktigare än andra (använd positionsviktning).
- Olika former av ett ord beskriver samma begrepp (använd stam eller grundform).
- Flerordsenheter betyder något annat än sina delar (engelska: sök efter sammansättningar).
- Delarna i en sammansättning är relaterade till de enskilda orden (svenska: dela upp sammansättningar).
- Begrepp är relaterade (använd tesaurer/ontologier/ordnät).

Översikt

Informationssökning

Grundmetoder

Systemarkitektur

Indexering

Inverterat index

Frågebearbetning/Sökning

Rankning

Utvärdering

Informationsextraktion

Textsammanfattning



# Inverterat index

Dok	Text
1	katt åt råtta
2	råtta åt katt
3	katt och råtta
4	katt och råtta, katt åt råtta

Index	1	2	3	4
katt	1	1	1	2
åt	1	1	0	1
råtta	1	1	1	2

Översikt

Informationssökning

Grundmetoder

Systemarkitektur

Indexering

Inverterat index

Frågebearbetning/Sökning

Rankning

Utvärdering

Informationsextraktion

Textsammanfattning



# Frågebearbetning/Sökning

- “Indexering” av fråga
- Frågeexpansion
  - Automatisk
    - Böjda former
    - Synonymer
    - Hyperonymer
- Frågeförfining
  - Återkoppling
    - Betydelsesdisambiguering
    - Hyponymer

Översikt

Informationssökning

Grundmetoder

Systemarkitektur

Indexering

Inverterat index

Frågebearbetning/Sökning

Rankning

Utvärdering

Informationsextraktion

Textsammanfattning



# Rankning

Dok	Text	katt	åt	råtta
fråga	katt åt mus	1	1	0
dok 1	katt åt råtta	1	1	1
dok 2	råtta åt katt	1	1	1
dok 3	katt och råtta	1	0	1
dok 4	katt och råtta, katt åt råtta	2	1	2

- Boolesk modell (binär viktning + sanningsvärdeslikhet)
- Vektorrummodell (tf.idf + vinkellikhet)
- Statistisk modell (binär viktning + relevansodds)

Översikt

Informationssökning

Grundmetoder

Systemarkitektur

Indexering

Inverterat index

Frågebearbetning/Sökning

Rankning

Utvärdering

Informationsextraktion

Textsammanfattning



# Utvärdering

- Uppgifter
  - Relevanta dokument
  - Rankning av relevanta dokument
- Mått
  - Noggrannhet
  - Fullständighet
  - F-mått
  - 11-punktsnoggrannhet
- TREC (Text REtrieval Conference, <http://trec.nist.gov/>)
- CLEF (Cross Language Evaluation Forum, <http://www.clef-campaign.org/>)

Översikt

Informationssökning

Grundmetoder

Systemarkitektur

Indexering

Inverterat index

Frågebearbetning/Sökning

Rankning

Utvärdering

Informationsextraktion

Textsammanfattning



Översikt

Informationssökning

**Informationsextraktion**

Informationsextraktion

Enheter

Attribut

Fakta

Händelser

Utvärdering

Textsammanfattning

# Informationsextraktion





# Informationsextraktion

- Ska passa ett väldefinierat informationsbehov.
  - Terroristattacker
  - Katastrofer
  - Omvärldsbevakning (sammanslagningar och uppköp)
- Bygger på domänmodellering.
  - Ontologier
  - E/R-diagram
- Söker t.ex. efter
  - namngivna enheter
  - koreferenser
  - fasta fraser med variabler

Översikt

Informationssökning

Informationsextraktion

**Informationsextraktion**

Enheter

Attribut

Fakta

Händelser

Utvärdering

Textsammanfattning





# Enheter

[Exempel från MUC]

Fletcher Maddox, former Dean of the UCSD Business School, announced the formation of La Jolla Genomatics together with his two sons. La Jolla Genomatics will release its product Geninfo in June 1999. Geninfo is a turnkey system to assist biotechnology researchers in keeping up with the voluminous literature in all aspects of their field.

Dr. Maddox will be the firm’s CEO. His son, Oliver, is the Chief Scientist and holds patents on many of the algorithms used in Geninfo. Oliver’s brother, Ambrose, follows more in his father’s footsteps and will be the CFO of L.J.G. headquartered in the Maddox family’s hometown of La Jolla, CA.

Persons:	Organizations:	Locations:	Artifacts:	Dates:
Fletcher Maddox	UCSD Business School	La Jolla	Geninfo	June 1999
Dr. Maddox	La Jolla Genomatics	CA	Geninfo	
Oliver	La Jolla Genomatics		Geninfo	
Oliver	L.J.G.			
Ambrose Maddox				

- Översikt

---

- Informationssökning

---

- Informationsextraktion

---

- Informationsextraktion

---

- Enheter

---

- Attribut

---

- Fakta

---

- Händelser

---

- Utvärdering

---

- Textsammanfattning



# Attribut

Fletcher Maddox, [former Dean of the UCSD Business School](#), announced the formation of La Jolla Genomatics together with his two sons. La Jolla Genomatics will release [its product](#) Geninfo in June 1999. Geninfo is a turnkey system to assist biotechnology researchers in keeping up with the voluminous literature in all aspects of their field.

Dr. Maddox will be [the firm's CEO](#). [His son](#), Oliver, is the [Chief Scientist](#) and holds patents on many of the algorithms used in Geninfo. [Oliver's brother](#), Ambrose, follows more in [his father's](#) footsteps and will be [the CFO of L.J.G.](#) headquartered in [the Maddox family's hometown](#) of La Jolla, CA.

Översikt

Informationssökning

Informationsextraktion

Informationsextraktion

Enheter

Attribut

Fakta

Händelser

Utvärdering

Textsammanfattning

Name:	Descriptor:	Category:
<a href="#">Fletcher Maddox</a> <a href="#">Dr. Maddox</a>	former Dean of the UCSD Business School his father the firm's CEO	PERSON
<a href="#">Oliver</a>	His son Chief Scientist	PERSON
<a href="#">Ambrose</a>	Oliver's brother the CFO of L.J.G.	PERSON
UCSD Business School		ORGANIZATION
La Jolla Genomatics L.J.G.		ORGANIZATION
<a href="#">Geninfo</a>	its product	ARTIFACT
<a href="#">La Jolla</a> CA	the Maddox family's hometown	LOCATION
		LOCATION



# Fakta

PERSON	Employee_of	ORGANIZATION
Fletcher Maddox	Employee_of	UCSD Business School
Fletcher Maddox	Employee_of	La Jolla Genomics
Oliver	Employee_of	La Jolla Genomics
Ambrose	Employee_of	La Jolla Genomics

ARTIFACT	Product_of	ORGANIZATION
Geninfo	Product_of	La Jolla Genomics

LOCATION	Location_of	ORGANIZATION
La Jolla	Location_of	La Jolla Genomics
CA	Location_of	La Jolla Genomics

- Översikt
- Informationssökning
- Informationsextraktion
- Informationsextraktion
- Enheter
- Attribut
- Fakta**
- Händelser
- Utvärdering
- Textsammanfattning



# Händelser

## COMPANY-FORMATION\_EVENT:

COMPANY: La Jolla Genomatics  
Fletcher Maddox  
PRINCIPALS: Oliver  
Ambrose

DATE:

CAPITAL:

## RELEASE-EVENT:

COMPANY: La Jolla Genomatics  
PRODUCT: Geninfo  
DATE: June 1999  
COST:

Översikt

Informationssökning

Informationsextraktion

Informationsextraktion

Enheter

Attribut

Fakta

Händelser

Utvärdering

Textsammanfattning



# Utvärdering

- Uppgifter
  - Namngivna enheter
  - Mallelementenheter (attribut)
  - Mallrelationer (employee\_of)
  - Scenariomallar (händelser)
  - Koreferenser
- Mått
  - Noggrannhet
  - Fullständighet
  - F-mått
- MUC (Message Understanding Conference,  
[http://www.itl.nist.gov/iaui/894.02/related\\_projects/muc/](http://www.itl.nist.gov/iaui/894.02/related_projects/muc/))

Översikt

Informationssökning

Informationsextraktion

Informationsextraktion

Enheter

Attribut

Fakta

Händelser

**Utvärdering**

Textsammanfattning



Översikt

Informationssökning

Informationsextraktion

**Textsammanfattning**

Kontextfaktorer

Informativ/Flerspråkig

Indikativ/Informativ

Indikativ/Evaluerande

Metoder

Extrakt

Abstrakt

Utvärdering

# Textsammanfattning





# Kontextfaktorer

The screenshot shows a window titled "Summarisation system" with several control panels:

- Compression:** A slider bar ranging from 0 to 100, with a marker at approximately 25.
- Source:** Radio buttons for  **Single document** and  **Multiple documents**.
- Focus:** Radio buttons for  **Generic** and  **Specific**.
- Function:** Radio buttons for  **Indicative**,  **Informative**, and  **Evaluative**.
- Type:** Radio buttons for  **Extract** and  **Abstract**.
- Information:** Radio buttons for  **Background** and  **Just the news**.
- Language:** A dropdown menu currently set to "English".

Översikt

Informationssökning

Informationsextraktion

Textsammanfattning

Kontextfaktorer

Informativ/Flerspråkig

Indikativ/Informativ

Indikativ/Evaluerande

Metoder

Extrakt

Abstrakt

Utvärdering



## Ryggtaavlans känsl felbeskriven i dermatomkartor



NILS DAHLGREN, överläkare,  
docent, smärtenheten, onkolo-  
giska kliniken, Universitets-

sjukhuset i Lund  
helgonavagen18@msn.com

Genom att studera patienter drabbade av herpes zoster har ryggtaavlans segmentella sensoriska innervation kartlagts. Resultatet avviker påtagligt från vad de flesta av gängse dermatomkartor framställer. Den enkla parallelliteten mellan den dorsala grenens sensoriska täckning och motsvarande segments revben, vilket vanligtast återges av litteraturens dermatomkartor, får inget stöd i föreliggande studie.

Enligt författaren har denna missvisande information medfört att fler än en patient underkastats en inadekvat handläggning av smärttillstånd lokaliserade till ryggtaavlans.

### **Dermatom och dermatomkartor**

Med kranialnerverna V, VII och X samt de segmentellt avgående spinala nervrötterna från ryggmärgen registrerar det cen-

### **SAMMANFATTAT**

Genom att studera patienter med herpes zoster har ryggtaavlans segmentella sensoriska innervation kartlagts. Resultatet avviker påtagligt från vad de flesta av gängse dermatomkartor framställer. Den enkla parallelliteten mellan den dorsala grenens sensoriska täckning och motsvarande segments revben,

som vanligtast återges av litteraturens dermatomkartor, får inget stöd i föreliggande studie. Huden över länden och korsbenets kraniella omfång innerveras av nedre hälften av bröstrygggradens nervrötter, vilket för till att de lumbala rötterna saknar dorsal täckning.

### **Summary**

Based on observed patients suffering from herpes zoster, a dermatomal chart of the back has been processed. It was found that the skin of the lumbar area and cranial part of the sacrum is innervated by nerves of a thoracic origin. Obviously lumbar nerve roots do not supply the skin of the back, contrary to what most charts indicate.

Översikt

Informationssökning

Informationsextraktion

Textsammanfattning

Kontextfaktorer

Informativ/Flerspråkig

Indikativ/Informativ

Indikativ/Evaluerande

Metoder

Extrakt

Abstrakt

Utvärdering





# Indikativ/Informativ

Översikt

Informationssökning

Informationsextraktion

Textsammanfattning

Kontextfaktorer

Informativ/Flerspråkig

Indikativ/Informativ

Indikativ/Evaluerande

Metoder

Extrakt

Abstrakt

Utvärdering

The screenshot shows a Google search interface. At the top right is a "Logga in" link. Below it are navigation links for "Nätet", "Bilder", "Grupper", and "Nyheter" (with a "Nytt!" badge). The search bar contains "språkteknologi" and a "Sök" button. To the right of the search bar are links for "Avancerad sökning" and "Inställningar". Below the search bar are radio buttons for "webben", "sidor på svenska", and "sidor från Sverige". The search results section is titled "Nätet" and shows "Resultat 1 - 100 av ungefär 66 600 vid sökning efter språkteknologi. (0,24 sekunder)". The first result is "Välkommen till Språkteknologi.se! — Språkteknologi.se", with a snippet: "Nationellt centrum för svensk språkteknologi. Information om produkter, teknik, konferenser, samt kontaktuppgifter till personer och organisationer och ...". Below the snippet are links: "sprakteknologi.se/ - 33k - Cachad - Liknande sidor". The second result is "Vad är språkteknologi? Läs mer om olika delområden — Språkteknologi.se", with a snippet: "Språkteknologi omfattar all tillämpning och användning av språkkunskap vid utveckling av datorsystem som kan känna igen, förstå, tolka, och generera ...". Below the snippet are links: "sprakteknologi.se/vad-aer-sprakteknologi - 38k - Cachad - Liknande sidor". At the bottom of the results is a link: "[ Fler resultat från sprakteknologi.se ]". On the right side of the search results, there is a section for "Sponsrade länkar" with a link for "Språkteknologiska" with the text "Bra Erbjudande På Språkteknologiska Spara Pengar - Köp På Nätet Idag!" and the URL "www.tradera.com".



# Indikativ/Evaluerande (BTJ-häftet 3/2003)

Översikt

Informationssökning

Informationsextraktion

Textsammanfattning

Kontextfaktorer

Informativ/Flerspråkig

Indikativ/Informativ

Indikativ/Evaluerande

Metoder

Extrakt

Abstrakt

Utvärdering

## Språkvetenskap

F.06

**Stalfelt, Sven Olov.**

Med språket som verktyg. Industrilitt.

243 s.

**BN 4505288-47**

521914

**031-03-038**

Mars

plastad mjuk pärm

318:–



Denna bok innehåller många råd och tips om användning av svenska i tal och skrift för en mängd olika ändamål. Ändamålen omfattar till exempel hur man talar inför en publik, hur man argumenterar eller utreder frågor, hur man gör en layout tilltalande, hur olika skriftgener är strukturerade (t.ex. offert, formellt brev), hur man sammanträder korrekt och effektivt. Dessa råd kan fungera väl som handledning för talare och skribenter som ska lära sig nya kommunikativa funktioner. Däremot är författarens presentation av mer teoretiska språkfrågor betydligt mer osäker. Det röriga och summariska grammatikavsnittet relateras bara svagt till de praktiska skriv- och talråden. Diskussionen om språket i samhället innehåller förlegade åsikter (exempelvis om halvspråkighet) och felaktiga förenklingar (exempelvis "språk är namn"). Bokens språk är lättfattlig och layouten trevlig. Sven Olov Stalfelt är handledare och lärare i muntlig och skriftlig kommunikation inom näringsliv och samhälle. – **Sally Boyd.**



# Metoder

- IR-metoder: meningsextraktion
  - + domänoberoende
  - - osammanhängande
- IE-metoder: faktaextraktion
  - + tar fram relevant information
  - - domänberoende
- Språkmodellering
  - meningskomprimering
  - rubrikgenerering
- Kunskapsintensiva metoder: språkförståelse, diskursanalys och generering



# Extrakt

## Viktiga meningar

- innehåller ord som förekommer ofta i texten
- innehåller ord som förekommer i titel och rubriker
- förekommer först eller sist i stycken
- förekommer i olika positioner beroende på genre
- innehåller bonusord, medan oviktiga meningar innehåller stigmaord
- och begrepp är de med flest länkar i en semantisk struktur av texten
- och oviktiga meningar är möjliga att ta ut från en diskursrepresentation av texten.

[Exempel]



# Abstrakt

## Kunskapsinsamling:

- Tänka-högt-protokoll
- Parallellkorpus med fullständig text–abstrakt

## Extraktion + reducering/korrigering:

- Reduceringsoperatorer och generering
  - stryk (Sara hade en röd ballong → Sara hade en ballong)
  - aggregera (Erik åt + Anna åt → Erik och Anna åt)
  - generalisera (banan + äpple → frukt)
- Statistiska reduceringsmodeller

Översikt

Informationssökning

Informationsextraktion

Textsammanfattning

Kontextfaktorer

Informativ/Flerspråkig

Indikativ/Informativ

Indikativ/Evaluerande

Metoder

Extrakt

Abstrakt

Utvärdering



# Utvärdering

Inomtextuell utvärdering: användares bedömning av

- sammanfattningens flyt
- hur väl sammanfattningen täcker vissa angivna huvudidéer
- sammanfattningens likhet med en idealisk sammanfattning

Utomtextuell utvärdering: bedömning av hur väl det går att

- besvara angivna frågor eller förstå sammanhang utifrån sammanfattningen
- avgöra dokumentets relevans för ett ämne.

Utvärderingsforum:

- SUMMAC (Text Summarization Evaluation Conference, [http://www.itl.nist.gov/iaui/894.02/related\\_projects/tipster\\_summac/](http://www.itl.nist.gov/iaui/894.02/related_projects/tipster_summac/))
- DUC (Document Understanding Conference, <http://duc.nist.gov/>)



UPPSALA  
UNIVERSITET

Översikt

Informationssökning

Informationsextraktion

Textsammanfattning

**Extraktexempel**

Exempeltext

Frekvensanalys

Ord i titel och rubriker

Kärnmeningar

Genrebaserad position

Signalord

Länkar i semantisk  
struktur

Diskursstruktur

# Extraktexempel





## Exempeltext

### VINSÄLJANDE HANDLARE BRÖT MONOPOLET OCH GREPS AV POLISEN SVALÖV TT

**ICA-handlaren Harry Franzen i Röstånga i Skåne bröt på nyårsdagen under stort tumult spritmonopolet. Men efter bara tio minuter kom polisen och tog vinet i beslag, stängde affären och grep Franzen, den förste livsmedelshandlaren i landet som försökt starta vinförsäljning.**

En flaska han dock Franzen sälja till högstbjudande innan polisen stoppade kommersen.

Resten av vinflaskorna påstod Franzen att han skänkte bort till kunderna. - En flaska per person. Tänk på nykterheten, sade han innan kunderna stormade in i affären. Harry Franzen, etablerad kommunpolitiker för centrumdemokraterna i Svalövs kommun, hade inte gjort någon hemlighet av att han tänkte bryta mot lagen den dag EES-avtalet började gälla i Sverige. Ett stort pressuppbåd och många bybor från den lilla skånska orten mitt i Skåne var också på plats när butiken öppnade klockan elva på förmiddagen.

Franzen vill få prövat i domstol om Systembolagets monopol bryter mot EES-avtalets konkurrensregler eller ej. - Jag kommer att vinna processen. Det är jag övertygad om. Men det kommer att ta tid, sade Franzen, för dagen finklädd med slips, väst och skärmmössa, innan han motades in i en civil polisbil för att köras till Landskrona för förhör.





# Frekvensanalys

Översikt

Informationssökning

Informationsextraktion

Textsammanfattning

Extraktexempel

Exempeltext

Frekvensanalys

Ord i titel och rubriker

Kärnmeningar

Genrebaserad position

Signalord

Länkar i semantisk  
struktur

Diskursstruktur

Frekvens	Ordform	Frekvens	Basform
13	Franzen	10	Franz
7	vin	9	vin
7	polisen	8	polis
4	sälja	5	sälja
4	flaska	5	affär
4	affären	4	Systembolag
3	Systembolaget	4	förhör
3	Röstånga	4	flaska
3	Landskrona	4	EES-avtal
3	Harry	3	Röstånga
3	förhör	3	Landskrona
3	EES-avtalet	3	Harry
2	/.../	3	gripa
		3	Franzen
		3	bryta
		2	/.../



## Ord i titel och rubriker

- VINSÄLJANDE **HANDLARE BRÖT MONOPOLET** OCH GREPS AV **POLISEN**
- FICK ÅKA HEM
- **BÖTER** ELLER **FÄNGELSE**

**ICA-handlaren** Harry Franzen i Röstånga i Skåne **bröt** på nyårsdagen under stort tumult **spritmonopolet**. Men efter bara tio minuter kom **polisen** och tog vinet i beslag, stängde affären och grep Franzen, den förste **livsmedelshandlaren** i landet som försökt starta vinförsäljning. Straffet för försäljning av rusdrycker är **böter**, eller om brottet anses grovt, **fängelse** i höst fyra år.

Översikt

Informationssökning

Informationsextraktion

Textsammanfattning

Extraktexempel

Exempeltext

Frekvensanalys

Ord i titel och rubriker

Kärnmeningar

Genrebaserad position

Signalord

Länkar i semantisk  
struktur

Diskursstruktur



# Kärnmeningar

Översikt

Informationssökning

Informationsextraktion

Textsammanfattning

Extraktexempel

Exempeltext

Frekvensanalys

Ord i titel och rubriker

Kärnmeningar

Genrebaserad position

Signalord

Länkar i semantisk  
struktur

Diskursstruktur

[ICA-handlaren Harry Franzen i Röstånga i Skåne bröt på nyårsdagen under stort tumult spritmonopolet.](#) Men efter bara tio minuter kom polisen och tog vinet i beslag, stängde affären och grep Franzen, den förste livsmedelshandlaren i landet som försökt starta vinförsäljning.

[En flaska han dock Franzen sälja till högstbjudande innan polisen stoppade kommersen.](#) Resten av vinflaskorna påstod Franzen att han skänkte bort till kunderna. - En flaska per person. Tänk på nykterheten, sade han innan kunderna stormade in i affären.

[Harry Franzen, etablerad kommunpolitiker för centrumdemokraterna i Svalövs kommun, hade inte gjort någon hemlighet av att han tänkte bryta mot lagen den dag EES-avtalet började gälla i Sverige.](#) Ett stort pressuppbåd och många bybor från den lilla skånska orten mitt i Skåne var också på plats när butiken öppnade klockan elva på förmiddagen.

[Franzen vill få prövat i domstol om Systembolagets monopol bryter mot EES-avtalets konkurrensregler eller ej.](#) - Jag kommer att vinna processen. Det är jag övertygad om. Men det kommer att ta tid, sade Franzen, för dagen finkläd med slips, väst och skärmmössa, innan han motades in i en civil polisbil för att köras till Landskrona för förhör.



## Genrebaserad position

ICA-handlaren Harry Franzen i Röstånga i Skåne bröt på nyårsdagen under stort tumult spritmonopolet. Men efter bara tio minuter kom polisen och tog vinet i beslag, stängde affären och grep Franzen, den förste livsmedelshandlaren i landet som försökt starta vinförsäljning. En flaska han dock Franzen sälja till högstbjudande innan polisen stoppade kommersen. Resten av vinflaskorna påstod Franzen att han skänkte bort till kunderna. - En flaska per person. Tänk på nykterheten, sade han innan kunderna stormade in i affären. Harry Franzen, etablerad kommunpolitiker för centrumdemokraterna i Svalövs kommun, hade inte gjort någon hemlighet av att han tänkte bryta mot lagen den dag EES-avtalet började gälla i Sverige. Ett stort pressuppbåd och många bybor från den lilla skånska orten mitt i Skåne var också på plats när butiken öppnade klockan elva på förmiddagen.

Översikt

Informationssökning

Informationsextraktion

Textsammanfattning

Extraktexempel

Exempeltext

Frekvensanalys

Ord i titel och rubriker

Kärnmeningar

Genrebaserad position

Signalord

Länkar i semantisk  
struktur

Diskursstruktur



# Signalord

+	-
sammanfattningsvis resultat understryker att slutsatsen är syftet är viktigaste bästa bättre /.../	för övrigt inom parentes exempelvis (...) i förbigående bl.a. inte han /.../

Översikt

Informationssökning

Informationsextraktion

Textsammanfattning

Extraktexempel

Exempeltext

Frekvensanalys

Ord i titel och rubriker

Kärnmeningar

Genrebaserad position

Signalord

Länkar i semantisk  
struktur

Diskursstruktur



# Länkar i semantisk struktur

## VINSÄLJANDE HANDLARE BRÖT MONOPOLET OCH GREPS AV POLISEN SVALÖV TT

ICA-handlaren **Harry Franzen** i **Röstånga** i **Skåne** bröt på nyårsdagen under **stort** tumult spritmonopolet. Men efter bara tio minuter kom **polisen** och **tog** **vinet** i beslag, **stängde** **affären** och **grep** **Franzen**, den förste livsmedelshandlaren i landet som försökt starta vinförsäljning.

En **flaska** han dock **Franzen** **sälja** till högstbjudande innan **polisen** stoppade kommersen. Resten av vinflaskorna påstod **Franzen** att han skänkte bort till **kunderna**. - En **flaska** per person. Tänk på nykterheten, sade han innan **kunderna** stormade in i **affären**.

**Harry Franzen**, etablerad kommunpolitiker för centrumdemokraterna i Svalövs kommun, hade inte gjort någon hemlighet av att han tänkte bryta mot **lagen** den dag **EES-avtalet** började gälla i Sverige. Ett **stort** pressuppbåd och många bybor från den lilla skånska orten mitt i **Skåne** var också på plats när butiken öppnade klockan elva på förmiddagen.

**Franzen** vill få prövat i **domstol** om **Systembolagets** monopol bryter mot EES-avtalets konkurrensregler eller ej. - Jag kommer att vinna **processen**. Det är jag övertygad om. Men det kommer att ta tid, sade **Franzen**, för dagen finklädd med slips, väst och skärmmössa, innan han motades in i en civil polisbil för att köras till **Landskrona** för **förhör**.

Harry Franzen polisen flaskan affären kunderna Landskrona processen Röstånga sälja tog vinet förhör stängde grep domstol lagen EES-avtalet Systembolagets

Översikt

Informationssökning

Informationsextraktion

Textsammanfattning

Extraktexempel

Exempeltext

Frekvensanalys

Ord i titel och rubriker

Kärnmeningar

Genre baserad position

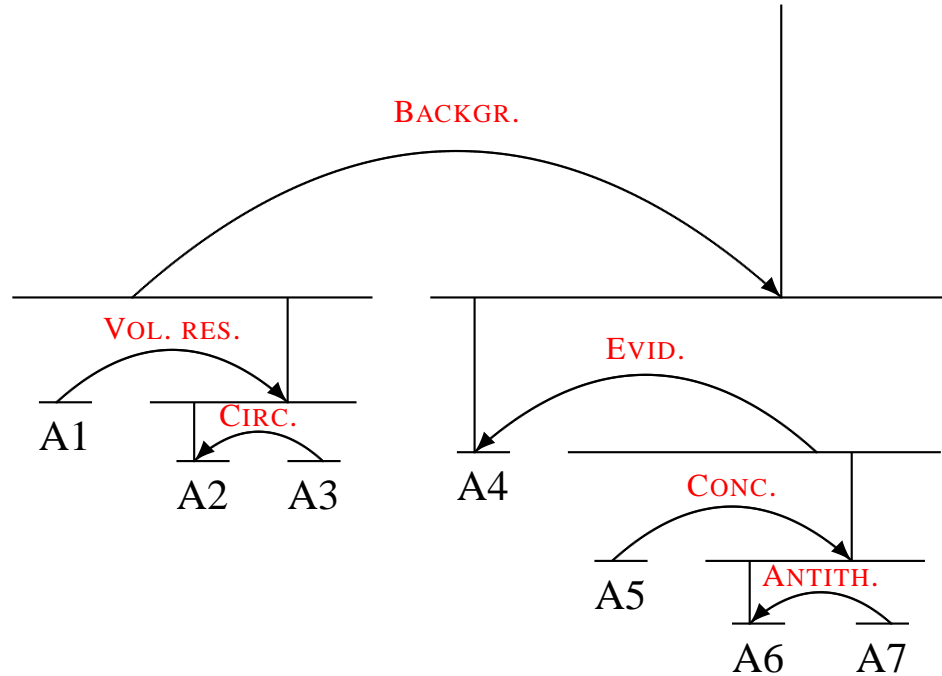
Signalord

Länkar i semantisk struktur

Diskursstruktur



# Diskursstruktur



[Farmington police had to help control traffic recently]<sup>A1</sup> [when hundreds of people lined up to be among the first applying for jobs at the yet-to-open Marriott Hotel.]<sup>A2</sup> [The hotel's help-wanted announcement – for 300 openings – was a rare opportunity for many unemployed.]<sup>A3</sup> [The people waiting in line carried a message, a refutation, of claims that the jobless could be employed if only they showed enough moxie.]<sup>A4</sup> [Every rule has exceptions,]<sup>A5</sup> [but the tragic and too-common tableaux of hundreds or even thousands of people snake-lining up for any task with a paycheck illustrates a lack of jobs,]<sup>A6</sup> [not laziness.]<sup>A7</sup>

(*The Hartford Courant*, editorial)