

# Er kan meer dan men doet



# n̂ | STEVIN

# Overzicht

1. Overzicht van het projekt
2. Structuur en inhoud van de database
3. Taaltechnologische mogelijkheden
4. Gebruikersscenario's
5. Toekomstscenario's

# Doel van Cornetto (STE05039)

- Een lexicale semantische database voor het Nederlands:
  - 40K lemma's
  - Generische en centrale woorden
- Data:
  - Lijkt op een combinatie van WordNet en FrameNet
  - Verticale en horizontale semantische relaties
  - Combinatorische lexicale constraints
  - Gekoppeld aan het English Wordnet
  - Uitgebreid met een ontologie
- Automatische acquisitie toolkit
- Startdatum is april-2006, einddatum is april-2008
- <http://www.let.vu.nl/onderzoek/projectsites/cornetto/start.htm>



relaties  
chronisch zieke, langdurig zieke,  
psychisch/geestelijk zieke



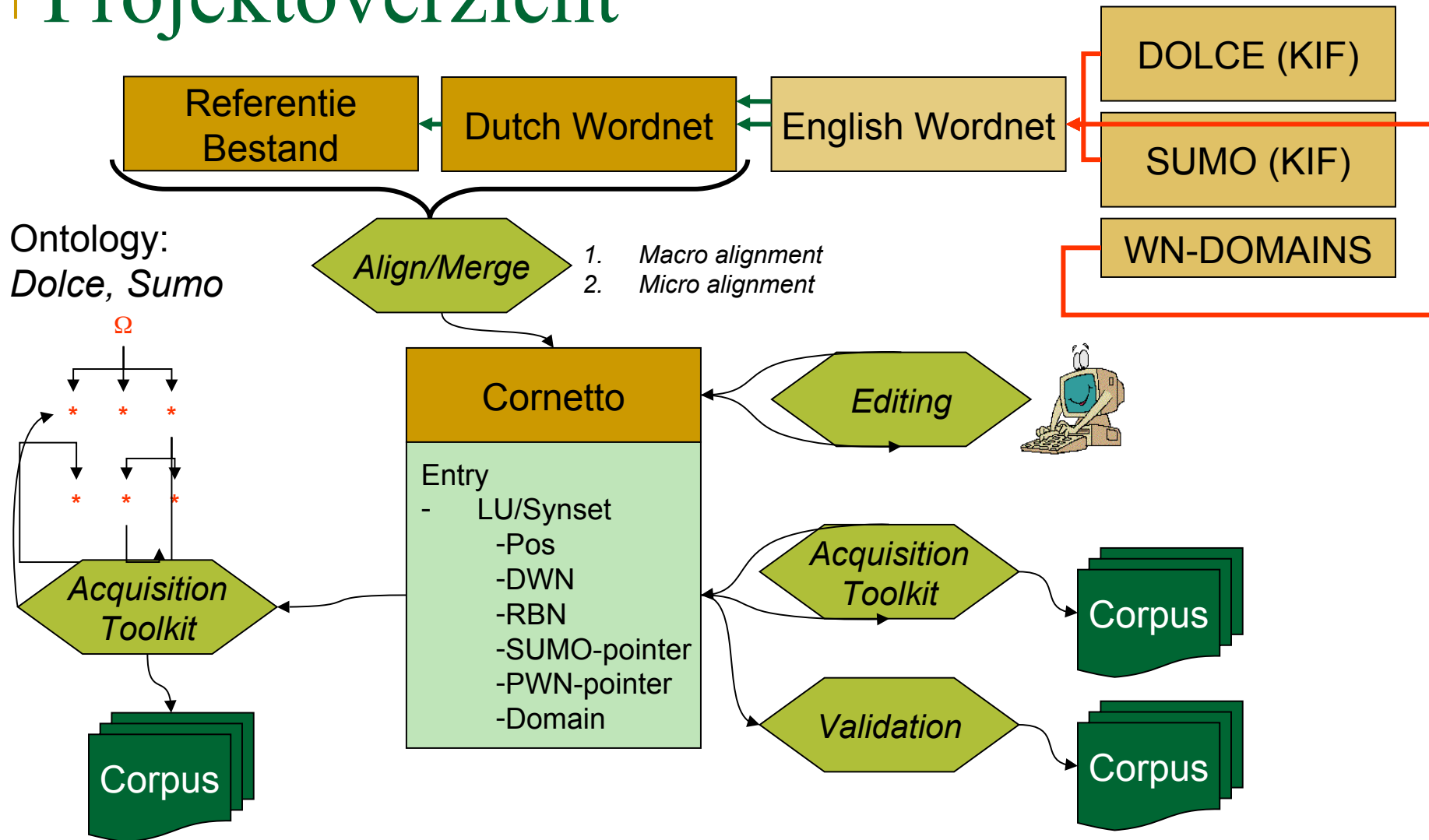
# Combinatorics

| slots  | fillers (lex/conc)                   | fillers (coll)                                                                 |
|--------|--------------------------------------|--------------------------------------------------------------------------------|
| action | <i>behandelen</i><br>(someone treat) | <i>iem. behandelen</i>                                                         |
| theme  | <i>patiënt</i><br>(a patient treat)  | <i>een patiënt behandelen</i>                                                  |
| state  | <i>ziekte</i>                        | <i>iem. behandelen <u>voor</u> een ziekte</i><br>(someone treat for a disease) |
|        | <i>behandelen</i>                    | <i>iem. <u>aan</u> zijn verwondingen</i><br>(somene at his injuries treat)     |
|        |                                      | <i>een ziekte behandelen (a disease treat)</i>                                 |

# Methode

- Combineren van de informatie van twee bestaande Nederlandse lexicale databases:
  - Het Dutch wordnet (DWN): synsets en lexicale semantische relaties
  - Het Referentiebestand Nederlands (RBN): morpho-syntactische informatie, semantiek, pragmatiek, framestructuren, lexicale functies en combinatorics
- Macro level alignment: welke betekenis uit het DWN hoort bij welke betekenis uit het RBN?
- Micro level alignment: wat zijn de afhankelijkheden van de informatie bij het DWN van de informatie uit het RBN
- Voorzien van een ontologisch raamwerk

# Projectoverzicht



# Consortium

- Vrije Universiteit Amsterdam, Faculteit der Letteren, General Linguistics Department, Onderzoeksgroep Lexicologie/Terminologie:
  - Willy Martin, Isa Maks, Hennie vd Vliet, Roxane Segers, Piek Vossen
- Universiteit van Amsterdam, Instituut voor Informatica:
  - Maarten de Rijke, Erik Tjong Kim Sang, Katja Hofmann
- Katholieke Universiteit Leuven, Interdisciplinair Centrum voor Recht en Informatica (ICRI):
  - Sien Moens, Jan de Beer
- Irion Technologies BV:
  - Joop van Gent, Hetty van Zutphen, Piek Vossen

# Andere partners

## ■ Gebruikersgroep:

- ❑ Polderland
- ❑ Knowledge Concepts
- ❑ LibRT
- ❑ Irion Technologies
- ❑ Van Dale Lexicografie
- ❑ Larcier-De Boeck

## ■ Ontologiegroep:

- ❑ Dr. W. Ceusters, Office Line Engineering nv
- ❑ Prof. F. van Harmelen, Vrije Universiteit Amsterdam
- ❑ Dr. P. Buitelaar, DFKI
- ❑ Dr. P. Monachesi, Universiteit van Utrecht

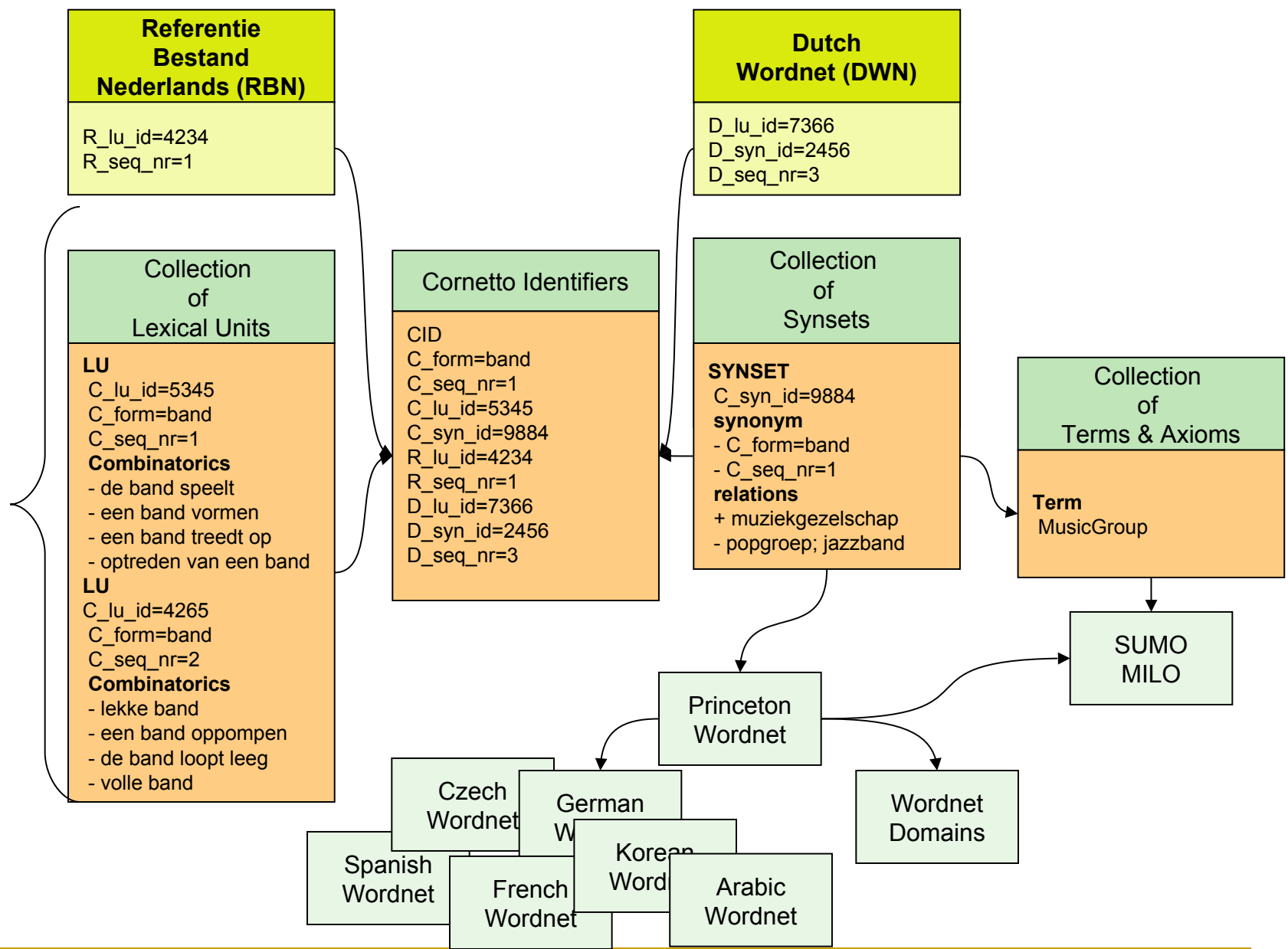
# Structuur en inhoud van de database

---

# Lexical Units, Synsets en Termen

- Lexical Unit = vorm-betekenis relatie, zodanig dat:
  - vorm = abstracte representatie van bepaalde realisaties;
  - zelfde woordsoort;
  - zelfde betekenis, aangegeven door de Synset waar het toebehoort;
- Synset = set van synoniemen (=Lus, bijv. *fiets* en *rijwiel*) die naar dezelfde entiteiten kunnen verwijzen in de meeste contexten.
  - Gedefiniëerd door lexicale semantische relaties;
  - Gedefiniëerd door een verwijzing naar de ontologie Termen of een KIF expressie met die Termen;
- Term = verzameling van klassen met formele definitie in ontologisch raamwerk

# Cornetto Database (CDB)



### canonical form

in een band spelen  
een band oprichten  
de band speelt

### canonical form

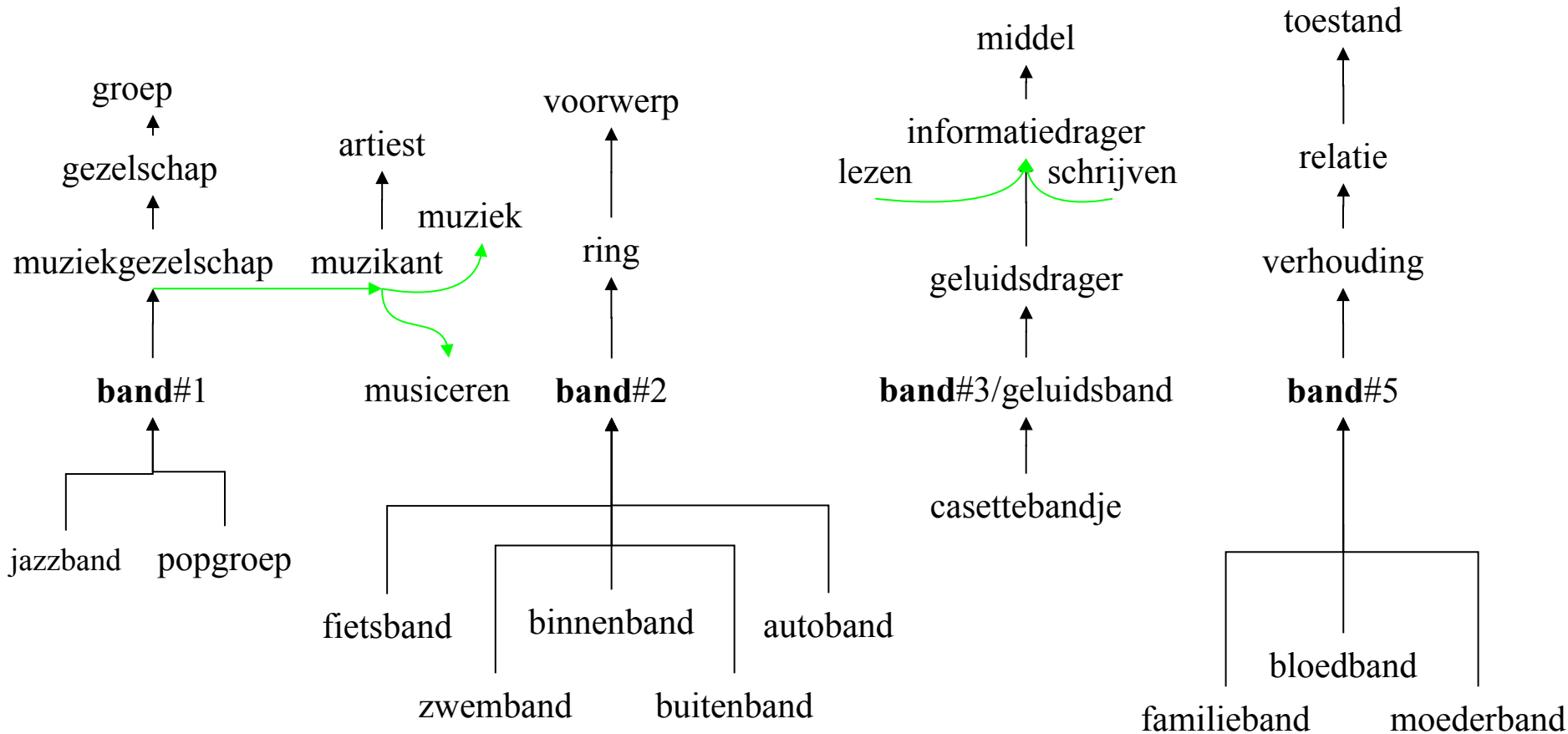
de band oppompen  
een band plakken  
een leuke band  
de band springt

### canonical form

de band starten  
op de band [zetten/opnemen/staan]  
de band afspelen  
iets op de band opnemen

### canonical form

een goede/sterke band  
de banden verbreken  
een band hebben met iemand



# Semantiek voor framestructuren

- Event structure voor werkwoorden in RBN:
  - ❑ E: behandelen <e0> action
  - ❑ A1: <?a1> pers
  - ❑ A2: <?a2>
  - ❑ C3: <?c3> prep
  - ❑ iemand aan [zijn verwondingen] behandelen
  - ❑ een patiënt voor [een nieraandoening/puistje/keelpijn] behandelen
  - ❑ iemand met [fysiotherapie/medicijnen]<sub>instrument</sub> behandelen
- DWN:
  - ❑ [causes] [v] genezen:2, beteren:1, herstellen:1
  - ❑ [involved\_agent] [n] arts:1; dokter:1 <?a1>
  - ❑ [involved\_patient] [n] zieke:1; patiënt:1 <?a2>
  - ❑ [involved\_instrument] [n] hart-longmachine:1 <?c3>
  - ❑ [involved\_instrument] [n] mitella:1, draagdoek:1 <?c3>
  - ❑ [involved\_instrument] [n] geneesmiddel:1; medicijn:1 <?c3>
  - ❑ etc...

# Wat hebben we aan de combinatie RBN en DWN?

- Informatie die ontbreekt in de een kan worden afgeleid uit de ander;
- Argumentstructuren en frames van RBN worden verrijkt met semantiek => **bepalen van de betekenis van woorden in teksten**;
- Hierarchie toegevoegd aan combinatoriek;
- Synsets in DWN worden verrijkt met contextuele informatie: => **bepalen van de betekenis van woorden in teksten**;
- Synsets in DWN worden verrijkt met lexicale en grammaticale selectiecriteria: => **tekstgeneratie**;

# Ontologisering van Cornetto

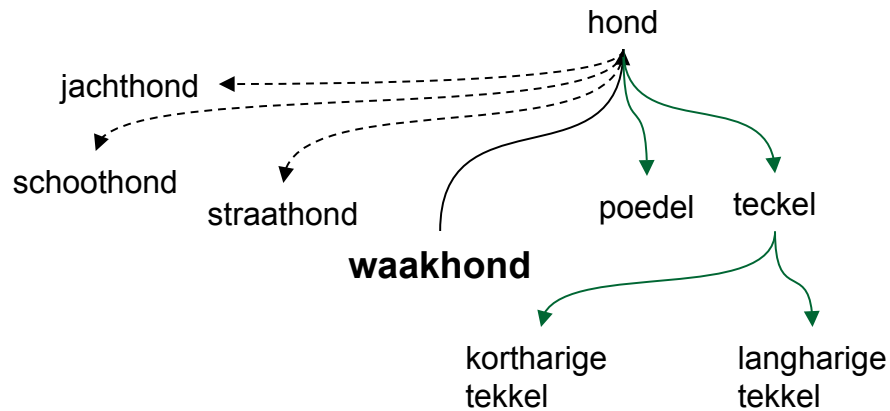
- Identiteitscriteria OntoClean (Guarino & Welty 2002), :
  - rigiditeit: in wat voor mate zijn eigenschappen van entiteiten waar in alle mogelijke werelden? Een *mens* ben je altijd, een *student* kun je tijdelijk zijn.
  - essentie: welke eigenschappen zijn essentieel voor een entiteit? *Vorm is essentieel voor een beeld en niet voor de klei waaruit een beeld bestaat.*
  - uniciteit: wat vormt een geheel en welke dingen zijn een onderdeel van een geheel? Een *zee* is een geheel maar *water* niet.
- Hyponiemen van *hond* in DWN:
  - bokser; corgi; loboor; mopshond; pekinees; pointer; spaniël;
  - pup; reu; teef
  - bastaard; straathond; blindengeleidehond; bullebijter; diensthond; gashond; jachthond; lawinehond; schapendoder; schoothondje; waakhond

# Toegepast op DWN

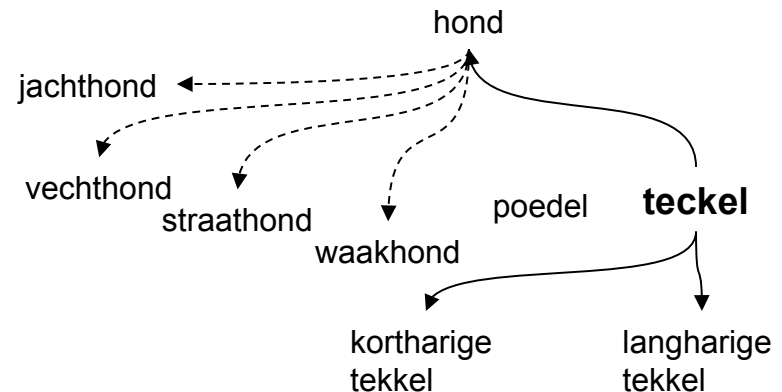
- (Semi-)rigide type hiërarchie in ontologie:
  - Canine => PoodleDog; NewfoundlandDog; DalmatianDog, etc.
- Wordnet bestaat dan uit namen voor (semi-)rigide types en andere woorden voor *honden* met rollen:
  - poedel = PoodleDog
  - jachthond (?CAN)
  - ⇒ (exists (?CAN ?EV)  
(and  
(instance ?CAN Canine)  
(instance ?EV Hunting)  
(agent ?CAN ?EV)))
- Type hiërarchie blijft compact en zuiver

# Expansie met zuivere hyponymierelaties

Expansie van rol naar types



Expansie van types naar rollen



# Wat hebben we aan de ontologisering van Cornetto?

- Meer expliciete definitie van hiërarchische structuren en rollen: => **betere semantische expansie**;
- Doorrekenen van semantische relaties levert meer exacte semantische database op;
- Maakt het mogelijk om met de database te **redeneren**;
- Maakt het mogelijk om domeinspecifieke databases af te leiden: => **nodig voor semantic web applicaties**;

# Taaltechnologische mogelijkheden

---

# Tekstanalyse

- Disambiguering van woordbetekenissen in teksten:
  - horizontale relaties
  - frame structuren met combinatorische constraints
  - voorbeelden
- Meer verfijnde expansie naar varianten op grond van gedifferentieerde hyponymie-structuren;
- Betere definitie van cohesie in teksten omdat lexicale inclusie wordt herkend:
  - knippen -> schaar
  - genezen -> ziekte, verwonding
- Maken van inferenties, bijvoorbeeld tijdsafhankelijkheden:
  - geboorte -> huwelijk -> scheiding -> sterven
  - oprichten -> muzikoptreden -> uit elkaar gaan

# Tekstgeneratie

- Selectie van gangbare combinaties uit alternatieven die worden geboden door een wordnet of de ontologie:
  - drank -> prepareren; maken; brouwen
  - limonade maken; koffie zetten; drankje brouwen
  - smaakvolle soep; pittig eten; sterke koffie;
- Realisatie van conceptuele relaties in syntactisch correcte constructies:
  - behandelen aan verwondingen
  - behandelen voor een ziekte

# Gebruikersscenario's

# Toepassingen in Cornetto

- Automatische acquisitie van nieuwe concepten en relaties van tekstcorpora;
- Uitbreiding van de database (woorden en ontologie) naar juridisch domein;
- Verbetering van performance van bestaande systemen;
  - Q&A
  - Multimedia en tekst retrieval
  - Dialoogsystemen

# Q&A systeem UVA

- Gebruikt wordnet synsets voor:
  - Classificatie van vragen;
  - Parafrasering van vragen;
- Leidt expansie via disjuncte types tot beter resultaat?
- Coverage belangrijker dan diepe analyse van rijke data;
- Exploitatie van de rijkdom in Q&A system is meer onderzoeksonderwerp voor PhD project

# CHOICE & MunCH

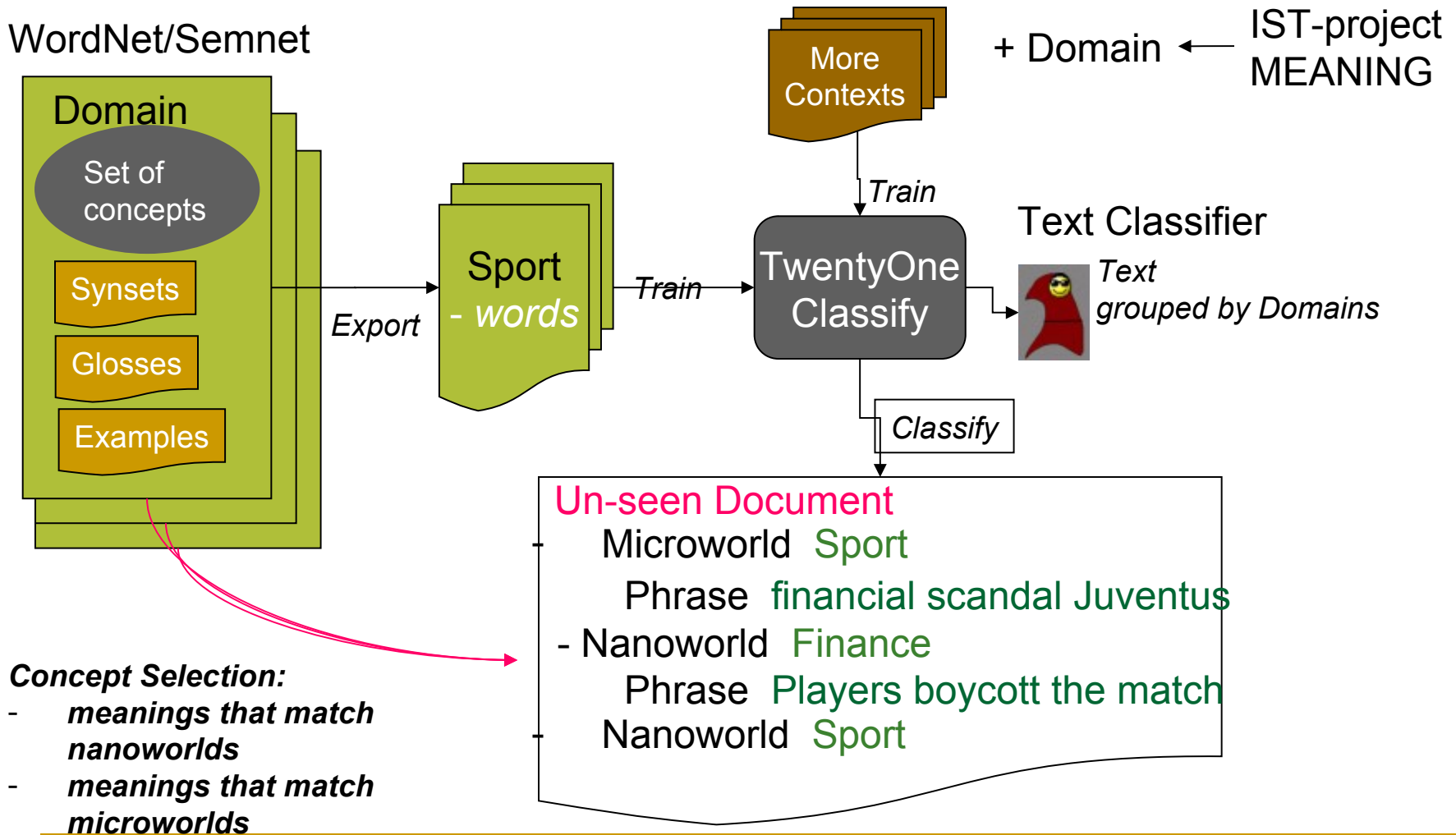
- Onderdeel van CATCH research programma;
- Toegankelijk maken van cultureel erfgoed;
- Semi-automatische semantische annotatie van gearchiveerde objecten zoals video, images en boeken;
- Categorieën van gestandaardiseerde metadata collecties: domein thesauri en ontologieën;
- Doel: ondersteunen van zoeken:
  - query disambiguering: "bed" om te slapen onderscheiden van "river beds";
  - query generalisatie of specialisatie: vinden van foto's van "crib" als je zoekt naar een "bed" om te slapen;
- Mediamill semantic video search engine
- <http://www.nwo.nl/CATCH/CHOICE>
- <http://ilps.science.uva.nl/munch/index.html>

# Text retrieval

- Evaluatie raamwerk van EU-project MEANING (IST-2001-34460)
- Queries met ambiguïteit en parafraseringen worden gextraheerd, e.g. "police cell" (jail), "cell phone" (mobile), "nerve cell" (neuron).
- Test op verschillende indexen:
  - Index van letterlijke tekst
  - Index van tekst geëxpandeerd met wordnet
  - Index van tekst geëxpandeerd met wordnet na bepalen van betekenis
- Betekenis wordt bepaald door woorden aan domeinen te koppelen en domeinen aan wordnet (Magnini et al 2000): hoge precisie en lage recall;
- Cornetto wordt gebruikt om de recall te vergroten: frame-achtige relaties als er geen domein van toepassing is;



# Domain-based WSD (*IRST-Trento, Magnini 2002*)



# Effectiviteit van Domein desambiguering

|                                     | Spanish          |               | English        |               |
|-------------------------------------|------------------|---------------|----------------|---------------|
| <b>total concepts</b>               | <b>2,769,753</b> |               | <b>403,124</b> |               |
| <b>disambiguated in microworlds</b> | <b>220,574</b>   | <b>7,96%</b>  | <b>18,541</b>  | <b>4,60%</b>  |
| <b>disambiguated in nanoworlds</b>  | <b>1,691,079</b> | <b>61,06%</b> | <b>31,4394</b> | <b>77,99%</b> |
| <b>unaffected concepts</b>          | <b>858,100</b>   | <b>30,98%</b> | <b>70,189</b>  | <b>17,41%</b> |

- 2<sup>nd</sup> Level domains(163 -> 57);
- NPs classified in a window of 10 NPs;
- Threshold was set to 60;

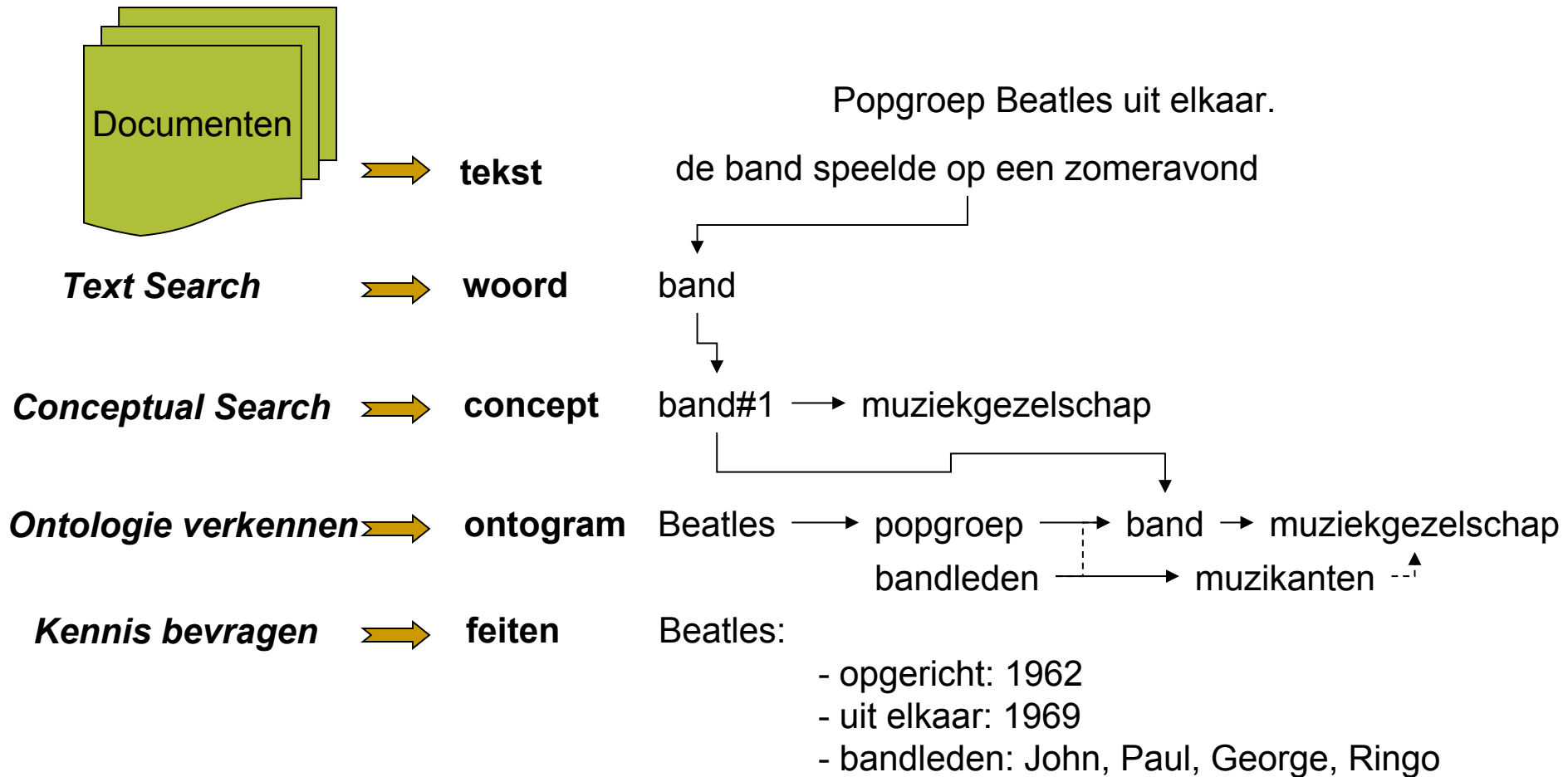
# Toekomstscenario's

---

# Ontluikende toepassingen

- Kennisontginning
  - Automatisch afleiden van ontologie
  - Controlled Inferencing
- Cooperatieve dialoogsystemen:
  - Relaties tussen informatie zijn belangrijk
  - Vaagheid en ambiguïteit worden
- Samenvattingen maken met betere cohesie
- Tekstgeneratie

# Kennisonontginning



TaxEdit

File Help

Projects

Burgerloket

Burgerloket\_1: 2006.11.28 AD at 08:13:25 PM CET, docs: 156, nPages=748:682 :Legal:

- 1#Aangifte overlijden.txt# nPages=3# Terms: nl:42 :Legal:0.89#Medical:0.82#;
- 2#Aangifte trouwen.txt# nPages=6# Terms: nl:54 :Legal:0.9#;
- 3#Aansprakelijkheid gemeente.txt# nPages=4# Terms: nl:33 :Legal:0.89#Finance:0.82#;
- 4#Aanwijzen monument.txt# nPages=3# Terms: nl:42 :Legal:0.84#Society:0.83#Architectuur:0.82#;
- 5#Adoptie.txt# nPages=5# Terms: nl:23 :Legal:0.88#Society:0.86#;
- 6#Adresgegevens.txt# nPages=3# Terms: nl:16 :Society:0.88#Government:0.82#Transport:0.82#;
- 7#Adresgegevens De Wegwerker.txt# nPages=10# Terms: nl:13 :Society:0.92#Transport:0.82#;
- 8#Adresgegevens gemeente Gilze.txt# nPages=2# Terms: nl:3 :Farming:0.82#Architectuur:0.82#;
- 9#Adresgegevens gemeentehuis.txt# nPages=5# Terms: nl:10 :Society:0.91#Architectuur:0.82#;
- 10#Adresgegevens milieustraat.txt# nPages=7# Terms: nl:9 :Environment:0.88#Society:0.82#;
- 11#Adresgegevens Villa Mol.txt# nPages=4# Terms: nl:7 :Government:0.88#Society:0.82#;
- 12#Afgifte rijbewijs.txt# nPages=3# Terms: nl:63 :Legal:0.88#;
- 13#Afval op riolering.txt# nPages=3# Terms: nl:23 :Environment:0.85#Legal:0.83#Society:0.82#;
- 14#Afval storten.txt# nPages=3# Terms: nl:18 :Environment:0.8#;
- 15#Afvalstoffenheffing.txt# nPages=6# Terms: nl:16 :Legal:0.83#Finance:0.82#;
- 16#Alarm installatie.txt# nPages=10# Terms: nl:22 :Legal:0.87#Technology:0.83#;
- 17#Asbestverwijdering.txt# nPages=3# Terms: nl:9 :Materials:0.81#Geology:0.81#;
- 18#Automatische incasso.txt# nPages=10# Terms: nl:17 :Finance:0.89#Legal:0.81#;
- 19#Avondwinkel.txt# nPages=3# Terms: nl:15 :Commerce:0.89#Legal:0.84#;
- 20#Belasting betalingsregeling.txt# nPages=20# Terms: nl:32 :Finance:0.88#Legal:0.82#;
- 21#Belasting kwijtschelding.txt# nPages=25# Terms: nl:70 :Finance:0.88#Legal:0.82#;
- 22#Beroepschrift.txt# nPages=5# Terms: nl:15 :Legal:0.92#;
- 23#Bewijs Nederlanderschap.txt# nPages=3# Terms: nl:31 :Legal:0.87#Government:0.82#;
- 24#Bewijs van in leven zijn.txt# nPages=3# Terms: nl:32 :Legal:0.89#Society:0.85#Finance:0.82#;
- 25#Bezwaarschrift belasting.txt# nPages=12# Terms: nl:44 :Legal:0.88#Finance:0.82#;
- 26#Bezwaarschrift tegen de gemeente.txt# nPages=3# Terms: nl:37 :Legal:0.89#Commerce:0.82#;
- 27#Bibliotheek.txt# nPages=3# Terms: nl:14 :Society:0.81#;
- 28#Bijstand voor zelfstandigen.txt# nPages=5# Terms: nl:43 :Society:0.86#Finance:0.82#;
- 29#Bijstandsuitkering.txt# nPages=4# Terms: nl:39 :Society:0.89#Finance:0.84#Legal:0.82#;
- 30#Bijzondere bijstand.txt# nPages=4# Terms: nl:20 :Legal:0.85#Society:0.85#Finance:0.82#;
- 31#Bingo organiseren.txt# nPages=3# Terms: nl:20 :Legal:0.84#Society:0.82#Finance:0.82#;
- 32#Bodemvervuiling.txt# nPages=8# Terms: nl:12 :Environment:0.86#;
- 33#Bouwvergunning.txt# nPages=4# Terms: nl:35 :Legal:0.86#Society:0.81#Architectuur:0.82#;

Burgerloket\_1

Legal:0.89#Medical:0.82# Tops:173, Nodes:682, Docs:1326

Import View Edit Save Sort Clear

vergunning:0:76

- gemeente vergunning:1:4:0.8:4
- waterschap vergunning:3:2:0.8:2
- ventenvergunning:2:1:0.8:1
- jachtvergunning:6:2:0.8:2
- inritvergunning:2:3:0.8:7
  - inritvergunning geweigerd:2:2:0.8:2
  - inritvergunning verleend:2:2:0.8:2
- dergelijke vergunning:3:4:0.8:4
- vergunning vermelde:2:2:0.8:2
- terrasvergunning:28:3:0.8:3
- ventvergunning:8:2:0.8:4
  - ventvergunning geldig:6:2:0.8:2
- collectevergunning:2:2:0.8:2
- milieuvergunning:3:3:0.19:3
- bouwvergunning:3:6:0.3:8
  - plan bouwvergunning:1:2:0.8:2
- loterijvergunning:7:1:0.8:1
- vergunning afgegeven:2:2:0.8:2
- afdeling Vergunningen:2:6:0.8:6
- kapvergunning:1:2:0.8:2
- gewijzigde vergunning:2:2:0.8:2
- gebruiksvergunning:3:3:0.8:3
- horecavergunning:10:3:0.8:3
- aanlegvergunning:6:3:0.8:3
- voorwaarden vergunning:2:2:0.8:2
- Gemeentelijke vergunning:3:2:0.75:2
- nieuwe veraanwinst:3:2:0.38:2

Done loading thesaurus...

100%

TaxEdit

File Help **IRION**

Projects

Scope Knowledge

Scope Knowledge\_1: 2006.10.31 AD at 10:25:30 AM CET, docs: 113, nPages=8298:64

1## nPages=0

2#alcnom# nPages=425# Terms: en:12694 :Chemistry:Physics;;

3#http://pubchem.ncbi.nlm.nih.gov# nPages=33# Terms: en:1102 :Government:Co

4#http://chemfinder.cambridgesoft.com/# nPages=59# Terms: en:222 :Communica

5#http://full.chemistry.uakron.edu/erd/index.html# nPages=3# Terms: en:16 :Commu

6#http://chem.sis.nlm.nih.gov/chemidplus/# nPages=23# Terms: en:236 :Chemistry

7## nPages=2595

8#http://chembank.broad.harvard.edu# nPages=1# Terms: en:5 :Education:Society;

9#http://www.citychemical.com/SearchResults.asp# nPages=7# Terms: en:18 :Com

10#http://www.novachem.kiev.ua/eng/wcp.php3?sb=C&pagenum=5&sf=# nPages=

11#http://www.dsl-chem.com/all\_products# nPages=2# Terms: en:4 :Commerce:Le

12#http://www.wychem.com/catalogues/anilinesamidesandamines.htm# nPages=2

13#http://webbook.nist.gov/chemistry/# nPages=80# Terms: en:1090 :Physics:Tech

14#http://www.omegachem.com/en/Product\_list.asp# nPages=33# Terms: en:113 :

15#http://www.alfa.com/alf/Product%20Indexes/Alfa\_Complete/C10\_idx.html# nPag

16#http://www.chembuyersguide.com# nPages=1# Terms: en:21 :Topography:Com

17#http://infochems.com/chemdb/chem\_list.asp# nPages=1# Terms: en:2 :Chemis

18## nPages=0

19## nPages=0

20#ChemSearch# nPages=6# Terms: en:367 :Chemistry:Medical:Bio;;

21## nPages=783

22## nPages=192

23## nPages=0

24## nPages=607

25#http://www.made-in-china.com/products/catlist/listsubcat/114/00/mic/Chemicals

26## nPages=0

27## nPages=0

28#http://www.osha.gov/SLTC/healthguidelines/index.html# nPages=3# Terms: en:

29## nPages=28

30## nPages=76

31## nPages=61

32## nPages=0

33## nPages=10

Scope Knowledge\_1

Generic# Tops:1041, Nodes:6446, Docs:13018

Import View Edit Save Sort Clear

acid::0:445

- acetoacetic acids;acetoacetic acid;acetoacetic acid shown:6:1:0.8:1
- acidic character of carboxylic acid;acidic character of carboxylic:4:2:0.8:2
- fluoroacetic acid acidic:2:2:0.8:2
- comparing acidities;comparing acid:3:3:0.8:3
- isomeric hydroxybenzoic acids:2:2:0.8:2
- stereoisomeric 2-butenedioic acids:2:2:0.8:2
- ascorbic acid;Ascorbic acid:4:2:0.8:2
- aldaric acid;Aldaric Acid;aldaric acids:6:1:0.8:1
- 2-Methylpentanoic Acid:2:1:0.8:1
- Benzoic Acid;Benzoic acid;benzoic acid;benzoic acids;benzoic acid's;BENZOIC ACID;Benzoic acid:4:
- Ginkgolic Acid:9:1:0.8:1
- aminoacids:2:1:0.8:1
- methyl decreases acidity of acid ten fold:2:2:0.8:2
- 2-ethylhexanoic acid;2-Ethylhexanoic Acid:3:2:0.8:2
- linoleic acid:2:2:0.8:2
- ortho-toluic acid:2:2:0.8:2
- Electron withdrawing substituents increase acidity of benzoic acid;Electron withdrawing substituents
- vigorous acid:2:1:0.8:1
- Hypobromous Acid;hypobromous acids;hypobromous acid:3:2:0.8:2
- N-Acetylneuraminic Acid:4:1:0.8:1
- dibasic acids:6:3:0.8:3
- malonic acid;malonic acids;Malonic acid:9:2:0.8:2
- aqueous acid;Aqueous acid:23:14:0.8:14
- aspartic acid;Aspartic Acid;aspartic acid shown:12:3:0.8:3
- aldonic acid;Aldonic Acid;aldonic acids:3:1:0.8:1
- myristic acid:2:2:0.8:2

Done loading thesaurus...

100%



# Beperkingen van gestructureerde data

- Taalgebruik van grote publiek past niet op het taalgebruik van de informatie:
  - Regelgeving:
    - **Voertuigen** op de **openbare weg** moeten worden voorzien van een geldig **kentekenbewijs** en duidelijke **markeringen**;
  - Gebruikersvraag:
    - Ik heb een **old-timer** die ik haast nooit gebruik. Moet die ook een **kentekenplaat** en **lichten** hebben als ik die op **straat** parkeer?
- Eindgebruikers willen niet een complexe indeling volgen/leren/kennen maar hun eigen perspectief

## Bacheloropleidingen

### Bacheloropleiding ACW

#### Bachelorprogramma

#### Beroepsperspectieven

De kunst van het vergelijken

In deeltijd

Interessante links

Introductie

Kosten

Medewerkers

Meer informatie

Nieuws

Onderzoek

Stages

Studiebegeleiding

Studievereniging

Toelating en inschrijving

### Bacheloropleiding Literatuur

#### Bachelorprogramma

#### Beroepsperspectieven

Contact

Frame

In deeltijd

Interessante links

Introductie

Kosten

Medewerkers

Meer informatie

Tijdschrift frame

Toelating en inschrijving

Vakkenoverzicht

## Masteropleidingen

### Masteropleiding ACW

#### Beroepsperspectieven

In deeltijd

Interessante links

Introductie

Kosten

Masterprogramma

Medewerkers

Meer informatie

Onderzoek

Stages

Studievereniging

Toelating en inschrijving

### Masteropleiding Literatuur

#### Beroepsperspectieven

Contact

In deeltijd

Interessante links

Introductie

Kosten

Masterprogramma

Medewerkers

Meer informatie

Onderzoek

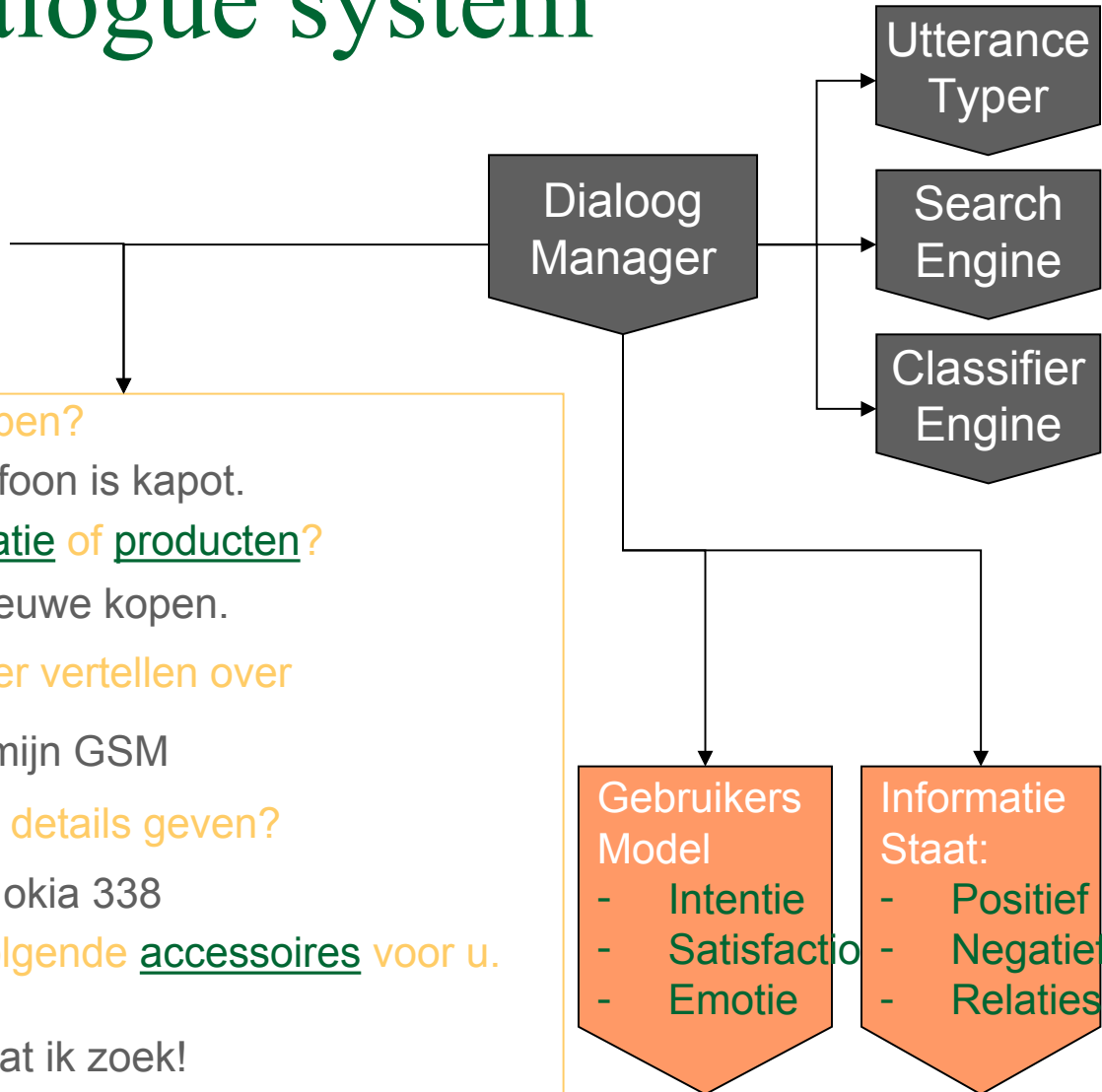
Tijdschrift frame

Toelating en inschrijving

# Communicatief dialoogsysteem

- Toegang tot informatie en diensten:
  - ongeacht woordkeuze
  - ongeacht de structurering van de informatie
  - eventueel met gebruikmaking van die structurering
- Werkt samen met een gebruiker:
  - Vraagt de gebruiker om help, instructies, bevestiging, uitleg
- Gebruikt 4 informatielagen:
  - De **intentie** van de gebruiker: *klacht indienen, producten kopen, support, informatie*
  - De mate waarin iemand tevreden is (**satisfaction rate**)?
  - De **emotionele staat** van de gebruiker; *is iemand boos, vrolijk, vriendelijk?*
  - De **informatie staat** gebaseerd op de inhoudelijke beschrijving die een gebruiker geeft van een informatiebehoefte; *Waar is iemand naar opzoek?*

# Dialogue system



Frases

Concepten

{ informatie 

{ producten  
mobiel  
accessoires  
koptelefoon 

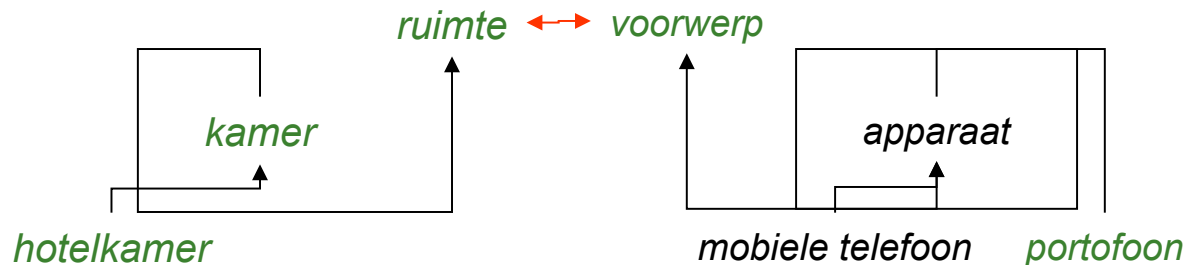
{ reparatie 

- Kan ik u helpen?
- Mijn koptelefoon is kapot.
- Wilt u reparatie of producten?
- Ik wil een nieuwe kopen.
- Kunt u meer vertellen over producten?
- Het is voor mijn GSM
- Kun je meer details geven?
- Het is een Nokia 338
- Ik heb de volgende accessoires voor u.
- Bekijk ze eens.
- Dat is niet wat ik zoek!

# Communicatief dialoogsysteem

## ■ Voorkomt deadlocks:

- Detecteert vaagheid en ambiguïteit (welke betekenis van *band*?)
- Detecteert veranderingen van onderwerp
- Gebruikt negatieve informatie: "*Geen muziekbond, ik zoek fietsbanden!*"
- Kan out-of-domain vragen aan:
  - "*We hebben geen hotelkamers maar wel elektronische apparaten*".
  - "*Nee, we hebben geen portofoons maar wel andere elektronische apparaten zoals mobiele telefoons*"





# The end....