

# Nederlandse Taalunie & Taal - en spraaktechnologie

Elisabeth D'Halleweyn  
Nederlandse Taalunie

- TST:wat is het?
- TST: wat kan je ermee?
- TST: hoe werkt het?
- TST: wat doet de Taalunie?
- TST-Centrale en IPR
- TST en dataleveranciers

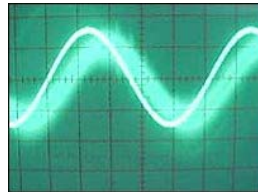
## TST: WAT IS HET?

taal- en spraaktechnologie (TST) maakt het mogelijk in natuurlijke taal te communiceren in een digitale omgeving

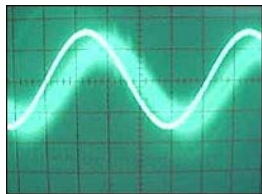
→ taal- en spraaktechnologie maakt communicatie mogelijk in de informatiemaatschappij

herkennen en reproduceren van gesproken taal

- spraaksynthese: tekst naar spraak



- spraakherkenning: spraak naar tekst



# INTELLIGENTE OMGEVING



<http://www.youtube.com/watch?v=z4WA-n5lqS8>

- schrijfhulp: spelling, grammatica, stijl, afbreken, thesaurus

- samenvatten

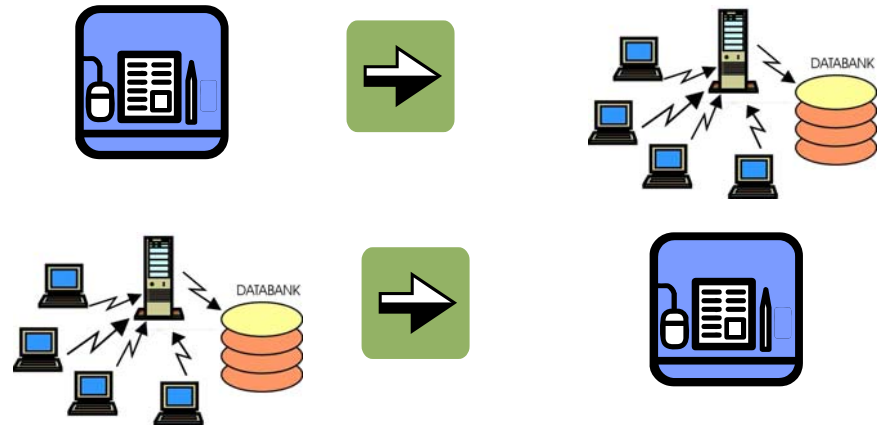


- vertalen

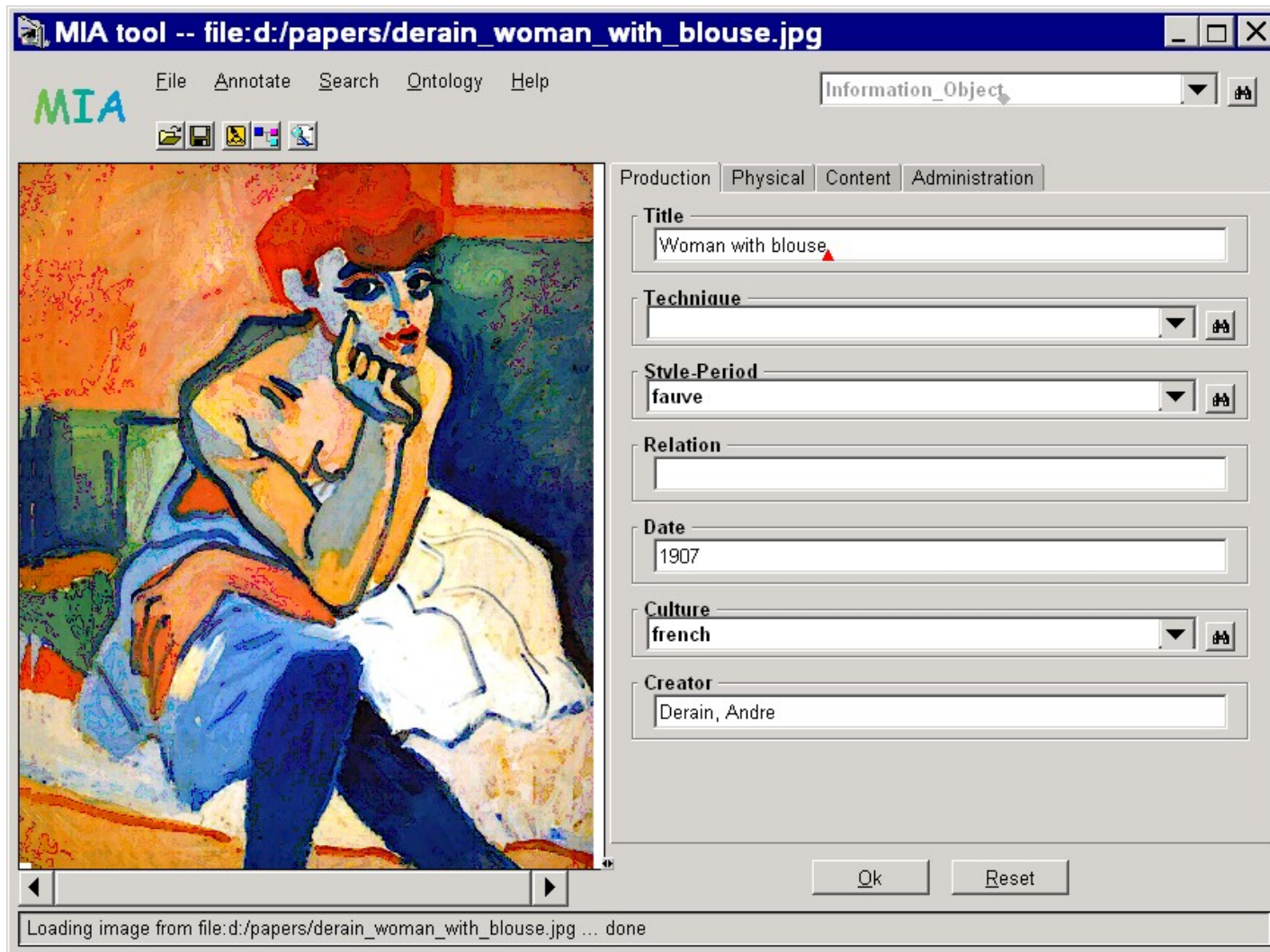


## geavanceerde systemen voor informatiebeheer

- categoriseren
- indexeren
- opzoeken
- extractie
- combineren

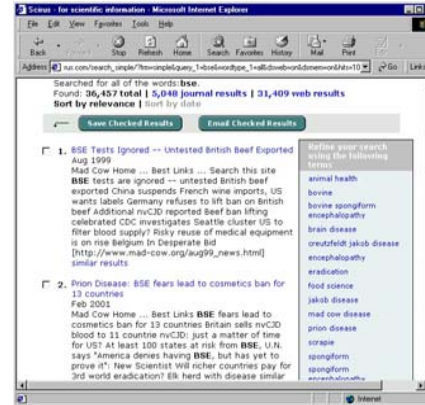


# TST: WAT KAN JE ERMEE? TAAALTECHNOLOGIE

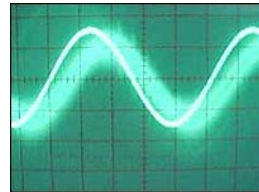
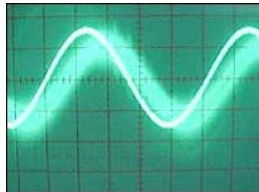


# TST: WAT KAN JE ERMEE? DIALOGEN

- vraag en antwoord



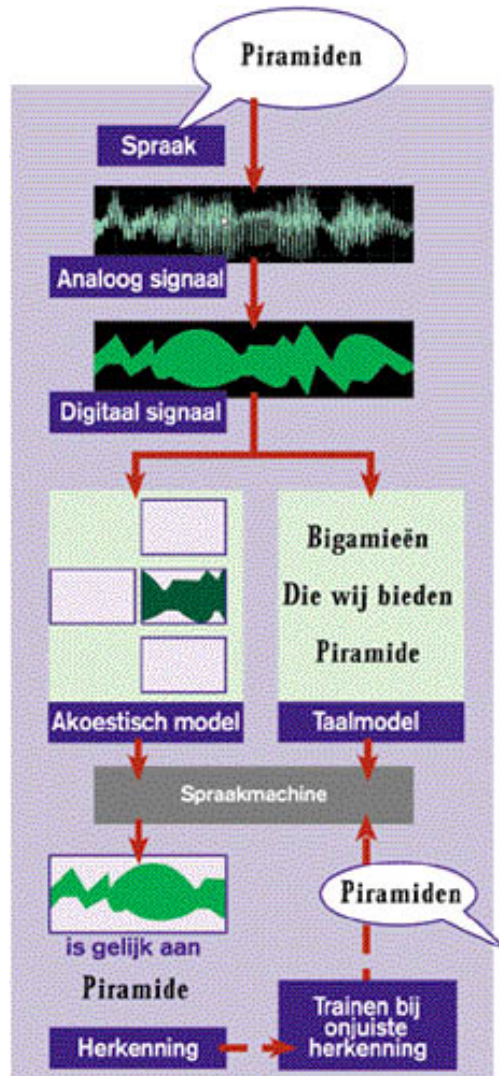
- gesproken dialogen



# TST: HOE WERKT HET? (1)

- wat heb je nodig?
  - data, heel veel data (corpora geschreven en gesproken taal, digitale woordenboeken enz...)
    - bv. spraak: hoe gevarieerder opgenomen gesproken materiaal, hoe sneller stem getraind kan worden.
    - evaluatiemateriaal
  - software (spraakherkenners, zinsontleders, grammatica's enz...)
- methode
  - Regels, taalmodel
  - statistiek (veel data nodig: bijv. frequentie woord, context woorden) voorspellingen maken.

# TST: HOE WERKT HET? (2)



Doc 1

Now is the time  
for all good men  
to come to the aid  
of their country

Doc 2

It was a dark and  
stormy night in  
the country  
manor. The time  
was past midnight

Term	Doc #
now	1
is	1
the	1
time	1
for	1
all	1
good	1
men	1
to	1
come	1
to	1
the	1
aid	1
of	1
their	1
country	1
it	2
was	2
a	2
dark	2
and	2
stormy	2
night	2
in	2
the	2
country	2
manor	2
the	2
time	2
was	2
past	2
midnight	2

motto: 'de gebruiker centraal'

- alle gebruikers van het Nederlands moet in alle communicatieve situaties de eigen taal kunnen gebruiken - dus ook in communicatie binnen de informatiemaatschappij.



## TAALUNIE EN TST (2)

- juiste voorwaarden creëren om de nodige toepassingen te ontwikkelen;
- basismaterialen (taalsoftware, corpora enz.) voor het Nederlands nodig;
- grote investeringen: in een relatief klein taalgebied is tussenkomst van de overheid nodig;
- initiatief voor STEVIN: groot stimuleringsprogramma voor het Nederlands in taal- en spraaktechnologie.

## STEVIN: financiers, middelen, structuur

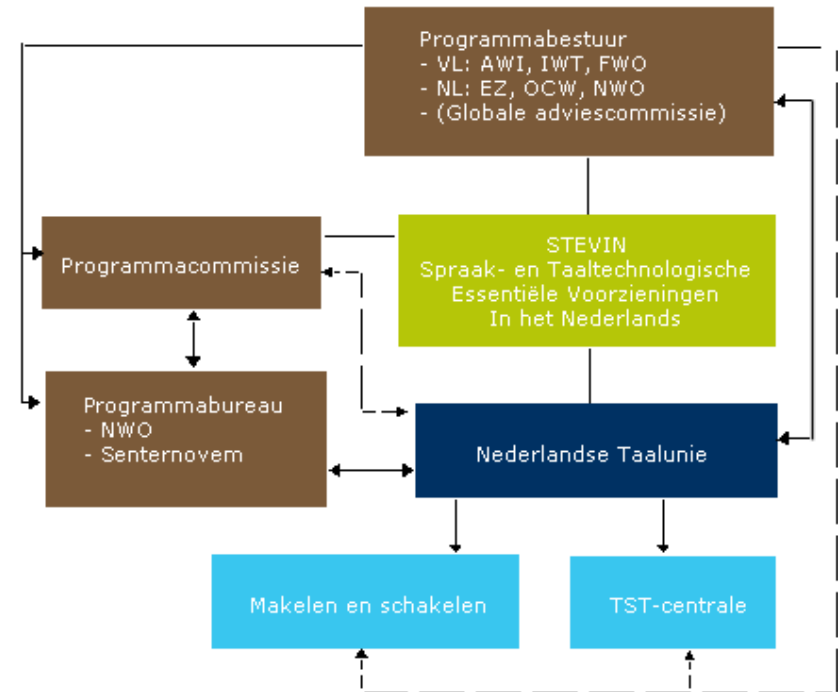
### VL-NL samenwerking

- NL: OCW, EZ, NWO
- VL: EWI, IWT, FWO
- NTU = coördinator

budget: €11,4 miljoen  
voor 5 jaar

gezamenlijke pot

looptijd: 2004-2009



# STEVIN: stand van zaken

- verschillende oproepen gelanceerd
- groot deel van veld betrokken:  
kennisinstellingen en bedrijven
- zie DIXIT

# TST-CENTRALE (1)

duurzaamheid digitale taalmaterialen: oprichting TST-centrale (centrale plaats voor onderhoud, beheer en distributie)

➤ Instituut voor Nederlandse Lexicologie (Leiden en Antwerpen)

- taken:
- archiveren
  - onderhouden en beheren
  - toegankelijk maken en houden
  - contracten afsluiten
  - gebruikers ondersteunen
  - behoeften signaleren

# Dank voor uw aandacht!

aanvullende informatie:

STEVIN: <http://taalunieversum.org/stevin>

TST-Centrale: <http://www.tst.inl.nl/>

# vragen?

