

Detecting And Exploiting Semantic Overlap



Site visit, Tilburg, November 8, 2007

Agenda

1. Inleiding
2. Stand van zaken: corpus data, tools en annotatie
3. Plannen voor komende periode van 6 maanden
4. Recente resultaten zinsfusie
5. Discussie

What is semantic overlap?

- Zangeres Christina Aguilera heeft eindelijk verteld waarom haar buik zo dik is. [NOS]
- Christina Aguilera heeft in het Amerikaanse tijdschrift Glamour bevestigd dat zij zwanger is. [AD]
- Christina Aguilera heeft eindelijk bevestigd wat de hele wereld al wist: ze is zwanger. [NOVUM]
- Iedereen wist het al, maar nu zou Christina Aguilera het zelf voor het eerst hebben toegegeven: ze is zwanger. [The Agenda]



Research Questions

- **Detecting semantic overlap**
How to determine to what extent different sentences express the same information?
- **Using semantic overlap**
How to use knowledge about semantic overlap for different NLP / IR applications?

About the Daeso Stevin project

- **Period:** October 2006 - October 2009

- **Partners:**
 - Tilburg University [2006 - 2009]: Erwin Marsi, Emiel Kraemer
 - Antwerp University [2007 - 2009]: Iris Hendricks, Walter Daelemans
 - University of Amsterdam [2009]: Maarten de Rijke
 - TextKernel [2009]: Jakub Zavrel

- **Rough planning:**
 - Year 1: Corpus collection and annotation
 - Year 2: Semantic overlap detection / Summarization application
 - Year 3: Semantic overlap detection / Summarization, IE, QA application