

## Status for arbejdet med danske sprogteknologitermer

Hanne Fersøe  
Københavns Universitet  
Center for Sprogteknologi  
hannef@hum.ku.dk



## Nordisk Sprogteknologisk Forskningsprogram 2000 - 2004

- Et dokumentationscenter for sprogteknologi i hvert nordisk land
  - Danmark – <http://www.cst.dk/dandokcenter/index.html> eller i dag <http://cst.ku.dk/videncenter/>
  - Finland – <https://kitwiki.csc.fi/twiki/bin/view/FiLT/FiLTWikiEn>
  - Island – <http://www.tungutaekni.is/english.html>
  - Norge – <http://www.norskdok.uib.no/>
  - Sverige <http://sprakteknologi.se/>
- Et netværk til fremme af samarbejdet mellem de nordiske dokumentationscentre
  - NorDokNet <http://cst.dk/nordoknet/>



## Arbejdet i NorDokNet

- Aktiviteter omkring dokumentationscentrene
  - Fælles informationstyper på hjemmesiderne
  - Fælles strukturering af informationstyper
  - Fælles nyheder
  - Fælles ontologi for sprogteknologi baseret på engelsk som pivot-sprog
  
- Andre aktiviteter
  - Samarbejde med DFKI, LT-World <http://www.lt-world.org/>
  - Samarbejde om EU-ansøgning
  - Formidling i de baltiske lande m.h.p. fremtidigt samarbejde



## Sprogteknologiske termer i Danmark

Fra EUROMAP/HOPE-projekterne fandtes et HLT Glossary:

- 90 enkelt- og flerordstermer fra emneområdet HLT (Human Language Technology)
- Sprog: engelsk, fransk, tysk, spansk, italiensk, dansk, finsk
- Alle termer har definitioner på alle sprogene. Definitionerne er de samme.

I NorDokNet blev der igangsat et eksperiment med automatisk uddragning af flersproglig terminologi på basis af tekst på flersproglige hjemmesider.

I Danmark arbejdede Jürgen Wedekind videre med resultaterne heraf i projektet Nordisk Netordbog:  
[http://cst.ku.dk/projekter1/projekter\\_slut\\_i\\_2007/netordbog/](http://cst.ku.dk/projekter1/projekter_slut_i_2007/netordbog/)



## Konklusion

Ingen forskere på CST arbejder generelt med at udvikle terminologi, men i specifikke finansierede projekter, som fx Nordisk Netordbog, arbejder vi med problemstillinger inden for terminologi.

I IATE-projektet arbejdede Lina Henriksen og Claus Povlsen med disse ting:

- Specifikation af funktionalitet i IATE termdatabasen og af rutiner til import af de eksisterende termbaser til IATE
- Analyse og specifikation af automatisk identifikation og fletning af dubletter i IATE-basen <http://iate.europa.eu>
- Udvikling af metoder og standarder i EuroTermBank-projektet <http://www.eurotermbank.com/>

Vi deltager gerne i drøftelser af hvordan en dansk terminologi for sprogteknologi kan udvikles eller videreudvikles i et nordisk regi.

