



Det Humanistiske Fakultet



Center for Sprogteknologi 25 år 1991-2016

Bente Maegaard
Center for Sprogteknologi, NFI

bmaegaard@hum.ku.dk



Forhistorien op til skabelsen af CST i 1991

EU (dengang EF) igangsatte et maskinoversættelsesprojekt EUROTRA, første møder 1978.

- For alle EU-sprog (7 da forslaget fremsattes)
- Forberedende fase 1979-1984
- Udførende fase 1984-1990

1989: Vi havde opbygget en forskningsgruppe i Danmark, lokaler på KU, gode leverancer til EU, omtale i pressen, osv.

Hvad skulle der ske med den opbyggede gruppe og dens ekspertise?

Styregruppen for EUROTRA-DK: denne ekspertise skulle fastholdes.

Forslag om et center der kunne videreføre arbejdet

Møder med ministeriet osv. Støtte fra erhvervslivet.

Høring af institutioner og råd.

Mange var positive, men nogle var negative.

Og: Del af universitetet eller ej? Kun maskinoversættelse eller bredere?

Bølgerne gik højt!

Artikel i Information, samt Sprog og Erhverv,
juni 1990

Skrevet af medarbejdere ved søsterinstitut
på KU

Fortielser om mislykket EF-forskningsprojekt

Ministeriet besluttede sig alligevel for at give os en chance

Oprettede CST for 5 år, 1991-1995

- Selvstændigt center under ministeriet
- Breder end bare maskinoversættelse
- Skrappe krav om ekstern indtjening

Bestyrelse, formænd

- Niels Ole Bernsen, RUC, 1991-1995
- Erik Kofod, Dansk Management Forum, 1996-2000
- Jens Langeland-Knudsen, Carpe Diem Investment 2000-2003

Medlemmer fra universiteter, forskningsråd, organisationer, erhvervsliv, mv.



Fra starten i 1991. For Lingtech: PaTrans – maskinoversættelse af patenter - byggende på EUROTRA.

Lingtech var glade: de tjente penge

CST var glade: vi opfyldte de store krav om ekstern indtjening.

EU var glade: EUROTRA virkede

Foto fra årsberetning 1993



rektor Viggo Hansen: sprog er det mest kulturbærende element overhovedet ...

CST 1993



En del af CST's stab: (bagest fra venstre) Bo Simonsen, Henrik Kjærsgaard, Connie Hughson, Claus Povlsen, Janne Madsen, Uffe Sonne Svenden, (midterst) Bente Mægaard, Annelise Bech, Anne Neville, Anne-Marie Eriksen, Costanza Navarretta, Marianne Dall, Christina Mazzi, (forrest) Lene Offersgaard, Lina Henriksen, Helene Buch Brøndel, Bolette Sandford Pedersen og Lise Damsgaard Jørgensen.

Resten af den organisatoriske historie

International evaluering 1994

Tidsbegrænsningen ophævedes 1994

Sektorforskningsinstitution 1996

Fusioneret med KU-HUM 2004, som institut

CST og NFI fusionerede 2015

Og nu fusionerer vi snart med Institut for Nordiske Sprog og Sprogvidenskab

Hele perioden har vi været afhængig af eksterne projekter.

Ca. 70 projekter støttet af forskningsråd, EU mv.

De fleste projekter er samarbejdsprojekter.

Opbygget og vedligeholdt netværk: danske, nordiske, europæiske, internationale

CST's opgaver

- Nationalt center for Sprogteknologi
- Forskning og udvikling
- Myndighedsbetjening af ministeriet – fremhævet i 2004 ved fusionen
- God sprogteknologi til danske brugere og andre brugere af dansk
- Undervisning, fra fusionen med KU i 2004
- Undervise i IT & Cognition fra 2007



Projekter gennem årene (EU, fonde mv.)

| Forskningsinfrastruktur | 17 |
|---|----|
| Sprogteknologiske anvendelser (MT, IR, mv.) | 19 |
| Sprogprocessering og sprogressourcer | 22 |
| Multimodal kommunikation | 4 |
| Formidling og dokumentation | 7 |

Se projekterne på vores hjemmeside cst.ku.dk

CST har været koordinator for 12 samarbejdsprojekter

Eksempler på hvordan vi har bidraget til god sprogteknologi

Lingtech - PaTrans, patentoversættelse

DR – hjulpet med valg af ordforråd til deres talegenkendelse til tekstning

Schultz Information – hjælp til anonymisering af domme mv.

Sensus/Tegnsprogsordbogen- Lingvistiske specifikationer for en online tegnsprogsordbog

Statens Information - Undersøgelser af automatisk kategorisering

Nordisk Ministerråd – automatisk fremfindning af relevante tekster på de nordiske sprog ud fra et søgeord på et af sprogene

Max Manus - Anonymisering af data til træning af talegenkendelsessystemer

Folketinget - Undersøgelse og anbefalinger ift. talegenkendelsessystemer, samt - Anbefalinger og strukturering af skrivevejledninger

Det hollandske forskningsråd – Validering af sprogressourcer (ikke dansk)



Klip fra 2007: Bladet Humanist

side 2 nr. 8

humanist



Ordleger og lektor i sprogteknologi, forfatter leder af Center for Sprogteknologi, Bente Maegaard.

Vi vil gerne brande Det Humanistiske Fakultet

Fakultetet skal være på omdrejningshøjde med den teknologiske udvikling: fx skal der forskes i at gøre computere meget bedre til at håndtere sprog, når borgere søger på nettet. Sådan lyder en af Bente Maegaards visioner. Hun er centerleder på Center for Sprogteknologi – et af de nyeste centre, der er kommet under KU's vinger

Det vil vi stadig gerne!

