



Det Humanistiske Fakultet



MUDS 12

9. okt. 2008

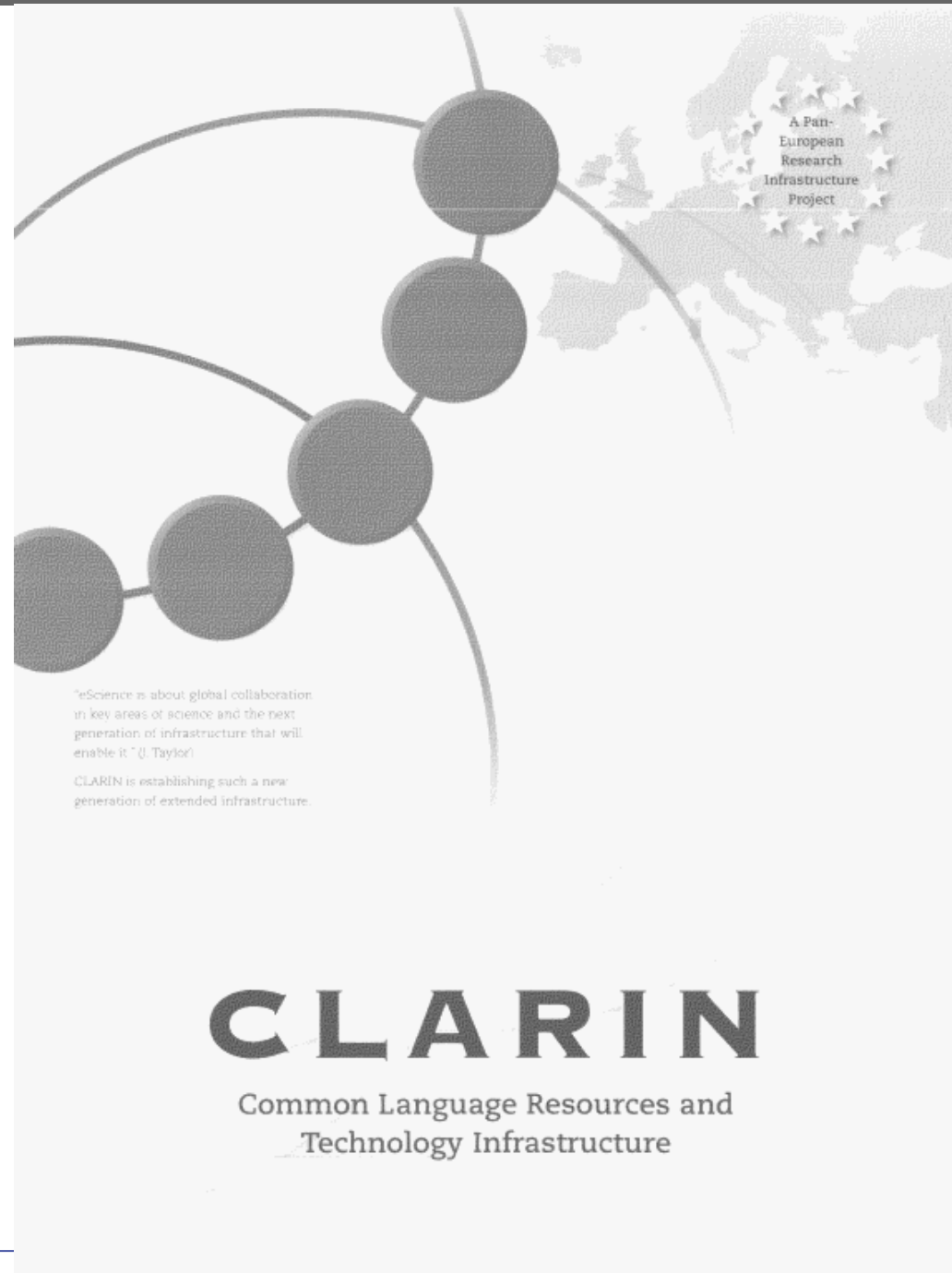
Hanne Ruus
Institut for Nordiske Studier
og Sprogvidenskab

Københavns Universitet

Fri og bunden forskning om CLARIN-DK
WP 2.3 Knowledge for everyman

MUDS 12 - 9. okt. 2008
Dias 1



The background of the slide features a stylized network of dark grey circles connected by thin lines, set against a light grey map of Europe. In the upper right corner, there is a circular emblem with twelve stars, containing the text "A Pan-European Research Infrastructure Project".

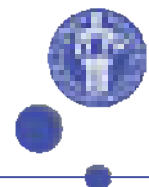
A Pan-European Research Infrastructure Project

"eScience is about global collaboration in key areas of science and the next generation of infrastructure that will enable it" (J. Taylor)

CLARIN is establishing such a new generation of extended infrastructure.

CLARIN

Common Language Resources and
Technology Infrastructure



CLARIN-DK

The Danish CLARIN project

The University of Copenhagen, heading the Danish CLARIN consortium, has been given a three year grant of 15 million DKK (appr. 2 million €) to construct a Danish research infrastructure for the humanities integrating written, spoken, and visual records into a coherent and systematic digital repository. The project runs from January 2008 until the end of 2010.

The infrastructure is named ***Centre for Danish Language resources and Technology Infrastructure for the Humanities***. The partners include eight leading Danish humanities institutions: four universities, a university library, a museum, and two government research institutions.



Parterne i CLARIN-DK

University of Copenhagen – KU - with three departments from the Faculty of Humanities:

- 1) Centre for Language Technology – CST – co-ordinator of the project**
- 2) Danish National Research Foundation Centre for Language Change in Real Time – LANCHART**
- 3) Department of Scandinavian Studies and Linguistics – INSS**
- 4) University of Southern Denmark – SDU**
- 5) University of Aarhus – AU**
- 6) Copenhagen Business School – CBS**
- 7) The Royal Library–KB**
- 8) The National Museum of Denmark– NatMus**
- 9) Society for Danish Language and Literature – DSL**
- 10) Danish Language Council – DSN**



CLARIN-DK – mål og vision

- Målet er at opbygge en forskningsinfrastruktur for humaniora i form af et digitalt bibliotek med tekster, talesprog , videooptagelser og billeder.
- Biblioteket tilbyder en række it-værktøjer til søgning i data, konvertering, lagring mm.
- Biblioteket henvender sig til humanioraforskere i alle fag.
- Bibliotek med værktøjer udvikles 2008-2010
- Biblioteket videreføres og udvides efter 2010



CLARIN-DK - organisation

5 arbejdspakker:

**WP1 Coordination and Technical management, incl.
copyright and privacy issues**

WP2 Basic written language resources

WP3 Spoken language resources and tools

WP4 Technological resources

WP5 Technical Infrastructure

4 perioder:

jan. 2008 til okt. 2008

okt. 2008 til juli 2009

juli 2009 til maj 2010

maj 2010 til jan. 2011

For hver periode:

specificeret arbejdsplan og delresultater



CLARIN-DK arbejdsapakke 2

- 2.1 Almensprogligt tekstkorpus på 45 mil. Ord
Reference corpus of general language
- 2.2 Fagsprogskorpus fra nutiden på 11. mil. ord
Corpus of sublanguage texts
- 2.3 Faglig formidling i historisk perspektiv
Knowledge for everyman 1500-1750
- 2.4 Digitalisering af Johs. V. Jensen med mærker og søgefaciliteter
Enhanced annotation and improved search possibilities to old literary texts
- 2.5 Nationalmuseets billeddatabase med forklaringer inkl. Kunstammerdata
Images, artefacts and texts from the National Museum of Denmark
- 2.6 Parallelttekster mellem dansk, engelsk mm. 20. mil. ord
Parallel multi-lingual text resources, with alignment



WP 2.3 Knowledge for everyman

Arbejde i første periode:

Forberedelse af lokal infrastruktur:

konto til bevilling

ansættelse af personale

Oplæring af medarbejdere i tekstmærkning

pilotarbejde med tekster fra DSST

www.duds.nordisk.ku.dk/tekstresurser

Oversigter over mulige tekster i perioden 1500-1750

valg af fagområder

Valg af xml-editor



WP 2.3 Knowledge for everyman

Delresultater for første periode

AC-fuldmægtig og to studentermedarbejdere oplært i tekstmærkning

Præliminær oversigt over mulige tekster 1500-1750

Fastlagt tekstskanningsprocedurer
(billeder og OCR med præliminær mærkning)

Personaleresurser for anden periode:

4/5 AC-fuldmægtig

1/2 forskningslektor

2 forskningslektormåneder

2 professormåneder

eksterne eksperter til faglig rådgivning:

boghistoriker, filologi- og literacy-ekspert, materialfilolog

påtænkt ansættelse af flere studentermedarbejdere



WP 2.3 Knowledge for everyman

Fagområder 1500-1600

Almanak

Bønnebog

Nyhedsformidling

Husholdning

Katekismus

Naturvidenskab

Pædagogik

Metafysik



Forside på et månedsblad 1722



Extraordinaire

Relationer

For
September Maaned 1722
Danmark

Den 29 udi afvigte Maaned ere Laugs=Artiklerne bestaaende af 19 Puncter for Steenfører Lauget i Kiøbenhavn aller=Naadigst givne, hvorefter det 1. er Steenførerne tillige med Pram=Mændene tilladt til Vands med egne Fartøje Grund og Broesteene sampt Muur og Broesand at indføre, og paa sine dertil anord= nede

