

Data Archiving and Networked Services

DANS

***Permanent access to research data from
the humanities and social sciences – from
the Dutch experience to a European
research infrastructure***

Peter Doorn, Director DANS

BOREAS Workshop on Data Management (“Cyber-
research Infrastructures”) in Humanities and Social
Science-based research, Paris, 19-20 February 2007



Overview

1. Data archiving in the social sciences and humanities
2. Research Infrastructures: in the natural sciences, social sciences, arts and humanities
3. The infrastructures for the social sciences and humanities on the ESFRI Roadmap
4. The DARIAH proposal: Digital Research Infrastructure for the Arts and Humanities



1. Data archiving in the social sciences and humanities



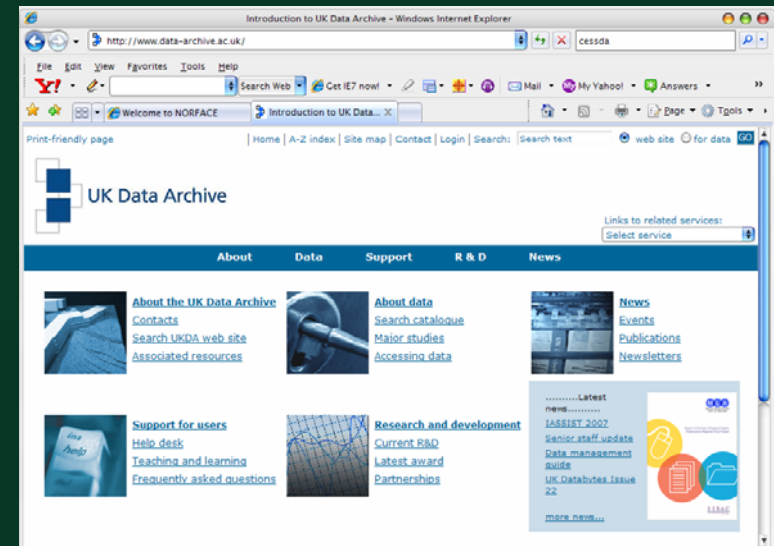
Long tradition of data archiving and networking

- Data archives in the social sciences since 1960s and 1970s
- Text archives since late 1970s
- Historical data archives since late 1980s
- International collaboration:
 - Council of European Social Science Data Archives (CESSDA)
 - International Federation of Data Organisations (IFDO)
 - International Association for Social Science Information Service and Technology (IASSIST)
 - Association for Computing in the Humanities (ACH)
 - Association for Literary and Linguistic Computing (ALLC)
 - Association for History and Computing (AHC)

Examples of Social Science Data Archives

Zentralarchiv für Empirische Sozialforschung (1960):

- archives primary material (data, questionnaires, code plans) and results of empirical studies in order to prepare them for secondary analyses and to make them available to the public



UK Data Archive (1967):

- centre of expertise in data acquisition, preservation, dissemination and promotion;
- curator of the largest collection of digital data in the social sciences and humanities in the UK
- provides resource discovery and support for secondary use of quantitative and qualitative data in research, teaching and learning



Example of Humanities Data Archives

The Arts and Humanities Data Service | tel: 020 7848 1988 | email: info@ahds.ac.uk - Enabling D - Windows Internet Expl...

http://ahds.ac.uk/

ahds arts and humanities data service

Select a subject area: Go

Search Collections
Creating Resources
AHRC Applicants
Depositing Resources
News and Events
About the AHDS
Projects
Search Site
Vacancies

[Printer friendly version](#)
[Text version](#)

AHDS Executive Address:
26 - 29 Drury Lane
3rd Floor
King's College London
LONDON, WC2B 5RL

Enabling Digital Resources for the Arts and Humanities

The Arts and Humanities Data Service (AHDS) is a UK national service aiding the discovery, creation and preservation of digital resources in and for research, teaching and learning in the arts and humanities. Currently, we cover five subject areas:

Archaeology History Visual Arts

Literature, Languages & Linguistics Performing Arts

Latest News

- [Scoping study](#) - on creating and using digital resources in performing arts
- [Saving history records](#) - migrating punch card records to digital resources
- [Archaeology Image Bank](#) - for finding and sharing archaeological images

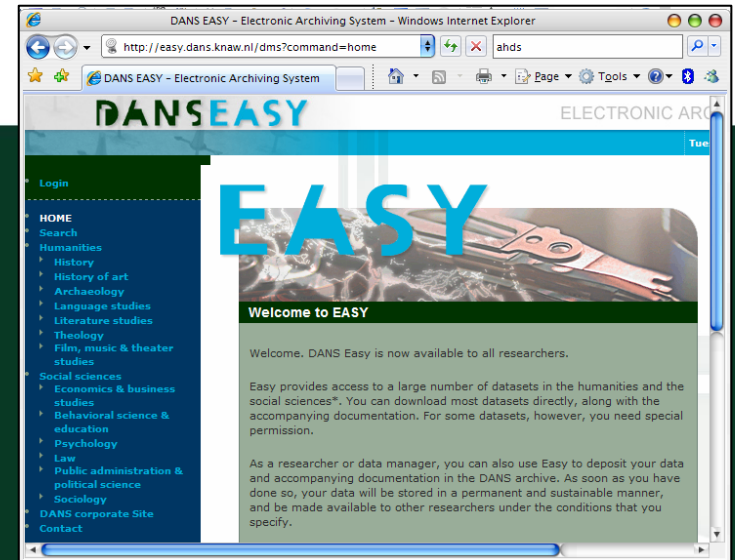
From the AHDS Blog

- [Software for doing fieldwork](#)
- [Users do not want what you expect](#)
- [Call for Submissions: Sharing Learning Resources Online](#)

DANS

What is DANS?

- DANS:
 - Dutch national organisation for access to research data
 - Guarantee long-term curation to such data
 - Works for humanities and social sciences
- Joint initiative:
 - Royal Netherlands Academy of Arts and Sciences (KNAW)
 - Netherlands Organisation for Scientific Research (NWO)
- Aim:
 - Boost collaboration with researchers and between researchers, both within and across disciplines
 - Open Access is the guiding principle
- DANS stimulates permanent access to research data at the European level





2. Research Infrastructures: in the natural sciences, social sciences, arts and humanities

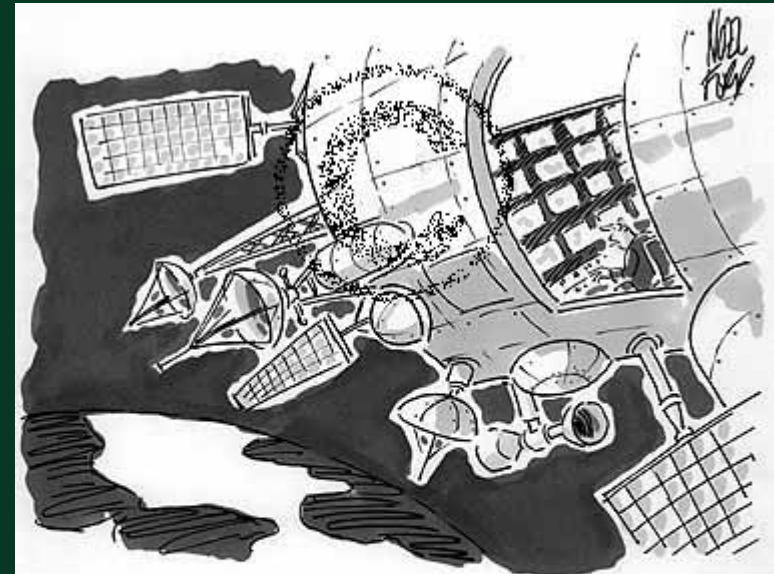


Research Infrastructures (R.I.)

- R.I. in general: **permanent** and **physical**
- R.I. for the natural sciences: ice breakers for polar research, satellites, telescopes, particle accelerators, laboratories
- R.I. for the social sciences and humanities?
 - Interview (survey) and registration data for the social sciences
 - Cultural heritage information for the humanities
 - Libraries and archives are the traditional “laboratories” for the humanities
- In the digital age, essential for innovative social science and humanities research is:
 - Access to social science data and digitised heritage data (data bases, text corpora, speech, image collections, g.i.s., etc.)
 - Tools to process this information

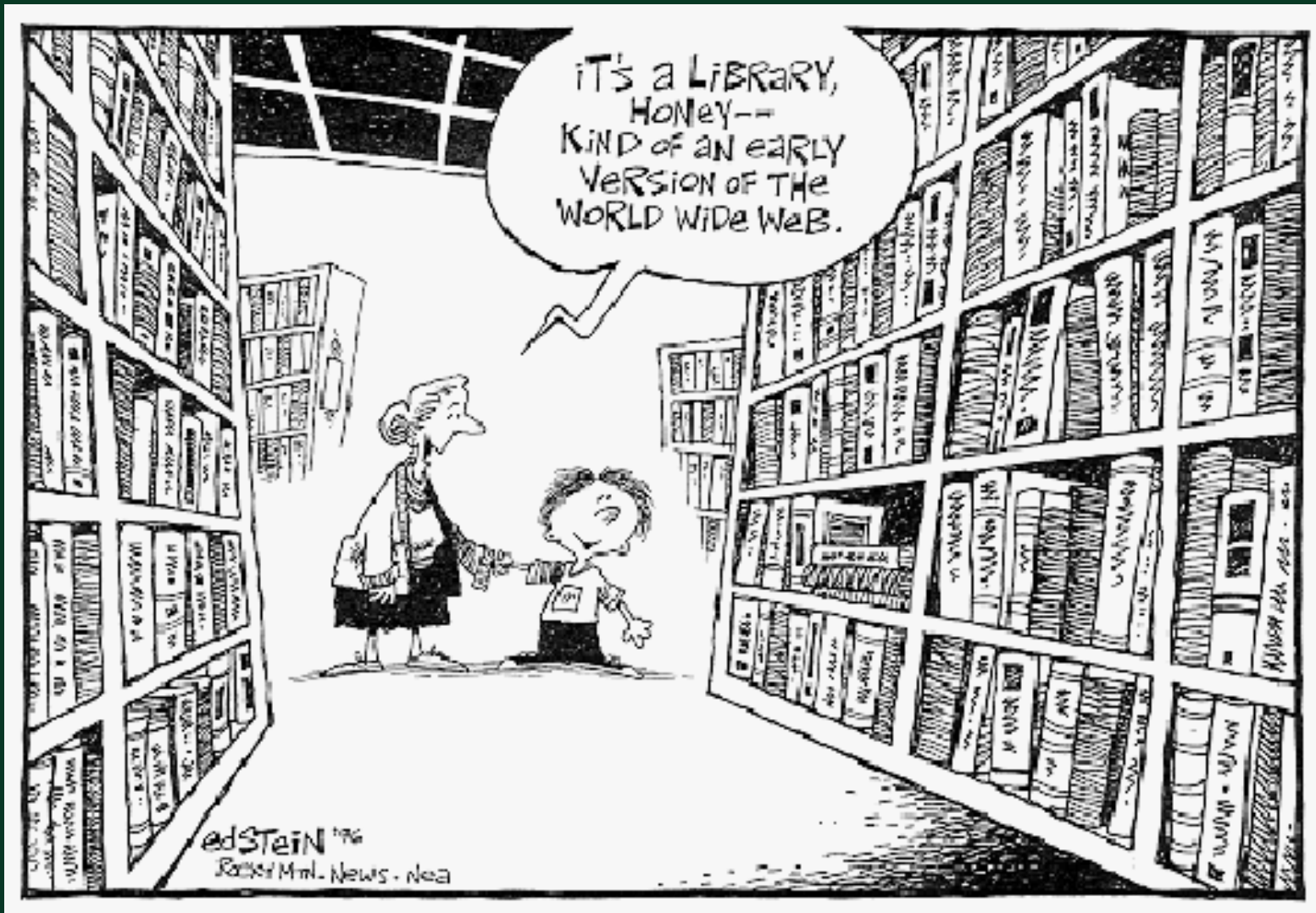
DANS

A digital research infrastructure for the social sciences and humanities is comparable to a virtual astrophysical observatory



"Jones - there's a new market emerging in Tunisia... Williams - the Balkan situation is improving... Simmons - why are you still propping up that bar in Shanghai..."

DANS





European infrastructure challenges

- In spite of some achievements, existing infrastructures are primarily national
- European activities are until now funded on a project basis and carried out as voluntary activities by national centres
- Stable, pan-European data infrastructures for the humanities and social sciences hardly exist
- Increasing internationalisation of humanities and social science research puts new requirements for such infrastructures



3. Infrastructures for the social sciences and humanities on the ESFRI Roadmap



ESFRI Roadmap

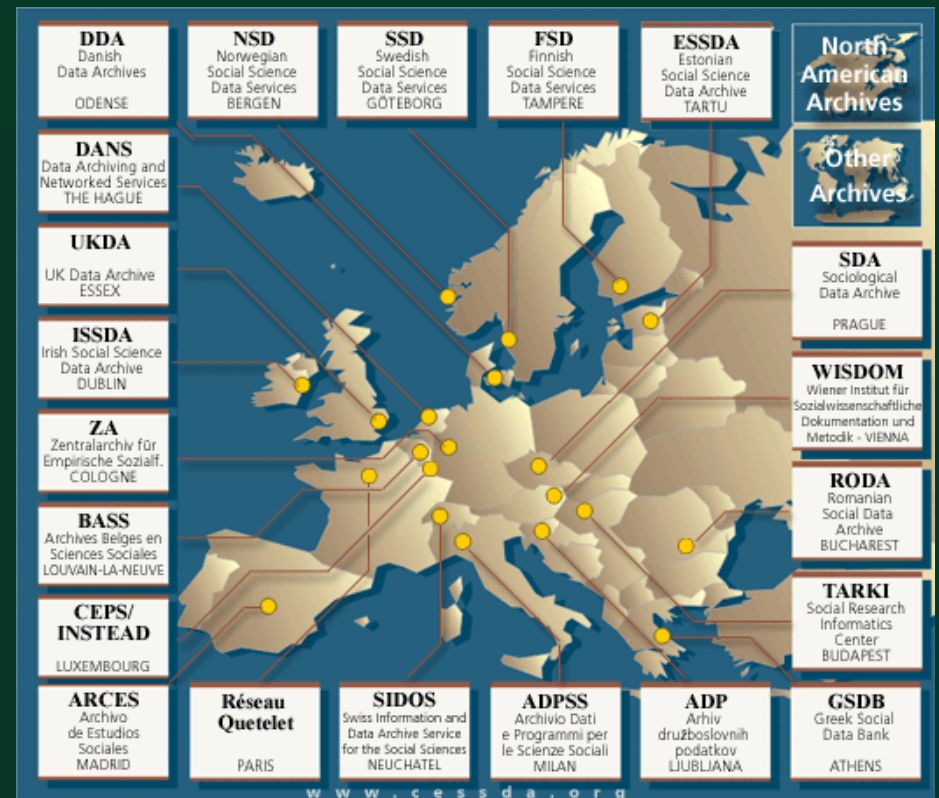
- Within ESFRI working group for Social Sciences and Humanities two 'expert groups':
 - EROHS: concentrating on social sciences
 - Cultural Heritage: concentrating on humanities
- About 90 proposals for SSH submitted
- Six proposals selected for 2006 ESFRI Roadmap

Projects (in alphabetical order per discipline)	Estimated Construction Cost (M€) *	First possible operations for users	Indicative Operational/ Deployment Cost (M€/year)
CESSDA	30	2008	6
CLARIN	108	2008	10
DARIAH	10	2008	4
EROHS	43	2008	12
ESS: European Social Survey	9	2007	9
SHARE	50	2007	<1



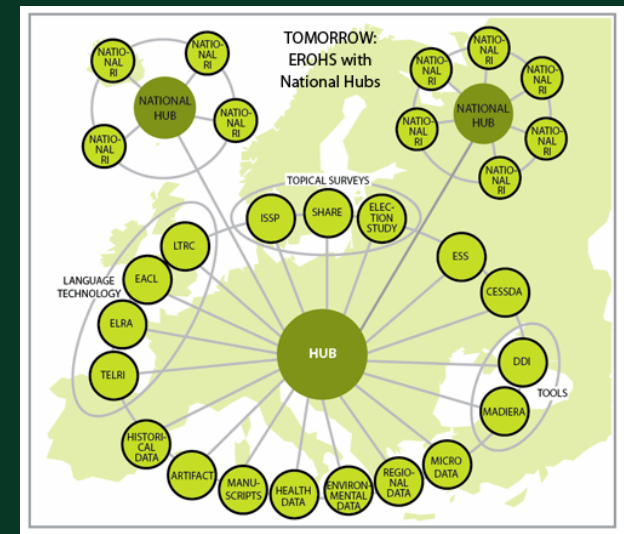
CESSDA: Council of Social Science Data Archives

- Now: Network of social science data archives across 21 countries in Europe, 15,000 data collections, serving over 20,000 researchers
- Ambitions :
 - major upgrade:
 - strengthening
 - widening
 - technological renewal
 - capacity building
 - in several respects comparable to DARIAH



EROHS (European Resource Observatory for the Humanities and Social Sciences)

- Will operate both as a central and distributed facility with a strong physical hub working in close conjunction with a number of spokes across Europe
- Coordinated yet decentralised network, aiming at organising the communication, coordination, documentation and sharing of information
- Oriented to social sciences and humanities



Adapted from the EROHS Roadmap proposal



ESS: European Social Survey

- Set up in 2001 to monitor long term changes in social values throughout Europe, now covering 27 countries
- Upgrade: stable funding for data collection over the next 10 years, improve survey methodology

SHARE (Survey of Health and Retirement in Europe)

- Data infrastructure for fact-based economic and social science analyses of the on-going changes in Europe due to population ageing.
- Upgrade: From the original 8-country survey to all Member States of the EU; stable funding for data collection and dissemination



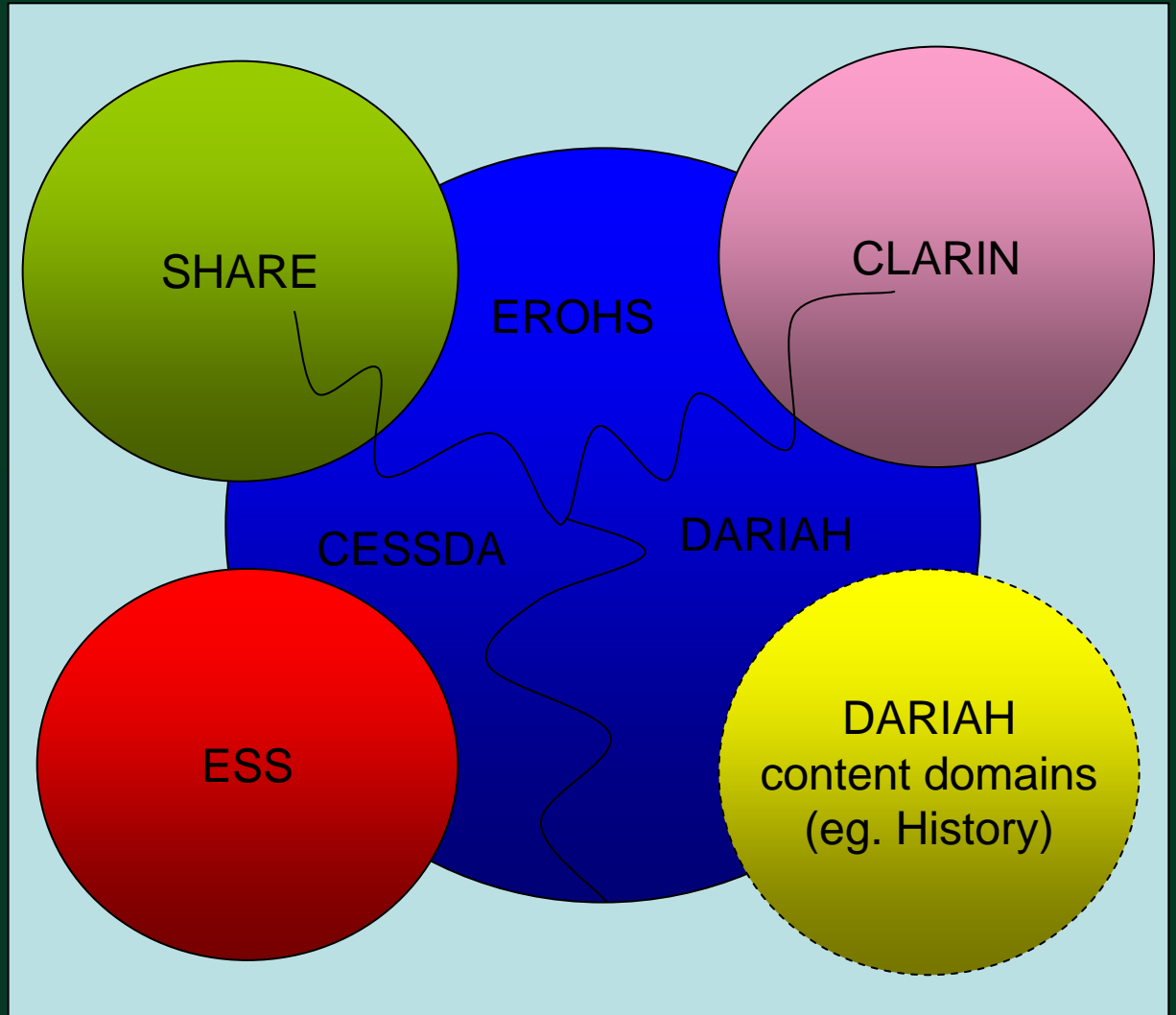
CLARIN: Common Language Resources and Technology Infrastructure

- Make language resources and technology available and useful to scholars of all disciplines, in particular the humanities and social sciences.





EROHS, CESSDA and DARIAH are to co-ordinate (or integrate) their functions very well and to supply networked services (data archiving, dissemination) to the content-oriented proposals (SHARE, ESS, CLARIN, etc.)



Interdependence of SSH Roadmap proposals

4. The DARIAH proposal: Digital Research Infrastructure for the Arts and Humanities

www.dariah.eu



Digital Research Infrastructure for the Arts and Humanities

The DARIAH idea

The Grand Vision:

- Provide access to European humanities and cultural heritage information across time
- A Research Infrastructure that can Coordinate, Catalyse, Enhance, Support

Digital research infrastructure for the humanities:

- Provide permanent access to data collected/digitised in European projects: providing continuity for discontinuous activities



Digital Research Infrastructure for the Arts and Humanities

- Support research networks in the humanities and social sciences

Outline of tasks of DARIAH

Digitise – Curate – Preserve

- Standards development and promotion
- Preservation and digitisation services
- R&D, technology platforms, tools development
- Legal services and advice

Discover – Access – Deliver

- Authentication and authorisation,
- Harvesting, aggregating, hosting
- User-friendly discovery and delivery

Connect – Collaborate – Use

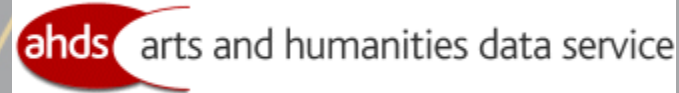
Digital Research Infrastructure for the Arts and Humanities

- Supporting communities of practice



Pan-European partnership

Four initiators:



- France, Germany, Netherlands, UK



Additional interest to participate from:

- Croatia, Cyprus, Denmark, Greece, Ireland, Romania, Slovenia, ...



Digital Research Infrastructure for the Arts and Humanities

Science Case

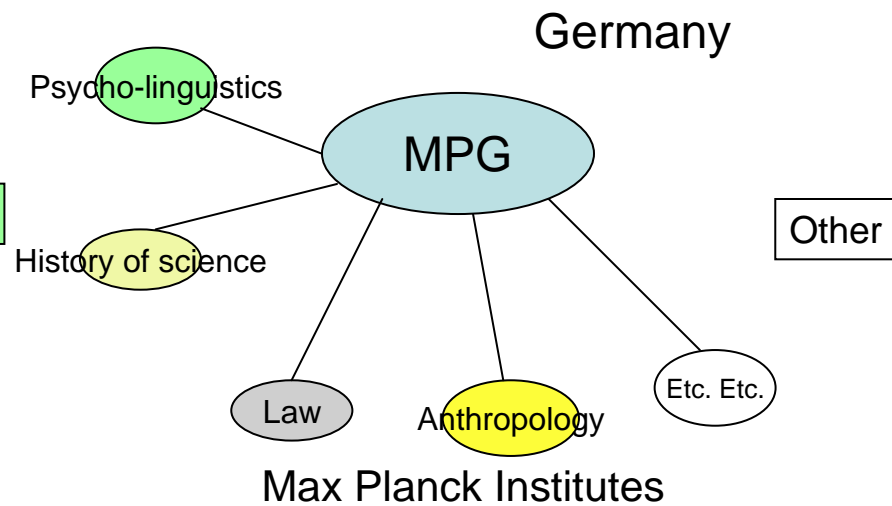
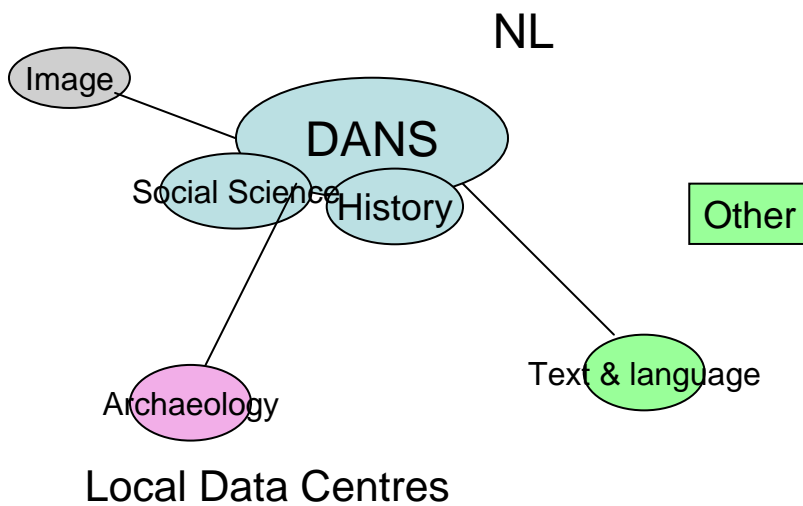
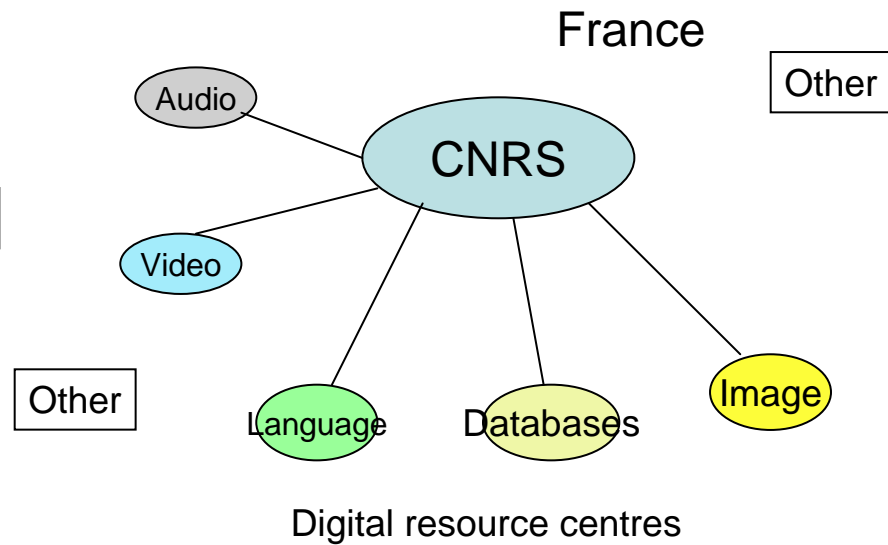
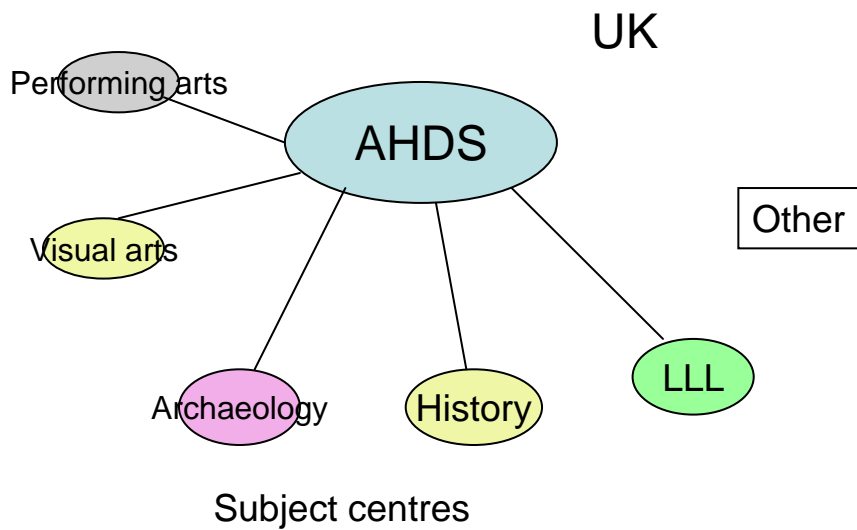
Changing research practice in a networked environment:

- Data (including text, images, and other media) is the laboratory of the scholar in the humanities
- Resources on the web are distributed (data grid)
- The scale of research goes up: networked projects
- New technologies and methods of analysis

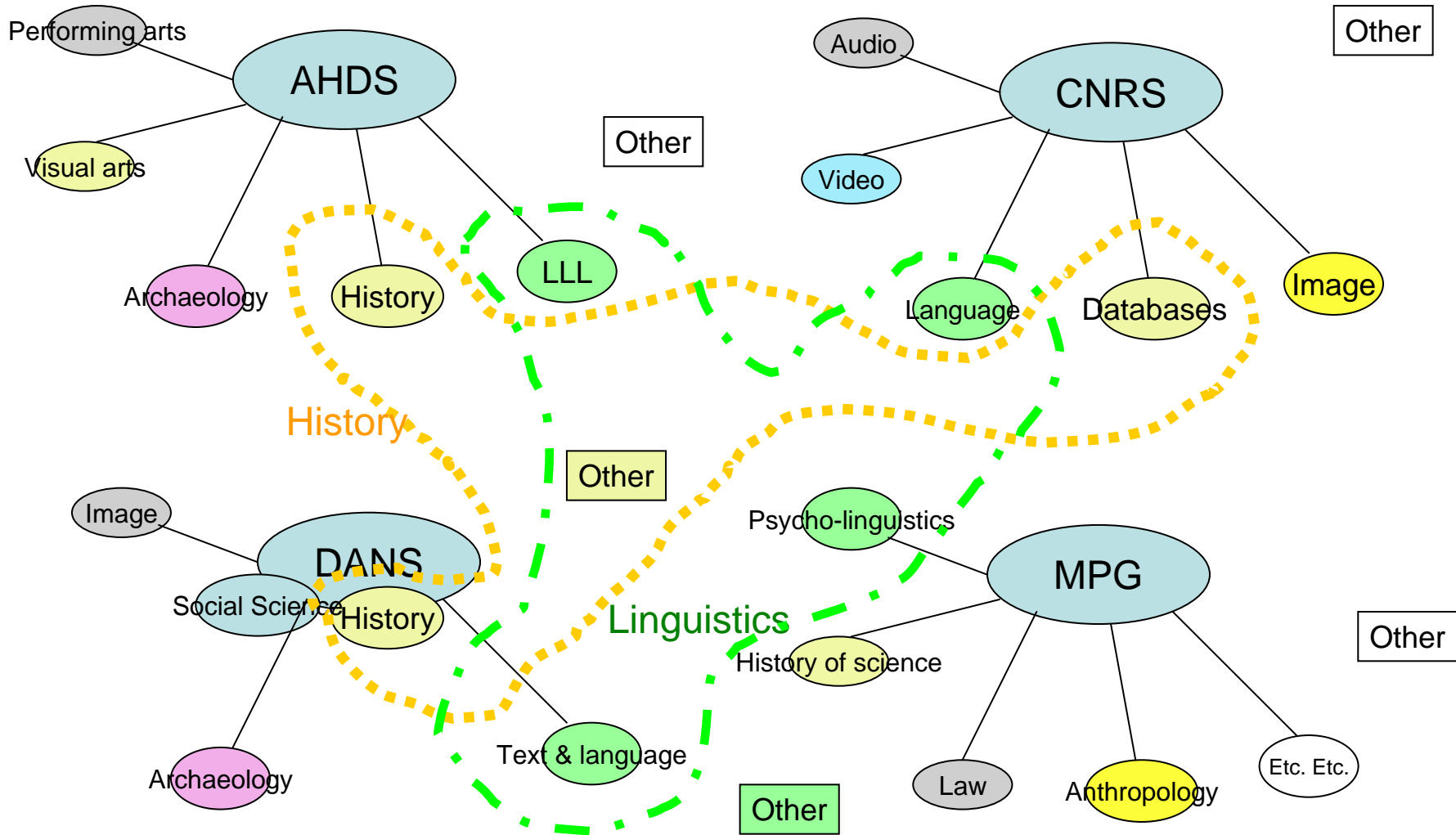


However, European projects have no continuity
The existing structures are too weak (ad hoc
networks, no permanence) and national in scope
Answer: strong European data infrastructure

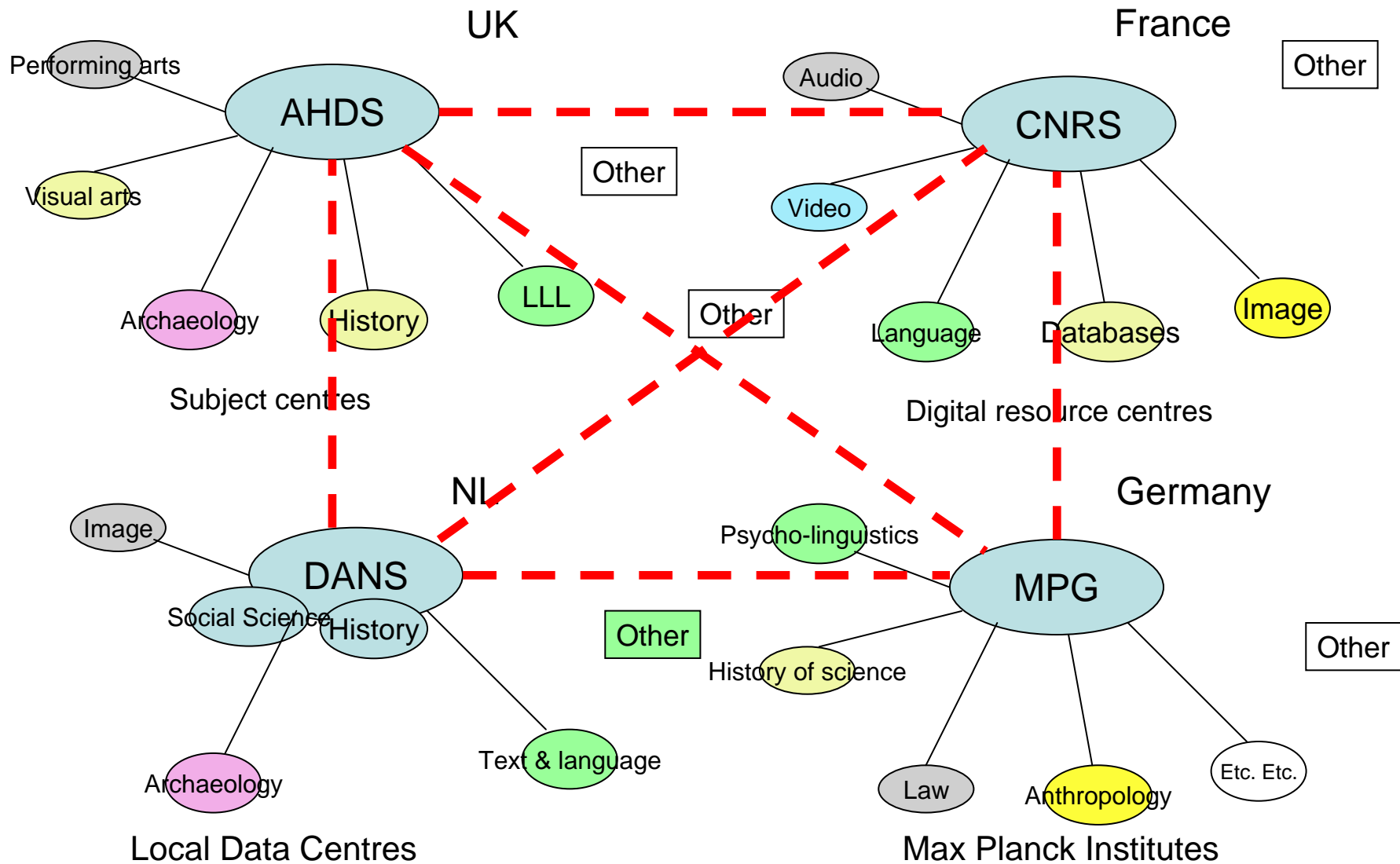
Present situation



Projects and temporary coalitions: examples in two domains



Coalition towards infrastructure



Three Tiered Model

European / International Layer:

- Enabling, coordinating, funding
- Best practice and standards
- Harvesting, harmonisation and combination

National Layer:

- Enabling, coordinating
- Service provision
- Best practice and standards

Local/thematic (domain) Layer:

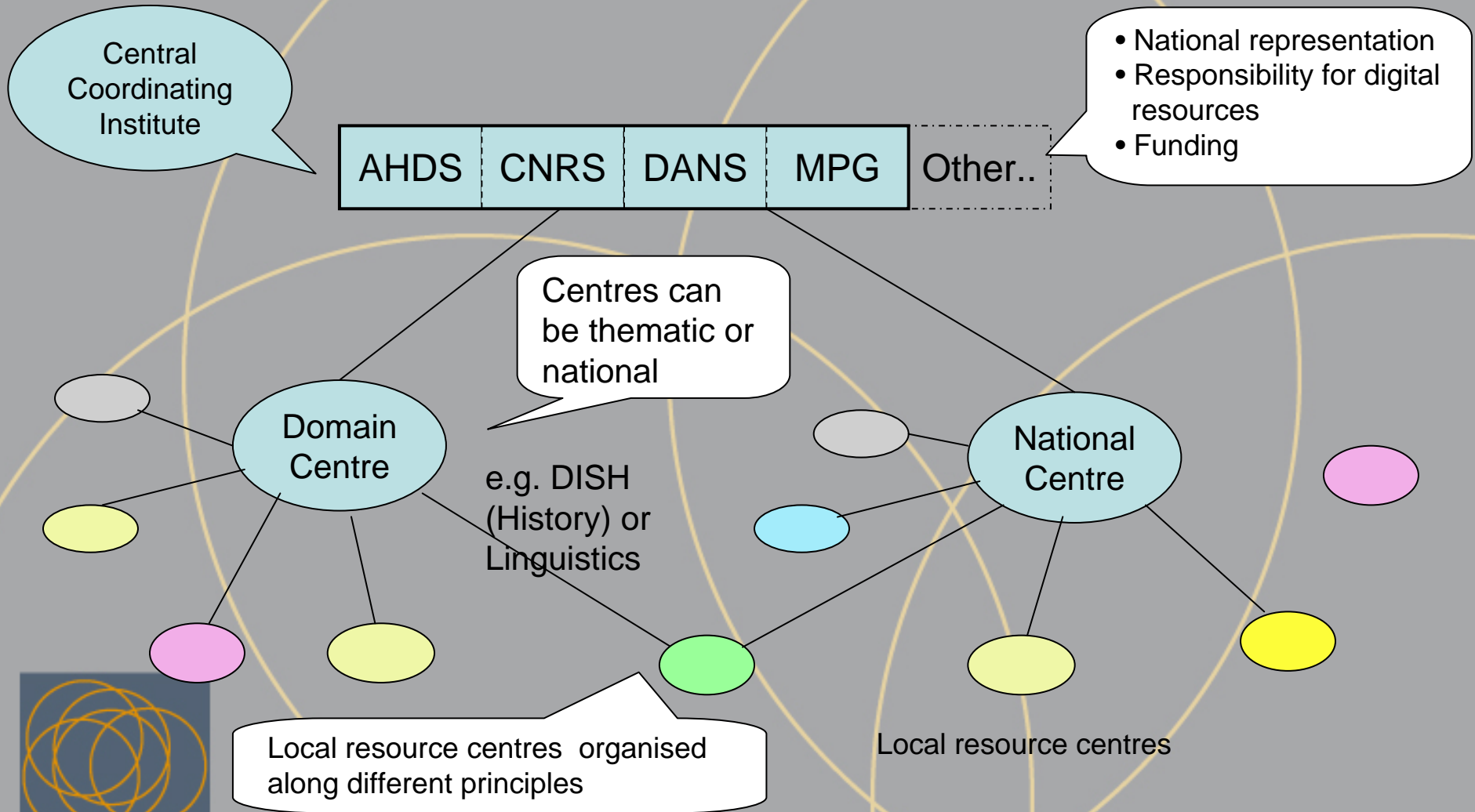
- Projects, communities of practice



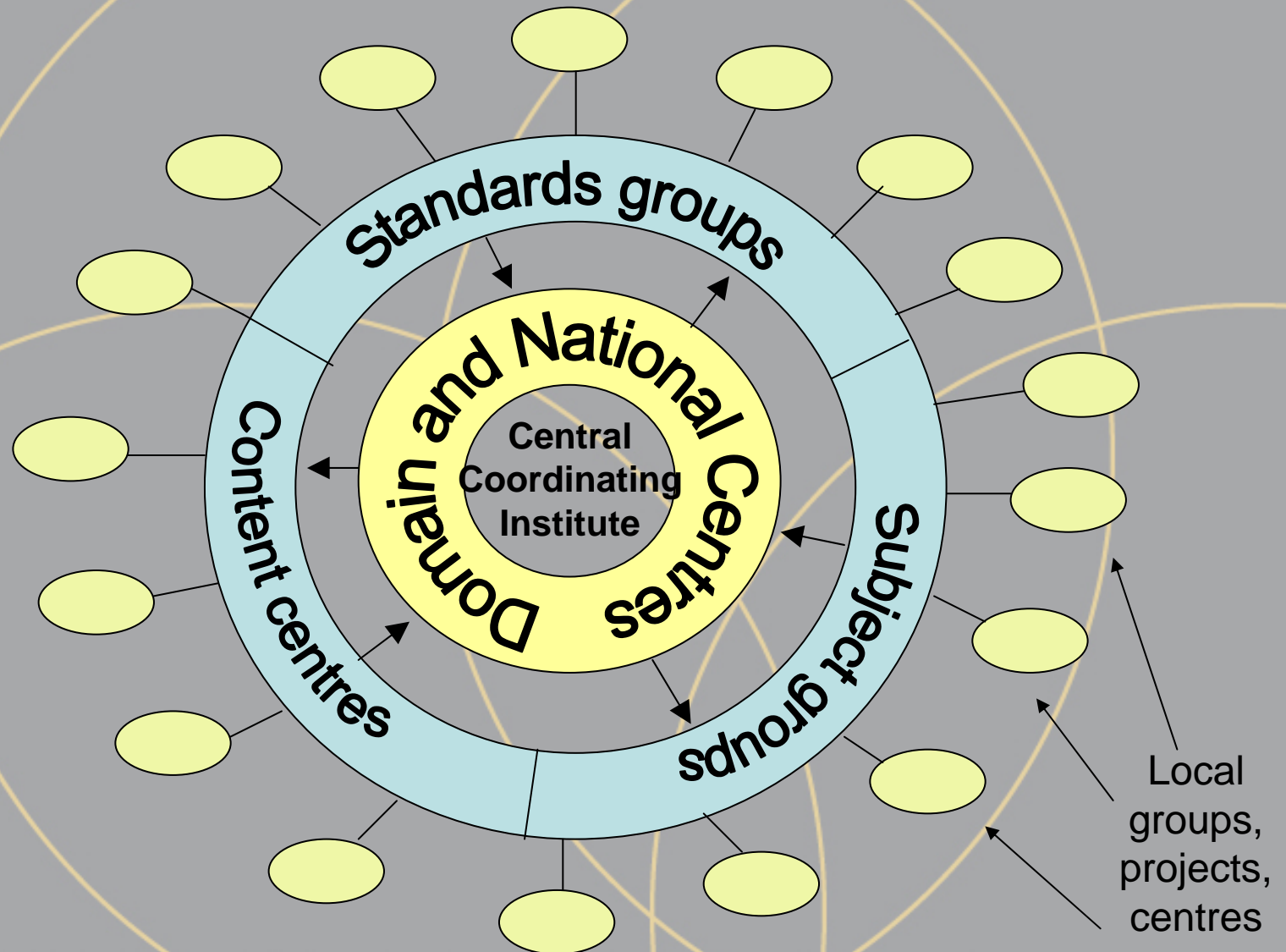
Digital Research Infrastructure for the Arts and Humanities
Infrastructure is to be.

- Flexible and responsive

Organising the infrastructure

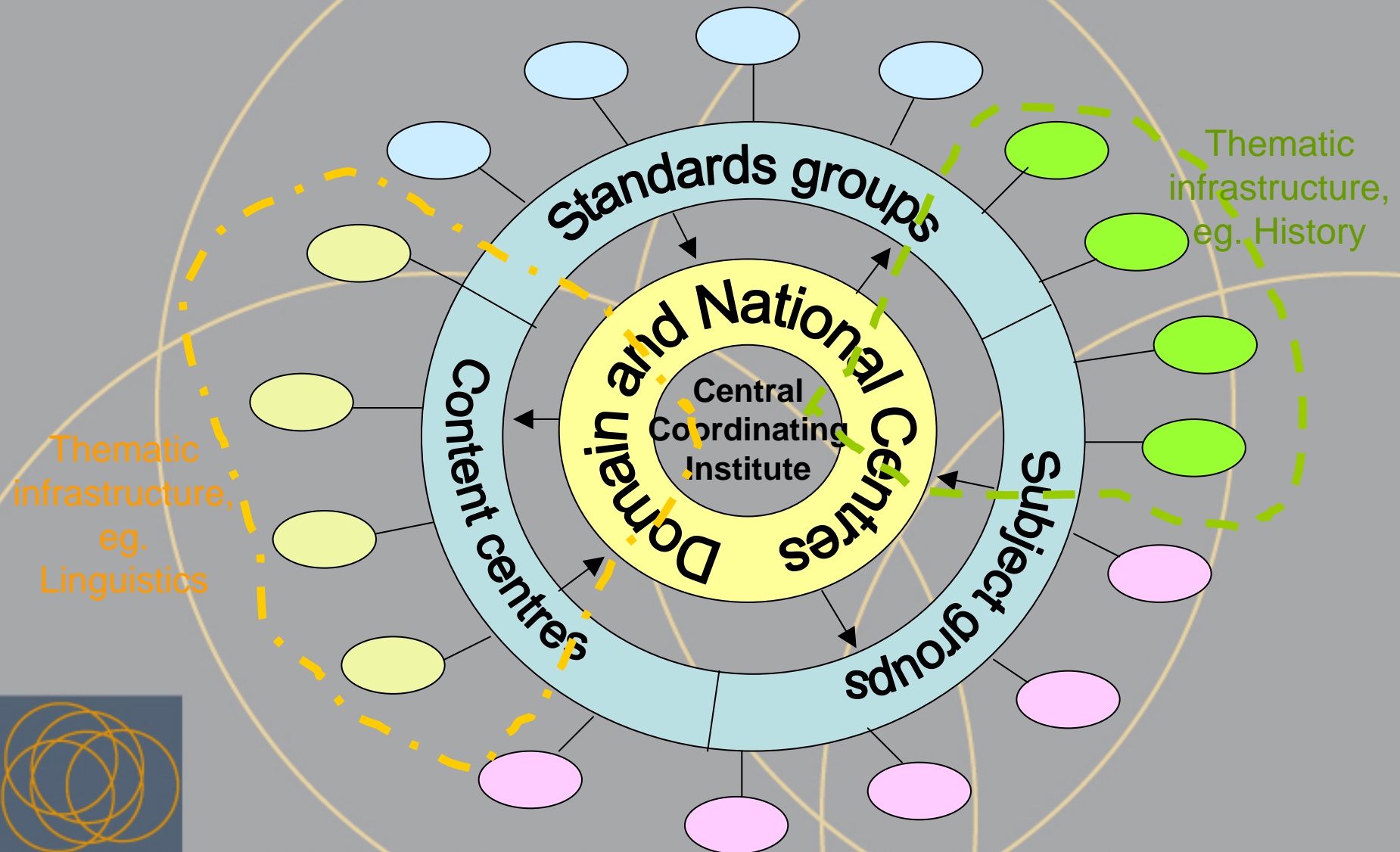


Expanding and consolidating the infrastructure



Digital Research Infrastructure for the Arts and Humanities

Expanding and consolidating the infrastructure



Digital Research Infrastructure for the Arts and Humanities

Summary

Mission: to enhance the European research infrastructure in the humanities

- linking (and upgrading) distributed digital resources and merging them into a grid-empowered architecture
- designing new facilities for pioneering research, preferably of an international and interdisciplinary nature

Structure: a single, core organisation that combines specialist knowledge of the fields with technology expertise in digital information and communication structures



DARIAH

Digital Research Infrastructure for the Arts and Humanities

Summary (continued)

Organising principle: a decentralised network;
a strong core in a cluster of satellites

- The core will bear responsibility for organising and supporting the network, for the basic infrastructure, and for the method and means of communication.
- The decentralised 'hubs' will bear responsibility for the specific thematic or disciplinary expertise. The hubs will be prominent institutes and research networks with a leading role within the European context. The model is an open one and will be able to embrace new, promising fields that are as yet unable to play such a leading role in Europe.





How to contact DANS?

www.dans.knaw.nl

info@dans.knaw.nl

Postal Address
P.O Box 93067
2509 AB The Hague

visiting address
Borneo Building
Anna van Saksenlaan 51
2593 HW The Hague

T +31 70 3494450

F +31 70 3494451

