Data Archiving and Networked Services

DANS

Towards a Digital Research Infrastructure for the Humanities in Europe

Peter Doorn
Director, DANS

peter.doorn@dans.knaw.nl
www.dans.knaw.nl
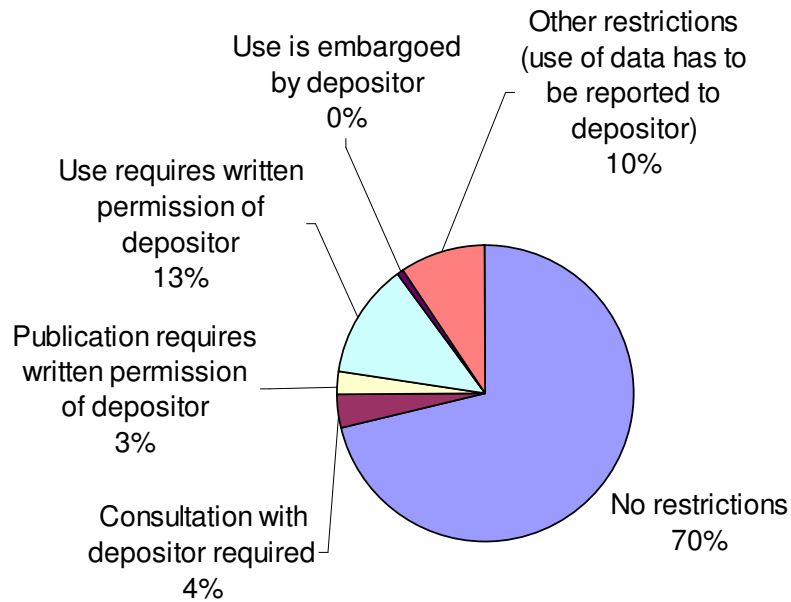
Open Access Conference, Berlin, 29 March 2006

Data Archiving and Networked Services

DANS

Contents

➢ DANS and open access to research data in the Netherlands

➢ The European context:

– Need for a European Digital Research Infrastructure for the Humanities

– Outline and model for its development

Data Archiving and Networked Services

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

Data Archiving and Networked Services
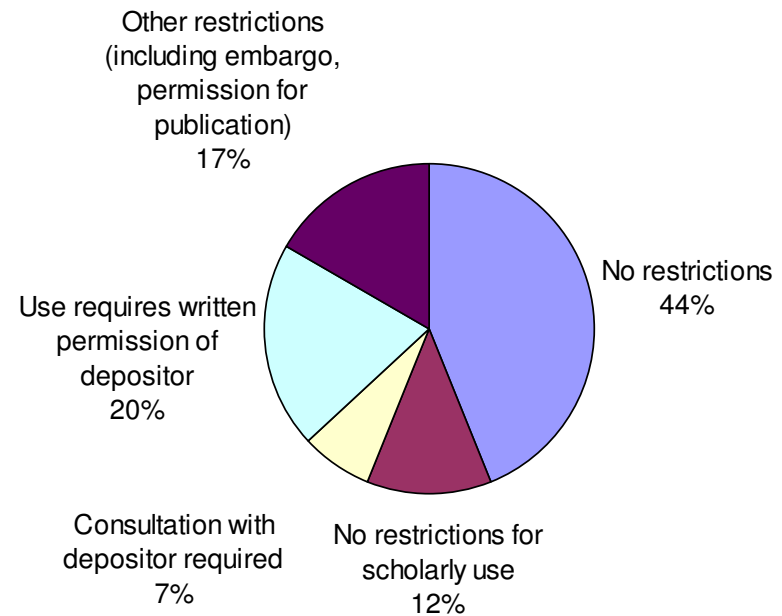
DANS and open access

- ➢ Dutch ministers of economic affairs and of education, culture and sciences favour open access (OECD agreement)
- ➢ KNAW and NWO signed Berlin declaration
- ➢ How DANS works on open access:
  - DANS abolished all subscriptions, fees and other payments for data in its digital archives
  - Code of conduct for researchers for secure access
  - Authorisation system for access to privacy sensitive data
- ➢ Practice is less favourable:
  - Researchers want some degree of protection and control on who accesses "their" data
  - Official data providers still expect DANS to continue past data contracts

Access restrictions in two data archives

Steinmetz Archive

Use is embargoed by depositor
0%

Other restrictions (use of data has to be reported to depositor)
10%

Use requires written permission of depositor
13%

Publication requires written permission of depositor
3%

Consultation with depositor required
4%

No restrictions
70%

Historical Data Archive

Other restrictions (including embargo, permission for publication)
17%

Use requires written permission of depositor
20%

Consultation with depositor required
7%

No restrictions for scholarly use
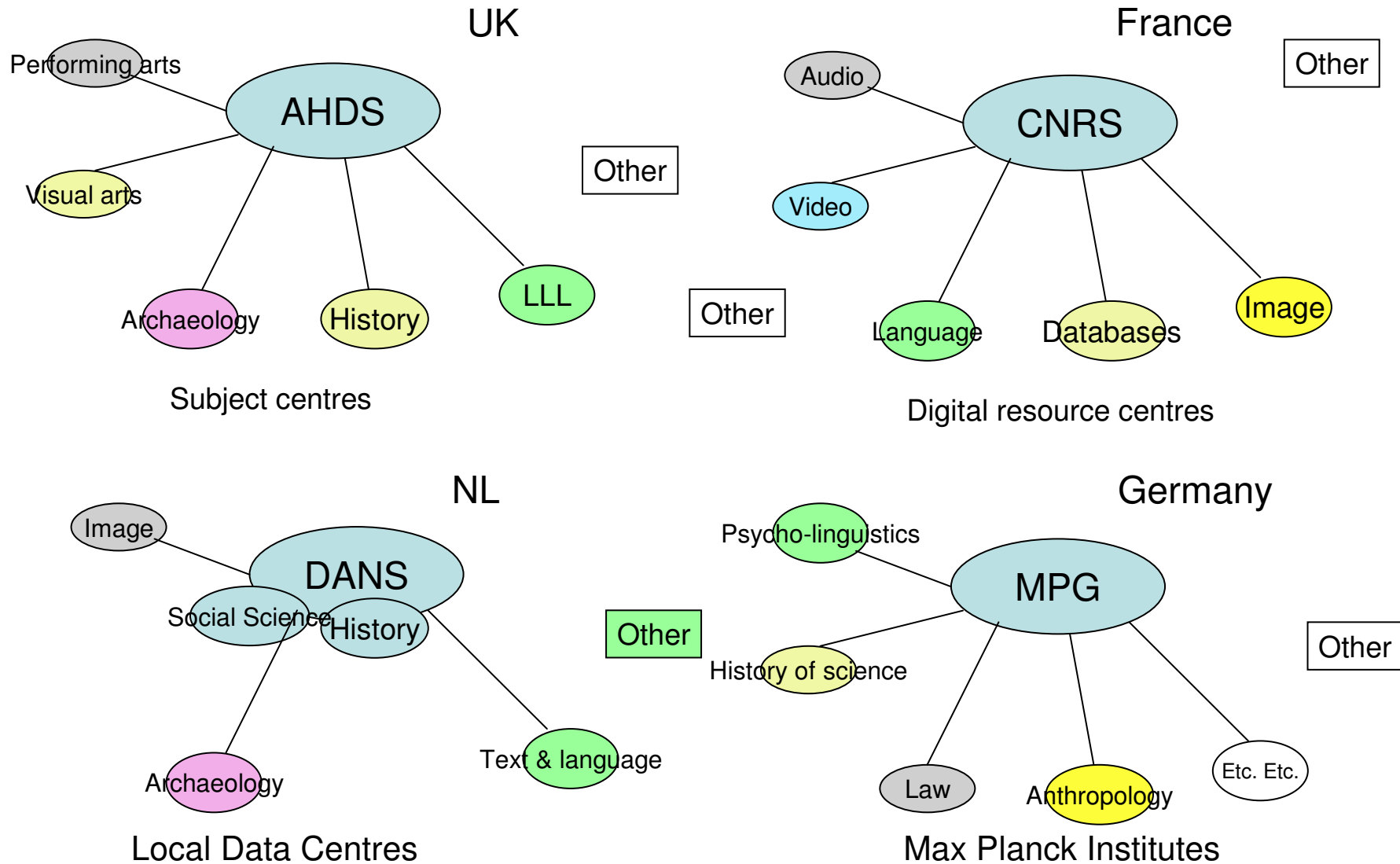12%

No restrictions
44%

DANS

European context

➤ Existing data organisations (like DANS) will function in the context of emerging European research infrastructures for the humanities (and social sciences)

➤ ESFRI Roadmap: European Strategy Forum for Research Infrastructures

  – Within ESFRI working group for Social Sciences and Humanities two Expert Groups:

  – EROHS: concentrating on social sciences

  – Cultural Heritage: concentrating on humanities

➤ Research infrastructures in general: **permanent** and **physical**

➤ Research infrastructures for the natural sciences: ice breakers for polar research, satellites, telescopes, particle accelerators, laboratories

➤ Cultural heritage in all forms is the main source of humanities research

➤ Libraries and archives are the traditional "laboratories" for the humanities

➤ In the digital age, essential for innovative humanities research is:

  – Access to digitised heritage data (data bases, text corpora, speech, image collections, etc.)

  – Tools to process this information

➤ The most important new research infrastructure for the humanities is therefore a digital one

➢ The Grand Vision:

  – Provide Access to all surviving European humanities and cultural heritage information across all of 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

  – Support research networks in the humanities and social sciences

  – Structure: a strong nucleus in a cluster of networked organisations and sattellites
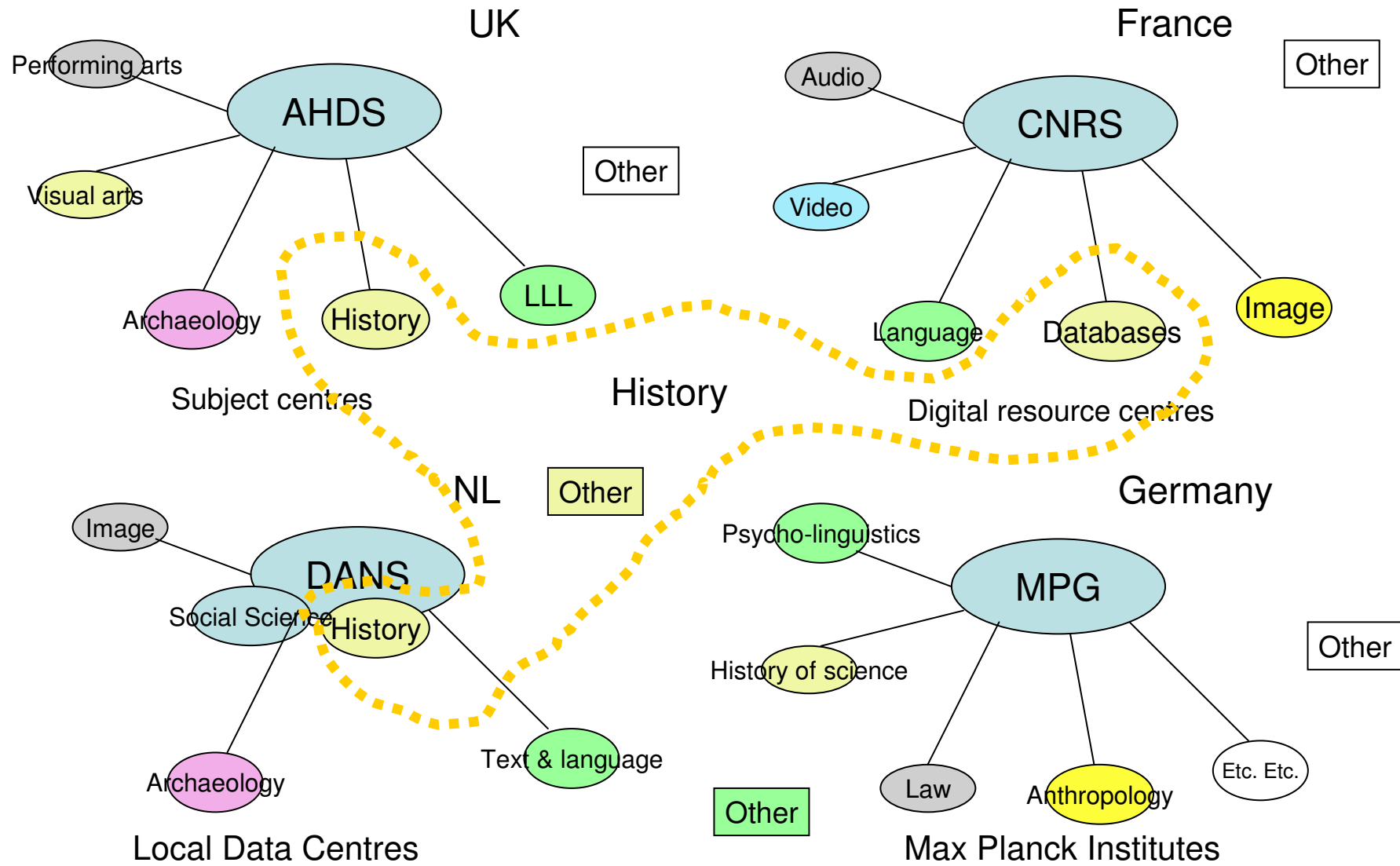
Data Archiving and Networked Services

DANS

Outline of tasks of European RI

➢ 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
- – Supporting communities of practice
- – Facilitating new research practices
- – Tools development and tools registries

Data Archiving and Networked Services

DANS

Science Case

➤ Changing research practice in a networked environment:

    – Data (including text, images, and other media) is the laboratory of the scholar in the social sciences and humanities

    – Distributed resources on the web (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 too national in scope

➤ Answer: strong European data infrastructure

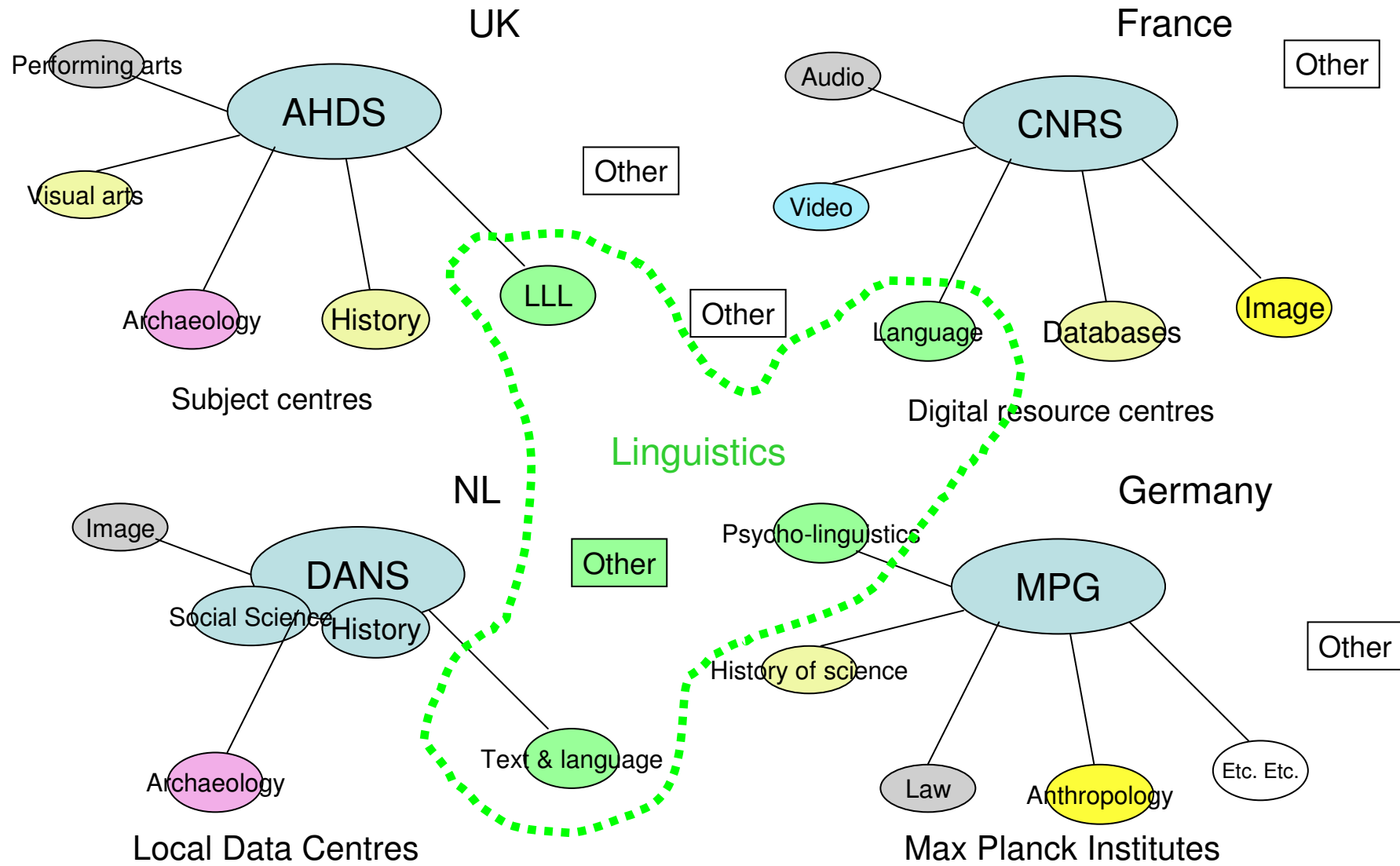➤ Start: four strong national organisations are forming the starting point for emerging pan-European RI
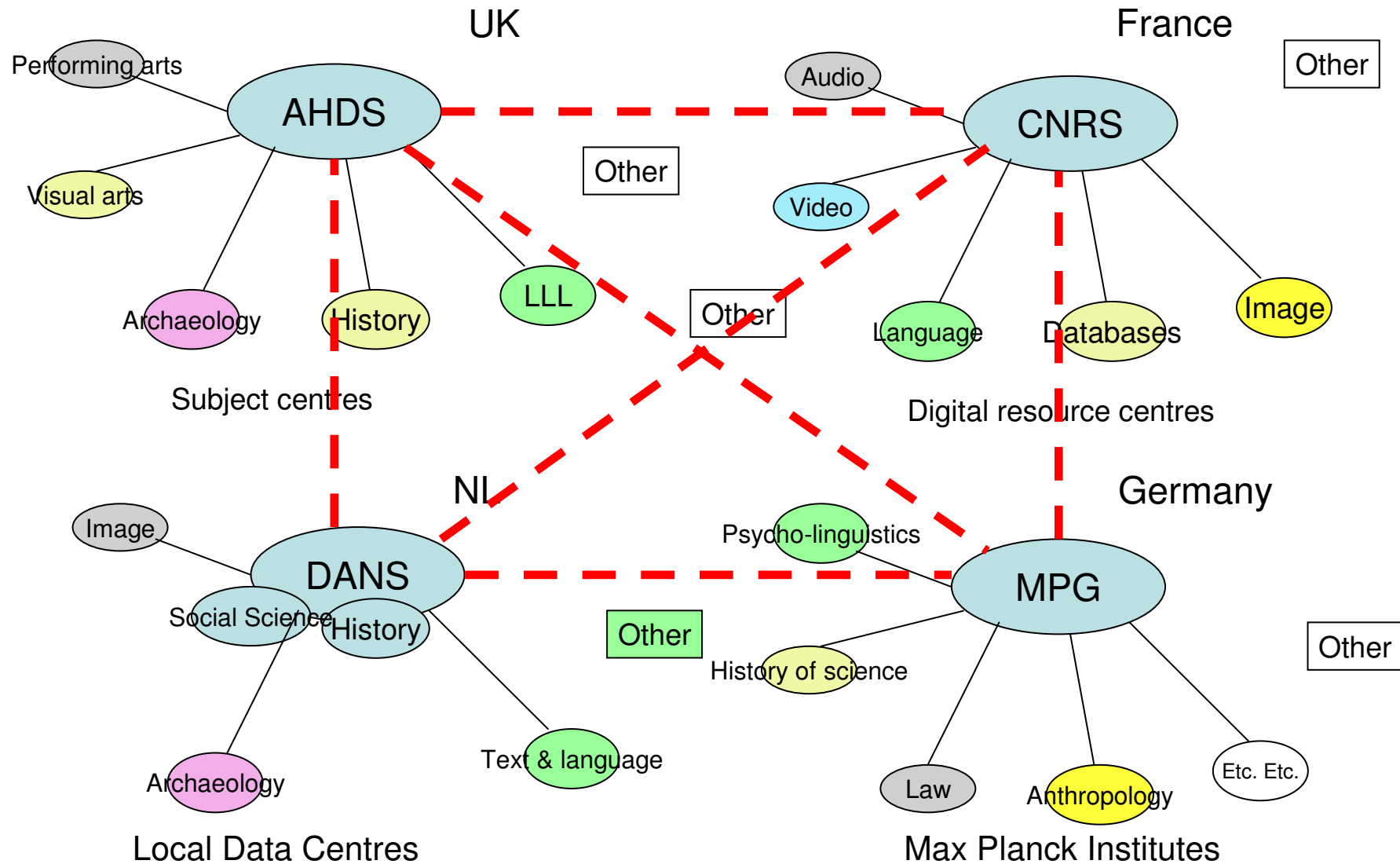
Present situation

UK

Performing arts

AHDS

Other

Visual arts

Archaeology   History   LLL

Subject centres

France

Audio

CNRS

Other

Video

Other

Language   Databases   Image

Digital resource centres

NL

Image

DANS

Social Science   History

Other

Archaeology

Text & language

Local Data Centres

Germany

Psycho-linguistics

MPG

History of science

Other

Law   Anthropology   Etc. Etc.

Max Planck Institutes

Projects and temporary coalitions: History example

UK

Performing arts

AHDS

Visual arts

Archaeology

History

LLL

Other

Subject centres

France

Audio

CNRS

Video

Language

Databases

Image

Other

Digital resource centres

History

Other

NL

Image

DANS

Social Science

History

Archaeology

Text & language

Local Data Centres

Other

Germany

Psycho-linguistics

MPG

History of science

Law

Anthropology

Etc. Etc.

Other

Max Planck Institutes

Projects and temporary coalitions: Linguistics example
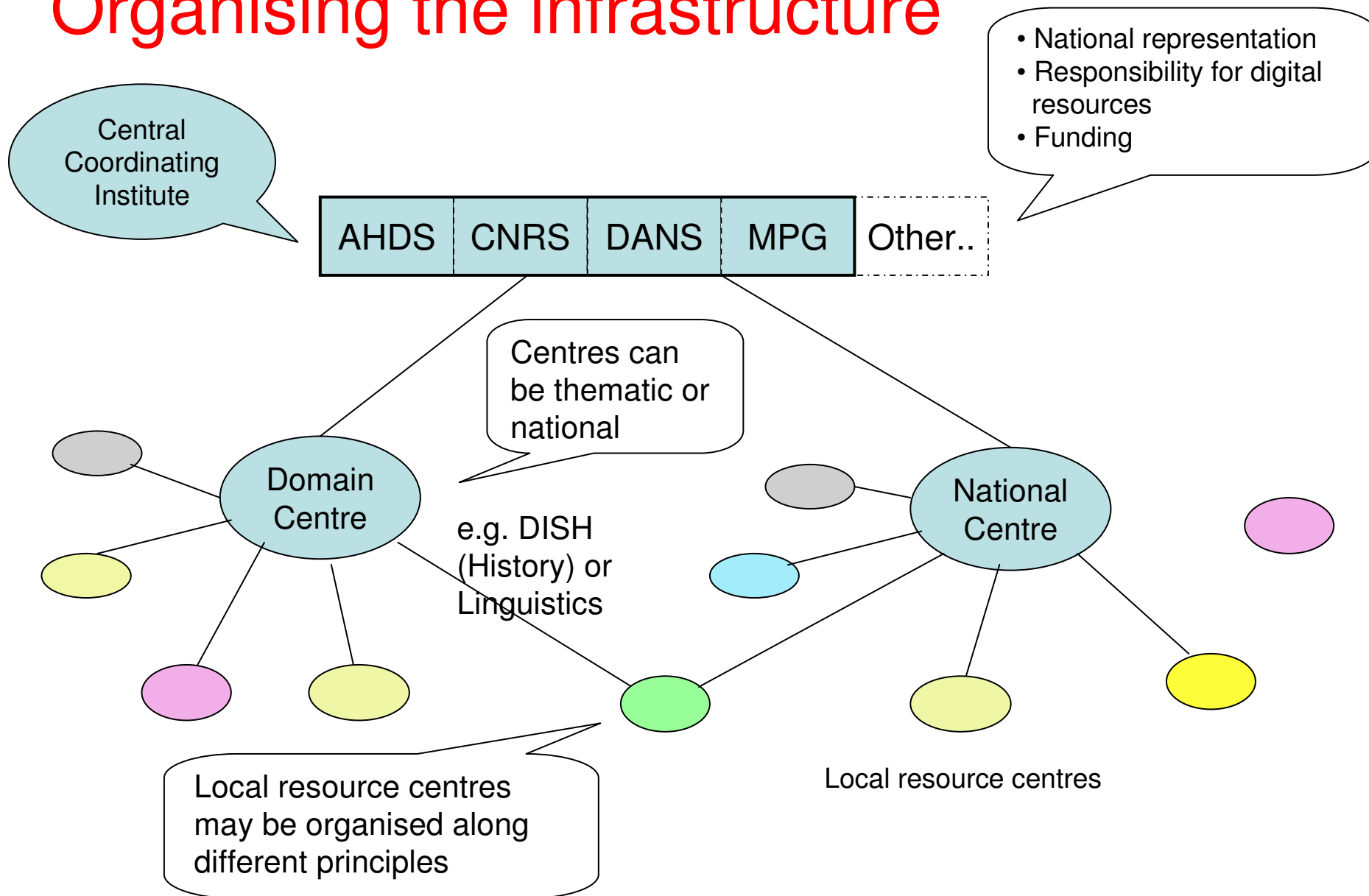
Coalition towards infrastructure
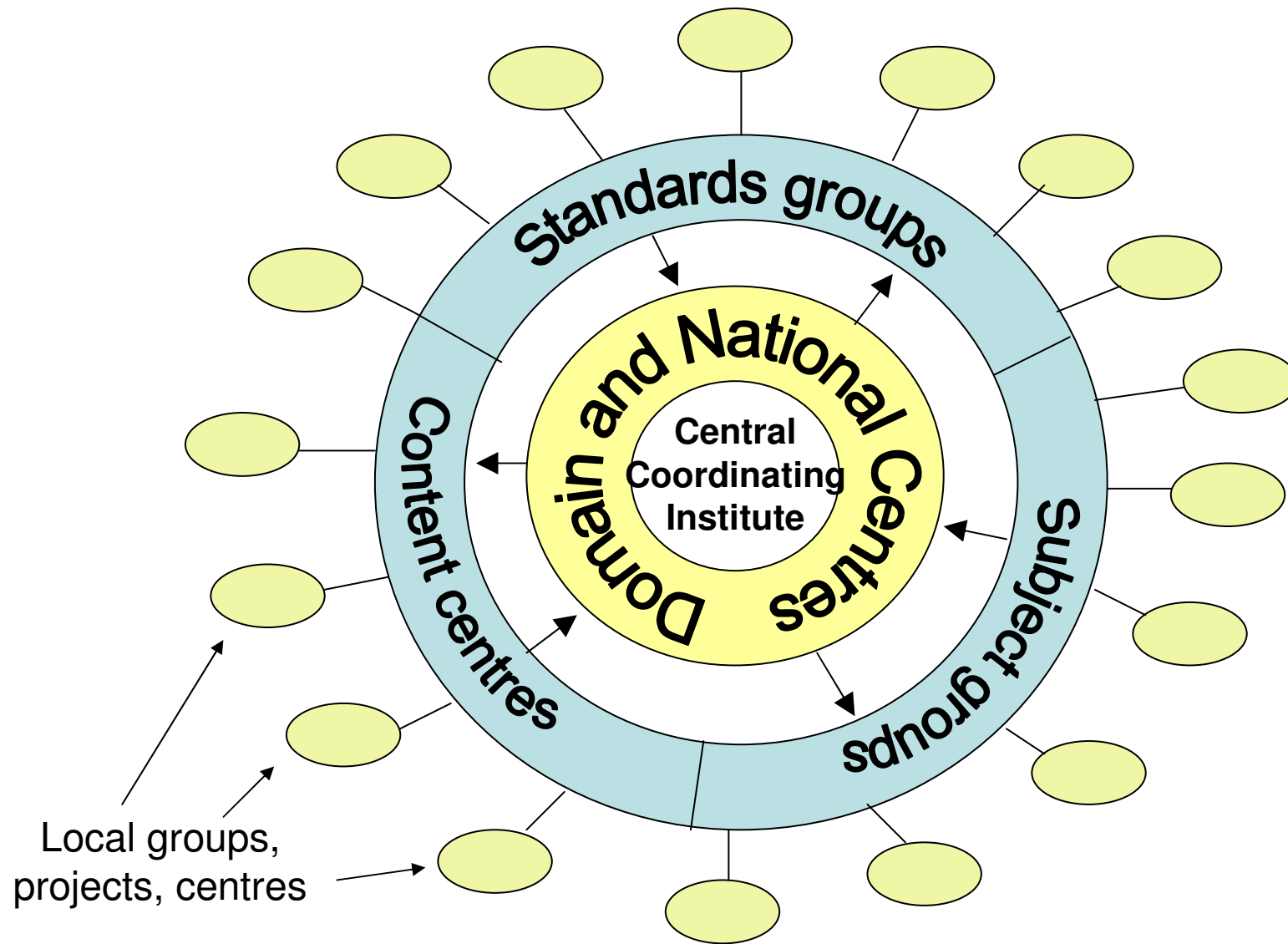
Data Archiving and Networked Services

DANS

➢ 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

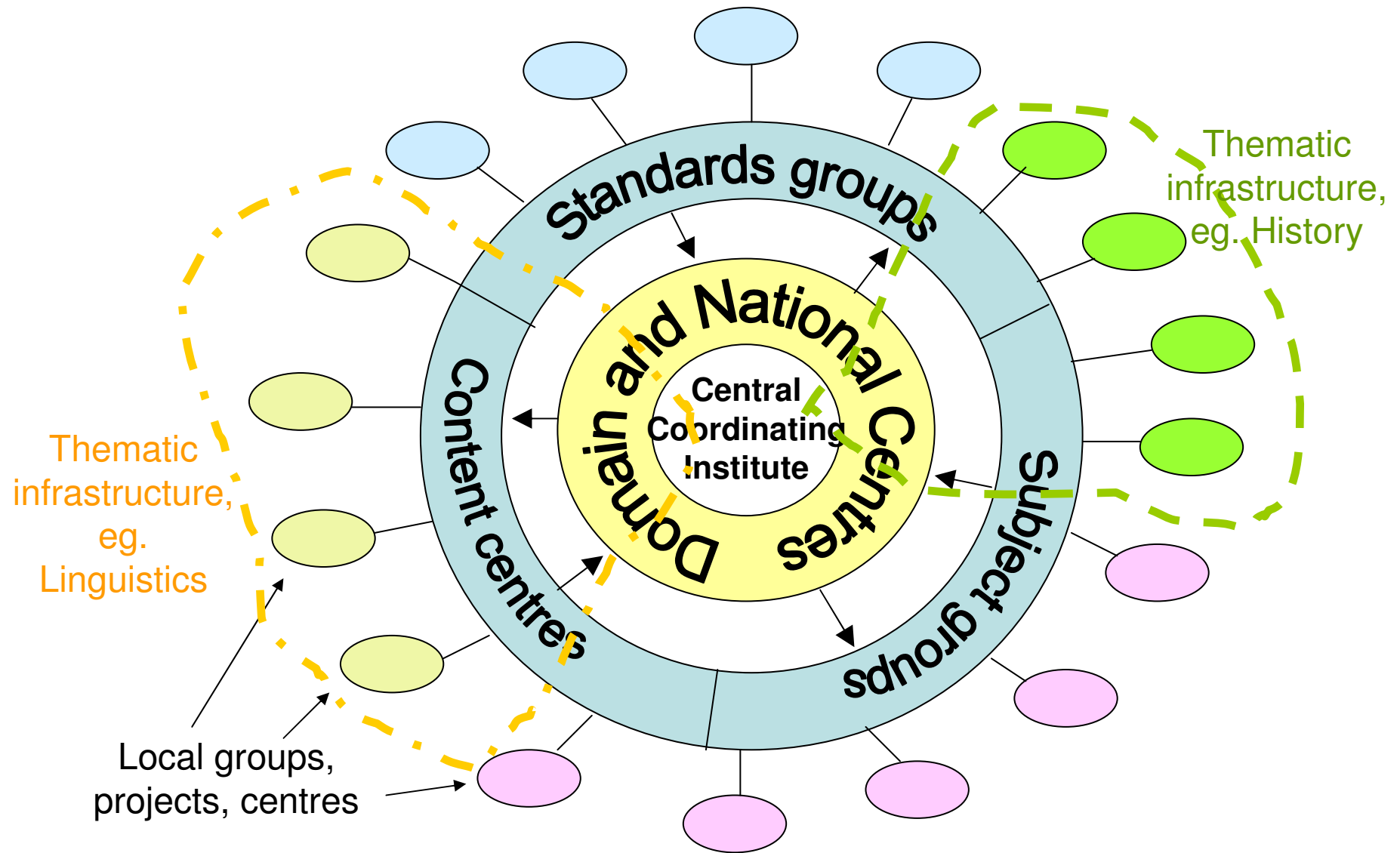Flexible and responsive; reciprocal relationships, subsidiarity

Organising the infrastructure

Expanding and consolidating the infrastructure

Standards groups

Content centres

Domain and National Centres

Subject groups

Central Coordinating Institute

Local groups, projects, centres

Expanding and consolidating the infrastructure

Thematic infrastructure, eg. History

Thematic infrastructure, eg. Linguistics

Standards groups

Content centres

Domain and National Centres

Central Coordinating Institute

Subject groups

Local groups, projects, centres

Summary

➢ <u>Mission</u>: to enhance the European research infrastructure in the humanities (and social sciences?)
  – 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
➢ <u>Structure</u>: a single, core organisation that combines specialist knowledge of the fields with technological expertise in digital information and communication structures
➢ <u>Organising principle</u>: 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.