



Digital Research Infrastructure
for the Arts and Humanities

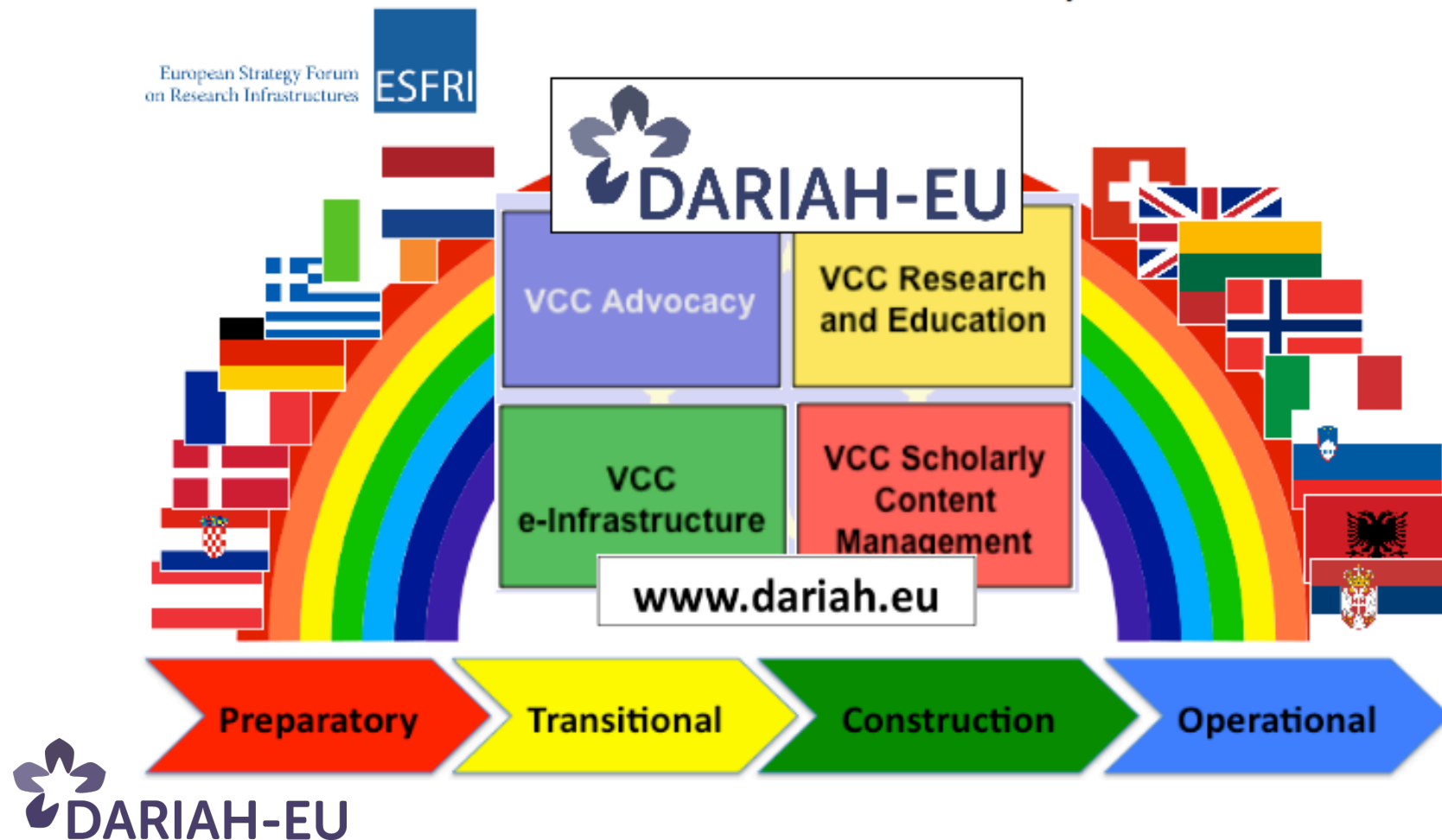
The Digital Research Infrastructure for the Arts and Humanities (DARIAH)

Background: ESFRI
Achievements to date
Organisation
Participation
Collaboration

www.dariah.eu

Digital Research Infrastructure for the Arts and Humanities (DARIAH)

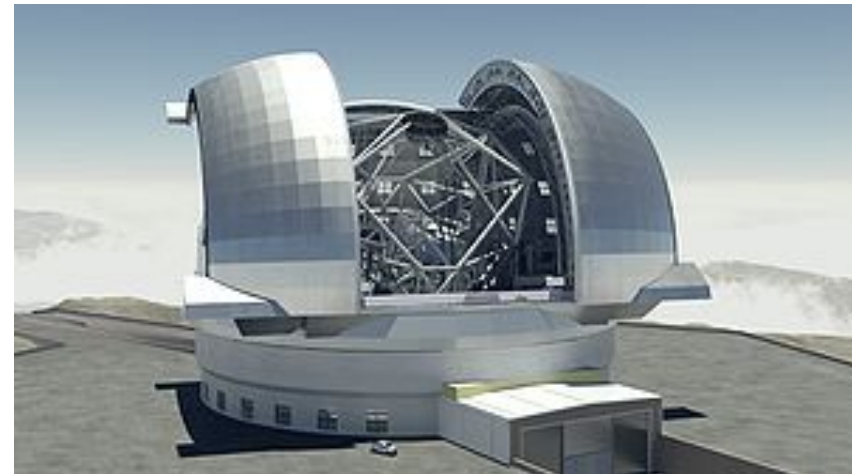
To enhance and support digitally-enabled research across the arts and humanities spectrum

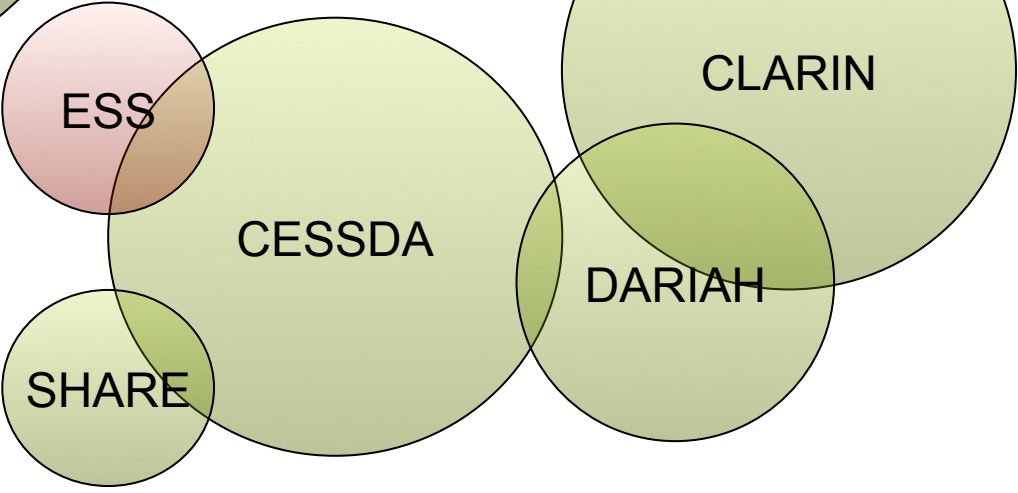
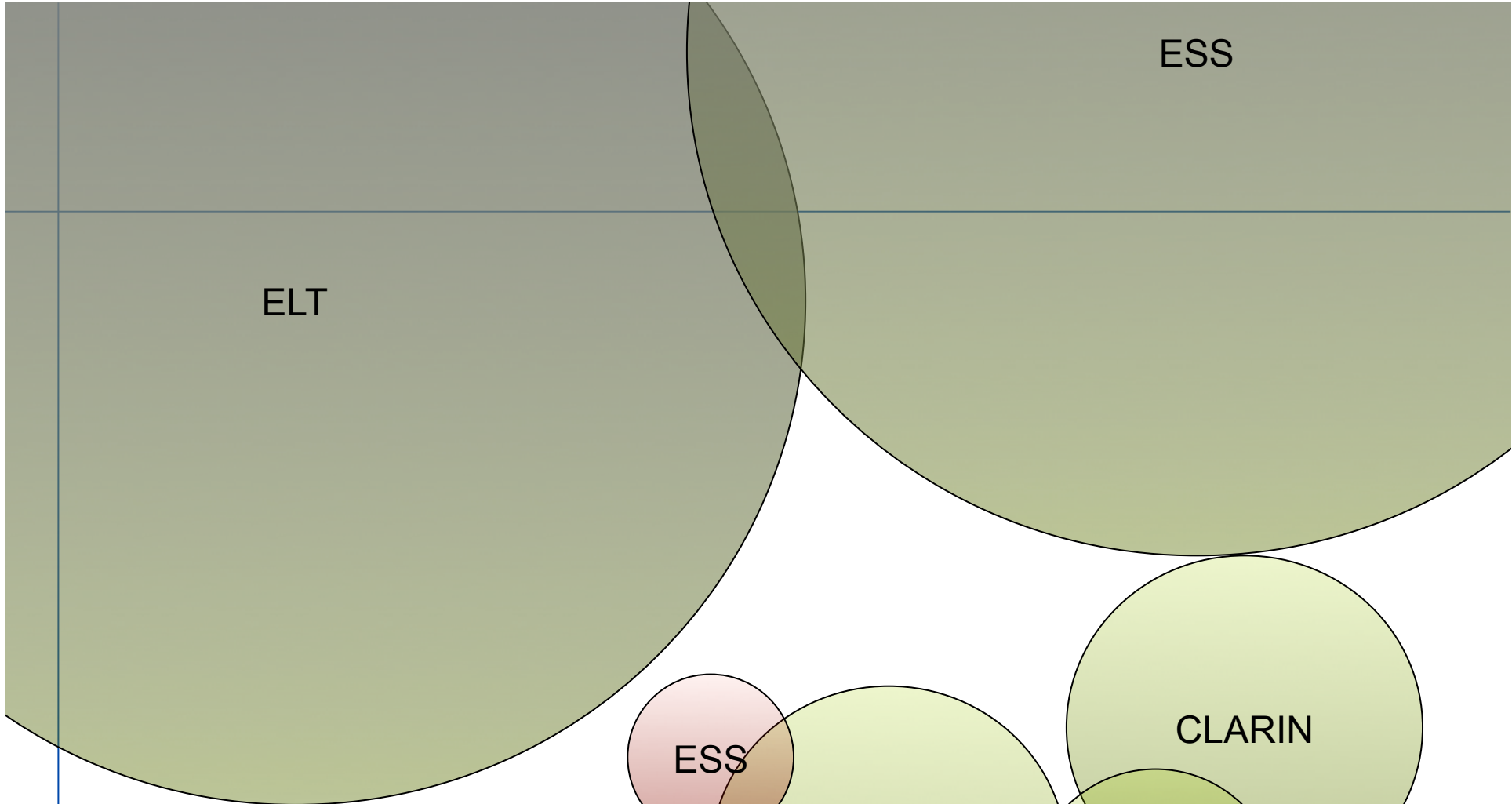


European Strategy Forum of Research Infrastructures (ESFRI)

ESFRI

Projects (in alphabetical order per discipline)	Estimated Construction Cost (M€) *	First possible operations for users	Indicative Operational/ Deployment Cost (M€/year)
ELT: The European Extremely Large Telescope	850	2018	40
Infrastructure for Clinical Trials and Biotherapy Facilities	36	2007	5





DARIAH on the ESFRI Roadmap

DARIAH

The facility

DARIAH (Digital Infrastructure for Humanities) will build on existing networks and services based on the Max Planck Society in the Netherlands, the United Kingdom and is however open to embrace new facilities from European countries. Online (ECHO), a Digital Infrastructure for cultural heritage.

Digital Research



Strategy Report
on Research Infrastructures

Roadmap 2010

- Multi-national collaborations
 - Funded until recently by the European Union to prepare the infrastructure
 - Need member states to pay fees for the sustained infrastructure
- Budget: €4m per year
- DARIAH: Several national projects with a relatively small European organisation (~10-20%)

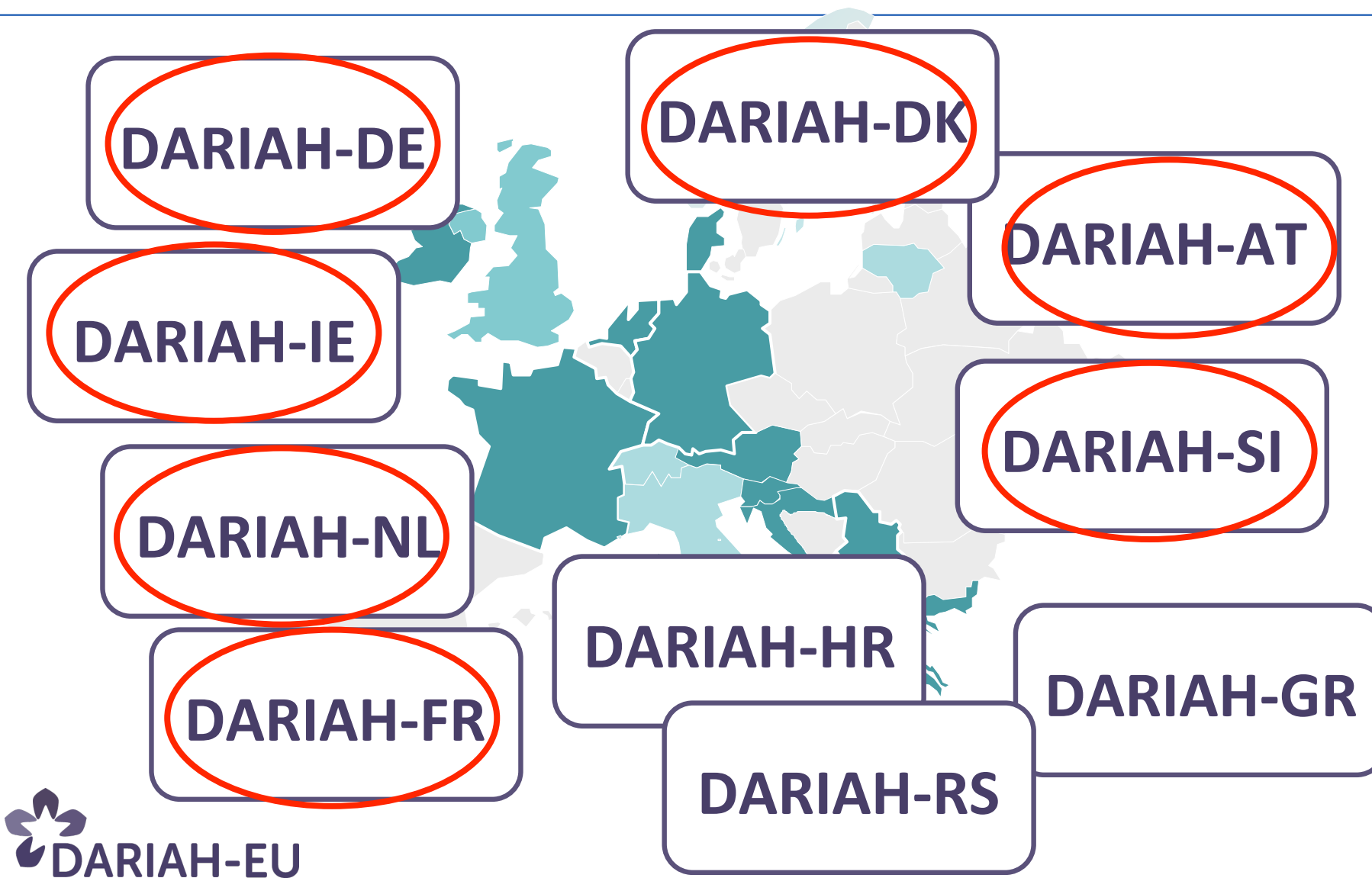


Manuscript studies (Copyright: DARIAH).

- Operations: 2.4 M€/year
- Decommissioning: not applicable

www.dariah.eu

National initiatives



Projects and initiatives in a larger circle

Open Aire

EGI

GRDI2020

Other SSH initiatives

EUDAT

Sibling initiatives

CLARIN

TEI

Affiliated projects

NeDiMAH

DASISH

BAMBOO

 DARIAH-EU

DESIRE

CENDARI



EHRI

DC-Net

CULTURA

ARROW





Achievements

To date and under progress

www.dariah.eu

September
11 2012



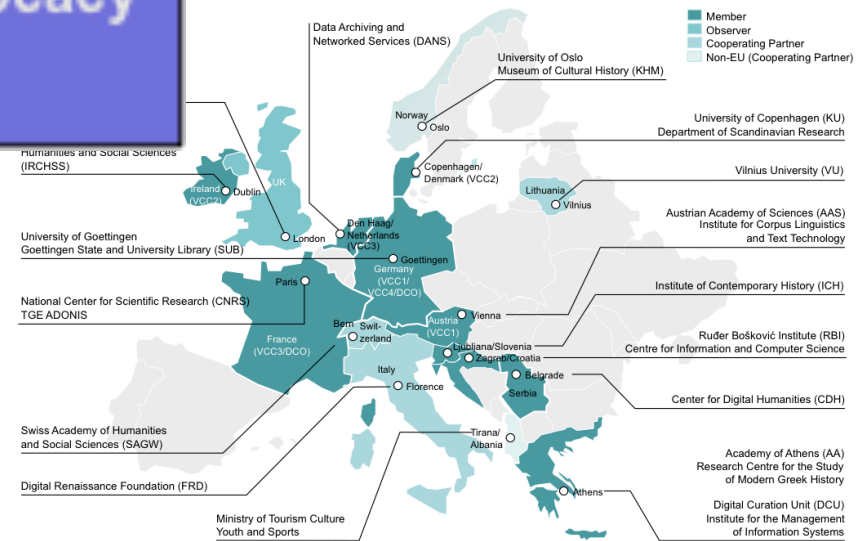
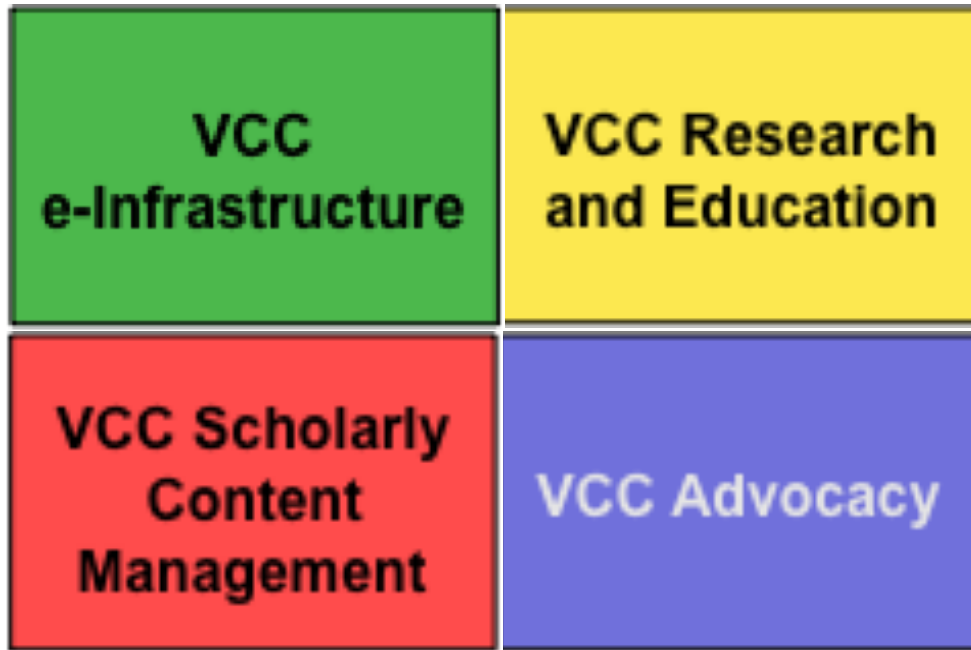
Digital Research Infrastructure
for the Arts and Humanities

Organisation

Forget project – think institution
Sustainable institutional
framework (under construction)

www.dariah.eu

DARIAH Virtual Competency Centres (VCCs)



DARIAH Virtual Competency Centres (VCCs)

VCC
e-Infrastructure



- To establish a shared technology platform for Arts and Humanities research

VCC Research
and Education



- To expose and share researcher's knowledge, methodologies and expertise

VCC Scholarly
Content
Management



- To facilitate the exposure and sharing of scholarly content

VCC Advocacy



- To interface with key influencers in and for the Arts and Humanities

Establishing an European Research Infrastructure Consortium (DARIAH-ERIC)

- **Submitted early September:** Statutes and scientific-technical description
- **Timeline to ERIC: Summer 2013**
 - Feedback on documents by commission and peer review
 - Revised version signed by all member states
 - ERIC commission
- **German/Dutch/French:** One of only three ERIC's submitted by France





Digital Research Infrastructure
for the Arts and Humanities

If you are interested in the formal structure of DARIAH, please come and ask me in the lunch break

There are many more abbreviations to learn: CIO, GS, GA, COO, ...

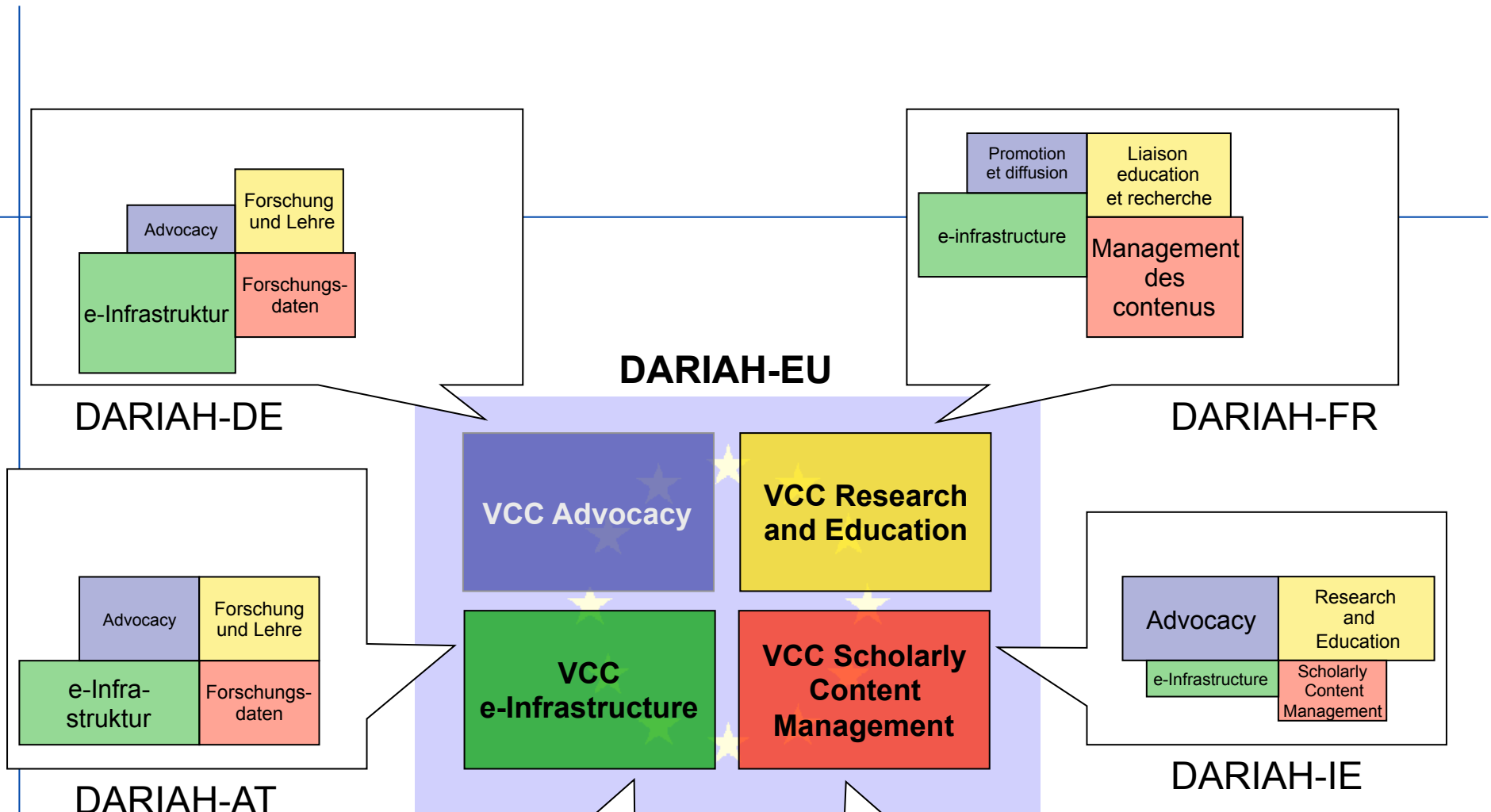
www.dariah.eu



Participation

Scaling up national projects

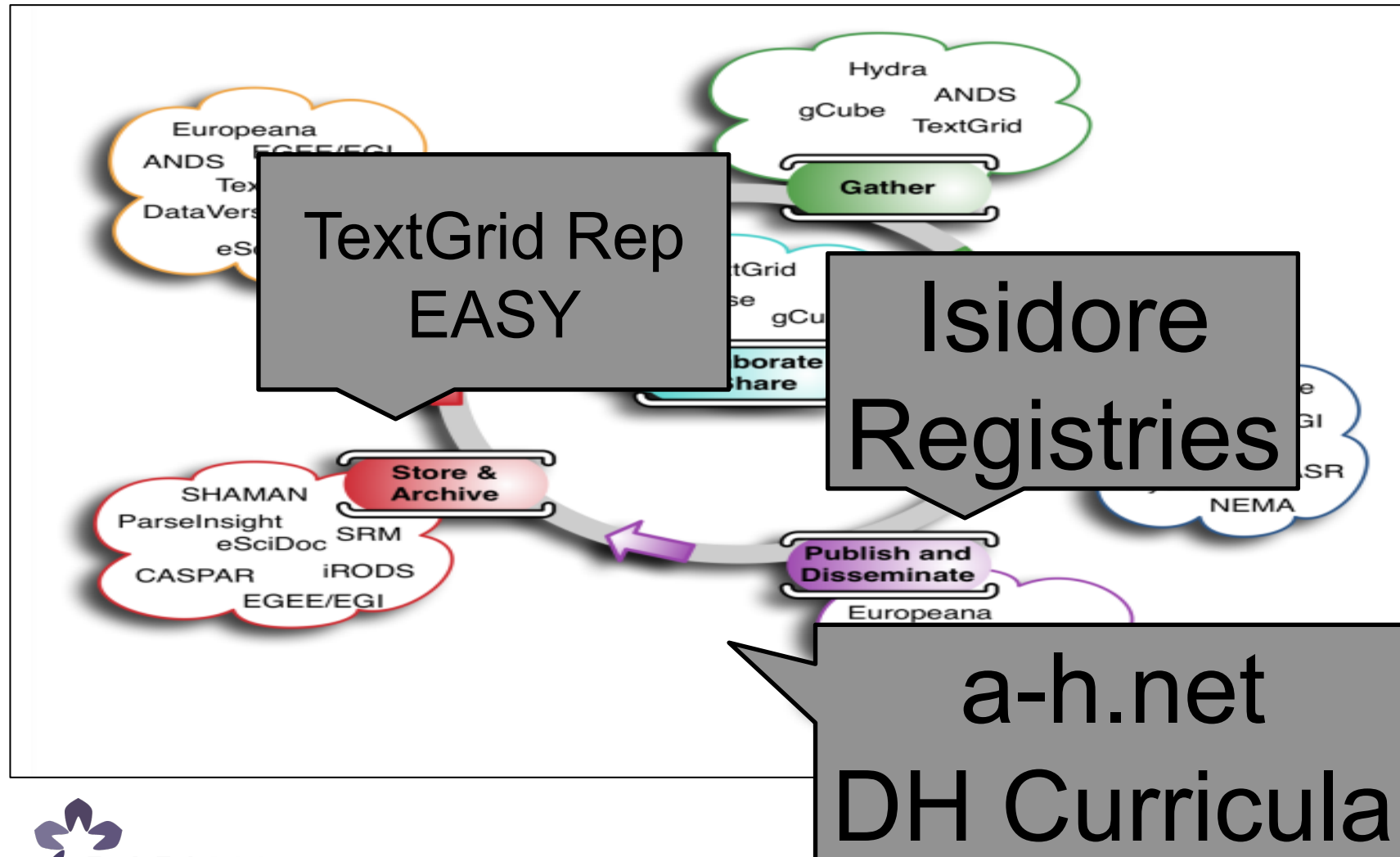
www.dariah.eu



Active since the beginning of 2012

- General VCC meeting in April 2012 in Utrecht, No 2 planned for Nov in Vienna
- VCC heads regular virtual meetings
- Individual VCC meetings
- First common services

Integrating Activity





Digital Research Infrastructure
for the Arts and Humanities

If you are interested in the content structure of DARIAH, please come and ask me in the lunch break

There are lots of existing collaborations, projects and no acronyms

www.dariah.eu

Collaboration

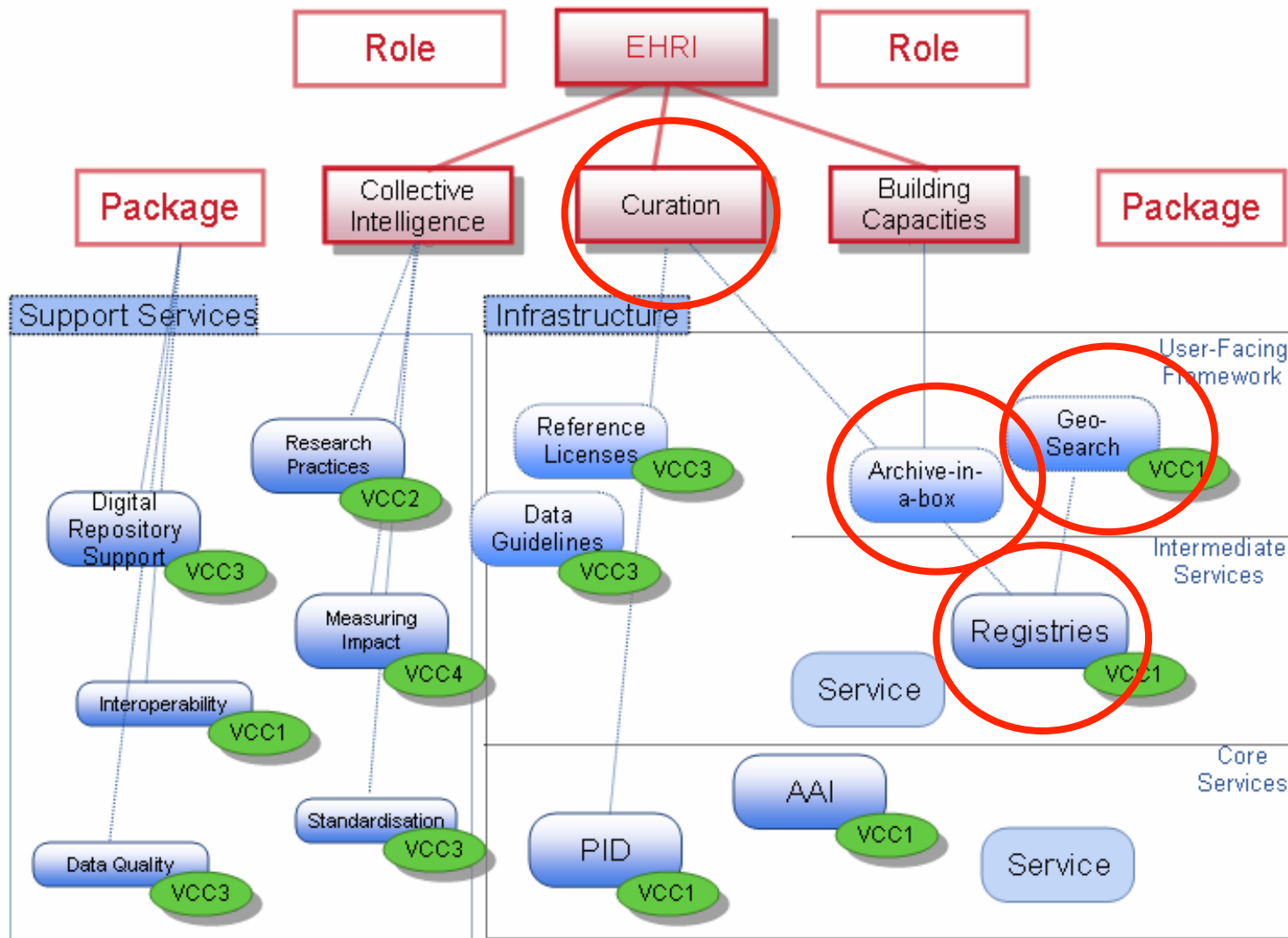
European Activities

DARIAH at work in Europe

- **Enable discipline-specific infrastructures**
 - Integrating activities
 - **EHRI and CENDARI**
 - **ARIADNE**
- **Research into e-Research**
 - Analysis of methods: **NeDiMah**
- **Collaborations with other Infrastructures**
 - Europeana Research Cloud
 - DM2E
 - DASISH



EHRI Packages



In summary

- ESFRI and its complications
- DARIAH as an organisation
- DARIAH as integrating activity
 - VCC's scaling up national activities
- DARIAH as a partner in European projects



Thank you

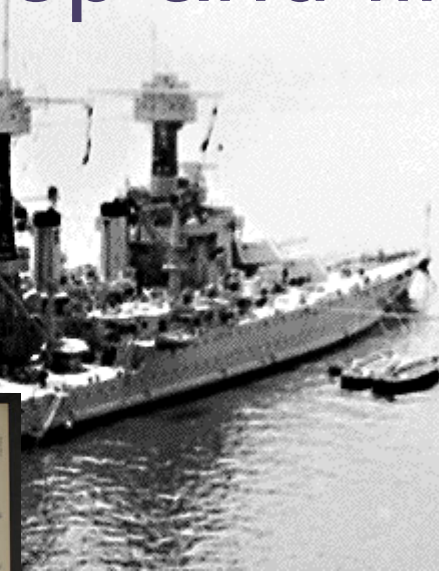
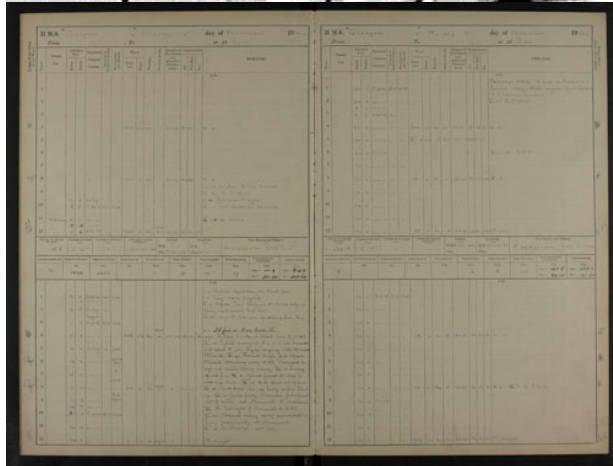
DARIAH-EU Coordination Office
dariah-info@dariah.eu

www.dariah.eu

dariah-info@dariah.eu

Digging deep and find

- Patient Records
- Historic Weather
- Place Names
- Stormont



Great Ormond
Street Children's Hospital



DARIAH-EU <http://kcl.ac.uk/cerch>

LARGE HADRON COLLIDER
13 PETABYTES (2010)
The proton collider, near Geneva, Switzerland, generates about 15 petabytes of data per year — even after rejecting 99.9995% of collisions.

HEAVY-DUTY DATA

The computer-storage space required to support projects in the digital humanities is now starting to rival that of big-science projects.

BIG SCIENCE

SLOAN DIGITAL SKY SURVEY 50 TERABYTES

The survey, begun in 1998 using a 2.5-metre telescope in New Mexico, has discovered nearly half-a-billion asteroids, stars, galaxies and quasars.

GENBANK 530 GIGABYTES

This database, which stores publicly available sequenced DNA, included 127 billion bases at the latest count.

BIG HUMANITIES

CULTUROMICS N-GRAMS VIEWER 300 GIGABYTES (English only)

The string of letters in this corpus of 5 million books is 1,000 times longer than the human genome.

YEAR OF SPEECH 1 TERABYTE

This database includes recordings from telephone conversations, broadcast news, talk shows and US Supreme Court arguments.

UNIVERSITY OF SOUTHERN CALIFORNIA SHOAH ARCHIVE 200 TERABYTES

This archive stores 52,000 videotaped interviews with Holocaust survivors from 56 countries.

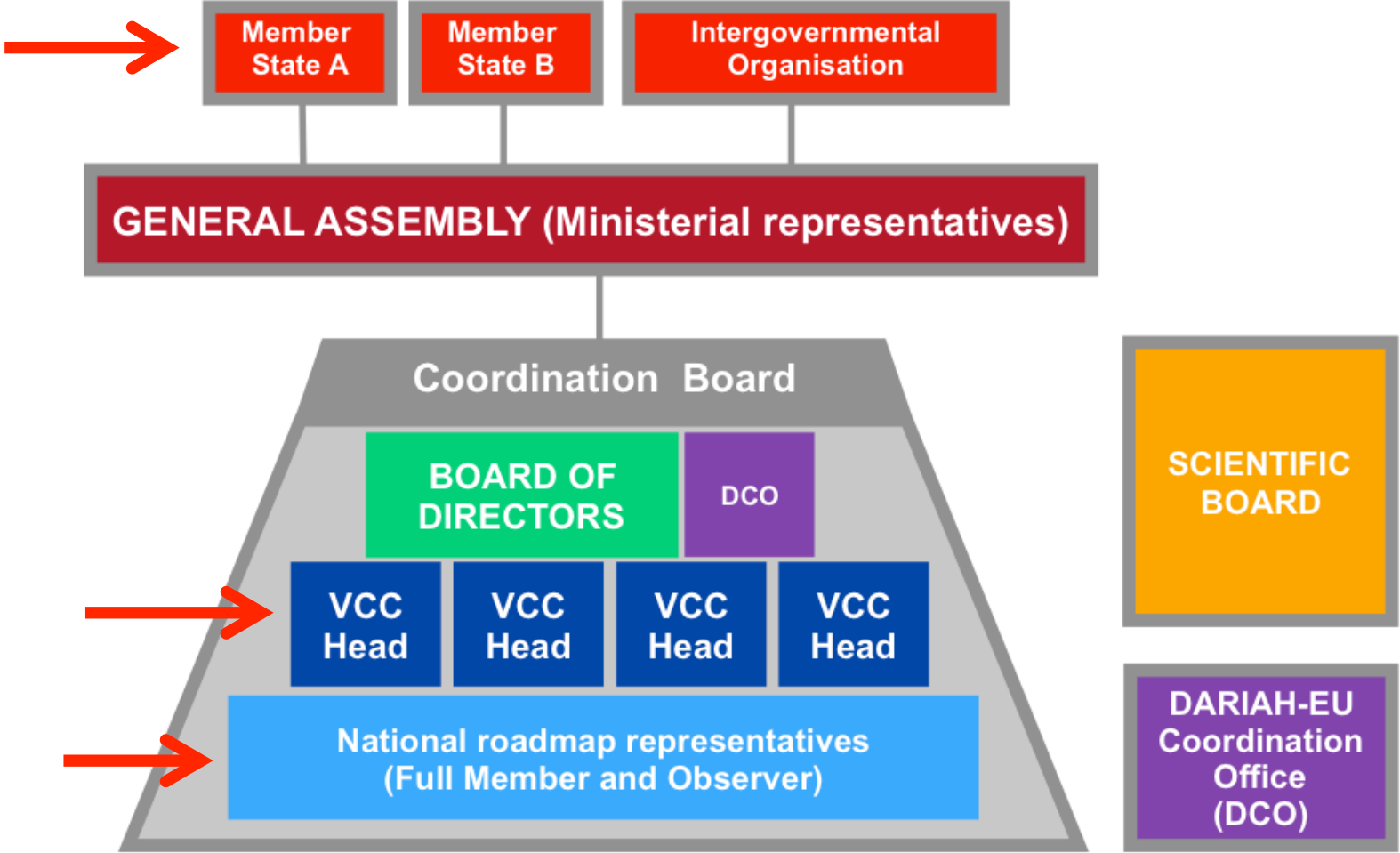
1 petabyte = 1,024 terabytes = 1,048,576 gigabytes

Big data

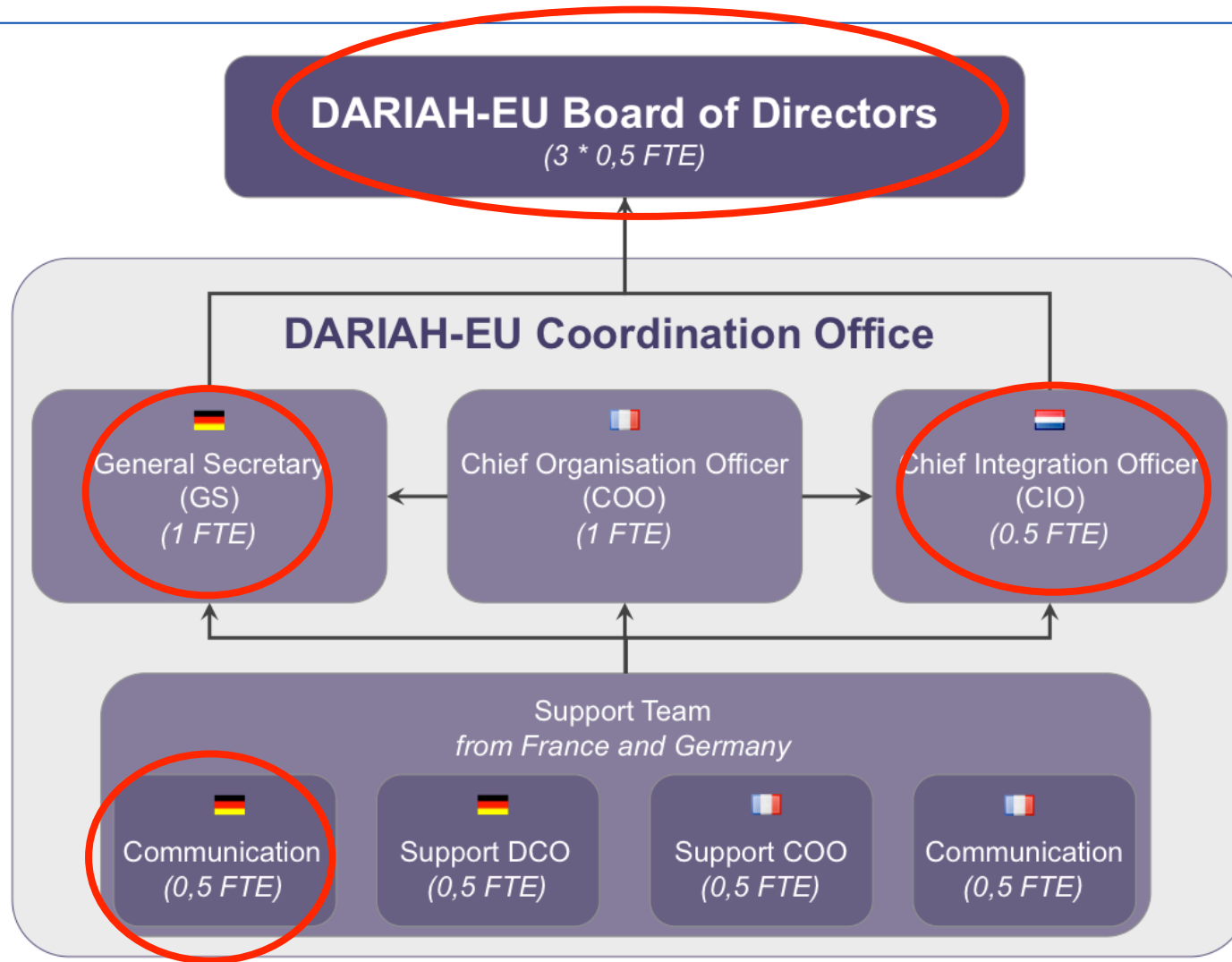
Arts and humanities are poised to achieve large benefit from e-science **methods and infrastructure as the human record becomes increasingly digitised and multimedia.** For example, [...] the **ability now to compute across enormous collections [...].** Copyright management for such work will remain an issue but for public domain and open-license materials, the field is now wide open.

Dan Atkins et al, RCUK e-Science Review

DARIAH governance structure



DARIAH-EU Coordination Office (DCO)



Collective Intelligence

Social Marketplace of Services



TEXTvire

TextGrid

building capacities

tech uptake + evolution

data curation



VCC2

VCC3

VCC4

VCC1

