



Technische Infrastruktur

DARIAH-DE Preconference Workshop

Dhd Konferenz, Passau, 26.3.2014

Tibor Kalman, GWDG, tibor.kalman@gwdg.de

(Beiträge von Thomas Kollatz, Harald Lordick, Peter Gietz)

Überblick

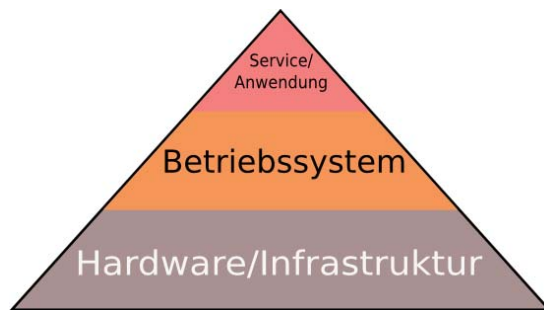
- Anwendungsfälle (Kollatz, Lordick)
- Technische Infrastruktur von DARIAH-DE
- Entstehung und Einbindung neuer Dienste
- Infrastruktur im (europäischen) Kontext
- Koordinierung, Entwicklung der Infrastruktur
- Sicherheit in DARIAH (Gietz)
- Datenhaltung in DARIAH (Gietz)
- Nachhaltigkeit, Geschäftsmodell (Gietz)

IT versus Fachwissenschaftler

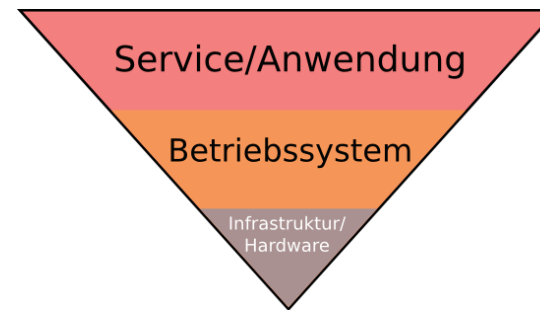
Rechenzentrumsbetrieb

Forschung

Interesse von RZ-Betrieb



Interesse von DARIAH



Überblick

- **Anwendungsfälle (Kollatz, Lordick)**
- Technische Infrastruktur von DARIAH-DE
- Entstehung und Einbindung neuer Dienste
- Die Infrastruktur im (europäischen) Kontext
- Sicherheit in DARIAH (Gietz)
- Datenhaltung in DARIAH (Gietz)
- Nachhaltigkeit, Geschäftsmodell (Gietz)

Anwendungsfälle

- Forschungsinfrastrukturen werden für WissenschaftlerInnen gebaut
 - Ergebnis DARIAH-DE 1: Dienste werden genutzt
- Anwendungsfall Nr.1:
 - Ein (assoziiertes) A+H Forschungsprojekt arbeitet auf der DARIAH-DE Infrastruktur (Kollatz)
- Anwendungsfall Nr.2:
 - Entstehung und Aufnahme eines neuen A+H Dienstes (Lordick)

Anwendungsfall Nr. 1

Anwendungsfall Nr.1:

- Ein (assoziiertes) A+H Forschungsprojekt arbeitet auf der DARIAH-DE Infrastruktur
- Relationen im Raum (Thomas Kollatz)

Anwendungsfall Nr. 2

Anwendungsfall Nr.2:

- Entstehung und Aufnahme eines neuen A+H Dienstes
- Orte der jüdischen Geschichte (Harald Lordick)

Überblick

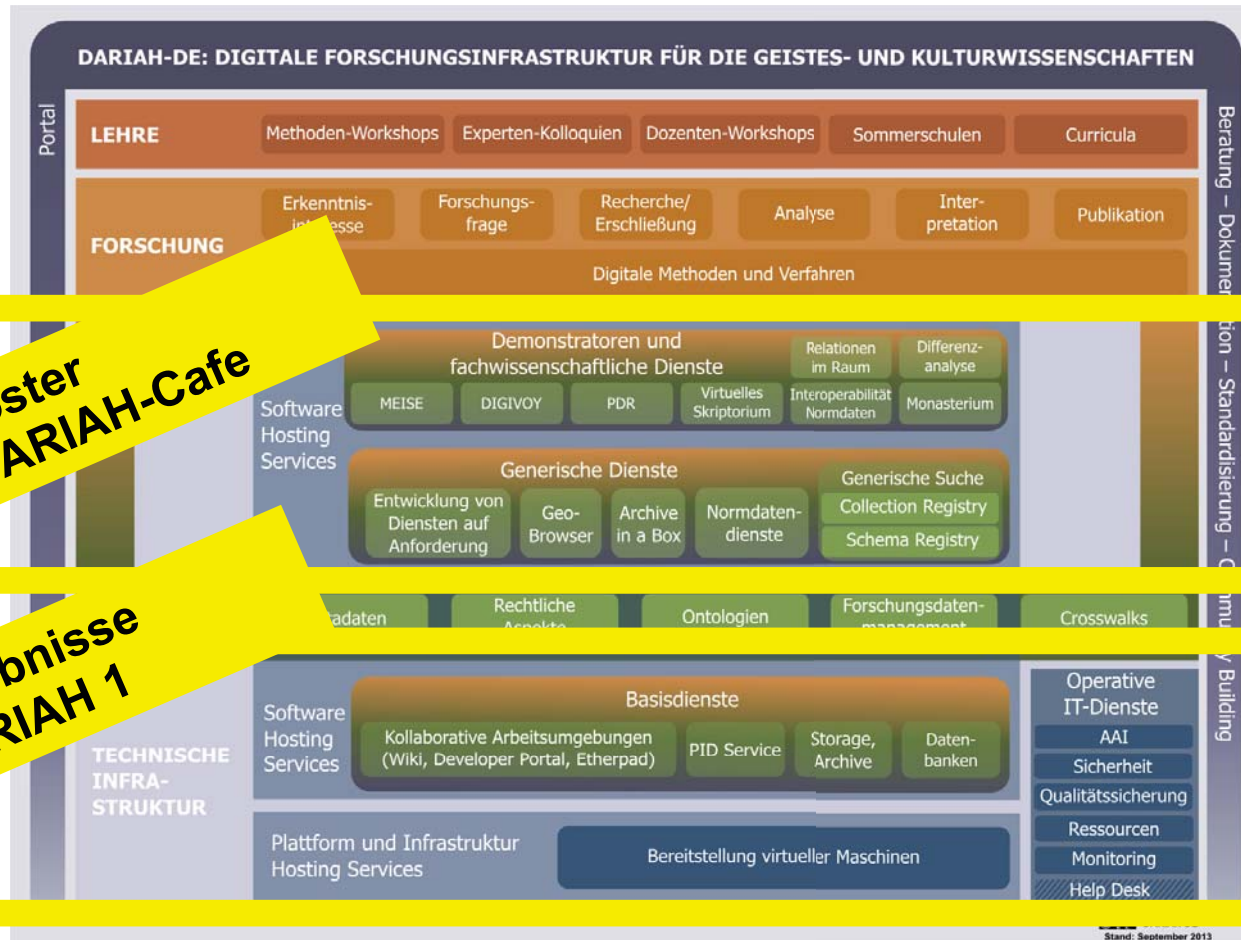
- Anwendungsfälle (Kollatz, Lordick)
- **Technische Infrastruktur von DARIAH-DE**
- Entstehung und Einbindung neuer Dienste
- Die Infrastruktur im (Europäischen) Kontext
- Koordinierung, Entwicklung der Infrastruktur
- Sicherheit in DARIAH (Gietz)
- Datenhaltung in DARIAH (Gietz)
- Nachhaltigkeit, Geschäftsmodell (Gietz)

Anwendungsfall Nr. 1

Anwendungsfall Nr.1: ein (assoziiertes) A+H Forschungsprojekt arbeitet auf der DARIAH-DE Infrastruktur

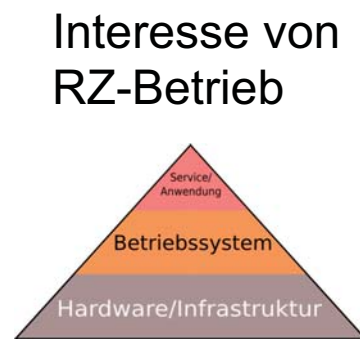
- Welche Dienste stehen für assoziierte Projekte zur Verfügung?
- Im Bereich der Technischen Infrastruktur werden
 - Software-,
 - Plattform- und
 - Infrastruktur-Hosting-Services
 - sowie operative IT-Diensteals **tragfähige** und flexibel **anpassbare** Lösungen entwickelt und bereitgestellt.

Fall Nr.1: Produktive Infrastruktur



Poster DARIAH-Cafe

Ergebnisse DARIAH 1



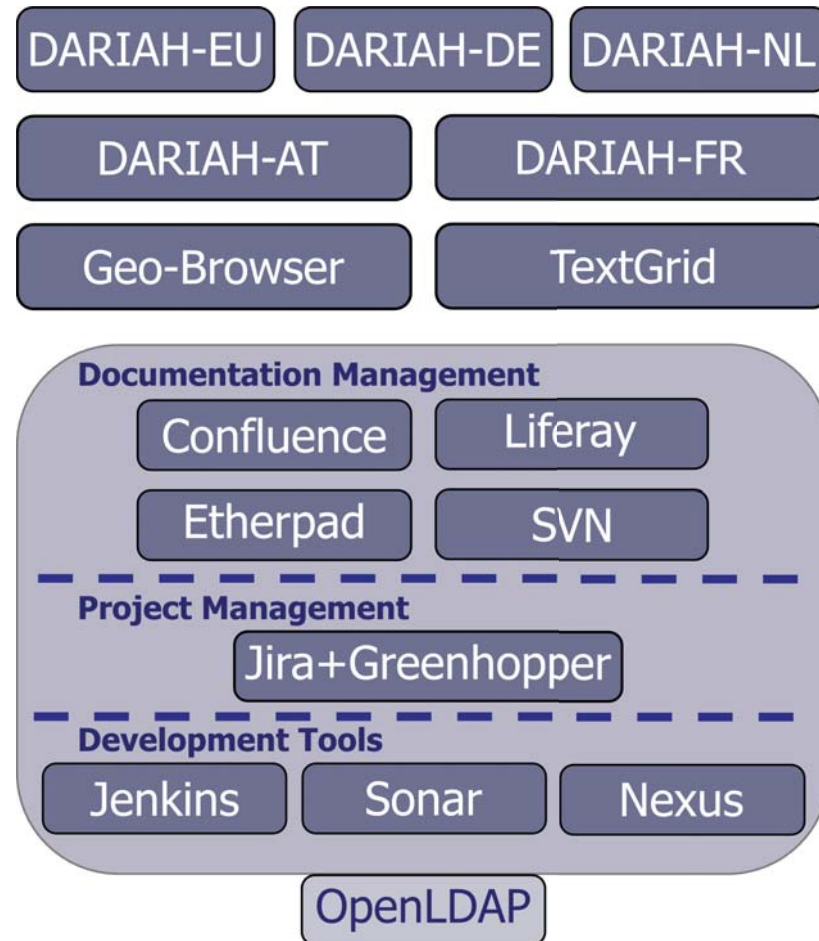
Produktive Infrastruktur

Ausgewählte Beispiele:

- Basisdienste
 - kollaborative Arbeitsumgebungen (Developer Portal etc.),
 - persistente Identifizierungsdienste für die Zitier- und Publikationsfähigkeit von Artikeln und Forschungsdaten/-sammlungen
 - Bereitstellung virtueller Maschinen.
- Operative IT Dienste
 - Als technische Qualitätssicherung findet u.a. ein Monitoring der DARIAH-DE und externen Dienste aus affilierten Projekten statt.
 - Aktuell erarbeitet DARIAH-DE Authentifizierungsdienste, die auch auf EU-Ebene eingesetzt werden sollen.

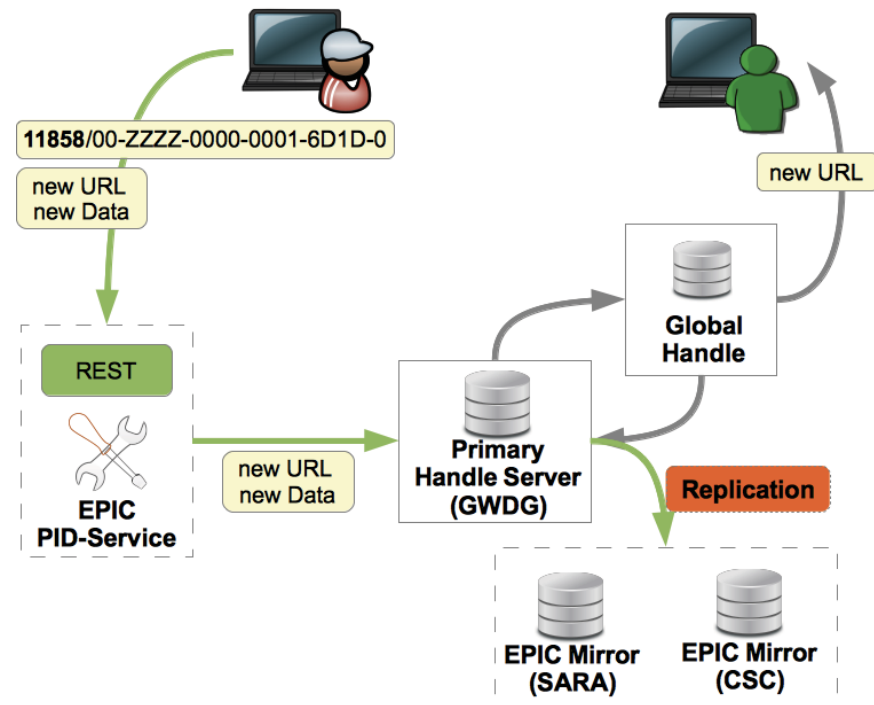
Kollaborative Arbeitsumgebung – Das Developer Portal

- Das DARIAH-DE Developer Portal bietet eine Reihe von Tools und Softwarekomponenten, die von EntwicklerInnen für die Realisierung von DH-Projekten genutzt werden können.
- Das Developer Portal basiert auf Standard-Entwicklertools.
- Diese Tools können von DARIAH-DE für Forschungs- und Entwicklungsprojekte der Digital Humanities "on demand" und flexibel eingerichtet werden.



PID-Service: Persistente Identifizierung

- PID: Persistent Identifier
- Dienst für persistente Identifizierung für die Zitier- und Publikationsfähigkeit von Artikeln und Forschungsdaten
- DARIAH setzt auf existierende Standards

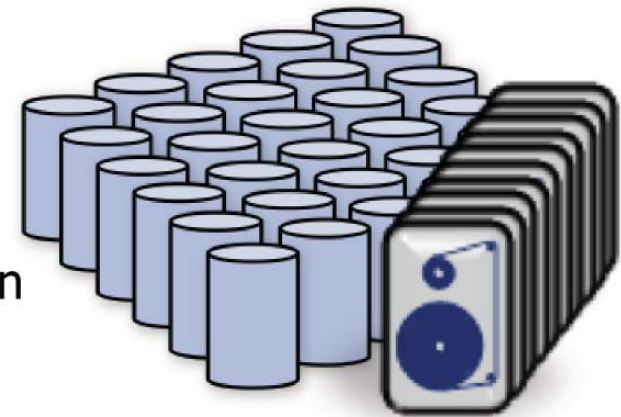


Tibor Kálmán, GWDG, 2012

Datenhaltung (Storage)

Die DARIAH-DE Infrastruktur bietet umfangreiche Speicherressourcen für geisteswissenschaftliche Forschungsdaten.

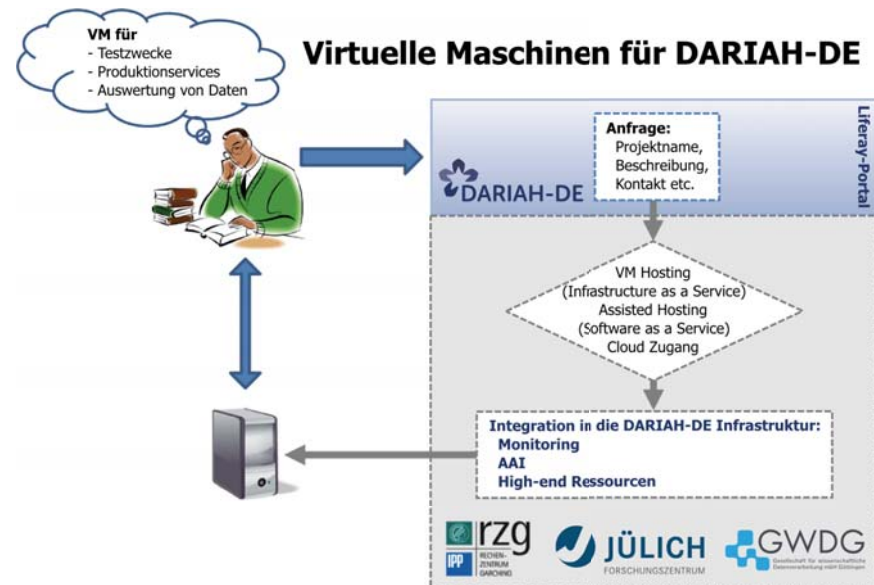
Durch die Einbindung verschiedener Rechenzentren kann eine vielfältige Auswahl angeboten werden.



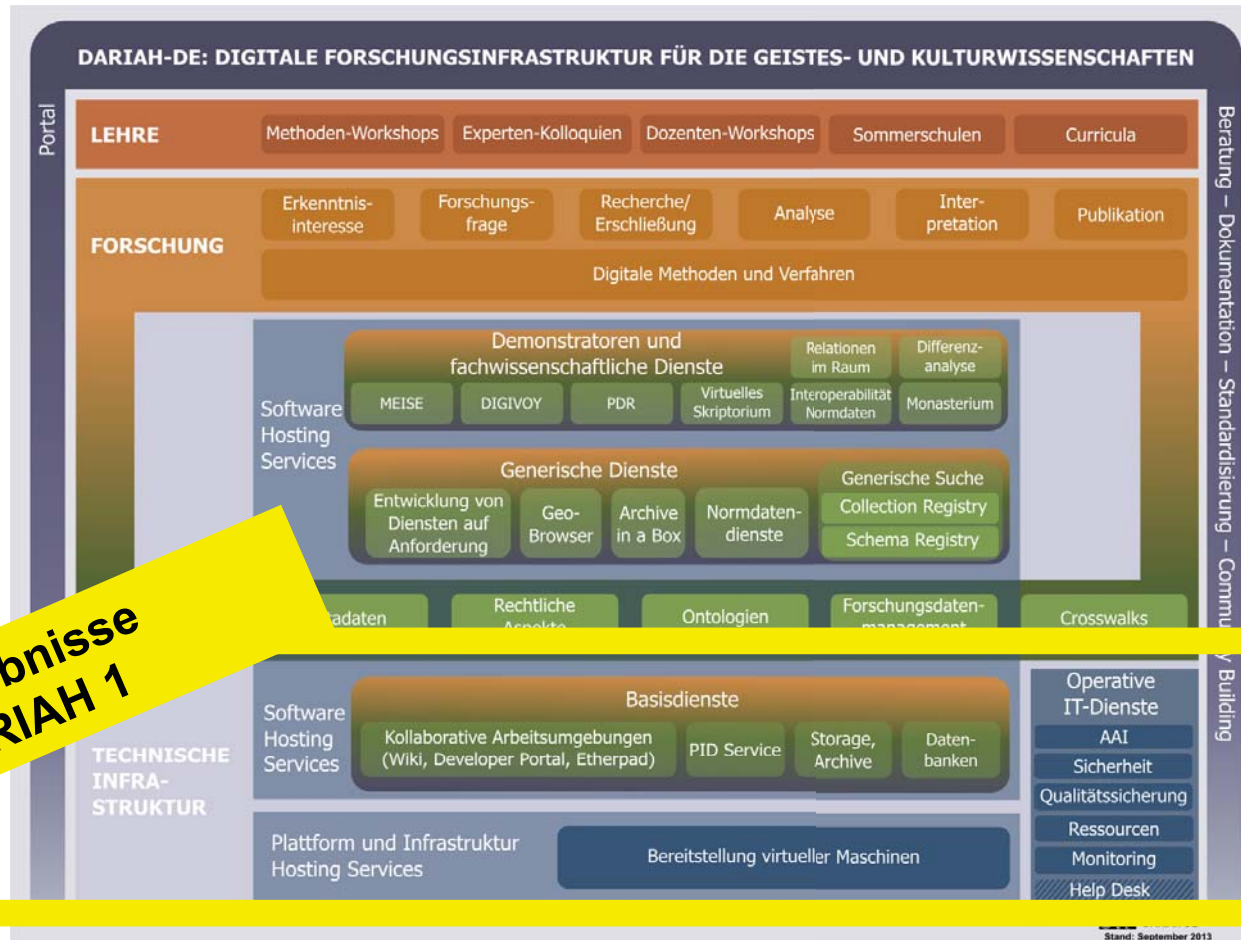
Darüber hinaus ist es möglich, eine virtuelle Maschine für die eigenen fachspezifischen Applikationen zu erhalten.

Bereitstellung virtueller Maschine (VM)

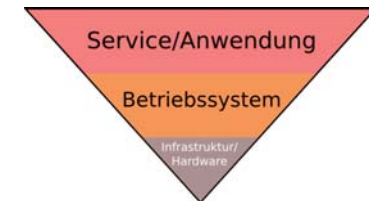
- Virtuelle Maschinen werden für:
 - Testzwecke
 - Produktionsservices
 - Auswertung von Daten
 - usw.benötigt.
- Ein Rechenzentrums-übergreifender Prozess wird in DARIAH 2 entwickelt.



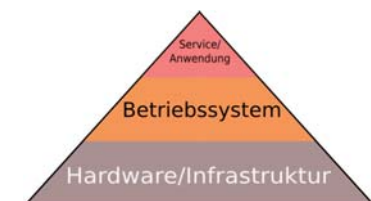
Operative IT-Dienste



Interesse von Forschung



Interesse von RZ-Betrieb

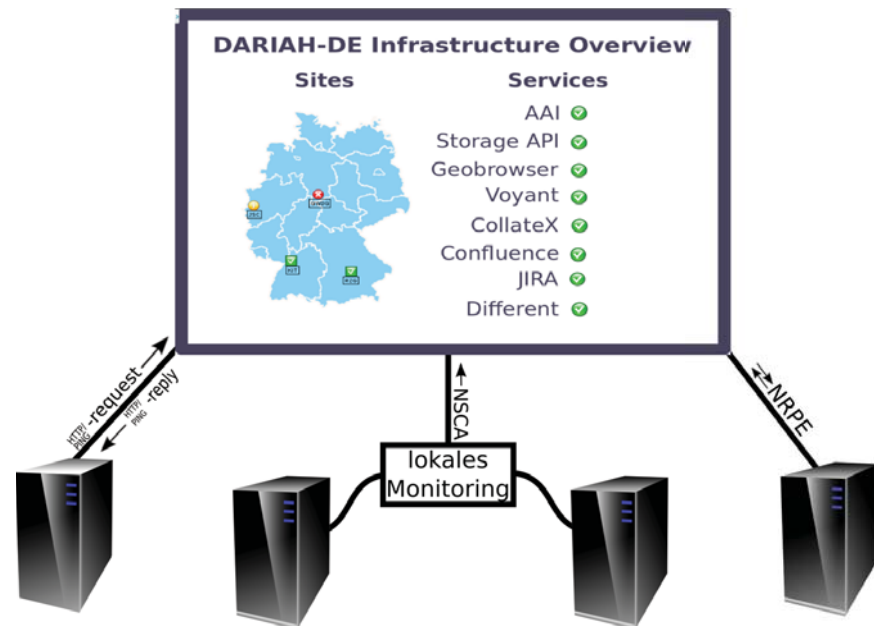


Operative IT-Dienste

- Operative IT-Dienste bilden die Grundlage von Entwicklungstätigkeiten in den Digital Humanities und den digital arbeitenden und forschenden Geistes- und Kulturwissenschaften.
- Auf ihrer Basis fußt der Aufbau von höherwertigen und fachwissenschaftlichen Diensten und Angeboten einer digitalen Forschungsinfrastruktur für die Geistes- und Kulturwissenschaften.
- Dazu zählen beispielsweise:
 - ein umfassendes und zentrales Monitoring von Diensten und Tools, aber auch eine Authentifizierungs- und Autorisierungsinfrastruktur, die den Zugriff und die Nutzung von Angeboten oder auch Forschungsdaten sicherstellt.

DARIAH Monitoring

- Ein zentraler Monitoring-Server ist in Betrieb
- Eigene Monitoring-Logik existiert für DARIAH
- Visualisierung (Zugriff über AAI)
- Die beteiligten Rechenzentren sind integriert
- Bietet funktionales Monitoring für AAI, StorageAPI
- Überwacht die Erreichbarkeit einer Vielzahl von Web-Services
- Monitoring der DARIAH-DE und externen Dienste aus affilierten Projekten statt



<https://dariah.zam.kfa-juelich.de/nagvis>

Überblick

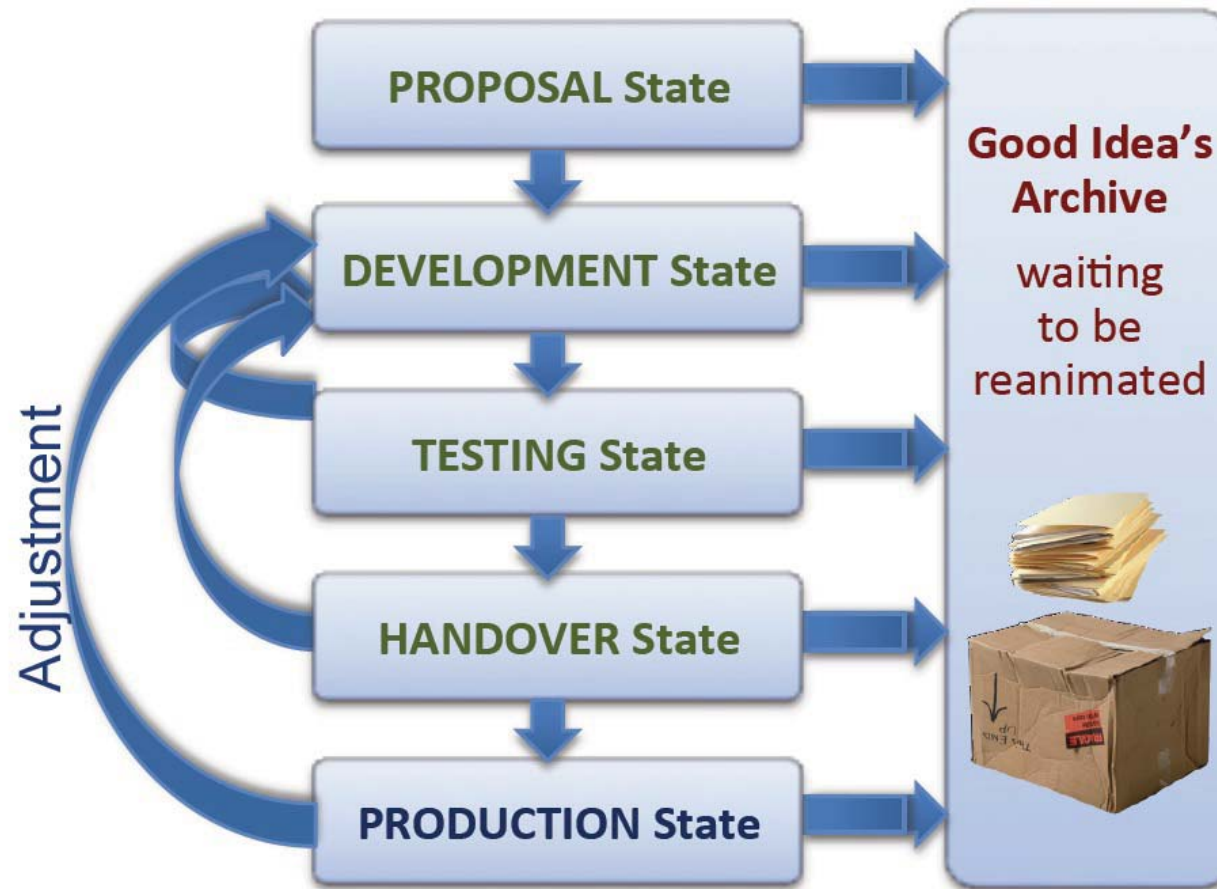
- Anwendungsfälle (Kollatz, Lordick)
- Technische Infrastruktur von DARIAH-DE
- **Entstehung und Einbindung neuer Dienste**
- Die Infrastruktur im (europäischen) Kontext
- Sicherheit in DARIAH (Gietz)
- Datenhaltung in DARIAH (Gietz)
- Nachhaltigkeit, Geschäftsmodell (Gietz)

Eigene Tools und Dienste über DARIAH-DE bereitstellen

- Eigene Tools und Dienste können über DARIAH-DE bereitgestellt werden.
- Dadurch ist:
 - die Nachhaltigkeit fachwissenschaftlicher Dienste gewährleistet
 - die Nachnutzung durch Dritte möglich.
- Schritte der Integration in die DARIAH-DE-Infrastruktur:
 - DARIAH Service Life-Cycle

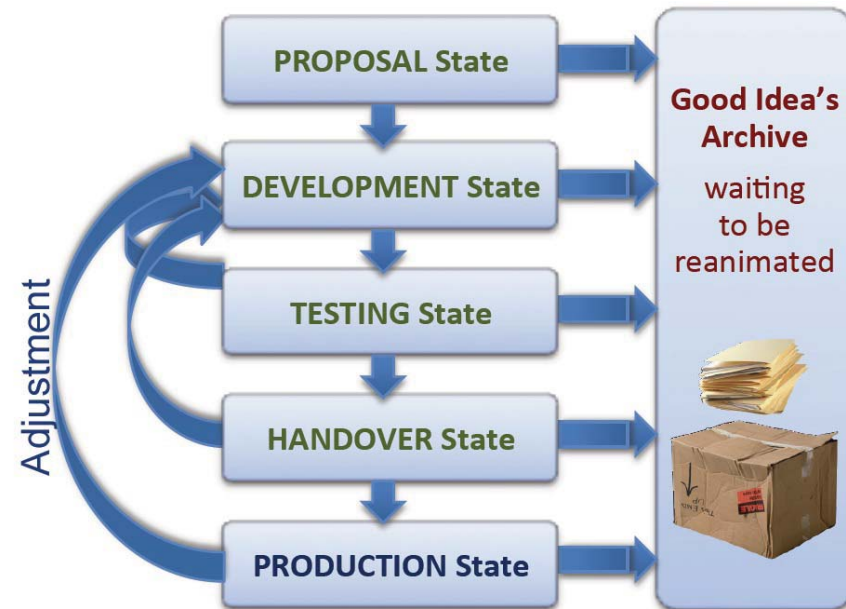
Fall Nr.2: Eigene Tools und Dienste

Schritte der Integration in die DARIAH-DE-Infrastruktur
(DARIAH-DE Service Life Cycle):



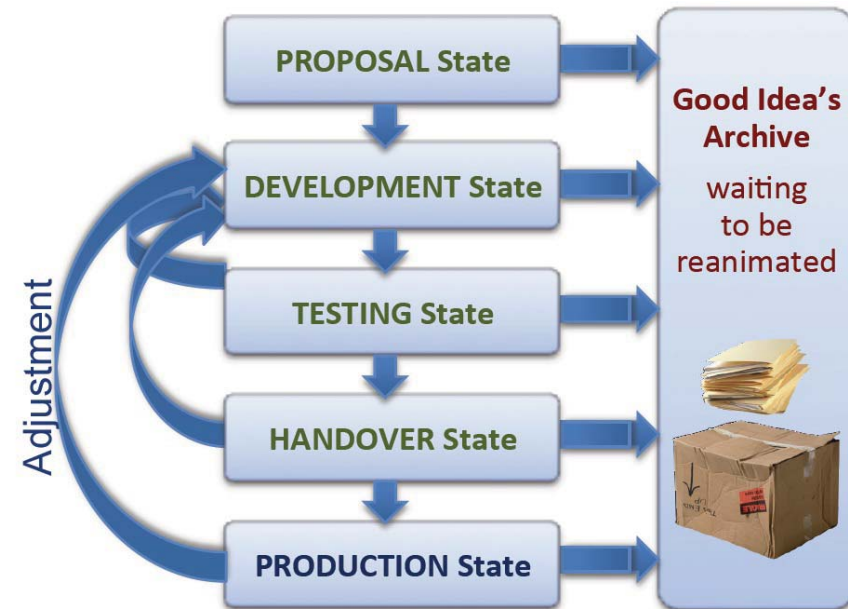
Fall Nr.2: Eigene Tools und Dienste

- In der ersten Phase (Proposal):
 - Mentoren werden benannt
 - Es wird evaluiert, ob der Vorschlag sich als DARIAH-Komponente integrieren lässt. (Begleitung durch den Integrationsprozess durch technische und fachwissenschaftliche Mentore:n.)
- Nach der positiven Entscheidung (Development):
 - Die Entwicklung beginnt.
 - Der betreffende Dienst und dessen Dokumentation weiterentwickelt und in die DARIAH-Infrastruktur eingebunden.



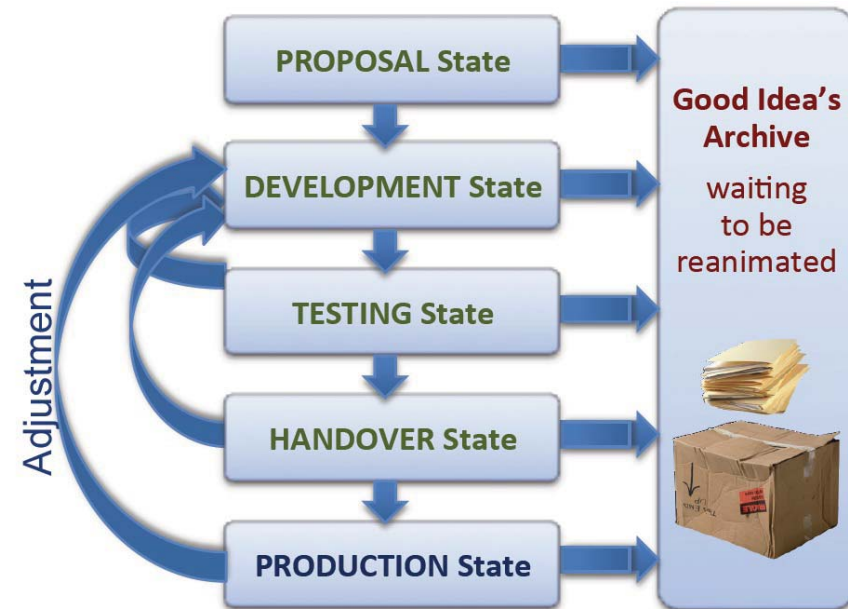
Fall Nr.2: Eigene Tools und Dienste

- In dem Testing State:
 - wird der Dienst von geisteswissenschaftlichen Nutzergruppen getestet und deren Rückmeldung eingeholt.
 - erfolgen ausführliche Tests, um eine hohe Qualität sicherzustellen.
 - Bei Bedarf wird der Dienst weiterentwickelt.
- Im Handover State:
 - Damit der entwickelte Dienst in die Produktionsphase übergehen kann,
 - müssen dessen Komponenten, z.B. Software, Daten und Dokumentation, dem DARIAH Service Hosting Team übergeben werden.



Fall Nr.2: Eigene Tools und Dienste

- Production State
 - In der letzten Phase sorgt DARIAH-DE für:
 - den nachhaltigen Betrieb,
 - die Pflege des Dienstes
 - sowie dessen Verbreitung.
 - Der Dienst steht allen Benutzern zur Verfügung, die die DARIAH Terms of Use akzeptieren.



Überblick

- Anwendungsfälle (Kollatz, Lordick)
- Die DARIAH-DE Technische Infrastruktur
- Entstehung und Einbindung neuer Dienste
- **Die Infrastruktur im (europäischen) Kontext**
- Sicherheit in DARIAH (Gietz)
- Datenhaltung in DARIAH (Gietz)
- Nachhaltigkeit, Geschäftsmodell (Gietz)

Was ist ESFRI?

- **E**uropean **S**trategy **F**orum on **R**esearch **I**nfrastructures
<http://ec.europa.eu/research/esfri/>
- ESFRI ermöglicht den Aufbau nachhaltiger, verteilter Forschungsinfrastrukturen (FI)
 - Für FIs sind strategische Richtlinien vorgesehen
- ESFRI stärkt der europäischen Wissenslandschaft
 - Durch die Identifizierung von Lücken in der Forschungslandschaft, die ESFRI Projekte schließen können
 - Synergien zw. ESFRI Projekten können genutzt werden
- Besonders wichtig: Roadmap 2010

3 Phasen der ESFRI-Projekte

ESFRI Phasen

Vorbereitungsphase (2–3 Jahre)

- Gefördert von der EU
- Implementierung des juristischen Rahmens (ERIC) und somit der Forschungsinfrastruktur

Aufbauphase (5–6 Jahre)

- Länderfinanziert

Betriebsphase (>15 Jahre)

- Nutzerfinanziert?

DARIAH-DE

DARIAH-DE 2008–2010

DARIAH-DE 2011–2015

DARIAH-DE 2016–...

Vorteile eines ERIC

- ERIC: European Research Infrastructure Consortium
 - ist eine juristische Körperschaft *sui generis*
 - leitet sich direkt von den ERIC-Regularien ab
 - ist weder eine Organisation nach internationalem Recht noch nach nationalem Recht
 - unterliegt der Kontrolle durch das EU-Parlament
- http://ec.europa.eu/research/infrastructures/index_en.cfm?pg=eric

DARIAH-ERIC

- bedeutet die Einrichtung einer neuen juristischen Entität nach europäischem Recht für DARIAH
- ist eine Alternative zu existierenden Rechtsformen
- wird als Rechtsperson von allen europäischen Mitgliedsstaaten anerkannt
- kann die Nachhaltigkeit von DARIAH garantieren
 - speziell auf pan-europäische FIs ausgerichtet
- flexibel an Erfordernisse von DARIAH anpassbar



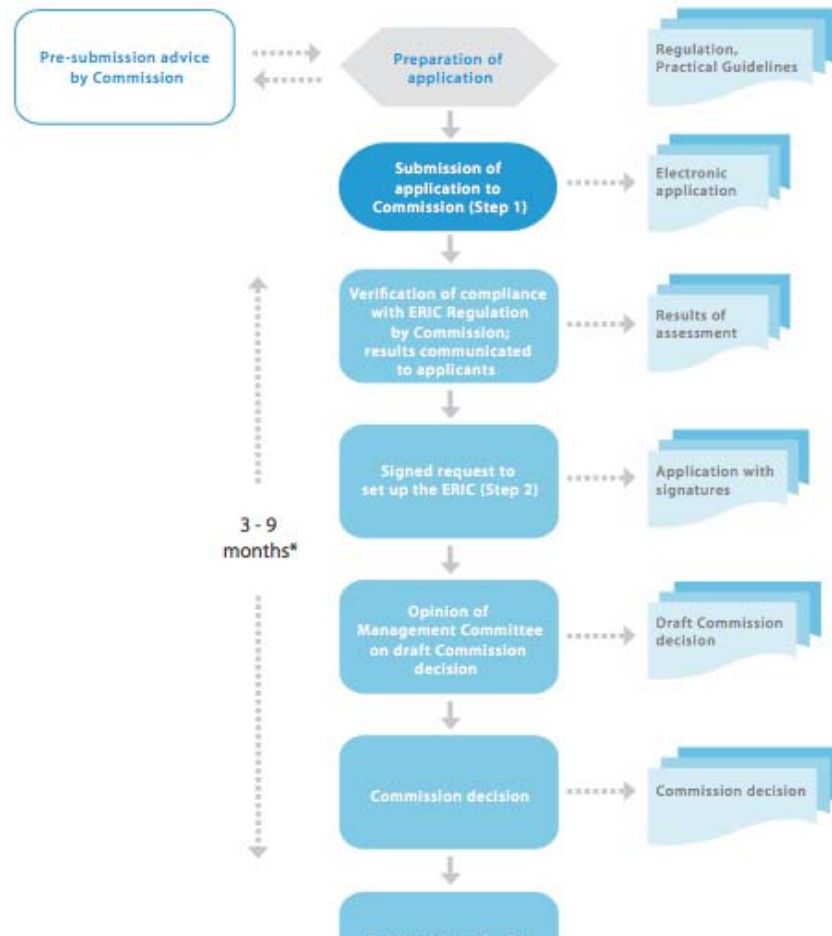
Mitglieder des DARIAH-ERIC

Founding Members	
Austria	
Denmark	
France	
Germany	
Greece	
Ireland	

Founding Members	
Italy	
Luxembourg	
Netherlands	
Serbia	
Slovenia	
...	

Candidate countries	
Belgium	
Croatia	
Lithuania	
Poland	
Portugal	
Spain	
Switzerland	
...	

Entstehung des DARIAH-ERIC



← Jetzt

← Mitte/Herbst 2014

Zusammenarbeit mit DARIAH assoziierten Projekten

- Assoziierte Projekte nutzen die DARIAH Infrastruktur:
 - ARIADNE (Advanced Research Infrastructure for Archaeological Dataset Networking in Europe)
 - CENDARI (Collaborative European Digital/Archival Infrastructure),
 - DASISH (Data Service Infrastructure for the Social Sciences and Humanities),
 - DiXiT (Digital Scholarly Editions Initial Training Network)
 - EHRI (European Holocaust Research Infrastructure),
 - NeDiMAH (Network for Digital Methods in the Arts and Humanities)
- Kollaboration mit FI Projekte
 - CLARIN (Common Language Resources and Technology Infrastructure)
 - TAB, Technische Infrastruktur, Rechenzentrumspartner

Überblick

- Anwendungsfälle (Kollatz, Lordick)
- Technische Infrastruktur von DARIAH-DE
- Entstehung und Einbindung neuer Dienste
- Die Infrastruktur im (europäischen) Kontext
- **Koordinierung, Entwicklung der Infrastruktur**
- Sicherheit in DARIAH (Gietz)
- Datenhaltung in DARIAH (Gietz)
- Nachhaltigkeit, Geschäftsmodell (Gietz)

Koordination, Weiterentwicklung der technischen Infrastruktur

- Wie wird die technische Infrastruktur koordiniert?
- Wie werden die Dienste weiterentwickelt und die technische Infrastruktur aufgebaut?

DARIAH-DE I.

Arbeitspaket 1 – e-Infrastruktur

Leitung: Staats- und Universitätsbibliothek Göttingen (SUB)

- Aufbau einer digitalen Informationsinfrastruktur zur Vernetzung fachwissenschaftlich relevanter Daten und Dienste unter Berücksichtigung generischer und fachspezifischer Standards.

Arbeitspaket 2 – Forschung und Lehre

Leitung: Universität Würzburg

- Stärkung und Zusammenführung der nationalen Entwicklungen im Feld der Digital Humanities sowie Einbettung in den europäischen und internationalen Kontext.

Arbeitspaket 3 – Forschungsdaten

Leitung: Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG)

- Forschungsdatenmanagement, digitale Langzeitarchivierung und Vernetzung mit anderen Archiven, Portalen etc.

Arbeitspaket 4 – DARIAH-DE Konsortium Management

Leitung: Staats- und Universitätsbibliothek Göttingen (SUB)

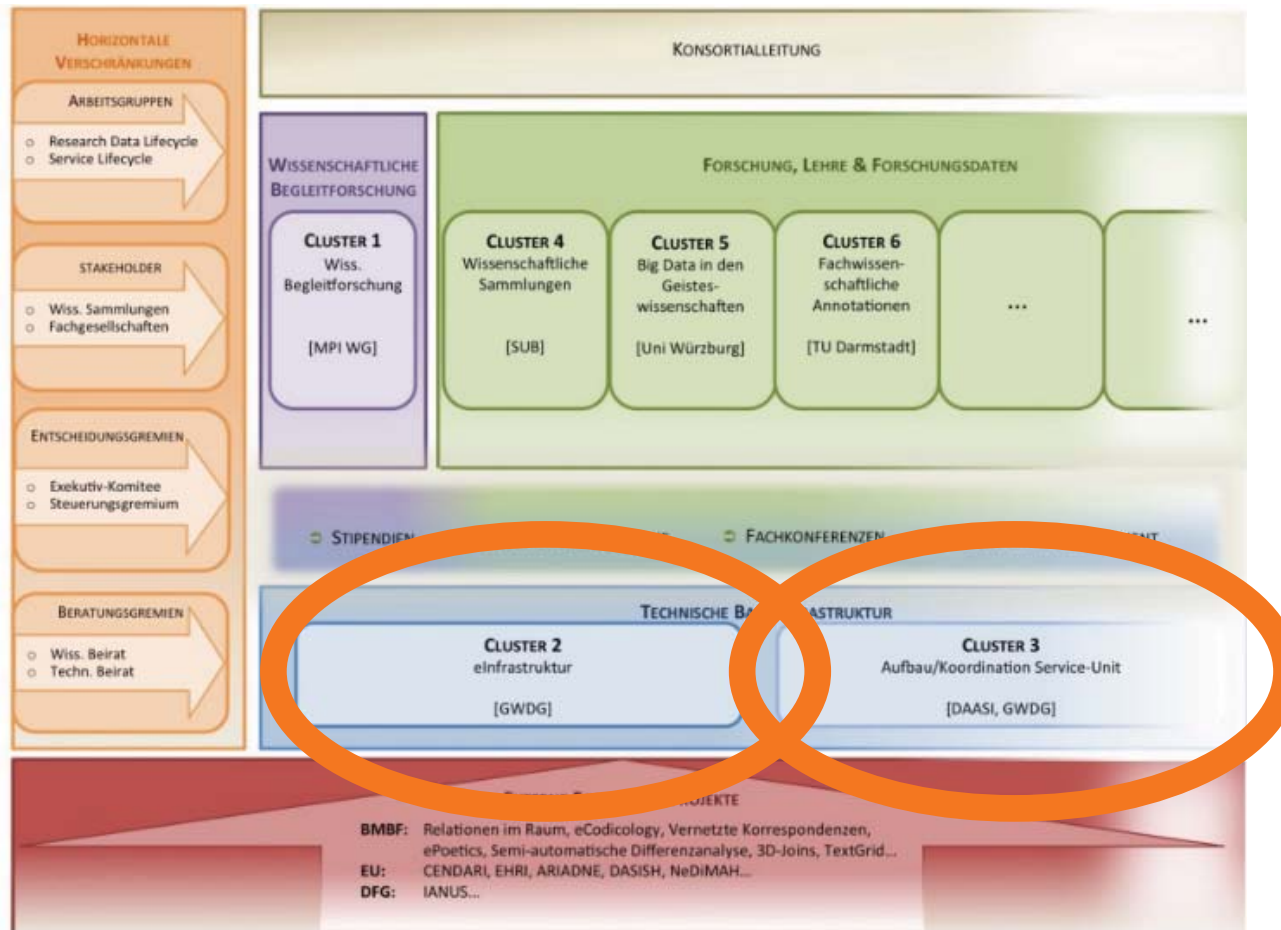
- Administrative Projektleitung und Moderation des Gesamtprojekts DARIAH-DE

Arbeitspaket 5 – Verantwortung für DARIAH-EU

Leitung: Göttingen Centre for Digital Humanities (GCDH)

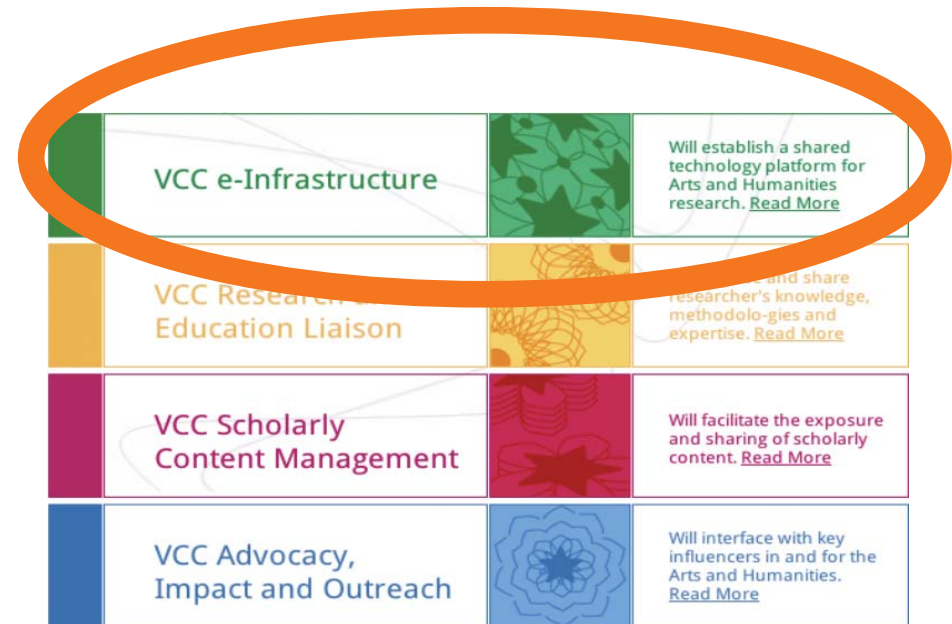
- Bündelung der Arbeiten auf DARIAH-EU-Ebene im Coordination Office (DCO) und im VCC 4.

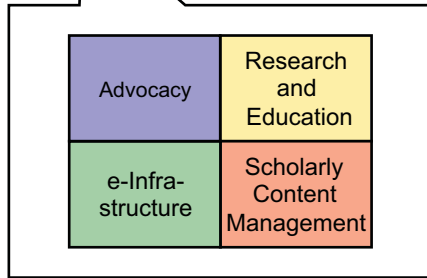
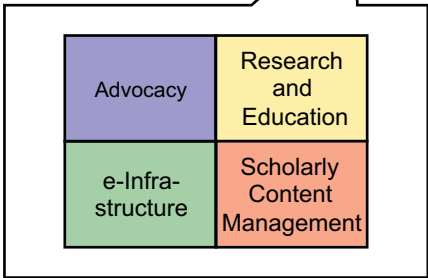
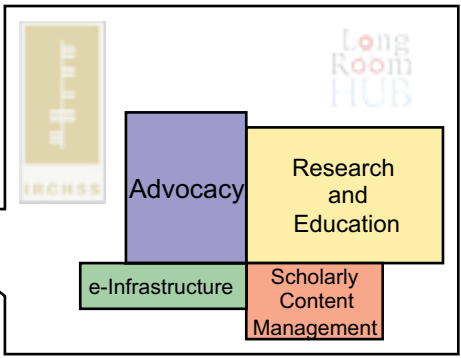
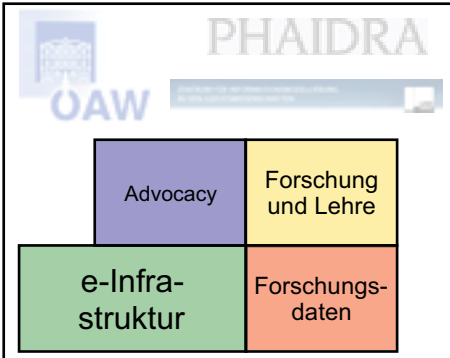
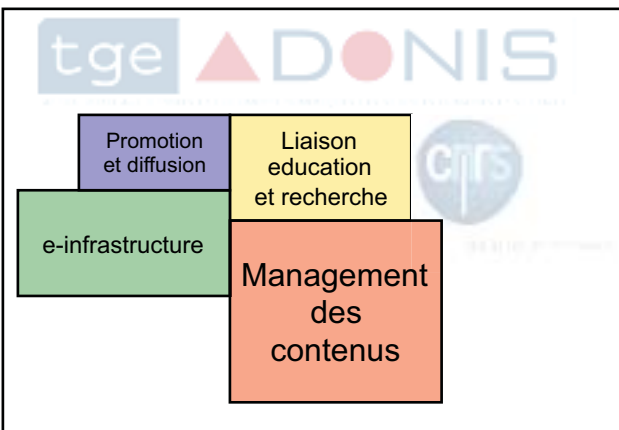
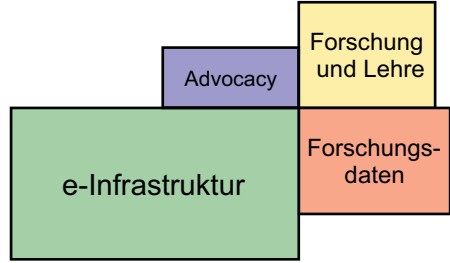
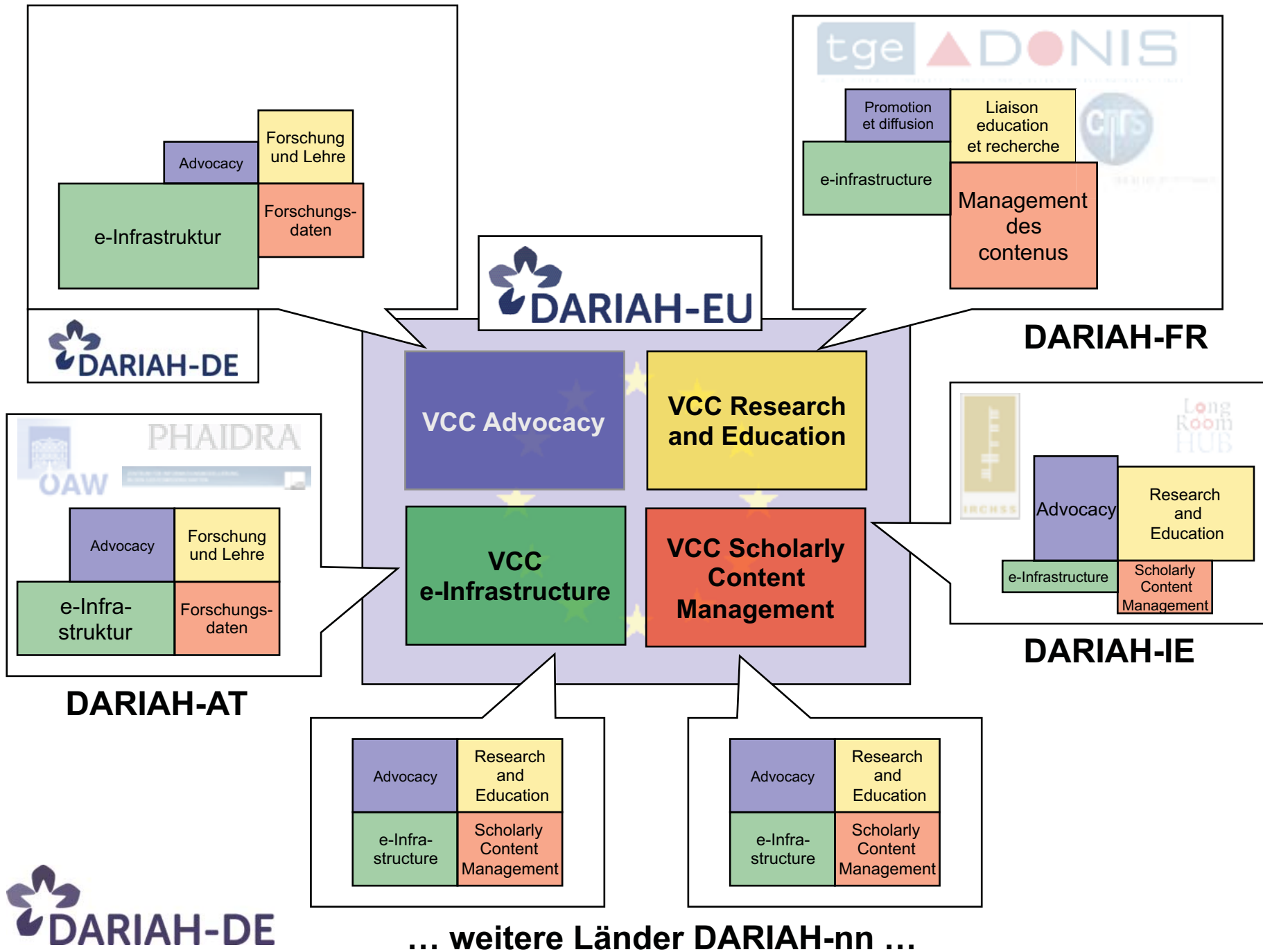
DARIAH-DE II.



DARIAH Virtual Competency Centers

- DARIAH-EU ist in VCCs (Virtual Competency Centres) organisiert
- Ein VCC vereint die Kompetenzen mehrerer Partnerinstitutionen
- Eine Partnerinstitution kann zu mehreren VCCs beitragen
- Bis zu zwei Partner können ein VCC leiten
- VCC1 e-Infrastructure [Germany / Austria]
 - „... to establish a shared technology platform for A+H research“





VCC 1: e-Infrastructure

- 1.1 Management+Coordination
- 1.2 A+H Infrastructure Services
- 1.3 Reference Software Packages
- 1.4 Preservation Infrastructure
- 1.5 Data federation and interoperability
- 1.6 Developer Community
- 1.7 A+H Service Environment
- 1.8 A+H Research environment demonstrators

... to establish a shared technology platform for A+H research ...



Vielen Dank!