

Agenda

Ort: SUB Göttingen
Zentralbibliothek
Platz der Göttingen Sieben 1
Großer Seminarraum (2. Etage) und Großes Konferenzzimmer (3. Etage)
37073 Göttingen
Termin: Freitag, 14. September 2012, 11:00-16:00 Uhr

Freitag, 14. September 2012

<i>Uhrzeit</i>	<i>Thema</i>
10:30-11:00	Registrierung (Großer Seminarraum, 2. Etage)
11:00-12:30	<ol style="list-style-type: none"> 1. Begrüßung und Vorstellungsrunde (Lossau, 15min) 2. Was ist DARIAH-EU (Blanke, 15min) Was ist DARIAH-DE (Neuroth, 15min) 3. Fachwissenschaftliche Perspektiven auf digitale Forschungsinfrastrukturen: Das Beispiel DARIAH-DE (Rapp, 30 min) 4. Architektur der DARIAH-DE-Forschungsinfrastruktur (Stotzka, 15 min)
12:30-13:30	Mittagspause (Großes Konferenzzimmer, 3. Etage)
13:30-15:00	<ol style="list-style-type: none"> 5. Einführung in die DARIAH-DE-Demonstratoren (Stotzka, 15min) 6. Parallele Vorstellung von 3 Anwendungsszenarien: DARIAH-DE-Demonstratoren (75 min) <ul style="list-style-type: none"> ○ Mit den Voyant Tools Textgrids Digitale Bibliothek entdecken (Schöch) ○ Der Geobrowser – ein Baustein der DARIAH-Infrastruktur (Kollatz) ○ Die Apokalypse in der Langzeitarchivierung (Tonne)
15:00- 16:00	<ol style="list-style-type: none"> 7. Diskussion (60 min) (Großer Seminarraum, 2. Etage) <ul style="list-style-type: none"> ○ Nachhaltigkeit von Forschungsinfrastrukturen ○ Schwerpunkte für die nächste Phase

Dokumentation im DARIAH-DE-Wiki: [Link](#)

Protokoll: Stefan Buddenbohm, Martina Kerzel

Teilnehmerliste

Wissenschaftlicher Beirat

Anwesend:

- Prof. Dr. Thomas **Bürger** (SLUB Dresden)
- Prof. Dr. Elisabeth **Burr** (Universität Leipzig, Französische / frankophone und italienische Sprachwissenschaft)
- Dr. Gabriele **Buschmeier** (Akademie der Wissenschaften und der Literatur Mainz)
- Prof. Dr. Günther **Görz** (Universität Erlangen, Institut für Informatik)
- Prof. Dr. Thomas **Gloning** (Universität Gießen, Zentrum für Medien und Interaktivität)
- Prof. Dr. Dr. Dr. h.c. Ernest W.B. **Hess-Lüttich** (Universität Bern, Institut für Germanistik)
- Prof. Dr. Reinhold **Kliegl** (Universität Potsdam, Department Psychologie)
- Dr. Peter **Leinen** (Universität Mannheim, Leitung Rechenzentrum)
- Prof. Dr. Wolfgang **Nagel** (TU Dresden, Direktion Zentrum für Informationsdienste und Hochleistungsrechnen)
- Ao.Univ.Prof. Dipl.-Ing. Dr. techn. Andreas **Rauber** (TU Wien, Information & Software Engineering Group)
- Frank **Scholze** (Karlsruher Institut für Technologie, Direktion Bibliothek)
- Ass.-Prof. Mag. Dr. Hubert **Stigler** (Universität Graz, Zentrum für Informationsmodellierung in den Geisteswissenschaften)
- Prof. Dr. Angelika **Storrer** (Universität Dortmund, Lehrstuhl für Linguistik der deutschen Sprache und Sprachdidaktik)
- Rüdiger **Eichel** (Ministerium für Wissenschaft und Kultur Niedersachsen)
- Dr. Markus **Klein** (Ministerium für Wissenschaft, Forschung und Kunst Baden-Württemberg)
- Hans **Nerlich** (PT-DLR-Geistes-, Sozial- und Kulturwissenschaften für das Bundesministerium für Bildung und Forschung)

Entschuldigt:

- Prof. Dr. Christian **Bischof** (TU Darmstadt, Leitung Hochschulrechenzentrum)
- Prof. Dr. Peter **Matussek** (Universität Siegen, Medienwissenschaftliches Seminar)
- Prof. Dr. Jan Christoph **Meister** (Universität Hamburg, Institut für Germanistik II)
- Prof. Dr. Dr. h.c. Gernot **Wilhelm** (Universität Würzburg, Lehrstuhl für Altorientalistik)

DARIAH-EU

- Dr. Tobias **Blanke** PhD, King's College London (DARIAH-EU)

Aus dem DARIAH-DE-Konsortium

- Prof. Norbert **Lossau** (SUB)
- Stefan **Buddenbohm** (SUB)
- Prof. Rainhard **Förtsch** (DAI)
- Christiane **Fritze** (SUB)
- Peter **Gietz** (DAASI)
- Prof. Fotis **Jannidis** (UWÜ)
- Martina **Kerzel** (SUB)
- drs Thomas **Kollatz** (STI)
- Dr. Wolf-Hagen **Krauth** (BBAW)
- Daniel **Kurzawe** (GWDG)
- Prof. Gerhard **Lauer** (GCDH)
- Dr. Heike **Neuroth** (SUB)
- Prof. Andrea **Rapp** (TUD)

- Dr. des. Stefan **Schmunk** (IEG)
- Dr. Christof **Schöch** (UWÜ)
- Dr. Ulrich **Schwardmann** (GWDG)
- Dr. Rainer **Stotzka** (KIT)
- Prof. Manfred **Thaller** (HKI)
- Danah **Tonne** (KIT)

Protokoll

TOP 1: Begrüßung und Einführung (Lossau)

- Betonung des Vernetzungs- und Nachhaltigkeitsgedankens. Im DARIAH-DE-Kontext können dabei von Interesse sein: ESFRI-Projekte, Europeana, OpenAIRE. Insbesondere größeren Einrichtungen kommt vor diesem Hintergrund die Rolle von Motoren und Motivatoren zu.
- Vorführung des Filmes „Virtuelle Forschungswelten“ (<http://youtu.be/JRBYR90A45Q>)

TOP 2: Was ist DARIAH-EU? Vortrag von Tobias Blanke ([Link](#))

- Diskussion und Fragen
 - Anregung von Peter Gietz zum Schaubild: Die Kreisgröße der geförderten Projekte ins reale Verhältnis zu den Fördermitteln setzen
 - DARIAH-DE sollte nicht als Projekt, sondern als Institution gedacht werden, um so auch semantisch Nachhaltigkeit und Wirkung zu demonstrieren

TOP 3: Was ist DARIAH-DE? Vortrag von Heike Neuroth ([Link](#))

- Handout “Übersichtsgrafik” ([Link](#))
- Kurzzusammenfassung der Diskussion
 - **Wie stehen die eHumanities-Projekte zur DARIAH-DE-Infrastruktur?** Eine Nutzung wird gewünscht und vom BMBF unterstützt, aber nicht zur Bedingung gemacht. Auf der DARIAH-DE-Webseite soll die Verknüpfung mit den eHumanities-Projekten ergänzt werden.
 - Eine **Usability-Evaluierung** der Infrastruktur ist erst für die nächste Förderphase von DARIAH-DE vorgesehen. In der aktuellen Förderphase wird dessen ungeachtet aber dennoch auf das Thema geachtet.
 - **Wie und wo sollen die Infrastrukturen nach der Förderphase betrieben werden?** Eine Verkettung mit den Rechenzentren ist für den nachhaltigen Betrieb der Infrastruktur essenziell und im Projektplan vorgesehen. Es findet eine Zusammenarbeit mit denselben RZ wie in CLARIN statt. Zusätzlich ist das KIT involviert.
 - Frage der **langfristigen Commitments** („Start als Dreimaster und Landgang als Ruderboot“), d.h. das Auseinanderklaffen von Anspruch und Wirklichkeit bleibt ein ungelöstes Problem: eine langfristige Finanzierung fehlt und die Empfehlungen des Wissenschaftsrates sind Wasser auf diese Mühlen. Allerdings muss auch festgestellt werden, dass wir im Kontext von DARIAH-DE und DARIAH-EU ein bisher ungekanntes Commitment haben, d.h. für Betriebsphase zusammen mit der Construction Phase 15 Jahre (ESFRI). Entscheidend ist es, von der Projektförderung wegzugehen und in institutionelle Förderung überzugehen bzw. die sich verschiebende Verantwortung bedenken: dies gilt für Bibliotheken ebenso wie für

weitere Infrastruktureinrichtungen (Rechenzentren) und Universitäten; eine Mischung ist notwendig.

TOP 4: Fachwissenschaftliche Perspektiven auf digitale Forschungsinfrastrukturen.

Vortrag von Andrea Rapp ([Link](#))

- Die Unterschiede zwischen den Arbeitsmethoden und dem Umgang mit Inhalten bedingen Unterschiede in den Infrastrukturen (analog und digital). Eine neue **Ästhetisierung von Daten** ist z.B. eine Folge davon: Bsp. DFG-Viewer zur Darstellung von Texten.
 - Ein wichtiger Aspekt hierbei ist die Menge des noch nicht digitalisierten Materials. In Deutschland liegt die **„Digitalisierungsquote“** bei unbefriedigenden 4%, d.h. dass im Zweifelsfall 96% des Materials nicht direkt berücksichtigt werden können (bzw. nicht direkt über digitale Infrastrukturen verfügbar sind).
- Fragen und Diskussion
 - Wie hat sich in DARIAH-DE das traditionelle **Mehrwertversprechen** entwickelt? Findet eine kritische Überprüfung und Evaluation statt, ob DARIAH den Zielen gerecht wird?
 - Das Mehrwertversprechen ist nicht für alle Bereiche geprüft und auch noch nicht für alle eingelöst. Dies hängt mit „kritischer Masse“ zusammen, Bsp.: Computerlinguistik, quantitative Methoden. Zum ersten Mal sind Massenphänomene ein möglicher Forschungsgegenstand (z.B. am Verbum, diese Forschung ist ohne digitale Methoden nicht möglich). Die **Verbindung von quantitativen mit qualitativen Methoden** ist ein entscheidender Schritt und eine wichtige Vermittlungsaufgabe. Die Wissenschaft muss in verschiedenen Schritten an die Möglichkeiten der neuen Technologien herangeführt werden. Hierfür sind insbesondere die Demonstratoren als „didaktisches Instrument“ von Bedeutung.

TOP 5: Architektur der DARIAH-DE Forschungsinfrastruktur, Vortrag von Rainer Stotzka ([Link](#))

- Vorstellung des DARIAH-DE-Schaubildes nach Ebenen gestaffelt und Verdeutlichung der methodischen Möglichkeiten nach Zielgruppen/ Forschungsfrage/ Demonstrator oder AP in DARIAH-DE einzuführen.

TOP 6: Einführung in die DARIAH-DE-Demonstratoren, Vortrag von Rainer Stotzka **Demonstratoren in paralleler Vorstellung ([Link](#), Demonstratoren werden zusätzlich individuell verlinkt)**

- **1.) „Mit den Voyant-Tools TextGrids Digitale Bibliothek entdecken“** (Kümmerer: Christof Schöch) ([Link](#))
 - Der Demonstrator verbindet die Digitale Bibliothek, die im TextGrid-Repository verfügbar ist, mit Voyant, einem Online-Tool zur Exploration und Analyse von Texten und Textsammlungen.
 - Anmerkungen/ Anregungen
 - Es wurde deutlich, wie Voyant den Texten des TGR Sichtbarkeit und Nutzbarkeit gibt; wie TextGrid Voyant leicht zugängliche, interessante Daten liefert; und wie der Demonstrator Textgrids TEI-Auszeichnung für verbesserte Recherchen in Voyant nutzt.

- Filteroptionen für TEI-Tags müssen erweitert werden
 - Verbesserung der Benutzerführung, bspw. visuelle Sequenzierung der Nutzerschritte
 - weitergehende Präprozessierung der Texte, mit Lemmatisierung und semantischer Kategoriensuche anzudenken
 - Bedürfnisse der Wissenschaft: flexible Tools, keine festen Vorgaben
- **2.) „Der Geobrowser als Baustein der DARIAH-DE-Infrastruktur“** (Kümmerer: Thomas Kollatz) ([Link](#), im weiteren Verlauf der Seite auch Testdaten zugänglich)
 - Anmerkungen/ Anregungen:
 - automatische ID-Konkordanz bestehender Geolocation-Gazetter (TGN, Perseus, etc.)
 - Einbindung ins DARIAH-Portal (Sichtbarkeit)
 - From/to relation (Migration, Briefcorpora)
 - Publish-Button: Material erst dann dauerhaft im DARIAH-Storage archivieren
 - Einbindung von historischem Kartenmaterial
- **3.) „Die Apokalypse der Langzeitarchivierung“** (Kümmerer: Danah Tonne) ([Link 1](#), Zugang zu 2. mit DARIAH-Account möglich)
 - Anregungen
 - PID-Konzept bei den Geisteswissenschaftlern weitgehend bekannt
 - Nachfragen zur Granularität, zur Nachhaltigkeit des Resolvers und zum Verfahren bei Formatänderung im Zuge der Content Preservation
 - AAI als spannendes Thema
 - Ausbau der Dienste in den darüber liegenden Schichten wichtig
 - Demonstratoren zum Ausprobieren im Netz verfügbar machen
 - Redundanzen mit anderen (PID-)Anbietern beachten
- **TOP 7: Diskussion (zusammengefasst nach Themen)**
- **1. Nachhaltigkeit von Forschungsinfrastrukturen**
 - DARIAH-DE befindet sich als geisteswissenschaftliche Initiative im **Wettbewerb um Ressourcen** mit anderen Initiativen und „Versorgungseinrichtungen“, bspw. für Maschinenbau, Ingenieurwesen, Physik oder Biologie. Während in den vorgenannten Disziplinen die Versorgungslage teilweise durchaus befriedigend ist, ist sie in den Geisteswissenschaften und anderen Disziplinen häufig prekär. Eine wichtige Frage in diesem Kontext ist das **überregionale Angebot von Infrastruktur**. Dieses müsste eigentlich auch eine überregionale Finanzierung nach sich ziehen. Eine regionale Beschränkung von Infrastruktur entspricht nicht den Anforderungen der Wissenschaft.
 - Das **Langfristigkeitsversprechen** gegenüber der Wissenschaft bringt Schwierigkeiten mit sich, bezogen auf die Weiterentwicklung von Technologien, Interfaces, der Portierung von Systemen oder ressourcenaufwändige Upgrades. Als verbesserungswürdig wird hier die Arbeit der Rechenzentren gesehen, die mehr Ressourcen für Innovation bereitstellen sollen. Eine Erhaltung des Status Quo bzw. die gesicherte Bereitstellung von Basisdiensten allein, sind nicht ausreichend.

- Interessante, aber bisher noch unterrepräsentierte Partner wie **Museen und Sammlungen** sollten verstärkt in den Blick genommen werden. Dies entspricht auch den Empfehlungen des Wissenschaftsrates, **mehr Raum für Experimente** zu öffnen.
- **2. Kostenanalysen und Geschäftsmodelle**
 - Die Arbeit als Infrastrukturdienstleister in einem geisteswissenschaftlichen Kontext zeigt, dass in der Wissenschaft häufig keine genauen Vorstellungen über die Kosten von Infrastrukturen bestehen. Das bedeutet im Umkehrschluss auch, dass von dieser Seite Argumente und Unterstützung im Bemühen um Nachhaltigkeit (s.o.) fehlen. Es ist daher wichtig, die einzelnen Dienste entsprechend ihrem Entwicklungs- und Betriebsaufwand mit „**Preisschildern**“ zu versehen.
- **3. DARIAH-DE als Infrastrukturanbieter**
 - Wie positioniert sich DARIAH-DE zu Akteuren wie bspw. **Dropbox oder Google, die generische Dienste anbieten**? Google ist auf einer sehr elementaren Ebene mit seinen generischen Diensten fester Bestandteil der Wissenschaft. Dazu kann und sollte DARIAH-DE nicht in den Wettbewerb treten. Ein Vorteil von DARIAH-DE liegt darin, die wissenschaftliche Suche in Kollektionen zu ermöglichen, die häufig von Google nicht erfasst werden. Nachteilig für DARIAH-DE ist, dass es als nicht-kommerzieller Akteur nicht derart unbefangen und unabhängig wie Google mit der Einstellung von Diensten verfahren kann (Optimierungsauslese). Als **wissenschaftsnaher/ bibliothekarischer Anbieter** genießt DARIAH-DE einen Vertrauensvorschuss und macht längerfristige Versprechen, die eingehalten werden müssen. Anbieter wie Dropbox oder Google bieten im Gegensatz zu DARIAH-DE keine Dienste für die „last mile“ an, d.h. es steht nicht zu befürchten, dass bspw. Google die geisteswissenschaftliche Informatik erdrückt. DARIAH-DE sollte es sich nicht zur Aufgabe machen, generische Dienste anzubieten, da diese von kommerziellen Anbietern in der Regel besser bereitgestellt werden können.
 - Problematisch für Initiativen wie DARIAH-DE ist der recht **große Abstand vieler Geisteswissenschaftler zum Entwicklungs- und Infrastruktur-Know-How** (anders als bspw. in der Physik). Dies ist eine Vermittlungsaufgabe, die auch an die oben erfolgte Diskussion des Wettbewerbs um Ressourcen anschließt.
- **4. DARIAH-DE und die wissenschaftliche Community**
 - Die **Community-Frage** muss v.a. auch als **Kommunikationsaufgabe** verstanden werden. Die Chance von DARIAH-DE ist es, überzeugend die Verschränkung von hermeneutischen Methoden und anderen Arbeitstechniken für die Geisteswissenschaften zu demonstrieren und auch offene und kritische Diskussionen anzustoßen, die interdisziplinär leider noch zu selten sind. Um diesen Punkt zu adressieren, plant DARIAH-EU ein **living review-Instrument** zu etablieren, das in Form eines internationalen Journals die Sichtbarkeit und Evaluation von Werkzeugen erhöhen soll.