



Digital Lab – Konzeptpapier

Manuel Burghardt und Eva Ommert, Oktober 2022

1 Einordnung und Selbstverständnis des Digital Lab

Die weitreichenden Effekte der Digitalisierung wirken sich auf sämtliche Bereiche des öffentlichen und privaten Lebens, die Politik, die Medien und im Zuge des Computational Turn (vgl. Berry, 2011) in zunehmendem Maße auch auf die Wissenschaft aus. In diesem Zusammenhang ergeben sich vor allem für die Geistes- und Sozialwissenschaften neue Herausforderungen bei der Deutung von gesellschaftlichen und kulturellen Phänomenen, die eine digitalisierte Welt mit sich bringt. Mit den Digital Humanities (DH) wird seit einigen Jahren dieser Entwicklungen Rechnung getragen und das Methodeninventar der Geistes- und Sozialwissenschaften um neue digitale Tools und Datenkollektionen erweitert (vgl. Burghardt & Wolff, 2014). Um das Thema nachhaltig an Hochschulen zu verankern, bieten sich unterschiedliche institutionelle Strukturen an, bspw. die Etablierung von einschlägigen Studiengängen (Sahle, 2013) oder die Gründung von DH-Zentren (Burghardt & Wolff, 2015). In zunehmendem Maße werden seit einiger Zeit auch sogenannte DH-Labs umgesetzt (Earhart, 2015; Lane, 2016)¹ und an einigen Stellen gar ein *laboratory turn* (Pawlicka-Deger, 2020) für die (Digital) Humanities ausgerufen.

Das Laborkonzept ist traditionell eng verknüpft mit natur- und lebenswissenschaftlichen Disziplinen. Die zentrale Rolle eines Labors in diesem Kontext fasst Hannaway (1986: 585) folgendermaßen zusammen:

[...] observation and manipulation of nature by means of specialized instruments, techniques, and apparatuses that require manual skills as well as conceptual knowledge of their construction and deployment“

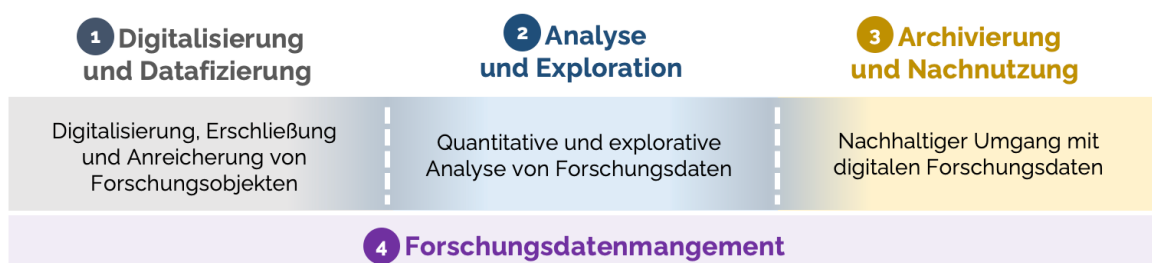
¹ Beispielfhaft zu erwähnen sind hier etwa das *Stanford Literary Lab* (Stanford), das *Digital Scholarship Lab* (University of Richmond), das *Humanities and Critical Code Studies Lab* (University of Southern California), das *DHLab* (Universität Basel), das *KNAW Digital Humanities Lab* (The Royal Netherlands Academy of Arts and Sciences) sowie das *HUMLab* (Umeå Universitet, Schweden).

Ersetzt man im obigen Zitat den Begriff *nature* durch *socio-cultural artifacts*, ergibt sich unmittelbar die grundlegende Agenda des Digital Lab: Hier werden geistes- und sozialwissenschaftliche Daten in Experimentalsettings so aufbereitet, transformiert und kodiert, dass sie der Untersuchung mithilfe von speziellen Instrumenten und Verfahren – also digitalen Tools und Methoden – zugänglich gemacht werden können. Der Einsatz solch neuartiger, digitaler Verfahren macht wiederum den Erwerb entsprechender Fertigkeiten zur Bedienung der digitalen Werkzeuge – also etwa grundlegende Kenntnisse im Bereich der Datenmodellierung, der statistischen Analyse und der Programmierung – erforderlich (*data literacy*). Gleichzeitig fließen bei der Entwicklung digitaler Analyseverfahren geistes- und sozialwissenschaftliches Domänenwissen mit ein. Die regelmäßige Reflexion des wechselseitigen Einflusses von digitalen Methoden und bestehenden Theorien und Epistemen eröffnet im Digital Lab zusätzlich die Möglichkeit erkenntnistheoretischer Experimente.

Im beschriebenen Experimentalkontext versteht sich das Digital Lab vor allem als Vermittler zwischen geistes- und sozialwissenschaftlichen Anwendungsbereichen und stärker informatischen und technischen Bereichen, die genuin mit der Entwicklung digitaler Tools und Methoden sowie sämtlichen Aspekten der Speicherung, Annotation und Analyse digitaler Daten beschäftigt sind. Durch das Zusammentreffen unterschiedlicher Disziplinen, etwa der Informatik und der Geschichtswissenschaft, in einem gemeinsamen Laborraum kann die von Snow (1959) beschriebene Kluft zwischen den „Two Cultures“ überwunden und stattdessen interdisziplinäre Forschungskooperation befördert werden².

2 Arbeitsfelder des Digital Lab

Das Digital Lab stellt sich den Herausforderungen geistes- und sozialwissenschaftlichen Forschens unter den Bedingungen einer fortschreitenden Digitalisierung von Forschungsdaten und wissenschaftlichen Workflows. Insgesamt ergeben sich drei größere Arbeitsfelder, die im Projektkontext oftmals fließend ineinander übergehen. Einen weiteren, übergreifenden Bereich stellt das Forschungsdatenmanagement (FDM) dar.



² Snow (1959) selbst schreibt hierzu „The clashing point of [...] two cultures [...] ought to produce creative chances“ – das Digital Lab ist dabei die ideale Umgebung um einen solchen „clash of cultures“ zu moderieren und die kreative Energie die daraus entsteht entsprechend zu kanalisieren.

Im Digital Lab können alle potenziell relevanten digitalen Forschungsaktivitäten entlang dieser Arbeitsfelder gebündelt und als interdisziplinäre Forschungsprojekte kanalisiert werden. Im Laufe der Zeit entstehen so spezialisierte „Instrumente, Techniken und Apparaturen“ (Hannaway, 1986) in allen genannten Arbeitsfeldern sowie *good practices* zu deren optimaler Bedienung, welche jeweils in verschiedenen Ausbildungs- und Qualifikationsprogrammen zurückfließen können. Dies gilt insbesondere für die Digital Humanities-Studiengänge (B. Sc. / M. Sc.) sowie auch für das Curriculum im Rahmen der strukturierten Promovierendenausbildung an der an ReCent-Globe angebotenen Graduate School Global and Area Studies (GSAGS). Auch das Zertifikatsprogramm für Postdocs der Graduiertenschule und die M. A.-Programme des Global and European Studies Institute (GESI) können hiervon profitieren.

2.1 Digitalisierung und Datafizierung

Mit dem zentralen Schritt der Digitalisierung erfolgt eine Übertragung von analogen Forschungsobjekten in digitale, maschinenlesbare Forschungsdaten. Der damit einhergehenden Datafizierung inhärent ist dabei, dass sozio-kulturelle Daten – die häufig durch ein hohes Maß an Ambiguität und Heterogenität gekennzeichnet sind – immer auch auf eine formale Art und Weise modelliert und repräsentiert werden müssen, um überhaupt erst computergestützten Analysen zugänglich gemacht zu werden. Dieser Übersetzungsschritt, der neben der Erschließung häufig auch eine Annotation und Anreicherung der Daten mit sich bringt, stellt eine zentrale Herausforderung dar, die in den Digital Humanities stark diskutiert wird.

Konkrete Themenbereiche:

- Digitale Aufnahme (Scan, Audio- und Videodigitalisierung)
- Semi-automatische Transkription (OCR, speech-to-text, image-to-text, etc.)
- Bereitstellung von verschiedenen Datenformaten und dazu passenden Viewern / Suchfunktionen
- Annotation und Anreicherung (Objekterkennung, Segmentierung, Feature-Extraktion für Ähnlichkeitssuche, etc.)
- Semi-automatische Katalogisierung und Verschlagwortung
- Clustering und Aggregation von Metadaten
- Verknüpfung von Datensätzen (untereinander / zu Normdaten)
- Qualitätskontrolle (Schemaebene, Plausibilitätschecks, Ausreißeranalysen...)

2.2 Analyse und Exploration

Nach der erfolgreichen Digitalisierung und Datafizierung von Forschungsobjekten ergeben sich in der Folge spannende Fragen zur Verwendung von digitalen Tools und Methoden, wenn es etwa darum geht, quantitative Muster in den vorher modellierten Daten zu finden und diese gegen qualitatives Domänenwissen von Expert*innen zu validieren. Zu erwarten sind in diesem Arbeitsfeld vor allem analytische Methoden mit quantitativ-empirischer Ausrichtung (etwa im Sinne der *cultural analytics*), aber auch innovative Ansätze im Spannungsfeld von *close* und *distant reading*, die häufig als

scalable reading bezeichnet werden. Bei Letzterem bieten sich vor allem Methoden der explorativen Datenanalyse sowie die Visual Analytics an.

Konkrete Themenbereiche:

- Überblick zu bestehenden digitalen Toolangeboten (*DH tool directories*)
- Anwendung und Entwicklung quantitativer Methoden im Sinne der *cultural analytics / computational social science / humanities data science*
- Analytische Ansätze im Sinne des *close / distant / scalable reading*
- Methodenadaption aus anderen Domänen (bspw. *cultural evolution*)
- Methodevaluierung (qualitativ und quantitativ)
- Methodenreflexion, epistemologische Implikationen und Theoriebildung
- Visualisierung / Visual Analytics
- Explorative Tools und Informationssysteme (Research Software Engineering)
- Immersive und interaktive Technologien für die Geistes- und Sozialwissenschaften (Humanist-Computer-Interaction, Immersive Humanities)

2.3 Archivierung und Nachnutzung

Für die Archivierung von Forschungsdaten steht für die am Zentrum angesiedelten Forschungsprojekte das Dataverse-basierte Repositorium „GlobeData“ zur Verfügung, welches Ende 2022 in den Regelbetrieb überführt wird. GlobeData steht darüber hinaus als fachliches Repositorium auch Forschenden anderer Institutionen und aus aller Welt offen. Gemäß des Ansatzes der „verteilten Archivierung“ können Metadaten zudem in fachlich ggf. einschlägigeren Repositorien abgebildet oder umgekehrt von dort via Harvesting in GlobeData eingespeist werden. Das Repositorium bietet zwischen der Veröffentlichung von Forschungsdaten und der reinen (geschlossenen) Archivierung zu Überprüfungs Zwecken auch abgestufte Varianten wie die gezielte Freigabe von Daten für bestimmte Personen oder etwa temporäre Embargos. Die koordinierte Kuration, Sammlung und Bereitstellung von Forschungsdaten via GlobeData bietet nicht zuletzt die Chance auf eine gezielte Nachnutzung von Daten in neuen Forschungsprojekten am Zentrum und dabei natürlich auch die Konzipierung neuer Projekte auf Basis vorhandener Bestände.

Konkrete Themenbereiche:

- Qualitätssicherung der Datensätze
- Sicherstellung einer nachvollziehbaren Dokumentation
- Auswahl geeigneter Dateiformate
- Beschreibung der Datensätze mit standardisierten Metadaten
- Auswahl einer geeigneten Lizenzierung

Die Aktivitäten in diesem Bereich sind eng mit dem umfassenderen Arbeitsbereich Forschungsdatenmanagements verknüpft, welcher im nächsten Abschnitt näher betrachtet wird.

2.4 Forschungsdatenmanagement

FDM erfährt seit einigen Jahren erhöhte Aufmerksamkeit in der akademischen Welt. Fächerübergreifend sind Wissenschaftler*innen aufgefordert, sich mit der Frage nach dem Verbleib ihrer Daten und Materialien nach Abschluss eines Projekts oder einer Publikation auseinanderzusetzen und sie ggf. für eine Überprüfung oder Nachnutzung zur Verfügung zu stellen. In den Geistes-, Kultur- und zumindest in den qualitativen Sozialwissenschaften gehört dies mehrheitlich nicht zur Fachtradition, die andere Formen der Überprüfung pflegt. Nicht zuletzt rechtliche Aspekte schränken zudem häufig eine Bereitstellung für andere ein. Grundsätzliches Ziel dieses Arbeitsbereichs ist daher die Unterstützung eines angemessenen Forschungsdatenmanagements der am Zentrum angesiedelten Verbünde und Projekte. Dabei werden stets Aufwand und Nutzen abgewägt und projektspezifische Lösungen gefunden. Besondere Relevanz besitzt FDM freilich für Forschungsprojekte, welche digitale und quantitative Methoden der Erschließung, Aufbereitung oder Analyse von Daten anwenden. Gleichwohl unterstützt das Digital Lab explizit auch Forschende mit anderem methodischen Hintergrund in Bezug auf FDM.

Konkrete Themenbereiche:

- Erstellung von Datenmanagementplänen in Forschungsprojekten
- Beratung zu verschiedenen FDM-Themen (z.B. Datendokumentation und -organisation, rechtliche Aspekte)
- Einrichtung technischer Dienste (z.B. Speicherung, Backup, Verschlüsselung)
- Archivierung und ggf. Veröffentlichung von Forschungsdaten- und Materialien im zentrumseigenen Repositorium GlobeData
- Beratung bzgl. FDM bei Antragsstellung

Nächster Schritt: Ein wesentliches Ziel des Digital Lab wird es in den nächsten Monaten sein, die konkreten Bedarfe entlang dieser Arbeitsfelder für laufende Forschungsprojekte oder künftige Projektideen zu konkretisieren und Tandems aus geistes- und sozialwissenschaftlich Forschenden und digitalen Expert*innen herzustellen. Dabei sollen u.a. Instrumente wie regelmäßige Projektsprechstunden und Methoden-Bazare erfolgen. Bei Erstgenanntem schildern Forschende aus den Geistes- und Sozialwissenschaften ihre Fragestellungen, Projektideen und Forschungsobjekte und erarbeiten mit Experten aus Informatik und Digital Humanities gemeinsame Konzepte für einen digital unterstützten Forschungsprozess. Bei den Methoden-Bazaren präsentieren umgekehrt die Akteur*innen aus stärker informatischen Kontexten ihre bestehenden Tools und digitalen Methoden und erschließen so ggf. neue Anwendungskontexte.

Ein Überblick zu relevanten Akteur*innen und Institutionen / Infrastrukturen erfolgt im nächsten Abschnitt.

3 Rahmenbedingungen und Einbindung des Digital Lab

3.1 ReCentGlobe

Die Rahmenbedingungen für ein Digital Lab unter den eingangs beschriebenen Gesichtspunkten sind an der Universität Leipzig geradezu ideal. So gibt es für die erwähnten geistes-, sozial- und raumwissenschaftlichen Anwendungsbereiche mit dem ReCentGlobe eine zentrale Einrichtung, welche vielfältige Disziplinen bündelt und unterschiedliche Forschungsfragen zu globalen Dynamiken in insgesamt vier unterschiedlichen Research Areas organisiert. Die Research Areas werden flankiert von einem Theory Forum sowie zwei weiteren Labs, die sich spezifisch den Aspekten Transfer und internationale Kooperation widmen. Wo immer sinnvoll, versucht das Digital Lab in enger Abstimmung mit diesen anderen ReCentGlobe-Infrastruktureinrichtungen zusammenzuarbeiten. Dabei bietet sich ein Austausch mit dem Theory Forum insbesondere bei der Verhandlung und Reflexion epistemologischer Konsequenzen, welche die Digitalisierung von Forschungsdaten und Analyseverfahren zwangsläufig mit sich bringt, an. Da das Digital Lab im Sinne eines *collaboratory* (vgl. Siemens & Siemens, 2012) u.a. auch den Aspekt der interdisziplinären Kollaboration unterstützen will, ist ein enger Austausch mit dem Cooperation Lab geplant, um insbesondere auch internationale Zusammenarbeit an digital unterstützten Forschungsvorhaben zu befördern. In Kooperation mit dem Transfer Lab sollen weiterhin Strategien entwickelt werden, um innovative digitale Methoden und vor allem Forschungsergebnisse (in Form von wiederverwendbaren Forschungsdaten) auch jenseits der Universität in einer breiteren Öffentlichkeit zu disseminieren.

3.2 FDHL

Wenngleich das Digital Lab institutionell primär im ReCentGlobe-Verbund angesiedelt ist, so ist es das ausdrückliche Ziel dieses Labs auch anderen geistes-, kultur-, und sozialwissenschaftlichen Forschenden offen zu stehen. Vor diesem Hintergrund ist eine enge Zusammenarbeit mit dem Forum für Digital Humanities Leipzig (FDHL)³ angedacht, welches seit 2018 die vielfältigen DH-Aktivitäten in Leipzig und der näheren Umgebung bündelt und regelmäßige Veranstaltungen in Form von Vorträgen und Workshops organisiert. Die Rolle des FDHL ist dabei primär die eines breit angelegten Netzwerks für alle Aspekte der DH, während das Digital Lab ein konkreter Arbeitsraum für die Konzeption und Umsetzung von digitalen Projektideen und methodischen Innovationen ist. Ein natürlicher Schnittbereich ergibt sich im Bereich der Vortrags- und Schulungsangebote, die künftig noch stärker aufeinander abgestimmt werden sollen um letztlich ein möglichst großes Publikum zu erreichen.

3.3 Partner*innen entlang der Arbeitsfelder

Für die Aspekte der **Digitalisierung und Datafizierung** (Arbeitsfeld 1) kann mit dem Universitätsarchiv und der Universitätsbibliothek – hier insbesondere der Abteilung

³ <https://fdhl.info/>

digitale Dienste – auf zwei etablierte Infrastrukturen der Universität Leipzig zurückgegriffen werden. Weiterhin sollen bestehende Kontakte mit der Sächsischen Akademie der Wissenschaften, die mit ihren zahlreichen Editionsprojekten hier umfangreiche Erfahrung vorweisen kann, weiter ausgebaut werden. In ihrer Funktion als Sächsische Landesbibliothek mit zahlreichen „Digitalen Sammlungen“ ist auch die SLUB ein naheliegender Partner in diesem Arbeitsfeld.

Im Bereich der computergestützten **Analyse und Exploration** sowie im eng damit verbundenen Bereich des Research Software Engineering (Arbeitsfeld 2) ist der Standort Leipzig ebenfalls hervorragend aufgestellt. Im Bereich der Infrastruktur (Server, Rechencluster, etc.) ist das Rechenzentrum der Universität Leipzig ein starker Partner. Weiterhin gibt es mit dem Institut für Informatik einen starken Partner für sämtliche Aspekte der Informatikforschung. Besonders hervorzuheben sind hier etwa die Professuren für „Bild- und Signalverarbeitung“, „Text Mining and Retrieval“ und „Computational Humanities“ sowie diverse KI-nahe Professuren, die u.a. am „Center for Scalable Data Analytics and Artificial Intelligence“ (ScaDS.AI)⁴ angesiedelt sind. Hervorzuheben ist weiterhin das Forschungslabor der Abteilung Computational Humanities, welches sich insbesondere mit dem Thema der immersiven und interaktiven Analyse und Exploration von Daten aus dem Anwendungsbereich der Spatial Humanities beschäftigt. Weiterhin beheimatet das Institut mit den Studiengängen „Informatik“ (B. Sc. / M. Sc.), „Digital Humanities (B. Sc. / M. Sc.) und „Data Science“ (M. Sc.) Ausbildungsprogramme auf Bachelor- und Masterniveau, die mit aktuell (Stand WiSe 2022/2023) etwa 1.500 Studierenden sehr stark nachgefragt sind. Damit ist ein wichtiger Grundstein für die Ausbildung von einschlägig qualifiziertem wissenschaftlichem Nachwuchs für künftige Forschungsprojekte innerhalb des Digital Labs gelegt. Niedrigschwellige Möglichkeiten der Kooperation mit Studierenden aus diesen Bereichen sind darüber hinaus im Rahmen von Praktikumsmodulen und Abschlussarbeiten gegeben.

Archivierung und Nachnutzung (Arbeitsfeld 3) und das damit eng verknüpfte **Forschungsdatenmanagement** (Arbeitsfeld 4) wird bisher vor allem durch das FDM-Pilotprojekt im SFB 1199 und der damit verbundenen Stelle bearbeitet (Ommert, 2022), wobei Erfahrungen und Workflows sukzessive auf das gesamte Zentrum ausgeweitet wurden/werden. Das FDM wird unterstützt von der AG Forschungsdaten der Universität Leipzig. Darin vertreten sind das Universitätsrechenzentrum, welches das ReCentGlobe mit technischen Diensten wie Speicherung, Backup und virtuellem Rechnen unterstützt, sowie die Universitätsbibliothek und das Dezernat 1 „Forschung und Transfer“, welche u.a. Beratung zu Rechtsfragen leisten. Ein weiterer starker Partner mit umfangreicher Projekterfahrung im Bereich FDM ist darüber hinaus die Sächsi-

⁴ <https://scads.ai/>

sche Akademie der Wissenschaften, hier insbesondere das „KompetenzwerkD: Sächsisches Forschungszentrum und Kompetenznetzwerk für Digitale Geisteswissenschaften und Kulturelles Erbe“⁵.

Literatur

- Berry, D. M. (2011). The computational turn: Thinking about the digital humanities. *Culture Machine*, 12, 1–22.
- Burghardt, M., & Wolff, C. (2014). Digital Humanities: Buzzword oder Strukturwandel in den Geisteswissenschaften? *Blick in Die Wissenschaft*, 23(29), 39–47.
- Burghardt, M., & Wolff, C. (2015). Zentren für Digital Humanities in Deutschland. *Information - Wissenschaft & Praxis*, 66(5–6), 313–326.
- Earhart, A. (2015). *The digital humanities as a laboratory*. MIT Press.
- Hannaway, O. (1986). Laboratory Design and the Aim of Science: Andreas Libavius versus Tycho Brahe. In *Isis – Journal of the History of Science Society* 77(4), 585-610.
- Lane, R. J. (2016). *The Big Humanities: Digital Humanities/Digital Laboratories*. Routledge.
- Langer, L., Burghardt, M., Borgards, R., Köhring, E. & Wirth, C. (to appear 2022). Digital Environmental Humanities – Zum Potential von „Computational and Literary Biodiversity Studies“ (CoLiBiS). Proceedings of the 8th annual conference of the DHd (Digital Humanities im deutschsprachigen Raum) Association.
- Ommert, E. (to appear 2022). Forschungsdatenmanagement im SFB 1199. Ein Erfahrungsbericht. *SFB 1199 Working Paper Series*.
- Pawlicka-Deger, U. (2020). The Laboratory Turn: Exploring Discourses, Landscapes, and Models of Humanities Labs. *DHQ: Digital Humanities Quarterly*, 14(3).
- Sahle, P. (2013). DH studieren! Auf dem Weg zu einem Kern- und Referenzcurriculum der Digital Humanities. In *DARIAH-D Working Papers*, 2013-1.
- Siemens, L. & Siemens, R. (2012). “Notes from the Collaboratory: An Informal Study of an Academic Lab in Transition”,
- Snow, C. P. (1959). *The Two Cultures*. London: Cambridge University Press.

⁵ <https://www.saw-leipzig.de/de/akademie-digital/akademie-digital/kompetenzwerk-d-saechsisches-forschungszentrum-und-kompetenznetzwerk-fuer-digitale-geisteswissenschaften-und-kulturelles-erbe>