

Harald Klinke (Hrsg.)

# #DigiCampus

Digitale Forschung und Lehre  
in den Geisteswissenschaften

readbox unipress · Open Publishing LMU

Harald Klinke (Hrsg.)

#DigiCampus

Digitale Forschung und Lehre in den Geisteswissenschaften

## **Open Publishing LMU**

Mit Open Publishing LMU unterstützt die Universitätsbibliothek der Ludwig-Maximilians-Universität München alle Wissenschaftlerinnen und Wissenschaftler der LMU dabei, ihre Forschungsergebnisse parallel gedruckt und digital zu veröffentlichen.



Universitätsbibliothek  
Ludwig-Maximilians-Universität München

# #DigiCampus

Digitale Forschung und Lehre  
in den Geisteswissenschaften

Herausgegeben von  
Harald Klink

Diese Publikation wurde maßgeblich finanziert durch das  
Department Kunstwissenschaften.

© für alle Texte bei den jeweiligen Autoren 2018  
Erstveröffentlichung 2018

**Bibliografische Information der Deutschen Nationalbibliothek**

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in  
der Deutschen Nationalbibliografie; detaillierte bibliografische  
Daten sind im Internet abrufbar über <http://dnb.dnb.de/>

Universitätsbibliothek der Ludwig-Maximilians-Universität  
Geschwister-Scholl-Platz 1  
80539 München

Herstellung über:  
readbox unipress  
in der readbox publishing GmbH  
Am Hawerkamp 31  
48155 Münster  
<http://unipress.readbox.net/>

Open-Access-Version dieser Publikation verfügbar unter:  
<http://nbn-resolving.de/urn:nbn:de:bvb:19-epub-41218-5>

978-3-95925-055-9 (Druckausgabe)  
978-3-95925-056-6 (elektronische Version)

# Inhaltsverzeichnis

- 1 **Vorwort: Die digitale Transformation in den Geisteswissenschaften**  
Harald Klinke
  
- 9 **Die Geisteswissenschaften und das Digitale**  
Ein Quantensprung oder *business as usual*?  
Hubertus Kohle
  
- 19 **Datenanalyse in der Digitalen Kunstgeschichte**  
Neue Methoden in Forschung und Lehre und der Einsatz des DHV*Lab* in der Lehre  
Harald Klinke
  
- 35 **DHLehre**  
Integration von Forschung und universitärem Unterricht  
Thomas Krefeld
  
- 47 **Das DHV*Lab* Analysecenter**  
Ein Tool zur datengestützten Forschung in den Geisteswissenschaften  
Stefanie Schneider
  
- 57 **Von der IT-Gruppe Geisteswissenschaften zu einem LMU Center for Digital Humanities?**  
Zu Entwicklung, Mehrwert und Nachhaltigkeit des Digitalen  
Christian Riepl
  
- 77 **Auf dem Weg zu einem DH-Curriculum**  
Digital Humanities in den Geschichts- und Kunstwissenschaften an der LMU München  
Julian Schulz

- 103 **Digital Humanities Sprachwissenschaften**  
Schwerpunkte bei der Entwicklung eines  
sprachwissenschaftlichen DH-Nebenfaches  
Markus Frank
- 115 **Digital Humanities an der Universität Regensburg**  
Geschichte – Projekte – Studiengang  
Manuel Burghardt, Daniel Isemann, Markus Kattenbeck, Bernd  
Ludwig und Christian Wolff
- 129 **Squirrel**  
Ein innovativer Editor zur Erstellung digitaler  
Editionen und sein Pilotprojekt  
Mark Hengerer, Isabella Hödl-Notter, Gerhard Schön

# Vorwort: Die digitale Transformation in den Geisteswissenschaften

Harald Klinke

Die digitale Transformation bezeichnet die Folgen der fortschreitenden Digitalisierung aller Lebensbereiche.<sup>1</sup> Die Digitalisierung hat bereits viele Zweige der Industrie – zum Teil disruptiv – verändert. Hierzu gehört beispielsweise die Kommunikation: Mit dem Smartphone ist man ständig mit dem Internet verbunden, kann Informationen abrufen und ist jederzeit erreichbar. Dies verändert beispielsweise die erwartete Reaktionsgeschwindigkeit auf Nachrichten und somit den beruflichen Tagesablauf und schafft die Möglichkeit neuer Arbeitszeitmodelle. Leistungsfähige mobile Endgeräte sind zudem die Voraussetzung für ein anderes Beispiel digitaler Transformation: die Mobilität. Ein Geschäftsmodell für Leihräder, wie zum Beispiel *Call-a-bike*, ist nur mit der dahinterstehenden Mobilfunk-Technologie und Smartphones denkbar. Eine App lokalisiert den Standort, zeigt Räder in der Umgebung an, entsperrt das Schloss und stellt am Ende die Fahrt in Rechnung. Beim *Car-Sharing* wird das Fahrtziel automatisch von der App an das Navigationsgerät des Autos weitergegeben und kann, dort angekommen, weiteren Mietern ggf. *on-the-fly* übergeben werden, was die Parkplatzsuche überflüssig macht. Unter dem Stichwort „Connected Car“ werden Daten über Sensoren erhoben, die mit anderen Fahrzeugen geteilt werden, um mehr Sicherheit zu gewährleisten.<sup>2</sup> So zeichnet das Unternehmen Tesla Fahr- und Situationsverhalten auf, um damit *Machine Learning*-Algorithmen zu trainieren und letztlich Autos autonom fahren lassen zu können. Diese

---

1 Vgl. Klauß, Thomas/ Mierke, Annika: Szenarien einer digitalen Welt – heute und morgen: wie die digitale Transformation unser Leben verändert, München 2017.

2 Vgl. Köhler, Thomas R./ Wollschläger, Dirk: Die digitale Transformation des Automobils: 5 Mega-Trends verändern die Branche, Pattensen 2014.

Entwicklungen haben Auswirkungen auf den Individualverkehr, die Parkplatzsituation, die Notwendigkeit individuellen Autobesitzes und die Rolle der Fahrzeughersteller als Dienstleister.<sup>3</sup> Weitere Beispiele für die Auswirkungen der Digitalisierung auf Geschäftsmodelle und sozio-ökonomische Zusammenhänge ist *Social Media*, die die politische Kommunikation in einen immer schnelleren Nachrichtenzyklus drängt.<sup>4</sup> Oder auch die Medizin, in der über das Sammeln großer Mengen an Daten über Körpersensoren nicht nur den Arzt bei der Therapie unterstützen, sondern auch den Patienten zunehmend zur Selbstdiagnose veranlassen.<sup>5</sup>

Letztlich betrifft die digitale Transformation nicht nur Unternehmen, auch die Wissenschaft ist Teil dieses Wandels und seiner Chancen. Die Naturwissenschaften haben längst begonnen, die Möglichkeiten der Datenerhebung und Datenverarbeitung systematisch in ihre Vorgehensweise einzubeziehen. So wäre die Sequenzierung des menschlichen Genoms ohne den Einsatz des Computers und der damit verbundenen Automatisierung und Auswertung nicht möglich gewesen.<sup>6</sup> An der Astronomie wird deutlich, dass der direkte Blick durch das Teleskop nicht mehr zum Arbeitsalltag eines Wissenschaftlers gehört, sondern die computergestützte Auswertung von Bilddaten vieler Frequenzbänder.<sup>7</sup> Und auch die Geisteswissenschaften haben eine lange Tradition des Computereinsatzes, um ihre wissenschaftlichen Forschungsziele zu erreichen. Unter dem Begriff der „Digital Humani-

---

3 Vgl. Martyn Briggs u. Karthik Sundaram: Environmentally Sustainable Innovation in Automotive Manufacturing and Urban Mobility, 2016, online abrufbar unter: <https://www.btplc.com/Purposefulbusiness/Ourapproach/Ourpolicies/cars2025report.pdf> (06.12.2017).

4 Vgl. Lehmann, Christian (u.a.): Geschäftsmodellinnovationen in der deutschen Medienwirtschaft – Zwischenbericht, in: Bayreuth Reports on Strategy (BaRoS), Nr. 9, Bayreuth 2015, S. 1–13.

5 Vgl. Müller, Martin U.: Wie das Handy den Arzt ersetzt, in: Der Spiegel, 29/2017, <http://www.spiegel.de/spiegel/handys-mit-medizin-apps-ersetzen-aerzte-und-teure-diagnosegeraete-a-1158365.html> (06.12.2017).

6 Vgl. Böhm, Kurt: High performance computing for the human genome project, in: Computer Methods and Programs in Biomedicine, Band 46(2), 1995, S. 107–112

7 Vgl. Malina, Roger: Yes again to the end of the Digital Humanities! Please!, 06.08.2015, <http://malina.diatrope.com/2015/08/06/yes-again-to-the-end-of-the-digital-humanities-please/> (06.12.2017).

ties“ – deren Gründungsvater meist Robert Busa bezeichnet wird, der bereits 1951 eine Konkordanz mithilfe von IBM-Rechnern aufbaute – werden Daten erhoben und mit Methoden der Statistik und Datenanalyse verarbeitet, um Zugriff, Findbarkeit und Analysen großer Korpora zu organisieren.<sup>8</sup>

Die digitalen Geisteswissenschaften stehen heute jedoch vor einer *zweiten Welle* von Einsatzmöglichkeiten der Technologie. Waren sie bisher von methodischen Fragen geprägt, wie der Computer bei der Auswertung von Daten den klassischen Fragestellungen beitragen können und wie neue Fragestellungen daraus entstehen, rückt jetzt zunehmend die Aufmerksamkeit auf andere Bereiche. Dies hat verschiedene Gründe:

(A) Die Menge an verfügbaren *Daten* steigt in den Geisteswissenschaften exponentiell. War Big Data bisher vorwiegend ein Thema der Naturwissenschaften, rücken große, heterogene und verteilte Datenmengen nun auch in den Mittelpunkt der Humanities.<sup>9</sup> Institutionen wie Universitäten, Bibliotheken oder Museen machen Bestände zunehmend digital verfügbar. Projekte, wie die geplante „Venice Time Machine“ machen es sich zur Aufgabe, niemals zuvor vorhandene Datenmengen in der Tiefenerschließung zugänglich zu machen.<sup>10</sup> Zudem werden diese Daten auch immer häufiger – beispielsweise über Linked Open Data – miteinander vernetzt. So bleiben die Repositorien verschiedener Institutionen nicht als Inseln separiert, stattdessen können diese Inseln miteinander verbunden und im Vergleich miteinander ausgewertet werden. Dies verspricht einen Quantensprung in dem Verständnis kultureller Prozesse und Zusammenhänge.<sup>11</sup>

---

8 Vgl. Thaller, Manfred: Geschichte der Digital Humanities, in: Jannidis, Fotis u.a. (Hrsg.): Digital Humanities: Eine Einführung, Stuttgart 2017, S. 3–12.

9 Vgl. Laney, Douglas: The Importance of ‚Big Data‘: A Definition, in: Gartner, 2012, <http://www.gartner.com/resId=2057415> (06.12.2017).

10 Vgl. Abbott, Alison: The ‘time machine’ reconstructing ancient Venice’s social networks, in: Nature, 2017, Band 546 (7658), S. 341, online abrufbar unter: <https://www.nature.com/articles/n-12446262> (06.12.2017).

11 Vgl. Schich, Maximilian: Figuring out Art History, in: International Journal for Digital Art History, Nr. 2, Okt. 2016, S. 40–67.

(B) Zunehmend stehen *Technologien* zur Verfügung, die noch vor wenigen Jahren außer Reichweite der Geisteswissenschaften lagen. Dazu gehört Open Source Software, die Auswertungsmethoden kostenlos verfügbar macht, die bisher nicht erschwinglich waren. Programmierplattformen wie Python oder R mit einer Entwicklercommunity, die für viele Zwecke Zusatzbibliotheken entwickeln, ermöglichen Datenanalysen, die bislang nur mit hohem Aufwand umsetzbar waren. Dazu gehört beispielsweise auch die Anwendung von Machine Learning, der zu Recht ein disruptives Potential zugeschrieben wird.

(C) Die interdisziplinäre *Zusammenarbeit* mit der Informatik ist heute ein übliches Element, das nicht mehr aus dem Wissenschaftsprozess wegzudenken ist. Der von der LMU organisierte internationale Hackathon „Coding Dürer“, der im März 2017 Kunsthistoriker und Informatiker in München zusammenbrachte, zeigte, welches Potenzial an der Schnittstelle zwischen den Geisteswissenschaften und Technologie liegt, das bisher weltweit noch kaum genutzt wird.<sup>12</sup> Durch solche Kooperation entstehen methodische und technologische Innovationen, von denen beide Seiten profitieren. Jetzt geht es nicht mehr nur um die Übernahme und Adaption von IT im wissenschaftlichen Prozess, jetzt besteht die Chance und die Notwendigkeit, die Geisteswissenschaften ganz neu zu definieren.

Aufgrund dieser genannten Treiber verändern sich die Geisteswissenschaften auf drei Ebenen:

(1) In der *Forschung* werden wissenschaftliche Erkenntnisse nicht mehr nur durch individuelles Studium von Wissensmedien erzeugt, sondern durch die Inanspruchnahme von Auswertungswerkzeugen der Datenanalyse, der kollaborativen Zusammenarbeit und des Publizierens von Daten. Dies führt zu neuen Arbeitsprozessen und der Notwendigkeit neuer Kompetenzen.

---

<sup>12</sup> Coding Dürer. A five days international and interdisciplinary Hackathon for Art History and Information Science, March 13–17, 2017, Munich, Germany, <http://codingdurer.de/> (06.12.2017).

(2) In der *Lehre* beginnt die Transformation der Fächer, denn zunehmend sind Nachwuchswissenschaftler mit Erfahrungen gefragt, die sie befähigt, sich an digitalen Forschungsprojekten zu beteiligen. Daher müssen diese Kompetenzen bereits früh in der Lehre vermitteln werden. Diese Lehrinhalte ersetzen nicht bisherige, sondern erweitern diese.

(3) Die neuen Geisteswissenschaften schaffen neue *Infrastrukturen*, die beispielsweise Cloudspeicher und Rechenleistungen zur Verfügung stellen, aber auch Fragen von Langzeitarchivierung und Support spielen eine wichtige Rolle. Dies wird mit der Verarbeitung wachsender Datenmengen immer relevanter, vor allem dann, wenn diese vernetzt zugreifbar sein müssen und umfassende Analysen erfordern, die nicht mehr auf einem lokalen Rechner durchführbar sind.

Im Freistaat Bayern bestehen besonders gute Voraussetzungen dafür, die digitale Transformation in den Geisteswissenschaften anzuführen. Als Technologieregion existiert hier ein großes Know-how im Bereich IT. Zahlreiche internationale Unternehmen und bedeutende Institutionen sind hier angesiedelt. Zudem hat die Landesregierung mit dem Masterplan BAYERN DIGITAL ein großes Interesse daran deutlich gemacht, die Chancen der Digitalisierung zu fördern und zu einem zentralen strategischen Element der Wissenschafts- und Forschungspolitik zu machen.<sup>13</sup> Dazu gehört beispielsweise die Gründung des Zentrums Digitalisierung Bayern (ZD.B).<sup>14</sup> Universitäten, Rechenzentren und kulturellen Institutionen spielen hier eine wichtige Rolle. Neben der Landeshauptstadt München sind vor allem Nürnberg-Erlangen, Regensburg und Passau Zentren der Innovation. Vor allem aber haben sich die digitalen Geisteswissenschaften in den letzten Jahren entscheidend vernetzt und organisiert. Dieser Band stellt

---

13 Im März 2017 wurde mit dem Masterplan BAYERN DIGITAL II ein umfangreiches Maßnahmenpaket mit einem Finanzvolumen von rund 500 Millionen Euro beschlossen: <https://www.stmwi.bayern.de/digitalisierung/bayern-digital/> (06.12.2017).

14 „Das Zentrum Digitalisierung Bayern (ZD.B) ist eine [...] Forschungs-, Kooperations- und Gründungsplattform, die als Impulsgeber in Zusammenarbeit mit Wirtschaft, Wissenschaft, Verbänden und öffentlichen Maßnahmen wirkt“: <https://zentrum-digitalisierung.bayern/> (06.12.2017).

die ersten Schritte in diese Richtung vor: Eine über mehrere Städte verteilte Lehre, eine serverbasierte Lehrplattform sowie Projekte, die zeigen, in welche Richtung der Weg der digitalen Geisteswissenschaften gehen sollte.

In der vorliegenden Publikation wird das Projekt „IT for all“ dargestellt, das zum Ziel hat, IT-Kompetenzen in allen Fächern zu integrieren. Das bezieht sich nicht nur auf Fragen des Curriculums, sondern auch auf die Lehrinfrastruktur und auf Projekte, die neue Möglichkeiten der Datenauswertung und digitalen Publikation vorstellen. Dazu hat am 19. Juni 2017 die Tagung „#DigiCampus. Digitale Forschung und Lehre in den Geisteswissenschaften“ an der LMU München stattgefunden, die von dem Institut für Kunstgeschichte und der IT Gruppe Geisteswissenschaften ausgerichtet wurde.<sup>15</sup> Dazu haben zahlreiche Referenten beigetragen, deren Vorträge hier in schriftlicher Form vorliegen. Dies sind erste Schritte, die deutlich machen, dass diese *zweite Welle* der Digitalisierung die Geisteswissenschaften stark verändern wird und zu weit mehr führt als nur zu Chancen für die Einzelfächer, sondern dass neue wissenschaftliche Erkenntnisse auch das Potenzial haben, umfangreichere gesellschaftliche Fragen zu beantworten und sich die Geisteswissenschaften in den Diskurs um Digitaltechnologie einzuschalten.<sup>16</sup> Sie bilden nicht nur Nutzer von Technologie aus, sie können die Entwicklungen – wie beispielsweise Big Data, Internet of Things, Machine Learning, Augmented/Virtual Reality – mit dem neuen methodischen Instrumentarium zudem epistemisch begleiten. So stellt sich auf der hier dargestellten Basis die Frage, was die nächsten Schritte der digitalen Geisteswissenschaften sind.

Historische und kulturelle Daten verfügbar zu halten, diese zu interpretieren, zu analysieren und für die Gegenwart nutzbar zu machen, ist der Kern der neuen Digital Humanities und wird zunehmend als

---

15 Videos der Vorträge auf dieser Veranstaltung sind abrufbar unter: <http://www.kunstgeschichte.uni-muenchen.de/forschung/digitalekg/digicampus/videos/> (06.12.2017).

16 Siehe dazu beispielsweise den Antrag auf ein DFG-Schwerpunktprogramm mit dem Titel „Das digitale Bild“ des Instituts für Kunstgeschichte der LMU München: <http://www.kunstgeschichte.uni-muenchen.de/forschung/digitalekg/digitales-bild1/> (06.12.2017).

„Cultural Analytics“ bezeichnet.<sup>17</sup> Um die Aufgaben in Zukunft weiter zu bündeln, zu koordinieren und in einem internationalen Rahmen zu vernetzen, ist auf regionaler Ebene eine Institution notwendig, die interdisziplinär die bayernweite Zusammenarbeit organisiert und unabhängig von bestehenden institutionellen Abhängigkeiten die Chancen der Digitalisierung für die Geisteswissenschaften vorantreibt. Ein Zentrum, das auf allen technischen Ebenen die Möglichkeiten der Technologie für die Fächer nutzt und gezielt organisiert, kann daher „*Center for Cultural Analytics*“ genannt werden. Die Gründung eines solchen Zentrums würde der Region einen Vorsprung sichern und zukunftssträchtige Projekte organisieren, die ohne eine solche organisatorische Bündelung nicht umsetzbar wäre. Gleichzeitig – daher der englischsprachige Titel – würde es eine internationale Vernetzung und Sichtbarkeit gewährleisten, die weit über die Region hinausweist. Die digitale Transformation liegt nicht vor uns, sondern ist bereits in vollem Gange. Jetzt geht es darum, derzeitige und zukünftige Entwicklungen aktiv zu gestalten und die Chancen der Veränderungsprozesse zu nutzen. Für die Geisteswissenschaften bedeutet dies neben der Erschließung neuer Forschungsfragen, die Disziplinen selbst neu zu definieren. Die Voraussetzungen, um hier eine progressiv-gestaltende Funktion einzunehmen, sind in Bayern derzeit ideal.

An dieser Stelle soll den Vortragenden und Autoren gedankt werden, die die Tagung „#DigiCampus. Digitale Forschung und Lehre in den Geisteswissenschaften“ am 19. Juni 2017 und die vorliegende Publikation mit ihren Beiträgen bereichert haben. Besonderer Dank gilt Philipp Hartmann, der für die Logistik und die Unterstützung des editorischen Prozesses verantwortlich war. Tagung und Publikation wurden ermöglicht durch die freundliche Unterstützung des Departments Kunstwissenschaften der LMU München.

---

17 „Cultural Analytics“ bezeichnet das Ergebnis der Kombination von Big Data, digitalem Mapping und konzeptioneller Analyse, um die Herausforderungen des durch das Digitale veränderten Verhaltens von Menschen vorherzusehen und sich darauf vorzubereiten. So wird der Begriff durch das CulturePlex definiert, einer von Juan-Luis Suarez geleiteten multidisziplinären Einrichtung an der University of Western Ontario: <http://www.cultureplex.ca/cultural-analytics/> (06.12.2017).

Video-Mitschnitte der Vorträge der Veranstaltung sind verfügbar unter <http://www.kunstgeschichte.uni-muenchen.de/forschung/digitalekg/digicampus/videos/index.html>

# Die Geisteswissenschaften und das Digitale

## Ein Quantensprung oder *business as usual*?

Hubertus Kohle

### I.

Das Digitale hat angefangen, das kollektive Bewusste in einer Weise zu bestimmen, wie nur wenige technische Phänomene zuvor. Manche halten es für den Agenten einer radikal gewandelten Zukunft und preisen sein Befreiungspotential, das den Menschen aus der Knechtschaft der Arbeit erlöst. Diskussionen, wie die um das bedingungslose Grundeinkommen, wären ohne die, von Software gesteuerte Automatisierung, weder möglich noch nötig. Andere drehen den Spieß um und erkennen im Digitalen ein Versklavungsinstrument, das dem Menschen die Selbstbestimmung nimmt und flächendeckende Kontrolle durchsetzt. Dabei sind die Rollen in der bundesdeutschen Presselandschaft klar verteilt: Im Wirtschaftsteil dominiert die erste, im Feuilleton die zweite Grunddisposition. Der Wirtschaftsteil beschwört die ökonomischen Chancen, die mit der Digitalisierung gerade in einem Hochindustrie- und Hochlohnland wie der Bundesrepublik erwachsen, und warnt vor Rückfall in die zweite Reihe, wenn wir diese Chancen missachten. Das Feuilleton, dem die Technologiekritik spätestens seit dem zweiten Weltkrieg zur *raison d'être* geworden ist, scheint nur darauf gewartet zu haben, dass NSA und Edward Snowden ihm die Argumente für eine Radikalskepsis an die Hand gaben. Die Geisteswissenschaften verbinden sich mit dem Feuilleton, insofern ist auch deren Rolle klar: Wie es sich in dieser Konstellation gehört, äußern sich ihre Vertreter gewöhnlich kritisch.

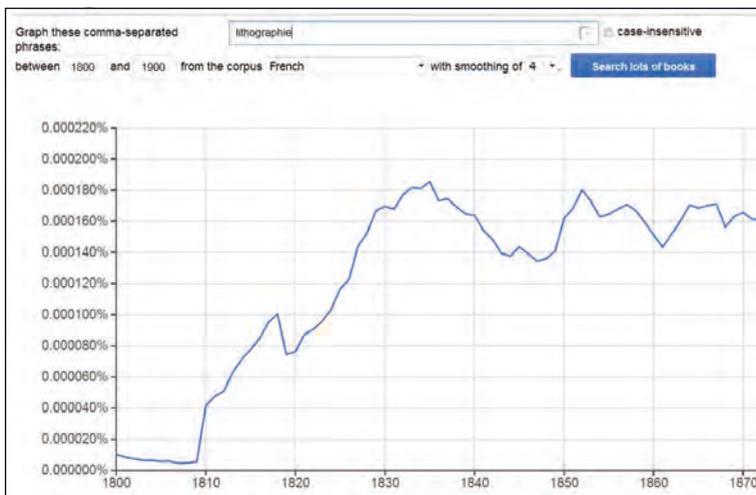
## II.

Auch die sogenannten Digital Humanities fristen eine Randexistenz, wobei sie allerdings zuletzt verstärkt von sich reden machen. Der quantifizierende Ansatz computationeller Methodik stößt in einer Fächerkultur auf Befremden, die im Qualitativen immer ihren eigentlichen Bestimmungsgrund gesucht und gefunden hat. Den Geist zu *messen*, das scheint sowohl denjenigen, die sich professionell mit ebendiesem befassen, als auch einer Öffentlichkeit, die in Geschichte, Kunst und Literatur eher das Andere der Technologie vermutet, als Sakrileg. Die Förderinitiativen des Staates im Bereich der digitalen Geisteswissenschaften, auch unternommen, um die eher lockere Verdrahtung der Absolventen im gesellschaftlichen Produktionsprozess zu straffen, stoßen bei den Vertretern traditionell hermeneutischer Ansätze auf Misstrauen. Sie werden gerne als Gängelung freier Wissenschaft beziehungsweise als aggressive Ökonomisierung denunziert.

Auch wenn ich an dieser Stelle für Offenheit plädieren möchte: Ein Blick auf die bisherigen Leistungen der Digital Humanities ist doch ein wenig ernüchternd. Er kann es, angesichts der vergleichsweise kurzen Geschichte dieser Disziplin, auch nur sein – wenn man von „Disziplin“ überhaupt reden will. Als „Digital Humanities“ figuriert der Forschungsbereich erst seit etwa einem Jahrzehnt, davor hieß er meist „humanities computing“, was schon im Begriff eine eher mangelhafte disziplinäre Identität verriet. Dabei fallen den Kritikern vor allem immer wieder die erzielten Ergebnisse im Kontrast zu den teilweise großspurigen Ankündigungen auf. Bei genauerem Hinsehen wird man allerdings entdecken, dass solche Ankündigungen den Digital Humanities eher unterstellt werden, als dass sie diese selber produziert, vermutlich von Fachvertretern, die bis heute die rechnerische Methodik nicht in den Kern ihrer jeweiligen Fachidentität haben vordringen lassen. Es lohnt sich in jedem Fall, die Anregungen Revue passieren zu lassen, die mit dem Rechner in das Gebiet des Geistes einziehen. Ich würde die Leistungen in drei Hauptbereiche einteilen. Da ist erstens das Digitale als analytisches Medium, zweitens als das der Publikation und drittens als das der Vermittlung.

### III.

Erstmal zur Analytik. Der überkommene Ausdruck für das, was heute vornehm verkürzend unter dem Begriff des „Digitalen“ läuft, also die gute alte elektronische Datenverarbeitung, bezeichnet trotz seiner mangelhaften up-to-dateness genau das, was die Digital Humanities betreiben: Sie analysieren *Daten*. Das macht sie einerseits verdächtig, weil es dem traditionellen Geisteswissenschaftler einfach nicht übers Herz gehen will, Goethes Gedichte aus dem West-Östlichen Diwan als simple *Daten* zu beschreiben. Andererseits ist *Datum* nur das „Gegebene“, und gegeben ist selbst Goethes dichterisches Werk. Gerade auf der Ebene des Gegebenen ist selbst dieses kein Zauberstück. Es besteht genauso aus Worten wie jedes andere sprachliche (Kunst-)Werk auch. Worte kann der Rechner zum Beispiel in Wortfrequenzanalysen thematisieren. Diese, scheinbar simple Unternehmung, liefert schon einige Aussagen über die historische und stilgeschichtliche Zuordnung von Kunstwerken. So ist es relativ einfach möglich, ein typisch klassizistisches von einem typisch romantischen Sprachkunstwerk abzugrenzen. Zugegeben, darin steckt noch nicht allzu viel Verstehen und sehr viel mehr Erklären, aber die Diltheysche Trennung und Entgegensetzung der beiden Bereiche ist vielleicht doch nicht so entschieden, wie



das gerne suggeriert wird. Gerade für das Problem der (historischen) *Genese* kann der Rechner einiges liefern, weniger wohl für das der (ästhetischen) *Geltung* – aber wer weiß. Vor allem kann er, nach meiner Erfahrung, auf erklärungsbedürftige Stellen hinweisen. In dem hier gezeigten Beispiel geht es um das Vorkommen des Begriffes „Lithographie“ in der französischen Publizistik des 19. Jahrhunderts (Abb.1).

Dieses setzt um 1800 ein, weil das graphische Reproduktionsverfahren genau zu diesem Zeitpunkt – übrigens hier, in Bayern – erfunden wird. Danach geht es erst einmal zurück, weil erst in der spät-napoleonischen Zeit mehr Gebrauch von ihm gemacht wird. Dann dreht die Kurve steil bergan, da das Verfahren in der frühen Restauration gerne für die oppositionelle politische Propaganda genutzt wird, allerdings nur bis zu den scharfen regierungsamtlichen Maßnahmen der späten 1810er Jahre, was sich in einer erneuten Abschwächung seiner Erwähnungshäufigkeit ausdrückt. Danach geht es wieder steil bergan und findet einen Höhepunkt vor 1835, der Zeit, in der Daumier und Kollegen ihre berühmten monarchiekritischen Karikaturen schaffen, die als Lithographien veröffentlicht werden. Auch den weiteren Höhepunkt nach der 1848er Revolution wird man erklären können, aber was ist mit der hohen Frequenz in den mittleren 1860er Jahren? Hier zeigt sich, dass der Rechner im vorhandenen Material erklärungs- und deuthungsbedürftige Stellen aufzeigt. Er bietet sich an als ein Recommender-System, das – und darin sehe ich das Entscheidende – die eigentliche Interpretation in Menschenhand belässt. Er zwingt mich, um bei dem Beispiel zu bleiben, den *peak* um 1865 zu erklären, für den ich zunächst einmal keine Deutungsmöglichkeit vorweisen kann.

Noch besser zu sehen ist das vielleicht bei einer datengenerierenden Anwendung wie Artigo, das wir hier an der LMU als eine *social-tagging-Anwendung* zur Einbindung der Laien in die Beschreibung von Kunstwerken betreiben. In 10 Jahren wurden hiermit 10.000.000 Annotationen gesammelt, mit denen in großen Bilddatenbanken nach Kunstwerken gesucht werden kann. Momentan nutzen wir die Daten zu einer Netzwerkanalyse, bei der wir den Computer eigenständig Cluster bilden lassen, mit denen er das umfangreiche Material

von 50.000 Bildern aus allen Epochen der Kunstgeschichte in Klassen einteilt. Und siehe da: Er kategorisiert die Werke auf der Basis der Tags in Gattungen – was an sich keine herausragende wissenschaftliche Leistung ist – aber doch dem Forschenden die Arbeit erleichtert. Schließlich bekommt dieser seine Werke gleich in sinnvollen Kategorien geliefert, mit denen er dann arbeitet. Dabei kann er die Verfeinerung der Kategorisierung fast beliebig weitertreiben. Das Verfahren hat übrigens tiefergehende Konsequenzen als es auf den ersten Blick scheinen mag. Einerseits können erheblich umfangreichere Bestände von Werken in die Analyse einbezogen werden, als das bei einer Meisterwerk-Kunstgeschichte üblich ist, andererseits ist das Material handhabbarer, da es jeweils präzise anzusteuern ist. Zu erwarten ist daher bei einer konsequenten Anwendung digitaler Analytiken eine unterschiedene Entkanonisierung bei gleichzeitiger Steigerung objektivierter historischer Befunde.

So eine Arbeit kann man als Kunsthistoriker natürlich nicht selbst machen, sodass ich mich der unschätzbaren Hilfe einer Statistikerin bediene. Bei dieser Gelegenheit möchte ich erwähnen, dass ich für die Heranbildung eines forschungsorientierten Nachwuchses auf dem Gebiet der Digital Humanities inzwischen den Eindruck habe, dass es lohnenswert wäre, Absolventen eines naturwissenschaftlichen Studienganges herüberzuziehen und in den Geisteswissenschaften mit einem entsprechenden Projekt zu promovieren. Interdisziplinarität und Kooperation bleiben in diesem Feld nicht bloß modische Forderung, sondern werden zur Notwendigkeit. Eine Tatsache, die die Geisteswissenschaften auch in ihrer Arbeitsweise an die Naturwissenschaften heranrückt.

Wenn ich vor allem bei dem zuletzt genannten Beispiel die Auswirkung des Computers eher als eine der Weiterentwicklung und nicht des Bruches beschreibe, so hängt das in den Augen der Techno-Utopiker, die sich in Deutschland vor allem aus der Kittler-Schule rekrutieren, mit der Tatsache zusammen, dass wir ihn bei Artigo nicht auf die Bilder selber, sondern auf deren Metadaten in Form von Worten zugreifen lassen. Denn hiermit binde ich den Computer an das Menschenwissen

zurück, dessen Überwindung sich diese Schule erhofft. Darüber hinaus kann man mit dem Rechner auch auf das Bild selbst zugreifen, indem man deren kleinste Einheiten, die sogenannten Pixel adressiert. Allerdings vertraue ich, im Gegensatz zu besagten Techno-Utopikern, die in der hiermit entstehenden Ordnung eine vollkommen neue, dem historischen Denkens entgegengesetzte beziehungsweise indifferente Ordnung erkennen wollen, darauf, dass das Geschichtliche als Kategorie im Horizont des Digitalen nicht einfach obsolet wird. Vielmehr spricht vieles dafür, dass sich Geschichtlichkeit nicht erst in der, vom Menschen vorgenommenen Benennung eines Phänomens, manifestiert, sondern dass diese – auf wie vermittelte Weise auch immer – bereits im Phänomen selbst angelegt ist. Gerne konzedere ich dabei andererseits, dass sich mit der direkten Bildadressierung eine neue Analytik ergibt, deren Aufschlusskraft für die Kunstgeschichte als historischer Wissenschaft noch gar nicht richtig abzuschätzen ist. Auch würde ich nicht ausschließen, dass sich hier Perspektiven ergeben, die das Fach deutlich stärker in eine systematische Wissenschaft verwandeln. Aber dann ist es eben eventuell keine historisch arbeitende Geisteswissenschaft mehr.

#### IV.

Der Computer hat sich – zweitens – als ideales Publikationsmedium erwiesen, wenn er auch als solches vor allem im Bereich der Geisteswissenschaften noch nicht flächendeckend akzeptiert wird. Dies gilt zunächst für die Darbietung des Ausgangsmaterials, also der historischen oder künstlerischen Quellen. Erst im Digitalen können Editionen so realisiert werden, dass sie komplexe historische Verläufe und Gestaltungsweisen zu repräsentieren in der Lage sind. Ganz schlagend ist auch der Vergleich mit den alten Zuständen in meinem Bereich: Früher hatte jedes kunsthistorische Institut eine mehr oder weniger gute Diasammlung der in der Lehre zu zeigenden Kunstwerke lokal vorrätig, für deren Produktion und Verwaltung ein erheblicher Personaleinsatz notwendig war. Heute greifen alle auf exponentiell größere Bestände zu, die zentral gespeichert beziehungsweise administriert werden. In einen sehr weiten Editions begriff gehören auch Rekon-

struktionen von früheren Zuständen, für die der Rechner ausgefeilteste Visualisierungsinstrumente zur Verfügung stellt. Das gilt für die Lesbarmachung von unleserlich gewordenen Textstellen, etwa durch Kontraststeigerung oder Streiflichtwirkungen genauso, wie für die CAD-gestützte Wiederauferstehung von zerstörten Gebäuden – einschließlich der Möglichkeit, sich in ihnen wie im Original zu bewegen. Ganz wichtig in den genannten Fällen: Aus den, in elektronischen Daten vorliegenden, Produkten können grundsätzlich alle möglichen anderen Erscheinungsformen realisiert werden, ein gedrucktes Buch als Verfestigung der digitalen Edition, ein mit dem 3D-Drucker gebautes Architekturmodell als Konversionsprodukt einer digitalen Rekonstruktion oder sogar ein herkömmliches Dia aus der Bilddatei – wenn Sie sich überhaupt noch daran erinnern, was es mit einem Diapositiv auf sich hat.

Ein besonders heißes Thema ist die wissenschaftliche Sekundärpublizistik. Gemeint ist die für den Leser kostenfreie und unbehindert mögliche Verwendung wissenschaftlicher Forschungsergebnisse über das Internet. Um hier ein wenig Abhilfe zu schaffen, sehen wir in einem, gemeinschaftlich von den Exzellenzuniversitäten Köln und München finanzierten, Projekt die hybride Publikation von herausragenden Doktorarbeiten vor und glauben, dass dies auch ein Aushängeschild für die Universitäten ist.<sup>1</sup> Hybrid meint hier, dass sie sowohl gedruckt als Buch, als auch *open access* im Netz erscheinen. Kostenlos ist *open access* nicht, und es wird alles darauf ankommen, dass die Sache so organisiert wird, dass sie nachher nicht etwa noch teurer wird, als das im Moment der Fall ist. Aber neben dem Kosten- ist vor allem das Reichweitenargument entscheidend. Online greifbare Bücher werden nachweislich häufiger gelesen als nur gedruckte. Darüber hinaus wichtig: erst im *open access* kann wissenschaftliche Forschung interoperabel werden, also in internet-gestützte, automatisierte und überlokale Analyseprozesse einbezogen werden.

---

1 <http://www.humanities-map.net/>

Wie ein Damoklesschwert hängt aber über diesem ganzen Bereich die augenblickliche Urheberrechtsgesetzgebung. Deren Liberalisierung im Bereich der wissenschaftlichen Anwendungen wird zwar zur Zeit, im Sommer 2017, durchgeführt, unter dem Trommelfeuer interessierter Verlage und einflussreicher Presseorgane scheint diese aber längst nicht so weit zu gehen, wie es eigentlich wünschenswert wäre. Ich habe manchmal den Eindruck, dass die unter Originalitäts-, Effizienz- und Konkurrenz Gesichtspunkten unbedingt notwendige Verlagerung der Wissenschaft ins Internet unter diesen Bedingungen ganz grundsätzlich behindert, wenn nicht unmöglich wird. Das sieht übrigens bei der flexibleren *fair-use*-Regelung in den USA ganz anders aus.

## V.

Das Internet hat, drittens, zu einer Explosion an Bildungsangeboten geführt – meine Überlegungen gehen jetzt gegen Ende deutlich über die Frage nach den eigentlichen Digital Humanities hinaus. Die Khan Academy produziert seit 2004 tausende von Kursen aus fast allen Bereichen der Bildung auf Englisch, Spanisch, Bengali und Hindi und verbreitet sie in Form von Videos über YouTube. Der Schwerpunkt liegt dabei allerdings auf den Naturwissenschaften. Die genannten Sprachen weisen darauf hin, dass hier die Vermittlung an Menschen im Vordergrund steht, die aufgrund ihres Wohnortes und ihrer Armut ansonsten kaum in der Lage sind, ihre Potentiale zu realisieren. Am Verbreitungskanal YouTube ist zu sehen, dass die Mauern zwischen Wissenschaft und Laien-Öffentlichkeit niedriger werden. Wer sich auf das Internet einlässt, ist immer nah an nur begrenzt wissenschaftlichen Angeboten dran. Ein eingeschworener Kritiker aus der Germanistik sagte mir einmal, er wolle doch seine Veröffentlichung nicht neben der Porno-Reklame sehen.

Auch Universitäten, die etwas auf sich halten, bieten online mannigfaltige Lehrveranstaltungen an. Sie wissen selbst, wie aktiv die LMU auf diesem Feld ist. So wie alle neuen Medien in ihrer Frühzeit die alten Medien nachahmen, gibt es auch in der online-Lehre anfänglich vor allem Veranstaltungstypen, die aussehen, als hätte man eine analoge

Lehrform eins zu eins ins Netz übertragen. Vor allem bei den eher informellen Vermittlungsformen scheint mir noch eine Menge Weitergehendes möglich. Als ich vor Jahren mal vorgeschlagen habe, die Fakultäten dieser Universität könnten über jeweils aktuelle Forschungsprojekte auf hohem Niveau, aber knapp und verständlich in einem gut organisierten und ästhetisch anspruchsvollen Weblog berichten, ist das nur bei wenigen auf Gegenliebe gestoßen. Gerade aber in einer politischen Großwetterlage, in der Experten und Eliten immer mehr unter den Verdacht der Selbstbedienungsmentalität geraten, gleichzeitig aber dringend auf Legitimation ihrer Unternehmungen gegenüber denjenigen angewiesen sind, die sie durch ihre Steuermittel letztendlich finanzieren, schiene mir ein solches Unternehmen noch immer ausgesprochen aussichtsreich. Es gibt inzwischen Universitäten, die ihr, im Übrigen äußerst anspruchsvolles Lehrprogramm, komplett ins Internet verlegt haben. Das scheint mir nur einen Vorgeschmack auf das zu geben, was die nicht nur universitäre Bildung in absehbarer Zukunft ausmacht. Ob die Institution in ihrer räumlichen Verfasstheit in der Mitte des Jahrhunderts überhaupt noch existieren wird, ist in keiner Weise ausgemacht. Zumindest wird sich die online-Lernform neben der in situ angebotenen als feste Größe etablieren.

## VI.

Gerade was die Geisteswissenschaften angeht, ist das Ergebnis dieser extrem kursorischen Überlegungen zwiespältig. In der Organisationsform kündigen sich radikale Veränderungen an, sowohl in der Art, wie wir unsere Ergebnisse publizieren, als auch in derjenigen, wie wir sie vermitteln. In der Frage nach den Inhalten habe ich den Eindruck, dass die Geisteswissenschaften weniger dramatischen Veränderungen entgegensehen: Wenn entsprechende Frageinteressen überhaupt noch eine Rolle spielen werden – manche Entscheidungen im politischen Feld, gerade auch im Ausland, mögen hier Skepsis erzeugen – dann steht die Spezifik dieser Interessen einer baldigen Automatisierung eher entgegen. Ich würde sogar meinen, dass geisteswissenschaftliche Fragestellungen ihrer Natur nach an den nicht substituierbaren menschlichen Geist gebunden bleiben.



# Datenanalyse in der Digitalen Kunstgeschichte

## Neue Methoden in Forschung und Lehre und der Einsatz des DHVLab in der Lehre

Harald Klinke

Ein Schwerpunkt am Kunsthistorischen Institut der LMU München liegt in der Digitalen Kunstgeschichte. Der Begriff bezeichnet die Entwicklung neuer Erkenntnismöglichkeiten für das Fach mittels des Einsatzes von Computern. Konkret geht es um die Erhebung, Verarbeitung und Visualisierung von Daten, um daraus neue, bisher nicht erkannte Zusammenhänge zwischen Objekten, Künstlern und Ideen herzustellen.<sup>1</sup> Dies ist für die Kunstgeschichte ein Ansatz, dessen Wurzeln bis in die 1980er Jahre zurück reichen<sup>2</sup> und sich als Teil der Digital Humanities versteht. Da es sich hier jedoch vorwiegend um visuelle Daten handelt und die wissenschaftliche Erkenntnis in erster Linie auf einem Bildverständnis aufbaut, hat das Thema erst in den letzten Jahren, vor allem durch die Inhaltserschließung mittels Systemen wie *ARTigo* oder Visual Computing zu einer dynamischen Entwicklung gefunden, die das Potential hat, das Fach und seinen Arbeitsmethoden umfangreich zu verändern.

Im Zentrum kunsthistorischer Forschung steht zunächst das einzelne Original. Kein Fach der Geisteswissenschaften betrachtet so systematisch Bilder als historische Dokumente wie die Kunstgeschichte. So kann man von diesem Fach von einer empirischen Wissenschaft sprechen, denn im ersten Schritt werden aufgrund einer Inaugenscheinnahme Daten vom Werk erhoben. Dies geschieht über eine

---

1 Vgl. Klinke, Harald: Digitale Kunstgeschichte, in: Akademie Aktuell, 01/2016, S. 68–71.

2 Vgl. Bentkowska-Kafel, Anna: Debating Digital Art History. International Journal for Digital Art History, Nr. 1, Juni 2015, S. 50–65.

Bildbeschreibung. Gesehenes in Worte zu bringen dient dazu, sich schrittweise bewusst zu machen, was dargestellt ist, um anschließend mithilfe von Bildvergleichen Gemeinsamkeiten und Unterschiede herauszuarbeiten, um beispielsweise Datierungen und Zuschreibungen vornehmen zu können und das Einzelwerk in einen kulturhistorischen Zusammenhang zu kontextualisieren.<sup>3</sup>

Doch gehen Kunsthistoriker längst nicht mehr allein vom Original aus, sondern nutzen umfangreiche Bestände medialer Repräsentationen von Kunstobjekten. Im 20. Jahrhundert war das wichtigste Medium das Diapositiv, das vor etwa 100 Jahren durch Heinrich Wölfflin in München in das Fach eingeführt wurde.<sup>4</sup> Diatheken an kunsthistorischen Instituten, die nicht selten mehrere hunderttausend Objekte enthielten, dienten dazu, sehr schnell auf Abbildungen von Objekten zugreifen und miteinander in einen Zusammenhang bringen zu können, die sich physisch nicht an einem Ort befinden oder inzwischen nicht mehr im Original vorhanden sind. Nicht nur entwickelte sich die dadurch mögliche Doppelpjektion zu einem wichtigen methodischen Element des systematischen Bildvergleichs, auch erlaubte der Zugriff auf jedes beliebige Werk zu einer universalen Vorstellung der Geschichte der Kunst, die vor Landes- und Kulturgrenzen keinen Halt macht.<sup>5</sup>

Inzwischen sind die Diatheken aus dem täglichen Arbeitsprozess von Forschung und Lehre verschwunden, nachdem sie zunächst aufwändig digitalisiert wurden und mit den Metadaten, die bisher auf den Dia-

---

3 Die Methode, systematisch von der Bilderbeschreibung zu einer schrittweisen Bedeutungszuschreibung und Kontextualisierung zu gelangen, wird von Erwin Panofsky beschrieben: Panofsky, Erwin: Ikonographie und Ikonologie, in: Kaemmerling, Ekkehard (Hrsg.): Bildende Kunst als Zeichensystem. Ikonographie und Ikonologie. Band 1: Theorien – Entwicklung – Probleme, Köln 1994, S. 207–225.

4 Vgl. Ratzeburg, Wiebke: Mediendiskussion im 19. Jahrhundert. Wie die Kunstgeschichte ihre wissenschaftliche Grundlage in der Fotografie fand, in: Kritische Berichte, Band 30, Nr. 1, 2002, S. 22–40 und Reichle, Ingeborg: Fotografie und Lichtbild. Die unsichtbaren Bildmedien der Kunstgeschichte, in: Zimmermann, Anja (Hg.): Sichtbarkeit und Medien: Austausch, Verknüpfung und Differenz von naturwissenschaftlichen und ästhetischen Bildstrategien. Hamburg University Press: Hamburg 2004, S. 177–191

5 Vgl. Malraux, André: Das imaginäre Museum, Frankfurt am Main 1987.

rähmchen standen, in Bilddatenbanken abgelegt wurden. Man stellte jedoch schnell fest, dass sich die Diatheken verschiedener kunsthistorischer Institute zu einem Großteil inhaltlich überlappen und nicht redundant gescannt werden müssen, wenn eine gemeinsame Datenbank aufgebaut würde. Im *Prometheus Bildarchiv*, einer im Jahr 2001 gegründeten Metasuchmaschine über die Bestände verschiedener universitärer Bilddatenbanken, befinden sich inzwischen etwa 1,7 Millionen Datensätze.<sup>6</sup>

Die Digitalisierung eröffnet weitaus mehr als nur die Möglichkeit, schneller einzelne Abbildungen zu suchen und zu finden. Liegen diese Bilddaten digital vor, können diese auch verarbeitet werden, um kunsthistorische Erkenntnisse zu erzeugen. Die Visualisierung großer Korpora stellt eine zusätzliche Lösung für die kunsthistorische Forschung dar, die Geschichte der Kunst auf neue Weise in ihrer Komplexität zu erfassen. Da der Computer mehr Bilder speichern und verarbeiten kann, als ein Mensch in seiner Lebensspanne sehen und wahrnehmen, wird in Anlehnung an den Begriff „Big Data“ hier von „Big Image Data“ gesprochen<sup>7</sup> oder auch in Anlehnung an den Begriff des „Distant Readings“ aus der Literaturwissenschaft von einem „Distant Viewing“.<sup>8</sup>

An einem Beispiel wird dies deutlich: Google hat seit dem Jahr 2004 in Zusammenarbeit mit Bibliotheken über 25 Millionen Bücher mit dem Ziel gescannt, jedes jemals publizierte Werk digital verfügbar zu machen.<sup>9</sup> Liegen diese Texte einmal digital vor, können Worthäufigkei-

---

6 Vgl. Dieckmann, Lisa: *Prometheus the distributed archive for research and education goes international*, in: EVA London 2008. Electronic Visualisation and the Arts. Conference Proceedings. London 2008, S. 61–67.

7 Vgl. Klinke, Harald: *Big Image Data within the Big Picture of Art History*. International Journal for Digital Art History, Nr. 2, Okt. 2016, S. 14–37.

8 Vgl. Moretti, Franco/Piazza, Alberto: *Graphs, Maps, Trees: Abstract Models for Literary History*, London 2007 und Zweig, Benjamin: *Forgotten Genealogies: Brief Reflections on the History of Digital Art History*, in: International Journal for Digital Art History, Nr. 1, Juni 2015, S. 38–49.

9 Vgl. Heyman, Stephen: *Google Books: A Complex and Controversial Experiment*, in: New York Times, 28. Okt. 2015, <https://www.nytimes.com/2015/10/29/arts/international/google-books-a-complex-and-controversial-experiment.html> (06.12.2017).

ten über die Zeit visualisiert werden. Das Tool *Google Ngram Viewer* zeigt so,<sup>10</sup> dass beispielsweise Albrecht Dürer in Publikationen mehr erwähnt wurde als Lucas Cranach (Abbildung 1). Im Speziellen wird sichtbar, dass in den 1930er Jahren mehr über den Künstler publiziert wurde, als zuvor, dass diese Erwähnungshäufigkeit nach dem Zweiten Weltkrieg eingebrochen ist und zur Nürnberger Dürer-Ausstellung zum 500. Dürer-Geburtstag wieder anstieg, um anschließend um einen mittleren Wert zu pendeln. Man könnte behaupten, dies bestätigt nur, was man ohnehin schon weiß – das Diagramm ließe sich nur durch bestehendes Wissen interpretieren. Jedoch werden die Ergebnisse erst durch die Relation zu anderen Ergebnissen, hier zu Cranach, interessant, denn es wird möglich, die Rezeption von Künstlern in der Literatur empirisch quantitativ zu untersuchen und zu vergleichen. Es handelt sich also um eine zusätzliche Methode, um Hypothesen zu prüfen und aufgrund der Daten neue Forschungsfragen zu entwickeln, denen qualitativ nachgegangen werden kann.

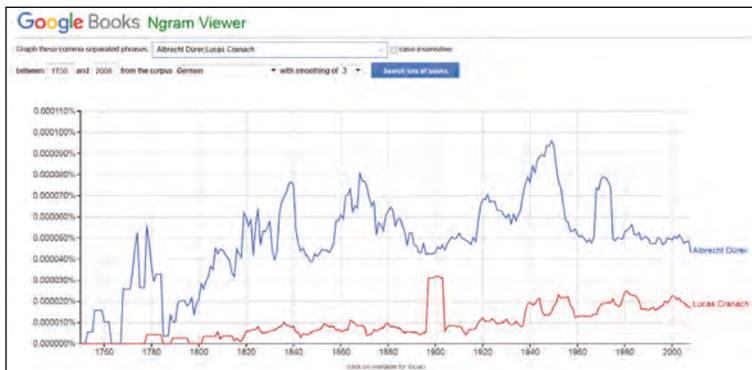


Abbildung 1: Häufigkeitsverteilung der Erwähnungen von „Albrecht Dürer“ und „Lucas Cranach“ im *Google Ngram Viewer*

Um für die Kunstgeschichte eine vergleichbare Datenbasis zu schaffen, sind nicht nur Buch-, sprich Textdaten, sondern auch Bilddaten relevant, denn Kunstwerke liegen, beispielsweise in den kunsthistorischen Instituten, vor allem als Bilddatenbanken bestehend aus Meta-

10 Ngram Viewer: <https://books.google.com/ngrams> (06.12.2017).

daten und fotografischer Abbildung repräsentiert vor. Am Beispiel der Datenbank *Artemis* des kunsthistorischen Instituts der LMU München wird deutlich, dass bereits die Metadaten durch Aggregation einen Überblick über die Sammlung geben kann (Abbildung 2). Summiert man die Anzahl der in der Datenbank vorhandenen Werke je Entstehungsjahr auf einer chronologischen Skala, wird das im Zusammenhang mit Forschung und Lehre dieses Instituts stehende Sammlungsinteresse ablesbar, hier vor allem das 19. Jahrhundert.

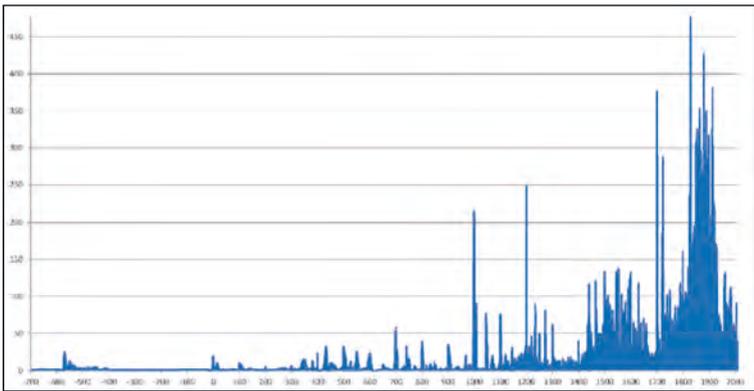


Abbildung 2: Zeit-Signatur der *Artemis*-Bilddatenbank

Das Fach Kunstgeschichte ist vor allem eine Bildwissenschaft mit einer historischen und visuellen Kompetenz, wodurch sie sich von allen anderen Geisteswissenschaften unterscheidet. „Digitale Kunstgeschichte“ bedeutet, diese Kompetenzen mit digitalen Methoden zu unterstützen. Daher kann es nicht nur um Metadaten gehen, sondern auch um die Erschließung von Bildern und ihren Inhalten. Dies ist jedoch eine technische Herausforderung, sind Digitalbilder strenggenommen nur Pixelraster aus RGB-Helligkeitswerten. Die relevante Information, das heißt das zu erkennende Dargestellte, liegt auf einer höheren semantischen Ebene. Aber alleine die Berechnung der Durchschnittshelligkeit jedes in der Datenbank vorhandenen Bildes und die Sortierung nach diesem Wert können einen ganz anders gearteten Überblick über einen Bild-Korpus liefern. Wieder am Beispiel der *Artemis*-Datenbank kann gezeigt werden, dass bereits auf diese Weise

inhaltliche Unterscheidungen deutlich werden (Abbildung 3). So finden sich auf der einen Seite der Skala Werke auf Papier, wie Zeichnungen oder Architekturdarstellungen mit hohem Weißanteil, und auf der anderen Seite vor allem Porträtgemälde und Darstellungen von Skulpturen vor schwarzem Hintergrund. Allein eine solche, einfach zu erstellende Darstellung kann beispielsweise ein Hilfsmittel sein, um einem Kurator einen Überblick über das Depot des Museums zu geben oder Besuchern im Web einen Eindruck der Inhalte der Sammlung.



Abbildung 3: 45.000 nach Helligkeit sortierte Objekte der Artemis-Bilddatenbank

Neben diesen Eigenschaften auf niedriger semantischer Ebene (*low level features*) ist das Ziel, auch auf Inhaltsdaten einer kunsthistorischen Datenbasis zugreifen zu können. Daher hat das Institut für Kunstgeschichte zusammen mit dem Institut für Informatik der LMU München im Jahr 2010 das Projekt *ARTigo.org* entwickelt.<sup>11</sup> In einen

11 Nach der Entwicklung eines Prototypen im Jahr 2007 wurde das Projekt im Rahmen von „play4science“ ab 2010 von der DFG gefördert, siehe Kohle, Hubertus/Bry, François/Kreffeld, Thomas/Riepl, Christian/Schulz, Klaus: Entwicklung sozialer Web-Plattformen zur Datengewinnung in den Geisteswissenschaften (KO 1091/4–2): [http://www.play4science.uni-muenchen.de/oeffentlichkeit/publikationen/abschlussbericht\\_2013\\_11\\_15.pdf](http://www.play4science.uni-muenchen.de/oeffentlichkeit/publikationen/abschlussbericht_2013_11_15.pdf) (06.12.2017) und Wieser, Christoph/Bry, François (u.a.): ARTigo. Building an Artwork Search Engine With Games and Higher-Order Latent Semantic Analysis, in: Proceedings of Disco 2013, Workshop on Human Computation and Machine Learning in Games at HComp, Palm Springs, CA, USA (6th-9th November 2013).

Crowdsourcing- und Gamification-Ansatz wird der User an einem Webfrontend dazu ermutigt, Schlagwörter zu Objekten aus der *Artemis*-Datenbank einzugeben, indem dieser zeitbegrenzt gegen einen virtuellen Gegenspieler antritt, Spielpunkte sammelt und in einer Bestenliste erscheint. Im Backend erzeugt dieses Verfahren Daten über die heutige visuelle Wahrnehmung von Digitalisaten historischer Bildwerke. Die Datenbasis von inzwischen über 9,3 Millionen Schlagwörtern zu etwa 80.000 Bildobjekten ermöglicht zunächst, Objekte besser zu finden. So liefert beispielsweise die Suchanfrage „Himmel-Wolken“ eine Liste an Werken zurück, die mit „Himmel“, aber nicht mit „Wolken“ getaggt wurden, also Bilder mit wolkenlosem Himmel. Ebenso ist – ähnlich wie beim *Google Ngram Viewer* – die Darstellung der Verteilung von Schlagwörtern über die Zeit möglich, um so beispielsweise zu erkennen, in welcher Zeit Häufungen an Objekten in dieser Sammlung auftreten, die von den Usern mit einem bestimmten Begriff versehen wurden.

Schaut man sich diese Begriffe in ihrer Verteilung genauer an, wird deutlich, dass einige Begriffe häufiger genannt werden als andere: „Schwarz“, „Weiß“, „Mann“, „Frau“ etc. sind Tags, die viele Objekte gemeinsam haben. Sie können deshalb nicht zur Unterscheidung der Objekte voneinander dienen. Schaut man sich die Schlagwörter zu zwei Werken, z. B. von Albrecht Dürer, an (Abbildung 4), wird klar, welche Begriffe beide Objekte teilen („Bart“, „Hut“ etc.) und in welchen sie sich unterscheiden („Schrift“, „Kette“ etc. zu der Darstellung Maximilians; „Bart“, „Signatur“ etc. zu der Phillips). Gemeinsamkeiten und Unterschiede zwischen zwei Werken festzustellen ist spätestens seit der durch Heinrich Wölfflin eingeführten Dia-Doppelprojektion eine fest etablierte kunsthistorische Methode (s. o.). Liegen nun eine große Anzahl von Objekten verbunden mit Schlagwortwolken digital vor, können diese nun der automatisierten Berechnung von inhaltlicher Nähe und Distanz nicht nur zweier Objekte, sondern von allen in der Datenbank vorhandenen Objekten dienen und so als Grundlage von gesamtheitlichen Bildvergleichen, Visualisierungen und Erschließungsmöglichkeiten dienen.



sich die Schlagworte (vor allem „Zeichnung“, „Holzschnitt“, „Renaissance“, „Monogramm“ etc.), darum herum radial gruppiert finden sich Illustrationen Dürers zu seinen Büchern „Proportionslehre“ und „Unterweisung der Messung“ (oben und links), die Druckwerke „Große Passion“ und „Kleine Passion“ (unten), die Gemälde (rechts unten) sowie die Aquarelle (rechts oben).<sup>12</sup>

Auch hier bietet es sich an, nicht nur Begriffe, sondern die Digitalisate selbst in dem Netzwerkdiagramm darzustellen.<sup>13</sup> In einer 3D-Visualisierung<sup>14</sup> wird deutlich, dass sich auf diese Weise die Objekte der *Artemis*-Datenbank nun besser als bei der Sortierung aufgrund von Helligkeitswerten trennen lassen (Abbildung 6): Porträtwerke clustern sich nun in Zusammenhängen abseits von Architekturzeichnungen, Landschaftsmalerei und der Abstraktion. Eine solche Darstellung könnte nun beispielsweise als *User Interface* dienen, um explorativ auf eine Bilddatenbank zuzugreifen.<sup>15</sup> Ziel eines solchen Systems kann es sein, auf der Ebene des *Distant Viewings* schnell durch einen Korpus zu navigieren, neue Zusammenhänge zu erkennen, indem benachbarte oder diametral entgegengesetzte Werke gefunden werden und das Einzelobjekt in seinen Big Data-Kontext gesetzt wird. Auf diese Weise kann ein neuer Blick auf die Geschichte der Kunst geworfen werden<sup>16</sup> – und gleichzeitig lässt sich das Einzelwerk entdecken, das ohne die Hilfe des Computers nicht gefunden worden wäre, und damit zum

12 Detailliertere Analysen sind über das Analytics Center des Instituts für Informatik der LMU (<http://analytics.pms.ifl.lmu.de>) sowie über das Analysecenter des DHVLab (<https://dhvlab.gwi.uni-muenchen.de/analysecenter/>) möglich.

13 Dass man mit Bildern über Bilder sprechen sollte forderte bereits Gottfried Boehm, vgl. Boehm, Gottfried: Die Bilderfrage, in: ders. (Hrsg.): Was ist ein Bild?, München 1994, S. 325–343.

14 Hierbei handelt es sich um ein Gemeinschaftsprojekt von Clements Schefels, Erich Schubert (beide LMU München) und Arthur Zimek (Universität Heidelberg), beschrieben auf dem ARTigo Blog: Data Analysis on ARTigo, <http://blog.artigo.org/2016/03/11/data-analysis-on-artigo/> (07.12.2017).

15 Beispiele für explorative Interfaces finden sich bei Glinka, Katrin (u.a.): Linking structure, texture and context in a visualization of historical drawings by Frederick William IV (1795–1861), in: International Journal for Digital Art History, Nr. 2, Okt. 2016, S. 198–213 oder Bernhard, Mathias: Gugelmann Galaxy: An Unexpected Journey through a collection of Schweizer Kleinmeister, in: Ebd., S. 94–113.

16 Vgl. Manovich, Lev: Manifesto for Democratic Art History, <http://lab.softwarestudies.com/2016/02/manifesto-for-democratic-art-history.html> (13. Februar 2016).

*Close Viewing*, letztlich dem Aufsuchen des Originals, zurückkehren. Der Computer wird also zum Hilfsmittel im digitalen Prozess kunsthistorischer Forschung.



Abbildung 6: Multidimensionale Skalierung der ARTigo-Daten<sup>17</sup>

Solche Methoden sind in der Einheit von Forschung und Lehre an die Studierenden weiterzugeben,<sup>18</sup> um ihnen die notwendigen Schlüsselkompetenzen zu vermitteln, die zur Mitarbeit an digitalen Projekten an der Universität, in Museen oder im Kunstmarkt notwendig sind. Diese Methoden sind aber ebenso mit ihnen weiterzuentwickeln. Denn viele Fragen sind noch unbeantwortet, beispielsweise welche weiteren Daten wie erhoben werden müssen, auf welche Weise die Datenanalyse zu neuen kunsthistorischen Erkenntnissen führen kann und wie diese Entwicklungen das Fach in seinem Selbstverständnis verändern werden.

<sup>17</sup> S. Anm. 14

<sup>18</sup> Siehe das Bildungsideal Wilhelm von Humboldts: Humboldt, Wilhelm von: Bericht der Sektion des Kultus und des Unterrichts an den König, Dezember 1809, in: Flitner, Andreas/Giel, Klaus (Hrsg.): Wilhelm von Humboldt – Werke in fünf Bänden. Band IV: Schriften zur Politik und zum Bildungswesen, Darmstadt 1982, 3. Aufl., S. 210–238 sowie ders.: Ueber die innere und äussere Organisation der höheren wissenschaftlichen Anstalten in Berlin (1810), in: Flitner (op. cit.), S. 255–266.

Aus diesem Grund werden am Kunsthistorischen Institut der LMU München Lehrveranstaltungen zur Digitalen Kunstgeschichte durchgeführt. In den letzten drei Jahren wurden unter anderem Seminare mit den Titeln „Digitale Methoden der Kunstgeschichte“, „Software für Kunsthistoriker“, „Dürer Digital“ oder „Datenanalyse für Kunsthistoriker“ angeboten.<sup>19</sup> Dabei besteht ein zentrales Problem darin, im Seminarzusammenhang die jeweilige Software auf unterschiedlichen Hardware- und OS-Kombinationen der Studierenden zu installieren. So kam beispielsweise die Tabellenkalkulation (*Open Office Clac*), Statistik (*RStudio*), Bildverarbeitung (*ImageJ*), Netzwerkanalyse (*Gephi*) oder visuelle Programmiersprachen (*Processing*) zum Einsatz, die zunächst zeitaufwändig auf den Einzelrechnern installiert werden musste, bevor begonnen werden konnte, gemeinsam damit zu arbeiten.

Die Lösung für dieses Problem ist eine Lehr- und Forschungsumgebung, die folgende Kriterien erfüllt. Ein solches System sollte

1. serverseitig **Software** vorinstalliert haben, auf die von verschiedenen Clients aus zugegriffen werden kann, um die Schwierigkeiten individueller Installationen zu vermeiden.
2. praktisch unbegrenzten **Speicherplatz** zur Verfügung stellen, um auf beliebig große Datenmengen Zugriff zu haben. Während auf dem lokalen Rechner die Kapazität der Festplatte das Limit darstellt, kann in einem cloudbasierten System in Zukunft auch Big Data verarbeitet werden, wenn auf dem Server entsprechende Datenbanken angeboten werden.
3. **Rechenleistung** zur Verfügung stellen, die ggf. weit über die eines lokalen Rechners hinaus gehen, denn wenn in Zukunft größere Datenmengen analysiert werden müssen oder Anwendungen wie *Machine Learning* oder *Augmented Reality* relevant werden, ist ein schwellenloser Zugang zu Computing-Cluster hilfreich.
4. das **Teilen** von Ergebnisdaten ermöglichen, um ein kollaboratives Arbeiten zu unterstützen. Teil dessen ist auch das Ma-

---

<sup>19</sup> Vgl. Klinke, Harald: Digitale Kunstgeschichte lehren, in: Kuroczyński, Piotr u.a. (Hrsg.): *Computing Art Reader. Einführung in die digitale Kunstgeschichte* (erscheint in Kürze).

nagement von **Langzeitarchivierung**, um auch später auf Forschungsdaten zugreifen zu können.

5. das **Publizieren** von Ergebnissen ermöglichen. Wurde die Kommunikation der Ergebnisse in bisherigen Seminaren über externe Blogs organisiert, sollte diese Funktion zentral in ein solches System integriert sein.

Anforderungen an eine Lehr- und Forschungsumgebung	Konventionelles Vorgehen	DHVLab
Serverseitig vorinstallierte Software	✘	✓
Unbegrenzt verfügbare Datenbasis	✘	✓
Serverseitige Rechenpower	✘	✓
Integriertes Teilen der Ergebnisdaten	Über externe Dienstleister	✓
Integriertes Publizieren der Ergebnisse	Über externe Dienstleister	✓

Abbildung 7: Übersicht von Kriterien an eine Lehr- und Forschungsumgebung

Um diesen Kriterien gerecht zu werden, wurde die virtuelle Lehr- und Forschungsumgebung *DHVLab* entwickelt und im Wintersemester 2016/17 erstmals in der Lehre eingesetzt.<sup>20</sup> Zur Verwendung des *Digital Humanities Virtual Laboratory* benötigen die Studierenden nur einen Laptop mit Browser und einem WLAN-Zugang (z.B. Eduroam) und können sich mit einem Passwort und Username am System anmelden, in dem sie auf einen Desktop eines virtuellen Rechners zugreifen können, auf dem die notwendige Software bereits installiert ist, die zu verwendenden Daten vorliegen und eigene Daten abgelegt und mit anderen Studierenden ausgetauscht werden können.

<sup>20</sup> Weitere Informationen dazu finden sich unter: <https://dhvlab.gwi.uni-muenchen.de/> sowie in den Vorträgen der Tagung vom 19.06.2017 unter <http://www.kunstgeschichte.uni-muenchen.de/forschung/digitalekg/digicampus/videos/>.

Der Lehrende kann in diesem System ein „Lab“ einrichten, das bedeutet eine Arbeitsgruppe von Studierenden, die sich registrieren können und Zugriff auf die virtuellen Maschinen erhalten. In diesem System können Softwarepakete definiert und Datenbanken konnektiert werden sowie über einen Blog zentral Ergebnisse in einen öffentlichen Diskurs gestellt werden.

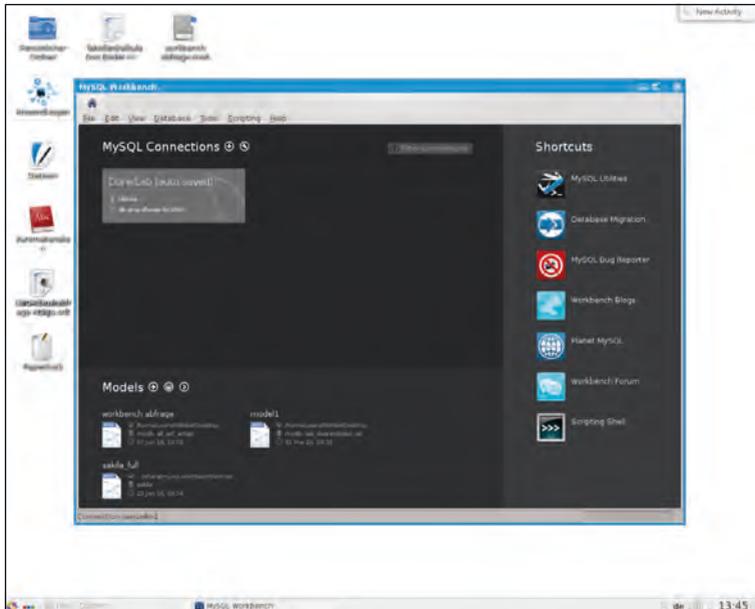


Abbildung 8: Screenshot des Desktops des DHVLab

Aus technischer Sicht ist das System folgendermaßen aufgebaut: Mittels *OVirt* werden neben dem Cloudspeicher virtuelle Maschinen organisiert, die auf verschiedene Aufgaben spezialisiert sind. Zu diesen Aufgaben gehört die Bereitstellung des Remote-Desktops (*CentOS* mittels *TigerVNC* und *Guacamole*) mit Anwendungen, des Datenbank- und Webservers (für den Blog und die Nutzerverwaltung), aber beispielsweise auch des *Shiny*-Servers oder des Dienstes *Jupyter Notebooks*. Auf diese Weise ist das System flexibel skalierbar und kann mit

weiteren Komponenten und Ressourcen je nach Erfordernis leicht erweitert werden.<sup>21</sup>

Praktisch ist ein Seminar auf dieser Basis wie ein Workshop organisiert. Das bedeutet, die Studierenden halten keine Referate, sondern leiten eine kleine Code-Academy, in der ihr eigener Bildschirminhalt per Beamer den anderen Studierenden sichtbar ist. Von der ersten Minute können dann Schritt für Schritt die vorgestellten Abläufe auf dem eigenen Rechner auf dem gleichen System nachvollzogen werden. Bei diesem Ansatz geht es darum, sowohl theoretisches Wissen zu erwerben, wie auch um die Möglichkeit, Erfahrungen mit dem Umgang mit Software zu machen. Dies dient auch der Vermeidung von Schwellenängsten, die weniger technikaffine Studierende vor neuen Software-Tools haben können. Wenn die vielfältigen Fehlerquellen einer Installation wegfallen und in einem solchen Workshop die Software zunächst spielerisch ausprobiert wird, werden die Studierenden an Themen und Konzepte leicht herangeführt. Haben sie ein solches Programm einmal verwendet und stoßen sie im Laufe des Studiums auf eine Forschungsfrage, mit digitalen Mitteln bearbeitet werden könnte, besteht so bereits eine Vorstellung vom Vorgehen und den Möglichkeiten der Datenverarbeitung und Visualisierung.

Nach den ersten erfolgreichen Einsätzen des *DHVLab* an der LMU München ist es nun das Ziel des Projektes, dieses System bayernweit zu expandieren und in weitere Einsatzszenarien einzubinden. Dazu werden interessierte Dozenten gesucht, die dieses System in ihre Lehre ebenfalls integrieren möchten.

Der oben skizzierte cloudbasierte Ansatz des *DHVLab* ist für die Zukunftsfähigkeit der Digitalen Kunstgeschichte von zentraler Bedeutung. Neue Anwendungen, die große Datenmengen und Rechenleistungen benötigen, zeichnen sich bereits ab. Beispielsweise wird Mustererkennung in großen Bilddatenbanken eine wichtige Rolle spielen,

---

<sup>21</sup> Detaillierte Informationen von Linus Kohl über die Systemarchitektur der Lehr- und Forschungsumgebung für die Ausbildung in den Geisteswissenschaften finden sich unter <https://dhvlab.gwi.uni-muenchen.de/Category:Architektur>.

um automatisch Vorzeichnungen Gemälden zuzuordnen oder ähnliche Gesten in anderen Bildern wiederzuerkennen.<sup>22</sup> Und gerade in Bereich der räumlichen Werke werden 3D-Technologien wie *Augmented* und *Virtual Reality* in der Forschung und Lehre zu Schlüsseltechnologien. Bereits jetzt werden in dem gemeinsamen Projekt der LMU München und der Bayerische Akademie der Wissenschaften „Corpus der barocken Deckenmalerei in Deutschland“ Bildwerke auf Wandoberflächen 3D-gescannt, wobei große Datenmengen entstehen.<sup>23</sup> Aber auch in der Rekonstruktion beispielsweise von Skulpturen werden zunehmend verlorene Elemente im Computer modelliert und mithilfe von 3D-Druckern ins Original eingesetzt.<sup>24</sup> Anwendungen wie diese benötigen eine neue Forschungsdaten-Infrastruktur, die auf die gesteigerten Anforderungen reagieren und es ermöglichen, die Digitale Kunstgeschichte auf einen neue Eben zu heben.

Bei der Digitalen Kunstgeschichte geht es aber nicht allein um technische Fragen, wie etwa um den Einsatz von IT für kunsthistorische Forschung und Lehre. Letztlich geht es auch darum, wie die historische und visuelle Kompetenz des Fachs ins digitale Zeitalter überführt wird, wie das Gegenwartsphänomen des digitalen Bildes zu bewerten ist<sup>25</sup> und schließlich, wie Geisteswissenschaftler die Digitalisierung der Welt urteilsfähig begleiten können. Die Kunstgeschichte war immer auch eine Wissenschaft, die sich mit der Veränderung des Bildes, seiner Medien und seiner Funktion auseinandergesetzt hat. Um die gegenwärtige digitale Transformation der Bildmedien hermeneutisch zu begleiten und historisch zu erklären, benötigt es eine Wissenschaft,

---

22 Ein solche Projekt findet sich beispielsweise bei Peter Bell u. Björn Ommer: Training Argus, in: Kunstchronik. Monatsschrift für Kunstwissenschaft, Museumswesen und Denkmalpflege, Band 68, 2015, S. 414--420

23 Weiteres zu diesem Projekt unter: Stephan Hoppe: Barocke Deckenmalerei in 3D, in: Akademie Aktuell, Ausgabe 02/2016, S. 65–75

24 Beispielhaft hierfür ist das Theatiner-Projekt der Technische Hochschule Deggendorf: Erdmann, Lisa (u.a.): Lukas aus der Asche. Auferstandenes Kulturerbe aus dem 3D-Labor, Lindenberg im Allgäu, 2016

25 Dazu hat das Institut für Kunstgeschichte der LMU München einen Antrag auf ein DFG-Schwerpunktprogramm mit dem Titel „Das digitale Bild“ gestellt. Mehr Informationen dazu auf der Webseite: <http://www.kunstgeschichte.uni-muenchen.de/forschung/digitalekg/digitales-bild1/>

die sowohl die Geschichte des Bildes bis zu zeitgenössischen Formen kennt, wie auch Methoden einsetzt und entwickelt, mit denen sie auf die digitale Bilderflut systematisch reagieren kann. Welches Fach könnte diese Aufgabe besser übernehmen als die Kunstgeschichte?

# DHLehre

## Integration von Forschung und universitärem Unterricht

Thomas Krefeld

### Forschungskommunikation

Die universitäre Lehre folgt – vom speziellen Fall der Didaktik abgesehen – der Forschung (und nicht etwa umgekehrt). Mit der Frage, was unter Forschung zu verstehen sei, brauchen wir uns nicht lange aufzuhalten. Einen guten Hinweis gibt bereits die Herkunft der englischen Bezeichnung *research*, die – wie jedes gute englische Wort – aus dem Fra. stammt, nämlich aus *recherche*, das wiederum vom Verb *rechercher* mit der ursprünglichen Bedeutung ‚wiederholt, dauerhaft suchen‘ abgeleitet wurde. In diesem Sinn lässt sich Forschung immer noch als Suche nach fehlenden, bzw. adäquaten Wissensbeständen umschreiben: Forschung sucht methodisch nachvollziehbare Antworten auf theoretisch begründete Fragen.

Der Drang nach Wissenvermehrung, d.h. nach Forschung ist zweifellos eine anthropologische Konstante, wie Michel de Montaigne (Bd. III, Kap. 13, *De l'expérience*) in Fortführung eines aristotelischen Prinzips<sup>1</sup> festgestellt hat:

„Il n'est desir plus naturel que le desir de cognoissance.  
(„Es gibt kein natürlicheres Verlangen als das nach Wissen.“)

Alles andere als konstant sind jedoch die Rahmenbedingungen der Forschung, denn sie sind massivem historischen Wandel unterworfen; dazu gehören die Institutionen, die Finanzierungswege und nicht

---

1 <<All men naturally desire knowledge.>> (Aristotle 1933, 980a22)

zuletzt die Formen der Kommunikation. Vor allem letztere wurden seit dem Aufkommen der Neuen Medien revolutioniert: Die Gutenberg-Galaxis (vgl. McLuhan 1962) strebt mit zunehmender, rasanter Geschwindigkeit auseinander, da die Gravitationskraft ihres Fixsterns, des gedruckten Buchs, immer schwächer wird; auch von der ehemaligen Leuchtkraft ist nurmehr ein matter Schimmer geblieben. In Forschungsprojekten, die mit Webtechnologie operieren, verläuft die Wissenschaftskommunikation jedenfalls nach vollkommen anderen Regeln (vgl. Krefeld 2017c), die hier nicht im Einzelnen ausbuchstabiert werden müssen. Es reicht im Hinblick auf die Lehre auf die vollkommen veränderte Realität der Publikation hinzuweisen. Die traditionelle Kommunikation von Forschungsergebnissen lässt sich wie folgt schematisieren:

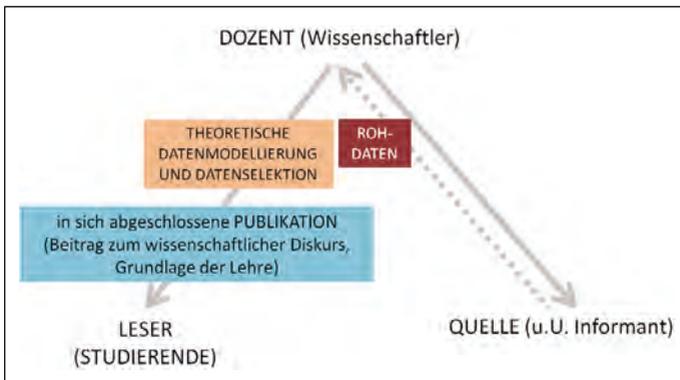


Abbildung 1: Forschungskommunikation auf der Grundlage gedruckter Texte

Publikationen, die für den Druck konzipiert werden, sind gewissermaßen dingliche, objektive Endprodukte eines Forschungsprozesses. Ihre (kommerzielle) Herstellung und Distribution liegt in der Hand Dritter, der Verlage, und ihre Form entspricht spezifischen Gattungen, wie zum Beispiel einem Wörterbuch, einer Abhandlung, einer Anthologie, einem Atlas oder auch einem Lehrbuch.

Unter den Bedingungen des Web lässt sich der wissenschaftliche Diskurs im Rahmen von freier Zugänglichkeit (*open access*) und offenen Quellcodes (*open source*) etwa wie folgt schematisieren:

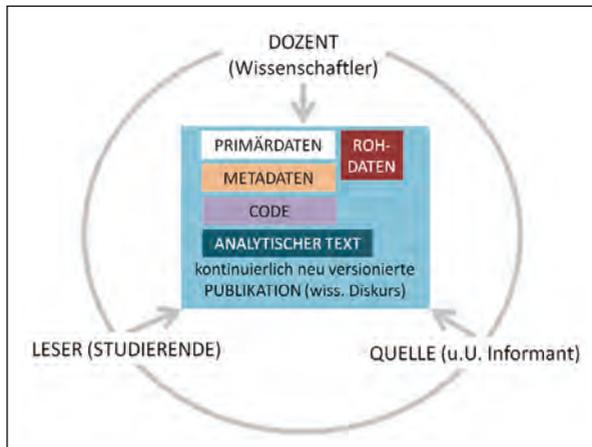


Abbildung 2: Webbasierte Forschungskommunikation

Unter diesen Bedingungen erweitert sich zunächst der Bereich des Publizierten, da außer dem analytischen Fließtext auch die Primär- und Metadaten also solche sowie speziell entwickelter Code veröffentlicht werden. Sodann sind die Gattungsgrenzen aufgehoben, allerdings ohne die sinnvolle Funktionalität der traditionellen Gattungen zu entwerfen. So sind im Fall des DFG-Projekts *Verba Alpina* zwar ein Wörterbuch (*Lexicon Alpinum*), eine Atlasdarstellung (*Interaktive Karte*) und analytischer wissenschaftlicher Text (*Methodologie*) als eigenständige Module angelegt; untereinander sind sie jedoch so verlinkt, dass von einer Funktion stets in eine andere gewechselt werden kann.<sup>2</sup>

Die teils beharrlichen Vorbehalte gegenüber Publikationen im Web werden durch ganz unterschiedliche Gründe getragen, die allesamt nicht substantiell, sondern akzidentiell zu sehen sind. Am schwersten wiegt das Argument der prekären Nachhaltigkeit (vgl. dazu Krefeld &

2 [https://www.verba-alpina.gwi.uni-muenchen.de/?page\\_id=162](https://www.verba-alpina.gwi.uni-muenchen.de/?page_id=162)

Lücke 2017), das durch ein Paket konzertierter Maßnahmen bewältigt werden kann:

- die Schaffung textstabiler und insofern dauerhaft zitierfähiger Versionen;
- die Auszeichnung mit Metadaten, die sowohl durch Suchmaschinen wie in Bibliothekskatalogen leicht auffindbar sind;
- die Ablage auf institutionell betriebenen Servern mit zuverlässiger Aussicht auf langfristige Verfügbarkeit (große Rechenzentren).

Ein Format, das diesen Ansprüchen genügen kann, wurde an der LMU in Zusammenarbeit mit der UB mittlerweile entwickelt. Es ist so angelegt, dass die jeweilige textstabile Version in der URL angezeigt wird und weiterhin durch die Nummer eines Abschnitts, auf den man sich bezieht, oder den man wörtlich zitiert, spezifiziert werden kann. Hier die empfohlene Zitierweise eines Artikels in der 7. Version (...v=7):

- Thomas Krefeld & Stephan Lücke (2017): Nachhaltigkeit – aus der Sicht virtueller Forschungsumgebungen. Korpus im Text. Version 7 (10.03.2017, 12:27). url: <http://www.kit.gwi.uni-muenchen.de/?p=5773&v=7>

sowie die URL seines ersten Abschnitts (...p:1):

- <http://www.kit.gwi.uni-muenchen.de/?p=5773&v=7 - p:1>

Der soeben zitierte Text wurde in unserer Publikationsplattform *Korpus im Text* (KiT) veröffentlicht, die nicht zuletzt für die Lehre relevant ist. Ein Lehrbuch, in das vielfältige Forschungserfahrungen eingeflossen sind, sowie mehrere wissenschaftliche Arbeiten und Artikel stehen hier bereits zur Verfügung; ein weiteres Lehrbuch, einige Dissertationen und andere, auch umfangreiche wissenschaftliche Vorhaben werden vorbereitet.<sup>3</sup> Darüber hinaus bietet auch DH-Lehre spezi-

---

3 [http://www.kit.gwi.uni-muenchen.de/?post\\_type=band](http://www.kit.gwi.uni-muenchen.de/?post_type=band)

fische Publikationsmöglichkeiten mit engem Bezug zur Lehre, nämlich für exzellente **Abschlussarbeiten** einerseits und für ausformulierte **Vorlesungen** andererseits.

## Ein neues Kooperationsprinzip

Im Hinblick auf die Lehre ist jedoch vor allem von Bedeutung, dass sich mit der webbasierten Forschungs- und Wissenschaftskommunikation neue Horizonte der akademischen Zusammenarbeit von lehrenden Forschern und Lernenden auftun. Ganz selbstverständlich ergeben sich Kooperationen zwischen Dozenten, bzw. Forschungsprojekten und Studierenden einerseits und zwischen Studierenden untereinander andererseits. Dazu drei Beispiele.

### 1. Kooperation zwischen Projekten und Studierenden

Im Rahmen einer Master-Arbeit wurden von Myriam Abenthum lexikologische Kommentare verfasst, die nach Revision durch die Projektleiter in das **Wörterbuchmodul** des Projekts *Verba Alpina* übernommen wurden.

### 2. Kooperation zwischen Doktoranden und BA-Studierenden zur Toolentwicklung

Etwa ein Drittel der sprachwissenschaftlichen DoktorandInnen unserer Graduiertenschule nimmt die Lehre der ITG in Anspruch und stützt ihre Dissertation auf relationale Datenbanken;<sup>4</sup> die Themenbereiche sind sehr unterschiedlich und die BetreuerInnen weit gestreut. Darunter ist ein ebenso innovatives wie methodisch richtungsweisendes Projekt von Sebastian Lasch zur Erfassung und Analyse von Graffiti im Stadtgebiet Roms. Bei der kartographischen Visualisierung der ca. 3000 Graffiti werden Instrumente zur Darstellung quantitativer Verhältnisse zum Einsatz kommen, die für *Verba Alpina* entwickelt wurden (insbesondere eine *heatmap*). Aber eine nicht geringe Schwierigkeit bei der Datenerfassung und -verarbeitung besteht darin, die unterschiedlichen Information so aufzunehmen, dass jedes Zeichen

---

4 <https://www.dh-lehre.gwi.uni-muenchen.de/?p=57328>

einzelnen mit spezifischen Metadaten und darüber hinaus auch im Kontext der Gesamtäußerungen und ihrer Abfolge wiedergegeben werden kann. Das folgende Foto zeigt:

- von einer Hand den grünen Text *Stalin c'è* („Stalin ist da“) mit zwei grünen Hammer & Sichel-Symbolen;
- von einer zweiten Hand zwei schwarze Zusätze, die den grünen Text negieren: *Stalin non c'è più* („Stalin ist nicht mehr da“);
- eventuell von einer weiteren Hand (?) die Übermalung der Hammer & Sichel-Symbole durch faschistische Kreis & Kreuz-Symbole;
- die Übermalung von *non* und *più* mit dem offenkundigen Ziel sie unkenntlich zu machen und die Negation aufzuheben;

dem Anschein nach zuletzt angebrachte weiße Übermalungen, ein weißes faschistisches Symbol über dem schwarz übermalten, ursprünglich grünen sowie eine unleserliche weiße Schrift über dem Namen Stalin.



Abbildung 3: Eine Wand mit Graffiti in Rom (Foto Sebastian Lasch)

Eine direkte Aufnahme dieser Daten in das relationale Modell, mit dem die ITG ausnahmslos arbeitet erschien möglich, jedoch reichlich umständlich. Just als diese Überlegungen angestellt wurden, tauchte ein Studierender der Informatik an unserem Institut (der Romanistik) auf, Daniel Pollithy, und bekundete sein Interesse, in seiner informatischen BA-Arbeit etwas zu entwickeln, das sich auf sprachliche Daten anwenden ließe. Ich skizzierte ihm das Problem und in Abstimmung mit der ITG und engem Austausch mit dem Doktoranden entstand ein außerordentlich nützliches Tool, das eine direkte Erfassung der Primärdaten und die Zuordnung von Metadaten in beliebige vielen Layern und ausgehend vom Foto selbst gestattet. Einen Eindruck der Bearbeitung (aus der Alpha-Version) gibt der folgende Screenshot:



Abbildung 4: Erfassung semiotischer Informationen auf Basis einer Photographie (Sebastian Lasch | Daniel Pollithy)

Das Datenschema zeigt, dass bei der Erfassung die syntaktische Abfolge (FOLGT) unterschiedlicher Zeichentypen (TOKEN, SYMBOL), die zeitliche Abfolge (ÜBERDECKT), die Assoziation mit semantischen bzw. kognitiven Feldern (AKTIVIERT) sowie diverse semiotische und linguistische Metadaten berücksichtigt werden.

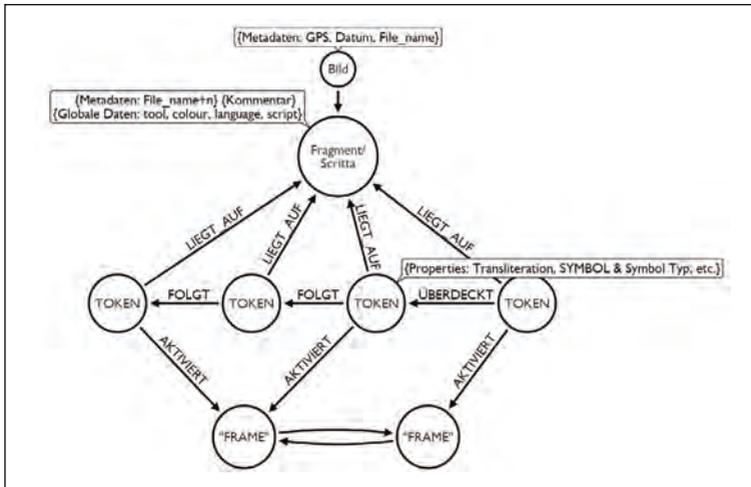


Abbildung 5: Datenmodell für die Erfassung von Graffiti  
(Sebastian Lasch | Daniel Pollithy)

Eine syntaktische spezifizierte Verknüpfung (,follows‘) zeigt der folgende Screenshot:

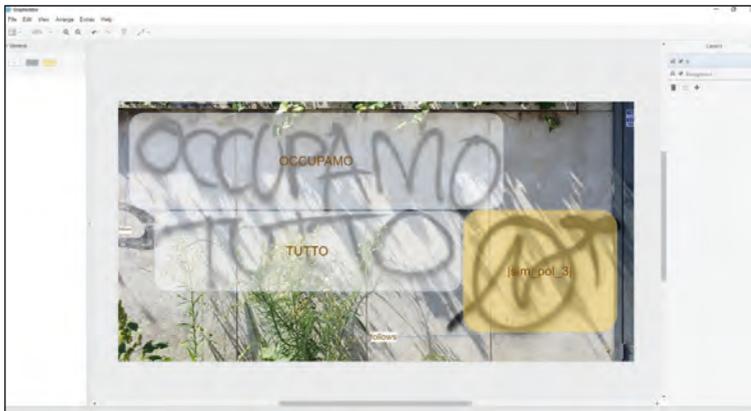


Abbildung 6: Syntaktische Relation zwischen semiotischen Einheiten  
(Sebastian Lasch | Daniel Pollithy)

Für das präsentierte Tool wurde zwar ein *graph database* verwandt (die Community-Edition der Neo4j GraphDB); es gestattet jedoch den Export der erfassten Daten in das übliche relationale Modell (MySQL) der ITG. Nach Fertigstellung der BA-Arbeit wird es dem DHVLab zur Verfügung gestellt und sich dort sicherlich als sehr nützlich für editorische, lexigraphische und andere Projekte erweisen.

### Kooperation von Forschung Lehre im Bereich der Korpora und Repositorien

Ein letztes Beispiel illustriert eine gleichermaßen strategische wie strukturelle Form der Kooperation unserer Lehr- und Forschungsinfrastruktur. Sie besteht darin, Korpora, die für Dissertationsprojekte konzipiert, erhoben und strukturiert werden, nach Abschluss der Dissertation der Lehre (und natürlich weitere Forschung) zur Verfügung zu stellen. Ein prototypisches Beispiel ist das italienische WhatsApp-Korpus, das Katharina Jakob im Rahmen einer Untersuchung über *Medienbedingte Variation am Beispiel italienischer WhatsApp-Chats* aufgestellt und mit einem soliden Datenbestand gefüllt hat, wie der folgende Screenshot zeigt:

Zeige Datensätze 0 - 99 (43583 insgesamt, Die Abfrage dauerte 0.0012 Sekunden.)

SELECT \* FROM `whatsapp\_chats`

1 > >> Anzahl der Datensätze: 100 Zeilen filtern: Diese Tabelle durchsuchen Nach Schlüssel sortieren

+ Optionen						
	msg_id	ora	data	alias	testo	Quelle wordcount
	1	12:08	25/set/2014	alessandra	La iellatura continua	chat001 3
	2	12:08	25/set/2014	alessandra	Oggi esco di casa e uscendo mi porto l'immondizia...	chat001 29
	3	12:08	25/set/2014	alessandra		chat001 0
	4	12:40	25/set/2014	Zoe	noooo	chat001 1
	5	12:40	25/set/2014	Zoe	<Media omesso>	chat001 2
	6	12:44	25/set/2014	alessandra	Una lingua sconosciuta	chat001 3
	7	12:45	25/set/2014	Zoe	<Media omesso>	chat001 2
	8	12:46	25/set/2014	Zoe	passagui significa una che passa i guai	chat001 7
	9	12:46	25/set/2014	alessandra	Sepp...	chat001 4

Abbildung 7: Allgemeine Ansicht der Datensätze im WhatsApp-Korpus von Katharina Jakob (Screenshot)

Diese spezifische Art des aktuellen Sprachgebrauchs zeichnet sich durch etliche, erst ansatzweise analysierte Besonderheiten aus, wie z.B. durch das systematische Fehlen des Apostrophs nach der Kurzform des definiten Artikels vor vokalischem beginnendem Folgewort: *il/la > l* anstatt der Standardform *l'*; das folgende Bild zeigt die Abfrage und den Beginn der Trefferliste.

The screenshot shows a phpMyAdmin interface with a SQL query box containing the following query:

```
SELECT * FROM whatsapp_chats WHERE (text LIKE '%il/la%' OR text LIKE '%l%' OR text LIKE '%la%' OR text LIKE '%l%');
```

The results table displays the following data:

msg_id	ora	data	alias	testo	Quelle	wordcount
3897	06:55	01dic/2014	Diletta	Bella l'idea	chat006	3
3944	20:45	03dic/2014	Diletta	ora mi sto scaldando spinaci l'unica cosa che	chat006	15
4269	23:43	11dic/2014	Diletta	Domani sera pensavo di fare un bar taur con la [no]	chat006	58
4612	18:42	10dic/2014	Diletta	Mi sn confusa con l'altro luc collega	chat006	8
4663	21:19	10dic/2014	Diletta	Come va l'aperitivo? Ho fatto un insalata e una b...	chat006	17
4693	19:31	11dic/2014	Diletta	L'artista potrebbe essere ubriaco. Guarda come una	chat006	12
4752	20:38	12dic/2014	Diletta	Si paga l'entrata al [vugo] stase quindi?	chat006	8
4841	17:41	13dic/2014	Diletta	Siamo nel duomo ho trovato per tema dove esce l'a	chat006	22
5134	14:24	16dic/2014	Diletta	E stato il unico tramezzino che c'era...messa chiu.	chat006	11
5258	02:28	21dic/2014	Diletta	Quindi se mi riesci a chiamarmi prova l'altro nume	chat006	10
5389	13:10	24dic/2014	Diletta	Dopo facciamo l'albero e ti mando una foto	chat006	9
5473	21:52	24dic/2014	Diletta	Ti piace l'albero??	chat006	4
5721	20:40	26dic/2014	Diletta	Come va l'ape tesoro?	chat006	5
6334	19:00	07gen/2015	Diletta	Ho mi sn stanca anche se mi ho dormito bene. Sopra	chat006	25
7441	00:23	15ott/2014	Matilda	lo nemmeno te l'aspetto	chat011	5
7476	00:35	15ott/2014	Matilda	ma spesso il odia ke noi abbiamo e diversa	chat011	5
8078	17:17	21ott/2014	Isabella	te piu ci e stata l'illusione	chat016	7
8137	11:24	25ott/2014	Isabella	Non ci sn soldi in nero. niente per l'a guardia di	chat016	13
8280	11:58	08nov/2014	Isabella	Pier l'Italia?	chat016	3
8575	10:21	14gen/2015	Isabella	A dopo, mi vedi l'ora?	chat016	6
8500	10:54	14gen/2015	Isabella	Nin ho l'ombrello...	chat016	4

Abbildung 8: WhatsApp-Korpus – Ausschnitt aus Trefferliste bei Abfrage von fehlendem Apostroph bei bestimmtem Artikel (Katharina Jakob)

Das Korpus dokumentiert einen attraktiven Bereich der sprachwissenschaftlichen Lehre; es wird zweifellos schnell zum Gegenstand anderer studentischer Arbeiten werden, die es dann nicht nur verwenden,

sondern gleichzeitig weiter ausbauen und womöglich differenzierter strukturieren sollen.

## Forschungskommunikation dient der Aufklärung

Digitale Forschungskommunikation dient also, genauer gesagt, der digitalen Aufklärung, dem eigentlichen Unterrichtsziel unseres Studiengangs. Frei nach der berühmten Definition des preußischen Aufklärers Immanuel Kant<sup>5</sup> ist digitale Aufklärung der Ausgang des Geisteswissenschaftlers aus seiner selbstverschuldeten digitalen Unorganisiertheit. Unorganisiertheit ist das Unvermögen, sich seiner eigenen Forschungsdaten ohne Abstimmung mit den Daten anderer zu bedienen. Selbstverschuldet ist diese Unorganisiertheit, wenn die Ursache derselben nicht am Mangel der Forschungsdaten, sondern der Bereitschaft liegt, sich seiner eigenen Daten ohne Abstimmung mit den Daten der Kommilitonen und Fachkollegen zu bedienen.

## Literatur

VerbaAlpina: <https://www.verba-alpina.gwi.uni-muenchen.de/>

Aristotle (1933): *Metaphysics*, Volume I. Translated by Hugh Tredennick, Loeb Classical Library 271. Harvard.

Dylan, Bob (1965): *Ballad of a thin man*, Highway 61 revisited. <https://bobdylan.com/songs/ballad-thin-man/> (04.01.2018).

Kant, Immanuel (1784): *Beantwortung der Frage: Was ist Aufklärung?*, *Berlinische Monatsschrift*, 481–494. [https://www.rosalux.de/fileadmin/rls\\_uploads/pdfs/159\\_kant.pdf](https://www.rosalux.de/fileadmin/rls_uploads/pdfs/159_kant.pdf) (04.01.2018).

Kant, Immanuel (1784): *Beantwortung der Frage: Was ist Aufklärung?*, *Berlinische Monatsschrift*, 481–494. [https://www.rosalux.de/fileadmin/rls\\_uploads/pdfs/159\\_kant.pdf](https://www.rosalux.de/fileadmin/rls_uploads/pdfs/159_kant.pdf) (04.01.2018).

---

5 Die Vorlage des folgenden Passus lautet im Original: «Aufklärung ist der Ausgang des Menschen aus seiner selbstverschuldeten Unmündigkeit. Unmündigkeit ist das Unvermögen, sich seines Verstandes ohne Leitung eines anderen zu bedienen. Selbstverschuldet ist diese Unmündigkeit, wenn die Ursache derselben nicht am Mangel des Verstandes, sondern der Entschließung und des Mutes liegt, sich seiner ohne Leitung eines andern zu bedienen.» (Kant 1784, 481)

- Krefeld, Thomas (2017b): Kooperation. In: VerbaAlpina-de 17/1, Methodologie. [https://www.verba-alpina.gwi.uni-muenchen.de/?page\\_id=493&letter=K&db=171-22](https://www.verba-alpina.gwi.uni-muenchen.de/?page_id=493&letter=K&db=171-22) (04.01.2018).
- Krefeld, Thomas (2017c): Wissenschaftskommunikation im Web. In: VerbaAlpina-de 17/1, Methodologie. [https://www.verba-alpina.gwi.uni-muenchen.de/?page\\_id=493&db=171&letter=W-62](https://www.verba-alpina.gwi.uni-muenchen.de/?page_id=493&db=171&letter=W-62) (04.01.2018).
- Krefeld, Thomas & Lücke, Stephan (2017): Nachhaltigkeit – aus der Sicht virtueller Forschungsumgebungen, Korpus im Text. <http://www.kit.gwi.uni-muenchen.de/?p=5773&cv=7> (04.01.2018).
- Lücke, Stephan (2017): Digitalisierung. In: VerbaAlpina-de 17/1, Methodologie. [https://www.verba-alpina.gwi.uni-muenchen.de/?page\\_id=493&db=171&letter=D-15](https://www.verba-alpina.gwi.uni-muenchen.de/?page_id=493&db=171&letter=D-15) (04.01.2018).
- McLuhan, Marshall (1962): The Gutenberg Galaxy. London.
- Montaigne, Michel de (1595): Les Essais, <http://www.bribes.org/trismegiste/montable.htm> (04.01.2018).

# Das DHVLab Analysecenter

## Ein Tool zur datengestützten Forschung in den Geisteswissenschaften

Stefanie Schneider

Digitale Forschungswerkzeuge breiten sich auf mannigfaltige Weise in den Geisteswissenschaften aus; seien es nun webbasierte geografische Informationssysteme oder digitale historisch-kritische Editionen.<sup>1</sup> Nichtsdestotrotz werden teils ausgefeilte und nützliche *Tools* kaum eingesetzt.<sup>2</sup> Eine Studie, in der Fred Gibbs und Trevor Owens überwiegend US-amerikanische Historiker befragten, welche Bedürfnisse *Tools* zu erfüllen haben, nannte dafür zwei wesentliche Gründe: Erstens gehen viele *Tools* nicht auf den traditionellen geisteswissenschaftlichen Anwender ein und stellen keine einfachen und selbsterklärenden Oberflächen bereit. Zweitens mangelt es ihnen häufig an Anleitungen mit praxisnahen Beispielen und an einer Dokumentation, die den methodologischen Wert des *Tools*, und seine Limitierungen, mit nichttechnischem Vokabular beschreibt.<sup>3</sup>

Das im Folgenden präsentierte *Analysecenter*<sup>4</sup> nimmt sich dieser Kritikpunkte an. Auf der einen Seite stellt es eine prädefinierte, flexible

- 1 *Digital Research Tools (DiRT)* bietet eine Übersicht verschiedenster Werkzeuge in den unterschiedlichsten Kategorien, beispielsweise „Analyze relationships between pieces of data“ und „Transcribe audio, video or manuscripts“. Siehe <http://dirtdirectory.org/> (09.08.2017).
- 2 „Only about six percent of humanist scholars go beyond general purpose information technology and use digital resources and more complex digital tools in their scholarship.“ Siehe Summit on Digital Tools in the Humanities: A Report on the Summit on Digital Tools, Charlottesville 2005, S. 4.
- 3 Siehe Gibbs, Fred und Owens, Trevor: Building Better Digital Humanities Tools: Toward Broader Audiences and User-Centered Designs, in: Digital Humanities Quarterly 6.2, 2012, <http://www.digitalhumanities.org/dhq/vol/6/2/000136/000136.html> (09.08.2017). Gibbs und Owens schreiben weiterhin (Abschnitt 7): „Tool interfaces must help more traditional historians feel more comfortable with new ways of visualizing, analyzing, and thinking about sources and about data.“
- 4 <http://dhvlab.gwi.uni-muenchen.de/app/analysis/> (09.08.2017).

Werkzeugpalette zur Verfügung, um einen Gegenstand, die *Small* bis *Large* oder *Biggish Data*,<sup>5</sup> mit einer intuitiven grafischen Benutzeroberfläche zu untersuchen und die Ergebnisse dieser Untersuchung, Grafiken und Tabellen, komfortabel zu exportieren. Auf der anderen Seite ist es als Modul in das *Digital Humanities Virtual Laboratory*<sup>6</sup> integriert und nutzt die dort bestehende Infrastruktur – darunter Manuale, die in die Statistik und Informatik mit konkreten, aus den Geisteswissenschaften stammenden Anwendungsfällen einführen –, auch in der Lehre.<sup>7</sup> Das *Analysecenter* hat zwei Ziele: erstens eine zuvor definierte Fragestellung zu beantworten, zweitens eine in diesem Zuge neue Fragestellung zu gewinnen; das Bekannte zu untermauern und das Unbekannte zu explorieren – mithilfe der Empirie, sowohl auf Mikro- als auch auf Makroebene.<sup>8</sup> Die Programmiersprache *R*<sup>9</sup>, das auf *R* aufsetzende Webapplikationspaket *shiny*<sup>10</sup> sowie die Erweiterungen *shinydashboard*<sup>11</sup> und *shinyjs*<sup>12</sup> bilden dafür das technische *Framework*. Das sogenannte *tidyverse*<sup>13</sup> unterstützt ferner einen kohärenten *Workflow* im *Backend*; es organisiert, strukturiert, transformiert und manipuliert den Gegenstand. Als *Data* des momentanen Prototyps fungieren rund sieben Millionen *Taggings* der kunsthistori-

- 5 Da *Big Data* nochmals andere Herausforderungen mit sich bringt, die das *Analysecenter* nicht berücksichtigt, wird der Begriff bewusst unterlassen.
- 6 Eine „digitale Lehr- und Forschungsinfrastruktur für die Datenanalyse in den Kunst-, Geschichts- und Sprachwissenschaften“ der Ludwig-Maximilians-Universität München; siehe die weiteren Beiträge dieses Bandes und ferner <http://dhvlab.gwi.uni-muenchen.de/> (09.08.2017).
- 7 Das *Analysecenter* war unter anderem Gegenstand der Seminare „Big Data in der Kunstgeschichte“ im Wintersemester 2016/2017 und „Automatische und manuelle Bildadressierung in der Kunstgeschichte“ im Sommersemester 2017 an der Ludwig-Maximilians-Universität München (beide Hubertus Kohle und Stefanie Schneider).
- 8 Dazu auch Weaver, Warren: Science and Complexity, in: American Scientist 36, 1948, S. 536–544.
- 9 R Core Team: R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, 2017, <https://www.r-project.org/> (09.08.2017).
- 10 Chang, Winston et al.: shiny: Web Application Framework for R, 1.0.3, 2017, <https://cran.r-project.org/package=shiny> (09.08.2017).
- 11 Chang, Winston und Borges Ribeiro, Barbara: shinydashboard: Create Dashboards With Shiny, 0.6.1, 2017, <https://cran.r-project.org/package=shinydashboard> (09.08.2017).
- 12 Attali, Dean: shinyjs: Easily Improve the User Experience of Your Shiny Apps in Seconds, 0.9.1, 2017, <https://cran.r-project.org/package=shinyjs> (09.08.2017).
- 13 Wickham, Hadley: tidyverse: Easily Install and Load Tidyverse, 1.1.1, 2017, <https://cran.r-project.org/package=tidyverse> (09.08.2017).

schen Spieleplattform *ARTigo*<sup>14</sup>, die Spielern digitale Reproduktionen von Kunstwerken präsentiert, um diese zu annotieren.<sup>15</sup> Die *Taggings* eines konkret interessierenden Werks werden in den nächsten Absätzen näher betrachtet, um einige ausgewählte Funktionen des *Analysecenters* vorzustellen.

Wenden wir uns exemplarisch Tizians „Venus mit Cupido, Orgelspieler und Hündchen“ (Abb. 1) und den diesem Werk ähnlichen Werken zu. Der erste Reiter des *Analysecenters*, Übersicht, zeigt links alle Künstler und die Anzahl ihrer in *ARTigo* annotierten Kunstwerke (Abb. 2). Tizian findet sich auf einer hinteren Seite, und seine „Venus mit Cupido, Orgelspieler und Hündchen“ schließlich im mittleren *Container*, wenn links die Zeile angewählt wird, die das *Label* „Titian“ trägt.<sup>16</sup> Ein Klick auf „Venus mit Cupido, Orgelspieler und Hündchen“ gibt rechts in einem weiteren *Container* die *Tags* und ihre absolute Häufigkeit, die sogenannte *Suchworddichte*, frei: darunter zunächst die elementaren Schlagwörter „Frau“, „Hund“, „Mann“ und „Orgel“, die jeweils viermal annotiert wurden, doch danach ebenso die aus kunsthistorischer Perspektive spezifischeren *Tags* „Putte“, zweimal, sowie „Venus“ und „Renaissance“, jeweils einmal (Abb. 3). Die quantitative Analyse beginnt im *Browsing*, in einer deskriptiven *Tour d’Horizon* durch den Gegenstand, der ihm anhaftenden Variablen und deren Ausprägungen.<sup>17</sup>

Rücken wir näher an die Fragestellung. Der dritte Reiter des *Analysecenters*, Ähnlichkeitsanalysen, fächert zwei Unterreiter, *1-Alle-Vergleich* und *Alle-Alle-Vergleich*, auf. Zugrunde liegt ihnen das im *Information Retrieval* populäre Modell, einen Text zu *tokenisieren*, ihn also in seine Bestandteile, Wörter – oder allgemeiner Terme –, zu zerlegen, und diese derart zu sortieren und zu indexieren, dass ein Vokabular

<sup>14</sup> <http://www.artigo.org/> (09.08.2017).

<sup>15</sup> Für eine detaillierte Beschreibung siehe Kohle, Hubertus: Kunstgeschichte Goes Social Media, in: *Aviso: Zeitschrift für Wissenschaft und Kunst in Bayern* 3, 2011, S. 38–43.

<sup>16</sup> Die alternative Schreibweise, Titian statt Tizian, ist zu beachten und in der Such- und Filtermaske über einen *regulären Ausdruck*, beispielsweise „Ti(z|t)ian“, abzufangen.

<sup>17</sup> Dazu Chang, Shan-Ju und Rice, Ronald E.: *Browsing: A Multidimensional Framework*, in: *Annual Review of Information Science and Technology* 28, 1993, S. 231–276.

entsteht, und mit ihm eine mathematische Struktur, die sich für die quantitative Analyse eignet.<sup>18</sup> Tizians „Venus“ wandelt sich auf diese Weise in einen Vektor, der – repräsentiere er lediglich die zuvor spezifizierten sieben Tags – durch

$$r = (w_1, w_2, w_3, w_4, w_5, w_6, w_7) = (4, 4, 4, 4, 2, 1, 1)$$

definiert wird, wobei  $r$  die sogenannte *Ressource* bezeichne, Tizians „Venus mit Cupido, Orgelspieler und Hündchen“, und  $w_1$  bis  $w_7$  die *Suchworddichte* der über Indizes referenzierten sieben *Tags* wiedergebe, „Frau“, „Hund“, „Mann“, „Orgel“, „Putte“, „Venus“ und „Renaissance“, aus denen sich die *Ressource* im mathematischen Sinne bildet. Dieses Modell ist aus zwei Gründen praktikabel: Zum einen spannt ein Vektor einen Raum auf,<sup>19</sup> der weitere Vektoren enthalten kann; ein Beispiel ist Redons „Geburt der Venus“. Zum anderen liegen nicht nur ähnlich konstellierte *Ressourcen*<sup>20</sup> in diesem euklidischen Raum nahe beieinander, sondern es ist auch ihre Nähe zu messen, und zwar über den Kosinus des zwischen ihnen liegenden Winkels. Hier setzt die uns interessierende Ähnlichkeit an.

Selektieren wir nun Tizians „Venus mit Cupido, Orgelspieler und Hündchen“ im Reiter *1-Alle-Vergleich*.<sup>21</sup> Die Resultate der ersten Seite (Abb. 4) illustrieren: Zuvorderst erscheint eine im *Museo del Prado* ausgestellte Wiederholung der „Venus“ – noch mit Orgel und Orgel-

18 Salton, Gerard, Wong, Anita und Yang, Chung-Shu: A Vector Space Model for Automatic Indexing, in: Communications of the ACM 18.11, 1993, S. 613.

19 Die mathematisch präzisere, doch weniger anschauliche Formulierung lautet: Ein Vektor bildet eine Basis eines Raums.

20 Ähnlich konstellierte meint in diesem Beispiel, aus vielen gleichen *Tags* zusammengesetzt, die ähnlich häufig annotiert wurden.

21 Unter *Lokale Gewichtung* wurde das Feld *Suchworddichte*, unter *Globale Gewichtung* das Feld *Keine Gewichtung* und unter *Ähnlichkeitsmaß* das Feld *Kosinus-Ähnlichkeit* markiert. Weiterhin wurde die Anzahl der *Tags*, die beide Kunstwerke, das prädefinierte feste und das mit jeder Iteration wechselnde, gemein haben müssen, auf zehn reduziert, um die Menge der möglichen Ergebnisse zu erhöhen. Eine Einführung in die unterschiedlichen Maße, einen Term, auch in Abhängigkeit anderer Terme, zu gewichten, liefern Salton, Gerard und Buckley, Christopher: Term Weighting Approaches in Automatic Text Retrieval, in: Information Processing and Management 24.5, 1988, S. 513–523.



Abbildung 1: Tizian, „Venus mit Cupido, Orgelspieler und Hündchen“, um 1550, Gemäldegalerie Berlin. Gemeinfrei.

 The screenshot shows the 'Übersicht' (Overview) page of the DHV Lab Analysecenter. The interface is dark-themed with a sidebar on the left containing navigation options like 'Deskriptive Analysen', 'Ästhetik- und Stilanalysen', 'Kontextanalyse', 'Direkte Bildadressierung', and 'Virtueller Raum'. The main content area features three summary cards at the top: 'Künstler' (18.442), 'Kunstwerke' (50.404), and 'Tagungen' (6.698.643). Below these is a table listing artists and their associated artwork counts.
 

Name	Kunstwerke
Turner, Joseph Mallord William	368
Dieter, Adrechi	338
Lüthi, Silo	315
Marc, Franz	303
Bombardieri van Rijn	290
Bellini, Gian Lorenzo	289
Sommer, Tobias	286
Cézanne, Paul	282
O'Keefe, Georgia	244
Chapman, John	244

 At the bottom of the table, there is a search bar and a checkbox labeled 'Nach Schlagwortern filtern'.

Abbildung 2: Reiter *Übersicht* des Analysecenters. Der Container links zeigt alle Künstler und die Anzahl der Kunstwerke des Künstlers, die in ARTigo annotiert wurden.

**DHV.ab**

Übersicht

- Deskriptive Analyse
- Ähnlichkeitsanalyse
- Kombinationsanalyse
- Direkte Bildadressierung
- Information

**18.442** Künstler

**50.404** Kunstwerke

**6.698.643** Schlagwörter

Name	Kunstwerke	Titel	Schlagwörter	Schlagwort	Anzahl
Tizian	23	Ländliches Konzert	350	Flucht %	5
Tizian	23	Actaeon und Diana	345	Frau %	4
Tizian	23	Parazzo di donna (La Favorita)	94	Hand %	4
Tizian	23	ecce Homo	88	Mann %	4
Tizian	23	Venus und Adonis	86	Vorhang %	4
Tizian	23	Venus mit Cupido, Orgelspieler und Hündchen	80	Engel %	4
Tizian	23	Doppelbildnis einer Dame mit ihrer Tochter	78	Wald %	3
Tizian	23	Venus mit Orgelspieler und Hündchen (Replik)	77	Landschaft %	3
Tizian	23	Lucretia à venere (G. Anon.)	71	Engel %	3
Tizian	23	Adam und Eva	71	Lager %	2

Abbildung 3: Reiter *Übersicht* des *Analysecenters*. Im linken *Container* wurde nun Tizian selektiert, im mittleren seine „Venus mit Cupido, Orgelspieler und Hündchen“. Der *Container* rechts zeigt alle für dieses Werk bislang hinterlegten *Tags*.

**DHV.ab**

Übersicht

- Deskriptive Analyse
- Ähnlichkeitsanalyse
- 1-Alle-Vergleich
- Alle-Alle-Vergleich
- Kombinationsanalyse
- Direkte Bildadressierung
- Information

**18.442** Künstler

**50.404** Kunstwerke

**6.698.643** Schlagwörter

**1-Alle-Vergleich**

Alle-Alle-Vergleich

Kombinationsanalyse

Direkte Bildadressierung

Information

**Venus mit Cupido, Orgelspieler und Hündchen**

Tizian (1550)

Tags: 80 (gesamt), 84 (verschieden)

Links: Artikel · Bild

**Bitte Kunstwerk auswählen**

Kleine Kärtchen anzeigen

Titel	Name	Jahr	G.Tags	Ähnl. %
Venus mit Orgelspieler und Hündchen (Replik)	Tizian	1550	30	85,08 %
Jupiter und Antiope	Caracci, Annibali	1592	11	83,22 %
Diana	Tizian	1544	16	81,78 %
Joseph und das Weib des Potifar	Mieris, Willem van	1700	15	81,73 %
Alexis und Pygmalion	David, Jacques-Louis	1817	20	80,26 %
Apelles malt Lampião	Winghe, Joos van	1595	15	79,80 %
Herkules und Olympio	Sprünker, Bartholomaeus	1530	12	49,75 %
Bildnis eines Chepari	Renaix, Giovanni	1640	13	49,57 %
Israel des Reich	Wastwey, Jean-Antoine	1738	15	49,41 %
Das homerische Laoköter	Corroth, Louis	1809	12	48,19 %

Abbildung 4: Reiter *1-Alle-Vergleich* des *Analysecenters*. Selektiert wurde Tizians „Venus mit Cupido, Orgelspieler und Hündchen“.



Abbildung 5: Tizian, „Venus mit Orgelspieler und Hündchen“, um 1550, Museo del Prado. Gemeinfrei. Vergleiche dazu auch Abb. 1.



Abbildung 6: Jacques Louis David, „Amor und Psyche“, um 1817, Cleveland Museum of Art (links), Lovis Corinth, „Das homerische Gelächter“, um 1909, Neue Pinakothek München (rechts). Beide gemeinfrei.

spieler, doch ohne Cupido und mit anderem Schoßhündchen (Abb. 5) –, die zweifellos auch unter traditionellen hermeneutischen Kriterien als zu der in der *Gemäldegalerie Berlin* ausgestellten „Venus“ ähnlich deklariert werden würde. Bereits an Position fünf und zehn folgen allerdings Davids „Amor und Psyche“ und Corinths „Das homerische Gelächter“, die allein ob ihres Titels, und, mehr noch, ihrer Datierung – „Amor und Psyche“ entstand um 1817, „Das homerische Gelächter“ um 1909 –, aus dem Raster des Naheliegenden fallen. Ein Blick auf die als sogenannte *Ausreißer* deklarierten Werke legitimiert ihr Auf-

treten teilweise. Beide zeigen ein nacktes Paar, Mann und Frau, auf einer Schlafstätte; die Venus nur zeigen sie beide nicht: Bei David ist es die schlafende Psyche, bei Corinth die in flagranti ertappte Aphrodite (Abb. 6). Weder Psyche noch Aphrodite werden jedoch, wenn es nach den Spielern von *ARTigo* geht, im bedeutsamen Maße erkannt. Nur ein *Tagging* weist „Psyche“ direkt aus, während sich „Aphrodite“ überhaupt nicht unter den hinterlegten *Taggings* findet.

Der durch die quantitative Analyse suggerierte Objektivismus ist kontextueller Natur.<sup>22</sup> Obgleich *gegeben*, interagiert der Gegenstand, Tizians „Venus mit Cupido, Orgelspieler und Hündchen“, mit einem Subjekt, das ihn erst als *gegeben* konstituiert. Da es sich bei *ARTigo* um ein Spiel handelt, ist hier der Spieler das Subjekt, und der Spieler ordnet sich den Regeln des Spiels unter, um es möglichst erfolgreich zu absolvieren: Die Punkte, die er für die Eingabe eines gültigen, weil in einer anderen als der aktuellen Spielrunde annotierten *Tags* erhält,<sup>23</sup> beeinflussen die *Tags*, die er tätigt. Hierbei entsteht ein *Bias*. Wissen wir jedoch um ihn, deuten wir Davids „Amor und Psyche“ und Corinths „Das homerische Gelächter“ als Produkte des sie generierenden Prozesses und als Produkte des Algorithmus, der sie transformiert. Der hier den Gegenstand konstituierende Prozess leitet den Spieler an, *Tags* „trivialerer“ Natur zu annotieren; schließlich birgt „Liege“ eine höhere Wahrscheinlichkeit mit Punkten belohnt zu werden als beispielsweise „Chaiselongue“. Der hier den Gegenstand transformierende Algorithmus wiederum sieht die einzelnen *Tags* als unabhängig zueinander. Eine Chaiselongue ist im semantischen Sinne aber eine Liege, und als solche sind sich auch die *Tags* „Chaiselongue“ und „Liege“ ähnlicher als beispielsweise die *Tags* „Chaiselongue“ und „Horizont“; zumindest aus der Perspektive eines Menschen. Der in diesem Beispiel verwendete Algorithmus weiß darum nicht: Für ihn sind sich alle *Tags* in gleicher

22 Lisa Gitelman und Virginia Jackson bemerken ebenso: „The point is not how to judge whether objectivity is possible [...] but how to describe objectivity in the first place.“ (Gitelman, Lisa und Jackson, Virginia: Introduction, in: Gitelman, Lisa (Hrsg.): „Raw Data“ is an Oxymoron, Cambridge 2013, S. 4.)

23 Das Prinzip des *Matching* dämmt Missbrauch ein. Ein einmal in Edvard Munchs Porträt des August Strindberg annotiertes „Micky Maus“ etwa wird selten ein zweites Mal annotiert werden.

Weise unähnlich.<sup>24</sup> Ihn gilt es ebenso zu reflektieren wie den Gegenstand, dem er sich nähert.<sup>25</sup>

Deskription, Exploration und Reflexion rahmen den multiperspektivischen *Workflow* des *Analysecenters*. Dieser interaktive und iterative Prozess des Entdeckens ist essenziell, beide Ziele – eine zuvor definierte Fragestellung zu beantworten und eine in diesem Zuge neue Fragestellung zu gewinnen – zu realisieren, und ist damit gleichwohl in die universitäre Lehre zu integrieren. Ein *Tool* wie das hier anhand einer exemplarischen Fragestellung umrissene *Analysecenter* steht letztlich niemals entkoppelt: Die Algorithmen, die es implementiert, sind zwar nicht mehr eigenhändig zu programmieren. Auf theoretischer Ebene zu verstehen sind sie aber dennoch, um die aus der Arbeit mit dem *Tool* resultierenden Ergebnisse entsprechend, auch auf qualitativer Ebene, interpretieren und mögliche Limitierungen einschätzen zu können.

## Literatur

- Attali, Dean: shinyjs: Easily Improve the User Experience of Your Shiny Apps in Seconds, 0.9.1, 2017, <https://cran.r-project.org/package=shinyjs> (09.08.2017).
- Chang, Shan-Ju und Rice, Ronald E.: Browsing: A Multidimensional Framework, in: Annual Review of Information Science and Technology 28, 1993, S. 231–276.
- Chang, Winston, Cheng, Joe, Allaire, Joseph J., Xie, Yihui und McPherson, Jonathan: shiny: Web Application Framework for R, 1.0.3, 2017, <https://cran.r-project.org/package=shiny> (09.08.2017).
- Chang, Winston und Borges Ribeiro, Barbara: shinydashboard: Create Dashboards With Shiny, 0.6.1, 2017, <https://cran.r-project.org/package=shinydashboard> (09.08.2017).

---

<sup>24</sup> An dieser Stelle sei angemerkt, dass es durchaus Algorithmen gibt, die semantische Relationen zwischen Wörtern adäquat modellieren können; beispielsweise *Latent Semantic Indexing*.

<sup>25</sup> Siehe Stalder, Felix: Algorithmen, die wir brauchen, 2017, <https://netzpolitik.org/2017/algorithmen-die-wir-brauchen/> (09.08.2017).

- Gibbs, Fred und Owens, Trevor: Building Better Digital Humanities Tools: Toward Broader Audiences and User-Centered Designs, in: *Digital Humanities Quarterly* 6.2, 2012, <http://www.digitalhumanities.org/dhq/vol/6/2/000136/000136.html> (09.08.2017).
- Gitelman, Lisa und Jackson, Virginia: Introduction, in: Gitelman, Lisa (Hrsg.): „Raw Data“ is an Oxymoron, Cambridge 2013, S. 1–14.
- Kohle, Hubertus: Kunstgeschichte Goes Social Media, in: *Aviso: Zeitschrift für Wissenschaft und Kunst in Bayern* 3, 2011, S. 38–43.
- R Core Team: R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, 2017, <https://www.r-project.org/> (09.08.2017).
- Salton, Gerard und Buckley, Christopher: Term Weighting Approaches in Automatic Text Retrieval, in: *Information Processing and Management* 24.5, 1988, S. 513–523.
- Salton, Gerard, Wong, Anita und Yang, Chung-Shu: A Vector Space Model for Automatic Indexing, in: *Communications of the ACM* 18.11, 1993, S. 613–620.
- Stalder, Felix: Algorithmen, die wir brauchen, 2017, <https://netzpolitik.org/2017/algorithmen-die-wir-brauchen/> (09.08.2017).
- Summit on Digital Tools in the Humanities: A Report on the Summit on Digital Tools, Charlottesville 2005.
- Weaver, Warren: Science and Complexity, in: *American Scientist* 36, 1948, S. 536–544.
- Wickham, Hadley: tidyverse: Easily Install and Load Tidyverse, 1.1.1, 2017, <https://cran.r-project.org/package=tidyverse> (09.08.2017).

# Von der IT-Gruppe Geisteswissenschaften zu einem LMU Center for Digital Humanities?

## Zu Entwicklung, Mehrwert und Nachhaltigkeit des Digitalen

Christian Riepl

Die sich in allen Gebieten der Forschung und Lehre in den Geisteswissenschaften ausbreitende Digitalisierung erfordert auch eine Anpassung und einen Ausbau bestehender IT-Infrastrukturen, ohne die digitale Forschung und Lehre nicht existieren kann. Der Titel dieses Beitrags ist bewusst als elliptischer Satz mit einem Fragezeichen formuliert. Das Fragezeichen lässt einen Bezug auf die indeterminierte Nominalgruppe, nach der es steht, oder auf den gesamten Sachverhalt zu. Durch das fehlende verbale Prädikat bleiben Handlungsart und Zeitbezug offen. Das fehlende Subjekt gibt Institutionen und Gremien in Wissenschaft und Politik die Möglichkeit zur Identifikation mit dem Handlungsträger, der ein Werden von einer bestimmten Größe aus zu einer noch unbestimmten Größe mit fraglicher Bezeichnung bewirkt oder bewirken wird. Letztlich läuft es auf die Frage hinaus: Werden alle Beteiligten, insbesondere die geisteswissenschaftlichen Fächer, Einsicht und Willen in ausreichendem Maß hervorbringen, um die Digitalisierung voranzutreiben?

Dieser Beitrag zeigt zunächst die Entwicklung des Digitalen in den Geisteswissenschaften an der LMU München, erläutert dann den Mehrwert des Digitalen und legt schließlich strategische und organisatorische Folgerungen zur Verwirklichung der Nachhaltigkeit des Digitalen dar.

## 1. Zur Entwicklung des Digitalen

An der LMU liegen die Wurzeln der heute sogenannten Digitalen Geisteswissenschaften etwa Anfang bis Mitte der 1980er Jahre.<sup>1</sup> In den Sprach- und Literaturwissenschaften wurden zunächst die Studienschwerpunkte „Elektronische Sprach- und Textanalyse“ (ESTA) und „Informationswissenschaftliche Sprach- und Literaturforschung“ (ISLIF) gebildet. Beteiligt waren primär die Phonetik sowie diejenigen Disziplinen, die heute als Computerlinguistik und Computerphilologie bekannt sind. Daraus konnte sich einerseits der CIP-Pool der Germanistik, das heutige IT-Zentrum Sprach- und Literaturwissenschaften, andererseits das Centrum für Informations- und Sprachverarbeitung (CIS) entwickeln. In den Kultur- und Geschichtswissenschaften wurde etwa Mitte der 1980er Jahre eine Arbeitsstelle eingerichtet, die die „kleinen Fächer“ an der LMU im Bereich der rechnergestützten wissenschaftlichen Arbeit unterstützen sollte. Während die Vielzahl der Fächer an der Lösung von Problemen der Textverarbeitung (insbesondere dem Generieren von nicht-lateinischen Zeichensätzen) interessiert war, wagten nur wenige sprach- oder sozialwissenschaftlich orientierte Disziplinen die Anwendung regelbasierter und statistischer Ansätze. Die Arbeitsstelle für rechnergestützte Forschung der Fakultät 12 und verwandte Fächer ging Ende der 1990er Jahre in die heutige IT-Gruppe Geisteswissenschaften (ITG) über. Etwa Mitte bis Ende der 1980er Jahre vollzog sich an der Fakultät für Mathematik der Ausbau des Instituts für Informatik zu einem eigenständigen Department, unter anderem mit dem Ziel, durch entsprechende Studienordnungen die Informatik stärker auf die Geisteswissenschaften auszurichten. Ungefähr im selben Zeitraum entstand in der Zentralen Universitätsverwaltung das Referat für EDV-Beratung, das bald zur Abteilung für Büro-, Informations- und Kommunikationstechnik ausgebaut wurde und heute im Dezernat VI für Informations- und Kommunikations-

1 Der geschichtliche Rückblick basiert zum einen auf Zeitzeugen, zum anderen auf Quellen des Archivs der IT-Gruppe Geisteswissenschaften, in dem sich sieben Aktenordner aus den Jahren 1984 bis circa 1992 mit Anträgen, Bestellungen, Protokollen und Korrespondenzen befinden, außerdem Vorlesungsverzeichnisse seit dem Sommersemester 1985, diverse Zeitschriftenberichte und Auszüge aus wissenschaftlichen Veröffentlichungen, ferner Protokolle, Dokumente und Jahresberichte von 1993 bis 2015.

technik sechs Referate umfasst, wobei es im Lauf seiner Geschichte alle Phasen der Computerisierung und Digitalisierung entscheidend geprägt hat. Nicht zuletzt zeigt sich hier, wie bedeutend die Informationstechnologie in allen Prozessen der Planung und Organisation des Wissenschaftsbetriebes geworden ist.

Im Rückblick auf die letzten 30 Jahre markieren folgende Meilensteine jeweils den Beginn einer neuen Phase bezüglich der Digitalisierung in den Geisteswissenschaften:

**1986** wird in den „kleinen Fächern“ die Beschaffung der ersten Computer geplant; zwei digitale Pilotprojekte werden mit Computern ausgestattet und beginnen ihre Arbeit; eine „Arbeitsstelle für rechnergestützte Forschung“ wird eingerichtet.

**1995** wird das Stammgelände der LMU flächendeckend vernetzt; in Informatik und Informationstechnologie sind zwei Themen vorherrschend: Multimedia und WWW.

**2000** wird auf der Grundlage eines fach- und fakultätsübergreifenden Konzeptes die IT-Gruppe Geisteswissenschaften gegründet; die geisteswissenschaftlichen Fakultäten erarbeiten und verabschieden ein gemeinsames Statut; 2001 konstituiert sich der Lenkungsausschusses als fakultätsübergreifendes Gremium. 2010 ist die Aufbauphase im Hinblick auf die personelle und technische Infrastruktur abgeschlossen.

**2013** setzt eine zunehmende Förderung der Forschungsrichtung „Digital Humanities“<sup>2</sup> ein.

**2016** wird die Förderung der „Digital Humanities“ im Rahmen des Programms „Digitaler Campus Bayern“<sup>3</sup> auf die Lehre ausgeweitet, wobei insbesondere die Vermittlung von IT-Inhalten an informatikferne Studiengänge im Mittelpunkt steht.

---

2 In das Jahr 2013 fällt auch die Gründung des Verbandes „Dhd - digital humanities im deutschsprachigen raum“, der seither auf nationaler und internationaler Ebene die Interessen der digitalen Geisteswissenschaften bündelt und fördert: <http://dig-hum.de/> (2.8.2017)

3 <https://www.km.bayern.de/studenten/digitalisierung/hochschule-digitaler-campus.html> (2.8.2017)

Seit dem ersten Antrag auf eine Ausstattung der geisteswissenschaftlichen Fächer mit Computern wächst der Betreuungsbereich der damals eingerichteten Arbeitsstelle an. Im Antragsjahr 1986 waren es zwölf zu betreuende Einrichtungen, 1988 bei den ersten Beschaffungen und der Erstbesetzung der Stelle bereits 15, 1993 bei der Neubesetzung der Stelle 20. Ab 1998 weitet sich die Arbeitsstelle zu einer Arbeitsgruppe und hat alle Fächer der beiden theologischen, der philosophischen sowie Teile der sprach- und literaturwissenschaftlichen Fakultäten an der LMU zu betreuen. 1999 kommen die Fächer des Historischen Seminars dazu. Bei der Gründung der IT-Gruppe Geisteswissenschaften im Jahr 2000 umfasst der Betreuungsbereich alle Einrichtungen der geisteswissenschaftlichen Fakultäten mit Ausnahme der Phonetik und des CIS. Erst 2002 werden die Kunst- und Musikpädagogik aus dem Betreuungsbereich der Psychologie übernommen. In den Jahren 2003 bis 2010 werden sukzessive das Zentrum Seniorenstudium, das Center for Advanced Studies, das Rachel Carson Center und das Munich Center for Mathematical Philosophy an die Betreuung durch die IT-Gruppe Geisteswissenschaften angehängt.

Der Aufbau einer gemeinsamen und einheitlichen informationstechnischen Infrastruktur und die Konsolidierung einer relativ heterogenen Hardware- und Softwarelandschaft vollzog sich – finanziert aus Mitteln der WAP- und CIP-Programme – über einen Zeitraum von ungefähr zehn Jahren (2000–2010). Aus einer 1988 als Server dienenden Workstation mit 15 PC-Arbeitsplätzen, deren hauptsächliche Anwendungen im Bereich der Textverarbeitung lagen, wurde bis zum Jahr 2017 eine hochkomplexe Serveranlage mit circa 60 Servern (auch virtuellen Maschinen) und 11 Storage-Einheiten, die circa 1800 Rechnerarbeitsplätze an 30 Standorten versorgen und den informationstechnischen Kern von über 330 Datenbanken und über 80 webbasierten Projekten bilden.

Folgendes Diagramm veranschaulicht die, seit 2013 steil nach oben verlaufende, Entwicklung im Bereich der Datenbanken und – meist drittmittelgeförderten – DH-Projekte seit 1998:

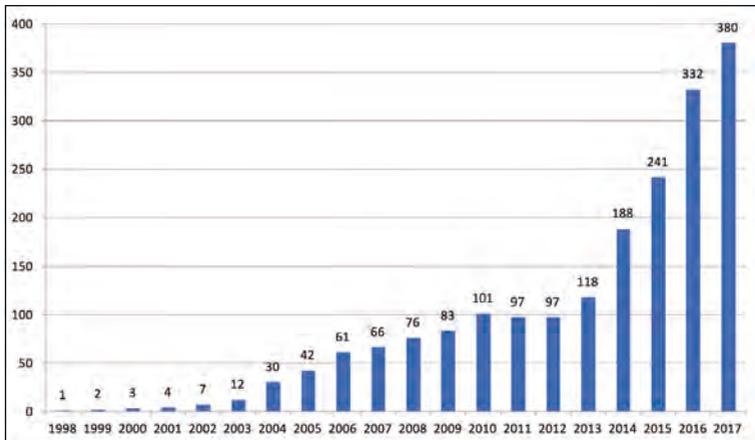


Abbildung 1: Entwicklung Datenbanken/DH-Projekte

Im Anschluss an wenige Prototypen MultiBHt<sup>4</sup> (1998), Artemis<sup>5</sup> und BMLO<sup>6</sup> (2001), AsiCA<sup>7</sup> (2005), Artigo<sup>8</sup> (2009), mehrsprach-o-mat<sup>9</sup> (2015), KiT<sup>10</sup> und dh-lehre<sup>11</sup> (2016) konnten sich eine Reihe web- und datenbankbasierter Systeme<sup>12</sup> entwickeln, zum Beispiel: RSQV<sup>13</sup>,

- 4 Biblia Hebraica transcripta: <http://www.bht.gwi.uni-muenchen.de/> (2.8.2017). Das älteste DH-Projekt der LMU war eines der beiden oben erwähnten Pilotprojekte, das 1986 von Wolfgang Richter initiiert wurde. Mit dem Ziel, den hebräischen Text des Alten Testaments sprach- und literaturwissenschaftlich zu analysieren, dafür Computer einzusetzen und mit der Informatik zu kooperieren, war damit – ganz in der Tradition eines Roberto Busa – auch an der LMU eine Anwendung aus der Theologie führend und wirkte inspirierend auf andere geisteswissenschaftliche Fächer. Zum Projekt Biblia Hebraica transcripta, siehe Riepl 2016 S. 295–311. Zu Roberto Busa und dem Gründungsmythos der Digital Humanities, siehe Thaller 2017 S. 3.
- 5 Artemis Bilddatenbank: <http://artemis.gwi.uni-muenchen.de/> (2.8.2017).
- 6 Bayerisches Musikerlexikon online: <http://bmlo.de/> (2.8.2017).
- 7 Atlante sintattico della Calabria: <http://www.asica.gwi.uni-muenchen.de/> (2.8.2017).
- 8 Der an der ITG entwickelte Prototyp wurde im Rahmen des Projektes „play4science“ weiterentwickelt zu Artigo mit verschiedenen Spielvarianten: <http://www.artigo.org/> (2.8.2017).
- 9 Mehrsprach-o-mat: <https://www.mehrsprach-o-mat.gwi.uni-muenchen.de/> (2.8.2017): Die Seite ist aus urheberrechtlichen Gründen Zugangsgeschützt.
- 10 Korpus im Text: <http://www.kit.gwi.uni-muenchen.de/> (2.8.2017).
- 11 Lehre in den Digital Humanities: <https://www.dh-lehre.gwi.uni-muenchen.de/> (2.8.2017).
- 12 Einen (nicht immer vollständigen) Überblick gibt <http://www.itg.uni-muenchen.de/projekte> (2.8.2017).
- 13 Richard Strauss Quellenverzeichnis: <http://rsqv.de/> (2.8.2017).

Richard Strauss Online Plattform<sup>14</sup>, play4science<sup>15</sup>, OUL/OUDB<sup>16</sup>, VerbaAlpina<sup>17</sup>, eDiAna<sup>18</sup>, DHVLab<sup>19</sup>, ARMEP<sup>20</sup> und kaiserhof<sup>21</sup>. Einen großen Anteil der Datenbanken machen die Sprach- und Textkorpora der Promovierenden der Graduiertenschule Language & Literature/Class of Language sowie Studierende der Sprachwissenschaften aus. Entscheidend ist die Einsicht, dass alle Daten und Systeme jeweils von Projektbeginn an und über die Förder- beziehungsweise Studien-/ Promotionszeit hinaus im Sinne der vielgeforderten Nachhaltigkeit weiter bestehen bleiben und auch in Zukunft zu warten, zu pflegen und zu tradieren sind.

In der Lehre nehmen seit 2003 vor allem auf dem Gebiet der Korpuslinguistik Interesse und Bedarf an der Vermittlung von IT-Grundlagen (Programmierung und Datenbanken) zu. Das Lehrangebot der IT-Gruppe Geisteswissenschaften wurde erst einsemestrig in zwei bis dreistündigen Lehrveranstaltungen organisiert und zunächst nur vereinzelt, auf freiwilliger Basis, wahrgenommen. Seit 2009 ist es für Doktoranden der Korpuslinguistik verpflichtend, seit 2013 ist es als optionales IT-Modul in die Curricula der MA-Studiengänge für Sprachwissenschaft integriert. Mittlerweile wächst auch schon bei BA-Studierenden der Sprach- und Geschichtswissenschaften die Einsicht in die Notwendigkeit, vertieftes IT-Wissen zu erwerben. Vielfach hat es sich schon gezeigt, dass einsemestrige Lehrangebote nicht mehr ausreichen.

---

14 Kritische Ausgabe der Werke von Richard Strauss (Online-Plattform): <http://www.richard-strauss-ausgabe.de/> (2.8.2017).

15 Projektseite: <http://www.play4science.uni-muenchen.de/> (2.8.2017).

16 Ob-Ugric Languages und Ob-Ugric Database: <http://www.babel.gwi.uni-muenchen.de/> (2.8.2017).

17 VerbaAlpina (Digitaler Online-Sprachatlas): <https://www.verba-alpina.gwi.uni-muenchen.de/> (2.8.2017).

18 Digital Philological-Etymological Dictionary of the Minor Ancient Anatolian Corpus Languages (eDiAna): <https://www.ediana.gwi.uni-muenchen.de/> (2.8.2017).

19 Digital Humanities Virtual Laboratory: <http://dhvlab.gwi.uni-muenchen.de/> (2.8.2017).

20 Ancient Records of Middle Eastern Polities: <https://www.armep.gwi.uni-muenchen.de/> (2.8.2017).

21 Personendatenbank der Höflinge der österreichischen Habsburger des 16. und 17. Jahrhunderts: <http://kaiserhof.gwi.uni-muenchen.de/> (2.8.2017).

Ein steigendes Interesse an Lehre in den Digital Humanities zeigt sich auch deutlich in zwei Projekten, die seit 2016 im Rahmen des Förderprogramms „Digitaler Campus Bayern“ in den Geisteswissenschaften umgesetzt werden. Das erste Projekt, „IT-for-all: Digitale Datenanalyse in den Geschichts- und Kunstwissenschaften“ (2016–2017), hatte zunächst die Gestaltung von zwei Lehrveranstaltungen im Blick. Bereits während des ersten Projektjahres entstanden Pläne für einen Zertifikatsstudiengang „Digital Humanities – Geschichte“. Das zweite Projekt, „BA-Nebenfach ‚Digital Humanities – Sprachwissenschaften‘“ (2017–2021), wird zunächst in einem fünfsemestrigen Lehrangebot die großen sprachwissenschaftlichen Disziplinen Anglistik, Germanistik und Romanistik/Italianistik bedienen. Auch hier besteht auf Seiten der kleineren sprachwissenschaftlichen Fächer sowie weiterer sprachwissenschaftlicher Disziplinen der kulturwissenschaftlichen Fakultät schon jetzt der Wunsch, in das Lehrangebot integriert zu werden.

Am Rande sei erwähnt, dass in den digitalen Geisteswissenschaften generell eine große Neugier nach informationstechnischem Wissen sowie ein intensives Bedürfnis nach interdisziplinären Austausch, Kooperation, Reflexion und Diskussion herrscht.<sup>22</sup>

Zuletzt soll kurz auf die Stellen- und Personalentwicklung an der IT-Gruppe Geisteswissenschaften eingegangen werden. Seit dem Abschluss der Aufbauphase (circa 2000–2010) stehen sieben unbefristete Stellen zur Verfügung, die in den Aufgabenbereichen (a) Konzeption, Planung, Organisation und Beschaffung, (b) Projekt- und Wissensmanagement (Beratung, Projekte, Lehre) sowie (c) IT-Management (Benutzerservice und Infrastruktur für Forschung und Lehre) tätig sind. Das Kernteam wird nach Bedarf durch zwei bis drei befristete Hilfskräfte verstärkt. Im Bereich Projekt- und Wissensmanagement kommen Drittmittelgeförderte und daher befristete Projektstellen hinzu, deren Anzahl seit 2013 stetig steigt. Aktuell sind an

---

22 Die Angebote an Workshops, Tagungen, Symposien, Summerschools etc. haben seit etwa 2013 beträchtlich zugenommen und sind fast schon unüberschaubar geworden.

der IT-Gruppe Geisteswissenschaften insgesamt 27 Mitarbeiterinnen und Mitarbeiter auf ca. 15 Stellen beschäftigt.

Anhand der gerade dargestellten Fakten, lässt sich folgendes festhalten:

- Die Entwicklung des Digitalen ist von Beginn an expansiv und attraktiv: Angebote und Ressourcen werden angenommen und genutzt.
- Die Entwicklung verläuft organisch, innovativ und synergetisch: Mit dem Bedarf wachsen IT-Infrastruktur und Angebote. Mit neuen Anwendungen ergeben sich neue Fragestellungen. Alle Prototypen haben Modellcharakter und werden fachübergreifend eingesetzt.
- Die Entwicklung wird sich nicht aufhalten oder umkehren lassen. Vielmehr wird die Digitalisierung in die Breite und in die Tiefe gehen und an Komplexität gewinnen.
- Die Entwicklung wird sich weiter beschleunigen.

Digitale Forschung und Lehre werden eine zunehmende Quantität und Qualität von Daten bewirken. Die Anzahl der Anwendungssysteme wird steigen, da sie in der Regel als „lebende Systeme“<sup>23</sup> konzipiert sind und als solche dauerhaft verfügbar sein sollten. Die Komplexität von Daten und Systemen wird wachsen.

Digitale Forschung und Lehre brauchen eine dauerhafte technische und personelle Infrastruktur, die Hochverfügbarkeit und Nachhaltigkeit von Forschungsdaten und Anwendungssystemen, ferner die Skalierbarkeit von Speicher- und Rechenkapazität gewährleisten kann. Bei der Entwicklung von Anwendungssystemen ist größtmögliche Effektivität (zum Beispiel durch Verwendung von Softwaremodulen und Schnittstellenstandards), im Umgang mit Forschungsprozessen und Forschungsdaten größtmögliche Transparenz anzustreben.

---

<sup>23</sup> Zur Nachhaltigkeit von Forschungsdaten in virtuellen Forschungsumgebungen, siehe Krefeld/Lücke 2017.

## 2. Zum Mehrwert des Digitalen

Das Buch „Imaginäre Größe“ von Stanislaw Lem ist 1973 erschienen und enthält fiktive Vorwörter zu fiktiven Werken. Eines davon ist „Vestrands Extelopädie in 44 Magnetbänden, Vestrand Books Co., New York, London, Melbourne, MMXI“<sup>24</sup>. Da die Artikel in Enzyklopädien schon bei ihrem Erscheinen veraltet sind, bietet der Verlag „Vestrand Books Co.“ als Lösung eine „EXTELOPÄDIE“, eine „EXTRAPOLATIVE TELEO-NOIMISCHE ENZYKLOPÄDIE“ an.<sup>25</sup> Dabei werden die Stichwörter und Sprachen der Zukunft in einem sehr aufwändigen und langwierigen Rechenprozess<sup>26</sup> unter anderem mit Hilfe von „achtzehntausend COMFUTER[n]“<sup>27</sup>, die in „BASCHHWELUKEN“<sup>28</sup> der „Komurbs“<sup>29</sup> organisiert sind und von einem „SUPERFUTER“<sup>30</sup> kontrolliert werden, vorausberechnet. Das Resultat<sup>31</sup> ist fehlerfrei, da menschenfrei. Die Bedienung<sup>32</sup> der „Extelopädie“ ist einfach: Sie wird in 44 „Magnetbänden“ geliefert und auf einer Stellage montiert. Der Benutzer positioniert sich davor, spricht ein Stichwort aus und der richtige „Magnetoband“ springt „selbstblättern“ in die ausgestreckte rechte Hand. Schwarze Stichworte geben sichere und unveränderliche Informationen wieder. Rote Stichworte beinhalten unsichere Informationen, bei denen die Zukunftserforschung noch im Gange ist. Hier steht der Text in „hologenetischer“ Fernverbindung mit der Chefredaktion. Ein Flimmern der Wörter zeigt an, dass sich die (kostenfreie) Korrektur und Aktualisierung gerade vollziehen.



24 Lem 1982, S. 85–108.

25 Siehe Lem 1982, S. 87 (Grund bzw. Problem), S. 89 (Termini).

26 Siehe Lem 1982, S. 89–93.

27 Siehe Lem 1982, S. 88: „futurologische Computer“.

28 Siehe Lem 1982, S. 89: „Batterien der Schwersten Lumenischen Comfuter“.

29 Siehe Lem 1982, S. 89: „Komurbs“ bzw. S. 92: „COMURBS“, „die Computerstadt“.

30 Siehe Lem 1982, S. 89: „elektronische Verkörperung des Mythos vom Übermensch“.

31 Siehe Lem 1982, S. 88.

32 Siehe Lem 1982, S. 93–96.

Ein Blick auf die heutige digitale Welt, in der Informationen aktuell und unabhängig von Raum und Zeit zugänglich sind, zeigt, dass die Utopie von Lem längst Wirklichkeit geworden ist. Bewirkt wurde dies durch die Fortschritte der Informationstechnologie, Informatik, Mathematik und Statistik, näherhin der Web- und Datenbanktechnologien, sowie der Medien-, Multimedia-, Programmier-, Rechen- und Speichertechnologien, ferner der Konzepte und Technologien, die zum Beispiel bei Human Computation und Crowdsourcing angewendet werden. Damit bestätigt sich auch für die Utopie von Stanislaw Lem und den Bereich der Informationstechnologie das Urteil von Eugen Biser<sup>33</sup>, nach welchem wir in einer Zeit leben, in der Utopien Wirklichkeit werden.

In der Utopie von Lem ist ein weiterer Punkt merkwürdig: Information und Wissen werden digital mit Geräten, wie den „Comfutern“ verarbeitet beziehungsweise berechnet und analog, beispielsweise in „selbstblättrenden Magnetbänden“ präsentiert. Lem sagt aber nichts dazu, wie Information und Wissen in die Computer gelangt, denn eine digitale Berechnung setzt eine digitale Abbildung der Wirklichkeit voraus.

Die Anwendung digitaler Techniken und damit Digitalisierung in den Geisteswissenschaften bedeuten im Wesentlichen einen Abbil-

---

33 Vgl. dazu Biser 1996, S. 250–262 und Biser 2004, S. 39–41, der im Rückgriff auf Sigmund Freud („Das Unbehagen in der Kultur“, 1930) und in Bezug auf die Rede vom Tod Gottes von Friedrich Nietzsche die Entwicklungen der modernen Technik kritisch betrachtet, insofern der Mensch die Technik nicht zur Daseinserleichterung nutzt, sondern zur Verwirklichung uralter Menschheitsträume, und insofern die Technik nicht nur positive Wirkungen, sondern auch Rückschläge hervorgebracht hat. Biser führt den Gedanken der Realisierung von Utopien theologisch weiter und sieht darin im Grunde den Übergang der Attribute Gottes (Allmacht, Allwissenheit, Allgegenwart, Schöpferum, Richtergewalt) auf den Menschen: Im Fall der in der Informationstechnologie weitergeführten Nachrichtentechnik wäre der Mensch dabei, sich die Gottesattribute der Allwissenheit und Allgegenwart, in gewisser Weise auch der Allmacht anzueignen. Die hoffnungsvollen und enttäuschenden Seiten der Digitalisierung und der Informationsgesellschaft hat auch Alexandra Borchardt in ihrem Festvortrag „Mensch 4.0 – Wie die Digitalisierung die Gesellschaft verändert“ zur Eröffnung des „technē - Campus-Netzwerk für Digitale Geistes- und Sozialwissenschaften“ am 27. April 2017 beleuchtet.

dungsprozess: Der Forschungsgegenstand muss erst digital abgebildet, kodiert werden. Die Abbildung bzw. Beschreibung muss logisch-struktural erfolgen, nicht (primär) graphisch. Somit steht die digitale Abbildung des Forschungsgegenstandes im Mittelpunkt digitaler wissenschaftlicher Arbeitsweise. Es handelt sich dabei primär um eine fachspezifische Auseinandersetzung<sup>34</sup> mit dem Gegenstand sowie den Methoden und Theorien. Der Abbildungsprozess kann für sich schon ein Erkenntnisgewinn in der Fachwissenschaft sein. Erst die logisch-strukturelle Abbildung des Forschungsgegenstandes ermöglicht als Mehrwert der digitalen Technik die exakte Beschreibung, Berechnung und Auswertung der Daten, sodann darauf aufbauend ihre Visualisierung, Konvertierung und Präsentation. Der Gesamtprozess kann als A-D-A-Transformation, eine Abbildung vom Analogen ins Digitale und wieder zurück, gesehen werden.

Zwei Beispiele mögen diesen zentralen Prozess der Abbildung veranschaulichen. Als Gegenstand dient der Titel des Vortrags. Den Einzelwörtern sollen Wortarten zugeordnet werden. Zuerst folgt die graphisch-pseudodigitale Abbildung, die ungefähr so aussehen könnte:

Von der IT-Gruppe Geisteswissenschaften zu einem LMU Center for Digital Humanities?

Die Wortartenanalyse sieht folgende Farbkodierung vor:

hellblau	Substantiv
schwarz	Eigenname
orange	Adjektiv
grün	Artikel
rot	Präposition

Das Fragezeichen, das unmittelbar dem vorausgehenden Substantiv folgt, sei hier der Einfachheit halber übergangen.

<sup>34</sup> Die oft für diesen Schritt von den Geisteswissenschaften vereinnahmte Informatik wird diese Abbildung nicht leisten können, kann aber dazu beitragen, indem sie zum Beispiel geeignete Datenstrukturen anbietet.

Eine Kodierung der Wortarten mit Farben wäre mit Hilfe geeigneter Office-Programme möglich und wird in der geisteswissenschaftlichen Praxis angewendet. Diese analog-graphische Welt der Office-Programme, die zum Beispiel auch bei PDF-Publikationen von wissenschaftlichen Veröffentlichungen und Textcorpora sehr verbreitet ist, stellt allerdings nur eine als pseudodigital zu bewertende Abbildung der Wirklichkeit dar. Die digitale Verarbeitung der so gespeicherten oder annotierten Daten ist (obwohl sie in gewisser Weise ja digital vorliegen) entweder gar nicht möglich oder extrem aufwändig. In der Regel geht der Workflow primär von der Textverarbeitung aus und bereitet von da aus entweder eine digitale Druckvorlage (PDF) oder in einem erneuten aufwändigen Arbeitsschritt die Vorlage für eine digitale Webpräsentation auf.

Als zweites folgt eine Möglichkeit für eine logisch-strukturelle Abbildung<sup>35</sup>:

ID	i_von	i_bis	Token	CS	Wortart
01	00	01	Von		Präp
02	01	02	der		Atk-def
03	02	03	IT-Gruppe		Sub
04	03	04	Geistesw.		Sub
05	04	05	zu		Präp
06	05	06	einem		Atk-idef
07	06	07	LMU		EN
08	07	08	Center	E	Sub
09	08	09	for	E	Präp
10	09	10	Digital	E	Adj
11	10	11	Humanities	E	Sub
12	11	12	?		.?

Die Einzelwörter und das Fragezeichen wurden in Tokens segmentiert und mit einem Referenzsystem (ID: absolute Identifikationsnummer;

<sup>35</sup> Zu Werkzeugen und Methoden im Allgemeinen, siehe Lücke/Riepl/Trautmann 2017. Zu den Möglichkeiten, die diese Abbildung mit Hilfe relationaler Datenbanken ermöglicht, siehe Lücke/Riepl/Trautmann 2017, S. 135–186.

i\_von und i\_bis: Positionsangaben als Intervalle) versehen, das in den Spalten ID, i\_von und i\_bis abgelegt ist. Einzelwörter und Satzzeichen befinden sich als Vertikaltext in der Spalte Token. Zusätzlich wurden die Spalten CS (Code-Switching) und Wortart angehängt. Die Tabelle lässt sich beliebig um Spalten für weitere Annotationen ergänzen. Durch die logisch-strukturelle Abbildung in Form einer Tabelle werden die Urteile, die vom Forscher getroffen werden, transparent und nachvollziehbar. Der Forscher ist sich bewusst, dass alle analytischen Entscheidungen, die in der Tabelle abgebildet werden, auf Regeln und Theorien beruhen, unbeschrieben davon, ob die Wortartenzuordnung wie in diesem Beispiel manuell oder automatisch zum Beispiel mit Hilfe eines Part-of-Speech-Taggers erfolgt.

Diese Art von Arbeit mit logisch-strukturalen Abbildungen und in der Folge auch der Anwendung von Auswertungsverfahren geht über die Möglichkeiten von Office-Programmen hinaus<sup>36</sup> und erfordert informatische und statistische Grundlagenkenntnisse in Bezug auf Datenstrukturen (Abbildung), Programmierung (Transformation, Analyse), Datenbanken (Verwaltung, Auswertung) und statistische Modellierung und Berechnungsverfahren (Auswertung, Visualisierung). Darüber hinaus sind Kenntnisse zur Anwendung bestimmter Programme (Editoren, Tagger, DBMS, R) oder fachspezifischer Anwendungen (GIS, CAD, Videoanalyse) erforderlich.

Erst hier zeigt sich der Mehrwert des Digitalen, denn die logisch-struktural abgebildeten Daten sind nach Quantität und Qualität modifizierbar und erweiterbar. Sie sind berechenbar mit Hilfe regelbasierter oder statistischer Verfahren, die halbautomatisch oder automatisch auf große strukturierte Datenmengen angewendet werden können. Sie sind langfristig haltbar, projekt- und fachübergreifend verknüpfbar und nachnutzbar, womit sie eine plurale Sicht auf einen Forschungsgegenstand ermöglichen. Die strukturierten und analytisch annotierten

---

36 Entscheidend ist dabei die Fähigkeit zu abstraktem, logischem, strukturalen Denken, das gefördert und angewendet werden muss. Eine Qualifikation als „Digital Native“ reicht nicht aus, da die „digitale Welt“, mit der dieser angeblich vertraut sei, meist ganz der analog-graphischen Welt verhaftet bleibt.

Daten sind subjektiv, exakt und transparent. Metasprachliche Transkriptionen helfen, die Urteile der Forscher darzulegen.

Die Notwendigkeit, Daten der analogen in die digitale Welt abzubilden, hat gezeigt, dass digitale Projekte nicht „menschenfrei“ sind. Im Mittelpunkt steht der Wissenschaftler, der die Wirklichkeit – genauer: eine (= seine) Sicht auf die Wirklichkeit – abbildet. Techniken und Methoden, die zum Einsatz kommen, beruhen auf theoretischen Voraussetzungen. Digitalisierung geschieht auch nicht zum Selbstzweck oder als Modeerscheinung. Sie muss von den forschenden und lehrenden Wissenschaftlern aus innerer, fachlicher Einsicht und methodischer Notwendigkeit gewollt und getragen werden.

Die Digitalisierung in Forschung und Lehre wirkt sich wiederum auf den Wissenschaftler aus. Ein digitaler Geisteswissenschaftler arbeitet nicht mehr für sich alleine. Er modelliert seine Daten im Hinblick auf Nachnutzbarkeit und Transparenz für andere Disziplinen. Er nutzt digitale Techniken zur Abbildung seines Forschungsgegenstandes, zur Analyse und Visualisierung seiner Forschungsergebnisse. Genauso wie sich die Astronomie des Teleskops oder die Biologie des Mikroskops bedient, werden in den Geisteswissenschaften mit Hilfe der oben genannten Technologien und Verfahren, insbesondere der Webtechnologie und Netzwerkanalyse, Instrumente wie ein „Syngeochronoskop“ für eine räumlich-zeitliche Synopse von Daten<sup>37</sup>, ein „Diageochronoskop“ für eine Sicht auf Daten im zeitlichen Verlauf<sup>38</sup> oder ein „Pangeochronoskop“ für weltweite Zusammenhänge unterschiedlichster Daten<sup>39</sup> denkbar und als Verfahren methodisch in das Fach integriert.

37 Siehe z.B. die georeferenzierte Darstellung von Sprachdaten bei VerbaAlpina, die beliebig auch mit anderen Daten (z.B. aus der Tabula Peutingeriana) überlagert werden können: [https://www.verba-alpina.gwi.uni-muenchen.de/?page\\_id=133](https://www.verba-alpina.gwi.uni-muenchen.de/?page_id=133) (2.8.2017)

38 Siehe z.B. die geo- und chronoreferenzierte Darstellung der Daten unter <http://www.theatrescapes.gwi.uni-muenchen.de/#theatres/maps/global> (2.8.2017) beim Theatrescapes Research Tool: <http://www.theatrescapes.gwi.uni-muenchen.de/> (2.8.2017)

39 Siehe die „Collection Places & Events“ mit den Navigationspunkten „Explore connections“, „Travel through time“ und „Traverse the globe“ unter <http://www.environmentandsociety.org/mml/collection/16233> (2.8.2017) beim Environment & Society Portal des Rachel Carson Centers: <http://www.environmentandsociety.org/> (2.8.2017)

Digitale Projekte sind ein ständiger und beidseitig fruchtbarer Lernprozess im Dialog mit der Informationstechnologie. Sie erfordern auf geisteswissenschaftlicher Seite eine Ausbildung, die diesen Dialog ermöglichen und fördern kann. Gerade für digitale Projekte sind Teamarbeit und Interdisziplinarität, Kooperation und Austausch, Arbeit und Diskussion am Gegenstand charakteristisch. Sie sehen sich eingebettet in größere Zusammenhänge, wie zum Beispiel Fachportalen, oder in einem umfassenderen Forschungsdatenmanagement, das das Auffinden von und Rechnen mit Daten über die unterschiedlichsten Disziplinen hinweg ermöglicht. Digitale Projekte folgen dem Prinzip „digital first“/„online first“ und fühlen sich dem Prinzip der offenen Lizenzen verpflichtet.

### **3. Zur Nachhaltigkeit des Digitalen**

Im Gesamtprozess der Digitalisierung spielt die IT-Gruppe Geisteswissenschaften eine zentrale Rolle. Als eigenständige Dienststelle an der LMU institutionalisiert, ist sie über einen akademischen Lenkungsausschuss mit Vertretern aller sechs geisteswissenschaftlichen Fakultäten dem wissenschaftlichen Bereich zugeordnet. Sie ist mit einer überschaubaren Anzahl von Planstellen ausgestattet, die ein dauerhaftes Kernteam bilden und für die Nachhaltigkeit des Digitalen in all ihrer Komplexität verantwortlich sind. Zur Nachhaltigkeit in diesem Sinne gehören die Wartung und Pflege der Softwaresysteme und Forschungsdaten ebenso wie die Bewahrung und Weitergabe von fachspezifischem IT-Wissen. Auf der Basis eines fach- und fakultätsübergreifenden Konzepts gelang es, eine allen geisteswissenschaftlichen Fakultäten gemeinsame IT-Infrastruktur aufzubauen, fachübergreifende Lösungen in Forschung und Lehre zu erarbeiten und damit Synergien auf allen Ebenen zu fördern. Die IT-Gruppe Geisteswissenschaften kooperiert eng mit den Referaten des Dezernats VI für Informations- und Kommunikationstechnik, der Universitätsbibliothek und der Unterrichtsmitschau sowie dem Leibniz-Rechenzentrum. Auf Münchner Ebene ist die LMU unter anderem über die IT-Gruppe Geisteswissenschaften im „dhmuc – Netzwerk für Digitale Geistes-

wissenschaften in München<sup>40</sup> mit den Kooperationspartnern Bayerische Akademie der Wissenschaften, Bayerische Staatsbibliothek und Deutsches Museum vernetzt. Auf bayerischer Ebene ist die IT-Gruppe Geisteswissenschaften zusammen mit DH-Einrichtungen der Universitäten München, Erlangen und Regensburg Partner im „technē – Campus-Netzwerk für digitale Geistes- und Sozialwissenschaften“<sup>41</sup>. Außerdem werden internationale Projektkooperationen in verschiedenen DH-Projekten gepflegt.

Die IT-Gruppe Geisteswissenschaften übt eine wichtige Vermittlungsfunktion an der Schnittstelle zwischen den Geisteswissenschaften und der Informationstechnologie aus. Sie schafft die Voraussetzungen für einen Dialog mit den statistischen und informatischen Disziplinen; sie bereitet die Grundlagen zur Integration digitaler Techniken und Verfahren in fachspezifische Methoden.

Burghardt/Wolff 2015 betrachten drei ausgewählte DH-Zentren<sup>42</sup> hinsichtlich (1) Struktur und Genese<sup>43</sup>, (2) Profil, Aufgaben und Dienste<sup>44</sup> sowie (3) Nutzer, Vernetzung und Projekte.<sup>45</sup> Die genannten Zentren zeichnen sich durch Kompetenzen in den Bereichen Beratung, Projekt/Methoden und Lehre aus.<sup>46</sup> Wichtig ist ferner die „Etablierung neuer Forschungsmethoden, Publikationskulturen in den digitalen Medienformen“<sup>47</sup> und die Wahrnehmung „strategischer Aufgaben“<sup>48</sup> (zum Beispiel in Netzwerken). Die bei Burghardt/Wolff 2015 genannten Charakteristika von DH-Zentren lassen sich nach dem bisher Ausgeführten hinsichtlich Konzeption, Organisa-

40 <http://dhmuc.hypotheses.org/> (2.8.2017)

41 Die Webseite des Netzwerks ist im Aufbau: <https://www.techne.gwi.uni-muenchen.de/> (2.8.2017).

42 Siehe Burghardt/Wolff 2015, S.314: Befragt wurden das Trier Center for Digital Humanities, das Cologne Center for eHumanities und das Göttingen Centre for Digital Humanities.

43 Siehe Burghardt/Wolff 2015, S. 315.

44 Siehe Burghardt/Wolff 2015, S. 315–316.

45 Siehe Burghardt/Wolff 2015, S. 316.

46 Vgl. dazu Burghardt/Wolff 2015, S. 315 und 317.

47 Siehe Burghardt/Wolff 2015, S. 315.

48 Siehe Burghardt/Wolff 2015, S. 316.

tion, Kompetenzen und Aufgaben auch bei der IT-Gruppe Geisteswissenschaften feststellen. Demnach hat die IT-Gruppe Geisteswissenschaften längst die Rolle eines Kompetenzzentrums für Digitale Geisteswissenschaften übernommen. Sie spiegelt in jeder Hinsicht die Situation von DH-Zentren in Deutschland wider.

Die weitere Entwicklung ist absehbar: Die Digitalisierung wird vorangetrieben werden, wofür Forschungsfördereinrichtungen wie die DFG oder auch der Rat für Informationsinfrastrukturen<sup>49</sup> sorgen werden. Die Anforderungen werden in allen Bereichen steigen und die Geschwindigkeit der Entwicklung wird weiter zunehmen. Um nachhaltige Strukturen zu verwirklichen, bedarf es Strategien. Eine davon ist, durch geschickte Organisation und komplexe Vernetzung eine Arbeitsteilung respektive eine Umverteilung von Aufgaben zu erreichen. Eine andere Strategie wäre, durch Investitionen eine Erhöhung der dauerhaften Personalkapazität herbeizuführen. Beide Strategien schließen sich nicht aus.

Die IT-Gruppe Geisteswissenschaften hat 2013 im Rahmen der eHumanities-Initiative des BMBF<sup>50</sup> einen Entwurf für ihren Ausbau zu einem DH-Zentrum<sup>51</sup> an der LMU vorgelegt, was sich jedoch schnell als Utopie erwies. Dabei wurde der Bedarf für die nachhaltige Sicherung der geförderten Strukturen auf vier Stellen geschätzt, die sich auf die Aufgabenbereiche (1) statistische Modellierung und Berechnung, (2) Datenbanken und Webschnittstellen, (3) Visualisierung und (4) Programmierung bezogen. Langfristig wurde als Ziel verfolgt, Kooperationen mit der Statistik, der Computerlinguistik und der Informatik auf- und auszubauen. Dennoch ist diese Utopie mittlerweile Wirklichkeit geworden, mit dem Unterschied, dass die Stellen für die oben genannten Aufgabenbereiche aus Drittmittelprojekten bereitgestellt

49 <http://www.rfii.de/> (2.8.2017).

50 <https://www.bmbf.de/foerderungen/bekanntmachung.php?B=804> (2.8.2017): Förderlinie 2 war für die Gründung oder der Ausbau von Zentren gedacht.

51 Die Bezeichnung einer solchen Einrichtung (Zentrum, Center, Centre oder dergleichen bzw. eHumanities, Digital Humanities oder dergleichen) ist dabei sekundär, sollte aber dennoch nicht vernachlässigt werden, zumal sie nach außen einen Anspruch und ein Niveau vermitteln kann.

wurden oder werden und organisatorisch für einen hohen Synergieeffekt gesorgt wird. Der Entwurf an sich aber hat sich im Alltag von Forschung und Lehre bewährt. Besonders die Vermittlung von Kenntnissen der statistischen Modellierung, Berechnung und Visualisierung und ihre informatische Vorbereitung und Umsetzung mittels Datenbanken und Webtechnologien stößt im Rahmen der aktuellen Projekte großes Interesse sowohl bei Lehrenden als auch bei Studierenden der Geisteswissenschaften. Eine Nachhaltigkeit über die Förderzeiten hinaus kann nur mit einem Ausbau vorhandener Strukturen erreicht werden. Die fortschreitende Digitalisierung zwingt geradezu alle an diesem Prozess beteiligten Institutionen und Gremien in Wissenschaft und Politik, bei allen zentralen DH-Einrichtungen und sogenannten Informationsinfrastrukturen dauerhafte personelle Infrastrukturen in enger Vernetzung mit Rechenzentren, Bibliotheken und geisteswissenschaftlichen Fächern zu schaffen. Auf dieser Grundlage kann die digitale Transformation der Geisteswissenschaften ein neues wissenschaftliches Niveau erreichbar machen.

## Literaturverzeichnis

- Biser, Eugen: Der Mensch – Das uneingelöste Versprechen. Entwurf einer Modalanthropologie, 2. Aufl., Düsseldorf 1996.
- Biser, Eugen: Orientierung im Spannungsfeld der Gegenwart, in: Aladár Gajáry (Hrsg.): Die Zeichen der Zeit erkennen. Theologische Zeitsignaturen am Beginn des 21. Jahrhunderts, 2004, S. 35–50.
- Burghardt, Manuel/Wolff, Christian: Zentren für Digital Humanities in Deutschland, in: Information – Wissenschaft & Praxis, 66 (5–6), 2015, S. 313–326.
- Krefeld, Thomas/Lücke Stephan: Nachhaltigkeit – aus der Sicht virtueller Forschungsumgebungen. Korpus im Text, Version 7 (10.03.2017, 12:27), <http://www.kit.gwi.uni-muenchen.de/?p=5773&v=7> (04.01.2018).
- Lem, Stanislaw: Imaginäre Größe. Frankfurt am Main 1982.

- Lücke, Stephan/Riepl, Christian/Trautmann, Caroline: Softwaretools und Methoden für die korpuslinguistische Praxis, in: Korpus im Text, Bd. 1, München 2017. Online abrufbar unter: <https://epub.ub.uni-muenchen.de/36308> (04.01.2018); <http://doi.org/10.5282/kit01> (04.01.2018).
- Riepl, Christian: Biblia Hebraica transcripta – Das digitale Erbe, in Rechenmacher, Hans (Hrsg.): In Memoriam Wolfgang Richter. Arbeiten zu Text und Sprache im Alten Testament 100, St. Ottilien 2016, S. 295–311.
- Thaller, Manfred: Geschichte der Digital Humanities, in: Jannidis, Fotis/Kohle, Hubertus/Rehbein, Malte (Hrsg.): Digital Humanities. Eine Einführung, Stuttgart 2017, S. 3–12.



# Auf dem Weg zu einem DH-Curriculum

## Digital Humanities in den Geschichts- und Kunstwissenschaften an der LMU München

Julian Schulz

Die Zahl der Studiengänge im Bereich der Digital Humanities (DH) im deutschsprachigen Raum hat in den vergangenen Jahren stetig zugenommen.<sup>1</sup> Ihre Etablierung an zahlreichen Universitätsstandorten ist eine Reaktion auf den voranschreitenden methodologischen und technischen Wandel in den geisteswissenschaftlichen Disziplinen. Die Institutionalisierung des weiten und zuweilen unscharf definierten Feldes, welches sich hinter dem Label „DH“ verbirgt, zeigt sich in der wachsenden Zahl an Professuren, die zumindest eine Teildenomination für DH besitzen.<sup>2</sup> Die variable Bezeichnung und Ausgestaltung betrifft auch die Lehrangebote selbst. Diese reichen von übergreifenden DH-Studiengängen (BA und MA) über disziplinentrierte Master bis hin zu den sogenannten „Bindestrich-Informatikern“<sup>3</sup>. Abgesehen von der unterschiedlichen Ausgestaltung besitzen diese Programme eine Gemeinsamkeit: Bei allen handelt es sich um explizite Angebote zur Ausbildung in den Digital Humanities.

Der digitale Wandel innerhalb der Geisteswissenschaften wirkt sich auf die Art und Weise aus, wie künftig in den einzelnen Fachdisziplinen geforscht wird. Digitalisierte beziehungsweise digitale Quellen erfordern die Kenntnis neuer Methoden für ihre Bearbeitung. Inno-

---

1 Patrick Sahle zufolge belief sich die Zahl im Jahr 2013 auf rund 70 nachgewiesene BA- und MA-Studiengänge im Bereich der DH; Sahle 2013, S. 4. Eine kartographische Darstellung, ohne Anspruch auf Vollständigkeit, bietet die DARIAH Course-Registry: <https://registries.clarin-dariah.eu/courses/> (04.01.2018).

2 Vgl. Sahle 2016-B, o. S.

3 Rapp/Bartsch/Borek 2016, S. 173.

vative Techniken bedürfen eines kritischen Verständnisses der ihnen zugrundeliegenden Algorithmen. Geisteswissenschaftliche Absolventen finden sich zunehmend in digitalen Projekten wieder, in denen sie im interdisziplinären Austausch mit Informatikern stehen. Vor diesem Hintergrund erscheint es notwendig, IT-Kenntnisse an künftige Geisteswissenschaftler nicht nur im Rahmen der eingangs genannten DH-Studiengänge zu vermitteln. Informationstechnologisches Basiswissen sollte vielmehr in das geisteswissenschaftliche Fachstudium integriert werden.<sup>4</sup>

Im Folgenden werden zunächst disziplinspezifische Gründe, die für eine IT-Grundausbildung in den Geisteswissenschaften, konkret in den Geschichts- und Kunstgeschichtswissenschaften, sprechen, knapp umrissen. Ein besonderer Schwerpunkt wird anschließend auf die, durch die Digitalisierung im Wandel begriffene, Arbeitsmarktsituation für (Kunst-)Historiker gelegt. Hierzu wurde eine Auswertung der Stellenangebote des Onlineportals H-Soz-Kult vorgenommen, die im Rahmen dieses Beitrags erstmals vorgelegt wird. Aufbauend auf die Bedarfsanalyse folgt im zweiten Teil des Beitrags die Vorstellung eines konkreten Ansatzes zur Einbindung des Digitalen in das Curriculum der Fächer Geschichte und Kunstgeschichte an der Ludwig-Maximilians-Universität München (LMU). Die Initiative zur Konzeption eines DH-Zertifikats für die beiden genannten Fächer wird durch Angehörige des Historischen Seminars, des Departments für Kunstgeschichte sowie der IT-Gruppe Geisteswissenschaften der LMU München (ITG) vorangetrieben. Dank des im Rahmen der Förderlinie „Digitaler Campus Bayern“ initiierten Pilotprojekts „Digitale Datenanalyse in den Geschichts- und Kunstwissenschaften“ und der langjährigen Erfahrung der ITG in der Betreuung digitaler Projekte und der Lehre in den DH kann auf eine solide Basis aufgebaut werden.

---

4 Diese Erkenntnis ist inzwischen bereits in die Tagespresse vorgedrungen, siehe bspw. Hannah Bethke: Arbeitswelt 4.0. Wo Bildung digital werden muss, in: FAZ, Beruf & Chance [28. Juni 2016], <http://www.faz.net/-gyl-8ir48> (04.01.2018).

## 1. *IT for all* – Gründe für die Integration einer IT-Grundausbildung in das Fachstudium

### 1.1 Digitale Alphabetisierung

Das Wissen und die Kompetenzen, welche im Rahmen eines Geschichts- oder Kunstgeschichtsstudiums erworben werden, sind für jede wissenschaftliche Beschäftigung mit Gegenständen der einzelnen Disziplinen von zentraler Bedeutung. Der Wert der Befähigung zur kritischen wissenschaftlichen Reflexion wird durch die Digitalisierung nicht geschmälert. Dennoch verändert der Einzug des Digitalen in die Wissenschaft die Art und Weise, wie Historiker und Kunsthistoriker zukünftig arbeiten werden. Interdisziplinären Teams, bestehend aus Geisteswissenschaftlern, Digital Humanists und Informatikern scheint die Zukunft zu gehören.

Damit eine fruchtbare Kommunikation zwischen der Informatik und den Geisteswissenschaften möglich ist, bedarf es einer grundlegenden und möglichst frühzeitigen „digitalen Alphabetisierung“<sup>5</sup> der Geisteswissenschaften in der Breite.<sup>6</sup> Wie beim Erlernen einer Fremdsprache kommt es nicht auf ihre perfekte Beherrschung an – für eine verständliche Kommunikation erscheinen rudimentäre Sprachkenntnisse ausreichend. Ziel soll es sein, Geisteswissenschaftlern bereits frühzeitig ein solides ‚Sprachniveau‘ zu vermitteln, damit sie den Diskurs um digitale Entwicklungen mitgestalten können. Die Förderung digitaler Kompetenzen sollte daher „zum Kernbestandteil akademischer Grundausbildung“<sup>7</sup> ausgebaut werden, um einem „digitalen Analphabetismus vorzubeugen“<sup>8</sup>.

- 5 Diese Begrifflichkeit geht auf das im Entstehen begriffene Deutsche Internet-Institut zurück; vgl. die Übersichtsliste, zusammengestellt von Mertens, Peter und Barbian, Dina: Digital... Inflation der Wortschöpfungen, 16.08.2016, [https://web.archive.org/web/20170918211256/http://wi1.uni-erlangen.de/sites/wi1d7.wi1projects.com/files/publications/D\\_Liste\\_16.08.2016.pdf](https://web.archive.org/web/20170918211256/http://wi1.uni-erlangen.de/sites/wi1d7.wi1projects.com/files/publications/D_Liste_16.08.2016.pdf), S. 3, Nr. 174.
- 6 Siehe bspw. Wettlaufer 2014, o. S., Dinsman 2016, o. S. sowie Rapp/Bartsch/Borek 2016, S. 174.
- 7 Fickers 2014, S. 27; so auch Krajewski 2015, S. 3. Benötigt werde eine „differenzierte Vermittlung der grundlegenden akademischen Kulturtechniken Lesen, Schreiben, Denken – und im 21. Jahrhundert eben auch: Programmieren“.
- 8 Krajewski 2015, S. 4. In die gleiche Richtung geht Ehrlicher 2016, S. 629, der – frei nach Kant – die Forschungscommunity zum Ausgang aus dem „Stand der [digitalen]

## 1.2 Digitale Quellenkritik

Ein weiterer Grund für eine IT-Grundausbildung geisteswissenschaftlicher Studierender liegt in der Befähigung zu einer digitalen Quellenkritik.<sup>9</sup> Dank umfangreicher Digitalisierungsunternehmungen stehen Quellenbestände zunehmend in digitaler Form zur Verfügung.<sup>10</sup> Neben digitalisierten und nachträglich in maschinenlesbaren Text umgewandelten Sammlungen nehmen durch die steigende Zahl an DH-Vorhaben auch *born-digital*-Datensammlungen stetig zu.<sup>11</sup>

So unterschiedlich die digital verfügbaren Quellen sind, so ähnlich sind die Herausforderungen, die sich an die Fachwissenschaften stellen: Mit der steigenden Zahl an digital verfügbaren (Primär-)Quellen steigt auch der Bedarf an methodologischer Sicherheit im Umgang mit ihnen. Sie erfordern die Verbindung vertrauter Methoden (z. B. *close reading*) mit innovativen Verfahren (*distant reading*). Nur wenn die als neuartig empfundenen Methoden (z. B. *text mining*, Bildähnlichkeitsanalysen) und der souveräne Umgang mit digitalen Quellen in der Breite der Fächer als Ergänzung zu traditioneller Quellenkritik gelehrt werden, erscheint eine Sicherung wissenschaftlicher Standards auf Dauer möglich.<sup>12</sup> Auf diese Weise wird das, was heute noch mit dem Label „Digital“ versehen wird, auf absehbare Zeit diesen Zusatz verlieren und zum „selbstverständlichen Methodenkanon“<sup>13</sup> der historischen Disziplinen gehören.

---

Unmündigkeit“ ermuntert.

9 Umfassend dargelegt wird dies bei Rehbein 2015-A, S. 18.

10 Für die Geschichtswissenschaften sei exemplarisch auf die digitalisierten und mit linguistischen Annotationen versehenen Bestände des Deutschen Textarchivs (DTA) verwiesen (<http://www.deustextarchiv.de/>). Für die Kunstgeschichte stellt „prometheus. Das verteilte digitale Bildarchiv für Forschung und Lehre“ (<http://www.prometheus-bildarchiv.de/>) eine zentrale digitale Quelle dar

11 Ein Beispiel aus der Kunstgeschichte ist das Play4Science-Projekt ARTigo (<http://www.artigo.org/>), im Rahmen dessen durch die Annotation von Kunstwerken neue Datensammlungen entstehen. In der Mediävistik wird beispielsweise mit [monasterium.net](http://monasterium.net/mom/home) (<http://monasterium.net/mom/home>) ein kollaboratives Urkundenarchiv geschaffen, das verschiedene Medien und Ressourcen bündelt und durch die Community sukzessive erweitert wird.

12 Siehe Schlotheuber/Bösch 2015, S. 2.

13 Rehbein 2015-A, S. 21; siehe auch Pannapacker 2013, o. S.

### 1.3 Digitale Algorithuskritik

Eng verbunden mit der Befähigung zum kritischen Umgang mit digital vorliegenden Quellenbeständen ist ein reflektierter Umgang mit technischen Verfahren, die zu ihrer Auswertung und Visualisierung eingesetzt werden.<sup>14</sup> Die Analyse (verhältnismäßig) großer Mengen an Daten und ihre anschließende visuelle Darstellung kann zur Aufdeckung historischer Phänomene beitragen und zu neuen Fragestellungen anregen. Ohne das kritische Hinterfragen der eingesetzten Software und Methodik können diese Ansätze jedoch positivistische Vorannahmen bewirken. Um keiner „Vereinfachungsmetaphorik“<sup>15</sup> zu unterliegen, gilt es die jeweils angewendete technische Herangehensweise an die Quellen nachzuvollziehen und zu verstehen. Nur so werden etwaige Unschärfen in der Analyse auffindbar.<sup>16</sup>

Studierenden geisteswissenschaftlicher Fächer sollte daher frühzeitig vermittelt werden, welcher Algorithmus für die Auswertung der vorliegenden Daten, welche Visualisierungsform für die Darstellung von Ergebnissen sinnvoll sein kann.<sup>17</sup> Die Erkenntnis, dass sich nicht jedes Tool in gleicher Weise für die Bearbeitung historischer Fragestellungen eignet, sollte in den Fokus der akademischen Grundausbildung rücken.

### 1.4 Der geisteswissenschaftliche Arbeitsmarkt im digitalen Wandel

Nach den skizzierten disziplinspezifischen Gründen, die für die Verankerung einer IT-Grundausbildung in der Breite der Fächer sprechen, widmen sich die folgenden Ausführungen der sich wandelnden Arbeitsmarktsituation in den Geisteswissenschaften.

Zunehmend wird auch in „klassischen“ geschichts- und kunstwissenschaftlichen Berufsfeldern (zum Beispiel in Archiven, Bibliotheken, Forschungseinrichtungen, Museen und Universitäten) Sicherheit im

<sup>14</sup> Siehe Klinke 2016, S. 31.

<sup>15</sup> Rehbein 2015-A, S. 17; siehe auch König 2016, o. S.

<sup>16</sup> Jakacki 2016, o. S.: „helping them to create rubrics and develop assessment techniques that emphasize critical engagement and process, and not fall prey to the pretty/shiny“, Blevins 2013, o. S. sowie Ehrlicher 2016, S. 633.

<sup>17</sup> Siehe Fickers 2014, S. 26f.

Umgang mit digitalen Methoden und Software vorausgesetzt. Die Aneignung dieses informationstechnologischen Grundverständnisses ausschließlich im Selbststudium erscheint vor diesem Hintergrund nicht mehr ausreichend.<sup>18</sup> Die Vermittlung dieser Kompetenzen gilt es vielmehr im Fachcurriculum zu implementieren.

Konkrete Zahlen gibt es für diese Annahme bisher nicht. Für den deutschen fachspezifischen Arbeitsmarkt liegen momentan nur Schätzungen vor. Die Zahl derjenigen, die im weiten Feld der Digital Humanities tätig sind, wird bei „derzeit rund 500 bis 1000“ Personen vermutet. Zudem werde heute kaum mehr ein größeres Forschungsvorhaben initiiert, welches keine digitale Komponente besitzt.<sup>19</sup> Um die These bezüglich der Auswirkungen des digitalen Wandels auf den geisteswissenschaftlichen Arbeitsmarkt messbar zu machen, lohnt sich die Auswertung von Stellenausschreibungen: Stellenanzeigen eignen sich in besonderer Weise für die Analyse von Anforderungsprofilen an Absolventen, da Arbeitgeber darin die gewünschten Qualifikationen in idealtypischer Form benennen. Dies verspricht, technologische Veränderungen auf dem Arbeitsmarkt abbildbar zu machen.<sup>20</sup>

Basierend auf den Ausschreibungen des Stellenportals der American Historical Association (AHA) werden derartige Analysen für den nordamerikanischen Raum alljährlich durchgeführt.<sup>21</sup> Um diese Lücke für die Arbeitsmarktsituation für (Kunst-)Historiker in Deutschland zu schließen, wurde eine Auswertung des größten deutschsprachigen Online-Portals für Geschichtswissenschaften, H-Soz-Kult, vorgenommen. Im Fokus der Analyse, die gemeinsam mit Tobias Englmeier (ITG, LMU München) durchgeführt wurde, stand die Entwicklung des Anteils derjenigen Stellen, die über IT-Kompetenzen im Anforderungsprofil verfügen.

---

18 So auch Spiro 2012, S. 332 (online: Absatz 4).

19 Sahle 2016-A, S. 80.

20 Siehe Hermes/Schandock 2016, S. 4.

21 Für die jüngste Auswertung bezüglich der Jahre 2015/2016 siehe Townsend/Swafford 2017. In die Analyse werden nur Ausschreibungen für PhD-Absolventen in vollem Stellenumfang einbezogen.

### 1.4.1 Methode

In einem ersten Schritt galt es, alle Stellenanzeigen von H-Soz-Kult seit Onlinegang des Portals im Januar 2002 zu beziehen. Da diese mit einem eindeutigen Identifier versehen sind,<sup>22</sup> konnten mit Hilfe eines Skripts sämtliche Stellen samt Ausschreibungstext automatisiert abgerufen und lokal abgespeichert werden. Anschließend wurden die Daten in eine Datenbank importiert und für die Analyse aufbereitet.<sup>23</sup>

Für den Zeitraum vom 9.1.2002 bis 30.6.2017 lassen sich insgesamt 12995 Einträge ermitteln. Zunächst wurden alle Fälle durchgesehen, bei denen die automatische Extraktion kein Veröffentlichungsdatum erfassen konnte. Die Datumsangabe wurde nachträglich ergänzt, sofern es sich beim jeweiligen Eintrag um eine Stellenausschreibung handelt. Nachfolgend wurden die datierten Fälle auf ihre Eigenschaft als Stellenangebot hin untersucht. Einträge, die offenkundig keine Stellenausschreibung darstellen (Hinweise auf neue BA-/MA-Studiengänge, Auslobung von Wissenschaftspreisen, Reisekostentstipendien, Call for Papers), wurden entsprechend gekennzeichnet. Nach dieser ersten Durchsicht verblieb eine Gesamtzahl von 11757 Datensätzen.

Im nächsten Schritt wurden Stellenangebote für Praktikanten und (wissenschaftliche) Hilfskräfte herausgefiltert, wodurch sich die Gesamtzahl auf 10658 verringerte. Schließlich wurde der Datenbestand auch hinsichtlich der Anstellungsdauer bereinigt: Kurzzeitstipendien, Vertretungsanfragen, Werkverträge und Visiting Fellowships mit einer Laufzeit von bis zu sechs Monaten wurden herausgefiltert, es sei denn, eine Option auf Verlängerung wurde im

---

22 Schema: [http://www.hsozkult.de/job/id/stellen-\[id\]](http://www.hsozkult.de/job/id/stellen-[id]). Dies ist ein entscheidender Vorteil gegenüber Stellenanzeigen, die z.B. auf Institutsseiten veröffentlicht und nach Stellenbesetzung häufig gelöscht werden; vgl. hierzu Sahle 2016-B, o. S.

23 Zur Nachvollziehbarkeit der Ergebnisse und für weiterführende Analysen (Open Database License: ODC-ODbL) stehen die Daten und Abfragen dauerhaft zur Verfügung: <https://doi.org/10.5282/ubm/data.110> (Open Data, Universitätsbibliothek LMU, 4. September 2017).

Ausschreibungstext explizit benannt. Als Ausgangswert für die nachfolgende Auswertung verblieben 10131 Stellen.<sup>24</sup>

Im Erkenntnisinteresse der Analyse stand nicht die Zahl der expliziten DH-Stellen; vielmehr sollten gerade (auch) diejenigen Stellen ermittelt werden, die sich an Absolventen geschichtswissenschaftlicher Fächer richten und dabei ein Mindestmaß an IT-Kompetenz wünschen oder sich zumindest ein IT-Bezug aus der Projektbeschreibung ergibt. Hierfür wurde ein Analyseset aus IT-Kompetenzen gebildet, die den DH zuzurechnen sind. Darunter fallen insbesondere die Bereiche Datenmanagement (Datenbanken, Langzeitarchivierung), Programmierung (Java, PHP, Python), Digitale Editionen (XML, TEI) sowie die Statistik. Einbezogen wurden auch Technologien zur Erstellung von Webseiten (HTML, JavaScript) sowie allgemeinere Formulierungen (u.a. Digital Humanities, Digital History, IT-Kenntnisse). Die Abfrageergebnisse wurden wiederum qualitativ überprüft, um Einträge herauszufiltern, die als Stellen mit IT-Kompetenz ausgegeben wurden, diesem Label jedoch nicht gerecht werden.<sup>25</sup>

Um die Entwicklung seit 2002 nachzeichnen zu können, wurde die Abfrage für jedes Jahr separat durchgeführt und die ermittelte Stellenzahl in Relation zu der jährlichen Gesamtstellenzahl gesetzt. Die vorgenommene Einordnung von Stellen mit IT-Kompetenz erhebt keinen Anspruch auf Vollständigkeit, sondern möchte generelle Tendenzen und Entwicklungen auf dem geisteswissenschaftlichen Arbeitsmarkt sichtbar machen. Für zukünftige Auswertungen wäre eine Präzisierung der Analysemethodik (beispielsweise mit Hilfe von *machine learning*) denkbar.

---

24 Die Zahlen weichen von den im Vortrag genannten Zahlen ab, da gemäß dem Motto „publish, then filter“ (Shirky 2008, S. 81) die Datenbereinigung in der beschriebenen Form erst im Zuge dieser Veröffentlichung vorgenommen wurde.

25 Dies betrifft beispielsweise Einträge, in denen unter „IT-Kenntnisse“ die Beherrschung gängiger Softwareprodukte wie der MS-Office-Produktpalette oder von Bildbearbeitungsprogrammen sowie die Betreuung von Webseiten mit Hilfe eines Content-Management-Systems verstanden wird. Für die komplette Schlagwortliste vgl. die Aufstellung unter dem in Anm. 23 angegebenen Link.

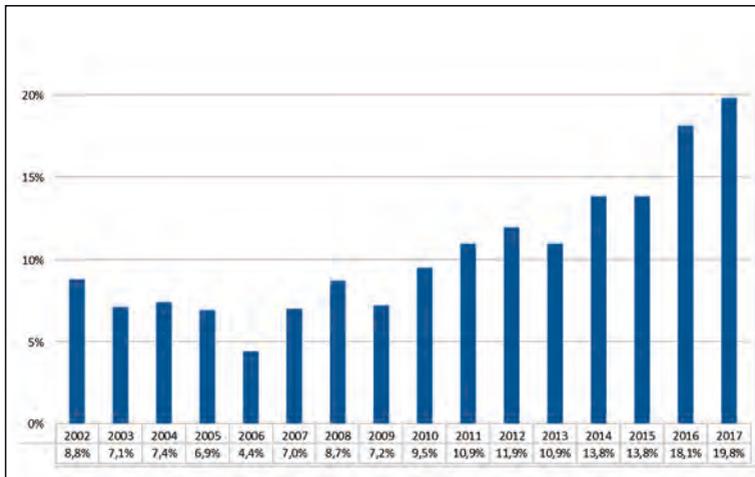


Abbildung 1: Auswertung der Stellenanzeigen des Portals H-Soz-Kult:  
Anteil der Stellen mit IT-Kompetenzen im Ausschreibungstext.

#### 1.4.2 Die Auswertung

Mit der geschilderten Heuristik konnte, aufgegliedert auf die einzelnen Jahre, folgendes Ergebnis ermittelt werden (vgl. Abb. 1): In der Anfangsphase des Jobportals (2002–2004) findet sich ein konstantes Niveau an Stellen mit IT-Kompetenz im Anforderungsprofil, welches zwischen 7,1% und 8,8% rangiert. Dies könnte darauf zurückzuführen sein, dass in der Anfangszeit des Portals insgesamt noch relativ wenige Angebote eingestellt wurden, darunter jedoch verstärkt solche, die nach IT-affinen Historikern suchten. In den nachfolgenden Jahren (2005–2009) geht der Prozentsatz zunächst leicht zurück, gelangt jedoch rasch wieder an das vorgenannte Niveau heran. Während die Gesamtstellenzahl durch die zunehmende Popularität des Portals stieg, wuchs der Anteil an IT-Kompetenzstellen in diesem Zeitraum nur langsam.

Die Zahlen stabilisieren sich bis zum Jahr 2010. Ab diesem Zeitpunkt, zu dem auch die DH zunehmend als Disziplin greifbar werden,<sup>26</sup> nimmt der prozentuale Anteil der „IT-Stellen“ gemessen an der Gesamtzahl stetig zu. Ein zwischenzeitlicher Rückgang im Jahr 2013 fällt nicht hinter das Niveau von 2010 zurück. Ein deutlicher Anstieg findet schließlich zwischen den Jahren 2015 und 2016 statt (13,8% auf 18,1% bzw. 202 von 1114 Datensätzen). Am Ende der sukzessiven Entwicklung steht für das 1. Halbjahr 2017 ein Wert von 19,8% (128 von 647 Datensätzen), d.h. jede fünfte Stelle für geschichtswissenschaftliche Absolventen führt mindestens eine IT-Kompetenz im Ausschreibungstext. Damit wird der hohe Wert von 2016 sogar noch leicht überschritten.

Diese Tendenz ist beachtlich, insbesondere vor dem Hintergrund, dass es sich bei dem über H-Soz-Kult ermittelten Wert nur um eine stichprobenartige Erhebung handelt. Es kann davon ausgegangen werden, dass die tatsächliche Zahl noch höher liegt: Erstens finden sich auf H-Soz-Kult nur Stellen, die auch offiziell ausgeschrieben und eingestellt wurden,<sup>27</sup> zweitens werden viele Stellen mit IT-Bezug, die auch für Fachhistoriker mit entsprechender Qualifikation in Frage kommen, ausschließlich über spezielle E-Mailverteiler, wie zum Beispiel die DHd-Mailingliste,<sup>28</sup> verbreitet. Drittens gilt es zu bedenken, dass Absolventen mit IT-Zusatzqualifikation auch bei Stellen, die diese nicht explizit im Anforderungsprofil benennen, einen Vorteil gegenüber ihren Mitbewerbern besitzen können. Schließlich sollte berücksichtigt werden, dass ein beachtlicher Teil der Absolventen geisteswissenschaftlicher Fächer einen Karriereweg außerhalb des traditionellen Arbeitsmarktes einschlägt. Der Erwerb grundlegender IT-Kompetenzen neben dem Fachstudium führt zu einer gesteigerten Anschlussfä-

---

26 Faktoren für eine Institutionalisierung sind u.a. die Einrichtung von Professuren und die Akkreditierung von Studiengängen, siehe Rapp/Bartsch/Borek 2016, S. 173.

27 Der Bericht des Wissenschaftsrats 2014, S. 29, verweist darauf, dass Stellenzuweisungen an Universitäten „teilweise selbst ohne interne Ausschreibung“ erfolgen.

28 Sämtliche Nachrichten, die über die DHd-Mailingliste versandt wurden, können über das Listenarchiv eingesehen werden: <https://mailman.rz.uni-hamburg.de/pipermail/dhd/>.

higkeit auf dem außeruniversitären Markt.<sup>29</sup> Diesen Fakt zu benennen sollte kein Tabu darstellen. Vielmehr ist an den Universitäten dafür Sorge zu tragen, Absolventen auf verschiedene Betätigungsfelder – innerhalb und fernab der Fachwissenschaft – bestmöglich vorzubereiten.<sup>30</sup>

Mit Hilfe der Auswertung von H-Soz-Kult konnte die sukzessive Zunahme von Stellen mit IT-Kompetenzen im Anforderungsprofil innerhalb der Fachwissenschaften dargelegt werden. Wenngleich die ermittelten Prozentzahlen aufgrund der heuristischen Methodik keine letztgültige Exaktheit besitzen, ist die Tendenz doch eindeutig: Auch von Kunsthistorikern und Historikern werden verstärkt informationstechnische Grundkompetenzen erwartet.

Die Vorteile für den Arbeitgeber liegen auf der Hand: IT-affine Geisteswissenschaftler besitzen ein gesteigertes Bewusstsein für die Bedeutung von Kommunikations- und Teamfähigkeit im interdisziplinären Kontext, sie sind mit Datenstrukturen und -formaten vertraut und analytische Prozesse gehören zu ihrem Handwerkszeug.<sup>31</sup> Sie besitzen einen geschulten Blick auf digitale Quellen und können Auswertungen und Visualisierungen kritisch einordnen. Dies leitet über zu der Frage, wie ein studienbegleitendes DH-Curriculum ausgestaltet werden sollte, um diesen Anforderungen bestmöglich zu entsprechen.

---

29 Diese Einschätzung findet sich bspw. bei Pannacker 2013, o. S. und Sahle 2013, S. 23f. Vgl. hierzu auch zwei im Jahr 2017 in der Süddeutschen Zeitung veröffentlichte Artikel von Tim Kummert: Endlich einer, der nicht nur Formeln anwenden kann, in: SZ, Karriere vom 20. August 2017 (<http://www.sueddeutsche.de/karriere/1.3623308>), sowie: Geisteswissenschaftler haben bessere Chancen denn je (Interview mit Cornelia Soetbeer), in: SZ, Karriere vom 21. August 2017 (<http://www.sueddeutsche.de/karriere/1.3623312>).

30 Diese Auffassung folgt Strobel/Welpe 2017, S. 317: „Hochschulen hingegen sollten in der Lage sein, sowohl auf dem Arbeitsmarkt nachgefragte Kompetenzen zu vermitteln [...] und einen reflektierten Umgang mit digitalen Technologien ermöglichen.“

31 Vgl. Sahle 2013, S. 21f.

## 2. Auf dem Weg zu einem DH-Curriculum

### 2.1 Pilotprojekt DHVLab

Eine Reaktion auf die eingangs angeführten Gründe für eine IT-Grundausbildung in der Breite geisteswissenschaftlicher Fächer, aber auch auf die eben benannten Veränderungen auf dem Arbeitsmarkt stellt die Förderlinie „Digitaler Campus Bayern“ dar. In deren ersten Förderphase wurde im Rahmen des Pilotprojekts „Digitale Datenanalyse in den Geschichts- und Kunstwissenschaften“ mit dem *DHVLab*<sup>32</sup> eine Infrastruktur für die praxisbezogene IT-Grundausbildung in den Fächern Geschichte und Kunstgeschichte geschaffen. Zum Einsatz kommt die Plattform derzeit in mehreren eigens konzipierten Seminaren, die von Mitarbeitern der IT-Gruppe Geisteswissenschaften in Kooperation mit Fachkollegen angeboten werden.<sup>33</sup>

Im Fokus steht die Verbindung fachwissenschaftlicher Fragestellungen mit digitalen Methoden. Dieser Ausgangspunkt aus der Mitte der Fachdisziplinen hinein in die digitale Welt wurde gewählt, um die Akzeptanz innerhalb der traditionellen Wissenschaften zu erhöhen.<sup>34</sup> Durch ihre praxisorientierte Ausrichtung besitzen die Seminare einen gewissen Laborcharakter, welcher charakteristisch für die Lehre in den Digital Humanities scheint:<sup>35</sup> Die Studierenden arbeiten im Team, gemeinsam werden Datensammlungen angelegt und strukturiert; diese dienen wiederum als Ausgangspunkt für erste eigenständige Gehversuche in der Visualisierung von Forschungsdaten.

Durch die sukzessive wachsende Plattform und die aktuell angebotenen Kurse wurde ein Fundament für eine Verankerung im fachwissen-

---

32 <https://dhvlab.gwi.uni-muenchen.de/>.

33 Eine Übersicht zu den Lehrveranstaltungen findet sich auf der Projekthomepage: [https://web.archive.org/web/20180223135942/https://dhvlab.gwi.uni-muenchen.de/Das\\_DHVLab\\_im\\_Einsatz](https://web.archive.org/web/20180223135942/https://dhvlab.gwi.uni-muenchen.de/Das_DHVLab_im_Einsatz). Bezüglich des Einsatzes der Plattform in der Lehre siehe den Beitrag von Harald Klink in diesem Band.

34 Die Notwendigkeit des fachwissenschaftlichen Ausgangspunktes betont bereits Wettlaufer 2014, o. S.

35 Diese Auffassung vertreten König 2016, o. S., Ehrlicher 2016, S. 625 und Rapp/Bartsch/Borek 2016, S. 174.

schaftlichen Curriculum geschaffen. Nun ist es erforderlich, über das Projektende hinaus Strukturen zu schaffen, die eine dauerhafte Fortführung dieses Ausbildungsangebotes in den Geschichts- und Kunstwissenschaften ermöglichen.

## 2.2 Bedarfsanalyse

### 2.2.1 Arbeitsmarkt und DH-Referenzcurriculum als Rahmen

Für die Institutionalisierung der angelaufenen IT-Grundausbildung in den genannten Fächern laufen derzeit Vorbereitungen zur Schaffung eines DH-Zertifikats. Dieses verfolgt das Ziel, Studierende bestmöglich auf die eingangs genannten Herausforderungen und den sich wandelnden Arbeitsmarkt vorzubereiten. Bei der Realisierung dieses Vorhabens kann neben den Eindrücken aus dem Pilotprojekt auf die langjährige Erfahrung der ITG im digitalen Projektmanagement und der Lehre einerseits und auf die Unterstützung durch das Multiplikatoren-Programm (Lehre@LMU) andererseits gebaut werden.<sup>36</sup>

Dass der Bedarf an Geisteswissenschaftlern mit grundlegenden IT-Kompetenzen hoch ist, wurde bereits gezeigt. Es stellt sich für die Planung allerdings die Frage, welche konkreten Fähigkeiten in der Breite der Fächer vermittelt werden sollten, um diesen Anforderungen gerecht zu werden. In einem ersten Schritt gilt es in Form einer Bedarfsanalyse zu ermitteln, was den Kern des IT-Basiswissens ausmacht, das jeder Absolvent beherrschen sollte.

Hier lohnt noch einmal der Blick auf die Stellenanalyse von H-Soz-Kult. Von den über den gesamten Zeitraum hin ermittelten 1187 Stellen mit IT-Profil besitzt die überwältigende Mehrheit (599) einen Bezug zu dem weiten Feld der Datenbanken. Hinzu kommen 80 Fälle, in denen die Arbeit mit fachspezifischen Datenbankanwendungen erwartet wird. IT-Kenntnisse im Zusammenhang mit Editio-

---

<sup>36</sup> Zur Entwicklung der ITG siehe den Beitrag von Christian Riepl in diesem Band. Für die Förderung der Zertifikatsentwicklung durch das Multiplikatoren-Programm vgl. <http://web.archive.org/web/20170815093146/http://www.multiplikatoren-projekt.peoplemanagement.uni-muenchen.de/multiplikatoren/geistes--und-kultur/fak-9/6-jahrgang-15161/index.html>.

nen sind in 141 Fällen gewünscht, gefolgt von 135 Ausschreibungen im Bereich Datenmanagement. Immerhin 127 Ausschreibungen führen Programmiererfahrung und 89 Statistikenkenntnisse an. Erfahrung in der Erstellung von Webseiten ist in 73 Ausschreibungen ein Kriterium (vgl. Abb. 2).<sup>37</sup>

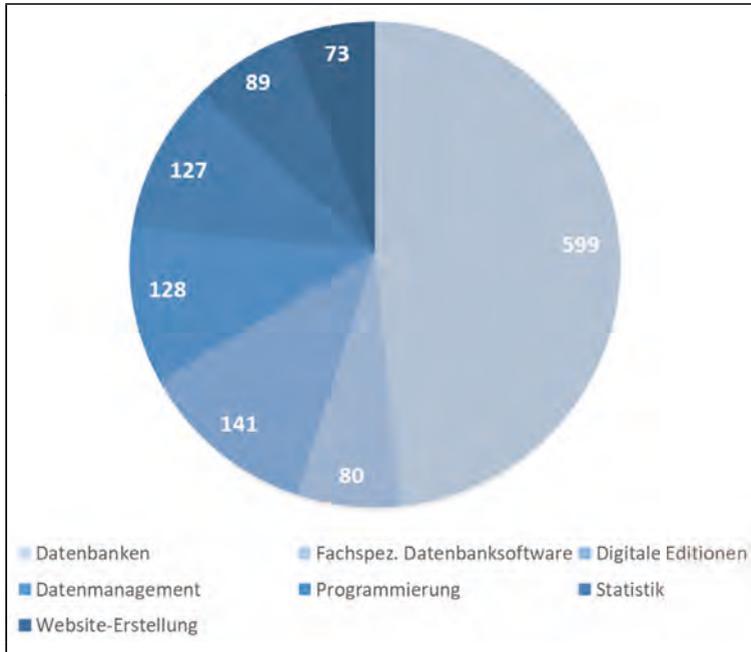


Abbildung 2: Stellen mit IT-Profil, aufgegliedert nach Kompetenzbereichen (Mehrfacherfassung möglich).

Dies gibt einen inhaltlichen Rahmen vor, in dem sich das DH-Curriculum bewegen wird. Einen wichtigen Ausgangspunkt für die inhaltliche und strukturelle Ausgestaltung von DH-Programmen stellt ferner das in fortwährender Weiterentwicklung befindliche DH-Referenzcurriculum dar. Dieses von der gleichnamigen DHD-Arbeitsgruppe

<sup>37</sup> Für die Ermittlung der Zahlen vgl. die entsprechenden SQL-Abfragen unter dem in Anm. 23 angegebenen Link.

betriebene Vorhaben zielt darauf ab, „im Sinne einer Referenzarchitektur“<sup>38</sup> Modelle zur Gestaltung von DH-Studienprogrammen vorzustellen und zu einer Vereinheitlichung des einhergehenden Vokabulars beizutragen. Dies verspricht die Vergleichbarkeit von Ausbildungsprogrammen zu verbessern, ohne die Flexibilität und den Gestaltungsspielraum vor Ort zu stark einzuschränken.<sup>39</sup> Nicht zuletzt dienen die IT-Zertifikate an den Universitäten zu Köln<sup>40</sup> und Passau<sup>41</sup> sowie „das digitale Archiv“ (DDA) in Stuttgart<sup>42</sup> als Vergleichsparameter für die Ausgestaltung des Curriculums an der LMU.

### 2.2.2 IT-Lernzielkatalog

In Form eines IT-Lernzielkataloges wird zunächst klar definiert, welche technischen Kenntnisse im Rahmen des Zertifizierungsprogramms vermittelt werden und welcher Grad des Kompetenzerwerbs erreicht wird. Dabei lohnt ein Rückgriff auf das Drei-Ebenen-Modell der Kompetenzvertiefung nach Sahle:<sup>43</sup> Eine Technologie, Softwarelösung oder ein Standard wird demnach A) gekannt, B) verstanden und angewendet, C) selbständig weiterentwickelt. Das Zertifikat wird sich in den beiden erstgenannten Niveaustufen (A und B) bewegen, da im Rahmen dieser Zusatzausbildung die Anleitung zu eigenständiger Software-(weiter)Entwicklung nicht realisierbar ist.

Die zu vermittelnden Gegenstände selbst lassen sich ebenfalls auf Grundlage des (vorläufigen) Referenzcurriculums definieren (vgl. Abb. 3):<sup>44</sup>

38 Thaller 2015-A, S. 4.

39 Vgl. Rapp/Bartsch/Borek 2016, S. 173f.

40 Programmwebseite der Universität zu Köln: <http://www.itzertifikat.uni-koeln.de/>. Eine kurze Beschreibung liefert Thaller 2015-B, S. 73f.

41 Programmwebseite der Universität Passau: <http://web.archive.org/web/20170815093525/http://www.phil.uni-passau.de/zertifikat-dh/>. Das Zertifikatsprogramm wird ausführlich beschrieben bei Rehbein 2015-B, S. 76–88.

42 Programmwebseite der Universität Stuttgart: <http://web.archive.org/web/20170815093706/http://www.uni-stuttgart.de/dda/>. Seit dem Wintersemester 2015/16 gibt es ebendort einen Masterstudiengang Digital Humanities; das DDA besteht weiterhin als Vorbereitungsprogramm.

43 Sahle 2013, S. 23.

44 Die Ausführungen folgen den Vorschlägen zur Klassifizierung bei Sahle 2013, S. 20f. sowie Schumacher et al. 2016, S. 15.



Abbildung 3: Aufbau des DH-Zertifikats für Geschichtswissenschaften und Kunstgeschichte an der LMU München (Entwurf).

1) In einem einführenden Modul werden aktuelle Entwicklungen und Definitionen der Digital Humanities ergründet sowie technisches Basiswissen<sup>45</sup> vermittelt (Niveaustufe B).

2) Ein zentrales Ausbildungselement stellt die Akquirierung, Strukturierung und Modellierung von Forschungsdaten dar: Inhaber des Zertifikats sollten in der Lage sein, Daten auf verschiedene Weise zu generieren, bestehende Datensammlungen kritisch zu bewerten und schließlich in ein bestimmtes Datenmodell zu überführen. Die Befä-

<sup>45</sup> Darunter fallen bspw. die Funktionsweise von Computern, Datensysteme, Algorithmen, das Bedienen der Kommandozeile, Zeichenkodierung und Reguläre Ausdrücke.

higung zur Operationalisierung und der sichere Umgang mit Datenbanken erscheint – auch mit Blick auf das Ergebnis der Stellenauswertung – als fundamentaler Kern der Ausbildung (Niveaustufe B).

3) Zur Akquirierung und Weiterverarbeitung von Daten (z. B. Tokenisierung, PoS-Tagging) ist die Kenntnis von Skriptsprachen erforderlich. Daher soll in einem weiteren Modul eine Einführung in anwendungsorientiertes Programmieren (Python) gegeben werden (Niveaustufe A-B).

4) Wenngleich in interdisziplinären Teams die technische Seite der Analyse und Visualisierung von Forschungsdaten vorzugsweise seitens der Informatik übernommen wird, so sind die Kenntnis und die Fähigkeit zur Begutachtung geeigneter Software und Webtechnologien – auch hinsichtlich eigener Forschungsvorhaben – hilfreich. Im Modul „Analyse - Auswertung - Visualisierung“ kann daher zwischen einer Einführung in die Datenanalyse (mittels der Statistiksoftware R) oder einem Kurs zu Webtechnologien gewählt werden (Niveaustufe A).

5) Ebenfalls besteht die Möglichkeit zur Wahl zwischen dezidiert fachspezifischen Kursen zur digitalen Geschichte (z.B. Auszeichnen in und Verarbeiten von TEI-XML), Kunstgeschichte (Ähnlichkeits- oder Kombinationsanalysen) oder anderen, an der Fakultät 09 angesiedelten Fächern, insb. Musik- und Theaterwissenschaften (Niveaustufe B).

6) Als Abschluss der Zertifikatsausbildung ist ein gemeinsamer Kurs mit Studierenden der Informatik vorgesehen. Im Fokus dieser Übung soll das teamorientierte Arbeiten im interdisziplinären Kontext an einem projektbezogenen Fallbeispiel stehen (Niveaustufe B).

### 2.2.3 Einbezug der beteiligten Akteure

Das geplante Vorhaben erscheint nur dann erfolgversprechend, wenn es an die Gegebenheiten vor Ort angepasst wird und alle Beteiligten in die Diskussion einbezieht.<sup>46</sup> Nach der Ausarbeitung des IT-Lern-

---

46 So bereits Strobel/Welpe 2017, S. 316, Jakacki 2016, o. S. und Sahle 2013, S. 24.

zielkataloges (inklusive detaillierter Definition der Niveaustufen A und B) und der Studienstruktur werden diese einerseits im Fachkollegenkreis vorgestellt, andererseits mit Studierenden der betroffenen Fächer diskutiert. Bereits im Vorfeld sollen die genannten Gruppen durch die Beantwortung von Fragebögen in den Entstehungsprozess eingebunden werden. So wird die Zielgruppe des geplanten DH-Zertifikats, die sich von BA- und MA-Studierenden über Lehramtskandidaten bis hin zu Doktoranden erstrecken soll, beispielsweise befragt, wie viele Semesterwochenstunden für die Zusatzausbildung aus ihrer Sicht realistisch erscheinen. Von Dozenten wird eine Angabe erbeten, ob sie künftig für den fachwissenschaftlichen Part in Lehrveranstaltungen mit IT-Inhalten zur Verfügung stehen und welche interessanten Datenbestände sich aus ihrem Fachbereich für derartige Kurse anbieten würden.

Nach dieser „internen“ Erhebung wird das Konzept anschließend im Arbeitskreis für digitale Geistes- und Sozialwissenschaften in München, kurz *dhmuc*, vor Vertretern Münchener Wissenschaftseinrichtungen und damit der potentiellen Arbeitgeberseite vorgestellt.<sup>47</sup> Flankierend zu dem quantitativen Analyseergebnis sollen bei dieser Gelegenheit die von Bewerbern erwarteten Kernkompetenzen aus dem Bereich der DH qualitativ diskutiert werden. Diese Feinjustierung des Zertifizierungsprogramms verspricht eine Steigerung der Attraktivität der Ausbildung.

### 2.3 Implementierung

Auf dem Weg hin zu einem DH-Zertifikat für die Geschichtswissenschaften und die Kunstgeschichte an der LMU gilt es neben inhaltlichen und strukturellen auch organisatorische und rechtliche Fragen zu klären. Neben einer ausgewogenen Studienstruktur besitzt die dauerhafte Gewährleistung des Zertifizierungsprogramms, also das fortwährende Angebot aller vorgesehenen Kurse, Priorität. Dafür erscheint es

---

<sup>47</sup> Dem Arbeitskreis (<http://dhmuc.hypotheses.org/>) gehören u. a. das Bayerische Hauptstaatsarchiv, die Bayerische Staatsbibliothek, das Deutsche Museum, seitens der LMU die Departments für Geschichte und Kunstgeschichte sowie die ITG und die Universitätsbibliothek an.

wichtig, das Programm von einzelnen Personen zu entkoppeln<sup>48</sup> und auf ein tragfähiges Fundament zu hieven, welches ohne größere Koordinations- und Verwaltungsmehrarbeit auskommt.

Andererseits bedarf es einer Evaluation des zur Verfügung stehenden Lehrdeputats. Das Zertifikatsprogramm möchte einerseits bestehende DH-Kurse am Institut für Kunstgeschichte und dem Historischen Seminar bündeln; andererseits wird darauf abgezielt, Synergien mit benachbarten Fächern zu schaffen. Dabei könnten beispielsweise einzelne Kurse aus dem neugeschaffenen optionalen DH-Nebenfach für die Sprachwissenschaften<sup>49</sup> mitgenutzt und in die Ausbildung integriert werden. Verhandlungen über Kooperationsvereinbarungen mit dem Institut für Informatik und der Statistik für das vorgesehene Praxismodul werden in Kürze aufgenommen.

In die Überlegungen zur strukturellen Ausgestaltung werden auch digitale Angebote einbezogen. Standardisierte Inhalte können einen ersten theoretischen Zugang vermitteln oder heterogene Ausgangsqualifikationen ausgleichen.<sup>50</sup> Hierzu eignen sich online verfügbare Kursmaterialien.<sup>51</sup> Während in den USA diese Verzahnung von Online-Angeboten und Präsenzveranstaltungen vorangeschritten ist, erscheint diese „Hybridisierung“<sup>52</sup> für den deutschen Hochschulraum derzeit noch wenig kompatibel. Daher kommt der polyvalenten Nachnutzung bestehender Präsenzveranstaltungen im Rahmen des

---

48 Zu möglichen Fallstricken bei der Curriculums-Entwicklung: Sahle 2013, S. 7 u. S. 26f.

49 Für die Vorstellung des DH-Nebenfaches siehe den Beitrag von Markus Frank in diesem Band.

50 Siehe Strobel/Welpe 2017, S. 317f.

51 Verwiesen sei exemplarisch auf das Angebot von DARIAH-DE: <https://de.dariah.eu/tutorials>. Vgl. hierzu Schumacher et al. 2016, S. 9 u. S. 13f.

52 Strobel/Welpe 2017, S. 318. Ein Report der FutureManagementGroup 2015, S. 16, zur digitalen Bildung geht davon aus, dass sich „beispielsweise von Udacity und Coursera angebotene ‚Massively Open Online Courses‘ oder kurz MOOCs [...] als ein Bildungsinstrument der Zukunft durchsetzen“ werden. MOOCs seien als Chance „für die deutsche Hochschullandschaft“ anzusehen.

Zertifikats eine tragende Rolle zu. Auf diese Weise wird das Angebot der Ausbildungsinhalte dauerhaft gewährleistet.<sup>53</sup>

Parallel zum Aufbau von Kooperationsvereinbarungen sind auch rechtliche Fragen zu klären, insbesondere in Bezug auf die Prüfbarkeit von DH-Anteilen. Als Prüfungsleistung würde sich in einigen der angeführten Arbeitsfelder die Anlage eines semesterbegleitenden Portfolios, in anderen Kursen der erbrachte Programmiercode, eine erstellte Edition oder die angefertigte Projektarbeit anbieten. Um Fragen wie diese zu klären, ist es sinnvoll, vor der Überführung des Programms in ein offizielles Universitätszertifikat zunächst einen Probelauf durchzuführen. Die gewonnenen Erfahrungen werden unmittelbar in die zu erstellende Zertifikatsordnung einfließen. Eine Anrechnung der erbrachten Leistungen dieser ersten Teilnehmerkohorte auf das offizielle Zertifikat gilt es zu gewährleisten.

### 3. Resümee

Das DH-Zertifikat für die Geschichtswissenschaften und die Kunstgeschichte an der LMU, welches sich derzeit in der Konzeptionsphase befindet, versteht sich nicht als konkurrierendes Angebot zu grundständigen Digital Humanities-Studiengängen, sondern vielmehr als deren komplementäre Ergänzung. Die Vermittlung von IT-Grundlagenwissen in der Breite geisteswissenschaftlicher Fächer wirkt einem Auseinanderdriften der DH und den Fachwissenschaften entgegen. Gleichzeitig wird auf den zunehmenden Bedarf an Kommunikationsfähigkeit in interdisziplinären Teams (Digitale Alphabetisierung), an Befähigung zu einer digitalen Quellenkritik und der darauf aufbauenden Algorithmenkritik reagiert. Nicht zuletzt wird der sich verändernden Arbeitsmarktsituation in den Geisteswissenschaften – wie sie mit Hilfe der exemplarischen Auswertung von H-Soz-Kult dargestellt wurde – Rechnung getragen. Der klassische geisteswissenschaftliche Arbeitsmarkt erwartet bereits jetzt von knapp 20% der Absolventen

53 Henrich 2015, S. 15: Natürlich wäre es „ideal, einen Studiengang nach fachlich und didaktischen Gesichtspunkten am Reißbrett planen zu können. In der Praxis sind die Gestaltungsmöglichkeiten aber geringer.“

informationstechnologisches Know-How. Auch für den Arbeitsmarkt jenseits der Wissenschaft ist diese Zusatzqualifikation von großem Vorteil.

Unabhängig von der finalen Ausgestaltung des Curriculums besitzen zwei Elemente grundlegende Bedeutung:

- 1) Der Fachbezug, das heißt die enge Verknüpfung der IT-Lernziele mit geisteswissenschaftlichen Fragestellungen, und
- 2) ihre praxisbezogene Vermittlung. Hierzu wird auf die bestehende Infrastruktur des *DHVLab* aufgebaut, die beide Aspekte in sich vereint.

Da die Plattform auch in anderen, im Rahmen des ‚Digitalen Campus Bayern‘ neu geschaffenen Ausbildungsvorhaben zum Einsatz kommen soll und nachhaltig optimiert wird, darf mit wertvollen Synergien zwischen den beteiligten Fächern gerechnet werden. Durch die Weiterführung und Institutionalisierung bestehender Ansätze innerhalb der genannten Fächer und die Nachnutzung polyvalenter Angebote anderer Fachbereiche wird der Grundstein für eine sukzessive Ausweitung der IT-Grundausbildung gelegt.

## Literaturverzeichnis

- Blevins, Cameron: Learning by doing: Labs and pedagogy in the Digital Humanities (<http://www.cameronblevins.org/posts/learning-by-doing/>), 5. Februar 2013.
- Dinsman, Melissa: The Digital in the Humanities: An interview with Marisa Parham, auf: LARB (<https://lareviewofbooks.org/article/digital-humanities-interview-marisa-parham/>), 19. Mai 2016.
- Ehrlicher, Hanno: Fingerübungen in Digitalien: Erfahrungsbericht eines teilnehmenden Beobachters der Digital Humanities aus Anlass eines Lehrexperiments, in: Romanische Studien 4 (2016), S. 623–636 (<http://www.romanischestudien.de/index.php/rst/article/view/88>).
- Fickers, Andreas: Der ultimative Klick? Digital Humanities, Online-Archive und die Arbeit des Historikers im digitalen Zeitalter, in: Forum für Politik, Gesellschaft und Kultur in Luxemburg 337 (2014), S. 25–29 (<http://hdl.handle.net/10993/21285>).
- FutureManagementGroup AG: Zukunftsmarkt ‚Digitale Bildung‘. Vom Bildungswesen zum Bildungsmarkt. Eltville 2015 (Market Foresights 07-2015) (37 S.; [http://www.futuremanagementgroup.com/wp-content/uploads/2017/01/MF\\_Zukunftsmarkt-Digitale-Bildung.pdf](http://www.futuremanagementgroup.com/wp-content/uploads/2017/01/MF_Zukunftsmarkt-Digitale-Bildung.pdf)).
- Henrich, Andreas: Die Bamberger Studiengänge zu Digital Humanities und ihr Bezug zum Referenzcurriculum, in: Schubert, Zoe (Red.): Digital Humanities als Beruf. Fortschritte auf dem Weg zu einem Curriculum. Akten der DHd-Arbeitsgruppe „Referenzcurriculum Digital Humanities“. Graz 2015, S. 7–22 ([https://www.digitalhumanities.tu-darmstadt.de/fileadmin/dhdarmstadt/materials/Digital\\_Humanities\\_als\\_Beruf\\_-\\_Stand\\_2015.pdf](https://www.digitalhumanities.tu-darmstadt.de/fileadmin/dhdarmstadt/materials/Digital_Humanities_als_Beruf_-_Stand_2015.pdf)).
- Hermes, Jürgen/Schandock, Michael: Stellenanzeigenanalyse in der Qualifikationsentwicklungsforschung. Die Nutzung maschineller Lernverfahren zur Klassifikation von Textabschnitten. Bonn 2016 (<https://www.bibb.de/veroeffentlichungen/de/publication/show/8146>).

- Jakacki, Diane: How we teach? Digital Humanities pedagogy in an imperfect world. Keynote at CDSH/SCHN Conference, Calgary (<http://dianejakacki.net/how-we-teach-digital-humanities-pedagogy-in-an-imperfect-world/>), 5. Juni 2016.
- Klinke, Harald: Big Image Data within the Big Picture of Art History, in: DAH 2 (2016), S. 15–36 (<http://journals.ub.uni-heidelberg.de/index.php/dah/article/download/33527/27215>).
- König, Mareike: Was sind Digital Humanities? Definitionsfragen und Praxisbeispiele aus der Geschichtswissenschaft, auf: DHDHI (<http://dhdhi.hypotheses.org/2642>), 19. Februar 2016.
- Krajewski, Markus: Programmieren als Kulturtechnik, in: H-Soz-Kult, 30.11.2015 (4 S.; <http://www.hsozkult.de/debate/id/diskussionen-2901>).
- Pannacker, William A.: Stop calling it „Digital Humanities“ and 9 other strategies to help liberal-arts colleges join the movement, in: The Chronicle of Higher Education (<http://www.chronicle.com/article/Stop-Calling-It-Digital/137325/>), 18. Februar 2013.
- Rapp, Andrea/Bartsch, Sabine/Borek, Luise: Aus der Mitte der Fächer, in: die Mitte der Fächer: Studiengänge und Curricula – Digital Humanities in der universitären Lehre, in: Bibliothek – Forschung und Praxis 40,2 (2016), S. 172–178 (<https://doi.org/10.1515/bfp-2016-0030>).
- Rehbein, Malte: Geschichtsforschung im digitalen Raum. Über die Notwendigkeit der Digital Humanities als historische Grund- und Transferwissenschaft (24 S.; im Druck. Preprint vom 16. April 2015: [http://web.archive.org/web/20170815095246/http://www.phil.uni-passau.de/fileadmin/dokumente/lehrstuehle/rehbein/Dokumente/GeschichtsforschungImDigitalenRaum\\_preprint.pdf](http://web.archive.org/web/20170815095246/http://www.phil.uni-passau.de/fileadmin/dokumente/lehrstuehle/rehbein/Dokumente/GeschichtsforschungImDigitalenRaum_preprint.pdf)).
- Rehbein, Malte: Digital Humanities an der Universität Passau – Eine Fallstudie, in: Schubert, Zoe (Red.): Digital Humanities als Beruf. Fortschritte auf dem Weg zu einem Curriculum. Akten der DHd-Arbeitsgruppe „Referenzcurriculum Digital Humanities“. Graz 2015, S. 76–88 ([https://www.digitalhumanities.tu-darmstadt.de/fileadmin/dhdarmstadt/materials/Digital\\_Humanities\\_als\\_Beruf\\_-\\_Stand\\_2015.pdf](https://www.digitalhumanities.tu-darmstadt.de/fileadmin/dhdarmstadt/materials/Digital_Humanities_als_Beruf_-_Stand_2015.pdf)).

- Sahle, Patrick: DH studieren! Auf dem Weg zu einem Kern- und Referenzcurriculum der Digital Humanities. Göttingen 2013 (DARIAH-DE Working Papers 1) (<http://resolver.sub.uni-goettingen.de/purl/?dariah-2013-1>).
- Sahle, Patrick: Digital Humanities als Beruf. Wie wird man ein „Digital Humanist“, und was macht man dann eigentlich?, in: Akademie Aktuell 56 (2016), S. 78–83 ([https://www.badw.de/fileadmin/pub/akademieAktuell/2016/56/0116\\_18\\_Sahle\\_V04.pdf](https://www.badw.de/fileadmin/pub/akademieAktuell/2016/56/0116_18_Sahle_V04.pdf)).
- Sahle, Patrick: Zur Professoralisierung der Digital Humanities, auf: DHd-Blog. Digital Humanities im deutschsprachigen Raum (<http://dhd-blog.org/?p=6174>), 23. März 2016.
- Schlotheuber, Eva/Bösch, Frank: Quellenkritik im digitalen Zeitalter Die Historischen Grundwissenschaften als zentrale Kompetenz der Geschichtswissenschaft und benachbarter Fächer, auf: VHD-Blog (4 S.; PDF-Version: [http://www.historikerverband.de/fileadmin/user\\_upload/vhd\\_journal\\_2015-04\\_beileger.pdf](http://www.historikerverband.de/fileadmin/user_upload/vhd_journal_2015-04_beileger.pdf)), 30. Oktober 2015.
- Schumacher, Mareike u.a. (Hg.): Big Data in den Geisteswissenschaften: Konzept für eine Lehr- und Lernmittelsammlung. Göttingen 2016 (DARIAH-DE Working Papers 15) (<http://nbn-resolving.de/urn:nbn:de:gbv:7-dariah-2016-1-2>).
- Shirky, Clay: Here comes everybody. The power of organizing without organizations. New York 2008.
- Spiro, Lisa: Opening up Digital Humanities education, in: Hirsch, Brett D. (Hg.): Digital Humanities pedagogy. Practice, principles and politics. Cambridge 2012 (Digital Humanities Series 3), S. 331–363 (<http://books.openedition.org/obp/1654>).
- Strobel, Maria/Welpe, Isabell M.: Hochschule 4.0. Die Zukunft der Hochschule erfinden, in: Forschung & Lehre 23,4 (2017), S. 316–318 (<http://www.wissenschaftsmanagement-online.de/beitrag/hochschule-40-die-zukunft-der-hochschule-erfinden-7648>).
- Thaller, Manfred: Digital Humanities als Beruf – Fortschritte auf dem Weg zu einem Curriculum, in: Schubert, Zoe (Red.): Digital Humanities als Beruf. Fortschritte auf dem Weg zu einem Curriculum. Akten der DHd-Arbeitsgruppe „Referenzcurriculum Digital Humanities“.

- Graz 2015, S. 3–5 ([https://www.digitalhumanities.tu-darmstadt.de/fileadmin/dhdarmstadt/materials/Digital\\_Humanities\\_als\\_Beruf\\_-\\_Stand\\_2015.pdf](https://www.digitalhumanities.tu-darmstadt.de/fileadmin/dhdarmstadt/materials/Digital_Humanities_als_Beruf_-_Stand_2015.pdf)).
- Thaller, Manfred: Informatik und Geisteswissenschaften in Köln – eine Fallstudie, in: Schubert, Zoe (Red.): Digital Humanities als Beruf. Fortschritte auf dem Weg zu einem Curriculum. Akten der DHd-Arbeitsgruppe „Referenzcurriculum Digital Humanities“. Graz 2015, S. 63–75 ([https://www.digitalhumanities.tu-darmstadt.de/fileadmin/dhdarmstadt/materials/Digital\\_Humanities\\_als\\_Beruf\\_-\\_Stand\\_2015.pdf](https://www.digitalhumanities.tu-darmstadt.de/fileadmin/dhdarmstadt/materials/Digital_Humanities_als_Beruf_-_Stand_2015.pdf)).
- Townsend, Robert B./Swafford, Emily: Conflicting signals in the academic job market for history, in: Perspectives on History 55,1 (2017) (<https://www.historians.org/publications-and-directories/perspectives-on-history/january-2017/conflicting-signals-in-the-academic-job-market-for-history>).
- Wettlaufer, Jörg: Digital Humanities: Eigene Disziplin oder Hilfswissenschaft?, auf: digihum.de (<http://digihum.de/2014/01/digital-humanities-eigene-disziplin-oder-hilfswissenschaft/>), 13. Januar 2014.
- Wissenschaftsrat: Empfehlungen zu Karrierezielen und -wegen an Universitäten. Dresden 2014 (<https://www.wissenschaftsrat.de/download/archiv/4009-14.pdf>).



# Digital Humanities Sprachwissenschaften

## Schwerpunkte bei der Entwicklung eines sprachwissenschaftlichen DH-Nebenfaches

Markus Frank

### 1. Eine Positionsbestimmung

Was die Vermittlung digitaler Methoden in den sprachwissenschaftlichen Studiengängen angeht, so ist die Situation im Jahr 2017 noch relativ ernüchternd: Obwohl besonders linguistische Spitzenforschung nicht mehr ohne fundamentale Kenntnisse digitaler Techniken betrieben werden kann, existieren keine systematischen Ausbildungs- und Studienprogramme zur Vermittlung der erforderlichen Kenntnisse. Sofern digitale Komponenten in sprachwissenschaftlichen Projekten vorhanden sind, werden diese zumeist entweder von genuin fachfremden Wissenschaftlern (Informatikern, Computerlinguisten, Statistikern) oder von Autodidakten konzipiert und umgesetzt<sup>1</sup>. Auch wenn fachwissenschaftliche Kooperationen mit anderen Wissenschaftsbereichen sich als äußerst gewinnbringend erweisen, besteht dennoch ein unmittelbarer Bedarf an Linguisten mit informationstechnischer Zusatzausbildung. Fehlen grundlegende Kenntnisse im Bereich der Programmierung, Datenstrukturierung und Datenanalyse, so erschwert dies nicht nur die Kooperation mit eben diesen fachfremden Wissenschaftlern, es besteht auch die Gefahr, dass gewinnbringende neue Methoden und Ansätze nicht entwickelt werden, da den bearbeitenden Linguisten das Verständnis für die Integration digitaler

---

<sup>1</sup> Siehe beispielsweise die Projekte Verba Alpina (<http://www.verba-alpina.gwi.uni-muenchen.de/>, Version 17/1 13.07.2017) oder eDiAna (Miller et al. 2017) an der Ludwig-Maximilians-Universität.

Techniken in ihren Forschungsalltag fehlt. Nur die Kenntnis beider Bereiche begünstigt einen gewinnbringenden Transfer und infolgedessen synergistische Effekte.

Dieser Umstand macht die Entwicklung eines sprachwissenschaftlichen Nebenfaches, welches die Vermittlung der benötigten digitalen Techniken als Kern seines Curriculums ansieht, zu einem besonderen Anliegen bei der Ausbildung zukünftiger Linguisten. Im Folgenden werden die Schwerpunkte eines solchen Curriculums kurz skizziert und im Anschluss das so entstehende Fach von der Computerlinguistik abgegrenzt, welche von ihren Forschungsinteressen als auch von den vermittelten Inhalten grundsätzlich anders gelagert ist als das hier entworfene Fach.

## 2. Schwerpunkte des Curriculums

Das Nebenfach „Digital Humanities - Sprachwissenschaft“ (DH-S) wird um eine fest definierte Menge curricularer Kernelemente herum entworfen, welche dem Studiengang sein spezifisches Profil verleihen:

1. Metareflexion
2. Projektorientierung
3. Die ‚technische Trias‘
4. Fachtransfer
5. Wahlfreiheit

### 2.1 Metareflexion

Der Bereich Metareflexion umfasst einerseits die Beschäftigung mit grundlegenden Theorien und thematischen Schriften der Digital Humanities, andererseits wird die Sonderstellung der Linguistik im geisteswissenschaftlichen Fächerkanon diskutiert. Von zentraler Bedeutung im Bereich der Metareflexion sind Fragestellungen zur Veränderung der geisteswissenschaftlichen Forschungslandschaft durch die Digital Humanities. Ermöglicht die Digitalisierung in diesem Bereich tatsächlich neue Forschungsperspektiven, oder handelt es sich lediglich um die Fortführung bereits etablierter Forschungspara-

digmen mit den Mitteln der digitalen Implementierung? Wie ist das Verhältnis der theoretischen Basis zu den neu entstehenden digitalen Objekten?

Humanists may learn from the content of such objects [digitized artefacts, M.F.], treated as research materials, as they always have. These objects will serve as more or less effective surrogates for their physical originals and may produce efficiencies of access and other practical benefits of one sort or another. But they have no contribution to make to humanities scholarship: they make no intervention, they leave no intellectual mark.<sup>2</sup>

Problemstellungen dieser Art werden ebenso behandelt, wie die komplexen Interaktionsphänomene von Informatik und geisteswissenschaftlicher Forschung: *“Interestingly, the humanities are already very much like computing [...], in that they primarily study human-created artifacts.”*<sup>3</sup>

Die Verwendung digitaler Methoden im Fachbereich der Linguistik wird im Kontext dieser Überlegungen verstärkt analysiert und diskutiert. Besonders in der Sprachwissenschaft ist zu beobachten, dass auf Kosten klassischer introspektiver Verfahren immer mehr datengetriebene Verfahren zum Einsatz kommen, sei es im Bereich der digitalen Korpuslinguistik<sup>4</sup>, der experimentellen Psycho- und Neurolinguistik oder, besonders prominent, in der Computerlinguistik. Große Teile der Sprachwissenschaft scheinen sich im Wandel, hin zu einer massiv auf digitale Methoden setzende, stark von den ‚exakten Wissenschaften‘ und der Psychologie beeinflussten Datenwissenschaft zu befinden.<sup>5</sup> Die Auseinandersetzung mit diesen Veränderungen wird als unerlässlicher Bestandteil der Metareflexion angesehen.

---

2 Flanders 2013, S. 210–211.

3 Rosenbloom 2013, S. 224.

4 Siehe hierzu exemplarisch Glynn & Fischer 2010.

5 Siehe Hajič 2004, S. 79–87.

## 2.2 Projektorientierung

Die primäre Arbeitsform in der Sprachwissenschaft stellt das Forschungsprojekt dar. Projekte reichen hierbei von einfachen Experimenten oder Fragebogenstudien, wie sie in wissenschaftlichen Qualifizierungsarbeiten üblich sind, über umfangreiche korpuslinguistische Studien bis hin zu sprachwissenschaftlichen Großprojekten mit mehrjähriger Laufzeit und großem personellem Aufwand. Mögen sich digitale Projekte in ihrer Dimension und inhaltlichen Zielsetzung zwar unterschiedlich gestalten, so sind sie dennoch von einer Reihe von Gemeinsamkeiten gekennzeichnet: Zu Beginn steht die fachwissenschaftliche Analyse des Forschungsbereiches, die Entwicklung von Arbeits- und Ablaufplänen, das Entwerfen der digitalen Infrastruktur und nicht zuletzt die Etablierung eines funktionsfähigen und effektiven Workflows. Eine komplexe Forschungsfrage muss in bearbeitbare Teilkomponenten zerlegt und das digitale Forschungsdesign so gewählt werden, dass sich die sprachlichen Untersuchungsgegenstände mithilfe der digitalen Techniken sowohl abbilden, als auch in der intendierten Weise untersuchen lassen. Zu diesem Zweck legt das Curriculum besonderen Wert auf den Transfer der theoretisch vermittelten Methoden in konkrete Forschungsprojekte: Die Studierenden konzeptionieren und realisieren im Rahmen des Studiums eigene digitale Forschungsvorhaben, deren wissenschaftlicher Anspruch und deren Umsetzbarkeit im Vorfeld von den Lehrenden überprüft werden.

Einen weiteren Grundstein im Rahmen der Projektorientierung stellt die Ausbildung an wechselnden großen sprachwissenschaftlichen Forschungsprojekten dar, wie sie an der LMU realisiert werden. Die Studierenden erhalten dabei sowohl Einblick in die fachwissenschaftliche Grundlage der vermittelten Forschungsprojekte, als auch in die konkrete technische Umsetzung und in den Planungsprozess. Im Planungsprozess wird darauf geachtet, dass die präsentierten Projekte aus verschiedenen sprachwissenschaftlichen Domänen stammen. So erweist sich beispielsweise eine Kombination aus Korpuslinguistik (NeoCrawler<sup>6</sup>), Geolinguistik (Verba Alpina) und digitaler Lexiko-

---

6 Kerremans, Stegmayr & Schmid 2012.

graphie (eDiAna) als geeignete Kombination, um die Themenbereiche einer Ringvorlesung zu bilden.

### 2.3 Die ‚technische Trias‘

Die Vermittlung fundamentaler technischer Kenntnisse ist für einen Digital Humanities Studiengang mit sprachwissenschaftlichem Schwerpunkt unerlässlich. Durch die Veränderung des Fachbereiches stellt die Kenntnis digitaler Techniken nicht lediglich eine wünschenswerte Zusatzqualifikation dar, sondern eine unerlässliche Kernkompetenz:

Thus it is not enough to be able to create a WWW site for a course; building the site is like being able to type a paper, it is the threshold of participation.<sup>7</sup>

Ohne ein grundlegendes Repertoire an digitalen Fähigkeiten werden Studierende zukünftig kaum noch in der Lage sein, sich in moderne Forschungsbereiche einzuarbeiten oder selbständig Forschung zu betreiben. Da der Umfang eines Nebenfach-Curriculums stark begrenzt ist, werden im informationstechnischen Bereich nur solche Inhalte zur Vermittlung ausgewählt, die sich in der konkreten sprachwissenschaftlichen Forschungspraxis als besonders nützlich erwiesen haben. Das Ergebnis des Auswahlprozesses ist eine Schwerpunktsetzung in drei informationstechnische Bereiche, welche in der Forschungspraxis eng mit einander verzahnt sind:

1. Prozedurale und objektorientierte Programmierung
2. Datenbanken und Markup-Sprachen
3. Statistische Datenanalyse und Visualisierung

Zur erfolgreichen Realisierung digitaler sprachwissenschaftlicher Forschung werden in der Regel alle drei Schwerpunktbereiche benötigt: Kenntnisse in der prozeduralen und objektorientierten Programmierung (Python) ermöglichen es den Studierenden, Korpora oder andere

---

<sup>7</sup> Rockwell 2013, S. 27f.

sprachwissenschaftliche Datensätze zu formatieren, in andere Formate zu konvertieren oder Parser anzuwenden. Sie dienen als Integrations-schnittstelle zwischen allen eingesetzten technischen Elementen des Forschungsvorhabens. Zur Speicherung der Daten und zur systematischen Abfrage und Evaluation derselben werden Datenbankformate (SQL) und Markup-Sprachen (XML) vermittelt. Die Kenntnisse gängiger Formate und Techniken in diesem Bereich sind sowohl im Hinblick auf zukünftiges Forschungsdatenmanagement als auch für die Gestaltung weitreichend kompatibler und evaluierbarer Datenformate unerlässlich. Fähigkeiten zur statistischen Datenanalyse und Visualisierung von sprachwissenschaftlichen Forschungsdaten runden den informationstechnischen Pflichtteil des Curriculums ab: Die moderne Sprachwissenschaft macht extensiven Gebrauch von statistischen Methoden, seien es probabilistische Analysen zur Identifikation von sprachlichen Mustern in Korpora, oder klassische teststatistische Methoden zur Hypothesenevaluation, wie sie in der Psychologie oder den empirisch forschenden Sozialwissenschaften bereits lange erfolgreich eingesetzt werden. Insofern haben auch die Sprachwissenschaftler, wenn sie empirisch forschen, einen großen Bedarf an statistischen Modellen und, damit verbunden, auch an Methoden zur grafischen Visualisierung der beobachteten Muster und Zusammenhänge.

Als integriertes Gesamtkonzept sollen die drei informationstechnischen Bereiche die Studierenden in die Lage versetzen, eigene Forschungsprojekte vollkommen eigenständig zu entwickeln, die Daten zu modellieren und schließlich die statistische Evaluation der Ergebnisse zu liefern.

## 2.4 Fachtransfer

Wie bereits in Kapitel 2.3 erläutert, besitzt das Nebenfach einen methodischen Schwerpunkt und bezieht seine konkreten Anwendungsbereiche und Forschungsinteressen direkt aus den sprachwissenschaftlichen Disziplinen, welchen es als Nebenfach zugeordnet ist. Angestrebt wird ein bidirektionaler Austausch zwischen den Inhalten des sprachwissenschaftlichen Hauptfaches und den Forschungsmethoden, die im Nebenfach vermittelt werden. Da das Nebenfach selbst

keine Forschungsinhalte liefert, müssen konkrete Fragestellungen und Forschungsgegenstände aus den sprachwissenschaftlichen Disziplinen exportiert werden. Im Rahmen des Nebenfaches lernen die Studierenden, sprachwissenschaftliche Forschungsfragen so zu formulieren, dass sie sich mithilfe digitaler Methoden erforschen lassen. Die so entwickelten digitalen Methoden und Umsetzungsformen exportieren ihre Ergebnisse im Anschluss zurück in die Fachwissenschaft, wo sie interpretiert und ausgewertet werden. Ziel dieser Methode ist die Heranbildung von Sprachwissenschaftlern, die ihre jeweilige Fachwissenschaft in enger Verzahnung zu den digitalen Forschungsmethoden sehen, welche sie in ihrer alltäglichen Arbeit anwenden.

Diese angestrebte enge Verzahnung zum sprachwissenschaftlichen Hauptfach unterscheidet den Nebenfach-Studiengang DH-S von anderen informationstechnischen Nebenfachoptionen wie der Informatik: Der Transfer von Wissen aus der Informatik in sprachwissenschaftliche Fragestellungen hinein gelingt den Studierenden erfahrungsgemäß nur in den seltensten Fällen, da beide Wissenschaftsbereiche, zumindest in derjenigen Form, wie sie im universitären Kontext vermittelt werden, nur geringe Überschneidungsbereiche besitzen. An dieser Stelle setzt das Nebenfach an.

## 2.5. Wahlfreiheit

Im Rahmen des Studiums erhalten die Studierenden die Möglichkeit, aus einer festgelegten Menge an Wahlpflichtbereichen nach eigenem Ermessen solche Bereiche auszuwählen, die ihren Interessen und Fähigkeiten am ehesten entsprechen. Neben der Ermöglichung von Wahlfreiheit dient dieses Angebot auch zur Verbreiterung des curricularen Inhaltsangebotes, welches durch den verhältnismäßig geringen Umfang eines Nebenfaches (60 ECTS) beschränkt wird. Die Studierenden wählen zwei aus vier Wahlpflichtbereichen, deren Inhalte in enger Absprache mit den Fachwissenschaften entwickelt und angeboten oder sogar aus diesen direkt importiert werden:

- **Spezialisierung Korpuslinguistik:** Korpuslinguistische Untersuchungen stellen in allen Bereichen der Linguistik eine primäre Forschungsmethode dar. Vermittelt werden der Umgang

mit vorhandenen Korpora geschriebener und gesprochener Sprache, die Erstellung eigener digitaler Korpora sowie die automatische, halbautomatische und manuelle Annotation von Korpusdaten. Auch theoretische Ansätze zur Korpuslinguistik (zum Beispiel ‚Corpus-Driven‘ versus ‚Corpus-Based‘<sup>8</sup>) sowie Auswertungsmethoden werden in diesem Spezialisierungsbereich vermittelt.

- **Spezialisierung Sprachtechnologie und Statistik:** Dieser Wahlpflichtbereich dient zur Spezialisierung in digitale linguistische Forschungsmethoden, wie sie primär in der kognitiven und der Neurolinguistik zum Einsatz kommen, angelehnt an die empirischen Forschungsmethoden der Psychologie und der empirisch forschenden Sozialwissenschaften. Vermittelt wird der Umgang mit technischen Hilfsmitteln (z.B. Eye-Trackern) zur Erfassung sprachwissenschaftlicher Daten, ebenso wie die Erstellung von Fragebögen oder linguistischen Experimenten, sofern deren Ergebnisse aus analysebedürftigen digitalen Forschungsdaten bestehen. Abgerundet wird dieser Bereich durch eine Vertiefung der statistischen Methoden, welche bereits im Pflichtcurriculum verankert sind.
- **Spezialisierung Sprachvariation und Sprachwandel:** Innerhalb dieses Wahlbereiches lernen die Studierenden den Umgang mit räumlich und/oder zeitlich repräsentierbaren Sprachdaten, welche in der Variationslinguistik, der historischen Linguistik und der Geolinguistik üblicherweise erhoben werden. Vermittelt wird neben der Verwendung bereits vorhandener Ressourcen auch die Erstellung, Kodierung, Abbildung und Analyse von linguistisch relevanten Geo- und Temporaldaten.
- **Spezialisierung Webtechnologie:** Viele große linguistische Forschungsprojekte werden heute als interaktive, browserbasierte Webapplikationen realisiert, da die HTML/CSS-Standards zusammen mit für den Webeinsatz optimierten Programmiersprachen wie JavaScript die Erstellung von

---

8 Siehe McEnery et al. 2006, 8–10.

Benutzeroberflächen fundamental erleichtern. Vermittelt werden in diesem Bereich die Fähigkeiten zur strukturierten Erstellung und Realisierung und Dokumentation von sprachwissenschaftlichen Webapplikationen.

### 3. Abgrenzung zur Computerlinguistik

Das Nebenfach grenzt sich in vielen Aspekten sowohl inhaltlich als auch in der methodischen Ausbildung von den klassischen Studieninhalten der Computerlinguistik ab. Einen Einblick in das Selbstverständnis der Computerlinguistik liefern diverse Fachprofile, welche von Universitäten und Gesellschaften für dieses Fach definiert wurden:

Computerlinguistik erforscht die maschinelle Verarbeitung natürlicher Sprachen. Sie erarbeitet die theoretischen Grundlagen der Analyse, Erkennung und Erzeugung geschriebener und gesprochener Sprache durch Maschinen.<sup>9</sup>

Computerlinguistik untersucht, wie menschliche Sprache von Computern verarbeitet und interpretiert werden kann. Sie erforscht die mathematischen und logischen Eigenschaften natürlicher Sprache und entwickelt algorithmische und statistische Verfahren zur automatischen Sprachverarbeitung.<sup>10</sup>

Computational linguistics is the scientific study of language from a computational perspective. Computational linguists are interested in providing computational models of various kinds of linguistic phenomena. These models may be “knowledge-based” (“hand-crafted”) or “data-driven” (“statistical” or “empirical”).<sup>11</sup>

---

9 Informationsmaterial des Studiengangs Computerlinguistik, Centrum für Informations- und Sprachverarbeitung (CIS), LMU München, <http://www.cis.uni-muenchen.de/download/flyer/cisflyer.pdf> (13.07.2017).

10 Informationsmaterial des Studiengangs Computerlinguistik, Institut für Computerlinguistik, Universität Heidelberg, [http://www.cl.uni-heidelberg.de/studies/was\\_ist\\_computerlinguistik.pdf](http://www.cl.uni-heidelberg.de/studies/was_ist_computerlinguistik.pdf) (13.07.2017).

11 What is computational linguistics?, Association for Computational Linguistics, <https://www.aclweb.org/portal> (13.07.2017).

Der Kern der Forschung liegt in der automatischen Informationsextraktion, der Informationssuche in Texten, der automatischen Vorschlagwortung und Textklassifikation, der maschinellen Übersetzung geschriebener und gesprochener Sprache sowie in der logisch-mathematischen Beschreibung von Sprache. Als Resultat dieser Schwerpunktsetzung steht unter anderem die Entwicklung und Optimierung von Sprachparsern für verschiedene Einsatzszenarien im Zentrum der computerlinguistischen Ausbildung und Forschung. In diesem Sinne ist die Computerlinguistik eher als Teilbereich der Informatik zu sehen, deren eigenständige Forschungsfragen in vielen Bereichen nicht mit den spezifischen Forschungsfragen der linguistischen Fachwissenschaften konvergieren.

Das Nebenfach DH-S hingegen beschäftigt sich nicht mit der Entwicklung automatischer Klassifizierungsverfahren für sprachliche Artefakte, auch wenn der Umgang mit Produkten computerlinguistischer Forschung, seien es POS-Tagger oder OCR-Verfahren, im Studiengang vermittelt wird. Das Nebenfach sieht sich vielmehr als primär methodenorientierter Studiengang, welcher Studierenden der linguistischen Fachwissenschaften – sofern diese Bereiche mit digitalen Datenstrukturen arbeiten – die nötigen digitalen Grundkenntnisse für breite linguistische Anwendungsbereiche vermitteln soll. Digitale sprachwissenschaftliche Forschungsdaten liegen in vielen Fällen als Ergebnisse von Fragebogenstudien, linguistischen Laborexperimenten (zum Beispiel Eye-Tracking und Reaktionszeitmessungen), in georeferenzierter Form oder auch als manuell annotierte Sprachkorpora vor. Diese Daten müssen digital aufbereitet, in viablen Datenstrukturen gespeichert, statistisch analysiert und visualisiert werden, auch wenn sie nicht mittels vollautomatischer Prozeduren gewonnen wurden. Das Nebenfach ist in dieser Konzeption nicht in der Lage, eigene Forschungsinteressen zu entwickeln, wie dies in der Computerlinguistik problemlos möglich ist. Alle vermittelten Inhalte speisen sich direkt aus den Bedürfnissen der sprachwissenschaftlichen Hauptfächer, sie liefern die Hypothesen und theoretischen Rahmensetzungen, im Nebenfach werden Methoden und Ansätze gelehrt, digi-

tale Werkzeuge zur Überprüfung dieser Fragestellungen einzusetzen und bei Bedarf eigenständig zu entwickeln. Dass es dabei auch zum Gebrauch computerlinguistischer Techniken kommt ist unvermeidbar und erwünscht, die Aneignung genuin computerlinguistischer Forschungsinteressen ist jedoch nicht Ziel des hier entworfenen Faches.

Wenn im DH-S-Nebenfach klassische Korpuslinguistik betrieben wird, so geschieht dies im Gegensatz zur Computerlinguistik stets unter einem Corpus-Based-Paradigma, nicht jedoch als Corpus-Driven Ansatz, wie er in der Computerlinguistik, unterstützt von Algorithmen des maschinellen Lernens und ohne Vorerwartung das Analyseergebnis betreffend, durchaus der Fall ist. Vermittelt wird der Einsatz von Textkorpora gesprochener und geschriebener Sprache als Werkzeug zur Überprüfung von Hypothesen, welche anhand linguistischer Theorien gebildet, analysiert und je nach Ergebnislage zur Unterstützung, partiellen Reformulierung oder Verwerfung der theoretischen Position genutzt werden sollen.

Zusammenfassend lässt sich folglich sagen: Das Nebenfach sieht sich in der Tradition klassischer empirischer Forschungsparadigmen, in deren Kern die Formulierung von Theorien und deren anschließende Überprüfung durch den Prozess der Beobachtung steht. Alle Forschungsdaten liegen in digitaler Form vor und benötigen profunde Kenntnisse im Bereich Programmierung, Datenstrukturierung und statistischer Datenanalyse, um sie zu erheben, nachhaltig zu speichern und letztendlich auszuwerten.

#### **4. Ausblick**

Mithilfe des Nebenfach-Curriculums soll nicht nur der fortschreitenden Digitalisierung der Fachwissenschaft bereits in der wissenschaftlichen Ausbildung Rechnung getragen werden, vielmehr ist auch die Stärkung der Digital Humanities im universitären Alltagsbetrieb ein Hauptanliegen. Die Erkenntnisse aus der Entwicklung eines Nebenfaches, welches zum Zweck der methodischen Ergänzung speziell für

die Bedürfnisse eines großen fachwissenschaftlichen Forschungsparadigmas entworfen wurde, werden sich für zukünftige ähnlich gelagerte Vorhaben als äußerst nützlich erweisen.

## Literaturverzeichnis

- Flanders, Julia: The Productive Unease of 21st-century Digital Scholarship, in: Terras, Melissa/Nyhan, Julianne/Vanhoutte, Edward (Hrsg.): *Defining Digital Humanities*, Farnham/Burlington 2013, S. 205–218.
- Glynn, Dylan/Fischer, Kerstin (Hrsg.): *Quantitative Methods in Cognitive Semantics: Corpus-Driven Approaches*, Berlin/New York 2010.
- Hajič, Jan: Linguistics Meets Exact Sciences, in: Schreibman, Susan/Siemens, Ray/Unsworth, John: *A Companion to Digital Humanities*, Malden/Oxford/Charlton 2004, S. 79–87.
- Kerremans, Daphné/Stegmayr, Susanne/Schmid, Hans-Jörg: The Neo-Crawler. Identifying and retrieving neologisms from the internet and monitoring on-going change, in: Allan, Kathryn/Robinson, Justyna (Hrsg.): *Current methods in historical semantics*, Berlin 2012, S. 59–96.
- McEnery, Tony/Xiao, Richard/Tono, Yukio: *Corpus-based language studies: An advanced resource book*, Abingdon 2006.
- Miller, Jared & Rieken, Elisabeth & Hackstein, Olav & Yakubovich, Ilya (eds.) 2017: eDiAna. <https://www.ediana.gwi.uni-muenchen.de/>
- Rockwell, Geoffrey: Is Humanities Computing an Academic Discipline?, in: Terras, Melissa/Nyhan, Julianne/Vanhoutte, Edward (Hrsg.): *Defining Digital Humanities*, Farnham/Burlington 2013, S. 13–34.
- Rosenbloom, Paul: Toward a Conceptual Framework for the Digital Humanities, in: Terras, Melissa/Nyhan, Julianne/Vanhoutte, Edward (Hrsg.): *Defining Digital Humanities*, Farnham/Burlington 2013, S. 219–236

# Digital Humanities an der Universität Regensburg

## Geschichte – Projekte – Studiengang

Manuel Burghardt, Daniel Isemann, Markus  
Kattenbeck, Bernd Ludwig und Christian Wolff

### 1. Einleitung

Die Effekte der Digitalisierung wirken sich auf alle Bereiche des öffentlichen und privaten Lebens, die Politik, die Gesellschaft und auch auf den akademischen Bereich aus (vgl. dazu den *Computational Turn*<sup>1</sup>). Während in den Lebens- und Naturwissenschaften der Einsatz quantitativ-empirischer, computergestützter Methoden eine lange Tradition hat, öffnen sich vor dem Hintergrund einer sich rasant entwickelnden Digitalisierung von kulturellen Artefakten in zunehmendem Maße auch die Geisteswissenschaften für den Einsatz digitaler Tools und Methoden. Diese Entwicklung in den traditionell eher ideographisch, also das Einzelne beschreibend, arbeitenden Geisteswissenschaften, hin zum Einsatz computergestützter Methoden wird gemeinhin unter dem Begriff der *Digital Humanities* (DH) zusammengefasst.

Konkret ergeben sich dabei Herausforderungen bei der digitalen Repräsentation von nicht unmittelbar quantifizier- und formalisierbaren Kulturobjekten<sup>2</sup> und bei der computerbasierten Analyse solcher digitalisierter Daten, die grundlegend neue Perspektiven für die jeweiligen Fachdisziplinen eröffnen (vgl. etwa den *Distant Reading*-Ansatz in der Literaturwissenschaft<sup>3</sup>). Darüber hinaus werden auf der fachwissenschaftlichen Metaebene in diversen Sammelbänden die epistemologi-

---

1 Vgl. Berry 2011.

2 Unsworth 2002.

3 Moretti 2000.

schen Auswirkungen digitaler, berechnender Methoden auf die traditionell hermeneutisch arbeitenden Geisteswissenschaften diskutiert<sup>4</sup>.

Ziel von Forschung und Lehre im Bereich Digital Humanities an der Universität Regensburg ist es einerseits, die digitale Modellierung und Analyse geisteswissenschaftlicher Forschungsobjekte zu bewerkstelligen, etwa durch die Entwicklung innovativer Software-Werkzeuge oder die Vermittlung technischer Kompetenzen in Lehrveranstaltungen mit primär geisteswissenschaftlich geprägtem Publikum. Andererseits soll mit einem dedizierten Masterstudiengang im Bereich DH und einer geplanten Institutionalisierung, beispielsweise in Form eines DH-Kompetenzzentrums, langfristig ein akademisches Umfeld geschaffen werden, in dem auch die Auswirkungen computerbasierter Ansätze in den Geisteswissenschaften kritisch reflektiert und neu verhandelt werden können.

## 2. Regensburg und die digitalen Geisteswissenschaften

Die Universität Regensburg feiert im Jahr 2018 ihren 50. Geburtstag. Trotz dieses jungen Alters – oder vielleicht gerade deswegen – kann der Standort Regensburg auf eine verhältnismäßig lange Tradition computergestützter geistes- und kulturwissenschaftlichen Disziplinen zurückblicken. So finden sich etwa erste Lehrveranstaltungen auf dem Gebiet der digitalen Geisteswissenschaften schon im Wintersemester 1971/72, damals durchgeführt vom späteren Professor für linguistische Informationswissenschaft in Regensburg, Jürgen Krause (vgl. Abb. 1).

---

<sup>4</sup> Vgl. Berry 2012, Gold 2012, Terras et al. 2013.

33 619	Intensivkurs: Allgemeine Einführung in die Linguistische Datenverarbeitung und in die Handhabung des Regensburger Textauswertungsprogrammes T 4 H 11. - 14. Okt., täglich 9–12 und 14–15	J. Krause
33 620	Einführung in die Programmierung linguistischer Probleme (Fortran IV) - Keine Vorkenntnisse 3 st., Zeit und Ort nach Vereinbarung	J. Krause

Abbildung 1: Ausschnitt aus dem Vorlesungsverzeichnis der Universität Regensburg, WS 1971/72.

Als Volluniversität deckt die Universität Regensburg mit ihren elf Fakultäten ein breites Spektrum an Disziplinen ab. Bemerkenswert ist dabei, dass es in Regensburg zu keinem Zeitpunkt eine Kerninformatik beziehungsweise ein Institut oder eine Fakultät für Informatik gab. Vielmehr ist der technisch-informatische Bereich durch eine Vielzahl dezentraler Informatikkompetenzen vertreten, die den jeweiligen Fakultäten zugeordnet sind. So gibt es im natur- und lebenswissenschaftlichen Bereich die Bio- und Medizininformatik, für die wirtschaftswissenschaftliche Fakultät analog dazu die Wirtschaftsinformatik. In der Fakultät für Sprach-, Literatur- und Kulturwissenschaften wird die Rolle der Fachinformatik von den Lehrstühlen für Informationswissenschaft und Medieninformatik übernommen. Beide Lehrstühle bieten mit ihren individuellen fachlichen Schwerpunktsetzungen, beispielsweise *Information Retrieval* und Informationsverhalten in der Informationswissenschaft<sup>5</sup> sowie *Software Engineering* und *Usability Engineering* in der Medieninformatik<sup>6</sup>, sehr gute Anknüpfungspunkte für transdisziplinäre Kooperationsprojekte auf dem Gebiet der Digital Humanities.

<sup>5</sup> Burghardt et al. 2015.

<sup>6</sup> Wolff 2009.

### 3. Digital Humanities in der Forschung

Vor dem oben beschriebenen Hintergrund konnte in den letzten Jahren ein breites Spektrum an Digital Humanities-Projekten umgesetzt werden.<sup>7</sup> Die so entstandenen Projekte decken dabei unterschiedliche Medientypen und Untersuchungsgegenstände (Text, Bild, Video, Musik, 3D-Objekte, Geoinformation) mit jeweils ganz unterschiedlichen geistes- und kulturwissenschaftlichen Bezugspfaltern (Philologien, Medienwissenschaft, Musikwissenschaft, Wissenschaftsgeschichte, Wirtschafts- und Sozialgeschichte, Kulturwissenschaft, Kunstgeschichte und -erziehung) ab (vgl. Abb. 2).

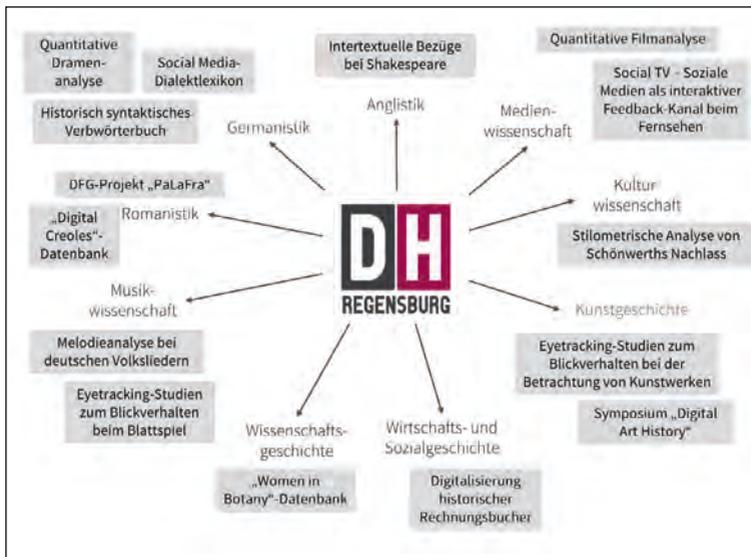


Abbildung 2: Überblick zu Forschungskooperationen der Arbeitsgruppe Digital Humanities.

So wird beispielsweise im aktuellen DFG/ANR-geförderten Projekt *PaLaFra* („Le passage du latin au français“) mit Kolleginnen und Kollegen aus der Romanistik an den Standorten Lille, Lyon, Tübingen

<sup>7</sup> Burghardt & Wolff 2014.

und Regensburg der Übergang vom Latein zum Französischen computergestützt durch den Aufbau und die Analyse eines digitalen Korpus spätlateinischer und altfranzösischer Texte untersucht.<sup>8</sup> Weitere Beispiele für Kooperationen im Bereich der Sprachwissenschaft finden sich in der Germanistik: In der Arbeitsgruppe „Historisch syntaktisches Verbwörterbuch des Deutschen“ (<https://histvw.wordpress.com/>) wird an verschiedenen Standorten (Graz, Heidelberg, Helsinki, Leipzig, Regensburg, Turku und Zürich) ein Projekt für ein Valenzwörterbuch des Alt-, Mittel- und Frühneuhochdeutschen vorbereitet. Aus Perspektive der Arbeitsgruppe Digital Humanities ist die zentrale Herausforderung dabei die formale Modellierung der Wörterbuchartikel und der Valenzinformationen, um diese später für automatische Analysen bezüglich diachroner Valenzänderungen und verwandter Phänomenen zugänglich zu machen.<sup>9</sup> In einem weiteren Projekt aus dem Bereich der Dialektologie wird die Erstellung von Dialekt-Lexika des Bairischen aus *Social Media*-Daten untersucht. Dabei werden aus öffentlichen Facebook-Gruppen, in denen Bairisch geschrieben wird, automatisch Dialektwörter extrahiert. Diese Wörter werden dann über einen Crowdsourcing-Ansatz von den Mitgliedern der Gruppe ins (Hoch-)Deutsche rückübersetzt.<sup>10</sup>

Zwei weitere Projekte sollen die Darstellung der Aktivitäten der Regensburger Arbeitsgruppe in ihrer methodischen und fachlichen Breite abrunden: Die Benutzung von Eye-Tracking-Geräten zur Aufzeichnung von Blicken hat in der Regensburger Informationswissenschaft bereits eine mehr als 10-jährige Tradition. In Zusammenarbeit mit namhaften Museen gibt es Kooperationen auf dem Gebiet der Analyse von Blickbewegungen. Im Zentrum steht dabei die Frage, wie Personen Kunstwerke betrachten. Die hier erzielten Ergebnisse können einerseits dazu dienen, die Vorannahmen von Künstlern und Künstlerinnen zu prüfen, andererseits stellen die Erkenntnisse einen ersten Schritt auf dem Weg zur personalisierten Ausstellungsweise Führung durch Kunstaussstellungen dar. Auch auf dem

---

8 Döhling et al. 2017.

9 Burghardt & Reimann 2016.

10 Burghardt et al. 2016.

Gebiet der Musikpädagogik ist die Verfolgung von Blickbewegungen eine lohnenswerte Datenerhebungsmethode. Bei der Analyse des Blickverhaltens beim Blattspiel/-singen ist unter anderem eine starke Korrelation mit dem Expertisegrad der betreffenden Personen festzustellen. Diese Erkenntnisse können beispielsweise dazu verwendet werden, weniger erfahrene Personen gezielt in ihrem Blickverhalten zu schulen, wodurch diese schneller zu Expertinnen und Experten avancieren.

Ein detaillierter Überblick zu weiteren aktuellen Forschungsprojekten findet sich auf dem Forschungsblog „Digital Humanities Regensburg“ ([www.dhregensburg.de](http://www.dhregensburg.de)).

#### 4. Digital Humanities in der Lehre

Im Rahmen der Masterstudiengänge Medieninformatik, Informatikwissenschaft und Medienwissenschaft wird seit dem Wintersemester 2012/13 ein eigenes Studienmodul „Digital Humanities“ angeboten, das von den Studierenden aller drei Fächer sehr gut nachgefragt ist. Daneben gibt es zahlreiche Einzelveranstaltungen, die zumeist mit Kolleginnen und Kollegen aus benachbarten geisteswissenschaftlichen Fächern durchgeführt werden. Prägnante Beispiele sind die Zusammenarbeiten mit der allgemeinen Sprachwissenschaft im Bereich der Korpuslinguistik, mit der Germanistik auf dem Gebiet der Sentiment-Analyse und mit der Historischen Sprachwissenschaft in der digitalen Lexikographie.

Seit dem Wintersemester 2017/18 kann erstmals ein eigener Masterstudiengang „Digital Humanities“ in Regensburg studiert werden. Der Master wird im Rahmen des Innovationsprogramms *Digitaler Campus Bayern* von der Bayerischen Staatsregierung gefördert und ist Teil des Campus-Netzwerks *téchnē* (τέχνη), das zusammen mit den Universitäten FAU Erlangen-Nürnberg und der LMU München gebildet wird. Der Masterstudiengang an der Universität Regensburg erweitert somit die bislang vergleichsweise überschaubare Landschaft universitärer DH-Masterstudiengänge in Deutschland (einschlägige Masterstu-

diengänge sind bisher nur in Bamberg, Darmstadt, Köln, Mainz, Trier und Würzburg etabliert).

#### 4.1 Studiengangsziele

Eine Vielzahl erfolgreicher Absolventen in geistes- und kulturwissenschaftlichen Studiengängen sehen sich derzeit mit der Situation konfrontiert, dass nur wenige nach einem Bachelorabschluss in qualifizierte und gut bezahlte Beschäftigungen übernommen werden. Dies betrifft auch Lehramtsstudenten, von denen gegenwärtig – ungeachtet sehr guter Examensergebnisse – nur wenige ins Lehramt übernommen werden können. Der Masterstudiengang in Regensburg richtet sich deshalb insbesondere an diese Zielgruppen aus den geistes- und kulturwissenschaftlichen Fächern und verfolgt eine doppelte Stoßrichtung: Einerseits eröffnet er zusätzliche Tätigkeitsfelder (*qualifikations-erweiternde* Funktion). Absolventinnen und Absolventen werden in die Lage versetzt, Tätigkeiten mit IT-Bezug und fachlicher Nähe zu den Studienfächern zu übernehmen. Andererseits bietet der Studiengang eine ergänzende Zusatzqualifikation (*qualifikationsergänzende* Funktion) im Bereich der digitalen Medien und der Informationstechnologie. Ergänzend zu einer geistes- oder kulturwissenschaftlich geprägten Tätigkeit werden so digitale Arbeitstechniken für ein breites Spektrum an Zusatzanwendungen oder auch für die Kommunikation mit IT-Experten vermittelt. Im Fall von Lehramtsstudierenden kann diese Zusatzqualifikation beispielsweise die Ausgestaltung des Unterrichts mit digitalen Methoden für ein breites Spektrum an Schulfächern unterstützen.

#### 4.2 Qualifikationsziele

Ziel des Masterstudiengangs in Regensburg ist es, Studierende, die bereits einen geistes- oder kulturwissenschaftlichen Abschluss haben, methodisch weiter zu bilden. Dabei werden keinerlei IT-Kenntnisse vorausgesetzt. Zur methodischen Ausbildung gehört die Vermittlung anwendungsorientierter Kenntnisse in Programmierung, Datenerhebung und -auswertung sowie in Sprach- und Texttechnologie. Aufbauend auf diesen Kenntnissen werden Studierende dazu befähigt, selbstständig Projekte durchzuführen. Ein Schwerpunkt ist dabei die

Fähigkeit, Textkorpora (zum Beispiel digital vorliegende historische Quellen, Sprachkorpora oder Texte aus den sozialen Medien) mit geeigneten digitalen Werkzeugen zu analysieren und die Ergebnisse der Analysen in geeigneter Form zu interpretieren, aufzubereiten und diese zu kommunizieren.

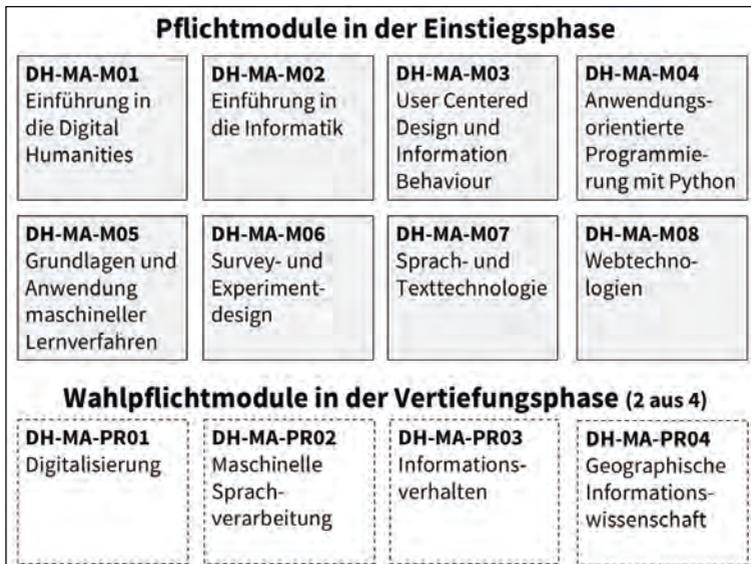


Abbildung 3: Überblick über die Module im Masterstudiengang Digital Humanities.

### 4.3 Studienmodule

Der idealtypische Studienverlauf des Masterstudiums Digital Humanities in Regensburg gliedert sich in ein Lehrjahr und ein Projektjahr (vgl. Abb. 3). Die ersten beiden Semester sehen eine intensive Einführungsphase in die erforderlichen IT- und Analyse-Themen vor (DH-MA-M01 bis DH-MA-M08), die der geistes- und kulturwissenschaftlich vorgebildeten Zielgruppe voraussetzungsfrei die theoretischen Inhalte und das methodische Rüstzeug für die Projektphase vermitteln. In dieser vertiefen die Studierenden dann ab dem dritten Semester ihre Kenntnisse durch Projektarbeiten. Durch die Schaf-

fung von Wahlpflichtmodulen (DH-MA-PR01 bis DH-MA-PR04) besteht die Möglichkeit zur Schwerpunktsetzung entsprechend der eigenen akademischen Vorbildung. Dabei ermöglicht das gut ausgebaute Netzwerk der Regensburger Medieninformatik und Informatikwissenschaft zu geistes- und kulturwissenschaftlichen Fakultäten, aber auch zu städtischen und Landesmuseen sowie zu weiteren Kultureinrichtungen, den Studierenden ihre Kenntnisse in verschiedenen Anwendungsfeldern praktisch zu erproben. Die so entstehenden Projektarbeiten führen wiederum zur Masterarbeit als wissenschaftlich fundierter Abschlussarbeit hin.

Tabelle 1 bietet einen Überblick über die Lerninhalte und -ziele der einzelnen Module im Masterstudiengang Digital Humanities.

<b>MODUL</b>	<b>LERNINHALTE/-ZIELE (AUSZUG)</b>
DH-MA-M01: Einführung in die Digital Humanities	Interdisziplinäre Reflexion über die Digitalisierung/digitale Gesellschaft verbunden mit fundierter Kenntnis der Genese der Digitalen Geistes- und Kulturwissenschaften.
DH-MA-M02: Einführung in die Informatik	Kenntnis grundlegender Konzepte der Informatik (z. B. Zahlensysteme, Rechnerarchitekturen, formale Sprachen).
DH-MA-M03: User Centered Design und Information Behaviour	Vermittlung von theoretischem und methodischem Wissen zur Erhebung von Informationsbedarfen und zur Umsetzung nutzerzentrierter Benutzerschnittstellen.
DH-MA-M04: Anwendungsorientierte Programmierung mit Python	Programmatische Lösung einfacher Problemstellungen in Python unter Verwendung einschlägiger Programmierbibliotheken (z. B. NLTK).
DH-MA-M05: Grundlagen und Anwendungen maschineller Lernverfahren	Selbstständige Datenauswertung mit Hilfe maschineller Lernverfahren (z. B. WEKA) und Interpretation der so erzielten Ergebnisse.
DH-MA-M06: Survey und Experimentdesign	Treffen begründeter Entscheidungen für die forschungsproblemspezifische Wahl von Methoden der Datenerhebung und Datenauswertung, verbunden mit der Vermittlung von Fähigkeiten zur praktischen Auswertung von Datensätzen mit Hilfe statistischer Software.

<b>MODUL</b>	<b>LERNINHALTE/-ZIELE (AUSZUG)</b>
DH-MA-M07: Sprach- und Texttechnologie	Kenntnis von Grundkonzepten (z. B. <i>Part-of-Speech Tagging</i> , <i>Named Entity Recognition</i> ) und Verfahren der Sprach- und Texttechnologie sowie Beurteilung derselben auf ihre jeweilige, fragestellungsbezogene Eignung.
DH-MA-M08: Webtechnologie	Kenntnisse zum grundsätzlichen Aufbau von Webanwendungen und Erstellung eigener, kontextbezogener Webanwendungen.
DH-MA-PR01: Digitalisierung	Mögliche Themenstellungen: Digitalisierung und Erschließung analoger Datenbestände; die Anwendung computergestützter Forschungsmethoden im geisteswissenschaftlichen Kontext (z. B. <i>Digital Art History</i> ); der nutzerzentrierte Entwurf von geisteswissenschaftlichen Arbeitswerkzeugen.
DH-MA-PR02: Maschinelle Sprachverarbeitung	Mögliche Themenstellungen: Fragestellungen der Stilanalyse (z. B. der automatischen Zuordnung von Texten zu Autoren); korpuslinguistische Problemstellungen (z. B. der Aufbau eines durchsuchbaren, digitalen Korpus einer bestimmten Sprache oder Sprachvarietät).
DH-MA-PR03: Informationsver- halten	Mögliche Themenstellungen: Quantitative Analyse großer Datensätze aus dem Umfeld der sozialen Medien (z. B. Twitter); qualitative Studien zum Finden von Information (z. B. an Informationskiosken in Museen).
DH-MA-PR04 Geographische Informationswis- senschaft	Mögliche Themenstellungen: Verarbeitung/Aufbereitung geographischer und linguistischer Daten, z. B. zur Unterstützung der räumlichen Orientierung und Navigation.

Tabelle 1: Kurzbeschreibung der Module aus dem Masterstudiengang Digital Humanities.

Das Regensburger Modell ist als nicht-konsekutiver Master angelegt, der sich an Absolventen mit grundständiger geistes- und kulturwissenschaftlicher Ausbildung wendet und diesen eine Entwicklung und Erweiterung ihrer Kompetenzen in Richtung digitaler und analytischer Verfahren anbietet. Es wendet sich damit bewusst nicht an substanziell informatisch und informationswissenschaftlich Vorgebildete. Modelle an anderen Standorten, mit anderer Ausrichtung und anderem Zielpublikum, haben dabei gleichermaßen ihre Berechtigung und stoßen ebenfalls auf reges Interesse. In einem emergenten und sich institutionell noch im Entstehen befindlichen Feld wie es die Digital

Humanities sind, wirft die Gestaltung der Lehre und Nachwuchsausbildung vielfältige Fragen auf. Unweigerlich sind in diesem Ökosystem verschiedene Modelle auf Bachelor- und Master-, Haupt- und Nebenfachebene entstanden und werden auch weiterhin entstehen und sich fortentwickeln.<sup>11</sup> Bei der weiteren Abstimmung eines DH-Curriculums muss somit die Heterogenität der gegenwärtigen DH-Landschaft (in Forschung, Lehre und konzeptionellem Selbstverständnis) mitbedacht werden und sowohl Zielgruppe als auch Qualifikationsziel des Studiums müssen klar umrissen und kommuniziert werden.

## **5. Ausblick: Institutionalisierung der Digital Humanities durch eine zentrale Infrastruktur?**

Wie eingangs gezeigt, besteht an der Universität Regensburg eine lange Tradition digital unterstützter Geistes- und Kulturwissenschaften, die auch aktuell in Forschung und Lehre ihre Fortsetzung findet. Mit der Arbeitsgruppe Digital Humanities am Institut für Information und Medien, Sprache und Kultur wurden in den letzten Jahren erste Versuche unternommen, die unterschiedlichen Forschungsaktivitäten an einer zentralen Stelle zu bündeln und damit interdisziplinäre Kooperationsformen zu befördern, was bislang überaus erfolgreich war. Darüber hinaus wurden Beratungs- und Lehrangebote für die angemessene Verwendung digitaler Medien und Methoden etabliert und gleichzeitig neue, digitale Tools und Ressourcen für die geistes- und kulturwissenschaftlichen Fächer entwickelt.

Langfristig spricht vieles dafür, diese und weitere wichtige Aufgaben von einzelnen Lehrstühlen zu entkoppeln und an eine zentrale Stelle auszulagern, die sich in die bestehende Infrastrukturlandschaft der Universität Regensburg – also etwa Rechenzentrum und Bibliothek – einfügt. Das Ziel ist dabei, ein eigenständiges Kompetenzzentrum für Digital Humanities nach dem Vorbild bestehender Zentren wie etwa in Trier, Köln oder Göttingen<sup>12</sup> zu etablieren. Ein solches Zentrum

---

11 Sahle 2013; DHd-AG „Referenzcurriculum Digital Humanities“ 2015.

12 Burghardt & Wolff 2015.

wäre dann zuständig für das Vorantreiben genuiner Forschung im Bereich Digital Humanities, für die systematische Etablierung computergestützter Methoden, für die Unterstützung bei der Drittmittelakquise und zuletzt auch bei der Bereitstellung einer technischen Infrastruktur und der Unterstützung beim Forschungsdatenmanagement und beim digitalen Publizieren. Zielpublikum eines solchen Kompetenzzentrums ist dabei das gesamte Spektrum geistes- und kulturwissenschaftlicher Fächer an der Universität Regensburg, denen es als Ansprech- und Kooperationspartner in sämtlichen Belangen digitaler Medien, Methoden und Werkzeuge zur Verfügung steht (vgl. Abb. 4).

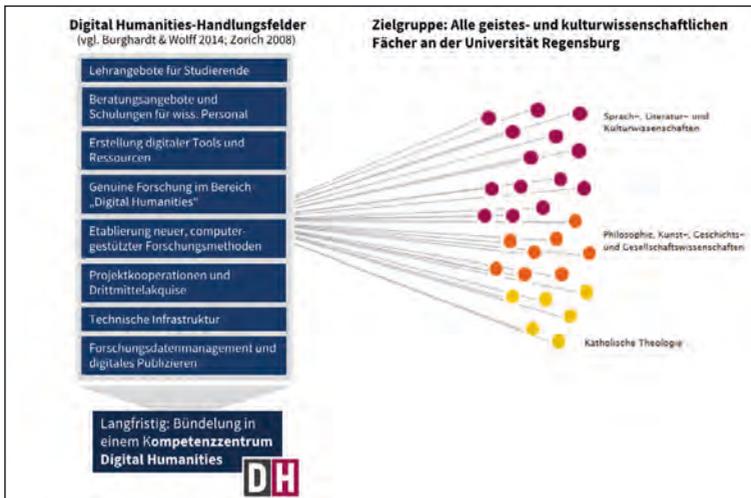


Abbildung 4: Digital Humanities-Handlungsfelder und potenzielle Zielgruppen eines Kompetenzzentrums Digital Humanities an der Universität Regensburg.

Die Finanzierung könnte – wie bei vielen anderen bestehenden Digital Humanities-Zentren – als sogenanntes *Profit Center* erfolgen. Derart organisierte Zentren finanzieren sich in erster Linie durch eingeworbene Drittmittel und werden parallel hinsichtlich grundlegender infrastruktureller Anforderungen durch die Trägerhochschule unterstützt. Weitere Möglichkeiten der Finanzierung ergeben sich etwa durch die

kostenpflichtige Beratung oder Schulung externer Universitäten, die über kein eigenes Kompetenzzentrum auf dem Gebiet der Digital Humanities verfügen sowie über die Lizenzierung von Softwarelösungen für Anwender außerhalb der Universität Regensburg.<sup>13</sup>

## 6. Literatur

- Burghardt, M. & Wolff, C.: Digital Humanities: Buzzword oder Strukturwandel in den Geisteswissenschaften? In: *Blick in die Wissenschaft*, 23(29) (2014), S. 39–47.
- Burghardt, M. & Wolff, C.: Zentren für Digital Humanities in Deutschland, in: *Information – Wissenschaft & Praxis*, 66(5–6) (2015), S. 313–326.
- Burghardt, M., Wolff, C. & Womser-Hacker, C.: Informationswissenschaft und Digital Humanities, in *Information – Wissenschaft & Praxis*, 66 (5–6) (2015), S. 287–294.
- Burghardt, M., Granvogel, D. & Wolff, C.: Creating a Lexicon of Bavarian Dialect by Means of Facebook Language Data and Crowdsourcing, in: *Proceedings of the 10th edition of the Language Resources and Evaluation Conference* (2016).
- Burghardt, M. & Reimann, S.: Möglichkeiten der elektronischen Aufbereitung und Nutzung eines historisch syntaktischen Verbwörterbuchs des Deutschen, in: Greule, A. & Korhonen, J. (Hrsg.): *Historisches syntaktisches Verbwörterbuch. Valenz- und konstruktionsgrammatische Beiträge*, Frankfurt am Main 2016, S. 301–322.
- Berry, D. M.: The Computational Turn: Thinking about the Digital Humanities, in: *Culture Machine*, 12 (2011), S. 1–22.
- Berry, D. M. (Hrsg.): *Understanding Digital Humanities*. Hampshire (UK) & New York (US) 2012.
- DHd-AG „Referenzcurriculum Digital Humanities“: Digital Humanities als Beruf. Fortschritte auf dem Weg zu einem Curriculum, in: *Akten der DHd-Arbeitsgruppe „Referenzcurriculum Digital Humanities“*, Graz, 24. bis 27. Februar 2015.

---

13 Zorich 2008.

- Döhling, L., Burghardt, M. & Wolff, C.: PaLaFra – Entwicklung einer Annotationsumgebung für ein diachrones Korpus spätlateinischer und altfranzösischer Texte, in: Book of Abstracts der DHd, Bern 2017.
- Gold, M. K. (Hrsg.): Debates in the Digital Humanities. University of Minnesota Press, 2012. Online verfügbar unter <http://dhdebates.gc.cuny.edu/>.
- Moretti, F.: Conjectures on world literature. *New Left Review* (2000), S. 54–68.
- Sahle, P.: DH studieren! Auf dem Weg zu einem Kern- und Referenzcurriculum der Digital Humanities, in: DARIAH-DE Working Papers 1 (2013). Online verfügbar unter <http://webdoc.sub.gwdg.de/pub/mon/dariah-de/dwp-2013-1.pdf>.
- Terras, M., Nyhan, J. & Vanhoutte, E. (Hrsg.): Defining Digital Humanities. Burlington 2013.
- Unsworth, J.: What is humanities computing (and what is not)? In: *Jahrbuch für Computerphilologie* 4 (2002), S. 71–84. Online verfügbar unter <http://computerphilologie.uni-muenchen.de/jg02/unsworth.html>.
- Wolff, C.: „embedded media computing“ – die Regensburger Ausrichtung der Medieninformatik, in: Workshop Medieninformatik in Forschung, Lehre und Praxis im Rahmen der GI-Fachtagung Mensch und Computer, Berlin 2009. Online verfügbar unter: <https://epub.uni-regensburg.de/14236/>.
- Zorich, D.: A Survey of Digital Humanities, in: Council on Library and Information Resources, Washington D.C. 2008.

# Squirrel

## Ein innovativer Editor zur Erstellung digitaler Editionen und sein Pilotprojekt<sup>1</sup>

Mark Hengerer, Isabella Hödl-Notter, Gerhard Schön

Die Erstellung von Transkriptionen und die Veröffentlichung von Editionen gehören zum Kernbereich der Geschichtswissenschaft. Beide Tätigkeiten sind oftmals schwierige und langwierige Prozesse und brauchen in der Lehre viel Zeit, individuelle Übung und professionelle Unterstützung.

Die kritische Edition handschriftlicher Quellen, einst von nicht wenigen als altmodisch und wenig kreativ belächelt, erfährt durch die neuen Möglichkeiten von Open-Access-Publikationen seit einigen Jahren eine neue, sich in zahlreichen digitalen Editionen äußernde Akzeptanz. Diese Entwicklung geht allerdings noch nicht überall mit einer Stärkung der grundwissenschaftlichen Ausbildung an den deutschsprachigen Universitäten einher; erst 2015 in einem Grundsatzpapier des Verbandes der Historiker und Historikerinnen Deutschlands wurde beklagt, dass entsprechende Ausbildungsangebote zum Umgang mit handschriftlichen Quellen rückläufig seien.<sup>2</sup>

- 
- 1 Für konstruktive Anregungen zu diesem Text und die großzügige Unterstützung des Projektes danken wir Dr. Christian Riepl (IT-Gruppe Geisteswissenschaften).
  - 2 „Grundlegende Kenntnisse und Fähigkeiten drohen deshalb nicht nur bei den Studierenden, sondern langfristig auch bei den Lehrenden in einem Maße abzunehmen, dass die kulturelle Überlieferung der Vergangenheit nicht mehr eigenständig erschlossen und beurteilt werden kann.“ Forum: E. Schlotheuber / F. Bösch: Quellenkritik im digitalen Zeitalter: Die Historischen Grundwissenschaften als zentrale Kompetenz der Geschichtswissenschaft und benachbarter Fächer, in: H-Soz-Kult, 16.11.2015, <http://www.hsozkult.de/debate/id/diskussionen-2866> (30. Oktober 2017).

Um die geschichtswissenschaftliche Kernkompetenz im Bereich Transkription und Edition zu stärken, startete im Jahr 2014 am Historischen Seminar der LMU München ein Editionsprojekt auch als informatisches, von der IT-Gruppe Geisteswissenschaften betreutes Pilotprojekt.<sup>3</sup> Der Editor, den wir mit Blick auf die Freuden und Leiden des wissenschaftlichen Edierens inzwischen *Squirrel* genannt haben, wurde im Wintersemester 2014/15 von Gerhard Schön (IT-Gruppe Geisteswissenschaften) in Zusammenarbeit mit den Studierenden des Editionsprojekts „Die Briefe Erzherzog Leopold Wilhelms an Kaiser Ferdinand III. aus dem Reichsarchiv Stockholm“ – Isabella Hödl-Notter, Franziska Honer, Julius Kiendl, Heike Nickel, Hildegard Renner, Alexandra Röckel, Christine Rogler und Pia Zachary unter Leitung von Mark Hengerer – entwickelt. Seit Frühjahr 2017 besteht zudem eine Kooperation mit dem Computerlinguisten Max Hadersbeck und seinem Team, Oksana Budurova, Azada Rustamova, Stefan Schwyter und Abuliezi Xieyidan. Ziel dieser vom Programm Lehre@LMU finanziell geförderten Kooperation ist die Verschränkung des Editors mit computerlinguistischen Analysetools (Unitex/GramLab). *Squirrel* trägt mithin dazu bei, zukunftssträchtige grundwissenschaftliche Forschungsstrukturen zu entwickeln und nachhaltig zu verankern.

Die IT-Gruppe Geisteswissenschaften wiederum bereitet auf einem fach- und fakultätsübergreifenden Konzept die Grundlagen für eine interdisziplinäre Zusammenarbeit von Geschichtswissenschaft, Informatik und Computerlinguistik, indem sie (a) dauerhaft die informationstechnische Infrastruktur (File-, Datenbank- und Webserver) zur Verfügung stellt, (b) Softwaremodule projektübergreifend konzipiert, entwickelt, wartet und pflegt sowie projektspezifisch adaptiert, (c) in Zusammenarbeit mit der Universitätsbibliothek und dem Leibniz-Rechenzentrum (LRZ) Einzelprojekte in ein umfassendes und nachhaltiges Forschungsdatenmanagement einbindet.

---

3 [http://www.fnz.geschichte.uni-muenchen.de/forschung/forsch\\_projekte/briefe-leopold/index.html](http://www.fnz.geschichte.uni-muenchen.de/forschung/forsch_projekte/briefe-leopold/index.html)

## 1. Squirrel. Der Editor

### 1.1 Technische Details

Die Zweispaltenansicht stellt dem zoomfähigen Scan der Briefseite ihre Transkription gegenüber, die als Freitext erfasst und nach einem vorgegebenen, aber adaptierbaren Schema annotiert werden kann (Abb. 1). Die Wörter aus dem Faksimile können per Textrahmen mit ihrem Transkript verknüpft werden. Zu jedem markierten Wort werden dann automatisch die Bilder von gleich und ähnlich transkribierten Schriftzügen als Hilfestellung eingeblendet (Abb. 2). Der Editor *Squirrel* stellt so beim Edieren visuelles Vergleichsmaterial für transkribierte Passagen zur Verfügung und ermöglicht ein permanentes Feedback während der Transkriptionsarbeit. Die Orientierung an Vergleichsmaterial kann zwar zeitweise auch Falschlesungen stabilisieren, doch hat die technisch stark unterstützte Arbeit mit visuellem Vergleichsmaterial solche Fehler bislang regelmäßig auch aufzudecken geholfen.

Alle diese Bildausschnitte sind auch über eine Suchfunktion erreichbar und jeweils mit einem Link zur Sofortkorrektur der Transkription versehen (Abb. 3).



Abbildung 1



Abbildung 2

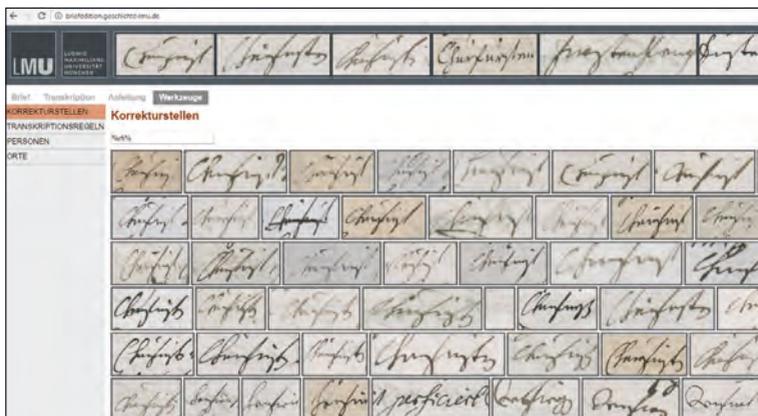


Abbildung 3

*Squirrel* basiert clientseitig vollständig auf HTML5 und erfordert auf Benutzerseite außer einem Browser und Internetverbindung keinerlei Zusatzsoftware oder Plugins. Serverseitig ist die Anwendung in PHP implementiert. Der Editionstext samt XML Annotationen wird auf dem Server vorgehalten (derzeit in geshredderter Form im relationalen Datenbanksystem MySQL) und kann über die Webseite als XML exportiert und durch XSLT in verschiedene Ausgabeformate transfor-

miert werden. Für die Vergleichsfunktionen hat sich der Einsatz von Trigrammanalysen bewährt.

Die Weiterentwicklung des technischen Bereichs von *Squirrel* verteilt sich seit dem Frühjahr 2017 auf drei Bereiche:

1. Zunächst implementieren wir die wichtigsten Teile des vom Deutschen Textarchiv (DTA) verwendeten Basisformats (DTABf), einer auch von der DFG und CLARIN-D empfohlenen Einschränkung der TEI Richtlinie P5 für historische Drucke. Hierzu werden bewährte Standards innerhalb der Geschichtswissenschaft (für die Epochen Mittelalter und Neuzeit) ermittelt und im Editor als Annotationsoptionen bereitgestellt. Die domänenspezifische Adaption aufwändiger Einzelprojekte lässt sich damit auf ein geringes Maß an Programmierarbeit reduzieren. Projekte innerhalb der klassischen Quellengruppen sollten mit dieser Vorarbeit ohne weitere oder wenigstens ohne aufwändige Adaption auskommen.

2. Zudem verschränken wir *Squirrel* mit dem UNITEX/GramLab (<http://unitexgramlab.org/>). Mit Elementen des GramLab wird eines der zentralen Probleme von Online-Publikationen gelöst, nämlich die automatisierte Bereitstellung eines Registers. Baut man (mit vertretbarem Arbeitsaufwand) ein Lexikon mit UNITEX/Gramlab auf, zeigt die Suche nach Grundformen (Lemmata) sämtliche Flektionsformen in der Trefferliste an. Zudem lässt sich eine Wortkonkordanz einfach erstellen.

3. Schließlich bereiten wir Schnittstellen zu wichtigen Publikationsformaten (*XHTML*, *LaTeX/PDF*) vor, welche hybride Veröffentlichungen erleichtern. Erforderlich ist die elektronische und typographische Adaptierung der durch unsere Auswahl aus dem DTA Basisformat sowie der in den verschiedenen Epochendisziplinen verwendeten weiteren Annotationstypen (Formatierungen) in für die Studierenden bereitzustellenden Annotationsoptionen (Formatvorlagen). Die besten Ergebnisse im Printbereich liefern derzeit Dateien in LaTeX/PDF, im Onlinebereich Dateien mit XHTML.

## 1.2 Workflow

Für die Arbeit mit *Squirrel* bietet sich somit der folgende Workflow an:

- Einrichtung eines Zugangs zum *Squirrel* Editor für Lehrende und Studierende, gegebenenfalls durch Anmeldung im Digital Humanities Virtual Laboratory (DHVLab, <https://dhvlab.gwi.uni-muenchen.de>).
- Anlegen eines Datensatzes pro Archivalieneinheit (Brief, Faszikel) und Eintragung der Metadaten zur Quelle.
- Hochladen der zu transkribierenden Digitalisate (zum Beispiel Scans der Handschriften im JPEG Format).
- Übungen im Transkribieren (optimalerweise Blended Learning), von *Squirrel* unterstützt durch Einblendung gleich und ähnlich transkribierter Buchstabenfolgen in der Oberleiste der Anwendung (Abb. 2). Es ist so der visuelle Abgleich möglich, welcher durch Nutzung der Funktion „Werkzeug“ intensiviert werden kann; das Programm erlaubt die Anzeige aller gleich oder ähnlich transkribierten Buchstabenketten (Abb. 3). Schreibvarianten werden dadurch gut sichtbar.
- *Optional* Fortentwicklung der Transkriptionen zur publizierbaren Edition, das heißt Fixierung der im Hinblick auf Rezipienten, Ressourcen und wissenschaftliche Adäquanz hin spezifizierten Grundsätze der Texterstellung; diese „Domänenspezifikation“ ist zwar nötig, wird aber durch das DTA Basisformat soweit möglich minimiert, um eine Anschlussnutzung zu erleichtern. Obschon Interoperabilität angestrebt wird, ist doch festzuhalten, dass absolut interoperable Datenbanken im Bereich Edition historischer Quellentexte bislang noch nicht vorliegen.
- Erstellung von Register und Konkordanz (Disambiguierung, Lexikon) sowie gegebenenfalls weiterer linguistischer Analysen mit UNITEX/GramLab.
- Publikation im Format XHTML und PDF sowie im Druck.<sup>4</sup>

---

<sup>4</sup> Warum Druckfassungen nach wie vor wichtig sind, macht die 2017 auf politischen Druck Chinas hin erfolgte Löschung von Artikeln aus der Online-Zeitschrift *China Quarterly* deutlich: siehe *New York Times*, 18. August 2017 „Cambridge University Press Removes Academic Articles on Chinese Site“. Die Kehrtwende des Verlages

## 2. Lehre. Das Pilotprojekt

### 2.1 Entwicklungsgeschichte

Die Entwicklung des Online-Editors *Squirrel* baut auf mehrjährigen Vorarbeiten auf. Am Anfang stand auf eine von Mark Hengerer am Historischen Seminar der LMU München durchgeführte Lehrveranstaltung zu frühneuzeitlichen Quellen im Wintersemester 2014/15. Als sich abzeichnete, dass die Motivation für die Bearbeitung eines längerfristigen Pilot- und Publikationsprojekts gegeben war, begann die Entwicklung des Online-Editors durch Gerhard Schön unter laufender Berücksichtigung der sich aus der Praxis ergebenden Wünsche und Anregungen der Teilnehmer.

Inzwischen wurde der Editor in der Lehre des Historischen Seminars an der LMU mehrmals eingesetzt, auch in einem mediävistischen Pilotprojekt zur Urkundenedition (Julian Schulz, Magdalena Weiler). Im Wintersemester 2017/18 bieten Max Hadersbeck, Mark Hengerer und Gerhard Schön gemeinsam eine Lehrveranstaltung an, in der die computerlinguistische Erweiterung von *Squirrel* an einem weiteren Editionsfallbeispiel erprobt wird („Textedition und Textanalyse digital – eine Einführung“).

Der Editor wurde inzwischen auch mehrfach öffentlich präsentiert und diskutiert, unter anderem beim Historikertag Hamburg 2016 (Postersession).

Zuletzt wurde der Editor mitsamt dem Lehrprojekt ausführlich bei einem Workshop des dh muc (<https://dhmuc.hypotheses.org/>) im November 2016 präsentiert.<sup>5</sup>

---

macht nur noch sichtbar, wie manipulierbar der Zugriff auf wissenschaftliche Publikationen ist; siehe The Guardian, 22. August 2017 „Cambridge University censorship U-turn is censored by China“.

5 <https://dhmuc.hypotheses.org/workshop-digitale-editionen-und-auszeichnungssprachen>

## 2.2 Lehre im Wandel

Neben den technischen Neuerungen bringt der Einsatz digitaler Methoden tiefgreifende Änderungen für die Lehre mit sich. Das notwendigerweise interdisziplinär strukturierte digitale Arbeitsumfeld erfordert nicht nur Weiterbildung, sondern verlangt auch ‚neue‘ soziale Fähigkeiten. Diese Entwicklung ist insofern neu, da die Geschichtswissenschaften trotz verstärkter Projektforschung lange Zeit vorrangig auf Einzelarbeit ausgerichtet waren.

Durch den Einzug digitaler Methoden in die Lehre erweist sich das Erlernen dieser sozialen Fähigkeiten als Herausforderung, denn Team- und echte Projektarbeit gewinnen für den Erfolg von Forschung und Lehre an fundamentaler Bedeutung. Die für Arbeit in Teams und Projekten nötigen Kompetenzen aber sind bislang nicht Teil des Kanons universitärer Lehre.<sup>6</sup> Anhand der verschiedenen oben genannten Pilotprojekte wurden die Herausforderungen und Potentiale sichtbar, die in den Bereichen Teamarbeit und Projektmanagement auf die Lehre zukommen.

Unsere Erfahrungen lassen sich ungefähr wie folgt fassen: Bedingung erfolgreicher Arbeit in Teams ist bewusst strukturierte Kommunikation. Diese muss den Spagat zwischen den Erfordernissen wissenschaftlicher Präzision, individueller Motivation und Effekten von Gruppenbildung bewältigen. Bewährt haben sich in dieser Situation

- die explizite Reflexion der verschiedenen Rollen und Motivationslagen der Teammitglieder – Zielkonflikte ließen sich so auf akzeptierte Ausgangsbedingungen zurückführen und konstruktiv lösen;
- strukturierte Interaktion, welche wegen der zeitintensiven dislozierten Arbeit im Browser nicht mehr selbstverständlich gegeben ist

---

<sup>6</sup> Entsprechend attraktiv sind externe Kursangebote, hier am Beispiel der LMU: Graduate Center (ab Promotion/Promotionsplanung), LMU Center for Leadership and People Management (ab Promotion), LMU Extra (unbeschränkt), LMU Plus (unbeschränkt), LMU Profil (nur Lehrende), Student und Arbeitsmarkt (unbeschränkt).

- das persönliche Kennenlernen der Teammitglieder, in unserem Falle bei einer gemeinsamen mehrtägigen Exkursion;
- regelmäßige persönliche Treffen im Wochentakt; Arbeitsfortschritte, Ziele, sachliche und soziale Probleme und Lösungsansätze ließen sich so am besten kommunizieren.

Nicht bewährt hat sich, vor allem bei der Kommunikation von Aufgaben, die ‚gute alte‘ E-Mail; in einem Parallelprojekt hat sich auch ein Weblog nicht bewährt.

Neben der ‚Arbeit am Team‘ nimmt Projektmanagement eine herausgehobene Position ein. Für die Motivation der Studierenden und einen erfolgreichen Projektabschluss ist es wichtig, Projekte den universitären Rahmenbedingungen entsprechend zu dimensionieren. So sollten Projekte den Zeitraum eines Semesters nicht überschreiten. Allerdings hat sich in anderen geschichtswissenschaftlichen Lehrveranstaltungen, in denen besondere IT-Kenntnisse wie PHP/MySQL am konkreten Forschungsgegenstand vermittelt wurden, gezeigt, dass ein Semester kaum ausreicht. Im Zweifel sind die Projekte daher lieber „klein“ zu dimensionieren.

Weiter sind Lösungen für das Problem der Ausweisung individueller Prüfungsleistungen innerhalb von Gemeinschaftsprojekten zu finden. Aus studentischer Perspektive sind prüfungsadäquate Leistungen oft schon durch die notwendigerweise besonders intensive Mitarbeit gegeben, etwa durch die Eingabe von Daten. Insbesondere dann, wenn als Projektziel die Option Publikation gewählt wird, ist damit umzugehen, dass Datenbereinigung (insbesondere die Kontrolle der Transkription) mindestens ebenso viel Zeit verbraucht wie die sorgfältige Datenaufnahme (Transkription). Der Faktor 2 scheint hier realistisch; rechnet man die ‚Publikationszielgerade‘ hinzu, sollte man mit dem Faktor 3 rechnen. Wird dies am Anfang der Projektarbeit klar kommuniziert, bleibt die Motivation bei den Studierenden erhalten – es bewährt sich dann die „Impftheorie“ (Dieter Frey).

### 3. Fazit und Ausblick

Der Online-Editor *Squirrel* ist ein Editionstool für Lehre und Forschung. Es basiert auf der interdisziplinären Zusammenarbeit von Geschichtswissenschaft, Informatik und Computerlinguistik. Er ermöglicht den Studierenden individuelles online-basiertes Lernen und verbindet dieses optimal mit Publikationsprojekten. Das Tool wurde speziell für den Einsatz für Editionsprojekte in der Lehre entwickelt. Es erfasst alle Projektphasen vom Basisniveau (als wissenschaftliche Übung) bis hin zur Publikation mit Druck- und Onlineschnittstelle (LaTeX/PDF, XHTML).

Durch die Implementation von Funktionen aus UNITEX/GramLab werden eine neuartige Registererstellung und bessere Durchsuchbarkeit der Texte ermöglicht (Disambiguierung, Lemmatisierung, Konkordanz) sowie Optionen für eine semantische Suche angeboten. Da es möglich ist, für neue Projekte auch das visuelle Vergleichsmaterial früherer Projekte verfügbar zu machen, vergrößert sich mit jedem neuen Projekt die Menge des Vergleichs- und Übungsmaterials.

<http://www.squirrel.uni-muenchen.de/>

## #DigiCampus

Die Tagung #DigiCampus hatte zum Ziel, die im Rahmen des Verbundes der LMU München mit der FAU Erlangen-Nürnberg und der Universität Regensburg entwickelte digitale Lehr- und Lerninfrastruktur für die Digital Humanities sowie die verschiedenen Komponenten des DHV Labs vorzustellen. Dabei wurde deutlich, wie eine IT-Ausbildung für die Geisteswissenschaften aussehen kann, wie sich weitere Institutionen an diesem Projekt beteiligen und wie diese das System in ihren Lehrbetrieb einbinden können.

## IT for All und DHV Lab

Das Projekt *IT for All* wird im Rahmen der Ausschreibung *Digitaler Campus Bayern* gefördert und hat eine Ausbildung zur digitalen Datenanalyse in den Geschichts- und Kunstwissenschaften zum Ziel. In diesem Zusammenhang ist das *Digital Humanities Virtual Laboratory* (DHV Lab) entstanden, das erstmalig in den Geisteswissenschaften eine serverbasierte Lehr- und Forschungsumgebung anbietet, die nun die Testphase hinter sich gelassen hat und in den Regelbetrieb sowie in die bayernweite Expansion gehen kann.

18,00 €

ISBN: 978-3-95925-055-9

ISBN (eBook): 978-3-95925-056-6

