

**Impulsvortrag: Einführung zum Workshop „Repräsentativität in digitalen Archiven“** (08.03.2022, virtueller Workshop auf der DHd 2022 Potsdam „Kulturen des digitalen Gedächtnisses“ vom 07.-11.03.2022)

*Dr. Corinna Dziudzia & Dr. Mark Hall*

Wir freuen uns, Sie alle heute virtuell begrüßen zu können und sind begeistert über das breite Interesse von Wissenschaftler\_innen, Archivar\_innen, Bibliotheksmitarbeiter\_innen, Kolleg\_innen aus dem Wissenschaftsmanagement, aus Deutschland und benachbarten Ländern, wie den Niederlanden, Luxemburg, Österreich und der Schweiz. Die Menge der Anmeldungen hat uns unmittelbar eine unserer Thesen bestätigt: das Thema „Repräsentativität in digitalen Archiven“ scheint ein übergreifendes, eines, das aus vielen verschiedenen Perspektiven relevant und für viele wichtig ist, das aber bislang, soweit wir sehen können, nur in Ansätzen problematisiert wird, es wenig Bewusstsein dafür gibt, dass Auswahlprozesse im Hinblick auf Digitalisierung diskutiert, reflektiert und transparent gemacht werden müssen, wobei doch gemeinsame Anstrengungen und der Austausch von Best-Practice-Beispielen hilfreich sein könnten und unseres Erachtens notwendig wären.

Bereits im März 2021 haben Mark und ich einen internationalen Workshop durchgeführt.<sup>1</sup> Es ist uns im letzten Jahr und nun erneut aufgefallen, wie breit dieses Feld ist und wer darauf aus den verschiedensten Perspektiven blickt. So sehr wir heute nicht zuletzt diese Breite in der Breite mit Ihnen allen adressieren wollen und deswegen auch einen sehr großen Archiv- und Repräsentativitätsbegriff zugrunde legen, in dem Sie sich idealerweise alle wiederfinden, haben wir doch eine entscheidende Einschränkung vorgenommen, die wir offen legen wollen: Uns geht es primär um historische Artefakte, wenn wir von Digitalisaten sprechen, nicht um ‚digital born artifacts‘. Vieles mag analog gelten und ist vermutlich übertragbar, es bildet nur nicht primär den Fokus unseres eigenen Nachdenkens, uns geht es

---

<sup>1</sup> Videos und Materialien sind auf der Homepage zum Projekt „Digitales Archiv und Kanon“. Dort gibt es auch eine wachsende Literaturliste zu dem Komplex, URL: <https://www.digitales-archiv-und-kanon.de/rda2022/bibliography>.

um die Digitalisierung der Artefakte, die bereits physisch in Archiven, Sammlungen, Bibliotheken und Museen liegen.<sup>2</sup>

Wir alle schauen aus unterschiedlichen Disziplinen kommend auf die Digital Humanities, sehen nicht selten vor allem das eigene Projekt und den eigenen kleinen Kosmos. In dieser kurzen Einführung wollen Mark und ich unsere Sichtweise thematisieren, nicht, weil wir die dominierenden Stimmen dieses Austauschs sein wollen, sondern um erste Fragen, Impulse und Ideen zu liefern. In den anschließenden Arbeitssessions gibt es dann die Gelegenheit, sich untereinander auszutauschen.

Mark als Informatiker und mich als Germanistin verbindet ein gemeinsames ‚Störgefühl‘: Auf der DHd-Konferenz 2019 in Frankfurt hat Mark quantitativ empirisch analysiert und herausgearbeitet, dass eine merkliche Anzahl von Abstracts der Teilnehmer\_innen Methoden und Tools der Digital Humanities an genau zwei Werken der deutschen Literatur erproben wollten, Lessings *Emilia Galotti* und Goethes *Faust*.<sup>3</sup> Das ist aus verschiedenen Gründen erstaunlich und ein Problem: Man kann zu den ohnehin schon regalmeterweise existierenden Arbeiten just zu diesen beiden Werken mit neuen, sich noch in der Erprobung befindenden Methoden nur schwer etwas zum Forschungsstand beitragen, im Gegenteil, die Versuche fielen nicht selten dahinter zurück, was nicht hilfreich ist, um die Digital Humanities in den ‚klassischen‘ Geisteswissenschaften als wissenschaftlich solide zu verankern. Zudem besteht die deutsche Literatur aus mehr als dem immergleichen, sehr kleinen Kanon zu dem auch Goethe und Lessing gehören. Aber warum wählten überhaupt so auffallend viele Beiträger\_innen ausgerechnet diese zwei Texte?

Vermutlich – neben einer Reihe anderer Gründe – weil just von diesen beiden Werken unproblematisch digitale Textfassungen verfügbar sind. *Emilia Galotti* und *Faust* sind in jedem Korpus, in jeder Sammlung, in jedem

---

<sup>2</sup> Diese Unterscheidung ist insofern wichtig, als dem bereits ein Selektionsprozess vorgeschaltet ist, der bei den ‚digital born artifacts‘ noch aussteht: Diesbezüglich muss noch entschieden werden, was in Archive, Bibliotheken und Sammlungen Aufnahme findet.

<sup>3</sup> Hall, Mark: “DH is the Study of dead Dudes.” In: DHd 2019 Digital Humanities multimedial und multimodal. 6. Tagung des Verbands „Digital Humanities im deutschsprachigen Raum“ (DHd 2019), Frankfurt am Main und Mainz, <https://doi.org/10.5281/zenodo.4622026>.

digitalen Archiv zu deutscher Literatur, manchmal sogar mehrfach, zu finden, nicht zuletzt, weil es sehr viele klassische Editionen von diesen Texten gibt, die Ausgangsbasis also schon eine sehr breite und umfangreiche ist.<sup>4</sup> *Faust* hat zudem eine eigene digitale Edition,<sup>5</sup> bei Lessing ist das geplant.<sup>6</sup> Es ist praktisch, diese Texte zu nehmen, und es ist einfach, sie liegen maschinenlesbar vor, sie sind qualitativ hochwertig – in mehr als einer Hinsicht.

Das Beispiel illustriert allerdings, wie sich eine Textauswahl in völlig neuen Kontexten und unter völlig anderen Rahmenbedingungen manifestieren kann, über die doch eigentlich seit mindestens den 1970er Jahren geklagt wird: Sie ist zu eng, zu homogen, zu männlich, zu protestantisch, zu weiß.<sup>7</sup> In den historischen Fachwissenschaften wird, mittlerweile seit Jahrzehnten, der Zusammenhang von Macht und Wissen thematisiert, über den großen Überlieferungsverlust und über Leerstellen im kulturellen Gedächtnis gesprochen, die ohnehin schon da sind, wird gefragt, wer entscheidet, was im Kanon und was im Archiv ist, was überliefert wird.<sup>8</sup> Daran muss sich nun notwendigerweise die Frage anschließen, wer entscheidet, was digitalisiert wird, was im digitalen Archiv ist und wie das bereit gestellt wird.<sup>9</sup> Der Fokus auf den Auswahlprozess, die Selektionen, konfrontiert uns mit einer Reihe von grundlegenden weiteren Fragen, die wir unter anderem in

---

<sup>4</sup> Vgl. dazu etwa das ZVDD, URL: <https://www.zvdd.de/startseite/>.

<sup>5</sup> URL: <http://faustedition.net/>.

<sup>6</sup> Vgl. URL: [https://www.ikfn.uni-osnabrueck.de/institut/mitarbeiterinnen\\_und\\_mitarbeiter/prof\\_dr\\_kai\\_bremer/aktuelle\\_forschungen/lessing\\_digital.html](https://www.ikfn.uni-osnabrueck.de/institut/mitarbeiterinnen_und_mitarbeiter/prof_dr_kai_bremer/aktuelle_forschungen/lessing_digital.html).

<sup>7</sup> Vgl. dazu etwa Ziolkowski, Theodore: „Zur Politik der Kanonbildung.“ In: Charlier, Robert, und Günther Lottes, (Hrsg.): Kanonbildung: Protagonisten und Prozesse der Herstellung kultureller Identität. 1. Aufl. Aufklärung und Moderne, Bd. 20. Hannover: Wehrhahn, 2009. S. 33-50.

<sup>8</sup> Vgl. hierzu u.a. Assmann, Aleida, und Jan Assmann (Hrsg.): Kanon und Zensur. Beiträge zur Archäologie der literarischen Kommunikation 2. München: W. Fink, 1987; Heydebrand, Renate v. (Hrsg.): Kanon, Macht, Kultur: theoretische, historische und soziale Aspekte ästhetischer Kanonbildung. Germanistische Symposien, Berichtsbände 19. Stuttgart: J.B. Metzler, 1998; Assmann, Aleida: Der lange Schatten der Vergangenheit: Erinnerungskultur und Geschichtspolitik. München: Beck, 2006; Haye, Thomas: Verlorenes Mittelalter: Ursachen und Muster der Nichtüberlieferung mittellateinischer Literatur. Mittellateinische Studien und Texte, volume 49. Leiden; Boston: Brill, 2016.

<sup>9</sup> Dziudzia, Corinna, und Hall, Mark: „Die Kanonfrage 2.0.“ In: DHd 2020 Spielräume: Digital Humanities zwischen Modellierung und Interpretation. 7. Tagung des Verbands Digital Humanities im deutschsprachigen Raum (DHd 2020), Paderborn, <https://doi.org/10.5281/zenodo.4621782>.

der heute am Ende des Workshoptages zu gründenden DHd-Arbeitsgruppe zu adressieren hoffen: Wie läuft dieser Prozess der Digitalisierung von historischen Artefakten ab? Wer entscheidet das, nach welchen Kriterien? Welche Kriterien gibt es überhaupt? Wie wird die Priorisierung festgelegt? Wird der aktuelle Stand der Forschung in diesen Prozess eingespeist? Nach welchen Kriterien erfolgt die Erschließung, bzw. werden Metadaten für das Digitalisierte vergeben? Wie steht es um die Findbarkeit, wie kann sie dauerhaft gewährleistet werden? Erfolgt der Digitalisierungsprozess nachhaltig?

Im Augenblick erscheint der Prozess intransparent, als Black Box, und ist in vielen Aspekten nicht zuletzt deswegen potentiell problematisch, das geht weit über die Frage nach *Emilia Galotti* und *Faust* hinaus. Es geht auch weit über die Anekdote hinaus, dass zuerst in der *Europeana*-Datenbank die *Mona Lisa*, wohl eines der bekanntesten Gemälde überhaupt, von den Nutzer\_innen nicht gefunden wurde: Das Gemälde von Leonardo da Vinci war unter seinem eigentlichen Titel abgelegt, *La Gioconda*. Obgleich schwer zu überprüfen, ob sich das wirklich so zugetragen hat, vermittelt die Anekdote ein unmittelbar intuitives Verständnis dafür, wie wichtig Metadaten sind – und wo wissenschaftliche Präzision mit einer breiteren Öffentlichkeit und heterogenen Nutzergruppen kollidieren.

Das sind letztlich aber vor allem Symptome, die man auf der Einzelfallebene leicht beheben kann und für die jeweils Lösungen existieren. Das Sammeln dieser Symptome und der Austausch über diese ‚anekdotische Evidenz‘ ist aber notwendig, um die Reflektion über das Große und Ganze zu motivieren und zu ermöglichen. Dem widmet sich der Vormittag des heutigen Workshops, dem Sammeln, dem Austausch als Grundlage eine gemeinsame Basis zu finden. Zudem ist das Diktum, unter dem wir versuchen zu operieren, dass es über diese Probleme der ‚anekdotischen Evidenz‘ hinaus gehen muss, es reicht nicht, Symptome zu beschreiben, zu kritisieren und auf der Einzelfallebene zu beheben. Zugleich brauchen wird diese Beispiele, denn die überblicken wir und verstehen sie, sie bilden die Puzzle-Teile für das größere Bild. Dieses wollen wir am heutigen Nachmittag beginnen gemeinsam zusammenzusetzen – zumindest erste Eckpunkte dafür zu sammeln.

Digitalisierung wird mittlerweile seit den 1980er Jahren vorangetrieben: exemplarisch etwa sei der Aufsatz von Lynch und Brownrigg von 1986 genannt, die als wesentliche Aufgaben der Digitalisierung „preservation and accessibility“ definierten, Bewahren und Zugänglichmachen.<sup>10</sup> Es begann zunächst damit, die bei den Nutzer\_innen wenig beliebten Microfiche zu ersetzen und war zentrale Aufgabe von Bibliotheken und Archiven. Best-Practice-Modelle hinsichtlich der technischen Umsetzung wurden damit begonnen zu etablieren und Standards definiert.

Aus dem Jahr 2009 findet sich eine aufschlussreiche Metastudie zur Frage der Selektion im Hinblick auf Digitalisierung. Finanziert wurde sie durch die niederländische Regierung: Bart Ooghe und Dries Moreels analysieren darin hunderte von in Europa, Nordamerika und Australien verfügbaren Berichten, Handreichungen, Umfragen usw. zur Frage von Selektionskriterien digitalisierender Institutionen.<sup>11</sup> Sie identifizieren als Ergebnis ihrer Untersuchung 25 Kriterien, sortiert in 6 Kategorien. In der Regel haben sie diese aus den jeweils sichtbar werdenden Praktiken abgeleitet. Als ein Aspekt findet sich darunter Repräsentativität vermerkt.<sup>12</sup> Insbesondere konstatieren die beiden Autoren als Ergebnis, dass Zentraleuropa in der Frage nach Selektionskriterien zurück liegt, die Frage oftmals gar nicht adressiert wird und für nicht so wichtig erachtet wird, die Digitalisierungspraxis vielmehr auf Ad-Hoc-Entscheidungen und den jeweils verfügbaren Finanzierungen, bzw. „on demand“ zu basieren scheint.<sup>13</sup> So zahlreich die zu beobachtenden Digitalisierungsvorhaben sind, so zahlreich scheinen gleichermaßen die ihnen jeweils unterliegenden Kriterien der Auswahl. Oftmals wird dabei nicht nach Kriterien der Auswahl und Gründen für die Digitalisierung unterschieden. Kommuniziert wird

---

<sup>10</sup> Lynch, Clifford A., und Edwin B. Brownrigg: „Conservation, Preservation, and Digitization“. *College & Research Libraries* 47, Nr. 4 (1. Juli 1986): 379–82, <https://doi.org/10.5860/crl.47.04.379>.

<sup>11</sup> Ooghe, Bart, und Dries Moreels: „Analysing Selection for Digitisation: Current Practices and Common Incentives“. *D-Lib Magazine* 15, Nr. 9/10 (September 2009), <https://doi.org/10.1045/september2009-ooghe>.

<sup>12</sup> Ooghe und Moreels beziehen den Begriff dabei augenscheinlich vor allem auf die einzelne Sammlung und ob diese repräsentativ im Digitalen dargestellt wird, also wenn eine Sammlung bspw. 80% Prosatexte umfasst und 20% Lyrik, ob dieses Verhältnis auch im Digitalen gewahrt bleibt.

<sup>13</sup> Im Verlauf des Workshop-Tages wurde das von einem Teilnehmer treffend als im Augenblick dominierendes „frühkapitalistisches“ Prinzip bezeichnet.

darüber in der Regel noch weniger,<sup>14</sup> obwohl es sich doch um einen hochkomplexen Prozess handelt, der bestimmt wird durch die einzelnen Rahmenbedingungen der Institution, dem jeweiligen Material, dessen physischer Beschaffenheit, den vorhandenen Metadaten, den Finanzen, dem Rechtsrahmen usw.

Das hat Folgen weit jenseits des Umstandes, dass sich kein gemeinsames Vokabular etablieren kann, wenn man nicht darüber kommuniziert, wie die Autoren schreiben: „This makes it difficult for end-users to interpret the relationships between analogue and digital collections, a matter of increasing importance as many users come to know collections only through their digital counterparts.“<sup>15</sup> Ooghe und Moreels konstatieren damit entsprechend ein Paradox: Obwohl sich doch gerade im digitalen Raum die Rahmenbedingungen massiv verändern und damit theoretisch eine größere Transparenz einhergeht, resultiert daraus nicht notwendig mehr Kenntnis, weder über die jeweiligen physischen Sammlungen noch über die digitalen. Den einzelnen Nutzer lässt es in der Regel unwissend der eigentlichen Komplexität gegenüber, was die jeweilige digitale Zusammenstellung umfasst und aus welchen Gründen sie so erscheint, wie sie sich jeweils präsentiert.<sup>16</sup>

Die Studie von Ooghe und Moreels zielte nicht zuletzt darauf, die impliziten Kriterien sichtbar zu machen, die Praktiken. Sie kamen zu dem Ergebnis, dass es im Grunde immer eine Vielzahl von miteinander interagierenden Kriterien sind, etwa inhaltlich getriebene Auswahlkriterien, oder Handlungsdruck aufgrund fragiler Materialität. Anderes ist bestimmt durch

---

<sup>14</sup> Laut des Enumerate Survey von 2017 haben 82% der befragten Institutionen eine digitale Sammlung, aber nur 42% eine verschriftlichte Strategie dazu. Am häufigsten wird darin die Frage der Langzeitarchivierung/Bereitstellung thematisiert, für welche rund die Hälfte der befragten Institutionen keine Lösung hat. URL: [https://pro.europeana.eu/files/Europeana\\_Professional/Projects/Project\\_list/ENUMERATE/deliverables/DSI-2\\_Deliverable%20D4.4\\_Europeana\\_Report%20on%20ENUMERATE%20Core%20Survey%204.pdf](https://pro.europeana.eu/files/Europeana_Professional/Projects/Project_list/ENUMERATE/deliverables/DSI-2_Deliverable%20D4.4_Europeana_Report%20on%20ENUMERATE%20Core%20Survey%204.pdf).

<sup>15</sup> Ooghe & Moreels 2009.

<sup>16</sup> Im Verlauf der DHd-Konferenz 2022 wurde von den Cultural-Heritage-Institutionen dabei immer wieder betont, dass der einzelne Nutzer am besten die Bestände vor Ort nutzt, nur ein Bruchteil online zu finden ist usw. Dies aber wird der Komplexität von Forschungsfragestellungen, die wohl nicht selten verstreute Bestände berücksichtigen möchten, genauso wenig gerecht wie die Frage nach den Ressourcen, die das erfordern würde, die ebenfalls in Überlegungen der Nachhaltigkeit mit einzuberechnen sind.

rechtliche Bedingungen. In der Regel implizit bleiben die zugrundeliegenden Priorisierungen, die Digitalisierung wahrscheinlicher machen, beispielsweise eine bereits erfolgte Aufnahme eines bestimmten Bestands in die Online-Kataloge; eine einfache Zugänglichkeit zum jeweiligen Artefakt; dessen Größe usw. Ihre zusammengetragenen 25 Kriterien sind entsprechend nicht handlungsleitend zu verstehen – wie man es machen sollte und was es zu berücksichtigen gilt –, sondern sind Zusammenschau dessen, was im Hinblick auf Selektionen bislang praktiziert wurde und vermutlich seit dem Fortsetzung fand.

Der Prozess der Selektion, was jeweils digitalisiert wird und weshalb, ist entsprechend komplex, wie auch Alexandra Mills schreibt, nicht zuletzt, weil die Bestände jeweils bereits zuvor Auswahlprozesse durchlaufen haben:

Determining which items from a carefully vetted and built analogue collection merit digitization is a complex process. As many institutions predominantly acquire materials to meet user needs, any or all of these collection items could presumably be digitized. However, there are other factors to consider. Among them, items may be selected because online availability would increase access and satiate user demand; digital copies would reduce the handling of original documents; the intellectual and evidential value of the original merits a digital access point; or physical characteristics that limit the use of originals would be circumvented by the presence of a digital copy.<sup>17</sup>

Welche Faktoren dabei in die jeweiligen Entscheidungen eingeflossen sind, erfährt der Nutzer in der Regel nicht. Nur manchmal stehen auf Homepages digitalisierender Institutionen allgemeine Informationen, dass bestimmte Bestände, aus bestimmten Zeiträumen und bestimmten Sammlungen digital verfügbar sind.<sup>18</sup> Nicht selten scheint der weit überwiegende Teil eines physischen Archivs oder Magazins noch nicht digitalisiert, bisweilen ist das

---

<sup>17</sup> Mills, Alexandra: „User Impact on Selection, Digitization, and the Development of Digital Special Collections“. *New Review of Academic Librarianship* 21, Nr. 2 (4. Mai 2015): 160–69, S. 162, <https://doi.org/10.1080/13614533.2015.1042117>.

<sup>18</sup> Es gibt dabei Projekte wie VD16-18, die das Schrifttum ganzer Jahrhunderte digitalisieren wollen, was sich entsprechend natürlich auf die Bestände der beteiligten Institutionen bezieht, der Prozess ist noch nicht abgeschlossen. Für VD17 findet sich online ein transparenter Masterplan hinsichtlich der vorgenommenen Priorisierung, nach Unika, regionalen Kriterien, ‚Mainstream-Literatur‘ und Nummernliste (Konvolute mit mind. 1000 Titeln), URL: <http://www.vd17.de/files/vd17-masterplan-1.pdf>. Über Google Books bereits zur Verfügung stehende Bücher werden einbezogen. Allein daran wird ersichtlich, welche wichtige Rolle Katalogisierung und Erschließung dabei zukommt.

auch gar nicht geplant. Anderes ist digitalisiert, aber nicht online zugänglich. Aufschlussreich erweist sich für diese Fragen der Bericht des *Enumerate Survey*, eine Befragung von europäischen Kulturinstitutionen aus 28 Ländern: Demnach haben 82% der Befragten eine digitale Sammlung; im Schnitt sind 48% der Sammlungen katalogisiert.<sup>19</sup> Im Schnitt sind 22% der Sammlungen digital reproduziert, für weitere 54% des Bestandes ist eine Digitalisierung vorgesehen, bei 24% wird das als unnötig angesehen. Von den 22% des digitalisierten Bestandes sind 42% nicht online verfügbar. Weit überwiegend werden Texte und 2D-Bildobjekte digitalisiert. Leerstellen ergeben sich entsprechend auf allen Ebenen: nicht alles ist katalogisiert, nicht alles ist im OPAC aufgenommen, nicht alles ist digitalisiert, nicht alles ist online usw.

Im Fall der Digitalisierung von Texten handelt es sich um eine Reproduktion dreidimensionaler Objekte in zweidimensionaler Form, es werden zumeist, etwa durch *Google Books*, Bilddateien erzeugt, die zunächst nicht maschinenlesbar sind. Just ohnehin tendenziell marginalisierte Artefakte finden sich (wenn überhaupt) nicht selten in schlechterer Qualität digitalisiert, ohne OCR, geschweige denn als TEI-Code, nicht annotiert und oftmals mit falschen oder keinen Meta-Daten versehen. Das ist die Kehrseite des Problems, dessen Auswirkungen in Marks eingangs erwähnter kleiner Studie mit *Emilia Galotti* und *Faust* aufgefallen war, denn der in der Regel prekär und zu kurz finanzierte Wissenschaftler, gerade der Nachwuchs, greift zu dem, was in besserer Qualität digitalisiert vorliegt, damit Qualifikationsschriften in kurzer Zeit angefertigt werden können. Es liegt jedoch in besserer Qualität digitalisiert tendenziell eher vor, was bereits mehr Aufmerksamkeit erhalten hat, es wird eher digitalisiert, was einfacher oder überhaupt technisch realisierbar ist, was leichter finanzierbar ist und was eher nachgefragt wird. Hinzu kommen Fragen der Annotation, der Metadaten usw., auch das bestimmt sich über die vorhandene Basis und die vorliegende Qualität, denn auch das ist bei kanonischen Werken besser und

---

<sup>19</sup> URL:

[https://pro.europeana.eu/files/Europeana\\_Professional/Projects/Project\\_list/ENUMERATE/deliverables/DSI-2\\_Deliverable%20D4.4\\_Europeana\\_Report%20on%20ENUMERATE%20Core%20Survey%204.pdf](https://pro.europeana.eu/files/Europeana_Professional/Projects/Project_list/ENUMERATE/deliverables/DSI-2_Deliverable%20D4.4_Europeana_Report%20on%20ENUMERATE%20Core%20Survey%204.pdf).

umfangreicher verfügbar. Wenn dann diese Digitalisate schließlich wieder stärker genutzt werden, dazu mehr aktuelle Forschung mit neuen Methoden erfolgt, bestätigt das – auf der Oberfläche – die zugrunde gelegten Kriterien. Zugleich verstärkt das die Leerstellen, es kann nicht beforscht werden (in Zeiten des Lockdowns und geschlossener Institutionen unmittelbar eindrücklich geworden), was nicht online sichtbar ist oder was gar nicht in digitalisierten Katalogen und digital reproduzierten Sammlungen enthalten ist.<sup>20</sup>

Jede dieser, dem eigentlichen Digitalisierungsprozess vorangegangenen, an sich kleinen und unbedeutend wirkenden Entscheidungen, die in sich sehr sinnvoll, praktisch und valide sein können, erscheinen auf einer abstrakten Metaebene von außen betrachtet ungesteuert und sind als gesamter Prozess in seiner Komplexität intransparent und schwer nachvollziehbar. Vermutlich unintendierte Konsequenzen werden damit nur manchmal sichtbar und können nur in den Extremen thematisiert werden. Nicht zuletzt das versucht der Workshop zu adressieren und einen Raum zu eröffnen, verschiedene Perspektiven zusammenzubringen und grundlegende Fragen zu stellen. Wie Andrew Prescott und Lorna Hughes noch 2018 feststellten und kritisierten, gibt es nach wie vor das, was sie einen, “significant gap in research on the impact of digitization on scholarship” nennen.<sup>21</sup> Was macht Digitalisierung eigentlich mit der Wissenschaft, wenn – und da hat *Google Books* natürlich große Anteile daran – Digitalisierung hochskaliert wird und millionenfach, schnell und viel, massenhaft Bücher gescannt und online gestellt werden? Prescott und Hughs schreiben dazu weiter:

But this advocacy for digitization has discouraged the development of critical and reflective discussions on the way in which digitization is undertaken. There is a risk that digitization programs, by focusing on making “treasures” more widely available, will reinforce existing cultural stereotypes and canonicities. The criteria used to select manuscripts for digitization and the way they are presented online are very poorly articulated and require wider discussion and debate.

---

<sup>20</sup> Das ist jenseits pandemischer Beschränkungen vor allem auch eine demokratische Frage, Wissenschaftler\_innen aus Ländern ohne die europäische Freizügigkeit in puncto Visabedingungen Bestände zugänglich zu machen. Zugleich ist es aber auch für mit Reisemitteln und -möglichkeiten ausgestatteten Forscher aus Nachhaltigkeitsgründen nicht effizient wegen einzelner Bestände weite Archivreisen zu unternehmen.

<sup>21</sup> Prescott, Andrew, und Lorna Hughes: „Why Do We Digitize? The Case for Slow Digitization“, September 2018, 15, URL: <https://eprints.gla.ac.uk/165726/7/165726.pdf>.

Since the advent of Google Books, many librarians and curators have been anxious to maximize digital coverage of their collections as quickly as possible. However, by seeking to rapidly digitize large numbers of books, manuscripts, and archives, archivists, librarians, and scholars may sacrifice many of the benefits that digital technologies offer for the exploration of manuscripts and books as textual artifacts. Too often, digitization is treated as a form of color microfilm, thereby offering distorted views of the manuscript and making it appear to be a simpler and more stable object than it really is. Digitization provides a constantly expanding toolbox for probing and analyzing manuscripts that goes beyond simple color imaging. Like archaeological artifacts, manuscripts should be explored gradually, using a variety of technical aids and methods, building a multifaceted digital archive of the manuscript.<sup>22</sup>

Prescott und Hughes geht es um handschriftliche Manuskripte, die sehr davon profitieren, qualitativ hochwertig digitalisiert zu werden, was aber einer großen Menge an Ressourcen und Expertise bedarf. Das wird entsprechend zunächst nur einem kleinen Bestand zuteil, tendenziell dem Kanon. Sie kritisieren, dass mit der Priorisierung der Digitalisierung der besonders wertvollen ‚Schätze‘, kulturelle Stereotype und kanonische Auswahlen verfestigt werden, dass die Kriterien und bisweilen vermutlich zugrunde liegenden ‚Sachzwänge‘ zudem aber auch gar nicht thematisiert werden würden. Es werde wenig darüber reflektiert, wie die Präsentation erfolge, in welcher Qualität und wie dieses Digitalisat dann wirke. Prescott und Hughes lassen ihre Kritik folglich in ein Plädoyer für langsame, im Grunde bereits erforschende Digitalisierung münden, fordern Klasse statt Masse. Vergessen dürfen wir aber eben nicht, was wir dann nicht (qualitativ hochwertig) digitalisieren und wie viel langsamer eine solche geforderte ‚slow digitization‘ ist. Grundsätzlich ist eine solche Entscheidung verständlich und kann sehr sinnvoll sein, muss aber transparent gemacht werden. Zudem braucht es ein Bewusstsein, was es bedeutet, wenn viele Ressourcen tendenziell auf wenige ‚Schätze‘ allokiert werden – sie fehlen dann an anderer Stelle.

Interessant ist das Zitat von Prescott und Hughes zudem, weil es darauf verweist, dass die Bedeutung der Digitalisierung längst über ein basales Bewahren und Zugänglichmachen hinausgeht, auch wenn das, etwa mit dem

---

<sup>22</sup> Prescott & Hughes 2018.

Brand in der Herzogin Anna Amalia Bibliothek wieder deutlicher als ebenfalls notwendige Aufgabe in den Vordergrund getreten ist. Zwangsläufig bedarf es einer Priorisierung, es kann nicht alles gleichzeitig erfolgen, und nicht selten profitieren zuerst die ‚Schätze‘. Nicht zuletzt erhöht sich mit deren Digitalisierung die Sichtbarkeit der jeweiligen Institution. Zugleich ist dafür eine Förderung einfacher zu erhalten, denn es entfällt die Begründungsnotwendigkeit: Natürlich muss das besonders Wertvolle bewahrt und auch digitalisiert werden. Nur wer bestimmt, was das ist? Wertvoll(er) im Hinblick worauf?

Zudem weisen Prescott und Hughes darauf hin, dass gerade die Kooperation mit externen Partnern dem eigentlich angestrebten Ideal einer demokratisch-offenen Zugänglichkeit konträr gegenüberstehen kann, etwa, wenn Digitalisate hinter Bezahlschranken verschwinden. Da ist paradoxerweise zu beobachten, dass bestimmte Angebote als an sich so wichtig erachtet werden, dass davon ausgegangen wird, dass Nutzer\_innen dafür zahlen würden, etwa für die Archive von Zeitungen und Magazinen. Das kann dann allerdings in der Konsequenz bedeuten, wie Laurel Brake es formuliert, dass das Gegenteil eintritt: digitized texts “are behind pay walls and that’s a real problem, and some of these are so successfully behind expensive paywalls that they are hardly seen at all.”<sup>23</sup> Man muss also jeweils die Frage stellen, für wen wird was digitalisiert und zur Verfügung gestellt? Für die breite Öffentlichkeit? Oder für eine kleine Gruppe an Experten mit spezifischen Forschungsfragestellungen? Wer genau soll zahlen? Jeder Master-Studierende, der eine Abschlussarbeit anfertigt oder jeder prekär finanzierte Doktorand? Oder die institutionelle Ebene? Worum soll es gehen, ein Service-Angebot zu bieten, eine Dienstleistung? Oder Forschung zu befördern? Es besteht eine realistische Gefahr, dass genau diese ‚Pay-Wall-Angebote‘ im Zweifel dann nicht berücksichtigt und stattdessen zu einem blinden Fleck werden.

---

<sup>23</sup> Brake, L. & Mussell, J.: “Digital Nineteenth-Century Serials for the Twenty-First Century: A Conversation”, 19: Interdisciplinary Studies in the Long Nineteenth Century. 2015(21), doi: <https://doi.org/10.16995/ntn.761>.

Der Impact von Digitalisierung ist dann tendenziell, durch Leerstellen und solche fast schon systematische blinde Flecken, Bias zu erzeugen.<sup>24</sup> Jon Coburn bezeichnet das in einem Aufsatz von 2020 als „digital selectivity“ und weist in seiner Befragungsstudie darauf hin, dass es bei den von ihm interviewten Historikern tendenziell die Regel war, diese Problematik zu ignorieren.<sup>25</sup> Dabei haben gerade die Historiker eine lange Tradition, Quellenkritik zu betreiben – im Hinblick auf die Digitalisierung muss dieser Impetus jedoch noch darüber hinaus gehen.<sup>26</sup> Nicht zuletzt verändert sich durch den gewaltigen Impact der Digitalisierung grundsätzlich, wie wir Fragen stellen, an Sammlungen, an Objekte, an Texte usw., das unterscheidet sich je nach Perspektive und Disziplin. Das scheint allerdings nicht immer deckungsgleich mit der Art und Weise, was, wie digitalisiert und präsentiert wird. Interessiert man sich etwa für handschriftliche Eintragungen in Büchern, finden sich diese seltener digitalisiert, denn oft wird durch digitalisierende Institutionen nur die reine Kopie, also das unbeschriebene Exemplar, digitalisiert. Man kann online verschiedene Digitalisate des gleichen Buchs finden, d.h. mehrfach die digitale Reproduktion einer ohnehin schon technisch reproduzierten Vorlage ansehen, aber just die Kopie mit einer Handschrift, das, was die Reproduktion dann wieder einzigartig macht, findet man eher nicht online. Lange haben diese ‚Gebrauchsspuren‘ und Einschreibungen die Forschung weniger interessiert, mittlerweile ist da allerdings ein Paradigmenwechsel zu beobachten,<sup>27</sup> der nun aber der gängigen und langjährig praktizierten Digitalisierungspraxis zuwiderläuft.

---

<sup>24</sup> Dabei gibt es ein großes Bewusstsein dafür, dass die akademische Forschung der wichtigste Grund für Digitalisierung ist, wie der Enumerate Survey ausweist. URL: [https://pro.europeana.eu/files/Europeana\\_Professional/Projects/Project\\_list/ENUMERATE/deliverables/DSI-2\\_Deliverable%20D4.4\\_Europeana\\_Report%20on%20ENUMERATE%20Core%20Survey%204.pdf](https://pro.europeana.eu/files/Europeana_Professional/Projects/Project_list/ENUMERATE/deliverables/DSI-2_Deliverable%20D4.4_Europeana_Report%20on%20ENUMERATE%20Core%20Survey%204.pdf)

<sup>25</sup> Coburn, Jon: „Defending the Digital: Awareness of Digital Selectivity in Historical Research Practice“. *Journal of Librarianship and Information Science* 53, Nr. 3 (1. September 2021): 398–410, <https://doi.org/10.1177/0961000620918647>.

<sup>26</sup> So etwa in Föhr, Pascal: *Historische Quellenkritik im Digitalen Zeitalter*. Glückstadt: VWH Verlag Werner Hülsbusch, 2019.

<sup>27</sup> Vgl. etwa den DFG SFB 933: *Materiale Textkulturen. Materialität und Präsenz des Geschriebenen in non-typographischen Gesellschaften*.

Genauso gibt es ein berechtigtes Interesse an Frauen als Autorinnen, Leserinnen, Künstlerinnen, Sammlerinnen usw., wenngleich dies viel zu lange marginalisiert worden ist und sich nun potentiell nach wie vor als Leerstelle zeigt: Man kann keine historisch-kritischen Textausgaben digitalisieren, die es nicht bereits schon in physischer Form gibt, wie dies etwa für Lessing und Goethe der Fall ist. Digitalisiert werden könnten allerdings etwa ungedruckte Manuskripte von Autorinnen, zu deren Herausgabe oder Druck es nie kam und zu denen die Forschung noch am Anfang steht. Digitalisierung (und deren Priorisierung), die diese Zusammenhänge und bisweilen strukturellen Gründe für ohnehin existierende Marginalisierungen nicht berücksichtigt und eindenkt, verfestigt eine alte Systematik, indem sie den Kanon nun auch im Digitalen festschreibt. Prescott und Hughes kritisieren entsprechend: “If digital libraries merely reiterate and reinforce long-standing cultural narratives and stereotypes, rather than enabling the exploration of forgotten and neglected collections, then they can become agents of cultural exclusion.”<sup>28</sup>

Zugleich diagnostizieren sie nach wie vor einen “lack of consensus as to what criteria should be adopted in selecting material for digitization.”<sup>29</sup> Das bildet denn auch die zugrunde liegende Kernfrage: Wer entscheidet, was, wie, in welcher Reihenfolge und für wen digitalisiert wird; was die Kriterien dafür jeweils sind und wie sie Anwendung finden? Sind sich die Akteure ihrer wichtigen und komplexen Rollen jeweils bewusst? Wie können wir die Prozesse der Digitalisierung, d.h. der Auswahl und der Priorisierung, sichtbar machen? Wie können wir sie idealerweise verwissenschaftlichen, sprich, objektiver, nachvollziehbarer, transparenter machen? Wie können wir sicherstellen, dass durch eine unreflektierte Priorisierung in der Digitalisierung nicht tendenziell problematische blinde Flecken entstehen, bzw. fortgeführt werden, die zudem dem Nutzer im Zweifel unbewusst bleiben? Wie können wir verhindern, dass Digitalisierung in der Folge und als Resultat einer Vielzahl von Auswahlentscheidungen alte Forschungsfragen zementiert, Paradigmenwechsel unterläuft und effektiv

---

<sup>28</sup> Prescott & Hughes 2018.

<sup>29</sup> Prescott & Hughes 2018.

Forschung damit behindert und nicht befördert, wenn das doch eines der dezidiert erklärten Ziele ist?<sup>30</sup>

Es gibt, neben der Kritik, und auch darauf verweisen Ooghe und Moreels in ihrer Metastudie bereits, wenngleich weniger in Zentraleuropa, sondern mehr im nordamerikanischen Raum, digitalisierende Institutionen, die sich für die Problematik sensibilisiert zeigen und Handlungsanweisungen daraus ableiten, etwa, wenn der Leiter einer kleinen texanischen Bibliothek proklamiert, seine Digitalisierungsauswahl bewusst als eine anti-rassistische Stellungnahme zu verstehen. Er schreibt: “By deciding what, among the many possible choices, to digitize in special collections and archives, we choose what narratives to promote, what history to highlight, and what legacies to further.”<sup>31</sup> Die einzelne Digitalisierung geht weit über die Sache an sich hinaus, das Digitalisat gewinnt potentiell eine große, globale Sichtbarkeit. Das wird, zuerst mit der Digitalisierung, und dann über die Online-Präsentation entschieden. Etwas zu digitalisieren ist, zugespitzt formuliert, ein eminent politischer Akt, denn es resultiert in die Nicht-Digitalisierung, bzw. Herunter-Priorisierung anderer Bestände.<sup>32</sup>

Die Konsequenz des Bibliothekars der texanischen Bibliothek ist es nun, vorrangig bislang marginalisierte Archivbestände zu digitalisieren, etwa einen Briefwechsel aus dem 19. Jahrhundert zwischen einem lesbischen Liebespaar, der in den eher pruden 1950ern und 1960er Jahren in seinem Karton ganz nach hinten in das Magazin gewandert war, oder darauf zu achten, dass grundlegend nicht nur der klassische, sehr oft weiße Kanon digitale Repräsentation erfährt, sondern just eine texanische Bibliothek in einer der US-amerikanischen Südstaaten an der Grenze zu Mexiko auch anderen Stimmen Sichtbarkeit geben muss. Das geht über die Frage von ‚Race‘ natürlich weit hinaus, betrifft Geschlecht, sexuelle Orientierung, Religion, politische Überzeugungen, Minderheiten usw. Es betrifft auf der

---

<sup>30</sup> URL:

[https://pro.europeana.eu/files/Europeana\\_Professional/Projects/Project\\_list/ENUMERATE/deliverables/DSI-2\\_Deliverable%20D4.4\\_Europeana\\_Report%20on%20ENUMERATE%20Core%20Survey%204.pdf](https://pro.europeana.eu/files/Europeana_Professional/Projects/Project_list/ENUMERATE/deliverables/DSI-2_Deliverable%20D4.4_Europeana_Report%20on%20ENUMERATE%20Core%20Survey%204.pdf)

<sup>31</sup> Ziegler, S.L.: “Digitization Selection Criteria as Anti-Racist Action”, in: Code{4}lib Journal, Issue 45, 2019, URL: <https://journal.code4lib.org/articles/14667>

<sup>32</sup> Für die Archivwürdigkeit gibt es etwa klare Kriterien – wäre das vergleichbar auch für die Digitalisierung nötig und möglich?

Ebene der Präsentation dieser Digitalisate natürlich auch die Frage von Provenienz, von Master-Narratives und von kultureller Hegemonie. Das scheint nicht nur eine Frage zu sein, ob Objekte, Texte, Sammlungen historischer Artefakte insgesamt ‚angemessen‘ repräsentiert werden, sondern auch eine danach, wie das nach wissenschaftlichen Kriterien erfolgen kann. Es geht dabei nicht um Aktivismus und darum, aus einem Prozent einhundert zu machen, sondern um historisch belegbare Heterogenität. Dazu gehört es, Repräsentativität und Sichtbarkeit von beispielsweise Minderheiten auch in den digitalen Sammlungen und Repositorien sicher zu stellen, soweit es geht, d.h. soweit es diese Artefakte gibt. Das beginnt entsprechend damit, kritisches Bewusstsein auszubilden und Zusammenstellungen zu hinterfragen, die auffallend homogen sind und Leerstellen aufweisen: Im Falle deutscher Literatur etwa fehlen sehr oft jüdische oder weibliche Autoren und es kann deswegen der Eindruck entstehen, dass dies schlicht nicht existiert, obgleich das nicht zutrifft.<sup>33</sup> Bei circa zehn Prozent jüdischer Bevölkerung in Deutschland um 1900 muss eine jede Textsammlung zu dieser Zeit stutzig machen, die keinen oder nur vereinzelt jüdisch geborene Autor\_innen enthält.<sup>34</sup> Sprich, was die digitalen Sammlungen zeigen und wie sie es präsentieren, verweist im Umkehrschluss darauf, was sie nicht zeigen, was aus ihnen entfernt worden ist, etwa aus antisemitischen oder misogynen Gründen. Das berührt eine grundsätzliche Frage danach, ob etwas wirklich fehlt, es z.B. keine briefeschreibenden, lesbischen Pärchen im 19. Jahrhundert gegeben hat, oder solche vereinfachte Annahmen unterlaufenden Beispiele eigentlich sehr wohl in den physischen Archiven liegen, sie sich aber nicht – aus welchen Gründen auch immer – in den digitalen finden, außer jemand erklärt es zur Priorität.

Die Problematik fällt etwa auch in den Netzwerkdarstellungen des Kalliope-Verbunds von Handschriften ins Auge: obgleich es sich beim Brief doch

---

<sup>33</sup> Ruth Whittle weist in ihrer Studie darauf hin, dass bereits am Ende des 19. Jahrhunderts jüdische und weibliche Autoren zunehmend aus der literaturgeschichtlichen Tradierung herausfallen. Vgl. Whittle, Ruth: *Gender, Canon and Literary History: The Changing Place of Nineteenth-Century German Women Writers (1835-1918)*. Berlin 2013.

<sup>34</sup> So gibt es etwa einen heute kaum bekannten jüdisch geborenen Literaturnobelpreisträger mit Paul Heyse oder den Herausgeber und Autor Julius Rodenberg, die Bestsellerautorin Fanny Lewald, um nur die wichtigsten zu nennen. In manchen Online-Repositorien findet sich auch mal ein einzelner Text von ihnen, das bildet aber bei weitem nicht die Breite ihres Schaffens ab.

eigentlich um eine ‚weibliche‘ Gattung handelt,<sup>35</sup> sollen bei teilweise hunderten von Korrespondenzpartnern alle nahezu ausschließlich Männer gewesen sein, wie etwa im Netzwerk von Wilhelm Bartholomaei ersichtlich, einem fleißigen Briefeschreiber aus Weimar, Mitte des 18. Jahrhunderts, einem Pfarrer.<sup>36</sup> Es mag sein, dass das so war – aber ist es wirklich plausibel? Ist es nicht viel logischer anzunehmen, dass die Briefe seiner Korrespondenzpartnerinnen vermutlich (noch) nicht erfasst und nicht ohne weiteres digitalisierbar vorliegen?<sup>37</sup> Bisweilen ist das zugrundeliegende Problem mangelnder Sichtbarkeit wohl weniger tatsächliche Nicht-Existenz, sondern eher, dass so manches noch unerschlossen, unerforscht und undigitalisiert in den Archiven liegt. Wenn nun aber eine Sammlung als Kriterium für die Digitalisierung ausgerechnet etwa die Nutzungshäufigkeit in Anschlag bringt, kann und wird sich das nicht ändern.

Problematisch ist es entsprechend, wenn – Stichwort ‚anekdotische Evidenz‘ – in Museen die Bilder von Malerinnen im Magazin stehen und gar nicht erst ausgepackt werden, weil man weiß, da müssten erstmal sehr viel Zeit, Expertise und Ressourcen hineingehen, bevor man über Digitalisierung überhaupt nur nachdenken kann; oder wenn es wenig Interesse daran gibt, ihre Urheberschaft in den Katalogen korrekt zu verzeichnen, weil das mit einem Wertverlust einhergehen würde, da das Bild bislang dem weit bekannteren Ehemann oder Vater zugeschrieben wurde.<sup>38</sup> Die bisweilen problematische Provenienz vor dem Hintergrund von Enteignungen in der deutschen Geschichte manch eines musealen Objekts mag da noch dazu kommen, dass einige Artefakte bewusst unbeleuchtet, metaphorisch und wörtlich, d.h. undigitalisiert in den hinteren Winkeln des Magazins

---

<sup>35</sup> Barthel, Katja: Gattung und Geschlecht: Weiblichkeitsnarrative im galanten Roman um 1700. Berlin; Boston: Walter de Gruyter, 2016. S. 155; Dampe-Jarosz, Renata, und Pawel Zarychta (Hg.): „... nur Frauen können Briefe schreiben“: Facetten weiblicher Briefkultur nach 1750. Berlin: New York: Peter Lang, 2019.

<sup>36</sup> URL: <https://kalliope-verbund.info/de/graph?q=Wilhelm%20bartholomaei&fq=ead.pers.index%3A%28%22Bartholomaei%2C%20Wilhelm%20Ernst%20%281697-1753%29%22%29&lastparam=true>

<sup>37</sup> Gerne wird ja an dieser Stelle das Pauschalargument genutzt, Frauen hätten nicht lesen oder schreiben gekonnt, in den damaligen thüringischen Herzogtümern ist das aber nachweislich der Fall, bereits im 16. und 17. Jahrhundert wird dort nahezu flächendeckend die Schulpflicht eingeführt.

<sup>38</sup> Etwa bei der Familie Tischbein, URL: <https://www.hna.de/lokales/frankenberghaina-kloster-ort317400/vergessene-schwester-ausstellung-rueckt-frauen-malerfamilie-tischbein-licht-6252864.html>

verbleiben. Dabei ist offenkundig, welches Potential die Digitalisierung bietet, mit so manch einer schiefen Vorstellung aufzuräumen, endlich, vielleicht erstmals, Sichtbarkeit herzustellen, sprich also letztlich, bessere Wissenschaft zu machen, mehr und genauer zu verstehen. Damit ist aber auch klar, welche Gefahr genauso darin liegen kann, Verzerrungen überhaupt erst zu produzieren oder zu verstärken. Das berührt systemische Fragen, zur Aufgabe und Funktion von Archiven und Bibliotheken – zu welchem Zweck diese digitalisieren sollen und nach welchen Kriterien. Es wäre doch etwa an sich auch ein solides Kriterium, gerade das zu digitalisieren und in der Priorität hochzustufen, was bislang keine Forschungsaufmerksamkeit gefunden hat, zu digitalisieren (und sei es zunächst ‚quick & dirty‘), um Forschung zu ermöglichen, gerade um marginalisierten Werken, Autoren und Künstlern mehr, bzw. überhaupt Sichtbarkeit zu verschaffen. Als Beispiel sei da etwa wieder auf die in der deutschen Literatur im 19. Jahrhundert viel gelesenen jüdisch geborenen Autoren und Autorinnen hingewiesen, deren Erinnerung dem wachsenden Antisemitismus des 19. Jahrhunderts in Deutschland zum Opfer gefallen ist und die heute nicht mehr im kulturellen Gedächtnis sind. Wenn sie auch nicht in das digitale Archiv Eingang finden, kann und wird sich daran nichts ändern. Es gibt andere, weit komplexere Beispiele marginalisierter Aspekte, Themen und Bereiche, deren Archivierung und Digitalisierung bisweilen noch viel komplizierter ist, Zeugnisse oraler Kulturen etwa, flüchtige Ereignisse usw. Insgesamt hilft es wohl zunächst nur, die komplexen Auswahlprozesse transparent zu machen und vor allem überhaupt zu thematisieren; die anekdotische Evidenz hinter uns zu lassen, auch wenn das womöglich bedeuten kann, dass dabei die eine oder andere unbequeme Systematik zum Vorschein kommt.

Zum Schluss der einführenden Bemerkungen geht es darum, von der konkreten Problematik der Repräsentativität auf eine allgemeinere Ebene zu wechseln. Digitalisierung ist ein ‚boundary object‘,<sup>39</sup> in dem die Frage der

---

<sup>39</sup> Star, Susan Leigh, und James R. Griesemer: “Institutional Ecology, ‘Translations’ and Boundary Objects: Amateurs and Professionals in Berkeley’s Museum of Vertebrate

Repräsentativität nur ein Aspekt ist, verbunden mit der übergeordneten Frage der Selektivität.<sup>40</sup> Der Komplex wird in der Forschung etwa auch unter Begriffen wie postcolonial computing,<sup>41</sup> Artificial Whiteness<sup>42</sup> oder Intersektionalität<sup>43</sup> problematisiert. Angesichts von Big Data und der großen Menge bereits digitalisierter Artefakte gibt es aufgrund allseits knapper Ressourcen unzweifelhaft eine Notwendigkeit der Kuratierung, es muss ausgewählt werden, es müssen Entscheidungen getroffen werden, hinsichtlich der Präsentation, der Priorisierung usw.<sup>44</sup> Zudem gibt es die jeweils in sich komplexen Probleme der Annotation, Metadaten und der Findbarkeit zu berücksichtigen. Es eröffnet sich das utopische Potential des Internets als demokratischem, alles Wissen umfassenden Orts, der allen zugänglich ist, so wie es idealerweise auch Bibliotheken, Museen und Universitäten sein sollten, aber es handelt sich dabei natürlich um eine Idealvorstellung. Das gilt nicht nur für autoritäre Regime, sondern jede Suchmaske kann zu einer nur für den Eingeweihten überwindbaren Barriere werden. Zugespitzt formuliert: Niemandem helfen Massen an unerschlossenen Digitalisaten, die nicht auffindbar sind.

Zudem haben wir noch als aktuell stärker wachsendes Bewusstsein die Frage der Nachhaltigkeit zu berücksichtigen: Einerseits als Frage, wie man Ressourcen schonen kann, etwa bei der Digitalisierung von Artefakten, aber

---

Zoology, 1907-39“, in: Social Studies of Science 19, Nr. 3, 1989: 387–420.

<https://doi.org/10.1177/030631289019003001>.

<sup>40</sup> Vgl. hierzu die Opening Keynote von Amalia S. Levi auf der DHd2022 sowie Zaagsma, Gerben: The Digital Archive and the Politics of Digitization. In: DHd 2022 Kulturen des digitalen Gedächtnisses. 8. Tagung des Verbands „Digital Humanities im deutschsprachigen Raum“ (DHd 2022), Potsdam, <https://doi.org/10.5281/zenodo.6328223>.

<sup>41</sup> Irani, L.; Vertesi, J.; Dourish, P., Philip K. und R.E. Grinter: “Postcolonial computing: a lens on design and development”, in: Proceedings of the SIGCHI conference on human factors in computing systems 2010: 1311-1320; Philip, Kavita; Lilly Irani, und Paul Dourish: “Postcolonial Computing: A Tactical Survey“. Science, Technology, & Human Values 37, Nr. 1 (1. Januar 2012): 3–29. <https://doi.org/10.1177/0162243910389594>; Kizhner, Inna [et.a.]: “Digital cultural colonialism: measuring bias in aggregated digitized content held in Google Arts and Culture“, in: Digital Scholarship in the Humanities 36, Nr. 3, 2021: 607-40. <https://doi.org/10.1093/llc/fqaa055>; Ali, Syed Mustafa: “A Brief Introduction to Decolonial Computing“. XRDS: Crossroads, The ACM Magazine for Students 22, Nr. 4 (13. Juni 2016): 16–21. <https://doi.org/10.1145/2930886>.

<sup>42</sup> Katz, Yarden: Artificial Whiteness: Politics and Ideology in Artificial Intelligence. Columbia University Press, 2020.

<sup>43</sup> Risam, Roopika: “Beyond the Margins: Intersectionality and the Digital Humanities“, in: digital humanities quarterly, Volume 9 Number 2, 2015, URL: <http://digitalhumanities.org:8081/dhq/vol/9/2/000208/000208.html>

<sup>44</sup> Poole, Alex H.: “Now is the Future Now? The Urgency of Digital Data Curation in the Digital Humanities“, In: Digital Humanities Quarterly 7 (2)

auch bei der Bereitstellung,<sup>45</sup> denn nicht alles muss von jeder Institution selbst digitalisiert und bereit gehalten werden, weit klüger ist oftmals ein Link.<sup>46</sup> Die frei werdenden Ressourcen könnte man dann etwa einsetzen, um die eigenen Digitalisate gründlicher zu erschließen. Andererseits berührt die Frage der Nachhaltigkeit die dauerhafte Sicherstellung der Zugänglichkeit, über das jeweilige Projektende hinaus, für die es entsprechende Infrastruktur braucht. Oftmals, und dem widmet sich ein Beitrag von Ciaran Trace der aktuellen Ausgabe der *DHquarterly*, müssen die Ressourcen allerdings in die – für den normalen Nutzer nur sichtbare, wenn sie nicht erfüllt wird – Aufgabe des Maintenance gesteckt werden, insofern digitale Archive eben auch Infrastruktur sind, die gepflegt, gewartet und erhalten werden müssen, während die Finanzierung eines Projekts die langfristige Pflege und Bereitstellung der Daten in der Regel nicht umfasst.<sup>47</sup>

Mit anderen Worten: Die Komplexität der multidimensionalen Problematik, wie sie etwa aus der Studie von Ooghe und Moreels spricht, umfasst strukturelle wie technische Organisation, und hat politische, rechtliche und ethisch-moralische Dimensionen, berührt Fragen der digitalen Systeme genauso wie die der Dokumentation und Reproduzierbarkeit. Dazu kommen die verschiedenen Funktionen oder Aufgaben von Digitalisierung, seien das demokratische, konservatorische oder didaktische. Vor diesem Hintergrund soll sich die neu zu gründende DHd-Arbeitsgruppe diesen und weiteren, hier nur skizzierten Fragen, Problemen und Aspekten widmen, Auswahl- und Priorisierungsprozesse bei der Digitalisierung historischer Artefakte zu diskutieren, reflektieren und mehr Transparenz herzustellen.

Wir freuen uns auf den Austausch!

---

<sup>45</sup> Barats, Christine: "Fading Away...The Challenge of Sustainability in digital studies", in: digital humanities quarterly, Volume 14 Number 3, 2020, URL: <http://www.digitalhumanities.org/dhq/vol/14/3/000484/000484.html>

<sup>46</sup> Dem nähern sich Stuart Lewis, Paul Gooding and Mike Furlough von der Frage ausgehend, ob man ein Global Digitisation Dataset erstellen könnte, URL: <https://www.digitales-archiv-und-kanon.de/dac2021/contributions/lewis-gooding-furlough>

<sup>47</sup> Trace, Ciaran B.: „Archives, Information Infrastructure, and Maintenance Work“. Digital Humanities Quarterly Vol. 16, Nr. 1 2022, URL: <http://digitalhumanities.org:8081/dhq/vol/16/1/000603/000603.html>