

Thorsten Wübbena

KNOTEN, KANTEN, KONTEXTE – EIN IMPULS ZU »SAMMLUNGEN ALS DATEN«

Ein Netz besteht ja nicht nur aus Knoten, sondern auch aus Löchern.
Peter Weibel¹

Sammlungen in Galerien, Bibliotheken, Archiven und Museen (GLAM) werden dort nicht nur physisch geordnet, sie werden zugleich auf der Ebene ihrer Beschreibung in definierte Wissensordnungen eingeordnet bzw. definieren Ordnungen selbst mit. Diese beschreibenden Informationen bilden in Form von (Meta-)Daten seit jeher einen wichtigen Baustein der wissenschaftlichen Forschung in den Geisteswissenschaften. Die Digitalisierungskampagnen im GLAM-Bereich brachten die digitalen Repräsentanten der Objekte und die zugehörigen Metadaten in einem neuen Raum zusammen – außerhalb der physischen Architekturen. Damit wurden auch die Voraussetzungen für den Zugang zu »Sammlungen als Daten«² geschaffen. Ohne ein ordnendes Narrativ, wie es der physische Raum erzeugen würde, können die Daten im digitalen Raum eine, wie Gregory Bateson es nannte, »zukünftige Information«³ bilden. Als Daten bleiben sie konstant und unveränderlich, doch in unterschiedlichen Konstellationen, gleichzeitig und wechselnd, zufällig und kreierte, wird ihre Ordnung veränderlich: »was verschieden ist, sind die Modalitäten ihrer Beziehungen«,⁴ wie es Hayden White in seinem Buch *Auch Klio dichtet* formuliert. Ihre prinzipielle Unabgeschlossenheit und Multiperspektivität sind charakteristisch für diese Sammlungen.

Dieser Beitrag fragt danach, auf welchem Weg, unter welchen Bedingungen und angesichts welcher Herausforderungen sich die Vernetzung von »Samm-

- 1 Gerhard Johann Lischka: Polylog. Für eine interaktive Kunst, in: *Kunstforum*, Band 103, 1989, S. 65–86, hier S. 73.
- 2 Dieser Ausdruck ist dem Projekt »Always Already Computational: Collections as Data« entnommen. Mehr zu diesem Vorhaben in: Thomas Padilla, Laurie Allen, Hannah Frost, et al.: *Final Report – Always Already Computational: Collections as Data*. 2019, <https://doi.org/10.5281/zenodo.3152935>.
- 3 Gregory Bateson: *Ökologie des Geistes*. Anthropologische, psychologische, biologische und epistemologische Perspektiven, Frankfurt am Main 1983, S. 488.
- 4 Hayden White: *Auch Klio dichtet oder Die Fiktion des Faktischen*. Studien zur Tropologie des historischen Diskurses, Stuttgart 1986, S. 119.

lungen als Daten« konstituiert. Angesichts seines Formats als Impuls⁵ werden hierfür relevante Punkte wie u. a. der Umgang mit Vagheit und Zweifel,⁶ die Aufhebung des exklusiven Zugangs zu verknüpften Informationen und Entgrenzung im Netzwerk nur angerissen und können komplementär zu den anderen Beiträgen in diesem Band gelesen werden.

Teilen

Zu dem Behalten und Bewahren von Sammlungen kommt seit der Digitalisierung die Aufgabe des »erschließenden Aufbereitens, Ordners und Teilens« hinzu, wie Thomas Stäcker es formulierte.⁷ Schaut man sich die Praxis des Teilens nun genauer an, so ist festzustellen, dass die Bereitstellung von »Sammlungen als (vernetzte) Daten«, die für alle Anforderungsszenarien – im Sinne einer »one size fits all«-Lösung – gedacht sind, oft niemandem dienen. Die spezifischen Bedürfnisse der Nutzenden erfordern hingegen mehrere, verschiedene Ansätze des Datenzugangs. Zunächst wird mit der Online-Bereitstellung der Digitalisate der ubiquitäre lesende Zugriff auf die zum Teil zuvor schwer zugänglichen Objekte (Rara, Depotverwahrung) ermöglicht. Diese Nutzenden, die durch »digitales Umblättern« die virtuelle Sammlung erkunden möchten, werden bei der Konzeption von Online-Angeboten in den Museen und Bibliotheken zumeist in den Fokus gestellt. Konzentriert sich die vorherrschende digitale Sammlungsentwicklung aber vielleicht zu sehr auf die Replikation traditioneller Rezeptionsformen mit Objekten in einem digitalen Raum? Gerade die Digitalisierung und Datafizierung⁸ ermöglicht einen multiperspektivischen und »rechnenden« Zugriff auf das Material. Eine Sammlung unter dem Datenimperativ bereitzustellen bedeutet auch, möglichst allen Nutzenden Mittel bereitzustellen, sich kritisch mit den

5 Der hier vorliegende Beitrag geht auf meinen Impuls für die Podiumsdiskussion der Tagung am 30. März 2022 zurück.

6 Mehr dazu bei Andreas Kuczera, Thorsten Wübbena und Thomas Kollatz (Hg.): Die Modellierung des Zweifels. Schlüsselideen und -konzepte zur graphbasierten Modellierung von Unsicherheiten, Wolfenbüttel 2019 (Zeitschrift für digitale Geisteswissenschaften, Sonderbände 4), <https://doi.org/10.17175/sbo04>.

7 Thomas Stäcker: Die Sammlung ist tot, es lebe die Sammlung!, in: BIBLIOTHEK – Forschung und Praxis 2019, 43(2), S. 304–310, hier S. 309. <https://doi.org/10.1515/bfp-2019-2066>.

8 Siehe zum Begriff »datafication« insb. Kenneth Neil Cukier, Viktor Mayer-Schoenberger: The Rise of Big Data, in: Foreign Affairs. May/June 2013. <https://www.foreignaffairs.com/articles/2013-04-03/rise-big-data>.

Sammlungen auseinanderzusetzen. Und zwar so umfassend wie möglich, der Komplexität ihrer Form entsprechend und in kritischem Bewusstsein der Möglichkeiten und Herausforderungen, die mit ihrer Nutzung einhergehen. Ebenso selbstverständlich wie Leseansichten braucht es die Verfügbarmachung umfassender Objektdaten⁹ per Schnittstelle (API).

Die Ermöglichung der Erforschung des kulturellen Erbes mit digitalen Methoden und Verfahren, um Fragen zu stellen und nach Mustern zu suchen, wird insbesondere im Licht der erweiterten, vernetzten (Sammlungs-)Datenbestände neue Potentiale und Erkenntnisse fördern. Das bedeutet für die sammlungsbewahrenden Gedächtnisinstitutionen eine neue Rolle als Daten-Hub – und ein Zusammendenken von Informations- und Forschungsinfrastrukturen. »HathiTrust« bietet für die computergestützte Analyse der bereitgestellten Werke sowohl Datensätze als auch Webtools an.¹⁰

Die bloße Freigabe von Daten bedeutet nicht, dass sie auch genutzt werden, und das Anbieten einer API bedeutet nicht, dass alle Bedürfnisse der Forschung befriedigt sind. Viele gängige Werkzeuge zur Weiterbearbeitung von Daten erfordern im Gegensatz zur Abfrage einer API keine Programmierkenntnisse. Die Entscheidung, neben dem API-Zugang zu den eigenen Daten auch weniger »glamouröse« Wege wie »flat files« anzubieten (CSV, JSON etc.), kann für Kultureinrichtungen den Unterschied zwischen Nutzung und Nichtnutzung machen.

Kanon

Vor der Sichtbarwerdung und Nutzungsmöglichkeit der Daten in Online-Szenarien haben in Sammlungen Selektions- und Ordnungsprozesse stattgefunden. Vor der Sammlungsvernetzung auf horizontaler Ebene (der Vernetzung mit Sammlungen aus anderen Gedächtnisinstitutionen) wird das Fundament hierfür auf vertikaler Ebene gelegt, also innerhalb einer Sammlung. So ist mit den Digitalisierungskampagnen (der Vergangenheit) zum Beispiel durchaus die Gefahr negativer Auswirkungen auf eine Kanonrevision bzw. -erweiterung verbunden. Denn Auswahlentscheidungen folgen neben technischen und urheberrechtlichen Aspekten zumeist auch entlang der »Highlights« einer Sammlung. Die nicht-digitalisierten Werke stehen im »digitalen Schatten« und laufen Gefahr, abseits der Online-Angebote umso mehr in Vergessenheit zu geraten. Auch reine Digitalisierungsprogramme verlieren daher nicht an

⁹ Umfasst die datafizierten Objekte samt Metadaten.

¹⁰ »HathiTrust Research Center Analytics«, <https://analytics.hathitrust.org/>.

Bedeutung, insbesondere kleinere Häuser sollten noch stärker partizipieren (können), um die verfügbare Materialvielfalt abzubilden.

Verantwortung

In den bewahrenden Institutionen ging der Weg der Daten oftmals von der analog erstellten Karteikarte über die initiale Digitalisierung zur Migration in die erste (Online-)Sammlungsdatenbank. Dieses Vorgehen birgt Herausforderungen für die Verfügbarmachung dieser tradierten Informationen in vernetzten Sammlungen. Oftmals für die interne Verwaltung konzipiert und erstellt, konnten in den Datenbanksystemen fehlende Spezifikationen, Mehrdeutigkeiten oder Unsicherheiten durch Kenntnisse von Expert:innen vor Ort kompensiert werden. Nur selten wurden Datenbestände im Zuge der Integration oder Überführung von bestehenden Systemen in technologisch neuere Umgebungen intellektuell bearbeitet, bereinigt oder ergänzt. Um Sammlungen als Daten außerhalb des Expert:innenkreises vor Ort nutzbar zu machen, müssten sowohl die Entstehungskontexte und Veränderungsprozesse lesbar gemacht werden.

Die genutzten Klassifikationssysteme sind hier mitgemeint. Derartige Wissensordnungen sind kulturelle, subjektiv geprägte Erzeugnisse einer spezifischen Zeit und weder gegeben, objektiv noch allgemeingültig. Eine Klassifizierung, die nicht aktualisiert und angepasst wird, steht in der Gefahr, Konzepte und Wertvorstellungen unreflektiert in neue Anwendungen hinein zu tradieren.¹¹

Vernetzen/Entnetzen

Sobald eine Verbindung zwischen zwei Knoten besteht, gibt es ein Netzwerk, und es kann durch die Verknüpfung mit neuen Knoten wachsen. Somit bietet die Netzform die Möglichkeit zu einer nahezu unendlichen Ausdehnung. Damit verbunden ist das innewohnende Versprechen auf Anschlussfähigkeit, mit der Option, dass ein Netzwerk sich in unerwartete Richtungen entwickeln und sich dabei verzweigen und auseinandergehen kann. Hierin enthalten ist auch die Fähigkeit, neue Anfänge zu erschaffen.

¹¹ Siehe hierzu insb. den Blogbeitrag von Alina Kühnl: Iconclass. Ein Klassifizierungssystem für Kunst und Mensch?, 6.10.2020 (aktualisiert: 3.11.2020). <https://thearticle.hypotheses.org/9773>.

Kann es im Sammlungskontext auch eine zu starke Vernetzung geben? Zum Netz(-werk) gehören wie die Knoten und verbindenden Kanten eben auch die Zwischenräume, die »Löcher«, wie es Peter Weibel beschreibt. Ein zu engmaschiges Netz könnte zu einer ungewollten Verengung im Hinblick auf die Beschäftigung mit den vorliegenden Informationen führen, so dass anstelle einer gewollten Inhaltstiefe ein Inhaltskollaps eintritt.

Bei der permanenten Erweiterung und Intensivierung und dem »natürlichen Hunger« von Netzwerken nach mehr Verbindungen wird das Augenmaß zu wahren zur Herausforderung. Zunächst nur eine technische Herausforderung, tangiert das Verbinden von Sammlungen in Netzwerken auch soziale und politische Fragen. So zum Beispiel finanzielle Aspekte der Infrastruktur – wer investiert und wer monetarisiert die offene Bereitstellung von Daten? –, aber auch die Möglichkeiten der Datenbeeinflussung.

Die Entwicklung und Bereitstellung von Sammlungen als Daten ist kein feststehender Prozess und kein abgeschlossenes Projekt, sondern eine fortlaufende Aufgabe, die die produktive Integration neuer und bestehender Technologien, Arbeitsabläufe und Dienstleistungsmodelle sowie neuer und bestehender wissenschaftlicher Erkenntnisse und gesellschaftlicher Entwicklungen sucht.