

GOETHE'S FARBENLEHRE ALS NETZWERK

VOM WERKKOMPLEX ZUM KNOWLEDGE GRAPH

Die Aufgabe, die umfassenden Meta- und Forschungsdatenbestände der Kulturerbe-Institutionen nicht nur digital zur Verfügung zu stellen, sondern ebenso die zugrundeliegenden semantischen Beziehungsnetzwerke erfassen und abfragen zu können, hat in den letzten Jahren mehr und mehr an Bedeutung gewonnen. Das Ziel ist es, die meist nur für sich bestehenden institutionellen Daten-Silos aufzubrechen und den Schritt hin zu einem sogenannten Semantic Web zu gehen, in dem Daten nicht nur aufeinander verweisen, sondern diese Verweise auch ihrer Bedeutung nach differenzierbar sind. Auf diese Weise kann das in den Datenbanken umgesetzte Wissen adäquater abgebildet werden.¹ Der viel beschworene Begriff des ›Knowledge Graph‹ ist ein weiteres Schlagwort, welches gern in diesem Zusammenhang verwendet wird.² In diesem Kontext steht auch das an der Klassik Stiftung Weimar (KSW) angesiedelte MWW-Verbundprojekt des *Virtuellen Sammlungsraums*.³ Die in der Fallstudie »Goethe Digital« erhobenen Daten zum Werkkomplex »Goethes Farbenlehre« dienen als ein Beispiel, an dem die Integration von Daten in ein semantisch ausdifferenziertes Netzwerk technologisch erprobt werden soll. Grundlage dafür sind 95 Datensätze zu unterschiedlichen Objekten aus dem Entstehungszusammenhang von Goethes Farbenlehre. Dabei soll sowohl in der Fallstudie als auch im vorliegenden Beitrag das Datennetzwerk nicht nur

- 1 Vgl. hierzu auch Brenda O'Neill und Larry Stapleton: Digital Cultural Heritage Standards. From Silo to Semantic Web, in: *AI & Society*, 37, 2022, S. 891–903, sowie Osma Ilmari Suominen und Nina Hyvönen: From MARC silos to Linked Data silos?, in: *obib. Das offene Bibliotheksjournal*, 4, 2017, Nr. 2, DOI: <https://doi.org/10.5282/o-bib/2017H2S1-13>. Sämtliche Internet-Links in diesem Beitrag wurden am 15.12.2023 überprüft.
- 2 Zum Konzept des Knowledge Graph im Kulturerbekontext siehe u.a.: Dominic Oldman und Diana Tanase: Reshaping the Knowledge Graph by Connecting Researchers, Data and Practices in ResearchSpace, in: *The Semantic Web – ISWC 2018. Proceedings Part II*, hg. von Denny Vrandečić, Kalina Bontcheva, Mari Carmen Suárez-Figueroa, Valentina Presutti, Irene Celino, Marta Sabou, Lucie-Aimée Kaffee, Elena Simperl, Cham 2018, S. 325–340, hier S. 330–331.
- 3 Für weitere Informationen zum Projekt siehe <https://vfr.mww-forschung.de/web/virtueller-sammlungsraum>.

auf der Ebene der Erfassung und Speicherung der Datensätze in einer Datenbank betrachtet, sondern ebenso der Frage nachgegangen werden, wie dieses Netzwerk den Nutzer*innen vermittelt werden kann.

Der Virtuelle Sammlungsraum

Die Geschichte des *Virtuellen Sammlungsraums* ist eng mit dem Konzept des *Sammlungserschließenden Katalogs* verbunden, der Titel, unter welchem das Projekt ursprünglich gestartet ist. Letzterer ist nicht nur als ein bloßes User Interface zum Durchsuchen von Beständen zu verstehen, sondern steht stellvertretend für ein »Recherche- und Wissensmodell«, welches »dynamische Relationen« darstellen sowie »Entstehungs- und Überlieferungszusammenhänge« abbilden kann.⁴ Die im selben Kontext gebrauchte Formulierung vom »Wissen der Sammlungen«⁵ lädt gerade dazu ein, die Brücke zu Ansätzen wie dem Knowledge Graphen zu schlagen. Für das Konzept des *Virtuellen Sammlungsraums* als einer möglichen Realisation eines solchen *Sammlungserschließenden Katalogs* hat dies ganz konkrete Folgen:

1) Wissen und Zusammenhänge werden nicht nur abgebildet, sondern erzeugt, vermittelt und erfahrbar gemacht.

2) Dafür muss die Idee des *Sammlungserschließenden Katalogs* über einen verzeichnenden Katalog hinausweisen.

Kataloge und Inventarlisten bzw. ihre digitalen Pendanten und Erweiterungen in Form von Portalen und Discovery-Systemen stellen Informationen in einer übersichtlichen und leicht verständlichen, gegliederten Art und Weise dar.⁶ Darüber hinaus sind sie für größere Datenmengen gut skalierbar, indem sie die Daten auf mehrere Bände (z.B. für Inventarbücher) oder Seiten verteilen bzw. bei Bedarf nachladen. Gedruckte Register oder digitale Filterfunktionen erlauben es, schnell Einträge nach bestimmten Kriterien zu durchsuchen und so die entsprechenden Informationen zu gewinnen. Doch

4 Reinhard Laube: Die Herzogin Anna Amalia Bibliothek im Jahr 2019, in: *Supra-Libros. Mitteilungen der Gesellschaft Anna Amalia Bibliothek e.V.*, 25, 2020, S. 1–5, hier S. 2.

5 Reinhard Laube: Das Wissen der Sammlungen. Die Zukunft der Archiv- und Forschungsbibliothek, in: *Zeitschrift für Bibliothekswesen und Bibliographie*, 67, 2020, Nr. 2, S. 6–14.

6 Discovery Systeme können insofern bereits als Erweiterung der Idee des traditionellen Katalogs gesehen werden, als in ihnen nicht nur Bestände verzeichnet werden, sondern ebenso weitere Datenräume angegliedert und suchorientierte Hinweise gegeben werden können.

gerade Aspekte wie dynamische Relationen und Überlieferungszusammenhänge lassen sich eher schwierig mit Listen- und Kachelansichten abbilden.

Beziehungen können hingegen sehr gut als Knoten und Kanten eines Netzwerks visualisiert werden. Die Möglichkeit, über Interaktivität Zusammenhänge nicht nur abzubilden, sondern gleichzeitig spielerisch erfahrbar machen zu können, ist ein weiterer Vorzug. So wird das Wissen, welches durch die dargestellten Netzwerke repräsentiert wird, nicht nur präsentiert, sondern gleichzeitig vermittelt, indem die Nutzer*innen damit interagieren und die Ansichten manipulieren können.⁷ Das Netzwerk wird hierdurch nicht nur zu einem Modell, mit dem Sammlungen modelliert und analysiert werden können, sondern ebenso zu einem Interaktionsraum, einem (virtuellen) Ort, an dem sich Sammlungen und Bestände durchsuchen und aus verschiedenen Blickwinkeln in Augenschein nehmen lassen.

Das Ergebnis dieser Überlegung ist ein prototypischer Workflow sowie die Erstellung einer Web-Anwendung als einer möglichen beispielhaften Realisierung eines virtuellen Sammlungsraumes. Dabei wurde versucht, das Netzwerk als Konzept der allseitigen Verknüpfung auf den verschiedenen technischen Ebenen der Anwendung zu integrieren. Im Backend bzw. dem Datenbankbereich wurde dies durch den Einsatz sogenannter Graphdatenbanken umgesetzt. Für die Schnittstellen wurden Ansätze mit dem Framework GraphQL und einer SPARQL-Schnittstelle getestet. Im Frontend, dem Bereich, welcher für die Nutzer*innen der eigentlich sichtbare Interaktionsbereich ist, wurde mit Visualisierungen gearbeitet, welche das dahinterliegende Datennetz darstellen und vermitteln sollen.

Mit Hilfe von Graphdatenbanken können – sehr vereinfacht ausgedrückt – Daten in Form eines Netzwerks bzw. im technisch-mathematischen Sinn als Graphen organisiert werden. Die verbreitetsten Ansätze sind die auf dem Resource Description Framework (RDF) als Datenmodell basierenden Triplestores sowie sogenannte Labeled Property Graph-Datenbanken. RDF ist eine der Schlüsseltechnologien des Semantic Web und von Linked Open Data. Daten werden nach dem Prinzip Subjekt-Prädikat-Objekt (Person – hat Name – Johann Wolfgang von Goethe) als sogenannte Triple kodiert. Subjekt und Objekt bilden dabei Knoten und das Prädikat die Kanten des

7 Zum Aspekt der Interaktion mit Visualisierungen siehe u.a. auch: Ji Soo Yi, Youn ah Kang, John Stasko und J.A. Jacko: Toward a Deeper Understanding of the Role of Interaction in Information Visualization, in: IEEE Transactions on Visualization and Computer Graphics, 13, 2007, Nr. 6, S. 1224–1231, sowie Viktoria Brüggemann, Mark-Jan Bludau und Marian Dörk: The Fold. Rethinking Interactivity in Data Visualization, in: Digital Humanities Quarterly, 14/3, 2020, <http://www.digitalhumanities.org/dhq/vol/14/3/000487/000487.html>.

Netzwerks. Auch im Labeled Property Graph-Modell besteht das Netzwerk aus Knoten und Kanten. Im Gegensatz zu RDF können aber sowohl Knoten als auch Kanten weitere Eigenschaften in Form von properties haben.⁸ Die oben modellierte Aussage könnte so auch als ein Personenknoten mit der Eigenschaft »Johann Wolfgang von Goethe« als Name modelliert werden. Während bei RDF jeder Informationsbaustein also ein eigener Knoten im Netzwerk ist, lassen sich mit Labeled Property Graphen kompaktere Netzwerke realisieren, indem manche Informationen den Knoten und Kanten direkt als Eigenschaften zugeordnet werden. Für den Prototypen des *Virtuellen Sammlungsraums* wurde die Graphdatenbank neo4j⁹ als Vertreter für das Labeled Property Graph-Modell verwendet. Ein auf RDF basierender Ansatz wurde als GraphDB¹⁰ von ontotext umgesetzt. Neben den bereits genannten Unterschieden hinsichtlich der Möglichkeiten der Datenmodellierung ist noch anzumerken, dass Abfragen bei neo4j mittels der Abfragesprache cypher geschehen, während GraphDB wie andere Triplestores auch über einen SPARQL-Endpunkt verfügt. SPARQL ist ein weit verbreiteter Standard, über den sich auch zahlreiche andere Dienste wie beispielsweise Wikidata¹¹ abfragen lassen.

Neben der Sicherung und Präsentation von Daten mit Hilfe von Technologien, die auf der Idee des Netzwerks bzw. Graphen aufbauen, wurden im Rahmen der Zusammenarbeit mit den Fallstudien des Forschungsverbunds MWW verschiedene Workflows getestet, welche die Ebenen Erschließung, Speicher und Darstellung möglichst eng und komfortabel miteinander verbinden sollen. Für die Erfassung der Daten wurden die sogenannten Dynamischen Datenlisten des *Virtuellen Forschungsraums* (VFR)¹² des MWW genutzt. Der VFR basiert wiederum auf der Portalsoftware Liferay.¹³ Mit den Dynamischen Datenlisten lassen sich über ein einfaches Drag 'n Drop-Verfahren sehr schnell flexible Eingabelisten zusammenstellen. Leider war es im Rahmen der Prototyp-Entwicklung nicht möglich, eine Schnittstelle zu einem Graphdatenbank-Backend und damit eine Verbindung zwischen einer einfachen Eingabemaske und einem dahinterliegenden Mapping auf

8 Für Verwirrung kann an dieser Stelle mitunter sorgen, dass die Kanten bzw. die Prädikate der RDF Triple-Aussagen auch als properties bezeichnet werden können. Indessen sind sie nicht mit den properties im Labeled Property Graphen zu verwechseln.

9 <https://neo4j.com/product/neo4j-graph-database/>.

10 <https://graphdb.ontotext.com/>.

11 <https://query.wikidata.org/>.

12 <https://vfr.mww-forschung.de/>.

13 <https://www.liferay.com/de/>.

eine Knoten- und Kanten-Systematik zu schaffen, wie es beispielsweise Softwarelösungen wie WissKI¹⁴ bereits anbieten. Für den Prototypen wurden daher die Daten aus dem VFR exportiert und unter Verwendung von OpenRefine¹⁵ entsprechend für den Import in die Graphdatenbanken transformiert. Für die Zukunft wäre eine direkte Anbindung der Datenbanken sowie Schnittstellen in die Verbundinstitutionen des MWW zur Abfrage bereits bestehender Daten der erstrebenswerte Weg.

Der Werkkomplex »Goethes *Farbenlehre*«: Zur Auswahl der Daten

Einen wichtigen Aspekt in der Erschließung stellten in diesem Zusammenhang die Relationen zwischen den Objekten dar, welche in den Quelldatenbanken von Museen der Klassik Stiftung Weimar, der Herzogin Anna Amalia Bibliothek und des Goethe- und Schiller-Archivs so bisher nicht erfasst waren. Dies hat letztlich sammlungsgeschichtliche Gründe. Zwar bilden die Sammlungen und Handschriften Goethes historisch den Grundstein der Institution, die heute Klassik Stiftung Weimar heißt. Allerdings verfügte Walther von Goethe, der letzte noch lebende Enkel des Dichters, in seinem 1883 abgefassten Testament eine Aufteilung des Erbes. Das Großherzogtum Sachsen-Weimar-Eisenach erhielt das Haus, die natur- und kunstgeschichtlichen Sammlungen sowie die persönliche Bibliothek des Dichters. Diese bilden den Grundstock der heutigen Museen der Stiftung, natürlich insbesondere des Goethe-Nationalmuseums am Weimarer Frauenplan. Großherzogin Sophie von Sachsen-Weimar-Eisenach erbte hingegen Goethes Handschriften,¹⁶ den Grundstock des heutigen Goethe- und Schiller-Archivs.

¹⁴ <https://wiss-ki.eu/de>.

¹⁵ <https://openrefine.org/>.

¹⁶ Vgl. Walther von Goethe: Testament, Weimar, 24. September 1883, in: Das Goethe-Nationalmuseum in Weimar. Dokumente, Bd. 1, hg. von Paul Kahl und Hendrik Kalvelage, Göttingen 2015, S. 694–696. Dem war ein jahrzehntelanger Konflikt zwischen Goethes Familie und verschiedenen öffentlichen Akteuren vorangegangen. Grundsätzlich ging es darum, Goethes Haus und seine verschiedenen Sammlungen von der Familie zu erwerben, um idealerweise eine Gedenkstätte zu errichten. Die Familie war allenfalls zum Verkauf der Sammlungen bereit, zumindest zeitweise. Die ambitionierteste Initiative des Deutschen Bundes in den Jahren 1842 bis 1845 scheiterte, danach blockte die Familie prinzipiell alle weiteren Vorstöße ab. Vgl. ausführlich Paul Kahl: Die Erfindung des Dichterhauses. Das Goethe-Nationalmuseum in Weimar. Eine Kulturgeschichte, Göttingen 2015, S. 49–122.

Schon damals wurden die originalen Sammlungszusammenhänge auseinandergerissen, insofern Handschriften von ihren Ursprungsorten entfernt und in das neue Handschriftenarchiv verbracht wurden. Später wurden weitere Institutionen Teil des Verbundes, unter anderem die ehemalige herzogliche Bibliothek, deren Oberaufsicht Goethe von 1797 bis zu seinem Tod bekleidet hatte und aus der er als Nutzer über 3.000 Bände entlieh. Nach 1919 bestand die Institution als Thüringische Landesbibliothek weiter. 1969 wurde sie als Zentralbibliothek der Deutschen Klassik den Gedenkstätten angegliedert und firmiert heute als Herzogin Anna Amalia Bibliothek (HAAB).

Museen, Archiv und Bibliothek entwickelten sich zu Säulen innerhalb der Weimarer Gedenkstätten, die unter einem gemeinsamen Dach eine gewisse Eigenständigkeit bewahren; heute sind sie separate Direktionen der Klassik Stiftung Weimar. Diese Säulen führten daher jeweils eigene Katalogsysteme in Übereinstimmung mit den jeweiligen disziplinären Gepflogenheiten, die in der Einleitung genannten Daten-Silos. Die ursprünglichen Zusammenhänge der Objekte, auch innerhalb desselben Bestandes, wurden dabei nicht oder nur fragmentarisch festgehalten. Die Einführung elektronischer Kataloge in den 1990er Jahren löste dieses Problem nicht, im Gegenteil: Museen und Archiv setzten miteinander inkompatible Oracle-Datenbanken auf, wobei die des Archivs grundsätzlich öffentlich zugänglich, die der Museen dem internen Gebrauch vorbehalten war. Die HAAB wurde dagegen Teil des Gemeinsamen Bibliotheksverbunds (GBV), der Universitäts- und Forschungsbibliotheken aus sieben Bundesländern vereinigt. Deren Katalog, der seit 2019 mit dem des Südwestdeutschen Bibliotheksverbundes (SWB) zum gemeinsamen Katalog *kioplus*¹⁷ zusammengeführt ist, operiert mit der Software PICA und folgt verbindlichen bibliothekarischen Regelwerken, derzeit vor allem dem Standard Resource Description and Access (RDA).¹⁸ Das Ergebnis sind auf den ersten Blick inkompatible Kataloge, die jedoch durch die gemeinsame Nutzung und die gemeinsame Erarbeitung von Normdaten potentiell auf Verknüpfung angelegt sind, derzeit umgesetzt in Pilotprojekten. Im Fall der Provenienz von einzelnen Objekten können Informationen in den Katalogen enthalten sein. Dies ist jedoch von den zeitlichen Ressourcen und dem Vorwissen der Bearbeitenden abhängig, zudem wurden Standards zur Bearbeitung von Provenienzen wie das an der HAAB mitentwickelte T-PRO zur Verzeichnung von Provenienzmerkmalen im bibliothekarischen Bereich¹⁹ erst in den letzten Jahrzehnten eingeführt, so dass früher angelegte

17 <https://www.bszgbv.de/services/kioplus/>.

18 <https://www.basiswissen-rda.de/>.

19 https://provenienz.gbv.de/T-PRO_Thesaurus_der_Provenienzbegriffe.

Datensätze allenfalls unter hohem zeitlichen Aufwand aktualisiert werden können.²⁰

Den ursprünglichen Zusammenhängen der Objekte untereinander wird dagegen in der Praxis nur ausnahmsweise Beachtung geschenkt; wenn, dann aber nur innerhalb des jeweils eigenen Katalogs, weil bislang schlicht kein übergreifendes System existiert. Das gilt sowohl für gleichartige (also zum Beispiel Manuskripte mit Manuskripten) als auch für heterogene Objekte (zum Beispiel Manuskripte mit Büchern, Graphiken oder Instrumenten). Ein virtueller Sammlungsraum nach unseren Vorstellungen muss also diese verschiedenen Objekte nicht nur auf einen einheitlichen Datenstandard bringen, sondern auch die Beziehungen der Objekte untereinander rekonstruieren und dokumentieren. Dafür bot es sich an, Objekte miteinander zu verknüpfen, die sich heute auf die verschiedenen Direktionen der Stiftung und deren Katalogsysteme verteilen. Paradigmatisch dafür ist der schon erwähnte Nachlass Goethes.²¹ Goethes Wohnhaus samt Interieur, Kunst- und Naturaliensammlungen gehört zu den Museen der Klassik Stiftung; Goethes

20 Tatsächlich unternimmt die Klassik Stiftung Weimar bereits Anstrengungen zur Erweiterung und Verknüpfung ihrer Kataloge. Zum einen setzt die Herzogin Anna Amalia Bibliothek ein eigenes Discovery-System für ein integriertes Informationsportal auf, das bisher in einer Beta-Version verfügbar ist (<https://portal.haab.klassik-stiftung.de/>). Dieser ausbaufähige Datenraum der Beta-Version umfasst bislang den normdatenbasierten Bestandskatalog, die Weimarer Bibliographien, Aufsatztitel aus den Online Contents, die Einbindung der Digitalen Sammlungen der HAAB sowie der Google-Buchsuche und normdatenbasierte Sucheinstiege in Sammlungen mit offenen Schnittstellen zu möglichen Netzwerken. Neben den Provenienzen ist es die weitere Arbeit an innovativen Normdaten für Sammlungen und Werktitel, die spartenübergreifende Verknüpfungen und Netzwerke, aber auch das Einspielen weiterer Datenräume ermöglicht. Auf der anderen Seite existiert seit 2021 die Forschungsdatenbank so:fie ([https://ores.klassik-stiftung.de/ords/f?p=900:1:1:::~:](https://ores.klassik-stiftung.de/ords/f?p=900:1:1:::)). Hier werden, verknüpft über die Personennormdatei der GND, Datensätze aus den Katalogsystemen des Archivs und der Museen zusammengeführt, so dass alle Bestände, je nach Grad der Erschließung auch Einzelobjekte, in diesen Beständen angezeigt werden, die jeweils mit dieser Person verknüpft sind. Obwohl dies einen großen Schritt bei der virtuellen Zusammenführung der Stiftungsbestände bedeutet, gibt es auch hier keine Informationen zur Verknüpfung der Objekte untereinander, und auch der Bibliothekskatalog der HAAB ist noch nicht angebunden. Die sammlungsbezogene Verknüpfung der Datenräume aus jeweils unterschiedlichen Perspektiven ist das Ziel der kommenden Arbeit.

21 Ebenso würde sich zum Beispiel der Nachlass Friedrich Nietzsches anbieten. Auch hier sind Handschriften, materielle Objekte und die Bibliothek des Philosophen auf die Museen, das Goethe- und Schiller-Archiv sowie die Herzogin Anna Amalia Bibliothek verteilt.

schriftlicher Nachlass, soweit noch im Besitz der Enkel vorhanden, liegt im Goethe- und Schiller-Archiv; bei Goethes Privatbibliothek mit ihren etwas über 7.000 Bänden handelt es sich zwar um Besitz der Museen, sie wurde aber in den letzten Jahren im Rahmen von MWW-Forschungsprojekten in der HAAB erforscht und dort im elektronischen Verzeichnis *Goethe Bibliothek Online* gemeinsam mit den Ausleihen des Dichters aus der herzoglichen Bibliothek in Weimar erschlossen und verzeichnet.²²

Tatsächlich handelt es sich bei Goethes Nachlass um eine riesige, vermutlich nicht einmal genau bezifferbare Zahl von Objekten. Allein die naturwissenschaftlichen Sammlungen Goethes umfassen nach einer neueren Zählung über 23.000 Datensätze.²³ Aus dieser Masse galt es also einen überschaubaren Bereich auszuwählen. Die Wahl fiel auf den Werkkomplex *Farbenlehre*, ein naturwissenschaftliches Arbeitsgebiet, auf dem Goethe über Jahrzehnte hinweg aktiv war, von ersten Experimenten um das Jahr 1790 über die zweibändige Monographie *Zur Farbenlehre* (1806/10) bis ins hohe Alter hinein. Von einem Werkkomplex *Farbenlehre* sprechen wir, um deutlich zu machen, dass es nicht nur um die bei Johann Friedrich Cotta verlegte zweibändige Monographie *Zur Farbenlehre*²⁴ gehen sollte, sondern um den gesamten thematischen Bereich, namentlich die wichtige Vorarbeit *Beyträge zur Optik*, die Goethe 1791 und 1792 in zwei kleineren Bänden veröffentlichte.²⁵ Ein Werkkomplex ist eine Aggregation mehrerer Einzelwerke, die für sich genommen eigenständig auftreten, aber thematisch oder vom Plot her eng aufeinander bezogen sind.²⁶ Goethes Werk ist reich an solchen Werkkom-

22 <https://opac.lbs-weimar.gbv.de/DB=2.5/>. Zu diesem Verzeichnis vgl. auch Stefan Höppner und Ulrike Trenkmann: *Goethe Bibliothek Online. Ein digitaler Katalog*, in: *Goethe-Jahrbuch* 134, 2017, S. 237–252. Ausführlich beschrieben ist Goethes Privatbibliothek in: Stefan Höppner: *Goethes Bibliothek. Eine Sammlung und ihre Geschichte*, Frankfurt am Main 2022.

23 Vgl. Diana Stört: *Goethes Sammlungsschränke. Wissen nach Maß*, Dresden 2020, S. 40.

24 Vgl. Johann Wolfgang Goethe: *Zur Farbenlehre*, 2 Bde. + 1 Tafelbd., Stuttgart/Tübingen 1806–1810.

25 Vgl. Johann Wolfgang Goethe: *Beyträge zur Optik. Erstes Stück. Mit XXVII Tafeln*, Weimar 1791; ders.: *Beyträge zur Optik. Zweytes Stück. Mit einer großen colorirten Tafel und einem Kupfer*, Weimar 1792.

26 Svetlana Efimova spricht vom Werkkomplex als »Zwischenstufe bzw. ein Bindeglied zwischen ›Einzelwerk‹ und ›Gesamtwerk‹«. Dieses ist »eine abgegrenzte Summe mehrerer Opera eines Autors und zugleich [...] ein ›Beziehungssystem‹ zwischen diesen Opera.« Svetlana Efimova: *Das Werk als Entgrenzung. Werkkomplex und Dynamik des Gesamtwerks*, in: *Zeitschrift für Germanistik* NF 31.1, 2021, S. 139–154, hier S. 140 bzw. 145.

plexen. Weitere Beispiele sind die beiden Teile des *Faust* (1808/33), die beiden *Wilhelm Meister*-Romane (1795/96 bzw. 1821/29) und die vielbändige Reihe *Aus meinem Leben*, die neben *Dichtung und Wahrheit* (4 Bde., 1811/33) auch die *Italienische Reise* (3 Bde., 1816/29) und die beiden Reiseberichte *Campagne in Frankreich* und *Belagerung von Mainz* (beide 1822) umfasst. In diesem Sinne soll auch der Werkkomplex *Farbenlehre* aufgefasst werden.

Der Inhalt und die Genese von Goethes Schriften zu diesem Thema sind nicht nur gut erforscht.²⁷ Die Sammlungen der Klassik Stiftung Weimar enthalten auch eine Vielzahl von dazu passenden und äußerst vielfältigen Objekten wie Briefe an und von Goethe, Lesenotizen des Autors, Nachweise über Buchausleihen, genutzte Bücher, Grafiken und wissenschaftliche Instrumente, insgesamt wohl eine niedrige vierstellige Zahl. Diese Kombination aus umfassender Dokumentation und Heterogenität der Objekte macht die *Farbenlehre* zu einem besonders geeigneten Beispiel für unsere Zwecke. Hier existiert eine Fülle potenziell passender Gegenstände aus den Beständen der Klassik Stiftung Weimar, so dass wiederum eine Auswahl getroffen werden musste. Die Selektion umfasste schließlich 95 Objekte aus der Museums- bzw. Archivdatenbank der KSW, der Goethe Bibliothek Online sowie aus einigen externen Einrichtungen wie zum Beispiel den Universitätsbibliotheken von Jena und Göttingen, die Goethe nachweislich für seine Forschungen nutzte, und dem Goethe-Museum Düsseldorf, in dem sich ebenfalls Objekte aus Goethes Nachlass befinden. Zum einen wurde dabei Wert auf eine möglichst große Heterogenität der Objekte gelegt. Es handelt sich um Briefe, Exzerpte und Lesenotizen, von Goethe genutzte Bücher aus verschiedenen Bibliotheken, ferner Zeichnungen, Druckgrafiken und wissenschaftliche Arbeitsinstrumente.

Zum anderen wurden vorzugsweise Objekte ausgewählt, zwischen denen ein Zusammenhang bekannt war, etwa ein Exzerpt Goethes, das sich auf ein bestimmtes Buch bezieht, aufeinander antwortende Briefe oder – wie im folgenden Beispiel – ein wissenschaftliches Instrument, das zur Vorlage

27 Goethes Arbeiten zur Farbenlehre sind zusammengefasst und kommentiert in: Johann Wolfgang Goethe: Die Schriften zur Naturwissenschaft, Abt. 1, Bd. 3–7, hg. von Rupprecht Matthaei, Weimar 1951–1958; die Entstehung der einzelnen Texte zur Farbenlehre ist dokumentiert in Momme Mommsen, Katharina Mommsen u. a.: Die Entstehung von Goethes Werken in Dokumenten, bisher 7 Bde., Berlin 1958ff. Da dieses Sammelwerk alphabetisch nach Werktiteln vorgeht und bisher erst beim Buchstaben I angelangt ist, sind noch nicht alle Einzeltexte aus dem Werkkomplex Farbenlehre erfasst. Der Entstehungsprozess von *Zur Farbenlehre* ist jedoch bereits akribisch dokumentiert in Bd. 4, S. 255–981; für die *Beyträge zur Optik* vgl. Bd. 1, S. 225–275.

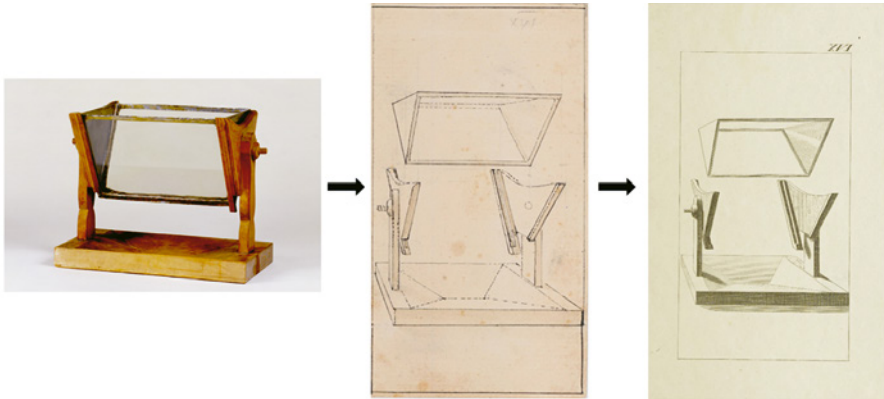


Abb. 1: Unbekannter Hersteller, Großes Wasserprisma
aus Goethes Besitz, 1791 (?)

für eine Zeichnung Goethes wird, welche wiederum als Grundlage für einen Kupferstich in den *Beyträgen zur Optik* und *Zur Farbenlehre* dient. Die Objekte verteilen sich dabei auf Goethes naturwissenschaftliche Sammlung im Goethe-Nationalmuseum, die Grafischen Sammlungen der KSW (beide von der Direktion Museen betreut und in der Museumsdatenbank erfasst) und die Digitalen Sammlungen der Herzogin Anna Amalia Bibliothek (im Bibliothekskatalog und in einer separaten Webpräsenz dokumentiert). Dieses Vorgehen stellt sicher, dass es sich um eine überschaubare Zahl von Objekten handelt, deren Beziehung untereinander im Ganzen gut dokumentiert ist.

Allerdings bringt dieses Vorgehen auch drei unumgängliche Nachteile mit sich: Zum einen setzt es bei den Bearbeiter*innen ein hohes Maß an Vorwissen zu Goethe und seiner wissenschaftlichen Arbeit voraus. Dabei gilt es gleichzeitig, die Datensätze so aufzubereiten, dass sie auch für weniger speziell ausgebildete Nutzer*innen verständlich und handhabbar bleiben, dass namentlich explorative Nutzungsweisen ermöglicht, ja gefördert werden. Zum anderen erfordert die Auswahl der Objekte und die Dokumentation ihrer Verknüpfungen ein hohes Maß an subjektiven Entscheidungen. Aufgrund des überschaubaren Rahmens für den Prototyp des *Virtuellen Sammlungsraums* kann nur eine relativ kleine Anzahl der potenziell verfügbaren Objekte ausgewählt werden. Andere Bearbeiter*innen hätten wahrscheinlich andere Objekte aus den Sammlungen der KSW bevorzugt. Entsprechend verhält es sich mit den Relationen der Objekte, die in den Datensätzen des *Virtuellen Sammlungsraums* festgehalten sind. Bestimmte Verknüpfungen

werden privilegiert, während andere dort gar nicht erst sichtbar werden. Und schließlich lassen sich unsere Datensätze aufgrund der unterschiedlichen technischen Grundlagen und Datenstandards der Quellsysteme nach jetzigem Stand nicht einfach *harvesten*, sprich in andere Systeme importieren. Stattdessen müssen sie auf Basis der dort verfügbaren Informationen, angereichert um eigene Recherchen (etwa zur Forschungsliteratur, aber auch zu den Relationen untereinander), für den *Virtuellen Sammlungsraum* eigens neu angelegt werden. Der damit verbundene Arbeitsaufwand macht es nach jetzigem Stand schwierig, mit wirklich großen Datenmengen zu arbeiten, von Big Data in einem ernst zu nehmenden Sinn ganz zu schweigen.

Zusammenhänge erfassen, speichern und visualisieren

Um die Verbindungen der Objekte untereinander zu dokumentieren, wurde insbesondere auf das LIDO-Vokabular zurückgegriffen, mit welchem sich Beziehungen wie »ist Reproduktion von«, »ist Version von« oder »ist dokumentiert in« etc. abbilden lassen.²⁸ Für die Relationstypen wurde ein kontrolliertes Vokabular verwendet, wobei darauf geachtet wurde, bestehende und bereits für LIDO dokumentierte Beziehungsarten nachzunutzen. Darüber hinaus wurden Normdaten für Personen, Institutionen, Orte, Werke und Sammlungen eingesetzt.

Die Unterschiede in der Art und Weise wie Labeled Property Graph-Datenbanken und auf RDF basierende Triplestores Daten abbilden, legten es auch nahe, die erfassten Daten entsprechend leicht unterschiedlich zu modellieren und somit verschiedene Ansätze auszuprobieren. Das Datenmodell für die Integration in neo4j kennt sieben verschiedene Label für Knoten: Objekte, Personen, Institutionen/Körperschaften, Schlagworte, Orte, Sammlungen und Werke. Für den Import der Daten in den Triplestore von GraphDB wurden die Daten entsprechend den classes und properties von CIDOC-CRM modelliert.²⁹ Die wichtigsten Unterschiede für beide Modelle ergeben sich dabei im Hinblick auf die für CIDOC-CRM wichtige Klasse der Events. Für CIDOC-CRM wird jedes Objekt im Laufe seines Lebenszyklus durch verschiedene Events geprägt. Johann Wolfgang von Goethe ist so mit einem von ihm verfassten Brief streng genommen durch zwei Events verbunden, in denen er die Rollen des Verfassers und des Absenders einnimmt. Das

28 <https://lido-schema.org/schema/latest/lido.html#relatedWorkRelType>, s.a. <http://terminology.lido-schema.org/relatedWorkRelType>.

29 https://cidoc-crm.org/sites/default/files/cidoc_crm_version_7.1.2.pdf.

Event stellt somit einen zusätzlichen Knoten im Graphenmodell dar, welcher in neo4j nicht mit integriert wurde. Hier wurde die Tatsache ausgenutzt, dass sowohl Knoten als auch Kanten Eigenschaften (properties im Sinne von Labeled Property Graphen) besitzen können. Brief und Verfasser sind direkt über eine Relation verbunden, welche die property »role«/»Rolle« mit dem Wert »Verfasser« oder »Absender« besitzt. Ähnlich wurde mit den Relationen zwischen zwei Objekten verfahren. In neo4j ist diese Verbindung direkt hergestellt und die Art der Verbindung als »type«-Eigenschaft umgesetzt. Im Triplestore werden zwei Objekte über ein Event vom Typ `Relation_Assignment` miteinander verbunden. Dabei wird einem Objekt die Beziehung zu einem anderen Objekt zugeordnet. Die Art der Beziehung wird dann als Typ des Events erfasst (siehe Abb. 2 für einen Vergleich der beiden Modellierungsmöglichkeiten). Ein Vorteil in der Verwendung von neo4j als Graphdatenbank ist die Möglichkeit, mit dem neo4j-Browser bereits ein mögliches Interface zur Hand zu haben, mit welchem Nutzer*innen Daten aus der Datenbank abfragen können und das Ergebnis in Form eines visualisierten Graphen erhalten. Voraussetzung hierfür ist mindestens eine grundlegende Vertrautheit mit der Abfragesprache cypher.

Für neo4j wurde GraphQL als technologische Grundlage für die Schnittstelle (API) gewählt. GraphQL erfreut sich einer seit Jahren wachsenden Popularität und ermöglicht es, APIs individuell auf die eigenen Datenbedürfnisse hin zu implementieren. Für neo4j existiert eine eigene GraphQL-Library, d.h. bestimmte vordefinierte Funktionen, welche eine Integration von neo4j als Backend und GraphQL als API erleichtern.³⁰ Ein Vorteil von GraphQL ist die Möglichkeit, Abfragen netzwerkartig in Form von hierarchisch miteinander verschachtelten Feldern anzulegen. Zudem sind leicht zu bedienende User Interfaces verfügbar, welche nicht nur die Implementierung der Schnittstelle erleichtern, sondern ebenso Nutzer*innen zugänglich gemacht werden können, um einen vereinfachten programmatischen Einstieg in die Netzwerkdaten zu ermöglichen.³¹ Für den Triplestore GraphDB ist eine auf SPARQL basierende Schnittstelle das Mittel der Wahl. SPARQL ist der etablierte Standard zur Abfrage von Daten im RDF Format. SPARQL bietet ebenfalls eine Vielzahl an Funktionen, um Daten speziell auf die eigenen Forschungsbedürfnisse hin abzufragen und kann zudem andere Schnittstellen als Services mit einbeziehen. So ließe sich eine gleichzeitige Abfrage

³⁰ Eine technische Einführung in die Implementierung von GraphQL mit neo4j bietet das Buch von Will Lyon: *Full Stack GraphQL Applications. With React, Node.js, and Neo4j*, Shelter Island, NY, 2022.

³¹ Vgl. hierzu zum Beispiel <https://github.com/graphql/graphiql>.

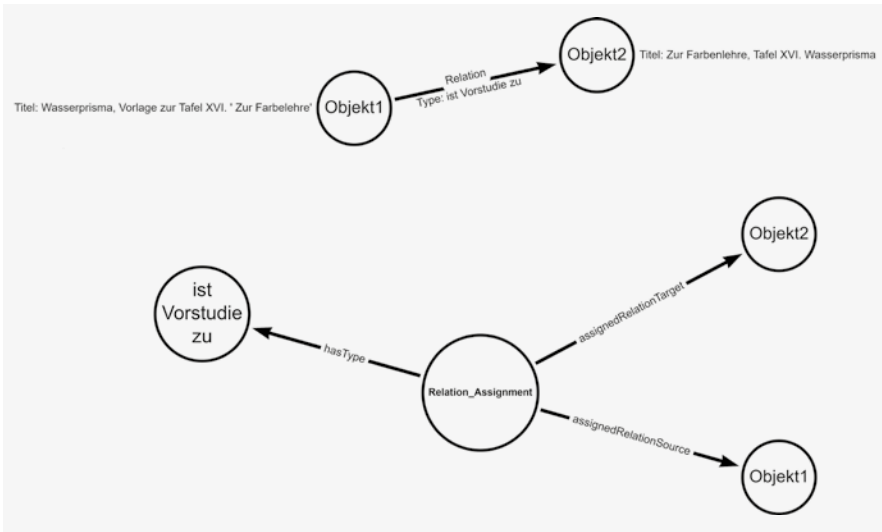


Abb. 2: Gegenüberstellung verschiedener Modellierungsmöglichkeiten von Objekt-Objekt-Relationen. Oben: Mit dem Labeled-Property-Graph Modell lassen sich Beziehungen sehr kompakt erfassen. Die Relationsarten können als properties direkt mit der Relation verknüpft werden. Unten: Ein an RDF angelehntes Modell. Hier wird die Zuschreibung der Relation als zusätzlicher Event-Knoten vom Typ »Relation_Assignment« eingeführt. Die Art der Relation ist dann wiederum ein eigener Knoten. Die höhere Komplexität dieses Modells ermöglicht zugleich eine differenziertere Erfassung, indem das Event mit einer Person verknüpft werden kann (wer hat diese Beziehung festgelegt) sowie weitere unterstützende Angaben (festgelegt auf Basis von Zitat xy) als eigene Knoten angebunden werden können.

der Goethe-Daten und beispielsweise von Wikidata als möglichem Anknüpfungspunkt realisieren. Auch ein User Interface (vergleichbar mit dem Query Interface von Wikidata) kann für SPARQL problemlos umgesetzt werden und somit Nutzer*innen neben dem programmatischen auch einen vereinfachten Zugang zu den Daten ermöglichen. Jedoch ist die Lernkurve bei SPARQL steiler als beispielsweise bei der Nutzung eines Interfaces für GraphQL. Weiterhin ist zu bemerken, dass die GraphQL-API besonders von dem einfacheren Graphenmodell und der Nutzung von properties als Eigenschaften für Knoten und Kanten profitiert, während hingegen SPARQL auch mit komplex und sehr granular modellierten Daten gut umgehen kann.

Im Frontend basieren beide Prototypen auf React und D3js. Letzteres ist technisch der De-facto-Standard für die Umsetzung komplexer interaktiver Visualisierungen in der Webentwicklung. An dieser Stelle treten besonders

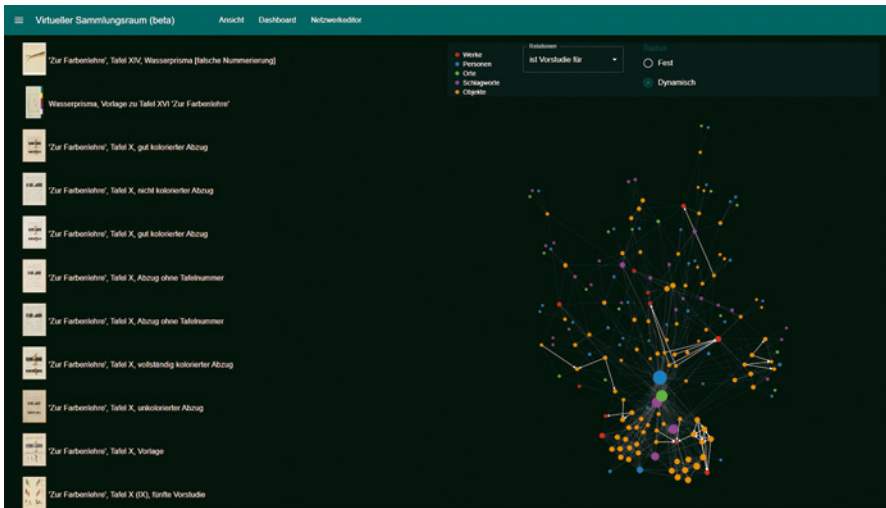


Abb. 3: Ansicht des User Interfaces des VSR-Prototypen

die beiden Leitgedanken des *Virtuellen Sammlungsraums* in den Vordergrund und das Ziel, Katalogdarstellungen mit visuellen Hilfsmitteln zu ergänzen. Im Kontext der MWW-Fallstudie »Goethe Digital« sollte zudem der Frage nachgegangen werden, inwiefern eine solche Darstellung neue Perspektiven auf den Werkkomplex *Farbenlehre* bzw. andere, möglicherweise in der Zukunft integrierte Datenbestände ermöglichen kann.

Im linken Bereich der Ansicht (Abb. 3) befindet sich die Katalogansicht, welche die erfassten Objekte als Ergebnisliste anzeigt. Rechts daneben sind die Daten als Netzwerk visualisiert. Den verschiedenen Knotentypen (Personen, Ort, Objekte etc.) werden verschiedene Farben zugeordnet. Mit Hilfe des Kontrollpanels über der Visualisierung lässt sich die Ansicht steuern. Im gezeigten Beispiel ist der Radius der Knoten entsprechend der anliegenden Relationen skaliert. Weiterhin sind nur Relationen vom Typ »ist Vorlage für« angezeigt. So lassen sich schnell Knoten entsprechend ihrer relativen Bedeutung im Netzwerk sowie bestimmte Konstellationen von Werk- und Entstehungszusammenhängen abbilden. Über ein Kontroll-Panel über dem Netzwerk stehen verschiedene Funktionen zur Verfügung, um bestimmte Aspekte des Netzwerks zu manipulieren (Kanten nach Relationstypen ein- und ausblenden, Radius der Knoten entsprechend der anliegenden Kanten, Knoten nach Kategorien ein- und ausblenden etc.). Ist ein Knoten angeklickt, wird die Ansicht auf diesen fixiert und man hat die Möglichkeit, ihn auszu-

blenden oder im Falle eines Objekts in die Detailansicht zu wechseln. Eine ausklappbare Seitenleiste am linken Rand verbirgt Filterfunktionen. Damit können sowohl die Ergebnisliste als auch die Netzwerkansicht nach bestimmten Schlagworten, Orten, Personen, Werknamen etc. gefiltert werden.

Somit stehen Nutzer*innen verschiedene Möglichkeiten zur Verfügung, um sich mit den Daten bzw. dem Werkkomplex Farbenlehre auseinanderzusetzen. Es ist möglich, einfach nur die Inhalte der Sammlung zu durchsuchen; oder aber mit Hilfe des Netzwerks zu untersuchen, wo sich Cluster befinden und welche Objekte eher peripher stehen. Damit stehen sowohl Funktionen zur schnellen Informationsrecherche sowie für eingehendere und explorative Beschäftigung mit den unterliegenden Daten zur Verfügung.

Wird ein Eintrag der Ergebnisliste angeklickt, erfolgt eine Weiterleitung in die verlinkte Detailansicht. Neben allen wichtigen Daten zum Objekt lassen sich hier auch verbundene weitere Objekte mit den spezifischen Objektrelationen in Netzwerkform anzeigen. Auch hier wird vom Paradigma der linearen Liste abgewichen, damit Objekte in ihrem semantischen Beziehungskontext gezeigt werden können. Das sei hier an den Objekten aus Abb. 1 verdeutlicht. Klickt man auf das Große Wasserprisma, wird auch die Relation zu Goethes Handzeichnung angezeigt. Nach jetzigem Stand treten hierbei jeweils nur die unmittelbaren Relationen zu anderen Objekten mit hervor, nicht aber die mittelbaren. Um die Relation zwischen Handzeichnung und Kupfertafel anzuzeigen, muss man also eigens eines dieser beiden Objekte anklicken. Die Optionen unserer Visualisierung sind aber noch stark ausbaufähig. Für die prospektive Entwicklung des *Virtuellen Sammlungsraums* wird neben der Integration zusätzlicher Sammlungen und Datenbestände die Verbindung der Katalogansicht mit weiteren Visualisierungstypen und der Ausbau der Interaktionsmöglichkeiten mit den Visualisierungen angestrebt.

Ein wichtiger Aspekt der angestrebten Entwicklung ist es aber auch, Nutzer*innen anzuregen, darüber zu reflektieren, was sie in dem Moment rezipieren, wenn sie mit den Daten, Katalog- und Netzwerkansichten interagieren. Die vorgestellten Bauteile sind für viele in Zeiten von Google und allgegenwärtigen Visualisierungsformen von Heatmaps zur Darstellung der aktuellen Covid-Inzidenzen hin zu Linecharts (Entwicklung des Gaspreises) so normal und scheinbar vertraut, dass die Frage, wie die präsentierten Ergebnisse zustande gekommen sind, nur noch selten gestellt wird. Doch genau diese Frage ist zum richtigen Verständnis essentiell. Oder wie Johanna Drucker es ausdrückt:

So naturalized are the Google maps and bar charts generated from spread sheets that they pass as unquestioned representations of »what is«. This is the hallmark of realist models of knowledge and needs to be subjected to a radical critique to return the humanistic tenets of constructed-ness and interpretation to the fore.³²

Entsprechend muss auch im Zusammenhang des vorgestellten Prototyps die Frage gestellt werden, was hier dargestellt wird. Ist es (frei nach Schopenhauer) der Werkkomplex *Farbenlehre* an sich? Leider nicht, denn jede Erschließung von Kulturgut und ebenso jede Form der visuellen Darstellung basiert auf ganz individuellen Entscheidungen bezüglich der Auswahl der Daten und der Erfassung bestimmter Relationen – und vielleicht auch des Weglassens anderer –, basiert auf Datenmapping, Design und der technischen Umsetzung. Was zu sehen ist, ist also nicht so sehr der Werkkomplex *Farbenlehre*, sondern eine auf vielen subjektiven Entscheidungen unsererseits basierende Perspektive auf Goethes Beschäftigung mit der *Farbenlehre*.³³ Die Sammlungen – und als solche kann man die Auswahl der Objekte zum Thema Goethes *Farbenlehre* durchaus betrachten – sind immer Produkte sozialer Interaktionen und sozialer Netzwerke.³⁴ Die Wahl der Objekte setzt zwar eine gewisse Kennerschaft unsererseits voraus, und zwar was Goethes Beschäftigung mit dem Thema und den vorhandenen Materialien angeht. Andere Forscher*innen würden aber vermutlich andere Objekte und Relationen privilegieren. Die im *Virtuellen Sammlungsraum* dokumentierten Objektrelationen untereinander sind den Objekten zudem nicht inhärent, sondern

32 Johanna Drucker: *Humanities Approaches to Graphical Display*, in: *Digital Humanities Quarterly*, 5/1, 2011, <http://www.digitalhumanities.org/dhq/vol/5/1/0091/000091.html#>.

33 Das gilt letztlich sogar für den Werkkomplex als solchen, denn auch die Entscheidung, welche Texte dazu zu zählen sind, ist im Detail individuell zu treffen. Für die Zwecke unseres Aufsatzes verstehen wir darunter die in Bd. 3 bis 7 der *Schriften zur Naturwissenschaft* (vgl. Anm. 27) aufgenommenen Texte.

34 In diesem Fall sind letztlich mindestens drei Ebenen von Sammlung zu unterscheiden: Goethe, der seinerseits Materialien zur *Farbenlehre* anhäuft und den dazugehörigen Handschriften eine eigene Abteilung im Repertorium einräumt, die er 1822 von seinem Bibliothekar und Sekretär Friedrich Theodor Kräuter anfertigen lässt; die Weimarer Institutionen, die diese Materialien aufbewahren und ihnen durch den Akt der (nach Direktion getrennten) Katalogisierung eine eigene Ordnung verleihen; und schließlich unser Projekt »Goethe Digital«, das aus diesen Katalogen bestimmte Objekte und ihre Relationen untereinander auswählt und privilegiert. Zum von Kräuter angelegten Repertorium vgl. Eva Beck: *Das erste Findbuch des Archivs – Kräuters »Repertorium über die Goethesche Repositur«*, in: *Manuskripte* 5, 2012, S. 47–62.

ein Produkt der Forschung, namentlich von eigens für den *Virtuellen Sammlungsraum* angestellten Recherchen. Ein mögliches Szenario dieser Interaktion verschiedener Akteure, Wissenschaftler*innen, Entwickler*innen und ebenso Nutzer*innen ist in dem Prototypen des *Virtuellen Sammlungsraums* realisiert – aber eben nur *ein* mögliches. Das »Wissen der Sammlungen« ist letztlich immer (auch) das Wissen und Wollen der Sammelnden, die die entsprechenden Datensätze anlegen. Nur dieses Wissen kann auch abgebildet werden. Das schließt aber nicht aus, dass der *Virtuelle Sammlungsraum* und vor allem die Netzwerkvisualisierung dabei Muster sichtbar machen, die den Bearbeiter*innen zuvor nicht bewusst waren.

Der *Virtuelle Sammlungsraum* – eine vorläufige Bilanz

Unser Vorhaben war ein interdisziplinäres: das Zusammenwirken einer vom *material turn* tingierten, aber letztlich traditionell ausgerichteten Goethe-Philologie und innovativer Methoden der Digital Humanities. Wie sieht nun der Ertrag aus diesen beiden Perspektiven und methodischen Zugängen aus?

Aus Sicht der Goethe-Forschung ist der Gewinn beträchtlich, wenn man den *Virtuellen Sammlungsraum* als Prototypen auffasst. Die Silo-Struktur der vorgängigen Quellkatalogsysteme wird für die bearbeiteten Objekte aufgehoben, übergreifende Zusammenhänge und Strukturen werden sichtbar und vor allem für eine explorative Nutzung angeboten. Damit eröffnen sich neue Perspektiven auf ein bestimmtes Gebiet im Schaffen Goethes, nämlich auf den Werkkomplex *Farbenlehre*. Durch das Prinzip der Interaktivität werden zudem die eingangs genannten Potenziale des *Virtuellen Sammlungsraums* als *Sammlungserschließender Katalog* eingelöst. Es handelt sich tatsächlich um ein exploratives Recherche- und Wissensmodell, das für die Nutzer*innen Entstehungs- und Überlieferungszusammenhänge sichtbar macht – wenn auch auf Basis des Wissens des- oder derjenigen, die den *Virtuellen Sammlungsraum* und die in ihm vorhandenen Datensätze anfertigen.

Statt der abgeschlossenen Texte Goethes werden so das Prozesshafte des Schreibens und des Herstellens der visuellen Materialien sichtbar. Ein wesentlicher Teil der Relationen ist nämlich temporaler Natur, zum Beispiel »ist Vorlage zu«, »basiert auf« etc., auch wenn diese zeitlichen Relationen auf die Simultaneität einer Netzwerkdarstellung abgebildet werden. Der Autor Goethe wird als Teil eines Netzwerks von Papieren, Instrumenten und anderen Materialien deutlich, wobei zu überlegen ist, ob man dabei nur den beteiligten Personen oder auch unbelebten Objekten eine Form von *Agency* zuschreiben will; die Konsequenzen aus dieser Erkenntnis wären mit Hilfe

der maßgeblich von Bruno Latour entwickelten Akteur-Netzwerk-Theorie zu vertiefen.³⁵ Indem er Objekte und ihre Beziehung untereinander sichtbar macht, lässt sich der *Virtuelle Sammlungsraum* hervorragend didaktisch im Literaturunterricht einsetzen. Da das Verständnis des Werkkomplexes Farbenlehre sehr voraussetzungsreich ist, eignet sich unser spezieller Prototyp vor allem für den universitären Bereich. Allerdings lässt sich der *Virtuelle Sammlungsraum* bei entsprechenden Inhalten auch im Schulunterricht oder bei der Vermittlung kulturwissenschaftlicher Sammlungen an ein allgemeines Publikum einsetzen.

Eine Schwierigkeit stellt dabei noch die Nutzung des Interfaces in seiner derzeitigen Form dar. Ein bloßes Nebeneinander von Visualisierung und Katalog lässt Nutzer*innen oft initial und intuitiv mit der Frage nach dem Mehrwert der Netzwerkdarstellung zurück. Eine gewisse Überforderung, welche im Umgang mit der begrenzten Auswahl an Objekten aus dem Farbenlehre-Kontext in einigen Fällen schon zurückgemeldet wurde, wäre im Falle einer noch größeren Anzahl an Knoten und Relationen sicher noch häufiger zu beobachten. Hinsichtlich des damit verbundenen Aspekts der Vermittlung wären thematische oder auf eine bestimmte Fragestellung hin konzipierte Einstiege in die Visualisierung denkbar, die zum einen vorerst nur eine begrenzte Anzahl an Objekten zeigen und zum anderen den möglichen Erkenntnisgewinn verdeutlichen, welcher durch die Nutzung der Visualisierung entstehen kann.

Skeptisch ist der *Virtuelle Sammlungsraum* zu beurteilen, wenn es darum geht, ob er an die Stelle traditioneller Katalogsysteme treten kann. Zumindest für die Datenhaltung der Klassik Stiftung Weimar ist das beim aktuellen Stand der Dinge mit einem vorsichtigen »Nein« zu beantworten. Nach jetzigem Stand ist die Erstellung und Bearbeitung der einzelnen Datensätze aufwändig, da sich die Quelldaten nicht oder nur unter großen Schwierigkeiten aus den miteinander inkompatiblen Quellsystemen importieren lassen; wir mussten stattdessen jeden einzelnen Datensatz für unseren Prototyp nachbauen. Die Zahl der Objekte und ihrer Relationen untereinander ist dabei momentan gering und voraussetzungsreich; Auswahl und Verknüpfung beruhen auf individuellen Entscheidungen der Mitarbeiter*innen. Die »Sammlung« des Sammlungserschließenden Katalogs ist also bislang nur so groß wie die Menge der dafür eigens erzeugten Datensätze. Er eignet sich bisher also eher zur detaillierten Darstellung und Exploration kleinerer Sammlungen

35 Zur Einführung vgl. Bruno Latour: Eine neue Soziologie für eine neue Gesellschaft. Einführung in die Akteur-Netzwerk-Theorie, übers. von Gustav Roßler, Frankfurt am Main 2010.

denn als Ersatz großer Katalogsysteme. Gelänge es, einen verlässlichen, automatisierten Datenimport aus den Quellsystemen durchzuführen, wobei die heterogenen Quelldaten auf einen gemeinsamen Datenstandard zu bringen wären, wäre das Potenzial natürlich erheblich größer. Aber auch dann wären die Relationen zwischen den Objekten händisch nachzutragen, weil sie in den vorliegenden Quellsystemen der KSW entweder gar nicht oder nur fragmentarisch dokumentiert sind. Natürlich wäre auch ein virtueller Sammlungsraum ohne die Abbildung solcher Relationen denkbar, und er wäre vermutlich einfacher und schneller zu erstellen. Damit würde ihm aber eine wesentliche Dimension fehlen, insofern die genannten und eigentlich zentralen »Entstehungs- und Überlieferungszusammenhänge« unter den Tisch fielen.

Die Relationen zwischen den Objekten mittels Künstlicher Intelligenz oder einer Öffnung des *Virtuellen Sammlungsraums* für die Datenerfassung durch Nutzer*innen auch außerhalb der aktuell beteiligten Institutionen anzureichern, wäre eine Möglichkeit, um an dieser Stelle die Dichte an verwertbaren Daten zu erhöhen. Dies ist aber in vieler Hinsicht noch weit entfernte Zukunftsmusik und bringt Schwierigkeiten ganz eigener Art mit sich. Zum einen müsste die Herkunft der so erfassten Daten transparent nachvollziehbar gemacht werden. Zum anderen wird eine gewisse redaktionelle Betreuung der Daten, ähnlich eines in anderen Fällen schon praktizierten Librarian-in-the-Loop,³⁶ weiterhin notwendig sein und folglich Personalressourcen binden, wenn ein gewisses Level an Datenqualität weiterhin gewährleistet bleiben soll. Die Beschäftigung mit dem Netzwerk zu Goethes Farbenlehre kann so betrachtet auch als Startpunkt für eine Diskussion über das damit eng verbundene soziale Netzwerk dienen, das die Produktion und Rezeption der Daten und digitalen Ansichten begleitet. Ausgehend von den bereits bestehenden Funktionen zur Interaktion und Manipulation der Netzwerkvisualisierung könnten jedoch auf diese Weise ganz neue und vielfältige Netzwerke und Sammlungen geschaffen werden. Das Ergebnis wäre eine Plattform, die ganz unterschiedliche Perspektiven aus verschiedenen Blickwinkeln und Einstiegen ermöglicht. Die Sammlung als Netzwerk würde damit in einem Netzwerk verschiedener Sammlungsperspektiven aufgehen und die vielen verschiedenen Kontexte greifbarer machen, in denen wir uns mit Sammlungen beschäftigen.

36 Lizhou Fan et al.: Librarian-in-the-Loop. A Natural Language Processing Paradigm for Detecting Informal Mentions of Research Data in Academic Literature, arXiv.org, 10. März 2022, DOI: <https://doi.org/10.48550/arXiv.2203.05112>.