

AR 3364

---

*Susanne Blumesberger*

# Repositorien als Tools für ein umfassendes Forschungsdatenmanagement. Am Beispiel von PHAIDRA an der Universitätsbibliothek Wien

**Zusammenfassung:** Der Beitrag soll zeigen, inwieweit Repositorien eine Rolle beim Forschungsdatenmanagement spielen und welche Möglichkeiten aber auch Herausforderungen damit verbunden sind. Der Umgang mit Daten hat sich in den letzten Jahren aufgrund der technischen aber auch der gesellschaftspolitischen Bedingungen geändert. Das hat auch Einfluss darauf, wie Repositorien gesehen und verwendet werden. Am Beispiel des Repositoriums PHAIDRA der Universität Wien kann diese Entwicklung von einem reinen Datenspeicher bis zu einem gut integrierten Tool in einer forschungsunterstützenden Umgebung nachvollzogen und beschrieben werden.

**Schlüsselwörter:** digitale Repositorien, Forschungsdatenmanagement, Bibliotheken, Langzeitarchivierung

## **Repositories as Tools for Comprehensive Research Data Management. The Example of PHAIDRA at the Vienna University Library**

**Abstract:** This article is intended to show the extent to which repositories play a role in research data management and the opportunities but also challenges associated with this. The handling of data has changed in recent years due to technical but also socio-political conditions. This also has an influence on how repositories are seen and used. Using the example of the repository of the University of Vienna, PHAIDRA, this development from a pure data storage to a well-integrated tool in a research supporting environment can be understood and described.

**Keywords:** digital repositories, research data management, libraries, long term preservation

### **1 Einleitung**

Die Krise um COVID-19 hat gezeigt, dass es sehr rasch zu einer nicht plan- und vorhersehbaren Änderung des täglichen Lebens kommen kann. Universitäten stellen auf Notbetrieb um, Bibliotheken schließen landesweit und es kommt plötzlich zu einer Reduktion der Informationsversorgung mit

analogen Medien. In dieser Situation ist es hilfreich auf digitale Medien, die sicher archiviert, rasch und möglichst kostenlos verfügbar sind, zugreifen zu können. Repositorien wurden und werden genau zu diesem Zweck aufgebaut. Dort sind im besten Fall Daten aller Art, wie Bücher, Texte, Video- und Audiofiles, Lehr- und Lernressourcen sicher gespeichert und mit entsprechenden Lizenzen versehen, damit andere Personen sie auch rechtskonform weiterverwenden und nachnutzen können. Abgesehen von diesen grundsätzlichen Anforderungen, sind Repositorien auch ein wichtiges Werkzeug im Forschungsdatenmanagement. Dafür benötigt man mehr als einfach ein funktionierendes technisches System, wichtig ist auch eine umfassende Unterstützung beim Datenmanagement für Forschende, passende Lösungen für unterschiedliche Anforderungen bei der Speicherung, Archivierung und Visualisierung anzubieten und als Schnittstelle zu anderen Bibliotheksbereichen zu fungieren, beispielsweise zu Forschungsinformationssystemen damit die Daten nicht mehrmals eingegeben werden müssen. Der Beitrag soll vor allem anhand des Repositoriums PHAIDRA<sup>1</sup> der Universität Wien zeigen, wie dies gelingen kann, aber auch, wo es Grenzen und Hürden gibt.

## 2 PHAIDRA, das Repository der Universität Wien

Als 2007 an der Universitätsbibliothek Wien ein Projekt startete, das sich zum Ziel gesetzt hatte, ein Repository mit Langzeitarchivierungsfunktionen für die gesamte Universität aufzubauen, war nicht abschätzbar, welche Tätigkeitsfelder sich daraus in Zukunft entwickeln würden. Nachdem für die technische Umsetzung die Entscheidung für das Open Source Produkt „Fedora“ gefallen war, folgte ein Jahr Planung der Universitätsbibliothek gemeinsam mit dem Zentralen Informatikdienst der Universität Wien. Gemeinsam erfolgte eine intensive Recherche nach Anforderungen, es wurden möglichst viele Universitätsmitarbeiter\*innen aus unterschiedlichen Fachdisziplinen eingebunden und die Software an die Bedürfnisse der Universität Wien, die mit ca. 100.000 Angehörigen - davon knapp 90.000 Studierende -, nicht nur sehr groß, sondern auch fachspezifisch sehr heterogen ist, angepasst. Genau ein Jahr nach Projektstart, am 16.4.2008, ging das Repository<sup>2</sup> PHAIDRA online. Von Anfang an gab es zwei klare Richtlinien:

- a) PHAIDRA sollte erstens allen Angehörigen der Universität Wien, also auch allen Studierenden offen stehen, alle durften Objekte, die für sie archivierungswürdig waren und für die sie die Rechte hatten, hochladen und
- b) es durfte zweitens nichts mehr gelöscht werden.

---

<sup>1</sup>Der Name PHAIDRA ist ein Akronym und bedeutet „Permanent Hosting, Archiving and Indexing of Digital Resources and Assets“

<sup>2</sup><https://phaidra.univie.ac.at/>

Die Metadaten, die Beschreibungen, die die Forscher\*innen selbst eintragen, bzw. bei einem Bulk-Upload hochladen lassen, können jedoch von den Eigentümer\*innen der Daten, d.h. von der Person, die das Objekt hochgeladen hat, immer wieder verändert und ergänzt werden. Bei Bedarf erhalten die Forscher\*innen Unterstützung beim Eintragen der Metadaten, die immer im Netz auffindbar bleiben, auch wenn die Objekte gesperrt sind.

Die Entscheidung, PHAIDRA auch für Studierende zu öffnen war zwar recht mutig und ungewöhnlich, stellte sich aber als praktikabel und richtig heraus, da bis heute keine missbräuchliche Verwendung des Repositoriums zu beobachten ist. Die zweite Richtlinie, die die Nichtlösbarkeit eines Objekts betrifft, wird, wie weiter unten beschrieben, zurzeit überdacht.

Kurze Zeit später folgte die Möglichkeit, Zweitveröffentlichungen von Publikationen über eine spezielle Eingabemaske hochzuladen, es entstand das Institutionelle Repositorium *u:scholar*<sup>3</sup>, das technisch gesehen, ein Teil von PHAIDRA ist, und das ebenfalls sämtlichen Angehörigen der Universität Wien zur Verfügung steht. Die vergebenen Lizenzen werden bei *u:scholar* vor der endgültigen Archivierung von Bibliotheksmitarbeiter\*innen einzeln geprüft und mit den Vorgaben der Verlage verglichen. So werden Rechtsverletzungen ausgeschlossen.

Die Erwartungen der Betreiber\*innen waren nach dem Go-Live von PHAIDRA entsprechend groß, wurden jedoch zum Teil zunächst enttäuscht. Weder Forscher\*innen noch Lehrveranstaltungsleiter\*innen nutzten das neue System vorerst sehr intensiv. Es dauerte eine Weile bis nach zahlreichen Schulungen, Präsentationen und Roadshows sowie durch Mundpropaganda PHAIDRA und *u:scholar* bekannter geworden waren. Erst nach einiger Zeit sahen die Forscher\*innen die Vorteile, die man mit einem Repositorium hat, das die eigenen Daten und Publikation mit einer permanenten Signatur versieht und sicher speichert. Nun war es plötzlich möglich, den eigenen Forschungsoutput und den der anderen jederzeit aufzurufen und zu zitieren, digitalisierte Objekte, Texte, Videos, Audiofiles, Bilder und Daten jeglicher Herkunft und Art zusammenzufügen und auch anderen – wenn gewünscht weltweit – zur Verfügung zu stellen. Im August 2020 sind in PHAIDRA mehr als 53500 Bilder gespeichert, über 9600 Bücher, mehr als 6400 Artikel und über 7300 Textdokumente. Die 1800 Videos und 660 Audiofiles wurden vermehrt in den letzten Jahren hochgeladen, und auch die bisher archivierten 1260 Forschungsdatensätze kamen erst in der letzten Zeit hinzu.

So heterogen die Universität Wien fachlich ist, so unterschiedlich haben auch die Wissenschaftler\*innen aufgrund ihrer speziellen Forschungsmethoden in den jeweiligen Disziplinen PHAIDRA in die Forschungsprozesse integriert. Stark genutzt wurde das Repositorium zunächst vor

---

<sup>3</sup><https://uscholar.univie.ac.at/>

allem von der Universitätsbibliothek Wien selbst. Denn durch das Projekt EOD<sup>4</sup>, e-Books on Demand, war schon eine beachtliche Anzahl an wertvollen historischen Büchern gescannt worden und konnte nun endlich auch entsprechend in einem eigens dafür kreierten Bookviewer<sup>5</sup>, der in PHAIDRA für hochgeladene Bücher angeboten wird, der Öffentlichkeit open access präsentiert werden. Nach einer intensiven Bewerbungswelle von PHAIDRA begannen einige Forschungsprojekte das Repositorium rege zu nutzen um einen großen Bestand an neu digitalisierten Objekten, wie etwa Glasplattendias, die bereits vom Verfall bedroht waren, langfristig zu sichern. Auch an anderen Institutionen Österreichs und auch im Ausland, allen voran an der Universität Padua, war man auf PHAIDRA aufmerksam geworden und bat um einen Kooperationsvertrag.<sup>6</sup> Daraus entwickelte sich nach und nach ein derzeit 20 Institutionen umfassendes Netzwerk.<sup>7</sup> Langsam erkannten die Wissenschaftler\*innen die Vorteile von PHAIDRA, das von Beginn an auch als Forschungsdatenrepositorium konzipiert war, es wurden im Laufe der Zeit vermehrt Forschungsdaten archiviert. Vor allem die erhöhte Sichtbarkeit und Sicherheit des eigenen Forschungsoutputs boten einen großen Anreiz, die Ergebnisse in PHAIDRA zu stellen.

Das System folgt der 2014 verabschiedeten Open Access Policy<sup>8</sup> der Universität Wien, jedes hochgeladene Objekt ist per Default sichtbar, kann aber sofort auf unterschiedliche Benutzer\*innenkreise eingeschränkt werden. Wenn zum Beispiel aus rechtlichen Gründen open access nicht in Frage kommt, können die Uploader\*innen jederzeit den Zugriff auf das Objekt auf die gesamte Universität, auf Institute, Personengruppen oder Einzelpersonen einschränken und auch wieder öffnen. Dann bleiben nur noch die Beschreibungen der Objekte, die Metadaten, weltweit offen und sichtbar. Wichtig war von Anfang an, dass PHAIDRA eine den heutigen FAIR-Prinzipien entsprechende Forschung ermöglicht. Danach müssen die Daten findable (auffindbar), accessible (zugänglich), interoperable (austauschbar) und reusable (wiederverwendbar) sein.<sup>9</sup> Um Daten wiederverwenden zu können, ist es unverzichtbar, sie mit Lizenzen zu versehen, die die Möglichkeiten der Nachnutzung eindeutig bestimmen. Derzeit wird in PHAIDRA neben den Creative-Common-Lizenzen – ab CCBY 2.0-Lizenz auch die Public Domain Marke und die GNU-Lizenz für Computerprogramme angeboten.<sup>10</sup> Die Nutzer\*innen wählen selbst die geeignete freie Lizenz aus, können sie jedoch aus juristischen Gründen im Nachhinein nicht mehr verändern. Es gibt aber die Möglichkeit „keine Lizenz“ auszuwählen, was bedeutet, dass das hochgeladene Objekt lediglich im

---

<sup>4</sup><https://books2ebooks.eu/de>

<sup>5</sup>Siehe etwa <https://phaidra.univie.ac.at/o:19958> Gaius Plinius Secundus' Plinius "Historia Naturalis" ist das älteste Buch im Bestand der Universitätsbibliothek Wien. Es handelt sich dabei um eine Inkunabel (Frühdrucke aus der Zeit 1450-1500)

<sup>6</sup><https://phaidra.cab.unipd.it/>

<sup>7</sup><https://phaidra.org/>

<sup>8</sup><https://openaccess.univie.ac.at/ueber-open-access/oa-policy-der-uni-wien/>

<sup>9</sup>Siehe <https://www.force11.org/group/fairgroup/fairprinciples>

<sup>10</sup>Siehe <https://datamanagement.univie.ac.at/forschungsdatenmanagement/juristisches/lizenzmodelle/>

Rahmen des Urheberrechts verwertet werden darf. Dieser Prozess der Lizenzierung wird regelmäßig in Schulungen und bei Einzelbesprechungen erklärt, bei speziellen Fragestellungen kann auch juristische Expertise herangezogen werden.

Im Nachhinein war es eine gute Entscheidung, für das Repositorium eine Open-Source Software zu nutzen und nach Anforderungsanalysen das Repositorium darauf aufzubauen, denn Änderungen, wie die derzeitige Umgestaltung der Eingabemaske, können recht rasch und unabhängig von kommerziellen Anbietern umgesetzt werden. Die Erfahrung beim Aufbau des Systems und die Software selbst, die auch open source zur Verfügung steht, werden bereitwillig mit Interessent\*innen geteilt.

### **3 PHAIDRA als Teil des Forschungsdatenmanagements an der Universität Wien**

Das Bibliotheksteam und die Kolleg\*innen vom Zentralen Informatikdienst haben früh begonnen auf PHAIDRA aufbauende Services zu entwickeln. Diese umfassen Hilfestellungen beim Hochladen von Objekten, z.B. über ein dem Projekt angepasstes Bulk-Upload-Tool, Unterstützung beim gesamten Datenmanagement inklusive der sonstigen Angebote am ZID der Universität Wien für das Speichern und Teilen von Dokumenten, wie Shares, und vor allem auch die Beratung beim Ausfüllen von Datenmanagementplänen. Dennoch war es nicht so einfach, diese Angebote auch rasch und flächendeckend zu kommunizieren. Die Größe und die Heterogenität der Universität Wien<sup>11</sup> erschweren eine rasche und direkte Kommunikation mit jenen Zielgruppen, die das System auf unterschiedliche Art und Weise nutzen. Grob kann man zwischen den Anwender\*innen aus der Verwaltung, aus der Lehre und jenen aus der Forschung unterscheiden, wobei auch hier noch eine genaue Differenzierung nötig ist. Zu den Daten aus der Verwaltung zählen beispielsweise historische Studierendenakte oder andere Bestände aus dem Universitätsarchiv, wie Senatsakten oder Studienkataloge, beispielsweise Informationen über promovierte Frauen vor 1923, die offen und frei zugänglich archiviert werden können.<sup>12</sup> Verwaltungsdokumente können aber auch Informationen enthalten, die nur für den internen Gebrauch gedacht sind, personenbezogene oder sensible Daten beinhalten und deshalb gesperrt werden müssen. In vielen Fällen dürfen hier nicht einmal die Beschreibungen der jeweiligen Objekte sichtbar sein, sämtliche Metadaten müssen ausgeblendet werden, eine Forderung, die außerhalb von PHAIDRA in einem internen System realisiert werden musste.

Bei manchen Objekten, zum Beispiel bei Daten aus einem laufenden Forschungsprojekt ist es wichtig, wie etwa auch bei Zeitschriften üblich, eine Embargofrist anzugeben, das heißt, dass Objekte

---

<sup>11</sup> <https://www.univie.ac.at/ueber-uns/auf-einen-blick/zahlen-daten-broschueren/>

<sup>12</sup> Siehe beispielsweise eine Collection des Universitätsarchivs der Universität Wien: <https://phaidra.univie.ac.at/o:527682>

beispielsweise aus rechtlichen Gründen für eine bestimmte Zeit für die Öffentlichkeit gesperrt werden müssen, danach aber frei zugänglich sein dürfen. Dies wurde in PHAIDRA von Anfang an mitbedacht, jedes Objekt kann von der Person, die das Objekt hochgeladen hat (Owner) jederzeit mit einer Frist versehen werden. Die Objekte sind nach Ablauf dieser Frist automatisch frei verfügbar.

Daten, die während des Forschungsprozesses entstehen und/oder für die Forschung benötigt werden, bedürfen einer genaueren Betrachtung. Grundsätzlich können als Forschungsdaten alle Objekttypen gelten, wie beispielsweise Bilder, Texte, Bücher, Videos, Audiofiles, 3D-Objekte usw. Dabei lassen sich wieder digitalisierte Objekte und born-digital -Objekte unterscheiden, die eventuell mit unterschiedlichen Metadaten versehen werden können. Dies hat wiederum Auswirkungen auf die Wahl des jeweiligen Metadatenschemas in den Repositorien. Für PHAIDRA wurde entschieden, Dublin Core, Datacite oai\_dc<sup>13</sup> oai\_openaire<sup>14</sup> und das LOM-Schema zu verwenden. Für den Metadatenaustausch mit OpenAIRE<sup>15</sup>, Europeana<sup>16</sup> und beispielsweise auch mit dem Kulturpool<sup>17</sup> werden die Pflichtmetadatenfelder aus Dublin Core herangezogen (dc:title, Titel des Objekts; dc:description, Beschreibung; dc:language, Sprache des Objekts; dc:contributor, Rolle/n; dc:rights, Lizenz). Eine offen zugängliche API-Schnittstelle erlaubt den Austausch von Metadaten mit nationalen und internationalen Datenbanken. Auch technisch gesehen, sind Forschungsdaten differenziert zu betrachten. Bei born-digital-Objekten gibt es zum Beispiel eine recht große Bandbreite an Formaten, die nur zum Teil für die langfristige Verfügbarkeit geeignet sind. Werden Objekte, wie etwa Karten, Briefe, Bücher oder Videoaufnahmen digitalisiert, ist es möglich, mehr Einfluss auf das empfohlene Format für die digitale Langzeitarchivierung zu nehmen und beispielsweise auch als Repositorienbetreiber\*in schon im Vorfeld der Digitalisierung Unterstützung bei der Formatauswahl anzubieten. Eine technische Herausforderung stellen auch die Speicherung und das Verfügbarhalten sehr großer Dateien im Repository dar oder aber auch das Verwalten einer Vielzahl an kleinen Einheiten. Für diese Fälle, sowie für die langfristige Verfügbarkeit von Datenbanken wird an Lösungen gearbeitet.

Studierende nehmen das Angebot PHAIDRA zu nutzen (meist) anfangs meist nur zögerlich an, das ändert sich jedoch, wenn sie selbst in ein Forschungsprojekt eingebunden werden, bzw. ihre Abschlussarbeit schreiben. Derzeit werden aus diesem Grund verstärkt Schulungen im Bereich Datenmanagement für Studierende angeboten.

---

<sup>13</sup><http://www.openarchives.org/OAI/openarchivesprotocol.html>

<sup>14</sup>[https://guidelines.openaire.eu/en/latest/data/use\\_of\\_oai\\_pmh.html](https://guidelines.openaire.eu/en/latest/data/use_of_oai_pmh.html)

<sup>15</sup><https://www.openaire.eu/>

<sup>16</sup><https://www.europeana.eu/de>

<sup>17</sup><http://www.kulturpool.at/display/kupo/Home>

### 3.1 Differenziertes Forschungsdatenmanagement

Für Forschende ist die nachhaltige Nutzung ihrer Daten sehr wichtig geworden, nicht zuletzt aufgrund der Fördergeber, die die finanziellen Mittel, die aus Steuergeldern stammen, sparsam einsetzen müssen und deshalb darauf achten, dass einmal erhobene Forschungsdaten nachgenutzt werden und nicht neu erhoben werden müssen. Dafür wird die passende Infrastruktur benötigt. Repositorien sind ein Teil davon. Neben technischen, rechtlichen und ethischen Fragen, die sich bei jedem Projekt, das forschungsrelevante Daten archivieren möchte, immer neu und immer anders stellen, ist auch laufend auf die gesellschaftlichen Entwicklungen zu achten, auf Veränderungen in der Lehre, in der Forschung, bei den Fördergebern, im Umgang mit Medien usw. Die Erfahrungen haben gezeigt, dass eine gut funktionierende Kommunikation und Interaktion zwischen den Bibliotheksmitarbeiter\*innen, den Techniker\*innen und den Anwender\*innen hinsichtlich des Forschungsdatenmanagements unerlässlich ist. Dazu benötigen sie nicht nur technisches Basiswissen, sondern auch einen Einblick in den aktuellen Wissenschaftsbetrieb. Hier kommt das Konzept von Datastewards und Datachampions zum Tragen, das an manchen Universitäten bereits umgesetzt wird und das an der Universität Wien derzeit konzipiert wird.<sup>18</sup>

### 3.2 Rechtliche und ethische Fragen im Forschungsdatenmanagement

Vor allem bei aktuellen Forschungsdaten gibt es hinsichtlich der rechtlichen Bestimmungen ein breites Spektrum wie und wem sie zur Verfügung gestellt werden dürfen. Bei Messdaten, bei denen es im juristischen Sinn keine Urheber\*innen gibt, bestehen in den seltensten Fällen rechtliche Bedenken gegen einen offenen und freien Zugang. Anders sieht es jedoch aus, wenn man beispielsweise Bilder frei über Repositorien anbieten möchte. Hier muss nicht nur darauf geachtet werden, dass die Rechte der Fotograf\*innen berücksichtigt werden, sondern auch die der eventuell abgebildeten Personen. Auch bei Texten müssen Rechte bei den Urheber\*innen, bzw. bei den Nachlassverwalter\*innen oder beim Verlag eingeholt werden. Bei Ton- und Videoaufnahmen kann es eventuell noch schwieriger werden, da hier in den meisten Fällen mehrere Personen am Entstehen des Werkes beteiligt waren. Nicht immer lassen sich diese rechtlichen Fragen eindeutig lösen, deshalb kann es vorkommen, dass Forschende ihre Objekte nicht frei zugänglich machen. Umfrageergebnisse enthalten oft sensible Daten und sind deshalb mit der nötigen Umsicht zu behandeln, es müssen Einverständniserklärungen eingeholt werden, die Daten müssen oft vor der Publikation pseudonymisiert, bzw. anonymisiert werden. Da aber die Policy<sup>19</sup> von PHAIDRA erlaubt, die Objekte auf einen selbst definierten Benutzer\*innenkreis einzuschränken, - wenn nötig auch mit einer Embargofrist – steht einer Archivierung in PHAIDRA in den meisten Fällen nichts entgegen.

---

<sup>18</sup>Ein Beispiel dafür ist die TU Delft: <https://www.tudelft.nl/en/library/current-topics/research-data-management/r/support/data-stewardship/>

<sup>19</sup><https://datamanagement.univie.ac.at/ueber-phaidra-services/policy-von-phaidra/>

Zu den rechtlichen Bestimmungen kommen aber zum Teil auch noch ethische Fragen hinzu. Wie geht man beispielsweise mit Interviews um, die vor 30 Jahren geführt wurden und für die man damals die Einverständniserklärung eingeholt hat – in einer Zeit ohne weltweite Verbreitungsmöglichkeit durch das Internet? Damals waren die Folgen der technischen Weiterentwicklung weder für den oder die Interviewer\*in noch für den oder die Interviewpartner\*in absehbar. Ist diese Person heute nicht mehr erreichbar, stellt sich die Frage, ob es ethisch vertretbar ist, das Interview frei ins Netz zu stellen. Auch Informationen, die andere Menschen zum Beispiel aufgrund ihrer politischen Einstellung, ihrer ethischen Zugehörigkeit, ihrer Vergangenheit usw. gefährden würden, müssen mit Vorsicht behandelt werden. Diskussionen gibt es beispielsweise auch bei der Archivierung von Bildern, die menschliche Überreste zeigen. Vor allem im Umgang mit sensiblen Daten aber auch bei rechtlichen Fragen ist in den letzten Jahren ein deutliches Mehr an Reflexion zu erkennen. In den Schulungen und Beratungsgesprächen bieten wir auch an bei solchen Fragen behilflich zu sein. Allerdings steht eine allgemeingültige Lösung heute noch aus. Der enge Kontakt mit den Forschenden und das Kennenlernen ihrer Bedürfnisse haben dazu geführt, dass es auch bei PHAIDRA ein Weiterdenken bzgl. der technischen Angebote für das Forschungsdatenmanagement gab, beispielsweise in Form der Bereitstellung von Speicherplatz oder das Nutzen von GitLab, einer Webanwendung zur Versionsverwaltung für (Software-)projekte. Diese zusätzlichen Angebote sollen mit den Anforderungen der Projekte mitwachsen, denn nahezu jedes Projekt hat einen anderen Bedarf und daher ist Flexibilität gefragt.

### **3.3 Repositorien für die Lehre und für Hochschulschriften**

Daten, die in der Lehre verwendet werden, können aus unterschiedlichen Zeiten, Fachgebieten und Kontexten stammen. Je nachdem können sie entweder mit freien Lizenzen und für alle nutzbar zur Verfügung gestellt werden oder müssen den Studierenden in einem geschlossenen System angeboten werden. Universitäten, wie beispielsweise auch die Universität Wien, fördern die Erstellung und Archivierung von Open Education Resources, die in PHAIDRA zur Verfügung gestellt werden. Wenn diese Offenheit aufgrund rechtlicher Bestimmungen nicht möglich ist, stehen den Studierenden die E-Learning Plattform moodle oder eine interne Datenbank namens UNIDAM zur Verfügung. UNIDAM basiert auf der Software easydb der Firma Programmfabrik und wird vom PHAIDRA-Team mitbetreut. Die Datenbank wird vor allem für Bilddaten verwendet, die mittels eines eigens vergebenen Accounts Lehrenden und Studierenden für Lehrzwecke bereit steht.

Hochschulschriften, die in Repositorien langfristig abrufbar sein sollen, bedürfen einer besonderen Aufmerksamkeit. Hier sind Universitätsgesetze zu beachten, die genau regeln, wie die Arbeiten archiviert werden dürfen, welche Schritte die Urheber\*innen unternehmen müssen und welche die Betreiber\*innen der Systeme. Abschlussarbeiten durchlaufen vor der Archivierung meist einen



Plagiatsprüfungsprozess, der in den Gesamtworkflow der Archivierung miteingebunden ist. Eventuell muss es den Urheber\*innen ermöglicht werden, Abstracts und andere Metadaten zu verstecken und die Arbeiten zu sperren. Die Gründe dafür können eine später geplante Publikation der Arbeit sein, aber auch Verträge mit Firmen oder mit der Abschlussarbeit verbundene Patente. Der Ablauf ist an österreichischen Universitäten unterschiedlich, an der Universität Wien beispielsweise wird derzeit die Migration der Hochschulschriften von einem externen Server auf PHAIDRA durchgeführt, eine Arbeitsgruppe, die sich aus Expert\*innen aus dem Bibliotheks- und IT-Bereich zusammensetzt, begleitet diesen Prozess.

#### **4 Neue Forschungsfelder – neue Herausforderungen für das Forschungsdatenmanagement**

Doch nicht nur die Datenmanagementpläne sensibilisieren für den Umgang mit Daten, sondern auch andere Faktoren, wie die technische Weiterentwicklung und die veränderte Arbeitsweise der Forschenden. Auch das Bearbeiten neuer Themen und damit die Produktion neuartiger Daten zählen dazu. Die rasante Entwicklung des Smartphones zu einem vollwertigen Computer, den man ständig mit sich trägt, macht das Filmen und Fotografieren einfacher und weniger aufwändig. Ob die Formate, in denen Bilder und Filme dann gespeichert werden, immer für die Langzeitarchivierung günstig sind, ist jedoch fraglich. Neue Forschungsfelder entstanden etwa in der Germanistik und in der Kommunikationswissenschaft unter anderem rund um diverse soziale Netzwerke. Für eine geplante Langzeitarchivierung diverser Postings sind eventuell technische Fragen relevant, zum Beispiel welches Datenformat man wählen soll, vor allem aber rechtliche und eventuell ethische Fragestellungen. Das Mediennutzungsverhalten von Kindern und Jugendlichen wird derzeit ebenfalls recht häufig untersucht.<sup>20</sup> Will man diese Daten nachnutzbar machen, muss man schon während der Vorbereitung der Studien einige Vorkehrungen treffen. All diese Entwicklungen haben einen signifikanten Einfluss auf die Angebote der forschungsunterstützenden Services, darunter eben auch auf die Repositorien.

#### **5 Unterschiedliche Anforderungen – verschiedene Lösungen**

Bei medizinischen Forschungen müssen Daten oft nach einem gewissen Zeitraum wieder gelöscht werden, das heißt PHAIDRA ist dafür nicht das geeignete System. Um jedoch weiterhin für alle Forschungsrichtungen attraktiv zu sein, wurde der Beschluss gefasst, PHAIDRA in Zukunft zweistufig zu betreiben und vor der endgültigen Langzeitarchivierung, bei der ein permanenter Link vergeben wird, eine Löschfunktion anzubieten.

---

<sup>20</sup>Siehe beispielsweise die Forschungsplattform „Research Platform Mediatized Lifeworlds Young people's narrative constructions, connections and appropriations“ <https://youthmedialife.univie.ac.at/>

Auch die Art und Weise, wie Daten archiviert werden, hat sich verändert, bzw. beschleunigt. Ursprünglich war PHAIDRA für den Einzelupload von Objekten gedacht. Dies entspricht heute nicht mehr dem gewohnten Umgang mit Daten, die im Drag and Drop-Verfahren in Systeme hochgeladen werden können. Auch hier musste sich PHAIDRA weiterentwickeln. Den Forschenden stehen derzeit mehrere Möglichkeiten des Uploads zur Verfügung. Personen oder Forschergruppen, die eine große Anzahl von Objekten in PHAIDRA laden möchten, erhalten maßgeschneiderte technische Unterstützung, wie z.B. einen Bulk-Upload, die derzeit für Angehörige der Universität Wien kostenlos in Anspruch genommen werden kann.

Um eine möglichst hohe Sichtbarkeit des Forschungoutputs zu erreichen, sind die Metadaten immer öffentlich sichtbar, auch wenn die Objekte selbst gesperrt sind. Diese Vorgehensweise entspricht z.B. auch den FAIR-Prinzipien. In einigen Fällen ist diese Vorgehensweise jedoch kontraproduktiv – vor allem wenn es sich um hochkompetitive Forschungsfelder handelt, oder wenn vertrauliche Daten archiviert wurden. In diesen Fällen werden den Anwender\*innen andere Speichermöglichkeiten, z.B. Server, die vom Zentralen Informatikdienst der Universität Wien (ZID) betreut werden, empfohlen. Die Erfahrung mit Repositorien hat gezeigt, dass nicht, wie anfangs fälschlicherweise angenommen, ein einziges Repository für alle Fälle ausreicht. Für Lehrzwecke wird zuzüglich zu Moodle und UNIDAM die Software GitLab angeboten. Für die Speicherung von Daten während eines Forschungsprozesses stehen universitätsinterne und den jeweiligen Bedürfnissen angepasste Cloudlösungen sowie Shares zur Verfügung. Sozialwissenschaftliche Daten sind bei AUSSDA<sup>21</sup> (The Austrian Social Science Data Archive) gut aufgehoben. Im Unterschied zu PHAIDRA wird dort kein Selfarchiving angeboten, die Daten werden kuratiert.

Permanente Identifikatoren haben in den letzten Jahren an Bedeutung gewonnen. Forschende benötigen für ihre Publikationen und nicht zuletzt für ihre Karriere und die Fördergeber Digital Object Identifier (z.B. DOI). Neben den automatisch vergebenen PHAIDRA- und Handle-Links<sup>22</sup> können User vor oder nach der Archivierung der Objekte in PHAIDRA einen DOI beantragen.<sup>23</sup> Aus mehreren Gründen, unter anderem weil Objekte, die in PHAIDRA hochgeladen werden, bereits einen DOI haben können, werden diese (noch) nicht automatisch vergeben.

Die Verwendung von ORCID setzt sich ebenfalls immer mehr durch. Für den österreichischen Wissenschaftsfonds FWF ist eine ORCID-ID für Antragsteller\*innen seit 2016 verpflichtend.<sup>24</sup> Auch

---

<sup>21</sup> <https://aussda.at/>

<sup>22</sup> <https://www.forschungsdaten.org/index.php/Handle>

<sup>23</sup> <https://doi-service.univie.ac.at/>

<sup>24</sup> <https://bibliothek.univie.ac.at/bibliometrie/orcid.html>

hier wurde in PHAIDRA die Möglichkeit geschaffen, diesen Identifier einzutragen und den Eintrag mit dem ORCID-Profil zu verlinken, eine automatische Datenübernahme von und nach ORCID ist geplant.

## 6 Zukünftige Herausforderungen

Die spezifischen Bedürfnisse jener, die die Daten zur Verfügung stellen, bzw. nutzen, führen zum Teil zu großen Herausforderungen bei den Repositorienbetreiber\*innen. Der Datenmanagementplan, der seit einiger Zeit von einigen Fördergebern verpflichtend eingeführt wurde, wie beispielsweise von der EU, hat dazu geführt, dass sich die Forschenden intensive Gedanken darüber machen, wie sie mit ihren Daten während und nach Abschluss der Forschungsprojekte umgehen. Konkrete Fragen nach Zuständigkeiten, nach der rechtlichen Klärung, nach eventuellen ethischen Fragen, aber auch nach technischen Anforderungen oder ob überhaupt für alle Daten eine langfristige Sicherung erforderlich oder überhaupt erlaubt ist, helfen auch den Betreiber\*innen der Repositorien. Diese erfahren im besten Fall dabei schon recht früh, welche Art von Daten und Metadaten in welchen Formaten und nach welchen Standards gespeichert werden soll, wie groß die Dateien sind, um welches Datenvolumen es sich handelt, wer zuständig für die Daten ist, ob es technischer Adaptierungen bedarf usw.

Je rascher die Repositorienbetreiber\*innen von den Forschungsvorhaben und den detaillierten Plänen für die Daten erfahren, desto besser können sie sich auf die Anforderungen einstellen und Vorbereitungen treffen. Die Aufgaben, die ein universitäres Repository heute erfüllen muss, erweitern sich ständig und müssen immer wieder an neue Vorgaben von außen, an Bedürfnisse der Forschenden und an andere Systeme angepasst werden.

Eine Herausforderung wird in Zukunft auch der Umgang mit der Datenflut sein, denn durch die fortlaufende Digitalisierung sind immer mehr Daten im Netz verfügbar. Wir alle, ob Nutzer\*innen, Anbieter\*innen der Daten oder beispielsweise Bibliothekar\*innen werden Wege finden müssen, uns in diesem großen Pool an Informationen zurechtzufinden. Suchmaschinen, bzw. Metasuchmaschinen über mehrere Repositorien hinweg, werden dabei immer wichtiger und sind auch für PHAIDRA geplant. Daten müssen rasch aufgefunden werden, das verlangt neben einer guten technischen Basis, auch hochqualitative und maschinenlesbare Beschreibungen der Daten.

Für die Umsetzung der FAIR-Data-Prinzipien, werden mehrere Komponenten benötigt, unter anderem die Maschinlesbarkeit der Daten und Beratung im Bereich der Metadaten, beispielsweise bei der Auswahl von geeigneten Metadatenschemata. Viele, die selbst Daten anbieten, unterschätzen den Wert der Metadaten, die für die Visualisierung der Daten, bzw. für den Austausch mit anderen Systemen, unverzichtbar sind. Die von den Fördergebern zum Teil bereits vorgeschriebenen Datenmanagementpläne werden von den Forschenden oft als Hürde gesehen,

erweisen sich in der Praxis aber meist als wirksames Instrument um sich den eigenen Forschungsprozess zu verdeutlichen und sich Strategien für das Datenmanagement zu überlegen. Die Sensibilität gegenüber der Frage was man wie veröffentlichen darf, bzw. worauf man achten sollte, ist glücklicherweise gewachsen, es besteht eine gesteigerte Nachfrage nach Unterstützung auf diesem Gebiet, zum Beispiel bei Fragen nach Pseudonymisierung oder Anonymisierung. Die Forschenden sind mehr denn je daran interessiert ihre Daten mit anderen auszutauschen, auch hier wird Unterstützung angeboten.

## **7 Mögliche Strategien für ein effektives Forschungsdatenmanagement**

Um fit für ein umfangreiches Forschungsdatenmanagement zu bleiben bedarf es einiger Überlegungen. Netzworkebildungen erleichtern auf jeden Fall den Austausch über Herausforderungen. Beispielsweise hat sich das Netzwerk für Repositorienmanager\*innen (RepManNet) beim Lösen von Problemen bzw. beim Gestalten von Workflows als hilfreich erwiesen.<sup>25</sup> Aber auch Netzwerke, die gemeinsam mit Forschenden aufgebaut wurden, erweisen sich als nützlich. So hat das Netzwerk „Nachhaltige Softwareentwicklung code4research“<sup>26</sup> an der Universität Wien dazu geführt, dass sich Forschende aus unterschiedlichen Fachbereichen, vor allem aus den Digital Humanities regelmäßig treffen und gemeinsam Strategien überlegen, wie man beispielsweise Datenbanken auf lange Sicht archivieren und zugänglich halten kann. Auch durch die insgesamt 20 Partnerinstitutionen von PHAIDRA<sup>27</sup> ist durch viele gemeinsame Erfahrungen neues Wissen entstanden. Durch internationale Projekte, wie beispielsweise Europeana und nationale wie etwa „e-infrastructures austria“<sup>28</sup> und „e-infrastructures austria plus“<sup>29</sup> wurde Expertise aufgebaut und mit anderen geteilt. Der kontinuierliche Kontakt mit den Fördergebern lässt frühzeitig Trends erkennen und ermöglicht es rasch auf Veränderungen bei den Vergaberichtlinien zu reagieren. Als sehr wichtig erweist sich immer mehr der enge Kontakt zu den Forschenden, denn durch sie gewinnen wir einen Einblick in ihre Arbeit und können unsere Angebote danach richten. Absolut grundlegend ist natürlich die stete Zusammenarbeit mit der IT. Nur so können miteinander rasch und effektiv Projekte umgesetzt und das Repository weiterentwickelt werden.

Zusammenfassend lässt sich sagen, dass Repositorien ein wichtiges Werkzeug für effizientes Datenmanagement sind. Sie müssen, ebenfalls wie die zugehörigen Services, an den Bedürfnissen der

---

<sup>25</sup>Siehe auch: Blumesberger (2019).

<sup>26</sup><https://datamanagement.univie.ac.at/rdm/netzwerk-nachhaltige-softwareentwicklung/>

<sup>27</sup><https://phaidra.org/>

<sup>28</sup><https://e-infrastructures.univie.ac.at/>

<sup>29</sup><https://www.e-infrastructures.at/de/>

Forschenden mitwachsen und sich entwickeln. Genau hier liegen die Herausforderungen. Was zeigen also unsere Erfahrungen?

- 1) **Zusammenarbeit:** Eine kontinuierliche Zusammenarbeit von Forschenden mit Techniker\*innen und Bibliothekar\*innen ist wichtig. Forscher\*innen müssen aktiv in Überlegungen rund um das Forschungsdatenmanagement und den Aufbau von Repositorien miteinbezogen werden.
- 2) **Ressourcen:** Es müssen genügend Personal - und finanzielle Ressourcen zur Verfügung stehen um ein stabiles und umfassendes Forschungsdatenmanagement garantieren zu können.
- 3) **Flexibilität:** Eine hohe Flexibilität der vorhandenen Systeme, wie beispielsweise der Repositorien ist nötig. Diese müssen bei Bedarf immer wieder angepasst werden können. Das betrifft die Technik und beispielsweise auch die Metadatenschemata
- 4) **Diversität:** Für unterschiedliche Anforderungen müssen verschiedene technische Lösungen bereitgestellt werden, neben der Langzeitarchivierung sollten Löschoptionen oder ausreichend Speicherplatz für Forschende vorhanden sein.
- 5) **Rahmenbedingungen:** Es müssen zusätzlich geeignete Rahmenbedingungen geschaffen werden, wie beispielsweise Policies, die Abläufe regeln.
- 6) **Anpassung:** Das Repository und die angebotenen Services müssen in die bereits vorhandene Systemlandschaft eingebettet werden
- 7) **Kommunikation:** vor allem in großen und heterogenen Institutionen ist es wichtig, die Services entsprechend zu bewerben
- 8) **Offenheit:** Offen zugängliche Schnittstellen ermöglichen den raschen Austausch mit anderen Systemen
- 9) **Netzwerke:** Netzwerke, die Zusammenarbeit mit Forschungsdatenmanager\*innen aus anderen Institutionen ermöglichen gemeinsame Aktivitäten und einen regen Austausch
- 10) **Usability:** Nicht zuletzt entscheidet die Usability, ob das Repository angenommen wird oder nicht. Das heißt aber auch, dass die Repositorienbetreiber\*innen bzgl. der technischen und gesellschaftlichen Entwicklungen am Laufenden bleiben müssen
- 11) **User\*innenorientierung:** Es ist wichtig zu wissen, was die User\*innen an Technik und an Services benötigen.
- 12) **Neugierde:** Interesse an den Forschungen in den jeweiligen Fachdisziplinen, an den neueren technischen und gesellschaftspolitischen Entwicklungen um passende Services anzubieten.
- 13) **Bodenhaftung:** Ein sehr wichtiger Punkt ist, einschätzen zu können, was wann mit welchen Mitteln umgesetzt werden kann, denn leere Versprechungen verärgern die User\*innen.

## Literatur

- Blumesberger, Susanne (2017): „Herausforderungen für Repositorien!“ Tagung anlässlich 10 Jahre Phaidra an der Universität Wien. In: *Mitteilungen der VÖB*, 70 (3/4), 373–76. Online verfügbar unter <https://doi.org/10.31263/voebm.v70i3.1960>
- Blumesberger, Susanne (2018): Bibliothek und Digitalität am Beispiel der Universitätsbibliothek Wien. In: *Germanistik digital. Digital Humanities in der Sprach- und Literaturwissenschaft*, hg. v. Ingo Börner, Wolfgang Straub und Christian Zolles. Wien: Facultas, 199–211.
- Blumesberger, Susanne (2018): Neue Anforderungen – viele offene Fragen. Zu den vielfältigen Rollen von Repositorien am Beispiel der UB Wien. In: *Repositorien in Österreich – Status quo und (zukünftige) Herausforderungen*, hg. v. Susanne Blumesberger, Andreas Ferus, Christian Kaier, Gertraud Novotny. VÖB Mitteilungen, 71 (1), 149–61. Online verfügbar unter <https://doi.org/10.31263/voebm.v71i1>
- Blumesberger, Susanne (2019): „Druckt ihr eh alles aus was in PHAIDRA ist?“ Die Abteilung Repositorienmanagement PHAIDRA-Services – ein Paradigmenwechsel. In: *Menschen im Aufbruch. Universitätsbibliothek und Archiv der Universität Wien im Selbstverständnis ihrer Mitarbeiter\_innen. Festschrift für Maria Seissl*, hg. v. Stefan Alker-Windbichler, Claudia Feigl, Christina Köstner-Pemsel, Thomas Maisel, Wolfgang Nikolaus Rappert, Pamela Stückler, Markus Stumpf. Göttingen: Vandenhoeck & Ruprecht, 269–72. Online verfügbar unter <https://doi.org/10.14220/9783737010986>
- Blumesberger, Susanne (2019): Das RepManNet – das österreichische Netzwerk für RepositorienmanagerInnen. In: *Mitteilungen der VÖB*, 72 (1), 16063. Online verfügbar unter <https://doi.org/10.31263/voebm.v72i1.2291>
- Blumesberger, Susanne; Ganguly, Raman (2019): Der Umgang mit heterogenen (Forschungs-)daten an einer wissenschaftlichen Bibliothek – Use Cases und Erfahrungen aus technischer und nicht technischer Sicht an der Universität Wien. In: *Forschungsdaten – Sammeln, sichern, strukturieren. 8. Konferenz der Zentralbibliothek, Forschungszentrum Jülich, WissKom 2019, Jülich, Germany, 4 Jun – 6 Jun 2019*, 193–200. Online verfügbar unter <http://hdl.handle.net/2128/22274>
- Eberhard, Igor; Kraus, Wolfgang (2018): Der Elefant im Raum. Ethnographisches Forschungsdatenmanagement als Herausforderung für Repositorien. In: *Mitteilungen der VÖB*, 71 (1), 41–52. Online verfügbar unter <https://doi.org/10.31263/voebm.v71i1.2018>
- Ganguly, Raman (2018): Nachhaltige Softwareentwicklung: code4research. In *Mitteilungen der VÖB*, 71 (1), 171–80. Online verfügbar unter <https://doi.org/10.31263/voebm.v71i1.2019>
- Hagmann, Dominik (2018): Überlegungen zur Nutzung von PHAIDRA als Repository für digitale archäologische Daten. In: *Mitteilungen der VÖB*, 71 (1), 53–69. Online verfügbar unter <https://doi.org/10.31263/voebm.v71i1.1974>
- Torggler, Andrea; Andrae, Magdalena (2018): Aus dem Leben einer/s Repman – ein Bericht aus dem österreichischen „Netzwerk für RepositorienmanagerInnen“. In: *Mitteilungen der VÖB*, 71 (1), 107–24. Online verfügbar unter <https://doi.org/10.31263/voebm.v71i1.1986>



**Dr. Susanne Blumesberger, MSc**

Leitung Abteilung Repositorienmanagement PHAIDRA-Services  
Bibliotheks- und Archivwesen  
Universitätsbibliothek  
Universitätsring 1  
A-1010 Wien  
Österreich  
**[susanne.blumesberger@univie.ac.at](mailto:susanne.blumesberger@univie.ac.at)**