

■ STRATEGISCHE OPTIONEN FÜR UNIVERSITÄRE REPOSITORIEN IN DEN DIGITAL HUMANITIES

von *Wolfgang Schmale*

Zusammenfassung: Ein ideales Repository bedient alle Interessen der unterschiedlichen beteiligten AkteurlInnen und ihrer Institution, es entspricht den Erfordernissen konkreter Fächer und Fächergruppen und unterstützt Strategien, die die Wissenschaft insgesamt und langfristig betreffen. Es lässt die für den User unkomplizierte Archivierung sämtlicher digitaler Objekte inklusive Datenbanken zu, spiegelt innovative Forschungsfelder der Universität wider und unterstützt insbesondere auch die Lehre. Es bietet Tools und Rechercheoptionen im Sinne von Big Data an.

Schlüsselwörter: Strategie für Repositorien; Datenbankenarchivierung; Big Data; Publikationsstrategie; Digital Humanities

STRATEGIC OPTIONS FOR UNIVERSITY REPOSITORIES IN THE DIGITAL HUMANITIES

Abstract: An ideal repository serves all the interests of the different actors involved and their institution, it meets the requirements of specific subjects and subject groups and supports strategies that affect scholarship in general and in the long term. It allows easy archiving of all digital objects including databases to the user. It reflects innovative fields of research at the university and in particular also supports teaching. It offers tools and research options in the sense of Big Data.

Keywords: Repository strategy; archiving of databases; big data; publication strategy; Digital Humanities

DOI: <https://doi.org/10.31263/voebm.v71i1.1966>



Dieses Werk ist lizenziert unter einer

Creative-Commons-Lizenz Namensnennung 4.0 International

Inhalt

1. Einleitung
2. Das ideale Repositorium
3. Akteursgruppe der WissenschaftlerInnen
4. Die Universität, die das Repositorium anbietet, als Akteur
5. Anforderungen an eine Strategie
6. Besondere Anforderungen der Digital Humanities
7. „Studienliteratur“ im Repositorium?

1. Einleitung

Der folgende Beitrag ist aus der Perspektive des Forschers an der Universität Wien geschrieben, der die von der Uni aufgebauten Repositorien (insbesondere [Phaidra](#) und [u:scholar](#)) aktiv und passiv nützen soll und auch will.¹ Als Nutzer stelle ich mir daher gewisse Dinge vor, die den Umgang mit einem Repositorium zeitökonomischer gestalten würden – und Zeitökonomie gehört zur Strategie –, die aber einen technischen Aufwand erfordern würden, der nicht ohne weiteres zu machen ist. Um Letzteres kümmere ich mich trotzdem nicht, sondern fordere es einfach ein, ich spreche also manches bewusst einseitig aus. Unterm Strich steht und fällt jede Strategiediskussion damit, ob der erforderliche personelle, finanzielle und technische Aufwand von der Universität gebilligt wird oder nicht. Vorab kann man aber wohl festhalten, dass die Uni Wien hier mehr investieren sollte.

2. Das ideale Repositorium

Ein ideales Repositorium bedient alle Interessen der unterschiedlichen beteiligten AkteurInnen und ihrer Institution, es entspricht den Erfordernissen konkreter Fächer und Fächergruppen, es unterstützt Strategien, die die Wissenschaft insgesamt und langfristig betreffen, die Nutzung ist passiv wie aktiv einfach und schnell, die Suchfunktion „erweiterte Suche“ ist breit ausdifferenziert und es werden Arbeitstools angeboten, mit denen mindestens einfache statistische Auswertungen vorgenommen werden können. Das ideale Repositorium funktioniert immer perfekt, egal, ob es sich bei den digitalen Objekten um Texte, Bilder, Video-, Audio- bzw. audiovisuelle Dateien oder Datenbanken handelt, wie sie häufig in Forschungsprojekten entstehen.

Zu jedem einzelnen der im Folgenden angesprochenen Aspekte kann es u.U. besondere rechtliche Erfordernisse geben. Hierauf einzugehen, würde aber zu weit führen.

3. Akteursgruppe der WissenschaftlerInnen

Zunächst soll ein Blick auf die Akteursgruppe der WissenschaftlerInnen geworfen werden. Auszugehen ist davon, dass Forschungsdokumentation in der Regel ein (gesetzliches) Muss ist. Deshalb stellt die Eingabe der entsprechenden Daten in das universitäre Forschungsdokumentationssystem den zentralen Ausgangspunkt dar. Hier sollte aus Sicht der individuellen WissenschaftlerInnen alles auf einmal erledigt werden können: Eingabe der erforderlichen Daten für die Forschungsdokumentation, Hochladen der für das Repositorium geeigneten bzw. vorgesehenen digitalen Objekte, zusätzlich, falls es ansteht, die Möglichkeit zur Langzeitarchivierung, soweit diese sich technisch vom Depot im Repositorium unterscheidet (beispielsweise Datenbanken).

Aus Sicht der WissenschaftlerInnen ist es problematisch, dass das Repositorium der Institution nur so lange aktiv befüllt werden kann, solange ein Arbeitsverhältnis besteht. Für WissenschaftlerInnen müssen Repositorien aber zur Schärfung und Vertiefung des eigenen wissenschaftlichen Profils beitragen, das heißt sie müssen den Forschungsoutput möglichst umfassend dokumentieren, und zwar unabhängig davon, wie lange und wo sie an einer Institution arbeiten. Deshalb nutzen viele lieber allgemeine akademische Repositorien, die nicht an eine bestimmte Institution gebunden sind, oder unterhalten eine eigene institutionenunabhängige Homepage, die zugleich als Repositorium dient und wo sie die alleinige und volle Entscheidungsgewalt haben. Überspitzt formuliert: Universitäre Repositorien haben ggf. klare Wettbewerbsnachteile vom Start weg, die durch anderweitige Attraktivität ausgeglichen werden müssen.

Noch problematischer ist jedoch, dass institutionelle Repositorien meistens keine Lösung für Datenbanken anbieten, die in einem Forschungsprojekt entstanden sind, aber nicht mehr aktiv verändert werden, da dieses mittlerweile abgeschlossen wurde und keine Personalressourcen mehr zur Verfügung stehen. Gerade dieser Aspekt wäre aber für die Digital Humanities von besonderem Interesse.

Hinzu kommt, dass sich ein universitäres Repositorium an eine allgemeine und internationale wissenschaftliche Öffentlichkeit richtet. Wer aber in Princeton, Oxford oder an der Sorbonne sollte überhaupt u: scholar

oder Phaidra aus welchen Gründen benutzen? Niemandem geht etwas ab, wenn er es nicht tut. Das muss sich ändern.

4. Die Universität, die das Repositorium anbietet, als Akteur

Wenden wir uns kurz dem anderen Akteur zu, der Institution, die das Repositorium betreibt. Für die Universität (das Forschungsinstitut oder auch Museum etc.) ist eine umfassende Dokumentation der Forschungsleistung, die die Zugänglichmachung der Forschungsergebnisse in Gestalt von Publikationen, Datenbanken etc. einschließt, mittlerweile ein Muss, und zwar ganz unabhängig davon, ob allfällige Gesetze das vorschreiben oder nicht.

Die Zugänglichmachung der Forschungsergebnisse selber, die über die Forschungsdokumentation im engeren Sinn hinausgeht, soll möglichst dem Open Access-Ideal folgen. Dem stehen bekanntermaßen viele rechtliche Hürden entgegen, die kommerzielle Interessen schützen. Lediglich in Bezug auf Zeitschriftenaufsätze und Beiträge zu Sammelbänden hat die österreichische Urheberrechtsnovelle von 2015 eine Erleichterung erbracht.

Ein ideales Repositorium wäre für die ForscherInnen so attraktiv, dass sie ihre Forschungsergebnisse vorzugsweise und original im Repositorium ihrer Universität etc. veröffentlichen würden.

Das geschieht aber so gut wie nicht, weil die in den Fächern üblichen und angesehenen Publikationsorgane attraktiver sind und Reputation garantieren. Im günstigsten Fall sind Repositorien zweite Wahl, vielfach finden sich dort aber in erster Linie Vorstadien der späteren und endgültigen Veröffentlichung, also Konferenzpaper, Vortragsmanuskripte usw.

Auch wenn man eine der größten europäischen Universitäten ist und ein insgesamt hohes Ansehen genießt, wie die Universität Wien, besteht hier ein schwer lösbares Dilemma, da gegen die Attraktivität der bewährten Fachorgane für Publikationen nicht anzukommen ist. Allerdings ist das auch nicht das, worauf es im Augenblick ankommt.

5. Anforderungen an eine Strategie

Erstens sollte es an erster Stelle darum gehen, den LeiterInnen von Forschungsprojekten eine Lösung für ihre im Projekt entstandenen Datenbanken anzubieten. Die Zusammenarbeit muss dabei zu einem frühestmöglichen Zeitpunkt beginnen, im Grunde vor Beginn des Projekts bzw. schon bei der Antragstellung, wenn es sich um ein Drittmittelprojekt handelt. Bei

Datenbanken in den Digital Humanities besteht weniger Konkurrenz durch angesehene Zeitschriften, das heißt, hier hat die Universität die Chance, mit dem Repositorium bzw. auch der Langzeitarchivierung einen echten Zuzugewinn in Gestalt von Reputation für die WissenschaftlerInnen anzubieten.

An und für sich wäre es wünschenswert, dass Forschungsprojektdatenbanken möglichst lange aktiv fortgeführt werden, aber aus Gründen, die allen bekannt sind, ist das nur selten möglich. Deshalb muss von Beginn an geplant werden, in welcher Weise die betreffende Datenbank im Repositorium faktisch archiviert werden kann. Wichtig ist, die Funktionalität zu erhalten. Das reine Archivieren der Forschungsdaten wäre nicht das Problem, aber der Mehrwert der Datenbank ginge verloren, wenn ihre Funktionalität nicht erhalten bliebe.

Gäbe es für wenigstens die einzelnen Fächer Datenbankstandards, die in den allermeisten Fällen den Zweck erfüllen würden, hätte man es mit einer überschaubaren Zahl von Systemen zu tun, für die sich die Institution zu einer Wartung über einen längeren Zeitraum entschließen könnte. Auch die Migration der Daten in neue Systeme oder Systemversionen könnte vorgesehen werden, aber dies und anderes setzt voraus, dass jemand vom ursprünglichen Projekt weiterhin zur Verfügung steht, um Fehler zu erkennen.

Allein realistisch ist daher ein Angebot zur Archivierung von Datenbanken mit Erhaltung deren Funktionalität. Allerdings dürfte es sich lohnen, die ForscherInnen von vorneherein, wenn sich die Frage der Anlage einer Datenbank in einem Forschungsprojekt stellt, entsprechend zu beraten und zu begleiten. Denn in nicht wenigen Fällen würde mittelfristig wahrscheinlich die Archivierung der eigentlichen Forschungsdaten ausreichen, da die Verknüpfungen der Daten miteinander, die die Datenbank leistet, von Fragestellungen abhängig sind, die sich mit der Zeit wandeln können.

Zweitens gilt: Qualität vor Quantität. Dieser Grundsatz erfordert eine Moderation des Repositoriums nach bestimmten Gesichtspunkten. Das bedeutet nicht zwangsläufig, dass jemand dasitzt und einen Annehmen- oder Ablehnen-Schalter betätigt.

Die Attraktivität von Repositorien kann erhöht werden, wenn die digitalen Objekte nach Fächerportalen bzw. Fächergruppenportalen gegliedert angeboten werden, um Fachprofile abzubilden. Natürlich kann man überall z. B. im Suchfeld „Thema“ den Namen eines Faches eingeben, aber es ist wirkungsvoller, zusätzlich ausgesprochene Fachportale als Binnenstrukturierung des Repositoriums anzubieten, weil das die aktiven NutzerInnen besser motiviert, die erforderlichen Metadaten, in dem Fall die Zuordnung zu einer oder mehreren Fachdisziplinen, einzugeben und weil es den Nutzungsgewohnheiten entgegenkommt.

Eine solche Oberflächenstruktur kann genutzt werden, um innovative Fächer oder auch Forschungsfelder, die sich an der Universität oder der Institution gebildet haben, besser sichtbar zu machen. Digital Humanities wäre ein gutes Beispiel dafür. Zwar sind diese auch nicht mehr blutjung, aber immer noch ein Innovationstreiber, der das Repositorium jeder Universität schmücken würde.

Darüber hinaus könnten Zuordnungen zu Forschungsfeldern oder -schwerpunkten, die zwar kleiner als ein einzelnes Fach, oft aber inter- bzw. transdisziplinär charakterisiert sind, einen beachtlichen Mehrwert ergeben, zumal sich auch darin Forschungsinnovationen einer Institution zeigen. Außerdem existieren viele Querschnittsmaterien, die in der Oberflächenstruktur sichtbar sein sollten. Mit Bezug auf die Uni Wien lassen sich benennen: Forschungsschwerpunkte in den Fakultäten, Forschungs-Plattformen auf Uni-Ebene, oder faktisch bestehende Querschnittsmaterien wie „18th-Century Studies“, die durch keine der Uni-Strukturen direkt abgebildet werden, aber dennoch bestehen und z. B. über eine wissenschaftliche Gesellschaft, die aber keine universitäre Einrichtung ist, organisiert werden, etc.

Dies wird unter das Stichwort von „Qualität vor Quantität“ gesetzt, weil es nur dann funktioniert, wenn die WissenschaftlerInnen die erforderliche Menge an Metadaten eingeben, die es braucht, um Fach- bzw. Fächergruppenportale, Forschungsfelder oder Querschnittsmaterien auf einen Klick hin zu generieren. Es wird ihnen ein nicht unerheblicher Zeitaufwand zugemutet, selbst wenn das Forschungsdokumentationsportal (Uni Wien u:cris) so eingerichtet ist, dass dort sämtliche Eingaben inklusive Hochladen des digitalen Objekts (Datenbanken bleiben sicher ein Sonderfall) in einem einzigen Arbeitsgang erledigt werden könnten und die jeweils zeitökonomischsten Eingabe- und Auswahlverfahren angeboten würden.

Die in u:cris eingebauten Felder nach ÖFOS und ASJC Scopus reichen nicht aus, weder für die Disziplinen noch für Forschungsfelder – schon allein hier entsteht ein Mehraufwand, sei es, dass eine weitere Auswahlliste angeboten wird, sei es, dass es frei zu ergänzende Felder gibt. Frei ausfüllbar sind schon jetzt die Schlagwortfelder, aber nicht jedes Schlagwort entspricht einem Forschungsfeld oder einer Querschnittsmaterie, insoweit sind diese tatsächlich eine eigenständige Kategorie.

Schließlich sollte es ein Angebot geben, das den tatsächlichen Publikationsprozessen im Zusammenhang mit einer laufenden Forschung entspricht und diese abbilden kann. Es wäre sowohl bequem als auch hochattraktiv für ForscherInnen, wenn folgendes Angebot bestünde: Jemand trägt in u:cris die erforderlichen Daten ein und lädt ein Paper hoch. Diese markieren oft den Anfang eines Prozesses, der über weitere Publikationen schlus-

sendlich bis hin zu einem Buch zum Thema des anfänglichen Papers führt. Dieser Prozess sollte automatisch, also ohne dass ein aktives Eingreifen der betreffenden UrheberInnen erforderlich wäre, erfasst werden, sodass das Follow-up allfälliger digitaler Objekte bis hin zur schlussendlichen Monographie und deren Überführung ins Repositorium, soweit rechtlich zulässig, automatisiert geschieht. Natürlich müssen UrheberInnen die Möglichkeit haben, die Automatisierung abschalten zu können, aber in den meisten Fällen wird diese Art von Service gerne angenommen. Dazu müssen in der Forschungsdokumentation mehr Metadaten eingegeben werden, damit das System die Zusammengehörigkeit digitaler Objekte zum selben Forschungsfeld bzw. Forschungsvorgang automatisch erkennen kann. Als praktisches Beispiel: In u:cris kann ich meinen neuen Eintrag mit früheren Einträgen unter „Verknüpfungen“ verbinden, aber ich muss es händisch machen, und wenn ich viel publiziere, ist die Liste lang und es kann dauern, bis man zur passenden Verknüpfung gelangt ist, auch wenn das Eingabefeld durch das Erkennen schon vorhandener Inhalte bei der Abkürzung hilft. Dennoch führt der Gesamtaufwand nur zur Abwägung, ob die Verknüpfung einen solchen Mehrwert erbringt, dass sich der Mehraufwand an Zeit lohnt, oder eben nicht. Dann spart man sich die Zeit lieber.

Als Zwischenbilanz: Ein institutionelles Repositorium, wie das der Uni Wien, könnte ein starkes Profil erhalten, wenn es Datenbanken und aktuelle Forschungsfelder sowie Querschnittsmaterien zu einem Schwerpunkt machen würde und zusätzlich weitere disziplinäre Zugänge zum Material anbieten würde.

6. Besondere Anforderungen der Digital Humanities

Haben die Digital Humanities spezielle Anforderungen? An der Uni Wien erhalten diese aktuell ein eigenes Profil, sie könnten also einen sinnvollen Schwerpunkt für das Repositorium bilden und dieses interessant machen. Datenbanken war bereits ein Stichwort und stellt diesbezüglich einen vorrangigen Aspekt dar. Das Repositorium muss für Datenbanken eingerichtet werden.

Dass Digital Humanities automatisch papierlos sind, kann man nicht sagen, trotzdem ist das Arbeiten mit und das Produzieren von digitalen Objekten aller Art vorrangig und stellt ein wichtiges Merkmal der DH dar, ohne exklusiv zu sein. In den DH haben in den vergangenen Jahren wissenschaftliche Blogs geradezu explosionsartig zugenommen und sich einen fixen Platz unter den Publikationsoptionen sowie Publikationsstrategien er-

obert. Ein universitäres Repository sollte diese durchaus fundamentale Veränderung wissenschaftlicher Publikationsgewohnheiten aufnehmen.

Der Mehrwert der DH zeigt sich auch in der Nutzung umfassender digitaler Forschungscorpora. Das Arbeiten mit Big Data wird in diesen Fächern immer wichtiger. Repositorien, gerade einer sehr großen Universität wie der Uni Wien, werden in dieser Beziehung als Datenpool interessant. ForscherInnen sollten daher die im Repository enthaltenen digitalen Objekte ihren individuellen Forschungsfragen entsprechend zu einem Forschungskorpus zusammenstellen können, das für sich durchsucht und ausgewertet werden kann. In dieser Beziehung gilt sozusagen Quantität vor Qualität, je mehr Objekte im Repository sind, desto größer können die individuellen Forschungscorpora werden, desto mehr bewegt man sich in Richtung Big Data als Forschungsgrundlage. Auch wenn es sich derzeit meistens bei den digitalen Objekten in den Fächern, die den DH zugrunde liegen, um Texte handelt, sind diese wesentlich mehr, Bild- und Audio-dateien oder eben gleich multimediale Datenbanken. Bei Bild- sowie Audioobjekten wäre es ein erheblicher Mehrwert, wenn die ForscherInnen auf die Arbeit von Algorithmen zurückgreifen könnten, die deren Inhalte automatisch erkennen und taggen.

Ein Repository zu einem Big Data-Pool zu entwickeln, erscheint unter den strategischen Optionen besonders wichtig, zumal dies wohl auch die Chance erhöht, dass ForscherInnen ihre digitalen Objekte ins Repository einspeisen. Deshalb gehört es zu dieser Strategie dazu, die ForscherInnen zu motivieren, *systematisch* das Repository neben dem Originalveröffentlichungsort als sekundären Veröffentlichungsort zu verwenden. Wenn im Zuge dessen digitale Objekte, die originär keinen DOI haben, wie in der Regel bei Einträgen in wissenschaftlichen Blogs, diese im Repository zugeteilt bekommen, entsteht eine zusätzliche Motivation.

7. „Studienliteratur“ im Repository?

Was die mit einer Veröffentlichung im Repository verbundene oder eben gerade nicht damit verbundene Reputation für die WissenschaftlerInnen angeht, stellt sich die Frage, warum die Uni Wien das Label „University Press“ nicht selber als Institution betreibt, vorzugsweise digital. Hier würden sich ambitionierte Ziele für das universitäre Repository mit einer Universitätsstrategie für die Veröffentlichung von Forschungsergebnissen verknüpfen lassen, die an ihr und mindestens teilweise mit ihren Ressourcen und Infrastrukturen erzielt wurden.

Abgesehen von Forschungsergebnissen entsteht an einer Universität auch ein breites Angebot an, sehr allgemein formuliert, Studienliteratur. Diese findet ihren Weg in Studienliteraturreihen wie UTB und andere und kommt dann auf diesem oder jenen Weg auszugsweise in die Lernplattform zurück, soweit dem nicht verwertungsrechtliche Aspekte entgegenstehen. In Europa sind derzeit national unterschiedliche Urheberrechtsnovellen im Gange, die dieses „Modell“ ohnehin verändern werden.

Wenn man einmal die Mehrzahl der Gründe, die bis zum bisherigen Modell geführt haben, außer Acht lässt und die Sache rein rational und ökonomisch betrachtet, wäre es logisch, das Repositorium zu einem Ort frei zugänglicher, regelmäßig aktualisierter Studienliteratur zu machen. Ob Open Access nur für die Lehrenden und Studierenden der Uni Wien gewährt wird oder mehr, wäre zu diskutieren. Natürlich zerstört das ein Geschäftsmodell, das ausnahmsweise sogar für geisteswissenschaftliche AutorInnen begrenzt geldwerte Vorteile mit sich bringt. Auch darüber wäre zu diskutieren, zumal für WissenschaftlerInnen an der Uni zwar gewiss eine Pflicht zur Forschung besteht, nicht aber eine Pflicht, Studienliteratur herzustellen. Forschen müssen alle können, wenn sie als WissenschaftlerInnen beschäftigt werden wollen, Studienliteratur aber sollten nur die verfassen, die so etwas gut können. Das zusätzlich zu bezahlen, ist sehr wohl vertretbar, auch für eine Uni, und würde im Übrigen dem Leistungswettbewerb im Betrieb dienlich sein.

Da Studienliteratur nicht simpel als „Buch“ zu übersetzen ist, sondern am besten medial hybrid ausgestaltet wird, wäre die digitale Publikation das Mittel der Wahl. Kurz und gut: Das stellt eine weitere Entwicklungsperspektive für Repositorien dar. Was es dazu braucht, auf allen Seiten, ist, die Vorstellung abzulegen, Repositorien seien dafür da, dort nur etwas abzulegen, sprich vor allem zu archivieren. Die Vorstellung muss also hybridisiert werden, und die Uni Wien gibt ja zumindest vor, das zu tun, wenn sie zu u:scholar ausführt, dies sei sowohl als „Publikationsservice“ wie auch als „elektronische Archivierung wissenschaftlicher Dokumente und Materialien“ wie auch als „institutionelles Repositorium“ wie auch als Teil der „Forschungsdokumentation“ der Uni Wien gedacht.

o. Univ.-Prof. Dr. Wolfgang Schmale
Universität Wien, Institut für Geschichte
E-Mail: wolfgang.schmale@univie.ac.at

1 Eine erste Version des Artikels wurde veröffentlicht unter: <https://wolfgang.schmale.eu/repositorien/>.