

# Archivieren – aber wie?

Eine neugebildete Arbeitsgruppe hilft Forschenden beim Management ihrer Daten.



Christian Fuhrer, Andrea Malits und Peter Kunszt (v. l.) wissen, wie Forschungsdaten zu verwalten sind.

Marita Fuchs

Forschungsdaten – sie sind der Treibstoff der Hochschulen. Der online erhältliche wissenschaftliche Ausstoss vermehrt sich so schnell wie nie: Alle neun Jahre verdoppelt er sich. Angesichts dieser Datenvielfalt wünschen sich viele Forschende eine langfristige und sorgfältige Archivierung ausgewählter Forschungsdaten, wie die Konferenz der Universitätsbibliotheken der Schweiz vor einiger Zeit festgestellt hat.

Doch wie dabei vorgehen? Zurzeit liegt es meist bei den Forschungsgruppen selber, wie sie ihre Daten archivieren. Sie entscheiden, welche Daten in den Papierkorb wandern und welche aufbewahrt werden sollen. Eine definierte Vorgehensweise gibt es nicht. Das hat viele Nachteile.

## Konsistente Vorgehensweise gesucht

Nehmen wir den Fall eines kurz vor der Emeritierung stehenden Professors. Er ist Geisteswissenschaftler, hat mit seinem Team viele Quellen ausgewertet, gleichzeitig auch selbst Daten erhoben und dokumentiert. Teilweise sind diese Daten für wissenschaftliche Publikationen verwendet worden, doch sein Material geht weit darüber hinaus und könnte für neue Studien genutzt werden.

Abgelegt sind alle Daten auf einem Institutsserver. Die Zuständigkeit für regelmässige Backups liegt beim Informatikverantwortlichen des Instituts oder bei den Mitarbeitenden des Professors. Was passiert, wenn der Gelehrte nicht mehr an der Universität arbeitet? Wie lange werden die Daten aufbewahrt, und wer kann und darf auf sie zugreifen? Gibt es ein Konzept, wie in solchen Fällen vorgegangen wird?

All diese Fragen will eine Arbeitsgruppe aus Vertretern der Zentralbibliothek Zürich, der Hauptbibliothek und der Zentralen Informatik der UZH anpacken. Federführend sind Andrea Malits von der Zentralbibliothek, Christian Fuhrer, Leitung Open Access an der Hauptbibliothek, und Peter Kunszt von der Zentralen Infor-

matik der UZH. Sie arbeiten im Auftrag des Prorektorats Rechts- und Wirtschaftswissenschaften sowie der Zentral- und der Hauptbibliothek der UZH. Gemeinsam wollen sie einen Vorschlag erarbeiten für eine konsistente Vorgehensweise bei der Verarbeitung, Archivierung und Publikation von Forschungsdaten. Sie liefern damit die Grundlage für ein umfassendes Forschungsdatenmanagement an der UZH.

## Daten für den Review-Prozess

Viele europäische Hochschulen arbeiten an neuen Archivierungssystemen. Handlungsbedarf besteht auch deshalb, weil immer mehr wissenschaftliche Journale Datenanhänge («supplemental data») oder ganze Datensets nicht mehr selbst publizieren wollen. Sie erwarten von den Forschenden, dass sie ihre Daten in sogenannten Repositories, das sind vertrauenswürdige und nachhaltige Datenserver, zugänglich machen. «Die Journale wollen auch für den Review-Prozess zunehmend Zugriff auf die Originaldaten, und zwar in aufbereiteter Form», sagt Christian Fuhrer.

Doch nicht nur die Journale, auch Förderungsinstitutionen sind zunehmend daran interessiert, dass die Datenspeicherung gesichert ist. So etwa beim EU-Flaggschiff «Horizon 2020»: Wenn Forscher, deren Fachgebiet dort unter den «pilot on open research data» fällt, ihre Anträge einreichen, müssen sie angeben, wie viele Daten in etwa anfallen, wie sie zwischengespeichert und wo sie schliesslich archiviert und zugänglich gemacht werden.

Das kann mitentscheidend sein, ob ein Antrag akzeptiert wird oder nicht. Auch der Schweizerische Nationalfonds arbeitet an einem neuen Forderungskatalog zur Datenspeicherung.

## Was Primär- von Sekundärdaten scheidet

Forschungsdaten sind je nach Fachgebiet sehr unterschiedlich. Grob lassen sie sich jedoch in primäre und sekundäre Daten einteilen. Primärdaten, die etwa beim Ex-

perimentieren und Messen, aus Quellenmaterial oder Umfragen entstehen, werden auf unterschiedliche Arten aufbereitet. Es können digitale Tabellen, Bilder, Filme, Ergebnisse einer Umfrage oder errechnete Resultate sein. Nach Analyse der Primärdaten entstehen nach deren Auswertung sogenannte Sekundärdaten.

## Datenmanagement für Forschende

Besonders wichtig für eine spätere Archivierung ist die Beschreibung aller Datensätze durch Metadaten, die die Daten nach bestimmten Kriterien strukturieren. Denn erst die Metadaten erlauben ein schnelles Wiederauffinden. Hinzu kommen die Formatauswahl und die Anonymisierung personenbezogener Daten sowie die Beschreibung der experimentellen Protokolle und Workflows, mit denen die Daten erzeugt und verwertet wurden.

Nur so können andere Forschende die Daten reproduzieren. Das gilt für alle Fachbereiche, auch für die Geisteswissenschaften, die häufig mit digitalisierten Quellen arbeiten. Doch mehrere Fragen stellen sich, sobald ein Forschungsteam sich entscheidet, die Daten zu archivieren.

Was muss archiviert werden? Und wie muss man die Daten speichern, damit auch andere darauf zugreifen können? Welche Daten sollen offen publiziert werden (Open Access zu Forschungsdaten, Open Data)? Managementpläne zu Forschungsdaten sollen Forschende dazu anregen, sich mit den Fragen auseinanderzusetzen.

Bei dieser Aufgabe will das Team Andrea Malits, Christian Fuhrer und Peter Kunszt die Forschenden unterstützen. Ihr Ziel ist es, ab 2016 eine Service-Infrastruktur für Forschende aufzubauen. «Wir helfen bei der Erstellung von Datenmanagementplänen», sagt Peter Kunszt. Dabei werden vor Beginn der Forschungsarbeiten Schlüsselprobleme identifiziert, die bei der Arbeit mit den Daten entstehen können. Auch sollen die Forschenden während des Forschungsprozesses unterstützt werden. «Manchmal gibt es neue Entwicklungen oder Wendepunkte in der Forschung, denen man gerecht werden muss», weiss Kunszt. Für die Erstspeicherung grosser Datensätze sei die UZH gerüstet, sagt der Informatiker. Seine 2014 eingerichtete Abteilung stellt die Soft- und Hardware zur Verfügung.

## Ungelöste Fragen

Die bisherigen Kontakte der Arbeitsgruppe haben gezeigt, dass einige der Kernfragen im Umgang mit Forschungsdaten an der UZH ungelöst sind. Offen ist vor allem die langfristige Archivierung von Forschungsdaten, aber auch die Veröffentlichung von Daten, wenn dazu kein geeignetes Fachrepository existiert, und die nachhaltige Pflege von Datenbanken, die Forschende selber angelegt haben.

Insgesamt resultiere aus dem Datenmanagement und den Archivierungsvorgaben für die Forschenden zunächst etwas mehr Aufwand, aber mit der Zeit werde sich ein greifbarer Mehrwert einstellen, sagt Andrea Malits.

## APPLAUS

**Martina Kohle**, Institut für Tierernährung der Vetsuisse-Fakultät, wurde für ihre Dissertationsarbeit zum Thema «Einfluss der Alpengasse auf den Vitamin-D- und Knochenstoffwechsel laktierender Schafe und Ziegen» von der H.-Wilhelm-Schaumann-Stiftung zur Förderung der Agrarwissenschaften in Hamburg für die beste Dissertation auf dem Gebiet der Tierernährung ausgezeichnet.

**Bea Latal**, Leitende Ärztin am Forschungszentrum für das Kind am Kinderspital Zürich, und **Magdalini Polymenidou**, SNF-Assistenzprofessorin am Institut für Molekulare Biologie, haben den renommierten Georg-Friedrich-Götz-Preis 2014 erhalten. Der Preis wird jährlich für besondere Leistungen in der medizinischen Forschung vergeben und ist mit insgesamt 30 000 Franken dotiert.

**Katrin Preller**, Postdoc an der Psychiatrischen Universitätsklinik Zürich, ist für ihre Forschung über die sozialen Auswirkungen des Kokainkonsums mit dem Pfizer-Forschungspreis im Bereich Neurowissenschaften und Erkrankungen des Nervensystems ausgezeichnet worden.

**Daniel Süss**, Extraordinarius ad personam für Publizistikwissenschaft, wurde vom Bundesrat als Mitglied in die Eidgenössische Kommission für Tabakprävention gewählt.

**Alexandre Theocharides**, Hämatologe am Universitätsspital Zürich, erhält für seine Forschung zu Blutstammzellerkrankungen den mit 20 000 Franken dotierten Ellermann-Preis Hämatologie 2014. Er teilt die Auszeichnung mit **Johanna A. Kremer Hovinga** von der Abteilung für Hämatologie am Inselspital Bern. Die Ellermann-Stiftung Bern zeichnet mit dem Preis Wissenschaftlerinnen und Wissenschaftler für bedeutende klinische oder experimentelle Arbeiten auf dem Gebiet der Hämatologie und deren Grenzgebieten aus.

**Rolf Weber**, Ordinarius für Privat-, Wirtschafts- und Europarecht, wurde vom International Center for Trade and Sustainable Development (ICTSD) und dem World Economic Forum in die E15 Expert Group on The Digital Economy berufen.

**Gregor Zünd**, ausserordentlicher Professor für Chirurgie, wurde an der Generalversammlung der Swiss Clinical Trial Organisation für die Amtsdauer von 2015 bis 2018 zum Präsidenten gewählt.

## Auszeit im Denklabor



Das Denklabor Villa Garbald im bündnerischen Castasegna (Bergell), erbaut von Gottfried Semper, hat im vergangenen Jahr sein 150-jähriges Bestehen gefeiert. Das historische Gebäude und der Neubau der Architekten Miller & Maranta sind dank der ruhigen Lage und der grosszügigen Gartenanlage ein idealer Rückzugsort für Forschende und Seminargruppen. Angehörigen der Universität Zürich steht die Villa zu Vorzugsbedingungen zur Verfügung. Gewährt wird ein Preisnachlass von bis zu 30 Prozent.

Reservation und weitere Informationen unter: [www.garbald.ch](http://www.garbald.ch)