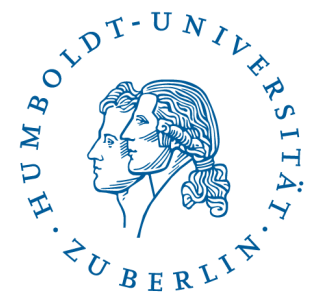


Forschungsdaten an der HU

*Eine gemeinsame Sitzung mit und für Kolleginnen
und Kollegen der Universitätsbibliothek*

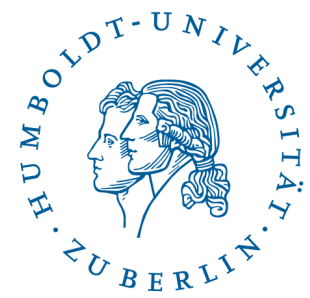
Elena Simukovic, Paul Vierkant, Dennis Zielke

12. September 2013



Ablauf

- › Einführung und Umfrage an der HU
(Elena Simukovic)
- › re3data.org - Register von Forschungsdaten-Repositoryen
(Paul Vierkant)
- › LAUDATIO - Repository für historische Korpora
(Dennis Zielke)
- › Diskussion



Einführung in das Thema und Umfrage an der HU

Elena Simukovic

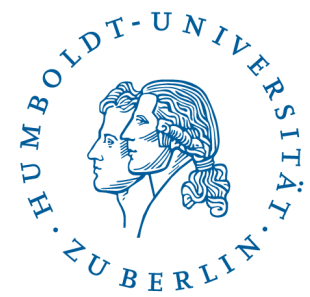
Forschungsdatenmanagement-Koordinatorin



Grundlegendes

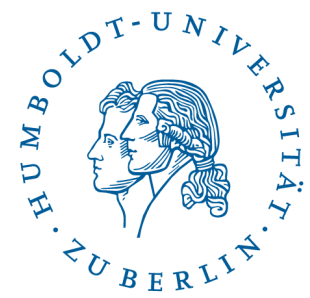
- **Forschungsdaten** – (eng. research data) – alle Daten, die im Laufe einer wissenschaftlichen Tätigkeit entstehen. Diese können sowohl qualitativer als auch quantitativer Art / Natur sein, sowie in einer digitalen und analogen Form vorliegen. Forschungsdaten stellen in der Regel die Grundlage einer wissenschaftlichen Publikation dar, können aber auch eigenständig publiziert werden.
- **Forschungsdatenmanagement** – (eng. research data management) – alle Aktivitäten, die mit der Aufbereitung, Speicherung, Archivierung und Veröffentlichung von Forschungsdaten verbunden sind.

Quelle: Simukovic, Elena; Kindling, Maxi; Schirnbacher, Peter (2013): Umfrage zum Umgang mit digitalen Forschungsdaten an der Humboldt-Universität zu Berlin. Umfragebericht (erscheint).



Wichtige Vorbemerkungen

- Forschungsdaten sind *kein neues* Phänomen
- Digital > analog
- Technik wird immer leistungsfähiger, Daten immer umfangreicher und komplexer
- Thema gewinnt an Bedeutung im Kontext der Open-Access-Bewegung
- Höhere Aufmerksamkeit auf nationaler und internationaler Ebene
- Forschungsdaten sind wesentlicher Bestandteil der Forschungsergebnisse
- Viele Akteure einbezogen: Forschungseinrichtungen (Hochschulen und außeruniversitäre), Förderorganisationen, Herausgeber von wissenschaftlichen Publikationen, Fachgesellschaften, Service-Einrichtungen, Repositorien-Betreiber, Wissenschaftspolitik



Beweggründe

- Forschungsdaten ermöglichen Nachprüfbarkeit (verify) und Nachnutzbarkeit (re-use) der Forschungsergebnisse
- Öffentliche Fördermittel → öffentlich zugängliche Forschungsergebnisse
- Gute wissenschaftliche Praxis
- Zugang zu Forschungsdaten fördert neue Fragestellungen, neue Kollaborationen und den Fortschritt der Wissenschaft
- Anstrengungen in Aufbereitung und Analyse von Forschungsdaten können belohnt und Ansehen herausragender Wissenschaftler gestärkt werden

Gute wissenschaftliche Praxis

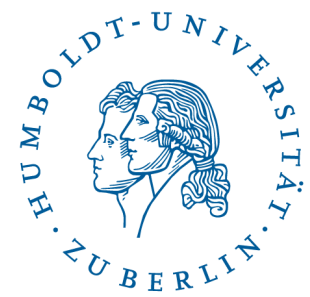
DFG

"Primärdaten als Grundlagen für Veröffentlichungen sollen auf haltbaren und gesicherten Trägern in der Institution, wo sie entstanden sind, für zehn Jahre aufbewahrt werden." (in: Empfehlungen zur Sicherung guter wissenschaftlicher Praxis (Empfehlung 7), 1998)

§ 5 Sicherung und Aufbewahrung von Primärdaten

- (1) Primärdaten als Grundlagen für Veröffentlichungen sollen auf haltbaren und gesicherten Trägern in der Arbeitsgruppe/ Einrichtung, wo sie entstanden sind, für zehn Jahre zugänglich bleiben.
- (2) Die Verantwortung für die Erstellung der Datenträger trägt der jeweilige Wissenschaftler oder die jeweilige Wissenschaftlerin; damit obliegt ihm bzw. ihr die Nachweispflicht für die ordnungsgemäße Protokollierung. Die Erstellung von Kopien ist dem jeweiligen Wissenschaftler oder der jeweiligen Wissenschaftlerin erlaubt.

Quelle: Satzung über die Grundsätze der Humboldt-Universität zu Berlin zur Sicherung guter wissenschaftlicher Praxis und über den Umgang mit Vorwürfen wissenschaftlichen Fehlverhaltens, S. 3-4 beschlossen vom Akademischen Senat am 25. Juni 2002. Online abrufbar unter: <http://www.amb.hu-berlin.de/2002/33/3320020>



Hintergründe

- **Deutsche Forschungsgemeinschaft (DFG):**

"Werden in dem geplanten Projekt systematisch Forschungsdaten erhoben, die für eine Nachnutzung geeignet sind, sollte der Antrag ein entsprechendes Konzept für die Überführung der Forschungsdaten in vorhandene Datenbanken oder Repositorien enthalten." (in: Leitfaden für die Antragstellung - Projektanträge (Punkt 2.4), 2010)

- **Allianz der deutschen Wissenschaftsorganisationen:**

"In Übereinstimmung mit wichtigen internationalen Organisationen auf dem Gebiet der Förderung und Durchführung von Forschungsaufgaben unterstützt die Allianz die langfristige Sicherung und den grundsätzlich offenen Zugang zu Daten aus öffentlich geförderter Forschung." (in: Grundsätze zum Umgang mit Forschungsdaten, 2010)

- **Kommission Zukunft der Informationsinfrastruktur (KII):**

"Forschungs-Daten sind als nationales Kulturgut anzusehen und sollten im Sinne einer öffentlichen Aufgabe dauerhaft gesichert und der (Fach-)Öffentlichkeit sowie zukünftigen (Forscher-)Generationen zur Nachnutzung zur Verfügung gestellt werden." (in: Gesamtkonzept für die Informationsinfrastruktur in Deutschland, 2011)

FDM-Initiative an der HU



Prof. Dr. Peter Schirmbacher / CMS



Prof. Dr. Peter A. Frensch / VPF



Dr. Ingmar Schmidt / Forschungsabteilung



Dr. Andreas Degkwitz / UB

- Eine dezidierte Stelle zum Forschungsdatenmanagement wurde auf Gemeinschaftsinitiative der Zentraleinrichtungen im Spätsommer 2012 geschaffen
- Federführung: Vizepräsident für Forschung
- Fachliche Begleitung: Institut für Bibliotheks- und Informationswissenschaft (IBI)

Bildquelle: Webseiten der HU

Kurze Rückschau

August
2012

- Forschungsdatenmanagement-Stelle besetzt
- Empfehlungsschreiben des VPF

Oktober
2012

- Projektseminar am IBI im Wintersemester 2012/13 gestartet

November
2012

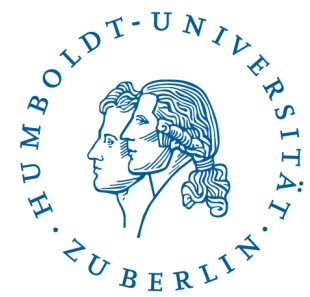
- Webseite der Initiative <http://www.cms.hu-berlin.de/dataman>
- Artikel in der Humboldt-Zeitung („[Das Erdöl von morgen](#)“)

Januar-
März 2013

- Umfrage an der HU

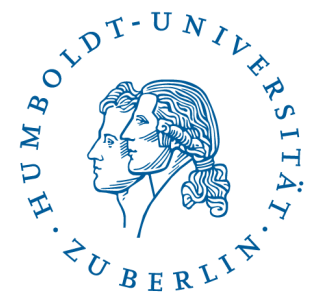
Laufend

- Präsentation der Ergebnisse (Vorträge beim Bibliothekartag, BBK, intern)
- Entwicklung der Policy zum Umgang mit digitalen Forschungsdaten an der HU + ergänzende Richtlinien (praktische Hinweise zur Umsetzung der Policy)
- Interviews mit Wissenschaftlern der HU
- Informationsveranstaltungen (z.B. Fachreferenten der UB, Open Access Week)
- Umsetzung der Umfrageergebnisse



Umfrage: Konzept & Ziel

- Baut auf vergleichbaren Umfragen an anderen Hochschulen (University of Cambridge, University of Glasgow, Imperial College London, ETH Zürich) und im PARSE.Insight-Projekt auf
- Ziel: Ermittlung des Ist-Zustands und der Anforderungen an zukünftige Serviceangebote
- Zielgruppe: wissenschaftliches Personal der HU



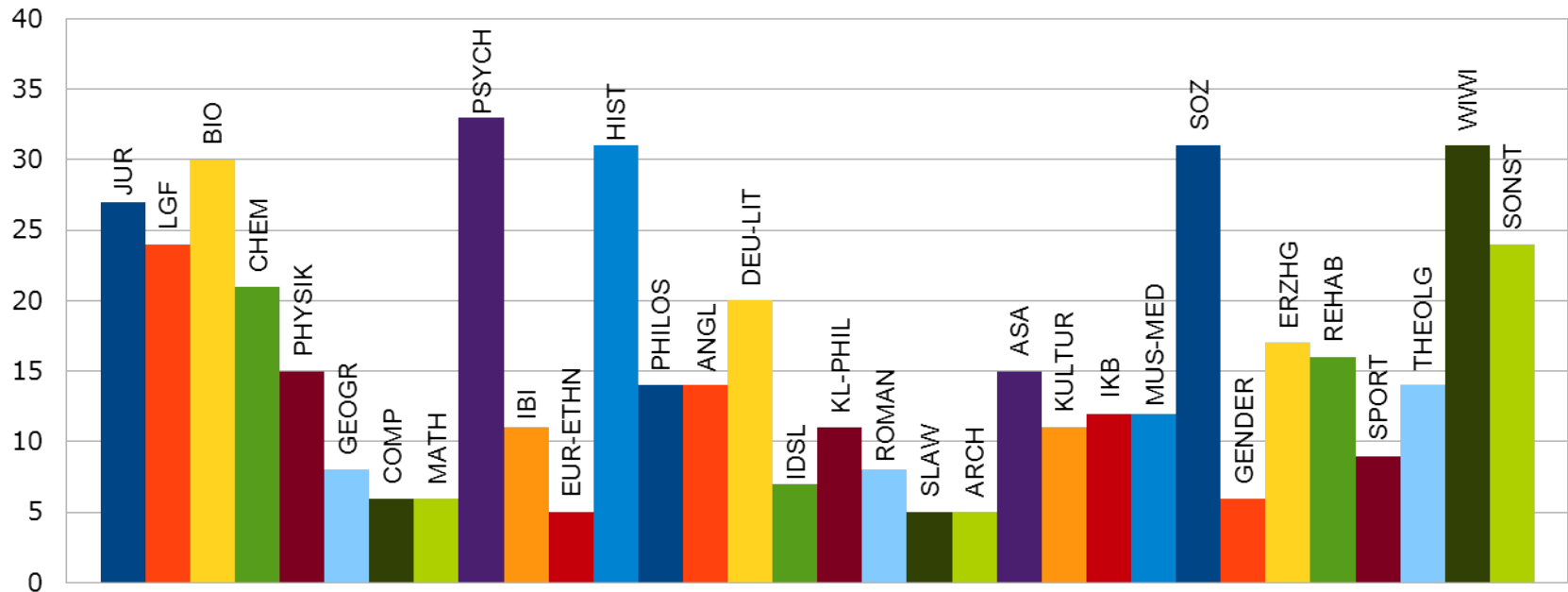
Umfrage: Eckdaten

- 24 Fragen, davon 6 Pflichtfragen
- ca. 10 Minuten zum Ausfüllen
- Versionen in zwei Sprachen (Deutsch und Englisch)
- Online-Fragebogen (Nutzung des Open-Source-Produktes LimeSurvey)
- 6 Wochen Laufzeit (24.01.-08.03.2013)

Rücklauf

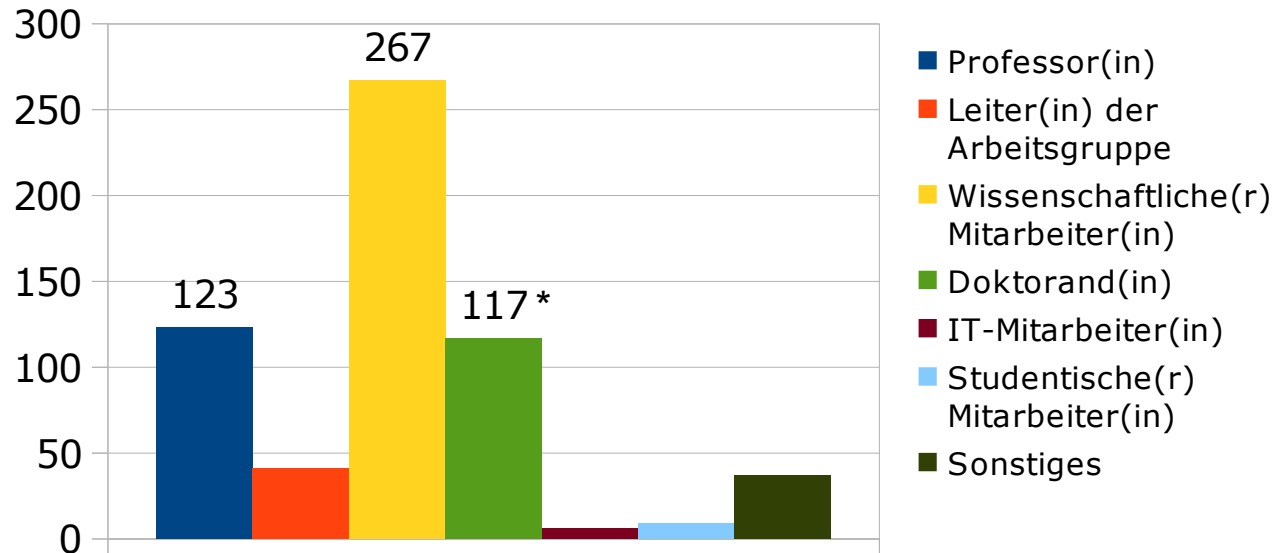
499 Antworten bzw. ~24% der Zielgruppe aus den meisten Einrichtungen:

Verteilung der Teilnehmer nach Einrichtungen
(n=499)



Position der Teilnehmer

Position der Teilnehmer



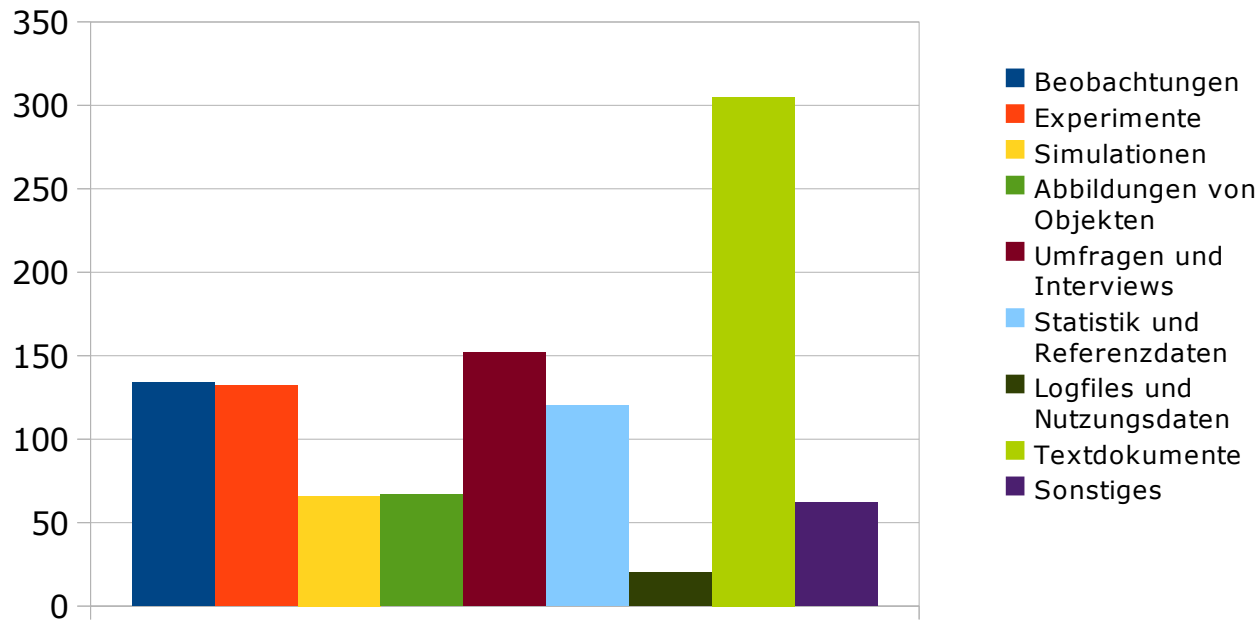
Entspricht etwa 29% der Professor/innen und 14% der wissenschaftlicher Mitarbeiter/innen der HU

Quelle: HU-Personalstatistik (Jahr 2013)

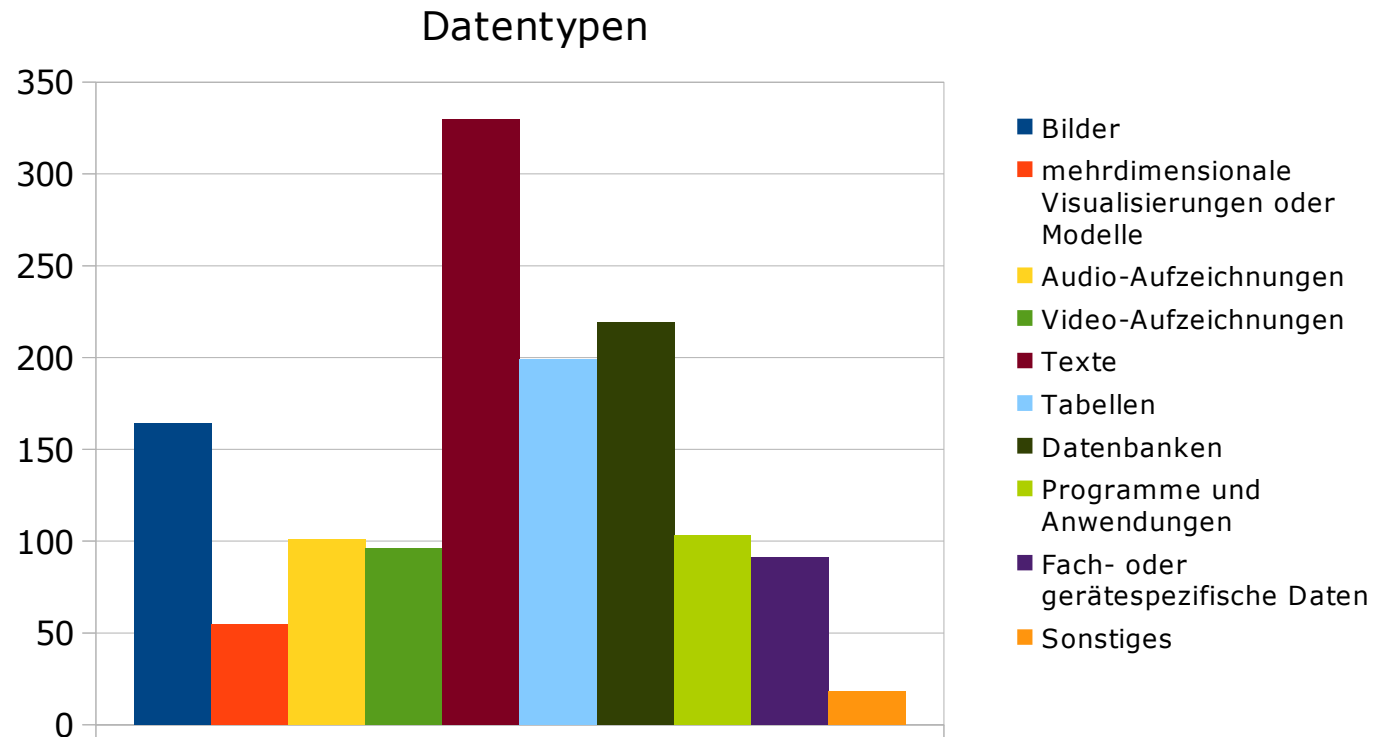
*Anm.: 63 Teilnehmer haben als Doktorand(in) und wissenschaftliche(r) Mitarbeiter(in) geantwortet (Mehrfachantworten erlaubt)

Aus welchen Quellen schöpfen Sie Ihre Forschungsdaten hauptsächlich?

Quellen von Forschungsdaten

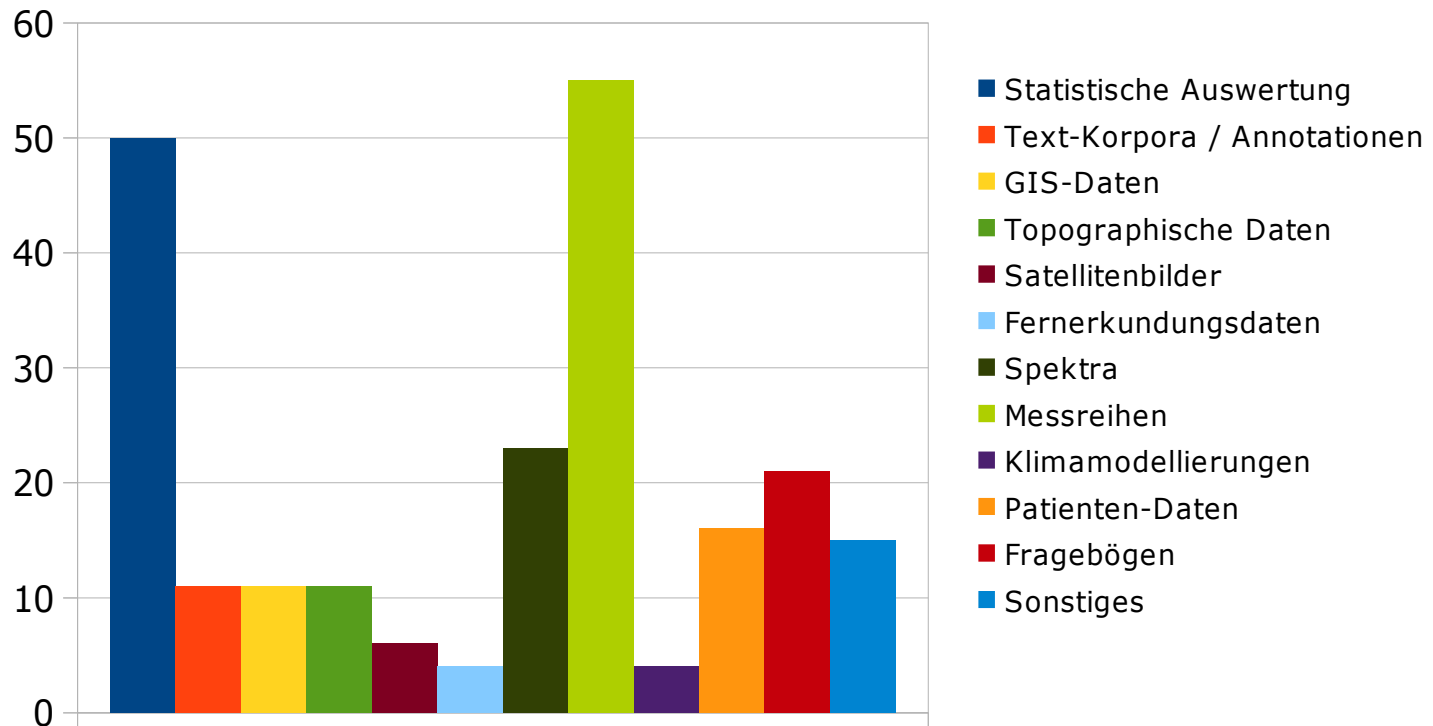


Um welche Datentypen handelt es sich?



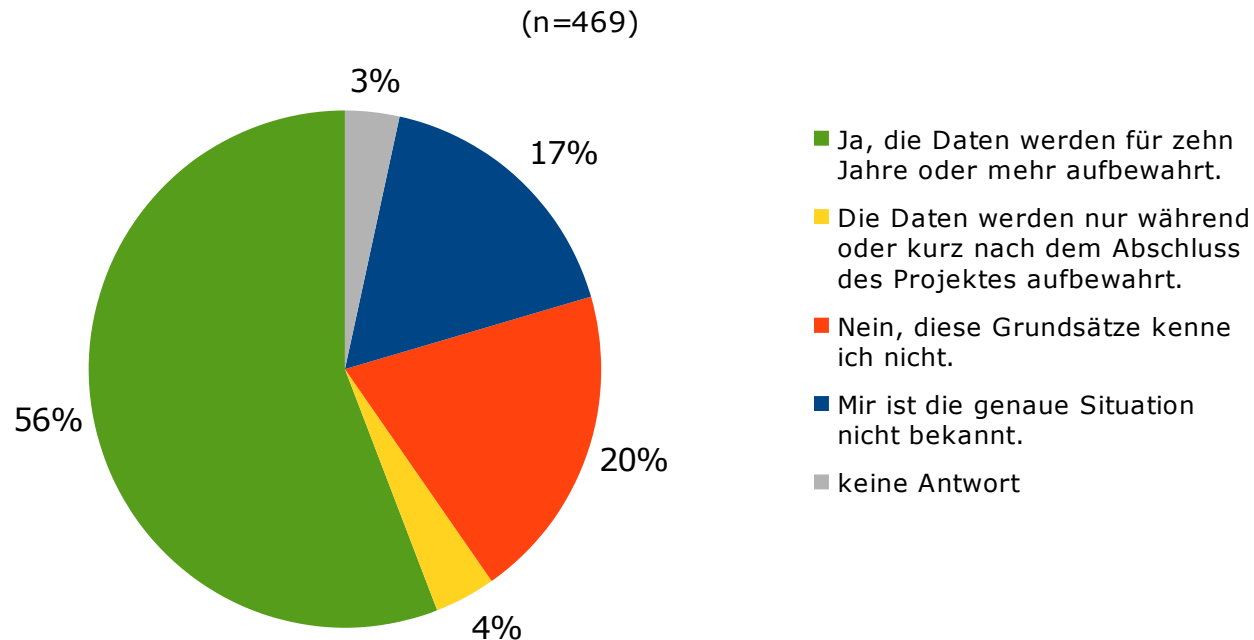
Um welche Art fach- oder gerätespezifische Daten handelt es sich?

Fach- oder gerätespezifische Daten



In den Grundsätzen zur Sicherung guter wissenschaftlicher Praxis der HU (2002) heißt es: "Primärdaten als Grundlagen für Veröffentlichungen sollen auf haltbaren und gesicherten Trägern in der Arbeitsgruppe/ Einrichtung, wo sie entstanden sind, für zehn Jahre zugänglich bleiben." Berücksichtigen Sie diese Grundsätze?

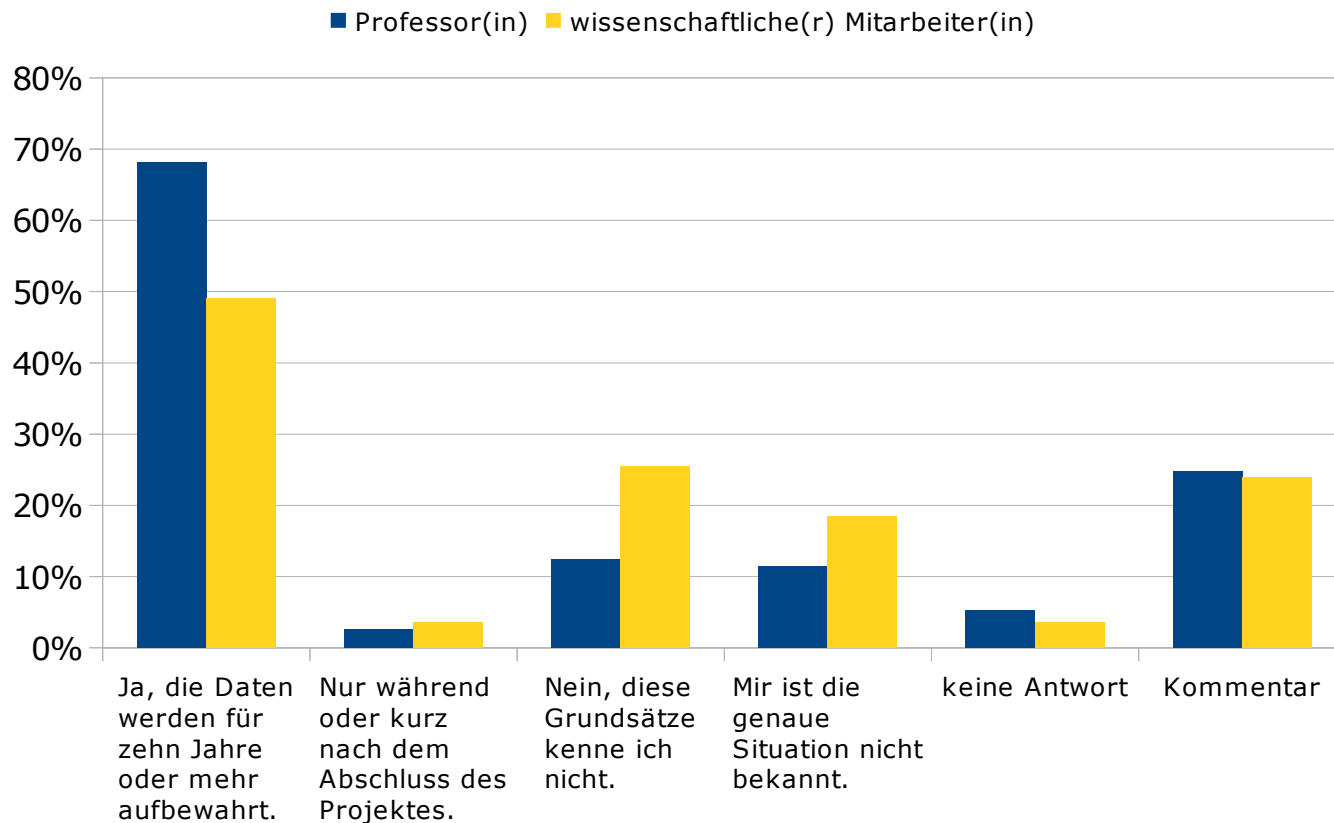
Sicherung guter wissenschaftlicher Praxis



Zudem zahlreiche Kommentare zur Anwendbarkeit der Grundsätze im eigenen Fachbereich, Einschränkungen durch kürzere Arbeitsverträge od. Ruhestand, Lizenzbestimmungen u.v.m.

Vergleich der Antworten zwischen Teilnehmer-Gruppen "Professor(in)" und "wissenschaftliche(r) Mitarbeiter(in)"

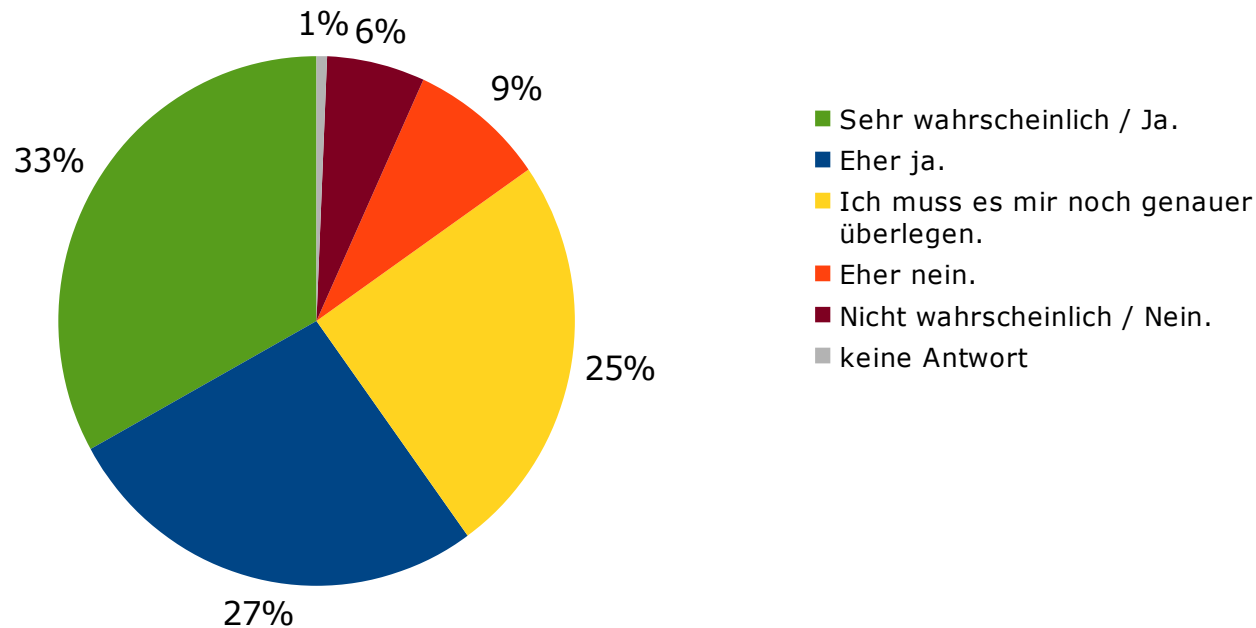
Sicherung guter wissenschaftlicher Praxis



Wären Sie grundsätzlich bereit, bestimmte Forschungsdaten in einem Datenarchiv abzulegen bzw. zugänglich zu machen?

Bereitschaft zur Zugänglichmachung

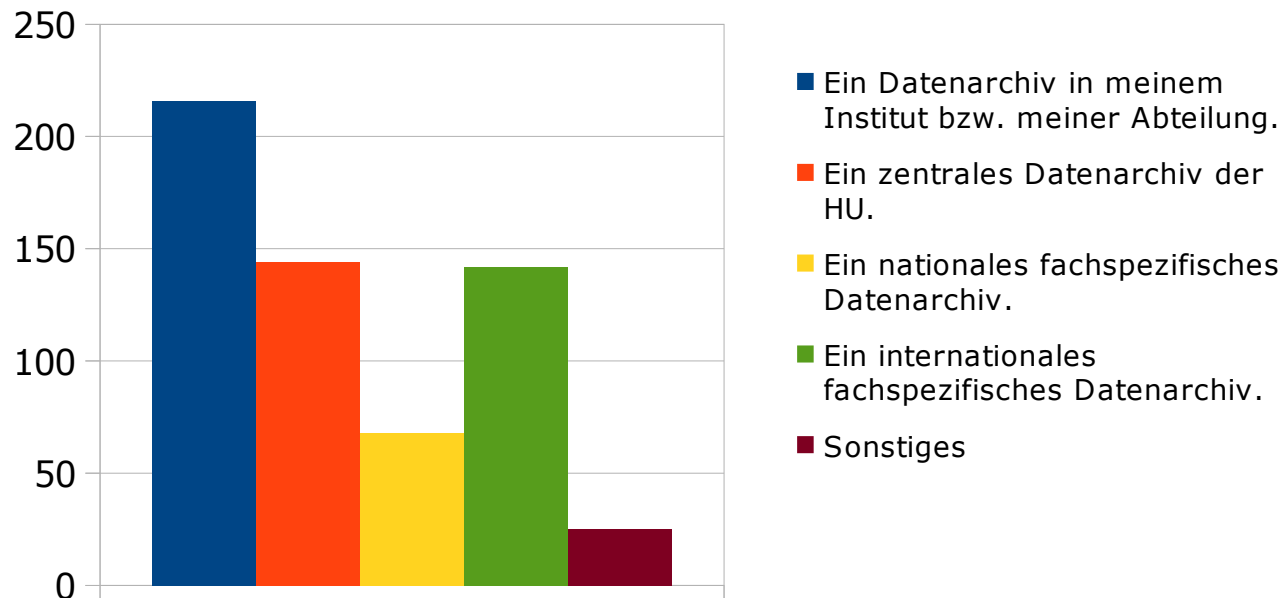
(n=457)



Ebenso zahlreiche Kommentare zu rechtlichen Bedenken, Aufwand für geeignete Aufbereitung, Anforderungen an das Datenarchiv aber auch weiteres Interesse und Begrüßung des Open-Access-Gedankens

Welches Datenarchiv würden Sie für die Ablage Ihrer Forschungsdaten am ehesten nutzen?

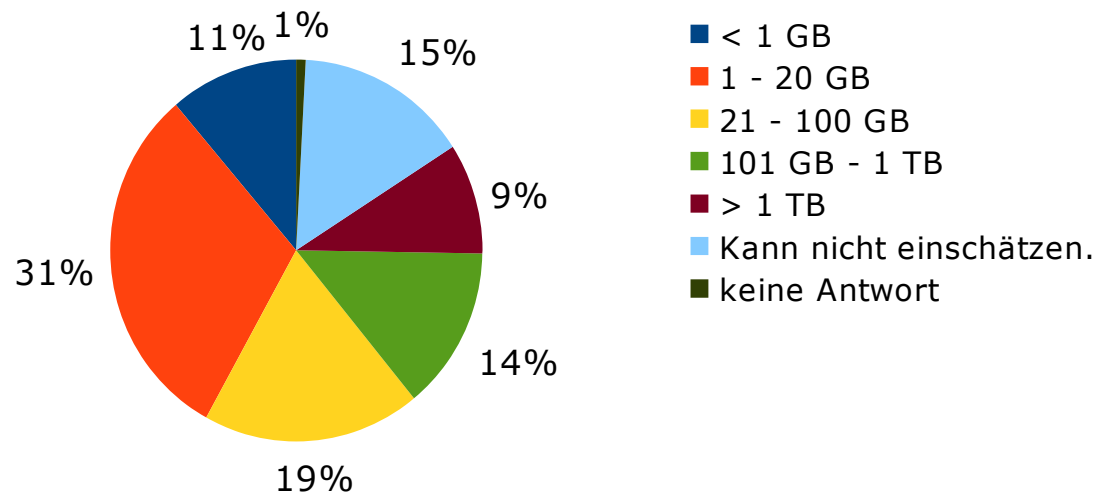
Bevorzugtes Datenarchiv



Tendenziell HU-eigene Lösungen oder international ausgerichtete Angebote der Fachcommunity bevorzugt

Bitte schätzen Sie die Gesamtgröße Ihrer Forschungsdaten bezogen auf den benötigten Speicherplatz.

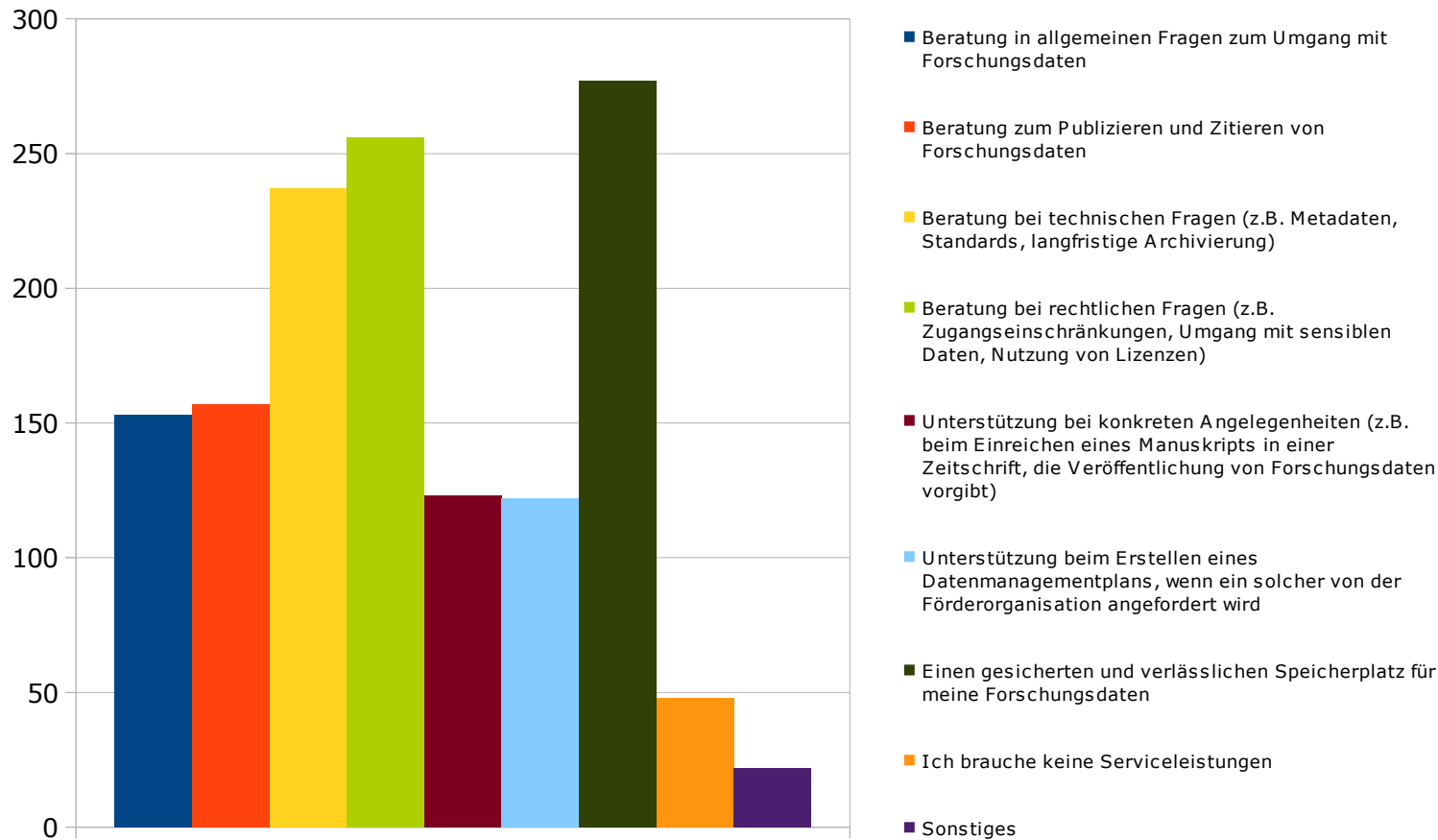
Gesamtgröße der Forschungsdaten
(n=491)



Dient als erster Versuch, das Aufkommen von Forschungsdaten an der HU zu beziffern, ist allerdings von der Definition der "Forschungsdaten" direkt abhängig

Welche Serviceleistungen würden Sie sich von der HU wünschen?

Gewünschte Unterstützung



UB und FDM

- Kompetenz im wissenschaftlichen Recherchieren und Publizieren
- Umfassende Kenntnisse von fachspezifischen Angeboten
- Expert/innen in Metadaten und Dokumentation
- Open Access in allen Facetten
- Historisch gewachsene Rolle in Unterstützung der Wissenschaftler
- Erweiterung und Stärkung des eigenen Berufsbildes
- Mitwirkung in der FDM-Initiative sehr willkommen!



Bildquelle: 3TU.Datacentrum and DANS. *Data Intelligence 4 Librarians*. III Technical skills, Citing data:
<http://dataintelligence.3tu.nl/en/iii-technical-skills/citing-data/>

Zu guter Letzt



DataCite

Helping you to find,
access, and reuse data

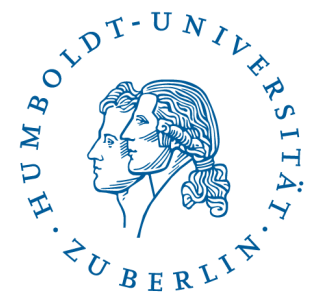
TIB | TECHNISCHE
INFORMATIONSBIBLIOTHEK



ZENODO is derived from Zenodotus, the first librarian of the Ancient Library of Alexandria and father of the first recorded use of metadata, a landmark in library history.

Zur Rolle der Bibliotheken im FDM:

- IATUL annual workshop, 5-6 Dec 2013 at Bodleian Libraries in Oxford: „Research Data Management: Finding our Role“
<http://www.bodleian.ox.ac.uk/science/iatul-workshop-2013>
- Data Information Literacy (DIL) Symposium at Purdue University, on September 23 and 24th 2013 <http://wiki.lib.purdue.edu/display/ste/Symposium>



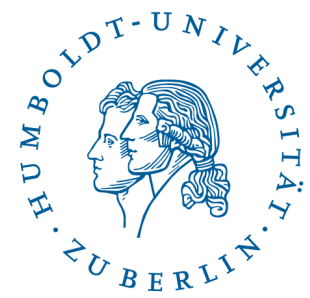
Lernmaterialien und Literatur

- Data Intelligence 4 Librarians designed by 3TU.Datacentrum and DANS
<http://dataintelligence.3tu.nl/en/home/>
- DIY Research Data MANTRA Training Kit for Librarians by EDINA and Data Library, University of Edinburgh: <http://datalib.edina.ac.uk/mantra/libtraining.html>
- Open Educational Resource for information professionals on Research Data Management, developed in RDMRose project (the libraries of the University of Leeds, Sheffield and York, with the Information School at Sheffield):
<http://rdmrose.group.shef.ac.uk/>
- RDM for librarians, Digital Curation Centre and University of Northampton:
<http://www.dcc.ac.uk/training/rdm-librarians>

- Zur Umfrage:

Simukovic, Elena; Kindling, Maxi; Schirnbacher, Peter (2013): Umfrage zum Umgang mit digitalen Forschungsdaten an der Humboldt-Universität zu Berlin. Umfragebericht (erscheint auf dem edoc-Server der HU).

Kindling, Maxi; Simukovic, Elena; Schirnbacher, Peter (2013): Forschungsdatenmanagement an Hochschulen: das Beispiel der Humboldt-Universität zu Berlin (erscheint in LIBREAS-Ausgabe #23).



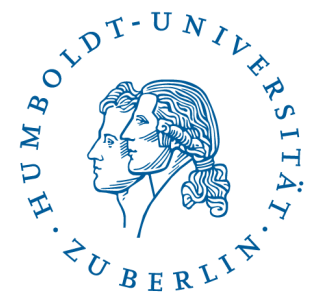
Vielen Dank!

Elena Simukovic

Forschungsdatenmanagement-Koordinatorin

elena.simukovic@cms.hu-berlin.de

<http://www.cms.hu-berlin.de/dataman>



Was sagen die Forscher?

"Ich würde meine Daten gerne länger sichern und freue mich, von dieser Möglichkeit zu hören!"

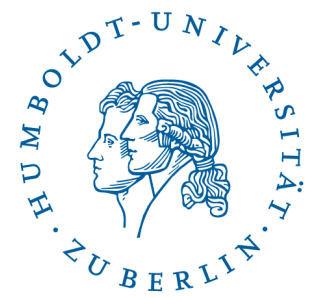
"Diese Grundsätze werden während der Ausbildung (Studium) nicht gelehrt."

"Ich kenne die Grundätze und halte sie "privat" soweit wie möglich ein, da die Verträge in meiner Karrierestufe jedoch immer befristet und tlw. nicht länger als 2 Jahre sind und übergreifenden Datenbanken bsiher kaum (gar nicht) genutzt werden ist die Praxis der Datenaufbewahrung undefiniert."

"Ich weiß nicht, was mit meinen auf dem Server gespeicherten Daten passiert, wenn ich den Lehrstuhl verlasse."

"Ich bin gerne bereit, meine Daten zur VErfügung zu stellen, wenn sie "gut behandelt" werden!"

"Die Daten werden aufbewahrt, die Grundsätze kannte ich nicht."



Was sagen die Forscher?

Einige Kommentare und Anregungen für Diskussion

Wo sind wir heute:

"Diesen Grundsatz [Gute wissenschaftliche Praxis] beachtet in meiner Gruppe sicherlich niemand. Ich wäre auch der einzige, der Erfahrung mit Haltbarkeit von Datenträgern wie CD etc. [hat]. Einerseits fehlt es an den Organisationsstrukturen dies zu erreichen. Andererseits am technischen Wissen, diese Haltbarkeit mit hinreichender Sicherheit zu erreichen."

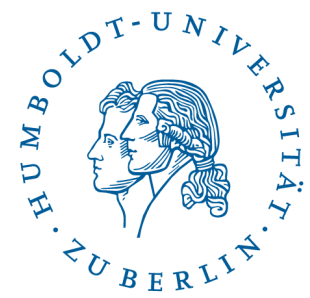
"Unfortunately there are many incidents of plagiarism in our department and thus I would be quite reluctant sharing data without personal control on who has access on them and why"

"Innovations- bzw. Konkurrenzdruck: man hält seine Datenfunde [...] lieber "geheim" und veröffentlicht stattdessen Aufsätze, Bücher etc., in denen die Daten zitiert werden; Ausnahme evt. Gruppenprojekte"

"Hauptziel meiner Arbeit sind Publikationen, die nichtveröffentlichten Arbeitsdateien gehen niemanden etwas an und würden auch niemanden interessieren."

"Es gibt keine geeigneten Kollaborationssysteme an der HU-Berlin, die zur Mitarbeit von Kollegen anderer Hochschulen animieren."

"Manchmal wäre Hilfe in den genannten Bereichen gut, aber ich kann mir auch wieder nicht vorstellen, wie das konkret laufen könnte. Letztendlich quält man sich dann doch wieder allein durch Aufgaben."



Was sagen die Forscher? (Fortsetzung)

Einige Kommentare und Anregungen für Diskussion

Wo wollen wir hin:

„Zumindest institutsintern würde ich meine statistischen Daten gerne Kolleg/innen und Studierenden zur Verfügung stellen“

"Ich würde gerne besser über die Möglichkeiten der Speicherung von Forschungsdaten in Datenarchiven Bescheid wissen und würde diese auch nutzen wollen."

"Ich bin für open Access und DATentransparenz, sofern die Urheberschaft mit Kennzeichnung des Projekts gewährleistet ist. Das unterstützt den Austausch, vermeidet Überschneidungen und fördert die Sichtbarkeit meiner wissenschaftlichen Arbeit"

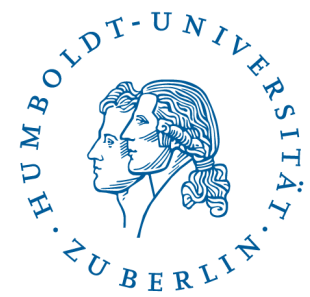
"Aber in manchen Fällen sind solche projektgebundenen Infrastrukturen nicht sinnvoll, ein projektübergreifendes System macht mehr Sinn, als in jedem Projekt eine halbgeheure Lösung."

"Es wäre schon gut, wenn es einen guten Cloud-Dienst gäbe oder die Netzinfrastruktur ZUVERLÄSSIG über VPN nutzbar wäre."

"Nicht nur Beratung sondern praktische Unterstützung"

"Einen Ansprechpartner für individuelle Fragen, den man anrufen/aufsuchen kann. Häufig sind meine Fragen eine bunte Mischung aus allen möglichen hier genannten Dingen."

"Bisher brauchte ich keine Serviceleistungen, aber es wäre sicher gut, sie zu haben."



Was sagen die Forscher? (Fortsetzung)

Einige Kommentare und Anregungen für Diskussion

Was wir zu bedenken haben:

"Der Professor unter dem ich als Doktorand arbeite hat hierbei allerdings das letzte Wort."

"Ich habe mich bisher noch wenig mit Datenarchiven befasst, so dass ich erst eine Nutzen/Risiko-Abwägung treffen müsste."

"Im Prinzip ja, ich fürchte allerdings, dass ich einige Zeit investieren müsste, um Materialien zu sortieren, bzw. so "aufzuräumen", dass Dritte damit arbeiten könnten oder das Material überhaupt in ein Datenarchiv eingespeist werden kann."

"will happily do it once I published my own research using the data."

"Der Begriff der Forschungsdaten ist so nicht tragfähig: Wenn alle Texte, mit denen ein Historiker zu tun hat, Forschungsdaten wären, gäbe es enorme urheberrechtliche Probleme mit der Ablage..."

"Ich habe Probanden zugesichert, dass niemand außer mir und meiner Arbeitsgruppe Zugriff auf die Daten haben wird."

"Quellen unterliegen teils dem Datenschutz/ Urheberrecht. Ich würde den Service nur nutzen, wenn sie vor unbefugtem Zugriff geschützt wären."