

## Bestands- und Bedarfserhebung zum Forschungsdatenmanagement an den BUA-Einrichtungen 2021/2022

**Forschungsvorhaben:** Projekt „Concept Development for Collaborative Research Data Management Services“ (kurz: BUA-FDM) der Berlin University Alliance (BUA)

**Förderkennzeichen:** 501\\_CRDMS

**Forschende:** Angela Ariza de Schellenberger (0000-0002-1005-5780), Evgeny Bobrov (0000-0002-2368-464X), Kerstin Helbig (0000-0002-2775-6751), Denise Jäckel (0000-0002-8720-6559), Monika Kuberek (0000-0002-1672-5271), Lea-Sophie Orozco Prado (0000-0001-8036-1479), Elisabeth Maria Schlagberger (0000-0002-2979-4997), Sibylle Söring (0000-0002-1698-3289), Britta Steinke (0000-0003-3320-5187)

**Kontext der Erhebung:** Das 2020-2022 von der Berlin University Alliance (BUA) geförderte Projekt „Concept Development for Collaborative Research Data Management Services“ hat sich die Konzeptentwicklung zum nachhaltigen Aufbau von Kompetenz und Expertise zum Thema FDM für Forschende und Multiplikator\*innen sowie die Stärkung von Services rund um das FDM innerhalb der BUA zum Ziel gesetzt, um bereits bestehende Ressourcen bedarfsgetrieben bestmöglich nutzbar zu machen, Parallelentwicklungen zu vermeiden und Synergieeffekte zu ermöglichen. Hierfür wurde im Zeitraum 11/2021 bis 01/2022 an den vier BUA-Einrichtungen (Freie Universität Berlin, Humboldt-Universität zu Berlin, Technischen Universität Berlin, Charité – Universitätsmedizin Berlin) eine Bestands- und Bedarfserhebung zum Umgang mit Forschungsdaten durchgeführt, deren Ergebnisse als Basis für die Entwicklung bedarfsorientierter standortspezifischer und standortübergreifender Beratungs-, Schulungs-, Kommunikations- und technischer Serviceleistungen dienen. Ziel der Befragungen war es, zu ermitteln, welche Services

- 1) aktuell an den jeweiligen Standorten bekannt sind und genutzt werden
- 2) an allen Standorten gewünscht, aber noch nicht angeboten werden
- 3) im Verbund als fruchtbar erachtet werden, um qualitätsvolle Forschung auch institutionsübergreifend zu ermöglichen.

**Zielgruppe der Erhebung:** Alle Forschenden der Einrichtungen (Professor\*innen, Doktorand\*innen, Wissenschaftliche Mitarbeitende) insbesondere auch Beschäftigte in Verbundvorhaben (z. B. Exzellenzcluster, Sonderforschungsbereiche) sowie forschungsunterstützendes Personal (z. B. technische Mitarbeiter\*innen), bei der Technischen Universität „Sonstige“.

**Erhebungsmethode:** Online-Befragung (insgesamt haben 975 Personen die Umfrage vollständig ausgefüllt, davon 135 Teilnehmende bei der Freien Universität Berlin, 162 Teilnehmende bei der Humboldt-Universität zu Berlin, 207 Teilnehmende bei der Technischen Universität Berlin und 471 Teilnehmende bei der Charité - Universitätsmedizin)

**Durchführende Einrichtungen:** Freie Universität Berlin, Humboldt-Universität zu Berlin, Technische Universität Berlin, Charité - Universitätsmedizin Berlin

**Erhebungszeitraum:** 15.11.2021 bis 07.01.2022 (Freien Universität Berlin und Technischen Universität Berlin), 02.12.2021 bis 07.01.2022 (Humboldt-Universität zu Berlin) und 07.12.2021 bis 19.01.2022 (Charité - Universitätsmedizin Berlin)

#### Verwendete Software:

- LimeSurvey (Version 5.3.13 u. Version 5.2.2)<sup>1</sup> zur Datenerhebung an der Freien Universität Berlin, Humboldt-Universität zu Berlin und Technischen Universität Berlin
- Sociolutions QUAMP<sup>2</sup> zur Datenerhebung an der Charité - Universitätsmedizin Berlin
- R (Version 4.0.3)<sup>3</sup> zur Datenanalyse und -aufbereitung
- Adobe Acrobat Pro DC (Version 2022.002.20191) zur Sicherung des Fragebogens
- Microsoft Office Excel 2016 für die Erstellung des Codebooks

#### Daten des generischen Datensatzes

Gemeinsamer Datensatz mit den Umfrageergebnissen aller vier BUA-Einrichtungen für die **generischen Fragen** der Kodierungen CON, REG, TOOL, SUP7, ICT, SOL, SER (identisch gestellte Fragen u. Antwortoptionen):

- BUA-FDM\_Umfragedaten\_generisch\_alle\_Institutionen\_2022\_Datensatz.csv
  - Anonymisierter und formatierter Datensatz der Umfrageergebnisse
  - Freitextantworten sind auf Deutsch und auf Englisch (aus datenschutzrechtlichen Gründen wurden die Freitextantworten der Charité durch Dummies ersetzt: „1“ Antwort gegeben, „0“ keine Antwort gegeben)
  - Die Variablen-Kodierung ist auf Englisch

Umfrageergebnisse aller vier BUA-Institutionen für die Frage nach dem **beruflichen Status** (jeweils ein Datensatz pro Institution, da sich die Antwortoptionen unterschieden haben). Die Variablen-Kodierung ist auf Englisch, alle Datensätze wurden anonymisiert und formatiert.

- BUA-FDM\_Umfragedaten\_STA\_FU\_2022\_Datensatz.csv
- BUA-FDM\_Umfragedaten\_STA\_HU\_2022\_Datensatz.csv
- BUA-FDM\_Umfragedaten\_STA\_TU\_2022\_Datensatz.csv
- BUA-FDM\_Umfragedaten\_STA\_CHA\_2022\_Datensatz.csv

Umfrageergebnisse aller vier BUA-Institutionen für die Frage nach der Zugehörigkeit zu einer **Organisationseinheit / Förderstruktur** (jeweils ein Datensatz pro Institution, da sich die Antwortoptionen unterschieden haben). Die Variablen-Kodierung ist auf Englisch, alle Datensätze wurden anonymisiert und formatiert.

- BUA-FDM\_Umfragedaten\_ORG\_FU\_2022\_Datensatz.csv
- BUA-FDM\_Umfragedaten\_ORG\_HU\_2022\_Datensatz.csv
- BUA-FDM\_Umfragedaten\_ORG\_TU\_2022\_Datensatz.csv

---

<sup>1</sup> Limesurvey GmbH. / LimeSurvey: An Open Source survey tool /LimeSurvey GmbH, Hamburg, Germany. URL <http://www.limesurvey.org> (zugegriffen am 30.09.22)

<sup>2</sup> <https://www.sociolutions.de/article/quamp-software.html> (zugegriffen am 30.09.2022)

<sup>3</sup> R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

- BUA-FDM\_Umfragedaten\_ORG\_CHA\_2022\_Datensatz.csv

Für die Auswertung der **Freitextfrage SER** wurden die Antworten **Kategorien** zugeordnet, die Verteilung ist in folgendem Datensatz abgebildet:

- BUA-FDM\_Kategorien-SER\_2022\_Datensatz.csv
  - Die Variablen-Kodierung ist auf Deutsch

### Analyseskript

- BUA-FDM\_Auswertung-Umfrage FDM\_2022\_R-Skript.Rmd
  - R-Code für die Analyse der Daten und die Erstellung entsprechender Abbildungen
  - Der R-Code ist auf Deutsch dokumentiert
  - Das Skript wurde mithilfe von `styler`<sup>4</sup> gemäß dem tidyverse-Styleguide formatiert
- `renv.lock`
  - Alle genutzten R-Pakete wurden mit `renv`<sup>5</sup> aufgezeichnet, damit die exakten Versionen reproduzierbar sind

### Zusätzliche Dokumentationsdateien

- BUA-FDM\_Codebook-Bedarfserhebung-FDM\_generische Fragen\_Tabelle\_2022.ods
  - Das Codebook erklärt die Variablen und ihre Werte, die als Spaltenbezeichnungen im Datensatz genutzt wurden (diese wurden äquivalent auch für die Kodierung der Variablen im R-Skript genutzt)
- BUA-FDM\_Fragebogen-Bedarfserhebung-FDM\_generisch\_2021.pdf
  - Der Fragebogen enthält sämtliche generische Fragen zur Nachnutzung

**Lizenz:** Alle Daten und Dateien sind frei verfügbar unter der Lizenz „Creative Commons Zero v1.0 Universal (CC0 1.0)“ (<https://creativecommons.org/publicdomain/zero/1.0/>).

---

<sup>4</sup> Müller K, Walthert L (2022). `styler`: Non-Invasive Pretty Printing of R Code. <https://github.com/r-lib/styler>, <https://styler.r-lib.org>.

<sup>5</sup> Ushey K (2022). `renv`: Project Environments. R package version 0.16.0, <https://rstudio.github.io/renv/>.