

UNEKE –

Vom USB-Stick zur NFDI – Entwicklung eines Kriterien-geleiteten Entscheidungsmodells für den Aufbau von Forschungsdaten-Infrastrukturen

Dr. Stephanie Rehwald

Projektteam UNEKE: B. Brenger, Dr. A. Lopez, Prof. S. Stieglitz, Dr. K. Wilms

RDA-DE Jahrestagung 13. Februar 2023

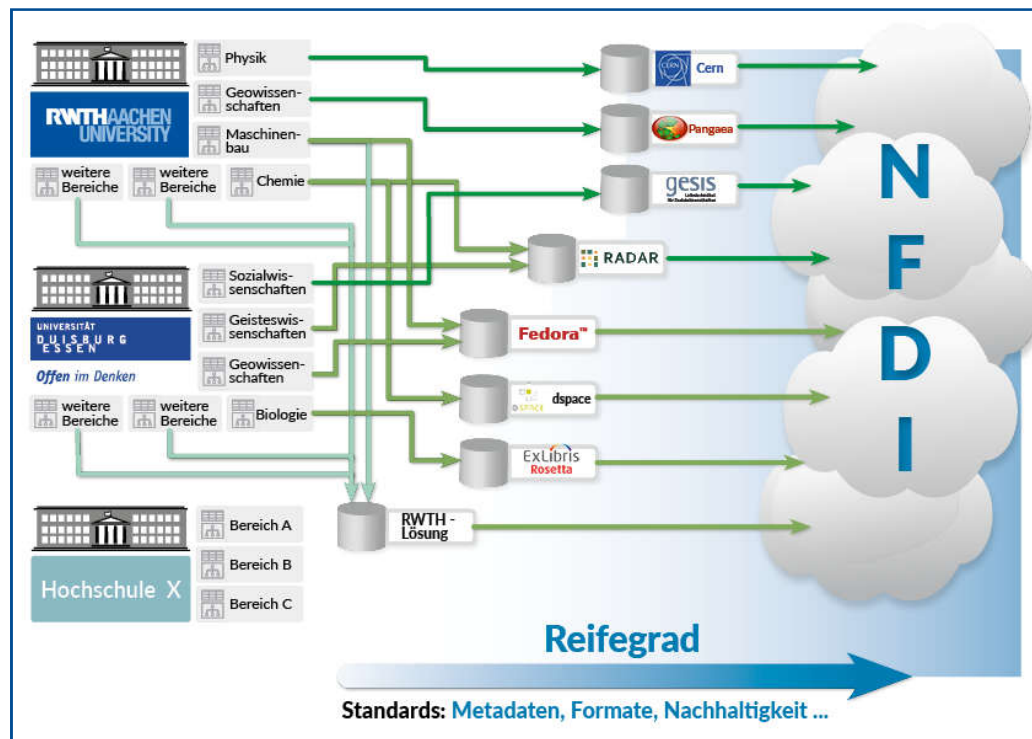


Die Präsentation ist lizenziert unter einer [CC BY 4.0-Lizenz](https://creativecommons.org/licenses/by/4.0/).
Logos und anders gekennzeichnete Symbole und Bilder sind von der Lizenzierung ausgenommen.

RWTHAACHEN
UNIVERSITY

Bildquelle: RWTH Aachen,
Alle Rechte vorbehalten

Dilemma: Institutionelle Speicherung von FD vs. nationale Strukturen (NFDI) – was müssen Hochschulen aufbauen?

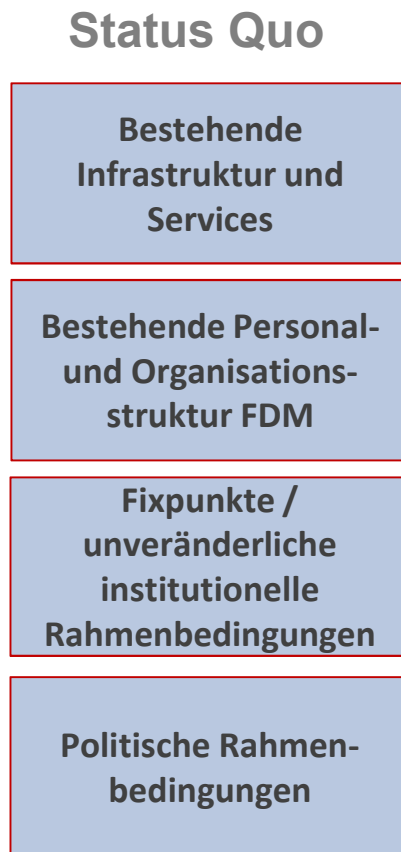


Idee:

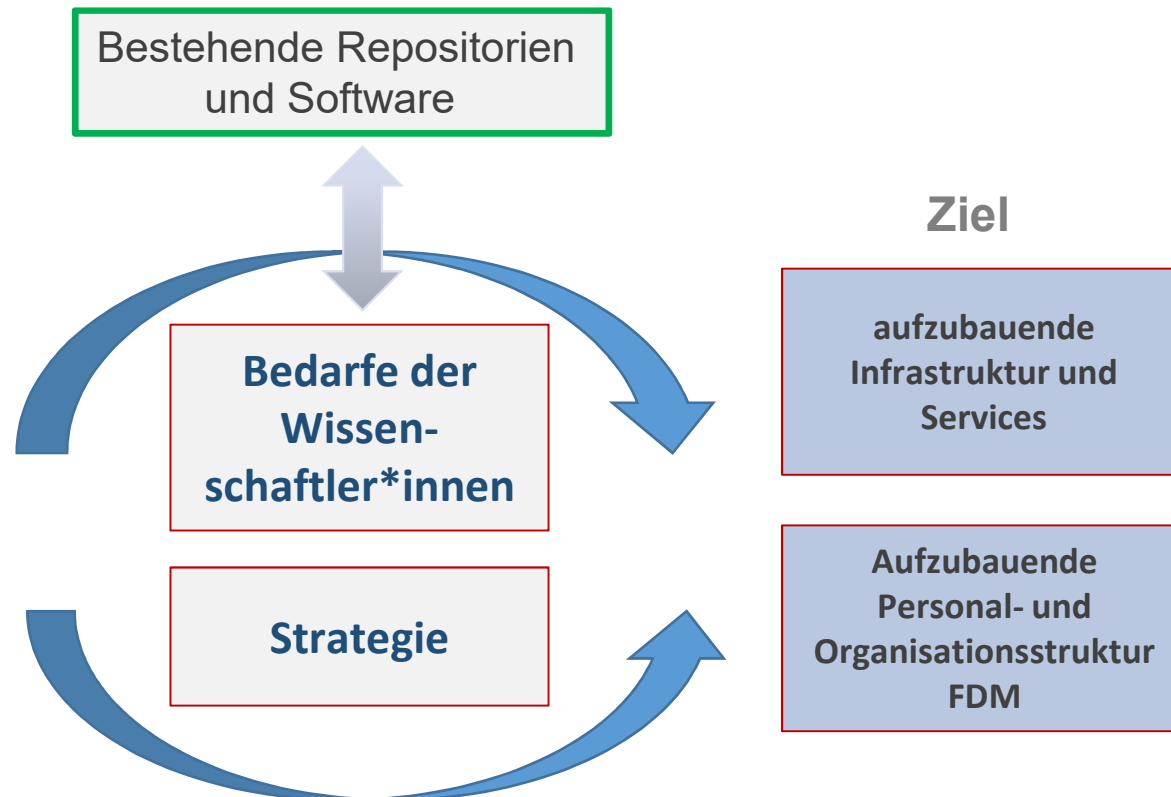
Entwicklung einer **Roadmap**, die entlang messbarer **Kriterien** als Grundlage für Hochschulen dienen kann, eine **Entscheidung** und Auswahl der passenden Speicher- und Repositorienlösung zu treffen.

Kontext der UNEKE: Entwicklung einer Roadmap zur FDM-Infrastruktur

1) Bestandsaufnahme und Bedarfsanalyse

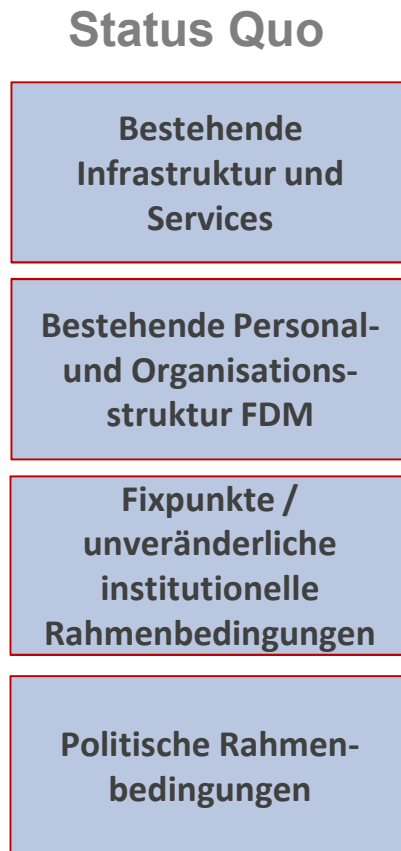


2) Bestandsaufnahme von Repositorien

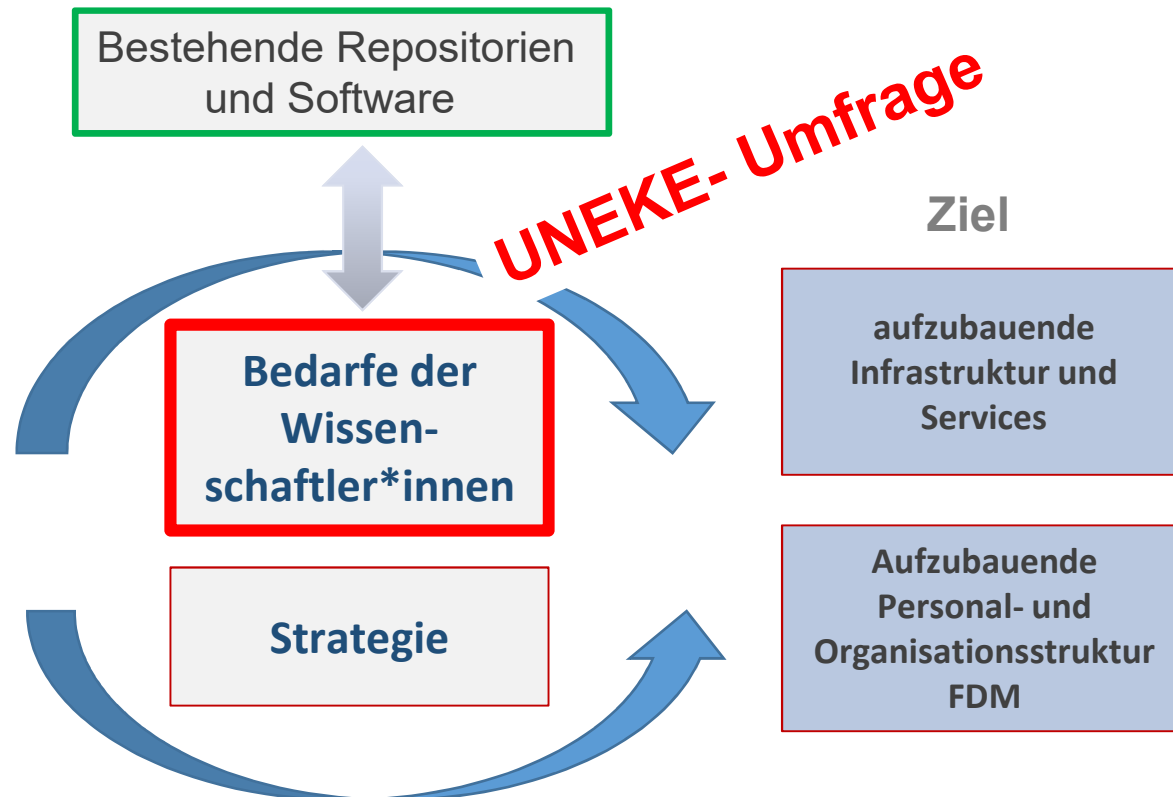


Kontext der UNEKE: Entwicklung einer Roadmap zur FDM-Infrastruktur

1) Bestandsaufnahme und Bedarfsanalyse

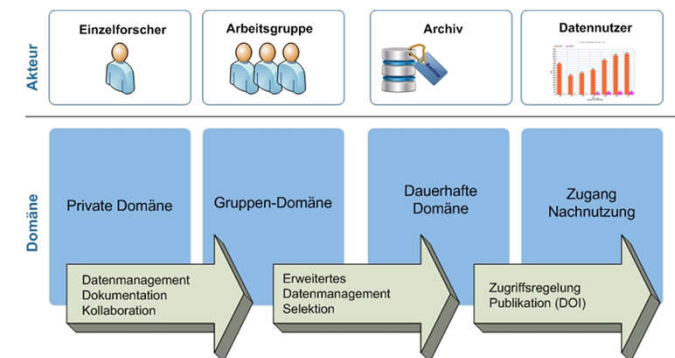


2) Bestandsaufnahme von Repositorien



Konzept der Umfrage:

- Basierend auf Workshops und Interviews
- Entwicklung des Fragebogens mit 6 Kooperationspartnern
- Wissenschaftlicher Teil:
„Daten teilen“- Hemmnisse und Anreize
- Anpassbarer deskriptiver Teil auf Basis einschlägiger Umfragen
- UNEKE: Schwerpunkt auf „Speicherung“
- Orientierung an Domänenmodell
- Nachnutzbar: Awareness-Materialien, Fragebogen, Daten



Quelle: Erweitertes Domänenmodell (Eigene Darstellung in Anlehnung an WissGrid: „Leitfaden zum Forschungsdaten-Management“, 2011)

Erhebung der Bedarfe der Forschenden



* Wilms, Konstantin, Brenger, Bela, Lopéz, Ania, Rehwald, Stephanie, & Stieglitz, Stefan (2020). UNEKE - Umfrage zur Speicherpraxis und Speicherbedarfen für Forschungsdaten. GESIS Datenarchiv, Köln. ZA6981 Datenfile Version 2.0.0, <https://doi.org/10.4232/1.13327>

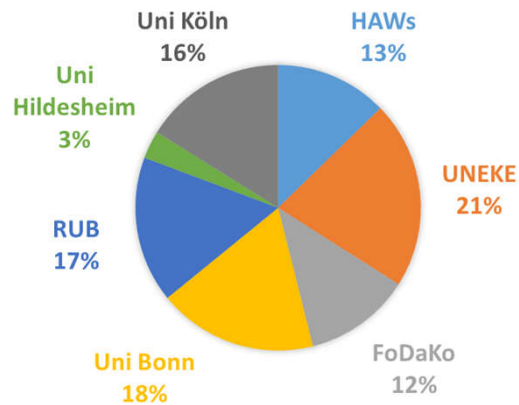
Ergebnisse – ein Einblick



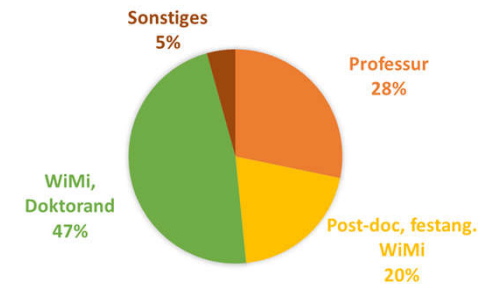
UNEKE-Umfrage 2018 - Eckdaten

Teilnehmende

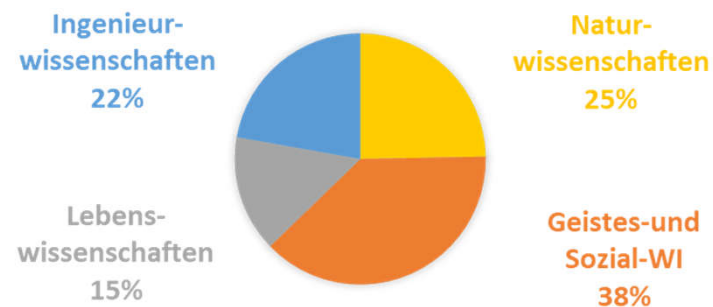
13 Hochschulen,
1684 TN



Statusgruppen



Fachdisziplinen



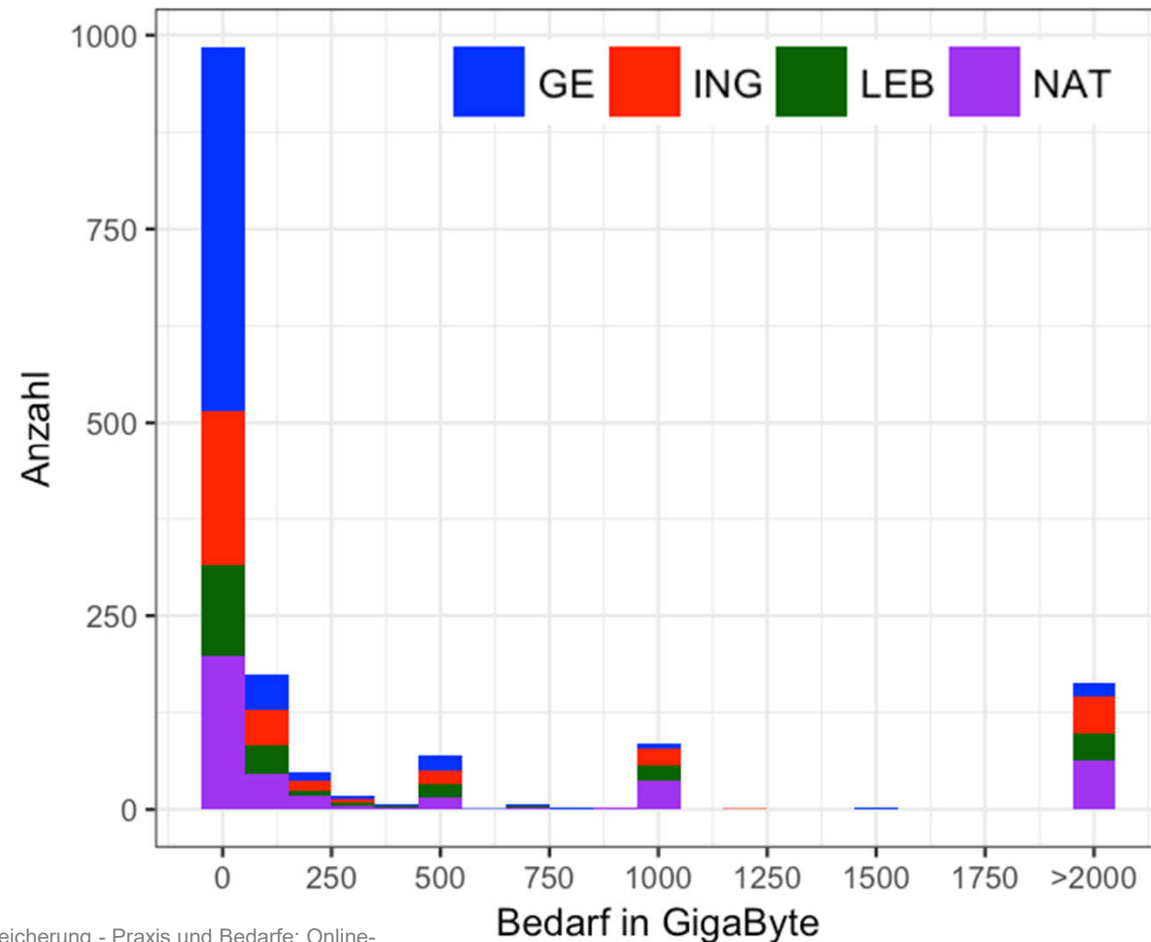
Quelle: Brenger, Bela et al. (2019): UNEKE: Forschungsdatenspeicherung - Praxis und Bedarfe: Online-Survey 2019. CC BY 4.0, Online unter: <https://doi.org/10.17185/dupublico/70259>

Speicherbedarf (pro Projekt)

Die meisten
Forschenden
kommen mit
<500GB aus

„Elefanten“
sind in allen
Disziplinen
vertreten

Speicherbedarf pro Forschungsprojekt

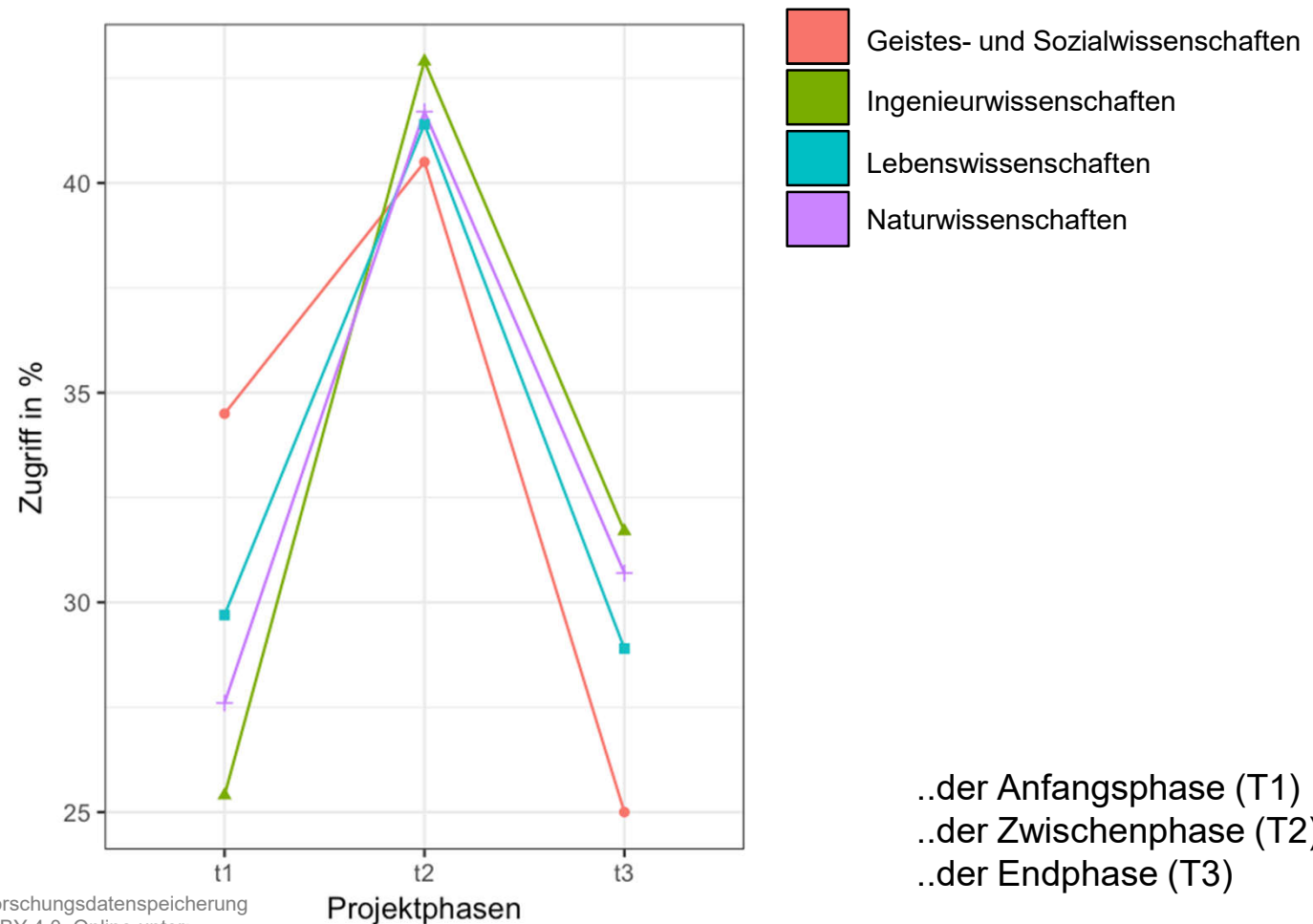


Quelle: Brenger, Bela et al. (2019): UNEKE: Forschungsdatenspeicherung - Praxis und Bedarfe: Online-Survey 2019. CC BY 4.0, Online unter: <https://doi.org/10.17185/dupublico/70259>

Projektphasen

Wieviel Prozent der Datenmenge fällt dabei im..
..ersten/zweiten/dritten Drittel des Forschungsprojektes an?

Daten-
erhebung
in der
mittleren
Phase

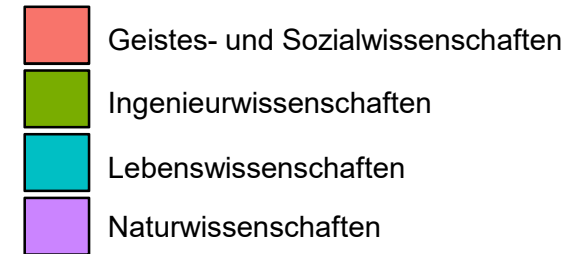
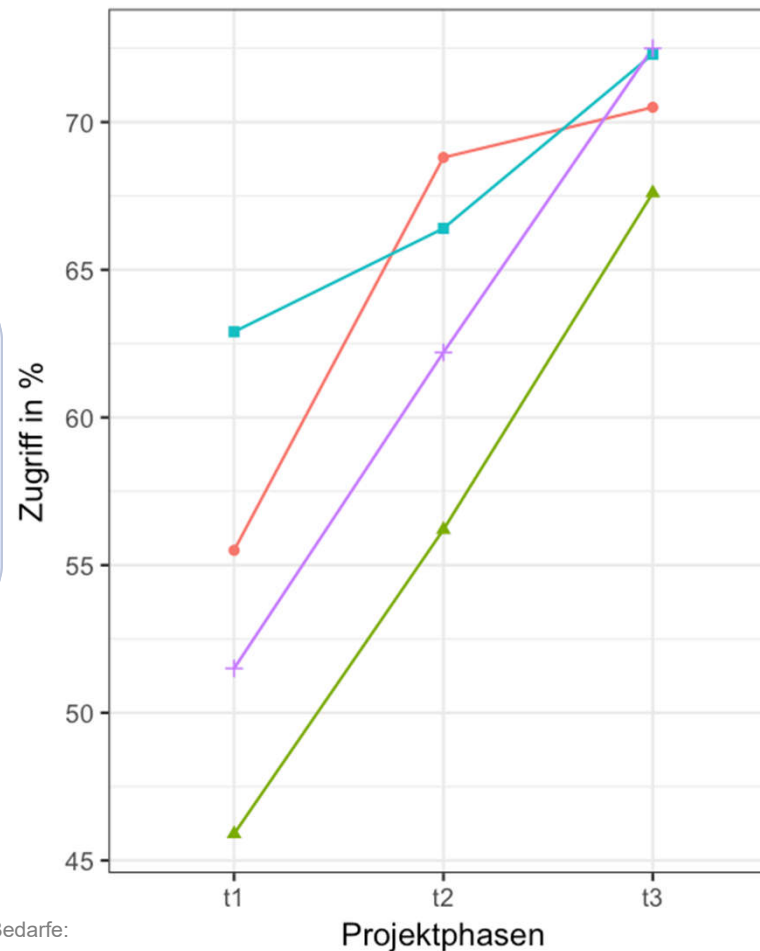


..der Anfangsphase (T1)
..der Zwischenphase (T2)
..der Endphase (T3)

Quelle: Brenger, Bela et al. (2019): UNEKE: Forschungsdatenspeicherung
- Praxis und Bedarfe: Online-Survey 2019. CC BY 4.0, Online unter:
<https://doi.org/10.17185/dupublico/70259>

Wieviel Prozent der Gesamtdaten müssen (schätzungsweise) im Sinne der guten wissenschaftlichen Praxis **10 Jahre gespeichert** werden?

Umfang der zu sichernden Daten steigt **kontinuierlich** bis auf **~70%**

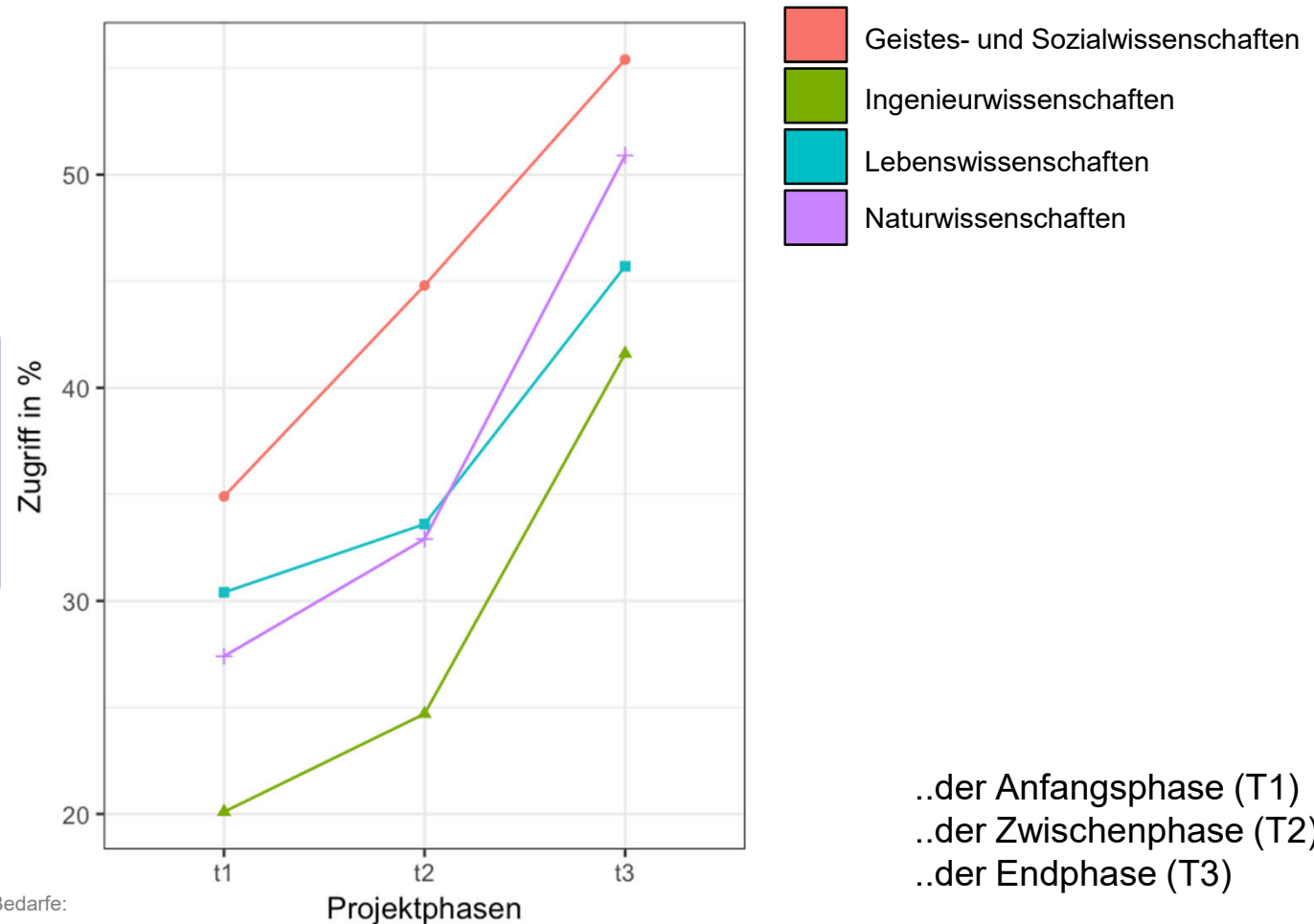


..der Anfangsphase (T1)
..der Zwischenphase (T2)
..der Endphase (T3)

Quelle: Brenger, Bela et al. (2019): UNEKE: Forschungsdatenspeicherung - Praxis und Bedarfe: Online-Survey 2019. CC BY 4.0, Online unter: <https://doi.org/10.17185/dupublico/70259>

Wie viel Prozent der Gesamtdaten, sollten aus Ihrer Sicht mindestens 25 Jahre, oder darüber hinaus archiviert werden?

Archivierung
verstärkt gegen
Projektende
und betrifft
40-50%
der Daten!



..der Anfangsphase (T1)
..der Zwischenphase (T2)
..der Endphase (T3)

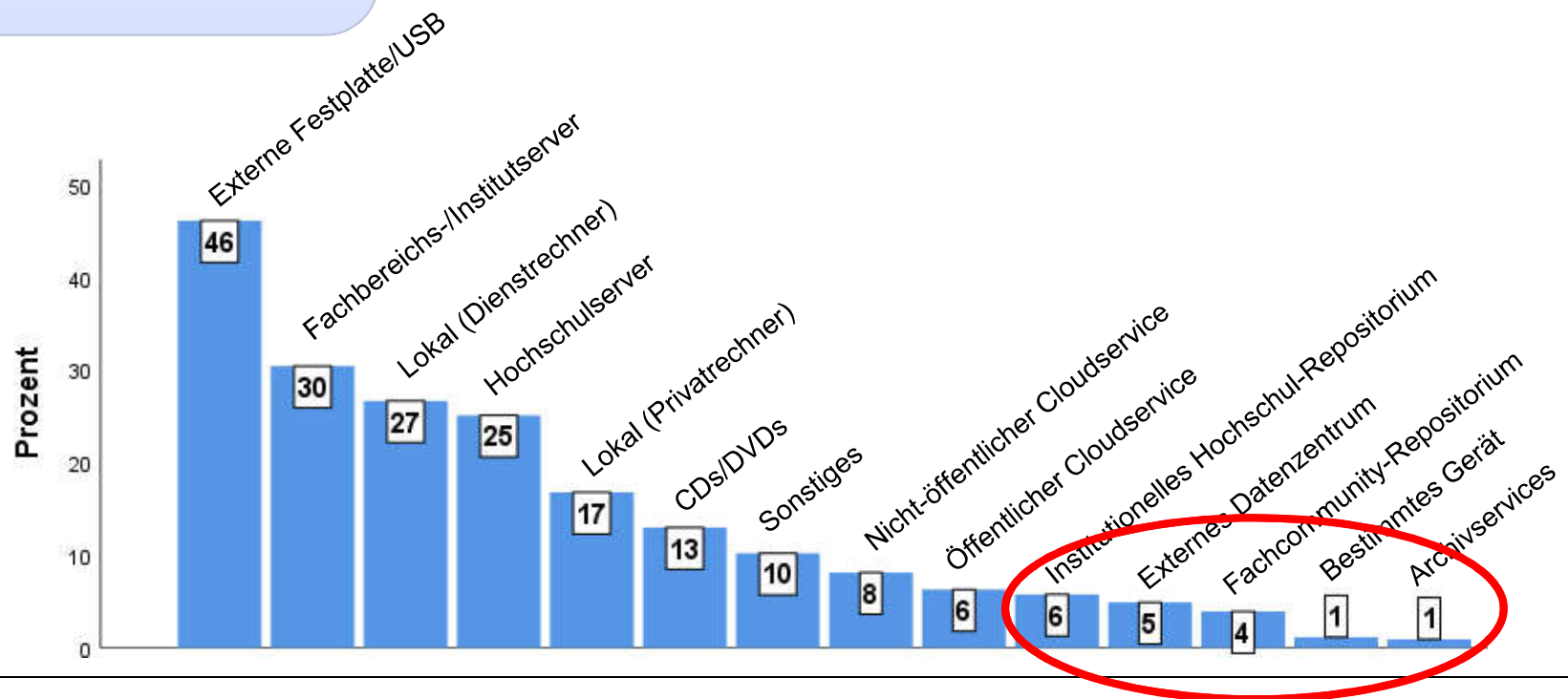
Quelle: Brenger, Bela et al. (2019): UNEKE:
Forschungsdatenspeicherung - Praxis und Bedarfe:
Online-Survey 2019. CC BY 4.0, Online unter:
<https://doi.org/10.17185/duepublico/70259>

Archivort:

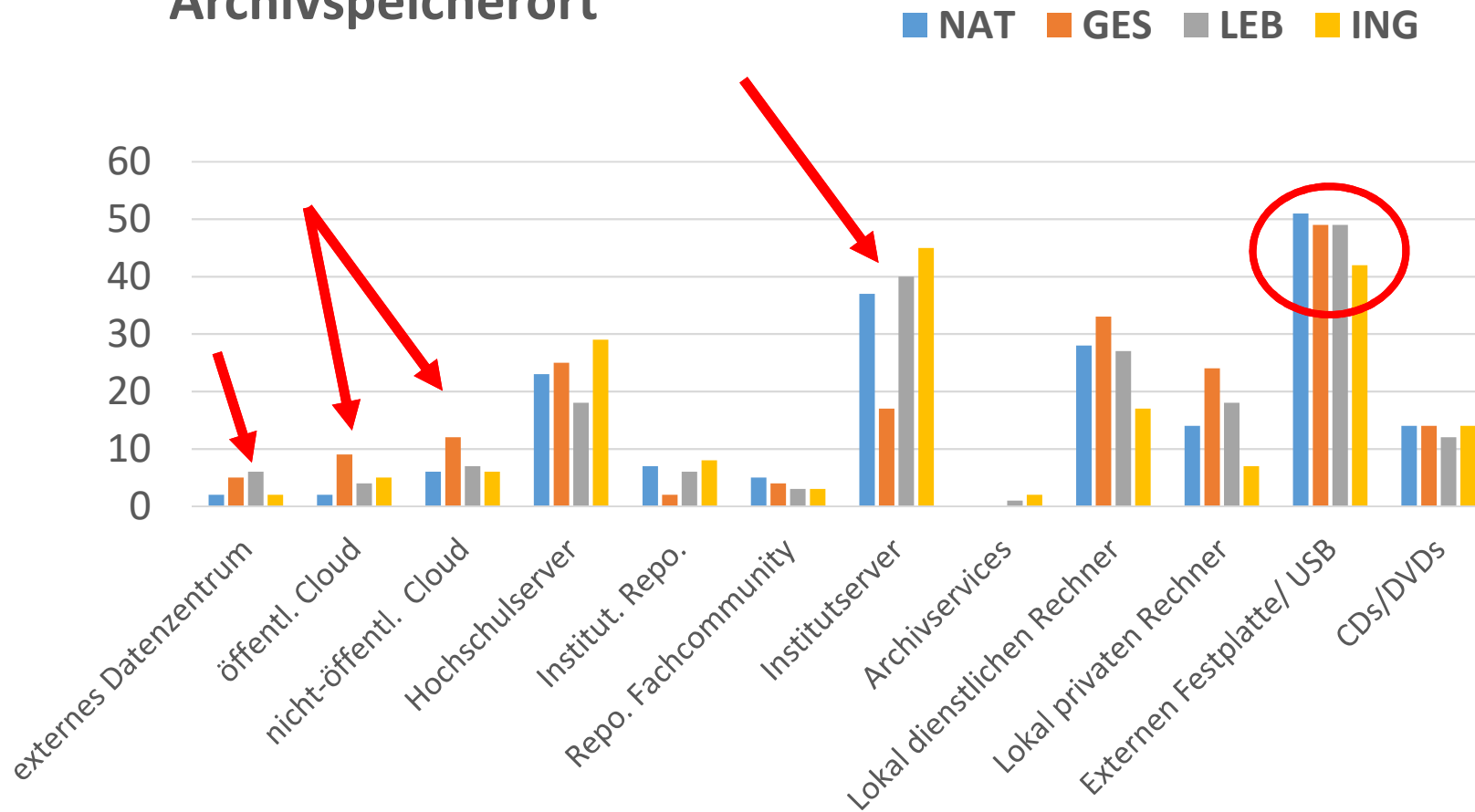
Wo/Wie archivieren Sie die Forschungsdaten (für mindestens 25 Jahre), die im Rahmen eines typischen Forschungsprojekts anfallen?

40-55% der Daten
in Archivierung
(25 Jahre)

Quelle: Brenger, Bela et al. (2019): UNEKE: Forschungsdatenspeicherung - Praxis und Bedarfe: Online-Survey 2019. CC BY 4.0, Online unter: <https://doi.org/10.17185/duerpublico/70259>



Archivspeicherort



Quelle: Brenger, Bela et al. (2019): UNEKE: Forschungsdatenspeicherung - Praxis und Bedarfe: Online-Survey 2019. [CC BY 4.0](https://doi.org/10.17185/duerpublico/70259), Online unter: <https://doi.org/10.17185/duerpublico/70259>

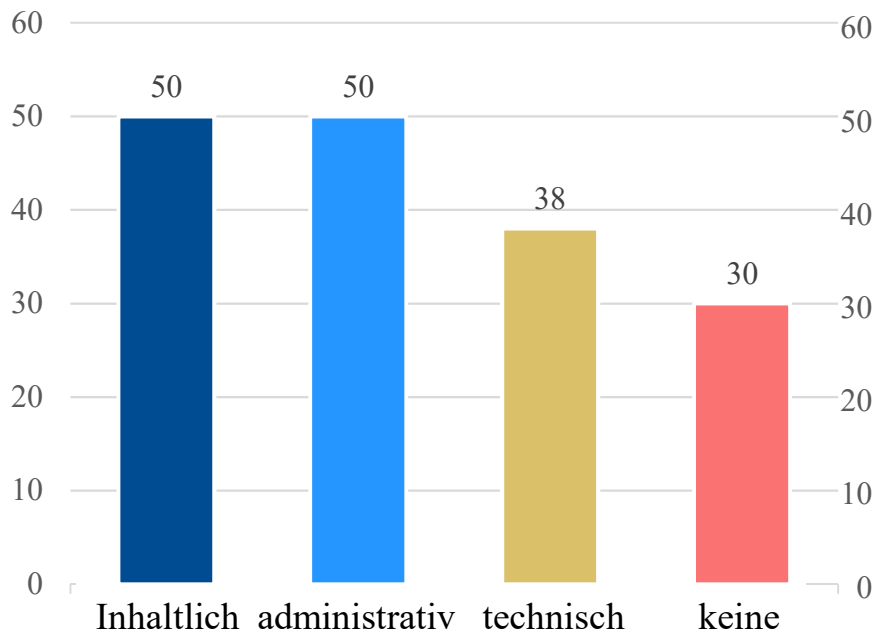
Metadaten:

Welche Metadaten erfassen Sie zur Wiederauffindbarkeit und Zuordnung Ihrer Daten?

Metadatenstandards:

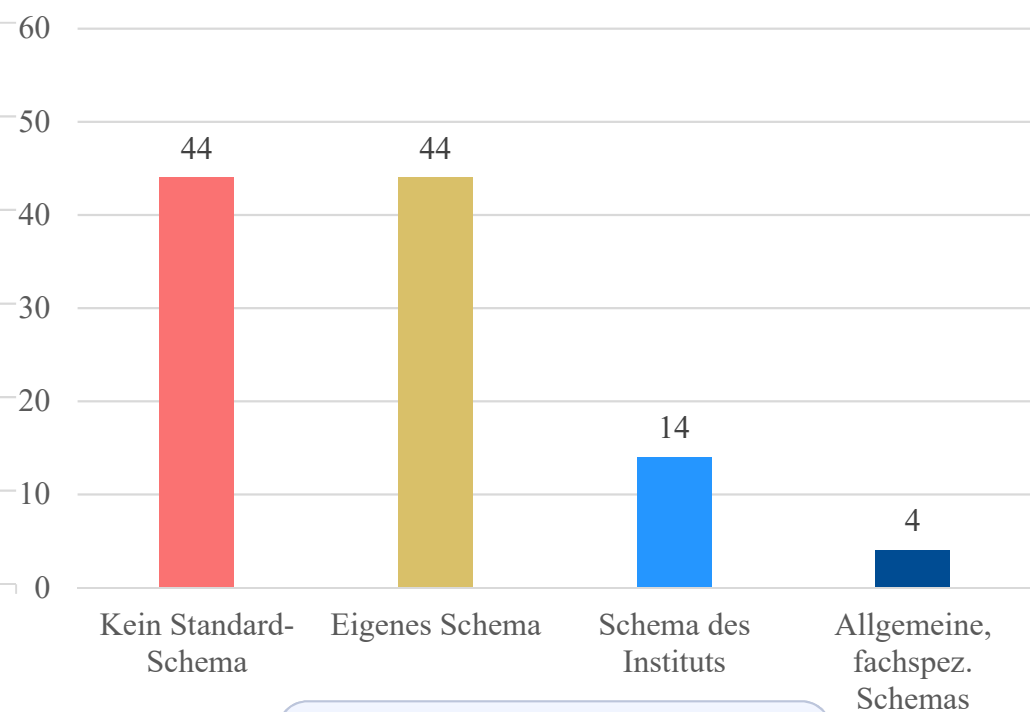
Nutzen Sie zur Erfassung der Metadaten ein existierendes Schema bzw. Standard?

Verwendete Metadaten



70 % der
Forschenden
erfassen Metadaten

Metadatenstandards



18 % der Forschenden
verwenden
Metadaten in Absprache

Quelle: Brenger, Bela et al. (2019): UNEKE: Forschungsdatenspeicherung - Praxis und Bedarfe: Online-Survey 2019. CC BY 4.0, Online unter: <https://doi.org/10.17185/dupublico/70259>

Workshops und
Umfrage liefern
komplementäre
Informationen

Hoher Beratungs-
bedarf & Ausbau
von Metadaten

Awareness
hochschulintern
& Kontaktaufbau
Forschende

Lessons learned

Anhaltspunkte für
FDM-Angebote &
Bedarf in Zahlen

UDE: Grundlage
für Gründung
der FDM-
Servicestelle



...WO stecken DEINE
Forschungsdaten
in 10 Jahren?!



...DEIN Backup?



...gute
wissenschaftliche
Praxis???

*Vielen Dank für Ihre
Aufmerksamkeit 😊
www.uneke.de*



Ordentliches
Forschungsdatenmanagement?!



...Forschungsdatenmanagement?



WO
liegen
DEINE
Daten...

...in
10 Jahren?

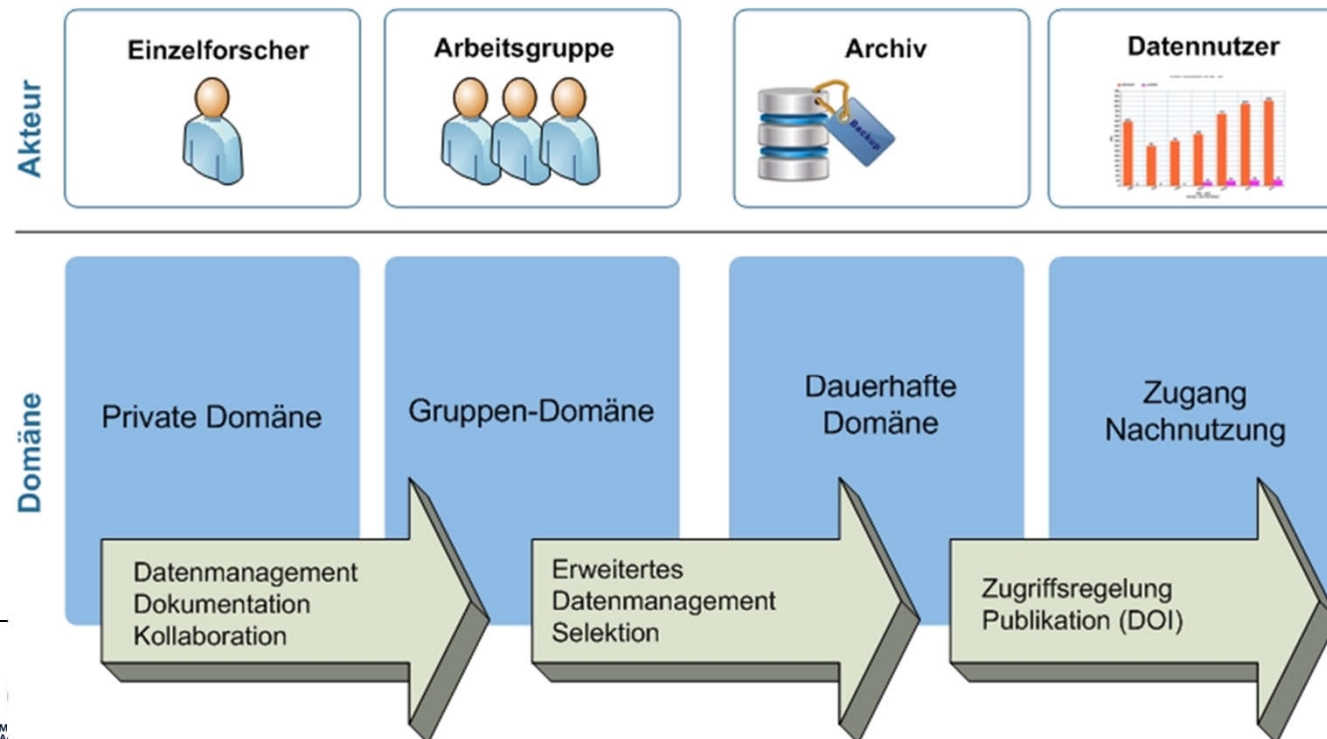


...DATEN FÜR DIE EWIGKEIT

- 70% der Forscher verwenden Metadaten
- Nur 18% verwenden dafür gemeinsame Schemata / Standards
- Speicherpraxis korreliert mit der zur Verfügung stehenden Infrastruktur
- Fachdisziplinäre Unterschiede sind weniger deutlich ausgeprägt als erwartet
- Über 50% sind nicht bereit Daten nach der Verwertung zu veröffentlichen und zu teilen obwohl Mehrwerte einer offenen Datenkultur bekannt sind
- nur 4% der Forscher nutzen Repositorien & Datenzentren
- ~ die Hälfte der Daten sollen archiviert werden
- Auf 40-50% der Daten müssen mehrere Personen zugreifen
- Hochschultypisch: hoher Datenaustausch an FH

Orientierung am Domänenmodell

- Aus der Perspektive der Speicherinfrastruktur
- Speicherbedarf und Verlauf während eines typischen Forschungsprojekts



Quelle:
Erweitertes
Domänenmodell
(Eigene
Darstellung in
Anlehnung an
WissGrid:
„Leitfaden zum
Forschungsdaten-
Management“,
2011

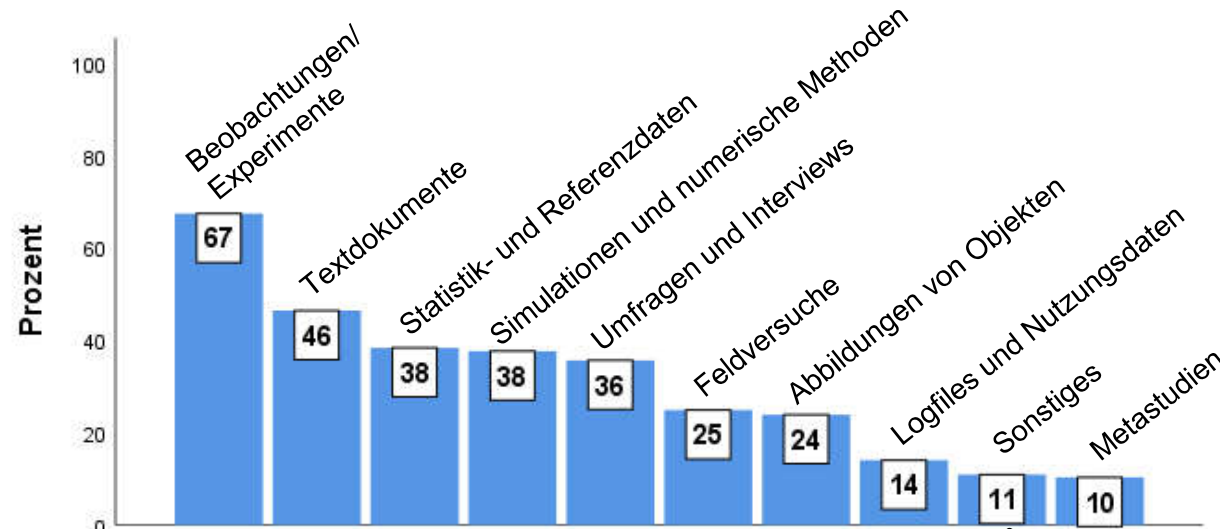
Datengenerierung, Speicherort und Zugriff



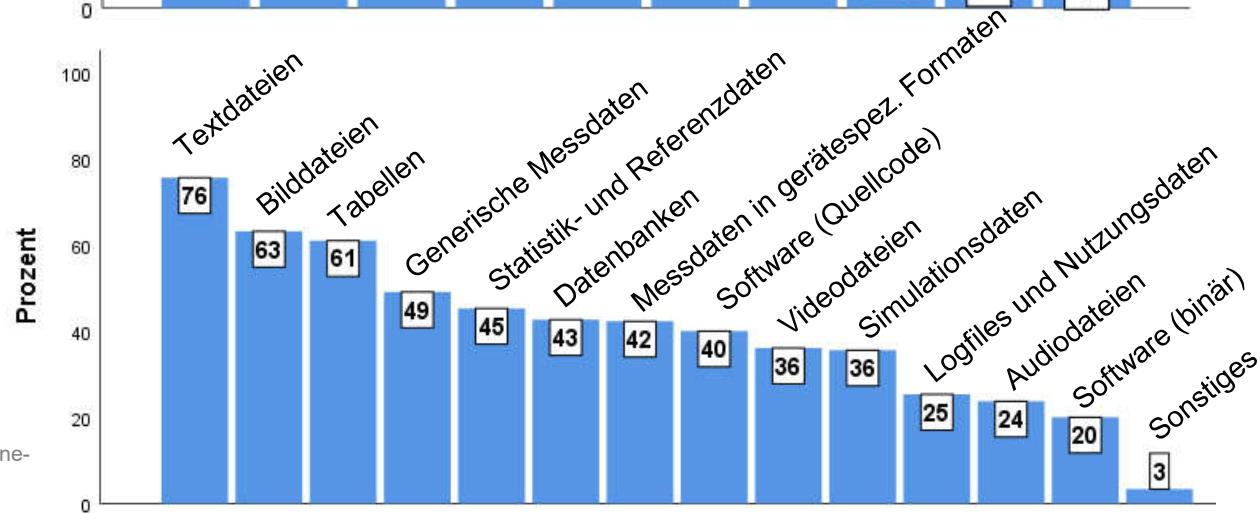
Datenquellen und Datenformate:

Auf welche Art und Weise bzw. aus welchen **Quellen** gewinnen Sie Ihre Forschungsdaten? Welche Datenformate fallen dabei an?

Vielfalt:
Quellen
von
Daten



Diverse
Formate
von
Daten

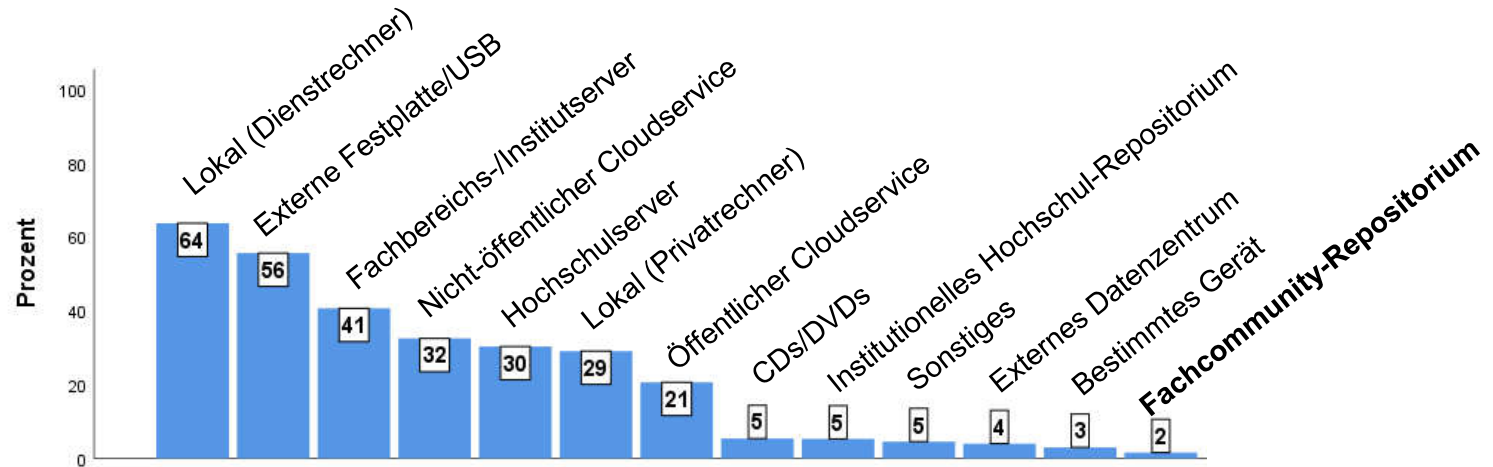


Quelle: Brenger, Bela et al. (2019): UNEKE: Forschungsdatenspeicherung - Praxis und Bedarfe: Online-Survey 2019. CC BY 4.0, Online unter: <https://doi.org/10.17185/dupublico/70259>

Speicherort:

Wo speichern Sie die Forschungsdaten eines typischen Projektes?

Geringe
Nutzung von
Datenzentren
und
Repositories

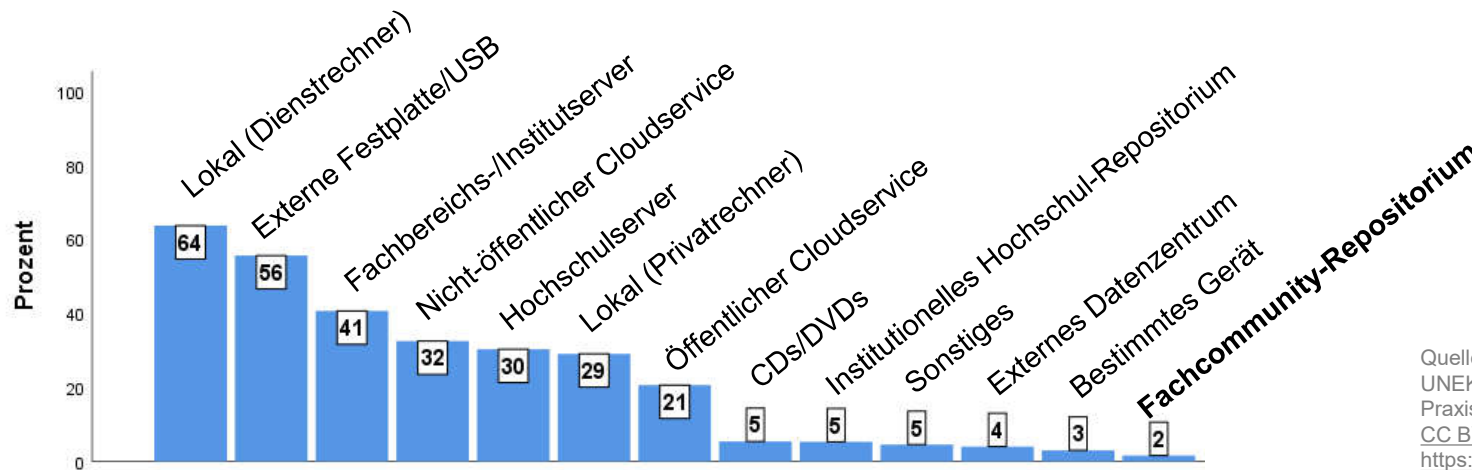
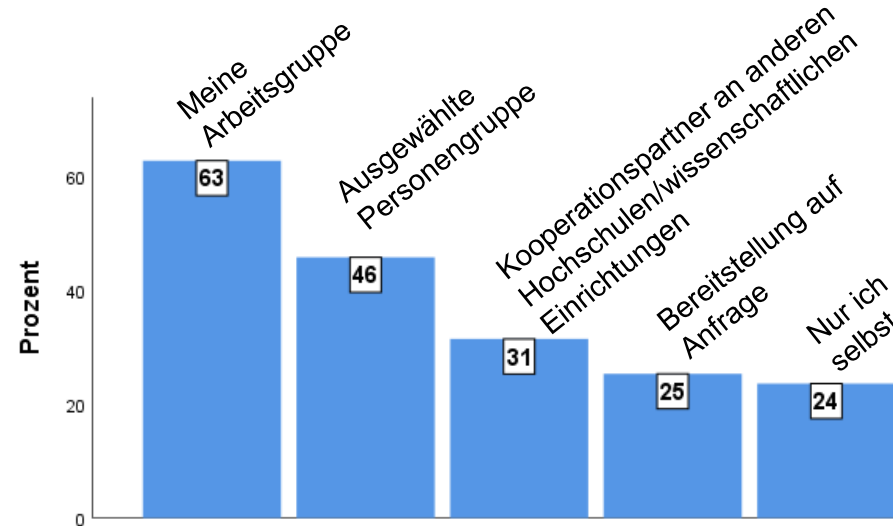


Quelle: Brenger, Bela et al. (2019): UNEKE: Forschungsdatenspeicherung - Praxis und Bedarfe: Online-Survey 2019. CC BY 4.0, Online unter: <https://doi.org/10.17185/dupublico/70259>

Speicherort und Zugriff

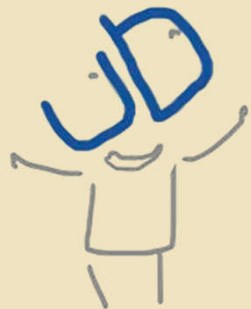
Wer darf bzw. kann in der Regel auf Ihre Forschungsdaten zugreifen?

Auf 40-50%
der Daten
Zugriff von
mehreren
Personen



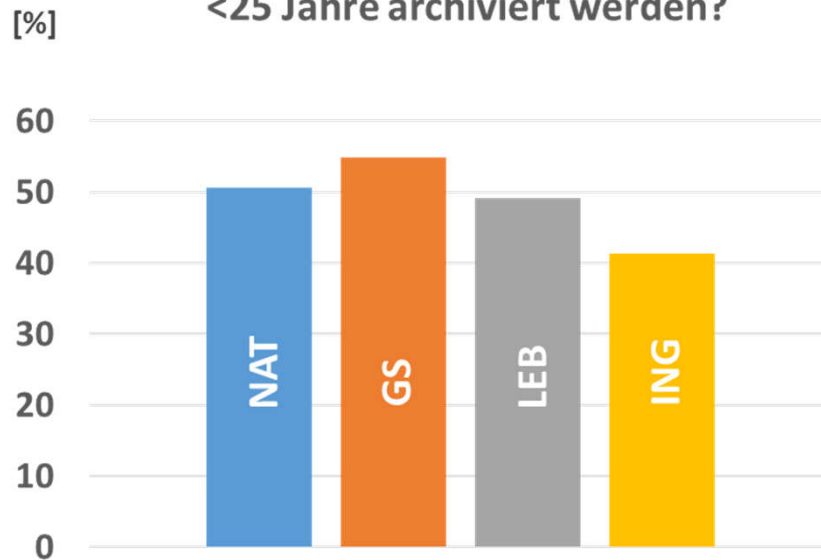
Quelle: Brenger, Bela et al. (2019):
UNEKE: Forschungsdatenspeicherung -
Praxis und Bedarfe: Online-Survey 2019.
CC BY 4.0, Online unter:
<https://doi.org/10.17185/duerpublico/70259>

Disziplinspezifische Unterschiede

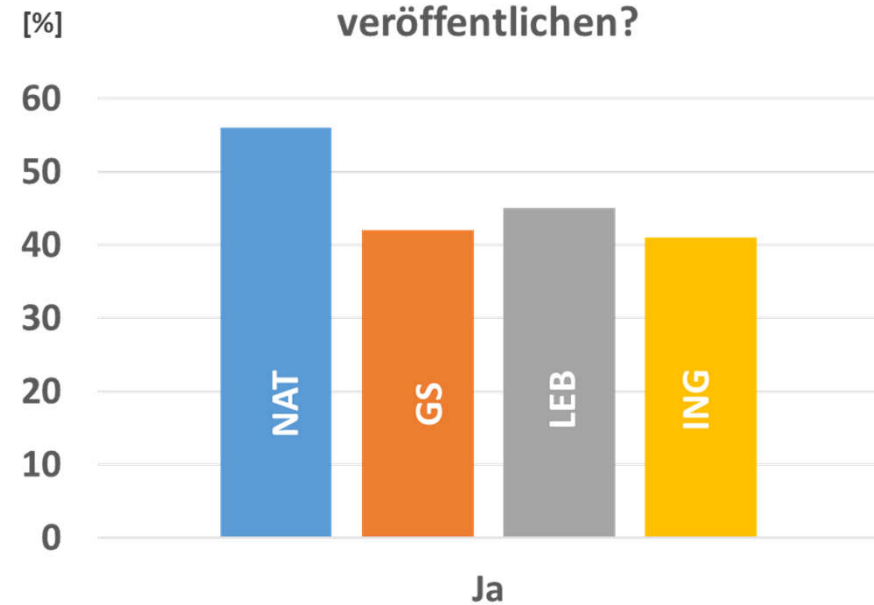


Archivierung und Publikation – disziplintypisch?

Welcher Anteil der Gesamtdaten soll <25 Jahre archiviert werden?



Rohdaten nach Erstverwendung veröffentlichen?



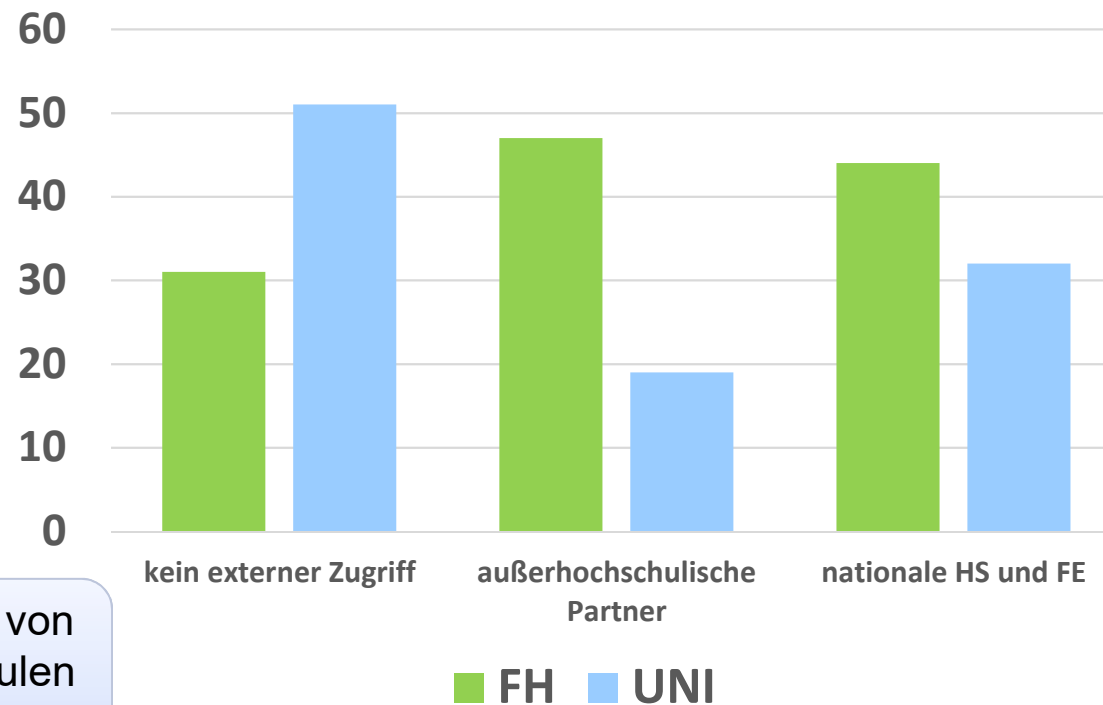
Gesamtdatensatz Stand November 2018 (11 Hochschulen ~ 1400 Teilnehmer)

Quelle: Brenger, Bela et al. (2019): UNEKE: Forschungsdatenspeicherung - Praxis und Bedarfe: Online-Survey 2019. CC BY 4.0, Online unter: <https://doi.org/10.17185/duerpublico/70259>

Hochschultypische Aspekte



Zugriff auf Daten



Unterschiede von
Fachhochschulen
und Universitäten

Gesamtdatensatz Stand November 2018 (11 Hochschulen ~ 1400 Teilnehmer)

Quelle: Brenger, Bela et al. (2019): UNEKE:
Forschungsdatenspeicherung - Praxis und Bedarfe: Online-
Survey 2019. CC BY 4.0, Online unter:
<https://doi.org/10.17185/dupublico/70259>

Ergebnisse der Begleitstudie

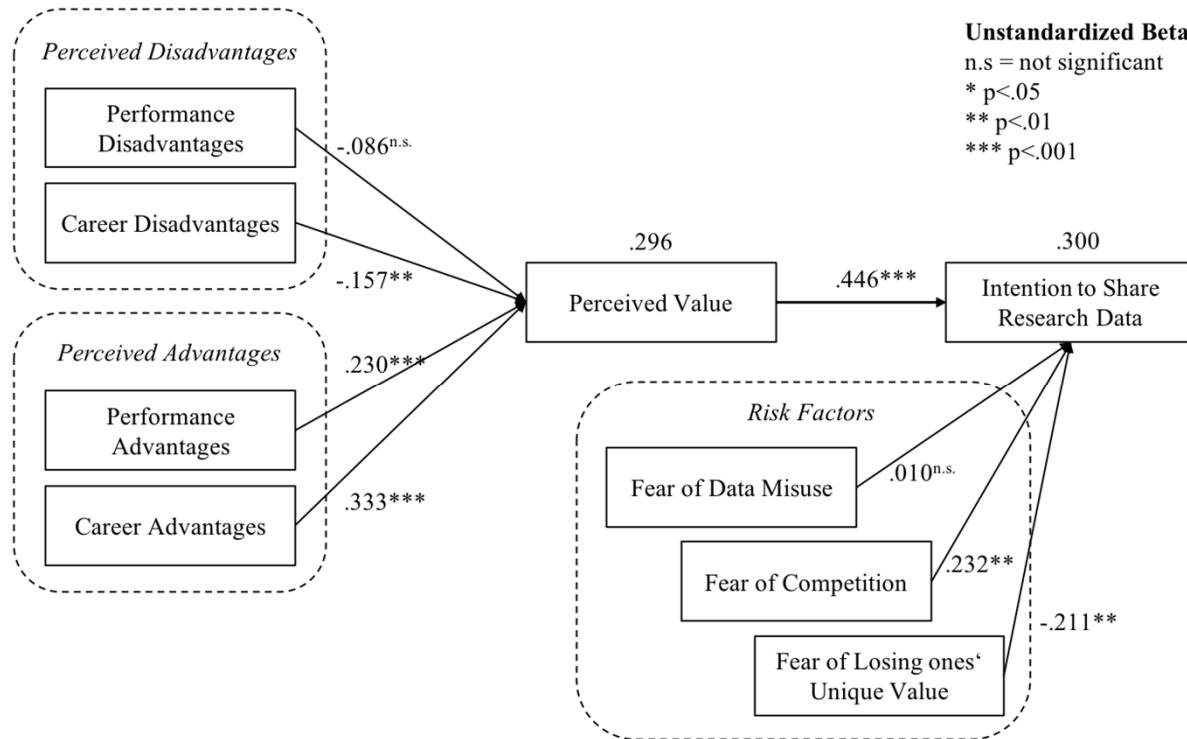


Kernaussagen der Begleitstudie

Mehrwerte die mit einer offenen Datenkultur einhergehen sind bekannt

Angst vor Datenmissbrauch und Konkurrenz führen zu Ablehnung.

Angst vor Konkurrenz kann je nach Disziplin treibender Faktor sein.



Quelle: Wilms, Konstantin (2019): A Multidimensional Perspective on the Acceptance of Organizational Communication and Collaboration Systems, S.167. Online unter: [10.17185/duerpublico/70204](https://doi.org/10.17185/duerpublico/70204)

Fokus sollte nicht sein die Vorteile zu verkaufen, sondern Sicherheiten zu geben.