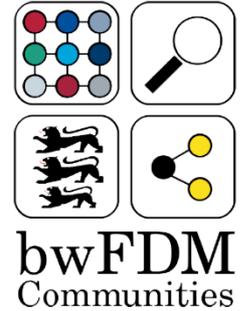


# bwFDM-Communities – Wissenschaftliches Datenmanagement an den Universitäten Baden-Württembergs

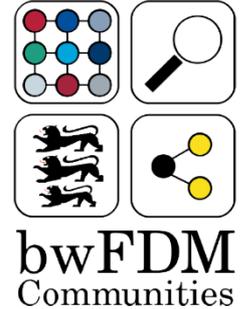


Forschungsdatenmanagement an den Universitäten des Landes Baden-  
Württemberg – Bestandsaufnahme und Empfehlungen

*Präsentation des Projektes „bwFDM-Communities“*

Frank Tristram (Projektleiter)

# „Abschlusspräsentation“ von bwFDM-Communities ist die Auftaktveranstaltung zur Ergebnisnutzung



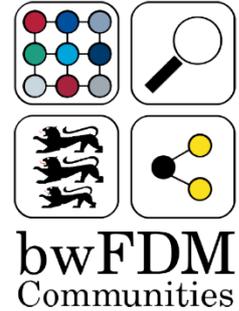
In diesem Vortrag werde ich nicht:

- Alle Resultate von 15 Mannjahren abhandeln
- Eine unterschriftsreife Strategie vorstellen

Was man dennoch aus dem Vortrag mitnehmen kann:

- **Richtungen** wo es brennt
- **Anregungen** zur Bedarfsdeckung
- Einen **Eindruck** was für Informationen in unseren **Daten** vorliegen (5%)

# Themenübersicht

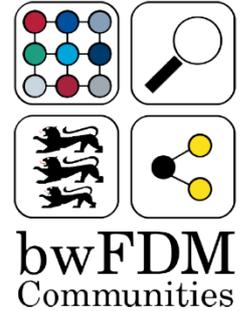


Kapitel	Unterthemen	User-Story-Anzahl
Strukturelle Rahmenbedingungen	3	782
Digitale Forschungsdatenverarbeitung	3	612
Datenerhebung und -austausch	3	303
Informationsbewahrung	3	529
Weitere Aspekte	6	1089

Etwa 20% der User-Stories sind mehrfach zugeordnet

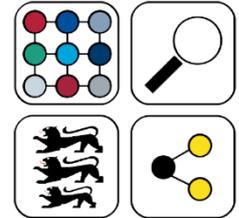
# Übersicht

## Strukturelle Rahmenbedingungen



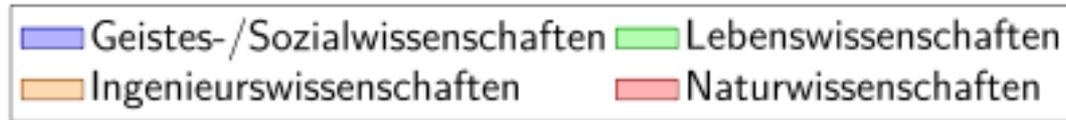
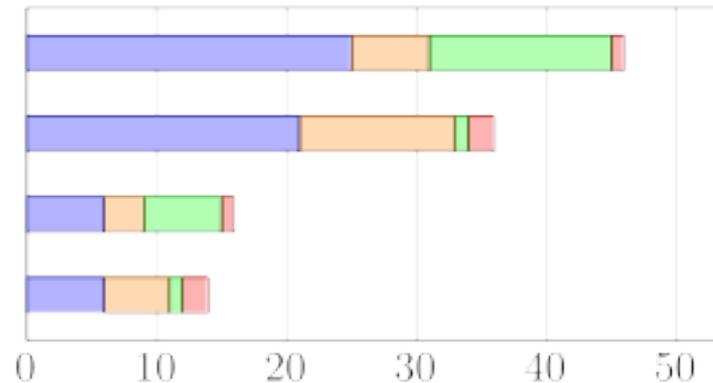
- Urheberrecht und Datenschutz (112)
- Informationsangebote (457)
- Anreize zur Datenpublikation (213)

# Urheberrecht und Datenschutz



bwFDM  
Communities

Klärung von Rechtsfragen (46)  
Rechtliche Bestimmungen als Forschungshindernisse (36)  
Datenschutzkonforme Infrastruktur (16)  
Schutz- und Vermarktungsrechte für produzierte Daten (14)



## ■ Datenschutz

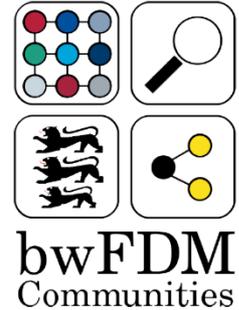
- Beratung, Zweckbindung, Infrastruktur, Abgabe

## ■ Urheberrecht

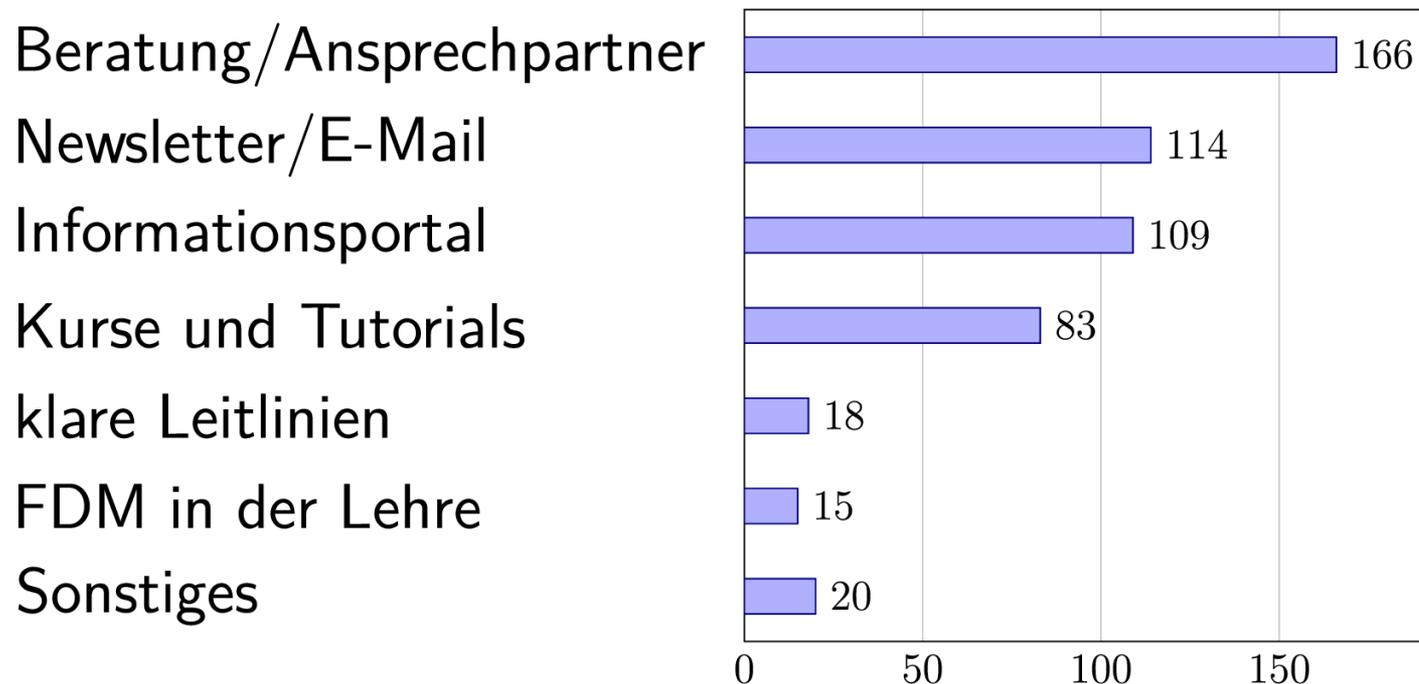
- Beratung, Big Data, Panorama, wem gehören Forschungsergebnisse?

# Informationsangebote

- Informationsangebote zählen zu den am stärksten nachgefragten Bedarfen

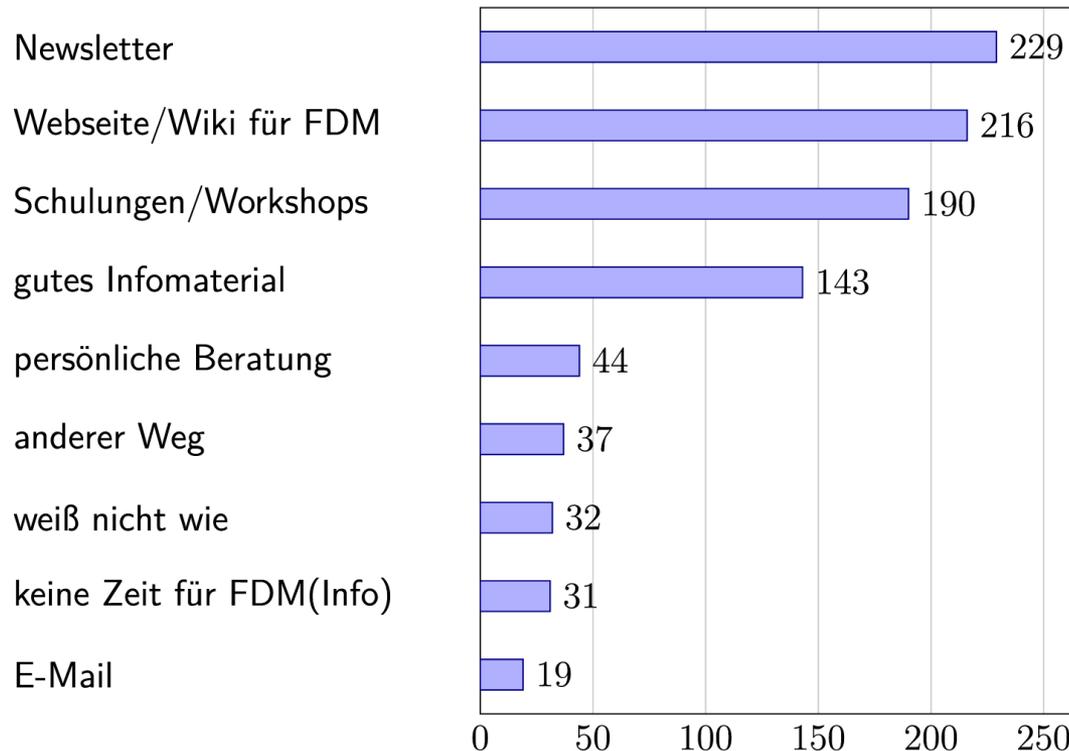
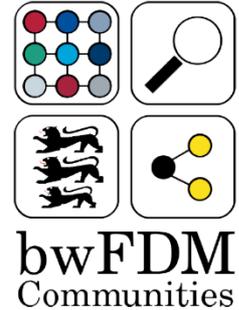


## Themenfelder innerhalb der User-Stories



# Informationsangebote

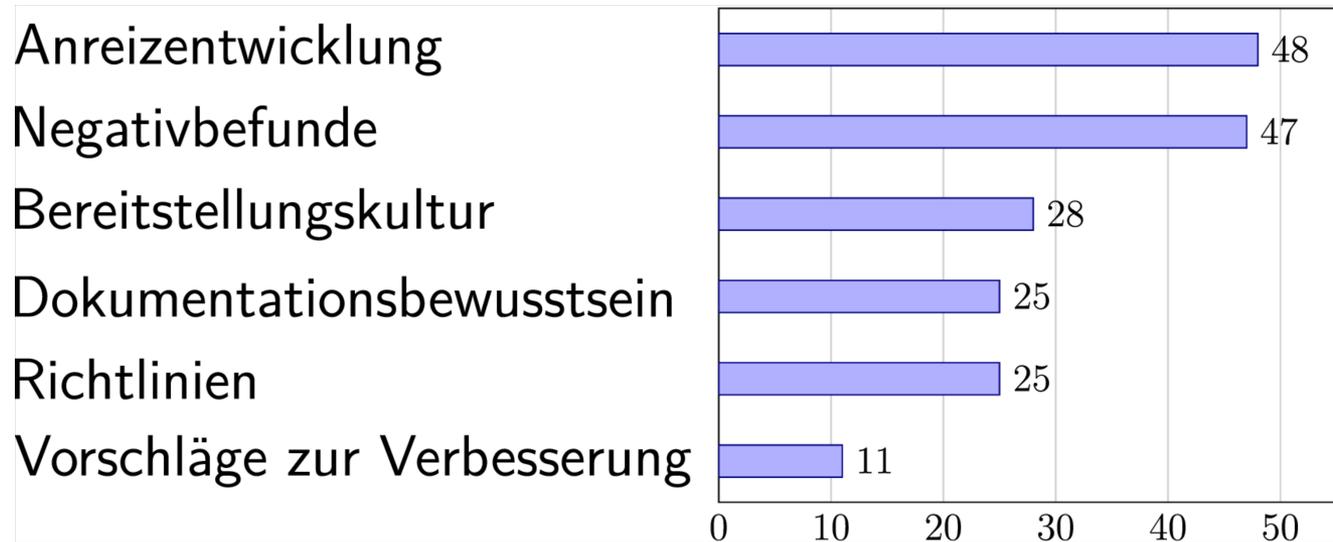
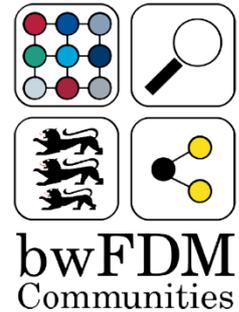
„Welche Informationsmethode würden sie bevorzugen?“  
(Frage 25)



- Die Interviewdaten enthalten Details darüber für welche Art von Information welcher Kanal bevorzugt wird.

# Anreize zur Datenpublikation

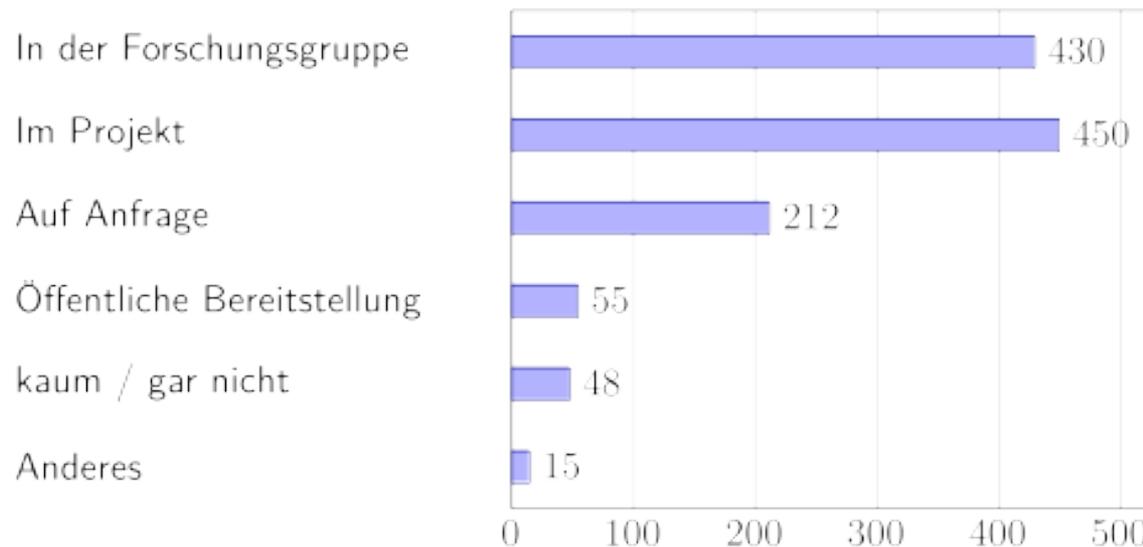
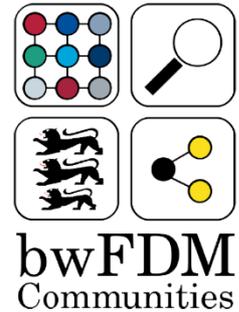
Themenübersicht anhand der User-Stories



- Es fehlen Anreize, z.B. bessere Berücksichtigung bei Berufungen und Mittelverteilung, um hochwertige Datenbereitstellung zu honorieren.
- Die Nachvollziehbarkeit von Publikationen durch die Veröffentlichung der zugrunde liegenden Daten zu sichern, wäre ein erstes Zwischenziel.

# Anreize zur Datenpublikation

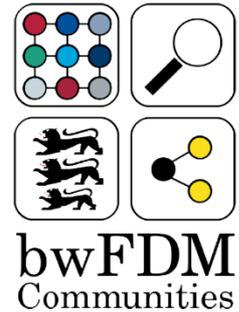
*In welchem Rahmen haben Sie schon eigene Primärdaten ausgetauscht? (Frage 20a)*



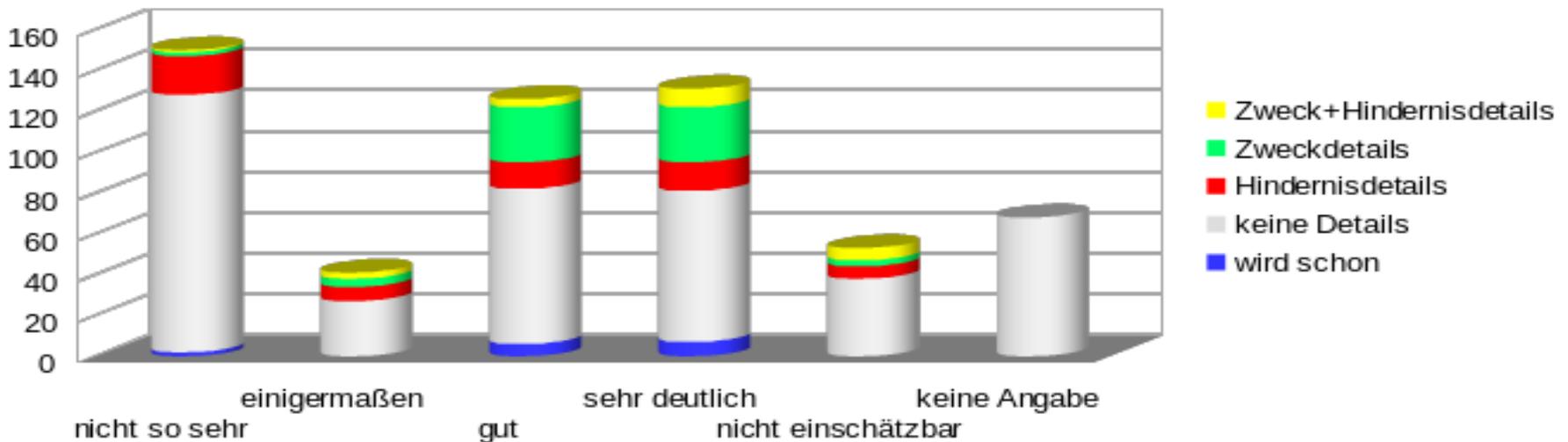
- Forschende, die allein arbeiten, sind die Ausnahme, die die Primärdaten zugänglich machen aber auch, obwohl solche Daten relevant sind.
- Fast 50% sagen, sie haben Daten, die auch für andere interessant sein könnten, geben diese aber aus verschiedenen Gründen nicht weiter.

# Anreize zur Datenpublikation

## 56 Jahre Publikationsbias\* und nur wenige Lösungen für „negative Ergebnisse“



Wie sehr würden Sie von einem Überblick über "Sackgassen" anderer Forscher profitieren?

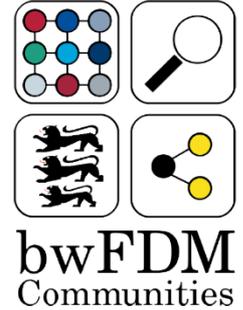


- Such- und Einarbeitungs- sowie Aufbereitungs- und Honorierungsproblem
- Lösungsvorschläge: Fachspezifische Journale, Diskussionseinladungstool, Vorabpublikation der Methode

\*Publication Decisions and Their Possible Effects on Inferences Drawn from Tests of Significance-Or Vice Versa (1959) [doi:10.2307/2282137](https://doi.org/10.2307/2282137)

# Übersicht

## Digitale Forschungsdatenverarbeitung

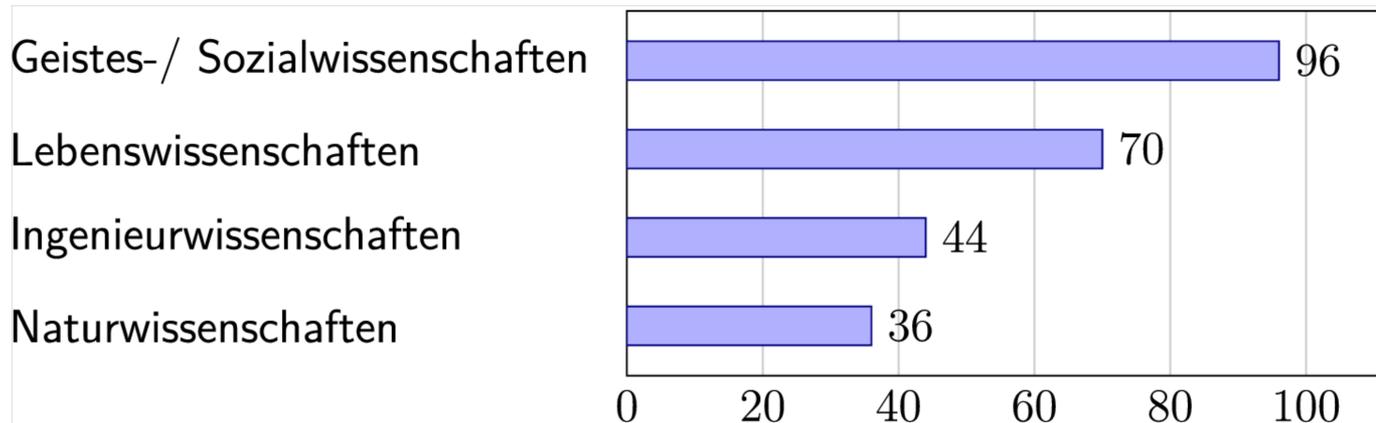
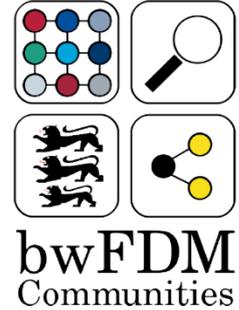


- Forschungsnahe Unterstützung (246)
- Wissenschaftliche Softwaretools (162)
- Formate und einheitliche Standards (204)

# Forschungsnahe Unterstützung

Wenn Forschende mit ihrem Know-How an digitale Grenzen stoßen und sich Hilfe wünschen.

*Web-Crawling, Datenkompression am Messinstrument, Konvertieren von Datenformaten, das Aufsetzen einer (Web-)Datenbank, statistisch komplexe Analysen ...*

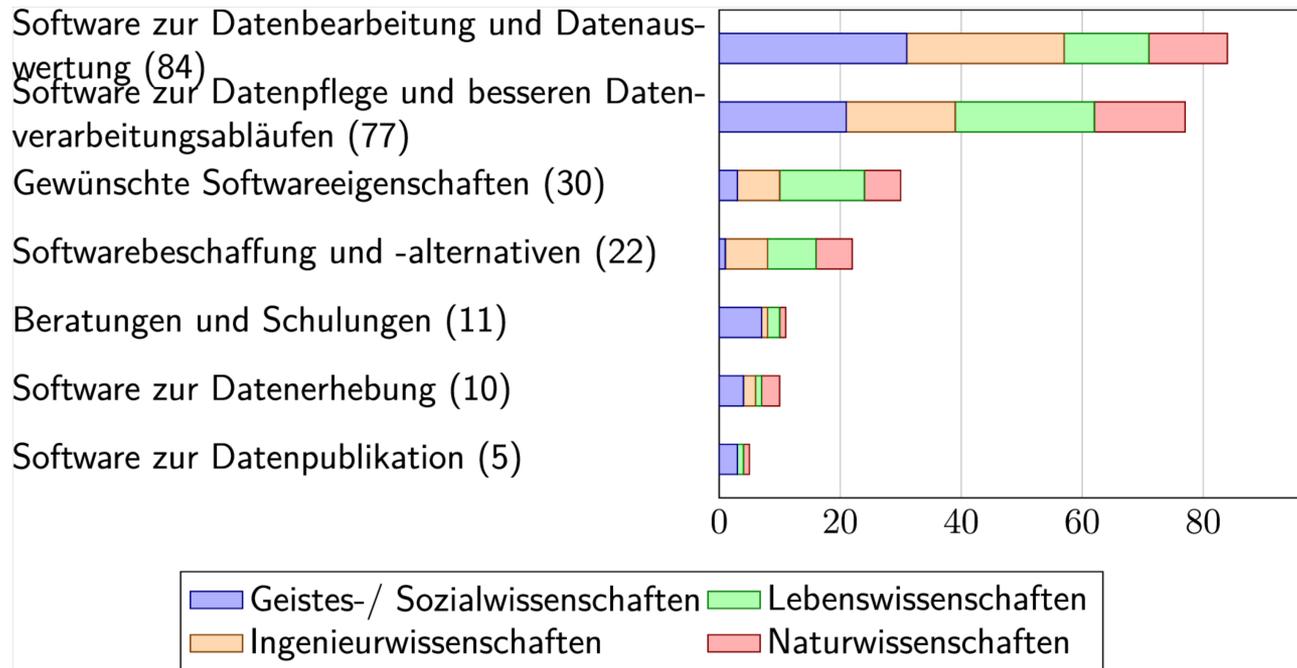
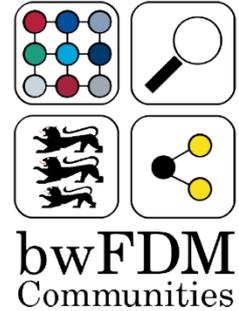


Lösungsvorschläge:

- Ausleihbare IT-Hilfskräfte
- Sience@home
- Verantwortliche für Forschungsdaten
- In der Lehre auf Forschungstätigkeit vorbereiten

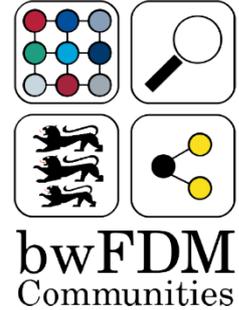
# Wissenschaftliche Softwaretools

In welchen Bereichen ein Mangel an Software den Datenumgang unvollkommen lässt



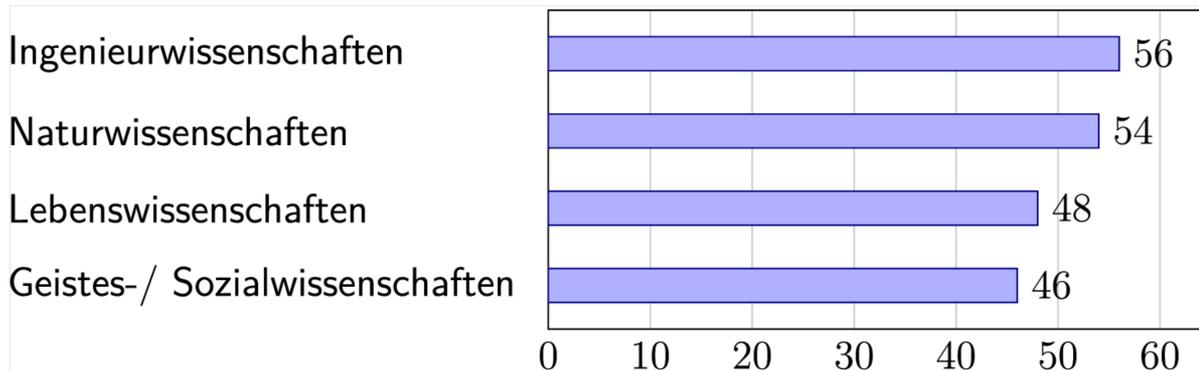
- Softwareentwicklung sollte mit Communitys zusammen stattfinden
- Bei Bedarf praxisnahe Schulungen zu Software (Kapitel 1 beachten)

# Formate und einheitliche Standards



## Wobei sollen Standards helfen?

- Erhöhung der Datenqualität, der Vergleichbarkeit und Verständlichkeit von Daten (28)
- Vereinfachung des Datenaustauschs (19)
- Erhöhung der Nachhaltigkeit, sowie Erleichterung der Datennachnutzung (16)
- Senkung des Aufwands für die Konvertierung, sowie Aufbereitung von Daten (15)
- Erleichterung der Interoperabilität der Formate (12)

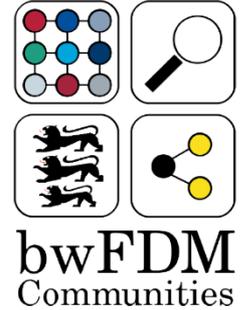


## Lösungsvorschläge

- Förderung der Nutzung von einheitlicher Software und Formate durch Landeslizenzen oder Open-Source-Förderung.
- Vermittlung von Richtlinien was gute Formate ausmacht.
- Einigung auf einheitliche Sprachen/Meta-Formate, um alle Formate abzubilden

# Übersicht

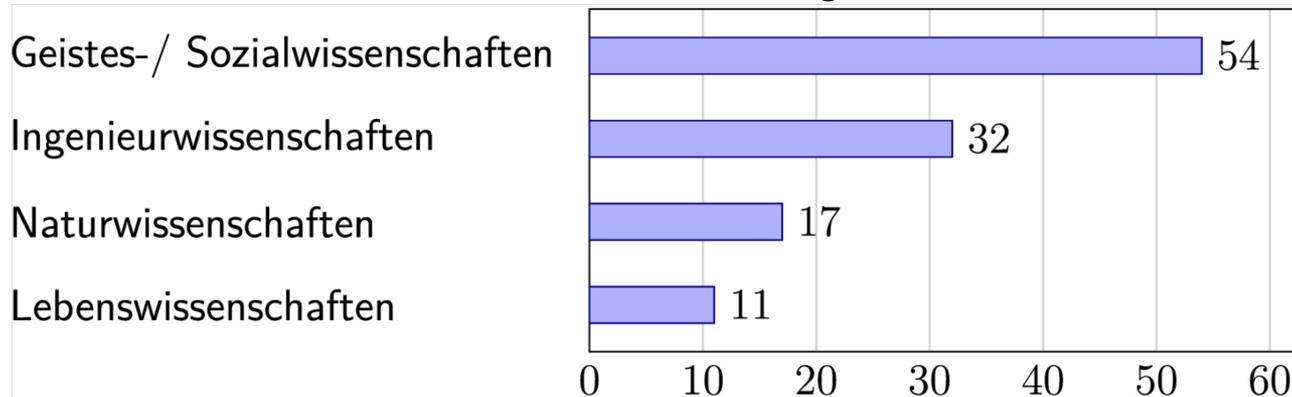
## Datenerhebung und -austausch



- Daten außeruniversitärer Einrichtungen (115)
- Erwartungen an die Digitalisierung (41)
- Wissenschaftliche Zusammenarbeit (147)

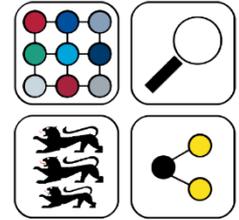
# Zugang zu Forschungsexternen Daten

Sozialwissenschaft, Wirtschaftswissenschaft und Architektur/Bauwesen sind Fachgebiete, die häufig externe Daten z.B. aus Behörden und Ämtern benötigen.



- Zugang zu offenen und geschlossenen Verwaltungsdaten aus Behörden und öffentlichen Einrichtungen würde durch zentrale Portale und einheitlichen Antragswegen erheblich verbessert.
  - Kontaktbildung und Anreize für die Industrie Daten bereitzustellen wäre erwünscht
  - Es wäre zu überlegen, ob gemeinsame Datenbereitstellungsdienste für Wissenschaft und öffentliche Verwaltung an manchen Stellen Synergien schaffen könnten.
- Nach Gespräch mit Open-Government-Initiative des Bundes sollen nun auch wissenschaftliche Bedürfnisse mehr Beachtung finden. Einige User-Stories werden als Wünsche direkt an die jeweils entsprechende Behörde weitergeleitet.*

# Wissenschaftliche Zusammenarbeit

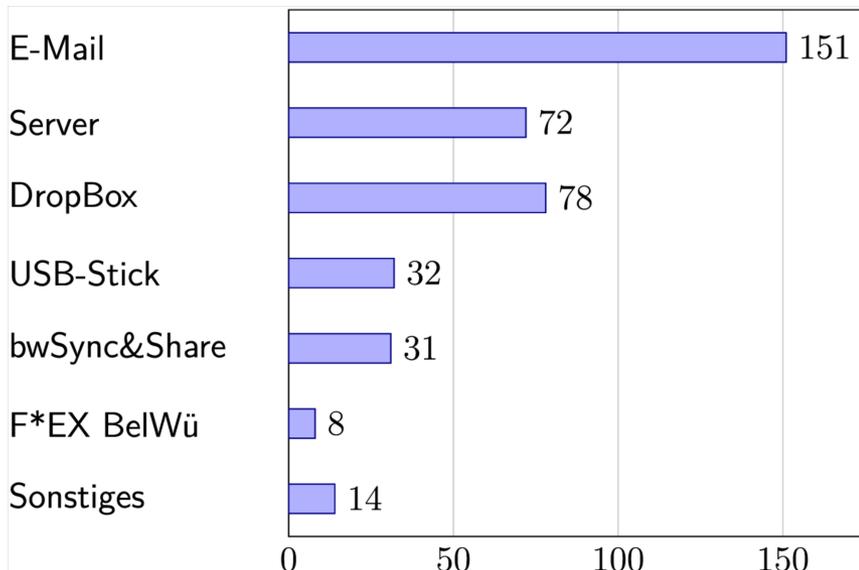


bwFDM  
Communities

## Praktische Datenverwaltung (54)

„Als Forschende, die größere Datenmengen mit Projektpartnern kollaborativ verarbeiten, wünschen wir uns eine sichere und zentrale Speicherlösung mit optionaler Metadatenstruktur und Versionsverwaltung sowie flexibler Rechtevergabe und Schnittstellen zu weiterer IT-Infrastruktur (z.B. HPC), um ortsunabhängig mit übersichtlichen und Datenschutzkonformen Systemen arbeiten zu können.“

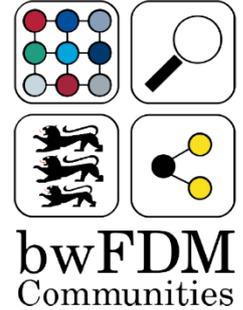
## Welche Programme oder Dienste nutzen Sie zum Austausch von Primärdaten?



- Erweiterungskonzept der aktuellen Datendienste anhand Bedarfen
- Einladende Infrastruktur zum Aufbau von Repositorien und virtuellen Forschungsumgebungen

# Übersicht

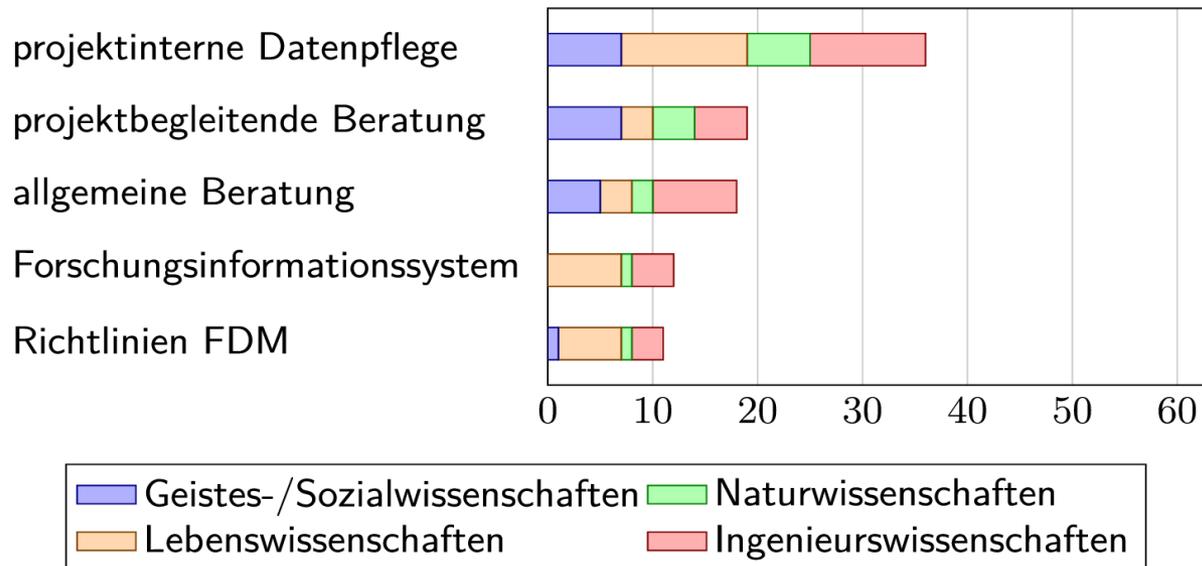
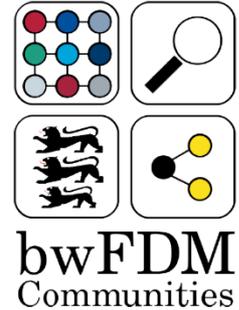
## Informationsbewahrung



- Projekt- und Datendokumentation (185)
- Repositorien für die Publikation von Forschungsdaten (237)
- Archivierung (107)

# Projekt- und Datendokumentation

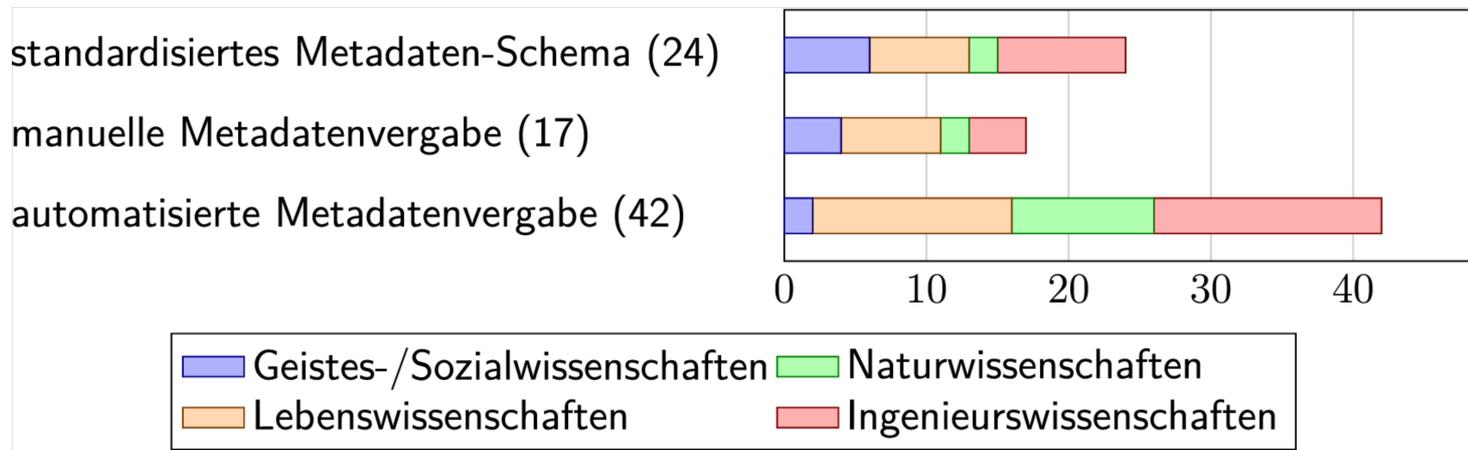
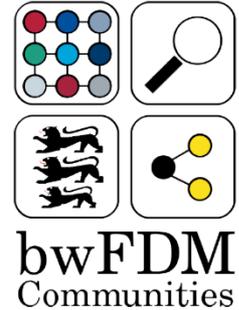
## Was fehlt für eine verlässliche Projektdokumentation?



- Die Mitarbeiterfluktuation macht es schwierig, auf vergangenen Projekten aufzubauen.
- Selbst für bemühte Institute fehlen Blaupausen für einheitliche Projektdokumentation, die nicht nur administrative, sondern auch nachnutzbar wissenschaftliche Inhalte erfassen.
- Halbautomatische Systeme und klare Verantwortlichkeiten sind unerlässlich, um auch in Zeiten hohen Arbeitsdrucks ein Dokumentationsgerüst zu erhalten.

# Projekt- und Datendokumentation

## Was fehlt für eine verlässliche Datendokumentation?



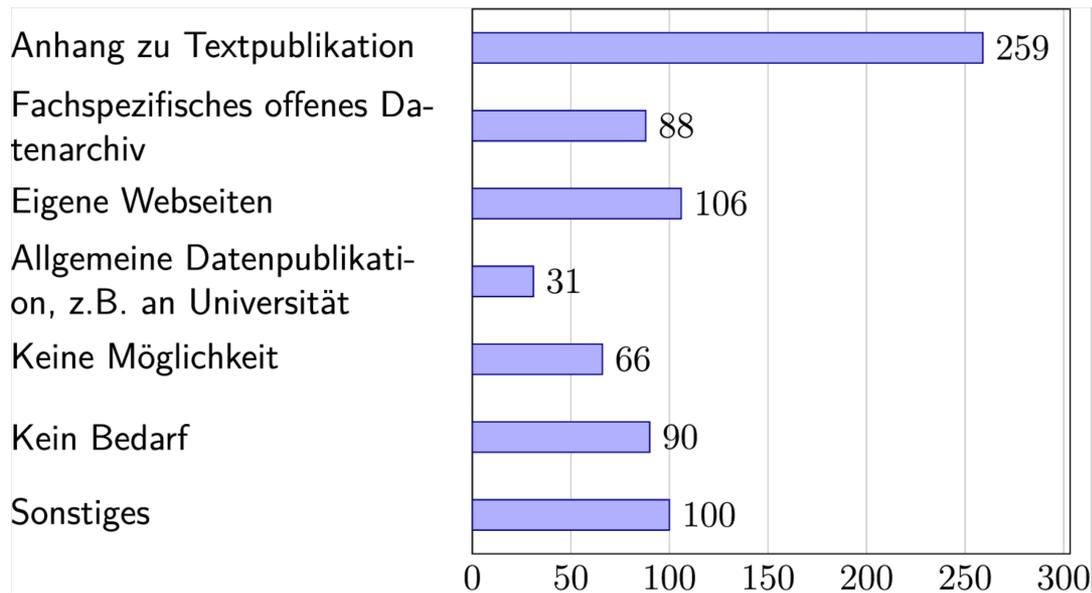
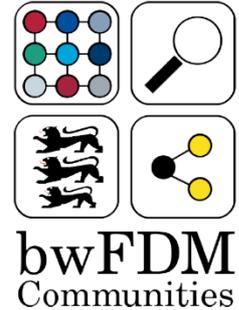
- Obwohl es in diesem Bereich einige Richtlinien zur Metadatenvergabe gibt, stellt die Umsetzung dieser Richtlinien offensichtlich eine Hürde im Forschungsalltag dar.
- In datenintensiven Disziplinen ist die automatisierte Metadatenvergabe besonders bedeutend.

### Für Projekt- und Datendokumentation:

- Workflows festlegen und etablieren, sodass jeder weiß was er zu tun hat.
- Best-Practices und Hilfen vermitteln und anbieten.
- Neue DFG-Ausschreibung „Forschungsdaten in der Praxis“

# Repositorien für die Datenpublikation

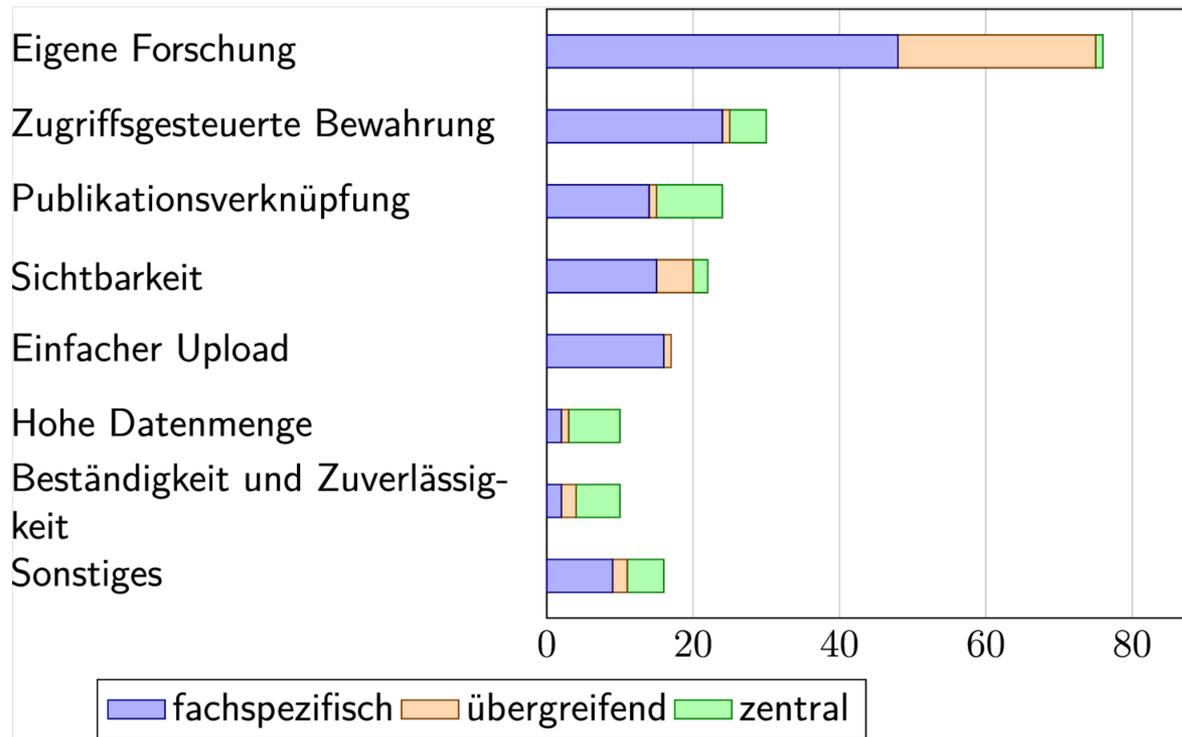
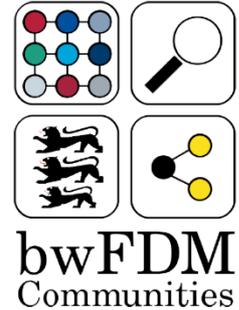
*Welches Medium nutzen Sie zur Veröffentlichung Ihrer Ergebnisdaten?*



- Weniger als 20% der Befragten sagen, sie haben keinen Bedarf Ergebnisdaten zu veröffentlichen.
- Auch Forschende die bisher andere Methoden nutzen wünschen sich Repositorien. (über 200 User-Stories zu Repositorien)

# Repositorien für die Datenpublikation

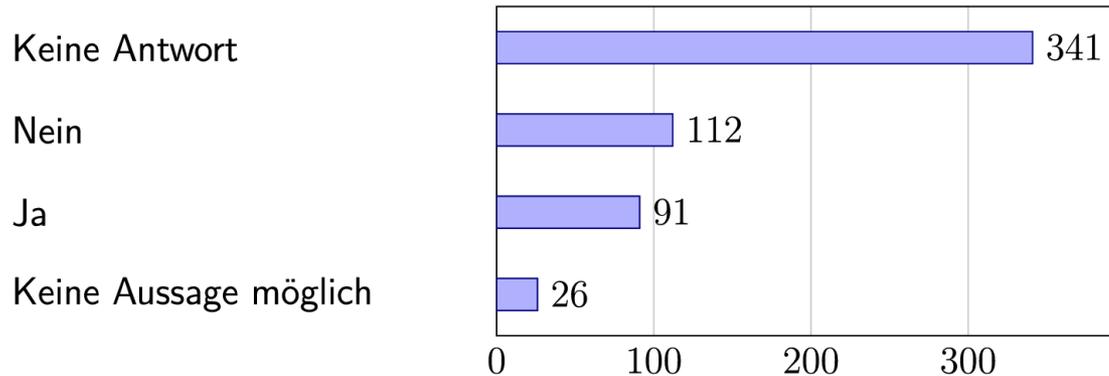
Wofür werden Repositorien gebraucht? (User-Stories)



- Fachübergreifende Repositorien werden als Datennutzer für die eigene Forschung gebraucht.
- Zentrale Repositorien sind als Datenbereitsteller aus verschiedenen Gründen gefragt.
- Ein ideales Repository berücksichtigt alle oberen Aspekte.

# Archivierung

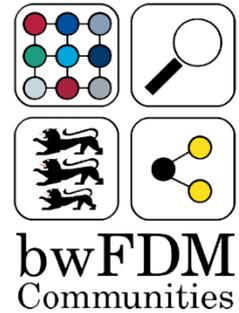
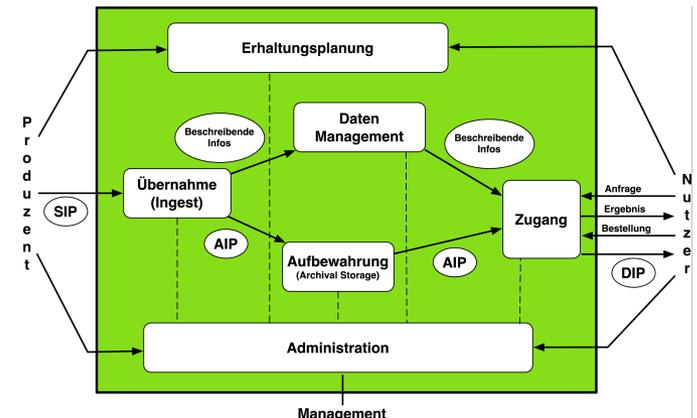
„Falls "Messungen" (auch Umfragen) an "realen Dingen" oder einmalige Beobachtungen vorgenommen werden: Könnte das für jemand in 100 Jahren interessant sein?“ (Frage 19b)



- Etwa 17% der Forschenden sagen, dass sie Daten haben, die womöglich nicht wiederholbar sind und auch noch in 100 Jahren relevant sein könnten.

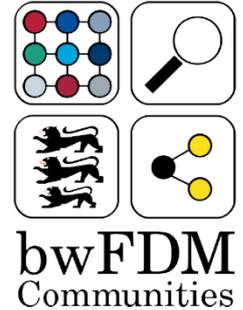
- Das OAIS\* Modell der NASA/ESA beschreibt welche funktionellen Einheiten eine Institution besitzen sollte, um erfolgreich archivieren zu können.

\*Open Archival Information System

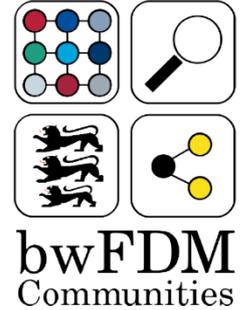


# FDM-ferne Aspekte

- Allgemeine IT-Infrastruktur (543)
- Allgemeiner IT-Support (220)
- Lizenzierung (168)
- Open Science (112)
- Finanzierung (60)



# Vorbehalte gegen FDM



## Übersicht

- Die meisten Vorbehalte drehen sich um die Anforderungen zu Aufwand und Effizienz.
- Die geforderte Datenqualität soll zur verfügbaren Zeit in einem gesunden Verhältnis bleiben.
- Es sollen fachspezifische Besonderheiten beachtet und Freiräume gewahrt bleiben.

## Beispielaussagen

- „Als Wissenschaftler dessen Fachbereich von zu vielen Veröffentlichungen geflutet wird, wünsche ich mir, dass die Entscheidungsvollmacht der Lehrstühle welche Daten sie publizieren nicht durch keine generelle Publikationspflicht ersetzt wird, um (im Gegensatz zur Praxis bei papern) zu gewährleisten dass nur wichtige und hochwertige Daten veröffentlicht werden.“
- „Als Wissenschaftler wünsche ich mir, dass ich mir keine Gedanken um FDM machen muss, um mich nur auf die Inhalte meiner Forschung konzentrieren zu können.“

# Ausblick Veröffentlichung

- User-Stories online durchsuchbar zur Verfügung stellen
- Interviewinhalte anonymisieren und Antworten auf Fragen in Kurzform zur Verfügung stellen
- Bis Projektende Ansprechpartner für alle Einrichtungen in Baden-Württemberg
- Öffentliche Präsentation der Ergebnisse und Vorstellung der E-Science Projekte am 30.9.2015 in Karlsruhe (letzter Projekttag)
- Übergabe der Informationen an E-Science Projekte

