

Handlungsempfehlungen

zur Umsetzung der Forschungs-
daten-Policy der TU Berlin

November 2022

Am 23. Oktober 2019 hat der Akademische Senat der Technischen Universität Berlin die [Forschungsdaten-Policy der TU Berlin](#) beschlossen. Damit will die TU Berlin ihren jetzigen und zukünftigen Forschenden eine Orientierung zum Umgang mit Forschungsdaten an die Hand geben. Die *Handlungsempfehlungen zur Umsetzung der Forschungsdaten-Policy* enthalten praktische Hinweise zum Forschungsdatenmanagement (FDM) in verschiedenen Phasen eines Forschungsvorhabens. Der Akademische Senat hat den Handlungsanweisungen am 23. Oktober 2019 zugestimmt. Sie werden vom Servicezentrum Forschungsdatenmanagement gepflegt und kontinuierlich aktualisiert.

Inhalt

Zentraler Kontaktpunkt an der TU Berlin	3
Empfehlungen zum Umgang mit Forschungsdaten	4
Vorab: Zehn Fragen zum FDM in der Antragsphase	4
I. Planungs- und Antragsphase (vor dem Forschungsvorhaben)	5
II. Durchführungsphase (während des Forschungsvorhabens)	6
III. Abschlussphase (nach dem Forschungsvorhaben)	7

Abkürzungen	
Abt. V	Abteilung V – Forschung und Technologietransfer
DFG	Deutsche Forschungsgemeinschaft
DMP	Datenmanagementplan
FAIR	Findable, Accessible, Interoperable, Reusable
FDM	Forschungsdatenmanagement
UB	Universitätsbibliothek
SZF	Servicezentrum Forschungsdatenmanagement
ZEEM	Zentraleinrichtung Campusmanagement

Zentraler Kontaktpunkt an der TU Berlin

Als zentraler Kontaktpunkt für alle Belange des FDM an der TU Berlin dient das **Servicezentrum Forschungsdatenmanagement** (Helpdesk: team@szf.tu-berlin.de).

Im Servicezentrum Forschungsdatenmanagement (SZF) arbeiten die Universitätsbibliothek (UB), die Zentraleinrichtung Campusmanagement (ZECM) und die Abt. V – Forschung und Technologietransfer (Abt. V) zusammen und bündeln ihre Kompetenzen, um die Forschenden der TU Berlin beim Umgang mit ihren Forschungsdaten zu unterstützen. Leitung und Koordination des SZF liegen bei der UB. Das SZF betreibt die Forschungsdaten-Infrastruktur der TU Berlin, die in die IT-Infrastruktur der TU Berlin integriert ist. Zentrale Infrastrukturdienste sind *DepositOnce*, das Repositorium für Forschungsdaten und Publikationen der TU Berlin, und *TUB-DMP*, ein Web-Tool zur Erstellung von Datenmanagementplänen (DMPs). Ein umfangreiches Schulungs- und Beratungsangebot ergänzt die technischen Dienste.

FDM und Forschungsdaten-Infrastruktur der TU Berlin sind nach den [FAIR-Prinzipien](#) ausgerichtet, internationalen Grundsätzen für die optimale Aufbereitung von Forschungsdaten, damit diese auffindbar (Findable), zugänglich (Accessible), austauschbar (Interoperable) und nachnutzbar (Reusable) sind.

Die [Website des SZF](#) dient als zentrale Plattform für das FDM in der TU Berlin. Hier finden Sie umfangreiche Informationen zu den FDM-Services der TU Berlin, einschließlich Ansprechpersonen.

Entsprechend der Aufgabenteilung im SZF gibt es **verteilte Zuständigkeiten** entlang des Forschungsdaten-Lebenszyklus (siehe Abbildung).

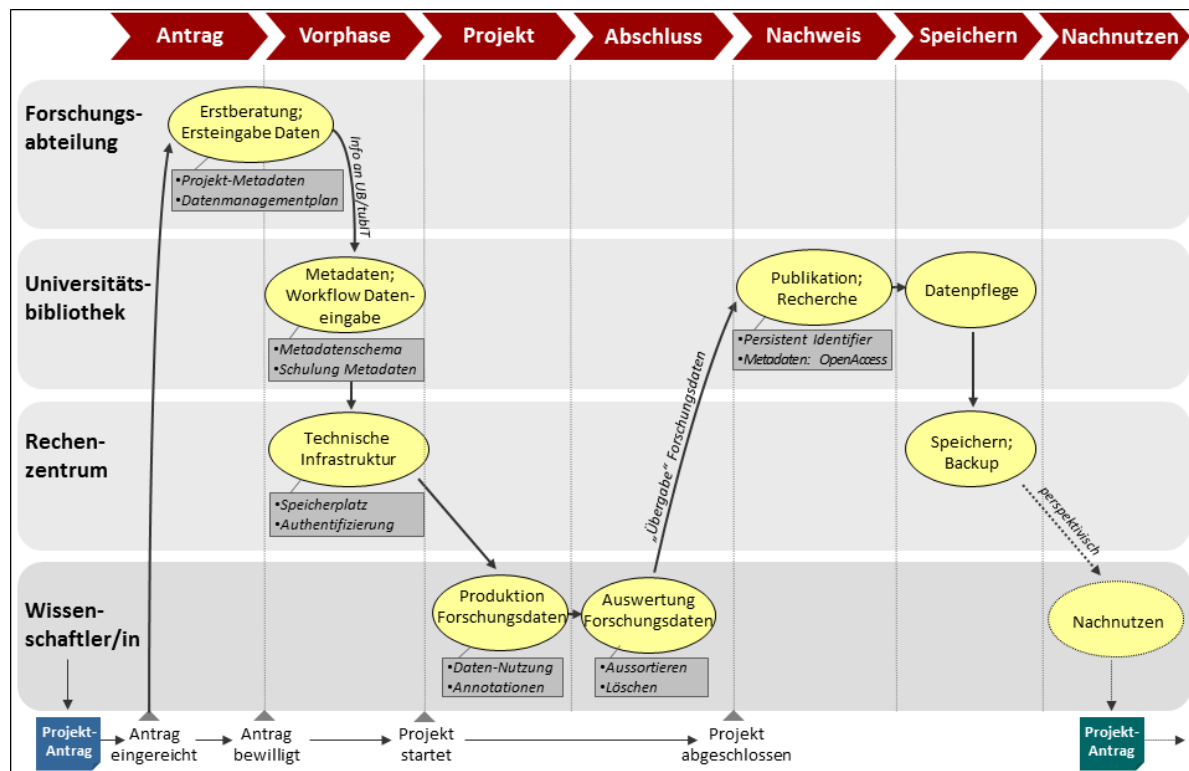


Abb. 1: Verteilte Zuständigkeiten im FDM der TU Berlin

Empfehlungen zum Umgang mit Forschungsdaten

Die frühzeitige Auseinandersetzung mit dem FDM hat nicht nur Vorteile in Bezug auf eine reibungslose Zusammenarbeit im späteren Projektverlauf. Sie ist auch wichtig in Bezug auf mögliche Kosten, die durch FDM entstehen. Personal für den Aufbau und die Betreuung von Forschungsinfrastruktur und die Infrastruktur selbst (z. B. [INF-Projekt in SFBs](#)) können rechtzeitig eingeplant und beantragt werden.

Vorab: Zehn Fragen zum FDM in der Antragsphase

1. Gibt es Vorgaben zum Umgang mit Forschungsdaten (z. B. vom Mittelgeber, der Forschungseinrichtung, dem Kooperations- oder Vertragspartner oder der Fachcommunity)?
2. Sind die Verantwortlichkeiten für das Forschungsdatenmanagement im Projekt geregelt?
3. Gibt es einen Datenmanagementplan für das Projekt?
4. Wird die Speicherinfrastruktur der TU Berlin genutzt? Wurde dazu schon Kontakt zur ZECM aufgenommen?
5. Werden fremde Forschungsdaten genutzt? Müssen hierfür ggf. Mittel beantragt werden?
6. Sollen externe Repositorien für die Speicherung und ggf. Veröffentlichung der Forschungsdaten verwendet werden? Wenn ja, welche?
7. Gibt es eine Publikationsstrategie für Forschungsdaten und Publikationen (z. B. Open Access)?
8. Entstehen Kosten für die Archivierung und/oder Publikation der entstehenden Forschungsdaten?
9. Wird Personal für das Management der Forschungsdaten und/oder die Betreuung der technischen Infrastruktur beantragt?
10. Wird sichergestellt, dass die Projektmitarbeiter*innen im Umgang mit Forschungsdaten geschult werden?

I. Planungs- und Antragsphase (vor dem Forschungsvorhaben)

- Wenn Sie Ihr **Forschungsvorhaben im Rahmen eines Drittmittelprojekts** durchführen, erkundigen Sie sich vorab nach eventuell geltenden Richtlinien der Förderorganisationen. Kosten für das FDM – z. B. für Personalmittel zur Datenaufbereitung oder zur Entwicklung projektinterner Workflows, für Publikationskosten oder auch für die langfristige Speicherung, die über die Grundausstattung der Universität hinausgehen – können und sollten Teil der beantragten Fördermittel sein. Nutzen Sie das Beratungsangebot der Abt. V, um sich über die Anforderungen der Förderorganisationen und Finanzierungsmöglichkeiten zu informieren und Ihren individuellen Bedarf zu ermitteln.
 - Beantragung von DFG- und BMBF-Projekten: [Servicebereich Forschung](#)
 - Beantragung von Graduiertenkollegs: [Center for Junior Scholars](#)
 - Beantragung von EU-Projekten: [EU-Büro](#)
 - Verhandlungen zu Forschungsverträgen: [Zentrum für geistiges Eigentum](#)
- Bei allen Forschungsvorhaben und auch bei der Veröffentlichung der Ergebnisse sind **rechtliche Rahmenbedingungen** zu beachten und möglichst schon im Vorfeld abzuklären. Bestimmte Forschungsdaten etwa in den Sozial- oder Lebenswissenschaften unterliegen strengen Auflagen zu **Datenschutz** und **Ethik**. Auch der Schutz des Urheberrechts und die berechtigten Interessen Dritter müssen gewährleistet sein.
 - Beratung zu Urheberrecht/Patentrecht: [Zentrum für geistiges Eigentum](#)
 - Beratung zu Datenschutz/Ethik allgemein: [Team Datenschutz der TU Berlin](#)
 - Beratung zu Ethikfragen in den Fakultäten:
 - [Ethik-Kommission der Fakultät III](#)
 - [Ethik-Kommission der Fakultät IV](#)
 - [Ethik-Kommission des Instituts für Psychologie und Arbeitswissenschaften](#) (Fakultät V)
 - [Ethik-Kommission der Fakultät VI](#)
 - [Ethik-Kommission der Fakultät VII](#)
- Um eine angemessene Unterstützung für Ihr neues Forschungsvorhaben zu erhalten, müssen Sie Ihr Projekt bei der Antragstellung in der Abt. V anmelden. Dies erfolgt über die [Elektronische Projektanzeige \(ePA\)](#).
 - Informationen zur Elektronischen Projektanzeige: [Servicebereich Forschung](#)
- **Forschungsdatenmanagement**

In jedem Forschungsvorhaben, in dem Daten generiert werden oder Daten die Grundlage Ihrer Forschung bilden, ist eine frühzeitige Auseinandersetzung mit den Anforderungen und Möglichkeiten des FDM dringend angeraten. Schon in der Antragsphase sollte eine **Strategie für die nachhaltige Speicherung und Verfügbarmachung** der Forschungsdaten bestimmt werden. Auch der rechtliche Status der entstehenden Daten sowie geeignete Schutzmaßnahmen für die Verwendung der Daten während und nach Abschluss des Forschungsvorhabens sind frühzeitig zu klären und festzulegen.

Diese Festlegungen sollten in einen **DMP** einfließen. Er dokumentiert den Entstehungsprozess von Forschungsdaten und wie sie regelgerecht gespeichert werden, so dass sie in späteren Jahren interpretierbar und nachvollziehbar, verfügbar, authentisch, zitierfähig und nachnutzbar sind. Zur Optimierung des FDM und als Basis für einen institutionellen Support sollte ein DMP möglichst vor Beginn eines Forschungsvorhabens erstellt und im Verlauf des Vorhabens aktualisiert werden (*living document*).

Fast alle Förderorganisationen erwarten die Vorlage eines DMP bei der Projektantragstellung. Zur Unterstützung bei der Erstellung Ihres DMP steht Ihnen das Web-Tool [TUB-DMP](#) zur Verfügung. Es enthält Vorlagen (Templates) mit relevanten Fragen, die Sie – sofern für Ihr Projekt relevant – in einem Schritt-für-Schritt-Workflow beantworten können.

 - Beratung zum FDM und zu TUB-DMP: [SZF-Team](#)

II. Durchführungsphase (während des Forschungsvorhabens)

- Für die **Speicherung und Verarbeitung** von Forschungsdaten sowie für die Kollaboration auf Basis dieser Daten sollten Verfahren verwendet werden, die dem aktuellen Stand der Technik entsprechen. Dazu gehört insbesondere die Beachtung von Datensicherheit hinsichtlich Verfügbarkeit, Integrität (Unverfälschtheit) und Authentizität. Dies impliziert z. B. die Nutzung der Datensicherung und die Verwendung sicherer Datenaustauschplattformen.
- In der Durchführungsphase eines Projekts unterliegen Datensätze meist einer Entwicklung über mehrere Stufen hinweg (z. B. durch Auswahl, Aggregation, Integration). Es hat sich als vorteilhaft erwiesen, die verschiedenen **Versionen** zu kennzeichnen, zu dokumentieren und zumindest während der aktiven Projektlaufzeit aufzubewahren. Insbesondere bei textbasierten Daten erleichtert die Verwendung von Versionierungstools, wie sie in der Softwareentwicklung üblich sind (z. B. GitLab, SVN), das Management der verschiedenen Versionen.
- Die ZECM stellt im Forschungsumfeld u. a. folgende Dienste zur Verfügung; nähere Beschreibungen finden Sie auf den [Webseiten der ZECM](#).
 - Nutzung von Netzwerkfilesystemen (inkl. Datensicherung)
 - Archivspeicherdienste auf Bandlaufwerken
 - Bereitstellung virtueller Root-Server (Serverhosting)
 - Unterbringung realer Server (Serverhousing)
 - Blockspeicherdienste für Server (virtuelle Festplatten über ein dezidiertes Speichernetz)
 - Datenaustauschdienste
 - Versionierungsdienste

Die Dienste werden entweder als Grundausstattung kostenfrei oder zum Selbstkostenpreis angeboten. Das Gleiche gilt für die [Werkzeuge zum kollaborativen Arbeiten](#), die die ZECM bereitstellt.

- Die verschiedenen Wissenschaftsdisziplinen und ihre Teilgebiete haben unterschiedliche Arbeitsweisen im Umgang mit Forschungsdaten, was eine fachübergreifende Empfehlung zur Nutzung konkreter Verfahren erschwert. Daher wird grundsätzlich empfohlen, sich im Vorhinein über die in der jeweiligen Fachcommunity etablierten **Datenformate, Software und Standards** zur Dokumentation und Annotation von Forschungsdaten wie z. B. Ontologien, kontrollierte Vokabulare oder Metadatenschemata zu informieren. Die Nutzung offener, nicht-proprietärer Datenformate unterstützt die Zugänglichkeit und Langzeitverfügbarkeit von Forschungsdaten.
- Von elementarer Bedeutung für die Nachnutzbarkeit von Forschungsdaten ist deren Beschreibung mit **Metadaten**. Metadaten sind Daten über Daten und geben den Kontext wieder, in dem diese erzeugt wurden. Als Faustregel gilt: Metadaten sollten die klassischen sechs W-Fragen beantworten: Wer? Was? Warum? Wie? Wann? Wo? Metadaten sind die Voraussetzung, um potenziellen Nachnutzenden das Auffinden und die Abschätzung der Eignung für die Nutzung zu ermöglichen. Idealerweise erfolgt die Beschreibung strukturiert und maschinenlesbar. Hierzu existieren in den meisten Fachdisziplinen Metadatenstandards und standardisierte Terminologien. Falls diese nicht existieren, sollten allgemeingültige Standards, wie z. B. [Dublin Core](#), zur Beschreibung verwendet werden. Sie werden von weltweiten Initiativen vorangetrieben und helfen, die Forschungsergebnisse besser nachvollziehbar und interoperabel zu machen.
- In Verbundprojekten oder bei großen Datenmengen ist der Einsatz dezidierter **Arbeitsumgebungen** und Portale zum Datenmanagement sinnvoll. Der Betrieb dieser Tools erfordert meist zusätzliche Ressourcen, bietet jedoch den Vorteil einer einheitlichen und zentralen Verwaltung der Forschungsdaten. Das Auffinden und der Austausch von Daten wird damit erleichtert, sollte jedoch im Projektkonsortium mit Hilfe einer projektspezifischen **Data Policy** geregelt werden.

→ Beratung zur technischen Infrastruktur und IT-Diensten: [ZECM](#)

→ Beratung zu Metadaten und -Standards: [SZF-Team](#)

III. Abschlussphase (nach dem Forschungsvorhaben)

- Im Sinne der guten wissenschaftlichen Praxis sind Forschungsdaten spätestens zum Projektende langfristig zu sichern und nach Möglichkeit zugänglich zu machen, sofern keine vertraglichen, ethischen oder gesetzlichen Bestimmungen dem entgegenstehen. Von vielen Förderorganisationen wird inzwischen besonderer Wert auf die **Zugänglichkeit** gelegt, um die Nachvollziehbarkeit von Forschungsergebnissen und die Nachnutzung der Daten zu ermöglichen. Gemäß ihrer Forschungsdaten-Policy unterstützt die TU Berlin den freien Zugang zu Forschungsdaten. Bei der Veröffentlichung von Forschungsdaten empfiehlt die TU Berlin, sich an dem Prinzip „Zugänglich, wenn möglich, eingeschränkt, wenn notwendig“ zu orientieren.
- Folgende Grundprinzipien hinsichtlich der **Veröffentlichung von Forschungsdaten** sollen beachtet werden:
 - Allgemeine Projekt-, Arbeitsgruppen-, Lehrstuhl- oder persönliche Webseiten sind kein geeigneter Ort für die Veröffentlichung von Forschungsdaten, da die langfristige Verfügbarkeit nicht gewährleistet und eine eindeutige Identifizierung (Stichwort: persistenter Identifikator) nur begrenzt möglich sind.
 - Bei der Auswahl der zu veröffentlichenden Daten empfiehlt die [DFG](#): „Forschungsdaten sollten in einer Verarbeitungsstufe (Rohdaten oder bereits weiter strukturierte Daten) zugänglich sein, die eine sinnvolle Nach- und Weiternutzung durch Dritte ermöglicht.“ Insbesondere Daten, die Grundlage wissenschaftlicher Artikel bilden, sollten – soweit keine (datenschutz)rechtlichen oder forschungsethischen Bedenken entgegenstehen – zugänglich gemacht werden.
 - Analog zu wissenschaftlichen Artikeln sollte auch Forschungsdaten bei der Veröffentlichung ein eindeutiger, persistenter Identifikator (PID) zugewiesen werden. Dadurch sind Forschungsdaten *für sich*, d. h. unabhängig von einer Publikation, auffindbar und eigenständig zitierbar. Bekannte Beispiele sind DOI (Digital Object Identifier) oder URN (Uniform Resource Name).
 - Um die Nutzungs- und Verwertungsrechte von Forschungsdaten zu regeln, sollten Daten immer mit einer entsprechenden Lizenz veröffentlicht werden. Die Wahl der Lizenz sollte zumindest einen offenen Zugang für wissenschaftliche Zwecke erlauben. Zu berücksichtigen sind ggf. spezielle Anforderungen der Förderorganisationen oder Repositorien. Bekannte freie Lizenzen aus dem Softwareumfeld sind die [GNU General Public License \(GPL\)](#), [MIT-Lizenz](#) oder [Apache-Lizenz](#). Für Texte, Bilder, Audio- und Videodateien haben sich [Creative-Commons-Lizenzen](#) durchgesetzt.
- Für die Veröffentlichung Ihrer Forschungsergebnisse (Forschungsdaten und Publikationen) steht Ihnen als Mitglied der TU Berlin und Ihren Kooperationspartnern das [Repositorium DepositOnce](#) der TU Berlin zur Verfügung. In DepositOnce werden Forschungsergebnisse gespeichert, d. h. konsolidierte Daten und alle Informationen (wie Skripte, Berechnungen usw.), die zur Reproduktion des Ergebnisses notwendig sind. Gemäß den Grundsätzen zur Sicherung guter wissenschaftlicher Praxis der TU Berlin werden die Forschungsdaten für die Dauer von mindestens 10 Jahren gespeichert.
 - Alle Daten in DepositOnce sind mit Metadaten (Standardformat Extended Dublin Core) versehen.
 - Alle Datensätze erhalten automatisch eine persistente Internetadresse (DOI).
 - Den Datensätzen können verschiedene freie Lizenzen zugewiesen werden.
 - Über den DOI können zusammengehörige Forschungsdaten und Publikationen miteinander verlinkt werden und verweisen dann aufeinander.
- Gemäß den Regeln guter wissenschaftlicher Praxis können veröffentlichte Forschungsdaten in DepositOnce nicht mehr geändert werden; so bleiben Zitierfähigkeit und Nachvollziehbarkeit erhalten. DepositOnce verfügt über eine Versionierung, bei der neue Versionen publiziert werden und gleichzeitig die vorhergehenden verfügbar bleiben. Jede neue Version erhält einen neuen DOI; vorhergehende und aktuelle Versionen werden automatisch miteinander verlinkt und verweisen aufeinander.

DepositOnce ist dem Open Access verpflichtet. Die Metadaten sind im Internet frei zugänglich und werden über Standardschnittstellen weit verbreitet und suchbar gemacht (Google Scholar u.a.). Die Forschungsdaten selbst können mit einer Sperrfrist (Embargo) versehen werden.

- Für viele Fachdisziplinen und Datentypen wurden und werden im Rahmen verschiedener Initiativen und Projekte geeignete fachspezifische Forschungsdaten-Infrastrukturen aufgebaut, an denen sich auch

Forschende der TU Berlin beteiligen. Inzwischen existiert daher eine große Anzahl **fachspezifischer Repositorien**. Diese können im Vergleich zu DepositOnce durchaus Vorteile bieten, wie z.B. fachspezifische Metadatenschemata und spezifische Suchmöglichkeiten. In den Fällen, in denen Sie als Mitglied der TU Berlin bereits ein Repositoryum ihrer Fachcommunity nutzen, sollten Sie dies weiterhin tun – ebenso, wenn es sinnvoll erscheint, die Forschungsdaten in einem fachspezifischen Repositoryum zu veröffentlichen und ein solches für Ihre Disziplin existiert. In einigen Disziplinen ist auch die Veröffentlichung von Daten als Supplement zum wissenschaftlichen Artikel etabliert. Diese Form der Datenpublikation hat jedoch den Nachteil, dass die Daten nur über den Artikel gefunden werden können und keine eigenständigen, zitierfähigen Publikationsobjekte bilden.

Bei der Wahl eines fachspezifischen Repositoryums sollten folgende Kriterien beachtet werden: Langzeitverfügbarkeit (mindestens 10 Jahre), Vergabe von persistenten Identifikatoren (z.B. DOI, URN), Lizenzen und Nutzungsrechte der Daten, Bekanntheit und Sichtbarkeit, Kosten. Auf der Suche nach einem geeigneten fachspezifischen Repositoryum für Ihre Forschungsdaten bietet das Portal re3data.org einen guten Überblick mit umfangreichen Such- und Filterfunktionen.

- Die **Veröffentlichung** von Forschungsdaten sollte so **zeitnah wie möglich** erfolgen. Liegen triftige Gründe vor, können Daten in DepositOnce mit einer Sperrfrist (Embargo) versehen werden. In diesem Fall werden nur die Metadaten veröffentlicht; die Daten selbst sind im Repositoryum gesichert, jedoch erst nach Ende der Sperrfrist zugänglich. Während der Sperrfrist können die Daten von Interessierten per E-Mail angefragt werden. Die Festlegung der Sperrfrist erfolgt durch die verantwortlichen Forschenden, wobei es gilt, die Vorgaben und Richtlinien der Forschungsförderer und Repositoryen zu beachten. Embargoperioden sollten eine Dauer von maximal 5 Jahren nach Projektende nicht überschreiten. Ein Embargo ist zu begründen, z.B. in einer Datei im Repositoryum, die auch das Ende des Embargos benennt.

→ Beratung zu Veröffentlichung von Forschungsdaten, Lizenzen, DepositOnce/Repositoryen: [SZF-Team](#)