

**Grundsätze zum Umgang mit digitalen
Forschungsdaten an der
Friedrich-Alexander-Universität Erlangen-Nürnberg**

Forschungsdaten-Policy

Von der Universitätsleitung beschlossen am 4. Dezember 2019.



Grundsätze zum Umgang mit digitalen Forschungsdaten an der Friedrich-Alexander-Universität Erlangen-Nürnberg

Forschungsdaten-Policy

[Von der Universitätsleitung beschlossen am 4. Dezember 2019.]

Präambel

Diese Forschungsdaten-Policy stützt sich auf die „Grundsätze zum Umgang mit Forschungsdaten“ (2010) der Allianz der Wissenschaftsorganisationen sowie die Empfehlungen zur „Sicherung guter wissenschaftlicher Praxis“ (2013), die „Leitlinien zum Umgang mit Forschungsdaten“ (2015), den Kodex „Leitlinien zur Sicherung guter wissenschaftlicher Praxis“ (2019) der Deutschen Forschungsgemeinschaft und die FAIR-Prinzipien¹. Daneben sind auch die weiteren Grundsätze der Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU) zu beachten, wie etwa die Open Access Policy und die Intellectual-Property-Policy sowie die „Satzung zur Sicherung guter wissenschaftlicher Praxis und zum Umgang mit wissenschaftlichem Fehlverhalten an der Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU)“.

Die FAU betrachtet die Sicherung und das Management von Forschungsdaten (FD) als essenziell für erfolgreiche und nachhaltige Forschung und wissenschaftliche Integrität. Der geplante und verantwortungsvolle Umgang mit Forschungsdaten ist unerlässlich und muss deshalb gefördert und gefordert werden, damit die Universität, ihre Angehörigen und die Öffentlichkeit davon profitieren können. Hierfür ist es entscheidend, das Bewusstsein für Forschungsdaten und für die FAIR-Prinzipien langfristig zu stärken.

Das Forschungsdatenmanagement (FDM) ist Teil der Informations- und Kommunikationstechnologie (IuK)-Strategie der Universität. Die FAU wird diesen fortlaufenden Prozess kontinuierlich durch geeignete Maßnahmen des Chief Information Officers (CIO) sowie der Serviceeinrichtungen (insbesondere Regionales Rechenzentrum Erlangen, FAU CRIS, Universitätsbibliothek) und der forschungsunterstützenden Stellen der Universitätsverwaltung begleiten.

Mit diesen *Grundsätzen zum Umgang mit digitalen Forschungsdaten* beabsichtigt die FAU einerseits, ihren Mitgliedern Leitlinien und Empfehlungen für den Umgang mit Forschungsdaten an die Hand zu geben. Andererseits soll die strategische Ausrichtung der Universität in dem sich schnell wandelnden Feld des wissenschaftlichen Datenmanagements weiter geschärft werden.

¹ Wilkinson et al., „The FAIR Guiding Principles for scientific data management and stewardship“, Scientific Data Vol. 3, 160018 (2016)

Verantwortlichkeit

Die vorliegenden Grundsätze richten sich an alle forschenden und lehrenden FAU-Angehörigen sowie die vorübergehend oder gastweise tätigen Wissenschaftlerinnen und Wissenschaftler. Die Leitenden von Forschungsvorhaben, aber auch selbstständig Forschende, tragen während der gesamten Dauer (von der Planung bis zur terminierten Löschung) die Verantwortung für das angemessene Management aller entstehenden digitalen Forschungsdaten (dFD). Ferner gestalten sie das FDM und die Qualitätssicherung in ihren Arbeitsgruppen und treffen Regelungen beim Wechsel der Mitglieder. Diese Regelungen umfassen einen Verbleib der Originaldaten am Entstehungsort, Vorkehrungen zur sachgerechten Weitergabe von Primärdaten und zur Klärung der Zugangsrechte sowie die Möglichkeit, bei Wechsel des Arbeitsplatzes ein Duplikat der Forschungsdaten zu erstellen.

Die Angehörigen der FAU haben bei allen Schritten des FDM die geltenden Rechtsvorschriften und forschungsethische Belange zu beachten. Dies betrifft insbesondere den Umgang mit personenbezogenen Daten und schutzfähigem geistigen Eigentum. Beim Umgang mit personenbezogenen Daten ist der Datenschutzbeauftragte immer zu beteiligen, sofern noch keine angezeigte Verarbeitungstätigkeit vorliegt. Im Zweifel sind die beratenden Stellen und Kontrollgremien der FAU (unter anderem Datenschutzbeauftragter, Rechtsabteilung, Universitätsbibliothek, Kontaktstelle für Wissens- und Technologietransfer oder die Ethikkommission) hinzuzuziehen.

Datenmanagementplan

Die FAU empfiehlt die Voraberstellung eines Datenmanagementplans, um einen systematischen Umgang mit den FD während des und nach dem Vorhaben zu ermöglichen, sowie gegebenenfalls die Beantragung der dafür notwendigen Fördermittel zu erleichtern. Hierbei werden Wissenschaftlerinnen und Wissenschaftler durch die Beratungsstellen der FAU unterstützt. Die Universität wird Beratung und Unterstützungswerkzeuge laufend optimieren, um sicherzustellen, dass aktuelle Standards im FDM bereits in der Planungs- und Antragsphase entsprechend berücksichtigt werden.

Dokumentation

Der Entstehungskontext der FD, Kontextinformationen zu Werkzeugen sowie die verwendete Software, die Analyseprotokolle und der Forschungsprozess an sich sind zu dokumentieren. Ebenso wird die Herkunft verwendeter oder nachgenutzter Materialien, Daten oder Software kenntlich gemacht. Es sollen dabei alle vertretbaren Anstrengungen unternommen werden, um Korrektheit, Unverfälschtheit und Vollständigkeit der Aufzeichnungen und der zugrundeliegenden Daten langfristig sicherzustellen. Neben der zuverlässigen Speicherung der Daten selbst ist die detaillierte Dokumentation eine zentrale Grundvoraussetzung für die Verifizierbarkeit geleisteter Forschungsarbeit und für die Nachnutzbarkeit der Daten im Rahmen neuer Forschungsfragen.

Speicherung

- a. Die Speicherung von dFD erfolgt nach Möglichkeit während des gesamten Forschungsprozesses in redundant gesicherten Storage-Systemen. Das Regionale Rechenzentrum Erlangen (RRZE) unterstützt und berät bei Datensicherung und Backup.
- b. Die langfristige Ablage von dFD erfolgt nach Möglichkeit in etablierten, fachspezifischen Repositorien oder Datenzentren. Es wird empfohlen, sich vorab über in Frage kommende Repositorien und gegebenenfalls deren spezielle Anforderungen für Archivierung oder Publikation der dFD zu informieren. Das RRZE bietet darüber hinaus die Möglichkeit, dFD auch vor Ort langfristig zu archivieren.
- c. Für das Format der dFD sind bevorzugt freie Standardformate oder in der Fachcommunity etablierte Formate zu wählen, um die Interoperabilität zu erleichtern und eine langfristige Lesbarkeit zu gewährleisten.
- d. Die für eine Publikation verwendeten dFD sind so lange aufzubewahren, wie es nach einschlägigen gesetzlichen oder vertraglichen Vorschriften, insbesondere nach dem Patentrecht oder nach einer Vorgabe des Drittmittelgebers erforderlich ist, mindestens jedoch für zehn Jahre. Im Fall der Weiterverarbeitung von dFD sollen auch die Primärdaten aufbewahrt werden, um bei Bedarf verfügbar zu sein. Sofern nachvollziehbare Gründe dafür existieren, bestimmte Forschungsdaten nicht aufzubewahren, dokumentieren die Wissenschaftlerinnen und Wissenschaftler dies.
- e. Die Qualität der Metadaten beeinflusst maßgeblich die Auffindbarkeit und Nachnutzbarkeit der Daten. Basale Informationen sind beispielsweise beschreibende Titel und Abstracts, beteiligte Forschende, Institution, Identifier, Ort und Zeitraum, Rechte, Formate etc. Die Nutzung standardisierter Metadatenschemata stellt eine möglichst einheitliche und nachvollziehbare Beschreibung sicher und unterstützt die Maschinenlesbarkeit und Interoperabilität.
- f. Die Auswahl, Aufbereitung und Dokumentation der Daten sowie die Speicherung über einen definierten Zeitraum benötigen ausreichende Ressourcen. Dies sollte bei bereits in der Planungsphase berücksichtigt werden; die FAU bietet eine entsprechende Beratung an.

Zugriffs- und Nutzungsbedingungen

- a. Es wird grundsätzlich empfohlen, dFD der (Fach-)Öffentlichkeit zur Nachnutzung zugänglich zu machen.
- b. Um die langfristige Zitierbarkeit der dFD sicherzustellen wird die konsequente Verwendung von persistenten Identifikatoren nachdrücklich empfohlen.
- c. In Abhängigkeit von rechtlichen und vertraglichen Anforderungen sind für die Veröffentlichung der dFD der geeignete Zeitpunkt, Umfang und passende Lizenzbedingungen zu wählen. Die Veröffentlichung soll unter dem Namen der Verantwortlichen erfolgen.
- d. Bei einer Übertragung von Nachnutzungs- oder Veröffentlichungsrechten soll darauf geachtet werden, dass die dFD für wissenschaftliche Zwecke frei verfügbar bleiben.
- e. Es soll Forschenden ermöglicht werden, ihre nach dieser Policy sowie aus Verträgen mit Drittmittelgebern und aus sonstigen Rechtsquellen bestehenden Rechte und Pflichten wahrzunehmen. Die Universität stellt für diejenigen Forschenden, in deren Fachdisziplinen sich noch keine Repositorien etabliert haben, einen Zugang zu Diensten und Einrichtungen für Lagerung, Sicherung und Aufbewahrung von dFD bereit.

- f. Die Universität erweitert bestehende Strukturen, wie etwa das Forschungsinformationssystem CRIS FAU, um alle, auch die dezentral in externen Repositorien gespeicherten dFD, hochschulweit dokumentieren und würdigen zu können.

Vermittlung von Kompetenzen

Die Vermittlung der Prinzipien guten Datenmanagements und von Grundkompetenzen im Umgang mit Daten ist für den wissenschaftlichen Nachwuchs unerlässlich. Forschungsdatenmanagement und Data Literacy sind Schlüsselqualifikationen der modernen Forschung, deren Grundlagen allen Graduierten vermittelt werden sollen.

Die Universität baut ihre bestehenden, allgemeinen Weiterbildungsangebote durch die zentralen Einrichtungen gezielt weiter aus. Eine curriculare Verankerung von fachspezifischen Methoden des FDM durch die Departments und Fakultäten in der Lehre wird ausdrücklich empfohlen und begrüßt.

Fachspezifische Regelungen

Die FAU erkennt an, dass in einzelnen Fachdisziplinen und für bestimmte Forschungsmethodologien detailliertere Richtlinien und Vorgaben notwendig sind, die über diese Policy hinausgehen. Die Forschenden sind daher angehalten, beim Datenmanagement bestehende fachspezifische Regularien und in der Community etablierte Vorgehensweisen zu berücksichtigen.²

Weiter unterstützt die FAU die Fakultäten beim Aufbau disziplinspezifischer Dienstleitungen für das FDM. Es wird insbesondere angestrebt, ein Netzwerk von an den Departments angesiedelten Ansprechpartnern aufzubauen, die bei Anliegen des Datenmanagements fachnahe Unterstützung bieten können.

Evaluation

Dieser Grundsätze werden regelmäßig evaluiert und spätestens nach 3 Jahren an aktuelle Standards angepasst.

Die Grundsätze zum Umgang mit digitalen Forschungsdaten an der Friedrich-Alexander-Universität Erlangen-Nürnberg wurden erstmals am 16.11.2016 von der Universitätsleitung verabschiedet. Geändert am 04.12.2019 durch Beschluss der Universitätsleitung.

Glossar

Ein **Datenmanagementplan** (DMP) dokumentiert den gesamten für die Daten vorgesehenen Lebenszyklus. Das heißt er beschreibt, wie mit den Forschungsdaten von der Planung der Erhebung bis zur langfristigen Archivierung oder gegebenenfalls planmäßigen Löschung umgegangen wird. Der Datenmanagementplan beantwortet dabei mindestens die Fragen:

² Eine Sammlung von Empfehlungen und Richtlinien von DFG und verschiedenen Fachgesellschaften zum Umgang mit Forschungsdaten findet sich unter https://www.dfg.de/foerderung/antrag_gutachter_gremien/antragstellende/nachnutzung_forschungsdaten/index.html (Zuletzt geprüft am 04.11.2019).

Was wird erhoben? Welche Stellen müssen vor der Erhebung hinzugezogen werden? In welcher Form und wo werden die Forschungsdaten in den unterschiedlichen Projektphasen gesichert? Wer hat ab welchem Zeitpunkt darauf Zugriff? Wer ist für die einzelnen Schritte zuständig / verantwortlich? Welche rechtlichen Vorgaben sind zu beachten?

Die **FAIR-Prinzipien** sind Anforderungen, die nachhaltige und wiederverwendbare Forschungsdaten erfüllen sollten. Das Akronym FAIR steht für **F**indable, **A**ccessable, **I**nteroperabel und **R**e-Usable. Die FAIR-Prinzipien werden von vielen Forschungsförderern (etwa EU, DFG oder SNF) als wichtige Voraussetzung für nachhaltige Forschung gesehen und auch entsprechend eingefordert. Als besonders wichtig für die Auffindbarkeit von Daten wird die Verwendung von persistenten Identifikatoren und detailreiche Metadaten gesehen. Der Einsatz von Standards – sowohl bei Schnittstellen als auch für Metadaten und Daten selbst – unterstützt neben der Zugänglichkeit auch die Interoperabilität der Daten. Die Wiederverwendbarkeit wird durch umfangreiche inhaltliche Metadaten und Dokumentationen sowie klare Nachnutzungsrechte erleichtert. FAIRe Daten müssen nicht notwendigerweise offen („open data“) sein. Lediglich die Metadaten sollten frei zugänglich sein.

Forschungsdaten sind alle Daten, die während eines Forschungsvorhabens mit einer wissenschaftlichen Arbeitsmethode entstehen – sowohl neu erzeugte als auch aufbereitete Daten. Dies ist unabhängig davon, ob diese Daten in eine Publikation einfließen oder nicht und ob sie in analoger oder digitaler Form vorliegen.

Forschungsdatenmanagement umfasst alle mit digitalen Forschungsdaten durchgeführten Aktivitäten (von der eigentlichen Nutzung abgesehen), insbesondere bezüglich der Dokumentation, der Sicherung und der langfristigen Verfügbarkeit.

Metadaten bezeichnen alle zusätzlichen Informationen, die zur Interpretation und (automatischen) Verarbeitung der eigentlichen Daten, etwa der digitalen Forschungsdaten, notwendig oder sinnvoll sind. Metadaten werden daher oft als ‚Daten über Daten‘ bezeichnet. Sie dienen dazu, die Eigenschaften digitaler Objekte zu kategorisieren und zu charakterisieren sowie weiterführende Informationen zur Verfügung zu stellen. Man unterscheidet verschiedene Kategorien von Metadaten: *Technische Metadaten* beinhalten beispielsweise Angaben zu Datenvolumen und Datenformat und sind für eine nachhaltige Datenspeicherung von zentraler Bedeutung. *Deskriptive Metadaten* (auch beschreibende oder Content-Metadaten genannt) geben Auskunft über die in digitalen Objekten enthaltenen Informationen und entscheiden damit über deren Auffindbarkeit, Referenzierung und Nachnutzbarkeit. Hierzu gehören auch Erläuterungen der zugrundeliegenden Messmethode, ein Abstract oder Schlagwörter. *Strukturelle Metadaten* beschreiben Beziehungen zwischen einzelnen Elementen eines Datensatzes oder die interne Struktur der Daten selbst. *Administrative Metadaten* beinhalten Informationen für die Qualitätssicherung der Daten (etwa Prüfsummen), Informationen zu Zugriffsrechten und Lizenzen oder Provenienz-Informationen.³

³ Adaptiert von „Metadaten“, [forschungsdaten.org](http://www.forschungsdaten.org), zuletzt geprüft am 09.09.2019, <http://www.forschungsdaten.org/index.php/Metadaten>.

Persistente Identifikatoren (PI) erlauben eine dauerhafte Adressierung von digitalen Ressourcen. Ein PI stellt einen eindeutigen *Namen* für Publikationen jeglicher Art (Ergebnisse, Daten, Software) dar. Dieser Name wird mit der Web-URL der digitalen Ressource verknüpft. Wenn sich die URL zu der Ressource ändert, muss nur die Adresse geändert werden, auf die der PI verweist, während der PI selbst gleich bleiben kann. Dies garantiert beispielsweise, dass eine einmal über einen PI zitierte Ressource auch dann noch gefunden werden kann, wenn sich ihr physikalischer Speicherort verändert hat. Beispiele für persistente Identifikatoren sind Digital Object Identifier (DOI), Uniform Resource Name (URN) und Handle.

Als **Repositorien** werden Infrastrukturen bezeichnet, in denen digitale Ressourcen dauerhaft gespeichert werden können und die die Auffindbarkeit und Sichtbarkeit der digitalen Objekte sicherstellen. Für Forschungsdaten existieren spezielle Datenrepositorien, die entweder auf die Bedürfnisse einzelner Fachdisziplinen zugeschnitten sind oder generische Dienstleistungen zur langfristigen Speicherung und Zugänglichmachung von Daten anbieten.