

Forschungsdaten-Richtlinie des Helmholtz Zentrums München

1. Präambel

Das Helmholtz Zentrum München, Deutsches Forschungszentrum für Gesundheit und Umwelt GmbH (im Folgenden Helmholtz Munich) erkennt die grundlegende Bedeutung von Forschungsdaten an und fordert alle Mitarbeitenden auf, die höchsten Standards bei der Organisation von Forschungsdaten anzuwenden (zur Definition von „Forschungsdaten“ siehe 8.1.3). Dies beinhaltet, dass korrekte und leicht abrufbare Forschungsdaten integraler Bestandteil jedes Forschungsprojekts am Zentrum sein müssen, damit sie die in der Gesellschaft und in der Wissenschaftsgemeinschaft anerkannte Wissenschaftskultur unserer Forschungsaktivitäten widerspiegelt. Eine allgemein verbindliche Forschungsdaten-Richtlinie ist nicht nur für die Verifizierung und Verteidigung von Forschungsergebnissen notwendig, sondern bildet auch die prinzipielle Voraussetzung für eine offene, transparente und vernetzte Wissenschaft (Open Science Policy).

Durch die Dynamik des Einsatzes neuer analytischer Instrumente können Forschungsdaten aus aktuellen und auch aus bereits abgeschlossenen Projekten einen zusätzlichen Wert für zukünftige wissenschaftliche Fragestellungen generieren. Dieser lässt sich aber nur realisieren und schützen, wenn bereits im Verlauf der Projekte die gewonnenen Daten vor Verlust geschützt werden und sie durch ihr Format, die Verknüpfung mit entsprechenden Metadaten und die Zuordnung persistenter Identifikatoren (DOI) auch für zukünftige Untersuchungen nachnutzbar bleiben. Die Sicherung der am Zentrum gewonnenen Forschungsdaten für eine Nachnutzung bildet eine essenzielle Voraussetzung, um durch zukünftige innovative (Meta)analysen z. B. zu den komplexen Wechselwirkungen zwischen Umweltfaktoren und der menschlichen Gesundheit neues Wissen zu gewinnen.

2. Ziel der Forschungsdaten-Richtlinie des Helmholtz Zentrums München

- Das Ziel, die am Zentrum gewonnenen Forschungsdaten für zukünftige Vorhaben nachnutzbar zu sichern wird im Bewusstsein verfolgt, dass die materiellen und finanziellen Ressourcen des Zentrums von der öffentlichen Hand bereitgestellt werden. Daher sieht sich das Zentrum verpflichtet, die damit erlangten Forschungsergebnisse mit dem größtmöglichen Potenzial für eine breite Anwendung in der Gesellschaft zu verwenden.
- Mit der Implementierung einer allgemein gültigen Forschungsdaten-Richtlinie inklusive entsprechender Regelungen zur Standardisierung sollen durch Schaffung der Grundlagen für die Verknüpfung von Datensätzen verschiedener Projekte maximale Synergie-Effekte erreicht werden.

- Insbesondere wird mit der Forschungsdaten-Richtlinie auch sichergestellt, dass für zukünftige wissenschaftliche Fragestellungen die am Zentrum bereits gewonnenen Forschungsdaten Nachnutzungs-fähig bleiben (1, 2).
 - Eine allgemein gültige Forschungsdaten-Richtlinie für das Helmholtz Zentrum München erleichtert Kooperationen mit Forschungseinrichtungen im In- und Ausland und erhöht durch die Prinzipien von Open Science die Sichtbarkeit des Zentrums in der internationalen Science Community (2, 3).

3. Geltungsbereich der Forschungsdaten-Richtlinie

Die nachstehende Richtlinie für den Umgang mit Forschungsdaten gilt für alle Mitarbeitenden des Helmholtz Zentrums München, unabhängig davon, wie die jeweiligen Forschungsvorhaben finanziert werden (Grundfinanzierung oder Sonderfinanzierung). Sie findet – im Rahmen der jeweiligen vertraglichen Vereinbarungen – ebenso Anwendung für Forschungsdaten, die im Rahmen von Kooperationen mit externen Partnern gewonnen wurden / werden.

Alle Mitarbeitenden des Zentrums haben sich im Umgang mit Forschungsdaten an dieser Richtlinie zu orientieren, unabhängig davon, auf welcher rechtlichen Grundlage (als festangestellte Forschende, Mitarbeitende an befristeten Projekten, Gastwissenschaftler:innen, Stipendiat:innen etc.) ihre Tätigkeiten erfolgen.

Die jeweils aktuellen Richtlinien des Zentrums zum Datenschutz und zur Informationssicherheit sowie die Richtlinien zur Speicherung von und die Zugriffsrechte auf Forschungsdaten haben Vorrang vor dieser Richtlinie. Insbesondere bedürfen Forschungsdaten als Träger geistigen Eigentums ggf. des Schutzes (IP-Richtlinie des Zentrums). Ihre Verwertungsrechte durch das Zentrum bzw. durch Drittmittel-Geber haben Vorrang vor ihrer Veröffentlichung (siehe 8.1.6) und dürfen durch eine Open-Data Kultur nicht gefährdet werden.

Ausnahmen von der Forschungsdaten-Richtlinie können nur in begründeten Fällen bewilligt werden. Sie bedürfen stets der Prüfung/Zustimmung durch die Abteilungen INNO und SPR.

4. Forschungsdaten - Management

Das Forschungsdatenmanagement umfasst den gesamten Lebenszyklus von Forschungsdaten, von der Erstellung eines Datenmanagementplanes (DMP) über ihre Erfassung und Verarbeitung bis hin zur dauerhaften Archivierung und möglichen Veröffentlichung.

Forschungsdaten sind nach den einschlägigen Standards der jeweiligen Forschungsdisziplin unter Angabe der zu ihrer Erzeugung und Verarbeitung verwendeten Werkzeuge und Methoden sowie des Entstehungskontextes zu erfassen und mit Metadaten, i.e. Beschreibungen des Datenkontextes langfristig zu speichern sowie ggf. zu veröffentlichen. Ihre Archivierung soll gemäß den vier Prämissen Wiederauffindbarkeit, Lesbarkeit, Austauschbarkeit und Nachnutzbarkeit – engl. FAIR (Findable, Accessible, Interoperable and Re-usable) – erfolgen (4). Die Verknüpfung mit ihren Metadaten soll unter Verwendung der etablierten und standardisierten Terminologie der betreffenden Disziplin erfolgen, wobei die Möglichkeit einer fächerübergreifenden Wiederverwendbarkeit und Interoperabilität berücksichtigt werden soll (5).

Der Umgang mit Forschungsdaten im Rahmen definierter Projekte folgt dem DMP, der bereits vor Projektbeginn erstellt wird (6). Im DMP sind die geplante Art der Forschungsdaten, ihre Gewinnung, ihre Aufbewahrung und Zugriffsmöglichkeiten auf die Daten festzulegen und die dafür jeweils Verantwortlichen zu benennen. Bei technisch oder methodisch bedingten Änderungen des Datenflusses im Verlauf eines Projektes ist der DMP anzupassen. Die für die Umsetzung des DMP anfallenden Kosten sind bei Drittmittel-geförderten Projekten als förderfähige Ausgaben bereits bei ihrer Beantragung mit zu

berücksichtigen. Hier berät der Bereich Projektförderung der Abteilung SPR Antragsteller:innen z. B. anhand bereits in anderen Projekten etablierter Datenmanagement-Pläne.

5. Verantwortlichkeiten und Zuständigkeiten

5.1 Jede:r Forschende trägt Verantwortung für:

- A. die Verwaltung und das Management von Forschungsdaten(sätzen) in Übereinstimmung mit den Grundsätzen und Anforderungen dieser Richtlinie und entsprechend den Regeln zur Sicherung guter wissenschaftlicher Praxis des Helmholtz Zentrums München (7),
- B. die Sammlung, Dokumentation, Archivierung und den Schutz vor Verlust der Forschungsdaten inklusive der forschungsbezogenen Aufzeichnungen (Metadaten). Hierzu ist der entsprechende DMP heranzuziehen, der Standard Operating Procedures vorgibt, die die Sammlung, Verwaltung, Integrität, Vertraulichkeit, Speicherung und Veröffentlichung der gewonnenen Daten definiert.
- C. die Einhaltung der allgemeinen Vorgaben des Zentrums, im Falle von Drittmittelprojekten zusätzlich der Vorgaben der Geldgeber sowie bei Kooperationsprojekten mit externen (Forschungs-)Partnern die diesbezüglichen vertraglichen Regelungen. Die besonderen Anforderungen des konkreten Einzelfalls sind im DMP auszuführen.

5.2 Jede:r Projektverantwortliche / -leitende (PL) trägt Verantwortung für:

- A. die Erstellung und Implementierung des Datenmanagementplans (DMP) für das jeweilige Projekt. Neben technischen Fragen zur Sammlung, Verwaltung, Integrität, Vertraulichkeit, Speicherung und Veröffentlichung der Daten sind auch die Einhaltung der allgemeinen Vorgaben des Zentrums und der Geldgeber zu berücksichtigen, um – wo immer möglich –
- B. die Forschungsdaten auch nach Projektende weiter nutzen zu können. Dies beinhaltet insbesondere die Definition von Datennachnutzungsrechten mit der Zuweisung entsprechender Lizenzen, sowie die Klärung der Datenspeicherung und Archivierung mit persistenten DOIs nach Abschluss des Projekts.
- C. eine ggf. festgelegte Vernichtung von Forschungsdaten und Forschungs-bezogenen Aufzeichnungen bei Daten, die als personenbezogene Daten (Patientendaten, Probandendaten (Kohorten)) in den Geltungsbereich der europäischen Datenschutz-GVO fallen. Dies gilt insbesondere auch für den Widerruf einer Einwilligung durch betroffene Personen.
- D. die Bereitstellung geeigneter Ressourcen (technisch / organisatorisch), um allen Mitarbeitenden innerhalb des Projektes ein FAIR-kompatibles Datenmanagement zu ermöglichen.

5.3 Helmholtz Zentrum München / Forschungsdatenmanagement (FDM) / HA DigIT

- A. Für das Forschungsdatenmanagement stellt das Zentrum durch FDM und weitere beteiligte Einheiten (INNO, RA, ZB, SPR-Projektförderung) Unterstützung in Form regelmäßiger Schulungen, gezielte Beratung (inkl. Fragen zu IP, Ethik etc.) zur Verfügung. Zur standardisierten Erstellung von DMP steht Software zur Verfügung.
Zur Schaffung valider und qualitativ hochwertiger Forschungs(meta)daten wird die Nutzung „elektronischer Laborbücher“ (engl. ELN) unterstützt.
- B. Die technischen Voraussetzungen zur Langzeit-Archivierung der Forschungsdaten und der zugehörigen Metadaten (Forschungsdatenrepositorien) werden vom Zentrum bereitgestellt.

6. Veröffentlichung von Forschungsdaten und daraus abgeleiteten wissenschaftlichen Erkenntnissen

Forschungsdaten sollen unmittelbar nach ihrer Erstverwertung durch Forschende des Zentrums öffentlich zugänglich gemacht werden (Publikationen, Patente etc.), sofern dem keine erheblichen Gründe entgegenstehen. Diese können beispielsweise gesetzliche oder vertragliche Regelungen sein oder die Verwendung von Daten aus anderen Forschungsvorhaben mit nicht-permissiver Lizenz, der Schutz

geistigen Eigentums bzw. die beabsichtigte Verwertung der Daten im Rahmen des Technologietransfers. Der Umfang der beabsichtigten Erstverwertung wird im jeweiligen DMP festgelegt. Bei Verträgen mit externen Kooperationspartnern und Gastwissenschaftler:innen ist darauf zu achten, dass keine vertraglichen Bindungen eingegangen werden, die einer Publikation von Ergebnissen mit Beteiligung des Zentrums im Wege stehen. Hierzu erfolgt ggf. eine Beratung durch SPR-Projektförderung und / oder die Abteilung LA.

Für die Zugänglichmachung der Forschungsdaten nach der Erstverwertung stehen Nachvollziehbarkeit und Nachnutzbarkeit der Daten entsprechend den FAIR Prinzipien (s. o.) im Vordergrund. Hierzu sind die Forschungsdatenrepositorien des Zentrums bzw. externe Repositorien zu nutzen. Das Helmholtz Zentrum München stellt dazu Werkzeuge zur Verfügung und berät die Projektleitenden bei deren Nutzung. Werden die Forschungsdaten öffentlich zugänglich gemacht, soll dies nach Möglichkeit unter CC-BY Lizenz geschehen. Im Rahmen der Machbarkeit sollen bislang nicht publizierte, aber publikationswürdige Daten nachträglich frei zugänglich gemacht und nach - Verknüpfung mit einem DOI - in Form von zitierbaren Datenpublikationen veröffentlicht werden. Die DOI-Vergabe erfolgt über die Bibliothek.

7. Anhang

7.1 Definitionen

7.1.1 Datenmanagementplan (DMP)

Um ein nachhaltiges Datenmanagement sicherzustellen, ist bereits vor Beginn eines Forschungsprojekts ein Datenmanagementplan (DMP) aufzustellen. Er soll einen systematischen und dauerhaften Umgang mit den im Rahmen eines Forschungsvorhabens / -projektes entstandenen Forschungsdaten sicherstellen. Darüber hinaus geht er auf die Urheber-, Nutzungs- und Zugangs- und Erstverwertungsrechte sowie die Aufbewahrung während und nach Ende des Forschungsvorhabens ein.

7.1.2 Wissenschaftler:innen

Wissenschaftler:innen sind alle in der Forschung aktiven Mitarbeitenden des Forschungszentrums mit abgeschlossenem Hochschulstudium einschließlich von Bachelor- und Masterstudent:innen sowie Gästen mit entsprechender Qualifikation.

7.1.3 Forschungsdaten

Als Forschungsdaten werden alle Daten bezeichnet, die während des Forschungsprozesses entstehen, dazu herangezogen werden oder dessen Ergebnis sind. Sie werden abhängig von der spezifischen Forschungsfrage und unter Anwendung verschiedener Methoden erzeugt bzw. gesammelt, beobachtet, simuliert oder abgeleitet, danach bearbeitet, validiert, weiterverarbeitet, analysiert, publiziert und schließlich archiviert. Forschungsdaten treten demnach in jeder Forschungsdisziplin und je nach Stadium ihres Lebenszyklus in unterschiedlichen Medientypen und Formaten, Aggregations- und Qualitätsstufen auf. Forschungsdaten, die die Grundlage einer Publikation bilden, zählen zum wissenschaftlichen Output der Wissenschaftler:innen des Zentrums.

7.1.4 Metadaten

Metadaten beschreiben Forschungsdaten. Zum einen umfassen sie Autorenschaft(en), Kontaktdaten der Urheber(organisation), Zeitpunkt der Erstellung, Lizenzen und Schlagworte. Sie definieren die verwendeten Datenformate und enthalten den Kontext, der zu den Daten geführt hat. Darüber hinaus können Metadaten Geräteeinstellungen, Umgebungsbedingungen (z. B. Temperaturen, Drücke), Kommentare und Messunsicherheiten enthalten. Metadaten sind essentiell für die Nachnutzung von Forschungsdaten. Diese Nachnutzung beinhaltet neben der Verwendung für weitere Forschungsarbeiten auch eine Verifikation von Forschungsergebnissen durch Dritte.

7.1.5 Forschungsdatenmanagement (FDM)

Der Begriff Forschungsdatenmanagement bezeichnet den Umgang mit Forschungsdaten von der Planung, ihrer Generierung und Verarbeitung bis zur Veröffentlichung, Langzeit-Archivierung und ggf. Löschung unter Beachtung der Regeln zur Sicherung der guten wissenschaftlichen Praxis und der europäischen DSGVO. Das Management von Forschungsdaten betrifft demnach den gesamten Lebenszyklus von Forschungsdaten.

Darüber hinaus beinhaltet der Begriff die fachspezifische Dokumentation der Prozesse im Kontext ihrer Erhebung. DMP erleichtern die Dokumentation dieser Prozesse und die Beschreibung der Daten.

7.1.6 Schutz geistigen Eigentums und Rechte zur Erstverwertung von Forschungsdaten

Die Schutzwürdigkeit und die Verwertungsrechte von geistigem Eigentum (engl. IP) sind im Arbeitsvertrag zwischen den Forschenden und dem Zentrum als Arbeits- bzw. Gastgeber festgelegt und in der aktuellen Richtlinie zu IP des Zentrums detailliert beschrieben. Urheber- und Verwertungsrechte an allen Formen geistigen Eigentums können durch weitere Vereinbarungen (z. B. Drittmittel-Verträge) erweitert bzw. spezifiziert werden.

Grundsätzlich hat Helmholtz Munich das Recht zu entscheiden, wie Forschungsdaten als Träger geistigen Eigentums veröffentlicht und geteilt werden können. Die Behandlung aller Forschungsdaten nach den FAIR Prinzipien stellt damit zwar die Grundlage von Veröffentlichungen und Datenaustausch dar, erzwingt letztere jedoch nicht. Die Schutzwürdigkeit an dem aus der Forschung des Zentrums hervorgehenden geistigen Eigentum hat Vorrang vor dem Datenaustausch im Rahmen einer zukünftigen Open Science Policy.

Die Erstverwertung von Forschungsdaten ist im Regelfall eine bzw. mehrere Textpublikation(en) in einer Fachzeitschrift, die auf diesen Daten basiert(en). Umfang und Dauer der Erstverwertung kann von den Wissenschaftler:innen im DMP festgelegt werden. Insbesondere soll das Ende der Erstverwertungsphase kontextabhängig und begründet festgelegt werden.

7.1.7 Datenveröffentlichung

Eine Veröffentlichung von Daten kann auf zwei Arten erfolgen:

- A. Eine reine Zugänglichmachung z. B. über eine Webschnittstelle oder eine zitierbare Datenpublikation. Letztere kann die Zugänglichmachung einschließen, muss das aber nicht. Stattdessen kann
- B. auch eine „Landing Page“ eingerichtet werden, über die Interessierte eine Anfrage beim jeweiligen dezentralen Datenmanager (s. o. unter „Verantwortlichkeiten und Zuständigkeiten“) zu Datenherausgabe stellen können.

Ein zentrales technisches Werkzeug zur Veröffentlichung von Forschungsdaten ist ein Datenrepositorium, ein Serverdienst, auf den Daten von ihren Erzeugern gespeichert werden können. Diese Daten erhalten eine weltweit eindeutige Kennung (z. B. einen DOI) und können durchsucht und ggf. heruntergeladen werden. Eine Einschränkung des Zugriffs durch den / die Datenerzeuger mittels Berechtigungseinstellungen ist in Abhängigkeit des jeweiligen Repositoriums möglich.

7.1.8 Lizenzen

Die Lizenz einer Veröffentlichung ist ein Vertrag zwischen veröffentlichenden und nachnutzenden Personen oder Institution, der die Bedingungen der Nachnutzung regelt. Viele Lizenzen“typen“ sind bereits vorformuliert über das Internet verfügbare, so dass hier auf sie verwiesen sei (8). So schränkt eine permissive Lizenz die Nachnutzung nur minimal ein. Bekannte Beispiele permissiver Lizenzen sind CC-0 und CC-BY der Creative Commons.

→ CC-0: <https://creativecommons.org/publicdomain/zero/1.0/deed.de>

→ CC-BY: <https://creativecommons.org/licenses/by/4.0/deed.de>

7.2 Rechtlicher Status von Daten

Ein Eigentum an Daten existiert in Deutschland nicht. Das Urheberrecht garantiert jedoch jedem den Schutz seiner geistigen Schöpfungen. Ob Forschungsdaten dem Schutz des Urheberrechtsgesetzes unterliegen, ist davon abhängig, ob die Anforderungen an die geistige Schöpfungshöhe oder die Voraussetzungen des Datenbankrechts erfüllt werden. Beides wird z. B. bei Routinemessungen regelmäßig nicht erfüllt. Je nach Art der Erhebung können Forschungsdaten eine schützenswerte geistige Schöpfungshöhe besitzen. Um für Rechtssicherheit zu sorgen, sollte stets von einer Schutzwürdigkeit ausgegangen werden, d. h. Nutzungs- und Verwertungsrechte sind in der Regel vertraglich mit externen Partner:innen zu klären. Ebenso ist auf sorgfältige Lizenzierung von veröffentlichten Daten zu achten.

Unabhängig vom Recht an den gewonnenen Daten, ist die Frage des Verwertungsrechtes zu sehen. Der Forschende bzw. die Forschungseinrichtung kann ein (zeitlich befristetes) Erstverwertungsrecht der im Rahmen ihrer Forschungsaktivitäten gewonnenen Daten in Form von Veröffentlichungen oder Patentschriften geltend machen.

Zitate

- (1)Strategiepapier des Forum Gesundheitsforschung des BMBF „Nutzbarmachung digitaler Daten für KI - und datengetriebene Gesundheitsforschung“, Berlin 2020
- (2)Berliner Erklärung über den offenen Zugang zu wissenschaftlichem Wissen, Helmholtz Open Science 2003, : <https://openaccess.mpg.de/Berliner-Erklaerung>
- (3)Positionspapier des Deutschen Wissenschaftsrates „Digitalisierung und Datennutzung für Gesundheitsforschung und Versorgung“, Köln, Juli 2022
- (4)Wilkinson, M. D. et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci. Data 3:160018. doi: 10.1038/sdata.2016.18 (2016).
- (5)LEARN Toolkit of Best Practice for Research Data Management. <http://learn-rdm.eu/en/research-data-management-toolkit-now-available/>
- (6)Empfehlungen für Richtlinien der Helmholtz-Zentren zum Umgang mit Forschungsdaten. In der 109. Mitgliederversammlung der Helmholtz-Gemeinschaft am 13.-14. September 2017 <https://doi.org/10.2312/os.helmholtz.002>
- (7)https://hint.helmholtz-munich.de/fileadmin/Intranet/Recht_und_Compliance/Interne_Regelungen/Dokumente_DE/Regeln_zur_Sicherung_guter_wissenschaftlicher_Praxis_06-10-2015.pdf
- (8) <https://creativecommons.org/licenses/>

München, 01.01.2023