Stiftung Tierärztliche Hochschule Hannover
University of Veterinary Medicine Hannover

Forschungsdatenmanagement

Dr. Max Hassenstein
forschungsdaten@tiho-hannover.de

Checkliste für Antragstellende zur Beschreibung des Umgangs mit Daten im Rahmen eines Forschungsvorhabens

In Anlehnung an die „Checkliste zum Umgang mit Forschungsdaten“ [1] der Deutschen Forschungsgemeinschaft (DFG) sowie die kommentierte Version der Uni Hamburg [2]

1. Datenbeschreibung

Auf welche Weise entstehen in Ihrem Projekt neue Daten?

Daten können Messungen, Befragungen, Foto- und Filmmaterial, Auswertungen, Software- und Analysecodes sein.
Bitte beschreiben Sie die genutzten Methoden zur Datenerzeugung umfassend und welche Art von Daten dabei entstehen (z. B. Datentabellen in CSV; Videomaterial in MP4) und wie Sie diese nutzen werden. Sofern zutreffend, geben Sie auch an, welche fachspezifischen Standards gelten und wie Sie diese erfüllen.

Nutzen Sie bereits existierende Daten (sog. Sekundärdaten), die nicht im Rahmes Ihres konkreten Forschungsvorhabens entstanden sind?

Wenn ja, nennen Sie bitte die Datenherkunft (inkl. DOI, sofern vorhanden), charakterisieren die Daten (z. B. Datentabellen in CSV; Videomaterial in MP4) und den Kontext ihrer Erzeugung. Geben Sie ggf. an, ob und wie Sie mehrere bereits verfügbare Daten miteinander verknüpfen. Bitten geben Sie an, ob Nutzungsbedingungen beachtet werden müssen.
Sofern keine Daten zur Weiterverwendung/Nachnutzung vorliegen, reicht ein Hinweis, das auf Basis einer Recherche keine adäquaten Daten zur Nachnutzung gefunden wurden oder die Begründung, warum sich existierende Daten zur Nachnutzung nicht für das Forschungsvorhaben eignen.

Wie verarbeiten Sie die generierten bzw. nachgenutzen Daten in Ihrem Projekt?

Bitte beschreiben Sie den Prozess der Datennutzung, Speicherung, langfristigen Ablage und ggf. der Datenpublikation. Dabei ist zu nennen, wer während der Projektlaufzeit Zugriff auf die Daten haben soll, aber auch ob und in welchem Umfang die erzeugten Daten publiziert werden sollen.
Des Weiteren geben Sie bitte an, wie Sie mit den Daten umgehen werden bzw. welche Methoden angewendet werden sollen. In diesem Zusammenhang sind auch direkte Aspekte wie die Dateibenennung und –ablage zu nennen.

In welchem Umfang fallen in Ihrem Projekt Daten an und können Sie abschätzen, wie viel Speicherplatz erforderlich ist?

Bitte charakterisieren Sie die Messungen wie z. B. die Anzahl der Messungen, Umfang der Foto-/Videomaterialien oder zu führenden Interviews. Sofern möglich, schätzen Sie bitte den erwarteten Speicherbedarf ein. Beispiel: „Wir planen Videoaufnahmen des Stalls an 10 konsekutiven Nächten zu je 6h. Pro Nacht entsteht Videomaterial von ca. 6 GB. Somit schätzen wir den Speicherbedarf der erzeugten Forschungsdaten auf ca. 60 GB für die Rohdaten und zusätzlich 60 GB für die Datensicherung."

Bitte geben Sie an, ob Sie den Speicherbedarf kurzfristig, oder über die gesamte Projektlaufzeit und ggf. auch zur Archivierung benötigen.

2. Dokumentation und Datenqualität

Verfolgen Sie Ansätze zur nachvollziehbaren Beschreibung der Daten (wie z. B. Dokumentations- und Metadatenstandards oder Ontologien)?

Die Beschreibung der Daten sollte genaue Informationen zum Ursprung, der Erhebungsmethode und den verwendeten Mess- und Endgeräten sowie Angaben zum Messzeitpunkt und den beteiligten Personen enthalten, z. B. in Form eines Protokolls.

In der Dokumentation ist stets die Messeinheit zu beschreiben um eine Nachvollziehbarkeit zu gewährleisten. Sollten Abkürzungen über allgemein bekannte Standardeinheiten hinausgehen (z. B. „min", „mL"), sollte idealerweise ein Abkürzungsverzeichnis angelegt werden.

Welche Maßnahmen treffen Sie, um die Qualität der Daten sicherstellen?

Bitte spezifizieren Sie die getroffenen Maßnahmen zur Sicherung der Datenqualität, bspw. die Eichung von Geräten, Validierungsmessungen, Plausibilitätskontrollen.

Sehen Sie Qualitätskontrollen der Daten vor und falls ja, in welcher Form?

Idealerweise überprüfen Sie die relevanten Dimensionen der Datenqualität, z. B. den Vollständigkeitsgrad, Genauigkeit, oder Einheitlichkeit (consistency).

Sind für die Datennutzung spezielle Methoden oder Tools (auch Software und Softwarecode) zur Nutzung erforderlich?

Bitten geben Sie an, was Sie zur Datennutzung benötigen bzw. bereits verwenden. Wenn Sie über Standardsoftware (z. B. SAS, MS Excel) hinaus Softwarecode und -pakete programmieren, dann sollte dazu eine umfassende Dokumentation gereicht werden, die eine Überprüfung und Nachnutzung ermöglicht. Messdaten aus Geräten mit proprietären Formaten sollten exakt dokumentiert werden und möglichst zusätzlich in ein universell und frei verwendbares Format (z. B. CSV oder im Falle von Bildmaterial JPG, PNG) überführt werden. Falls Sie keinen Datenverlust bei der Konvertierung zwischen Dateiformaten ausschließen können, behalten Sie zusätzlich die Originaldatei.

3. Datenspeicherung und -sicherung im Projektverlauf

Wie werden die Daten im Projektverlauf gespeichert und gesichert?

Bitte beschreiben Sie, auf welchen Geräten bzw. Netzwerken und Verzeichnissen Sie die Forschungsdaten ablegen. Sofern die Daten nicht auf Netzwerklaufwerken der TiHo liegen (kön-

nen), sorgen Sie eigenständig für eine redundante Datensicherung (Backup), welche idealerweise nicht am selben Ort lagert. Bitte beachten und kommentieren Sie ggf. notwendige Aufbewahrungspflichten.

Sofern Sie mit sensiblen Daten im Rahmen Ihres Projekts arbeiten, ist die Sicherheit über Zugriffs- und Nutzungsverwaltung gewährleistet?
Bitte legen Sie fest und beschreiben Sie, welcher Personenkreis in welchem Umfang über die Daten während der Projektlaufzeit verfügen, diese einsehen oder darüber bestimmen darf. Beschreiben Sie die Maßnahmen, damit Unbefugte keinen Zugang zu den Daten erlangen (z. B. Datenverschlüsselung im Falle der Verwendung externer Datenträger wie USB-Sticks).

4. Rechtliche Rahmenbedingungen und Verpflichtungen

Bestehen im Kontext der von Ihnen im Projekt erzeugten oder (weiter-)genutzten Daten rechtliche Besonderheiten?
Rechtliche Besonderheiten entstehen oftmals im Umgang mit personenbezogenen Daten, Geschäftsdaten und Patenten, durch Urheberrecht geschützte Daten, mit Einschränkungen bereitgestellte Daten. Zudem ist besonders auf Daten zu achten, deren Informationsgehalt eine Gefahr birgt (z. B. Standortangaben gefährdeter Arten oder wertvoller Objekte).

Sind Restriktionen in Bezug auf die Datenpublikation oder den Datenzugang durch Dritte nach Projektabschluss zu erwarten?
Sofern Einschränkungen mit Bezug auf rechtliche oder sonstige Aspekte vorliegen, beschreiben Sie diese bitte möglichst transparent.

Inwiefern berücksichtigen Sie Eigentumsfragen sowie nutzungs- und urheberrechtliche Aspekte?
Über die Veröffentlichung von Daten, die z. B. dem Urheberrecht unterliegen (z. B. geistige Leistungen, Werke oder geistige Schöpfungen) und Sie nicht eigenständig geschafften haben, dürfen nicht ohne Weiteres veröffentlicht werden. Klären Sie möglichst zu Beginn des Forschungsvorhabens, wer welche Nutzungsrechte an den Daten erwirbt. Legen Sie, sofern Sie selbst die Daten generiert haben und keine rechtlichen Aspekte dagegensprechen, eine möglichst offene Lizenz fest, welche eine einfache Datennachnutzung ermöglicht. Von Ihnen nachgenutzte bereits veröffentlichte Daten müssen nicht als Duplikat veröffentlicht werden, gleichwohl ist ein Verweis auf diese Daten notwendig (z. B. über den DOI). Falls sie die nachgenutzten Daten für Ihre eigenen Zwecke modifiziert haben, empfiehlt sich die Publikation der Modifikation einschließlich der Dokumentation der durchgeführten Schritte, z. B. zusammen mit geschriebenen Softwarecodes.

Müssen im Rahmen Ihrer Forschungstätigkeiten bedeutende Normen oder Kodizes berücksichtigt werden?

Bitte prüfen Sie die Empfehlungen der relevanten Fachgesellschaften und Forschungseinrichtungen. Auch Empfehlungen des Mittelgebers können hier zum Tragen kommen.

5. Datenaustausch und dauerhafte Datenzugänglichkeit

Welche der im Projekt erzeugten Daten bieten sich für eine Nachnutzung an, auch im Zusammenhang anderer Fachgebiete?
Durch die Nachnutzungsmöglichkeiten der erzeugten Daten über den Fachbereich hinaus eröffnen sich große Potenziale. Allerdings sind in diesem Zuge die umfassende Dokumentation und Beschreibung der Daten erforderlich, um potenzielle Hürden im Vorfeld zu vermeiden. Nennenswert sich hier die FAIR-Leitprinzipien (s. auch DFG-Leitlinien zur Sicherung guter wissenschaftlicher Praxis [3]).

Welches sind die Kriterien, nach denen Sie bestimmen, welche Forschungsdaten publiziert bzw. zur Nachnutzung öffentlich bereitgestellt werden?
Nennen Sie die von Ihnen festgelegten Kriterien. Bitte nennen Sie auch Hürden, Hemmnisse und Barrieren.

Sofern Sie die Datenarchivierung in einer Infrastruktur (z. B. Repositorium) absehen, welche planen Sie wie zu nutzen?
Bitte prüfen Sie, ob die Infrastruktur für die langfristige Datenarchivierung geeignet ist. Idealerweise begründen Sie zudem Ihre Auswahl (z. B. Renommee im Fachgebiet, hohe Reichweite). Als allgemeines Repositorium bietet sich bspw. Zenodo an.
Sofern Sie kein adäquates Repositorium ausfindig machen können oder Ihre Daten sich nicht zur Datenpublikation in den Repositorien anbieten, stellt sich die eLib der TiHo (https://elib.tiho-hannover.de) als Alternativangebot dar, in dem zukünftig auch Forschungsdaten abgelegt werden können.
Wenn keine Datenpublikation oder –archivierung in öffentlich zugänglichen Ressourcen vorgesehen ist, begründen Sie dies bitte. Explizit von der Empfehlung zur Datenpublikation ausgenommen sind Daten, die in identischer Form bereits publiziert und referenzierbar sind.

Sofern Ihre Forschungsdaten Sperrfristen unterliegen, ab welchem Zeitpunkt sind diese aufgehoben, damit die Daten durch Dritte nutzbar sind?
Sperrfristen und Embargos können in ausgewählten Repositorien eingestellt werden. Somit können Sie die Datenpublikation vorbereiten und die erforderlichen Arbeiten abschließen, ohne dass die Daten direkt veröffentlicht werden.

6. Verantwortlichkeiten und Ressourcen

Welche Person(en) sind im Rahmen des Projekts für einen adäquaten Umgang mit den Forschungsdaten verantwortlich?

Bitte beschreiben sie die Rollen der betroffenen Personen. Zwar ist jede Person zunächst für die „eigenen“ Daten zuständig, bei größeren Projekten ist es allerdings sinnvoll zu erwägen, ob eine oder mehrere Personen eine explizite Verantwortung für den Umgang mit Forschungsdaten übernehmen. Somit kann ein guter Umgang mit Forschungsdaten leichter gewährleistet werden.
Die Zusammenarbeit mit einer Beratungsstelle zum Forschungsdatenmanagement kann im Rahmen eines Projektantrags positiv hervorgehoben werden.

Welche Ressourcen (z. B. Zeit, Kosten) sind erforderlich, um einen angemessenen Umgang mit Forschungsdaten in Ihrem Projekt umzusetzen?
Das Forschungsdatenmanagement sollte nicht als Nebenangelegenheit abgetan werden. Die Transparenz, Nachvollziehbarkeit und Reproduzierbarkeit sind im Rahmen des Forschungsprozesses sehr bedeutend. Die DFG bietet, insbesondere in Verbundprojekten, Förderungen an.

Wer ist nach Ablauf der Projektlaufzeit für die archivierten bzw. publizierten Daten verantwortlich?
Bitte beschreiben sie die Rollen der betroffenen Personen. Zwar ist jede Person zunächst für die „eigenen“ Forschungsdaten verantwortlich, ggf. auch eine zuvor im Rahmen des Projekts bestimmte Person. Mit dem Abschluss des Projekts und der Datenpublikation beginnt ein neuer Abschnitt, der aber an sich deutlich geringere und idealerweise kaum Ressourcen erfordert: Unter Umständen sind nachträgliche Änderungen oder Korrekturen am Datensatz oder der Dokumentation erforderlich. Hier sollte festgelegt werden, welche Person oder Arbeitsgruppe (wichtig z. B. bei Ausscheiden einer Person aus einer Einrichtung) für die Datenpflege verantwortlich ist. Auch sollte dringend daran gedacht werden, intern zu dokumentieren, wo Daten liegen und wie diese über welchen Zugang ggf. angepasst werden können.

Referenzen

1. Deutsche Forschungsgemeinschaft (2021) Umgang mit Forschungsdaten. Checkliste für Antragstellende zur Planung und zur Beschreibung des Umgangs mit Forschungsdaten in Forschungsvorhaben. Abrufbar unter: https://www.dfg.de/de/grundlagen-themen/grundlagen-und-prinzipien-der-foerderung/forschungsdaten (Abgerufen am 22.08.2024)
2. Deutsche Forschungsgemeinschaft und ZFDM der Uni Hamburg (2022) Umgang mit Forschungsdaten. Mit Kommentaren und Hinweisen ergänzte DFG-Checkliste. Abrufbar unter: https://www.fdm.uni-hamburg.de/ueber-uns/a-nachrichten/fdm-dfg-checklist.html (Abgerufen am 22.08.2024)
3. Deutsche Forschungsgesellschaft (2022) Guidelines for Safeguarding Good Research Practice. Code of Conduct. DOI: 10.5281/zenodo.3923601