
ROADS to Reusable Datasets: Potenziale und Herausforderungen bei der Datennachnutzung in der Hochschullehre

Anna Daudrich  ¹

¹ ZHAW Zürcher Hochschule für Angewandte Wissenschaften, Schweiz

Die Wiederverwendung von Forschungsdaten bietet die Möglichkeit, wissenschaftliche Ressourcen langfristig effizienter und nachhaltiger zu nutzen. In der Praxis stehen diesem Potenzial jedoch zahlreiche Hürden entgegen. Der vorliegende Beitrag untersucht diese Herausforderungen und formuliert die Hypothese: Um die Datennachnutzung erfolgreich zu fördern, sollte deren Integration bereits im Studium gezielt unterstützt werden. Frühzeitige Erfahrungen mit der Nachnutzung von Forschungsdaten ermöglichen es Studierenden, zentrale Kompetenzen in den Bereichen Data Literacy und Forschungsdatenmanagement zu entwickeln – und schaffen somit die Grundlage für eine nachhaltige und effektive Nutzung von Daten in der späteren wissenschaftlichen Praxis.

Diese Hypothese stützt sich auf Erkenntnisse aus dem Projekt ROADS (Reusing Openly Accessible Research Data for Student Theses) an der ZHAW Zürcher Hochschule für Angewandte Wissenschaften. Eines der Projektziele ist es, Chancen und Herausforderungen der Datennachnutzung zu identifizieren – basierend auf Anwendungsfällen, in denen Studierende aus verschiedenen Disziplinen bestehende Forschungsdaten nachnutzten. Eine der herausfordernden Aufgaben für Studierende war die Datenrecherche und -interpretation – was auf die mangelnde Datenkompetenz zurückzuführen ist. Die Vielzahl an Datenquellen, die heterogene Datenlandschaft sowie die häufig unzureichende Dokumentation veröffentlichter Datensätze erschweren den Prozess der Datensuche. Die Bewältigung dieser Herausforderungen in den studentischen Anwendungsfällen erforderte einen hohen Zeit- und Arbeitsaufwand.

Trotz dieses Aufwands zeigt sich: Die Datennachnutzung bietet einen erheblichen Mehrwert. Wird sie systematisch in die Lehre integriert, können Studierende (und teilweise auch ihre Betreuenden) essenzielle Kompetenzen im Umgang mit Daten erwerben und

Publiziert in: Vincent Heuveline, Philipp Kling, Florian Heuschkel, Sophie G. Habinger und Cora F. Krömer (Hrsg.): E-Science-Tage 2025. Research Data Management: Challenges in a Changing World. Heidelberg: heiBOOKS, 2025. DOI: <https://doi.org/10.11588/heibooks.1652.c23923> (CC BY-SA 4.0).

ein Bewusstsein für effizientes Forschungsdatenmanagement entwickeln. Langfristig kann dies zu ressourcenschonenderen und effizienteren Forschungspraktiken beitragen.

Keywords: Datennachnutzung, Datenkompetenz, Datenqualität, Forschungsdaten, FAIR

1 Einleitend: Ressourcenschonung durch Datennachnutzung – Chance oder Herausforderung?

Nachhaltigkeit und ein effizienter Umgang mit Ressourcen gewinnen zunehmend an Bedeutung. Auch im Forschungskontext ist dieses Thema präsent. Zahlreiche Hochschulen und Forschungsinstitutionen haben Nachhaltigkeitsstrategien inzwischen fest verankert.¹ Ein sorgfältig geplantes Forschungsdatenmanagement (FDM) sowie Open Research Data (ORD) können maßgeblich dazu beitragen, ressourcenschonende Forschungspraktiken zu fördern – ein Aspekt, der auch in den Richtlinien der Forschungsförderer häufig hervorgehoben wird.²

Insbesondere die Nachnutzung bereits vorhandener Daten birgt großes Potenzial: Anstatt zeit- und kostenintensiv neue Daten zu erheben, können Forschende auf bestehende Datensätze zurückgreifen. Dies soll helfen, Doppelerhebungen von Daten sowie redundante wissenschaftliche Untersuchungen zu vermeiden (Hollstein und Strübing 2018). Gerade bei Langzeitstudien oder bei schwer erreichbaren Personengruppen erweist sich diese Praxis als besonders vorteilhaft. Die Nachnutzung von Daten aus quantitativen Befragungen oder qualitativen Interviews minimiert zudem datenschutzrechtliche Herausforderungen, da bereits anonymisierte Daten verwendet werden können und keine neuen Einwilligungen von Befragten erforderlich sind (Netscher und Trixa 2019).³ Darüber hinaus entfallen Auf-

1 Siehe dazu beispielsweise die Nachhaltigkeitsstrategien der folgenden Hochschulen: die ZHAW Zürcher Hochschule für Angewandte Wissenschaften (<https://www.zhaw.ch/de/fokusthemen/zhaw-sustainable/nachhaltigkeitsstrategie>; *Besucht am 30. März 2025*), die Universität Zürich (<https://www.sustainability.uzh.ch/de/policy-reporting.html>; *Besucht am 30. März 2025*) sowie die Friedrich-Schiller-Universität Jena (<https://www.uni-jena.de/207933/nachhaltigkeitsstrategie>; *Besucht am 30. März 2025*).

2 Siehe dazu beispielsweise die Nationale Schweizer Strategie für Open Research Data (<https://www.swissuniversities.ch/themen/open-science/open-research-data/nationale-strategie>; *Besucht am 30. März 2025*) sowie Grundlagen zur nachhaltigen Forschung bei der Deutschen Forschungsgemeinschaft (<https://www.dfg.de/de/grundlagen-themen/entwicklungen-im-wissenschaftssystem/nachhaltigkeit-im-forschungsprozess>; *Besucht am 30. März 2025*).

3 Anonymisierte Daten gelten nicht mehr als personenbezogene Daten. Sie unterliegen daher nicht den Bestimmungen des Datenschutzrechts. Das bedeutet, dass sie grundsätzlich für andere Zwecke verwendet werden dürfen – vorausgesetzt, eine Re-Identifizierung ist ausgeschlossen, und in der informierten Einwilligung (Informed Consent) wurde über eine mögliche Weiternutzung und/oder Weitergabe der Daten informiert. Zu beachten ist weiter, dass einer Weiternutzung und / oder Weitergabe die Ver-

gaben wie die Konzeption von Interviewleitfäden oder die Akquise von befragten Personen (Bendel 2021).

Obgleich die Nachnutzung von Forschungsdaten Chancen für eine ressourcenschonende Forschung bietet, stehen ihrer umfassenden Ausschöpfung zahlreiche Hürden entgegen. Im Folgenden werden diese Herausforderungen näher beleuchtet. Die Darstellung stützt sich auf Ergebnisse des Projekts ROADS („Reusing Openly Accessible Research Data for Student Theses“) an der ZHAW Zürcher Hochschule für Angewandte Wissenschaften.⁴ Zudem werden zentrale Begrifflichkeiten im Zusammenhang mit der Datennachnutzung erläutert – insbesondere im Hinblick auf das ‚R‘ (Reusable) der FAIR-Prinzipien⁵. Hierbei wird der Beitrag aufzeigen, unter welchen Bedingungen Daten als nachnutzbar gelten. Es ist dabei wichtig zu betonen, dass sich die Schlussfolgerungen dieses Beitrags auf Fallbeispiele stützen und daher nicht als allgemeingültige Aussagen, sondern als empirisch fundierte Hypothesen verstanden werden sollten.

2 Fallbeispiel: Projekt ROADS

2.1 Ausgangslage

Das Projekt ROADS wurde an der ZHAW Zürcher Hochschule für Angewandte Wissenschaften ins Leben gerufen, vor dem Hintergrund, dass in der Forschung inzwischen eine hohe Vielzahl an Datensätzen generiert wurde – sei es in Form von Statistiken, Laborexperimenten, Bildern, Transkripten von Interviews usw. Zahlreiche dieser Datensätze stehen bereits den Forschungscommunities und teils auch der breiten Öffentlichkeit zur Verfügung, da sie etwa in generischen, fachspezifischen oder institutionellen Repositorien als ORD veröffentlicht sind. Damit das Potenzial offener Daten voll ausgeschöpft werden kann, ist es entscheidend, dass sie die FAIR-Kriterien erfüllen, also auffindbar, zugänglich, interoperabel und nachnutzbar sind.⁶

Trotz zahlreicher Initiativen zur Förderung von ORD werden weiterhin nicht alle Datensätze veröffentlicht und verbleiben oft innerhalb der Institutionen, in denen sie generiert wurden.⁷ Diese Datensätze erfüllen zwar nicht alle Kriterien der FAIR-Prinzipien – sie

traulichkeit der Daten oder vertraglichen Regelungen entgegenstehen können (siehe dazu Art. 2 Abs. 1 DSG sowie Art. 5 lit. a DSG).

4 Projektlaufzeit: Dezember 2023 – Juni 2025, Co-Finanzierung: swissuniversities. Nähere Informationen zum Projekt: <https://www.zhaw.ch/de/forschung/projekt/74318>; *Besucht am 30. März 2025*.

5 Nähere Informationen zu den FAIR-Prinzipien: <https://www.go-fair.org/fair-principles/>; *Besucht am 30. März 2025*.

6 <https://www.unesco.org/en/open-solutions/open-data>; *Besucht am 30. März 2025*.

7 Dies bestätigen unter anderem die Ergebnisse der Umfrage des Schweizerischen Nationalfonds (SNF): <https://www.snf.ch/de/iyEnUsdV0AiHk9wF/news/open-research-data-ein-erster-blick-auf-die-aktuelle-praxis>; *Besucht am 30. März 2025*.

sind beispielsweise weder durch Suchmaschinen auffindbar noch für die (breite) Öffentlichkeit zugänglich –, dennoch besitzen sie weiterhin Potenzial zur Nachnutzung. Insbesondere die Datenautor:innen und andere Beteiligte an den Projekten, in deren Rahmen die Daten ursprünglich erhoben wurden, haben grundsätzlich die Möglichkeit, auf solche unveröffentlichten Datensätze zurückzugreifen und sie in neuen Forschungskontexten wiederzuverwenden.

Daraus ergibt sich die zentrale Frage: Sind all diese (öffentlich) verfügbaren Datensätze tatsächlich nachnutzbar, und was bedeutet „nachnutzbar“ überhaupt in diesem Kontext? Diese Überlegungen bildeten die Ausgangslage für das Projekt ROADS. Die Nachnutzbarkeit von Daten sowie die damit verbundenen Chancen und Herausforderungen wurden anhand konkreter Anwendungsfälle untersucht. Studierende der ZHAW auf verschiedenen Stufen (Bachelor, Master, Doktorat) und aus unterschiedlichen Disziplinen nutzten dabei bereits vorhandene Datensätze für ihre Abschlussarbeiten, Seminararbeiten und weitere studentische Projekte. Die Datennachnutzung erfolgte unter der Betreuung von Dozierenden aus verschiedenen ZHAW-Departementen. Zusätzlich wurde dieser Prozess durch Data Stewards aus der Hochschulbibliothek und den ZHAW Services Research Data unterstützt.

In jedem Anwendungsfall konnten die Studierenden und ihre Betreuenden selbst entscheiden, welche Daten genau und in welchem Kontext nachzunutzen sind. Zur Auswahl standen dabei vor allem offene Forschungsdaten. Aber auch unveröffentlichte Daten, etwa aus abgeschlossenen oder laufenden Forschungsprojekten der Betreuenden, boten sich grundsätzlich für eine Nachnutzung an. Da die Betreuenden in der Regel Zugang zu solchen Daten haben, konnten sie diese ihren Studierenden zu einer weiteren Auswertung zur Verfügung stellen – vorausgesetzt, dies war rechtlich zulässig. Tabelle 1 auf Seite 175 zeigt eine Übersicht der Daten, die zur Nachnutzung ausgewählt wurden.

Wie die Tabelle 1 zeigt, handelt es sich in den meisten Anwendungsfällen um nicht publizierte quantitative Daten. Die Entscheidung zugunsten quantitativer statt qualitativer Daten ist in diesem Fall wenig überraschend. Qualitative Daten sind in der Regel sehr spezifisch. Ihre Erhebung wird gezielt an den jeweiligen Forschungsgegenstand angepasst, so dass sie mit ihrem sozialen sowie kontextuellen Entstehungsrahmen eng verbunden sind (Gebel und Meyermann 2021). Die Entscheidung für nicht veröffentlichte Daten scheint jedoch im Widerspruch zu den Zielen von ORD und den FAIR-Prinzipien zu stehen: Wenn offene Forschungsdaten zugänglich und nachnutzbar sein sollen, warum werden dann stattdessen nicht veröffentlichte Daten wiederverwendet?

Diese und weitere Erkenntnisse aus den Anwendungsfällen im Rahmen von ROADS werden im Folgenden dargestellt. Dabei werden diese sowohl aus der Perspektive der Studierenden als auch der Betreuenden erläutert. Diese Erkenntnisse wurden auf unterschiedlichen Wegen erfasst: Die Studierenden reflektierten die Nachnutzung im jeweiligen Anwendungsfall durch schriftliche Ausführungen, die Teilnahme an Online-Umfragen sowie an

qualitativen Interviews. Die Betreuenden trugen ihre Reflexionen zur Datennachnutzung im Rahmen regelmäßiger Austauschgespräche und Projekt-Workshops bei.

Tabelle 1: Datennachnutzung in studentischen Arbeiten im Rahmen des Projekts ROADS.

Forschungsbereich	Datenarten	Verfügbarkeit	Kontext der Nachnutzung
Physiotherapie	Quantitative Bewegungsdaten, Qualitative Interviews	unveröffentlicht, veröffentlicht	Bachelor- & Masterarbeiten
Datenjournalismus	Quantitative Umfragen	unveröffentlicht, veröffentlicht	Seminararbeiten
Psychologie	Quantitative Umfragen	unveröffentlicht	Bachelor- & Masterarbeiten
Computergestützte Lebenswissen.	Quantitative Lebensmittel- & Gesundheitsdaten	unveröffentlicht, veröffentlicht	Seminararbeiten
Materialwissen.	Quantitative Daten zu Perowskit-Solarzellen	veröffentlicht	wissenschaftliche Veröffentlichung

2.2 Perspektive der Studierenden

Die Studierenden berichteten mehrheitlich, entweder keine oder nur geringe Erfahrung in der Nachnutzung von Daten zu haben. Ebenso war vielen – insbesondere auf der Bachelor-Stufe – die Erhebung eigener Daten noch nicht vertraut, wobei diese insgesamt häufiger praktiziert wurde als die Datennachnutzung. Zudem zeigten sich bei einigen Studierenden, vor allem zu Beginn des Bachelorstudiums, nicht nur Defizite im Umgang mit Daten, sondern auch Lücken in grundlegenden wissenschaftlichen Arbeitsmethoden. Aufgrund dieser mangelnden Erfahrung traten bereits zu Beginn der Anwendungsfälle zahlreiche Herausforderungen auf.

Ein zentrales Problemfeld stellte die Recherche und das Finden eines geeigneten Datensatzes dar. Viele Studierende verfügten über keine oder nur geringe Kenntnisse darüber, wo und wie sie nach relevanten Datensätzen suchen sollten. Diese Schwierigkeiten resultierten nicht nur aus mangelnder Erfahrung, sondern auch aus dem Fehlen spezifischer Schulungen zur Datenrecherche. Während in zahlreichen Studiengängen Kompetenzen in der Literaturrecherche – häufig vermittelt durch Bibliothekspersonal⁸ – gefördert werden,

⁸ Siehe beispielsweise das Angebot der ZHAW Hochschulbibliothek zur Rechercheberatung: <https://www.zhaw.ch/de/hochschulbibliothek/recherchehilfe-kurse#c5338>, *Besucht am 31.03.2025*.

existieren vergleichbare Angebote für die Datenrecherche in der Regel nicht. Dies führte dazu, dass Datensätze oft zufällig, unsystematisch oder mit wenig zielführenden Suchstrategien identifiziert wurden, was die Eignung der Daten für die jeweilige Forschungsfrage erheblich einschränkte.

Eine zusätzliche Herausforderung ergab sich aus der häufig fehlenden Kompetenz im Umgang mit Software-Programmen, die für die Bearbeitung und Analyse der Daten erforderlich sind. Selbst bei gängiger Software wie Excel mussten die Studierenden oft zusätzliche Zeit investieren, um sich die notwendigen Kenntnisse anzueignen.

Neben unzureichenden Kompetenzen in der Datenrecherche und in der Bedienung von Software stellte die Qualität der veröffentlichten Datensätze eine weitere erhebliche Herausforderung dar. Häufig fehlten essenzielle Informationen in der Dokumentation, beispielsweise Erläuterungen zu zentralen Variablen oder eine detaillierte Beschreibung der Laborbedingungen, die für das Verständnis und die Interpretation der Daten erforderlich sind. Diese Dokumentationslücken erschwerten es den Studierenden erheblich, die Daten korrekt zu kontextualisieren und ihre Relevanz für die jeweilige Forschungsfrage adäquat zu bewerten.

Ein weiteres Hindernis war die mangelnde Standardisierung in der Datenerhebung. So werden Messdaten oft unter unterschiedlichen Laborbedingungen oder ohne die Anwendung standardisierter Messverfahren erfasst, was die Zusammenführung und den Vergleich verschiedener Datensätze erschwert. Ohne ein solides Verständnis wissenschaftlicher Standards fiel es vielen Studierenden schwer, diese Mängel zu erkennen und entsprechend kritisch zu bewerten.

Schließlich ist der Zugang zu publizierten Daten nicht immer uneingeschränkt gewährleistet. In einigen Fällen werden lediglich die Metadaten veröffentlicht, während der vollständige Zugriff auf die eigentlichen Daten erst nach der Kontaktaufnahme mit den Datenautor:innen gewährt wird – ein Prozess, der zu Verzögerungen führen kann. Gerade bei zeitlich beschränkten Vorhaben – etwa bei studentischen Arbeiten – kann dies problematisch sein, wenn eine zeitnahe Analyse der Daten erforderlich ist.

Angesichts dieser Herausforderungen sahen sich die Studierenden häufig gezwungen, nach alternativen Datensätzen zu suchen, selbst wenn der ursprünglich gefundene Datensatz auf den ersten Blick geeignet erschien. In einigen Fällen war es sogar notwendig, die ursprüngliche Forschungsfrage anzupassen oder sogar vollständig zu ändern, da keine geeigneten Datensätze für die ursprünglich formulierte Frage gefunden werden konnten. Aus diesen Gründen wurde die Recherche nach geeigneten Datensätzen von den Studierenden als eine der größten und zeitaufwendigsten Herausforderungen im Kontext der Datennachnutzung wahrgenommen.

2.3 Perspektive der Betreuenden

Da die Recherche nach geeigneten offenen Datensätzen für die Studierenden eine anspruchsvolle und zeitaufwändige Aufgabe darstellte, mussten die Betreuenden entsprechend reagieren, insbesondere um den vorgegebenen zeitlichen Rahmen für die studentischen Arbeiten einzuhalten. Eine pragmatische Lösung bestand darin, den Studierenden nicht veröffentlichte Daten aus ihren laufenden oder abgeschlossenen Forschungsprojekten für die Anwendungsfälle bereitzustellen. In diesem Fall waren die Daten bereits verfügbar, sodass weder zusätzlicher Aufwand für die Suche noch für die Interpretation erforderlich war. Ein weiterer Vorteil war, dass die Betreuenden in der Regel selbst in die Datenerhebung involviert waren und somit als Datenautor:innen bei Unklarheiten oder fehlender Information in der Dokumentation unmittelbar zur Verfügung standen. Dies führte zwar zu einer Zeitersparnis für die Studierenden, brachte jedoch zugleich neue Herausforderungen für die Betreuenden mit sich.

Ein zentraler Aspekt bestand in der Prüfung der rechtlichen Rahmenbedingungen bezüglich der Datennutzung. Insbesondere musste bei den personenbezogenen Daten sichergestellt werden, dass die Einwilligungen der Proband:innen klare und rechtlich gültige Regelungen dazu enthalten. In der Praxis zeigte sich jedoch, dass viele Einwilligungserklärungen unpräzise oder mehrdeutig formuliert sind. Beispielsweise wurde in einigen Fällen lediglich die Nachnutzung für „weitere Forschungszwecke“ erlaubt, ohne eine klare Definition dieses Begriffs. Dies war problematisch, da unklar blieb, ob Forschungszwecke mit Lehr- und Lernzwecken gleichzusetzen sind oder ob eine Masterarbeit als Forschungsarbeit gewertet werden kann. Die Klärung solcher Fragen erforderte zusätzlichen Aufwand, der nicht nur von den Betreuenden, sondern auch durch den Rechtsdienst der ZHAW getragen wurde, der in diesen Fällen konsultiert werden musste.

Zusätzlich stellte die technische Komplexität vieler Datensätze eine Herausforderung dar. Oft erforderten die bereitgestellten Daten spezifische Softwarekenntnisse für ihre Bearbeitung und Analyse, über die die Studierenden nicht immer verfügen. Da die Aneignung dieser Kenntnisse innerhalb des begrenzten Zeitrahmens einer studentischen Arbeit nicht realistisch war, bestand eine weitere Aufgabe der Betreuenden darin, die Daten in einfachere und zugänglichere Formate – beispielsweise als Excel-Tabellendokumente – zu überführen, um deren Nachnutzung zu erleichtern. Dieser Transformationsprozess erwies sich in einigen Fällen als besonders anspruchsvoll und erforderte zusätzlichen Ressourcenaufwand vonseiten der Betreuenden.

Auch wenn die rechtliche Nachnutzung von Daten letztendlich geklärt und die Daten in geeigneter Form den Studierenden bereitgestellt wurden, ergab sich ein weiteres Problem: die Bewertung der studentischen Leistung im Kontext der Datennachnutzung. Die Nachnutzung von nicht publizierten Daten aus Forschungsprojekten der Betreuenden führte zwar zu einer Reduktion des Zeit- und Arbeitsaufwands aufseiten der Studierenden im Vergleich zu einer eigenständigen Datenrecherche, da die Daten bereits vorlagen und deren Qualität sichergestellt war. Dennoch blieb unklar, welche Anforderungen die Betreuenden

in diesem Zusammenhang an die Studierenden stellen sollten – insbesondere im Hinblick auf die ECTS-Leistungspunkte. Für eine studentische Arbeit ist eine bestimmte Anzahl an ECTS vorgesehen, die einen entsprechenden Arbeitsaufwand widerspiegelt. In Fällen der Datennachnutzung entfiel jedoch teilweise ein erheblicher Teil dieses Aufwands, was zu einer potenziellen Ungleichbehandlung gegenüber Studierenden führte, die die benötigten Daten eigenständig recherchieren oder erheben mussten.

Nicht zuletzt stellte sich für die betreuenden Lehrpersonen die Frage, ab welchem Qualifikationsniveau – etwa bereits zu Beginn des Bachelorstudiums, in dessen fortgeschrittenem Verlauf oder erst auf Master- bzw. Promotionsniveau – die Datennachnutzung sinnvoll erfolgen kann. Ebenso war zu klären, unter welchem Maß an Komplexität die dafür vorgesehenen Daten sowie der jeweilige Kontext der Arbeiten ausgestaltet sein sollten, um eine effektive Wiederverwendung zu ermöglichen. Denn insbesondere in den frühen Phasen des Bachelor-Studiums mangelt es Studierenden sowohl an grundlegenden Datenkompetenzen als auch an Kenntnissen wissenschaftlicher Arbeitsmethoden.

Um den unterschiedlichen Kompetenzniveaus der Studierenden angemessen zu begegnen, wurden seitens der betreuenden Lehrpersonen unterschiedliche didaktische Ansätze verfolgt. Während die Nachnutzung von Daten bei fortgeschrittenen Studierenden primär im Rahmen von Abschlussarbeiten stattfand – wobei sowohl die Komplexität der Analyseverfahren als auch die Tiefe der Forschungsfragen insbesondere auf Master-Niveau gesteigert wurden – erfolgte sie in den frühen Phasen des Bachelorstudiums in einem vereinfachten Kontext, etwa im Rahmen einer Aufgabe mit dem Ziel, einen wissenschaftlich fundierten Blogbeitrag zu verfassen.

Solche komplexreduziertere Aufgabenformate zur Datennachnutzung ermöglichten zudem Gruppenarbeiten, in denen Studierende ihre unterschiedlichen Vorkenntnisse einbringen und etwaige Defizite im Bereich der Softwareanwendung oder methodischen Kompetenz wechselseitig ausgleichen konnten. Auch die eigenständige Recherche geeigneter Datensätze gestaltete sich im Rahmen solcher kooperativer Arbeitsformen häufig effizienter. Allerdings stellt sich in solchen Fällen die Frage, ob die von den Studierenden selbst recherchierten Datensätze einer zusätzlichen Überprüfung durch die betreuenden Lehrpersonen hinsichtlich ihrer Eignung für den jeweiligen Untersuchungskontext bedürfen. Dies erscheint insbesondere deshalb relevant, weil häufig unklar ist, inwieweit Studierende – insbesondere in den frühen Studienphasen – in der Lage sind, Datenquellen kritisch zu hinterfragen und die Qualität der Daten angemessen einzuschätzen.

Die genannten Herausforderungen unterstreichen die Notwendigkeit, standardisierte Leitlinien für die Nachnutzung von Daten im akademischen Lehr- und Lernkontext zu entwickeln und zu implementieren. Ohne derartige Vorgaben fehlen klar definierte Prozesse, die eine zielgruppenspezifische Ansprache sowie eine verlässliche Einschätzung des notwendigen Ressourcen- und Arbeitsaufwands ermöglichen.

3 Begriffsdefinitionen aus der Fallstudie: Nachnutzbarkeit, Datenqualität, Data Literacy

Die im vorherigen Abschnitt dargelegten Reflexionen von Studierenden und Betreuenden führen zu der Schlussfolgerung, dass die Nachnutzbarkeit von Daten im Wesentlichen von zwei Faktoren abhängt: der Datenqualität und der Datenkompetenz (siehe Abbildung 1). Diese beiden Faktoren sind eng miteinander verknüpft – ohne das Eine bleibt das Andere wirkungslos. Dies wurde unter anderem am Beispiel des „F“ in den FAIR-Prinzipien deutlich vor Augen geführt: Selbst wenn Daten durch Befolgen der guten FDM-Praxis bei der Dokumentation und Veröffentlichung technisch auffindbar sind, bleiben sie oft praktisch nicht auffindbar, wenn das Wissen über geeignete Suchstrategien und Quellen fehlt. Sind die Daten unauffindbar, können sie nicht nachgenutzt werden – auch wenn sie das Potenzial zur Wiederverwendung haben. Die Nachnutzbarkeit oder das „R“ in den FAIR-Prinzipien kann also nur gegeben sein, wenn die vorhergehenden drei Komponenten F-A-I allesamt erfüllt sind.



Abbildung 1: Was bedeutet die Nachnutzbarkeit von Daten?

Die Analyse der Anwendungsfälle im Projekt ROADS legt nahe, dass Datenqualität aus der Perspektive der Datennachnutzung ebenfalls als eine Zusammensetzung mehrerer Komponenten zu verstehen ist, nämlich aus inhaltlichen, technischen und rechtlichen Aspekten (Abbildung 2).⁹ Inhaltliche Aspekte beziehen sich auf die gewährleistete Interpretierbarkeit der Daten, wobei der vorhandenen Dokumentation eine zentrale Bedeutung zukommt. Diese muss sämtliche relevanten Kontextinformationen enthalten, die erforderlich sind, um die Daten verstehen und interpretieren zu können. Technische Aspekte umfassen die Bereitstellung der Daten in Formaten, die idealerweise das Öffnen und Bearbeiten mit Open-Source-Software ermöglichen. Rechtliche Aspekte beziehen sich auf die explizite Festlegung der Bedingungen, unter denen eine Nachnutzung der Daten gestattet ist.

Basierend auf den Erkenntnissen aus dem Projekt RODS ist in Bezug auf Data Literacy im Kontext der Datennachnutzung im akademischen Lehr- und Lernumfeld zwischen

⁹ Detaillierte Ausführung zur Qualitätsprüfung der Daten bei der Datennachnutzung in der Checkliste: Mindestkriterien zur Nachnutzung von Forschungsdaten: <https://doi.org/10.5281/zenodo.11385458>
Diese Checkliste wurde im Rahmen des ROADS-Projekts erarbeitet.

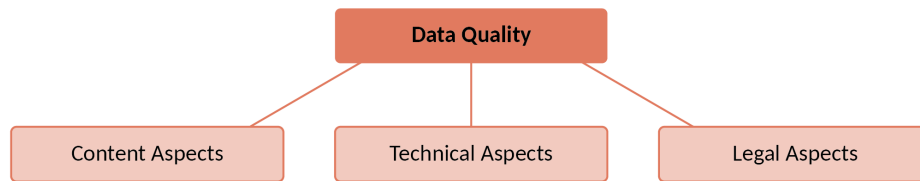


Abbildung 2: Was bedeutet die Datenqualität im Kontext der Nachnutzung?

zwei Arten von Kompetenzen zu unterscheiden: Einerseits die Kompetenzen, die Studierende für die Wiederverwendung von Daten benötigen, und andererseits die Kenntnisse, die Lehrende zur Vermittlung und Förderung dieses Prozesses erfordern (Abbildung 3). Dabei sind die Kompetenzen in beiden Fällen eng mit den Merkmalen der Datenqualität verknüpft.

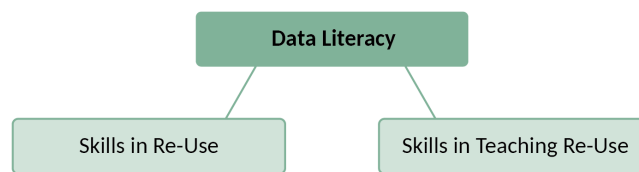


Abbildung 3: Was bedeutet Data Literacy im Kontext der Datennachnutzung im akademischen Lehr- und Lernumfeld?

Für Studierende umfasst Data Literacy unter anderem Kenntnisse über relevante Datenquellen sowie die Fähigkeit, diese gezielt zu recherchieren. Ebenso entscheidend ist der kritische Umgang mit Daten, insbesondere die Bewertung ihrer Zuverlässigkeit und der Angemessenheit ihrer Dokumentation. Zudem sollten Studierende mit einschlägiger Software vertraut sein und über rechtliche Kompetenzen im Umgang mit Daten verfügen, etwa hinsichtlich der Nachnutzungsbedingungen.

Für Lehrende liegt der Fokus auf methodisch-didaktischen Kompetenzen. Dazu gehört das Verständnis, welche Arten von Daten – in Abhängigkeit von Zielgruppe und Lehrkontext – für eine Wiederverwendung im akademischen Lehr- und Lernumfeld eignen und wie diese mit vertretbarem Aufwand in die Lehre integriert werden können. Ebenso relevant ist die Frage, wie die Datenkompetenz im Hinblick auf Datennachnutzung vermittelt werden sollte – ob als Bestandteil eines Curriculums oder als zusätzliches Lehrangebot.

Die beiden zentralen Komponenten – Datenqualität und Datenkompetenz – sind essenziell, um die Nachnutzbarkeit von Daten nachhaltig zu gewährleisten. Erst wenn beide Aspekte berücksichtigt werden, kann eine effektive Nachnutzung sichergestellt werden.

4 Ausblickend: Datennachnutzung als ein Schritt zur ressourcenschonenden Forschung

Nach der Darstellung der vielfältigen Herausforderungen im Zusammenhang mit der Datennachnutzung stellt sich die zentrale Frage, ob die Integration dieser Praxis in die akademische Lehre tatsächlich einen Mehrwert bietet oder ob die bestehenden Hürden die potenziellen Vorteile überwiegen. Der vorliegende Beitrag formuliert die Hypothese, dass Datennachnutzung im akademischen Lehrkontext weniger als Herausforderung, sondern vielmehr als Investition in eine nachhaltige, ressourcenschonende Forschung sowie in die langfristige Sicherung der Nachnutzbarkeit von Daten zu verstehen ist.¹⁰

Durch die gezielte Auseinandersetzung mit konkreten Problemstellungen werden Studierende frühzeitig für ein effizientes Forschungsdatenmanagement (FDM) sowie für die Relevanz der FAIR-Prinzipien sensibilisiert. Praktische Erfahrungen mit Herausforderungen wie unvollständigen Einwilligungserklärungen, die eine rechtssichere Nachnutzung verhindern, oder mit lückenhaften Dokumentationen, die die Interpretation der Daten erschweren, vermitteln Studierenden ein tiefgreifendes Verständnis dafür, welche Faktoren die Nachnutzbarkeit von Forschungsdaten wesentlich beeinflussen. Die FAIR-Prinzipien werden so nicht nur theoretisch erlernt, sondern im Sinne einer forschungspraktischen Kompetenzentwicklung verinnerlicht. Auf diese Weise entwickeln die Studierenden frühzeitig ein Bewusstsein für die Bedeutung nachhaltiger Datenpraktiken und wenden diese Prinzipien bei der Erhebung und Aufbereitung eigener Forschungsdaten – als künftige Generation von Open Scientists – selbstverständlich an.

Den Grundstein auf dem Weg zu einer nachhaltigen Forschung durch Datennachnutzung können unter anderem folgende Schritte bilden:

- **Schulungen zur Datenrecherche:** Spezialist:innen für Forschungsdaten sowie wissenschaftliche Bibliothekar:innen können gezielte Schulungen zur Datenrecherche anbieten – analog zu den bereits etablierten Kursen zur Literaturrecherche.
- **Praktische Anwendungsmöglichkeiten:** Lehrende und Forschende können den Studierenden vermehrt die Gelegenheit geben, bestehende Forschungsdaten wiederzuverwenden. Dies fördert nicht nur die Kompetenzentwicklung der Studierenden, sondern unterstützt auch eine Reflexion über nachhaltige Datenpraktiken innerhalb der wissenschaftlichen Gemeinschaft.
- **Entwicklung von Workflows und Richtlinien:** Es ist essenziell, klare Prozesse und Richtlinien für die Nachnutzung von Forschungsdaten im akademischen Lehr- und Lernkontext zu schaffen und zu etablieren.

¹⁰ Auf Bedeutung der curricularen Integration von Open-Science-Praktiken wird unter anderem auch in diesem Beitrag eingegangen: <https://doi.org/10.7554/eLife.89736>.

Die erfolgreiche Umsetzung dieser Schritte erfordert eine enge Zusammenarbeit zwischen Forschung, Lehre und forschungsunterstützenden Einrichtungen, wie beispielsweise Hochschulbibliotheken (Kohrs u. a. 2023).

Danksagungen

Ein besonderer Dank gilt swissuniversities für die Mitfinanzierung des Projekts. Ein ausdrücklicher Dank geht zudem an das ROADS-Projektteam: <https://www.zhaw.ch/de/forschung/projekt/74318>¹¹.

Interessenkonflikte

Die Autorin ist Koordinatorin des Projekts ROADS, dessen Ergebnisse die Grundlage des vorliegenden Beitrags bilden.

Literaturverzeichnis

- Bendel, Alexander. 2021. „Von der Forschungsfrage zum Ergebnis: Eine Rekonstruktion der sekundäranalytischen Auswertung qualitativer Daten am Beispiel einer Masterarbeit“. Kap. Von der Forschungsfrage zum Ergebnis: Eine Rekonstruktion der sekundäranalytischen Auswertung qualitativer Daten am Beispiel einer Masterarbeit in *Qualitative Sekundäranalysen: Daten der Sozialforschung aufbereiten und nachnutzen*, herausgegeben von Caroline Richter und Katharina Mojescik, 37–54. Springer Fachmedien Wiesbaden. ISBN: 9783658328511. <https://doi.org/10.1007/978-3-658-32851-1>.
- Gebel, Tobias und Alexia Meyermann. 2021. „Sekundäranalysen qualitativer Interviews: Eine Metaanalyse zur Praxis sekundäranalytischer Forschung zu Arbeitsorganisationen“. In *Qualitative Sekundäranalysen: Daten der Sozialforschung aufbereiten und nachnutzen*, herausgegeben von Caroline Richter und Katharina Mojescik, 17–36. Springer Fachmedien Wiesbaden. ISBN: 9783658328511. <https://doi.org/10.1007/978-3-658-32851-1>.

¹¹ Besucht am 1. Juli 2025 durch die Herausgeber*innen.

- Hollstein, Betina und Jörg Strübing. 2018. „Archivierung und Zugang zu Qualitativen Daten“. In *Archivierung und Zugang zu Qualitativen Daten. Working Paper Series des Rates für Sozial- und Wirtschaftsdaten (RatSWD)*, herausgegeben von Regina T. Riphahn, 1–13. German Data Forum (RatSWD).
- Kohrs, Friederike E., Susann Auer, Alexandra Bannach-Brown, Susann Fiedler, Tamarinde Laura Haven, Verena Heise, Constance Holman u. a. 2023. „Eleven strategies for making reproducible research and open science training the norm at research institutions“. *eLife* 12. <https://doi.org/10.7554/elife.89736>.
- Netscher, Sebastian und Jessica Trixa. 2019. „Forschungsdatenmanagement in der Sekundäranalyse“. In *Forschungsdatenmanagement sozialwissenschaftlicher Umfragedaten*, herausgegeben von Uwe Jensen, Sebastian Netscher und Kathrin Weller, 135–150. Verlag Barbara Budrich. <https://doi.org/10.3224/84742233.09>.