

FDM-LERNKARTEN Algorithmus CC-BY	FDM-LERNKARTEN Automatisierungsprozesse CC-BY
FDM-LERNKARTEN Strukturierte Daten CC-BY	FDM-LERNKARTEN Analoge Forschungsmaterialien CC-BY
FDM-LERNKARTEN Anonymisierung CC-BY	FDM-LERNKARTEN API CC-BY
FDM-LERNKARTEN Archiv CC-BY	FDM-LERNKARTEN Archivierungsdauer CC-BY
FDM-LERNKARTEN Authentifizierung CC-BY	FDM-LERNKARTEN Sicherheitsmerkmale CC-BY

In der Regel ist die Basis für Automatisierungsprozesse die Nutzung von Algorithmen.	Ein Algorithmus ist eine nachvollziehbare Schrittfolge zur Lösung eines Problems.
Analoge Forschungsmaterialien sind physische Objekte, wie zum Beispiel Proben, Fotos auf Papier, handschriftliche Notizen, Bücher, Audio-Kassetten, Gemälde, Fossilien, Bohrkerne oder Architekturmodelle.	Strukturierte Daten sind nach einer Struktur aufgebaute Daten, z.B. Normdaten und Metadaten.
Das Akronym API steht für "Application Programming Interface" (Programmierschnittstelle) und ist ein Teil eines Programms, das nach außen für andere Programme zugreifbar ist. Diese Anbindung wird im Kontext von Forschungsdaten z.B. für die automatische Abfrage von Datensätzen über das Web genutzt.	Die Anonymisierung von personenbezogenen Daten in der Wissenschaft gehört zur guten wissenschaftlichen Praxis und bedeutet, dass diese so verändert sind, dass „die Einzelangaben über persönliche oder sachliche Verhältnisse nicht mehr oder nur mit einem unverhältnismäßig großen Aufwand an Zeit, Kosten und Arbeitskraft einer bestimmten oder bestimmbaren natürlichen Person zugeordnet werden können“.
Bei Forschungsdaten hat sich eine Archivierungsdauer von zehn Jahren etabliert.	Unter einem Archiv ist allgemein eine Sammlung an Dokumenten zu verstehen, die zeitlich unbegrenzt aufbewahrt werden.
Zur Authentifizierung werden gelegentlich Sicherheitsmerkmale genutzt, also Schlüsseldateien, biometrische Merkmale oder Hardware-Token.	Authentifizierung bezeichnet den Zugang zu einem System mit vorher festgelegten Markmalen, z.B. Nutzernamen und Passwort, Freischaltungscode, Schlüsseldatei oder eine Kombination dieser Merkmale. Voraussetzung ist ein funktionierendes Identitätsmanagement (IdM) zur Speicherung und Verwaltung der genannten Merkmale.

FDM-LERNKARTEN Single-Sign-on-Verfahren CC-BY	FDM-LERNKARTEN Backup CC-BY
FDM-LERNKARTEN Best Practice CC-BY	FDM-LERNKARTEN Bitstream Preservation CC-BY
FDM-LERNKARTEN Born digital CC-BY	FDM-LERNKARTEN CARE-Prinzipien CC-BY
FDM-LERNKARTEN Collective Benefit (Care) CC-BY	FDM-LERNKARTEN Authority to Control (Care) CC-BY
FDM-LERNKARTEN Responsibility (Care) CC-BY	FDM-LERNKARTEN Ethics (Care) CC-BY

Backup oder Datensicherung bezeichnet die bewusste, redundante Speicherung von Daten und Dateien. Damit entsteht eine Sicherungskopie, mit der ein Speicherstand wiederhergestellt wird. Empfohlen ist die 3-2-1-Regel: Dreifache Speicherung, zwei verschiedene Speicherorte und einmal an einem anderen physischen Ort (z.B. über eine Cloudspeicherung)	Single-Sign-on-Verfahren (wörtlich: Einmalanmeldung) bedeutet, dass Nutzende sich nur einmal bei einem Identity Provider authentifiziert und dann Zugriff auf einen oder mehrere Dienste mit einer Anmeldung erhalten kann. Shibboleth ist ein bekanntes Beispiel für dieses Verfahren.
Digitale Daten bestehen aus einer festgelegten Abfolge von Bits, die den Wert 0 oder 1 haben und auf Datenträgern gespeichert werden. Diese Abfolge wird Bitstream genannt. Bitstream Preservation ist die Fähigkeit, den Bitstream über Technologiewechsel hinaus zu erhalten. Dies ist eine Grundvoraussetzung für digitale Langzeitarchivierung.	Mit Best Practice werden bewährte Methoden bezeichnet. Ein Beispiel hierfür wäre im Forschungsdatenkontext das Vorgeben von Metadatenschemata durch Repositorienrichtlinien.
Die CARE-Prinzipien wurden als Ergänzung zu den FAIR-Prinzipien von der Research Data Alliance International Indigenous Data Sovereignty Interest Group erarbeitet und durch die Global Indigenous Data Alliance (GIDA) veröffentlicht. Auf Grundlage der CARE-Prinzipien werden Forschende dafür sensibilisiert, im Rahmen der Bemühungen um Open Data und Open Science die Rechte und Interessen indigener Gemeinschaften zu wahren.	Als Born digital werden Daten und Materialien bezeichnet, die in digitaler Form entstanden sind. Auf analoge Materialien, die erst später digitalisiert werden, trifft dieser Begriff nicht zu.
Authority to Control (Care) ist ein Teil des Akronyms der Care-Prinzipien, in dem auf Rechte und Interessen an Indigenous Data hingewiesen wird, den betroffenen Gemeinschaften wird Governance über die Daten eingeräumt.	Collective Benefit (Care) ist ein Teil des Akronyms der Care-Prinzipien und beschreibt, dass die betroffenen Gemeinschaften einen Nutzen aus dem Datenökosystem ziehen können.
Ethics (Care) ist ein Teil des Akronyms der Care-Prinzipien und besagt, dass die Rechte und das Wohlergehen von indigenen Völkern in allen Stufen des Datenlebenszyklus im Mittelpunkt stehen.	Responsibility (Care) ist ein Teil des Akronyms der Care-Prinzipien. Damit ist gemeint, dass sich Forschende ihrer Verantwortung gegenüber von Gemeinschaften, die von ihrer Datenerhebung betroffen sind, bewusst sein sollen.

FDM-LERNKARTEN Datensouveränität CC-BY	FDM-LERNKARTEN Cloud CC-BY
FDM-LERNKARTEN Creative-Commons-Lizenzen CC-BY	FDM-LERNKARTEN CC0 CC-BY
FDM-LERNKARTEN BY (CC) CC-BY	FDM-LERNKARTEN SA (CC) CC-BY
FDM-LERNKARTEN NC (CC) CC-BY	FDM-LERNKARTEN ND (CC) CC-BY
FDM-LERNKARTEN Data Curation Profile CC-BY	FDM-LERNKARTEN Data Literacy CC-BY

In einer Cloud können abseits lokaler Rechner und Netzwerke Daten gespeichert sowie geteilt werden, sodass auch ein kollaboratives Arbeiten an Daten ermöglicht wird. Da Zugriff von überall möglich ist, sind diese Speicherlösungen ein wichtiger Bestandteil für die wissenschaftliche Arbeit mit Daten – sowohl für einzelne Forschende als auch für Forschungsgruppen.	Datensouveränität (auch: Datenhoheit) bezeichnet die Verfügungsgewalt über die Nutzung von Daten, einschließlich ihrer Verbreitung. Besondere Bedeutung hat dies beispielsweise im Kontext der Forschung an indigenen Bevölkerungsgruppen, deren traditionellen Wissen und ganz allgemein im Bezug auf personenbezogene Daten.
CC0 steht für "no Copyright" und ist ein Zusatz der Creative-Commons-Lizenzen. Diese Form erlaubt alle Bearbeitungen und Weiterverwendungen und entspricht in etwa dem US-amerikanischen Public Domain.	Um eine maximale Nachnutzbarkeit von Forschungsdaten zu gewährleisten, empfiehlt sich die explizite Vergabe von Nutzungsrechten, z. B. durch Lizenzen. Die Verwendung liberaler Lizenzmodelle, insbesondere der weltweit anerkannten Creative-Commons-Lizenzen (CC).
Das Kürzel SA (CC) der Creative-Commons-Lizenzen steht für Share Alike. So gekennzeichnete Inhalte dürfen nur unter gleichen Bedingungen weitergegeben werden.	Der Zusatz BY (CC) der Creative-Commons-Lizenzen steht für das englische Wort by und bedeutet, dass eine Namensnennung erfolgen muss.
Das Kürzel ND (CC) der Creative-Commons-Lizenzen steht für No Derivatives, es darf also keine Bearbeitung erfolgen.	Das Kürzel NC (CC) der Creative-Commons-Lizenzen steht für Non-Commercial und verbietet die kommerzielle Nutzung.
Data Literacy (dt. Datenkompetenz) beschreibt die Fähigkeit, entlang des gesamten Datenlebenszyklus planvoll mit Daten umzugehen. Dies umfasst das Erheben und Auswählen von Daten sowie deren Verarbeitung, Analyse, Darstellung und nachhaltige Nutzung. Das schließt auch eine kritische Reflexion des Umgangs mit Daten (z.B. gesellschaftliche Fragen) ein.	Ein Data Curation Profile beschreibt die ‚Geschichte‘ eines Datensatzes oder einer Datensammlung, d. h. Herkunft und Lebenszyklus eines Datensatzes innerhalb eines Forschungsprojekts. Das von den Purdue University Libraries entwickelte Profil sowie das zugehörige Toolkit ist sowohl ein Werkzeug als auch eine eigene Sammlung von Datensätzen.

FDM-LERNKARTEN Datenkompetenz CC-BY	FDM-LERNKARTEN Data Mining CC-BY
FDM-LERNKARTEN Data Steward CC-BY	FDM-LERNKARTEN Dateiformat CC-BY
FDM-LERNKARTEN Proprietäres Format CC-BY	FDM-LERNKARTEN Offenes Format CC-BY
FDM-LERNKARTEN Datenbankschutzrecht CC-BY	FDM-LERNKARTEN Datenjournal CC-BY
FDM-LERNKARTEN Datenkompetenzzentrum CC-BY	FDM-LERNKARTEN Datenkuration CC-BY

Unter diesem Begriff versteht man die Anwendung automatisierter, statistischer Analysemethoden auf große Datenbestände. Die Auswertung wird computergestützt meist mit Hilfe von komplexen Algorithmen in Datenbanken vorgenommen. Im Rahmen von Data Mining kommen häufig auch künstliche Intelligenz (KI) und zunehmend Visualisierungstechniken zum Einsatz. Das Ziel ist, bislang unbekannte Muster im Datenmaterial zu entdecken und automatisiert Hypothesen zu generieren.	Data Literacy (dt. Datenkompetenz) beschreibt die Fähigkeit, entlang des gesamten Datenlebenszyklus planvoll mit Daten umzugehen. Dies umfasst das Erheben und Auswählen von Daten sowie deren Verarbeitung, Analyse, Darstellung und nachhaltige Nutzung. Das schließt auch eine kritische Reflexion des Umgangs mit Daten (z.B. gesellschaftliche Fragen) ein.
Ein Dateiformat (manchmal auch Dateart, Dateityp oder Datenformat genannt) werden beim Speichern von Dateien erzeugt, sie beinhalten Informationen über die Struktur der in einer Datei vorhandenen Daten, sowie über deren Zweck und Zugehörigkeit. Mithilfe enthaltenen Informationen können Anwendungsprogramme Daten interpretieren und Inhalte verfügbar machen. Diese Formate können in der Regel an einer entsprechenden Endung erkannt werden, die an den eigentlichen Dateinamen angehängt wird.	Ein Data Steward ist Expert*in im Forschungsdatenmanagement. Diese Expert*innen werden schwerpunktmäßig an Forschungseinrichtungen eingesetzt, um Forschende beim nachhaltigen Umgang mit Daten zu unterstützen. Zu ihren Aufgaben gehören vor allem Beratungstätigkeiten, Schulungen, Bedarfsermittlungen und Anforderungsmanagement.
Ein Offenes Format (wie zum Beispiel .html, .csv, .jpg, .mp3, .gif) ermöglicht es im Gegensatz zu proprietären Formaten Dateien mit Software verschiedener Hersteller zu öffnen und zu bearbeiten. In wissenschaftlichen Bereich sollte in Sachen Dateiformat vor allem auf Kompatibilität, Eignung zur Langzeitarchivierung und verlustfreie Konvertierung in alternative Formate geachtet werden. Beim Veröffentlichen und Archivieren von Daten sollten aus den genannten Gründen, wo möglich, die offenen Dateiformate bevorzugt werden.	Ein Proprietäres Format erlaubt das Öffnen von Dateien nur mit den dazugehörigen Anwendungs-, Hilfs- oder Systemprogrammen öffnen, bearbeiten und speichern (zum Beispiel .doc/.docx, .xls/.xlsx).
Ein Datenjournal ist ein Publikationsmedium für Datensätze. Diese Medien bieten umfassende Beschreibungen von Datensätzen, ohne diese zu interpretieren, und unterstützen damit die Nachnutzung von Forschungsdaten sowie deren Anerkennung als wissenschaftliche Leistung. Gleichzeitig verfolgen sie das Ziel, die Transparenz von wissenschaftlichen Methoden und Ergebnissen zu verbessern und unterstützen damit gute Praktiken des Datenmanagements.	Das deutsche Datenbankschutzrecht (§§ 87a–87e Urheberrechtsgesetz, UrhG) schützt eine Datenbank für die Dauer von 15 Jahren vor ungenehmigter Nutzung und Vervielfältigung, sofern für ihre Erstellung eine „wesentliche Investition“ von Geld, Zeit, Arbeitskraft usw. notwendig war. Verboten ist dabei nicht die Vervielfältigung der gesamten Datenbank, sondern nur „wesentlicher Teile“, wobei im Einzelfall unklar sein kann, was damit gemeint ist.
Datenkuration umfasst alle erforderlichen Handlungen, um Forschungsdaten auszuwählen, (nachhaltig) zu pflegen, nachnutzbar zu machen und langfristig zu erhalten. Der Kurationsprozess beginnt mit der gezielten Auswahl von Daten, die aufgrund ihrer wissenschaftlichen Relevanz, Qualität oder rechtlichen Anforderungen aufzubewahren sind. Damit ist sie eine zentrale Bedingung für das Auffinden und Abrufen von Daten als auch für deren Qualitätserhalt, Mehrwert und perspektivische Nachnutzbarkeit.	Ein Datenkompetenzzentrum sich mit mehrheitlich regionalem Fokus der Vermittlung von Datenkompetenz für Forschende unterschiedlicher Fachrichtungen. Die Zentren schaffen didaktische Angebote und stellen diese bereit, bauen unterstützende Services auf und organisieren Veranstaltungen sowie Plattformen zum Austausch und zur Vernetzung. All dies verfolgt das Ziel, den Kulturwandel hinsichtlich der Nutzung von Forschungsdaten mitzugestalten.

FDM-LERNKARTEN Datenlebenszyklus CC-BY	FDM-LERNKARTEN Planen (Datenlebenszyklus) CC-BY
FDM-LERNKARTEN Erheben (Datenlebenszyklus) CC-BY	FDM-LERNKARTEN Aufbereiten und Analysieren (Datenlebenszyklus) CC-BY
FDM-LERNKARTEN Teilen und Publizieren (Datenlebenszyklus) CC-BY	FDM-LERNKARTEN Archivieren (Datenlebenszyklus) CC-BY
FDM-LERNKARTEN Nachnutzen (Datenlebenszyklus) CC-BY	FDM-LERNKARTEN Datenmanagementplan CC-BY
FDM-LERNKARTEN DMP CC-BY	FDM-LERNKARTEN Datenmapping CC-BY

Der erste Schritt im Datenlebenszyklus ist Planen (Datenlebenszyklus). Darunter fallen die Planung des Umgangs mit Daten, die Lokalisation bereits vorhandener Daten und die Vorbereitung des späteren Teilens von Daten unter dem Aspekt der Zustimmung zur Erhebung.	Das Modell des Datenlebenszyklus umfasst alle Phasen, die Forschungsdaten von der Erhebung bis zur Nachnutzung durchlaufen können.
Im Schritt Aufbereiten und Analysieren (Datenlebenszyklus) des Datenlebenszyklus werden Daten nachbearbeitet, also beispielsweise eingegeben, digitalisiert, gesichert, verwaltet, beschrieben oder interpretiert. Als letzter Schritt muss zudem die Datenhaltung vorbereitet werden.	Nach der Planung folgt im Datenlebenszyklus das Erheben (Datenlebenszyklus) von Daten. Es werden in diesem Schritt also aktiv neue Daten generiert, wobei ggf. eine Nutzungserlaubnis eingeholt werden muss.
Archivieren (Datenlebenszyklus) im Datenlebenszyklus ermöglicht das langfristige Bereitstellen von Forschungsdaten. Hierfür muss eine langzeitarchivierungsfähige Speicherform mit passenden deskriptiven Metadaten gewählt werden.	Das Teilen und Publizieren (Datenlebenszyklus) von Forschungsdaten im Datenlebenszyklus umfasst vier konsekutive Schritte: Zunächst das Festlegen des Urheberrechts, dann Zugangskontrolle, Teilen der Daten und zuletzt Bekanntmachung.
Ein Datenmanagementplan (DMP) beschreibt den Umgang mit Forschungsdaten, die im Rahmen eines Forschungsprojekts produziert oder (nach-)genutzt werden.	Das Nachnutzen (Datenlebenszyklus) von Daten ist im Datenlebenszyklus explizit vorgesehen. In diesem Schritt werden bereits erhobene und publizierte Daten wiederverwendet, z.B. in einem Folgeprojekt oder für die Lehre.
Datenmapping ist der Prozess der Überführung von Daten(elementen) von einem Datenmodell zu einem anderen. Dies ist der erste Schritt zur Integration von fremden Informationen in ein eigenes Informationssystem.	Das Akronym DMP steht für Datenmanagementplan.

FDM-LERNKARTEN Datennutzungsgesetz CC-BY	FDM-LERNKARTEN Datennutzungsvereinbarung CC-BY
FDM-LERNKARTEN Forschungsdatenpolicy CC-BY	FDM-LERNKARTEN Datenrichtlinie CC-BY
FDM-LERNKARTEN Datenschutz CC-BY	FDM-LERNKARTEN Datensparsamkeit CC-BY
FDM-LERNKARTEN Datensicherung CC-BY	FDM-LERNKARTEN Datenzentrum CC-BY
FDM-LERNKARTEN Digitalisat CC-BY	FDM-LERNKARTEN DINI-Zertifikat CC-BY

Der Begriff Datennutzungsvereinbarung ist mehrdeutig und kann u.a. bezeichnen: Eine Vereinbarung zur Weitergabe und Nachnutzung von Forschungsdaten, die nicht unter offener Lizenz stehen, Eine Vereinbarung (Vertrag) zwischen den Forschenden eines Forschungsvorhabens, in der möglichst frühzeitig (bestenfalls noch vor Beginn der Datenerhebung) festgelegt wird, wer zu welchem Zeitpunkt welche Befugnisse im Umgang mit den Daten haben wird, selten auch: datenschutzrechtliche Einwilligungserklärung für Webseiten.	Das Datennutzungsgesetz ist ein Gesetz in Deutschland, das vorschreibt, in welcher Weise Daten in öffentlicher Hand (auch Forschungsdaten aus öffentlich finanzierter Forschung) veröffentlicht werden müssen.
Datenrichtlinie ist der synonyme Begriff zur Forschungsdatenpolicy.	Eine Forschungsdatenpolicy ist eine an einer Einrichtung allgemeingültige institutionelle Vorschrift (an Hochschulen häufig in Form einer Richtlinie), die sich an Forschende richtet und den Umgang mit Forschungsdaten während und nach einem Forschungsvorhaben mehr oder weniger verbindlich regelt.
Mit Datensparsamkeit ist ein Konzept aus dem Datenschutz. Gemeint ist, dass für eine spätere Verarbeitung nur so viele personenbezogene Daten erhoben werden wie für den Erhebungszweck unbedingt nötig sind.	Unter Datenschutz werden technische und organisatorische Maßnahmen gegen den Missbrauch von personenbezogenen Daten verstanden, die z. B. in der Medizin oder der sozialwissenschaftlichen Forschung erhoben werden. Geregelt ist das Thema u.a. in der Datenschutz-Grundverordnung der EU (DSGVO).
Ein Datenzentrum (auch: Rechenzentrum, engl. Data Center) ist ein Ort zur Bündelung von IT-Infrastrukturen. Der Begriff bezeichnet sowohl Räumlichkeiten, an denen sich Server und Netzwerktechnik befinden, als auch Stellen, an denen Daten verwaltet, gespeichert und ausgewertet werden.	Die bewusste, redundante Speicherung von Daten wird als Datensicherung oder Backup bezeichnet. Ziel ist die Wiederherstellbarkeit des gesicherten Bearbeitungsstandes im Falle eines Datenverlustes.
Das DINI-Zertifikat (Deutsche Initiative für Netzwerkinformationen) ist ein weithin anerkanntes Gütesiegel für Repositorien.	Ein Digitalisat ist das Endprodukt einer Digitalisierung, d.h. der Überführung eines analogen Gegenstands (Buch, Handschrift, Bild, Plastik etc.) in digitale Werte zur elektronischen Speicherung. Vorteile der Digitalisierung sind u.a. die leichtere Zugänglichkeit und maschinelle Verarbeitbarkeit.

FDM-LERNKARTEN Digital Object Identifier CC-BY	FDM-LERNKARTEN DOI CC-BY
FDM-LERNKARTEN Elektronische Laborbücher CC-BY	FDM-LERNKARTEN ELNs CC-BY
FDM-LERNKARTEN Embargo CC-BY	FDM-LERNKARTEN Enhanced Publication CC-BY
FDM-LERNKARTEN Fachinformationsdienste CC-BY	FDM-LERNKARTEN FID CC-BY
FDM-LERNKARTEN FAIR-Prinzipien CC-BY	FDM-LERNKARTEN Findable (Fair) CC-BY

Das Akronym DOI steht für Digital Object Identifier.	Digital Object Identifier (DOI) sind eines der gebräuchlichsten Systeme zur persistenten Identifikation digitaler Objekte. Konkret ist ein DOI eine Online-Referenz, die einer digitalen Ressource (z. B. einem Artikel in einer Zeitschrift oder einem Forschungsdatensatz) zugewiesen wird, um ihr eine eindeutige und dauerhafte Referenz im Internet zu geben.
Das Akronym ELNs steht für Electronic Laboratory Notebook (Elektronisches Laborbuch).	Elektronische Laborbücher (engl. Electronic Laboratory Notebooks, kurz: ELNs) spielen im Bereich der Dokumentation der Planung, Durchführung und Auswertung wissenschaftlicher Experimente eine zentrale Rolle. Sie treten häufig an die Stelle konventioneller analoger Laborbücher, gehen in ihren Funktionalitäten allerdings darüber hinaus.
Als Enhanced Publication wird in der Wissenschaft eine elektronische Publikation verstanden, die zusätzlich zu einer Textpublikation auch damit verbundene Materialien wie beispielsweise Forschungsdaten, Videos, Tabellen oder Präsentationen enthält.	Ein (zeitliches) Embargo definiert im Forschungsdatenmanagement einen Zeitraum, in dem lediglich die deskriptiven Metadaten von Forschungsdaten einsehbar sind. Die damit verbundene Datenpublikation ist währenddessen nicht zugänglich.
Das Akronym FID steht für Forschungsinformationsdienste. Viele dieser Dienste bieten disziplinspezifische Services zum Forschungsdatenmanagement an.	Das DFG-Förderprogramm Fachinformationsdienste (FID) für die Wissenschaft verfolgt das Ziel, forschungsbezogene Informationsinfrastrukturen zu stärken und zu verbessern. Es ist das Nachfolgeprogramm der seit 1949 geförderten Sondersammelgebiete an wissenschaftlichen Bibliotheken in Deutschland.
Findable (Fair) (=Auffindbarkeit) bezeichnet im Kontext der FAIR-Prinzipien, dass Daten zur Nachnutzung gefunden werden können. Metadaten und Daten sollten sowohl maschinenlesbar als auch für Menschen nutzbar sein.	Das Hauptziel der FAIR-Prinzipien liegt in einer optimalen Aufbereitung von Forschungsdaten, die gemäß den Prinzipien auffindbar (findable), zugänglich (accessible), interoperabel (interoperable) und wiederverwendbar (reusable) sein sollen. Zentral ist dabei das Beschreiben von Forschungsdaten mit Metadaten sowie hinreichende dokumentarische Schritte.

FDM-LERNKARTEN Accessible (Fair) CC-BY	FDM-LERNKARTEN Interoperable (Fair) CC-BY
FDM-LERNKARTEN Reusable (Fair) CC-BY	FDM-LERNKARTEN Forschungsdaten CC-BY
FDM-LERNKARTEN Forschungsdatenmanagement CC-BY	FDM-LERNKARTEN FDM CC-BY
FDM-LERNKARTEN Gute wissenschaftliche Praxis CC-BY	FDM-LERNKARTEN Forschungsddatenzentrum CC-BY
FDM-LERNKARTEN GND CC-BY	FDM-LERNKARTEN Harvesting CC-BY

Interoperable (Fair) (=Interoperabilität) bezeichnet generell das (möglichst) nahtlose Zusammenspiel unterschiedlicher Systeme, um z. B. Daten effizient austauschen zu können, sie maschinell zu interpretieren und automatisiert mit anderen Datensätzen zu kombinieren. Im Bereich des Forschungsdatenmanagements spielen neben kompatiblen Dateiformaten interoperable Metadaten eine zentrale Rolle.	Accessible (Fair) (=Zugänglichkeit) ist ein Teil der FAIR-Prinzipien. Für Forschungsdaten soll in diesem Zusammenhang deutlich gemacht werden, wie (und ob) man auf diese zugreifen kann. Als Mindeststandard sollten die Metadaten stets zugänglich bleiben, auch wenn die Daten nicht verfügbar sind.
Forschungsdaten sind digitale oder analoge Daten, die während wissenschaftlicher Tätigkeit (z.B. durch Messungen, Befragungen, Quellenarbeit) entstehen, entwickelt oder ausgewertet werden bzw. auf diesen basieren. Sie bilden eine Grundlage wissenschaftlicher Arbeit und dokumentieren deren Ergebnisse.	Reusable (Fair) (=Wiederverwendbarkeit) ist eines der FAIR-Prinzipien. Um die Nachnutzung von Daten zu ermöglichen, müssen diese nach diesem Prinzip mit ausreichenden, deskriptiven Metadaten publiziert werden.
Mit FDM wird der Begriff Forschungsdatenmanagement abgekürzt. Forschungsdatenmanagement umfasst nicht nur das Veröffentlichen von Daten (Open Data), sondern auch Maßnahmen entlang der vorangehenden Schritte des gesamten Datenlebenszyklus sowie der Datenarchivierung und -nachnutzung.	Der Begriff Forschungsdatenmanagement bezeichnet strukturierte Maßnahmen im Kontext der Arbeit mit Forschungsdaten, die u. a. darauf abzielen, Daten unabhängig von den an der Erhebung beteiligten Personen langfristig nutz- bzw. nachnutzbar zu machen und somit die Effizienz der Forschung zu steigern. Weitere Ziele sind die Umsetzung rechtlicher Vorgaben und guter ethischer Praktiken.
Ein Forschungsdatenzentrum ist ein Ort, an denen Daten archiviert und für eine breite wissenschaftliche Nutzung zugänglich gemacht werden.	Unter Gute wissenschaftliche Praxis versteht man Schritte und Maßnahmen im Kontext wissenschaftlicher Prozesse zur Sicherstellung angemessener wissenschaftlicher Arbeit.
Als Harvesting bezeichnet man das automatische ‚Einsammeln‘ von Daten bzw. von Metadaten aus Archiven und Repositorien über sogenannte Data-Provider (wie bspw. BASE, OAIster oder Scientific Commons).	Das Akronym GND steht für die Gemeinsame Normdatei, in der Normdaten zu Entitäten aus kulturellen und wissenschaftlichen Sammlungen beschrieben werden. Geführt wird sie durch die Deutsche Nationalbibliothek, allen deutschsprachigen Bibliotheksverbänden und weiteren Einrichtungen.

FDM-LERNKARTEN Hochleistungsrechnen CC-BY	FDM-LERNKARTEN HPC CC-BY
FDM-LERNKARTEN Informierte Einwilligung CC-BY	FDM-LERNKARTEN Informed Consent CC-BY
FDM-LERNKARTEN Ingest CC-BY	FDM-LERNKARTEN Institutionelles Repository CC-BY
FDM-LERNKARTEN JSON CC-BY	FDM-LERNKARTEN JavaScript Object Notation CC-BY
FDM-LERNKARTEN Kontrolliertes Vokabular CC-BY	FDM-LERNKARTEN Künstliche Intelligenz CC-BY

Das Akronym HPC steht für High-Performance Computing (Hochleistungsrechnen). Darauf hin ausgerichtete Rechencluster sind auf eine möglichst schnelle Anbindung und extrem kurze Antwortzeiten der jeweiligen Computereinheiten untereinander angewiesen.	Hochleistungsrechnen (englisch: High-Performance Computing) umfasst einen Bereich des computergestützten Rechnens zur Bearbeitung komplizierter oder umfangreicher Berechnungen bzw. Datenanalysen. Hier stellen sich in der Regel große Anforderungen an Rechenleistung und Speicherkapazität.
Informed Consent ist der englische Begriff für Informierte Einwilligung. Das Einholen einer informierten Einwilligung ist nach Art. 6 DSGVO nicht nur rechtlich verpflichtend, sondern auch forschungsethisch geboten, wenn in einem Forschungsprojekt personenbezogene Daten erhoben werden.	Informierte Einwilligung (auch Informed Consent) bezeichnet einen Prozess, bei dem Forschende Forschungsteilnehmer*innen (z. B. Interviewpartner*innen) über das Projekt, in dem die Forschungsteilnehmer*innen beteiligt sind, informieren. Dies schließt die im Projekt erhobenen Daten ebenso ein wie die Frage, was mit diesen Daten geschieht. Auf dieser Basis sollen die Teilnehmer*innen freiwillig und informiert entscheiden können, ob sie die Zusammenarbeit akzeptieren oder ablehnen.
Ein Institutionelles Repositorium ist ein Repositorium, das von einer Institution gepflegt und geführt wird. Typische Beispiele sind Repositorien einzelner Universitäten.	Als Ingest (Übernahme) bezeichnet man das Aufnehmen von Forschungsdaten in ein Archiv oder Repositorium.
Das Akronym JSON steht für JavaScript Object Notation, dieses Format ist ein weit verbreitetes Austauschformat für API-Anfragen u.ä.	JSON (JavaScript Object Notation) ist ein kompaktes, leicht lesbares und softwareunabhängiges Datenformat für den Datenaustausch zwischen Anwendungen. Es wird insbesondere bei Webanwendungen zur Übertragung von strukturierten Daten eingesetzt, um diese in andere Systeme bzw. Anwendungen zu integrieren.
Künstliche Intelligenz (KI oder englisch: AI) bezeichnet ein Teilgebiet der Informatik, in dem u. a. menschliche Entscheidungsstrukturen mittels computationeller Methoden nachgebildet werden.	Ein Kontrolliertes Vokabular ist ein normierter Sprachgebrauch, in dem ein Wort oder eine Phrase genau eine festgelegte Bedeutung hat. Solche Vokabulare werden zum Beispiel beim Vergeben von Schlagwörtern in Metadaten verwendet, um digitale Objekte zu beschreiben.

FDM-LERNKARTEN KI CC-BY	FDM-LERNKARTEN Langzeitarchivierung CC-BY
FDM-LERNKARTEN Lizenz CC-BY	FDM-LERNKARTEN Maschinell verwertbare Daten CC-BY
FDM-LERNKARTEN Metadaten CC-BY	FDM-LERNKARTEN Metadatenschema CC-BY
FDM-LERNKARTEN Metadatenstandard CC-BY	FDM-LERNKARTEN NFDI CC-BY
FDM-LERNKARTEN Nationale Forschungsdateninfrastruktur CC-BY	FDM-LERNKARTEN Normdaten CC-BY

Langzeitarchivierung verfolgt das Ziel, über lange Zeiträume Zugang zu archivierten Daten zu bieten. Im Bereich Forschungsdatenmanagement hat sich eine Aufbewahrungsfrist von mindestens zehn Jahren als Standard für Forschungsdaten etabliert.	Das Akronym KI kürzt den Begriff Künstliche Intelligenz ab.
Maschinell verwertbare Daten können von Computersystemen automatisch gefunden und verwendet werden, ganz ohne oder mit nur minimaler menschlicher Unterstützung. Voraussetzung für eine maschinelle Nutzbarkeit ist eine einheitliche Datenstruktur. Maschinen oder Computer, die diese Daten lesen und nutzen sollen, werden auf der Grundlage dieser Struktur programmiert.	Eine Lizenz ist ein vertraglich vereinbartes Nutzungsrecht. Damit erlauben Rechteinhaber*innen (Lizenzgeber*innen) ihren Vertragspartner*innen (Lizenznehmer*innen), ein Werk auf verschiedene Arten zu nutzen (z. B. zu kopieren, zu speichern oder digital zugänglich zu machen).
Metadatenschema organisieren die Struktur von Metadaten. Sie legen fest, welche Elemente zur Beschreibung von Objekten wie z. B. Forschungsdaten verwendet werden und welche Informationen in welcher Form angegeben werden sollen. Damit einher geht eine erhöhte Interoperabilität und Maschinenlesbarkeit.	Metadaten beinhalten strukturierte Informationen über Daten (z. B. Forschungsdaten) oder andere Ressourcen und deren Merkmale. Sie werden entweder unabhängig von oder zusammen mit den Daten, die sie beschreiben, abgespeichert. Meist wird zwischen fachlichen (z. B. verwendete Messinstrumente) und technischen bzw. administrativen (z. B. DOI) Daten unterschieden.
NFDI steht als Abkürzung für Nationale Forschungsdateninfrastruktur, deren Ziel der strategische Ausbau des Forschungsdatenmanagements in Deutschland ist.	Ein Metadatenstandard dient einer inhaltlich sowie strukturell gleichförmigen Beschreibung ähnlicher Daten. Über sogenannte Mappings ist häufig die Überführung in andere Standards möglich.
Normdaten sind semantisch strukturierte, kontrollierte und gepflegte Datensätze zur eindeutigen Benennung und Verknüpfung von Inhalten (z. B. Personen, Schlagwörter, Körperschaften, Werke).	Die Nationale Forschungsdateninfrastruktur, abgekürzt NFDI, verfolgt das Ziel, Datenbestände von Wissenschaft und Forschung systematisch zu erschließen, diese zu vernetzen und nachhaltig sowie qualitativ nutzbar zu machen.

FDM-LERNKARTEN OAI-PMH CC-BY	FDM-LERNKARTEN Open Access CC-BY
FDM-LERNKARTEN Goldener Weg CC-BY	FDM-LERNKARTEN Grüner Weg CC-BY
FDM-LERNKARTEN Diamond Open Access CC-BY	FDM-LERNKARTEN Open Data CC-BY
FDM-LERNKARTEN ORCID CC-BY	FDM-LERNKARTEN Persistenter Identifikator CC-BY
FDM-LERNKARTEN PID CC-BY	FDM-LERNKARTEN Personenbezogene Daten CC-BY

Open Access bezeichnet den kostenlosen und möglichst barrierefreien Zugang zu digitalen wissenschaftlichen Inhalten. In der Regel erhalten Nutzende umfangreiche Nutzungsrechte und einfache Zugangswege. Je nach Definition sind mehrere Publikationsformen wie Texte, Forschungsdaten (Open Data) oder Bildungsmaterialien (Open Educational Resources) eingeschlossen.	OAI-PMH, das auf XML basierende Protocol for Metadata Harvesting (PMH) der Open Archives Initiative (OAI) unterstützt das automatisierte Harvesting von Metadaten über sogenannte Data und Service Provider, indem es die in Repositorien abgelegten Daten besser zugänglich macht.
Der Grüner Weg bezeichnet in der Textpublikation eine öffentliche Zweitveröffentlichung über z. B. Repositorien.	Der Goldener Weg ist ein Begriff aus der Publikation von Texten und bezeichnet die Veröffentlichung in Open-Access-Medien wie z.B. Open-Access-Zeitschriften gegen Publikationsgebühren.
Open Data (= "Offene Daten") bezeichnet im Internet offen zugängliche Daten, die von Dritten (nach-)genutzt werden dürfen. Nutzungsbezogene Einschränkungen existieren ausschließlich dahingehend, den Ursprung und die Offenheit der Daten zu wahren.	Diamond Open Access ist eine Form der Text Publikation, die Open Access in z. B. Open-Access-Zeitschriften ohne Publikationsgebühren erfolgt.
Ein Persistenter Identifikator bezeichnet dauerhafte digitale Identifikatoren, bestehend aus Ziffern und/oder alphanumerischen Zeichen, die Personen oder wissenschaftlichen Objekten (z. B. Textpublikationen, Datensätzen, Proben, Organisationen) zugeordnet werden und direkt auf diese verweisen.	Die Open Researcher and Contributor ID (ORCID) ist ein institutionsunabhängiger, international anerkannter persistenter Identifikator zur eindeutigen Identifizierung von Personen. Alle natürlichen Personen können für sich selbst ORCIDs anlegen und im Profil Informationen wie z. B. einen Lebenslauf oder eine Publikationsliste führen.
Die Datenschutz-Grundverordnung (DSGVO) der EU definiert Personenbezogene Daten als „alle Informationen, die sich auf eine identifizierte oder identifizierbare natürliche Person (im Folgenden „betroffene Person“) beziehen, als identifizierbar wird eine natürliche Person angesehen, die direkt oder indirekt, insbesondere mittels Zuordnung (...) identifiziert werden kann“ (Artikel 4 Nr. 1).	Die Abkürzung PID steht für Persistenter Identifikator. Im Gegensatz zu anderen seriellen Identifikatoren wie z. B. URL-Adressen verweist ein persistenter Identifikator auf das Objekt selbst und nicht auf dessen Standort im Internet.

FDM-LERNKARTEN Platform-as-a-service CC-BY	FDM-LERNKARTEN Primärdaten CC-BY
FDM-LERNKARTEN Prüfsumme CC-BY	FDM-LERNKARTEN Pseudonymisierung CC-BY
FDM-LERNKARTEN ReadMe-Datei CC-BY	FDM-LERNKARTEN Rechenzentrum CC-BY
FDM-LERNKARTEN Rechte an Daten CC-BY	FDM-LERNKARTEN Replikationsstudie CC-BY
FDM-LERNKARTEN Replizierbarkeit CC-BY	FDM-LERNKARTEN Repositoryum CC-BY

Primärdaten (auch bekannt als Rohdaten oder Originaldaten) bezeichnen direkt aus Datenerhebungen gewonnene beziehungsweise aus Untersuchungen oder Beobachtungen erhobene Daten (z. B. unbearbeitete, ungeprüfte oder unkommentierte Messdaten oder Audio- und Videoaufzeichnungen). Daraus abgeleitete bzw. prozessierte Daten werden als Sekundärdaten bezeichnet.	Platform-as-a-service (PaaS) bezeichnet die cloudbasierte Bereitstellung von Laufzeitumgebungen im Kontext der IT-Entwicklung, meist gegen ein Nutzungsentgelt. Die Vor- und Nachteile von PaaS sind vergleichbar mit denen von Software-as-a-Service (SaaS): Auslagerung der Bereitstellung und Wartung der benötigten IT-Infrastruktur an externe Dienstleister sowie Entstehung einer potentiellen Abhängigkeit von diesem Dienstleister, der Erfordernis einer permanenten und stabilen Internetverbindung und eventuelle Hürden im Umgang mit sensiblen Daten.
Bei der Pseudonymisierung personenbezogener Daten werden direkte oder indirekte Indikatoren (z. B. Name, Alter, Wohnort) durch Pseudonyme (z. B. Buchstaben- und/oder Zahlencode) ersetzt, um die Identifizierung der den Daten zugrundeliegenden Personen zu erschweren oder auszuschließen. Schlüsseldateien, in denen die originalen Werte den jeweils ersetzenden Werten zugeordnet werden, erlauben es, dies rückgängig zu machen. Anders als anonymisierte Daten unterliegen pseudonymisierte Daten nach wie vor Datenschutzbestimmungen.	Eine Prüfsumme ist ein aus einem digitalen Objekt (z. B. Datei oder Datensatz) berechneter (Zahlen-)Wert zur Kontrolle der Integrität des digitalen Objekts.
Rechenzentrum können aus zwei Blickwinkeln definiert werden. Aus Sicht der Forschenden sind dies Entscheidungsbefugnisse über die Daten, die sich aus der Erzeugung ergeben. Aus Sicht von Nutzerinnen und Nutzern handelt es sich um die Rechte, welche bei Nachnutzung von Daten zu beachten sind. Für die Nachnutzung von Daten gelten mindestens die Regeln der guten wissenschaftlichen Praxis.	Eine ReadMe-Datei enthält kompakte und strukturierte Informationen zu digitalen Medien wie Forschungsdaten und Software, um bei der Orientierung über diese Medien sowie bei deren Nutzung zu unterstützen. Zu diesem Zweck werden in dieser Art von Dateien z. B. zentrale Metadaten, Informationen zu verwendeten Benennungsstandards, Ordnerstrukturen, Abkürzungen und Normdaten und Änderungen sowie Versionierungen von Forschungsdaten aufgeführt.
Eine Replikationsstudie ist eine wissenschaftliche Untersuchung, die überprüft, ob die Ergebnisse einer vorangegangenen Untersuchung reproduzierbar ist.	Rechte an Daten können aus zwei Blickwinkeln definiert werden. Aus Sicht der Forschenden sind dies Entscheidungsbefugnisse über die Daten, die sich aus der Erzeugung ergeben. Aus Sicht von Nutzerinnen und Nutzern handelt es sich um die Rechte, welche bei Nachnutzung von Daten zu beachten sind. Für die Nachnutzung von Daten gelten mindestens die Regeln der guten wissenschaftlichen Praxis.
Der Begriff Repositorium bezeichnet einen Speicher- und Publikationsort für digitale Objekte. Generell wird zwischen Lösungen für Textpublikationen und für Forschungsdaten unterschieden – es existieren allerdings auch hybride Systeme, die beide Publikationstypen aufnehmen.	Der Begriff Replizierbarkeit bezeichnet in der Wissenschaft den Prozess, bei dem eine Studie mit neuen Daten wiederholt wird, wobei jedoch das ursprüngliche Forschungsdesign und die ursprüngliche Analyse verwendet werden, um ähnliche Ergebnisse zu erzielen.

FDM-LERNKARTEN Richtlinien CC-BY	FDM-LERNKARTEN Schöpfungshöhe CC-BY
FDM-LERNKARTEN Semantisches Web CC-BY	FDM-LERNKARTEN Semantic Web CC-BY
FDM-LERNKARTEN Software-as-a-Service CC-BY	FDM-LERNKARTEN Softwaremanagementplan CC-BY
FDM-LERNKARTEN SMP CC-BY	FDM-LERNKARTEN Urheberrecht CC-BY
FDM-LERNKARTEN URN CC-BY	FDM-LERNKARTEN Uniform Resource Name CC-BY

Bei der Schaffung eines Objektes bzw. Projektes kann die Persönlichkeit des Urhebers im Ergebnis zum Ausdruck kommen. Das Ausmaß davon definiert die Schöpfungshöhe (auch Gestaltungshöhe oder Werkhöhe), welche für die Schutzwürdigkeit nach deutschem Urheberrecht entscheidend ist. Dabei kommt es wesentlich darauf an, dass kein äußerer Zwang (Zielstellung, Funktionalität, Objektivität usw.), sondern der Urheber das Ergebnis individuell gestaltet. Deswegen sind Forschungsdaten fast nie vom Urheberrecht geschützt.	Richtlinien für Forschungsdatenmanagement sind eine häufige Form der allgemeingültigen institutionellen Forschungsdatenpolicy. Der genaue Inhalt, die Rechtsverbindlichkeit sowie die Länge einer Policy kann zwischen einzelnen Einrichtungen stark variieren.
Das Semantic Web (deutsch: Semantisches Web) bezeichnet Möglichkeiten, im Internet Daten um eine semantische Schicht zu erweitern. Um Maschinen den Kontext eines Begriffes zu vermitteln, wird auf maschinenlesbare Metadatenstandards zurückgegriffen.	Semantisches Web (englisch: Semantic Web) bezeichnet Aktivitäten mit dem Ziel, das World Wide Web um eine semantische Schicht zu erweitern. Diese ermöglicht den maschinellen Austausch von Informationen und beinhaltet die Kontextualisierung zentraler Begriffe einer Webseite um Metadaten, sodass bspw. klar wird, ob es sich bei ‚Berlin‘ um die Hauptstadt Deutschlands, eine andere Stadt oder einen Namen handelt. Um Maschinen den Kontext eines Begriffes zu vermitteln, wird auf maschinenlesbare Metadatenstandards zurückgegriffen.
Laut Definition der DINI/nestor AG Forschungsdaten beinhaltet ein Softwaremanagementplan (SMP) allgemeine und technische Informationen zum Softwareprojekt, Angaben zur Qualitätssicherung, zum Release und zur öffentlichen Verfügbarkeit sowie rechtliche und ethische Aspekte, die die Software betreffen. Zweck eines SMPs ist zunächst die Nachvollziehbarkeit sowie ggf. die langfristige Nutzbarkeit der Software.	Software-as-a-Service (SaaS) bezeichnet Software, die meist gegen ein Nutzungsentgelt durch einen IT-Dienstleister in einer Cloudinfrastruktur bereitgestellt wird und die der Kunde üblicherweise über einen Webbrowser nutzt.
In Deutschland gilt das deutsche Urheberrecht für die Nutzung von literarischen, künstlerischen und wissenschaftlichen Werken, die die dort genannten Schutzvoraussetzungen erfüllen. Ob Forschungsdaten diesem Schutz unterliegen oder nicht ist davon abhängig, ob die Anforderungen an die geistige Schöpfungshöhe oder die Voraussetzungen des Datenbankschutzrechts erfüllt werden. Um eine maximale Nachnutzbarkeit wissenschaftlicher Forschungsdaten zu gewährleisten sollte die Vergabe von zusätzlichen Nutzungsrechten in Betracht gezogen werden.	Das Akronym SMP steht für Softwaremanagementplan. Dieser fasst Informationen zusammen, die die Erstellung, Dokumentation, Speicherung, Versionierung, Lizenzierung, Archivierung und/oder Veröffentlichung der in einem Projekt erzeugten oder verwendeten Software hinreichend beschreiben und dokumentieren.
URN (Uniform Resource Name) ist die Bezeichnung eines Identifizierungs- und Adressierungssystems. Dieses wird ähnlich wie ein DOI zur persistenten Identifikation digitaler Objekte (z. B. Textpublikationen oder Datensätze) verwendet.	Das Akronym URN steht für Uniform Resource Name.

FDM-LERNKARTEN URI CC-BY	FDM-LERNKARTEN Uniform Resource Identifier CC-BY
FDM-LERNKARTEN Versionierung CC-BY	FDM-LERNKARTEN Virtuelle Forschungsumgebung CC-BY
FDM-LERNKARTEN VFU CC-BY	FDM-LERNKARTEN XML CC-BY
FDM-LERNKARTEN XSL CC-BY	FDM-LERNKARTEN Zertifizierung CC-BY
FDM-LERNKARTEN Validität CC-BY	FDM-LERNKARTEN Zitation CC-BY

Ein Uniform Resource Identifier (kurz: URI) dient der Identifikation von Ressourcen. Beispiele für URIs sind Adressen von Webseiten oder E-Mailadressen.	Das Akronym URI steht für Uniform Resource Identifier.
Virtuelle Forschungsumgebung (VFU) sind Softwarelösungen bzw. Plattformen zur ortsunabhängigen Zusammenarbeit zwischen Wissenschaftler*innen. Hierbei handelt es sich primär um anwendungsorientierte Dienste institutioneller Infrastruktureinrichtungen für bestimmte Forschungsverbünde oder -communities. VFUs vereinen vor allem fachspezifische Tools, Toolsammlungen und Arbeitsumgebungen.	Da sich Daten verändern, wenn man mit ihnen arbeitet, empfiehlt sich die Kennzeichnung einzelner Arbeitsstände und damit verbundener Veränderungen mithilfe von Versionierung. Hierfür sollte ein vorab mit weiteren Beteiligten abgesprochenes, festgelegtes sowie einfach verständliches Schema (z. B. Version 1.3 oder Version 2.1.4) genutzt werden.
XML ist eine Auszeichnungssprache zur Speicherung hierarchisch strukturierter Informationen als einfache Textdatei. Diese Sprache wird hauptsächlich im plattformunabhängigen Datenaustausch zwischen Anwendungen bzw. Computern benutzt. Die Kodierung ist sowohl maschinen- als auch menschenlesbar - dadurch kann Gestalt und Inhalt kodierter Informationen sehr genau beschrieben werden.	Das Akronym VFU steht für Virtuelle Forschungsumgebung.
Zertifizierung bezeichnet im Forschungsdatenmanagement in der Regel das Kennzeichnen von bestimmte Standards erfüllenden Repositorien. So ist von außen ersichtlich, dass ein Repository vertrauenswürdig und qualitativ hochwertig ist – die Abwesenheit eines Zertifikats bedeutet allerdings nicht automatisch, dass ein Repository qualitativ minderwertig ist.	XSL (XML Stylesheet Language) ermöglicht es, in XML gespeicherte Informationen zu interpretieren, zu verändern, zu portieren und in andere Dateiformate umzuwandeln.
Je nach Fachbereich und Forschungsdisziplin bestehen für das Referenzieren wissenschaftlicher Datenpublikationen unterschiedliche Zitationsgewohnheiten bzw. -vorgaben. Für die Zitation von Forschungsdaten existiert derzeit (noch) kein einheitlicher Standard, allerdings hat die Data Citation Synthesis Group in ihrer Joint Declaration of Data Citation Principles fachübergreifende Empfehlungen formuliert. Entscheidend ist beispielsweise der Einbezug persistenter Identifikatoren in Zitationen.	Die Validität von Forschungsdaten kann beispielsweise über Prüfsummen nachgewiesen werden. Dieses Vorgehen ermöglicht es, nachzuvollziehen, dass Daten unverändert (absichtlich oder unabsichtlich, wie etwa durch technische Fehler) übertragen worden sind.