


SCIENTIFIC DATA

OPEN

KOMMENTAR

Die TRUST-Prinzipien für digitale Repositorien

Dawei Lin¹ , Jonathan Crabtree², Ingrid Dillo³, Robert R. Downs⁴, Rorie Edmunds⁵, David Giaretta⁶, Marisa De Giusti⁷, Hervé L'Hours⁸, Wim Hugo⁹, Reyna Jenkyns¹⁰, Varsha Khodiyar¹¹, Maryann E. Martone¹², Mustapha Mokrane¹³, Vivek Navale¹⁴, Jonathan Petters¹⁵, Barbara Sierman¹⁶, Dina V. Sokolova¹⁷, Martina Stockhause¹⁸ & John Westbrook¹⁸

Da Informations- und Kommunikationstechnologien in unserer Gesellschaft allgegenwärtig geworden sind, sind wir zunehmend sowohl von digitalen Daten abhängig als auch auf Repositorien angewiesen, die den Zugang zu diesen Ressourcen und ihre Nutzung ermöglichen. Repositorien müssen sich das Vertrauen der Communitys, denen sie dienen sollen, verdienen und unter Beweis stellen, dass sie zuverlässig und in der Lage sind, die in ihnen vorgehaltenen Daten adäquat zu verwalten.

Nach einer jahrelangen öffentlichen Debatte und auf der Basis eines bestehenden Konsenses¹ der Community haben mehrere Stakeholder, die verschiedene Bereiche der Gemeinschaft digitaler Repositorien repräsentieren, gemeinsam eine Reihe von Leitprinzipien für den Nachweis der Vertrauenswürdigkeit („Trustworthiness“) digitaler Repositorien aufgestellt und gebilligt. **T**ransparency (Transparenz), **R**esponsibility (Verantwortung), **U**ser focus (Nutzerfokussierung), **S**ustainability (Nachhaltigkeit) und **T**echnology (Technologie): die TRUST-Prinzipien bieten einen gemeinsamen Rahmen, der die Erörterung und Umsetzung von Best Practices in der digitalen Bestandserhaltung durch alle Stakeholder fördern soll.

Kontext und Hintergrund

Seit mehr als sechzig Jahren sind die Verwaltung und Bewahrung digitaler Daten ein zentraler Bestandteil des Auftrags wissenschaftlicher Einrichtungen wie Bibliotheken, Archive und bereichsspezifische Repositorien², an denen viele weitere Akteure beteiligt sind, darunter Forscher, Geldgeber, Infrastrukturanbieter und Dienstleister. Wissenschaftliches Datenmanagement erfährt innerhalb und außerhalb der wissenschaftlichen Gemeinschaft eine wachsende Aufmerksamkeit, insbesondere im Rahmen des aktuellen Open Science-Diskurses. Es bildet sich allmählich ein Konsens über ‚gute‘ Datenmanagementpraktiken heraus, wenngleich es in einigen wissenschaftlichen Bereichen nach wie vor an einer effizienten Umsetzung mangelt.

Die FAIR-Prinzipien³ verdeutlichen die Notwendigkeit bewährte Verfahren zu übernehmen, indem sie wesentliche Merkmale von Datenobjekten definieren um sicherzustellen, dass die Daten von Menschen und Maschinen wiederverwendet werden können: sie sollten **F**indable (auffindbar), **A**ccessible (zugänglich), **I**nteroperable (interoperabel) und **R**eusable (wiederverwendbar), d. h. FAIR sein. Um Daten FAIR zu machen und sie gleichzei-

¹Division of Allergy, Immunology and Transplantation, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Maryland, USA. ²HW Odum Institute for Research in Social Science, University of North Carolina at Chapel Hill, North Carolina, USA. ³Data Archiving and Networked Services (DANS), The Hague, Netherlands. ⁴Center for International Earth Science Information Network (CIESIN), The Earth Institute, Columbia University, New York, USA. ⁵World Data System of the International Science Council (WDS), WDS International Programme Office, Tokyo, Japan. ⁶PtAB Ltd, Dorset, UK. ⁷Universidad Nacional de La Plata, Comisión de Investigaciones Científicas de la Provincia de Buenos Aires, La Plata, Argentina. ⁸UK Data Archive, UK Data Service, University of Essex, Colchester, UK. ⁹South African Environmental Observation Network, Cape Town, South Africa. ¹⁰Ocean Networks Canada, University of Victoria, Victoria, Canada. ¹¹Springer Nature, London, UK. ¹²University of California, San Diego, California, USA and SciCrunch Inc., San Diego, USA. ¹³Center for Information Technology, National Institutes of Health, Maryland, USA. ¹⁴Data Services, University Libraries, Virginia Tech, Virginia, USA. ¹⁵KB National Library of the Netherlands, The Hague, The Netherlands. ¹⁶University Libraries, Columbia University, New York, USA. ¹⁷German Climate Computing Center (DKRZ), Hamburg, Germany. ¹⁸RCSB, Protein Data Bank, Rutgers, The State University of New Jersey, Institute for Quantitative Biomedicine at Rutgers, New Jersey, USA. ✉ E-Mail: dawei.lin@nih.gov

Box 1 Die TRUST-Prinzipien

Prinzip	Leitlinien für Repositorien
Transparency (Transparenz)	Transparenz hinsichtlich der speziellen Dienste und Datenbestände des Repositoriums, die anhand öffentlich zugänglicher Nachweise überprüfbar sind.
Responsibility (Verantwortung)	Verantwortung für die Sicherstellung der Authentizität und Integrität der Datenbestände und für die Zuverlässigkeit und das Fortbestehen der Dienste.
User Focus (Nutzerfokussierung)	Sicherstellen, dass die Datenverwaltungsnormen und die Erwartungen der anvisierten Nutzer-Communitys erfüllt werden.
Sustainability (Nachhaltigkeit)	Aufrechterhaltung der Dienste und langfristige Bewahrung der Datenbestände.
Technology (Technologie)	Bereitstellung der Infrastruktur und der Ressourcen um sichere, dauerhafte und zuverlässige Dienste aufrechtzuerhalten.

tig langfristig zu bewahren, bedarf es jedoch vertrauenswürdiger digitaler Repositorien („Trustworthy Digital Repositories“, TDRs) mit nachhaltigen Rahmenbedingungen für Verwaltung und Organisation, einer zuverlässigen Infrastruktur und umfassender Richtlinien zur Umsetzung gemeinschaftlich vereinbarter Verfahrensweisen. TDRs, mit ihrem klaren Auftrag, Daten sowohl bei technologischen Veränderungen als auch veränderten Anforderungen der Stakeholder aktiv zu bewahren, spielen eine wichtige Rolle bei der Erhaltung des Wertes der Daten. Sie genießen das Vertrauen ihrer Nutzer, da sie Verantwortung für die Datenadministration übernehmen. Zur Erfüllung dieser Aufgabe, müssen TDRs auf Dauer unerlässliche Ressourcen aufweisen, um den Communitys, denen sie dienen, Zugang und Nachnutzung der Daten langfristig zu ermöglichen. TDRs unterstützen die Pflege von Daten und die Bewahrung von Datenbeständen mit unterschiedlichen Graden der Wiederverwendbarkeit. In bestimmten Fällen können Daten von geringerer Qualität, die sich nicht mit angemessenen Mitteln verbessern oder interoperabler machen lassen, dennoch von großem Wert für ihre Nutzer-Communitys sein und daher eine vertrauenswürdige Verwaltung erfordern. Ein TDR muss die von der Community anerkannten Kriterien identifizieren und anstreben sie einzuhalten und das erreichte Niveau der Datenqualität bekannt machen.

Das *Open Archival Information System* (OAIS)-Referenzmodell⁴ liefert Empfehlungen zur Einrichtung von Archiven, die die langfristige Bewahrung von und Zugänglichkeit zu Informationen (insbesondere digitaler Informationen) gewährleisten und zur Erstellung von Archivierungspaketen. Es bietet einen einheitlichen und umfassenden Rahmen von Prinzipien und terminologischen Begriffen für die Verwaltung von Archivinformationssystemen. Dem OAIS-Referenzmodell zu genügen ist jedoch kein Garant für Vertrauenswürdigkeit. Um die Vertrauenswürdigkeit des Repositoriums zu bewerten, müssen vielmehr auch andere Aspekte einschließlich der erforderlichen Governance-Struktur, der Ressourcen und der Sicherheit angesprochen werden. Da es sich bei dem OAIS um ein Referenzmodell handelt, das keine detaillierten Richtlinien für eine Implementierung vorsieht, gibt es unterschiedliche Interpretations- und Implementierungsmöglichkeiten, die Prüfungs- und Zertifizierungsmechanismen erforderlich machen, wie sie in dem aus dem Jahr 1996 stammenden Bericht *Preserving Digital Information*⁵ verstanden werden. Die Autoren dieses Berichts empfahlen, dass „Repositorien, die für sich in Anspruch nehmen eine Archivierungsfunktion erfüllen zu wollen, in der Lage sein müssen, nachzuweisen, dass sie sind, was sie vorgeben zu sein, indem sie die Standards und Kriterien eines unabhängig verwalteten Programms für die Zertifizierung von Archiven erfüllen oder übertreffen“.

Vertrauenswürdigkeit lässt sich durch Beweise belegen, die von Transparenz abhängen, und daher müssen Repositorien transparente, ehrliche und überprüfbare Nachweise für ihre Vorgehensweise liefern. Auf diese Weise können Stakeholder sicher sein, dass Repositorien über einen langen Zeitraum hinweg die Integrität, Authentizität, Korrektheit, Zuverlässigkeit und Zugänglichkeit der Daten gewährleisten. Vertrauenswürdigkeit ist keine einmalige Errungenschaft; sie kann nicht als selbstverständlich vorausgesetzt werden ohne eine regelmäßige Überprüfung und Zertifizierung.

Zertifizierung leistet einen objektiven und wichtigen Beitrag zum Vertrauen der verschiedenen Stakeholder eines Repositoriums. Um die Qualität ihrer professionellen Vorgehensweise zu bewerten und zu verbessern, stützen sich Repositorien auf eine Reihe von internationalen Zertifizierungsstandards, die eine grundlegende, erweiterte oder formale Zertifizierung vorsehen. Solche Standards wie etwa die CoreTrustSeal⁶-Kriterien oder die Normen DIN 31644/NESTOR⁷ und ISO 16363⁸ legen den Fokus auf vier Hauptbewertungsbereiche: Organisation, Verwaltung digitaler Objekte, technische Infrastruktur und Management von Sicherheitsrisiken. Die Standards unterscheiden sich hinsichtlich der Anzahl und Komplexität ihrer Anforderungen, wobei die Intensität des Prüfverfahrens von einer Peer-Review, einer Selbstevaluierung bis hin zu einer umfangreicheren Vor-Ort-Prüfung durch ein externes Auditteam reichen kann. Die Auswahl des Zertifizierungsmechanismus richtet sich nach der Notwendigkeit, der Bereitschaft und der Fähigkeit eines Repositoriums, in seine weitere Professionalisierung und Vertrauenswürdigkeit zu investieren. Die Übernahme der Trustworthy Data Repositories Requirements von CoreTrustSeal durch zahlreiche Datenrepositorien, ist ein Beispiel für die Verbesserungen, die vorgenommen wurden, um sicherzustellen, dass ihr Leistungsvermögen den Merkmalen der TRUST-Prinzipien⁶ entspricht. Viele Datenrepositorien haben die CoreTrustSeal-Zertifizierung erhalten und sind Mitglieder des World Data System (WDS) des International Science Council. Der Erwerb einer Zertifizierung und die Durchführung von Evaluierungen verdeutlichen das Bestreben

zahlreicher digitaler Repositorien, als vertrauenswürdig angesehen zu werden.

Manager von Repositorien und ihre Teams sind die primäre Zielgruppe für das bestehende OAIS-Referenzmodell und die oben erörterten Mechanismen zur Zertifizierung der Vertrauenswürdigkeit. In einem Open-Science-Kontext erwarten wir jedoch, dass eine breitere Zielgruppe, einschließlich Geldgebern und Nutzern von Repositorien, von dem durch die TRUST-Prinzipien vorgegebenen Rahmen profitieren wird, insbesondere angesichts der zunehmenden Aufmerksamkeit, die wissenschaftliches Data Stewardship erfährt (Box 1).

Transparenz

Bei der Auswahl des für einen bestimmten Anwendungsfall am besten geeigneten Repositoriums ist es für alle potenziellen Nutzer von Vorteil, wenn sie leicht Informationen über den Umfang, die Zielnutzer-Community, die Richtlinien und die Möglichkeiten eines Datenrepositoriums finden und auf diese zugreifen können. Transparenz in diesen Bereichen bietet die Möglichkeit, mehr über das Repositorium zu erfahren und dessen Eignung für spezielle Bedürfnisse einschließlich Datenablage und -archivierung und Datenermittlung (Data Discovery) zu prüfen. Um diesem Grundsatz zu entsprechen, sollten Repositorien sicherstellen, dass zumindest Leitbild und Aufgabenbereich des Repositoriums eindeutig beschrieben werden. Darüber hinaus empfiehlt es sich, folgende Punkte transparent bekannt zu machen:

- Nutzungsbedingungen für Repositorium und Datenbestände;
- Mindestzeiträume für die digitale Archivierung der Datenbestände;
- alle relevanten zusätzlichen Funktionen oder Dienste, wie zum Beispiel die Möglichkeit, sensible Daten verantwortungsvoll zu verwalten.

Eine klare Kommunikation der Richtlinien des Repositoriums und insbesondere der Bestimmungen für die Nutzung der Datenbestände, informiert die Nutzer über alle Beschränkungen, die die Nutzung der Daten oder des Repositoriums begrenzen könnten. Ebenso würde die Möglichkeit leicht beurteilen zu können, ob ein Repositorium mit sensiblen Daten verantwortungsvoll umgehen kann, ihre Entscheidung beeinflussen, ob sie die verfügbaren Datendienste nutzen.

Verantwortung

Vertrauenswürdige (,TRUST'-)Repositorien übernehmen Verantwortung für die Verwaltung ihrer Datenbestände und die Betreuung ihrer Nutzergemeinschaft. Verantwortung zeigt sich dabei

- in der Beachtung der Metadaten- und Datenpflegestandards der jeweiligen Community bei gleichzeitiger Verwaltung der Datenbestände, z. B. bei der technischen Validierung, Dokumentation, Qualitätskontrolle, dem Authentizitätsschutz, der langfristigen Verfügbarkeit,
- in der Bereitstellung von Datendiensten, z. B. Portal- und Maschinenschnittstellen, Datendownload, serverseitige Verarbeitung,
- in der Verwaltung der geistigen Eigentumsrechte der Datenerzeuger, dem Schutz sensibler Informationsquellen und der Sicherheit des Systems und seiner Inhalte.

Die Nutzer eines Repositoriums sollten darauf vertrauen können, dass Datengeber alle Metadaten unter Beachtung der Normen der Community zur Verfügung stellen müssen, da dies die Auffindbarkeit und den Nutzen der Daten erheblich steigert. Zu wissen, dass ein Repositorium die Integrität der verfügbaren Daten und Metadaten überprüft, überzeugt potenzielle Nutzer, dass die Datenbestände eher mit anderen relevanten Datensätzen interoperabel sind. Sowohl Datengeber als auch Datennutzer benötigen die Gewissheit, dass die Daten langfristig verfügbar bleiben und somit in wissenschaftlichen Publikationen zitiert und referenziert werden können.

Verantwortung kann rechtlich geregelt sein (Recht auf Aufbewahrung) oder die Form einer freiwilligen Einhaltung bestimmter Normen (ethische Standards) annehmen.

Nutzerfokussierung

Ein vertrauenswürdige (,TRUST'-)Repositorium muss sich darauf konzentrieren seiner Zielnutzer-Community zu dienen. Jede Nutzer-Community hat wahrscheinlich andere Erwartungen an die Repositorien der Community, was zum Teil davon abhängt, wie ausgereift die Gemeinschaft in Bezug auf Verwaltung und Austausch von Daten ist. Ein vertrauenswürdige (,TRUST'-)Repositorium ist in die Datenpraktiken seiner Zielnutzer-Community eingebettet und kann daher auf die sich entwickelnden Anforderungen der Community reagieren. Wir verstehen den Begriff „Nutzer-Community“ breit gefasst, so kann diese neben Nutzern, die Daten hinterlegen oder auf Daten zugreifen, denen, die elektronisch auf Datenbestände zugreifen, auch indirekte Stakeholder wie Geldgeber, Herausgeber von Zeitschriften, sonstige institutionelle Partner oder Bürger einschließen.

Nutzung und Nachnutzung von Forschungsdaten stellen einen zentralen Bestandteil des wissenschaftlichen Prozesses dar, weshalb vertrauenswürdige (,TRUST'-)Repositorien ihre Community in die Lage versetzen sollten, ihre Datenbestände im Hinblick auf eine mögliche (Nach-)Nutzung aufzufinden, zu erkunden und zu verstehen. Repositorien sollten ihre Nutzer dazu anhalten, die Daten zum Zeitpunkt der Hinterlegung vollständig zu beschreiben, und ein Feedback zu etwaigen Problemen (z. B. Qualität, Verwendbarkeit) ermöglichen, die nach der Bereitstellung der Daten zutage treten.

Repositorien spielen bei der Anwendung und Durchsetzung der Normen und Standards der Zielnutzer-Community eine entscheidende Rolle, da Compliance die Interoperabilität und Wiederverwendbarkeit von Daten fördert. Zu den

Datenstandards, die vertrauenswürdige („TRUST“-)Repositorien umsetzen sollten, gehören Metadatenschemata, Dateiformate für Daten, kontrollierte Vokabulare, Ontologien und sonstige innerhalb der Community gebräuchliche Semantiken. Ein vertrauenswürdiges („TRUST“-)Repositorium kann die Einhaltung dieses Grundsatzes unter Beweis stellen, indem es

- relevante Datenmetriken implementiert und diese den Nutzern zur Verfügung stellt,
- Community-Kataloge bereitstellt (oder zu deren Erstellung beiträgt), um das Auffinden von Daten zu erleichtern,
- verändernde Erwartungen der Community beobachtet und identifiziert und erforderlichenfalls auf diese Veränderungen reagiert.

Nachhaltigkeit

Die Sicherstellung der Nachhaltigkeit eines vertrauenswürdigen („TRUST“-)Repositoriums ist notwendig, um gegenwärtigen und zukünftigen Nutzer-Communities einen ununterbrochenen Zugriff auf die wertvollen Datenbestände zu ermöglichen. Für einen dauerhaften Datenzugriff muss das Repositorium in der Lage sein, Dienste auf Dauer bereitzustellen und auf sich verändernde Anforderungen der Nutzer-Community mit neuen oder verbesserten Diensten zu reagieren.

Ein vertrauenswürdiges („TRUST“-)Repositorium kann die Nachhaltigkeit seiner Bestände unter Beweis stellen, indem es

- ausreichende Planungen zur Risikominderung, zur Aufrechterhaltung des Betriebs, zur Notfallwiederherstellung und zur Nachfolge macht,
- die Finanzierung sicherstellt, um eine fortlaufende Nutzung zu ermöglichen und die erwünschten Eigenschaften der Datenressourcen, mit deren Archivierung und Verbreitung das Repositorium beauftragt wurde, zu erhalten,
- den organisatorischen Rahmen für die erforderliche langfristige Archivierung der Daten schafft, damit diese auch in der Zukunft auffindbar, zugänglich und nutzbar bleiben.

Technologie

Ein Repositorium ist auf das Zusammenspiel von Menschen, Prozessen und Technologien angewiesen, um sichere, dauerhafte und zuverlässige Dienste zu ermöglichen. Seine Aktivitäten und Funktionen werden von Software, Hardware und technischen Diensten unterstützt. Zusammen stellen diese die Werkzeuge für die Umsetzung der TRUST-Prinzipien bereit.

Ein vertrauenswürdiges („TRUST“-)Repositorium kann die Zweckmäßigkeit seiner technischen Möglichkeiten demonstrieren, indem es

- die relevanten und geeigneten Standards, Werkzeuge und Technologien für die Verwaltung und Pflege der Daten implementiert,
- über Pläne und Mechanismen verfügt, um Bedrohungen der Cybersicherheit und der physischen Sicherheit zu verhindern, aufzudecken und darauf zu reagieren.

Einfluss der TRUST-Prinzipien

Die in einer abstrakten, nicht technischen Sprache formulierten TRUST-Prinzipien erleichtern die Kommunikation und beeinflussen so Stakeholder innerhalb und außerhalb der Datennutzer-Community gleichermaßen. Wenn Datenrepositorien, Geldgeber und Datenerzeuger die FAIR-Prinzipien übernehmen und die TRUST-Prinzipien umsetzen, profitieren die Nutzer von Repositorien direkt durch andauernde und verbesserte Möglichkeiten zur effizienten und effektiven Nutzung von Daten. Gemeinsam leisten die Stakeholder der TRUST-Prinzipien einen Beitrag zu einem kulturellen Wandel in der Forschung hin zu einem Daten- und Informationsökosystem, das sich im Informationszeitalter herausgebildet hat, aber schon seit Jahrhunderten ein wesentlicher Bestandteil des wissenschaftlichen Prozesses ist.

Verschiedene Studien haben herausgefunden, dass Transparenz mit dem Vertrauen in digitale Repositorien verbunden wird⁹. So ist beispielsweise für Nutzer von Videodaten „Transparenz der Praktiken eines Repositoriums, und insbesondere Transparenz der Datenpflegepraktiken, wichtig für Vertrauen“¹⁰. Donaldson et al.¹¹ untersuchten die Erkenntnisse der Mitarbeiter von Datenrepositorien bei der Zertifizierung des Repositoriums und stellten fest, dass der Prozess zum Erwerb der Zertifizierung neben anderen Vorteilen auch zur Transparenz des Repositoriums beitrug.

Das OAIS-Referenzmodell beschreibt die Aufgaben von Archivinformationssystemen, die mit der Pflege von Informationsressourcen betraut sind. Peng et al.¹² beschreiben die Herausforderungen eines effizienten Data Stewardship und legen dar, dass „die Definition von Rollen und Verantwortlichkeiten auf jeder Ebene des Stewardship und in jeder Phase des Lebenszyklus von Datenprodukten helfen, diese Herausforderung leichter zu bewältigen“. Eine Untersuchung von Forschungsdatenpraktiken über den gesamten Datenlebenszyklus hinweg brachte Kowalczyk¹³ zudem zu der Erkenntnis, dass „[d]ie Wahrscheinlichkeit eines langfristigen Datenmanagements im Falle von Forschungssammlungen gering ist, wenn die laufende Verantwortung bei einem einzelnen Forscher oder Doktoranden liegt“.

Yoon¹⁴ untersuchte, wie Nutzererfahrungen das Vertrauen in Datenrepositorien beeinflussen und fand heraus, dass „die bewusste Wahrnehmung der Nutzer von Rollen oder Funktionen eines Repositoriums ein Faktor für die Entwicklung des Vertrauens der Nutzer sein kann“. Nutzer vertrauen Repositorien oft aufgrund ihrer eigenen Erfahrungen, der Praktiken und des Rufes des jeweiligen Repositoriums und der Erfahrungen der übrigen

Mitglieder der Community^{9,14,15}. Das Vertrauen der Nutzer in Daten hängt auch mit ihrem Vertrauen in das Archiv, aus dem die Inhalte stammen, zusammen¹⁶.

Der Bericht einer Studie über die Nachhaltigkeit digitaler Repositorien, die von der Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (OECD) durchgeführt wurde, kam zu dem Schluss, dass „Forschungsdatenrepositorien ein wesentlicher Bestandteil der Infrastruktur für Open Science sind ...“ [und dass es] „wichtig ist, die Nachhaltigkeit von Forschungsdatenrepositorien sicherzustellen“¹⁷. Die Bedeutung der Nachhaltigkeit der Forschungsdateninfrastruktur wurde bereits in Studien festgestellt, die die Bedürfnisse von Archäologen beschreiben^{9,18}. Ohne effiziente Nachhaltigkeitsstrategien und Kontinuitätspläne könnten Datenrepositorien und ihre Bestände, wie viele frühere Biologiedatenbanken¹⁹, verschwinden. Ironischerweise merken York et al.²⁰ an, dass „wir trotz der Vielzahl an Datenrepositorien, Stewardship-Initiativen und Richtlinien über die Forschungsdatenlandschaft hinweg nur relativ wenig über die Gesamtmenge, die Merkmale oder die Nachhaltigkeit gepflegter Forschungsdaten wissen“.

Die Ausgestaltung technischer Möglichkeiten sollte im Hinblick auf die Organisations-, Betriebswirtschafts- und Verwaltungsmöglichkeiten, die eine fortlaufende Nutzung der in einem Repository vorgehaltenen Datenbestände ermöglichen, ergänzt werden^{10,21}. In ihrer Beschreibung der Notwendigkeit das Vertrauen der Öffentlichkeit in Gesundheitsdaten zu stärken, fordern Van Staa et al.²² Möglichkeiten, die „neuen Technologien mit einer klaren Rechenschaftspflicht, transparenten Abläufen und öffentlichem Vertrauen [zu] verbinden ...“, wobei sie feststellen, dass „es bei Data Stewardship nicht nur um physische und digitale Sicherheit geht, sondern auch um Mitarbeiterschulungen, Standardbetriebsverfahren und die Kompetenzen und Einstellungen der Mitarbeiter ebenfalls wichtig sind“²².

Schluss

Die TRUST-Prinzipien sollen eine Merkhilfe darstellen, um die Stakeholder von Datenrepositorien an die Notwendigkeit zu erinnern, die Infrastruktur zu entwickeln und zu pflegen, um eine kontinuierliche Pflege von Daten zu fördern und die Nutzung ihrer Datenbestände in der Zukunft zu ermöglichen. Die TRUST-Prinzipien stellen jedoch keinen Selbstzweck dar, sondern vielmehr ein Mittel die Kommunikation mit allen Stakeholdern zu erleichtern und den Repositorien Leitlinien an die Hand zu geben für Transparenz, Verantwortung, Nutzerfokussierung, Nachhaltigkeit und Technologie.

Erhalten: 6. März 2020; Angenommen: 22. April 2020;

Online veröffentlicht: 14. Mai 2020

Übersetzung

Diese Übersetzung entstand im Auftrag des Verbundprojektes EcoDM.

DOI: [10.5281/zenodo.6256222](https://doi.org/10.5281/zenodo.6256222)

ecoDM

GEFÖRDERT VOM



Förderkennzeichen 16DWWQP

Übersetzungsfirma

Dialecta
Grünberger Str. 26
10245 Berlin

Lektorat und Redaktion

Claus Spiecker
Christine Burkart
Jasper Bothe

Literaturverzeichnis

1. RDA/WDS Certification of Digital Repositories IG. The TRUST Principles for Trustworthy Data Repositories – An Update. *Research Data Alliance (RDA)*, <https://www.rd-alliance.org/trust-principles-trustworthy-data-repositories—update> (2019).
2. Mokrane, M. & Parsons, M. Learning from the International Polar Year to Build the Future of Polar Data Management. *Data Sci. J.* **13**, IFPDA–15 (2014).
3. Wilkinson, M. D. et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* **3**, 160018 (2016).
4. Consultative Committee for Space Data Systems. Reference Model for an Open Archival Information System (OAIS). Recommended Practice CCSDS 650.0-M-2. *Consultative Committee for Space Data Systems*, <https://public.ccsds.org/Pubs/650x0m2.pdf> (2012).
5. Waters, D. & Garrett, J. *Preserving Digital Information, Report of the Task Force on Archiving of Digital Information*. 1400 16th St., NW, Suite 740, Washington, DC 20036-2217. 59 pp, <https://www.clir.org/pubs/reports/pub63/> (1996).

6. CoreTrustSeal. CoreTrustSeal Certified Repositories. *CoreTrustSeal*, <https://www.coretrustseal.org/why-certification/certified-repositories/> (2020).
7. Harmsen, H. *et al.* Erläuterungen zum nestor-Siegel für vertrauenswürdige digitale Langzeitarchive. *Nestor-Arbeitsgruppe Zertifizierung*, <http://nbn-resolving.de/urn:nbn:de:0008-2013100901> (2013).
8. Audit and Certification of Trustworthy Digital Repositories. ISO 16363/CCSDS 652.0-M-1, <https://public.ccsds.org/Pubs/652x0m1.pdf> (2011).
9. Yakel, E., Faniel, I. M., Kriesberg, A. & Yoon, A. Trust in Digital Repositories. *Int. J. Digit. Curation* **8**, 143–156 (2013).
10. Frank, R. D., Chen, Z., Crawford, E., Suzuka, K. & Yakel, E. Trust in qualitative data repositories. In *Proceedings of the Association for Information Science and Technology* **54** 102 – 111 Association for Information Science and Technology (2017).
11. Donaldson, D. R., Dillo, I., Downs, R. & Ramdeen, S. The Perceived Value of Acquiring Data Seals of Approval. *Int. J. Digit. Curation* **12**, 130–151 (2017).
12. Peng, G. *et al.* A Conceptual Enterprise Framework for Managing Scientific Data Stewardship. *Data Sci. J.* **17**, 15 (2018).
13. Kowalczyk, S. T. Modelling the Research Data Lifecycle. *Int. J. Digit. Curation* **12**, 331 – 361 (2017).
14. Yoon, A. End users' trust in data repositories: Definition and influences on trust development. *Arch. Sci.* **14**, 17 – 34 (2014).
15. Downs, R. & Chen, R. Organizational needs for managing and preserving geospatial data and related electronic records. *Data Sci. J.* **4**, 255–271 (2006).
16. Donaldson, D. R. Trust in Archives – Trust in Digital Archival Content Framework. *Archivaria* **88**, 50 – 83 (2019).
17. OECD. *Business models for sustainable research data repositories*. **58**, <https://doi.org/10.1787/302b12bb-en> (2017).
18. Williams, J. P. & Williams, R. D. Information science and North American archaeology: Examining the potential for collaboration. *Inf. Res.* **24**, Papier 820. Quelle: <http://InformationR.net/ir/24-2/paper820.html> (Archiviert von WebCite® unter <http://www.Webcitation.Org/78mnvhrti>) (2019).
19. Attwood, T. K., Agit, B. & Ellis, L. B. M. Longevity of Biological Databases. *EMBnet. Journal* **21**, 803 (2015).
20. York, J., Gutmann, M. & Berman, F. What Do We Know about the Stewardship Gap. *Data Sci. J.* **17**, 19 (2018).
21. Corrado, E. M. Repositories, Trust, and the CoreTrustSeal. *Tech. Serv. Q.* **36**, 61 – 72 (2019).
22. Staa, T.-P., van, Goldacre, B., Buchan, I. & Smeeth, L. Big health data: the need to earn public trust. *BMJ* **354**, i3636 (2016).

Danksagung

Die Autoren bedanken sich für die Verbesserungsvorschläge der nicht als Autoren beteiligten Mitglieder des CoreTrustSeal Standards and Certification Board, der Teilnehmer der 13. Research Data Alliance-Vollversammlung „Build TRUST to be FAIR - Emerging Needs of Certification in Life Sciences, Geosciences and Humanities“, die von der Certification of Digital Repositories Interest Group der RDA und des WDS einberufen wurde, und der Teilnehmer des vom Office of Data Science Strategy der NIH gesponserten NIH-Workshops „Trustworthy Data Repositories for Biomedical Sciences“ (NIH-Workshop 2019), bei dem das TRUST-Rahmenwerk erstmalig zur Diskussion über vertrauenswürdige Datenrepositorien verwendet wurde. Wir sind dankbar für die anregenden Gespräche mit Shelley Stall, Robert S. Chen, Mark Conrad, Peter Doorn, Eliane Fankhauser, Elizabeth Hull, Siri Jodha Singh Khalsa, Micky Lindlar, Limor Peer, Philipp Conzett und Rachel Drysdale. Schließlich geht unser Dank an Anupama Gururaj für das Korrekturlesen des Artikels.

Interessenkonflikte

V. K. K. arbeitet für Springer Nature, den Herausgeber von *Scientific Data*. Bis Februar 2020 nahm V. K. K. redaktionelle Aufgaben bei *Scientific Data* wahr. Die Autoren erklären, dass V. K. K. nicht am Redaktions- und Begutachtungsprozess für dieses Manuskript mitgewirkt hat. Einige der Autoren sind an den im Manuskript erörterten Norm- und Zertifizierungsbemühungen beteiligt. Hierzu gehören D. L., J. C., I. D., R. R. D., R. E., H. L. H., W. H., R. J. und M.M., die dem Standards and Certification Board von CoreTrustSeal angehören, und D. G. als Mitglied des Primary Trustworthy Digital Repository Authorization Board (PTAB). Alle anderen Autoren erklären, dass sie keine Interessenkonflikte haben.

Weitere Informationen

Schriftverkehr und Materialanfragen sind an D. L. zu richten.

Informationen zu Nachdrucken und Genehmigungen finden sich unter www.nature.com/reprints.

Anmerkung des Herausgebers Springer Nature bleibt im Hinblick auf gerichtliche Behauptungen auf veröffentlichten Karten und institutionelle Zugehörigkeiten neutral.



Open Access Dieser Artikel wird unter den Bedingungen der Creative Commons Attribution 4.0 International License veröffentlicht, die die Nutzung, Weitergabe, Anpassung, Verbreitung und Vervielfältigung in jedem Medium oder Format gestattet, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons License einfügen und angeben, ob Änderungen vorgenommen wurden. Die Bilder oder anderes Drittmaterial in diesem Artikel sind in der Creative Commons License des Artikels enthalten, sofern in einer Quellenangabe zu dem Material nicht etwas anderes erwähnt wird. Wenn Materialien nicht in der Creative Commons License enthalten sind und der von Ihnen beabsichtigte Verwendungszweck gesetzlich nicht zulässig ist oder die zulässige Nutzung überschreitet, müssen Sie direkt die Genehmigung des Urheberrechtsinhabers einholen. Ein Exemplar dieser Lizenz kann unter <http://creativecommons.org/licenses/by/4.0/> aufgerufen werden.

Dieser Artikel ist ein Werk der US-Regierung und unterliegt in den USA nicht dem urheberrechtlichen Schutz. Ausländische urheberschutzrechtliche Bestimmungen können anwendbar sein. 2020