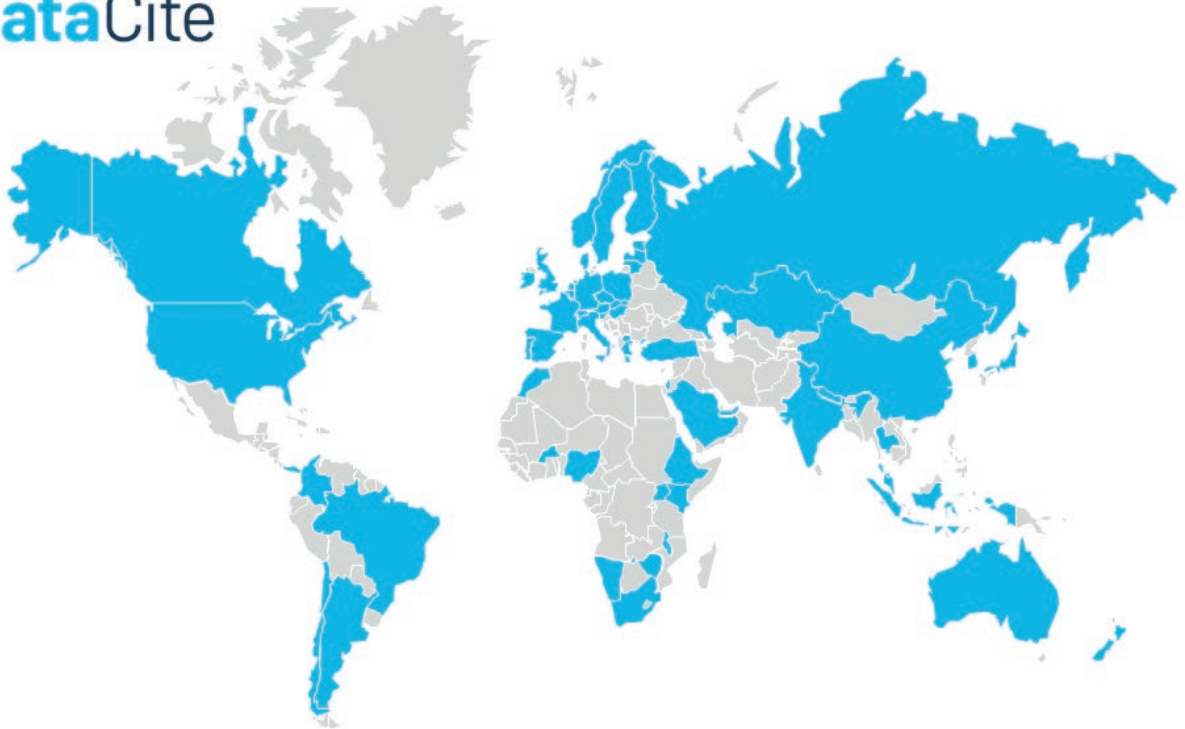


15 Jahre DataCite: ein Rückblick und ein Blick nach vorn

Britta Dreyer



DataCite hat sich zu einem globalen Netzwerk entwickelt – mit Mitgliedern auf der ganzen Welt. Grafik: TIB

Abstract

Seit seiner Gründung im Jahr 2009 hat sich DataCite von einer kleinen Initiative zu einer globalen Organisation entwickelt, die mit über 82,7 Millionen vergebenen DOIs (Digital Object Identifier) den Zugang zu Forschungsdaten nachhaltig verbessert. Als Schlüsselakteur im Bereich Open Science fördert DataCite die Auffindbarkeit, Zitierbarkeit und Nachvollziehbarkeit wissenschaftlicher Daten und arbeitet eng mit seinen Mitgliedern wie der TIB zusammen. Durch Initiativen wie die Collaborative Metadata Enrichment Taskforce (COMET) und Make Data Count (MDC), strategische Partnerschaften und die Weiterentwicklung von Metadatenstandards treibt DataCite die Umsetzung der FAIR-Prinzipien und die globale Vernetzung offener Forschungsinfrastrukturen voran.

Since its founding in 2009, DataCite has evolved from a small initiative into a global organization that has sustainably improved access to research data with over 82.7 million assigned DOIs (Digital Object Identifier). As a key player in the field of Open Science, DataCite promotes the discoverability, citability, and traceability of scientific data and works closely with its members, such as the TIB. Through initiatives like Collaborative Metadata Enrichment Taskforce (COMET) and Make Data Count (MDC), strategic partnerships, and the advancement of metadata standards, DataCite drives the implementation of the FAIR principles and linked open research infrastructures.

Am 1. Dezember 2009 wurde der Verein DataCite gegründet. Ziel war es, den Zugang zu und die Verfügbarkeit von Forschungsdaten auf globaler Ebene zu verbessern. Heute, 15 Jahre später, ist DataCite eine international etablierte Organisation, deren Mission es ist, wissenschaftliche Daten nachvollziehbarer, sichtbarer und eindeutig zitierbar zu machen.

Doch wie begann alles? Welche Hürden mussten überwunden werden? Und welche Visionen hat DataCite für die Zukunft?

Die Anfänge und die Entwicklung von DataCite

Nachdem 2005 in einem von der Deutschen Forschungsgemeinschaft (DFG) geförderten Projekt¹ der erste Digital Object Identifier (DOI) für Forschungsdaten registriert wurde, war klar, dass dies international von Bedeutung ist. Die Gründung² von DataCite mit sieben Mitgliedern aus sechs Ländern war der erste Schritt in Richtung einer

¹ <https://www.forschungsdaten.org/index.php/STD-DOI> [26.01.2025]

² <https://doi.org/10.5438/g4x3-r450> [11.03.2025]

Über DataCite

Was die ISBN für Bücher ist, ist der Digital Object Identifier (DOI) für Forschungsdaten oder andere digitale Objekte wie digitale Publikationen, 3D-Modelle, Grafiken und Videos. Ein DOI ermöglicht es, auf diese Objekte nachhaltig und eindeutig zuzugreifen sowie wissenschaftliche Ergebnisse zuverlässig zu zitieren.

Die TIB, die British Library (BL), das Technical Information Centre of Denmark (DTU), die TU Delft Library, das National Research Council Canada (NRC-CNRC), die California Digital Library und die Purdue University sind die sieben Gründungsmitglieder der Non-Profit-Organisation DataCite, die am 1. Dezember 2009 gegründet wurde. Ihr Ziel: Sie wollten den Online-Zugang zu Forschungsdaten erleichtern und die Akzeptanz digitaler Objekte als eigenständige und zitierfähige wissenschaftliche Ergebnisse fördern, sodass die zur Reputation von Forschenden beitragen.

15 Jahre nach der Gründung von DataCite sind DOIs für Forschungsdaten etabliert und aus der Wissenschaft nicht mehr wegzudenken, auch wenn zu Beginn Überzeugungsarbeit bei Wissenschaftler:innen nötig war. Die Zahlen sprechen für sich: In den vergangenen 15 Jahren wurden allein über DataCite 82,7 Millionen DOIs für Forschungsdaten und andere digitale Objekte vergeben. Die 312 DataCite-Mitglieder, davon 65 Konsortien kommen aus 58 Ländern und arbeiten bei der DOI-Registrierung mit 1.567 Forschungseinrichtungen der ganzen Welt zusammen. Die Geschäftsstelle von DataCite wird an der TIB in Hannover geführt.

Mehr Informationen zu DataCite <https://datacite.org/>

globalen Initiative. In den darauffolgenden Jahren beteiligte sich DataCite aktiv an EU-Projekten, die sich mit der Nutzung von Persistenten Identifikatoren (PIDs) für die Auffindbarkeit, Vernetzung und Nachnutzung von Forschungsinformationen befassen. Gleichzeitig wuchs die Organisation stetig. Heute umfasst die DataCite-Community mehr als 1.567 Einrichtungen aus 58 Ländern. Genau dieser Community-Ansatz ist es, was essenziell für den Erfolg von DataCite ist.

Etablierter Akteur im Bereich Open Science

DataCite hat sich als Schlüsselakteur im Bereich Open Science etabliert. Initiativen wie das Global-Access-Programm³ unterstützen gezielt Regionen, die bisher unterrepräsentiert waren. Auch auf strategischer Ebene spielt DataCite eine wichtige Rolle, etwa im Kontext der „Barcelona Declaration on Open Research Information“⁴, die sich für offene Forschungsinformationen als Standard einsetzt, oder der Nationalen Forschungsdateninfrastruktur (NFDI)⁵ in Deutschland, die das Potenzial von Forschungsdaten als wertvolle Ressource für Wissenschaft und Forschung nutzbar machen will, indem sie sie systematisch erschließt und nutzbar macht.

³ Mejias, G., Cousijn, H. 2023

⁴ <https://barcelona-declaration.org/>

⁵ <https://www.nfdi.de/>

⁶ <https://base4nfdi.de/>

⁷ <https://ev.igsn.org/>

⁸ <https://raid.org/>

Es ist für DataCite von großer Bedeutung, eng mit der Nutzer-Community zusammenzuarbeiten, um die Services und das standardisierte Metadatenschema bedarfsgerecht anzubieten. In diesem Zusammenhang bietet die Kooperation mit den Forschungsgruppen der NFDI im Rahmen der Base4NFDI-Services⁶ und des Projekts PID4NFDI, gemeinsam mit der TIB und dem Helmholtz Open Science Office, eine hervorragende Gelegenheit, die auch der internationalen Forschungscommunity zugutekommt.

Strategische Partnerschaften

Ergänzend zu Projekten hat DataCite strategische Partnerschaften aufgebaut, unter anderem mit Organisationen wie IGSN (International Generic Sample Number)⁷ und RAiD (Research Activity Identifier).⁸ Diese Partnerschaften erweitern nicht nur die Reichweite, sondern unterstützen auch den Aufbau von Communities of Practice. Durch die Zusammenarbeit mit Partnern, die spezialisierte Identifier anbieten, wird die Nachvollziehbarkeit und Verknüpfung wissenschaftlicher Informationen weiter gestärkt. Darüber hinaus zählen zu den strategischen Partnern von DataCite Integratoren und Dienstleister, die die Tools von DataCite in Workflows und nachgelagerte Systeme integrieren, um die Interoperabilität zu verbessern. Ebenso gehören Organisationen dazu, die gemeinsam mit DataCite Open Science und Innovation vorantreiben, sowie einflussreiche Akteure, die Forschungsinfrastruktur und -politik gestalten. Die strategischen Partnerschaften bilden die Grundlage, um komplexe Forschungsprozesse weiter zu optimieren und globale Kooperationen zu fördern.

In den vergangenen 15 Jahren ist aus einem kleinen Verein mit einer visionären Idee eine globale Community entstanden, die aktiv die Prinzipien von Open Science vorantreibt – für diesen Erfolg steht DataCite mit seinen Mitgliedern und den vielen Akteuren, die sich dort engagieren.

Herausforderungen auf dem Weg zur globalen Infrastruktur

Die letzten 15 Jahre waren nicht ohne Hürden. Eine der größten Herausforderungen war die Sicherstellung der finanziellen Nachhaltigkeit. Die Entwicklung und der Betrieb globaler Infrastrukturen erfordern erhebliche Ressourcen – sowohl technisch als auch personell. Ohne stabile finanzielle Grundlagen wären viele der heutigen Erfolge nicht möglich gewesen.

Ein weiteres Hindernis war die Vermittlung des Mehrwerts von Persistenten Identifikatoren, insbesondere an nicht-technische Zielgruppen. DataCite bietet einen 'non-tangible Service' an, also etwas, das nicht physisch greifbar ist, sondern im digitalen Raum der Wissenschaftskommunikation funktioniert. Den Mehrwert dieser Dienste – insbesondere von PIDs – sowohl technischen als auch nicht-technischen Zielgruppen klar zu vermitteln, war und ist immer noch eine Herausforderung. Auch die globale Reichweite stellte eine Herausforderung dar. DataCite musste Lösungen entwickeln, die weltweit einsetzbar sind und gleichzeitig lokale Besonderheiten berücksichtigen. Dabei halfen Prinzipien wie die „Principles of Open Scholarly Infrastructure“ (POSI)⁹, die Transparenz, Nachhaltigkeit und Offenheit sicherstellen sollen. Rückblickend haben DataCite diese Herausforderungen geholfen, zu priorisieren und DataCite zu einem vertrauenswürdigen Partner der wissenschaftlichen Community zu machen.

Die Zukunft von DataCite: Innovation und Zusammenarbeit

Trotz der Erfolge ruht sich DataCite nicht aus. Die Zukunft bringt neue Herausforderungen und Chancen. Ein zentrales Ziel ist die Verbesserung der Metadatenqualität, denn Metadaten sind der Schlüssel, um Forschungsdaten auffindbar und nutzbar zu machen.

Die Qualität der Metadaten von PIDs spielt eine zentrale Rolle bei der Umsetzung der FAIR-Prinzipien¹⁰ im Kontext von Open Science.¹¹ Sie dienen nicht nur dazu, die

» Was 2009 als kleiner Verein mit nur sieben Gründungsmitgliedern begann, ist heute ein globales Netzwerk, das weltweit von der akademischen Gemeinschaft getragen wird. DataCite ist ein etablierter und unverzichtbarer Baustein für die Umsetzung von Open-Science-Prinzipien und wir als TIB sind stolz, dass die Geschäftsstelle bei uns in Hannover angesiedelt ist. «

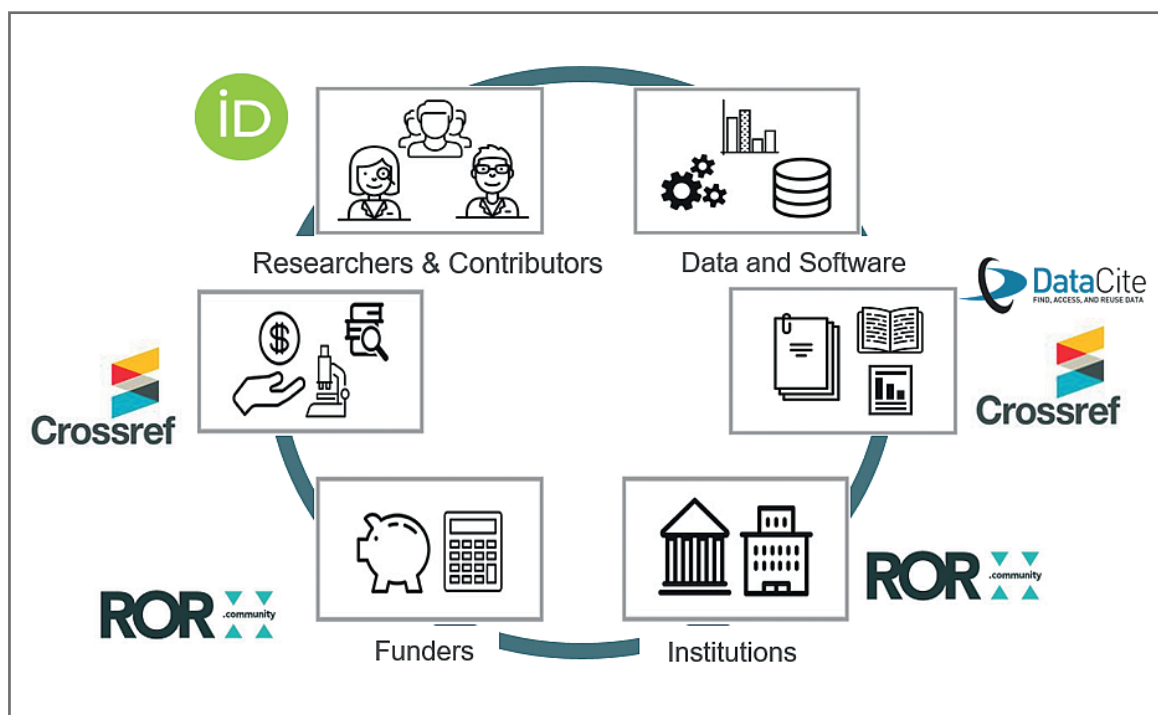
Dr. Irina Sens, stellvertretende Direktorin der TIB

Ressource selbst präzise zu beschreiben, sondern auch Verknüpfungen zu den beteiligten Akteuren herzustellen und die Ressource innerhalb der Forschungsprozesse sinnvoll zu kontextualisieren. Damit schaffen sie die Grundlage für offene Forschungsinfrastrukturen, die den Zugang zu Wissen demokratisieren.

Optimierungspotenziale nutzen

Um das volle Potenzial von PIDs auszuschöpfen, ist es zudem essenziell, die Metadaten kontinuierlich durch zusätzliche PIDs anzureichern – etwa durch die Verknüpfung mit den affilierten Einrichtungen, Datenmanagementplänen, Projekten, Förderinformationen und Zitationen.

Hier gibt es erhebliche Optimierungspotenziale, die nur durch eine gemeinschaftliche und koordinierte Zusammenarbeit der beteiligten Akteure realisiert werden kön-



⁹ <https://openscholarlyinfrastructure.org/>

¹⁰ Wilkinson et al. 2016.

¹¹ Hagemann-Wilholt et al. 2024.



nen. Die derzeitige Praxis, in der Anreicherungen von Metadaten häufig in verteilten und teilweise geschlossenen Systemen erfolgen, behindert den Fortschritt und den Nutzen offener Forschungsinfrastrukturen. Diese Fragmentierung muss überwunden werden.

Eine zentrale Herausforderung besteht darin, Anreicherungen, die sowohl technologische Ressourcen als auch spezialisierte fachliche Expertise erfordern, nicht nur systematisch durchzuführen, sondern auch nachhaltig zugänglich zu machen. Dies erfordert eine stärkere Vernetzung und Harmonisierung der beteiligten Akteure, um sicherzustellen, dass diese angereicherten Metadaten für die globale Forschungscommunity vollständig und offen verfügbar sind. Nur durch solche gemeinschaftlichen Anstrengungen kann das Potenzial von PIDs als Grundlage für FAIR-konforme und offene Wissenschaftsinfrastrukturen vollständig ausgeschöpft werden.

DataCite plant, die Prozesse zur Metadaten-Kuratierung zu zentralisieren und gemeinsam mit der California Digital Library (CDL) ein communitybasiertes Modell zur Metadatenanreicherung zu entwickeln. Diese Prozesse werden durch eine globale Initiative namens COMET (Collaborative Metadata Enrichment Taskforce)¹² vorangetrieben. COMET fördert die Zusammenarbeit innerhalb der Community und setzt auf innovative Ansätze, um Metadaten qualitativ zu verbessern und global zu harmonisieren. Das erhöht nicht nur die Datenqualität, sondern auch eine kollaborative Kultur, die den offenen Austausch von Wissen fördert.

Langfristig will DataCite technologisch und strategisch auch weiterhin Innovationen vorantreiben. Ziel ist es, Forschungsinfrastrukturen an die sich wandelnden Anforderungen der Wissenschaft anzupassen und gleichzeitig offen und nachhaltig zu bleiben.

Ein weiterer Schwerpunkt der Zukunft ist die Intensivierung strategischer Partnerschaften. DataCite wird weiterhin mit relevanten Akteuren zusammenarbeiten, um Forschungsinformationen in die offene Forschungsinfrastruktur zu integrieren und innovative Lösungen für komplexe wissenschaftliche Arbeitsprozesse zu entwickeln. Zudem sollen die Communitys rund um diese Partnerschaften weiter gestärkt werden, um den Austausch und die Weiterentwicklung zu fördern.

Darüber hinaus spielt DataCite eine aktive Rolle in der Initiative „Make Data Count“, die Open-Science-Metriken bereitstellt. Diese ermöglichen es, den Erfolg von Investitionen

in Open Science durch Institutionen, Förderer und Policy Maker messbar zu machen. Gleichzeitig leisten sie einen essenziellen Beitrag zu einem Research Assessment, das Open Science fördert und Forschungsleistungen entsprechend würdigt. Auch hier spielen PIDs eine bedeutende Rolle.

Eines steht sicher fest: Die kommenden Jahre werden spannend – DataCite und alle involvierten Akteure werden diesen Weg gemeinsam mit der Community gestalten und Open Science weltweit weiter voranzubringen.

Fazit

In den vergangenen 15 Jahren hat sich DataCite von einer kleinen Initiative zu einem globalen Akteur entwickelt, der die wissenschaftliche Community maßgeblich prägt. Mit Blick auf die Zukunft zeigt sich: Die Organisation bleibt ihrem Ziel treu, den Zugang zu Forschungsdaten zu verbessern und die Prinzipien von Open Science zu fördern. Dabei setzt sie auf Zusammenarbeit, Innovation und nachhaltige Lösungen – eine Erfolgsgeschichte, die gerade erst begonnen hat. Ergänzt wird diese Entwicklung durch die Arbeit des PID-Kompetenzzentrums der TIB, das die internationalen Aktivitäten rund um PIDs (Persistente Identifikatoren) mit den wissenschaftlichen Einrichtungen in Deutschland implementiert und so einen wichtigen Beitrag zur Integration globaler Standards in die deutsche Forschungslandschaft leistet. ■

Bibliographie

- Hagemann-Wilholt, S., El-Gebali, S., Dreyer, B., Vierkant, P.: Persistent Identifier, ihre Metadaten und FAIRness Eine Analyse zur verbesserten Auffindbarkeit von Forschungsressourcen, b.i.t. online, <https://www.b-i-t-online.de/heft/2024-06-fachbeitrag-hagemann-wilholt.pdf>
- Mejias, G., Cousijn, H.: DataCite launches Global Access Program with support from CZI, <https://doi.org/10.5438/3gs5-3j26>
- Wilkinson, M., Dumontier, M., Aalbersberg, I. et al.: The FAIR Guiding Principles for scientific data management and stewardship, Sci Data 3, <https://doi.org/10.1038/sdata.2016.18>



Britta Dreyer

leitet den PID- und Metadatenservices an der TIB – Leibniz-Informationszentrum Technik und Naturwissenschaften und Universitätsbibliothek und unterstützt dort Forschungseinrichtungen bei der Implementierung von Persistent Identifiers wie DOI und ORCID. Seit 2015 ist sie als Director of Operations & Finance bei DataCite tätig. Dreyer studierte Wirtschaftswissenschaften (MBA) an der Pepperdine University und der Hochschule Hannover. britta.dreyer@tib.eu
<https://orcid.org/0000-0002-0687-5460>

¹² <https://www.cometadadata.org/taskforce> [26.01.2025]