



# The Relationship Between the Publication of Open Research Data in Low and Lower-Middle Income Countries and Equitable Science

---

Dorothea Strecker

EasyChair preprints are intended for rapid dissemination of research results and are integrated with the rest of EasyChair.

September 22, 2023

# Der Zusammenhang zwischen der Publikation offener Forschungsdaten in low and lower-middle income countries und equitable science

**DOROTHEA STRECKER**, Humboldt-Universität zu Berlin, Institut für Bibliotheks- und Informationswissenschaft, Deutschland

Einige Stakeholder sehen einen direkten Zusammenhang zwischen der Anwendung von Open-Science-Praktiken und der gleichberechtigten Teilhabe an wissenschaftlichen Erkenntnissen und Erkenntnisprozessen („equitable science“), darunter die UNESCO in den „Sustainable Development Goals“(SDG). Allerdings stehen einige Forschende aus „low and lower income countries“(LLMIC) Open-Science-Praktiken zum Teil skeptisch gegenüber, obwohl sie laut UNESCO eigentlich in besonderem Maß davon profitieren sollten. In diesem Beitrag wird untersucht, welche Perspektive Forschenden aus LLMIC auf Veröffentlichung von Forschungsdaten haben und wie Forschungsdatenrepositorien darauf einwirken können. Anschließend wird die geografische Verteilung von Forschungsdatenrepositorien und traditioneller Publikationsinfrastrukturen auf Kontinente und Länder verschiedener Einkommensgruppen untersucht. Der Beitrag zeigt, dass der in den SDG dargestellte Zusammenhang zwischen Open Science und equitable science nicht in dieser direkten Form besteht. Das liegt daran, dass Open-Science-Praktiken in einem Wissenschaftssystem wirken, das bereits von Ungleichheiten geprägt ist, und diese nicht notwendigerweise beseitigt. Die Bereitschaft von Forschenden aus LLMIC, ihre Forschungsdaten zu veröffentlichen, wird durch die ungleiche Verteilung von Ressourcen, Misstrauen sowie Diskurse, die stark von „high income countries“(HIC) geprägt sind, negativ beeinflusst. Forschungsdatenrepositorien vor Ort können diese Probleme zwar nicht beseitigen, aber sie können die Publikation von Forschungsdaten vereinfachen. Allerdings sind sie global sehr ungleich verteilt – diese Verteilung folgt der traditioneller Publikationsinfrastrukturen, ist also stark auf HIC und Länder des Globalen Nordens konzentriert. Insgesamt leistet Open Science zwar einen Beitrag für mehr Teilhabe, allerdings sollte zukünftig neben der Verfügbarmachung von Forschungsergebnissen auch die Beteiligung von Forschenden aus LLMIC an Forschungsprozessen stärker in den Blick genommen werden.

Additional Key Words and Phrases: Open Science ; equitable science ; Sustainable Development Goals ; Forschungsdaten ; Publikationsinfrastrukturen

## 1 EINLEITUNG

In den „Recommendations on Open Science“beschreibt die UNESCO den Zusammenhang zwischen Open Science und dem Erreichen der „Sustainable Development Goals“(SDG) [43]. Open Science umfasst nach diesem Dokument verschiedene Praktiken, die darauf abzielen, wissenschaftliche Erkenntnisse und Prozesse der Wissensproduktion für eine breite Öffentlichkeit zugänglich zu machen. In dieser Vision der UNESCO führt Open Science zu einer Reduktion von Ungleichheiten im Wissenschaftssystem und unterstützt damit auch die Erfüllung der SDG.

Allerdings gibt es auch abweichende Positionen, was die Wirkung von Open-Science-Praktiken in Bezug auf „equitable science“, also die gleichberechtigten Teilhabe an wissenschaftlichen Erkenntnissen und Erkenntnisprozessen, und damit auch die SDG, angeht. Auffällig ist, dass auch Forschende aus „low and lower-middle income countries“(LLMIC), die laut UNESCO eigentlich von Open-Science-Praktiken in besonderem Maß profitieren sollten, diesen teilweise kritisch gegenüberstehen.

---

Author's address: [Dorothea Strecker](mailto:dorothea.strecker@hu-berlin.de), dorothea.strecker@hu-berlin.de, Humboldt-Universität zu Berlin, Institut für Bibliotheks- und Informationswissenschaft, Unter den Linden 6, Berlin, 10099, Deutschland.

## 1.1 Fragestellung

In diesem Beitrag wird darum untersucht, inwiefern sich eine bestimmte Open-Science-Praktik, die Publikation offener Forschungsdaten, auf die Teilhabe an wissenschaftlichen Erkenntnissen und Prozessen der Wissensproduktion in LLMIC auswirkt. Insbesondere wird die Rolle unterstützender Informationsinfrastrukturen beleuchtet. Der Beitrag orientiert sich an drei Leitfragen:

- (1) Welche Perspektiven haben Forschende aus LLMIC auf die Veröffentlichung von Forschungsdaten?
- (2) Welchen Einfluss können Forschungsdatenrepositorien in LLMIC auf die Veröffentlichung von Forschungsdaten haben?
- (3) Wie sind Forschungsdatenrepositorien geografisch und auf Kontinente und Länder verschiedener Einkommensgruppen verteilt, und wie unterscheiden sich diese Verteilungen von der traditioneller Publikationsinfrastrukturen?

Die Bearbeitung der ersten beiden Fragen stützt sich auf eine Auswertung der Literatur. Mit der dritten Frage wird die Verteilung von Forschungsdatenrepositorien und traditionellen Publikationsinfrastrukturen (wissenschaftliche Verlage) auf Kontinente und Länder verschiedener Einkommensgruppen untersucht. Zu diesem Zweck werden offene Datenquellen analysiert und visualisiert.

## 2 HINTERGRUND

### 2.1 Offene Forschungsdaten in LLMIC

Regelmäßig werden Hoffnungen geäußert, mit offenen Forschungsdaten könnten Lösungen für globale Herausforderungen gefunden werden, beispielsweise die Gewährleistung von Nahrungsmittelsicherheit, die Abwendung von Folgen der Klimakrise, oder die Sicherung von Biodiversität [35]. Aus diesem Grund haben viele Stakeholder ein Interesse an offenen Forschungsdaten, die in LLMIC erhoben wurden und lokale Verhältnisse beschreiben, darunter Forschende und Vertreter:innen von Regierungen oder Nichtregierungsorganisationen weltweit. Gleichzeitig berichten Forschende aus LLMIC, die diese Daten vor Ort erfassen, dass ihre Leistungen oft nicht ausreichend anerkannt werden [16]. Die mangelnde Anerkennung von Datenpublikationen kritisieren zwar auch Forschende aus „high-income countries“ (HIC) [21], allerdings verschärfen globale Ungleichheiten diese Diskurse aus der Perspektive von Forschenden aus LLMIC [38]. Dies zeigt ein grundlegendes Problem auf: Open-Science-Praktiken wirken in einem bestehenden Wissenschaftssystem, das durch grundlegende Ungleichheiten beim Zugang zu wissenschaftlichen Erkenntnissen und der Teilhabe an wissenschaftlichen Erkenntnisprozessen gekennzeichnet ist. Open Science adressiert diese strukturellen Ungleichheiten nicht notwendigerweise [3].

Die Ungleichheiten gehen häufig auf Unterschiede in der Verfügbarkeit von Ressourcen zurück [37]. Obwohl der Forschungoutput von LLMIC in den vergangenen Jahren deutlich angestiegen ist, verfügen sie im Vergleich zu HIC noch immer über weniger Ressourcen und Infrastrukturen [8]. Darum machen sich globale Ungleichheiten bereits bei der Datenerhebung bemerkbar [28] – etwa, wenn Forschende aus LLMIC, die Daten für Kollaborationen mit HIC erheben, unter prekären Bedingungen arbeiten [16]. Die Finanzierung von Forschungsaktivitäten stellt eine wesentliche Herausforderung für Forschende aus LLMIC dar. Dabei ist jedoch nicht allein der Mittelumfang entscheidend; auch das Aufbringen vergleichsweise kleiner Ausgaben, etwa für Mitgliedschaften oder Softwarelizenzen, kann schwierig sein, da diese oft nicht über projektgebundene Mittel finanziert werden können [9].

Der Ressourcenmangel hat weitreichende Folgen für Forschende aus LLMIC. Unterschiede in der Ausstattung führen etwa zu variierenden Geschwindigkeiten bei der Durchführung von Forschungsvorhaben in LLMIC und HIC. Der

Zeitaufwand für die Erhebung von Daten ist in LLMIC oft höher, etwa wenn eine zuverlässige Versorgung mit notwendigen Ressourcen in Laboren nicht gewährleistet ist oder Arbeitsmaterial geteilt werden muss. Eine Untersuchung in Laboren in Kenia zeigte, dass die Bereitschaft von Forschenden, ihre Daten zu veröffentlichen, wesentlich von dieser aufgebrauchten Zeit abhängt, da sie fürchten, ihr wissenschaftliches Kapital zu verlieren [7]. Mangelnde Ausstattung führt außerdem dazu, dass Forschende in LLMIC weniger von der Sekundärnutzung offener Forschungsdaten profitieren können [6]. So kommt ein wesentlicher Vorteil offener Forschungsdaten in LLMIC oft nicht zum Tragen, und ein Gefühl von Gegenseitigkeit bei der Publikation offener Forschungsdaten stellt sich nicht ein.

Neben mangelnden Ressourcen wirkt sich ein weiterer Faktor wesentlich auf die Bereitschaft aus, Forschungsdaten zu veröffentlichen: Misstrauen – genauer, die Angst davor, ausgenutzt zu werden [4]. Das Gefühl von Misstrauen kann in LLMIC durch die oben beschriebene ungleiche Ressourcenverteilung und Kolonialismuserfahrungen noch verstärkt werden: „Differences in languages and cultures, as well as imbalances in trust and power, breed suspicion about motives, impeding data access and sharing.“ [1, S. 334] Das Misstrauen von Forschenden aus LLMIC lässt sich beispielsweise auf zwei befürchtete Verhaltensmuster zurückführen, die in der Literatur beschreiben werden:

- „parachute science“: Die Extraktion von Daten aus LLMIC ohne angemessene Anerkennung der Forschenden vor Ort [1, 39]
- „scooping“: Die Angst davor, dass Forschende aus HIC Ergebnisse basierend auf Daten, die Forschende aus LLMIC erhoben und veröffentlicht haben, schneller publizieren können [8]

Die Angst vor „scooping“ knüpft an die oben beschriebenen variierenden Geschwindigkeiten von LLMIC und HIC an. Forschende aus LLMIC befürchten, dass ihnen Forschende aus HIC bei der Auswertung der von ihnen erhobenen Daten zuvorkommen könnten und sie so für ihre Arbeit nicht adäquat anerkannt werden. Eine weitere Befürchtung, die zur Bildung von Misstrauen gegenüber Veröffentlichung von Forschungsdaten beiträgt, ist die mangelnde Berücksichtigung der Interessen von Proband:innen aus LLMIC bei der Veröffentlichung von Daten [6, 24].

Das Misstrauen wird im Spannungsfeld zwischen lokalen Gegebenheiten und einer angestrebten Internationalisierung von Open-Science-Initiativen weiter verstärkt. Häufig werden beispielsweise Anforderungen von außen an Forschende aus LLMIC herangetragen, denn Vorstellungen von Open Science und guter wissenschaftlicher Praxis in internationalen Leitlinien sind oft von HIC geprägt. Aufgrund spezifischer lokaler Gegebenheiten sind sie jedoch nicht ohne weiteres auf LLMICs übertragbar, sondern sollten sorgfältig angepasst werden [1, 3, 6, 8, 10].

Spannungen, die bei der Verbreitung und Nutzung von Forschungsdaten entstehen können, werden auch als „data friction“ bezeichnet; Diskurse um offene Forschungsdaten in HIC sind stark beeinflusst von der Idee, diese Reibung möglichst zu minimieren [18]. Reibungen können jedoch auch ein wichtiges Signal sein, das auf ungeklärte Fragestellungen oder Konflikte unter Interessengruppen hinweist, die dringend adressiert werden sollten [27, 41].

## 2.2 Forschungsdatenrepositorien in LLMIC

Neben den oben beschriebenen Hürden während des Forschungsprozesses können auch fehlende Forschungsdatenrepositorien die Publikation offener Forschungsinfrastrukturen behindern. Forschungsdatenrepositorien sind Infrastrukturen, die auf die Kuratierung, Erhaltung und Verbreitung von Forschungsdaten spezialisiert sind [23]. Die Anzahl von Repositorien sowie der Umfang ihrer Sammlungen wachsen etwa in Afrika kontinuierlich an [13, 19], allerdings langsamer als in anderen Kontinenten [31]. Forschungsdatenrepositorien werden häufig mit dem Ziel etabliert, Zugang zu Forschungsergebnissen der betreibenden Einrichtung zu ermöglichen [12] und diese langfristig zu sichern [26]. Verschiedene Effekte tragen dazu bei, dass Informationen oft einseitig weitergereicht werden – von HIC an LLMIC.

Repositorien können dieser Tendenz entgegenwirken, indem sie die Teilhabe von Forschenden aus LLMIC am wissenschaftlichen Publikationssystem realisieren [25]. Zwar können sie Probleme, die während des Forschungsprozesses auftreten, nicht beseitigen; sie können aber Hürden bei der Publikation von Forschungsdaten verringern. Das ist möglich, da sie einerseits in globale Kontexte eingebunden sind, und andererseits eine starke lokale Wirkung entfalten, beispielsweise indem sie Forschende vor Ort unterstützen und so Expertise etablieren [13, 40]. Zudem sichern sie Wissensbestände über lokale Gegebenheiten, fördern Austausch unter Forschenden vor Ort und können die Ungleichheit in der Verfügbarkeit von Informationen über LLMIC gegenüber HIC verringern [30]. Der Betrieb von Repositorien in LLMIC ermöglicht es außerdem, die Rahmenbedingungen, unter denen Forschungsdaten veröffentlicht werden, zu definieren [3, 24]. So können Forschungsdatenrepositorien in LLMIC nicht nur die Interessen der lokalen Bevölkerung wahren, sondern auch die der Forschenden. Die Befugnis zur Kontrolle über Daten durch Betroffene wird beispielsweise auch im Kontext indigener Bevölkerungsgruppen hervorgehoben [11].

Der Betrieb von Repositorien kann viele Vorteile für LLMIC bieten, diesen stehen jedoch auch einige Herausforderungen gegenüber. Dazu zählen Schwankungen in der Internetverbindung und Stromversorgung, ungesicherte Finanzierung, fehlende personelle Ressourcen oder institutionelle Unterstützung durch die betreibende Einrichtung [2, 13, 17, 22]. Das hat zur Folge, dass das Netz von Forschungsdatenrepositorien in LLMIC gegenüber HIC weniger ausgeprägt ist. Im Folgenden werden diese globalen Unterschiede näher betrachtet und durch eine Analyse der Verteilung traditioneller Publikationsinfrastrukturen (wissenschaftlicher Verlage) kontextualisiert.

### 3 VORGEHEN

Um die Verteilung von Infrastrukturen auf Länder verschiedener Einkommensgruppen zu analysieren, werden offene Datenquellen genutzt.

Es gibt mehrere Verzeichnisse, die Forschungsdatenrepositorien nachweisen; ihre Bestände überlappen sich teilweise [5]. Für Angaben zu Forschungsdatenrepositorien und die betreibenden Einrichtungen wird das Verzeichnis von Forschungsdatenrepositorien re3data [34] genutzt, da es aktuell die umfangreichste Sammlung von Metadaten zu Forschungsdatenrepositorien ist. Der Auswertung traditioneller Publikationsinfrastrukturen liegt eine Liste der 100 größten wissenschaftlichen Verlage zugrunde, gemessen an der Anzahl der Zeitschriften, die sie herausgeben [32]. Beide Samples werden um die World Bank Income Groups ergänzt [44].

Im Folgenden wird die Datenerhebung im Detail beschrieben.

#### 3.1 Datenerhebung

*re3data*. Das Verzeichnis re3data wurde am 06.06.2023 über die offene API abgefragt.<sup>1</sup> Dabei wurden basierend auf dem re3data Metadata Schema [42] einige Einschränkungen vorgenommen, um Infrastrukturen auszuwählen, die der Definition von Forschungsdatenrepositorien, die diesem Beitrag zugrunde liegt [23], entsprechen. Die vorgenommene Einschränkungen sind:

- Infrastrukturen, die (auch) Daten bereitstellen und nicht nur Metadaten aggregieren (r3d:providerType = dataProvider)
- Infrastrukturen, die den Upload von Daten grundsätzlich ermöglichen (r3d:dataUploadType != closed)

<sup>1</sup>re3data API: <https://www.re3data.org/api/doc>

Abgefragt wurde das Feld `r3d:institutionCountry`. In diesem Feld werden Angaben zum Geschäftssitz der am Betrieb von Forschungsdatenrepositorien beteiligten Einrichtungen gemacht. Die resultierenden geografischen Informationen zu Forschungsdatenrepositorien sind demnach an die Einrichtungen geknüpft, die an deren Betrieb beteiligt sind. Dass die Lokalität eines Forschungsdatenrepositoriums, das ein virtueller Dienst ist, schwer zu bestimmen ist, wurde bereits an anderer Stelle dargelegt [36]; der Geschäftssitz der betreibenden Einrichtungen ist hingegen präziser. In dieser Analyse können einem Forschungsdatenrepositorium mehrere geografische Bezüge zugeordnet werden, wenn es von Institutionen aus mehreren Ländern gemeinsam betrieben wird.

*3.1.1 Liste der größten wissenschaftlichen Verlage.* Die Auswertung der geografischen Verteilung traditioneller Publikationsinfrastrukturen basiert auf einer Untersuchung, bei der verschiedene Datenquellen zusammengeführt wurden, um die 100 größten wissenschaftlichen Verlage zu ermitteln, gemessen an der Anzahl der Zeitschriften, die sie herausgeben [32]. 29 Verlage, die in der Datenquelle als „predatory“ deklariert wurden, wurden von der Untersuchung ausgeschlossen. Das Land des Geschäftssitzes der übrigen 71 Verlage wurde über die Verlagswebseiten ermittelt.

*World Bank Income Groups.* Die Aufteilung der Länder in Einkommensgruppen basiert auf den World Bank Income Groups, Stand 2021-2022 [44]. Die World Bank unterteilt Länder basierend auf ihrem Bruttonationaleinkommen pro Kopf in vier Einkommensgruppen:

- low income countries (LIC)
- lower-middle income countries (LMIC)
- upper-middle income countries (UMIC)
- high income countries (HIC)

### **3.2 Beschreibung der Samples**

Zum Zeitpunkt der Datenerhebung am 06.06.2023 verfügte re3data über 3.122 Einträge. Davon erfüllten 1.351 die oben genannten Einschlusskriterien, diese stellen das finale Sample dar.

Das Sample für traditionelle Publikationsinfrastrukturen umfasst die 71 größten wissenschaftlichen Verlage, die laut der nachgenutzten Liste keine „predatory“ Verlage sind.

## **4 ERGEBNISSE**

### **4.1 Forschungsdatenrepositorien**

*4.1.1 Geografische Verteilung.* Im untersuchten Sample sind insgesamt 1.351 Forschungsdatenrepositorien vertreten, an deren Betrieb Institutionen aus 66 Ländern beteiligt sind (siehe Abbildung 1).

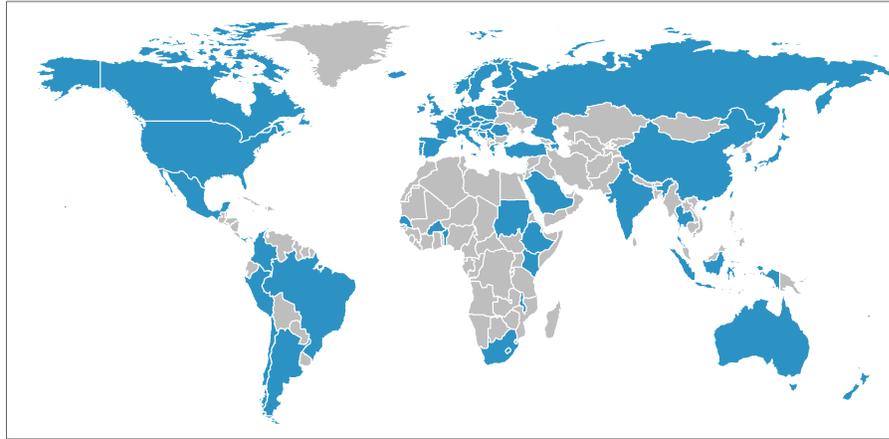


Fig. 1. Sitz der Einrichtungen, die am Betrieb von Forschungsdatenrepositorien beteiligt sind (Quelle: re3data)

Institutionen aus Nordamerika sind am Betrieb von 639 Forschungsdatenrepositorien beteiligt. Im Vergleich dazu sind es 602 aus Europa, 109 aus Asien<sup>2</sup>, 57 aus Ozeanien, 35 aus Südamerika und 17 aus Afrika (siehe Abbildung 2a).

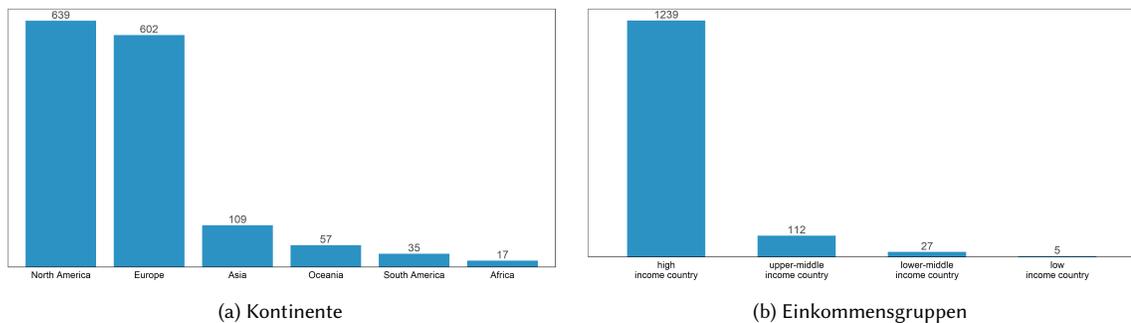


Fig. 2. Verteilung der am Betrieb von Forschungsdatenrepositorien beteiligten Einrichtungen auf Kontinente und Einkommensgruppen (Quelle: re3data / World Bank)

Mehrfachnennungen sind auf kontinentübergreifende Kollaborationen zurückzuführen. Diese sind jedoch insgesamt selten – nur 6,9 % (93) der Forschungsdatenrepositorien im Sample werden von Einrichtungen aus mindestens zwei Kontinenten betrieben. Dabei sind Kollaborationen zwischen Einrichtungen aus Europa und Nordamerika (52) noch am häufigsten zu beobachten, andere Kollaborationen sind deutlich weniger verbreitet.

**4.1.2 Verteilung auf Einkommensgruppen.** Im Sample sind 40 HIC vertreten, sowie 17 UMIC, 5 LMIC und 4 LIC. Diese ungleiche Verteilung auf Einkommensgruppen spiegelt sich auch an der Anzahl der Einrichtungen, die am Betrieb der Forschungsdatenrepositorien im Sample beteiligt sind, wider: Einrichtungen aus HIC sind am Betrieb von 1.239 Repositorien beteiligt; demgegenüber sind deutlich weniger Einrichtungen aus UMIC (112), LMIC (27) oder LIC (5) vertreten (siehe Abbildung 2b).

<sup>2</sup>Russland und Türkei wurden Asien zugerechnet.

Mehrfachnennungen sind auch hier möglich, wenn ein Repositorium von Einrichtungen aus Ländern verschiedener Einkommensgruppen gemeinsam betrieben wird. Kollaborationen unter Ländern verschiedener Einkommensgruppen sind jedoch noch seltener als interkontinentale Kollaborationen, sie bilden die Grundlage für nur 2,3 % (31) der Forschungsdatenrepositorien im Sample.

#### 4.2 Traditionelle Publikationsinfrastrukturen

Die 71 untersuchten wissenschaftlichen Verlage haben ihren Geschäftssitz in 22 Ländern auf 4 Kontinenten. 32 der größten Verlage haben ihren Geschäftssitz in Europa, 16 in Nordamerika, 15 in Asien, und 8 in Südamerika. Verlage aus Afrika und Ozeanien sind im Sample nicht vertreten (siehe Abbildung 3a).

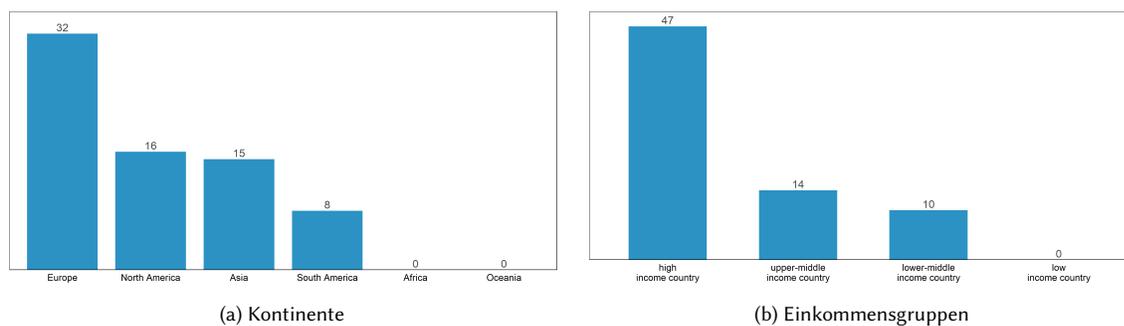


Fig. 3. Verteilung der größten wissenschaftlichen Verlage auf Kontinente und Einkommensgruppen (Quelle: [32] / World Bank)

47 Verlage haben ihren Geschäftssitz in einem HIC, deutlich weniger Verlage operieren aus UMIC (14) und LMIC (10). LIC sind im Sample nicht vertreten (siehe Abbildung 3b). Diese schiefe Verteilung auf Länder verschiedener Einkommensgruppen wird noch extremer, wenn man die Anzahl der Zeitschriften in der zugrundeliegenden Datenquelle betrachtet: 87,3 % (20.654) der Zeitschriften werden von Verlagen in HIC herausgegeben.

## 5 DISKUSSION

### 5.1 Verteilung von Forschungsdatenrepositorien und traditionellen Publikationsinfrastrukturen auf Kontinente und Einkommensgruppen

Forschungsdatenrepositorien vor Ort können zwar globale Ungleichheiten während des Forschungsprozesses nicht ausräumen, aber sie können die Bereitschaft von Forschenden aus LLMIC erhöhen, ihre Forschungsdaten zu veröffentlichen. Die Analyse verwies allerdings auf eine ungleiche Verteilung von Forschungsdatenrepositorien auf Kontinente und Einkommensgruppen. Institutionen aus LLMIC und dem Globalen Süden sind im untersuchten Sample systematisch unterrepräsentiert. Diese Beobachtung spiegelt existierende Ungleichheiten wider: Die Auswertung zeigte, dass die ungleiche Verteilung von Forschungsdatenrepositorien auf Kontinente und Länder verschiedener Einkommensgruppen der Verteilung traditioneller Publikationsinfrastrukturen folgt. Auch die größten wissenschaftliche Verlage sind stark im Globalen Norden konzentriert, und ein Großteil der Zeitschriften im Sample wird von Verlagen mit Geschäftssitz in HIC herausgegeben. Bestehende Muster in der Landschaft von Publikationsinfrastrukturen wiederholen sich also beim Aufbau von Diensten, die sich auf Open-Science-Praktiken spezialisiert haben.

Mechanismen, die globale Ungleichheiten im wissenschaftlichen Publikationswesen aufrecht erhalten, werden bereits

untersucht [15]. Um dieses grundlegende Problem nicht bei der Verfügbarmachung von Forschungsdaten zu wiederholen, sollte zukünftige Forschung kritisch prüfen, inwiefern sich diese Mechanismen auch auf Forschungsdatenrepositorien auswirken können.

Der Ausbau lokaler Infrastrukturen wäre ein Ansatzpunkt, um Open-Science-Praktiken in LLMIC zu etablieren. Forschungsdatenrepositorien wirken sowohl global als auch lokal: Sie fördern neben der internationalen Sichtbarkeit von Forschungsergebnissen auch die Sicherung von Wissensbeständen vor Ort. Die lokale Wirkung kann sich beispielsweise im Angebot niedrigschwelliger und an die spezifische Situation angepasste Beratungsangebote in der Sprache der Forschenden zeigen. Außerdem können die Rahmenbedingungen für die Veröffentlichung von Forschungsdaten lokal festgelegt werden. Dadurch kann das Vertrauen von Forschenden aus LLMIC in den Vorgang bestärkt werden, da ihre Interessen in internationalen Kontexten in der Vergangenheit nicht immer berücksichtigt wurden.

Wie die Analyse zeigte, konzentrieren sich neben anderen Infrastrukturen des wissenschaftlichen Publikationswesens auch die größten Forschungsdatenrepositorien im Globalen Norden. Viele dieser Forschungsdatenrepositorien nehmen zwar keine Einschränkungen in Bezug auf die Herkunft der Datengebenden vor, sie stehen also grundsätzlich auch Forschenden aus LLMIC offen. In Zukunft sollte jedoch näher untersucht werden, inwiefern internationale Forschungsdatenrepositorien die hier beschriebene lokale Wirkung entfalten können, denn allein die Verfügbarkeit von Open-Science-Infrastrukturen reicht nicht aus: Sie müssen auch vielfältigen Bedürfnissen, Fähigkeiten und Erkenntnisprozessen gerecht werden, wenn sie die beabsichtigte Wirkung erzielen sollen [33].

## 5.2 Open Science und equitable science

Bei genauerer Betrachtung zeigt sich, dass der Zusammenhang zwischen Open Science und equitable science nicht so direkt besteht wie in den SDG dargelegt. Das liegt daran, dass Open Science in einem Wissenschaftssystem wirkt, das bereits von Ungleichheiten geprägt ist. Diese strukturellen Probleme werden von Open-Science-Initiativen nicht zwangsläufig adressiert.

Außerdem werden Differenzen im Anspruch an Open Science sichtbar. Der Sammelbegriff Open Science kann als Projektionsfläche für teilweise sehr unterschiedliche Konzepte dienen [20], die verschiedene Erwartungen wecken. Dienen Open-Science-Initiativen primär dem Teilen wissenschaftlicher Erkenntnisse, oder geht es um ein umfassenderes Verständnis von Teilhabe? Auch die Wissenschaftsphilosophin Sabina Leonelli beobachtet diese abweichenden Auffassungen [29]: Sie beschreibt, dass Open-Science-Initiativen aktuell stark auf die Verfügbarmachung von Forschungsergebnissen ausgerichtet sind, und weniger darauf, die Beteiligung an Forschungsprozessen zu ermöglichen. Diese Unterscheidung in die Teilhabe an Forschungsergebnissen und die Teilhabe an Forschungsprozessen ist zentral, was den Zusammenhang zwischen Open Science und equitable science betrifft. Wie oben beschrieben wird die Bereitschaft von Forschenden aus LLMIC, ihre Forschungsdaten zu veröffentlichen, durch die ungleiche Verteilung von Ressourcen, Misstrauen von Forschenden sowie Diskurse, die stark von HIC geprägt sind, negativ beeinflusst. Dies zeigt, dass Forschende sich mehr gleichberechtigte Teilhabe an Forschungsprozessen wünschen – sie wollen nicht nur auf Forschungsergebnisse zugreifen, sondern selbst auch aktiv Forschung betreiben. Open-Science-Initiativen sollten klar kommunizieren, was genau sie unter Open Science verstehen, um keine falschen Euerwartungen zu wecken. Sie sollten auch die beschriebene Unterscheidung reflektieren und idealerweise nicht nur Forschungsergebnisse, sondern unbedingt auch Forschungsprozesse für Forschende aus LLMIC zugänglicher machen. Auch in anderen Gesellschaftsbereichen wächst unter dem Label „critical data studies“ ein kritisches Bewusstsein dafür, wie Datenflüsse Ungleichheiten erzeugen oder verstärken können [14]. Open-Science-Initiativen, die die Teilhabe von Forschenden aus LLMIC verbessern wollen, könnten von diesen Perspektiven profitieren.

Das Verhältnis zwischen Open Science und equitable science ist also komplexer als in den SDG dargestellt. Differenzierter könnte es beschrieben werden als: Open Science kann zwar einen Beitrag für mehr Teilhabe an wissenschaftlichen Erkenntnisprozessen leisten, ist aber kein Garant dafür.

## 6 FAZIT

Aktuell halten einige Barrieren Forschende aus LLMIC davon ab, ihre Forschungsdaten zu veröffentlichen. Da sich diese Barrieren besonders auf die mangelnde Teilhabe an Forschungsprozessen beziehen, sollten Open-Science-Initiativen ihren Fokus auch darauf ausweiten. Forschungsdatenrepositorien können dazu beitragen, Open-Science-Praktiken in LLMIC zu etablieren. Zu den Ungleichheiten im globalen Wissenschaftssystem zählt jedoch auch die Konzentration von Einrichtungen des Publikationswesens in HIC, die sich auch in der Landschaft von Forschungsdatenrepositorien fort schreibt. Verbreitete Open-Science-Praktiken reichen darum nicht aus, um mehr Teilhabe von Forschenden aus LLMIC zu ermöglichen, vielmehr müssen historisch gewachsene Muster aufgebrochen werden.

Der in den SDG dargestellte Zusammenhang zwischen Open Science und der gleichberechtigten Teilhabe an wissenschaftlichen Erkenntnissen und Erkenntnisprozessen besteht nicht in dieser direkten Form. Open-Science-Praktiken sollten darum eher als Beitrag zu equitable science angesehen werden, und nicht als Strategien mit Erfolgsgarantie.

## LITERATUR

- [1] Rediet Abebe, Kehinde Aruleba, Abeba Birhane, Sara Kingsley, George Obaido, Sekou L. Remy, and Swathi Sadagopan. Narratives and Counter-narratives on Data Sharing in Africa. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, pages 329–341. Association for Computing Machinery, 2021. doi: 10.1145/3442188.3445897.
- [2] Usman Ahmed Adam and Kiran Kaur. Institutional repositories in Africa: Regaining direction. *Information Development*, 38(2):166–178, 2022. doi: 10.1177/02666669211015429.
- [3] Denisse Albornoz, Maggie Huang, Issra Marie Martin, Maria Mateus, Aicha Yasmine Touré, and Leslie Chan. Framing Power: Tracing Key Discourses in Open Science Policies. In *Proceedings of the 22nd International Conference on Electronic Publishing*, 2018. doi: 10.4000/proceedings.elpub.2018.23.
- [4] Marlene Sophie Altenmüller, Matthias Fligge, and Mario Gollwitzer. Among us: Fear of exploitation, suspiciousness, and social identity predict knowledge hiding among researchers. *Social Psychological Bulletin*, 18:e10011, 2023. doi: 10.32872/spb.10011.
- [5] Miriam Baglioni, Andrea Mannocci, Gina Pavone, Michele De Bonis, and Paolo Manghi. (Semi)automated disambiguation of scholarly repositories. In *Proceedings of the 19th The Conference on Information and Research science Connecting to Digital and Library science*, pages 47–59, 2023. URL <https://ceur-ws.org/Vol-3365/paper2.pdf>.
- [6] Karen I. Barnes, Julio Arturo Canario, Sunil S. Vernekar, Shivaprasad S. Goudar, Roberto Espinal, Laura Merson, and Phaik Yeong Cheah. Equitable data sharing: challenges and suggestions for ways forward. *Wellcome Open Research*, 4:172, 2019. doi: 10.12688/wellcomeopenres.15425.1.
- [7] Louise Bezuidenhout. Technology Transfer and True Transformation: Implications for Open Data. *Data Science Journal*, 16(0):26, 2017. doi: 10.5334/dsj-2017-026.
- [8] Louise Bezuidenhout and Ereck Chakauya. Hidden concerns of sharing research data by low/middle-income country scientists. *Global Bioethics*, 29(1):39–54, 2018. doi: 10.1080/11287462.2018.1441780.
- [9] Louise Bezuidenhout, Ann H. Kelly, Sabina Leonelli, and Brian Rappert. ‘\$100 Is Not Much To You’: Open Science and neglected accessibilities for scientific research in Africa. *Critical Public Health*, 27(1):39–49, 2017. doi: 10.1080/09581596.2016.1252832.
- [10] Jeff Camkin, Susana Neto, Basundhara Bhattarai, Hemant Ojha, Shahbaz Khan, Ai Sugiura, Jiaying Lin, Fitrie Atviana Nurritasari, and Joseph Muiruri Karanja. Open Science for Accelerating the Sustainable Development Goals: Status and Prospects in Asia and the Pacific. *Frontiers in Political Science*, 4, 2022. doi: 10.3389/fpos.2022.878761.
- [11] Stephanie Russo Carroll, Ibrahim Garba, Oscar L. Figueroa-Rodríguez, Jarita Holbrook, Raymond Lovett, Simeon Materechera, Mark Parsons, Kay Raseroka, Desi Rodriguez-Lonebear, Robyn Rowe, Rodrigo Sara, Jennifer D. Walker, Jane Anderson, and Maui Hudson. The CARE Principles for Indigenous Data Governance. *Data Science Journal*, 19:43, 2020. doi: 10.5334/dsj-2020-043.
- [12] Colleen Takaingenhamo Chisita and Blessing Chiparasha. An Institutional Repository in a Developing Country: security and Ethical Encounters at the Bindura University of Science Education, Zimbabwe. *New Review of Academic Librarianship*, 27(1):130–143, 2021. doi: 10.1080/13614533.2020.1824925.
- [13] Elisha R.T. Chiware. Open research data in African academic and research libraries: a literature analysis. *Library Management*, 41(6/7):383–399, 2020. doi: 10.1108/LM-02-2020-0027.

- [14] Jonathan Cinnamon. Data inequalities and why they matter for development. *Information Technology for Development*, 26(2):214–233, 2020. doi: 10.1080/02681102.2019.1650244.
- [15] Fran M Collyer. Global patterns in the publishing of academic knowledge: Global North, global South. *Current Sociology*, 66(1):56–73, 2018. doi: 10.1177/0011392116680020.
- [16] Renato A. F. de Lima, Oliver L. Phillips, Alvaro Duque, J. Sebastian Tello, Stuart J. Davies, Alexandre Adalardo de Oliveira, Sandra Muller, Euridice N. Honorio Coronado, Emilio Vilanova, Aida Cuni-Sanchez, Timothy R. Baker, Casey M. Ryan, Agustina Malizia, Simon L. Lewis, Hans ter Steege, Joice Ferreira, Beatriz Schwantes Marimon, Hong Truong Luu, Gerard Imani, Luzmila Arroyo, Cecilia Blundo, David Kenfack, Moses N. Sainge, Bonaventure Sonké, and Rodolfo Vásquez. Making forest data fair and open. *Nature Ecology & Evolution*, 6(6):656–658, 2022. doi: 10.1038/s41559-022-01738-7.
- [17] Ncamsile Nombulelo Dlamini and Maritha Snyman. Institutional repositories in Africa: obstacles and challenges. *Library Review*, 66(6/7):535–548, 2017. doi: 10.1108/LR-03-2017-0021.
- [18] Paul N. Edwards, Matthew S. Mayermik, Archer L. Batcheller, Geoffrey C. Bowker, and Christine L. Borgman. Science friction: Data, metadata, and collaboration. *Social Studies of Science*, 41(5):667–690, 2011. doi: 10.1177/0306312711413314.
- [19] Ifeanyi J. Ezema and Omwoyo Bosire Onyancha. Open access publishing in Africa: advancing research outputs to global visibility. *African Journal of Library, Archives and Information Science*, 27(2), 2017. URL <https://www.ajol.info/index.php/ajlais/article/view/164661>.
- [20] Benedikt Fecher and Sascha Friesike. Open Science: One Term, Five Schools of Thought. In Sönke Bartling and Sascha Friesike, editors, *Opening Science*, pages 17–47. Springer, Cham, 2014. doi: 10.1007/978-3-319-00026-8\_2.
- [21] Kathleen Gregory, Anton Ninkov, Chantal Ripp, Emma Roblin, Isabella Peters, and Stefanie Haustein. Tracing data: A survey investigating disciplinary differences in data citation. *Quantitative Science Studies*, pages 1–51, 2023. doi: 10.1162/qss\_a\_00264.
- [22] Jerome Idiegbeyan-Ose, Goodluck Ifijeh, Juliana Iwu-James, Julie Ilogho, Jerome Idiegbeyan-Ose, Goodluck Ifijeh, Juliana Iwu-James, and Julie Ilogho. Management of Institutional Repositories (IR) in Developing Countries. In *E-Discovery Tools and Applications in Modern Libraries*. IGI Global, 2016. doi: 10.4018/978-1-5225-0474-0.ch016.
- [23] Lisa R. Johnston, Jacob Carlson, Cynthia Hudson-Vitale, Heidi Imker, Wendy Kozlowski, Robert Olendorf, and Claire Stewart. How Important is Data Curation? Gaps and Opportunities for Academic Libraries. *Journal of Librarianship and Scholarly Communication*, 6(1), 2018. doi: 10.7710/2162-3309.2198.
- [24] Jaranit Kaewkungwal, Pornpimon Adams, Jetsumon Sattabongkot, Reidar K. Lie, and David Wendler. Issues and Challenges Associated with Data-Sharing in LMICs: Perspectives of Researchers in Thailand. *The American Journal of Tropical Medicine and Hygiene*, 103(1):528–536, 2020. doi: 10.4269/ajtmh.19-0651.
- [25] Julia Kelly and Linda Eells. Global scholarship: The role of subject repositories in advancing research from the developing world. *College & Research Libraries News*, 76(5), 2015. doi: 10.5860/crln.76.5.9313.
- [26] Iryna Kuchma and Brian Rosenblum. Report on Open Repository Development in Developing and Transition countries. Technical report, EIFL, 2010. URL <https://www.eifl.net/resources/report-open-repository-development-developing-and-transition-countries>.
- [27] Sebastián Lehuédé. When friction becomes the norm: Antagonism, discourse and planetary data turbulence. *New Media & Society*, page 146144482211084, 2022. doi: 10.1177/14614448221108449.
- [28] Sebastián Lehuédé. The coloniality of collaboration: sources of epistemic obedience in data-intensive astronomy in Chile. *Information, Communication & Society*, 26(2):425–440, 2023. doi: 10.1080/1369118X.2021.1954229.
- [29] Sabina Leonelli. *Philosophy of Open Science*. Cambridge University Press, Cambridge, MA, 2023. doi: 10.1017/9781009416368.
- [30] Robert Molteno. Why African digital repositories for storing research writings are so important, 2019. URL <https://www.internationalafricaninstitute.org/repositories/why>.
- [31] Joseph Mwalubanda. The development of institutional repositories in East Africa countries: A comparative analysis of Tanzania, Kenya, and Uganda. *IASSIST Quarterly*, 45(3-4), 2021. doi: 10.29173/iq1012.
- [32] Andreas Nishikawa-Pacher. Who are the 100 largest scientific publishers by journal count? A webscraping approach. *Journal of Documentation*, 78(7):450–463, 2022. doi: 10.1108/JD-04-2022-0083.
- [33] Angela Okune, Rebecca Hillyer, Denisse Albornoz, Alejandro Posada, and Leslie Chan. Whose Infrastructure? Towards Inclusive and Collaborative Knowledge Infrastructures in Open Science. In *Proceedings of ELPUB 2018*, 2018. doi: 10.4000/proceedings.elpub.2018.31.
- [34] Heinz Pampel, Paul Vierkant, Frank Scholze, Roland Bertelmann, Maxi Kindling, Jens Klump, Hans-Jürgen Goebelbecker, Jens Gundlach, Peter Schirnbacher, and Uwe Dierolf. Making Research Data Repositories Visible: The re3data.org Registry. *PLOS ONE*, 8(11):e78080, 2013. doi: 10.1371/journal.pone.0078080.
- [35] Claudio Persello, Jan Dirk Wegner, Ronny Hansch, Devis Tuia, Pedram Ghamisi, Mila Koeva, and Gustau Camps-Valls. Deep Learning and Earth Observation to Support the Sustainable Development Goals: Current approaches, open challenges, and future opportunities. *IEEE Geoscience and Remote Sensing Magazine*, 10(2):172–200, 2022. doi: 10.1109/MGRS.2021.3136100.
- [36] re3data COREF. Mapping the global repository landscape, 2022. URL <https://coref.project.re3data.org/blog/mapping-the-global-repository-landscape>.
- [37] Tony Ross-Hellauer, Stefan Reichmann, Nicki Lisa Cole, Angela Fessl, Thomas Klebel, and Nancy Pontika. Dynamics of cumulative advantage and threats to equity in open science: a scoping review. *Royal Society Open Science*, 9(1):211032, 2022. doi: 10.1098/rsos.211032.

- [38] Amber Hartman Scholz, Matthias Lange, Pia Habekost, Paul Oldham, Ibon Cancio, Guy Cochrane, and Jens Freitag. Myth-busting the provider-user relationship for digital sequence information. *GigaScience*, 10(12):giab085, 2021. doi: 10.1093/gigascience/giab085.
- [39] David Serwadda, Paul Ndebele, M. Kate Grabowski, Francis Bajunirwe, and Rhoda K. Wanyenze. Open data sharing and the Global South—Who benefits? *Science*, 359(6376):642–643, 2018. doi: 10.1126/science.aap8395.
- [40] Ina Smith and Susan Veldsman. Data Driving Sustainability—the African Open Science Platform Project. In *Proceedings of ELPUB 2018*, 2018. doi: 10.4000/proceedings.elpub.2018.25.
- [41] Vicki Squire and Modesta Alozie. Coloniality and frictions: Data-driven humanitarianism in North-Eastern Nigeria and South Sudan. *Big Data & Society*, 10(1):20539517231163171, 2023. doi: 10.1177/20539517231163171.
- [42] Dorothea Strecker, Roland Bertelmann, Helena Cousijn, Kirsten Elger, Lea Maria Ferguson, David Fichtmüller, Hans-Jürgen Goebelbecker, Maxi Kindling, Gabriele Kloska, Thanh Binh Nguyen, H. Pampel, Vivien Petras, Rouven Schabinger, Edeltraud Schnepf, Angelika Semrau, Margarita Trofimenko, Robert Ulrich, Arne Upmeier, Paul Vierkant, Nina Leonie Weisweiler, Yi Wang, and Michael Witt. Metadata Schema for the Description of Research Data Repositories: version 3.1, 2021.
- [43] UNESCO. UNESCO Recommendation on Open Science, 2021.
- [44] World Bank. The World by Income and Region (2021-2022), 2022. URL <https://datatopics.worldbank.org/world-development-indicators/the-world-by-income-and-region.html>.