



Data At Your Fingertips: Wegmarken einer exzellenten und vertrauenswürdigen Datenlandschaft

EcoDM-Empfehlungen

Christoph Bruch

HELMHOLTZ OPEN SCIENCE OFFICE

Heike Neuroth

FACHHOCHSCHULE POTSDAM

Henriette Senst

DEUTSCHES ARCHÄOLOGISCHES INSTITUT

ZITIERHINWEIS:

Bruch, Christoph; Neuroth, Heike; Senst, Henriette (2022).
Data At Your Fingertips: Wegmarken einer exzellenten und vertrauenswürdigen
Datenlandschaft. Empfehlungen des EcoDM-Projekts. DOI: [10.5281/zenodo.5837408](https://doi.org/10.5281/zenodo.5837408).

COPYRIGHT:

© 2022 Die Autor*innen. Dies ist ein Open-Access Artikel, der unter den Bedingungen der Creative Commons Attribution 4.0 International License (CC-BY 4.0) zur Verfügung gestellt wird, die die uneingeschränkte Nutzung, Verbreitung und Vervielfältigung in jedem Medium erlaubt, sofern die ursprünglichen Autor*innen und die Quelle genannt werden. Siehe: <http://creativecommons.org/licenses/by/4.0/>.

GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung

FÖRDERKENNZEICHEN:

16DWWQP

Data At Your Fingertips:

Wegmarken einer exzellenten und vertrauenswürdigen Datenlandschaft

EcoDM-Empfehlungen¹

Vision

Exzellente und vertrauenswürdige Daten für Fragestellungen aller Art vereinen sich zu einer Datenlandschaft. Diese Landschaft formt sich durch Wegmarken wie Teilhabe, Ethik, Rechtssicherheit, Standardisierung, Interoperabilität und Wettbewerbsfähigkeit. Landschaftsgestaltende kommen aus dem Kreis der Wissenschaft, des Public Sector, der Wirtschaft, der Gesellschaft und der Privat-Interessierten.²

Das BMBF-geförderte Projekt EcoDM³ hat erforscht, welche Herausforderungen, Chancen und Hindernisse sich im Bereich des rasanten Wachstums digitaler Daten ergeben und wie Rahmenbedingungen aussehen können, Daten systematisch und FAIR (Findable, Accessible, Interoperable, Reusable) nutzen und teilen zu können. Den zentralen Teil der Untersuchung bildeten Landscape- und Gap-Analysen sowie Leitfaden-gestützte Interviews mit Expert*innen aus den Bereichen Wissenschaft, Wirtschaft, Public Sector und Qualifizierung. Der Forschungsansatz basierte auf zwei zentralen Hypothesen zu den Rahmenbedingungen, die mindestens erfüllt sein müssen, um Potenziale der Datennutzung und Datenteilung effektiver und effizienter auszuschöpfen:

1. Daten müssen auffindbar, zugänglich, interoperabel und nachnutzbar gemacht werden (FAIRification).
2. Relevante Akteursgruppen müssen zur Freigabe eigener und Nachnutzung von Daten Dritter motiviert bzw. dabei unterstützt werden.

Ausgehend von diesen beiden Hypothesen erfolgte die Erstellung von Landscape- und Gap-Analysen in den Bereichen Wissenschaft, Public Sector, Wirtschaft und Qualifizierung. Darüber hinaus wurden vom Projekt den einzelnen Bereichen zugeordnete Expert*innen mittels Leitfaden-gestützter Interviews befragt, um Bedingungen zu erforschen, welche

- a) die FAIRifizierung von Daten erleichtern oder erschweren und
- b) die Motivation für das Teilen bzw. Nachnutzen von Daten beeinflussen.

¹ Unkommentierte Version, die für die Podiumsdiskussion am 22.02.2022 zur Verfügung gestellt wurde.

² Basierend auf: Franke, M., Heinzl, S., Mauer, R., Neumann, J., Neuroth, H., Pfeiffenberger, H., Senst, H., Sieberg, A., Sitek, D., Spiecker, C., Wambsganß, J., Weber, D., Winkler-Nees, S. (2015): Positionspapier „Research data at your fingertips“ der Arbeitsgruppe Forschungsdaten, Potsdam : Deutsches GeoForschungsZentrum GFZ, 4 p. <https://doi.org/10.2312/allianzfd.001>.

³ Gefördert vom BMBF (Bundesministerium für Bildung und Forschung): Förderkennzeichen 16DWWQP; Projektlaufzeit 2019-2022.

In der vorliegenden Publikation werden die wichtigsten Empfehlungen auf dem Weg zu einer exzellenten und vertrauenswürdigen Datenlandschaft vorgestellt sowie die relevanten Erkenntnisse zusammengefasst.⁴ Die folgenden Wegmarken für diese Datenlandschaft sind als Beitrag zu den aktuellen Diskussionen zu sehen. Dies umfasst Entwicklungen und Vorhaben, wie sie zum Beispiel in Deutschland mit dem Aufbau von Dateninfrastrukturen, Instrumenten wie Datentreuhändern, Datendrehscheiben etc. gemeinsam mit Wirtschaft, Wissenschaft und Zivilgesellschaft auf den Weg gebracht werden sollen, um die Potenziale von Daten für alle zu heben.⁵ International steht die angebahnte Zusammenarbeit von GAIA-X und EOSC beispielhaft für diese Entwicklungen.⁶

Nachdem für die vier Bereiche Wissenschaft, Wirtschaft, Public Sector und Qualifizierung sowohl die jeweiligen Empfehlungen als Ergebnis aus einer Landscape- und Gap-Analyse als auch aus den Leitfaden-gestützten Interviews übergreifend analysiert wurden (vgl. EcoDM Bericht 2022), haben sich die folgenden übergreifenden EcoDM-Empfehlungen herauskristallisiert.

Harmonisierung und Kooperation

1. Die Ausgestaltung der in Europa angestrebten **digitalen Souveränität** sollte mit dem Interesse am Erhalt der internationalen Kooperationsfähigkeit ausbalanciert werden. Dabei darf der Schutz der Bürger*innenrechte nicht zur Disposition gestellt werden.
2. Der **Datenschutz** in Deutschland sollte ausschließlich durch Bundesrecht geregelt werden, um die praktische Umsetzung zu vereinfachen.
3. Internationale Aktivitäten zur **Harmonisierung der Standards** für Metadaten und Datenformate, z. B. European Open Data Portal, sollten unterstützt werden.
4. Die **technische Harmonisierung** sollte durch die Stärkung zentral betriebener Infrastrukturen unterstützt werden. Diese sind aufgrund ihrer Größe (ausreichende personelle, finanzielle und qualitätsgesicherte nachhaltige Ausstattung) in der Lage, mit der schnellen technischen Entwicklung Schritt zu halten und sich intensiv an internationalen Prozessen zur Entwicklung und Harmonisierung von Standards zu beteiligen und dort Perspektiven aus Deutschland einzubringen.
5. Plattformen sollten für ihre Daten offene Standards nutzen, um deren **Interoperabilität** (technisch, syntaktisch, semantisch) zu maximieren und somit einen Beitrag zur Nachnutzung von Daten zu leisten.
6. **Leistungserbringung** sollte haushaltsrechtlich über Institutions- und Ländergrenzen hinweg ermöglicht bzw. erleichtert werden, um gemeinsame Daten-Dienstleistungen bundesweit und gegebenenfalls auch international leichter anbieten zu können.
7. Informationsinfrastrukturen in der Wissenschaft sollten immer als **Teil einer europäischen Forschungsinfrastruktur** verstanden werden und dementsprechend internationale

⁴ Der zugrunde liegende Projektbericht beschreibt ausführlich die Fragestellungen, Methoden und einzelne Ergebnisse je Bereich.

⁵ Bundesregierung: Koalitionsvertrag 2021. - <https://www.bundesregierung.de/breg-de/service/gesetzesvorhaben/koalitionsvertrag-2021-1990800>.

⁶ EOSC Association (23.07.2021): Aligning priorities between EOSC and Gaia-X, https://www.eosc.eu/news/aligning-priorities-eosc-gaia-x?utm_source=newsletter&utm_medium=email&utm_campaign=20210803_eosc_gaia-x

Nutzungen ermöglichen. Es sollten Doppelungen im Aufbau und Betrieb von Beginn an vermieden werden.

8. Die Daten der NFDI-Konsortien sollten über eine **zentrale Plattform** zugänglich gemacht werden, um einer Versäulung innerhalb der Wissenschaftsdisziplinen entgegen zu wirken.
9. Bei der **Nutzung von Produkten von Hyperscalern** oder der *Kooperationen mit* ihnen sollte eine kritische und transparente Auseinandersetzung über Chancen und Risiken erfolgen.

Anreizsysteme und Kompetenzentwicklung

10. Das Teilen von Daten, einschließlich benötigter Software und Services, sollte durch geeignete **Anreizsysteme** honoriert werden.
11. **Anreiz- und Anerkennungsinstrumente** für das Teilen von Daten sollten z. B. durch die Schaffung neuer Berufsfelder, Karrierewege und Qualifizierung gestärkt werden.
12. Der **systematische Aufbau von Datenkompetenz** sollte auf allen beruflichen Ebenen unterstützt werden.
13. Der **Kulturwandel** sollte im Sinne einer Data Literacy Charta gefördert werden, d. h. in allen Bildungsbereichen beginnend mit der Elementarbildung.
14. Der Erwerb von Basiswissen zum Umgang mit **rechtlichen und ethischen** Aspekten bei der Zugänglichmachung und Nachnutzung von Daten sollte unterstützt werden.
15. Organisationen sollten ihre **Weiterbildung** stärker und regelmäßiger auf ihre Praxisrelevanz, z.B. für den Umgang mit Daten, überprüfen. In allen **Studiengängen** sollten die Themen Datenkompetenz (data literacy) und Datenmanagement stärker berücksichtigt werden.
16. Neue Berufsfelder mit **Schnittstellenfunktion** sollten sowohl innerhalb einer Einrichtung als auch zwischen Bereichen wie Wirtschaft und Wissenschaft etabliert werden.
17. Beratungsangebote zu FAIRem Datenmanagement sollten sowohl für wissenschaftsunterstützendes Personal wie auch für Forschende sichergestellt werden. Hierfür sollten institutionsinterne und -externe **Beratungsangebote vernetzt** (OER) werden.
18. Der **Nutzen** der Zugänglichmachung von Verwaltungsdaten sollte verwaltungsintern besser erläutert werden, um die Motivation zur Priorisierung dieses Zieles zu erhöhen.

Ethische Grundsätze und öffentliches Interesse

19. Ist für die Realisierung **herausragender gesamtgesellschaftlicher Interessen** der Zugriff auf bestimmte Daten notwendig, sollten bestehende rechtliche Möglichkeiten zur Gewährleistung dieses Zugriffes genutzt und falls notwendig eine Stärkung dieser Zugriffsrechte durch den Gesetzgeber (wie im aktuellen Koalitionsvertrag angesprochen) geprüft werden.
20. Bei der Nutzung von Daten in **(teil-)automatisierten Entscheidungsprozessen** muss durch Regeln zu Transparenz, einschließlich Auskunftsansprüchen und Qualitätssicherung, gewährleistet werden, dass die Entscheidungsfindung nachvollziehbar ist und Diskriminierungen ausgeschlossen werden.

Vertrauenswürdigkeit

21. Deutschland bzw. Europa **sollte rechtlich sowie technisch vertrauenswürdige Datenräume** schaffen, da diese nicht von einzelnen Datenanbietenden realisiert werden können. Dies gilt insbesondere für den Bereich sensibler Daten.
22. **Datentreuhandssysteme** müssen durch öffentlich-rechtliche Institutionen und nicht durch private Unternehmen kontrolliert werden.
23. Daten sollten im Zuge ihrer Freigabe zur Nachnutzung - möglichst liberal - „**lizenzier**“ werden, damit klar kommuniziert wird, wie die Daten genutzt werden dürfen. Dies betrifft auch die Kontexte rund um Daten, wie z.B. Datenbankrechte, Softwarelizenzen, Persönlichkeitsrechte etc. Die Entwicklung standardisierter, global verständlicher Sets von „Erlaubnis- und Verbotmodulen“ für diese Lizenzierungen, die optimalerweise maschinell interpretierbar sind, ist zu prüfen.

Vermarktung

24. Bei der Förderung der Nachnutzung von Daten sollte der **Aufbereitung** der Daten mehr Aufmerksamkeit geschenkt werden. Dort können neue Geschäftsfelder und damit Potenziale für Ausgründungen entstehen.
25. Die Förderung der öffentlichen Zugänglichmachung von Daten sollte die Förderung der **Nachnutzung** einschließen. Das kann beispielsweise über eine Ideen-Plattform passieren, über die Unternehmen ihre datengetriebene Fragestellung beschreiben. Dies unterstützt die Kommunikation zwischen Datenanbietenden und -nutzenden.
26. Unternehmen sollten Potenzialen der **Datennachnutzung** durch Dritte mehr Aufmerksamkeit schenken. Hier liegen große wirtschaftliche Chancen brach.
27. Die Nachnutzung von Knowhow und Daten aus der Forschung scheitert oft an Kompetenzlücken zwischen Forschung und Wirtschaft. Hier werden **vermittelnde Stellen** z. B. für die Datenaufbereitung, -analyse etc., benötigt. Für die Finanzierung dieser Mittler fehlen Finanzierungs- bzw. Geschäftsmodelle.
28. **Usability** muss intuitiv sein, z. B. in Richtung nutzerzentrierte, deviceoptimierte, personalisierbare User Experience.

Wirkung von Open Data Strategien

29. **Folgen der Offenheit** von Daten sollten genauer untersucht werden, z. B. Innovationspotenzial, Schnelligkeit von Entwicklungen, Fragestellungen zur Privacy, Wirtschaftlichkeit, Reziprozität.
30. Der **gesamtgesellschaftliche** (inklusive volkswirtschaftliche) **Nutzen** der Bereitstellung von Daten (z. B. durch die Verwaltung) sollte begleitend erhoben und evaluiert werden.
31. Es ist zu prüfen, inwieweit die öffentliche Zugänglichmachung von Daten im Rahmen bestimmter **Geschäftsmodelle** die Kooperationswilligkeit behindert, z. B. bei Verbundprojekten zwischen Wissenschaft und Wirtschaft.

*Beteiligte und Inputgeber*innen*

Jasper Bothe FACHHOCHSCHULE POTSDAM, Christine Burkart FACHHOCHSCHULE POTSDAM,
Maria Chlastak FACHHOCHSCHULE POTSDAM, Christian Djeffal TECHNISCHE UNIVERSITÄT MÜNCHEN,
Kirsten Endres FACHHOCHSCHULE POTSDAM, Reinhard Messerschmidt HELMHOLTZ OPEN SCIENCE OFFICE,
Florian Müller FACHHOCHSCHULE POTSDAM, Janina Richter HELMHOLTZ OPEN SCIENCE OFFICE,
Laura Rothfritz FACHHOCHSCHULE POTSDAM, Wolfgang Schmidle DEUTSCHES ARCHÄOLOGISCHES INSTITUT,
Esther Schneidenbach DEUTSCHES ARCHÄOLOGISCHES INSTITUT, Claus Spiecker FACHHOCHSCHULE
POTSDAM, Paul Walter DEUTSCHES ARCHÄOLOGISCHES INSTITUT.

Danksagungen

Wir bedanken uns bei den nachfolgenden Personen und Einrichtungen für ihre Unterstützung, die konstruktiven Diskussionen und die Bereitstellung ihrer Expertise:

- Reiner Mauer GESIS für die methodische Expertise
- Raphael Ritz MAX PLANCK COMPUTING AND DATA FACILITY als Inputgeber für den Projektantrag
- alle Interviewten für die Bereitschaft mit uns zu sprechen und die spannenden Einblicke
- den Podiumsgästen und dem Moderator Jan-Martin Wiarda für ihre Diskussionsfreude und die neuen Erkenntnisse
- dem Projektträger VDI/VDE Innovation + Technik GmbH für die Begleitung und Unterstützung in schwierigen Zeiten
- dem Bundesministerium für Bildung und Forschung für die Förderung

Weitere Informationen zum EcoDM-Projekt

www.ecodm.de