

EcoDM – ÖKOSYSTEM DATENMANAGEMENT

Analysen – Empfehlungen – FAIRifizierung

EcoDM-REPORT

GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung



EcoDM – Ökosystem Datenmanagement

Analysen – Empfehlungen – FAIRifizierung

Herausgeber*Innen

Christoph Bruch Helmholtz Open Science Office
Heike Neuroth Fachhochschule Potsdam
Henriette Senst Deutsches Archäologisches Institut

Autor*Innen

Gesamtreport

Claus Spiecker Fachhochschule Potsdam
Janina Richter Helmholtz Open Science Office
Paul Walter Deutsches Archäologisches Institut
Maria Chlastak Fachhochschule Potsdam
Christine Burkart Fachhochschule Potsdam
Esther Schneidenbach Deutsches Archäologisches Institut
Reinhard Messerschmidt Zukunft – Umwelt – Gesellschaft (ZUG) gGmbH (ehemals Helmholtz Open Science Office)
Kirsten Endres Fraunhofer-Institut für Produktionsanlagen und Konstruktionstechnik (ehemals Fachhochschule Potsdam)

Rechtsgutachten (Kapitel 3, Gesamtreport)

Christian Djeffal Technische Universität München
Eva Nürnberger Technische Universität München

Beteiligte

Jasper Bothe Fachhochschule Potsdam
Tanja Friedrich DLR (ehemals GESIS)
Reiner Mauer GESIS
Florian Müller Universität Kassel (ehemals Fachhochschule Potsdam)
Raphael Ritz Max Planck Computing And Data Facility
Laura Rothfritz Humboldt-Universität zu Berlin (ehemals Fachhochschule Potsdam)
Wolfgang Schmidle Deutsches Archäologisches Institut

FH;P Fachhochschule Potsdam
University of
Applied Sciences



DEUTSCHES
ARCHÄOLOGISCHES INSTITUT
ZENTRALE

HELMHOLTZ
Open Science

Impressum

Online veröffentlicht über das Open-Access-Repository
Zenodo: <https://doi.org/10.5281/zenodo.6256398>

Zitierhinweis

Bruch, C., Neuroth, H., Senst, H. (Hrsg.). (2022). EcoDM –
Ökosystem Datenmanagement: Analysen – Empfehlungen
– FAIRifizierung. <https://doi.org/10.5281/zenodo.6256398>.

Copyright

© 2022 Projekt EcoDM, vertreten durch das Deutsche Archäologische Institut. Dies ist eine Open-Access-Veröffentlichung, die unter den Bedingungen der Creative Commons Attribution 4.0 International License (CC-BY 4.0) zur Verfügung gestellt wird, welche 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/>. Die Lizenz für den EcoDM-Bericht gilt nicht für die Illustrationen, an denen individuelle Lizenzen vermerkt sind.

GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung

Förderkennzeichen

16DWWQP

Inhaltsverzeichnis

Executive Summary	I
Abbildungsverzeichnis	III
Tabellenverzeichnis	V
Akronym- und Abkürzungsverzeichnis	VI
1. Einleitung	2
1.1 Hintergrund und Ziele	2
1.2 Die Untersuchungsbereiche	3
1.3 Definitionen	4
1.3.1 Daten	4
1.3.2 Daten-Ökosystem	5
1.3.3 FAIR-Prinzipien	6
2. Methodik	10
2.1 Literatur-basierte Analyse: Landscape- und Gap-Analyse	12
2.2 Semi-strukturierte Expert*innen-Interviews	13
2.3 Erarbeitung von Handlungsempfehlungen	15
2.4 Community Engagement und Outreach	16
2.4.1 Community Engagement im Web	16
2.4.2 Einbindung von Nachwuchsforschenden	16
2.4.3 Community Engagement über Konferenzen und Workshops ...	16
2.4.4 Einbindung von Übersetzungen	18
3. Externe Rechtsexpertise	22
3.1 Einleitung: Problemaufriss	22
3.2 Rechte an Daten	22
3.2.1 Datenschutzrecht	22
3.2.2 Recht des geistigen Eigentums	23
3.3 Rechte auf Daten	27
3.3.1 Wettbewerbsrecht	27
3.3.2 Das Recht offener Daten in der öffentlichen Verwaltung im „zersplitterten“ Mehrebenensystem	29
3.4 Gesetzgeberische Initiativen	30
3.4.1 Freiwilliges Teilen von Daten zwischen privaten Akteur*innen	30
3.4.2 Pflichten zum Teilen von Daten zwischen privaten Akteur*innen	31
3.4.3 Das Teilen von Daten in der Verwaltung	32
3.5 Zusammenfassung	33

4. Literatur-basierte Analyse:	
Landscape- und Gap-Analyse	36
4.1 Wissenschaft	36
4.1.1 Quellengrundlagen	36
4.1.2 Rahmenbedingungen	36
4.1.3 Infrastruktur	39
4.1.4 Wissen und Renommee im Wandel	41
4.1.5 Defizite und Handlungsbedarf	42
4.1.6 Zusammenfassung der Kernpunkte	44
4.2 Public Sector	45
4.2.1 Quellengrundlagen	45
4.2.2 Rahmenbedingungen	45
4.2.3 Infrastrukturen und Stakeholder	49
4.2.4 Öffentlicher Auftrag und gesetzliche Vorgaben	51
4.2.5 Defizite und Handlungsbedarf	53
4.2.6 Zusammenfassung der Kernpunkte	55
4.3 Wirtschaft	55
4.3.1 Quellengrundlagen	55
4.3.2 Rahmenbedingungen	56
4.3.3 Infrastrukturen	59
4.3.4 Plattformökonomie	60
4.3.5 Defizite und Handlungsbedarf	63
4.3.6 Zusammenfassung der Kernpunkte	64
4.3.7 Exkurs: Kritik dominanter Narrative zum Verhältnis von Markt, Staat und Innovationen der Digitalwirtschaft	65
4.4 Qualifizierung	66
4.4.1 Quellengrundlagen	66
4.4.2 Rahmenbedingungen und Vorgaben	67
4.4.3 Infrastruktur	69
4.4.4 Qualifizierungsbedarf	69
4.4.5 Defizite und Handlungsbedarf	75
4.4.6 Zusammenfassung der Kernpunkte	76
5. Interview-basierte Analyse:	
Expert*innen-Interviews	80
5.1 Wissenschaft	80
5.1.1 Einstieg	80
5.1.2 Geschäftsmodell	80
5.1.3 Vorgaben und rechtliche Rahmenbedingungen	81
5.1.4 Datenbewusstsein	82
5.1.5 Kompetenzaufbau	83
5.1.6 Datenqualität und Standards	83
5.1.7 Vernetzung	84
5.1.8 Infrastruktur	84
5.2 Public Sector	85
5.2.1 Einstieg	85
5.2.2 Geschäftsmodell	86
5.2.3 Vorgaben und rechtliche Rahmenbedingungen	87
5.2.4 Datenbewusstsein	87
5.2.5 Kompetenzaufbau	88

5.2.6	Datenqualität und Standards	89
5.2.7	Vernetzung	89
5.2.8	Infrastruktur	90
5.3	Wirtschaft	91
5.3.1	Einstieg	91
5.3.2	Geschäftsmodelle	91
5.3.3	Vorgaben und rechtliche Rahmenbedingungen	92
5.3.4	Datenbewusstsein und Datenkompetenz	92
5.3.5	Datenqualität und Standards	93
5.3.6	Vernetzung	94
5.3.7	Infrastruktur	94
5.4	Qualifizierung	95
5.4.1	Einstieg	95
5.4.2	Rahmenbedingungen und Infrastruktur	95
5.4.3	Datenbewusstsein und Kulturwandel	96
5.4.4	Kompetenzprofile, Fähigkeiten und Skills	96
5.4.5	Medienkompetenzvermittlung und Qualifikationswege	97
5.4.6	Berufsbilder und Karrierewege	98
6.	Bereichsspezifische EcoDM-Empfehlungen	102
6.1	Wissenschaft	102
6.2	Public Sector	104
6.3	Wirtschaft	106
6.4	Qualifizierung	108
7.	Bereichsübergreifende EcoDM-Empfehlungen	114
8.	Fazit und Ausblick	118
9.	Literatur- und Quellenverzeichnis	124
10.	Anhang	134
10.1	Forschungsdatenmanagementplan	134
10.2	Vorlage der Einverständniserklärung	139
10.4	Bereichsübergreifender Interview-Leitfaden	143
10.5	Verzeichnis der Projektbeteiligten	148

Executive Summary

In die Nutzung von Daten, die durch die Digitalisierung in bisher nicht gekannten Mengen erzeugt und verarbeitet werden können, werden große Hoffnungen gesetzt. Dies gilt sowohl hinsichtlich der mit der Erzeugung und Nutzung erwarteten Steigerung der wirtschaftlichen Leistungsfähigkeit als auch bezüglich der Steigerung der Fähigkeit zur Lösung von Problemen und der stärkeren Unterfütterung von Entscheidungsprozessen mit Evidenz. Daten werden daher vielfach als der wichtigste Rohstoff der digitalisierten Gesellschaft betrachtet. In einem weltweit bekannt gewordenen Artikel aus dem Jahr 2017 in der Zeitschrift *The Economist* wurden Daten deshalb als das neue Öl bezeichnet („The world’s most valuable resource is no longer oil, but data“, 2017).

Die Metapher charakterisiert den Rohstoff Daten natürlich hinsichtlich einer zentralen Eigenschaft falsch. Anders als Öl werden Daten durch ihre Nutzung nicht verbraucht oder entwertet. Im Gegenteil: Der Wert von Daten kann durch die Herstellung großer Sammlungen und die Entwicklung der Kapazität, diese vielfältig auszuwerten, dauerhaft gesteigert werden. Dieser Umstand befeuert gleichzeitig das Interesse am Geheimhalten und Nachnutzen von Daten. Interesse am Nachnutzen von Daten besteht in allen Bereichen der Gesellschaft. Unternehmen fokussieren dabei auf den wirtschaftlichen Nutzen. Der Public Sector und die Politik möchten ihre Steuerungsfähigkeit verbessern und ihre Entscheidungsprozesse stärker evidenzbasiert ausrichten. In der datenhungrigen Wissenschaft verspricht man sich von der Nutzung digitaler Daten die Beschleunigung von Forschungsprozessen, die Entwicklung neuartiger interdisziplinärer Fragestellungen sowie die Ermöglichung neuer Forschungsmethoden und von der Veröffentlichung der Daten, deren Auswertung Forschungsergebnisse wesentlich begründen, eine Verbesserung der Qualitätssicherung von Forschungsprozessen. Die breite Gesellschaft hat ebenfalls Interesse an datenbasierten Informationen. Das illustriert z. B. die Entstehung der neuen Profession der Datenjournalist*innen. Die Wissenschaft hat mit der schnellen Entwicklung von Corona-Impfstoffen das bisher augenscheinlichste Beispiel für den Mehrwert des Teilens von Daten gegeben.

Die Datenflut, insbesondere wenn es sich um personenbezogene Daten handelt, begründet auch berechtigte Ängste. Unternehmen, die für Konsumenten Services anbieten, um im Kontext deren Nutzung Profile über die Nutzer*innen anzulegen und möglicherweise individualisierte Werbung einzublenden, handeln grundsätzlich mit dem Versprechen, Personen beeinflussen zu können. Die Glaubwürdigkeit dieses Versprechens zeigt sich in den Gewinnen, die diese Unternehmen u. a.

mit dem Schalten individualisierter Werbung generieren können. Die Macht, welche Unternehmen akkumulieren, die eine große Anzahl von Kund*innen haben, die mit ihren persönlichen Daten für die Nutzung eines Service zahlen, wird deutlich, wenn beispielsweise Amazon Vorteile bei der Prognose von Marktentwicklungen hat oder Einfluss darauf nehmen kann, von welchem Händler die Besucher dieses Handelsplatzes ein Produkt erwerben. Die zunehmende Bedeutung automatisierter und individualisierter Kommunikation im Kontext von Wahlen zu politischen Ämtern ist ein weiteres Szenario, das Vorbehalte gegen Big Data begründet.

Für wissenschaftliche Institutionen, im Projekt *EcoDM – Ökosystem Datenmanagement* repräsentiert durch das Deutsche Archäologische Institut, die Fachhochschule Potsdam und das Helmholtz Open Science Office, besteht ein besonderes Interesse an der Förderung des Teilens von Daten, sowohl innerhalb der Wissenschaft, als auch zwischen den Bereichen Public Sector, Wirtschaft und Wissenschaft. Aus diesem Interesse erklärt sich die Leitfrage für die Konzeption des Forschungsprojektes *EcoDM*: Welche Bedingungen fördern oder behindern in den Bereichen *Public Sector*, *Wirtschaft* und *Wissenschaft* das Teilen und Nachnutzen von Daten. Darüber hinaus wurden – ebenfalls für die drei genannten Bereiche – die Qualifizierungsbedarfe hinsichtlich des Managements und der Kuratierung von Daten erforscht.

Für die Verfolgung beider Fragestellungen wurde mithilfe von Literaturrecherchen für alle drei Untersuchungsbereiche je eine Landscape- und Gap-Analyse erstellt (vgl. Kapitel 4). Diese Analysen wurden mit der Durchführung und Auswertung bereichsspezifischer, semistrukturierter Expert*innen-Interviews ergänzt (vgl. Kapitel 5). Um eine möglichst große Offenheit der Expert*innen zu erreichen, wurde a) vereinbart, die Auswertungen der Interviews im vorliegenden Bericht nicht konkret mit einzelnen interviewten Expert*innen zu assoziieren, und b) darauf verzichtet, die Transkripte der Interviews mit dem *EcoDM*-Bericht zu veröffentlichen. Die Transkripte wurden jedoch zur Ermöglichung von Nachprüfungen beim Deutschen Archäologischen Institut archiviert. Die Literaturliste und -analyse wie auch die Erhebung und Auswertung der Expert*innen-Interviews wurden an der Grounded Theory orientiert. Die im Projekt angewendeten Methoden werden in Kapitel 2 erläutert. Aus Vorgesprächen im Kontext der Konzeption des Forschungsvorhabens war deutlich geworden, dass die rechtlichen Rahmenbedingungen des Teilens und Nachnutzens von Daten für die Akteur*innen in den drei Untersuchungsbereichen von zentraler Bedeutung sind. Weil für deren Analyse im Projekt die Expertise

fehlte, wurde diese Fragestellung in Form eines externen Gutachtens in das Projekt integriert (vgl. Kapitel 3). Die Forschungsergebnisse des Projektes gliedern sich in a) die bereichsspezifische Darstellung typischer Herausforderungen im Umgang mit Daten sowie Interessen und Aspekte, die sich auf Entscheidungen zum Teilen und zur Nachnutzung von Daten auswirken (vgl. Kapitel 4 und 5) und b) die Zusammenstellung bereichsspezifischer (vgl. Kapitel 6) und bereichsübergreifender Empfehlungen (vgl. Kapitel 7). Sowohl die fast 100 bereichsspezifischen als auch die 31 bereichsübergreifenden EcoDM-Empfehlungen sind vollständig aus den Landscape- und Gap-Analysen sowie den Expert*innen-Interviews abgeleitet. Die Leistung des Forschungsprojektes EcoDM besteht in der Identifikation, Gliederung und teilweisen Abstraktion der Empfehlungen.

Die bereichsübergreifenden EcoDM-Empfehlungen waren Thema einer Podiumsdiskussion von vier Expert*innen, die im Rahmen der Abschlussveranstaltung des Projektes durchgeführt wurde (vgl. Kapitel 8). Aus der Perspektive der Initiator*innen des Forschungsprojektes EcoDM illustrieren die folgenden drei Schlaglichter die Einschätzungen, welche die bereichsspezifischen und bereichsübergreifenden EcoDM-Empfehlungen prägen:

- Das Bewusstsein für die Chancen, die sich aus der Nutzung von Daten ergeben können, ist noch unzureichend ausgebildet.

EcoDM empfiehlt, den benötigten Bewusstseinswandel durch praxisnahe Projekte zu fördern, die den Mehrwert der Nutzung von Daten illustrieren.

- Die Fähigkeit zum Sammeln und Auswerten großer Datenmengen birgt auch Risiken.

EcoDM empfiehlt, begründeten Vorbehalten konstruktiv zu begegnen.

- Es fehlt an finanziellen und personellen Ressourcen für die Ausgestaltung eines leistungsfähigen Datenökosystems.

EcoDM empfiehlt, bei der Förderung des Teilens von Daten korrespondierenden Geschäftsmodellen bzw. deren Entwicklung mehr Aufmerksamkeit zu schenken, damit aus Kosten Investitionen werden, sowie Qualifizierung in den Fokus zu nehmen.

Die in der Podiumsdiskussion diskutierten 31 bereichsübergreifenden EcoDM-Empfehlungen wurden zusammen mit ausgewählten Informationen zum Projektkontext in einer separaten Publikation mit dem Titel *„Data At Your Fingertips: Wegmarken einer exzellenten und vertrauenswürdigen Datenlandschaft“*¹ veröffentlicht.

¹ <https://doi.org/10.5281/zenodo.5837408>

Abbildungsverzeichnis

Abbildung 1:

Methodisches Vorgehen mit der Darstellung der einzelnen Prozessschritte (Quelle: EcoDM)

Abbildung 2:

Interview-Prozessdarstellung zur Visualisierung der zeitlichen Abfolge und des zeitlichen Aufwands auf Basis von 20 Interviews und 330 Seiten Interview-Transkripten (Quelle: EcoDM)

Abbildung 3:

Belastbarkeit des Wissensstandes (Quelle: DZHW Wissenschaftsbefragung 2019/20, n=6.946, nach: Ambrasat & Heger, 2020, S. 14.; Lizenz: CC BY-S. 3.0 DE, <https://creativecommons.org/licenses/by-sa/3.0/de/>)

Abbildung 4:

Open Data – Bedeutung und eigene Bereitstellungspraxis (Quelle: DZHW Wissenschaftsbefragung 2019/20, n=1.584, modularisiert, nach: Ambrasat & Heger, 2020, S. 33.; Lizenz: CC BY-SA 3.0 DE, <https://creativecommons.org/licenses/by-sa/3.0/de/>)

Abbildung 5:

Modell eines Datenlebenszyklus (Quelle: RfII – Rat für Informationsinfrastrukturen, 2019a, A-24; Lizenz: CC BY-SA 4.0, <https://creativecommons.org/licenses/by-sa/4.0/>)

Abbildung 6:

Struktur des Kodex „Leitlinien zur Sicherung guter wissenschaftlicher Praxis“ (Quelle: Deutsche Forschungsgemeinschaft, 2019, S. 2; Lizenz: CC BY 4.0 international, <https://creativecommons.org/licenses/by/4.0/>)

Abbildung 7:

Gesamtwerte der Data Maturity der 34 untersuchten Länder 2021 (Quelle: van Hesteren et al., 2022, S. 5; Lizenz: CC BY 4.0 international, <https://creativecommons.org/licenses/by/4.0/>)

Abbildung 8:

Entwicklung der Data Maturity Deutschlands 2021 (Quelle: *Country Fact-Sheet Germany* 2021, S. 2; Lizenz: CC BY 4.0 international, <https://creativecommons.org/licenses/by/4.0/>)

Abbildung 9:

Entwicklung der Data Maturity Deutschlands 2019 (Quelle: *Country Fact-Sheet Germany* 2019, S. 2; Lizenz: CC BY 4.0 international, <https://creativecommons.org/licenses/by/4.0/>)

Abbildung 10:

Ergebnisse Deutschlands in den vier Untersuchungsschwerpunkten 2021 (Quelle: *Country Fact-Sheet Germany* 2021, S. 3; Lizenz: CC BY 4.0 international, <https://creativecommons.org/licenses/by/4.0/>)

Abbildung 11:
Ergebnisse Deutschlands in den vier Untersuchungsschwerpunkten 2019 (Quelle: *Country Fact-Sheet Germany* 2019, S. 3; Lizenz: CC BY 4.0 international, <https://creativecommons.org/licenses/by/4.0/>)

Abbildung 12:
Darstellung der an GovData beteiligten Bundesländer (Quelle: EcoDM)

Abbildung 13:
Schlagworte, die einer FAIRen Datennutzung entgegenstehen und somit die Auffindbarkeit, Zugänglichkeit, Interoperabilität und Wiederverwendbarkeit erschweren (Quelle: EcoDM)

Abbildung 14:
Schematische Darstellung zum Einsatz von Datenintermediären (Quelle: Thomas et al., 2021, S. 6; Copyright: Christiane Wendehorst)

Abbildung 15:
Cloud-Risiken (Quelle: Gentemann & Heidkamp, 2021, S. 10; Copyright: Bitkom Research im Auftrag von KPMG)

Abbildung 16:
Dimensionen eines Geschäftsmodells (Quelle: *Lernende Systeme – Die Plattform für künstliche Intelligenz*, 2020, S. 7; Copyright: Plattform Lernende Systeme)

Abbildung 17:
Vergleichende Darstellung der 100 wichtigsten Plattformen weltweit nach Börsenwert 2021 und Zugehörigkeit zu Weltregion sowie nach Anteil am Gesamtwert in Prozent; Stand Dezember 2021 (Quelle: Schmidt & Hosseini, 2021, unveröffentlicht; Copyright: Dr. Holger Schmidt)

Abbildung 18:
Data Professionals Skills Gap (Quelle: European Commission. Directorate General for Communications Networks, Content and Technology. et al., 2020, S. 17; Lizenz: CC-BY 4.0 international, https://creativecommons.org/licenses/by/4.0)

Abbildung 19:
FAIR4S Core Skills Groups (Quelle: Whyte & Ashley, 2017, S. 43; Lizenz: CC-BY 4.0 international, https://creativecommons.org/licenses/by/4.0)

Abbildung 20:
FAIR4S Key Skills (Quelle: *Key Skills – EOSC FAIR4S*. 2019; Lizenz: CC-BY 4.0 international, https://creativecommons.org/licenses/by/4.0)

Abbildung 21:
FAIR4S Skills Tabelle – Data Advisors und Data Service Providers (Quelle: Whyte et al., 2019, S. 80, Lizenz: CC-BY 4.0 international, https://creativecommons.org/licenses/by/4.0)

Abbildung 22:
Auswahl von Stakeholdern für die Interview-basierte Analyse als zweite Untersuchungsphase (Quelle: EcoDM)

Tabellenverzeichnis

Tabelle 1:

Übersicht der Poster-Beiträge zum und über das Projekt EcoDM auf themenverwandten Veranstaltungen (Quelle: EcoDM)

Tabelle 2:

Übersicht über die von EcoDM durchgeführten Veranstaltungen (Quelle: EcoDM)

Tabelle 3:

Übersicht über die von EcoDM übersetzten wichtigen Schlüsseldokumente ins Deutsche (Quelle: EcoDM)

Tabelle 4:

Bundesländer mit Open-Access-Strategien. In der Tabelle sind die Open-Access- bzw. die Open-Science-Strategien der Bundesländer aufgelistet, sofern vorhanden (Quelle: EcoDM)

Akronym- und Abkürzungsverzeichnis

A

B

B2B: Business-to-Business

B2C: Business-to-Consumer

B2G: Business-to-Government

BAköV: Bundesakademie für öffentliche Verwaltung

BMBF: Bundesministerium für Bildung und Forschung

BMWi: Bundesministerium für Wirtschaft und Energie (seit Dezember 2021 BMWK: Bundesministerium für Wirtschaft und Klimaschutz)

C

CARE-Prinzipien: Collective Benefit, Authority to Control, Responsibility und Ethics

CC-Lizenz: Creative-Commons-Lizenz

CCOD: Kompetenzzentrum Open Data

D

DANS: Data Archiving and Network Services in den Niederlanden

DDG: Gesetz über Digitale Dienste

DFG: Deutsche Forschungsgesellschaft

DGG: Daten-Governance-Gesetz

DM: Datenmanagement

DMG: Gesetz über digitale Märkte

DMP: Datenmanagementplan

DNG: Datennutzungsgesetz

DOI: Digital Object Identifier

DSGVO: Datenschutzgrundverordnung

E

EFTA-Staaten: European Free Trade Association; dt.: Europäische Freihandelsassoziation

EGovG: E-Government-Gesetz

EOSC: European Open Science Cloud

EU: Europäische Union

EU-HLEG: European Commission – High Level Expert Group

F

FAIR: Findable, Accessible, Interoperable, Reusable

FAIR + Q: kuratierte Datensammlungen auf Grundlage der FAIR-Prinzipien

FDM: Forschungsdatenmanagement

FES: Friedrich-Ebert-Stiftung

FRAND: Fair, Reasonable And Non-Discriminatory

G

GAFAs: Google, Apple, Facebook und Amazon (US-amerikanischen Internetkonzerne)

GeschGehG: Gesetz zum Schutz von Geschäftsgeheimnissen

GDI(-DE): Geodateninfrastruktur Deutschland

GovData: nationales Metadaten-Portal für öffentliche Daten von Bund, Ländern und Kommunen

GWB: Gesetz gegen Wettbewerbsbeschränkungen

H

h-Index: Hirsch-Index, Hirschfaktor, Hirsch-Koeffizient (Publikationsfaktor)

I

ICS: International Science Council; dt.: Internationaler Wissenschaftsrat

IFG: Informationsfreiheitsgesetz

INSPIRE: Infrastructure for spatial information in Europe; dt.: Europäische Geodateninfrastruktur für eine gemeinschaftliche Umweltpolitik

IPR: Intellectual Property Rights; dt.: Rechte am geistigen Eigentum

J

K

KI: Künstliche Intelligenz

KMU: Kleine und mittlere Unternehmen

L

LGA: Landscape- und Gap-Analyse(n)

M

N

NFDI: Nationale Forschungsdateninfrastruktur

O

P

PID: Persistent Identifier; dt.: Persistente Kennung (Bsp.: DOI)

PSI-Richtlinie: EU-Richtlinie 2003/98/EG über die Weiterverwendung von Informationen des öffentlichen Sektors; engl.: Public Sector Information

Q

R

RDA DE: Research Data Alliance Deutschland

RfII: Rat für Informationsinfrastrukturen

S

T

TVL: Tarifvertrag der Länder

TVöD: Tarifvertrag für den öffentlichen Dienst

TRUST-Prinzipien: Transparency, Responsibility, User Focus, Sustainability, Technology

U

UrhG: Urheberrechtsgesetz

URL: Uniform Resource Locator; dt.: einheitlicher Ressourcenanzeiger

V

W

X

Y

Z

EINLEITUNG



1. Einleitung

1.1 Hintergrund und Ziele

Digitale Daten werden in großen und rasant steigenden Mengen erzeugt. In Zusammenhang mit den technischen Entwicklungen, die dieses Wachstum verursachen bzw. ermöglichen, hat sich auch die Leistungsfähigkeit zur Übermittlung und Analyse von Daten gesteigert (vgl. Europäische Kommission, 2020a). Zugriffsmöglichkeiten auf Daten bzw. deren Nutzung werden deshalb von vielen Ökonom*innen als wichtiger, eventuell sogar als wichtigster Faktor für den Erfolg von Unternehmen und Volkswirtschaften gesehen („The world’s most valuable resource is no longer oil, but data“, 2017). In der Wissenschaft ist die Erzeugung und Nutzung digitaler Daten mit der Vorstellung verbunden, daraus neues Wissen generieren zu können (vgl. Kapitel 4.1). In der öffentlichen Verwaltung kann die Digitalisierung dazu beitragen, Prozesse effizienter zu gestalten (vgl. Kapitel 4.2). Open Data führt zu mehr Transparenz und die Bereitstellung von Daten kann Impulse für die Wirtschaft geben (vgl. Kapitel 4.3). Aktuell wird in der breiten Öffentlichkeit insbesondere die schnelle Entwicklung von Impfstoffen gegen Corona-Infektionen als Beleg für die Sinnhaftigkeit des Teilens von Daten wahrgenommen. Zudem kommt es durch die Pandemie zu einem nicht absehbaren Vorschub in der Digitalisierung in allen Bereichen. Mit der Förderung des Teilens und Nachnutzens von Daten ist die Hoffnung verbunden, die Problemlösungsfähigkeiten in der Wissenschaft zu erhöhen, Verwaltungsprozesse zu optimieren sowie das Wachstum der Volkswirtschaft zu steigern.

Daten werden von unterschiedlichen Akteur*innen in verschiedenen Kontexten produziert und genutzt. Neben der aktiven, zielgerichteten Erzeugung fallen beispielsweise digitale Daten auch beiläufig, als Nebenprodukt, an. Das geschieht unter anderem beim Surfen im Internet oder durch die Digitalisierung von Prozessen, wie etwa bei der Navigation im Auto oder der digitalen Unterstützung des Fahrens mit Geschwindigkeitsreglern oder Bremsassistenten. Der Einbau von digitaler Technologie in Autos ist auch ein gutes Beispiel für die Entstehung neuer Konfliktfelder: Wer soll auf Daten, die durch das Zusammenwirken vieler Akteur*innen entstehen, zugreifen bzw. über deren Nutzung entscheiden dürfen? Beim Beispiel Auto konkurrieren Zulieferer, Hersteller, Dienstleister und nicht zuletzt die Besitzer*innen der Autos um die Daten.

Mega-Unternehmen wie Alphabet (Google), Amazon, Apple oder Meta (Facebook) schöpfen von ihren Nutzer*innen fortlaufend riesige Mengen an Nutzungsdaten ab. Dies wirft sowohl persönlichkeitsrechtliche als auch wettbewerbsrechtliche Fragen auf (vgl.

Kapitel 3). Hier wird deutlich, wie stark Szenarien der Nutzung und des Teilens von Daten auch von gesetzlichen Regelungen beeinflusst werden. Bestimmte Unternehmen wie Internetprovider, E-Mail-Services oder Cloud-Anbieter sind technisch in der Lage, bei ihnen durchlaufende und abgelegte Daten für eigene Zwecke auszuwerten und zu speichern. Die Sicherheitsbehörden – einschließlich der Geheimdienste – verfügen ebenfalls über diese Möglichkeit, sei es durch eigene technische Anlagen oder über Zugriffsrechte, die Unternehmen zur Kooperation mit ihnen verpflichten. Die legalen Möglichkeiten von Unternehmen zum Abschöpfen von Kund*innendaten sowie die technischen Möglichkeiten zum illegalen Zugriff und zur Übernahme der Kontrolle über das Internet haben bei vielen Akteur*innen den Ruf nach der Schaffung vertrauenswürdiger Datenräume aufkommen lassen (vgl. Europäische Kommission, 2020a). Auf die Kritik am legalen Abgreifen von personenbezogenen Daten haben Gesetzgeber weltweit durch den Beschluss von Regelungen zum Datenschutz reagiert (vgl. Kapitel 3). Bisher ist es jedoch nicht gelungen, Konflikte um ein angemessenes Maß an Datenschutz zu befrieden. Dies gilt auf nationaler wie auf internationaler Ebene. Zudem variieren die gesetzlichen Regelungen zum Datenschutz international erheblich. Dies wird z. B. durch den seit Jahren andauernden Konflikt um das legale Übertragen von Daten zwischen der EU und den USA illustriert. Auf die Problematik illegaler Angriffe auf digitale Systeme und Netzwerke hat die Bundesregierung z. B. mit dem Beschluss der „Cybersicherheitsstrategie für Deutschland 2021“ reagiert (Bundesministerium des Innern, für Bau und Heimat, 2021). Trotz fortlaufenden Mahnungen zur Erhöhung der Cybersicherheit und ungeachtet des Beschlusses entsprechender Strategien sind Cyberangriffe nicht nur alltäglich, sondern vielfach auch erfolgreich (Engels, 2021). Die Bereitschaft zum Teilen bzw. Nachnutzen von Daten hängt deshalb wesentlich von der Existenz vertrauenswürdiger Datenräume ab. Dies gilt insbesondere, wenn die Bereitschaft zum Teilen von Daten von der technischen Absicherung zur Einhaltung der an die Datennutzung geknüpften Bedingungen bestimmt wird.

Neben den oben angesprochenen rechtlichen und sicherheitstechnischen Rahmenbedingungen hängt die Bereitschaft zum Teilen von Daten bzw. das Interesse zur Nutzung angebotener Daten auch davon ab, ob die Daten mit vorhandener Software gelesen bzw. bearbeitet werden können und ob die Daten um Informationen zu ihrem semantischen Gehalt und den Umständen ihrer Erhebung ergänzt werden. Diese zusätzlichen Informationen werden über beschreibende

Daten, sogenannte Metadaten, gegeben. Daher werden die Möglichkeiten zur Nachnutzung von Daten z. B. stark von der Verwendung verbreiteter oder offener Formate und der Bereitstellung von umfangreichen Metadaten beeinflusst. Vielfach werden keine offenen Formate genutzt und die Auswahl der Metadaten orientiert sich an Gewohnheiten sowie Vorgaben spezifischer Nutzer*innen-Gruppen (z. B. Forscher*innen im selben Forschungsfeld, Verwaltungseinheiten mit gleichen und ähnlichen Aufgabenstellungen oder Unternehmen einer Branche), so dass diese Metadaten für Dritte weniger oder gar nicht verständlich sind. Ein zentraler Aspekt der Überlegungen, welchen Daten ein zentraler Wert für Wirtschaft und Gesellschaft zugemessen wird, ist die Hoffnung, dass Daten unabhängig vom Kontext ihrer Erzeugung nachgenutzt werden können. Dafür ist es erforderlich, Metadaten möglichst weitgehend zu standardisieren. Wo dies nicht der Fall ist, wird eine Nachnutzung von Daten über den originalen Verwendungszweck hinaus deutlich erschwert.

Eine weitere Herausforderung ist die Gewährleistung der Auffindbarkeit der Daten, die für eine Nachnutzung bereitgestellt werden. Neben den technischen Voraussetzungen wie beispielsweise der Zugänglichkeit zu Datenbanken sind auch hier Metadaten wichtig, da sie Informationen zum semantischen Gehalt von Daten geben können.

Mit der Einführung der FAIR-Prinzipien¹ wurde versucht, einen geordneten Umgang mit Daten zu gewährleisten. Die Abkürzung FAIR steht dabei für die Begriffe Findable, Accessible, Interoperable und Reusable (vgl. Kapitel 1.3.3). Die FAIRifizierung von Daten ist in vielen Fällen Voraussetzung für deren Nachnutzung ohne individuelle Erläuterung durch die Datenproduzent*innen. Dies gilt insbesondere, wenn Daten in Kontexten nachgenutzt werden sollen, die fachlich weit von denen der Erhebung entfernt liegen. Vor diesem Hintergrund ist das Ziel von EcoDM, Bedingungen zu identifizieren, die das Teilen und die Nachnutzung von Daten fördern bzw. verhindern. Darauf aufbauend werden für die vier verschiedenen Untersuchungsbereiche *Wissenschaft*, *Public Sector*, *Wirtschaft* und *Qualifizierung* spezifische (vgl. Kapitel 6) und bereichsübergreifende Handlungsempfehlungen (vgl. Kapitel 7) zum Fördern des Teilens und Nachnutzens von Daten formuliert. Adressat*innen der Empfehlungen sind die Gruppen der Akteur*innen der jeweiligen Untersuchungsbereiche sowie bereichsübergreifende Stakeholder aus Wissenschaft, Gesellschaft, Wirtschaft und Politik.

1.2 Die Untersuchungsbereiche

EcoDM konzentriert sich methodisch auf vier Untersuchungsbereiche, die aufgrund ihrer zentralen Bedeutung bei der Förderung des Teilens und Nachnutzens von Daten sowie dem Umgang mit Daten ausgewählt wurden: *Wissenschaft*, *Public Sector*, *Wirtschaft* und als Querschnittsthema *Qualifizierung*. Die vier Untersuchungsbereiche umfassen jeweils unterschiedliche Akteur*innen sowie Gruppen von Akteur*innen, die hier als „aktiv handelnd“ bzw. als „aktiv gestaltend“ zu verstehen sind. Sie zeichnen sich durch jeweils eigene Vernetzungs- bzw. Governance-Strukturen und gemeinsame Interessen aus. Neben bereichsspezifischen Akteur*innen werden in EcoDM auch bereichsübergreifende Stakeholder mit in die Ergebnisse eingebunden. Dies ergibt sich aus dem übergeordneten Ziel von EcoDM, die Bedarfe des digitalen Datenmanagements nicht nur für den Austausch innerhalb der Untersuchungsbereiche zu ermitteln, sondern auch für die Nutzung zwischen den verschiedenen Daten-produzierenden Akteur*innen. Stakeholder beschreiben Anspruchspersonen oder -gruppen (Röttger et al., 2018, S. 67) mit einem berechtigten Interesse am Organisationshandeln bzw. jene, die direkt davon tangiert werden (ebd., S. 7). Die Bestimmung der Stakeholder erfolgte entsprechend der gleichzeitigen Einbeziehung zweier Anspruchsgruppenkonzepte (Rüegg-Stürm, 2004, S. 65 – 141): einerseits das strategische Anspruchsgruppenkonzept nach R. E. Freeman (1984) und andererseits die Berücksichtigung von Stakeholdern entsprechend des normativ-kritischen Anspruchsgruppenkonzeptes nach P. Ulrich (2001). Demnach wurden sowohl Stakeholder entsprechend der Wirkmächtigkeit ihrer Anspruchsgruppe identifiziert als auch entsprechend ihrer „potentiellen oder faktischen“ Beeinflussung unabhängig ihrer de facto Wirkmächtigkeit.

Zu der Gruppe von Akteur*innen des Untersuchungsbereichs **Wissenschaft** zählen beispielsweise Forschungseinrichtungen in Deutschland, wie sie in der Allianz der Wissenschaftsorganisationen² zusammengeschlossen sind, aber auch Fachgesellschaften und Dachorganisationen wie unter anderem die Hochschulrektorenkonferenz³. Darüber hinaus gehören ebenso relevante deutsche und europäische Initiativen (wie NFDI⁴, EOSC⁵) sowie deren Vertreter*innen in Deutschland zu den Gruppen der Akteur*innen. Die Wissenschaft produziert eine große Mengen an Forschungsdaten, welche auch in den anderen Bereichen nachgenutzt werden können. Zur Generierung von

¹ <https://www.go-fair.org/fair-principles/>

² <https://www.leibniz-gemeinschaft.de/ueber-uns/strategie-und-wissenschaftspolitik/allianz-der-wissenschaftsorganisationen.html>

³ <https://www.hrk.de/>

⁴ <https://www.nfdi.de/>

⁵ <https://www.eosc-portal.eu/>

neuem Wissen hat die Wissenschaft großes Interesse an den Daten aus den anderen Bereichen.

Im Untersuchungsbereich **Public Sector** entstehen digitale Daten in allen Teilen der Gesellschaft (z. B. in der öffentlichen Verwaltung). Die Gruppe der Akteur*innen dieses Untersuchungsbereiches ist durch ihre Zugehörigkeit zu Gebietskörperschaften besonders klar strukturiert und zeichnet sich darüber hinaus durch spezifische Handlungsspielräume aus, die sich aus ihren hoheitlichen Aufgaben ergeben.

Der *Public Sector* wird auch öffentlicher Sektor oder umgangssprachlich öffentliche Hand genannt. Eine allgemein-verbindliche Definition des Begriffs *Public Sector* gibt es nicht. Im Rahmen von EcoDM wird der Begriff *Public Sector* wie folgt abgegrenzt: Er umfasst sogenannte haushaltsorientierte Gebietskörperschaften. Zu diesen gehören Bund, Länder, Kreise und Kommunen sowie Verbände dieser Gebietskörperschaften und außerdem Anstalten und Körperschaften des öffentlichen Rechts mit eigener Abgaben- und Steuerhoheit. Sie unterstehen „der staatlichen Aufsicht [...] und verwalten und organisieren sich auf selbstständiger Basis im Rahmen der Selbstverwaltung und Selbstorganisation“ („▷ Gebietskörperschaft“, o. J.). Zu den Einrichtungen des öffentlichen Sektors gehören Regierungen, Behörden und Verwaltungen auf Bundes- und Landesebene sowie öffentliche Unternehmen und Bildungs- wie auch Forschungseinrichtungen. Letztere werden im Rahmen von EcoDM dem Untersuchungsbereich *Wissenschaft* zugeordnet.

Die Gruppe von Akteur*innen im Untersuchungsbereich der **Wirtschaft** umfasst Unternehmen aller Größen und sämtlicher Entwicklungsphasen von kleinen Startups bis zu internationalen Konzernen, die ein wirtschaftliches Interesse an Daten haben und bei denen die Erhebung, Verarbeitung, Speicherung und Vermarktung sowie das Teilen und Nachnutzen von Daten an Geschäftsmodelle gebunden sind. Sie organisieren sich ähnlich wie die Wissenschaft über branchenspezifische Fachverbände, für die z. T. sogar ein Mitgliedszwang besteht, und über branchenübergreifende Zusammenschlüsse wie beispielsweise dem Bundesverband der Deutschen Industrie⁶ und dem Deutschen Industrie- und Handelskammertag⁷.

Im Fokus der Gruppe der Akteur*innen des Untersuchungsbereichs **Qualifizierung** stehen neben dem Qualifizierungsbedarf im Umgang mit Daten Angebote und Maßnahmen zu dessen Deckung in den Untersuchungsbereichen *Wissenschaft*, *Public Sector* und *Wirtschaft*. Als Akteur*innen sind hier zunächst Anbieter von Qualifizierungsangeboten zu nennen. Zu ihnen zählen in erster Linie Einrichtungen im Bereich Aus- und Weiterbildung (z. B. Hochschulen, Anbieter

von Weiterbildungen und ihre Dachorganisationen) wie auch die in ihnen oder unabhängig von diesen in der Vermittlung tätigen Personen. Eine weitere Gruppe von Akteur*innen bilden alle Adressat*innen von Qualifizierungsangeboten im Rahmen von beruflichen Ausbildungen, von Hochschulstudiengängen und von Weiterqualifizierungen im Beruf. Zum Untersuchungsbereich *Qualifizierung* gehören aus der Perspektive von EcoDM außerdem alle Akteur*innen, die Rahmenbedingungen in diesem Bereich durch (gesetzliche) Vorgaben auf nationaler und internationaler Ebene bestimmen (z. B. Regierungen und Kommissionen) oder anderweitig Einfluss nehmen (z. B. Fachgesellschaften, Berufsverbände und ähnliche Vereinigungen).

1.3 Definitionen

Dieses Kapitel beinhaltet Definitionen der für das Projekt EcoDM zentralen Begriffe: *Daten*, *Daten-Ökosystem* und *FAIR-Prinzipien*. Diese Begriffe werden in der ausgewerteten Literatur, im Internet bzw. von verschiedenen Akteur*innen und Stakeholdern mit teilweise unterschiedlichen Bedeutungen verwendet. Die hier gegebenen Definitionen sollen klarstellen, mit welchem Bedeutungsgehalt die Begriffe in diesem Bericht verwendet werden.

1.3.1 Daten

Daten lassen sich definieren als Informationen bzw. als Zeichen, die Informationen darstellen. Laut internationaler Normung in der Informationstechnologie sind sie die „interpretierbare Darstellung von Informationen in formalisierter Art, geeignet zur Kommunikation, Interpretation oder Verarbeitung“ (ISO/IEC 2382, 2015, Def. 2121272). Dort werden digitale Daten zusätzlich definiert als „durch Ziffern dargestellte Daten, möglicherweise zusammen mit Sonder- und Leerzeichen“ (ebd., Def. 2121038), was bedeutet, dass Informationen maschinenlesbar beschrieben werden. Für die EU-Kommission sind Daten „jede digitale Darstellung von Handlungen, Tatsachen oder Informationen sowie jede Zusammenstellung solcher Handlungen, Tatsachen oder Informationen auch in Form von Ton-, Bild- oder audiovisuellem Material“ (Europäische Kommission 2020b, Art. 2, Pkt. 1). Analoge Daten sind nicht Teil dieser Betrachtung und werden daher in der Definition nicht berücksichtigt.

Was Daten genau sind, ist letztlich von dem Kontext abhängig, für den oder in dem sie erhoben, verarbeitet und archiviert werden. Für das Projekt EcoDM

⁶ <https://bdi.eu/>

⁷ <https://www.dihk.de/>

lassen sich Daten als diejenigen digital vorliegenden Informationen definieren, mit denen in den Untersuchungsbereichen gearbeitet wird bzw. die für die Akteur*innen in den Untersuchungsbereichen Wert haben. Nach dieser Definition sagen Begriffe wie Forschungs-, Verwaltungs- oder Wirtschaftsdaten nichts über die so bezeichneten Daten aus, sie verweisen lediglich auf den Kontext, in welchem diese Daten genutzt werden. Daten können gleichzeitig Forschungs-, Verwaltungs- und Wirtschaftsdaten sein. Je nachdem, wo diese Daten frei verfügbar vorliegen, spricht man von Open Research Data, Open Government Data und Open Business Data (vgl. Putnigs, 2021). Offene Daten (Open Data) sind „nach dem allgemeinen Verständnis Daten in einem offenen Format, die von allen zu jedem Zweck frei verwendet, wieder verwendet und weitergegeben werden können“ (Richtlinie (EU) 2019/1024, 2019, (16)).

In der *Wissenschaft* werden unter Daten in der Regel Forschungsdaten verstanden. Sie dienen in diesem Zusammenhang als Beleg für Forschungsergebnisse. Die DFG bezeichnet als Forschungsdaten „u. a. Messdaten, Laborwerte, audiovisuelle Informationen, Texte, Surveydaten oder Beobachtungsdaten, methodische Testverfahren sowie Fragebögen. Korpora, Software und Simulationen können ebenfalls zentrale Ergebnisse wissenschaftlicher Forschung darstellen und werden daher ebenfalls unter den Begriff Forschungsdaten gefasst“ (Deutsche Forschungsgemeinschaft, 2021, S. 1). Forschungsdaten entstehen also im gesamten Forschungsprozess und durch standardisierte Verfahren auch darüber hinaus. Außerdem kommen alle qualitätsgesicherten Daten für eine Nachnutzung infrage, unabhängig davon, ob sie gleichzeitig eine Funktion als Belege für Forschungsergebnisse haben. Diese Daten werden nicht immer veröffentlicht bzw. zur Verfügung gestellt. Sie sind daher nur eingeschränkt für andere zugänglich und nachnutzbar.

Die Daten des *Public Sector* sind breit gefächert und stammen aus verschiedenen Bereichen des gesellschaftlichen Lebens oder haben einen direkten Bezug dazu. Dazu zählt die EU beispielsweise „Soziales, Politik, Wirtschaft, Recht, Geographie, Umwelt, Meteorologie, Seismizität, Tourismus, Geschäftsleben, Patentwesen und Bildung“ (Richtlinie (EU) 2019/1024, 2019, (8)). Daten werden bei den „öffentlichen Stellen der Exekutive, Legislative oder Judikative“ auf mehreren Verwaltungsebenen wie der kommunalen, regionalen oder nationalen Ebene generiert, gesammelt und verarbeitet (ebd.). Die EU sieht diese Daten als einen „umfassenden, vielfältigen und wertvollen Fundus an Ressourcen, der der Gesellschaft zugutekommen kann. Die Bereitstellung dieser Informationen [...] in einem gängigen elektronischen Format ermöglicht es Bürgern und juristischen Personen, neue Möglichkeiten für deren Nutzung zu finden und neue, innovative

Produkte und Dienstleistungen zu schaffen“ (ebd.). Seitens der EU werden ihre Mitgliedsstaaten dazu angehalten, Daten aus dem *Public Sector* „nach dem Grundsatz ‚konzeptionell und standardmäßig offen‘ (open by design and by default)“ zu behandeln (ebd., (16)).

Beispiele für Daten der *Wirtschaft* sind Daten, die typischerweise durch Unternehmen generiert werden und damit all jene Daten, welche einen ökonomischen Mehrwert, sowohl für einzelne Unternehmen als auch den gesamten Wirtschaftsraum, haben können. Darunter fallen alle Daten, die von einem Wirtschaftsunternehmen erhoben werden, z. B. Kund*innen-, Nutzer*innen- und Prozessdaten. Darüber hinaus spielen auch die Daten aus der *Wissenschaft* und dem *Public Sector*, die im Rahmen von wirtschaftlichem Handeln von Interesse sind, eine wichtige Rolle. Erhebung, Verarbeitung, Verwendung, Speicherung sowie Verbreitung von Daten werden durch die jeweiligen Geschäftsmodelle bestimmt. Sie bilden die Stützpfiler eines gewinnorientierten Unternehmens. So sieht das BMWK Unternehmen „vor der Herausforderung und Chance, eine immer größer werdende Menge an Daten effizient einzusetzen“ und postuliert: „Daten sind im digitalen Zeitalter der Rohstoff für wirtschaftliche Wertschöpfung – ihre intelligente Nutzung ist die Quelle von Innovation und Wachstum“ (Bundesministerium für Wirtschaft und Klimaschutz, o. J.).

1.3.2 Daten-Ökosystem

Beim Daten-Ökosystem handelt es sich um ein organisatorisches und koordinatives Netzwerk rund um die Ressource ‚Daten‘, das aus verschiedenen Nutzungs- und Interessensgruppen besteht, die ein gemeinschaftliches Ziel verfolgen und dabei komplementäre Fähigkeiten wie auch Kompetenzen insgesamt verstärken (Lis et al., 2019, S. 12).

Der Begriff Ökosystem verweist auf das komplexe Zusammenspiel einer größeren Anzahl von Akteur*innen mit jeweils spezifischen Fähigkeiten, das den einzelnen Akteur*innen Handlungsspielräume eröffnet, die sie autonom agierend nicht hätten und welches der Umwelt dieser Akteur*innen Stabilität verleiht. Im Gegenzug bedeutet dies, dass der Wegfall einzelner Akteur*innen einen Zusammenbruch des gesamten Ökosystems zur Folge haben kann. Je höher die Bedeutung solcher Daten-Ökosysteme eingeschätzt wird, desto stärker muss diese systemische Perspektive Orientierungspunkt für die Formulierung von Maßnahmen zur Förderung des Teilens bzw. Nachnutzens von Daten werden.

Daten stellen grundsätzlich eine gemeinsame Ressource dar, weil man sie beliebig oft teilen kann. Sie sind die Basis für Informationen und Wissen, fungieren aber auch als Ware. Das Ökosystem bietet die Infrastruktur, in der diverse Akteur*innen zusammenkommen,

um diese Daten „zu produzieren, anzubieten, zu finden und zu ‚konsumieren‘ (d. h. nachzunutzen, zu verarbeiten, anzureichern, zu archivieren, zu publizieren, Entscheidungen darauf zu fällen etc.)“ (Putnings, 2021, S. 7). Neben dieser Infrastruktur bietet das System aber auch Anwendungen zur Sammlung und Auswertung von Informationen, z. B. für wissenschaftliche oder wirtschaftliche Zwecke. So entstehen Schnittstellen zum Datenaustausch oder gemeinsamen Arbeiten, wodurch einzelne Akteur*innen strategische Kooperationen eingehen können. Neben der Zusammenarbeit lassen sich auch datenbasierte Innovation, Datenkompetenz und digitale Souveränität sowie Wertschöpfungs- und Einsparungspotential als übergeordnete Ziele festmachen, die eng mit einem Daten-Ökosystem verknüpft sind.

Das Ökosystem Daten muss resilienter werden. Es befindet sich im stetigen Wandel, wobei durch deutsche und europäische Projekte große Verbesserungen für die kommenden Jahre geplant sind. Die Gestaltung des Ökosystems kann nie vollständig abgeschlossen sein, weil es mit der Zeit kontinuierlich verbesserungswürdig bleibt (vgl. Wissenschaftsrat, 2021, S. 65 ff.). So schreibt beispielsweise auch Putnings, dass „die Funktionsfähigkeit, d.h. die supportiven Fähigkeiten des Daten-Ökosystems zur Produktion und (Nach-)Nutzung von Daten sowie der potentiellen und konkreten Wertschöpfung daraus“ von den „politischen und förderrechtlichen Maßnahmen“, dem „technologischen Umfeld“, den „rechtlichen Rahmenbedingungen“, den Vorgaben durch das „vorherrschende ethische und Wertesystem“ sowie den „Akteuren selbst und deren Wertschöpfungsmöglichkeiten bzw. -aktivitäten“ abhängt (Putnings 2021, S. 8 f.). In Europa verfolgt man einen wertebasierten Ansatz, in dem unter anderem Freiheit und ein selbstbestimmter Umgang mit Daten in den Vordergrund gerückt werden.

1.3.3 FAIR-Prinzipien

Die Abkürzung FAIR steht für die Begriffe Findable (auffindbar), Accessible (zugänglich), Interoperable (interoperabel) sowie Reusable (wiederverwendbar)⁸ und meint, dass Daten und die zugehörigen Metadaten über eben diese Eigenschaften verfügen sollten, um einen universellen Austausch und damit eine erweiterte Nutzung sowohl für Menschen als auch Computersysteme zu ermöglichen.

Es handelt sich bei den FAIR-Prinzipien um Grundsätze des Datenmanagements. Sie sollen einen einfacheren und effizienteren interdisziplinären Zugriff auf und Austausch von Daten ermöglichen. Des Weiteren sind sie als eine mögliche Antwort auf die Replikationskrise,

d. h. der nicht möglichen Reproduzierbarkeit von Ergebnissen, in der Wissenschaft zu verstehen, die durch die Digitalisierung noch deutlicher sichtbar wurde. Damit ist die Einhaltung der FAIR-Prinzipien ein neues Kriterium für die Qualität von Daten, das sich europaweit etabliert.

Die zunehmende Notwendigkeit, einheitliche Standards zu definieren, um das Teilen und Nutzen von Daten zwischen verschiedenen Interessens- und Nutzungsgruppen zu erleichtern, kann als Grund für die Erarbeitung der Prinzipien genannt werden. Zudem sollten im Rahmen eines effizienteren Datenmanagements die Datenbestände selbst über Repositorien als Quellen für Innovationen und Entwicklungen nutzbar gemacht werden. Dabei sollten beim Umgang mit Daten sowohl Produzent*innen als auch Nutzer*innen und die mit ihrer Verwaltung befassten Akteur*innen von Anfang an mit einbezogen werden (Wilkinson et al., 2016). Erstmals postuliert wurde die Abkürzung bzw. das Konzept FAIR im Januar 2014 auf einem Workshop des Lorentz Centers an der Universität Leiden.⁹ Nach weiterer Präzisierung erfolgte die Veröffentlichung des Konzepts im März 2016 unter dem Titel „The FAIR Guiding Principles for scientific data management and stewardship“ (ebd.). Seitdem konnten sich die Prinzipien auch über den wissenschaftlichen Kontext hinaus etablieren und ihre Anwendung wird von Deutschland und der EU für alle Bereiche des Daten-Ökosystems angestrebt. So sind sie seit 2019 beispielsweise auch im Kodex „Leitlinien zur Sicherung guter wissenschaftlicher Praxis“ der DFG verankert (Deutsche Forschungsgemeinschaft, 2019).

Die einzelnen Aspekte der FAIR-Prinzipien lassen sich wie folgt beschreiben:

Um die Auffindbarkeit (Findability) von Daten sicherzustellen, müssen sie durch einen Persistent Identifier (PID) – z. B. einen Digital Object Identifier (DOI) – eindeutig und dauerhaft zugeordnet werden können. Darüber hinaus sollten Daten mit einer ausreichenden Menge an Metadaten beschrieben werden, die einen Persistent Identifier einschließen. Zudem sollten die Metadaten in einer durchsuchbaren Ressource registriert sein. Zentral für Findability ist ‚Eindeutigkeit‘, denn nur wenn Daten klar und unmissverständlich zugeordnet werden können, lassen sie sich auch problemlos finden (Wilkinson et al., 2016 und PARTHENOS et al., 2019).

Zugänglichkeit (Accessibility) wird verbessert, wenn die Daten und ihre Metadaten an einem Aufbewahrungsort langfristig und frei verfügbar gehalten werden und durch ein standardisiertes Kommunikationsprotokoll abrufbar sind, das möglichst offen, frei sowie universell einsetzbar ist. Zudem sollte das Protokoll

⁸ <https://www.go-fair.org/fair-principles/>

⁹ <https://www.lorentzcenter.nl/jointly-designing-a-data-fairport.html>

gegebenfalls ein Authentifizierungs- und Autorisierungsverfahren ermöglichen, wenn einzelne Datensätze Zugangsbeschränkungen unterliegen (müssen). Darüber hinaus ist darauf zu achten, dass die Metadaten auch dann noch zugänglich sind, wenn die Verfügbarkeit der Daten selbst nicht mehr gegeben ist. (Wilkinson et al., 2016 und PARTHENOS et al., 2019).

Interoperabilität (Interoperability) lässt sich am besten erzielen, wenn die Daten sowohl von Menschen als auch Computersystemen einfach und möglichst automatisiert mit anderen Datensätzen verknüpft werden können. Dies ist dann der Fall, wenn die Metadaten eine klare, gemeinsame und breit anwendbare Sprache nutzen und die anderen FAIR-Prinzipien berücksichtigt werden. Die Verwendung von Standards erleichtert den interdisziplinären Austausch von Daten (Wilkinson et al., 2016 und PARTHENOS et al., 2019).

Wiederverwendbarkeit (Reusability) von Datensätzen kann durch präzise Dokumentation und Beschreibung ermöglicht werden. Die Metadaten sollten mit einer Vielzahl von genauen und relevanten Attributen

ausgestattet sein und zumindest fachspezifische Standards erfüllen. Eine eindeutige Lizenzierung – z. B. eine Creative Commons (CC) Lizenz¹⁰ – und klare Benennung der Daten erleichtert es, sie zu späteren Zeitpunkten nach- oder weitzunutzen zu können (Wilkinson et al., 2016 und PARTHENOS et al., 2019).

Für das Projekt EcoDM bilden die FAIR-Prinzipien eine wichtige Grundlage, weil durch ihre konsequente Anwendung eine Basis zum vernetzten Arbeiten und Datenaustausch geschaffen wird. So werden sie zu einem Schlüsselthema für die Entwicklung von Daten-Ökosystemen (vgl. Kapitel 1.3.2). Außerdem werden sie bei der Interview-basierten Analyse der einzelnen Untersuchungsbereiche *Wissenschaft*, *Public Sector*, *Wirtschaft* und *Qualifizierung* berücksichtigt und tragen so zur Formulierung der Empfehlungen bei. Darüber hinaus werden die FAIR-Prinzipien auch auf das Projekt selbst angewandt und dementsprechend die eigenen Daten in einem Datenmanagementplan dokumentiert.

¹⁰ <https://creativecommons.org/licenses/?lang=de>

METHODIK



2. Methodik

Die Methodik des Forschungsteams orientiert sich am pragmatisch veranlagten Methodenrepertoire der Grounded Theory (Strübing, 2014). Kodierungen bzw. Codes sind in diesem Kontext Codes, Kommentare, Exzerpte, Notizen und Memos, die bei der Bearbeitung des Datenmaterials und in diskursiven Prozessen entstehen. „Die Leitidee des Kodierprozesses ist die Methode des ständigen Vergleichens (‘constant comparative method’) der Daten miteinander“ (Strübing, 2014, S. 15). Offene Codes wirken generativ auf den Theoriebildungsprozess und dienen zur ersten Einordnung. Axiale Codes stellen Zusammenhänge zwischen Kodierungen her, während selektive Codes komplexe, oft multidimensionale Zusammenhänge zwischen dem Datenmaterial in Bezug auf die Kernkategorien beschreiben (vgl. Strübing, 2014, S. 16 f.). Durch „theoretisches Sampling“ ist man hinsichtlich Literaturswahl, bei der Auswahl der Interviewpartner*innen und im Rahmen der Erstellung der Interviewleitfäden und Empfehlungen um entsprechende breite sowie multidimensionale „konzeptuelle Repräsentativität“ bemüht: Gegensätzliches wird aktiv in den Fokus genommen und auf Gemeinsamkeiten überprüft (vgl. Strübing, 2014, S. 31). Dieser Prozess wird wiederholt, bis kein neuer Aspekt mehr hinzugefügt werden kann und so eine theoretische Sättigung erreicht wird (vgl. ebd.). Die Methoden der Grounded Theory sind „weniger als präskriptives ‚Verfahren‘ [...] dem haargenau zu folgen wäre“, sondern als „konzeptuell verdichtete, methodologisch begründete und in sich konsistente Sammlung von Vorschlägen“ konzipiert, die „sich für die Erzeugung gehaltvoller Theorien über sozialwissenschaftliche Gegenstandsbereiche als nützlich erwiesen haben“ (Strübing, 2014, S. 5). Ihr Kern ist ein zyklisches, interpretatives Vorgehen, in welchem Forschende „nie allein neutrale Beobachter, sondern zwangsläufig [...] Interpretieren ihrer Daten und als Entscheider über den konkreten Gang der theoretischen Argumentation immer auch Subjekte des Forschungsprozesses“ sind (ebd., S. 12). Dieses zunächst liberal wirkende Methodenverständnis darf aber „nicht als Freibrief für ein ‚anything goes‘ [...] missverstanden werden“, sondern die „Methode des ständigen Vergleichens“ und Kategorisierens führt vielmehr zu einer zyklischen und mehrstufigen empirischen Auswertung (vgl. ebd.). Grundsätzlich gilt mit Blick auf eine mögliche „Überbetonung des induktivistischen Erkenntnismodus in der Selbstdarstellung der Grounded Theory“ und damit „mangelnde Berücksichtigung der Rolle von Kreativität im Forschungsprozess“ sowie „die ambivalente Beurteilung des Stellenwerts von theoretischem Vorwissen“ (ebd., S. 55): „Jeder Versuch, theoretische Konzepte allein aus den Daten emergieren zu lassen,

kann letztendlich nur dazu führen, dass man hilflos einer großen Menge unstrukturierter Datenmaterials gegenübersteht“ (Kelle & Kluge, 2010, S. 21). Um dem entgegenzuwirken, wird im methodischen Zugang dieses Projekts das „Vorwissen“ des Forschungsteams integrativ berücksichtigt und die Menge des Materials durch Codes auf strukturiertes Datenmaterial reduziert. In der entsprechenden Strukturierung und Fokussierung der Inhalte wird diskursiv mit dem Material und Vorwissen gearbeitet. Dabei ist mit Kelle und Kluge (ebd., S. 26) festzuhalten, dass die Formulierung und Prüfung guter Schlussfolgerungen sowohl vom Vorwissen abhängt als auch von Offenheit und dementsprechend der Fähigkeit, fortlaufend das gesamte bisherige Wissen zu hinterfragen, bis sich ein entsprechender Sättigungseffekt einstellt, ab dem aus dem Material keine weiteren Erkenntnisse gewonnen werden können.

Das methodische Vorgehen wird im Laufe des Forschungsprozesses kontinuierlich an das Forschungsziel – den bereichsübergreifenden Empfehlungen – angepasst. Dies ermöglicht die präzise Ausarbeitung gehaltvoller Analysen und Empfehlungen, die Unterschiede und Überschneidungspunkte der Untersuchungsbereiche *Wissenschaft*, *Public Sector* und *Wirtschaft* sichtbar machen. Um die allgemeinen Projektziele mit Blick auf die Förderung des Teilens und Nachnutzens von Daten zu realisieren (vgl. Kapitel 1.1), ist das methodische Vorgehen in drei übergeordnete Phasen unterteilt, die parallel begleitet werden vom Community Engagement (vgl. Kapitel 2.4):

(1) Eine umfangreiche Analyse (methodisches Vorgehen in Kapitel 2.1) der aktuellen und einschlägigen Literatur für die Landscape- und Gap-Analyse (LGA): In den Untersuchungsbereichen *Wissenschaft*, *Public Sector* und *Wirtschaft* bedeutet das die Darstellung von nationalen und internationalen Akteur*innen sowie die Identifizierung relevanter Infrastrukturen, Datenbestände und Handlungsfelder. Mit dem Fokus auf übergreifende FAIRe Datenproduktion und -nachnutzung sowie Identifizierung von Hindernissen und Hürden beim Datenteilen wird die LGA individuell für jeden der drei Untersuchungsbereiche ausgearbeitet. Im Querschnittsthema *Qualifizierung* werden Rahmenbedingungen, Vorgaben und Infrastrukturen betrachtet. Der inhaltliche Schwerpunkt liegt auf Qualifizierung im Umgang mit Daten und der Rolle, die Datenbewusstsein, Data Literacy, Kompetenzprofile, Datenkompetenzvermittlung und neue Berufsfelder spielen und auf den sich daraus ergebenden Bedarfen.

(2) Aufbauend auf der LGA werden leitfadengestützte Expert*innen-Interviews mit repräsentativen Vertreter*innen für die jeweiligen Untersuchungsbereiche

Methodisches Vorgehen EcoDM

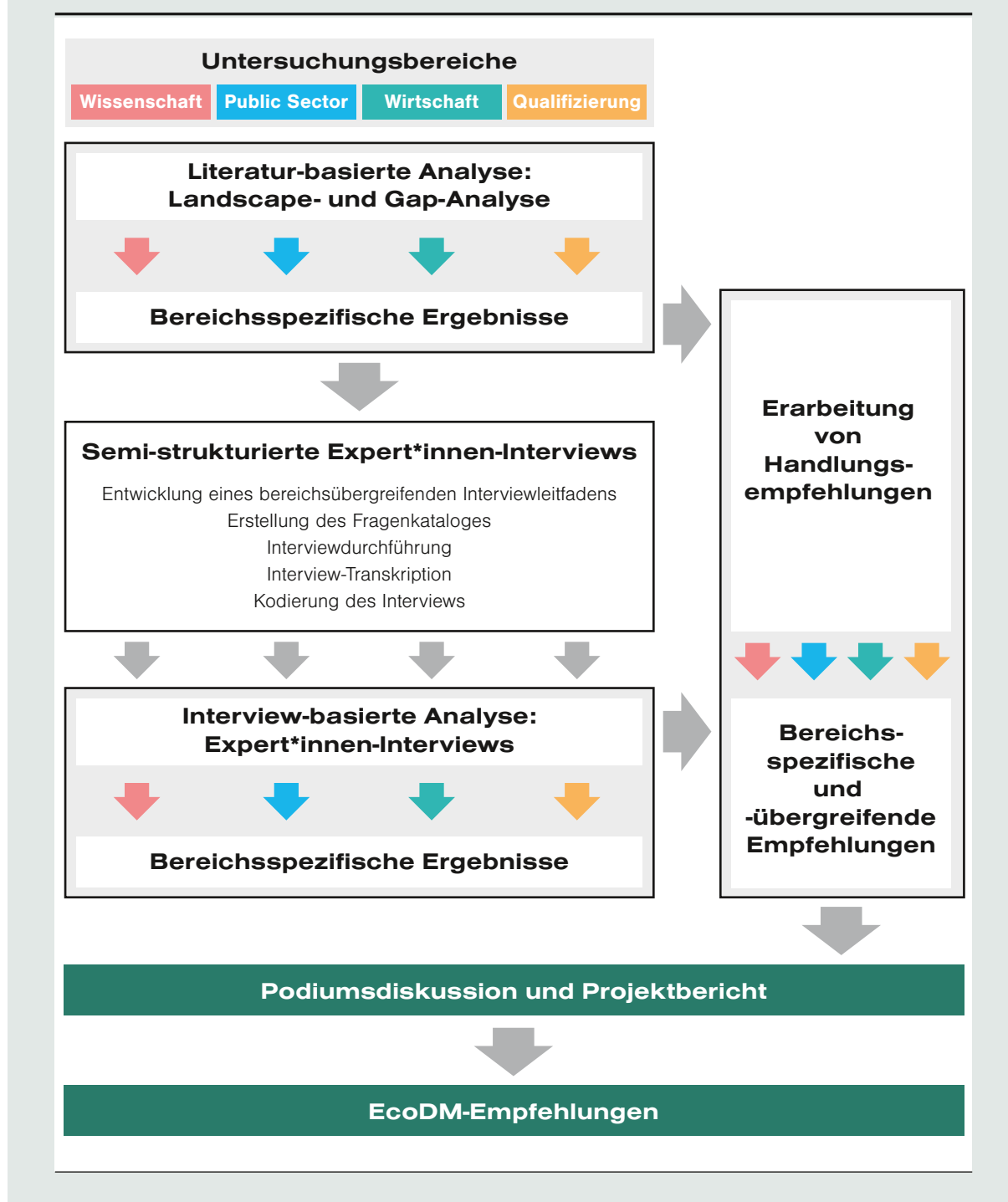


Abbildung 1: Methodisches Vorgehen mit der Darstellung der einzelnen Prozessschritte (Quelle: EcoDM)

durchgeführt bzw. ausgewertet (methodisches Vorgehen in Kapitel 2.2), um die Landscape- und Gap-Analyse zu konkretisieren und Wissenslücken zu füllen. Dies dient zur Identifizierung spezifischer Bedarfe von Gruppen von Akteur*innen und Unterstützungsmöglichkeiten für diese. Die Interviews ermöglichen die Konkretisierung der Hürden beim Datenteilen und tragen zur Erweiterung der Empfehlungen bei, die das

Ziel verfolgen, sowohl die Produktion als auch Nachnutzung von Daten zu steigern. Insbesondere durch die Auswahl der Expert*innen wird auf die aktuelle Entwicklung einschlägiger Diskurse reagiert und Wissenslücken bzw. spezifische Erkenntnisinteressen werden adressiert.

(3) In der dritten Phase werden bereichsspezifische Handlungsempfehlungen von den ersten beiden

Arbeitsphasen abgeleitet (vgl. Kapitel 6), die als Basis für die Erarbeitung der bereichsübergreifenden Handlungsempfehlungen dienen (methodisches Vorgehen in Kapitel 2.3). Diese werden wiederum in einer Handreichung und in Form einer Podiumsdiskussion mit Expert*innen der Öffentlichkeit vorgestellt, diskutiert und die Ergebnisse dokumentiert (vgl. Kapitel 8). Diese Vorgehensweise erlaubt es, der dynamischen Entwicklung des Diskurses zum Thema FAIRes Datenmanagement bzw. Nachnutzung von Daten gerecht zu werden und so in einem offenen und diskursiven Prozess als Ergebnis die EcoDM-Empfehlungen zu erarbeiten (vgl. Kapitel 7).

Parallel dazu ist die Entstehung, Bearbeitung und Speicherung der im Projekt gewonnenen Daten im separaten Forschungsdatenmanagement-Plan dokumentiert (Anhang 10.1). Die einzelnen Prozessschritte des methodischen Vorgehens sind in der Abbildung 1 dargestellt. Darüber hinaus liefert die externe Rechtsexpertise (vgl. Kapitel 3) einen Überblick über rechtliche Aspekte des Teilens von Daten und die aktuellen Entwicklungslinien im Recht zu Daten, z. B. zu Strategien und Gesetzesentwürfen.

2.1 Literatur-basierte Analyse: Landscape- und Gap-Analyse

Zu Projektbeginn erarbeitet das Forschungsteam ein gemeinsames Verständnis der Untersuchungsbereiche und des Forschungsziels auf Basis des vorliegenden Forschungsantrags. Relevante Literatur wird ausgetauscht und diskutiert. In einer gemeinsamen Zotero-Bibliothek wird Material gesammelt und offen kodiert. Die theoretischen Annahmen, auf denen die Codes beruhen, und damit auch das Verständnis des Untersuchungsbereichs werden kontinuierlich geschärft. So sind die Codes weder vollständig, noch werden veraltete Codes verbessert, sondern das neue Wissen wird sukzessive vergleichend auf neues Material angewandt und findet sich in neueren Codes (vgl. Strübing, 2014, S. 19). Die analytische Perspektive des Teams wird durch gemeinsame Besprechungen auf Basis des Materials kontinuierlich neu justiert. Diese ständige Neujustierung geschieht nicht ganz freiwillig und ist u. a. der Corona-Pandemie geschuldet, die viel Planung und etliche Ideen, gerade im Bereich Community Engagement, vom Tisch gewischt hat. So kommt es, dass die ausgearbeitete Landscape- und Gap-Analyse (LGA) zum größten Teil auf Literaturrecherche und Diskussionen im Team beruht und nicht etwa auf online zugespielte Materialien aus der Forschungs-Community. Konzeptionell wird der LGA eine gewisse, auf mehrere Dimensionen bezogene „Kontinuität“ unterstellt: „Alltagspraxis und Wissenschaft, Subjekt und Objekt/Umwelt, aber auch Handeln und Reflexion sind jeweils als

differenzhaltiges Kontinuum miteinander verbunden, statt – wie in der Tradition der analytischen Wissenschaftsphilosophie – als voneinander getrennte Entitäten aufgefasst zu werden“ (ebd., S. 50). Die Grounded Theory wirkt multiperspektivisch: Handeln und Wirklichkeit sind prozesshaft und aktivistisch, insofern es keinen neutralen Beobachter gibt (ebd.).

Die LGA basiert auf zwei zentralen Hypothesen, die auch als Axiome gelten: (1) Daten müssen auffindbar, zugänglich, interoperabel und nachnutzbar gemacht werden (FAIRifizierung). (2) Relevante Akteursgruppen müssen zur Freigabe eigener und Nachnutzung von Daten Dritter motiviert bzw. dabei unterstützt werden.

Es werden fortlaufend „generative Fragen“ erarbeitet und diskutiert (vgl. ebd., 2014, S. 87). In der Erhebungs- und Auswertungsphase wird sich beispielsweise an folgenden bereichsübergreifenden Fragen orientiert:

- Welche Anstöße braucht es zur Integration unterschiedlicher Gruppen von Akteur*innen in die Prozesse und Diskussionen rund um FAIRes digitales Datenmanagement?
- Was bedeutet es konkret, Daten FAIR zu machen, und wo existieren bereits Erfahrungen, die sich auf andere Bereiche übertragen lassen?
- Welche Voraussetzungen und Unterstützungen führen zur Umsetzung der FAIR-Prinzipien?
- Wie werden die Daten interoperabler (zwischen den Bereichen)?
- Wo gibt es konkrete Handlungsbedarfe innerhalb und zwischen Gruppen von Akteur*innen?
- Welche Fähigkeiten und Kompetenzen sollen im Rahmen von Aus- und Weiterbildung besondere Aufmerksamkeit erhalten?
- Welche Vorschläge zur Verbesserung für deutsche und europäische Dateninfrastrukturen (z. B. Nationale Forschungsdateninfrastruktur, European Open Science Cloud und Gaia-X) gibt es?

Die Literatursammlung konzentriert sich auf Quellen ab dem Jahr 2016. Wenn die Theorie im Kontext der Arbeit weiterhin aktuell ist, werden auch ältere Werke erwähnt. Hauptsächlich befasst sich die LGA jedoch mit möglichst neuen Quellen bis ins Jahr 2022. Die Textgattung ist nur im Sinne der Qualität beschränkt: So werden peer-reviewte Artikel, wissenschaftliche Bücher, Reporte, Policy-Papers, Strategietexte und Umsetzungsempfehlungen, Gesetzestexte, Qualifizierungsarbeiten oder Enzyklopädien, aber auch einschlägige Websites, Blogs und Social Media Posts, Newsletter, Präsentationen sowie Erfahrungen aus Konferenzen mit anderen Forscher*innen auf Deutsch und Englisch miteinbezogen.

Die Reduktion des Materials erfolgt zyklisch und bereichsspezifisch, um die Untersuchungsbereiche perspektivisch eindeutig herauszuarbeiten. So werden

Gemeinsamkeiten und Unterschiede in den Quellen pro Untersuchungsbereich deutlich und manche Quellen werden mehrfach – aus den unterschiedlichen Perspektiven von *Wissenschaft, Public Sector, Wirtschaft und Qualifizierung* – beleuchtet. Generell werden Quellen auf der Makroebene des jeweiligen Untersuchungsbereichs ausgewertet: Die erwähnte Literatur hat einen klaren Bezug zum jeweiligen Untersuchungsbereich als Ganzem, doch die jeweiligen Subbereiche, wie beispielsweise Forschungsdatenmanagement in der Mechatronik oder die Praxis eines konkreten Amtes im *Public Sector* werden nicht genauer betrachtet.

Mit der Methode des theoretischen Samplings wird „konzeptuelle Repräsentativität“ in der Fallauswahl angestrebt (Strübing, 2014, S. 31). Dieses Konzept ist in der Grounded Theory verankert und eng verbunden mit der Praxis des ständigen Vergleichens (ebd.). Es sollen sämtliche auf den Untersuchungsgegenstand bezogene Eigenschaften und Dimensionen erfasst werden, um eine ausreichend differenzierte und dichte Analyse zu gewährleisten (vgl. Strübing, 2014, S. 24). Das Wissen wird dabei zunächst auf die Homogenität hin untersucht, um in der anschließenden Phase nach heterogenen Elementen zu suchen, die mit den bisherigen Annahmen im Widerspruch stehen und diese ergänzen bzw. aktualisieren (vgl. Glaser und Strauss 1998, S. 62 f.). Dieser Vorgang wird wiederholt, bis sich die „theoretische Sättigung“ eingestellt hat: Der Untersuchungsbereich wird so lange durchleuchtet, bis keine relevante Information oder Quelle mehr hinzugefügt werden kann (vgl. Strübing, 2014, S. 32). So wird die LGA den gesamten Forschungsprozess über mit neuem Material ergänzt und anderes Material wiederum wird verworfen. Die Annahmen des Forschungsteams und das zusammengetragene Material werden konsequent und zyklisch hinterfragt, wodurch ein vielstufiger Analyseprozess durchlaufen wird.

Auch die Einigung auf Kernkategorien erfolgt durch theoretisches Sampling im Team. Im Vorfeld wird das Datenmaterial offen und dann bereichsspezifisch dimensionalisiert und kodiert. Durch gemeinsame Diskussion im Forschungsteam in repetitiven Prozessen werden bereichsübergreifende Kernkategorien entwickelt und aktualisiert, die sich in den Überschriften in Kapitel 4 widerspiegeln. Darüber hinaus wird eine juristische Rechtsexpertise von Expert*innen dargelegt, um die Kernkategorie Recht bereichsübergreifend zu klären (vgl. Kapitel 3).

Nachdem zu Beginn des Forschungsprozesses meistens offen kodiert wird, lassen sich im Projektverlauf verstärkt „in vivo“ Kodierungen (vgl. Strauss, 1987, S. 33) finden. Dies bedeutet, dass überwiegend direkte Zitate aus verschiedenen Quellen verglichen werden. Gegen Ende der Projektlaufzeit wird fast ausschließlich axial und selektiv kodiert, indem übergreifende

Zusammenhänge und Vorschläge eingebracht werden. Die Schlüsse, die aus dem Material gezogen werden, sind so von Induktion und Abduktion geleitet (vgl. Strübing, 2014, S. 31). Damit soll ein gegenstandsbezogenes Bild der aktuellen Landschaft dargelegt werden, um Lücken und Bedarfe zu identifizieren. Mittels der Neustrukturierung des Materials durch die verschiedenen Untersuchungsbereiche *Wissenschaft, Public Sector, Wirtschaft und Qualifizierung* wird ein Beitrag zu einem gemeinsamen Problemverständnis zwischen den untersuchten Bereichen geleistet, aus dem sich schließlich auch die EcoDM-Empfehlungen ableiten.

2.2 Semi-strukturierte Expert*innen-Interviews

Die Auswahl und Rekrutierung von Interviewpartner*innen erfolgt auf Basis der durch die Literaturanalyse (vgl. Kapitel 2.1) gewonnenen Einblicke in aktuelle Diskurse und Problemlagen. Um durch theoretisches Sampling ein möglichst breites Spektrum abbilden zu können, wird bei der Auswahl der Expert*innen auf die Abdeckung folgender Aspekte geachtet: themenbezogene Expertise und Nähe zu aktuellen Diskursen und/oder Praktiken, unterschiedliche Stakeholder-Kreise mit komplementären oder divergierenden Kompetenzen, Erwartungen und Interessen. Für die Akquise von möglichen Interviewpartner*innen wird ein Informationsblatt (Anhang 10.3) mit Kurzinformationen zum Projekt, zum Interviewablauf, zum weiteren Vorgehen sowie zur Verwendung der gewonnenen Informationen erstellt. Dieses Dokument wird durch eine Einverständniserklärung (Anhang 10.2) ergänzt, die in Zusammenarbeit mit dem Datenschutzbeauftragten der Fachhochschule Potsdam erarbeitet wurde und die Zustimmung zur Aufzeichnung, externen Transkription, Verwendung der Inhalte und für die nicht-öffentliche Speicherung der Transkription für zehn Jahre im Rahmen der guten wissenschaftlichen Praxis beinhaltet. Diese Materialien wurden den potenziellen Interviewpartner*innen bei der Kontaktaufnahme zur Verfügung gestellt.

Pro Themenbereich werden nach dem Prinzip „So offen wie möglich, so strukturierend wie nötig“ (Helfferich, 2019, S. 676) fünf semistrukturierte, auf einen entsprechend teilweise offenen Leitfaden gestützte Expert*innen-Interviews (vgl. Gläser & Laudel, 2010) durchgeführt. Hierfür wird basierend auf den Erkenntnissen aus der LGA bereichsübergreifend ein Leitfaden für die Interviews (Anhang 10.4) entwickelt, in dem sowohl Bedarfe und Hindernisse identifiziert als auch bereichsübergreifende Fragen formuliert werden. Der Leitfaden dient dazu, den Gesprächsverlauf und die entsprechenden Themenfelder vorzustrukturieren. Er

Interview und Analyseprozess EcoDM

Interviews in Zahlen

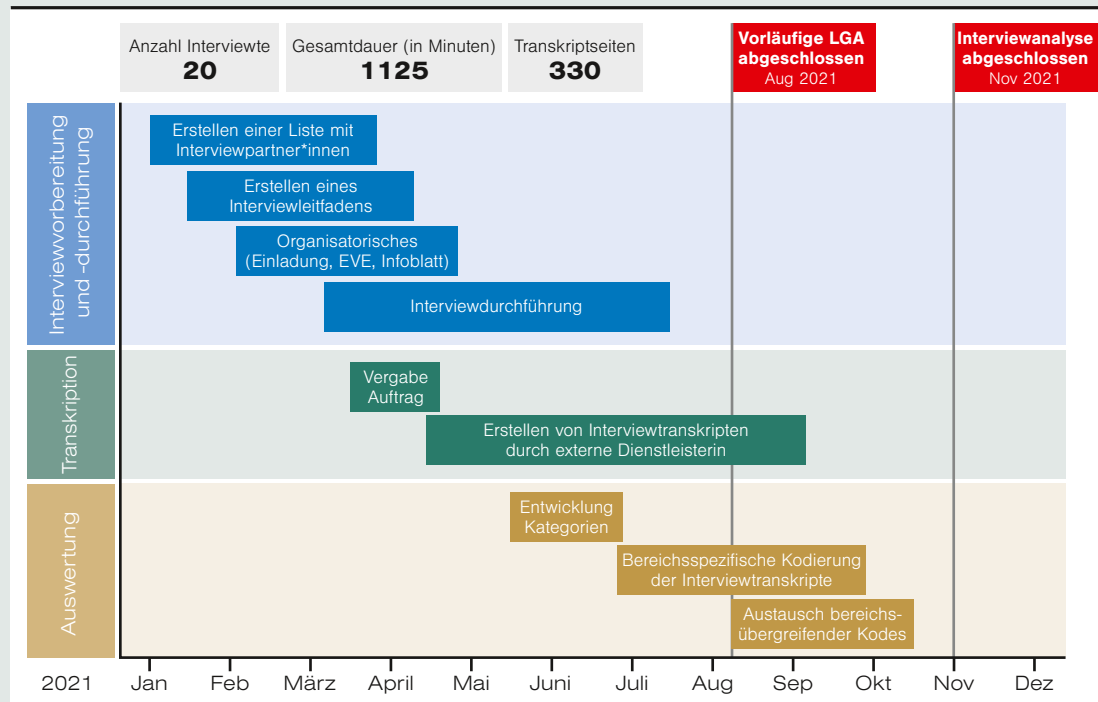


Abbildung 2: Interview-Prozessdarstellung zur Visualisierung der zeitlichen Abfolge und des zeitlichen Aufwands auf Basis von 20 Interviews und 330 Seiten Interview-Transkripten (Quelle: EcoDM)

stellt dem/der Interviewer*in einen auf das Erkenntnisinteresse zugeschnittenen Fragenkatalog mit allen relevanten Fragen zur Verfügung, auf dessen Basis eine bereichsspezifische Auswahl erfolgt. Auch wenn der Leitfaden inhaltlich möglichst umfassend und eindeutig ist, bleibt so genügend Raum und Freiheit, diesen je nach Expert*in und Gesprächsverlauf bedarfsgerecht zuzuspitzen und auszulegen.

Für den Fragenkatalog werden zunächst als übergeordnete Dimensionen „Status quo“, „Voraussetzungen/Bedingungen“, „Herausforderungen/Hürden“ und „Mittel/Lösungsstrategien“ identifiziert. Für eine inhaltliche Strukturierung und weitere Differenzierung – auch mit Blick auf die unterschiedlichen Expertisen der Expert*innen – werden, aufbauend auf den bisherigen Erkenntnissen, folgende übergeordnete thematische Blöcke gebildet: „Kultur des Datenteilens/Nachnutzung“, „Wertebasis/normativer Anspruch/Wertvorstellung“, „Infrastruktur“, „Chancen und Risiken“, „Interoperabilität“, „Standards“, „Anreize“, „Rechtliche Aspekte“, „Empfehlungen“ und „Vision/Zukunftsperspektive“. Für den Querschnittsbereich Qualifizierung werden die thematischen Blöcke um „Data Literacy“, „Datenkompetenz in der Lehre“,

„Qualifikationswege“, „(Neue) Berufsfelder/Karrierewege“ und „Kompetenzprofile/Fähigkeiten“ ergänzt. Für die einzelnen Interviews erfolgt die Erstellung individueller Fragenkataloge – angepasst an die Vorerfahrungen und Expertise der befragten Interviewpartner*innen. Diese werden nicht veröffentlicht, um die Privatsphäre der Interviewten zu wahren. Dabei wird darauf geachtet, dass Daten zu bereichsübergreifenden Gemeinsamkeiten auch in ähnlicher Weise erhoben und ausgewertet werden können.

Bei der Durchführung der Interviews (zum zeitlichen Ablauf, vgl. Abbildung 2) wird die Videokonferenz-Software Zoom¹ verwendet. Das Gespräch wird mit der Aufnahme-Funktion der Software aufgezeichnet. Jeweils zwei Projektmitarbeiter*innen sind während des Interviews anwesend, wobei ein/e Mitarbeiter*in das Interview führt und der/die andere Mitarbeiter*in als Backup für den Fall des Auftretens technischer Probleme bereitsteht und im Bedarfsfall inhaltlich unterstützt. Als Gesprächsdauer werden 45 – 60 Minuten pro Interview angestrebt. Daher sollen gemäß Erfahrungswerten im Leitfaden nicht mehr als fünf Themenblöcke adressiert werden (vgl. ebd., S. 144). Mit einem Fragenkatalog auf Basis

¹ <https://zoom.us/>

des Leitfadens wird ein Pretest durchgeführt, um sowohl die inhaltliche Durchführbarkeit, als auch die Gesprächsdauer zu verifizieren. Nach Beendigung der Interviews erfolgt die Transkription der Videoaufzeichnungen durch ein externes Unternehmen. Aus Datenschutzgründen werden die Aufzeichnungen am Ende des Projekts gelöscht. Archiviert werden ausschließlich die Transkripte. Die Transkripte werden nach der Erstellung projektinterner einer Qualitätssicherung unterzogen. Dabei werden transkriptionsbedingte Fehler bereinigt, Lücken gefüllt und ggf. Schreibweisen korrigiert.

Die Auswertung der Interviewdaten erfolgt auf Grundlage der Transkripte mit Hilfe von Software zur qualitativen Datenanalyse (MaxQDA²) durch Kodierungen. Dabei wird das Material systematisch soweit erschlossen, dass mittels fortlaufend vergebener und bei Bedarf weiterentwickelter Codes und Kategorien, „die erst sukzessive aus der kontinuierlich vergleichenden Analyse dieser Daten entwickelt werden müssen“ (Strübing, 2014, S. 16), die für das Erkenntnisinteresse des Projekts wesentlichen Aussagen erfasst werden. Im Verlauf des ständigen Vergleichens lassen sich drei Stufen des Kodierprozesses differenzieren, die aber keiner festen Sequenz folgen und auch als Mischform realisiert werden können: (a) offenes Kodieren, das dem „analytische[n] Herauspräparieren einzelner Phänomene und ihrer Eigenschaften“ dient, (b) axiales Kodieren, das qualifizierte Konzeptbeziehungen im Material herausarbeitet und (c) selektives Kodieren, das die Beziehungen verschiedener gegenstandsbezogener Konzepte zu den Kernkategorien klärt (vgl. ebd., S. 16f.). Da das Ziel der Expert*innen-Interviews vorwiegend in einer Ergänzung der LGA und der Herausarbeitung von Zielkonflikten und Empfehlungen liegt, wird primär eine Mischform aus offenem und axialem Kodieren angewandt, wobei bei Revisionen des Kategoriensystems auch selektives Kodieren verwendet wird. Dabei handelt es sich jedoch nicht um eine Korrektur vermeintlich falscher Kodierungen, sondern die fortlaufende Anpassung und Justierung der analytischen Perspektive an den zunehmenden Erkenntnisgewinn.

Trotz der wesentlichen Rolle induktiver Schlussverfahren bei der Entwicklung des Kategoriensystems und seiner Kategorien kann nicht von einem ausschließlich induktiven Vorgehen gesprochen werden. Um eine bereichsübergreifende Vergleichbarkeit zu gewährleisten, wird im Vorfeld deduktiv aus dem Fragenkatalog ein grobes Kategorienschema abgeleitet, welches eine erste Orientierung für die Kodierung bietet, aber durch das primär induktive Vorgehen fortlaufend bereichsspezifisch erweitert wird. Anschließend wird die

Kodierung der Interviews pro Bereich durchgeführt (primäres Ziel), wobei zugleich relevante Aussagen für andere Bereiche (sekundäres Ziel) ebenfalls kodiert werden. Erkenntnisse zu anderen Bereichen oder zu bereichsübergreifenden Aspekten werden an die anderen Projektmitarbeiter*innen kommuniziert.

2.3 Erarbeitung von Handlungsempfehlungen

Die Handlungsempfehlungen werden aufbauend auf den ersten beiden Phasen erarbeitet. Für die bereichsspezifischen Handlungsempfehlungen werden die Ergebnisse aus den jeweiligen bereichsspezifischen Literatur-basierten Analysen (methodisches Vorgehen in Kapitel 2.1) und die mit Hilfe von MaxQDA³ aus den 20 Expert*innen-Interviews (methodisches Vorgehen in Kapitel 2.2) gewonnenen Erkenntnisse verwendet. Die bereichsspezifischen Empfehlungen für *Wissenschaft, Public Sector, Wirtschaft* und *Qualifizierung* werden in Kapitel 6 dargestellt. Alle abgeleiteten Empfehlungen werden als „gleichwertig“ angesehen, ohne dabei eine Differenzierung der einzelnen Empfehlungen hinsichtlich ihres methodischen Ursprungs oder ihrer Häufigkeit vorzunehmen. Eine thematische Strukturierung erfolgt anhand der jeweils ermittelten Kernkategorien und ist daher in jedem Untersuchungsbereich individuell. Durch das theoretische Sampling (vgl. Kapitel 2.1) im Entstehungsprozess weisen die Empfehlungen eine hohe multi-dimensionale inhaltliche Dichte auf. Adressat*innen der bereichsspezifischen Empfehlungen sind die Akteur*innen der jeweiligen Untersuchungsbereiche.

Die bereichsspezifischen Empfehlungen werden im nächsten Schritt genutzt, um darauf aufbauend bereichsübergreifende Empfehlungen zu formulieren – nach demselben Prinzip wie bei den bereichsspezifischen Empfehlungen. Dabei werden die wichtigsten Empfehlungen aus den einzelnen Untersuchungsbereichen auf dem Weg zu einer exzellenten und vertrauenswürdigen Datenlandschaft zusammengetragen und die relevanten Erkenntnisse in Kategorien zusammengefasst. Diese sind als Beitrag zur aktuellen Debatte der digitalen Medienlandschaften zu betrachten. Die Kernkategorien der EcoDM-Empfehlungen sind „Harmonisierung und Kooperation“, „Anreizsysteme und Kompetenzentwicklung“, „ethische Grundsätze und öffentliches Interesse“, „Vertrauenswürdigkeit“, „Vermarktung“ und „Wirkung von Open Data Strategien“ (vgl. Kapitel 7). Die so erarbeiteten übergreifenden EcoDM-Empfehlungen werden unter dem Titel „Data At Your Fingertips: Wegmarken einer exzellenten und

² <https://www.maxqda.de/>

³ <https://www.maxqda.de/>

vertrauenswürdigen Datenlandschaft“ von ausgewählten Expert*innen in Form einer moderierten Fokusgruppendifkussion im Rahmen der RDA DE Tagung 2022⁴ diskutiert und so zugleich einer breiteren Öffentlichkeit zugänglich gemacht.

Die Auswahl der Expert*innen für die Podiumsdiskussion orientiert sich beim theoretischen Sampling an den bisherigen Erkenntnissen aus den Untersuchungsbereichen *Wissenschaft, Public Sector, Wirtschaft* und *Qualifizierung* sowie an ihrer jeweiligen Expertise zum Thema Daten. Zur Vorbereitung werden die übergreifenden EcoDM-Empfehlungen den Podiumsteilnehmer*innen und der moderierenden Person im Vorfeld der Fokusgruppendifkussion zur Verfügung gestellt.

Die Moderation erfolgt durch eine versierte externe und damit unabhängige Person. Das gewählte Format eröffnet die Möglichkeit, zusätzlich zu der Expertise der Diskutierenden auch die Expertise des Fachpublikums der Konferenz zur Prüfung der Thesen zu nutzen. Dies wird über die Chat- und F&A-Funktionen der Videokonferenz-Software Zoom, über die die Teilnehmer*innen die Diskussion kommentieren und Fragen stellen können, realisiert. Außerdem werden die Eco-DM-Empfehlungen über ein EtherPad⁵, das während der Podiumsdiskussion und drei Tage darüber hinaus zur Kommentierung online gestellt ist, diskutiert. Diese Diskussionsmöglichkeit wurde über die EcoDM-Webseite, einen Blogbeitrag und Twitter verbreitet. Details zu Teilnehmenden, Ablauf und Ergebnissen der Podiumsdiskussion sind in Kapitel 8 dargestellt.

2.4 Community Engagement und Outreach

2.4.1 Community Engagement im Web

Die Einbindung und Partizipation der Community erfolgt vor allem durch den Auf- und Ausbau der Projektwebseite (www.ecodm.de), die Beispielung der Twitter-Präsenz (@ecodm_projekt) wie auch die Betreuung des Informations- und Kommunikationsportals (<https://www.ecodm.de/informationsportal/>).

Nach der Neuorientierung des Projekts erfolgte der inhaltliche Relaunch der Webseite im Februar 2021, zugleich ging das Informations- und Kommunikationsportal online. Die Projektseite wird als umfangreiche Webpräsenz genutzt. Es werden nicht nur Neuigkeiten und Publikationen über den Blog kommuniziert, sondern auch Veranstaltungsteilnahmen

und die Durchführung von Workshops, inklusive verwendeter Materialien, dokumentiert. Das bietet für die Community die Möglichkeit, Informationen auch nachträglich einzusehen. Alle Projektergebnisse werden Open Access zur Verfügung gestellt, so dass Zwischenergebnisse diskutiert werden können. Der umfangreiche Veranstaltungskalender zu den Themen Open Data, Open Science, Open Access, Forschungsdatenmanagement, FAIR-Prinzipien und zahlreicher verwandter Themen wird seit Februar 2021 kontinuierlich gepflegt und listet im Januar 2022 mehr als 300 Veranstaltungen. Die Community hat zudem die Möglichkeit, über ein Eingabeformular aktiv Veranstaltungen zu ergänzen. Um die Aufmerksamkeit für den Veranstaltungskalender zu erhöhen, werden über Twitter Veranstaltungshinweise mit Verlinkung auf den Kalender verbreitet.

Die Informationsplattform bietet interessierten Leser*innen eine erste Orientierung mit Literaturlisten für alle vier Bereiche. Diese Listen werden im Zuge des Projektfortschritts kontinuierlich erweitert.

2.4.2 Einbindung von Nachwuchsforschenden

Zur Einbindung der Nachwuchsforschenden wurden im Rahmen der RDA DE Tagung 2020 fünf Reisestipendien vergeben. Die Erfahrungsberichte der ausgewählten Teilnehmenden wurden im Nachhinein auf dem Blog veröffentlicht und griffige Statements aus den Berichten auf Twitter geteilt, um die Diskussion über aktuelle Forschungsergebnisse und Entwicklungen auszuweiten.

2.4.3 Community Engagement über Konferenzen und Workshops

Ein wichtiger Pfeiler bei der Einbindung der Community besteht in der Teilnahme an fremden und in der Durchführung eigener Workshops und Schulungen (Tabelle 2) sowie in der Einreichung von Abstracts und Postern bei themenverwandten Konferenzen (Tabelle 1). Durch die Vorstellung des Projektes bzw. die Darstellung von Zwischenergebnissen konnten im Dialog mit interessierten Konferenzteilnehmer*innen Problemfelder thematisiert und Resultate anderer Forschenden ins Projekt aufgenommen werden. Es wurden Beiträge akzeptiert und vorgestellt bei der RDA DE⁶ Tagung 2020, den Open-Access-Tagen 2020⁷ (1. Posterpreis), den E-Science-Tagen 2021⁸ und der DINI⁹-Jahrestagung 2021.

⁴ <https://www.rda-deutschland.de/events/rda-de-2022>

⁵ <https://yopad.eu>

⁶ <https://www.rda-deutschland.de/>

⁷ <https://open-access-tage.de/>

⁸ <https://www.e-science-tage.de/>

⁹ <https://dini.de/>

Projektpublikationen			
Poster	RDA Deutschland Tagung 2020, 25. – 27.02.2020	EcoDM – Ökosystem Datenmanagement: Analysen – Empfehlungen – FAIRification	DOI: http://doi.org/10.5281/zenodo.3698088
Poster	Open-Access-Tage 2020, 15. – 17. 09. 2020	FAIRification von Public Sector Daten für die Wissenschaft	DOI: http://doi.org/10.5281/zenodo.4015576
Poster	E-Science-Tage 2021, 04. – 05.03.2021	EcoDM – Sektorübergreifende Daten FAIRifizierung: Analyse der Ausgangslage und praxisorientierte Entwicklung von Handlungsempfehlungen	DOI: http://doi.org/10.5281/zenodo.4643232
Poster	DINI-Jahrestagung 2021, 05. – 06.10.2021	EcoDM Werkstattbericht – Erkenntnisse aus den Interviews	DOI: http://doi.org/10.5281/zenodo.5550395

Tabelle 1: Übersicht der Poster-Beiträge zum und über das Projekt EcoDM auf themenverwandten Veranstaltungen (Quelle: EcoDM)

Projektveranstaltungen (organisiert / durchgeführt von EcoDM)			
Workshop	RDA DE Tagung 2020, 25. – 26.02.2020	Data Carpentry	Bericht: https://www.ecodm.de/2020/03/01/edodm-rda-de-tagung-2020/
Session	RDA DE Tagung 2020, 26.02.2020	Was bedeutet FAIR?	Bericht: https://www.ecodm.de/2020/03/01/edodm-rda-de-tagung-2020/
Workshop	RDA DE Tagung 2021, 22.02.2021	Mapping the landscape of visualization tools	Bericht: https://www.ecodm.de/2021/03/04/workshop-rda-2021/
Vorträge	RDA DE Tagung 2021, 24.02.2021	Förderung der europäischen Digitalwirtschaft: Welche Ziele sollen realisiert werden? Wie ist eine praktische Umsetzung möglich?	Bericht: https://www.ecodm.de/2021/03/04/rda-de-2021-digitalwirtschaft/
Session	12. govcamp vienna, 26.11.2021	EcoDM – Ökosystem Datenmanagement	Bericht: https://www.ecodm.de/2021/12/01/ecodm-beim-govcamp-vienna/
Workshop	Satellitenveranstaltung RDA DE Tagung 2022, 17.02.2022	Einführung in den Datenschutz	Bericht: https://www.ecodm.de/2022/02/22/rueckblick-datenschutz-workshop-fuer-startups/
Workshop	Satellitenveranstaltung RDA DE Tagung 2022, 17.02.2022	Datenschutz bei der Datenerhebung in der Forschung	
Podiums-diskussion	RDA DE Tagung 2022, 22.02.2022	EcoDM-Abschlussveranstaltung	Bericht: https://www.ecodm.de/2022/02/24/data-at-your-fingertips-podiumsdiskussion-zu-den-ecodm-empfehlungen/

Tabelle 2: Übersicht über die von EcoDM durchgeführten Veranstaltungen (Quelle: EcoDM)

Übersetzungen

Übersetzung 1	Die sechs Säulen für die Einbindung von Forschenden in das Forschungsdatenmanagement	DOI: https://doi.org/10.5281/zenodo.5795858
Übersetzung 2	FAIR Data Maturity Report. Spezifikation und Leitlinien	DOI: https://doi.org/10.5281/zenodo.5834115
Übersetzung 3	Von der Konzeptionalisierung zur Implementierung: FAIR-Evaluation von Forschungsdatenobjekten	DOI: https://doi.org/10.5281/zenodo.5877746
Übersetzung 4	Die CARE-Prinzipien für indigene Data Governance	DOI: https://doi.org/10.5281/zenodo.5995059
Übersetzung 5	Praxisbericht: Die CARE-Prinzipien für indigene Data Governance	DOI: https://doi.org/10.5281/zenodo.6007531
Übersetzung 6	Die FAIR-Prinzipien für wissenschaftliche Datenmanagement und Data Stewardship	DOI: https://doi.org/10.5281/zenodo.6247015
Übersetzung 7	FAIR-Prinzipien: Integrationen und Überlegungen zur Umsetzung	DOI: https://doi.org/10.5281/zenodo.6255623
Übersetzung 8	Die TRUST-Prinzipien für digitale Repositorien	DOI: https://doi.org/10.5281/zenodo.6256222
Übersetzung 9	Ethik und Datenschutz	DOI: https://doi.org/10.5281/zenodo.6259754

Tabelle 3: Übersicht über die von EcoDM übersetzten wichtigen Schlüsseldokumente ins Deutsche (Quelle: EcoDM)

Auch die Vernetzung mit Mitarbeiter*innen von NF-DI-Konsortien oder regionalen Einrichtungen im Forschungsdatenmanagement erfolgt über den (digitalen) Austausch während gemeinsam besuchter Veranstaltungen. Mit der Teilnahme und Durchführung einer Session beim govcamp vienna¹⁰ konnte die Einbeziehung der österreichischen Community gelingen. Interessierte werden bei Projektende über die Ergebnisse per E-Mail informiert.

Im Bereich Workshops und Schulungen wurde die Einbindung der Community durch den Einsatz von Etherpads und Q&A-Sessions im Anschluss an die Vorträge möglich gemacht, wodurch wichtige Impulse auch für zukünftige Workshop-Formate mitgenommen werden konnten.

2.4.4 Einbindung von Übersetzungen

Im Untersuchungsbereich Qualifizierung ist ein Meilenstein des Projektes die Übersetzung wichtiger Schlüsseldokumente aus dem Englischen ins Deutsche (Tabelle 3). Sie sollen im Sinne des Wissenstransfers

eine höhere Interoperabilität und Zugänglichkeit des Wissens ermöglichen und zudem eine bessere Akzeptanz der präsentierten Ergebnisse in der deutschen (Forschungs-)Community erreichen. Die Auswahl der Fachartikel erfolgte auf Basis der Landscape- und Gap-Analyse sowie im Dialog mit ausgewählten Vertreter*innen der untersuchten Bereiche. So konnten Texte identifiziert werden, die für die Community einen hohen Nutzen haben. Die Übersetzungen erfolgten durch externe Dienstleister*innen und wurden nach Qualitätsprüfung unter der CC-BY 4.0-Lizenz auf das Open-Access-Repository Zenodo¹¹ hochgeladen und erhielten so einen DOI. Im nächsten Schritt wurden die Übersetzungen auf der EcoDM-Webseite verlinkt und jeweils durch einen Blogartikel vorgestellt. Um die Reichweite zu erhöhen, wurde der Blogartikel und die dazugehörige Übersetzung auf Twitter geteilt. Die Einbindung der ursprünglichen Autor*innen der Beiträge sorgte für eine zusätzliche Verbreitung der Übersetzungen über deren eigene Web- und Social Media-Kanäle.

¹⁰ <https://barcamps.eu/govcamp-vienna-2021/>

¹¹ <https://zenodo.org>

EXTERNE RECHTS- EXPERTISE



3. Externe Rechtsexpertise¹

3.1 Einleitung: Problemaufriss

Das Teilen von Daten ist zu einer wichtigen Frage geworden, die sowohl technisch als auch gesellschaftlich und politisch kontrovers verhandelt wird. Diese Diskussionen betreffen sehr unterschiedliche Akteur*innen und Interessen in grundsätzlich verschiedenen Konstellationen. Sie können das Prinzip der offenen Datenhaltung der öffentlichen Verwaltung betreffen, das Zahlen mit Daten von Konsument*innen oder das Zusammenfassen von Datenbeständen der Automobilindustrie zur Schaffung entsprechender Trainingsdatensätze zur Entwicklung von Anwendungen Künstlicher Intelligenz. Diese unterschiedlichen Perspektiven, aber auch die steigende gesellschaftliche Wichtigkeit des Teilens von Daten schlagen sich auch im Recht und in der Rechtspolitik nieder.² Entsprechend den Erwartungen verschiedener Akteur*innen ist die Rechtsordnung dabei in manchen Aspekten bereits weit entwickelt, während andere Fragen gerade im Zentrum der regulatorischen Diskussion stehen. So kennt das geltende Recht viele Normen, die für das Teilen von Daten relevant sind, wie etwa den Datenschutz oder Informationsfreiheits- und Open-Data-Pflichten. Im bisherigen Recht gibt es aber weder eine konsolidierte Definition noch eine einheitliche Regelung des Teilens von Daten. Regulatorische Vorhaben auf europäischer Ebene sind aber im Begriff, das zu ändern. Diese Gesetzgebungsvorhaben fokussieren das Teilen von Daten und stellen es in den Mittelpunkt.

Im Folgenden geht es um rechtliche Aspekte des Teilens von Daten. Dabei sollen wesentliche Regelungsstrukturen und die wichtigsten Interpretationsfragen erläutert werden, wobei auch auf die wichtigsten offenen Interpretationsfragen eingegangen werden soll. Auf dieser Grundlage werden Entwicklungslinien im Recht dargestellt, die sich bereits heute auf der Grundlage von Strategien und Gesetzesentwürfen nachvollziehen lassen.

3.2 Rechte an Daten

Rechte an Daten können dem Teilen von Daten entgegenstehen oder es unter bestimmte Voraussetzungen stellen. Wie sich auch aus der gesetzlichen Wertung des Informationsfreiheitsgesetzes ergibt, sind die relevantesten Rechte der Datenschutz sowie der Schutz des geistigen Eigentums und der Betriebsgeheimnisse.³ Bereits seit vielen Jahren wird darüber hinaus eine Diskussion über weitergehende Rechte an Daten unter dem Schlagwort Dateneigentum geführt. Weil ein solches Konzept bisher nicht rechtlich umgesetzt wurde, soll hier lediglich auf die Debatte um das Dateneigentum verwiesen werden.⁴ Neben den Rechten an Daten gibt es auch öffentlich-rechtliche Gründe, die dem Teilen entgegenstehen können, wie etwa der Geheimnisschutz, der in § 4 Sicherheitsüberprüfungsgesetz geregelt ist.⁵

3.2.1 Datenschutzrecht

3.2.1.1 Zweck und Grundsatz

Das Datenschutzrecht ist der datenbezogene Rechtsbereich mit der längsten Tradition und der momentan größten Praxisrelevanz. Generell ist das Datenschutzrecht in der Datenschutzgrundverordnung (DSGVO)⁶ geregelt, wobei für die Fälle der präventiven und repressiven Polizeitätigkeit und in Justizangelegenheiten die nationalen Umsetzungen der JI-Richtlinie⁷ ebenso wie einige spezialgesetzliche Regelungen gelten. Wie sich aus Art. 1 der Datenschutzgrundverordnung ergibt, hat das Datenschutzrecht einen doppelten Zweck. Zum einen betrifft es den Schutz personenbezogener Daten durch Regulierung, wobei diese grund- und menschenrechtliche Gewährleistungen realisieren soll, wie sich aus Art. 1 Abs. 1, 2 der DSGVO ergibt. Andererseits sieht Art. 1 Abs. 3 auch vor, dass der freie Verkehr personenbezogener Daten in der Union aus Gründen des Schutzes natürlicher Personen bei der Verarbeitung personenbezogener Daten weder eingeschränkt noch verboten werden darf. Schon im Zweck der Datenschutzgrundverordnung ist also angelegt, dass die Regulierung auch den Verkehr und mithin

¹ Wir bedanken uns bei den Autor*innen Christian Djeflal und Eva Nürnberger für die rechtliche Einordnung und Expertise. Christian Djeflal ist Tenure Track Assistant Professor für Recht, Wissenschaft und Technik an der Technischen Universität München. Eva Nürnberger ist studentische Mitarbeiterin der Professur.

² Die Bundesregierung – Bundeskanzleramt, Datenstrategie der Bundesregierung. Eine Innovationsstrategie für gesellschaftlichen Fortschritt und nachhaltiges Wachstum (27. Januar 2021); Open-Data-Strategie der Bundesregierung (Juli 2021); Deuring, Datenmacht. Zugang als Herausforderung (2021); Hartl/Ludin, Recht der Datenzugänge, MMR 2021, 534.

³ Siehe dazu §§ 5 und 6 des IFG, die ausdrückliche Regelungen für diese Fälle beinhalten; Informationsfreiheitsgesetz vom 5. September 2005 (BGBl. I S. 2722), das zuletzt durch Artikel 44 der Verordnung vom 19. Juni 2020 (BGBl. I S. 1328) geändert worden ist (IFG).

⁴ Siehe hierzu etwa die Beiträge in Stiftung Datenschutz, Dateneigentum und Datenhandel (2018); Thouvenin/Weber/Früh, Data ownership: Taking stock and mapping the issues, in Dehmer/Emmert-Streib (Hrsg.), Frontiers in data science (2018) 111.

⁵ Sicherheitsüberprüfungsgesetz vom 20. April 1994 (BGBl. I S. 867), das zuletzt durch Artikel 4 des Gesetzes vom 5. Juli 2021 (BGBl. I S. 2274) geändert worden ist (SÜG).

⁶ Verordnung (EU) 2016/679 des Europäischen Parlaments und des Rates vom 27. April 2016 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten, zum freien Datenverkehr und zur Aufhebung der Richtlinie 95/46/EG (Datenschutz-Grundverordnung), ABl. L 119 vom 04.05.2016, S. 1 – 88.

⁷ Richtlinie (EU) 2016/680 des Europäischen Parlaments und des Rates vom 27. April 2016 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten durch die zuständigen Behörden zum Zwecke der Verhütung, Ermittlung, Aufdeckung oder Verfolgung von Straftaten oder der Strafvollstreckung sowie zum freien Datenverkehr und zur Aufhebung des Rahmenbeschlusses 2008/977/JI des Rates, ABl. L 119 vom 04.05.2016, S. 89 – 131.

auch das Teilen von Daten ermöglichen soll. Hintergrund dieser Regelung ist das Ziel der Europäischen Union der Verwirklichung eines Binnenmarkts, also gemäß Art. 26 Abs. 2 des Vertrags über die Arbeitsweise der Union (AEUV)⁸ eines Raums „ohne Binnengrenzen, in dem der freie Verkehr von Waren, Personen, Dienstleistungen und Kapital gemäß den Bestimmungen der Verträge gewährleistet ist.“ Während Daten hier nicht ausdrücklich erwähnt sind, greift schon Art. 1 Abs. 3 DSGVO den Grundgedanken des Binnenmarkts auf und stellt den Fluss von Daten in diesen Kontext. Aus dem folgt, dass das Datenschutzrecht gerade nicht nur den Schutz von Daten bezweckt, sondern auf der Grundlage eines hohen Schutzniveaus den freien Datenfluss im Binnenmarkt ermöglichen will.⁹

3.2.1.2 Regelungstechnik

Das Datenschutzrecht betrifft gemäß Art. 2 Abs. 1 DSGVO grundsätzlich jede Verarbeitung personenbezogener Daten. Wenn sich Daten auf ein Subjekt beziehen lassen und dieses Subjekt mithin identifiziert oder identifizierbar ist, fordert die DSGVO in formeller Hinsicht eine Rechtfertigung, wie etwa eine Einwilligung gemäß Art. 6 Abs. 1 Uabs. a & Art. 7 DSGVO, und in materieller Hinsicht die Einhaltung gewisser Voraussetzungen, wie etwa Auskunftsrechte aus Art. 15 DSGVO. Das Gesetz selbst kennt den Begriff des Teilens von Daten nicht. Vielmehr setzt das Teilen voraus, dass eine verarbeitende Stelle die Daten der anderen verarbeitenden Stelle zugänglich macht. Dabei erfasst die DSGVO „die Offenlegung durch Übermittlung, Verbreitung oder eine andere Form der Bereitstellung“ in Art. 4 Nr. 1 DSGVO. Entscheidend ist, dass der Vorgang des Teilens einer Verarbeitung gleichkommt, die sowohl formal gemäß Art. 6 DSGVO und ggf. gemäß Art. 9 DSGVO gerechtfertigt werden und zudem den übrigen materiellen Bestimmungen der DSGVO entsprechen muss. Teilen zwei Organisationen Daten, sind sie in der Regel beide in vollem Umfang verantwortlich.¹⁰ Ein weiterer wichtiger Aspekt ist, dass Daten nur zu den Zwecken verarbeitet werden dürfen, zu denen sie erhoben wurden. Daten müssen also bereits zum Zweck des Teilens erhoben werden.¹¹ Eine besondere materielle Anforderung beim Teilen von Daten ist die Informationspflicht gemäß Art. 14 DSGVO. Sie greift immer dann, wenn personenbezogene Daten nicht „bei der Person erhoben“ wurden,

wenn sie also ohne Kenntnis und Mitwirkung der Person erhoben wurden.¹² Datensubjekte haben gemäß Art. 15 Abs. 3 DSGVO auch das Recht, eine Kopie der über sie gespeicherten Daten zu verlangen. Personenbezogene Daten müssen folglich auch wieder mit dem Datensubjekt geteilt werden.

3.2.1.3 Sonderfall Datenportabilität

Art. 20 DSGVO sieht ein Recht auf Daten Übertragbarkeit vor. Dieses Recht ist insofern systemfremd, weil es nicht nur im Schutz der Daten als solchem verankert ist, sondern die selbstbestimmte Verwendung der Daten und Aspekte eines fairen Wettbewerbs adressiert.¹³ Wurden die Daten auf der Basis einer Einwilligung oder eines Vertrags erhoben und erfolgt die Verarbeitung automatisiert, so hat das Datensubjekt das Recht, seine Daten in einem strukturierten, gängigen und maschinenlesbaren Format zu erhalten und diese auch mit einem anderen Verantwortlichen zu teilen.¹⁴ Gemäß Art. 20 Abs. 2 DSGVO kann auch eine direkte technische Übertragung an den Dritten erwirkt werden. Das Recht auf Datenportabilität betrifft einen Sonderfall des Teilens von Daten. Hier gehen die Daten auf Geheiß des Datensubjekts von einer verarbeitenden Stelle auf die nächste über.

3.2.1.4 Zusammenfassung

Im Ergebnis stellt das Datenschutzrecht also Voraussetzungen für das Teilen von Daten auf und hemmt es so. Gleichsam definiert es aber bestimmte Voraussetzungen, unter denen das Teilen von Daten zu akzeptieren ist. So trägt das Datenschutzrecht zum freien Verkehr von Daten bei. Nach der Regelungstechnik der DSGVO ist auch das Teilen als Offenlegung von Daten zu qualifizieren. Als Offenlegung ist das Teilen von Daten zu rechtfertigen, etwa durch eine Einwilligung. Nach dem Grundsatz der Zweckbindung muss sich die jeweilige Rechtsgrundlage auch auf das Teilen erstrecken.

3.2.2 Recht des geistigen Eigentums

Ein Recht an Daten kann durch das Recht des geistigen Eigentums vermittelt werden. Ausgangspunkt bildet hier der Gedanke, dass Daten mittlerweile einen wichtigen Vermögenswert im Portfolio vieler Unternehmen darstellen, sich deren Schutz jedoch besonders schwierig gestaltet. Denn ein Eigentums- bzw. Ausschließlichkeitsrecht an Daten existiert bisher in

⁸ Vertrag über die Arbeitsweise der Europäischen Union (konsolidierte Fassung), ABl. C 326 vom 26.10.2012, S. 1 – 12.

⁹ Borges/Hilber, BeckOK IT-Recht4 (2021) Rz 14; Paal/Pauly/Ernst, Datenschutz-Grundverordnung, Bundesdatenschutzgesetz3 (2021) Rz 14; Botta, Die Datenverkehrsfreiheit – Ein Beitrag zur Schutzgutebete im Datenschutzrecht, Das Deutsche Verwaltungsblatt (DVBl.) 2021, 290 – 296.

¹⁰ Art. 26 DSGVO sieht eine gemeinsame Verantwortlichkeit vor. Wird die zweite Stelle nur nachrangig und „im Auftrag“ der ersten aktiv, etwa um bestimmte Dienstleistungen zu erbringen, spricht man gemäß Art. 26 von einem Auftragsverarbeiter.

¹¹ Nur unter engen Voraussetzungen wie etwa in Art. 6 Abs. 4 DSGVO kann der Zweck nachträglich geändert werden.

¹² Schmidt-Wudy in Brink/Wolff, BeckOK Datenschutzrecht38 (1.11.2021) DS-GVO Art. 13 Rz 30.

¹³ Jülicher/Röttgen/Schönfeld, Das Recht auf Datenübertragbarkeit, ZD 2016, 358; Djefal, „Privatheit 4.0“ im Spiegel von Recht und künstlicher Intelligenz, in Behrendt/Loh/Matzner (Hrsg.), Privatsphäre 4.01 (2019) 177 (182 ff).

¹⁴ Zu Einzelheiten siehe Art. 29 Gruppe, Leitlinien zum Recht auf Datenübertragbarkeit (Stand 13. 12. 2016).

Deutschland nicht.¹⁵ Unternehmen sehen sich deshalb mit dem Problem konfrontiert, wie sie die Kontrolle über ihre Daten und damit deren wirtschaftlichen Wert bewahren können, sobald sie diese mit Dritten geteilt haben. Diese Unsicherheit hemmt das Teilen von Daten und mithin den freien Verkehr der Daten.¹⁶ Auf der einen Seite kann das Recht des geistigen Eigentums durch das Einräumen von Rechtspositionen Sicherheiten zur wirtschaftlichen Verwertbarkeit bieten, die Dateninhaber*innen zum Teilen der Daten animieren. Der Schutz als geistiges Eigentum kann das Teilen von großen Datenmengen aber auch verkomplizieren und nahezu unmöglich machen, wenn dabei jeweils eine große Zahl von Rechteinhaber*innen befriedigt werden muss. Im Folgenden soll untersucht werden, wie weit der Schutz des geistigen Eigentums nach dem geltenden Recht reicht. Für das Teilen von Daten sind besonders zwei Schutzmöglichkeiten relevant: der Schutz von Datenbanken aus dem Urheberrechtsgesetz (UrhG)¹⁷ sowie der Schutz der Daten als Geschäftsgeheimnisse.

3.2.2.1 Urheberrecht

Im Rahmen des Urheberrechts können Rechte an Daten im hier relevanten Kontext auf zwei Arten erlangt werden: einerseits als Teil eines Datenbankwerks im Sinne des § 4 Abs. 2 UrhG, andererseits als Teil einer Datenbank gemäß den §§ 87a ff. UrhG. Die Definition einer Datenbank ist dabei für beide Komplexe weitestgehend identisch.¹⁸ Benötigt wird zunächst eine Sammlung von unabhängigen Elementen, die systematisch und methodisch angeordnet sind. Neben dem Aggregieren der Daten ist es außerdem notwendig, dass diese „einzeln mit Hilfe elektronischer Mittel oder auf andere Weise zugänglich sind“ (§§ 4 Abs. 2, 87a Abs. 1 UrhG). Dieses Element verdeutlicht einen Hintergedanken des Schutzes: die Datenverwertung soll durch einen effizienten Zugang zu den einzelnen Elementen verbessert werden.¹⁹ Dadurch wird deren Nutzbarkeit für Dritte, mit denen die Daten geteilt werden, sichergestellt.

Folgerichtig ist der Schutzgegenstand des § 4 Abs. 2 UrhG nicht der Inhalt der Datenbank, also die Daten selbst, sondern deren Struktur.²⁰ Denn darin liegt die schützenswerte persönliche geistige Schöpfung, die für das Eingreifen des Schutzes aus § 4 Abs. 2 notwendig

ist und die nur von einem Menschen ausgehen kann: in der systematischen und methodischen Anordnung und Auswertung der Inhalte (§§ 2 Abs. 2, 4 Abs. 2 UrhG).²¹ Deshalb liegt ein Verstoß gegen die Verwertungsrechte des Datenbank Urhebers nur dann vor, wenn eben diese Strukturleistung übernommen wird, nicht hingegen beim Teilen nur einzelner Elemente einer Datenbank.²² So lehnte der BGH bezüglich Sammelwerken einen Eingriff in das Recht aus § 4 Abs. 1 UrhG ab, wenn aus mehreren thematisch konzeptionierten Bildbänden Fotos entnommen und anschließend in einer neuen Sammlung zusammengestellt werden, die einem völlig neuen Themenkonzept folgt.²³ Das gleiche Prinzip gilt für einen Eingriff in § 4 Abs. 2 UrhG.

Diese „persönliche, geistige Schöpfung“ ist gerade keine Voraussetzung der §§ 87a ff. UrhG, weshalb man durch sie auch kein Urheberrecht erlangt, sondern ein Recht sui generis. Dies war mit Blick auf die Datenmengen, die in einer digitalen Wirtschaft geteilt werden, notwendig. Denn die individuelle Auswahl oder Anordnung der Daten durch einen Menschen dürfte dabei in den wenigsten Fällen gegeben sein. Damit der Datenbankhersteller auch diese Datenbanken wie auch die darin enthaltenen Daten mit Dritten teilen kann, ohne seine Investition in die Strukturierung und effiziente Gestaltung des Zugriffs zu verlieren, wurden die §§ 87a ff. UrhG geschaffen. Deren Schutzgut ist die Investition in die Datenbank, die „nach Art und Umfang wesentlich“ sein muss (§ 87a Abs. 1). Relevant sind dabei lediglich Kosten, die bei der Beschaffung, Überprüfung oder Darstellung der Datenbank angefallen sind.²⁴ Die Investitionen, welche zur Generierung der Daten angefallen sind, bleiben hingegen unberücksichtigt. In der Praxis führt dies zu erheblichen Abgrenzungsschwierigkeiten, die zu einer Vielzahl an Gerichtsentscheidungen geführt haben.²⁵ Besonders im Bereich der Industrie 4.0 bleiben zahlreiche Unschärfen, welche Investitionen eines Datenbankherstellers von Bedeutung sind und welche nicht.²⁶

Als Datenbankhersteller gilt dabei, wer die Investitionen vorgenommen hat (§ 87a Abs. 2). Der oben angeschnittene Streit darum, welche Investitionen Beachtung finden, ist folglich auch von hoher praktischer Relevanz. Dem Hersteller steht das ausschließliche Recht zu, die Datenbank insgesamt oder einen wesentlichen Teil zu vervielfältigen, zu verbreiten sowie

15 Schur, Teil 6.8 Know-how- und Geheimnisschutz von Daten, in Leupold/Wiebe/Glossner (Hrsg.), IT-Recht – Recht, Wirtschaft und Technik der digitalen Transformation, Rn. 1.

16 Vgl. bspw. Krüger/Wiencke/Koch, Der Datenpool als Geschäftsgeheimnis, GRUR 2020, 578, 579.

17 Urheberrechtsgesetz vom 9. September 1965 (BGBl. I S. 1273), das zuletzt durch Artikel 25 des Gesetzes vom 23. Juni 2021 (BGBl. I S. 1858) geändert worden ist.

18 Sattler, § 2 Schutz von maschinengenerierten Daten, in Sassenberg/Faber (Hrsg.), Rechtshandbuch Industrie 4.0 und Internet of Things, Rn. 29.

19 Sattler, § 2 Schutz von maschinengenerierten Daten, in Sassenberg/Faber (Hrsg.), Rechtshandbuch Industrie 4.0 und Internet of Things, Rn. 30.

20 Schrickler/Loewenheim/Leistner, Urheberrecht, UrhG § 4, Rn. 59.

21 Schrickler/Loewenheim/Leistner, Urheberrecht, UrhG § 4, Rn. 1; Sattler, § 2 Schutz von maschinengenerierten Daten, in Sassenberg/Faber (Hrsg.), Rechtshandbuch Industrie 4.0 und Internet of Things, Rn. 28.

22 Dreier/Schulze/Dreier, Urheberrechtsgesetz, Vorbemerkung zu § 87a ff. UrhG, Rn. 4; Wiebe, Teil 3. Schutz von Datenbanken, Computerspielen und Webangeboten, in Leupold/Wiebe/Glossner (Hrsg.), IT-Recht – Recht, Wirtschaft und Technik der digitalen Transformation, Rn. 17.

23 BGH GRUR 2013, 1213 Rn. 57 – SUMO.

24 Dreier/Schulze/Dreier, Urheberrechtsgesetz, UrhG § 87a Begriffsbestimmungen, Rn. 11.

25 Vergleiche hierzu die Darstellung inklusive Fußnoten in Schrickler/Loewenheim/Vogel, Urheberrecht, UrhG § 87a, Rn. 44 ff.

26 Sattler, § 2 Schutz von maschinengenerierten Daten, in Sassenberg/Faber (Hrsg.), Rechtshandbuch Industrie 4.0 und Internet of Things, Rn. 35 ff.

öffentlich wiederzugeben (§ 87b Abs. 1). Im Gegensatz zu dem Schutz aus § 4 Abs. 2 UrhG werden – zumindest indirekt – auch die Daten, also der Inhalt der Datenbank, geschützt. Denn selbst dann, wenn die verschiedenen Elemente aus der Datenbank entnommen und anschließend neu zusammengestellt veröffentlicht werden, liegt ein Verstoß gegen die Verwertungsrechte des Datenbankherstellers vor.²⁷ Denn Schutzgegenstand ist die Investition in die Datenbank, die auch durch eine solche Verwendung betroffen ist.²⁸ Konkurrenten des Herstellers könnten die durch ihn geteilten Daten mit erheblich geringeren Kosten in eine eigene Datenbank überführen und so den wirtschaftlichen Wert der Datenbank erheblich schmälern.

Sowohl § 4 Abs. 2 UrhG als auch §§ 87a ff. UrhG beziehen sich lediglich auf den Schutz der Struktur der Datenbank, nicht auf den Schutz der Daten. Damit soll ein Anreiz zur Aggregation, Aufbereitung und effektiven Zugänglichmachung von Daten gesetzt werden. Die einzelnen Elemente der Datenbanken, also die Daten selbst, sind nur indirekt geschützt, ein Recht an den Daten selbst besteht nicht.

3.2.2.2 Schutz von Geschäftsgeheimnissen

Ein solches Recht an Daten könnte jedoch deren Einordnung als Geschäftsgeheimnis vermitteln. Daten können ein Geschäftsgeheimnis im Sinne des Gesetzes zum Schutz von Geschäftsgeheimnissen (GeschGehG)²⁹ darstellen, falls sie die Voraussetzungen des § 2 Nr. 1 GeschGehG erfüllen. Ob ein bestimmter Datensatz als Geschäftsgeheimnis einzuordnen ist und damit einen Schutz vor dem ungerechtfertigten Teilen mit Dritten genießt, ist in jedem Einzelfall gesondert festzustellen.³⁰ Zu beachten ist dabei, dass dieser Schutz nur so lange besteht, wie das Geheimsein der Information, die in den Daten verkörpert ist, sichergestellt ist. Es besteht folglich kein Ausschließlichkeitsrecht an den Daten, weshalb die Geschäftsgeheimnis-Richtlinie³¹, auf der die deutsche Umsetzung beruht, den Geheimnisschutz lediglich als Ergänzung zu den Rechten des geistigen Eigentums sieht, die Exklusivitätsrechte verleihen.³²

Das GeschGehG definiert in den §§ 1 – 5 zunächst, wann ein Geschäftsgeheimnis vorliegt und was in

Bezug darauf erlaubt, was verboten und wo Ausnahmen liegen. Die §§ 6 – 12 GeschGehG beinhalten die zivilrechtlichen Ansprüche der Inhaber*innen des Geheimnisses gegen die Rechtsverletzenden, § 23 eine Strafvorschrift. Für das Gerichtsverfahren, das auf eine Klage gegen einen Verstoß gegen das GeschGehG folgt, sieht §§ 15 – 22 Verfahrensvorschriften vor, die insbesondere dem Interesse an Geheimhaltung Rechnung tragen.

Von besonderem Interesse für das Teilen von Daten in diesem Zusammenhang sind dabei drei Komplexe: erstens inwieweit Daten als Information iSd GeschGehG angesehen werden können. Zweitens wann eine Information als geheim anzusehen ist und was für Maßnahmen die Inhaberin / der Inhaber ergreifen sollte, um diesen Zustand sicherzustellen. Und drittens wer der rechtmäßige Inhaber / die rechtmäßige Inhaberin dieser Daten ist und damit die Ansprüche bei Rechtsverletzungen (§§ 6 ff. GeschGehG) geltend machen kann. Unstreitig ist, dass Daten grundsätzlich unter den Begriff der „Information“ fallen können.³³ Der Schutzzweck des GeschGehG ist jedoch klar auf den Inhalt der Daten ausgerichtet, also auf der semantischen Informationsebene anzusiedeln. Die syntaktische Informationsebene, das Verständnis von Daten lediglich als Folge von Bits, wird hingegen nicht erfasst.³⁴ Einzelne Datenpunkte, die in keinerlei Kontext stehen, sind folglich keine Information im Sinne des GeschGehG. Abgesehen davon sind die Anforderungen an den Informationsgehalt von Datensammlungen sehr niedrig anzusiedeln.³⁵ So sind beispielsweise bereits maschinengenerierte Daten³⁶ und Rohdatensammlungen³⁷ als Informationen i. S. d. § 2 Nr. 1 GeschGehG einzuordnen.

Der zweite Themenkomplex betrifft die Eigenschaft einer Information als geheim. Das ist dann der Fall, wenn sie „weder insgesamt noch in der genauen Anordnung und Zusammensetzung ihrer Bestandteile den Personen in den Kreisen, die üblicherweise mit dieser Art von Informationen umgehen, allgemein bekannt oder ohne Weiteres zugänglich ist“. Dabei muss entweder die Information selbst, die in den Daten verkörpert ist, geheim sein, oder deren genaue Anordnung und Zusammensetzung ihrer Bestandteile. Der

27 BGH, CR 2011, 498 Rn. 31 – Zweite Zahnarztmeinung II; BGH, K&R 2011, 641 Rn. 38 – Automobil-Onlinebörse.

28 Wiebe, Teil 3. Schutz von Datenbanken, Computerspielen und Webangeboten, in Leupold/Wiebe/Glossner (Hrsg.), IT-Recht – Recht, Wirtschaft und Technik der digitalen Transformation, Rn. 27.

29 Gesetz zum Schutz von Geschäftsgeheimnissen vom 18. April 2019 (BGBl. I S. 466) (GeschGehG).

30 Schur, Teil 6.8 Know-how- und Geheimnisschutz von Daten, in Leupold/Wiebe/Glossner (Hrsg.), IT-Recht – Recht, Wirtschaft und Technik der digitalen Transformation, Rn. 12.

31 Richtlinie (EU) 2016/943 des Europäischen Parlaments und des Rates vom 8. Juni 2016 über den Schutz vertraulichen Know-hows und vertraulicher Geschäftsinformationen (Geschäftsgeheimnisse) vor rechtswidrigem Erwerb sowie rechtswidriger Nutzung und Offenlegung, ABl. L 157 vom 15.06.2016, S. 1 – 18.

32 Schur, Teil 6.8 Know-how- und Geheimnisschutz von Daten, in Leupold/Wiebe/Glossner (Hrsg.), IT-Recht – Recht, Wirtschaft und Technik der digitalen Transformation, Rn. 57, 59; Erwägungsgrund 2, 16 Geschäftsgeheimnis-RL.

33 Vgl. bereits § 4 Abs. 1 Nr. 1 GeschGehG, der „elektronische Dateien“ aufführt; bezüglich Kund*innendaten vgl. BGH, Aneignung von Kundendaten als Verletzung des Geschäftsgeheimnisses, GRUR 2006, 1044; Schur, Teil 6.8 Know-how- und Geheimnisschutz von Daten, in Leupold/Wiebe/Glossner (Hrsg.), IT-Recht – Recht, Wirtschaft und Technik der digitalen Transformation, Rn. 12, m. w. N.

34 Schur, Teil 6.8 Know-how- und Geheimnisschutz von Daten, in Leupold/Wiebe/Glossner (Hrsg.), IT-Recht – Recht, Wirtschaft und Technik der digitalen Transformation, Rn. 11, 12.

35 Ohly, Das neue Geschäftsgeheimnisgesetz im Überblick, GRUR 2019, 441, 442.

36 Hessel/ Leffer, Rechtlicher Schutz maschinengenerierter Daten, MMR 2020, 647, 649.

37 Krüger/ Wiencke/ Koch, Der Datenpool als Geschäftsgeheimnis, GRUR 2020, 578, 580.

deutsche Wortlaut, der beide Teile geheim sehen will, ist insoweit irreführend.³⁸ Das Geheimsein der Information stellt in der heutigen vernetzten, datengetriebenen Wirtschaft, insbesondere in der Industrie 4.0, eine immer größere Herausforderung für den rechtmäßigen Inhaber / die rechtmäßige Inhaberin der Daten dar. Dieser muss den Schutz der Daten durch angemessene Geheimhaltungsmaßnahmen sicherstellen. Denn das Geschäftsmodell vieler Unternehmen beruht mittlerweile darauf, unternehmensübergreifend Daten zu teilen und eng mit verschiedenen Geschäftspartner*innen in „Wertschöpfungsnetzwerken“ zusammen zu arbeiten.³⁹ Dabei gleichzeitig das Geheimsein der eigenen Daten und damit die Rechte an den Daten aus dem GeschGehG sicherzustellen, steht in einem offensichtlichen Spannungsverhältnis dazu. Relevant für den Status als Geheimnis ist „die praktische Zugangsmöglichkeit für einen bestimmten Personenkreis.“⁴⁰ Dieser kann auch aus einer größeren Anzahl an Dritten bestehen, mit denen der Inhaber / die Inhaberin die Daten teilt, solange angemessene Geheimhaltungs-Maßnahmen unternommen wurden, die Information also immer noch vertraulich ist.⁴¹ Wenn eine Information geheim ist, führt auch – anders als noch unter altem Recht – deren rechtswidrige Offenlegung nicht zwingend zu einem Wegfall dieser Einstufung.⁴² Vielmehr sind Dritte, die ein Geschäftsgeheimnis erlangt, genutzt oder offengelegt haben und dabei wussten, dass die Daten rechtswidrig geteilt wurden, gemäß § 4 Abs. 3 GeschGehG selbst als Rechtsverletzer anzusehen.

Welche Maßnahmen angemessen sind, um das Geheimnis zu schützen, kann nicht pauschalisiert werden. Im Rahmen einer Einzelfallprüfung müssen vielmehr die Art des Geschäftsgeheimnisses und dessen wirtschaftlicher Wert, die Wirtschaftskraft des Inhabers / der Inhaberin sowie die konkreten Umstände, in denen es genutzt wird, berücksichtigt werden.⁴³ Der Inhaber / die Inhaberin der Daten muss also geeignete organisatorische, technische und vertragliche Maßnahmen ergreifen, die aufgrund der gesteigerten Konnektivität der Unternehmen und ihrer Daten einen Balanceakt zwischen dem einerseits gewünschten Teilen von Daten und andererseits der Sicherung aller Schnittstellen, an denen Zugang zu eigenen Daten gewährt wird, darstellen.⁴⁴

Als letzter Aspekt ist von Interesse, wer genau als Inhaber*in von Daten im Rahmen des GeschGehG gilt und damit diese nutzen, zugänglich machen sowie

teilen darf, ohne mit den Konsequenzen der §§ 6 ff. GeschGehG rechnen zu müssen. Grundsätzlich gilt: Wer die rechtmäßige Kontrolle über ein Geschäftsgeheimnis hat, ist deren Inhaber*in (§ 2 Nr. 2). Das kann sowohl eine natürliche als auch juristische Person sein. In der Datenwirtschaft ist diese Bestimmung aufgrund des regen Austauschs von Daten problematisch. Denn der Begriff „rechtmäßige Kontrolle“ bezieht sich weder darauf, wer die Daten ursprünglich aggregiert hat, noch ist das Eigentum an Datenträgern relevant. Kontrolle ist vielmehr tatsächlich zu verstehen und setzt die faktische Herrschaft über die Daten voraus. Folglich ist derjenige Inhaber*in der Daten im Sinne des GeschGehG, der selbst Zugang zu diesen hat sowie bestimmen kann, wer Zugang enthält, und dies notfalls durch Zwang durchzusetzen vermag.

Zu welchen Abgrenzungsschwierigkeiten das vor allem in der Industrie 4.0 führen kann, zeigt das Beispiel eines Sensors, der in ein Smart Car eingebaut ist: Als Dateninhaber kommen sowohl der Eigentümer des Autos, der Hersteller des Sensors sowie der Hersteller des Autos in Betracht.⁴⁵ Wer genau Inhaber*in ist, lässt sich nur im Einzelfall bestimmen. Den faktischen Zugriff auf die generierten Daten wird in solchen Konstellationen regelmäßig der Fahrzeughersteller haben. Um die Rechtmäßigkeit dieser Kontrolle zu gewährleisten, sollte er jedoch zusätzlich sein Recht an diesen Daten mit dem Auto-Eigentümer vertraglich vereinbaren. Solche Datennutzungsverträge mit exklusiven Nutzungsrechten der generierten Daten gehören bereits jetzt zum Standardrepertoire vieler Autohersteller.⁴⁶

Das Gesetz zum Schutz von Geschäftsgeheimnissen vermittelt kein exklusives Recht an Daten, sondern knüpft deren Schutz vielmehr an die Bedingung, dass diese geheim sind. Dies steht in einem direkten Spannungsverhältnis zu dem Wunsch des Inhabers / der Inhaberin, Daten mit anderen Personen zu teilen. Um den Schutz seiner Daten nicht zu verlieren, ist der Inhaber / die Inhaberin gezwungen, ein umfassendes Konzept an Geheimhaltungs-Maßnahmen zu entwickeln sowie zu implementieren.

3.2.2.3 Zusammenfassung

Sowohl das Urheberrecht als auch das Geheimnisschutzrecht sind in der Lage, dem Rechteinhaber / der Rechteinhaberin das Teilen seiner Daten zu ermöglichen, ohne seine Rechte an den Daten zu verlieren. Dabei weisen die beiden Schutzmechanismen jedoch unterschiedliche Schwächen auf: Das Urheberrecht

38 Ohly, Das neue Geschäftsgeheimnisgesetz im Überblick, GRUR 2019, 441, 443; Schur, Teil 6.8 Know-how- und Geheimnisschutz von Daten, in Leupold/Wiebe/Glossner (Hrsg.), IT-Recht – Recht, Wirtschaft und Technik der digitalen Transformation, Rn. 14.

39 Schur, Teil 6.8 Know-how- und Geheimnisschutz von Daten, in Leupold/Wiebe/Glossner (Hrsg.), IT-Recht – Recht, Wirtschaft und Technik der digitalen Transformation, Rn. 18.

40 Ohly, Das neue Geschäftsgeheimnisgesetz im Überblick, GRUR 2019, 441, 443.

41 Ohly, Das neue Geschäftsgeheimnisgesetz im Überblick, GRUR 2019, 441, 443; BGH, GRUR 2018, 1161 Rn. 38 – Hohlfasermembranspinnanlage II.

42 Krüger/Wiencke/Koch, Der Datenpool als Geschäftsgeheimnis, GRUR 2020, 578, 581.

43 Schur, Teil 6.8 Know-how- und Geheimnisschutz von Daten, in Leupold/Wiebe/Glossner (Hrsg.), IT-Recht – Recht, Wirtschaft und Technik der digitalen Transformation, Rn. 24 f.

44 Wiebe/Schur, Protection of Trade Secrets in a Data-driven, Networked Environment – Is the update already out-dated?, GRUR Int. 2019, 746, 749 ff.

45 Beispiel nach Krüger/Wiencke/Koch, Der Datenpool als Geschäftsgeheimnis, GRUR 2020, 578, 582.

46 Krüger/Wiencke/Koch, Der Datenpool als Geschäftsgeheimnis, GRUR 2020, 578, 582.

vermittelt kein direktes Recht an den Daten, sondern schützt lediglich deren Strukturierung in einer Datenbank. Das Geheimnisschutzrecht ist an die Voraussetzung geknüpft, dass die Daten geheim sind, was einen kontinuierlichen Balanceakt des Rechteinhabers / der Rechteinhaberin erfordert, falls er seine Daten mit weiteren Personen teilen will. Die beiden Schutzmechanismen zusammen führen zwar zu einer gewissen Sicherheit des Rechteinhabers / der Rechteinhaberin, sollten aber in jedem Fall durch vertragliche Vertraulichkeitspflichten unterfüttert werden, um ihre Schwächen auszugleichen. Denn keiner der beiden Mechanismen ist auf das vielfache Teilen von Daten, das im Rahmen der gesteigerten Vernetzung einer datengetriebenen Wirtschaft noch weiter wachsen wird, ausgelegt. Ferner folgt aus den durch das geistige Eigentum vermittelten Rechten noch nicht, dass diese auch auf eine Weise kommerzialisiert werden können, die den einzelnen Rechteinhaber*innen auch wirklich nützt. Die Einräumung von Rechten an viele Akteur*innen führt zu einem hohen Verwaltungsaufwand in der Wahrnehmung der Rechte. Das Recht des geistigen Eigentums vermittelt zwar Rechtspositionen, trägt aber weniger dazu bei, nachhaltige Geschäftsmodelle und Infrastrukturen der Kommerzialisierung von Daten zu schaffen.

3.3 Rechte auf Daten

Rechte auf Daten betreffen Verpflichtungen zum Teilen von Daten. Diese sind am weitesten im Verhältnis von Rechten Privater gegen den Staat entwickelt. Insbesondere im Bereich des Wettbewerbsrechts werden Rechte auf Daten zwischen Privaten geregelt und diskutiert.

3.3.1 Wettbewerbsrecht

Wird die Frage des Datenzugangs aus einer wettbewerbsrechtlichen Perspektive behandelt, bewegt man sich in einem Spannungsfeld entgegengesetzter Interessen: Auf der einen Seite stehen die Dateninhaber*innen, die aufgrund der wettbewerbsrechtlichen Anreizstrukturen Investitionen in ihren Märkten tätigen und Daten generieren, aggregieren sowie teilen. Auf der anderen Seite steht das Bedürfnis nach einem

funktionsfähigen Leistungs- und Innovationswettbewerb, der durch die Monopolisierung der Daten bei einigen Dateninhaber*innen nicht zu der bestmöglichen Verbesserung der angebotenen Dienstleistungen und Produkte führt. Doch nur dadurch kann das Gemeinwohl stetig gesteigert werden.⁴⁷ Dieser Konflikt wurde durch die 10. Novellierung des Gesetzes gegen Wettbewerbsbeschränkungen (GWB)⁴⁸ aufgegriffen, die am 19.01.2021 in Kraft getreten ist, und schlug sich in der Schaffung von zwei expliziten Datenzugangsrechten nieder: einerseits gem. § 19 Abs. 2 Nr. 4 GWB im Rahmen der Essential-Facilities-Doktrin, andererseits nach § 20 Abs. 1a GWB bei einer relativen Marktmacht des Unternehmens. Auch vor dieser Neuerung wurde ein Zugang zu Daten aus Marktmissbrauch theoretisch angedacht, war jedoch weder unumstritten noch kodifiziert.⁴⁹ Durch die Neuregelung soll nun die praktische Geltendmachung von solchen Zugangsansprüchen gefördert werden.⁵⁰ Daneben stellt auch das Europarecht mit Art. 102 AEUV eine Möglichkeit dar, eine Missbrauchskontrolle auf europäischer Ebene durchzuführen.

Der Gesetzgeber hat im Rahmen der Gesetzesänderung bewusst darauf verzichtet, die für §§ 19 Abs. 2 Nr. 4, 20 Abs. 1a GWB relevanten Daten zu definieren. Grundsätzlich kann folglich jegliche Art von Daten durch den Zugangsanspruch erfasst sein.⁵¹ Dabei wird besonders die Beziehung zwischen den oben dargestellten Rechten an Daten aus der DSGVO, dem GeschGehG, dem UrhG sowie zusätzlich die Schranke des § 1 GWB⁵² von praktischer Bedeutung sein.⁵³

§ 19 Abs. 2 Nr. 4 GWB sanktioniert den Missbrauch einer marktbeherrschenden Stellung durch ein Unternehmen und hat die Anpassung des deutschen Rechtsrahmens an die europäische Entwicklung sowie die Wettbewerbserhaltung zum Ziel.⁵⁴ Damit die Missbrauchskontrolle aktiviert wird, muss ein Unternehmen marktbeherrschend im Sinne des § 18 GWB sein, wobei explizit auch der Zugang zu wettbewerbsrelevanten Daten zu dieser Stellung führen kann (§ 18 Abs. 3 Nr. 3, Abs. 3a Nr. 4 GWB). Die Daten müssen außerdem eine *essential facility*⁵⁵ darstellen. Das ist der Fall, wenn der Zugang zu diesen Daten „objektiv notwendig ist, um auf einem vor- oder nachgelagerten Markt tätig zu sein und [...] die Weigerung [des Unternehmens] den wirksamen Wettbewerb auf diesem

47 Schweda/ von Schreitter, Ran an die Datenschätze? Datenzugangsansprüche nach der 10. GWB-Novelle, WuW 2021, 145, 145.

48 Gesetz gegen Wettbewerbsbeschränkungen in der Fassung der Bekanntmachung vom 26. Juni 2013 (BGBl. I S. 1750, 3245), das zuletzt durch Artikel 10 Absatz 2 des Gesetzes vom 27. Juli 2021 (BGBl. I S. 3274) geändert worden ist (GWB).

49 Schweda/ von Schreitter, Ran an die Datenschätze? Datenzugangsansprüche nach der 10. GWB-Novelle, WuW 2021, 145, 148.

50 Käseberg/ Brenner/ Fülling, Das GWB-Digitalisierungsgesetz im Überblick, WuW 2021, 269, 271.

51 Brenner, Kapitel 1: Wettbewerb und Digitalwirtschaft sowie weitere materielle rechtliche Änderungen, in Bien/Käseberg/Klumpe/Körper/Ost (Hrsg.), Die 10. GWB-Novelle, 2021, Rn. 82.

52 § 1 GWB Verbot wettbewerbsbeschränkender Vereinbarungen: „Vereinbarungen zwischen Unternehmen, Beschlüsse von Unternehmensvereinigungen und aufeinander abgestimmte Verhaltensweisen, die eine Verhinderung, Einschränkung oder Verfälschung des Wettbewerbs bezwecken oder bewirken, sind verboten.“

53 Schweda/ von Schreitter, Ran an die Datenschätze? Datenzugangsansprüche nach der 10. GWB-Novelle, WuW 2021, 145, 150.

54 Deutscher Bundestag, Drucksache 19/23492, S. 72; Schweda/ von Schreitter, Ran an die Datenschätze? Datenzugangsansprüche nach der 10. GWB-Novelle, WuW 2021, 145, 151.

55 Vergleiche zur Entwicklung des Begriffs und dessen europarechtlichen Bezügen Brenner, Kapitel 1: Wettbewerb und Digitalwirtschaft sowie weitere materielle rechtliche Änderungen, in Bien/Käseberg/Klumpe/Körper/Ost (Hrsg.), Die 10. GWB-Novelle, 2021, Rn. 92 ff.

Markt auszuschalten droht“ (§ 19 Abs. 2 Nr. 4 GWB). Objektive Notwendigkeit erfordert, dass die Daten für denjenigen, der den Zugang begehrt, „nicht mit angemessenem Aufwand reproduzierbar sind“, wobei es auf den Einzelfall ankommt.⁵⁶ Doch selbst wenn dies der Fall ist, kann die Weigerung immer noch sachlich gerechtfertigt sein.

§ 20 Abs. 1a GWB ist auf Konstellationen relativer Marktmacht anwendbar. Ein Unternehmen, das „für die eigene Tätigkeit auf den Zugang zu Daten angewiesen ist, die von einem anderen Unternehmen kontrolliert werden“, kann sich demnach in einer solchen Datenbedingten Abhängigkeit befinden, dass „[d]ie Verweigerung des Zugangs zu solchen Daten gegen angemessenes Entgelt [...] eine unbillige Behinderung [...] darstellen [kann].“ Irrelevant ist dabei, ob die Daten bereits zuvor verfügbar waren, „ein Geschäftsverkehr für diese Daten [also] bislang nicht eröffnet ist.“ (S. 3). Unter Vertragspartner*innen kann diese kartellrechtliche Korrektur der Vertragsfreiheit dann zum Einsatz kommen, wenn ein Unternehmen seine übermächtige Markt- oder Verhandlungsmacht ausspielt, um seinem Vertragspartner*innen den Zugang zu Daten und damit den Zugang zu vor- oder nachgelagerten Märkten vorzuenthalten.⁵⁷ Auch zwischen Unternehmen, die noch keine vertraglichen Beziehungen aufgenommen haben, ist eine datenbedingte Abhängigkeit möglich. In solchen Konstellationen ist jedoch die Billigkeitsprüfung von größerer Bedeutung.⁵⁸

Beide Zugangsansprüche verbindet, dass sie eine umfassende Abwägung der Interessen einerseits der Dateninhaber*innen, andererseits des Zugangssuchenden erfordern: § 19 über die sachliche Rechtfertigung, § 20 im Rahmen einer Unbilligkeitsprüfung. Der Gesetzgeber hat zahlreiche Kriterien in seiner Gesetzesbegründung aufgeführt, die dabei relevant werden können.⁵⁹ Diese detailreiche Prüfung muss dabei die konkreten Umstände des Einzelfalls berücksichtigen, u. a. die Form der Daten, deren Verarbeitbarkeit sowie deren Aufbereitung. Außerdem ist entscheidend, zu welchem Zeitpunkt ein Zugang gewährt werden muss: Reicht ein einmaliges Abschöpfen der Daten oder ist eine Echtzeitübertragung inklusive entsprechender Schnittstelle erforderlich? Welche Nutzungsmöglichkeiten müssen dem Zugangssuchenden eingeräumt werden?⁶⁰ Aufgrund dieser Vielzahl an

möglichen Fallgestaltungen, die auf der Rechtfertigungsebene unterschiedlich zu behandeln sind, bleibt abzuwarten, ob die §§ 19 Abs. 2 Nr. 4, 20 Abs. 1a GWB tatsächlich mehr Rechtssicherheit mit sich bringen, welche die Bildung von Daten Kooperations begünstigt.⁶¹

Sie führt außerdem dazu, dass die Rechtsfolgen einer Zugangsgewährung sehr unterschiedlich ausfallen können, je nach konkreten Bedürfnissen des Zugangssuchenden.⁶² Der Anspruch daraus ergibt sich aus § 33 GWB, der eine Beseitigung der Beeinträchtigung, also des Hindernisses beim Zugang zu den Daten, statuiert. Er ist so ausgestaltet, dass er von den Unternehmen als zivilrechtlicher Anspruch einklagbar ist.⁶³ Von großer praktischer Relevanz könnte dabei ein Unterschied zwischen den beiden Ansprüchen bezüglich der Beweislast werden, falls die bisherigen Maßstäbe übertragen werden: Während unter § 19 Abs. 2 Nr. 4 GWB die Dateninhaber*innen ihr überwiegendes Interesse darlegen und beweisen müssen, verpflichtet § 20 Abs. 1a GWB den Zugangssuchenden dazu, eine unbillige Behinderung darzulegen.⁶⁴

Zu beachten sind auch die Akteur*innen, die im Rahmen des Missbrauchsverbots von Relevanz sind. Eine Neuerung der GWB-Novelle war es, dass eine relative Marktmacht und damit ein Abhängigkeitsverhältnis gem. § 20 Abs. 1a GWB in Verbindung mit § 20 Abs. 1 GWB auch zwischen großen Unternehmen bestehen kann. Die Beschränkung auf kleine oder mittlere Unternehmen fiel weg.⁶⁵ Außerdem wurde das Konzept der Intermediationsmacht eingeführt. Digitale Märkte sind auch dadurch geprägt, dass statt eines klassischen Anbieter-Abnehmer-Verhältnisses Unternehmer*innen oftmals mit Intermediär*innen bzw. Vermittler*innen auf mehrseitigen Märkten in wirtschaftlichen Beziehungen stehen. § 18 Abs. 3b sowie § 20 Abs. 1 GWB greifen diese Entwicklung auf und stellen klar, dass auch die „erbrachten Vermittlungsdienstleistungen“ eine Marktmacht etablieren können.⁶⁶ Außerdem wurde die Position der Bundeskartellbehörde durch die Novellierung gestärkt, indem sie u. a. mehr Ermittlungsbefugnisse erhalten hat und die Sanktionsmöglichkeiten sowohl verschärft als auch konkretisiert wurden (vgl. §§ 81 ff. GWB).⁶⁷

Die europäische Ebene des Schutzes vor Marktmissbrauch findet über den Art. 102 AEUV statt und

56 Schweda/ von Schreitter, Ran an die Datenschütze? Datenzugangsansprüche nach der 10. GWB-Novelle, WuW 2021, 145, 147.

57 Deutscher Bundestag, Drucksache 19/23492, S. 80 f.

58 Deutscher Bundestag, Drucksache 19/23492, S. 81.

59 Vgl. Deutscher Bundestag, Drucksache 19/23492, S. 81.

60 Brenner, Kapitel 1: Wettbewerb und Digitalwirtschaft sowie weitere materiellrechtliche Änderungen, in Bien/Käseberg/Klumpke/Körper/Ost (Hrsg.), Die 10. GWB-Novelle, 2021, Rn. 83 ff.

61 Schweda/ von Schreitter, Ran an die Datenschütze? Datenzugangsansprüche nach der 10. GWB-Novelle, WuW 2021, 145, 150 f.

62 Brenner, Kapitel 1: Wettbewerb und Digitalwirtschaft sowie weitere materiellrechtliche Änderungen, in Bien/Käseberg/Klumpke/Körper/Ost (Hrsg.), Die 10. GWB-Novelle, 1. Auflage 2021, Rn. 83.

63 Schweda/ von Schreitter, Ran an die Datenschütze? Datenzugangsansprüche nach der 10. GWB-Novelle, WuW 2021, 145, 154.

64 Schweda/ von Schreitter, Ran an die Datenschütze? Datenzugangsansprüche nach der 10. GWB-Novelle, WuW 2021, 145, 150.

65 Deutscher Bundestag, Drucksache 19/23492, S. 78 f.

66 Käseberg/ Brenner/ Fülling, Das GWB-Digitalisierungsgesetz im Überblick, WuW 2021, 269, 271.

67 Käseberg/ Brenner/ Fülling, Das GWB-Digitalisierungsgesetz im Überblick, WuW 2021, 269, 274; Deutscher Bundestag, Drucksache 19/23492, S. 124 ff.

erfolgt – ebenso wie § 19 Abs. 2 Nr. 4 GWB – über die Essential-Facilities-Doktrin.⁶⁸ Der Anspruch auf Datenzugang hat dabei jedoch noch keine große praktische Bedeutung erlangt.⁶⁹ Dem Europarecht steht zudem durch das Daten-Governance-Gesetz, dem Gesetz über digitale Märkte sowie dem Gesetz über digitale Dienste unter anderem eine grundlegende Neugestaltung der wettbewerbsrechtlichen Aspekte des Zugangs zu Daten bevor, auf die im letzten Kapitel dieses Beitrags näher eingegangen wird. Aus diesen Gründen wird hier Art. 102 AEUV als Zugangsanspruch nicht näher vertieft.

3.3.2 Das Recht offener Daten in der öffentlichen Verwaltung im „zersplitterten“ Mehrebenensystem

Ein Bereich, in dem die Rechte auf Daten weit entwickelt sind, ist die öffentliche Verwaltung.⁷⁰ Seit Jahren haben sich die Erwartungen an Staat und Verwaltung grundlegend gewandelt, auch die Rechtsordnung hat immer mehr Anforderungen dafür formuliert, dass Staat und Verwaltung ihre Daten mit Bürger*innen teilen. Die Herausforderungen der föderalistischen Staatsorganisation und kommunaler Selbstverwaltung führen allerdings dazu, dass es erhebliche Unterschiede zwischen den Teil Rechtsordnungen des Bundes und der Länder gibt. Ein sichtbares Zeichen dafür ist, dass es derzeit 99 verschiedene Open-Data-Portale gibt, die quer durch den deutschen Verwaltungsaufbau reichen.⁷¹

Das jeweilige Leitbild wandelte sich im Laufe der Zeit vom „Amtsgeheimnis“ – häufig abwertend als Arkanprinzip bezeichnet – und dem anschließend herrschenden „Prinzip der beschränkten Aktenöffentlichkeit in Deutschland“ über den Grundsatz der Transparenz hin zum Leitbild eines offenen Regierungs- und Verwaltungshandelns. Institutionell unterfüttert wurde diese Entwicklung dadurch, dass die meisten Datenschutzbeauftragten auch als Transparenz- oder Informationsfreiheitsbeauftragte fungieren und Belange von Offenheit und Transparenz aktiv einbringen. Informations- und

Auskunftsansprüche wurden durch das Europarecht zuerst im Umweltrecht eingeführt und dann auch in das allgemeine Verwaltungsrecht aufgenommen. Wegen dem europarechtlich verankerten Grundsatz der beschränkten Einzelermächtigung und der institutionellen und verfahrensmäßigen Autonomie der Mitgliedsstaaten kann das Europarecht diese Fragen aber nur für einzelne Materien regeln, für die ein Kompetenztitel besteht. Da neben Spezialmaterial wie dem Umweltrecht nur allgemeine Kompetenzen für Maßnahmen zum Funktionieren des Binnenmarktes bestehen, musste die PSI-Richtlinie⁷² auf diese Kompetenzen gestützt werden. Dieser Bezug zum Binnenmarkt erklärt auch die rechtliche Unterscheidung rechtlicher Regime von Zugang und Nutzung der Daten.

Der Zugang zu Daten ist bisher unterschiedlich ausgestaltet. Das Informationsfreiheitsgesetz gewährt den Zugriff auf bestimmte Informationen, die einzeln herausverlangt werden müssen. Im Hinblick auf den Zugang kristallisieren sich bisher unterschiedliche Modelle heraus. Dabei zeichnen sich verschiedene rechtliche Umsetzungsmodelle insbesondere bei den Ländern ab. Zum einen gibt es Rechtsordnungen, die keine rechtlichen Regelungen kennen, wobei sie teilweise trotzdem entsprechende Datenportale errichten.⁷³ Eine zweite Gruppe von Rechtsordnungen verpflichtet zwar nicht zur Veröffentlichung von Daten, beschreibt aber verpflichtende oder nicht verpflichtende Voraussetzungen für die Veröffentlichung.⁷⁴ Eine dritte Gruppe von Rechtsordnungen statuiert eine objektive Rechtspflicht zur Veröffentlichung von Daten.⁷⁵ Dies bedeutet, dass die Behörden zwar eine Pflicht trifft, Bürger*innen aber kein Recht dadurch vermittelt wird, so dass sie den Zugang nicht einklagen können. Eine vierte Gruppe von Rechtsordnungen statuiert nicht nur eine objektive Pflicht zur Veröffentlichung bestimmter Datensätze, sondern räumt Personen des Privatrechts auch ein subjektives Recht ein, dies geltend zu machen.⁷⁶

Durch die Regelung der PSI-Richtlinie ist die Nutzung von Daten stärker vereinheitlicht. Sie wurde für die Bundesverwaltung durch das sog. Datennutzungs-

68 Zuletzt in Bezug auf Schnittstelleninformationen EuG 17.9.2007 – T-201/04, BeckRS 2007, 70806 – Microsoft; ausführlich Schmidt, Zugang zu Daten nach europäischem Kartellrecht, 2020, S. 382ff.

69 Brenner, Kapitel 1: Wettbewerb und Digitalwirtschaft sowie weitere materielle Änderungen, in Bien/Käseberg/Klumpke/Körber/Ost (Hrsg.), Die 10. GWB-Novelle, 2021, Rn. 97 f.

70 Dieser Abschnitt basiert auf den aktualisierten Erkenntnissen, die auch in folgendem Text niedergelegt sind: Djeflal, § 25 Öffentliche Verwaltung und Smart Cities, in Bräutigam/Kraul (Hrsg.), Internet of Things (2021) 925.

71 <https://www.offenesdatenportal.de/dataset/ubersicht-der-open-data-angebote-in-deutschland> (26.1.2021). Zu den Portalen zählen etwa Städte wie Moers und Dormagen, aber auch Frankfurt und Stuttgart, die Kreise Kleve und Viersen, Bundesländer wie Bayern und Hamburg, der Regionalverband Ruhr, die Deutsche Bahn, das Bundesamt für Seeschifffahrt und Hydrographie und das Bundesministerium für Verkehr und Digitale Infrastruktur.

72 Richtlinie (EU) 2019/1024 des Europäischen Parlaments und des Rates vom 20. Juni 2019 über offene Daten und die Weiterverwendung von Informationen des öffentlichen Sektors PE/28/2019/REV/1.

73 Dazu gehören Bayern, wobei hier wohl eine gesetzgeberische Initiative vorbereitet wird, siehe dazu mit Verweis auf entsprechende Gesetze Brandenburg, Hessen, Sachsen-Anhalt, wobei in den letzteren beiden Ländern der entsprechende Gesetzesentwurf noch nicht angenommen wurde, sowie Schleswig-Holstein, siehe Heckmann, Digitale Verwaltung, in Heckmann (Hrsg.), Juris PraxisKommentar Internetrecht7 (2021) 29.

74 § 10 Abs. 1 EgovG BW; § 10 Abs. 1 Gesetz zur Förderung der elektronischen Verwaltung in Bremen; § 8 Abs. 1 EgovG MV; § 16 EgovG NRW; § 17 Abs. 1 E-GovG SL; § 22 Abs. 1 ThürEGovG.

75 § 13 Abs. 1 EgovG Bln; § 8 Abs. 1 SächsEGovG, 12a EGovG Bund.

76 § 1 Abs. 2 HmbTG.

gesetz umgesetzt.⁷⁷ Darin ist jedenfalls der Grundsatz der uneingeschränkten Datennutzung statuiert, der besagt, dass zugängliche Verwaltungsdaten auch genutzt werden dürfen, wobei Nutzung unter Lizenz möglich bleibt. Erfasst sind vom Anwendungsbereich gemäß § 2 DNG auch öffentliche Unternehmen und die Forschung.

Die im DNG verankerten Gestaltungsvorgaben sind von besonderer Bedeutung für das Teilen von Daten. Gemäß § 1 Abs. 1 DBG sollen Daten „soweit möglich nach dem Grundsatz ‚konzeptionell und standardmäßig offen‘ erstellt werden“.⁷⁸ Daraus ergibt sich bereits eine Gestaltungspflicht, deren Voraussetzungen in § 7 DNG weiter erläutert werden. Auch diese Gestaltungspflicht kann eine große Auswirkung auf das Teilen von Daten haben, weil sie die technischen Voraussetzungen des Teilens von Daten adressiert, die oft faktische Hindernisse darstellen.

3.4 Gesetzgeberische Initiativen

Im EU-Recht gab es eine Reihe von Gesetzgebungsiniciativen, die große Auswirkungen auf das Teilen von Daten vorsehen. Zu nennen sind hier insbesondere die Vorschläge eines Daten-Governance-Gesetzes (DGG)⁷⁹, eines Gesetzes über Digitale Dienste (DDG)⁸⁰ und eines Gesetzes über digitale Märkte (DMG)⁸¹. Sie unterstützen einerseits das freiwillige Teilen von Daten insbesondere durch die Regulierung besonderer Organisationen. Andererseits sehen sie im Bereich von Plattformen aber auch bedeutende zwingende Datenzugangsrechte vor. Zudem werden die Verpflichtungen der öffentlichen Verwaltung weiterentwickelt. Die gesetzgeberischen Vorhaben betreffen besonders Konstellationen des freiwilligen Teilens von Daten zwischen Privaten. Allerdings sind auch einige gesetzliche Zugangsrechte vorgeschlagen worden. Im öffentlichen Bereich spielt das Teilen von Daten der Verwaltung die wichtigste Rolle. Demgegenüber werden Zugangsrechte des Staates gegen Private selten behandelt und hier entsprechend nicht näher erläutert.⁸²

3.4.1 Freiwilliges Teilen von Daten zwischen privaten Akteur*innen

Zu den bemerkenswerten Neuerungen des DGG gehört, dass es das Teilen von Daten direkt adressiert und das Teilen nicht regulieren, sondern fördern will. Dem Recht kommt dadurch die Funktion zu, Technik im Hinblick auf Innovation, Entwicklung und Einsatz zu fördern.⁸³ Das DGG möchte dies insbesondere durch die Schaffung von Infrastrukturen aber auch die Förderung bestimmter Akteur*innen erreichen. Aus den Vorarbeiten und der Gesetzesfolgenabschätzung ergibt sich, dass das Gesetz dadurch dazu anregen will, die Potenziale zu heben, die in der Analyse von Massendaten (Big Data) liegen.⁸⁴ Die wesentlichen Probleme, die durch die Regulierung für das Teilen von Daten Privater adressiert werden sollen, liegen im fehlenden Vertrauen privater Akteur*innen in das Teilen von Daten.⁸⁵ Dieses Problem geht das DGG dadurch an, dass es Organisationen durch ihre rechtliche Regulierung so ausgestaltet, dass ihnen ein höheres Vertrauen entgegengebracht werden kann. Im Falle der Datenintermediär*innen hebt das DGG faktisch ein neues Geschäftsmodell aus dem Boden, das in verschiedenen Akteurskonstellationen das Teilen von Daten erleichtern soll. Im Falle des Datenaltruismus werden gemeinnützige Organisationen geregelt, die ebenfalls das Teilen von Daten erleichtern sollen.

3.4.1.1 Datenintermediär*innen

Mit den Datenintermediär*innen definiert das DGG ein eigenes Geschäftsmodell, das sich gemäß Art. 9 DGG auf drei Konstellationen bezieht: erstens auf eine Vermittlung zwischen den Dateneinhaber*innen und Datennutzer*innen (a), zweitens auf eine Vermittlung zwischen Datensubjekten im Sinne des Datenschutzrechts und ihren personenbezogenen Daten oder natürlichen Personen, die andere Daten verfügbar machen (b), oder drittens Datengenossenschaften (c). In jedem der Fälle übernehmen die Datenintermediär*innen eine Vermittlungsfunktion durch seine

⁷⁷ Datennutzungsgesetz vom 16. Juli 2021 (BGBl. I S. 2941, 2942., 4114) (DNG).

⁷⁸ Diese Gestaltungsvorgabe setzt damit die gleichlautende Aufforderung aus Art. 5 Abs. 2 der PSI-Richtlinie um, „Dokumente nach dem Grundsatz ‚konzeptionell und standardmäßig offen‘ (open by design and by default) zu erstellen und zur Verfügung zu stellen“.

⁷⁹ Vorschlag für eine VERORDNUNG DES EUROPÄISCHEN PARLAMENTS UND DES RATES über europäische Daten-Governance (Daten-Governance-Gesetz) COM/2020/767 final Hierbei berücksichtigen wir die vorläufige Version des interinstitutionellen Kompromisses vom 15.12.2021, siehe <https://data.consilium.europa.eu/doc/document/ST-14606-2021-INIT/en/pdf>.

⁸⁰ Vorschlag für eine VERORDNUNG DES EUROPÄISCHEN PARLAMENTS UND DES RATES über einen Binnenmarkt für digitale Dienste (Gesetz über digitale Dienste) und zur Änderung der Richtlinie 2000/31/EG COM/2020/825 final. Dieses Gesetzgebungsverfahren läuft zum Zeitpunkt der Fertigstellung des Gutachtens noch, daher wird vorwiegend auf den Entwurf der Kommission Bezug genommen. Zur Position des Rats der Europäischen Union siehe Interinstitutional File: 2020/0374(COD). Zur Position des Parlaments siehe Amendments adopted by the European Parliament on 15 December 2021 on the proposal for a regulation of the European Parliament and of the Council on contestable and fair markets in the digital sector (Digital Markets Act) (COM(2020)0842 – C9-0419/2020 – 2020/0374(COD)).

⁸¹ Vorschlag für eine VERORDNUNG DES EUROPÄISCHEN PARLAMENTS UND DES RATES über bestrebbare und faire Märkte im digitalen Sektor (Gesetz über digitale Märkte) COM/2020/842 final. Dieses Gesetzgebungsverfahren läuft zum Zeitpunkt der Fertigstellung des Gutachtens noch, daher wird vorwiegend auf den Entwurf der Kommission Bezug genommen.

⁸² Siehe dazu etwa Richter, Zugang des Staates zu Daten der Privatwirtschaft, ZRP 2020, 245 Eine Ausnahme bildet hier zum Beispiel der Zugang von Forscher*innen zu Daten sehr großer Onlineplattformen gemäß Art. 31 DDG, der gemäß Abs. 2 zum Zweck der „Durchführung von Forschungsarbeiten, die zur Ermittlung und zum Verständnis systemischer Risiken... beitragen“, gewährt und erzwungen werden kann.

⁸³ Diefftal, Normative Leitlinien Für Künstliche Intelligenz in Regierung und Verwaltung, in Mohabbat Kar/Thapa/Parycek (Hrsg.), »(Un)Berechenbar? Algorithmen und Automatisierung in Staat und Gesellschaft« (2018) 493 (503 f.).

⁸⁴ Europäische Kommission, Impact Assessment Report. Accompanying the document Proposal for A Regulation of the European Parliament and of the Council on European data governance (Data Governance Act) (25.11.2020) 1 – 2, 8.

⁸⁵ Europäische Kommission, Impact Assessment Report 11 ff.

besondere Vertrauensstellung. Regulatorisch wird dieses Vertrauen durch ihre Notifizierungsverfahren unterstützt, in dessen Folge die zuständige Behörde das Vorliegen der Notifizierung (Abs. 7) und nach weiterer Prüfung (Abs. 7a = Abs. 8) das Vorliegen auch der weiteren Voraussetzungen für Datenintermediär*innen bestätigt. Nach Prüfung aller Voraussetzungen darf sich das Unternehmen als anerkannter Datenintermediationsdienstleister in der Europäischen Union betätigen.⁸⁶ Alle notifizierten Unternehmen werden ferner von der Europäischen Kommission in einem Register öffentlich gemacht. Damit wird sowohl Publizität hergestellt als auch ein offizielles „Vertrauenssiegel“ vergeben.

Das Geschäftsmodell und seine besonderen Vertrauensgesichtspunkte lassen sich aus den rechtlichen Anforderungen des Art. 11 ablesen. Datenintermediation soll auch bei größeren Unternehmen ein „alleinstehendes Angebot“ sein. Dienste dürfen die geteilten Daten nicht zu anderen Zwecken verwenden (Nr. 1), auch die Verwendung der Nutzungsdaten ist auf das Notwendige zur Erbringung und Verbesserung der Dienste beschränkt (Nr. 2). Es dürfen etwa keine Preisnachlässe mit der Nutzung anderer Dienste eines Dienstleisters gewährt werden. Ferner bedürfen Zusatzdienste, die das Teilen erleichtern, wie Zwischenspeicherung oder Anonymisierung, der Zustimmung der Dateninhaber*innen und -subjekte. Der Zugang soll fair, transparent und nichtdiskriminierend sein (Nr. 4a), ferner soll kontinuierlicher Zugang gewährleistet werden, wobei selbst im Insolvenzfall für Dateninhaber*innen und Datensubjekte noch Zugang zu ihren Daten bestehen soll. Die Datenintermediär*innen treffen auch besondere Schutz- und Verhaltenspflichten. Sie müssen aktiv den Missbrauch ihrer Dienste verhindern (Nr. 5), im Rahmen ihrer Pflicht zur Gewährleistung der IT-Sicherheit für sensitive Daten sogar das „höchste Sicherheitslevel“ bieten (Nr. 8). Im Kontext von personenbezogenen Daten sollen sie gar „im Interesse“ der Datensubjekte handeln (Nr. 10). Aus diesen Pflichten folgt, dass Datenintermediär*innen nicht nur kommerziellen Interessen folgen, sondern in bedeutendem Umfang auch öffentliche Interessen wahrnehmen sollen.⁸⁷ Dabei bietet Art. 13 DGG den Behörden verschiedene Durchsetzungsmöglichkeiten bis hin zu Bußgeldern oder einem Lizenzentzug. Diese Durchsetzungsmöglichkeiten können aber nicht darüber hinwegtäuschen, dass das angedachte Geschäftsmodell nur funktionieren kann, wenn sich genügend Unternehmen finden, die es auch mit Leben füllen wollen.

3.4.1.2 Datenaltruismus

Der in Art. 14a – 22 DGG geregelte Datenaltruismus betrifft das freiwillige, nicht-kommerzielle Teilen von Daten aus Gründen des Allgemeininteresses. Das DGG will auch hier Organisationen mit entsprechendem Zweck rechtlich ausgestalten. Diese Organisationen sollen wiederum das Teilen von Daten durch „Spenden“ erleichtern. Auch im Falle des Datenaltruismus soll mittels eines Registers Publizität hergestellt werden, allerdings sind für diesen Fall nationale Register wichtiger. Die Publizität wird gemäß Art. 15 Abs. (3=2) durch ein Kennzeichen mit Bezeichnung, Logo, eigenem Design und einem QR-Code hergestellt. Die nicht-kommerzielle Operation der jeweiligen Organisation ist beim Datenaltruismus bereits Voraussetzung für eine Registrierung, wie sich aus Art. 16 (b)&(c) ergibt. Datenaltruismusorganisationen treffen insbesondere umfassende Transparenzpflichten gemäß Art. 18, die sich auf alle wichtigen Geschäftsprozesse und technischen Verfahren beziehen. Diese reichen bis zu einer Pflicht zur Veröffentlichung der Nutzer*innen der Daten (c) und einer Zusammenfassung der Ergebnisse der Datenverarbeitung (d). Was die Prozesse angeht, wird die Kommission durch Art. 19a (=20) DGG ermächtigt, ein Regelwerk zu erstellen, das technische Anforderungen im Hinblick auf Datenschutz, Technik, IT-Sicherheit und Kommunikation aufstellt und Empfehlungen hinsichtlich Interoperabilitätsstandards erstellt. In die gleiche Richtung geht eine standardisierte und modularisierte Einwilligungserklärung, die von der Europäischen Kommission gemäß Art. 22 erstellt werden soll. Im Bereich des Datenaltruismus macht der europäische Gesetzgeber also detailliertere Vorgaben.

3.4.2 Pflichten zum Teilen von Daten zwischen privaten Akteur*innen

Insbesondere zur Gewährleistung eines fairen Wettbewerbs und zu Zwecken des Verbraucherschutzes führt das DMG zahlreiche Pflichten zum Datenteilen zwischen privaten Akteur*innen ein. Grundsätzlich betreffen diese Pflichten zentrale Plattformdienste, die gemäß Art. 2 Nr. 2 phänomenologisch nach bestimmten Erscheinungsformen definiert werden.⁸⁸ Das DMG ist auf Gatekeeper*innen anwendbar, die grundsätzlich nach den drei qualitativen Kriterien des Art. 3 Abs. 1 DMG definiert werden. Diese Gatekeeper*innen treffen einerseits bestimmte unmittelbare Pflichten nach Art. 5 und ggf. noch näher auszuführende Pflichten des Art. 6. Manche Pflichten beziehen sich

⁸⁶ Dies ist eine eigene Übersetzung des autoritativen englischen Wortlauts.

⁸⁷ Allerdings ist bereits Kritik laut geworden, die dem DGG vorwirft, dass ihre Vereinbarkeit mit der Datenschutzgrundverordnung nicht ausreichend geklärt ist. European Data Protection Board, EDPB-EDPS Joint Opinion 03/2021 on the Proposal for a regulation of the European Parliament and of the Council on European data governance (Data Governance Act) Version 1.1, https://edpb.europa.eu/system/files/2021-03/edpb-edps_joint_opinion_dga_en.pdf (abgefragt am 31. 1. 2022).

⁸⁸ Dazu gehören „Online-Vermittlungsdienste, Online-Suchmaschinen, Online-Dienste sozialer Netzwerke, Video-Sharing-Plattform-Dienste, nummernunabhängige interpersonelle Kommunikationsdienste, Betriebssysteme, Cloud-Computing-Dienste, Werbedienste, einschließlich Werbenetzwerken, Werbebörsen und sonstiger Werbewerstattungsdienste...“.

dabei direkt auf das Teilen von Daten. Art. 5 enthält für Gatekeeper*innen im Wesentlichen ein Verbot des Datenteilens innerhalb des Konzerns in lit. a). In Art. 6 lit. h) werden Gatekeeper*innen zur Ermöglichung der Datenportabilität gezwungen. Zentral für Fragen des Datenteilens sind die folgenden Pflichten für Gatekeeper*innen

„i) gewerblichen Nutzern sowie von einem gewerblichen Nutzer zugelassenen Dritten kostenlos einen effektiven, hochwertigen und permanenten Echtzeit Zugang gewähren und die Nutzung aggregierter oder nicht aggregierter Daten ermöglichen, die im Zusammenhang mit der Nutzung der relevanten zentralen Plattform Dienste durch diese gewerblichen Nutzer und die Endnutzer, die die Produkte oder Dienste dieser gewerblichen Nutzer in Anspruch nehmen, bereitgestellt oder generiert werden; den Zugang zu und die Nutzung von personenbezogenen Daten nur ermöglichen, soweit dies unmittelbar mit der Nutzung der Produkte oder Dienstleistungen durch den Endnutzer im Zusammenhang steht, die der betreffende gewerbliche Nutzer über den relevanten zentralen Plattformdienst anbietet, und sofern der Endnutzer dem Datenaustausch durch eine Einwilligung im Sinne der Verordnung (EU) 2016/679 zugestimmt hat;

j) Dritten, die Online-Suchmaschinen betreiben, auf deren Antrag hin zu fairen, zumutbaren und diskriminierungsfreien Bedingungen Zugang zu Ranking-, Such-, Klick- und Anzeigedaten in Bezug auf unbezahlte und bezahlte Suchergebnisse gewähren, die von Endnutzern in Online-Suchmaschinen des Gatekeepers generiert werden, vorbehaltlich der Anonymisierung der Such-, Klick- und Anzeigedaten, bei denen es sich um personenbezogene Daten handelt;“

Dies sind Beispiele, wie Gatekeeper*innen durch das Teilen von Daten Wettbewerb ermöglichen sollen. Bezeichnend sind bereits die Anforderungen eines „effektiven, hochwertigen und permanenten Echtzeit Zugangs“ in zeitlicher und qualitativer Hinsicht.

3.4.3 Das Teilen von Daten in der Verwaltung

Das DGG adressiert im Hinblick auf die Verwaltungen unter anderem⁸⁹ das Problem, dass das Teilen von Daten auch durch rechtliche Gestaltung unmöglich gemacht werden kann, insbesondere, wenn Rechte an den Daten bestehen oder durch die Verwaltung eingeräumt werden. Dies umfasst geschäftliche Geheimhaltung, statistische Geheimhaltung, geistiges Eigentum

und den Schutz personenbezogener Daten. Art. 4 DGG beinhaltet das grundsätzliche Verbot der Einräumung exklusiver Datennutzungsrechte an Dritte und regelt darüber hinaus die Ausnahmen zu diesem Verbot. Dadurch soll für die Zukunft verhindert werden, dass exklusive Nutzungsrechte vereinbart werden, wo das nicht strikt notwendig ist. Aus der Vorschrift folgt in Verbindung mit § 59 Abs. 1 VwGO i. V. m. § 134 BGB⁹⁰, dass anderslautende Ausschließlichkeitsvereinbarungen nichtig sind. Art. 5 DGG stellt für die Fälle, in denen der Schutz nach Art. 3 Abs. 1 wirksam ist, ein Regime von Voraussetzungen auf, unter denen Daten trotzdem geteilt werden dürfen. Art. 5 Abs. 3 DGG sieht dabei die Pflicht für die Verwaltung vor, eine Nutzung auch geschützter Daten durch eine entsprechende technische Gestaltung zu ermöglichen. Exemplarisch erwähnt werden in Art. 5 Abs. 3 DGG eine Modifikation der Daten zum Schutz der Rechte Dritter an den Daten, die Schaffung einer sicheren Nutzungsumgebung für den Fernzugriff auf die Daten oder die sichere Nutzung der Daten an ihrem Speicherort. Hierbei ist ausdrücklich ein sehr hohes IT-Sicherheitsniveau zu gewährleisten. Ist das unmöglich, soll die Verwaltung bei der Erlangung der entsprechenden Genehmigungen, Einwilligungen oder Lizenzen der Rechteinhaber*innen assistieren, ohne dabei deren Anonymität zu gefährden.

Insbesondere die Anforderungen des Art. 5 Abs. 3 DGG stellen öffentliche Verwaltungen in ihrer Breite vor große Herausforderungen. Um dem zu begegnen, sollen gemäß Art. 7 Abs. 1 DGG Zuständigkeiten zur Unterstützung in diesen Prozessen festgelegt werden. Dabei müssen die zuständigen Stellen unterschiedliche Unterstützungsleistungen vorhalten, die so weit gehen, dass die entsprechenden Stellen selbst technisch als „Datenintermediär*innen“ tätig werden, indem sie geschützte Verwaltungsdaten Dritten zugänglich machen. In Erwägungsgrund 19 wird darüber hinaus noch die Idee formuliert, dass „besonders sensitive Daten auch besonders geschützt werden müssten“. Als Beispiel werden dabei besondere Gesundheitsdaten genannt. Um hier ein hohes Schutzniveau zu garantieren, wird ein „Europäischer Gesundheitsdaten Raum“ ins Spiel gebracht. Insbesondere vor dem Hintergrund des deutschen Föderalismus ist es relevant, dass Art. 8 DGG eine*n einheitliche*n Ansprechpartner*in vorsieht, um Informationen über den Zugang bereitzustellen. Gemäß Erwägungsgrund 21 kann dies auch dahingehend ausgestaltet werden, dass entsprechende Datensätze automatisiert zur Verfügung gestellt werden, etwa durch ein Open-Data-Portal.

⁸⁹ Hervorgehoben werden daneben aber auch technische und kulturelle Hindernisse, siehe Europäische Kommission, Impact Assessment Report 13 ff.

⁹⁰ Zur Anwendbarkeit dieser Norm siehe Bonk/Neumann/Siegel in Stelkens/Bonk/Sachs, Verwaltungsverfahrensgesetz. Kommentar 9. Auflage 2018 (2018) VwVfG § 59 Nichtigkeit des öffentlich-rechtlichen Vertrags Rz 9; Ramsauer in Ramsauer/Kopp, Verwaltungsverfahrensgesetz. Kommentar 22 (2021) § 59 Rz 9.

3.5 Zusammenfassung

In der Zusammenschau der Ergebnisse dieses Gutachtens sollten folgende Punkte festgehalten werden:

Die aktuelle Rechtslage definiert Rechte an Daten insbesondere im Datenschutzrecht und im Recht des geistigen Eigentums. Diese Rechte an Daten sind auch auf das Teilen von Daten anwendbar. Sowohl das Datenschutzrecht als auch das Recht des geistigen Eigentums verfolgen jedenfalls auch den Zweck, das Teilen von Daten zu ermöglichen. Allerdings enthalten sie keine spezifischen Regelungen in dieser Hinsicht. Rechte auf Daten finden sich im geltenden Recht insbesondere im Verwaltungsrecht, das aufgrund des föderalen Verwaltungsaufbaus zersplittert ist, sich aber in zunehmendem Maße am Leitbild eines offenen Regierungs- und Verwaltungshandelns ausrichtet. Erst seit Kurzem erstarkt dieser Bereich auch im Wettbewerbsrecht.

Auf der Ebene der Europäischen Union hat die Europäische Kommission einen Reformprozess angestoßen und eine Reihe von Gesetzesvorschlägen eingebracht, die im Falle ihrer Verabschiedung große Auswirkungen auf das Teilen von Daten hätten. Hervorzuheben sind hierbei insbesondere das Daten-Governance-Gesetz und das Gesetz über Digitale Märkte. Beide Gesetze zielen direkt und unmittelbar darauf ab, das Teilen von Daten zu ermöglichen. Dabei werden Datenzugangsrechte sowohl für Gatekeeper*innen im Bereich von Plattformdiensten als auch für die öffentliche

Verwaltung formuliert. Daneben versucht insbesondere das Daten-Governance-Gesetz neue Akteur*innen zu schaffen, die besonders vertrauenswürdig sind und so das Teilen von Daten fördern. Das Recht befindet sich folglich in einem Wandlungsprozess. Neben der Regelung der Rechte an Daten will insbesondere das Daten-Governance-Gesetz sozio-technische Strukturen schaffen, die das Teilen von Daten fördern.

Hervorzuheben ist im Hinblick auf das Daten-Governance-Gesetz, dass das Recht dabei weder mit Instrumenten der Technikregulierung arbeitet noch unmittelbare Rechte an den Daten einräumt. Das Recht wird also nicht in seiner begrenzenden, sondern in seiner Technik ermöglichenden Funktion eingesetzt. Bewusst gefördert werden Informationsintermediär*innen in der Wirtschaft, der Verwaltung und der Zivilgesellschaft, die selbst nicht direkt an der Analyse der Daten beteiligt sein sollen. Bezeichnend ist auch, dass sowohl der Gesetzgeber als auch die im Gesetz vorgesehenen Behörden für einen aktiveren Ansatz der staatlichen Begleitung der Regulierung stehen. Der Gesetzgeber setzt nicht mehr nur Rahmen, sondern greift direkt in Organisations- und Geschäftsmodelle ein. Ob die vorgesehenen Mechanismen funktionieren und sich die Erwartungen an das Potential von Massendaten tatsächlich realisieren, kann nicht beurteilt werden. Jedoch zeigt sich deutlich, dass das Recht durch ein breites Arsenal verschiedener Instrumente zum Einsatz der Technik sowie zur Schöpfung ihrer Potentiale beitragen soll.

LITERATUR-
BASIERTE
ANALYSE:
LANDSCAPE-
UND
GAP-ANALYSE



4. Literatur-basierte Analyse: Landscape- und Gap-Analyse

Dieses Kapitel umfasst die Literatur-basierten Landscape- und Gap-Analysen (LGA) für die vier Untersuchungsbereiche *Wissenschaft*, *Public Sector*, *Wirtschaft* und *Qualifizierung*. Für die ersten drei LGA gilt die im nächsten Absatz beschriebene Darstellung der einzelnen Kapitel. Diese Überschriften stellen die diskursiv gewonnenen, bereichsübergreifenden Kernkategorien dar. Der Untersuchungsbereich *Qualifizierung* weicht als Querschnittsbereich von den anderen ab, da hier der Fokus auf Qualifizierung im Umgang mit Daten statt auf Datenteilen und -nutzen liegt.

Die LGA wurden in sechs Unterpunkte gegliedert. Unter (1) Quellengrundlagen ist die Auswahl der Quellen erklärt, die sich aufgrund der dynamischen Entwicklung des Themas „Datenteilen“ vor allem auf jüngere Veröffentlichungen stützt. In Unterkapitel (2) Rahmenbedingungen werden Strategiepapiere und Vorgaben untersucht, die sich mit Datenteilen und -nachnutzen befassen bzw. dazu animieren. Der Unterpunkt (3) Infrastruktur diskutiert bestehende Strukturen und Netzwerke, in denen Daten geteilt und genutzt werden. Punkt (4) trägt in allen Kapiteln eine andere Bezeichnung. Hier handelt es sich um eine „bereichsspezifische Rationalität“, die das Datenteilen im jeweiligen Untersuchungsbereich beeinflusst. Unter (5) Defizite und Handlungsbedarf werden Hürden und Empfehlungen mit Blick auf übergreifende FAIRe Datenproduktion und -nachnutzung sowie der Förderung zum Teilen und Nachnutzen von Daten beschrieben. Im letzten Unterkapitel (6) sind die Kernpunkte aus der Landscape- und Gap-Analyse zusammengefasst.

4.1 Wissenschaft

4.1.1 Quellengrundlagen

Im Untersuchungsbereich Wissenschaft gibt es eine Vielzahl an Artikeln, Büchern und Empfehlungen zur guten wissenschaftlichen Praxis und zum Datenmanagement. Sie alle im Rahmen von EcoDM zu analysieren, ist schlichtweg nicht möglich, weshalb an dieser Stelle eine gewisse Abstraktion vorgenommen wird. Während in der *Wirtschaft* beispielsweise zwischen Wirtschaftssektoren unterschieden wird, sind es in der *Wissenschaft* die verschiedenen Disziplinen, die jeweils gewisse Autonomie in der Ausgestaltung ihres Fachgebiets besitzen. Eine Literaturanalyse aller fachspezifischen Richtlinien und Empfehlungen zum Datenmanagement ist an dieser Stelle nicht sinnvoll,

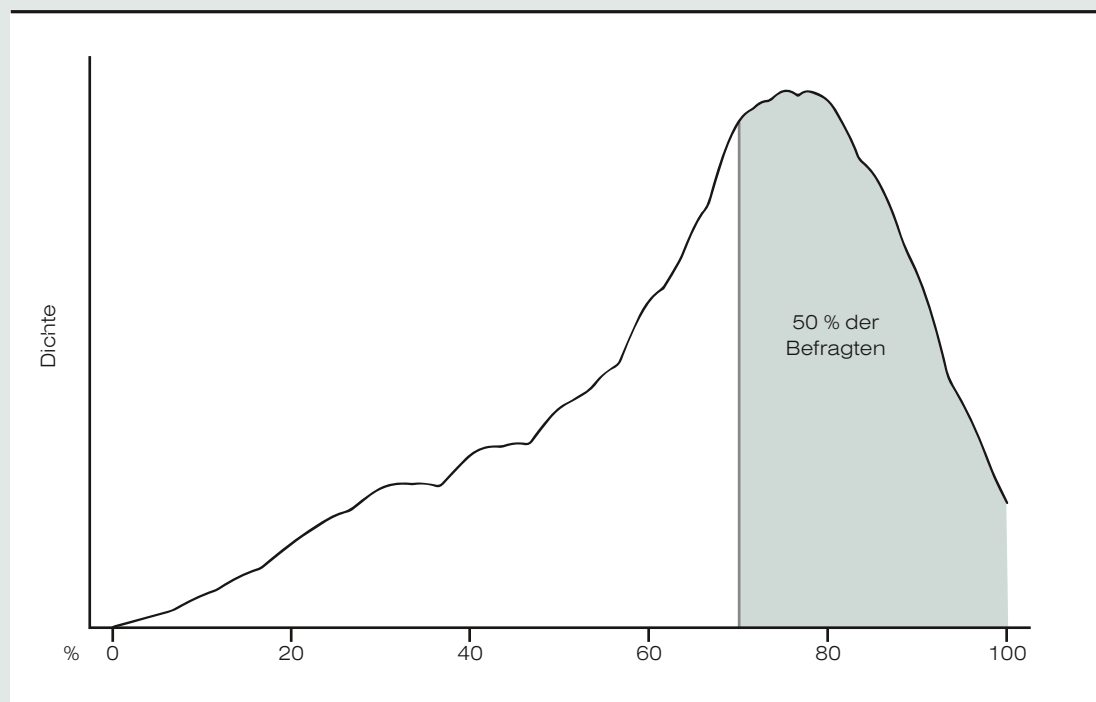
weil diese wegen ihres Umfangs den Blick auf die Gemeinsamkeiten und Unterschiede mit den anderen Untersuchungsbereichen *Public Sector*, *Wirtschaft* und *Qualifizierung* trüben würde. Um Vergleichbarkeit zwischen den Bereichen herzustellen, werden ausschließlich aktuelle Publikationen erwähnt, welche sich mit der Makroebene befassen: Es werden Werke analysiert, die sich mit Datenmanagement und guter wissenschaftlicher Praxis für das gesamte Wissenschaftssystem auseinandersetzen, ohne tief in einzelne Disziplinen einzutauchen. Insbesondere werden aktuelle Richtlinien und Empfehlungen ab 2016 aus Deutschland und der EU berücksichtigt.

4.1.2 Rahmenbedingungen

Die Digitalisierung hat das Wissenschaftssystem grundlegend verändert und durchdrungen. Die Covid-19-Krise beschleunigt diese Transformationsprozesse (Wissenschaftsrat, 2021, S. 7). Wissenschaftliche Recherche, das Publizieren und auch die Lehre sind ohne das Internet undenkbar. Immer mehr Daten sind rund um die Uhr abrufbar und kollaboratives Arbeiten wird auch digital zur Norm. „Erstens ist ein stetiger und rasanter Zuwachs an Forschungsdaten zu verzeichnen, und zweitens ergibt sich – damit eng verbunden – eine stark wachsende Nutzung von Software“ (Deutsche Forschungsgemeinschaft, 2020, S. 5). Im Forschungsbetrieb werden immer mehr verschiedene digitale Infrastrukturen und Programme verwendet. Beispielsweise werden digitale Office-Lösungen und Betriebssysteme von Microsoft, Google oder anderen Stakeholdern verwendet, dazu kommen wissenschaftsspezifische Archivierungsinfrastrukturen wie Zenodo und diverse andere Software wie etwa MaxQDA oder F4, die den Forschungsprozess unterstützen. Plattform- und Softwarebetreiber entscheiden, wie welche Datensätze erstellt, geteilt, verarbeitet und verteilt werden können, wobei eine Wechselwirkung zwischen Angebot und Nachfrage besteht: Privatwirtschaftliche Akteur*innen wirken als „De-Facto-Regulierer“ in Bereichen des Datenmanagements (Wissenschaftsrat, 2021, S. 73). Die spezifischen Angebote der Infrastrukturbetreiber*innen von Software- und Clouddiensten bestimmen, wie wir mit Daten umgehen können. Anders als bei einem vorhandenen Buch reicht es nicht aus, lesen zu können, um zu verstehen. Digitale Daten benötigen Programme, die sie erstellen und ins Lesbare aufschlüsseln. Die Hardware, Betriebssysteme und oft auch die Software sind von privatwirtschaftlichen Akteur*innen gestaltet und betrieben, auch wenn es Ausnahmen

Belastbarkeit des Wissensbestandes

alle Wissenschaftler/innen



Fragetext: In einigen Bereichen wird die Zuverlässigkeit des wissenschaftlichen Wissens z. B. unter dem Stichwort „Replikationskrise“ diskutiert. Wenn Sie auf Ihr Fachgebiet schauen: Welchen Umfang des Wissensbestandes halten Sie prinzipiell für belastbar?
Quelle: DZHW Wissenschaftsbefragung 2019/20, n=6.946

Abbildung 3: Belastbarkeit des Wissensstandes (Quelle: DZHW Wissenschaftsbefragung 2019/20, n=6.946, nach: Ambrasat & Heger, 2020, S. 14.; Lizenz: CC BY-SA 3.0 DE, <https://creativecommons.org/licenses/by-sa/3.0/de/>)

gibt (Liste von Open-Source-Hardware für Computer¹, linux.org, f-droid.org, ...).

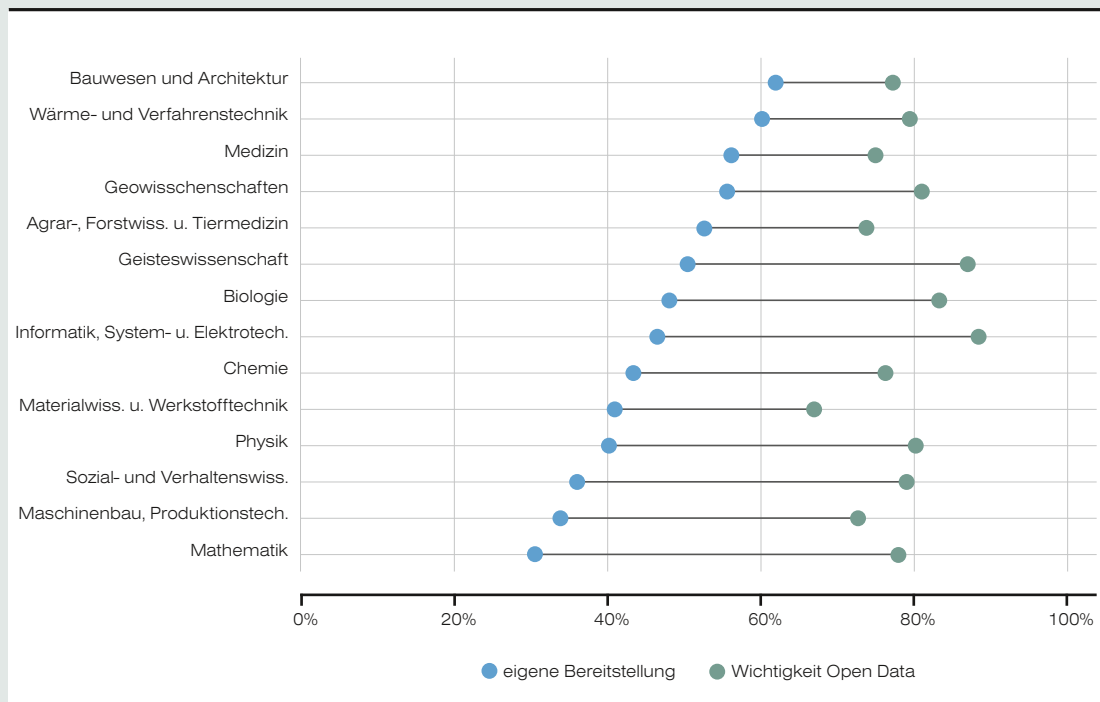
Das Internet hat den Zugang zu Information demokratisiert und traditionelle Gatekeeper*innen von Wissen obsolet gemacht (International Science Council, 2020, S. 2). Es ist wie eine unendliche Bibliothek, die sich selbst konstituiert, ohne traditionelle Autoritäten beachten zu müssen. Wissenschaftliche Fakten und privates Alltagswissen sind im selben Netz nebeneinander verfügbar, wobei der Zugang zu Alltagswissen oft leichter zu finden ist. Deshalb muss sich die Wissenschaft stärker in der Öffentlichkeit engagieren, um nicht übersehen zu werden und weiterhin zur menschlichen Entwicklung beitragen zu können (ebd.). Jede*r kann online Daten veröffentlichen, wodurch auch eine neue Unübersichtlichkeit entsteht. Daten und Ergebnisse, die nicht zugänglich sind, laufen andererseits Gefahr, ignoriert zu werden. Mit dieser neuen Realität ist es für das Wissenschaftssystem schwieriger, am gesellschaftlichen Fortschritt mitzuwirken. Desinformation trübt das Vertrauen gegenüber der Wissenschaft, wobei man dem durch kritische Medienkompetenz entgegenwirken kann. Diese sollte bei Wissenschaftler*innen besonders ausgeprägt

sein, weil sie ähnliche Fähigkeiten nutzen, wenn sie die Studien ihrer Kolleg*innen bewerten. Gleichzeitig sind wissenschaftliche Informationen oft nur eben dieser wissenschaftlichen Teilöffentlichkeit zugänglich. Wissenschaftskommunikation an eine größere Öffentlichkeit wird in den unterschiedlichen Disziplinen verschieden stark forciert. Der Ansatz der Citizen Science (Bürger*innenwissenschaften) kann auch als vertrauensbildende Maßnahme dienen, der Bürger*innen proaktiv und kooperativ in die Forschung miteinbezieht. Wie Citizen Science auf Augenhöhe stattfinden kann, ohne Verlust von Datenqualität, wird in diversen Disziplinen diskutiert. Exemplarisch seien hier Beispiele aus den Bibliothekswissenschaften (Wiederkehr, 2021) und aus der Biologie (Anhalt-Depies et al., 2019) genannt. Im Feld der Citizen Science gibt es starke Überschneidungsmöglichkeiten mit dem *Public Sector*, da durch Daten von Bürger*innen und Wissenschaft auch Verwaltungsprozesse effizienter sowie nachhaltiger gestaltet werden können (vgl. Hodgkinson et al., 2021). Doch das Vertrauen in die Informationsqualität wissenschaftlicher Publikationen ist auch innerhalb der Wissenschaft erschüttert (Abbildung 3).

¹ https://en.wikipedia.org/wiki/List_of_open-source_computing_hardware

Open Data – Bedeutung und eigene Bereitstellungspraxis

nur Befragte mit eigenen Forschungsdaten



Bedeutung: Relative Häufigkeit der Antworten „wichtig“ und „sehr wichtig“.

Quelle: DZHW Wissenschaftsbefragung 2019/20, n=1.584, modularisiert.

Abbildung 4: Open Data – Bedeutung und eigene Bereitstellungspraxis (Quelle: DZHW Wissenschaftsbefragung 2019/20, n=1.584, modularisiert, nach: Ambrasat & Heger, 2020, S. 33.; Lizenz: CC BY-S. 3.0 DE, <https://creativecommons.org/licenses/by-sa/3.0/de/>)

Es gibt Unterschiede zwischen den Fachdisziplinen, doch die Hälfte der Befragten aus dem Wissenschaftsbetrieb halten höchstens 70 % oder weniger des Wissensbestands für belastbar. Höhere Werte finden sich beispielsweise in der Mathematik, während man sich in der Psychologie mit nur 53,4 % am Unsichersten zeigt (Ambrasat & Heger, 2020, S. 15). Das fehlende Vertrauen in die Belastbarkeit der akademischen Wissensbestände lässt sich nicht auf einen einzelnen Faktor zurückführen. Ein wesentliches Problem ist, dass viele Forschungsergebnisse nicht replizierbar und damit nicht nachvollziehbar sind. Die in Kapitel 1.3.3 beschriebenen FAIR-Prinzipien lassen sich als mögliche Antwort aus der Wissenschaft auf diese Replikationskrise deuten, die durch die Digitalisierung nochmal verstärkt wird (vgl. auch, Wilkinson et al., 2016). Die gravierenden Veränderungen in der Informationsverarbeitung lassen auch die Politik Maßnahmen für die bessere Qualität von Forschungsdaten ergreifen (vgl. Kapitel 3.4). Um in Zukunft bessere datenbasierte Entscheidungen treffen zu können, werden vermehrt Daten gesammelt und ausgewertet. Das betrifft nicht nur Verwaltungsdaten, sondern Daten aus allen Bereichen des gesellschaftlichen Lebens, wobei ständig neue Bereiche erschlossen werden. Kuhlmann et al. (2021) sehen digitale Forschungsdaten als die essenzielle

Grundlage für wissenschaftliche Politikberatung. Sowohl die deutschen als auch die europäischen Pläne zur Verwendung von Daten sind ambitioniert. Die Datenstrategien (Bundesregierung, 2021a; Europäische Kommission, 2020a) sehen Maßnahmen im Bereich der Infrastruktur und Qualifizierung vor, aber auch die Erschließung von neuen Datenräumen, insbesondere im öffentlichen Bereich. All diese Veränderungen sind auch für die Wissenschaft hoch relevant, da man aus den öffentlichen Datenbeständen neue Forschungsfragen und -daten generieren kann. Die FAIR-Prinzipien wurden von diversen Stakeholdern zur besseren Datenanalyse aufgegriffen und damit sind sie das neue Kriterium für Datenqualität, das es durchzusetzen gilt. Die Zusammenarbeit mit Wissenschaftler*innen aus der Data Science wird weiter forciert. Der jährliche Verlust im europäischen Wissenschaftssystem durch Daten, die nicht den FAIR-Prinzipien entsprechen, wird auf 26.2 Milliarden Euro geschätzt (Directorate-General for Research and Innovation (European Commission) & PwC EU Services, 2018, S. 4). Eine der derzeit gültigen Maßnahmen ist die Richtlinie über offene Daten und die Weiterverwendung von Informationen des öffentlichen Sektors (Richtlinie (EU) 2019/1024, 2019). Diese Richtlinie sieht vor, dass Forschungsdaten „so offen

wie möglich, so geschlossen wie nötig“ publiziert werden (ebd., Art. 10(1)). Dieser Paradigmenwechsel wird in den wissenschaftlichen Disziplinen unterschiedlich assimiliert. In der oben erwähnten Wissenschaftsbefragung zeigt sich, dass der Großteil der Forschenden offene Wissenschaft ideell unterstützt, doch die offene Bereitstellung der eigenen Daten ist damit noch nicht garantiert (Abbildung 4).

Der Grad der Transparenz und Offenheit variiert stark und ab wann die Schutzwürdigkeit von Daten eine Rolle spielt, ist oft schwer zu bestimmen (vgl. Kapitel 3.2). Ethische, rechtliche und technische Gegebenheiten müssen berücksichtigt werden. In Forschungsprojekten der EU, wie Horizon Europe, ist deshalb ein Datenmanagementplan anzulegen (Burgelman et al., 2020, S. 3). Selbst wenn nicht alle Forschungsdaten offengelegt werden, sind diese im Datenmanagementplan erwähnt. Schon bei der Planung des Forschungsvorhabens soll erfasst werden, welche Daten und Metadaten erhoben werden sollen und wie es um deren Nachnutzung steht (ebd.). Es wird festgestellt, dass der Publikationsmarkt die Bereitstellung von Forschungsdaten bremst (Oßwald, 2021, S. 284). Allerdings werden offene Publikationen mehr zitiert und auch öfter in den Medien erwähnt (McKiernan et al., 2016, S. 1 ff.). Die Öffnung der Forschungsdaten ermöglicht außerdem die Aufdeckung falscher Behauptungen und Ungenauigkeiten sowie die Durchführung von Tests (Burgelman et al., 2020, S. 2). Wissenschaftliche Fakten müssen für eine Überprüfung zugänglich sein, deshalb sollen Forschungsdaten offen und frei verfügbar sein (Bundesregierung, 2021a; Deutsche Forschungsgemeinschaft, 2019; RfII – Rat für Informationsinfrastrukturen, 2021; International Science Council, 2020).

Bis 2030 sollen Hemmschwellen und Barrieren abgebaut und ein Forschungssystem basierend auf geteiltem Wissen etabliert sein (Directorate-General for Research and Innovation (European Commission) et al., 2020, S. 23). Dieser Fortschritt ist mit erheblichen technischen sowie infrastrukturellen Herausforderungen verbunden. Bisher ist noch nicht ganz klar, wie das offene Forschungssystem aussehen wird, doch befindet es sich in der Umsetzung: Im System sollen voller Datenschutz und Datensicherheit garantiert sein, mit umfassender Offenheit und Transparenz im Forschungskreislauf (RfII – Rat für Informationsinfrastrukturen, 2016; European Commission, 2021; Verordnung (EU) 2016/ 679, 2016). Ob das gewährleistet werden kann, hängt davon ab, inwiefern die Souveränität über die Forschungsdaten behalten wird. Diese Fragen der Datensouveränität sind komplex, insbesondere, wenn Daten online geteilt werden. Bis zu welchem Grad welche Anonymisierungs-Verfahren funktionieren, ist beispielsweise selbst noch Teil der Forschung (vgl. u. a. Chen et al., 2020; Gong et al.,

2020; Weinhardt, 2021). Das ist eine Hürde beim Teilen personenbezogener Forschungsdaten. Eine Datenveröffentlichung setzt in der Regel die Einwilligung der Beforschten voraus (vgl. u. a. RfII – Rat für Informationsinfrastrukturen, 2016; European Commission, 2021; Verordnung (EU) 2016/ 679, 2016). Es braucht spezielle Trainingsangebote und Expertise für alle Forschenden, um Datentransfer und -sicherheit nach aktuellen Standards gewährleisten zu können.

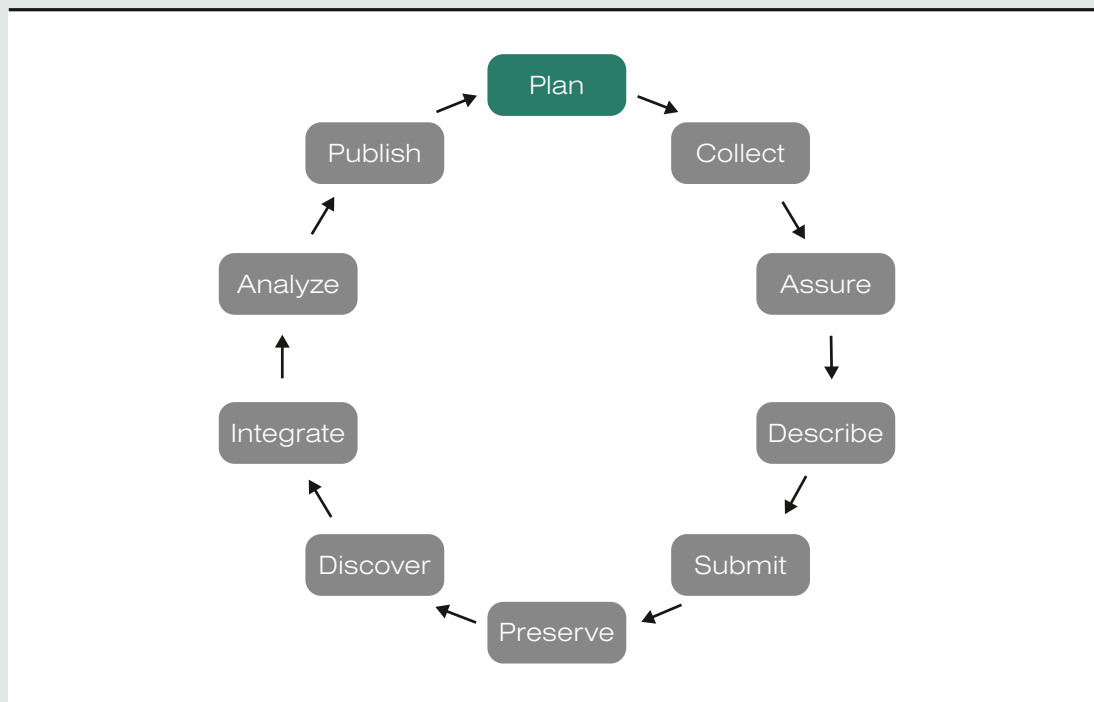
4.1.3 Infrastruktur

Die Digitalisierung stellt die Wissenschaft vor große finanzielle Hürden: „Die für die Forschung unabdingbaren digitalen Infrastrukturen müssen langfristig angelegt sein und zugleich eine schnelle Anpassung an technische Veränderungen ermöglichen, ohne dass sie an Verlässlichkeit, Sicherheit oder Stabilität verlieren“ (Deutsche Forschungsgemeinschaft, 2020, S. 10). In einem neueren Impulspapier der Deutschen Forschungsgemeinschaft aus der COVID-19-Krise werden die infrastrukturellen Herausforderungen für das Wissenschaftssystem im Kontext der sogenannten Resilienz-Orientierung zusammengefasst (Wissenschaftsrat, 2021, S. 65 ff.). Dies bedeutet, dass sichere, innovationsfähige Systeme entwickelt, weiterentwickelt und in Stand gehalten werden müssen, um auch auf unbekannte Krisen bzw. Entwicklungspotenziale vorbereitet zu sein.

Laut den „Leitlinien zur Sicherung guter wissenschaftlicher Praxis“ sind „Hochschulen und außerhochschulische Forschungseinrichtungen“ für die Archivierung von Forschungsdaten zuständig (Deutsche Forschungsgemeinschaft, 2019, S. 22, Leitlinie 17). Im Allgemeinen ist es Wissenschaftler*innen selbst überlassen, welches Publikationsorgan sie zur öffentlichen Bereitstellung der Forschungsdaten wählen – die Qualität des Beitrags bleibt davon unberührt (ebd., S. 21, Leitlinie 15). In kollaborativen Arbeitsszenarien müssen Konzepte erarbeitet werden, wie welche Infrastruktur genutzt wird, um die gemeinsamen Projektdaten zu schützen und zu sichern (vgl. Engelhardt & Kusch, 2021, S. 13). Scholze et al. (2021, S. 170) gehen davon aus, „dass rund 68 Prozent der Forschungsdatenrepositorien mit selbstentwickelter Software betrieben wird, die zwar auf Standardkomponenten oder Frameworks aufbaut, trotzdem jedoch keine Vorhersage zu Interoperabilität bzw. Schnittstellen für die Machine-to-Machine-Kommunikation erlaubt“. Dazu kommt, dass 54 % der Repositorien keine Persistenten Identifikationssysteme (PID) für Datensätze verwenden (ebd.). Ohne PID, wie Digital Object Identifier (DOI), lassen sich Daten nur schwer finden und die Langzeitverfügbarkeit ist nicht garantiert.

Im gesamten Forschungskreislauf fallen unterschiedliche Datensätze an, die für die Replikation einer Studie

Modell eines Datenlebenszyklus



Quelle: German Federation for Biological Data (2019).⁷⁴

Abbildung 5: Modell eines Datenlebenszyklus (Quelle: RfII – Rat für Informationsinfrastrukturen, 2019a, A-24; Lizenz: CC BY-SA 4.0, <https://creativecommons.org/licenses/by-sa/4.0/>)

und/oder für Nachnutzung benötigt werden (Abbildung 5). Es braucht die passende Infrastruktur, um diese Daten entsprechend hinterlegen zu können und damit abrufbar zu machen.

Die Daten verfügen über unterschiedliche Datei-Formate, die oft nur mit proprietärer Software ausgelesen werden können. Offene, zugängliche Standards und Dateiformate gewinnen durch die FAIR-Prinzipien an Bedeutung, doch werden diese oftmals sehr unterschiedlich ausgelegt, wenn sie überhaupt bekannt sind. Forschungs- und Informationsinfrastrukturen sind auch für die Qualitätssicherung des Datenbestandes unabdingbar (RfII – Rat für Informationsinfrastrukturen, 2019a S. 92 ff.). Sie könnten kontrollieren, inwiefern wissenschaftliche Datensätze den FAIR-Prinzipien entsprechen, und die Daten kuratieren, sodass andere Forschende leichter Zugriff darauf erhalten (RfII – Rat für Informationsinfrastrukturen, 2021, S. 66). Der Rat empfiehlt demnach Infrastrukturen, die Daten nach den FAIR-Prinzipien kuratieren und veröffentlichen (FAIR + Q = „gelingender Zugriff“) (ebd.).

Um ein koordiniertes Forschungsdatenmanagement in Deutschland aufzubauen, wurde am 12.10.2020

der Verein Nationale Forschungsdateninfrastruktur (NFDI) gegründet (Kraft et al., 2021, S. 2). Neben dem Aufbau und der Organisation des Vereins fokussiert sich die NFDI auf das Projekt FAIR Data Spaces², das im Mai 2021, zusammen mit der wirtschaftlich ausgerichteten Dateninfrastruktur Gaia-X³, gestartet ist. Zahlreiche Forschungseinrichtungen beteiligen sich an der NFDI und an den FAIR Data Spaces. Das Direktorat ist das Exekutivorgan der NFDI (ebd., S. 6). Die NFDI vertritt Deutschland im europäischen Projekt EOSC – European Open Science Cloud. Durch dieses Projekt soll langfristig ein „Web of FAIR Data and services“ geschaffen werden (European Open Science Cloud (EOSC), o. J.). Die Kommerzialisierung des Datenmarktes ist derzeit, anders als im Publikationsmarkt, nur wenig fortgeschritten, weshalb nicht unmittelbar mit wirtschaftlichen Gewinnen zu rechnen ist (Scholze et al., 2021, S. 172). Bisher spielte die (Re-)Finanzierung bzw. die Monetarisierung wissenschaftlicher (Online-)Angebote eine untergeordnete Rolle (RfII – Rat für Informationsinfrastrukturen, 2021, S. 66). Es sind private Firmen, die systemrelevante Funktionen in der Wissenschaft erfüllen. Der

² <https://www.nfdi.de/fair-data-spaces/>

³ <https://www.data-infrastructure.eu/>

Rat für Informationsinfrastrukturen empfiehlt daher auch, dass es „öffentlich garantierte Rückfalloptionen“ geben soll, wenn Unternehmen ausfallen (ebd., S. 73). Aufgrund marktwirtschaftlicher Betrachtungen gehen andere Vorschläge noch weiter und fordern, dass Marktführer in datengetriebenen Märkten ihre Daten offenlegen müssen (vgl. Prüfer, 2020). Sollten sich derartige Datenteilungspflichten im privatwirtschaftlichen Bereich durchsetzen, sodass beispielsweise Amazon die Daten offenlegt, liegt es an der Wissenschaft, diese Datenflut infrastrukturell verarbeiten und analysieren zu können. Durch die NFDI und die EOSC könnten derartige Optionen entstehen. Gerade befinden sich die Großprojekte in der Anfangsphase, sodass derartige Entwicklungen noch nicht absehbar sind.

4.1.4 Wissen und Renommee im Wandel

Zweck von Wissenschaft ist die Generierung von gesichertem Wissen. Es sind nicht die reichsten Wissenschaftler*innen, die am erfolgreichsten sind, auch nicht jene, die am meisten publizieren oder lesen. Diese Faktoren hängen zusammen, doch es kommt letztendlich auf die Qualität des Wissens an, die Forschende produzieren und wie dieses im Wissenschaftssystem aufgenommen wird. Die Bemessung des Erfolgs von Wissenschaftler*innen ist sehr komplex. Der h-Index ist ein bekanntes quantitatives Bewertungsschema, das Forschende auf Basis der Zitationen durch andere Forschende nach ihrer Reichweite beurteilt. Erfolg ausschließlich auf die Reichweite wissenschaftlicher Arbeiten zu beschränken wäre stark verkürzt (vgl. „San Francisco Erklärung Zur Forschungsbewertung“, o. J.). Gutes Forschungsdatenmanagement und frei publizierte Daten werden so nicht weiter bei der Bewertung berücksichtigt. Es gibt kein einfaches numerisches Maß, welches den Erfolg von Wissenschaftler*innen zufriedenstellend abbildet. Deshalb wurde der Kerndatensatz-Forschung⁴ entwickelt, der seit 2016 vom Wissenschaftsrat zur Evaluation von wissenschaftlichen Organisationen empfohlen wird und auch Datenpublikationen berücksichtigt. Der Begriff *Renommee* soll dieses vielschichtige Geflecht aus sozialen Erfolgsfaktoren für Wissenschaftler*innen beschreiben. Ein Synonym ist Reputation. *Renommee* beschreibt mehr als nur den guten Ruf, den eine Person haben kann, weil dieser zu schnell gewonnen und verloren ist. Zudem beschreibt der Ruf nicht unbedingt die Qualität von Arbeitsleistungen. Weiters ist der Begriff Renommee treffender als Ansehen oder Prominenz, weil die beiden Begriffe meist im Kontext von Öffentlichkeit gedacht sind, die renommierte Wissenschaftler*innen in unbeachteten Sparten kaum erreicht. Das schließt nicht aus, dass ein renommierter

Mensch einen guten Ruf und Ansehen genießt, es muss aber nicht so sein. Weiters ist mit Renommee nicht der Wert oder Rang von Wissenschaftler*innen gemeint, denn es gibt in Gesellschaft und Wissenschaft diverse Wert- und Rangsysteme. Renommee verhält sich ähnlich wie Geld: Man kann damit arbeiten und es ansparen. Trotzdem folgt es einer anderen Logik: Es gibt keine bestimmte Einheit, in der es berechnet werden könnte. Der persönliche und gesellschaftliche Wert von einzelnen Arbeiten, Preisen, Ehrungen, Positionen, akademischen Titeln oder Soft Skills (...) kann nicht durch eine Zahl zusammengefasst werden, doch der Wert lässt sich durch den Begriff Renommee begreifen. Ideale, Anreize und Aufstiegsmöglichkeiten werden im historisch gewachsenen Wissenschaftssystem aktiv gesetzt und gestaltet.

Die Digitalisierung, die von Teilen in der Wissenschaft vorangetrieben wird, ist gleichzeitig der Faktor, der das System nachhaltig verändert. Im Impulspapier identifiziert die Deutsche Forschungsgemeinschaft (2020, S. 5 ff.) drei Typen des digitalen Wandels auf die Wissenschaft: den transformativen Wandel, den ermöglichenden Wandel und den substituierenden Wandel. Werden beispielsweise analoge Daten digitalisiert, spricht man vom transformativen Wandel. Der digitale Wandel schafft neue Möglichkeiten für datenintensive Berechnungen, was ihn ermöglichend wirken lässt. Genauso kann der Wandel auch substituierend wirken, weil analoge Verarbeitungspraxen teilweise obsolet werden (ebd.). Die Digitalisierung gestaltet auch die Denkräume größer, wobei sie praktisch von finanziellen Möglichkeiten sowie ethischen und rechtlichen Rahmenbedingungen beschränkt werden.

Die Deutsche Forschungsgemeinschaft und der International Science Council (ISC) setzen Leitlinien, Vorgaben und Ideale, doch es gibt Barrieren, die Forschende daran hindern die gemeinsamen Ideale zu befolgen. Derzeit hinkt die Wissenschaft ihren eigenen Idealen weit hinterher. Der ISC (2020, S. 4) sieht Wissenschaft als weltweit verfügbares öffentliches Gut, welches unabhängig von „geography, gender, political boundaries, ethnicity or ability to pay“ erreichbar sein soll. Wissenschaftler sollten, laut dem Arbeitspapier, nicht als Dateneigentümer fungieren, sondern als deren Verwalter im Auftrag der Öffentlichkeit. Die Universitäten sollten Strukturen schaffen, um ihre Forschenden bei der Erfüllung der beschwerlichen Aufgabe des Datenmanagements zu unterstützen (ebd., S. 2 f.). Mit den Leitlinien der Deutschen Forschungsgemeinschaft (2019a, S. 18, Leitlinie 13) wird diese Vision unterstützt und im deutschen Sprachraum verankert. Die Leitlinien zur guten wissenschaftlichen Praxis geben nicht nur eine Richtschnur beim Forschungsdatenmanagement, sondern auch beim redlichen Verhalten im

⁴ <https://www.kerndatensatz-forschung.de/>

Wissenschaftssystem: Es gibt einen starken Berufsethos, der von Normen und Werten geprägt ist (ebd., S. 9, Leitlinie 2). Konsequenzen für Fehlverhalten sind angedacht (ebd., S. 23 ff.).

Die genaue Auslegung der Leitlinien obliegt der einzelnen wissenschaftlichen Institution (ebd.). Diese gemeinsamen Richtlinien, denen sich alle Wissenschaftler*innen verpflichten, haben erhebliche Auswirkungen auf das Verständnis von Renommee. Für besondere Leistungen innerhalb des Systems verdient man Renommee, je nach Fehlverhalten schwindet es. Dieser Prozess ist nicht symmetrisch. Über Jahrzehnte erarbeitete Renommee kann in verhältnismäßig kurzer Zeit abgesprochen werden, wenn Fehlverhalten vorliegt oder auch nur vermutet wird. In besonders gravierenden Fällen kann ein Fehlverhalten zu Berufsverboten führen. Niemand denkt an solche Maßnahmen, wenn Forschende vergessen, ihre Daten ordnungsgemäß zu veröffentlichen. Doch es werden Anreize und Maßnahmen gesetzt, welche die FAIR-Prinzipien umfassender in allen Disziplinen integrieren sollen. Scholze weist darauf hin, dass die Forschenden in den meisten Fachdisziplinen derzeit nicht auf die Veröffentlichung und Nachnutzung von Forschungsdaten vorbereitet sind (Oßwald, 2021, S. 292). Alle relevanten Akteur*innen sollten Maßnahmen zum besseren Datenmanagement ergreifen und nutzen, um die Qualität der Wissenschaft weiter zu steigern.

4.1.5 Defizite und Handlungsbedarf

In diesem Abschnitt werden verschiedene Defizite und Handlungsbedarfe im Untersuchungsbereich der *Wissenschaft* analysiert. Diese lassen sich in zwei Handlungsfelder gliedern. Erstens gibt es Herausforderungen innerhalb der *Wissenschaft*. Dies sind Defizite, die innerhalb des Wissenschaftssystems erkannt werden, und deren Auflösung ebendort angestrebt wird. Zweitens gibt es Handlungsfelder, die von außen, also von der Gesellschaft, an die *Wissenschaft* herangetragen werden. Damit sind Wünsche und Problemstellungen gemeint, die aus den verschiedenen Bereichen der Gesellschaft entstehen. Beispielsweise die Lösungen von Gesundheits- und Klimakrisen, innovative Datenräume und andere technische oder soziale Innovation. Empirisch lassen sich diese beiden Felder kaum voneinander trennen: Das Wissenschaftssystem ist in die Gesellschaft eingebettet und wird dadurch von ihr beeinflusst und gestaltet. Trotzdem behält das Wissenschaftssystem gewisse Autonomie und spielt eine spezifische Rolle in der Gesellschaft, wodurch eine theoretische Unterscheidung zwischen den Handlungsfeldern sinnvoll ist. In der erwähnten Literatur gibt es zahlreiche Problemanalysen und Empfehlungen im Umgang mit Daten, die das

Wissenschaftssystem betreffen. Die großen Infrastrukturinitiativen, die NFDI und EOSC, können als Projekte begriffen werden, die beide Felder zu verknüpfen versuchen.

Ein offenes und resilientes Wissenschaftssystem erscheint als dominante Zielvorstellung. Jedoch ist es notwendig und hilfreich, Offenheit als Prozess der Öffnung zu verstehen, dessen Endpunkt nicht klar definiert ist (vgl. Schrögel et al., 2020). Ähnlich verhält es sich auch mit dem Begriff der Resilienz. Er bietet eine Orientierung, aber keinen Endpunkt. So ist schon jetzt klar, dass derartige Prozesse niemals abgeschlossen sein können. Darüber hinaus erscheint es in diesem Kontext sinnvoll zu erwähnen, dass sich die Ausdrücke offene Wissenschaft – open science – und offene Innovation – open innovation –, unterscheiden (vgl. Hightech-Forum, 2020a, S. 2). Im Kontext dieser Arbeit ist das wichtigste Unterscheidungsmerkmal, dass offene Innovationen insbesondere auch von nicht-wissenschaftlichen Akteur*innen geschaffen werden (ebd.). Offene Innovationen gehen also über das Wissenschaftssystem hinaus und finden hier keine besondere Aufmerksamkeit.

Bedarfe innerhalb der Wissenschaft

Offenheit, also die transparente Darlegung der Forschungsdaten, ist ein Prozess, der von den Disziplinen unterschiedlich umgesetzt wird. Klar ist: Der Zugang zu Forschungsdaten und Software muss gewährleistet sein (Deutsche Forschungsgemeinschaft, 2020, S. 10). Dies bedeutet, dass es nachhaltige Forschungsinfrastrukturen braucht, die technische Innovationen antizipieren können, ohne dabei an Verlässlichkeit, Sicherheit oder Stabilität zu verlieren (ebd.). Daten aus dem gesamten Forschungskreislauf sollen einfach zu publizieren und mit anderen Arbeiten vernetzt abrufbar sein. Dazu muss die Sicherung und Archivierung gewissenhaft erfolgen, sodass das Arbeiten mit Daten unterschiedlicher Herkunft möglich ist. Der RfII – Rat für Informationsinfrastrukturen (2021, S. 66) – empfiehlt dazu wissenschaftlich kuratierte Datensammlungen: Daten, die zumindest den FAIR-Prinzipien entsprechen, werden kuratiert zur Verfügung gestellt (FAIR + Q). Will man diese Prinzipien erreichen, wird noch mehr Initiative zur Offenheit benötigt. Auch der RfII „bekräftigt den Grundsatz, dass mit öffentlichen Mitteln erzeugte wissenschaftliche Daten auch öffentlich zugänglich bleiben müssen“ (ebd., S. 75). Das betrifft eben auch Code und Software, mit denen Daten generiert werden. Um diesen Ansprüchen gerecht zu werden, braucht es eine dauerhafte Finanzierung dieser Infrastrukturen (Deutsche Forschungsgemeinschaft, 2020, S. 10). Dies könnte die Replikations- und Vertrauenskrise insofern mildern, als dass Forschungsdaten zugänglicher und nachvollziehbar werden. Hürden, wie beispielsweise Embargos im Publikationswesen,

Struktur des Kodex „Leitlinien zur Sicherung guter wissenschaftlicher Praxis“

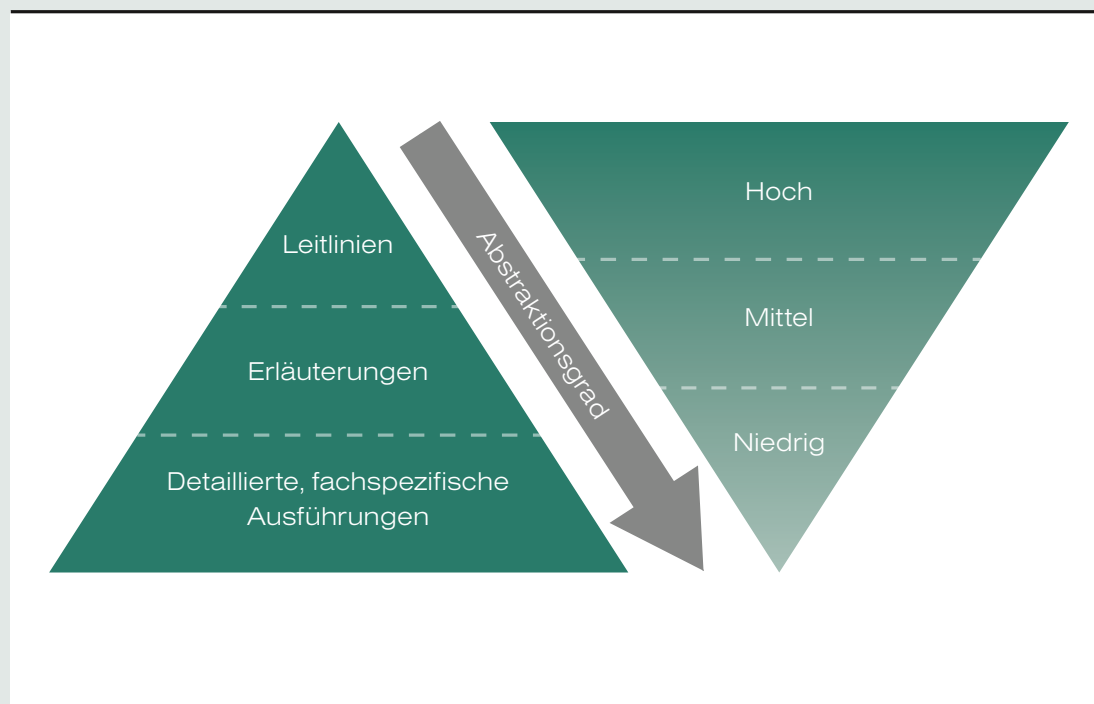


Abbildung 6: Struktur des Kodex „Leitlinien zur Sicherung guter wissenschaftlicher Praxis“ (Quelle: Deutsche Forschungsgemeinschaft, 2019, S. 2; Lizenz: CC BY 4.0 International, <https://creativecommons.org/licenses/by/4.0/>)

werden in den meisten Publikationen nicht adressiert. Transparentes Arbeiten und offene Publikationen sind noch nicht überall die Norm, wobei sich die FAIR-Prinzipien als Standard etablieren. Es gibt ein Bedürfnis nach rechtskonformen und sicheren Lösungen, die das Forschen erleichtern.

Seitens der Deutschen Forschungsgemeinschaft werden Leitlinien vorgegeben (vgl. Abbildung 6). Die fachspezifischen Implikationen und detaillierte Ausführungen können nur innerhalb der verschiedenen Disziplinen formuliert werden. Während Leitlinien bewusst einen hohen Abstraktionsgrad aufweisen, sind die fachspezifischen Ausführungen konkret und direkt anwendbar. Diese Hierarchie von Vorgaben der Deutschen Forschungsgemeinschaft ist in Abbildung 6 veranschaulicht.

Durch die Leitlinien sowie deren Erläuterungen ergeben sich diverse fach- und institutsspezifische Handlungsfelder, die hier nicht im Detail beleuchtet werden könnten. Wesentlich ist, dass Vorgaben so gestaltet sind, dass man sie auf der jeweiligen Ebene gut anwenden kann, wobei gesetzliche Vorgaben niemals die Funktion von Leitlinien oder detaillierten, fachspezifischen Ausführungen ersetzen könnten.

Van den Eynden und Bishop (2014, S. 15) setzen mit ihrer vergleichenden Studie bei der Motivation von Forschenden an. Was versprechen sich Forschende,

wenn sie ihre Daten offen publizieren? Welche Normen und Erwartungen sind von den Disziplinen und Geldgebern vorgeschrieben? Was bedeutet es für die Karrieren der Forschenden, wenn sie sämtliche Forschungsdaten transparent machen? Welche Rolle spielt offenes wissenschaftliches Arbeit bei der Stellenbesetzung im Wissenschaftsraum? Zielgerichtete leistungsorientierte Incentivierung könnte die Motivation der Forschenden heben (ebd., S. 32).

Handlungsfelder für die Wissenschaft

Im zweiten Feld geht es um Handlungsbereiche, die aus der Gesellschaft an die Wissenschaft herangetragen werden. Das betrifft vor allem neue Forschungsfelder im Bereich der Digitalisierung. Privatwirtschaftliche Akteur*innen und staatliche Institutionen wollen das Potential von Daten verstärkt nutzen. Zahlreiche Strategiepapiere deuten darauf hin, dass überaus hohe Erwartungen in die wissenschaftlichen Daten gelegt werden. In den folgenden Absätzen wird die europäische Datenstrategie (Europäische Kommission, 2020a) im Kontext der DSGVO (Verordnung (EU) 2016/ 679, 2016) (vgl. Kapitel 3.2) für den Bereich Wissenschaft analysiert, um Handlungsbedarfe zu identifizieren, mit denen das Wissenschaftssystem konfrontiert ist. Die Europäische Datenstrategie (Europäische Kommission, 2020a) klassifiziert zehn Sektoren, die jeweils

einen gemeinsamen Datenraum aufbauen sollen (ebd., S. 30 ff.). Zudem soll Forschung erleichtert werden (ebd., S. 15). Einer dieser Datenräume, der sich mit offener Wissenschaft befasst, ist die EOSC (ebd., S. 39). Die Handlungsaufforderung diesen zu gestalten und mit Wissen zu befüllen ist explizit. Doch auch die anderen Räume, der Mobilitätsdatenraum, der Gesundheitsdatenraum, der Finanzdatenraum, der Energiedatenraum, der Kompetenzdatenraum, der Industriedatenraum, der Agrardatenraum, die Datenräume für die öffentliche Verwaltung und der Datenraum für den European Green New Deal, fordern wissenschaftliche Kompetenz. Erstens will man bei allen Datenräumen mit „höchsten bestehenden Cyber-sicherheitsstandards“ (ebd., S. 25 f.) agieren, die von wissenschaftsnahen Akteur*innen weiterentwickelt und implementiert werden müssen. Zweitens implizieren die Datenräume, dass eine saubere Datenerhebung und -analyse möglich sein soll, was nicht ohne wissenschaftliche Kompetenz zu erreichen sein wird. Ob die Wissenschaft frei an den Daten aus allen Datenräumen forschen darf, steht nicht fest.

Die Europäische Datenschutzgrundverordnung (Verordnung (EU) 2016/ 679, 2016) soll den Schutz personenbezogener Daten gewährleisten, ohne dabei die freie Wissenschaft einzuschränken (ebd., Art. 5(1) b). Darin ist das Prinzip der Datenminimierung festgeschrieben: Personenbezogene Daten müssen „dem Zweck angemessen und erheblich sowie auf das für die Zwecke der Verarbeitung notwendige Maß beschränkt sein“ (ebd., Art. 5(1)c). Welche personenbezogenen Daten für einen Zweck angemessen und erheblich sind, ist nicht näher definiert. Alle relevanten Daten aus dem gesamten Forschungskreislauf sollten zur Weiterverwendung (Reuseability) offengelegt werden, dennoch sind personenbezogene Daten seitens der Verantwortlichen möglichst zu minimieren. Aus ethischen Überlegungen hat die Europäische Kommission den Grundsatz der Datenminimierung nochmals bestärkt (European Commission, 2021, S. 11). Forschende können weiterhin frei forschen und Daten erheben, aber Transparenz darüber, welche Daten wozu erhoben werden, ist vorgeschrieben (Verordnung (EU) 2016/ 679, 2016, Art. 5(1)a). Es gibt nur wenige spezielle Fälle, die Ausnahmen zulassen (ebd., Art. 49). Jede Person hat das Recht zu erfahren, welche Daten über sie gesammelt wurden und die Möglichkeit, deren Löschung zu veranlassen (ebd., Art. 15; Art. 17). Die rechtlichen Rahmenbedingungen bieten Raum für Prozess-Innovationen, welche es Forschenden erleichtern, diesen rechtlichen und technischen Herausforderungen gerecht zu werden. Hier braucht es auch technische Innovationen, die diesen rechtlichen Ansprüchen genügen und die individuelle Datensouveränität der Beforschten entsprechend stärken. Bis dahin wird versucht, den datenschutzrechtlichen

Ansprüchen nach besten Wissen und Gewissen gerecht zu werden. In einer aktuellen Handreichung stellt beispielsweise der Rat für Sozial- und Wirtschaftsdaten (2020) Informationen zum Datenschutz bereit.

Handlungsbedarfe innerhalb des Wissenschaftssystems können mit genügend Zeit und Ressourcen aufgelöst werden. Ob durch die Wissenschaft und offene Daten die gesellschaftlichen Innovationen hervorgebracht werden, die es braucht, um die großen gesellschaftlichen Herausforderungen wie die Klima- oder Gesundheitskrise zu lösen, sollte an dieser Stelle bezweifelt werden, weil noch kein systematischer Beweis erbracht werden konnte. Wie weit die Wissenschaft zum gesellschaftlichen Fortschritt beitragen kann, ist von vielen verschiedenen Faktoren abhängig, die in einem eigenen Forschungsprojekt entsprechend beleuchtet werden könnten.

4.1.6 Zusammenfassung der Kernpunkte

Die gelesenen Abschnitte präsentieren die Ergebnisse der Literaturanalyse gegliedert nach den Themen *Quellengrundlage, Rahmenbedingungen, Infrastruktur, bereichsspezifische Rationalitäten sowie Defizite und Handlungsbedarfe*.

Der Abschnitt *Quellengrundlage* beschreibt, welche Literatur in der Analyse verwendet wird.

Im Abschnitt *Rahmenbedingungen* ist das Arbeiten mit Forschungsdaten in einen Kontext gesetzt. Die Digitalisierung und die damit verbundene Vertrauens- und Replikationskrise machen Neuerungen wie die FAIR-Prinzipien notwendig. Diese Prinzipien werden von verschiedenen Akteur*innen unterschiedlich stark forciert, wobei sie sich sukzessive als Standard durchsetzen. Beim Ideal der Offenheit, das im Zuge der Digitalisierung neuen Aufwind erhält, erscheint das nicht ganz so sicher. Es erfährt großen Zuspruch, aber aus unterschiedlichen Gründen wird es kaum in der Publikationstätigkeit umgesetzt.

Die wissenschaftliche *Infrastruktur* steht aufgrund der Digitalisierung vor enormen Herausforderungen. Es braucht sichere, resiliente Strukturen, die technische Innovationen adaptieren können, ohne dabei an Verlässlichkeit zu verlieren. Daten aus dem gesamten Forschungskreislauf sollen nachgenutzt werden können. Derzeit ist das Wissenschaftssystem von privaten Softwarelösungen abhängig. Einerseits möchte man sich diesem Einfluss mit offenen Standards und freier Software entziehen. Andererseits gibt es auch Software, die sich etabliert hat und deshalb nur schwer substituiert werden kann. Es braucht Rückfalloptionen, falls einzelne Dienstleister*innen ausfallen. Sofern es aus wettbewerbsrechtlichen Gründen zu Datenteilungspflichten für große privatwirtschaftliche Akteur*innen kommen sollte, wäre es wichtig, dass das Wissenschaftssystem

die Fähigkeiten entwickelt, so zugänglich werdende große Datenmengen nutzen zu können.

Im Abschnitt *bereichsspezifische Rationalitäten* wird versucht, die spezielle Logik des Wissenschaftssystems grob zu umreißen, um es klar von den anderen Bereichen abzugrenzen. Im Mittelpunkt des Wissenschaftssystems steht die Generierung von gesichertem Wissen. Der digitale Wandel legt *Handlungsbedarfe und Defizite* offen, zu denen sich das Wissenschaftssystem verhält. Hier wird unterschieden zwischen Handlungsfeldern innerhalb des Systems und Handlungsbedarfen, die von außen an das System herangetragen werden. Die systeminhärenten Handlungsbedarfe lassen sich mit genügend Zeit und Ressourcen lösen. Wie weit man mit wissenschaftlichem Datenmanagement zum gesellschaftlichen Fortschritt beitragen kann, ist von vielen Faktoren abhängig.

4.2 Public Sector

4.2.1 Quellengrundlagen

Aufgrund des umfangreichen Untersuchungsspektrums durch die weite Definition des *Public Sectors* ist es nicht möglich, auf alle Teilbereiche im Detail einzugehen. Daher bezieht sich diese Analyse vor allem auf Gebietskörperschaften und die öffentliche Verwaltung, aber auch auf öffentliche Unternehmen. Der Stand zum Umgang mit digitalen Daten fällt in diesen Bereichen sehr unterschiedlich aus. So haben beispielsweise noch nicht alle Bundesländer Strategiepapiere zum digitalen Datenmanagement vorgelegt (vgl. Kapitel 3.3.2). Allerdings gibt es neben einer Vielzahl von Veröffentlichungen über den Wert von und den Umgang mit Daten auch spezifische Strategiepapiere einzelner Akteur*innen. Daher liegt der Fokus der vorliegenden Untersuchung auf einzelnen grundlegenden Veröffentlichungen wie der Datenstrategie des Bundes⁵, und Publikationen von strategiewerkzeugen Akteur*innen, beispielsweise dem mFUND⁶ (vgl. Kapitel 4.2.5). Des Weiteren muss der zeitliche Aspekt der Veröffentlichungen bei einem sich derart dynamisch entwickelnden Thema wie digitale Daten beachtet werden. So können nur Grundsatzentscheidungen und grundlegende Regelungen der Vergangenheit kurz vorgestellt werden. Daneben wurde bei der Untersuchung ein Fokus auf aktuellere Veröffentlichungen ab 2016 gelegt. Das erscheint als sinnvoll, da dies den Zeitraum seit der Veröffentlichung der FAIR-Prinzipien umfasst (vgl. Kapitel 1.3.3). Zudem werden die Ergebnisse der Sichtung von Veröffentlichungen wie

Handreichungen, Umfragen und Artikel von Dritten wiedergegeben und ergänzt durch Erkenntnisse aufgrund der Teilnahme an Fachveranstaltungen wie der Tagung des Arbeitskreises Public Data des mFund⁷ sowie der Jahrestagung des RatSWD⁸. Darüber hinaus sind Erkenntnisse aus den veröffentlichten Surveys und Auswertungen der Bundesregierung zu Behörden Daten eingeflossen. Außerdem fanden Strategiepapiere zu Open Data von Gemeinden, Kommunen und Behörden sowie Berichte und Veröffentlichungen verschiedener Stakeholder-Gemeinschaften und -gruppierungen wie Verbänden und Arbeitsgruppen des *Public Sectors* Eingang in die Untersuchung.

4.2.2 Rahmenbedingungen

In Europa gab und gibt es seit gut 30 Jahren Strategien, die eine Digitalisierung befördern sollen. 1999 wurde die Initiative eEurope der Europäischen Kommission ins Leben gerufen mit dem Ziel, Europas Internetnutzung auszubauen und wirtschaftspolitisch besser zu nutzen. In diesem Sinne legte die Europäische Kommission im März 2000 auf dem Treffen des Europäischen Rates in Lissabon einen Bericht vor, aus dem der Aktionsplan eEurope 2002 zur Etablierung einer europäischen Informationsgesellschaft erwuchs (vgl. Kommission der Europäischen Gemeinschaften, 2000). Mit dem zweiten Aktionsplan eEurope 2005 wurden Maßnahmen in den Aktionsplan aufgenommen, die zudem einen Schwerpunkt „auf [die, Anmerkung der Autorin / des Autors] Sicherheit von Netzen und Information sowie auf eGovernment, eLearning, eHealth und eBusiness“ legen sollten (Europäischer Rat, zitiert nach Kommission der Europäischen Gemeinschaften, 2002, S. 7). Im Jahr 2005 wurde die Lissabon-Strategie als i2010 neu aufgelegt und breiter aufgestellt. Ein Punkt war demnach eine Verbesserung von öffentlichen Diensten und Lebensqualität u. a. durch den Zugang zu Webdiensten und elektronischen Behördendiensten (vgl. Europäische Kommission, 2005). Im Nachfolgeprogramm Europe 2020 war die „Digitale Agenda“ dann nur noch einer von vielen Aspekten für ökonomisches Wachstum und gesellschaftliche Weiterentwicklung (vgl. Europäische Kommission, 2010). Im Sinne des Aktionsplans eEurope 2002 verabschiedeten im November 2003 das Europäische Parlament und der Europäische Rat die sogenannte PSI-Richtlinie (Richtlinie 2003/98/EG) über die Weiterverwendung von Informationen des öffentlichen Sektors. 2005 wurde diese auf den gesamten Europäischen Wirtschaftsraum ausgedehnt. In Deutschland erfolgte die Umsetzung der europäischen Vorgabe 2006 durch das Gesetz

⁵ <https://www.bundesregierung.de/resource/blob/992814/1845634/f073096a398e59573c7526feaaadd43c4/datenstrategie-der-bundesregierung-download-bpa-data.pdf>

⁶ <https://www.bmvi.de/DE/Themen/Digitales/mFund/Ueberblick/ueberblick.html>

⁷ 5. Arbeitsforum mFund-Begleitforschung des WIK: Open Data und Behörden, Online-Konferenz, 2020

⁸ 8. Konferenz für Sozial- und Wirtschaftsdaten, Gesellschaft braucht Wissenschaft. Wissenschaft braucht Daten, 2020

The overall open data maturity scores of the 2021 assessment

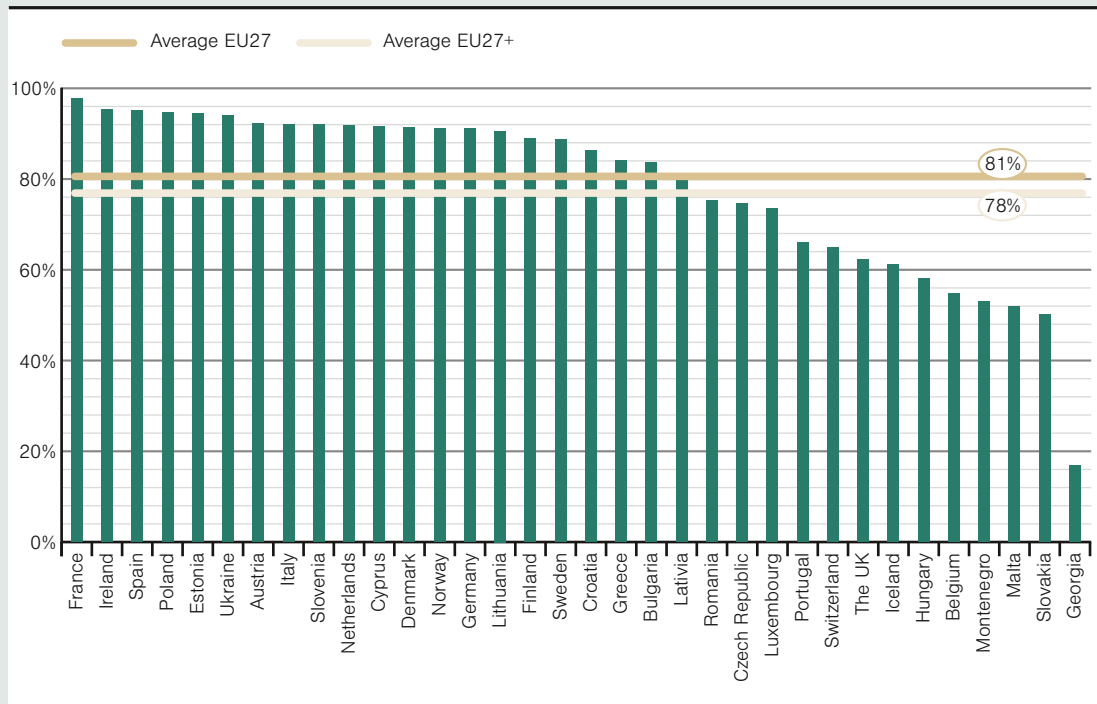


Abbildung 7: Gesamtwerte der Data Maturity der 34 untersuchten Länder 2021 (Quelle: van Hesteren et al., 2022, S. 5; Lizenz: CC BY 4.0 international, <https://creativecommons.org/licenses/by/4.0/>); Lizenz: CC BY 4.0 international, <https://creativecommons.org/licenses/by/4.0/>)

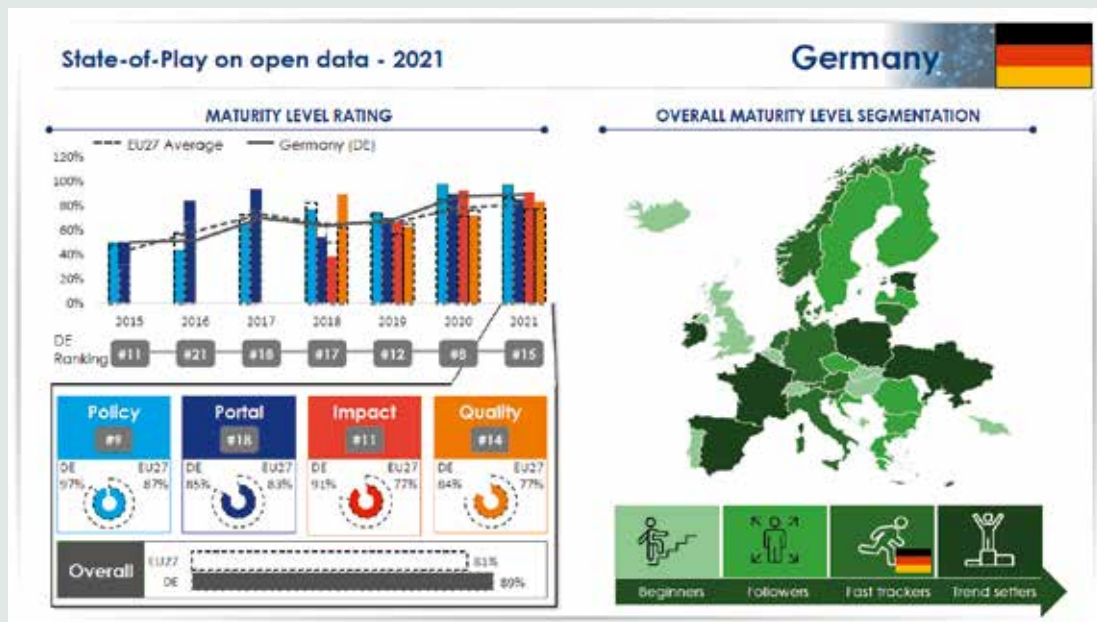


Abbildung 8¹⁰: Entwicklung der Data Maturity Deutschlands 2021 (Quelle: *Country Fact-Sheet Germany 2021*, S. 2; Lizenz: CC BY 4.0 international, <https://creativecommons.org/licenses/by/4.0/>)

über die Weiterverwendung von Informationen öffentlicher Stellen (Informationsweiterverwendungsgesetz – IWG). Die Überarbeitung der Richtlinie 2013/37/EU wurde zum 17. Juli 2021 von der Richtlinie (EU) 2019/1024 ersetzt. In Deutschland wurde daraufhin das Gesetz zur Änderung des E-Government-Gesetzes und zur Einführung des Gesetzes für die Nutzung von Daten des öffentlichen Sektors verabschiedet, das ebenfalls am 17. Juli 2021 in Kraft trat.

Im Gesetz zur Regelung des Zugangs zu Informationen des Bundes (Informationsfreiheitsgesetz – IFG) von 2009 wird der Zugang zu amtlichen Informationen geregelt (vgl. Kapitel 3.3.2). Allerdings sieht dieses Gesetz nur den Anspruch auf Zugang vor, nicht aber eine von der Behörde ausgehende öffentliche Bereitstellung der Daten. Dies regeln auf Bundesebene die oben erwähnten Gesetze: das 2021 geänderte E-Government-Gesetz (EGovG) sowie das neu geschaffene Gesetz für die Nutzung von Daten des öffentlichen Sektors (DNG).

Auf Landesebene sind die gesetzlichen Regelungen sehr unterschiedlich, doch lassen sich die Länder grob in drei Gruppen einteilen. Die erste Gruppe bilden die Bundesländer, in denen es kein eigenes IFG gibt. Ein Rechtsanspruch auf den Zugang zu amtlichen Informationen fehlt in Bayern, Niedersachsen und Sachsen (vgl. Schweigel et al., 2020). Die zweite Gruppe bilden Länder, die ein eigenes IFG vorweisen können. Der Zugang zu behördlichen Daten für Bürger*innen ist in folgenden Ländern eigens geregelt: Baden-Württemberg, Berlin, Brandenburg, Hessen, Mecklenburg-Vorpommern, Nordrhein-Westfalen, Saarland, Sachsen-Anhalt, Schleswig-Holstein und Thüringen (vgl. ebd.). Die dritte Gruppe bilden die Bundesländer, in denen es ein Transparenzgesetz gibt, das die „proaktive Veröffentlichung von Behördendaten“ vorsieht: Diese Gruppe bilden Bremen, Hamburg und Rheinland-Pfalz (ebd., S. 30). Unabhängig von der Landespolitik zeigen vor allem größere Städte und Kommunen – auch in Bundesländern ohne IFG – ein zunehmendes Interesse und Handeln am Themenbereich Open Public Data und schaffen sich über kommunale Informationsfreiheitssatzungen eine gewisse Handlungsfreiheit (vgl. ebd.).

Seit 2015 werden im Data Maturity Report die Entwicklungen im Bereich der Open-Data-Bereitstellung in den EU-Staaten einschließlich der EFTA-Staaten Liechtenstein, Norwegen und der Schweiz untersucht. Der Report untersucht die Data Maturity (Datenreife) der Länder anhand der Hauptbereiche Policy (Regularien), Impact (Wirkung), Portal (Portal) und Quality (Qualität). Der Report 2021 wurde erstmals auf data.europa.

eu⁹ veröffentlicht, das im Frühjahr 2021 aus dem Zusammenschluss des European Data Portal und des European Union Open-Data-Portal entstand. Es wurden 34 Länder untersucht und verglichen (Abbildung 7).

Auf europäischer Ebene kommt der Bericht zu dem Ergebnis: „the scores of the countries are consistently high and are gradually approaching the 100 %“ (van Hesteren et al., 2022, S. 103). Da die Spitze der Bewertungsskala somit erreicht ist, soll die Methodik voraussichtlich 2022 einer Revision unterzogen werden, um neue Kriterien für die Bewertung des Umgangs mit offenen Daten in Europa zu evaluieren (vgl. ebd.).

Im Projekt EcoDM – Ökosystem Datenmanagement wurden die Veröffentlichungen der Jahre 2019 bis 2021 untersucht, da diese dem Projektzeitraum entsprechen (vgl. Blank, 2019; van Knippenberg, 2020; van Hesteren et al., 2022). Im Ranking des Maturity Levels der Open Public Sector Daten aller evaluierten Länder ist Deutschland 2021 im Vergleich zu 2020 auf einen ähnlichen Stand wie 2019 abgerutscht. 2019 mit Rang 13 noch als „Follower“ eingestuft, stieg Deutschland 2020 auf Rang 8 und damit an die Spitze der „Fast Tracker“, um 2021 auf Rang 15 abzufallen (Abbildungen 8 und 9). Damit bleibt Deutschland zwar in der Gruppe der „Fast Tracker“, bildet dort aber nun das Schlusslicht (*Country Fact-Sheet 2021 Germany*, 2021).

Seit 2021 liegt Deutschland in allen vier Hauptuntersuchungsschwerpunkten Policy, Portal, Impact und Quality über dem ermittelten Durchschnitt aller 27 EU-Staaten (Detailldarstellung in Abbildungen 8 – 11). Vor allem im Bereich Quality konnte sich Deutschland im Vergleich zum Jahr 2020 verbessern. Leichte Rückgänge sind dagegen in den Bereichen Portal und Impact festzustellen. Dass Deutschland trotz besserer Performance im Vergleich zu den Vorjahren in der Gesamtwertung zurückgefallen ist, liegt daran, dass andere Staaten schneller besser wurden (van Hesteren et al., 2022).

Am besten schneidet Deutschland im Bereich Policy ab (Abbildung 10), wo es den 5. Platz belegt. In diesem Zusammenhang werden das E-Government-Gesetz (EGovG) wie auch das Datennutzungsgesetz (DNG) als positiv erwähnt, da sie die Zugänglichkeit und Nutzung von öffentlich finanzierten (offenen) Daten verbessern sollen. Auch die Datenstrategie der Bundesregierung, das Kompetenzzentrum Open Data (CCOD) und das Datenportal GovData finden Beachtung. Darüber hinaus wird der Austausch mit Repräsentant*innen der Zivilgesellschaft zum Nationalen Aktionsplan im Rahmen der Teilnahme an der Open Government Partnership genannt (vgl. ebd., S. 12 ff.).

⁹ <https://data.europa.eu/>

¹⁰ Anmerkung zur Abbildung 8: In dem Report 2021 wird der Rang Deutschlands 2019 mit 12 angegeben, jedoch lag dieser im Report 2019 bei Rang 13.

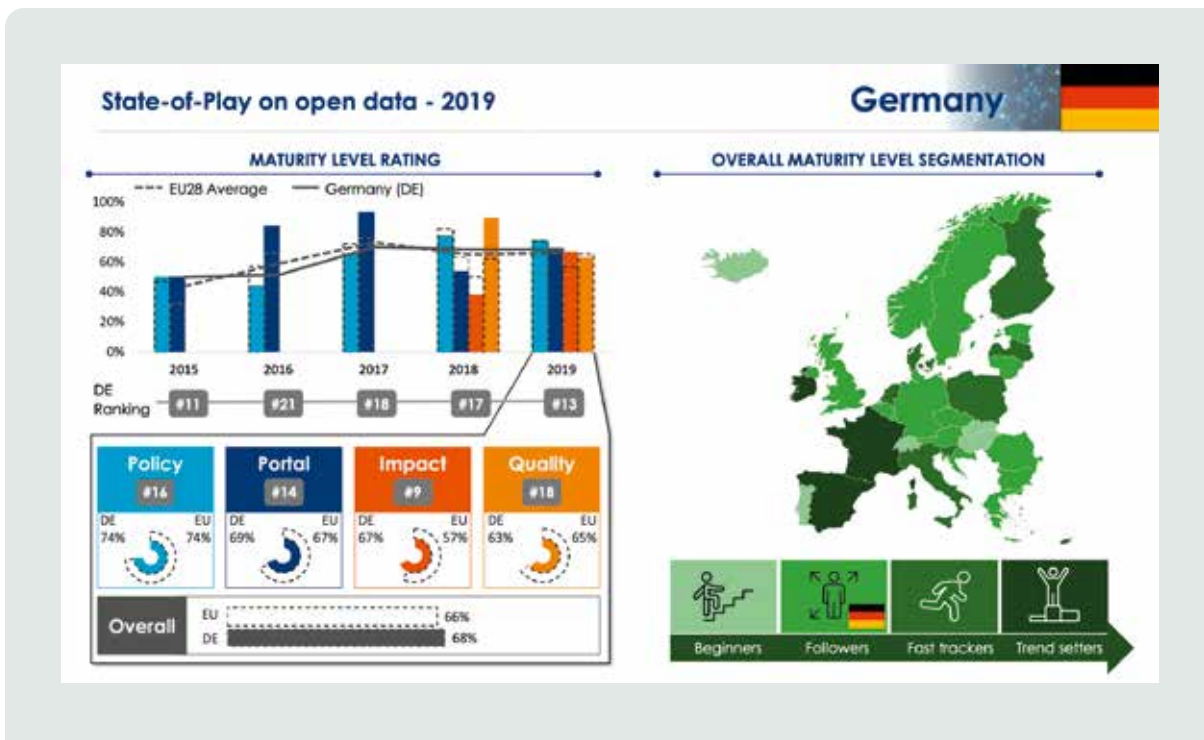


Abbildung 9: Entwicklung der Data Maturity Deutschlands 2019 (Quelle: Country Fact-Sheet Germany 2019, S. 2; Lizenz: CC BY 4.0 international, <https://creativecommons.org/licenses/by/4.0/>)

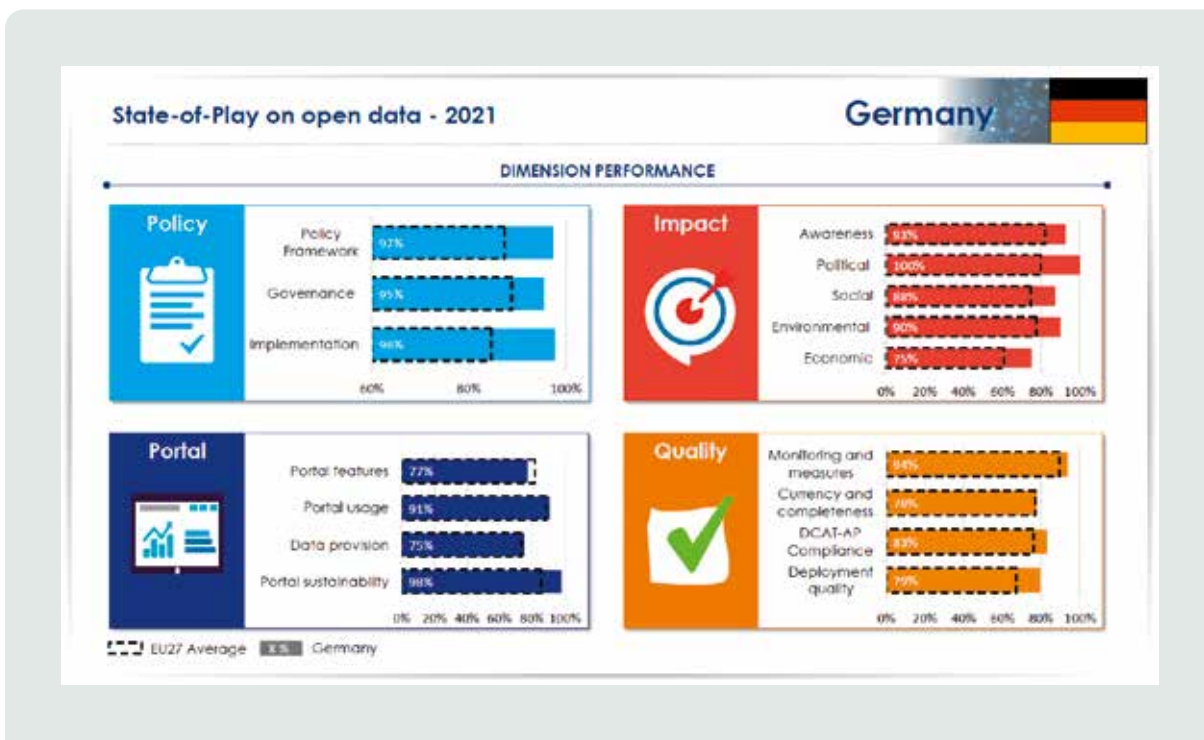


Abbildung 10: Ergebnisse Deutschlands in den vier Untersuchungsschwerpunkten 2021 (Quelle: Country Fact-Sheet Germany 2021, S. 3; Lizenz: CC BY 4.0 international, <https://creativecommons.org/licenses/by/4.0/>)

Unter dem Untersuchungspunkt Impact werden u. a. die Open Knowledge Labs¹¹ erwähnt, die sich im Netzwerk Code for Germany ehrenamtlich für nachhaltigen digitalen Wandel in Politik und Verwaltung

einsetzen. Außerdem kommt der #SmartDevelopmentHack zur Sprache, ein vom Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung sowie der Europäischen Kommission

11 <https://www.codefor.de/about/>

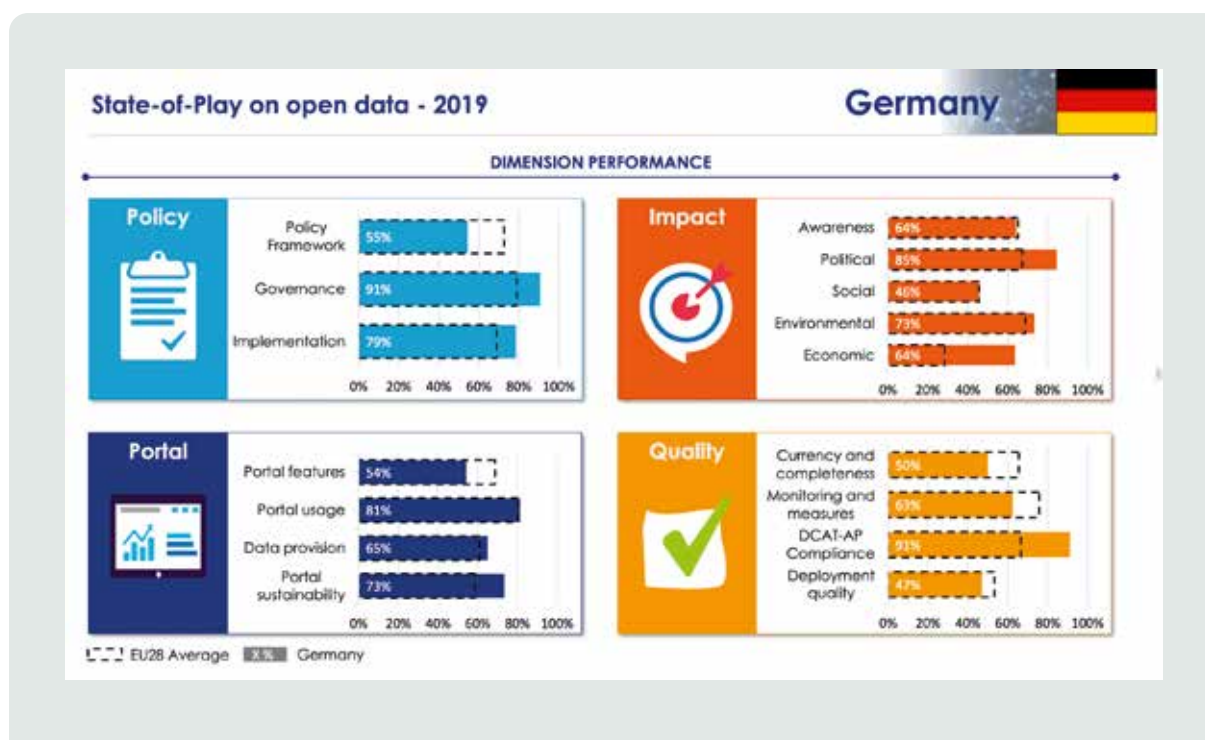


Abbildung 11: Ergebnisse Deutschlands in den vier Untersuchungsschwerpunkten 2019 (Quelle: *Country Fact-Sheet Germany* 2019, S. 3; Lizenz: CC BY 4.0 international, <https://creativecommons.org/licenses/by/4.0/>)

ausgerufener Wettbewerb zur Suche digitaler Ansätze für die Herausforderungen in der Coronavirus-Pandemie. Schließlich wird auf eine Studie vom Bundesministerium für Wirtschaft und Energie hingewiesen, welche die Bedingungen und Potentiale für die Bereitstellung und Wiederverwendung von Daten aus dem *Public Sector* untersucht hat (vgl. ebd., S. 30 ff.).

Nicht zur vollen Zufriedenheit schneidet Deutschland im Bereich Portal ab. Dies liegt daran, dass Deutschland das einzige der untersuchten Länder war, auf dessen nationalem Portal für Open Data es keine Möglichkeit für Nutzer*innen gibt, Daten anzufragen, die (noch) nicht veröffentlicht sind. Außerdem verzeichnet nur das maltesische Portal noch weniger Zugriffe als das deutsche. Der Report erwähnt auch, dass noch nicht alle Bundesländer über eigene Portale verfügen und sich noch immer einige Verwaltungen wenig kooperativ bis ablehnend zum Thema Veröffentlichung von Open Data zeigen (vgl. ebd., S. 49 ff.).

Im Bereich Quality wird über die zweite Version der Datenlizenz Deutschland¹² berichtet. Damit gehört Deutschland zu den 26% der untersuchten Staaten, die eine nationale Lizenz eingeführt haben. Die Mehrheit der Staaten empfiehlt aber die Nutzung von CC-Lizenzen. Darüber hinaus wird die Bedeutung der DCAT-AP Konformität für die nationalen Datenportale unterstrichen. Die meisten Staaten untersuchen die Ursachen für den Mangel an Konformität.

In Deutschland wären dies beispielsweise fehlende Kenntnis über Linked Open Data, Probleme bei der Verwendung korrekter URLs, Schwierigkeiten hinsichtlich der Zuordnung der richtigen Themen zu den Datensätzen und sowie fehlende Upgrades auf neuere Versionen von DCAT-AP und DCAT-AP.de¹³ (vgl. ebd., S. 67 ff.).

4.2.3 Infrastrukturen und Stakeholder

Um die Interessenträger im *Public Sector* systematisch zu ermitteln, ist es notwendig, die Individuen und Organe zu identifizieren, die entweder Einfluss auf die FAIRifizierung des Datenmanagements der Public-Sector-Daten haben oder davon beeinflusst werden – direkt oder indirekt. Die Stakeholderanalyse diene dazu, eine optimierte Interaktion in den Untersuchungsphasen des Projektes mit den relevanten Bezugsgruppen herzustellen.

Zu den Stakeholdern des Untersuchungsbereiches *Public Sector* zählen einerseits die Wissenschaft sowie alle dort beteiligten Akteur*innen wie Wissenschaftler*innen, Forschungseinrichtungen, Universitäten und Infrastrukturen wie die NFDI in Deutschland oder die europäische Wissenschaftscloudinfrastruktur EOSC. Diesbezüglich gibt es Überschneidungspunkte mit dem Untersuchungsbereich *Wissenschaft* im Projekt EcoDM (vgl. Kapitel 4.1). Andererseits sind als

¹² <https://www.govdata.de/dl-de/by-2-0>

¹³ <https://www.dcat-ap.de/>

Bislang neben dem Bund an GovData beteiligte Länder

An GovData beteiligte Länder:

- Baden-Württemberg
- Bayern
- Berlin
- Brandenburg
- Bremen
- Hamburg
- Hessen
- Mecklenburg-Vorpommern
- Niedersachsen
- Nordrhein-Westfalen
- Rheinland-Pfalz
- Sachsen
- Schleswig-Holstein
- Thüringen



Abbildung 12: Darstellung der an GovData beteiligten Bundesländer (Quelle: EcoDM)

Stakeholder direkte Akteur*innen des *Public Sectors* zu identifizieren wie Bund, Länder, Kommunen sowie bestimmte Körperschaften des öffentlichen Rechts (vgl. Definition in Kapitel 1.2) und öffentliche Unternehmen. Desweiteren gehören auch private Unternehmen und somit Akteur*innen aus der Wirtschaft zu Stakeholdern des *Public Sectors* sowie die wirtschaftsfokussierte Infrastruktur Gaia-X (vgl. Kapitel 4.3), die auch Interessen des *Public Sectors* widerspiegelt. Darüber hinaus ist sowohl aktiv als auch passiv die Öffentlichkeit bzw. die Gesellschaft als Stakeholder des *Public Sectors* von der Umsetzung des digitalen Datenmanagements betroffen.

Stakeholder im *Public Sector* sind:

- Förderer: Bundesministerium für Bildung und Forschung (BMBF), Deutsche Forschungsgemeinschaft (DFG), Stiftungen
- Bundesbehörden, Ministerien
- Länder und Kommunen sowie ihre Verwaltungen
- strategische Planungsorgane und -einrichtungen (Gremien)
- Infrastrukturen Inland: Einrichtungen des Bundes sowie der Länder zur Datenbereitstellung, also Datenportale wie GovData, Länderportale und NFDI
- Infrastrukturen europäische Ebene: EOSC, Gaia-X

- öffentliche Unternehmen
- Forschungseinrichtungen und -projekte
- Wissenschaftler*innen
- Bürger*innen (breite Öffentlichkeit)
- Angestellte des öffentlichen Dienstes

Im Jahr 2013 wurde GovData als nationales Metadaten-Portal initiiert, auf dem öffentliche Daten von Bund, Ländern und Kommunen eingestellt werden sollen. Im Jahr 2023 wird GovData in das Produktportfolio des IT-Planungsrats überführt und damit von der FITKO (Föderale IT-Kooperation) verwaltet werden. Das Portal dient als Katalog, in dem die Daten zentral verzeichnet und nach Möglichkeit offen zugänglich sind. Gespeichert werden die Daten weiterhin dezentral auf bestehenden Websites oder Portalen. Durch ein einheitliches Metadatenschema wird sichergestellt, dass die Daten bestmöglich auffindbar sind. Allerdings unterstützen die Länder Sachsen-Anhalt und Saarland das Portal nicht (Abbildung 12) und die Menge an eingestellten Daten ist vergleichsweise überschaubar, weshalb GovData bisher hinter seinen Möglichkeiten zurückbleibt (vgl. Schweigel et al., 2020, S. 36 ff.).

Darüber hinaus hat der Bund mit dem Kompetenzzentrum Open Data (CCOD)¹⁴ eine Beratungsstelle

¹⁴ https://www.bva.bund.de/DE/Services/Behoerden/Beratung/Beratungszentrum/OpenData/opendata_node.html

eingrichtet, die Bundesbehörden bei der Umsetzung von Open Data unterstützen sowie als Kontakt für Open-Data-Stellen der Länder fungieren soll.

Neben Plattformen auf Bundesebene verfügen auch einzelne Bundesländer oder Kreise und Kommunen über Open-Data-Portale. Es wird jedoch keine offizielle Übersicht über diese Portale z. B. über GovData angeboten. Eine solche Liste wurde aber von Thomas Tursics erstellt und ist sowohl über das OK Lab Berlin¹⁵ als auch über das offene Datenportal der Stadt Moers einsehbar.¹⁶ Dieses Beispiel illustriert das Entwicklungspotential, welches bei den öffentlichen Verwaltungen in Deutschland hinsichtlich der öffentlichen Zugänglichmachung von Daten besteht.

4.2.4 Öffentlicher Auftrag und gesetzliche Vorgaben

Es gibt diverse Gründe, die dafür sprechen, Daten, deren Erhebung mit öffentlichen Mitteln finanziert wurde, offen zur Verfügung zu stellen. Zunächst tragen sie zur Verwaltungstransparenz bei. Dies macht die öffentliche Verwaltung bürger*innennäher und fördert die Partizipation. Wichtig ist auch die Nutzung von aus der öffentlichen Verwaltung stammenden Daten durch die Wissenschaft. Hier kann vielfach eine Reziprozität angenommen werden, weil Forschungsergebnisse, die auf der Auswertung von Daten der Verwaltung beruhen, regelmäßig für eben diese Verwaltungen interessant sein werden. Darüber hinaus sprechen wirtschaftliche Erwägungen für das Datenteilen, denn von einer freien Verfügbarkeit öffentlicher Daten können Unternehmen profitieren. Dass den Daten des *Public Sectors* eine Schlüsselposition zukommt, lässt sich zudem daran ablesen, dass die angesprochenen Nachnutzungen auch innerhalb des *Public Sectors* erfolgen. Mehrere Studien über die Nutzung von Behördendaten zeigen, dass viele Zugriffe von anderen Behörden erfolgen, also die Kommunikation und Zusammenarbeit innerhalb der Verwaltung maßgeblich gestärkt und vereinfacht wird (vgl. Schweigel et al., 2020, S. 31).

Wesentlicher Treiber der Digitalisierung sowie der Bereitstellung von Daten durch den *Public Sector* ist die Politik, die durch Strategien und Gesetze die ihr nachgeordneten Verwaltungen mit wachsendem Druck zur öffentlichen Zugänglichmachung eigener Datenbestände verpflichtet. Innerhalb der Europäischen Union sind hier an erster Stelle die Richtlinie 2003/98/EG über die Weiterverwendung von Informationen des öffentlichen Sektors und ihre Nachfolgerinnen aus den Jahren 2013 und 2019 zu nennen. In vielen Kommunen wurden in Ergänzung zur Erfüllung von Vorgaben der genannten EU-Richtlinien bzw. der deutschen Gesetze,

mit welchen diese Richtlinien in deutsches Recht überführt wurden, zusätzliche Anstrengungen zur Bereitstellung von Daten unternommen. Vielfach geschah dies als Reaktion auf Forderungen, die von der lokalen Zivilgesellschaft formuliert wurden. Besonders deutlich lässt sich dieser Zusammenhang am Beispiel von Umweltdaten illustrieren. Hier war und ist das Interesse so groß, dass in Form der Umweltinformationsgesetze eigenständige gesetzliche Regelungen geschaffen wurden, die sogar durch eine internationale Konvention als eine Art globaler Standard verankert wurden (vgl. Richtlinie (EU) 2019/1024). Aus der Begründung der jüngsten Novellierung der PSI-Richtlinie wird deutlich, dass der europäische Gesetzgeber weiterhin erheblichen Handlungsbedarf bei der Zugänglichmachung von Daten des *Public Sectors* sieht. So heißt es dort in Erwägungsgrund drei, „dass Maßnahmen auf Unionsebene erforderlich sind, um die verbleibenden und neu entstehenden Hemmnisse zu beseitigen, die einer breiten Weiterverwendung von öffentlichen und öffentlich finanzierten Informationen in der gesamten Union im Wege stehen, um den Rechtsrahmen auf den neuesten Stand der digitalen Technik zu bringen und um weitere digitale Innovationen [...] zu fördern.“ (ebd., (3)). Neben Gesetzen formulieren sowohl die EU als auch der Bund und die Länder Strategien zur digitalen Entwicklung (vgl. Kapitel 4.2.2).

Im März 2021 rief die EU eine digitale Dekade für Europa aus. In Bezug auf den *Public Sector* sind die Ziele bis 2030 demnach, dass 100 % der wesentlichen öffentlichen Dienste digitalisiert sind, 80 % der Bürger*innen der EU über eine digitale ID verfügen und 100 % der Bürger*innen Zugang zu ihren Patientenakten haben (vgl. Europäische Kommission, 2021, S. 5).

Die Datenstrategie der Bundesregierung, die im Januar 2021 beschlossen wurde, definiert ihre Ziele weniger konkret. Sie will die „innovative und verantwortungsvolle Datenbereitstellung und Datennutzung insbesondere in Deutschland und Europa signifikant erhöhen – in der Wirtschaft, der Wissenschaft, der Zivilgesellschaft und der Verwaltung“ (Bundesregierung, 2021a, S. 6.). Sie gründet sich auf vier Säulen: a) die Ausgestaltung leistungsfähiger und nachhaltiger Dateninfrastrukturen, b) das Steigern einer innovativen und verantwortungsvollen Datennutzung, c) das Steigern der Datenkompetenz bzw. das Etablieren einer Datenkultur sowie d) die Entwicklung des Staates zum Vorreiter der angestrebten Datenkultur. Dabei bezieht sich vor allem die letzte Säule auf den Ausbau digitaler Kompetenzen und Netzwerke in (Bundes-)Behörden. Explizit genannt wird auch die Umsetzung der FAIR-Prinzipien bei hochwertigen Datensätzen (vgl. ebd., S. 91). Die Bundesländer wiederum arbeiten an eigenen

¹⁵ <http://opendata.tursics.de/?page=Data&level=all&dataset=portals&country=DE&lat=51.1928&lng=9.0629&zoom=6>
¹⁶ <https://www.offenesdatenportal.de/dataset/ubersicht-der-open-data-angebote-in-deutschland>

Land	Strategie
Baden-Württemberg	<p>PDF: Ministerium für Wissenschaft, Forschung und Kunst Baden-Württemberg, E-Science. Wissenschaft unter neuen Rahmenbedingungen. Fachkonzept zur Weiterentwicklung der wissenschaftlichen Infrastruktur in Baden-Württemberg, Mai 2014. https://mwk.baden-wuerttemberg.de/fileadmin/redaktion/m-mwk/intern/dateien/pdf/Forschung/066_PM_Anlage_E-Science_Web.pdf</p> <p>Homepage: https://mwk.baden-wuerttemberg.de/de/forschung/forschungslandschaft/e-science/</p>
Berlin	<p>PDF: Senat von Berlin, Open-Access-Strategie für Berlin, Juli 2015. http://www.open-access-berlin.de/_dokumente/Open-Access-Strategie_fuer_Berlin_2015-07-10.pdf Kindling et al., Open-Access-Landesstrategien und der Weg zur offenen Wissenschaft – Möglichkeiten und Herausforderungen am Beispiel Berlin, September 2021: https://zenodo.org/record/5535754#.YcBZfFkxk2w</p> <p>Homepage: http://www.open-access-berlin.de/strategie</p>
Brandenburg	<p>Ellen Euler, Open-Access-Strategie des Landes Brandenburg, August 2019. https://zenodo.org/record/3757920#.YcBaS1kxk2w</p> <p>Homepage: https://open-access-brandenburg.de/</p>
Hamburg	<p>PDF: Mitteilung des Senats an die Bürgerschaft. Haushaltsplan 2017/2018. Einzelplan 3.2 der Behörde für Wissenschaft, Forschung und Gleichstellung Programm Hamburg Open Science (HOS), 26.09.2017. https://www.buergerschaft-hh.de/ParlDok/dokument/59359/haushaltsplan-2017-2018-einzelplan-3-2-der-behoerde-fur-wissenschaft-forschung-und-gleichstellung-programm-hamburg-open-science-hos-.pdf</p> <p>Homepage: https://openscience.hamburg.de/de/startseite-hamburg-open-science/</p>
Niedersachsen	<p>PDF: Landeshochschulkonferenz Niedersachsen & Niedersächsisches Ministerium für Wissenschaft und Kultur, Leitlinien zur Transparenz in der Forschung. Gemeinsame Position der Landeshochschulkonferenz Niedersachsen und des Niedersächsischen Ministeriums für Wissenschaft und Kultur, Februar 2015. https://www.mwk.niedersachsen.de/download/94171/Leitlinien_zur_Transparenz_in_der_Forschung.pdf</p>
Nordrhein-Westfalen	<p>PDF: Die Landesregierung Nordrhein-Westfalen, Strategie für das digitale Nordrhein-Westfalen 2.0. Teilhabe ermöglichen – Chancen eröffnen, November 2021. https://www.digitalstrategie.nrw/digitalnrw/de/home/file/fileId/882/name/2021-11-16a%20Digitalstrategie-Update-final.pdf</p> <p>Homepage: Vorprojekt zur Open-Access-Strategie NRW: https://openaccess.nrw/</p>
Schleswig-Holstein	<p>PDF: Schleswig-Holstein Ministerium für Soziales, Gesundheit, Wissenschaft und Gleichstellung, Strategie 2020 der Landesregierung Schleswig-Holstein für Open Access, November 2014. https://www.schleswig-holstein.de/DE/Fachinhalte/H/hochschule_allgemein/Downloads/open_access_strategie.pdf</p> <p>Homepage: https://www.schleswig-holstein.de/DE/Fachinhalte/H/hochschule_allgemein/OpenAccess.html</p> <p>Open-Access-Infopoint: https://oa-info.sh/</p>
Thüringen	<p>PDF: Thüringer Ministerium für Wirtschaft, Wissenschaft und Digitale Gesellschaft, Thüringer Strategie für die Digitale Gesellschaft, Januar 2020. Update 2021 https://www.digitalagentur-thueringen.de/thueringer-digitalstrategie/</p> <p>Homepage: https://www.digitalagentur-thueringen.de/</p>

Tabelle 4: Bundesländer mit Open-Access-Strategien. In der Tabelle sind die Open-Access- bzw. die Open-Science-Strategien der Bundesländer aufgelistet, sofern vorhanden (Quelle: EcoDM)

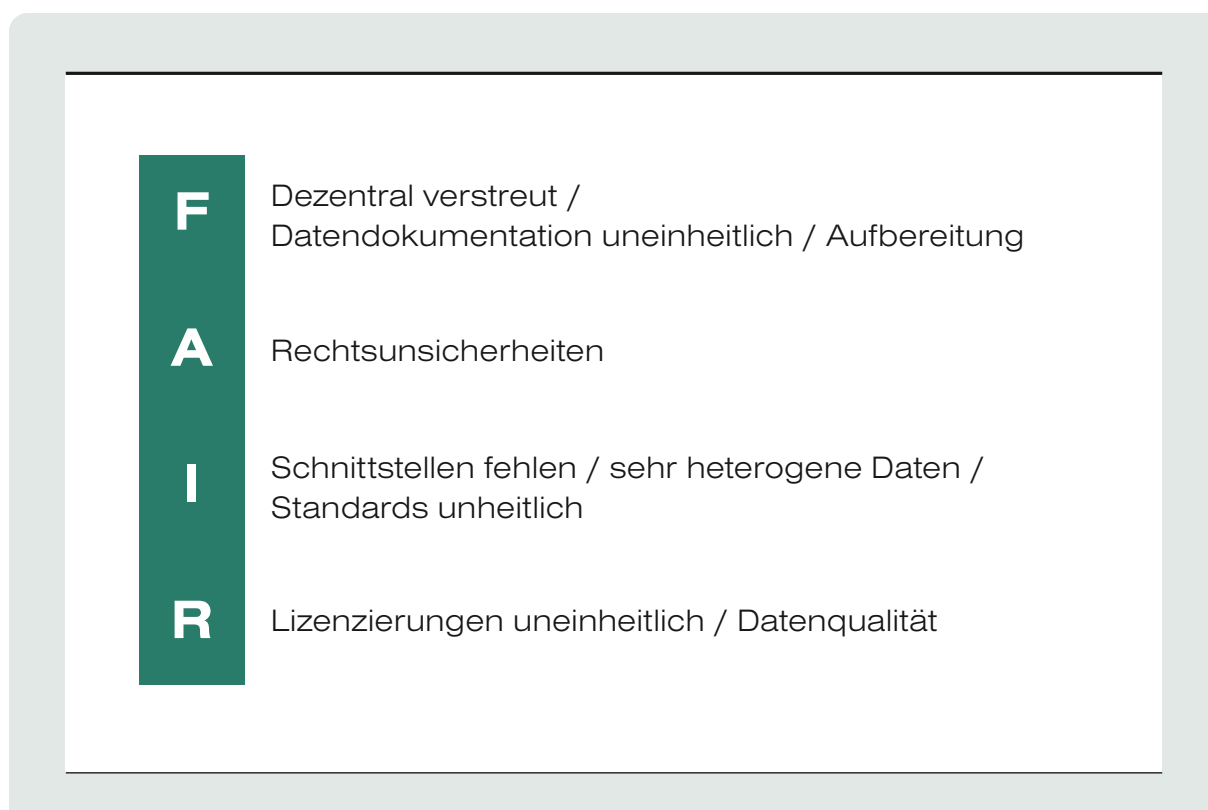


Abbildung 13: Schlagworte, die einer FAIRen Datennutzung entgegenstehen und somit die Auffindbarkeit, Zugänglichkeit, Interoperabilität und Wiederverwendbarkeit erschweren (Quelle: EcoDM)

Digitalisierungsstrategien. Dabei sind die Unterschiede, vor allem im Bezug auf die Bereitstellung öffentlicher Daten, allerdings enorm (vgl. Kapitel 3.3.2). So haben bisher nur einige Länder Open-Science- und Open-Access-Strategien entwickelt. Tabelle 4 gibt dazu einen Überblick.

4.2.5 Defizite und Handlungsbedarf

Die Defizite in der Umsetzung von Open Data bzw. der FAIR-Prinzipien im *Public Sector* sind teilweise schon angesprochen worden. In Abbildung 13 werden Problemfelder anhand der FAIR-Prinzipien zusammengefasst. Anschließend folgt eine Zusammenfassung der Hindernisse und Bedarfe bei der Bereitstellung von Daten des *Public Sectors* sowie eine Sammlung von Herausforderungen für die Forschung bei der Nutzung von Public-Sector-Daten und daraus abgeleitete Vorschläge. Diese wurden im Rahmen des 5. Arbeitsforums der mFUND-Begleitforschung des WIK unter dem Titel „Open Data und Behörden“¹⁷ im April 2020 erarbeitet und stehen beispielhaft für die Defizite und den Handlungsbedarf im *Public Sector*. Der mFUND wird gefördert vom Bundesministerium für Digitales und Verkehr und sorgt neben der

finanziellen Förderung von Projekten im Rahmen der Mobilität-4.0-Initiative für die Vernetzung zwischen Politik, Wirtschaft sowie Forschung. Eines der Arbeitsforen beschäftigte sich mit dem Thema *Open Data und Behörden*. Die Analyse der veröffentlichten Protokolle und die Teilnahme der EcoDM-Projektmitarbeiterin an einem Arbeitsforumtreffen dienten der Ermittlung von Hindernissen und Bedarfen aus der Sicht von Projekten, die auf Behörden- und Verwaltungsdaten zugreifen und diese nutzen wollen.

Neben informativen Protokollen der einzelnen Treffen veröffentlichte das Arbeitsforum auch einen Leitfaden für Behördendaten „Zugang zu Behördendaten für Digitalisierungsprojekte des mFUND“.¹⁸

Die mFUND-Arbeitsforumtreffen dienten dazu, einen Austausch für assoziierte Projekt zu ermöglichen. Verschiedene mFUND-Projekte konnten in diesem Rahmen ihre Ziele und ihr Vorgehen präsentieren sowie Kontakte knüpfen und Erfahrungen miteinander teilen. Während des Arbeitsforumtreffens im April 2020 kam es zu einem regen Austausch zwischen den Teilnehmenden zur Nutzung von Behördendaten. Es wurden Hindernisse, denen sich die Projekte gegenübergestellt sehen, und Bedarfe für die wissenschaftliche Nutzung von Behördendaten geschildert.

¹⁷ <https://www.wik.org/mfund/arbeitsforum-behoerden>

¹⁸ https://www.wik.org/fileadmin/mFUND_VF/mFUND-WIK-Studie-Zugang_Behoerndaten_AF_Open_Data_und_Behoerden.pdf

Hindernisse und Bedarfe bei der Bereitstellung von Daten des Public Sectors

Folgende Hindernisse und Bedarfe, denen sich der *Public Sector* bei der Bereitstellung von Daten aus der Sicht der Forschungsprojekte gegenüber sieht, können aus den Protokollen und der Teilnahme an den Arbeitsforentreffen festgehalten werden:

Hindernisse

1. *Ein teilweise fehlendes Bewusstsein der Behörden und Kommunen:*
 - a. *zum Umgang mit den eigenen Daten*
 - b. *für den Nutzen der Daten und deren Nutzbar-machung*
 - c. *hinsichtlich des Potentials der Kommerzialisierung bzw. Mehrwertschaffung*
2. *Ein hoher Aufwand für die Datenaufbereitung und -pflege für mittelgroße oder kleinere Städte*
3. *Fehlende Datenkompetenz*
4. *Konkurrenz- und Wettbewerbsängste*
5. *Rechtliche Unsicherheiten, besonders Datenschutz-bedenken*

Bedarfe

1. *Das Know-How, gerade bei Kleinkommunen, muss erarbeitet bzw. ausgebaut werden.*
2. *Eine Handbuch zur Mehrwertschaffung und Nutzung der eigenen Daten wird benötigt.*
3. *Es fehlen notwendige Expert*innen an den Einrichtungen wie etwa Data Scientists.*

Es wurde auch die Frage gestellt, auf welcher Ebene die Datenkompetenz vorhanden sein sollte und der Vorschlag geäußert, dass die Forschungsprojekte selbst auf den Nutzen der Behördendaten aufmerksam machen könnten.

Es wurde jedoch nicht nur die Problematik mit behördlichen Daten angesprochen, sondern auch die Nutzung von Daten durch private Anbieter diskutiert. So wurde geschildert, dass ein großes Hindernis der Datenveröffentlichung auch Konkurrenzängste darstellen. Dadurch werden Daten häufig nur intern zur Verfügung gestellt.

In der Diskussion während des 5. Arbeitsgruppentreffens wurde darauf verwiesen, dass auch die mFUND-Projekte selbst eine Strategie bezüglich der Öffnung nicht nur von Rohdaten, sondern auch ihrer prozessierten Daten entwickeln müssen. Der Mehrwert der Öffnung dieser Daten für die Öffentlichkeit und andere Forschungsprojekte wurde hierbei deutlich hervorgehoben.

Herausforderungen für die Forschung bei der Nutzung von Public-Sector-Daten

Aus den Berichten und Diskussionen der Projekte konnten im Rahmen von EcoDM Herausforderungen identifiziert werden, denen sich Forschungsprojekte gegenübersehen, die für ihre Untersuchungen auf Daten des *Public Sectors* angewiesen sind. Für die Forschung ist ein strukturiertes Datenmanagement der Public-Sector-Daten von der Erhebung über die Zusammenführung bis zur Veröffentlichung von außerordentlicher Bedeutung. Open Public Data und die Umsetzung der FAIR-Prinzipien sind hier hervorzuhebende Schlagworte.

1. *Hoher Aufwand durch verteilte Informationsquellen*
2. *Zugriff auf Daten ist häufig erschwert: Rechtsbedenken*
3. *Zugriffsregelungen sind unklar und uneinheitlich: Nutzungsbedingungen müssen individuell ausgehandelt werden*
4. *Hohe Investitionsmittel für Lizenzgebühren, Aufbereitungskosten etc. notwendig*
5. *Kostenabschätzung im Vorfeld schwierig*
6. *Die Beschaffenheit der Datenqualität, -konsistenz sowie Inhalte ist sehr unterschiedlich*
 - a. *Metadateninformationen sind oft unzureichend etc.*
 - b. *Dateiformate sind uneinheitlich*
=> *Angleichungen sind notwendig*
=> *Aufbereitung aufwendig*
 - c. *Unsicherheit über die Datenqualität oder Inhalte*
=> *Schwierigkeit der Kostenplanung*
 - d. *Inhalt und Aussage der Daten oftmals schwierig zu ermitteln*

Die Projekte sehen sich scheinbar oftmals konfrontiert mit einem schwierig einzuschätzenden Zeit- und Finanzrahmen. Begründet ist dies in den verstreuten Datenquellen und der Unsicherheit, wo relevante Daten eventuell noch vorgehalten sein könnten. So ist zwar beispielsweise das Portal GovData ein Anlaufpunkt, über das auf offene Daten aus Bund, Ländern und Kommunen verwiesen wird. Es sind jedoch bisher weder alle offenen Verwaltungsdaten dort durch ihre Metadaten verzeichnet, noch sind alle Daten tatsächlich ohne Zugriffsbeschränkungen einsehbar (siehe dazu auch die Informationen zum Portal auf der Internetpräsenz).¹⁹

Wir schlussfolgern daraus, dass mit GovData als Metadaten-Portal für Verweise auf (möglichst) offene Verwaltungsdaten ein Schritt in die Richtung eines zentralen Anlaufpunktes für die Datenrecherche eingerichtet worden ist. Im Sinne der Förderung der Forschung, deren Effektivitätssteigerung und auch aus

¹⁹ <https://www.govdata.de/web/guest/hilfe>

ökonomischen Gründen ist jedoch der weitere Ausbau zu einem zentralen Metadaten-Portal unter tatsächlicher Beteiligung aller Länder dringend notwendig.

Ein häufiger Grund für eingeschränkte oder verweigernde Zugriffsrechte auf Public-Sector-Daten ist laut der Projekte die oftmals undurchsichtige Rechtslage für deren Nutzung. Es werden insbesondere Datenschutzbedenken angeführt. Die Auslegung der Rechtsgrundsätze und damit auch die Verantwortung liegt bei den einzelnen Daten-haltenden Einrichtungen. Daher wird der Zugriff auf manche Public-Sector-Daten vorsichtshalber nicht genehmigt. Die Definition, welche Daten geschützt sind, scheint nicht eindeutig genug, um den Zugang für die Forschung einheitlich zu regeln.

Da die Nutzungsbedingungen individuell mit den Datenlieferanten ausgehandelt werden müssen, ist für die Projekte oft ein langer Atem notwendig. Die Ausarbeitung von Nutzungsverträgen hat somit sowohl zeitliche als auch finanzielle Auswirkungen und ist im Vorfeld bei der Projektplanung schwierig abzuschätzen. Hier hat der mFUND in seinem Leitfaden für den Zugang zu Behördendaten eine kurze Checkliste erstellt.²⁰ Vorschläge aus dem Arbeitsgruppentreffen:

- OGP (Open Government Partnership) Initiative sollte auch Länder- und Kommunenebene einbeziehen bzw. auch für diese sollte ähnliches existieren. Das OGP betrifft formell die nationale Regierungsebene. Seit zwei Jahren existieren erste Piloten zu subnationalen Teilnahmen und seit dem 2. Nationalen Aktionsplan wird versucht, auch die Länder / Kommunen miteinzubeziehen.
- Standardverträge zu Nutzungsbedingungen sollten erstellt und beworben werden. So hat mFUND z. B. Standardverträge als Vorlage erstellt und im Erfahrungsbericht veröffentlicht. Solche Materialien sollten an zentraler Stelle gesammelt und zur Verfügung gestellt werden.
- Digitale Angebote für bessere Datennutzung sollten für einen erweiterten Nutzen der Daten geschaffen bzw. ausgebaut werden.

4.2.6 Zusammenfassung der Kernpunkte

Die Literaturanalyse zeigt auf, dass der Umgang mit digitalen Daten in der Verwaltung in Deutschland sehr unterschiedlich ist. Die Aktualisierung des E-Government-Gesetzes und die Einführung des Datennutzungsgesetzes 2021 auf Bundesebene sind begrüßenswert. Allerdings sind die Länder und Kommunen in ihrer Entwicklung beim Umgang mit Daten und deren öffentlichen Zugänglichkeit sehr unterschiedlich weit vorangeschritten. Während einige Bundesländer die proaktive Veröffentlichung von Behördendaten

festgeschrieben haben, verfügen andere noch nicht einmal über ein Informationsfreiheitsgesetz, das betroffenenunabhängige Ansprüche auf Zugang zu Daten bei öffentlichen Stellen regelt.

Ähnlich sieht es bei Datenstrategien aus. So steckt beispielsweise die EU teilweise sehr ambitionierte Ziele ab. Allerdings kann sie aufgrund mangelnder Kompetenz und fehlender Sanktionsmöglichkeiten bei Nicht-Erfüllung die Umsetzung dieser Ideen nicht garantieren. Auch der Bund möchte mittels seiner Datenstrategie mit gutem Beispiel vorangehen und hat seinen Behörden mit dem Kompetenzzentrum Open Data (CCOD) einen Anlaufpunkt zur Hilfestellung eingerichtet. In Form von GovData ist zudem ein zentraler Katalog von Behördendaten von Bund und Ländern initiiert, in dem Metadaten standardisiert vorliegen, sodass sich die FAIR-Prinzipien zum Standard entwickeln können. Allerdings wird diese Infrastruktur noch nicht voll ausgeschöpft.

Open Public Data haben das Potential für Synergieeffekte in Wirtschaft, Wissenschaft und Gesellschaft, weil sich an ihnen Entwicklungen analysieren und Trends prognostizieren lassen. Es fehlt aber noch an dem nötigen Bewusstsein zum Umgang und für die Nutzung von Daten sowie oftmals an Kompetenz und Know-how.

4.3 Wirtschaft

4.3.1 Quellengrundlagen

Im Untersuchungsbereich *Wirtschaft* gibt es neben unterschiedlichen politischen Strategiedokumenten auch diverse Artikel und Bücher verschiedener Experten*innen sowie Veröffentlichungen durch Stiftungen und Verbände. Sie alle versuchen, zum Teil eine ökonomische, ökologische und rechtliche Grundlage für den Datenaustausch innerhalb der Wirtschaft (Business-to-Business, *B2B*) und mit anderen Bereichen (z. B. Wissenschaft) zu ermöglichen. Zudem stecken sie Rahmenbedingungen, Handlungsempfehlungen und Strategien ab, wobei eine Realisierung und Umsetzung dieser oftmals in den verschiedenen Wirtschaftssektoren unterschiedlich erfolgreich bzw. möglich ist. Eine separate Betrachtung und Auswertung einzelner Wirtschaftssektoren wird an dieser Stelle und im Rahmen des Projekts nicht vorgenommen. Wie in den Untersuchungsbereichen *Wissenschaft*, *Public Sector* und *Qualifizierung* werden ausschließlich aktuelle Publikationen aus Deutschland und der EU berücksichtigt, die sich übergreifend und allgemeingültig mit dem Untersuchungsbereich *Wirtschaft* auseinandersetzen.

²⁰ https://www.wik.org/fileadmin/mFUND_VF/mFUND-WIK-Studie-Zugang_Behoerndaten_AF_Open_Data_und_Behoerden.pdf, Checkliste dort auf S. 16.

4.3.2 Rahmenbedingungen

Laut ifo-Forschungsbericht (Falck & Koenen, 2020) zum volkswirtschaftlichen Nutzen bei der Datenbereitstellung sind Mitte 2018 die sieben Unternehmen mit der global höchsten Marktkapitalisierung Internet- und Technologiekonzerne gewesen. Deren Geschäftsmodelle basieren „zu bedeutenden Teilen auf dem Sammeln sowie der Analyse und Nutzung von Daten“. Zudem führen aktuelle technologische Entwicklungen dazu, „dass sowohl die Menge der vorhandenen digitalen Daten als auch ihr potentieller wirtschaftlicher Wert rasant ansteigen“ (ebd., S. 2). Die International Data Corporation (vgl. Reinsel et al., 2018) schätzt, dass sich bis zum Jahr 2025 der Wert der europäischen Datenwirtschaft (EU-27) nahezu verdreifachen (von 301 Mrd. € im Jahr 2018 auf 829 Mrd. €), das weltweite Datenvolumen mehr als verfünffachen (von 33 Zettabyte im Jahr 2018 auf 175 Zettabyte) und sich die Zahl der Datenexperten*innen ungefähr verdoppeln (von 5,7 Mio. im Jahr 2018 auf 10,9 Mio.) wird. Die Bertelsmann Stiftung (Pawelke, 2020, S. 6) wertet diese Schätzung als Beleg für die zunehmende Relevanz des Teilens von Daten im wirtschaftlichen Kontext.

Oftmals lenkt die Fokussierung auf die wirtschaftlichen Potentiale der Nutzung von Daten die Aufmerksamkeit von weiteren positiven sozialen und gesellschaftlichen Aspekten ab, die sich durch Teilen von Daten ergeben können (vgl. Pawelke, 2020, S. 6). Die Bertelsmann Stiftung (ebd., S. 6) betont, dass das Teilen von Daten längst auch eine zutiefst gesellschaftspolitische Frage sei, da „wissenschaftlicher Fortschritt, prosperierende Wirtschaft oder gesellschaftlicher Zusammenhalt“ letztlich davon abhängen, „wer Daten sammelt, mit wem sie geteilt werden, welche Ziele wir bei der Datennutzung verfolgen und wie wir die enormen Mengen von Daten bei der Entwicklung algorithmischer Systeme nutzen“. Mit Blick auf den Wettbewerb um Daten als wirtschaftliche Ressource zählt nach Einschätzung der Bertelsmann Stiftung das mangelnde Vertrauen insbesondere in die Logik der bestehenden Medienökonomie zu den zentralen Risiken und Herausforderungen (ebd., S. 12).

Aufgrund der stark wachsenden Bedeutung der Datenwirtschaft gewinnt vor allem die Frage, wer die Kontrolle über die Daten und somit auch über deren wirtschaftlichen Wert hat, immer stärker an Relevanz. Laut ifo-Forschungsbericht (vgl. Falck & Koenen, 2020) ist die hierbei zunehmende Bedeutung der Medienökonomie für die zukünftige deutsche und europäische Wirtschaftsleistung unbestritten, wenngleich eine zugrundeliegende Studie „verschiedene Faktoren identifiziert, die zum Versagen von Märkten von Daten führen können“, und es sei „daher zweifelhaft, dass der Markt alleine zu einer effizienten Allokation von Daten und zu optimalem Zugang und ausreichender Teilhabe an den

gesammelten Daten führen kann“ (Falck & Koenen, 2020, S. 25). Mit Blick auf dieses konstatierte Marktversagen wird gegenwärtig „mit hoher Dringlichkeit an Lösungsansätzen gearbeitet, wobei staatliche Eingriffe wie ein Recht auf Datenzugang mit privatwirtschaftlichen Lösungen, wie Data Trust oder Plattformen parallel verfolgt werden“ (ebd., S. 5) und angesichts der komplexen Problematik sei „ein Wettbewerb um Lösungen begrüßenswert“ (ebd., S. 25).

Unabhängig davon werden von der Europäischen Kommission (2020a) in ihrer Datenstrategie eine Reihe von Problemen benannt, welche die Datenwirtschaft in der EU an der Ausschöpfung ihrer Potentiale hindern. Neben unterschiedlichen rechtlichen Rahmenbedingungen in den Mitgliedstaaten werden vor allem die mangelnde Verfügbarkeit von Daten, ungleiche Marktmacht sowie Defizite mit Blick auf Interoperabilität, Datenqualität, Datengovernance, Dateninfrastrukturen und Datentechnologien als Hindernisse in der Datenwirtschaft aufgeführt (ebd., S. 7). Im Zuge dessen spricht sich die Europäische Kommission mit ihrer Datenstrategie dafür aus, dass „Daten für das öffentliche Wohl“ eingesetzt werden, denn sie „werden von der Gesellschaft hervorgebracht und können dazu dienen, Notlagen wie Überschwemmungen und Waldbrände zu bewältigen, den Menschen ein längeres und gesünderes Leben zu ermöglichen, öffentliche Dienstleistungen zu verbessern und die Umweltzerstörung und den Klimawandel zu bekämpfen“ (ebd., S. 7). Dementsprechend sollten nicht nur „vom öffentlichen Sektor erzeugte Daten und der damit verbundene Mehrwert“, sondern auch Daten „aus dem privaten Sektor“ dem Gemeinwohl dienen und entsprechend eingesetzt werden (vgl. Kapitel 3.4.1.2.) (ebd., S. 8). Auch die Datenstrategie der Bundesregierung (2021a, S. 6 ff.) zielt auf eine bereichsübergreifende, signifikante Erhöhung innovativer wie auch verantwortungsvoller Datenbereitstellungen und -nutzungen ab. Diese sollen auf Basis „der europäischen Werte [...] gerechte Teilhabe sichern, Datenmonopole verhindern und zugleich Datenmissbrauch konsequent begegnen“. Das Strategiepapier ist die Reaktion darauf, dass nach Einschätzung der Bundesregierung Daten in Deutschland bisher zu wenig genutzt und geteilt werden. Das liege sowohl an unzureichenden Kenntnissen über den Wert und effizienten Einsatz von Daten als auch an fehlenden Anreizen zur Investition in das Nutzen und Teilen von Daten (vgl. Falck & Koenen, 2020). Allerdings gibt es im ökonomischen Kontext auch Gründe bzw. Befürchtungen, die gegen das Datenteilen sprechen. An erster Stelle sind hier Ängste vor Kontrollverlusten zu nennen, die je nach Kontext die Existenz eines Unternehmens gefährden können (ebd.).

Daten haben in Wertschöpfungsketten von Unternehmen, die in die Sammlung, Verwaltung und Nutzung von Daten investieren, eine zentrale Bedeutung (vgl.

Fraunhofer-Institut für Software- und Systemtechnik ISST, 2019). Diese Unternehmen schrecken deshalb vielfach vor dem Teilen von Daten, insbesondere mit potentiellen Wettbewerbern, zurück. Zudem ist für ein Unternehmen oftmals schwer einschätzbar, ob es durch Datenteilen wirtschaftlichen Nutzen für sich generieren kann (ebd.). Aus diesem Grund werden tendenziell weniger Daten geteilt, als dies auch unter Berücksichtigung der ökonomischen Interessen der einzelnen Unternehmen möglich wäre. Dem aus volkswirtschaftlicher Perspektive bestehendem Problem einer Unternutzung von Daten steht aus bürgerrechtlicher Perspektive, soweit es sich um personenbezogene Daten handelt, die Befürchtung gegenüber, dass jede Nutzung mit dem Risiko der Verletzung von Persönlichkeitsrechten einhergeht. Dies gilt insbesondere, weil auch eine umfassende Anonymisierung personenbezogener Daten nicht sicher vor deren Deanonymisierung schützt. Deshalb setzt jede Form des Datenteilens (auch aggregiert, anonymisiert oder pseudonymisiert) ein gewisses Maß an Vertrauen derjenigen voraus, deren Daten geteilt oder weitergegeben werden, was angesichts von subjektiven wie objektiven Kontrollverlusten gegenwärtig weiterhin fehlt (vgl. Pawelke, 2020).

Laut der Europäischen Kommission (2020a, S. 8) sind nach aktuellem Stand, trotz langjähriger Open-Data-Politik, hochwertige Datensätze des öffentlichen Sektors häufig nicht EU-weit unter gleichen Bedingungen verfügbar. Die dadurch zusätzlich notwendig werdenden Aufwände für den Erwerb bzw. das Sammeln von Daten wirken sich vor allem nachteilig auf kleinere und mittlere Unternehmen (KMU) aus. Zudem werden sensible Daten (z. B. Gesundheitsdaten) aus öffentlichen Datenbanken Unternehmen teilweise nicht zugänglich gemacht, da dort nicht garantiert werden kann, dass die intendierte Nutzung der Daten durchgehend unter rechtskonformen Bedingungen erfolgt (vgl. Busch, 2019; Europäische Kommission, 2020a, S. 8).

Der Bereich B2G-Data (Business-to-Government) ist durch einen Mangel an Datenangebot gekennzeichnet. Das Ziel, durch eine Steigerung der Nutzung von Daten aus Unternehmen Entscheidungsprozesse bei öffentlichen Stellen stärker mit Fakten zu untermauern und damit zu verbessern, wird deshalb nur in Ausnahmefällen erreicht (vgl. Europäische Kommission, 2020a, S. 9). In diesem Kontext wird durch die Europäische Kommission (ebd.) auf die Empfehlung der EU-HLEG (vgl. European Commission. Directorate General for Communications Networks, Content and Technology, 2020) zur „Schaffung nationaler Strukturen für die gemeinsame Datennutzung bei B2G und die Entwicklung geeigneter Anreize für die Schaffung der gemeinsamen Datennutzung“ sowie Prüfung der „Möglichkeiten der Schaffung eines EU-Rechtsrahmens“ verwiesen. Als Beispiele werden im Bericht der

EU-HLEG (ebd., S. 39) steuerliche Anreize zum Teilen von Daten oder öffentlichen Investitionen zur Unterstützung der Entwicklung vertrauenswürdiger Tools sowie Anerkennungssysteme aufgelistet. Hierbei wird betont, dass Datenzugriffe nicht immer kostenlos gewährt werden müssten, sondern ggf. auch ein Gegenwert für die Bereitstellungs- oder Aufbereitungskosten verlangt werden könnte (ebd., S. 39). Für ein nachhaltiges B2G-Datenteilen sollte ein wechselseitiger Nutzen gewährleistet werden, wobei von kostenfreiem Zugang bis hin zu marktüblichen Preisen in verschiedenen Kontexten unterschiedliche Optionen möglich sind (vgl. European Commission. Directorate General for Communications Networks, Content and Technology, 2020).

In diesem Zusammenhang schlägt unter anderen die EU-HLEG (ebd., S. 37 f.) den Einsatz von Datenintermediär*innen vor. Der Oberbegriff Datenintermediär*innen, unter den beispielsweise Begriffe wie „Datentreuhänder*innen“, „data steward“, „data trustee“ fallen, hat sich im Rahmen des Digitalisierungsdiskurses als Analogie für Stellen herausgebildet, die „in besonderer Weise als Intermediäre zwischen einem oder mehreren Datengebern und einer Nachfrageseite agieren“ (Abbildung 14; vgl. Kapitel 3.4.1.1) (RfII – Rat für Informationsinfrastrukturen, 2020a, S. 8). In unterschiedlichen Anwendungsfeldern sollten sie neben den technischen, datenbezogenen Aufgaben (Datenqualität, Standards, etc.) zusätzliche Funktionen beim Datenteilen (data collaboration und coordination) übernehmen (vgl. European Commission. Directorate General for Communications Networks, Content and Technology, 2020, S. 37 f.). Im Rahmen des aktuellen Diskurses um den Aufbau von Datentreuhänder*innen hat der Rat für Informationsinfrastrukturen (RfII) im April 2020 mit einer Stellungnahme reagiert, in der dargestellt ist, dass „noch kein öffentliches oder politisches Einverständnis darüber“ besteht, „wie Datentreuhänder beziehungsweise Daten Treuhandstellen aufgebaut sein sollen“ (RfII – Rat für Informationsinfrastrukturen, 2020a, S. 8). Allerdings müssen Anstrengungen unternommen werden, um die „verantwortungsvolle Bereitstellung und Nutzung von Daten durch Personen und Institutionen [...] signifikant zu steigern“ (Bundesregierung, 2019, S. 1). Dabei gilt es, zukünftig Datenmonopole zu verhindern und Datenpools „in strategischen Sektoren und Bereichen von öffentlichem Interesse“ (Europäische Kommission, 2020a, S. 25) einzurichten, um den Datenzugang für verschiedene Akteur*innen zu erleichtern. Darüber hinaus sollen Datenintermediär*innen „auch datenschutzrechtliche Interessen und Gestaltungsrechte für eine Vielzahl von Verbraucherinnen und Verbraucher geltend machen“ (Bundesregierung, 2021a, S. 110).

Laut der Datenstrategie der Bundesregierung (2021a, S. 33) teilen Unternehmen „Daten überwiegend mit

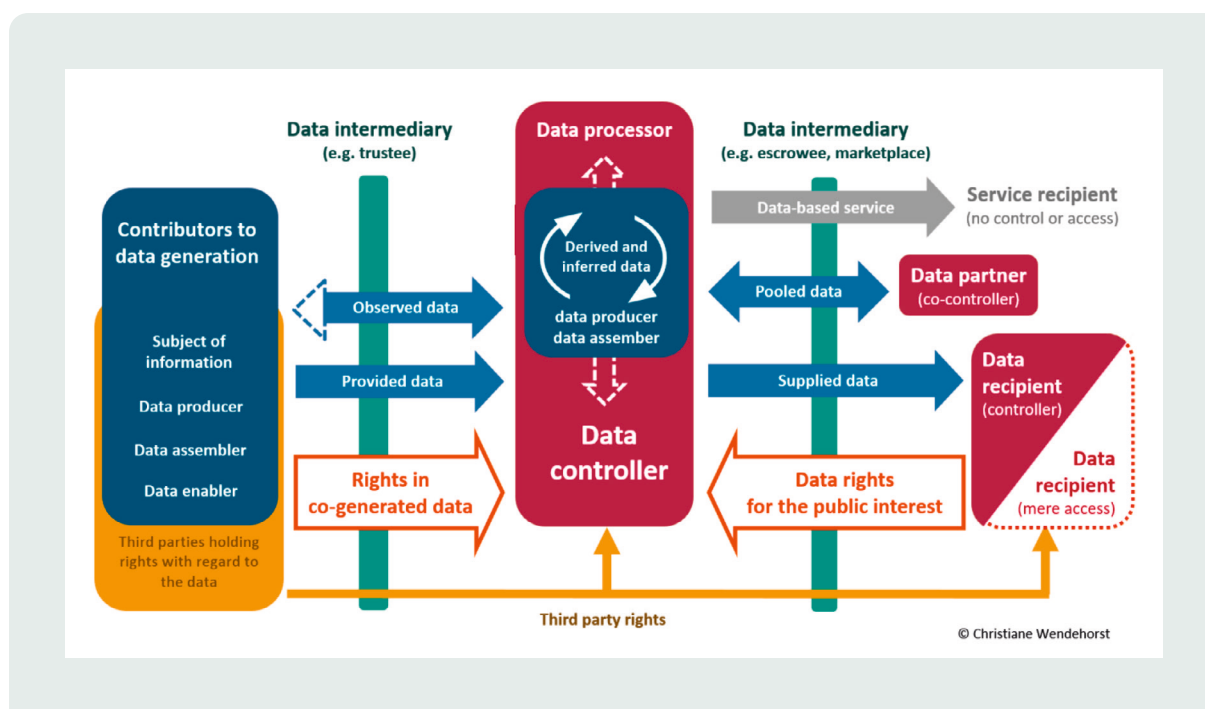


Abbildung 14: Schematische Darstellung zum Einsatz von Datenintermediär*innen (Quelle: Thomas et al., 2021, S. 6; Copyright: Christiane Wendehorst)

Kundinnen und Kunden (83 %), Lieferantinnen und Lieferanten (53 %), viel seltener mit anderen brancheneigenen oder branchenfremden Unternehmen (21 %) oder gar Wettbewerberinnen und Wettbewerbern (15 %).“ KMU sind hinsichtlich des Teilens von Daten mit Unternehmen noch zurückhaltender als große Unternehmen: Je nach Studie tauschen sich 50 % bis 80 % der großen Unternehmen mit anderen Unternehmen aus (ebd., S. 33). Laut Datenstrategie der Bundesregierung (2021) gaben lediglich 41 % der KMU, die sich an einer Umfrage zum Thema Daten teilen beteiligt hatten, an, zu planen, künftig mit externen Partnerinnen und Partnern vermehrt Daten auszutauschen (ebd., S. 33).

Zwischen Unternehmen (B2B) scheitert das Datenteilen trotz seines wirtschaftlichen Potentials oftmals aufgrund fehlender ökonomischer Anreize, Rechtsunsicherheiten, mangelndem Vertrauen, ungleicher Verhandlungspositionen sowie Befürchtungen, Wettbewerbsvorteile einzubüßen und/oder durch missbräuchliche Nutzung der geteilten Daten durch Dritte geschädigt zu werden (Europäische Kommission, 2020a, S. 8). In diesem Zusammenhang entstehen Marktungleichgewichte. In den letzten Jahren wurden in der Datenwirtschaft starke Konzentrationstendenzen beobachtet. Wenige große Unternehmen konnte große Mengen an Daten erzeugen und sammeln sowie ihre Marktmacht durch deren Auswertung weiter stärken (Bundesregierung, 2021a, S. 11). Diese bei wenigen Online-Plattformen (z. B. Amazon) angehäuften Datenmengen führen nicht nur zu Wettbewerbsvorteilen, sondern können teilweise den freien Marktzugang

über die Plattformen hinaus beeinträchtigen. „Daten Vorteile“ ermöglichen großen Marktteilnehmer*innen, Märkte zu dominieren und Regeln zu ihrem Vorteil für den Zugang und die Nutzung von Daten durchzusetzen (Europäische Kommission, 2020a, S. 9 f.). Die Marktmacht von Unternehmen, die besonders große Mengen an Daten kontrollieren und exklusiv nutzen können, beeinträchtigt nicht nur den freien Wettbewerb zwischen Unternehmen. Beeinträchtigt wird außerdem auch die Möglichkeit der Staaten zur Rechtsdurchsetzung (Bundesregierung, 2021a, S. 11). Ungleichgewichte bei der Chancengleichheit zwischen Unternehmen entstehen auch in anderen Kontexten, wie beispielsweise beim Zugriff auf Daten, die im Kontext des Internet der Dinge erzeugt werden. Hier werden Daten in Geräten erzeugt, die von mehreren Herstellern gebaut wurden, und in dem Zeitraum, in dem die Daten erzeugt werden, Dritten gehören. Aus dieser Konstellation entstehen komplexe Konkurrenzen bezüglich des Zugriffs auf Daten. Weitere wichtige Aspekte sind die z. B. die Interoperabilität entsprechender Geräte und Fragen hinsichtlich der Qualität und Authentizität (Fälschungssicherheit) gewonnener Daten (Europäische Kommission, 2020a, S. 10). Daher ist eine andauernde Weiterentwicklung der rechtlichen Rahmenbedingungen und korrespondierender Institutionen erforderlich, um fairen Wettbewerb und offene Märkte als Voraussetzung für „Wohlfahrt für Alle“ und die Beachtung von Grundrechten sicherzustellen. Aus den Zielen, freie Märkte zu garantieren, Verbraucherrechte durchzusetzen, die Einhaltung der Grundrechte zu sichern und zur Erhaltung wie auch Herstellung der Steuerungsfähigkeit demokratisch

legitimierter Regierungen ergeben sich auch Anforderungen an die Dateninfrastruktur. Diese Zusammenhänge werden aktuell unter dem Schlagwort „digitale Souveränität“ thematisiert (Datenstrategie der Bundesregierung, 2021a, S. 11).

4.3.3 Infrastrukturen

Anknüpfend an die Debatten zu International Data Spaces (vgl. Otto et al., 2019; Otto et al., 2016) sowie im Kontext der Plattform Industrie 4.0 betonten die Teilnehmer*innen der Digital-Gipfel-Plattform 2 „Innovative Digitalisierung der Wirtschaft“ (Bundesministerium für Wirtschaft und Energie, 2020, S. 9): Es sollte das strategische Ziel sein, die Innovationskraft unserer Volkswirtschaft zu stärken und „durch eine bewusste Öffnung von Innovationsprozessen“ die Lasten sinnvoll zu verteilen, „ohne dabei einen Know-how-Abfluss befürchten zu müssen.“ Gleichzeitig können Unternehmen auch aus einer Nachfragesperspektive stärker agieren. „Setzt die Wirtschaft intensiver auf Infrastrukturen, die die Resilienz und Digitale Souveränität stärken, [...], trägt sie selbst wiederum zu mehr Nachhaltigkeit bei [...]“ und kann so Produkte und Dienstleistungen gezielter anbieten und diese gleichermaßen nachfragen. In diesem Kontext erschließt sich die Notwendigkeit zur Schaffung datenschutzkonformer Clouddienste aufbauend auf europäischen Werten. Die Ergebnisse einer Umfrage des IT-Verbandes Bitkom (Abbildung 15) zeigen die wachsende Zunahme der Bedenken deutscher Unternehmen gegenüber ausländischen Cloud-Diensten von 2017 bis 2019.

Als Antwort auf diese Diskurse lässt sich der gegenwärtige Aufbau „einer vernetzten, offenen

Dateninfrastruktur auf Basis europäischer Werte“ unter dem Projektnamen Gaia-X verstehen. „Als Wiege eines offenen digitalen Ökosystems, in dem Daten sicher und vertrauensvoll verfügbar gemacht, zusammengeführt und geteilt werden können“ verfolgt Gaia-X primär drei Ziele (Bundesministerium für Wirtschaft und Energie, 2019, S. 2): Neben der Konkretisierung der technischen und wirtschaftlichen Konzeption sollen „ein gemeinsames Ökosystem von Anwendern und Anbietern aus Organisationen der öffentlichen Verwaltung, des Gesundheitswesens, Unternehmen und wissenschaftlichen Einrichtungen“ geschaffen sowie unterstützende Rahmenbedingungen und Strukturen etabliert werden. Die von Deutschland und Frankreich gestartete Gaia-X Cloud-Initiative für die EU soll neue Innovationen sowie neue datengetriebene Dienste und Anwendungen hervorbringen (vgl. ebd.).

Die Einschätzung, inwieweit sich dabei Anspruch und Wirklichkeit – insbesondere bzgl. digitaler Souveränität und „europäischer Werte“ – im Einklang befinden, variiert in verschiedenen Stakeholderkreisen erheblich. Nachdem bekannt wurde, dass sich nicht nur große US-amerikanische und chinesische Cloud-Anbieter (Hyperscaler) wie Amazon Web Services, Google, Microsoft und Alibaba, sondern mit Palantir auch ein höchst umstrittener Anbieter von Überwachungs- und Spionagesoftware mit an Bord befinden, fragte die in Deutschland führende IT-Zeitschrift c't in ihrem Editorial: „Gaia-X: Kentern unausweichlich“. In dem Editorial wird beklagt, das Projekt verkomme „unterm Strich [...] zu einer Farce“ und liefere verbunden mit erheblichen Kosten: „nun wahrscheinlich keine unabhängige europäische Dateninfrastruktur, sondern vielmehr ein willkommenes Feigenblatt für amerikanische

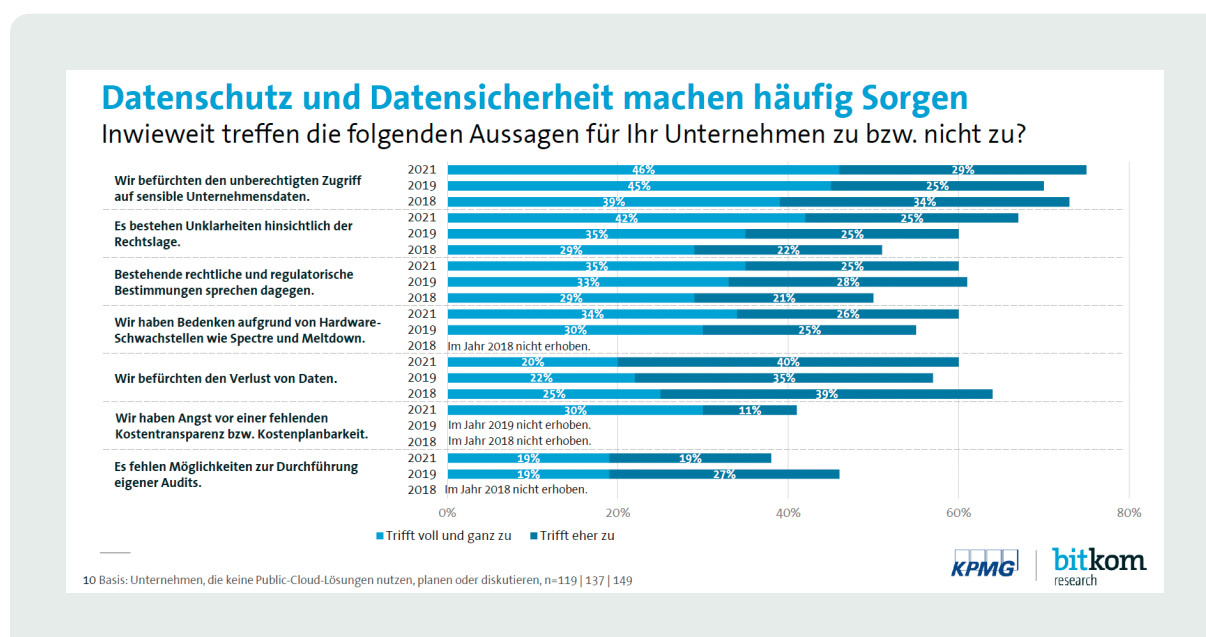


Abbildung 15: Cloud-Risiken (Quelle: Gentemann & Heidkamp, 2021, S. 10; Copyright: Bitkom Research im Auftrag von KPMG)

Datensammler und Gewohnheitsmonopolisten“ (vgl. Schnurer, 2021). Auch die European Cloud Industrial Alliance kritisiert an der Gaia-X Cloud-Initiative, dass Gaia-X „zu sehr von nicht Europäischen (sic!) Anbietern unterlaufen“ sei und „das einzige verbliebene Europäische an Gaia-X [...] inzwischen die Kunden“ seien (vgl. Volkert, 2021).

Trotzdem haben sich seit Mitte September 2020 im Rahmen von Gaia-X Organisationen und Konzerne vor allem aus Deutschland und Frankreich, darunter Atos, BMW, Bosch, SAP und Telekom, am Aufbau einer vertrauenswürdigen, souveränen digitalen Infrastruktur für Europa beteiligt. Mit dem Fokus auf Unabhängigkeit und Datenschutz als Kernthemen, reagiert Gaia-X auf die Sorge vieler deutscher Unternehmen um den unberechtigten Zugriff auf ihre Daten bei der Nutzung der ausländischen Anbieter. Erste Partner*innen haben inzwischen konkrete Gaia-X-kompatible Produkte angekündigt. So wollen beispielsweise die Unternehmen Telekom, Hoster OVHcloud, Plusserver und weitere Anbieter ein Public-Cloud-Angebot für alle Branchen schaffen, in dem Datensouveränität und DSGVO-Konformität eine bedeutende Rolle spielen. Zu den Gewinnern des Förderwettbewerbes zu „innovativen Anwendungen im digitalen Ökosystem Gaia-X“ des Bundeswirtschaftsministeriums gehört das Leuchtturmprojekt „Autowerkstatt 4.0“ zur Schaffung einer branchenspezifischen Datenplattform, an dem unter anderem der eco-Verband, Auto-Teile-Unger und das Deutsche Forschungszentrum für Künstliche Intelligenz beteiligt sind. Ebenfalls ausgezeichnet wurde das Projekt „cooperants“, welches unter anderem von Airbus Defence and Space zur Beschleunigung der Digitalisierung der Luft- und Raumfahrtbranche ins Leben gerufen wurde, und „dataLOFT“, das mithilfe von Gaia-X transparente cloud-basierte Anwendungen in hoch relevanten Gebieten der Gesundheitsversorgung entwickeln möchte (vgl. Wilkens, 2021). Doch tritt Europa mit der Cloud-Initiative Gaia-X gegen eine übermächtige Konkurrenz an (vgl. Wölbart, 2020): Die großen Hyperscaler, die größtenteils aus den USA und China kommen. Dem Marktforschungsunternehmen Synergy Research Group (2020)²¹ zur Folge dominieren drei Hyperscaler (Amazon, Microsoft und Google) etwa 50 % des weltweiten Cloud-Markts.

Befürworter der Gaia-X-Initiative betonen die Offenheit der Architektur und dass es in einer global vernetzten Weltwirtschaft nicht um Ausschluss, sondern maximale Vernetzung und die Möglichkeit der Etablierung klarer Regeln zum Datenaustausch sowie die Schaffung von mehr Auswahlmöglichkeiten gehen muss. Aus der Open Source Business Alliance kommt

zudem, gefördert von SPRIN-D²² sowie dem Bundeswirtschaftsministerium, Unterstützung durch den Sovereign Cloud Stack²³ (SCS) als „Alternative zu proprietär verfügbarer Cloud-Technologie“ (*OSB Alliance entwickelt Sovereign Cloud Stack (SCS) für das GAIA-X Projekt*, o. J.). Die SCS hat das Ziel, „einen vollständig freien Software Stack für eine freie, föderierte Cloud- und Container Infrastruktur zu definieren, umzusetzen und ein Ökosystem von vernetzten Anbietern aufzubauen“. Sie stellt „diese als Unterbau bereit, indem es die Entwicklung von Standards, aber auch Koordination von Software und Anbietern übernimmt“ (ebd.), was Interoperabilität, Transparenz und Unabhängigkeit von Rechten Dritter und damit Schutz vor wirtschaftlicher oder politischer Einflussnahme garantiert (vgl. *Sovereign Cloud Stack*, o. J.). Darüber hinaus befindet sich, angetrieben durch die EU-Mitgliedstaaten, die Parallelinitiative IPCEI-CIS (Important Project of Common European Interests – Cloud Infrastructure and Services) als europäisches Großprojekt für Cloud-Infrastrukturen und -dienste seit Anfang 2021 im Aufbau (vgl. Bundesministerium für Wirtschaft und Klimaschutz, 2021). Diese soll „den Austausch und die Verarbeitung sehr großer Datenmengen in dezentralen Systemen in Echtzeit gewährleisten“ (vgl. Bundesministerium für Wirtschaft und Klimaschutz & Ministère de l'Économie, des Finances et de la Relance, 2021), was zum Ziel hat, unternehmerische Rechenwolken (Edge-Clouds) zunehmend abzubauen und so Anwendungen wie die Industrie 4.0 und das autonome Fahren voranbringen. Mit der IPCEI-CIS sollen damit neue Maßstäbe hinsichtlich Skalierbarkeit, Interoperabilität und Vertrauenswürdigkeit für den Datenaustausch in Europa gesetzt werden (vgl. Krempel, 2021). Als Ergänzung zu Gaia-X wurde im Mai 2021 das durch das Bundesministerium für Bildung und Forschung geförderte Verbundprojekt „FAIR-Data Space“ mit dem Ziel des Aufbaus eines gemeinsamen Cloud-basierten Datenraums für Wissenschaft und Wirtschaft mittels Verknüpfung von Gaia-X und der Nationalen Forschungsdateninfrastruktur (NFDI) gestartet. „Das langfristige Ziel besteht in der Erweiterung der nationalen Architektur zu einer europaweiten (bzw. globalen) Infrastruktur zum Datenaustausch und zur Datenverarbeitung unter Beachtung europäischer Normen und Werte“ (vgl. Bundesministerium für Bildung und Forschung, 2021).

4.3.4 Plattformökonomie

Es besteht Einigkeit in der Einschätzung, dass für die Entwicklung der Wirtschaftsleistung in Deutschland und Europa die Plattform- und Medienökonomie

21 <https://www.srgresearch.com/>

22 <https://www.sprind.org/de/>

23 <https://scs.community/>

von zentraler Bedeutung ist (Falck & Koenen, 2020, S. 1). Mit dem Begriff Plattform sind Datenräume gemeint, in welchen Produkte und Dienste realisiert werden können. Im Zuge der Digitalisierung der Märkte entstanden als neue Gattung digitale Plattformen (vgl. Baums, 2015), welche „sich im Verlauf der vergangenen fünfzehn Jahre zum wichtigsten digitalen Geschäftsmodell entwickelt“ haben (Schmidt, 2018, S. 108). Somit sind digitale Plattformen strategische Geschäftsmodelle mit vielfältigen Märkten, Ökosystemen und Netzwerkeffekten, die in der Regel Marktmechanismen verändern – kurzum Plattformökonomie (vgl. ebd.). Digitale Plattformen nehmen erheblichen Einfluss auf wirtschaftliche oder gesellschaftliche Prozesse und greifen durch die leichte Verfügbarkeit von Gütern oder Informationen in die Dynamik dieser Prozesse ein (Mohabbat Kar et al., 2020, S. 7). Sie „etablieren Standards im technischen als auch im Sinne gesellschaftlicher Konventionen“, indem sie Bedingungen der Kommunikation und Interaktion definieren, wodurch sie das Möglichkeitsspektrum reduzieren und kontrollieren (ebd., S. 5). „Konsumenten wählen den Standard, der von den meisten Benutzer*innen konsumiert wird“, was zu positiven Netzwerkeffekten führt (ebd., S. 5). Mit dem Begriff Netzwerkeffekt wird der Umstand bezeichnet, dass die Attraktivität einer digitalen Plattform mit der Anzahl ihrer Nutzer*innen ansteigt (Nemitz & Pfeffer, 2020, S. 81). „Durch die Etablierung von Allianzen in [...] dynamischen Ökosystemen“ entstehen beispielsweise neue „Werteströme, welche zur Verdrängung der klassischen Anbieter vom Markt“ führen können (vgl. Schmidt, 2019). Darüber hinaus ermöglicht die Sichtbarkeit der Nutzer*innen einer Plattform und die Sichtbarkeit ihrer Vernetzung innerhalb der Plattform den Aufbau neuer Beziehungen. Durch die gezielte Beobachtung sowie Auswertung dieser Vernetzungen und der Kommunikation der Nutzer*innen einer Plattform können Profile von Einzelpersonen und Gruppen erstellt werden, die personalisierte Werbung und Beeinflussung z. B. durch Wahlwerbung ermöglichen. Diese Möglichkeiten zur Beeinflussung der Nutzer*innen einer Plattform ist neben den vordergründigen Services der jeweiligen Plattform ein zentrales und in vielen Fällen das wichtigste Produkt der digitalen Plattform (Nemitz & Pfeffer, 2020, S. 81). Die exklusiven Entscheidungsmöglichkeiten der Plattformbetreiber*innen, wer und in welchem Umfang Zugang zu diesem Nutzen erhält, führt dazu, dass digitale Plattformen zu Konkurrenten für Staaten und Institutionen werden können. Das Kompetenzzentrum für öffentliche IT (Mohabbat Kar et al., 2020, S. 5) führt hier als Beispiel Änderungen auf der Ebene von AGBs an, die das gesellschaftliche Zusammenleben in ähnlicher Weise beeinflussen können wie staatliche Regulierungen. In den verschiedenen Märkten lassen sich durch

negative Netzwerkeffekte Entwicklungen beobachten, die auf Marktversagen hindeuten könnten. Dazu zählen beispielsweise „Monopolisierungstendenzen, Abschottung und Einrichtung von „Datensilos“, oder fehlender Zugang zu Daten und damit Barrieren beim Markteintritt für neue Marktakteure“ (Falck & Koenen, 2020, S. 1).

Digitale Plattformen können durch die Vereinfachung und Strukturierung der Kommunikation zwischen den Nutzer*innen – ob Personen oder Organisationen – helfen, benötigtes Wissen über Wertschöpfungsnetzwerke bzw. Allianzen innerhalb digitaler Ökosysteme aufzubauen (Abbildung 16). Die durch digitale Plattformen ermöglichte Kommunikation und Koordination ermöglicht innovative Leistungsangebote auf Grundlage datengetriebener Geschäftsmodelle über eine flexible, dynamische und automatisierte Interaktion bzw. Kollaboration verschiedener Akteur*innen (vgl. Schössler, 2018). „Der wichtigste Erfolgsfaktor eines datenbasierten Wertschöpfungsnetzwerks ist klare und direkte Nutzversprechen für jeden einzelnen Beteiligten“ (Lernende Systeme – Die Plattform für künstliche Intelligenz, 2020, S. 40), und die Voraussetzung dafür bildet ein „sicherer, weitgehend offener Datenzugang bzw. Datenaustausch aus unterschiedlichsten Quellen über Unternehmens- und Branchengrenzen hinweg“ (ebd., S. 12). Datenbasierte Wertschöpfungsketten sind je nach Branche und Anforderungen der Nutzer*innen unterschiedlich aufgebaut, doch die Herausforderungen, die bei einer wirtschaftlichen Nutzung entstehen, sind vergleichbar: „Neben technischen Fragen der Schnittstellen, Datenformate und Interoperabilität der beteiligten Systeme sind das Vertrauen in die beteiligten Partner und die Sicherheit der Systeme von entscheidender Bedeutung. Die technischen Herausforderungen sind aber prinzipiell gut lösbar“ (ebd., S. 40). Um zukünftig neue Wertschöpfungsmodelle zu eröffnen, braucht es eine klare Datenstrategie, welche Qualität, Relevanz und Verfügbarkeit für die Wertschöpfung benötigter Daten festlegt. Zusätzlich braucht es den Aufbau strategisch ausgerichteter Know-how-Kooperationen mit Kompetenzträger*innen aus dem wirtschaftlichen oder akademischen Bereich und „in Bezug auf die weitere Entwicklung der Medienökonomie [...] die Gewährleistung der technischen Interoperabilität von Daten, der Schaffung von Referenzarchitekturen und gültigen Standards“, sowie „zu regulatorischen Themen wie dem Wettbewerbsrecht“ (ebd., S. 45).

Digitale Plattformen gelten derzeit als die wichtigsten digitalen Geschäftsmodelle (Schmidt, 2018, S. 106). Gemessen an ihrem Börsenwert ist global ein deutliches Ungleichgewicht hinsichtlich der geographischen Verortung der Zentralen der wertvollsten Betreiber*innen digitaler Plattformen erkennbar (Lernende Systeme – Die Plattform für künstliche Intelligenz, 2020). So sitzen 78,5 % der weltweit wertvollsten

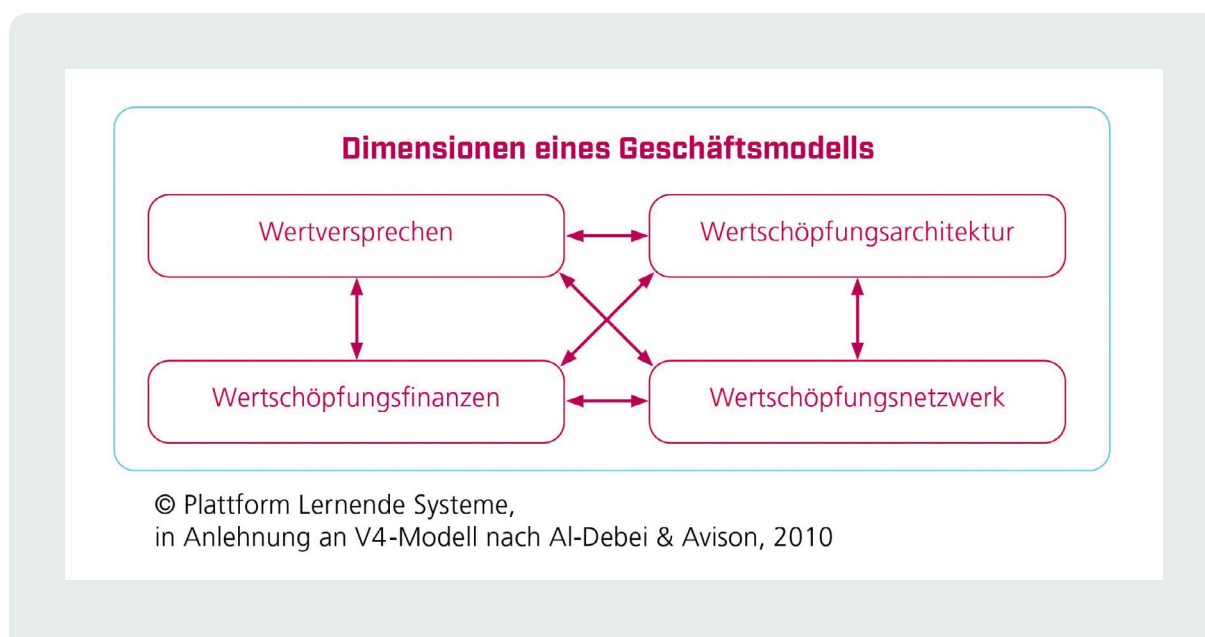


Abbildung 16: Dimensionen eines Geschäftsmodells (Quelle: Lernende Systeme – Die Plattform für künstliche Intelligenz, 2020, S. 7; Copyright: Plattform Lernende Systeme)

Plattformen in den Vereinigten Staaten, 17,3 % im asiatisch-pazifischen Raum, 2,8 % Europa und 1,4 % in Afrika (Abbildung 17). Der Umstand, dass der kumulierte Wert der in Europa und Afrika angesiedelten Plattform-Unternehmen fast gleich ist, kann aus europäischer Perspektive nur erschrecken. Auch bei Investitionen in KI stehen die USA und China Europa deutlich aus (vgl. Schössler, 2018). Deutsche Unternehmen sind sehr verspätet in dieses Geschäftsfeld eingestiegen und haben vor allem auf den B2C-Märkten (Business-to-Consumer) den Anschluss größtenteils verloren. Die Zukunft der deutschen Wirtschaft kann in dem B2B-Markt liegen (vgl. ebd.), denn gerade dort besteht durch die breite Kompetenz im Mittelstand enormes Potenzial für Deutschland (vgl. ebd.). Wenn es Unternehmen aus Deutschland schaffen, Standards und Schnittstellen zu definieren und ihre Fachkompetenzen intelligent zu neuen digitalen Diensten und Ökosystemen zu bündeln, bietet sich eine große Chance, eine entscheidende Rolle im B2B-Geschäft im Kontext von KI zu spielen (Deiningner, 2018, S. 113). Denn hier gilt dasselbe Prinzip der Plattformökonomie wie in den B2C-Märkten. Allerdings zeigen sich deutsche Unternehmen immer noch sehr zögerlich, ihre Produkte auf europäischen Plattformen anzubieten, und leider verzerrt der exportbedingte Erfolg der deutschen Wirtschaft in den vergangenen Jahrzehnten oft den Blick auf die Notwendigkeit der wirtschaftlichen Transformation. Daher nehmen viele Unternehmen die Bedrohung durch Plattform-Modelle, insbesondere im wachsenden B2B-Markt, nicht ernst genug (Schmidt, 2018, S. 106; Deiningner, 2018, S. 113). In den letzten Jahren haben sich Daten als neuer Produktionsfaktor mit enormer und steigender

Bedeutung etabliert. Demgegenüber treten die klassischen Produktionsfaktoren zunehmend in den Hintergrund. In der Plattformökonomie gilt oftmals das Prinzip „The winner takes it all“. Was bedeutet, dass das Plattform-Unternehmen (z. B. Amazon und Facebook) mit den meisten Nutzenden sich auf Dauer durchsetzen und kleinere Plattformbetreiber sowie plattformunabhängig agierende Unternehmen vom Markt verdrängen (Junker, 2018, S. 98). Die Situation von in der Volksrepublik China angesiedelten Plattform-Unternehmen (z. B. Tencent und Alibaba) stellt sich anders dar, weil sie durch die protektionistische Politik der chinesischen Regierung im heimischen Markt vor ausländischen Konkurrenten geschützt werden (vgl. ebd.). Ein großes Problem digitaler Geschäftsmodelle der Plattformökonomie ist der Umstand, dass sie es Unternehmen erleichtern, sich etablierten rechtlichen Normen und Regelungen sozialer Wirtschaftsordnungen zu entziehen. So sehen Plattformbetreibende sich selbst oftmals nur als Vermittelnde, nicht als Dienstleister*innen, und deshalb auch nicht als Verantwortliche für den Ausgleich negativer externer Effekte (Beispiel: Mindestlohn für auf Plattformen vermittelte Haushaltshilfen). Sie umgehen damit etablierte Regularien, denen „traditionelle“ Unternehmen unterliegen. Verschiedene Ereignisse zeugen vom Konfliktpotenzial der Geschäftsmodelle der Plattformriesen mit beispielsweise den Werten der deutschen und europäischen Gesellschafts- und Rechtsordnung (Beispiele: der Facebook-Datenskandal im Jahr 2018, die Wohnraum-Verknappung durch professionelle Vermietung statt Sharing auf Airbnb, die 4,3 Mrd. € Kartellstrafe, die von der Europäischen Union im Jahr 2018 gegen Google aufgrund des Missbrauchs

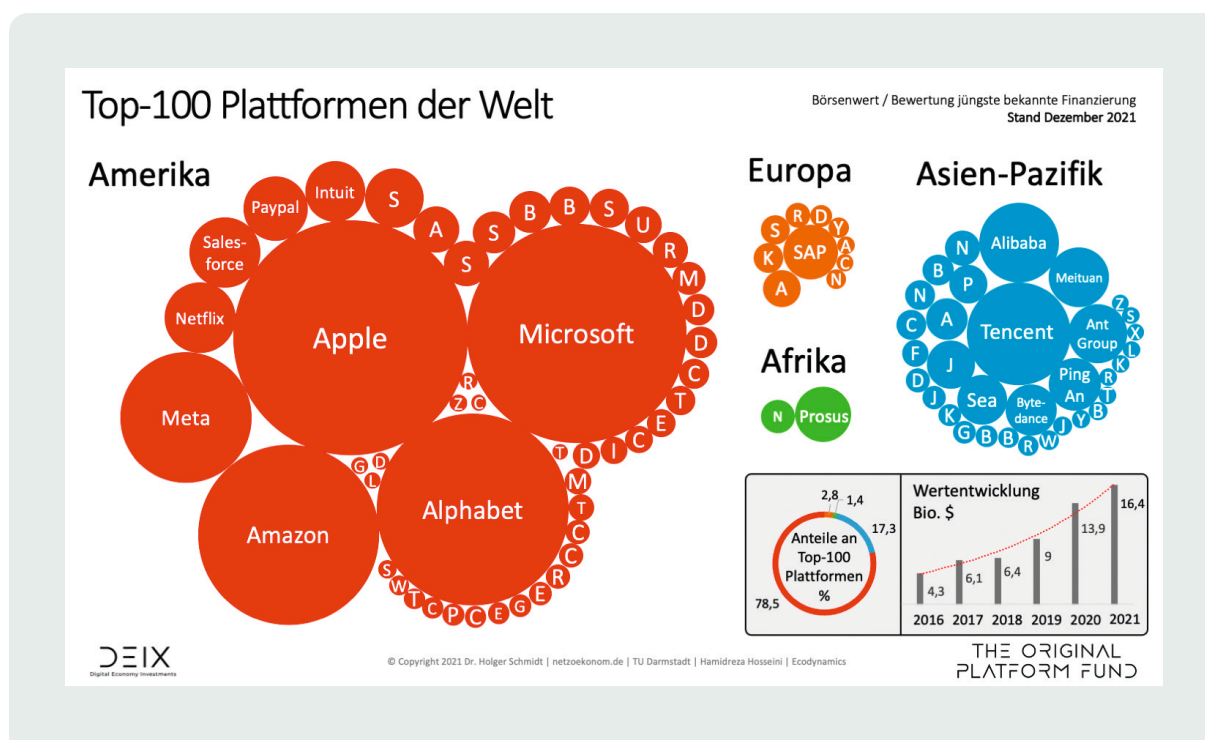


Abbildung 17: Vergleichende Darstellung der 100 wichtigsten Plattformen weltweit nach Börsenwert 2021 und Zugehörigkeit zu Weltregion sowie nach Anteil am Gesamtwert in Prozent; Stand Dezember 2021 (Quelle: Schmidt & Hosseini, 2021, unveröffentlicht; Copyright: Dr. Holger Schmidt)

seiner Marktmacht verhängt wurde). Das derzeit geltende „System der Selbstständigen kann nicht die Lösung bieten“ (ebd., S. 98), wenn die deutsche und europäische Gesellschafts- und Rechtsordnung eingebüßt und KMU vom Markt verdrängt werden (vgl. Busch, 2019). Eine Alternative könnte das Organisationsprinzip der „Platform Cooperativism“ aufzeigen. Diese wird von den beteiligten Unternehmen in genossenschaftlicher Form betrieben (Junker, 2018, S. 98). Kooperative Plattformen könnten beispielsweise Monopolbildung, Marktverdrängung und Lohndumping eindämmen. Genossenschaftlich organisierte digitale Plattformen haben höhere Betriebskosten als Plattformen, die von einem einzelnen Unternehmen betrieben werden. Darüber hinaus wird der Erfolg dieses Ansatzes auch von den rechtlichen Rahmenbedingungen beeinflusst (ebd., S. 98). Der vorzeitig geleakte Entwurf des Data Governance Acts (vgl. Kapitel 3.5) enthielt Vorgaben zur Ermöglichung von Daten-Genossenschaften innerhalb Europas, doch der aktuelle Entwurf sieht keine Möglichkeit zur kollektiven Ausübung von Dateirechten mehr vor (Lanier, 2021).

4.3.5 Defizite und Handlungsbedarf

Zur Umsetzung der Vision „Innovation durch strategische Offenheit“ (Leimüller et al., 2018) in der deutschen und europäischen Wirtschaft bedarf es handelnder Akteur*innen. Dabei sollte es das Ziel sein, Praktiken zu fördern, „die einen erhöhten sozialen und ökonomischen Nutzen versprechen und die die

Wettbewerbsfähigkeit [...] stärken“ (ebd., S. 59). Die Politik setzt hierfür Eckpfeiler, um zu mehr Datenteilen durch Wirtschaftstreibende zu gelangen. Dazu werden in diversen Governance-Prozessen gemeinsame Regeln erarbeitet, um unter der Berücksichtigung europäischer Werte die Hürden beim Datenteilen zu überwinden. Aufgrund der Diversität des europäischen Wirtschaftsraums – angefangen bei unterschiedlichen Wirtschaftsfeldern über geltende rechtliche Regularien bis hin zu Praktiken und Sprachen – erweist sich eine Standardisierung von Prozessen und Daten als besondere Herausforderung. Europäische Initiativen wie Gaia-X zur Schaffung neuer Innovationen sowie neuer datengetriebener Dienste und Anwendungen in Europa sollen vor allem europäische Unternehmen stärken (vgl. Bundesministerium für Wirtschaft und Energie, 2020). Fraglich bleibt jedoch, ob man mit dieser Herangehensweise den Wettbewerbsvorsprung der großen Daten-Monopolisten zu egalisieren vermag. Da datengetriebene Märkte zur Monopolisierung tendieren (vgl. Prüfer, 2020), sind politische Interventionen notwendig, um die Monopolmacht einzelner Unternehmen zu brechen und dadurch einen freien Markt zu garantieren.

Bestehende Dienste können allerdings nicht einfach verboten oder substituiert werden, daher müssen diverse Strategien in Betracht gezogen werden. Die Kommission Wettbewerbsrecht 4.0 des BMWi (2019) skizzierte hierfür drei Ziele: (1) Der Wettbewerb soll als Treiber von Innovationen geschützt werden. (2) Die Wahlfreiheit von Konsument*innen muss

gewährleistet sein – es braucht „offene Märkte mit unverfälschten Wettbewerb“. (3) Die Regelungsstrukturen sollen an den digitalen Wandel angepasst werden (ebd., S. 22 ff.). Um diese Ziele zu erreichen, braucht es eine Stärkung des „Datenzugang[s] und den selbstbestimmten Umgang mit Daten“ (ebd., S. 35 ff.) sowie eine Verbesserung der Datenportabilität und -interoperabilität. Darüber hinaus wird neben der Schaffung „klarer Verhaltensregeln für marktbeherrschende Plattformen“ (ebd., S. 49 ff.) auch eine bessere Rechtsdurchsetzung benötigt. Diesbezüglich gibt es weiterhin deutliche Unklarheiten bei vielen Rechtsfragen rund um das Teilen und Nachnutzen von Daten. Zudem zieht die Kommission Wettbewerbsrecht 4.0 (ebd.) ebenfalls die Möglichkeit der Einrichtung von Datentreuhänder*innen-Modellen in Betracht (vgl. Kapitel 4.3.2). Auch die Friedrich-Ebert-Stiftung formuliert in einer Veröffentlichung unter dem Titel „Mehr Fairness für KMU auf digitalen Märkten“ von Christoph Busch (2019) ähnliche rechtspolitische Handlungsoptionen. Zudem fordert er ein klares Verbot von Exklusivbindungen und damit einhergehend die Verpflichtung marktmächtiger Plattformen zu Neutralität und zum Datenteilen (ebd.), um so vor allem auch KMU zukünftig eine wirtschaftliche Perspektive in den digitalen Märkten zu ermöglichen. Möglichkeiten zur Realisierung einer „Datenteilpflicht“ werden in einer gleichnamigen Veröffentlichung der Friedrich-Ebert-Stiftung von Jens Prüfer (2020) erörtert. Der Fokus des Papiers liegt auf der Analyse der Dysfunktionalitäten, die durch Monopolbildung in Datenmärkten erzeugt werden, und auf der Entwicklung von Regulierungsansätzen, mit welchen gezielt diesen Dysfunktionalitäten entgegengewirkt werden kann. Diese Fokussierung ist eine offensichtliche Voraussetzung für eine verfassungskonforme Gestaltung von Eingriffen in die Datenhoheit privater Akteur*innen. Prüfer denkt beispielsweise daran, Verpflichtungen zum Teilen von Daten an die Größe von Unternehmen bzw. deren Datensammlungen zu binden (vgl. Prüfer, 2020, S. 11 f.). Ein weiterer Ansatz sind Verpflichtungen zu Interoperabilität, so fordert er, soziale Netzwerke (z. B. Facebook) zur Verwendung offener Standards zu zwingen (d. h. Interoperabilität und Portabilität). Auch diese Form der Intervention will er mit der Größe von Unternehmen bzw. deren Marktmacht verknüpfen (vgl. Gegenhuber, 2020, S. 9 f.). Ein stärkerer Wettbewerb zwischen den sozialen Plattformen würde Innovationen fördern und neue Angebote für Konsument*innen schaffen, daher sollte gelten: „Plattformen, die soziale Netzwerke betreiben, dürfen kein anderes soziales Netzwerk kaufen“ (vgl. ebd.). Der Stifterverband (vgl. Leimüller et al., 2018) setzt auf die Entwicklung eines nationalen politischen Rahmens für strategische Offenheit, einer Erweiterung des nationalen Forschungs- und Innovationsmonitorings sowie

auf die Förderung einer Einbindung unüblicher Akteur*innen in Forschungs- und Innovationsprojekte. Zudem sollte Wissenstransfer durch strategische Offenheit und Kooperationszentren in beide Richtungen gedacht werden: „Transfereinrichtungen sollten nicht nur wissenschaftliche Erkenntnisse für die Nutzung in Wirtschaft und Gesellschaft, sondern auch Wissen aus Wirtschaft und Gesellschaft für die Generierung von Forschungsfragen in der Wissenschaft bündeln“ (ebd., S. 7). Risiken, wie etwa in Bezug auf geistiges Eigentum, müssten gleichzeitig klar benannt werden und durch diverse Anreizstrukturen sollten Akteur*innen zur Beteiligung motiviert werden (vgl. ebd.).

Nach Einschätzung des Kompetenzzentrums Öffentliche IT (Mohabbat Kar et al., 2020, S. 5) gehen bisher realisierte Regulierungen digitaler Plattformen nicht weit genug und erreichen deshalb die intendierte Wirkung nur teilweise. Das Kompetenzzentrum Öffentliche IT (ebd., S. 5) fordert den Staat auf, der „proaktiven [...] Nutzung von Plattformprinzipien mehr Beachtung zu schenken“, beispielsweise beim Datenaustausch im öffentlichen Sektor. Darüber hinaus sollten „die durch Plattformen geschaffenen Organisationspotenziale für alternative, gemeinwohldienliche Ziele [...]“ eingesetzt werden (ebd.). Auch das Hightech-Forum (2020b, S. 2 f.) fordert ein Neudenken der „Rolle und Verantwortung des Staates in der Datenökonomie“, wobei „neue staatliche Strategien für eine Leitbildorientierung, für Regulierung und für Infrastrukturinvestitionen“ entwickelt werden müssen (ebd., 2. f.). Zudem braucht es zur Gestaltung der digitalen Transformation der Wirtschaft eine gezielte Förderung von digitalen Geschäftsmodellen „unter Berücksichtigung von immateriellen Produktionsfaktoren und Nachhaltigkeitsaspekten“ (ebd., S. 3). Allen voran sei die Evaluation der Datenschutz-Grundverordnung und verwandter Gesetzgebung zum Schutz der Persönlichkeitsrechte überfällig. Ihre Anwendung bestimme oftmals die Hürden beim Datenteilen. Hier müsse die Balance zwischen individuellen Rechten an Daten und den berechtigten Interessen Dritter noch gefunden werden.

4.3.6 Zusammenfassung der Kernpunkte

Mit dem Ziel vor Augen, national und EU-weit zu einer „bereichsübergreifenden, signifikanten Erhöhung innovativer und verantwortungsvoller Datenbereitstellung und -nutzung“ (Bundesregierung, 2021a, S. 6 ff.) zu gelangen, braucht es einen gezielten Abbau von Defiziten und Hürden:

Zu der Reihe von Problemen, welche die EU daran hindert, ihr Potential in der Datenwirtschaft auszuschöpfen, zählen im *Public Sector* Defizite bei einer bedarfsgerechten Zugänglichkeit von Daten. In der *Wirtschaft* bzw. bei vielen Unternehmen wird der

Analyse der Potentiale zur besseren Nutzung eigener Daten, sei es für unternehmensinterne Wertschöpfung oder für das Teilen mit Dritten oder von externen Daten vielfach zu wenig Aufmerksamkeit geschenkt. Von der *Wissenschaft* werden Daten in großem und steigenden Umfang öffentlich zugänglich gemacht. Der Fokus liegt hier jedoch meist auf der Ermöglichung der Nachnutzung innerhalb der Wissenschaft. Die Akteur*innen in allen drei genannten Bereichen können demnach mehr tun, um den immer wieder beschworenen Kulturwandel beim Teilen und Nachnutzen von Daten zu befördern. Der Handlungsspielraum einzelner Akteur*innen darf jedoch nicht überschätzt werden. Sollen z. B. Daten zur Förderung ihrer Nachnutzung aufwändiger kuratiert werden, bedarf es hierfür Fachkräften und den finanziellen Mitteln für deren Bezahlung. Die Marktmacht einer kleinen Gruppe sehr großer Unternehmen erschwert den Markteintritt neuer Unternehmen. Besonders im Kontext des Internets der Dinge sind viele Fragen hinsichtlich des Bestehens von Zugriffs- und Verwertungsrechten noch ungeklärt. Daraus resultierende Unsicherheiten wirken sich investitionshemmend aus. Etablierte Ansätze zur Regulierung von Märkten wie der Schutz von Verbraucher- oder Persönlichkeitsrechten oder die Besteuerung funktionieren vielfach nur noch eingeschränkt, weil die Digitalisierung die Möglichkeit für Unternehmen, sich Regulierungen zu entziehen erhöht.

Trotz des regelmäßigen Postulierens des Mantras vom wirtschaftlichen Potential, das durch Datenteilen entfesselt werden könne, wird dieses durch Unternehmen vielfach weder im Verhältnis B2B, noch im Verhältnis B2G realisiert. Letztlich fehlen vielen Unternehmen ökonomische Anreize zum Teilen von Daten. Um zukünftig neue Wertschöpfungsmodelle zu eröffnen, braucht es Strategien, die die Entwicklung derzeit fehlender Nachnutzungsszenarien bzw. Geschäftsmodelle unterstützen. Zu diesen Strategien zählen auch Maßnahmen zum Aufbau datennutzungsbezogener Know-how-Kooperationen.

Immer wieder werden auch Verunsicherungen angesprochen, die sich aus Kontrollverlusten, welche mit dem Teilen von Daten verbunden sein können, ergeben. Diese Problematik, die Teil der Diskussion um digitale Souveränität ist, verweist primär auf eine Aufgabenstellung für die Gesetzgeber, auch wenn am Ende die Kooperation aller betroffenen Akteur*innen notwendig ist. Zusätzlich zu gesetzlichen Regelungen müssen auch technische und organisatorische Voraussetzungen für vertrauenswürdige Datenökosysteme entwickelt werden. Dafür bieten europäische Innovationen wie die komplementären Plattformen Gaia-X und IPCEI-CIS erste Lösungsansätze. Sie setzen neue Maßstäbe hinsichtlich Skalierbarkeit, Interoperabilität und Vertrauenswürdigkeit auf europäischer Ebene und ermöglichen durch die Offenheit ihrer Architektur

eine maximale Vernetzung. Darüber hinaus eröffnen sie die Möglichkeit der Etablierung bzw. technischen Unterfütterung klarer Regeln zum Datenaustausch.

4.3.7 Exkurs: Kritik dominanter Narrative zum Verhältnis von Markt, Staat und Innovationen der Digitalwirtschaft

Mariana Mazzucato (2019, S. 352) – eine mit Blick auf europäische wie nationale F&I-Politik und deren Orientierung an „Missionen“ (European Commission, Directorate General for Research and Innovation, 2018; Bundesministerium für Bildung und Forschung, 2018; Mazzucato, 2021) aktuell sehr einflussreiche Ökonomin – vertritt, ebenso wie die französische, hierzulande erst seit wenigen Jahren und nur in einigen Disziplinen rezipierte „Economie des conventions“ (Salais, 1989; Diaz-Bone & Salais, 2012; Diaz-Bone, 2015) die Perspektive, Märkte als gesellschaftlich geformte „Ergebnisse“ von Prozessen zahlreicher Akteur*innen zu betrachten. Sie sind somit als Resultate sozialer Praktiken historisch wie kontextbezogen variabel und politisch gestaltbar. Dies ermöglicht einen gegenüber etablierten ökonomischen Narrativen alternativen Blick – insbesondere auf die Herausbildung der Digitalwirtschaft, wo zumeist die Rolle des Staats ausgeblendet bleibt (Mazzucato, 2011) und Machtasymmetrien fast schon naturgegeben und unausweichlich erscheinen.

Im Ergebnis der durch intensivierte Menge und Verarbeitung von Daten bei gleichzeitiger Markt- und Machtkonzentration gekennzeichneten letzten zwei Dekaden der Digitalisierung hat sich eine neue Struktur „proprietärer Märkte“ herausgebildet: „Die Leitunternehmen des kommerziellen Internets sind weniger Produzenten, die auf Märkten agieren als Märkte, auf denen Produzenten agieren. Daraus speist sich zu großen Teilen ihre Macht. Als Marktbesitzer wachen sie in wachsenden Bereichen der Ökonomie darüber, wer Zugang zum Markt erhält und unter welchen Bedingungen dies geschieht (Zugangskontrolle). Ihr Profitmodell liegt in der Erwirtschaftung ökonomischer Renten aus eben diesem Privileg. Die Kombination proprietärer Infrastruktur mit der Skalierbarkeit digitaler Güter ermöglicht den theoretisch infiniten Ausbau der Marktfunktion zu marginalen Grenzkosten. Profite aus Marktbesitz werden so tendenziell zu leistungslosen Einkommen. Die dominanten Investitionsfelder der jüngsten Vergangenheit und Gegenwart – Cloud, KI und Fintech – erweitern und vervollständigen die Marktfunktion“ (Staab, 2019, S. 223).

Mazzucato (2019, S. 343 ff.) betont diesbezüglich, dass im Sinne von „Adam Smith' Definition des ‚freien Markts‘ als eines Markts frei von Renten“ eine wesentliche Aufgabe der modernen Wirtschaft darin bestehe, „das Maß an Renten einzuschränken, die aus jedem

nicht kollektiven Ansatz der Wohlstandsschaffung entstehen“. Durch von dieser Denkweise geleitete ökonomische wie politische Praktiken könnten zahlreiche wichtige, weder komplett private noch komplett staatliche Einrichtungen erheblich profitieren, denn „[i]m neuen Diskurs [...] wäre es weithin akzeptiert, dass die beiden Sektoren und alle Institutionen dazwischen einander in Verfolgung eines gemeinsamen Ziels volkswirtschaftlicher Wertschöpfung nähren und sich gegenseitig verstärken“ (ebd., S. 343 ff.). Insofern zeigt sich als Ausgangspunkt der Landscape-Analyse für den Untersuchungsbereich Wirtschaft in der aktuellen Literatur einerseits ein starkes Momentum für einen Umbruch der bislang vorwiegend einseitig durch große digitale Konzerne geprägten Datenmärkte. Andererseits bestehen von Seiten eben dieser durch starke Lobbyinteressen flankierte Beharrungskräfte, so dass die Vision eines auf Reziprozität beruhenden Ökosystems statt eines parasitären „Räuber-Beute-Systems“ und somit neuer funktionierender Partnerschaften zwischen Wirtschaft, Staat und Zivilgesellschaft keine geringe Herausforderung darstellt: „Das ist keine leichte Aufgabe, und man kann noch nicht einmal mit ihr beginnen, ohne zunächst alle Akteur*innen neu, das heißt im Zentrum eines kollektiven Wertschöpfungsprozesses, zu positionieren. Kurzum, nur indem der Staat anders und in großem Stil denkt, kann er Wert und Hoffnung schaffen“ (ebd., S. 352).

Ausgehend von der in Zeiten großer gesellschaftlicher Herausforderungen drängenden Frage, welche Art öffentlich-privater Partnerschaften einer Gesellschaft die gewünschten Resultate beschieren, plädiert Mazzucato dafür, diese ebenso wie bisherige Auffassungen staatlicher Politik neu zu denken. Die Notwendigkeit eines neuen Diskurses, der bestehende Spannungsfelder produktiv unterläuft, wird besonders mit Blick auf die Überschneidung der Sektoren Wissenschaft und Wirtschaft deutlich, denn „[d]ie profitabelsten Anwendungen von Big-Data-Methodologie, KI und Quantum liegen gerade nicht im Bereich der fragmentierten und durch menschliches Verhalten begrenzten und oft unpräzisen persönlichen Daten, sondern in den Forschungsdaten und in den Wirtschaftsdaten, die nicht persönliche Daten sind, also den Marktdaten, den Daten des Maschinenbetriebs, der Physik und Technik, der Biologie und den Erdwissenschaften“ (Nemitz & Pfeffer, 2020, S. 68). Bislang beruhen dominante Geschäftsmodelle auf Netzwerk- und Lock-In-Effekten in geschlossenen Systemen: „Man beteiligt sich am Netzwerk, weil alle anderen schon dort sind. Und man verlässt es nicht mehr, weil das zu viel Aufwand bedeuten würde“ (ebd., S. 71). Die FAIR-Prinzipien allein werden zwar nicht den nötigen Kulturwandel zu einem möglichst offenen datenbasierten Innovationsökosystem (Directorate-General for Research and Innovation (European Commission) et al., 2020;

Hightech-Forum, 2020) einleiten, sind aber dafür eine notwendige Bedingung. Besonders aus wissenschaftlicher Sicht bestehen mit Blick auf Datenqualität zudem unterschiedliche Ambitionsniveaus, denn „[m]ehr noch als gute Algorithmik wird die Datenkultur dasjenige sein, was digitale Wissenschaft von anderen – zum Beispiel kommerziellen und journalistischen – Formen digitaler Datenverwendung unterscheidet. Die international anerkannten FAIR-Prinzipien, denen zufolge Daten findbar, zugänglich, interoperabel und nachnutzbar sein müssen, sind in dieser Richtung ein wichtiger, aber kein zureichender Orientierungspunkt. Streng genommen sorgt FAIR für Prozessierbarkeit und damit für intensive Datennutzung, überlässt die Qualitätsfrage aber der Nutzerin und dem Nutzer“ (Gehring & Lange, 2020). Diese lässt sich wiederum in unterschiedlichen Kontexten und Sektoren verschieden beantworten, was entsprechend mit Blick auf Interoperabilität zu berücksichtigen ist. Insofern gilt es sowohl unterschiedlichen Datenkulturen gerecht zu werden, als auch, wenn nötig, diese zu transformieren, um gesamtgesellschaftlich dysfunktionale Tendenzen zu revidieren. Aktuell werden mit dem Aufbau neuer Daten-Ökosysteme Pfadentscheidungen getroffen, die entweder den Status quo zementieren oder einen sektorübergreifenden Kulturwandel befördern können. Der Wirtschaft und der Frage danach, wie künftige digitale Märkte durch einen „gestaltenden Staat“ (Wissenschaftlicher Beirat der Bundesregierung Globale Umweltveränderungen, 2011) gestaltet werden, kommt hierbei eine Schlüsselrolle zu – nicht zuletzt, da sie im Sinne einer fortgeschrittenen „Ökonomisierung des Sozialen“ (Bröckling, 2007) im Gewand der Plattformökonomie verstärkt in andere gesellschaftliche Teilbereiche hinein wirkt und dabei in teils erheblicher Spannung zu dortigen Werten und Funktionslogiken stehen kann. Aus (wirtschafts-)ethischer Sicht geht es dabei um nichts geringeres als ein über das monetäre hinaus erweitertes Werteverständnis, bei dem das Gemeinwohl im Zentrum steht und „nicht die Rendite oder irgendein Automatisierungsindex“ (Spiekermann, 2019, S. 65). Inwieweit dieses unter den gegenwärtigen Rahmenbedingungen jedoch Resonanz finden kann, ist allerdings fraglich, wie die Expert*innen-interviews (Kapitel 5) gezeigt haben.

4.4 Qualifizierung

4.4.1 Quellengrundlagen

Im Untersuchungsbereich Qualifizierung gibt es politische Strategiedokumente, Fachartikel und zahlreiche auf Projekten und Untersuchungen beruhende Publikationen auf nationaler und EU-Ebene, die sich intensiv mit Daten als Qualifizierungsthema

auseinandersetzen. Eine Analyse der gesamten Veröffentlichungen aller möglichen Aspekte zu diesem Thema liegt außerhalb des Rahmens des Projekts. Daher liefert dieses Kapitel einen Überblick über Rahmenbedingungen, Bedarfe, Inhalte und Vermittlungsformen sowie den Aufbau von Datenkompetenzen und den sich entwickelnden Berufen in den unterschiedlichen Bereichen eines zunehmend digitalisierten Alltags. Wo auf der Basis der Quellen möglich, werden Unterschiede in den Untersuchungsbereichen *Wissenschaft*, *Public Sector* und *Wirtschaft* berücksichtigt und die einzelnen Themen anhand von Beispielen dargestellt. Wie in den anderen Untersuchungsbereichen werden aktuelle Veröffentlichungen verwendet, die einen übergreifenden oder exemplarischen Blick im Bereich Qualifizierung und Daten liefern.

4.4.2 Rahmenbedingungen und Vorgaben

Die zunehmende Digitalisierung der Gesellschaft führt zu einem steigenden Bedarf an Qualifizierung gerade auch im Umgang mit digitalen Daten. So weist die Bundesregierung beispielsweise auf einen stetig steigenden Bedarf an Datenkompetenz bei Beschäftigten in der Wirtschaft hin (Bundesregierung, 2021a, S. 45). Laut European Data Market Monitoring Tool der EU lag der Fachkräftemangel im Bereich der

Datenexpert*innen („Data Professional Skills Gap“) im Jahr 2019 in Deutschland bei rund 114.000 unbesetzten Stellen und für das Jahr 2025 wird die Zahl abhängig von der Konjunktorentwicklung auf 126.000 („Baseline Scenario“), 141.000 („Challenge Scenario“) bis 192.000 („High Growth Scenario“) steigen (IDC Italia srl & The Lisbon Council, 2020, S. 134, Table 47). „Für die EU wird die Zahl aller Datenexpert*innen im Jahr 2025 auf 8,4 Millionen geschätzt, was einer Zunahme von 1,8 Millionen im Zeitraum von 2020-2025 entspricht. Dennoch wird erwartet, dass 484.000 Stellen nicht besetzt werden können, was 5,7% des Bedarfs entspricht, da die Nachfrage schneller als das Angebot wächst“ (European Commission. Directorate General for Communications Networks, Content and Technology. et al., 2020, S. 9). Vergleiche hierzu auch die folgende Abbildung 18, in der die geschätzte Zahl der Datenexpert*innen für 2025 mit 8,5 Millionen („Challenge Scenario“) etwas höher liegt.

In die gleiche Richtung geht eine Meldung von Heise-Online im April 2021, „dass Data Scientists dringend gesucht werden“ und „laut einer Studie Serviceanbieter zunehmend eigene Mitarbeiter ausbilden, um dem Fachkräftemangel zu begegnen“ (vgl. Ungerer, 2021).

Auf der (wissenschafts-)politischen Ebene wird versucht, dem Mangel durch (gesetzliche) Vorgaben und

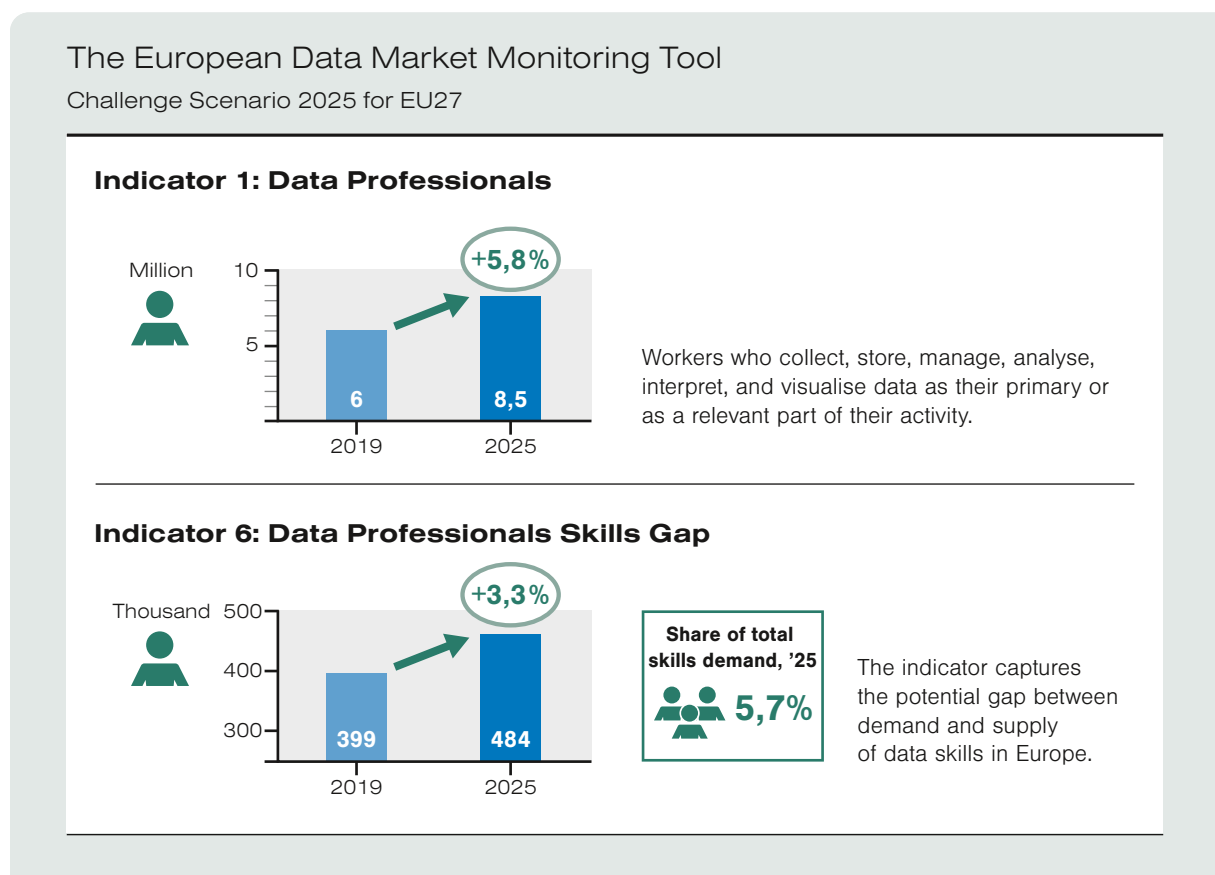


Abbildung 18: Data Professionals Skills Gap (Quelle: European Commission. Directorate General for Communications Networks, Content and Technology. et al., 2020, S. 17; Lizenz: CC-BY 4.0 international, <https://creativecommons.org/licenses/by/4.0>)

finanzielle Förderung entgegen zu wirken. Hinweise, wie erkannte Qualifizierungsdefizite im Bereich Datenkompetenz behoben werden sollen, finden sich vor allem in den nationalen und europäischen Datenstrategien. So erklärt die Bundesregierung 2021 in ihrem Strategiepapier „Datenkompetenz erhöhen und eine neue Medienkultur in Deutschland etablieren“ zu einem von vier Handlungsfeldern (Bundesregierung, 2021a, S. 8, 41f.). Sie beschreibt den aktuellen Stand der Qualifizierung im Hinblick auf Datenkompetenz und formuliert als Handlungsbedarf die Erhöhung der Bildung und den Aufbau von digitaler Kompetenz im öffentlichen Bereich sowie in der Wirtschaft und ordnet den Datenaspekt in den größeren Kontext ihrer Digitalstrategie ein (Bundesregierung, 2021b, S. 10 f.). Beispielsweise sollen neben der „Schaffung und Durchsetzung von verlässlichen, verständlichen und transparenten Regelungen und sicheren Infrastrukturen“ alle Bürger*innen „Grundkenntnisse erwerben, um informiert und souverän“ mit ihren eigenen Daten umgehen zu können (vgl. Bundesregierung, 2021a, S. 41). „Weitergehende Kenntnisse“ sollen in „Berufsausbildungen, Studiengängen und fachspezifischen Weiterbildungen“ vermittelt werden und Expert*innen als Fachkräfte aus- und fortgebildet werden (vgl. ebd.). Für die Bereiche Bildung und Ausbildung werden als konkrete Ziele formuliert, dass „Datenkompetenz in den von den Ländern verantworteten Lehrplänen verankert und altersgerecht aufbereitet“ (ebd., S. 44) wird. Darüber hinaus soll sichergestellt werden, dass „allen, die eine Ausbildung oder ein Studium in Deutschland abgeschlossen haben, ein Mindeststandard an Datenkompetenz vermittelt wurde“ (ebd.). Als eine Maßnahme für den Untersuchungsbereich *Wirtschaft* sollen z. B. für eine „effektive Bildungsberatung und -begleitung [...] 200 betriebliche Vertrauensleute in bis zu 100 Betrieben zu innerbetrieblichen Weiterbildungsmentorinnen und -mentoren“ qualifiziert werden, um „dadurch eine arbeitnehmersnahe innerbetriebliche Weiterbildungsberatung aufzubauen“ (vgl. ebd., S. 46). Für den Untersuchungsbereich *Public Sector* sind die „Weiterbildungsangebote im Bereich Datenkompetenz, insbesondere Datenauswertung und deren Visualisierung“, für das „bestehende Verwaltungspersonal“ noch in der Entwicklung (vgl. ebd., S. 59). Jedoch bot die „Bundesakademie für öffentliche Verwaltung (BAkÖV)“ 2021 u. a. eine Fortbildungsmaßnahme mit dem Titel „Data Literacy – Datenkompetenz in der öffentlichen Verwaltung“ an (ebd.). Die EU plant auf europäischer Ebene in ihrer Datenstrategie nicht nur, für „geeignete Rechtsvorschriften“ und eine „angemessene Führung zu sorgen, um die Verfügbarkeit von Daten sicherzustellen“, sondern auch in „Standards, Instrumente und Infrastrukturen

sowie in Kompetenzen für den Umgang mit Daten“ zu investieren (vgl. Europäische Kommission, 2020a, S. 5). Es wird betont, dass ein europäischer Datenraum nur funktionieren kann, wenn neben einer Förderung der Infrastruktur auch in digitale Kompetenzen wie Datenkompetenz investiert wird (ebd., S. 6).

Auf politischer Ebene ist der im Rahmen der „Digitalen Agenda der Bundesregierung 2014 – 2017“ initiierte Rat für Informationsinfrastrukturen (RfII), der als wissenschaftspolitisches Beratungsgremium Bund und Länder in Fragen der Weiterentwicklung wissenschaftlicher Informationsinfrastrukturen berät, mit seinen Positionen und Empfehlungen als wichtiger Akteur zu nennen. Für den Bereich Qualifizierung finden sich bereits wesentliche Empfehlungen und Erkenntnisse in „Digitale Kompetenzen – dringend gesucht“ (RfII – Rat für Informationsinfrastrukturen, 2019b). So etwa Feststellungen zu den für Qualifizierung relevanten Rahmenbedingungen an der Schnittstelle von Forschung und Lehre. Demnach ist der Untersuchungsbereich *Wissenschaft* an den Hochschulen durch „eine enge Verzahnung der Forschung mit der Lehre bzw. der Qualifizierung von wissenschaftlichem Nachwuchs“ (ebd., S. 2) gekennzeichnet, jedoch erfolgt der Erwerb von digitalen Kompetenzen aufgrund von befristeten Verträgen und Projektarbeit eher unsystematisch (vgl. ebd., S. 4). Die Digitalisierung des Forschungsprozesses lässt die Arbeitsteilung zwischen Forschung und Infrastruktur verschwimmen und „Aufgaben in Forschung und Lehre rücken im Zuge digitaler Forschungsprozesse näher heran an solche Aufgaben, die klassisch in zentralen Infrastruktureinrichtungen angesiedelt sind“ (ebd., S. 10). Dies hat einen erhöhten Qualifizierungsbedarf auch im Bereich Infrastruktur zur Folge.

Es lässt sich insgesamt festhalten, dass der Bedarf an Fort- und Weiterbildungsangeboten für das vorhandene Personal an wissenschaftlichen Einrichtungen hoch ist und einem geringen Angebot an einschlägigen wissenschaftlichen Weiterbildungen gegenübersteht.

Generell kann für alle die Vermittlung von Datenkompetenz an allen Bildungseinrichtungen von Schule über Ausbildungsstätten aller Art bis hin zu Hochschulen erfolgen. Darüber hinaus gibt es auch zahlreiche Weiterqualifizierungsangebote von kommerziellen Anbieter*innen. Zu unterscheiden ist bei Vermittlung von Datenkompetenz zwischen grundlegenden und fortgeschrittenen Kompetenzen. Je nach Anbieter*in gibt es ein Abschlusszeugnis, ein Zertifikat oder eine Teilnahmebestätigung. Ein Gesamtüberblick über alle Angebote zu bekommen ist derzeit kaum möglich. Forschungsdaten.org²⁴ listet im Januar 2022 beispielsweise nur zehn Hochschulen in Deutschland auf, die Angebote haben, bei denen der adäquate Umgang

24 <https://www.forschungsdaten.org>

mit Forschungsdaten Teil der Hochschulausbildung ist (siehe: *Ausbildung und Qualifikation – Forschungsdaten.org*, o. J.). Dagegen liefert eine Suche nach entsprechenden Studiengängen beim „Hochschulkompass der Hochschulrektorenkonferenz“ zum gleichen Zeitpunkt je nach Suchbegriff unterschiedliche Ergebnisse: für „Daten“ erhält man 299, für „Datenmanagement“ 33 und für „Forschungsdaten“ 1 Treffer (siehe: *Studieren in Deutschland – Studium in Deutschland – Hochschulkompass*, o. J.). Bei genauerer Betrachtung der Suchergebnisse zeigt sich, dass bei vielen gefundenen Studienangeboten das Thema Daten nur ganz am Rande eine Rolle spielt und sich beispielsweise bekannte Studienangebote mit Datenschwerpunkt gar nicht unter den Treffern finden. Damit lässt sich keinerlei belastbare Aussage über das tatsächliche Angebot machen, da die Datenbank dieses (noch) nicht adäquat abbildet. Die Ergebnisse geben vielmehr einen Hinweis darauf, dass sich aktuell noch neue Studiengänge mit dem Schwerpunkt Daten(-kompetenz) herausbilden, sei es durch komplett neue Angebote oder durch veränderte Inhalte von bestehenden Studienangeboten, die sich noch nicht in ein Raster einordnen lassen. Ein sehr ähnliches Bild ergibt eine Suche nach Weiterbildungsangeboten. Auch hier wäre eine Übersicht, die eine Vergleichbarkeit ermöglichen würde, wünschenswert.

4.4.3 Infrastruktur

Als relevante Infrastrukturen sind in diesem Bereich u. a. die NFDI und die EOSC anzusehen. So beobachtet der RfII „starke Engpässe und fehlende Qualifizierungen für Fachkräfte an der Schnittstelle Wissenschaft/Information“ und sieht die „Forschungsinfrastrukturen hier in einer besonderen Rolle“ (RfII – Rat für Informationsinfrastrukturen, 2020b, S. 4). Zudem fordert er, dass sich „insbesondere die Konsortien der NFDI [...] in dieser Hinsicht auch als Treiber für die Heranbildung von Fachkräften begreifen [sollten]“ (ebd.). Beim Aufbau der NFDI wird das Thema Qualifizierung Forschungsdatenmanagement (FDM) als ein Querschnittsthema diskutiert und unter Handlungsfeldern für eine NFDI-übergreifende Infrastrukturentwicklung als ein Bereich aufgefasst, der vorrangig berücksichtigt werden sollte. Neben „Training, Ausbildung, Beratung, Curricula, data literacy“ wird explizit auch die Entwicklung von „neuen Berufsfeldern“ als Thema genannt (vgl. Bierwirth et al., 2020, S. 3). Statt für die Querschnittsthemen ein eigenes Konsortium einzurichten, soll jetzt im Rahmen von sogenannten Sektionen und damit disziplinübergreifend gearbeitet werden. Laut Satzung der NFDI entsenden die Konsortien

hierfür „Mitarbeitende, um dort konsortialübergreifende Themen zu behandeln“ (Kraft et al., 2021, S. 8). Aktuell sind die Sektionen noch im Aufbau.

Auf europäischer Ebene wird als Infrastruktur die EOSC erstellt. Im Zuge des Aufbaus wurde im EOSCpilot²⁵ u. a. FAIR4S²⁶ als „skills and capability framework“ („Modell/Rahmenwerk für Kompetenzen und Fertigkeiten“) entwickelt. In FAIR4S werden unterschiedliche Rollen definiert und ihnen die erforderlichen Kompetenzen und Fähigkeiten abhängig vom Level der jeweiligen Expertise zugeordnet, die sie dazu befähigen, *Data Stewardship* („Datenverantwortung“) in allen Phasen des Data Life Cycles optimal zu übernehmen.²⁷ Das *Framework* stellt so einen wichtigen Meilenstein hinsichtlich der exakten Beschreibung der Aufgaben für den Umgang mit Daten im Forschungsprozess und für die Entwicklung von neuen Berufsbildern in diesem Bereich dar.

4.4.4 Qualifizierungsbedarf

Wie bereits oben (vgl. Kapitel 4.4.2) dargelegt, gibt es einen erheblichen Mangel an Qualifizierung hinsichtlich des Umgangs mit Daten. Der Qualifizierungsbedarf und Ansatzpunkte für dessen Deckung stehen im Zentrum des Untersuchungsbereichs *Qualifizierung*. Zum besseren Verständnis der wichtigsten Aspekte von Qualifizierung sollen in den folgenden fünf Kapiteln näher betrachtet werden:

- Datenbewusstsein
- Data Literacy (Datenkompetenz)
- Kompetenzprofile
- Datenkompetenzvermittlung
- (Neue) Berufsfelder/-bilder

Der aktuelle Stand wird, wo möglich, auch spezifisch für die einzelnen Untersuchungsbereiche *Wissenschaft*, *Public Sector* und *Wirtschaft* dargestellt.

4.4.4.1 Datenbewusstsein

Bereits 2016 postuliert der RfII eine „neue Datenkultur“ und weist darauf hin, dass sich mit der Digitalität von Forschungsdaten neue Herausforderungen an die Qualitätssicherung ergeben (RfII – Rat für Informationsinfrastrukturen, 2016, S. 52 f.). „Eine ausreichende Datenqualität über den gesamten Datenlebenszyklus zu sichern“ erfordert eine „gemeinschaftliche Anstrengung“ von „Produzenten, Verarbeiter[n]“ und „diverse[n] Nachnutzer[n] von Forschungsdaten“ (vgl. ebd., S. 52). Dies bedinge „eine neue Datenkultur der Offenheit und des Teilens (Data Sharing), aber auch die Verständigung auf Standards“ (ebd., S. 52). Um das Ziel eines „gemeinsamen Qualitäts-Diskurses aller Akteure“ zu erreichen, sieht der RfII als Zwischenschritte u. a.

²⁵ <https://www.eoscpilot.eu/>

²⁶ <https://eoscfair4s.github.io/>

²⁷ Vgl. EOSC FAIR4S, o. J., Materialien und Deliverables 7.1 – 7.5

den Aufbau von kollaborativen Strukturen und die „Durchsetzung gemeinsamer (Qualitäts-)Standards und Werkzeuge“ (vgl. ebd.). Darüber hinaus empfiehlt der RfII Fachverbänden, wissenschaftlichen Vereinigungen und ausgewiesenen Infrastrukturbetreibern, „fachlichen Standards und Qualitätskriterien große Aufmerksamkeit“ zu widmen und „[f]ächerübergreifende Kooperationsmöglichkeiten“ sowie die „Zusammenarbeit mit anderen Disziplinen und Forschungsformen [...] im Rahmen von Datenplattformen und virtuellen Forschungsumgebungen zu berücksichtigen“ (vgl. ebd.). Als weitere Elemente werden u. a. „auf Interoperabilität ausgerichtete Metadaten“, „gute Dokumentationen“, die Konzeption von „geeignete[n] Review-Verfahren“ sowie der frühzeitige „Kontakt zu (möglichen) Datennutzern und fachtypischen übergreifenden Verbänden angeführt (vgl. ebd., S. 52 f). Weitere Empfehlungen sind „Datenmanagementpläne (DMP) als Planungselemente verbindlich einzufordern“ und durch „Mechanismen einer regelmäßigen externen Qualitätssicherung“ für „die wissenschaftlichen Nutzer und Förderer Transparenz herzustellen“ (vgl. ebd., S. 53). Es wird deutlich, dass – nicht zuletzt bedingt durch die fortschreitende Digitalisierung – eine Veränderung der wissenschaftlichen Forschung insbesondere im Umgang mit Daten im Gange ist (vgl. Kapitel 4.1.2). Daten und ihr möglichst einfacher Austausch rücken in den Fokus und erfordern einen Bewusstseinswandel aller am Datenlebenszyklus mittel- und unmittelbar Beteiligten. Es handelt sich nicht nur um Einzelaspekte, sondern um den ganzen Prozess, der für Datenkultur steht.

Dieser Wandel hin zur Datenkultur betrifft nicht nur die Wissenschaft, sondern die ganze Gesellschaft. So wird in der Datenstrategie der Bundesregierung unter Datenkultur „ein offenes Verständnis von Daten als Ressource für die Wissenschaft und gesamtgesellschaftliche Teilhabe“ (Bundesregierung, 2021a, S. 41) verstanden. Dieses Verständnis impliziert, „dass Erkenntnisse aus Daten gewonnen, aber auch kritisch hinterfragt werden können“ und dass eine „aktiv gelebte Datenkultur“ ermöglicht, „evidenzbasierte Entscheidungen zu treffen und im Lichte europäischer Werte zu spiegeln“ (ebd.).

Für das Datenbewusstsein und den angestrebten Kulturwandel in den Untersuchungsbereichen spielt die jeweilige „bereichsspezifische Rationalität“, d. h. die jeweilige Motivation und die vorrangig verfolgten Ziele eine wichtige Rolle, da sie den Umgang mit Daten beeinflusst. Ein Antrieb in der Wissenschaft ist es, gesichertes Wissen zu generieren (vgl. Kapitel 4.1.4). Das Verständnis der Daten ermöglicht es, neue Kenntnisse zu gewinnen und sie zu nutzen, um Ergebnisse überprüfbar zu machen. Im Public Sector wird der Umgang mit Daten durch gesetzliche Vorgaben und den öffentlichen Auftrag bestimmt (vgl. Kapitel 4.2.4).

Daten offen zur Verfügung zu stellen hat das Ziel, ihre Nachnutzung zum Wohl der Gesellschaft zu ermöglichen und zu fördern. Hier liegt der Fokus auf der Bereitstellung und den damit verbundenen Anforderungen. Die Wirtschaft verfolgt mit den sich neu entwickelnden Geschäftsmodellen der Plattformökonomie (vgl. Kapitel 4.3.4) das Ziel der Gewinnerzielung. Diese Plattformen werden sehr von der Verwendung von Daten bestimmt. Der Einsatz und die Verwendung von Daten werden so zu einem ökonomischen Faktor. Für den Umgang mit Daten in den Untersuchungsbereichen bedeutet dies, dass sie unter verschiedenen Blickwinkeln betrachtet werden müssen: als wissenschaftlicher Beleg, als etwas, das verfügbar gemacht werden muss, und als wichtige ökonomische Ressource. Entsprechend unterscheidet sich auch das notwendige Bewusstsein für die Daten in den einzelnen Bereichen.

Wie weiter oben dargelegt, werden für eine neue Datenkultur die Aspekte „Offenheit und Teilen“, d. h. „data sharing“ bzw. Datenaustausch als wichtig definiert. Das „Support Centre for Data Sharing“ der EU (*Was ist Datenaustausch? | Support Centre for Data Sharing*, o. J.) sieht in diesem Zusammenhang drei Faktoren, die den Informationsaustausch (= data sharing) im digitalen Zeitalter verändert haben und so zur Veränderung des Bewusstseins beitragen. Zunächst die „verbesserte Verfügbarkeit und Qualität der Daten und die einfachere und günstigere Art und Weise sie zu speichern, zu verarbeiten und zu teilen“ (vgl. ebd.). Als zweiten Faktor erkennt es den Beginn eines Kulturwandels, der sich bei Einzelpersonen, Organisationen und Regierungen in einem besseren Verständnis von Daten zeigt. Dieser ist mit ihrer Bereitschaft verbunden, Daten als Ressource zu betrachten, in die investiert werden sollte (vgl. ebd.). Als dritter Faktor wird die Einbeziehung von politischen Entscheidungsträger*innen gesehen. Ihre Entschlossenheit, diesen Bereich zu regulieren, im Bewusstsein für Chancen und Risiken einer gemeinsamen Nutzung von Daten als wesentlicher Bestandteil dieses Prozesses, wird hervorgehoben (vgl. ebd.).

Eine neue Kultur des Datenteilens soll das Bewusstsein für Daten bei den Beschäftigten verbessern und so dazu beitragen, dass Organisationen und Behörden Daten in der Praxis auf sichere und faire Weise untereinander austauschen und dabei die Rechte derer beachten, welche die Daten betreffen. Außerdem wird der Mehrwert für Dienste und Serviceangebote betont, der in einer Kombination von Daten aus verschiedenen Quellen liegt. Als weitere auf dem Austausch von Daten beruhende Vorteile werden die Verbesserung der Zusammenarbeit und einer datengesteuerten Entscheidungsfindung gesehen (vgl. ebd.).

Dementgegen werden als Hindernisse und Herausforderungen beispielsweise ein fehlendes Verständnis für

die Bedeutung von Daten, fehlende Kenntnisse, wie Daten geteilt werden können, eine grundsätzlich fehlende Bereitschaft zum Datenteilen, fehlende (technische) Möglichkeiten, Daten zu teilen und die noch verbreitete, oft nur auf die eigene Disziplin fokussierte Sichtweise („Silodenken“) gesehen (vgl. ebd.).

4.4.4.2 Data Literacy (Datenkompetenz)

Wie bereits dargelegt, erfasst der Kulturwandel durch digitale Daten über die Wissenschaft und Wirtschaft hinaus die ganze Gesellschaft. Katharina Schüller und Paulina Busch bezeichnen Data Literacy als „eine Schlüsselkompetenz des 21. Jahrhunderts, die in der modernen Gesellschaft und Arbeitswelt unerlässlich sein wird“ (Schüller et al., 2021, S. 11). Ähnlich formulieren Jens Heidrich et al., dass „Data Literacy [...] eine zentrale Kompetenz für die Digitalisierung und die globale Wissensgesellschaft in allen Sektoren und Disziplinen“ ist und sich „[a]ngesichts der zunehmenden Menge und der Verfügbarkeit von Daten [...] die Herausforderung [stellt], mit den Daten Wissen zu generieren und fundiert Entscheidungen treffen zu können“ (Heidrich et al., 2018, S. 106). Die Bundesregierung schreibt zum Thema Data Literacy, zu „wissen, was Daten sind, zu verstehen, welche Konsequenzen sich aus dem eigenen Umgang mit Daten ergeben; mögliche Folgen eines unvorsichtigen Umgangs absehen zu können; datenbasierte Berichterstattung kritisch hinterfragen zu können; zu verstehen, welche Regeln im Umgang mit personenbezogenen und nicht-personenbezogenen Daten gelten und wie man seine Daten sichern kann“ sind „wichtige Fähigkeiten für ein selbstbestimmtes Handeln im digitalen Zeitalter“ (Bundesregierung, 2021a, S. 41). Das gelte „ebenso für ein grundlegendes Verständnis über eine zunehmend datenbasierte Ökonomie“ (ebd.). Dies verdeutlicht, dass Data Literacy eine hohe Bedeutung beigemessen wird.

Um das angestrebte Ziel einer besseren Qualifizierung im Umgang mit Daten zu erreichen, muss zunächst einmal geklärt werden, welche Kompetenzen vermittelt werden müssen, d. h. was unter „Data Literacy“ zu verstehen ist. Für die Definition des Begriffs Data Literacy wird zumeist auf den Report von Chantel Ridsdale et al. Bezug genommen, in dem diese schreiben: „Data literacy is the ability to collect, manage, evaluate, and apply data, in a critical manner“ (Ridsdale et al., 2015, S. 8). Ähnlich definiert die Bundesregierung: „Datenkompetenz oder Data Literacy beschreibt u. a. die Fähigkeit, mit Daten in technischer, ökonomischer, ethischer und rechtlicher Hinsicht kompetent und werteorientiert umzugehen“ (Bundesregierung, 2021a, S. 108) und setzt damit Data Literacy mit Datenkompetenz gleich.

Um Data Literacy zu fördern und voranzubringen, wurde auf Initiative des Stifterverbands²⁸ und zahlreichen Partner*innen im Januar 2021 die Data-Literacy-Charta verfasst (vgl. Schüller et al., 2021). In ihr wird postuliert, dass Datenkompetenzen für alle Menschen in einer durch Digitalisierung geprägten Welt wichtig sind und dass sie unverzichtbarer Bestandteil der Allgemeinbildung sind. Die Charta enthält fünf Leitprinzipien: (1) Data Literacy muss allen Menschen zugänglich sein. (2) Data Literacy muss lebenslang in allen Bildungsbereichen vermittelt werden. (3) Data Literacy muss als transdisziplinäre Kompetenz fachübergreifend vermittelt werden und zwar aus (a) anwendungsbezogener (Was ist zu tun?), aus (b) technisch-methodischer (Wie ist es zu tun?) und aus (c) gesellschaftlich-kultureller (Wozu ist es zu tun?) Perspektive. (4) Data Literacy muss den gesamten Prozess der Erkenntnis- und Entscheidungsfindung mit Daten systematisch abdecken. (5) Data Literacy muss Wissen, Fähigkeiten und Werthaltungen für einen bewussten und ethisch fundierten Umgang mit Daten umfassen (vgl. ebd.). Zu den Unterzeichnern der Charta gehören neben dem Stifterverband zahlreiche Verbände, Hochschulen, Bibliotheken und Einzelpersonen aus Wissenschaft, Wirtschaft und Politik. Die Unterzeichner*innen der Data-Literacy-Charta verpflichten sich, Maßnahmen zu ergreifen, das Verständnis von Data Literacy zu verbreiten und die dazugehörigen Kompetenzen weiter zu stärken (vgl. Koch, 2021).

4.4.4.3 Kompetenzprofile

Ziel bei der Erstellung von Kompetenzprofilen ist es, berufsrelevante Kompetenzen zu beschreiben, um diese beispielsweise mit Anforderungsprofilen aus der Praxis abgleichen zu können. Daher geben sie einerseits wichtige Hinweise auf den Qualifizierungsbedarf für den Umgang mit Daten in einem spezifischen Kontext. Andererseits können sie auch dazu beitragen, die in diesem Bereich entstehenden neuen Spezialisierungen bzw. neuen Berufe besser zu definieren und ggf. untereinander klar abzugrenzen.

Ausgehend von ihrer Definition von Data Literacy als der Fähigkeit „Daten zu sammeln, zu verwalten, zu bewerten und anzuwenden“ (siehe oben) entwickeln Ridsdale et al. in ihrer Studie ein Kompetenzraster. Auf Basis einer thematischen Analyse von peer-reviewten Artikeln bestimmen sie wichtige Bestandteile von Data Literacy und definieren darauf aufbauend 23 Kompetenzen und 64 damit verbundene erforderliche Fähigkeiten, Kenntnisse und Aufgaben. Diese wiederum ordnen sie den Kategorien *Conceptual Framework*, *Data Collection*, *Data Management*, *Data Evaluation* und *Data Application* auf drei unterschiedlichen Niveaustufen (*Conceptual Competencies*, *Core*

28 <https://www.stifterverband.org>

FAIR4S Core Skills Groups



Abbildung 19: FAIR4S Core Skills Groups (Quelle: Whyte & Ashley, 2017, S. 43; Lizenz: CC-BY 4.0 international, <https://creativecommons.org/licenses/by/4.0>)

Skills group	Key skill
Plan and design	Plan stewardship and sharing of FAIR outputs
Capture and process	Reuse data from existing sources
Integrate and analyse	Use or develop open research tools / services
Appraise and preserve	Prepare and document for FAIR outputs
Publish and release	Publish FAIR outputs on recommended repositories
Expose and discover	Recognise, cite and acknowledge contributions
Govern and assess	Develop open research strategy and vision
	Apply policies to comply legal requirements, ethical & FAIR principles
Scope and resource	Secure funding for open science / support
Advise and enable	Lead good practice by example

Abbildung 20: FAIR4S Key Skills (Quelle: Key Skills – EOSC FAIR4S, 2019; Lizenz: CC-BY 4.0 international, <https://creativecommons.org/licenses/by/4.0>)

FAIR4S competences and capabilities		Recommended expertise by professional group and service role								
		Data Advisors				Data service providers				
		Data Steward	Research Mgr	Ethics & DP	Comm. & IP	Data Librarian	Research Swr Eng	Service Mgr	Service Architect	Archivist
Plan and design	Plan stewardship and sharing of FAIR outputs	●	◐	◐	◐	◐	◐	○	○	○
	Model data structures and define database needs	◐	○	○	○	◐	●	○	●	◐
	Specifying metadata and persistent id. standards	◐	○	○	○	○	◐	○	◐	●
	Analyse requirements for services or software	○	○	○	○	◐	●	○	●	◐
	Appraise and select repositories for FAIR sharing	●	○	○	○	●	○	○	○	●
Capture and process	Reuse data from existing sources	◐	○	○	○	●	◐	○	○	●
	Manage databases	◐	○	○	○	◐	●	○	●	◐
	Software prototyping	◐	○	○	○	○	●	○	●	○
	Set up and document workflows	◐	○	○	○	◐	◐	○	◐	◐
	File naming and organisation	◐	○	○	○	●	◐	○	◐	●
	Data cleaning, processing and software versioning	◐	○	○	○	◐	●	○	◐	●
Integrate and analyse	Use or develop open research tools / services	◐	◐	◐	○	◐	●	◐	◐	◐
	Critical thinking and theory building	○	○	○	○	○	◐	○	●	○
	Creative problem solving, flexibility	◐	◐	○	○	◐	◐	○	●	◐
	Math and statistical knowledge application	○	○	○	○	○	◐	○	○	○
	Data transformation and integration	◐	○	○	○	●	●	○	◐	●
	Data mining, querying, interpretation	◐	○	○	○	◐	◐	○	◐	◐
	Predictive modelling and analytics	◐	○	○	○	○	◐	○	○	○
	Machine learning methods	○	○	○	○	○	◐	○	○	○

Legend: ○ = basic, ◐ = inmediate, ● = expert, - = none

Abbildung 21: FAIR4S Skills Tabelle – Data Advisors und Data Service Providers (Quelle: Whyte et al., 2019, S. 80, Lizenz: CC-BY 4.0 international, <https://creativecommons.org/licenses/by/4.0>)

Competencies und *Advanced Competencies*) zu (vgl. Ridsdale et al., 2015, S. 3). Auf dieser Systematik bauen viele weitere Studien auf (vgl. Bandtel et al., 2021, S. 397).

Auch das im Rahmen des Aufbaus der EOSC entwickelte „Skills and Capability Framework“ FAIR4S²⁹ orientiert sich daran. FAIR4S wurde mit dem Ziel entwickelt, Forschungseinrichtungen und Communities

bei der Umsetzung von *Research Data Management* und *FAIR Stewardship* im Open-Science- und Data-Science-Kontext zu unterstützen und soll hier als Beispiel dienen. In diesem *Framework* (Modell) werden die vier Rollen *Researcher* („Forscher*innen“), *Data Scientist* („Datenwissenschaftler*innen“), *Data Advisor* („Datenratgeber*innen“) und *Data Service Provider* („Datendienstleister*innen“) unterschieden

²⁹ <https://eosc-fair4s.github.io/>

sowie die je nach Level der jeweiligen Expertise notwendigen Kompetenzen und Fähigkeiten zugeordnet, um *Data Stewardship* („Datenverantwortung“) in allen Phasen des Data-Life-Cycle-Prozesses zu übernehmen.

Die Abbildung 19 baut auf diversen anderen Modellen wie z. B. dem Curation Lifecycle des Digital Curation Centers auf. Dargestellt werden die *Skill Groups*, wobei der innere Ring *Data Stewardship* Aktivitäten auf Projektebene beschreibt und der äußere Ring projektübergreifende Support-Aktivitäten nennt. In Abbildung 20 sind den *Skills Groups* die *Key Skills* zugeordnet.

Allen *Skills Groups* werden in Tabellen jeweils Kompetenzen und Fähigkeiten zugeordnet und mit einer der mit den vier definierten Rollen verbundenen Aktivitäten verknüpft (Abbildung 21). Diese wiederum werden innerhalb der Rollen einzelnen „Berufen“ zugeordnet und verbunden mit einer Einschätzung des Levels der notwendigen Expertise. In weiteren Übersichten werden die Aufgaben detailliert nach Expert*innenlevel beschrieben.

Am dargestellten Framework FAIR4S wird deutlich, dass Kompetenzprofile für einzelne Berufstätigkeiten im vernetzten System betrachtet werden müssen. Sie beschreiben einzelne (Teil-)Aufgaben, die dafür jeweils erforderlichen Fähigkeiten und das notwendige Level an Expertise. Kompetenzen und Kenntnisse von einzelnen Beschäftigten z. B. mit unterstützenden Aufgaben allein reichen jedoch für eine funktionierende Zusammenarbeit beim Umgang mit Daten nicht aus. Es müssen auch adäquate Kompetenzen und Kenntnisse im Gesamtsystem vorhanden sein, d. h. beispielsweise, dass jede*r Beschäftigte ein Grundverständnis für die Zusammenhänge über den eigenen Verantwortungsbereich hinaus haben muss und dass alle notwendigen Kompetenzen und Kenntnisse an zumindest einer Stelle im System vorhanden sein müssen.

4.4.4.4 Datenkompetenzvermittlung

Es herrscht „große Einigkeit in der Literatur“, dass „mit der Vermittlung von Kompetenzen im Bereich Data Literacy möglichst früh begonnen werden sollte“, betonen Heidrich et al. in ihrer Studie *Future Skills* (Heidrich et al., 2018, S. 8). Für notwendig halten sie es auch, dass „sowohl Studierenden als auch Organisationen“ ein „Bewusstsein für die Wichtigkeit“ vermittelt wird und dass das „Angebot an verschiedene Bildungsniveaus und disziplinspezifisch an den Kontext, die Terminologie und den Arbeitsablauf der Problemstellung angepasst werden“ muss (ebd., S. 8).

Für den Hochschulbereich empfiehlt der RfII bereits 2016 „Module zur Vermittlung von Informationskompetenz und Datenmanagement Kenntnissen in das

gesamte Spektrum bestehender Studiengänge zu integrieren“ und neue „Vollstudiengänge“ einzuführen, „die auf neue Berufsbilder wie den (digitalen) *Dokumentar*, den *Data Librarian*, den *Datenarchivar* oder den *Data Scientist* mit Spezialisierung in den jeweiligen Fächern hinführen“ (RfII – Rat für Informationsinfrastrukturen, 2016, S. 50 f.).

Eine weitere Chance der Medienkompetenzvermittlung sieht der RfII darin, dass „übergreifend entwickelte Kompaktkurse zur Methodenkompetenz für Graduierte“ auch „zum Einstieg in neue Rollen wie Datenkurator, Datenmanager oder Datenanalyst befähigen“ können. „Wissenschaftsorganisationen“ sollten „Graduiertenschulen mit digitalen Profilen als gemeinsame, interdisziplinäre Kompetenzzentren für die Ausbildung des digital qualifizierten wissenschaftlichen Nachwuchses“ einrichten und diese dann mit anderen „gemeinschaftlich organisierten Fort- und Weiterbildungsaktivitäten verzahnt“ werden (RfII – Rat für Informationsinfrastrukturen, 2019b, S. 26). Speziell für den Untersuchungsbereich *Wissenschaft* empfiehlt der RfII Medienkompetenzvermittlung auch über „informelle Formate wie interne Praktika“, „befristeten Stellen Austausch und Auslandsaufenthalte“ anzubieten und zu prüfen, ob „Personal im Rahmen von Verbundstrukturen“ weiterqualifiziert werden kann (ebd., S. 22).

Eine hohe Bedeutung sollte auch der beständigen Weiterbildung von Lehrenden beigemessen werden. So empfiehlt der RfII „reguläre Fort- und Weiterbildungen zu Fragen digitaler Forschungsprozesse auch für Professorinnen und Professoren sowie Führungskräfte im Wissenschaftsmanagement“ zu forcieren (vgl. ebd.). Nicht unerwähnt sollte an dieser Stelle bleiben, dass in der Praxis die Weiterqualifizierung von Mitarbeitenden in Unternehmen, Behörden und anderen Einrichtungen vielfach über standardisierte Weiterbildungsmöglichkeiten erfolgt. Hier ist darauf zu achten, dass Angebote für die Medienkompetenzvermittlung in den bestehenden Weiterbildungskatalog aufgenommen werden.

In der Praxis sind bereits einige der oben gemachten Empfehlungen umgesetzt worden. Wie unter „Rahmenbedingungen und Vorgaben“ (vgl. Kapitel 4.4.2) dargelegt, gibt es zwar keine vollständige Übersicht über alle (neuen) Studienangebote, welche die Vermittlung von Datenkompetenz beinhalten. Es wurden jedoch einige Masterstudiengänge an deutschen Hochschulen in den letzten Jahren eingerichtet. Unter anderen seien hier exemplarisch der Masterstudiengang *Digitales Datenmanagement*³⁰ der Humboldt Universität zu Berlin und der Fachhochschule Potsdam, der Masterstudiengang *Data Science*³¹ der

30 DDM: <https://www.ddm-master.de>

31 Masterstudiengang Data Science: <https://www.uni-bielefeld.de/einrichtungen/zest/master-data-science>

Universität Bielefeld und der Data Science (M.Sc.)³² der Berliner Hochschule für Technik (BHT) genannt. Von den Wissenschaftsorganisationen bietet die Fraunhofer-Gesellschaft Fortbildungen mit der Zertifizierung als Data Scientist mit Spezialisierungen in Datenmanagement³³ oder Datenanalyse³⁴ an. Auch die Helmholtz Information & Data Science Academy (HIDA)³⁵ der Helmholtz-Gemeinschaft offeriert zahlreiche Trainingsangebote zum Thema Data Science. Darüber hinaus wurden in einigen öffentlich geförderten Projekten Schulungsprogramme entwickelt und Trainingsressourcen zusammengetragen. Hier sind beispielhaft die zwischen 2015 – 2019 im gleichnamigen EU-Projekt entwickelte PARTHENOS Training Suite³⁶, das im Rahmen des BMBF-Projekts FDMentor³⁷ zum Thema Forschungsdatenmanagement (FDM) entwickelte Train-the-Trainer-Programm (vgl. Biernacka et al., 2021) und die Ressourcen-Sammlung des DARIAH-CAMPUS³⁸ zu nennen.

In der Datenstrategie der Bundesregierung wird außerdem darauf hingewiesen, dass der „Erwerb von Digital Kompetenzen“ bei den meisten Bürger*innen „vorrangig durch eigenes Ausprobieren oder durch Hilfe aus dem Familien- und Freundeskreis“ erfolgt. Das gilt auch gerade für die „Menschen, die keine Weiterbildungsmaßnahmen am Arbeitsplatz oder in der Schule wahrnehmen können, sondern auf informelle und non-formelle Bildungsveranstaltungen“ angewiesen sind (Bundesregierung, 2021a, S. 41).

4.4.4.5 (Neue) Berufsfelder/-bilder

Beim Thema Berufsfelder ist zu unterscheiden, ob ganz neue Berufe entstehen oder ob bestehende Berufe um neue Inhalte und Schwerpunkte ergänzt werden. Die Akzeptanz und das Verständnis eines neuen Berufes, der aus einem bestehenden weiterentwickelt wurde, ist leichter zu erreichen. In diese Richtung zielt bereits oben erwähnte (vgl. Kapitel 4.4.4.4) Empfehlung des RfII zu neuen Berufsbildern wie z. B. dem „(digitalen) Dokumentar“, dem „Data Librarian“, dem „Datenarchivar“ oder dem „Data Scientist“, als Spezialisierung in den jeweiligen Fächern (vgl. RfII – Rat für Informationsinfrastrukturen, 2016, S. 50 f.).

Wie bereits oben dargelegt (vgl. Kapitel 4.4.4.3), werden im FAIR4S Framework die vier Rollen *Researcher* („Forscher*innen“), *Data Scientist* („Datenwissenschaftler*innen“), *Data Advisor* („Datenratgeber*innen“) und *Data Service Provider* („Datendienstleister*innen“) unterschieden und jede dieser Rollen mit drei bis vier Hauptkompetenzen verbunden (vgl. *Roles*

of professional groups – EOSC FAIR4S, o. J.). Bei der Beschreibung wird eine kurze Charakterisierung gegeben und u. a. die zugeordneten Kompetenzen genannt. Zugleich wird angegeben, welche alternativen Berufsbezeichnungen in anderen Modellen auf der Basis der gleichen zugeordneten Kompetenzen gewählt werden. So wird der Data Scientist in einem anderen Modell auch Data Analyst oder Data Steward sein. Der Data Service Provider könnte alternativ Archivist, Data Librarian, Data Service Architect, Research Software Engineer, Service Manager oder Project Manager sein. Die inhaltlichen Vorstellungen der verschiedenen Modelle für ein neues Berufsfeld sind zwar ähnlich, aber nicht ganz deckungsgleich. Dies veranschaulicht, dass sich trotz erster Ansätze neue Berufsbilder noch nicht eindeutig etabliert haben und dass einige bereits etablierte Berufe um Kompetenzen im Umgang mit Daten ergänzt wurden und so eine Spezialisierung erfahren haben.

Die stark zunehmenden Angebote im Bereich der Weiterqualifizierung machen deutlich, dass die vermittelten Datenkompetenzen in der Praxis eine hohe Nachfrage bedienen. Diese Angebote tragen dazu bei, die neu entstehenden Berufsbezeichnungen in naher Zukunft zu etablieren. Aktuell werden in Stellenanzeigen, bei denen Mitarbeitende mit Datenkompetenzen gesucht werden, diese eher über die Tätigkeitsbeschreibungen als über die Berufsbezeichnungen identifiziert. Als Beleg dafür dient die Auswertung von 162 Stellenanzeigen im Zeitraum vom 01.03.2020 bis zum 22.11.2020 im Rahmen einer Qualifizierungsarbeit, bei der knapp unter 10 % eine Berufsbezeichnung mit „data ...“ aufwies (vgl. hierzu: Zeunert & Schneemann, 2021, S. 154, Tabelle B.3 Berufsbezeichnungen).

4.4.5 Defizite und Handlungsbedarf

Der Fachkräftemangel zeigt deutlich, dass es insgesamt einen hohen Qualifizierungsbedarf gibt. Noch ist nicht bei allen Beschäftigten in den Untersuchungsbereichen *Wissenschaft*, *Public Sector* und *Wirtschaft* ein ausreichendes Datenbewusstsein vorhanden. Es fehlt teilweise an technischen, rechtlichen und organisatorischen Kenntnissen. Oft fehlt das Verständnis für den Wert von Daten und das Potenzial, das in ihnen steckt, wenn beispielsweise Datensätze verbunden werden. Datenteilen wird nicht als Aufgabe oder Chance, sondern als lästige Pflicht wahrgenommen. Fehlende Kompetenz führt zu Verunsicherung und aus Angst,

³² Data Science (M.Sc.): <https://www.bht-berlin.de/m-ds>

³³ <https://www.bigdata-ai.fraunhofer.de/de/data-scientist/schulungssuche/certified-data-scientist-specialized-in-data-management.html>

³⁴ <https://www.iais.fraunhofer.de/de/data-scientist-schulungen/schulungssuche/certified-data-scientist-specialized-data-analytics.html>

³⁵ <https://www.helmholtz-hida.de>

³⁶ <https://training.parthenos-project.eu/>

³⁷ <https://www.forschungsdaten.org/index.php/FDMentor>

³⁸ <https://campus.dariah.eu/resources>

Fehler zu machen, werden keine Daten geteilt. Um passende Qualifizierungsangebote bereitzustellen, muss zunächst in jedem Einzelfall geklärt werden, worin die vorrangigen Bedarfe bestehen. Je nach Ergebnis können technische, rechtliche oder organisatorische Kenntnisse und Fähigkeiten über entsprechende individuelle Weiterbildungen vermittelt werden. Für ein besseres Verständnis von Zusammenhängen wären daher beispielsweise auch Workshops für ein ganzes Team anzubieten.

Datenkompetenzen werden noch nicht flächendeckend vermittelt, da sie teilweise noch nicht Teil von Ausbildungsinhalten oder von Curricula-Vorgaben sind. An dieser Stelle sind die ausbildenden Einrichtungen und die mit der Aufstellung von Lehrplänen befassten Stellen in der Pflicht, ggf. in enger Zusammenarbeit mit Personen aus der Praxis, passende Inhalte zu erarbeiten und zu integrieren.

Generell gibt es einen hohen Informationsbedarf in vielen Bereichen. So gibt es nach wie vor keinen vollständigen Überblick über Ausbildungsangebote und Studiengänge, welche die Vermittlung von Datenkompetenz beinhalten. Das Gleiche trifft auf Weiterbildungsangebote zu. Hinzu kommt die fehlende Einschätzbarkeit des Wertes der Angebote, da es an geprüften und zertifizierten Abschlüssen und Weiterqualifikationen fehlt. Unklar ist, an welcher Stelle eine Zertifizierung erfolgen kann.

Neue Berufe sind im Entstehen, aber oftmals ist nicht etabliert, was sie genau beinhalten. Einige Tätigkeiten mit unterschiedlichen Bezeichnungen gehen fließend ineinander über, gleich bezeichnete Berufe beinhalten in der Praxis völlig andere Tätigkeiten. Hier wären eine Professionalisierung sowie eine klare Abgrenzung und Einordnung wünschenswert. In einem weiteren Schritt gilt dies auch für die Etablierung von möglichen Karrierewegen.

4.4.6 Zusammenfassung der Kernpunkte

Der Bedarf an Qualifizierung zu Datenkompetenzen ist europaweit sehr hoch und kann nach wie vor nicht gedeckt werden. Wenn man den Prognosen zum weiter steigenden Fachkräftemangel folgt, wird sich das Qualifizierungsdefizit noch weiter verstärken.

Trotz einer zunehmenden Digitalisierung aller Lebensbereiche sowohl in der breiten Bevölkerung als auch bei den Beschäftigten in den analysierten

Untersuchungsbereichen *Wissenschaft*, *Public Sector* und *Wirtschaft* ist ein Bewusstsein für Daten, das für einen Kulturwandel steht, abgesehen von einigen positiven Beispielen, noch nicht flächendeckend verbreitet. Es fehlt oft das Verständnis für den Gesamtzusammenhang von der Erhebung, Verarbeitung, Speicherung, Veröffentlichung, Nachnutzung bis hin zur Langzeitverfügbarkeit von Daten, in dem diese betrachtet werden müssen, um an der jeweiligen Stelle souverän mit den Daten umgehen zu können.

Damit eine geeignete Qualifizierung erreicht wird, müssen zunächst die Bedarfe, d. h. die notwendigen Fähigkeiten und Kompetenzen bestimmt werden. Im nächsten Schritt muss festgestellt werden, wie und an welcher Stelle diese am besten vermittelt werden können. Es ist eine große Herausforderung, dass die Anforderungen vor Ort oft sehr individuell sind. Eine übergreifende Sicht auf Gesamtzusammenhänge, wie sie etwa das Framework FAIR4S bietet, erleichtert es, bedarfsgerechte Qualifizierungsziele für Ausbildung, Hochschule und Weiterbildungsangebote festzulegen. Die Vermittlung von grundlegenden Kompetenzen im Umgang mit Daten beginnt bestenfalls bereits an den Schulen, sollte jedoch spätestens durch die Ausbildungsstätten und Hochschulen erfolgen. Bei der Entwicklung von Curricula geschieht dies bereits durch die Integration von Angeboten zum Erwerb von Datenkompetenzen in bestehende Studiengänge oder durch die Konzeption von neuen Studiengängen mit dem Schwerpunkt Daten. Hier ist eine dynamische Entwicklung festzustellen, weil immer mehr Angebote hinzukommen.

Allen bereits in der Arbeitswelt stehenden Beschäftigten sollten bedarfsgerechte Weiterbildungsangebote gemacht werden. Hier fehlt es einerseits an vielen Stellen noch daran, dass die genauen Qualifizierungsbedarfe unbekannt sind und andererseits ist das Angebot an Weiterqualifizierungen zwar sehr vielfältig, aber auch sehr unübersichtlich.

Es ist festzustellen, dass sich neue Berufe und Berufsbilder im Datenbereich entwickeln. Dies geschieht einerseits durch Spezialisierung oder Zusatzqualifikation innerhalb eines bereits bestehenden Berufsbildes wie etwa Data Librarian, Data Archivist bzw. Data Scientist. Andererseits entstehen neue spezialisierte Berufe wie der Data Steward, Data Analyst oder Data Manager. Die Übergänge sind fließend, die Berufsbilder nur teilweise etabliert und ebenso wie mögliche Karrierewege noch im Aufbau.

INTERVIEW-
BASIERTE
ANALYSE:
EXPERT*INNEN-
INTERVIEWS



5. Interview-basierte Analyse: Expert*innen-Interviews

In diesem Kapitel sind die Auswertungen aus den Expert*innen-Interviews dargestellt. Die individuell auf die Gruppen von Akteur*innen angepassten Fragenkataloge wurden auf Grundlage diskursiver Prozesse sowie auf der Basis des gemeinsam erarbeiteten bereichsübergreifenden Interview-Leitfadens (vgl. Anhang 10.4) erstellt. Die interviewten Expert*innen repräsentieren neben bestimmten Akteur*innen der einzelnen Untersuchungsbereiche auch bereichsübergreifende Stakeholder. Sie wurden mit der Methode des theoretischen Samplings ausgewählt, nachdem sich das Team mithilfe der LGA (vgl. Kapitel 4) einen Überblick über den jeweiligen Untersuchungsbereich erarbeiten konnte. Konzeptuelle Repräsentativität wird angestrebt (vgl. Strübing, 2014, S. 31). Im Folgenden sind alle Ergebnisse aus den jeweils fünf Expert*innen-Interviews für die einzelnen Untersuchungsbereiche in anonymisierter Form bzw. in Einzelfällen nach Absprache in namentlicher Nennung in thematischen Abschnitten beschrieben und zusammengefasst. Die Reihenfolge und Darstellung der Interviewergebnisse sind in diesem Kapitel nach Kernkategorien – unter der Berücksichtigung von thematischen Zusammenhängen und der Häufigkeit ihrer Nennung – angeordnet.

5.1 Wissenschaft

5.1.1 Einstieg

Die fünf Interviewten im Untersuchungsbereich *Wissenschaft* sind ausgewiesene Expert*innen in ihrem Bereich und somit prädestiniert, Einblicke über die derzeitige Beschaffenheit des Datenmanagements zu geben. Es wurde Expertise von bedeutenden Forschungseinrichtungen und zur Nationalen Forschungsdateninfrastruktur (NFDI) eingeholt. Beispielsweise wurden Prof. Dr. Achim Streit und Ingrid Dillo Ph.D. befragt. Beide haben ihr Einverständnis gegeben, namentlich in diesem Report erwähnt zu werden, weshalb ihre Expertise an dieser Stelle exemplarisch beschrieben wird:

Prof. Dr. Achim Streit ist Direktor des Steinbuch Centre for Computing (Informationstechnologie-Zentrum) und Professor am Karlsruher Institut für Technologie (KIT), der Forschungsuniversität in der Helmholtz-Gemeinschaft. Er ist verantwortlich für große Mengen an Forschungsdaten und beschäftigt sich mit der Analyse großskaliger Daten und datenintensivem Rechnen.

Ingrid Dillo Ph.D. bringt eine europäische Perspektive in die Arbeit ein. Sie startete ihre Karriere in den

Niederlanden im Ministerium für Bildung, Kultur und Wissenschaft in der Wissenschaftsabteilung. Inzwischen ist sie stellvertretende Direktorin des *Data Archiving and Network Services (DANS)*, dem nationalen Datenarchiv der Niederlande. In ihrer Position ist sie verantwortlich für die Expertise im Bereich Forschungsdatenmanagement und Datenkuration wie auch für die Zertifizierung von Repositorien. Diese Tätigkeit umfasst auch koordinative Aufgaben auf europäischer Ebene wie für das Projekt *FAIRsFAIR* oder für die Zertifizierungsstelle *CoreTrustSeal*. Als Co-Vorstand der *Research Data Alliance* und des wissenschaftlichen Komitees von *World Data System* sowie als Vorstandsmitglied bei *DataCite* ist sie auch in internationalen Organisationen sehr engagiert. Ingrid Dillo ist damit eine einflussreiche Expertin für Daten und deren Wiederverwendung.

Die zwei genannten Interviewpartner*innen werden auch im Folgenden namentlich zitiert. Die Expertise der anderen Befragten fließt gleichwertig in die Interviewauswertung mit ein, wobei auf die namentliche Nennung einzelner Personen verzichtet wird. Durch die Pluralität der Perspektiven aus dem Wissenschaftsbetrieb werden verschiedene Themenbereiche divers beleuchtet und besser verständlich.

In den folgenden Abschnitten werden die Aussagen aus den Interviews nach den Themen *Geschäftsmodell, Vorgaben und rechtliche Rahmenbedingungen, Datenbewusstsein, Kompetenzaufbau, Datenqualität und Standards, Vernetzung* und *Infrastruktur* sortiert sowie auf Ähnlichkeiten und Dissonanzen analysiert. Wissenschaftsspezifische Aussagen aus den Expert*innen-Interviews der anderen Untersuchungsbereiche finden in den Themenblöcken ebenfalls Platz. Die folgenden Themenbereiche werden zur besseren Vergleichbarkeit mit den anderen Bereichen getrennt analysiert. Empirisch ist keine derartig klare Grenzziehung möglich, weil die einzelnen Themen in hohem Maß voneinander abhängig sind.

5.1.2 Geschäftsmodell

Die Funktionsweise der *Wissenschaft* ist von den anderen Untersuchungsbereichen zu unterscheiden. Anders als in der *Wirtschaft* steht nicht die Gewinnmaximierung, sondern die Generierung von Wissen im Fokus. Während die Ziele der Verwaltung von der Politik abhängig sind, bewahrt der Untersuchungsbereich der *Wissenschaft* durch ihre Disziplinen eine gewisse Autonomie, um die Freiheit der Wissenschaft zu gewährleisten. Die freie Generierung von wissenschaftlichen Daten ist somit die oberste Priorität im Wissenschaftssystem, wobei nicht direkt mit

monetärem Gewinn zu rechnen ist. Trotzdem müssen Wissenschaftler*innen finanziert sein, um leben zu können und so das System möglichst nachhaltig zu sichern. Weiters muss es möglich sein, besondere wissenschaftliche Leistungen zu belohnen. Die Belohnungsmechanismen für Wissenschaftler*innen sind vielfältig und stehen meistens im Zusammenhang mit finanziellen Anerkennungen. Gemeint sind (Re-)Publikationen, Preise, Titel, Einfluss oder ehrenvolle Positionen in Gremien, um einige Beispiele zu nennen. All diese nicht-monetären Belohnungen, die Wissenschaftler*innen dazu motivieren, qualitativ hochwertiges Wissen zu produzieren, werden wie in Kapitel 4.1.4 mit Renommee zusammengefasst. Ähnlich wie Geld wird Renommee von Individuen gesammelt, akkumuliert und ausgespielt, um gewisse Ziele zu erreichen. Finanzierung und Renommee sind für die Personen im Wissenschaftsbetrieb untrennbar miteinander verknüpft. Wenn ein*e Wissenschaftler*in viel Renommee besitzt, ist die Jobsuche einfach. Sofern für die Finanzierung gesorgt ist, fällt es leichter Renommee aufzubauen. Ohne Renommee und ohne Finanzierung, ist es sehr schwierig, auch nur eines davon zu akkumulieren, um damit einen individuellen Karrierevorteil zu erreichen. In den Interviews wurden beide Begriffe, Finanzierung und Renommee, im Bereich Open Data verortet.

Arbeiten im Bereich Open Data sollen Renommee bringen. Universitäten können diesbezüglich Anreize schaffen, indem sie Wettbewerbe für die besten Open-Data-Publikationen gestalten oder Open-Data-Publikationen anderweitig bekannter machen. Weiters sollten Qualifikationen im Bereich Open Data bei der Stellenvergabe berücksichtigt werden. Teilweise wird das in wissenschaftlichen Einrichtungen schon so betrieben, aber es ist nicht die Regel. Berufungskommissionen sollen speziell für diese Thematik geschult sein. Dabei gilt es nicht unbedingt, den h-Index oder ein anderes quantitatives Bewertungsschema neu zu erfinden. Die Arbeiten im Bereich Open Data lassen sich kaum in einer Zahl abbilden, da sich die Einheiten, Granularitäten und Dimensionen zu stark unterscheiden. Viel eher könnte sich die Mitarbeit an Open Data als gewünschter Wert niederschlagen. In den Begutachtungsprozessen von wissenschaftlichen Texten könnte Open Data berücksichtigt werden. In diesem Prozess könnte sich auch der Re-Use, also die Weiterverwendung von Daten, positiv auswirken. Geldgeber*innen könnten Projekte unter der Bedingung finanzieren, dass schon vorab ein Datenmanagementplan vorgelegt wurde, der offene Datenpublikationen vorsieht. Zur Förderung der offenen Wissenschaft sollten die erwähnten Möglichkeiten nicht nur informell zur Anwendung kommen, sondern in entsprechenden Regelungen und Policies der Forschungseinrichtungen festgeschrieben sein.

Der freie Wissenschaftsbetrieb muss ausreichend finanziert sein. Hierbei empfiehlt es sich, zwischen Forschungs- und Infrastrukturprojekten zu unterscheiden. Die Finanzierung von Forschungsprojekten kann, wie eben erwähnt, von freien Datenpublikationen abhängig gemacht werden. Infrastrukturprojekte sind insofern gänzlich von Forschungsprojekten zu unterscheiden, als dass sie langfristig, sicher und resilient bestehen sollen. Es muss eine Möglichkeit geben, Prozesse zu schaffen, die das garantieren. Das beinhaltet entfristete Stellen, Handlungsspielraum und Zeit. Es ist das Ziel, Infrastrukturen für offene Daten mit hoher Qualität hervorzubringen, daher müssen gezielt Daten- und Infrastrukturprojekte gefördert werden, die an diesen Prozessen arbeiten. Da die Wissenschaft keine eigenen Mittel generiert, um Unternehmen bezahlen zu können, braucht es andere Wege, um einen freien Informationsfluss aus der Wirtschaft zu erschließen.

5.1.3 Vorgaben und rechtliche Rahmenbedingungen

Ethik- und Rechtsfragen haben als Querschnittsthema in der Wissenschaft einen überproportional hohen Stellenwert, da sich Wissenschaftler*innen auch in rechtlichen Graubereichen bewegen. In den Interviews kamen vor allem fachspezifische Vorgaben und europäisches Recht im Bereich Datenteilen zur Sprache. Auffällig ist, dass die fachspezifischen Vorgaben zum Teilen wissenschaftlicher Daten sehr unterschiedlich ausgestaltet sind. Sie unterscheiden sich in Tiefe und Art, sodass eine Standardisierung unmöglich ist. Teilweise werden die FAIR-Prinzipien schon lange implizit angewandt und teilweise müssen die Vorgaben um den einen oder anderen (Schwer-)Punkt ergänzt werden. Es besteht großes Einvernehmen darüber, dass Forschungsdaten so offen wie möglich publiziert und somit nach den FAIR-Prinzipien geteilt werden sollten. Dass dies immer noch nicht fächerübergreifend passiert, lässt tiefere Problemfelder in den Strukturen des Wissenschaftssystems erkennen. Dazu zählen, neben den fehlenden Sanktionen, vor allem rechtliche Unsicherheiten. Diese treten insbesondere im Bereich des Datenschutzes auf, wo sich rechtliche und ethische Fragen manifestieren. Obwohl die Datenschutzgrundverordnung wissenschaftliche Freiräume bewahrt und die Gesetze für offene Daten mehr Datenfluss, insbesondere auch aus der Verwaltung, garantieren sollen, scheinen die rechtlichen Verordnungen doch mehr Verwirrung als Vorteil bei den Forschenden zu stiften. Der Wunsch nach klaren, anwendbaren Regeln wird laut, insbesondere dort, wo es keine genauen fachspezifischen Vorgaben gibt. Mit der neuen Menge an digitalisierten, vernetzten, personenbezogenen Daten werden ethische Abwägungen neu getroffen.

Beispielsweise wird überlegt, inwiefern die Gesundheit der Bevölkerung mit persönlichen Gesundheitsdaten verbessert werden kann. Informierte Zustimmung der Beforschten scheint gesellschaftlich nicht mehr unbedingt nötig, obwohl dieses Einverständnis mehrfach in wissenschaftlichen und rechtlichen Vorgaben ange-dacht ist. Auch wenn nicht gerade ein Notfall, wie eine Pandemie, die Gesellschaft heimsucht, sollen Gesundheitsdaten geteilt werden, um Leben zu retten. Die Beforschten müssen über Vorteile und Risiken in Kenntnis gesetzt werden und dem Projekt entsprechend informiert zustimmen, vor allem, wenn besonders schutzbedürftig Daten vernetzt werden. Anonymisierungsverfahren verlieren bei zunehmender Vernetzung mit externen Daten an Wirksamkeit. Das wirft in der Medizin ethische Fragen auf: Wie schützenswert können Daten sein, wenn damit nachweislich Menschenleben gerettet werden? Derartige gesellschaftlichen Möglichkeiten, Fragestellungen und Wünsche werden durch digitales Forschungsdatenmanagement (FDM) sichtbar. Man scheint auf privatsphäre-freundliche, nachhaltige Innovationen und auf Konkretisierungen im Recht zu hoffen. In der Zwischenzeit bleibt weiter darauf hinzuweisen, dass die informierte Zustimmung der Beforschten zur Datenverwendung für Forschungszwecke verpflichtend vorgesehen ist. Das gilt für alle Forschungsdaten. Freies Datenteilen und der Schutz von geistigem Eigentum müssen rechtlich und technisch möglich sein.

Ein Datenmanagementplan soll Transparenz schaffen und von den Geldgeber*innen gefordert sein. Parallel dazu sollten von der Wissenschaft fachspezifische Vorgaben zum Datenteilen entwickelt werden, sofern das nicht schon umfassend passiert ist. Wobei auch darauf hingewiesen wird, dass es bei den derzeitigen rechtlichen Unsicherheiten vermessen scheint, verbindliche Vorgaben zu fordern. Die *Leitlinien zur Sicherung guter wissenschaftlicher Praxis*¹ der Deutschen Forschungsgemeinschaft können als Basis für fachspezifische Guidelines genutzt werden.

5.1.4 Datenbewusstsein

In diesem Abschnitt werden die zentralen Themen Zugriff, Vertrauen und Offenheit analysiert, die mehr oder weniger kontrovers zur Sprache kommen. Die unterschiedlichen Sichtweisen kann man, unter anderem, auf spezifische Machtverhältnisse im Wissenschaftsbetrieb zurückführen. Hier gilt es zu betonen, dass es große disziplinäre Unterschiede im Datenbewusstsein gibt.

Uneinigkeit herrscht bei der Frage, ob es eine zentrale Suchmaschine, wie Google, für den Zugriff auf Forschungsdaten braucht. Obwohl die

Recherchemöglichkeiten für vorliegende Datenbestände derzeit weit unterentwickelt scheinen und Verbesserungen notwendig sind, ist nicht garantiert, dass eine Suchmaschine der Problematik gerecht wird. Um neues Wissen zu generieren, ist es nicht notwendig, alle bestehenden Daten zu finden. Man braucht die passenden Daten, um der wissenschaftlichen Fragestellung gerecht zu werden.

Das Wissenschaftssystem ist mit seinen Daten- und Wissensbeständen dezentral aufgebaut. Dadurch gibt es wenig Standardisierung, was das Finden von Daten schwieriger gestaltet. Diese Dezentralität garantiert aber, dass der Wissenschaftsbetrieb funktioniert, selbst wenn einzelne wissenschaftliche Institutionen vorübergehend oder dauerhaft in ihrer Arbeit beeinträchtigt sind oder diese gänzlich einstellen müssen. Diese Funktionsweise schafft gesellschaftliches Vertrauen in wissenschaftliche Daten und Fakten, weil es keine zentrale Instanz gibt, die bestimmen kann, wie und was Wissenschaft ist und darf.

Doch Vertrauen beginnt, wo Gewissheit fehlt. In den Interviews wurde Künstliche Intelligenz als Beispiel genannt. Ein weiteres Beispiel ist das Wahlsystem: Wie weit will man einem elektronischen Wahlprogramm vertrauen? Führen analoge, lokale bzw. einfache, nachvollziehbare Prozesse zu mehr Vertrauen? Ähnliche Vertrauensprobleme gibt es beim Datenteilen. Dillo sieht Vertrauen als das zentrale Schlüsselement, welches in beide Richtungen notwendig ist: Menschen, die ihre Daten offenlegen, müssen darauf vertrauen, dass diese Daten nicht missbraucht werden. Auf der anderen Seite ist Vertrauen in die Qualität der Daten notwendig, sobald man sie nachnutzen will. Kann sich nicht auf die Datenbasis verlassen werden, gibt es auch keine seriösen Schlüsse, die man daraus ziehen könnte. Bringt das Teilen von Daten persönliche oder institutionelle Nachteile, wird das ein Hemmnis für den freien Fluss von Daten und Wissen. Es braucht mehr vertrauensbildende Maßnahmen zwischen Datenanbieter*innen und Datennutzer*innen, um offene Daten zu fördern. Es gibt auch noch andere Faktoren im Datenbewusstsein, außer fehlendes Vertrauen, welche die Öffnung von Forschungsdaten erschweren.

Offenheit wird als Ideal anerkannt, dennoch setzt es sich nicht durch. Man ist teilweise sehr zurückhaltend, wenn es um Open Data geht. Geheimhaltungspflichten, Embargos für Begleitartikel und die Sorgen um das geistige Eigentum gehören zu den Faktoren, die Forschende als Hemmnis oder gar Hinderungsgrund wahrnehmen. Zudem besteht wenig Klarheit, welche Daten offengelegt werden sollen. Gerade, wenn es darum geht, Daten im laufenden Forschungsprozess offenzulegen, stellt sich die Frage, welche Daten relevant sind und wie sich derartige Offenlegungen auf das

¹ <https://zenodo.org/record/3923602>

Forschungsprojekt auswirken. Es geht darum, weitgehend die Kontrolle zu behalten. Wenn Forschungsgruppen Daten zu früh aus der Hand geben, könnten Dritte mit Analyse und Auswertung schneller sein und der Forschungsgruppe das Renommee rauben. Der Begriff der Offenheit, weit ausgelegt, könnte ebenfalls bedeuten, dass auch Bürger*innen im Rahmen von Citizen Science an Datensätzen mitwirken. Es gibt kein einfaches Verständnis von Open Data, vielmehr existieren diverse Interpretationsansätze. Offenheit wird als kontinuierlicher Prozess verstanden.

5.1.5 Kompetenzaufbau

Forschungsdatenmanagement und die FAIR-Prinzipien wurden bisher nicht integral in der wissenschaftlichen Lehre verankert. Dillo, Streit und andere Befragte sind sich jedoch einig, dass man möglichst früh in der universitären Ausbildung beginnen soll, derartiges Wissen zu vermitteln. Ein Befragter meint, dass er sich Forschungsdatenmanagement in Master-Ausbildungen wünschen würde. Derzeit fehlt es an Wissen auf jeder wissenschaftlichen Qualifikationsstufe, was auch Weiterbildungen für Professoren notwendig macht. Es geht nicht darum, alle Forschenden zu Datenwissenschaftler*innen zu machen, doch ein gewisses Grundverständnis scheint notwendig, um Forschungsdaten richtig archivieren und nutzen zu können. Die Vermittlung von Data Literacy ist in diesem Kontext auch integral. Weiters soll Fachexpertise vermittelt werden: Es braucht Datenkurator*innen, Datenverwalter*innen und Weiterbildungen für Verwaltungsmitarbeiter*innen. Spezielle Infrastruktur erfordert spezifische Fähigkeiten, Innovationen sollten verstanden und antizipiert werden können. Auch juristische Expertise ist notwendig, um den gesetzlichen Rahmenbedingungen gerecht zu werden. Technische Möglichkeiten, Standards und Gesetze sind kontinuierlichen Veränderungen unterworfen, wodurch es sinnvoll ist, dass sich Fachpersonal beruflich mit den Themen beschäftigt. Dazu sind disziplinspezifische, kontinuierliche Beratungsmöglichkeiten für Forschende wünschenswert, um einen kontinuierlichen Wissensaustausch zu ermöglichen und damit der Aufwand zum adäquaten Forschungsdatenmanagement gering bleibt. Letztendlich soll das Engagement im Kompetenzaufbau und der Beratung die Datenqualität wissenschaftlicher Daten erhöhen.

5.1.6 Datenqualität und Standards

Es haben sich disziplinspezifische Standards entwickelt, die Datenqualität anzeigen. Es besteht kein Bestreben, diese zu vereinheitlichen, weil das schlichtweg nicht sinnvoll möglich ist. Stattdessen befasst man sich in der NFDI mit der Konzeptentwicklung, die

Interoperabilitäten zwischen den Ontologien der verschiedenen Fächer aufzeigen soll. Ähnlich wie bei der Softwareentwicklung, um ein Beispiel aus der Informatik zu nennen, ist die Software nie ganz abgeschlossen. Es gibt Verfahren und Techniken, sodass verschiedene Entwickler*innen Funktionen hinzufügen oder ungewollten Code entfernen können, ohne dass die Funktionsfähigkeit der Software eingeschränkt ist. So muss man Software mit der Zeit anpassen und verbessern. Analog ist das Wissenschaftssystem und das FDM, gerade in Hinblick auf die Digitalisierung, stark verbesserungswürdig.

Die FAIR-Prinzipien haben sich fächerübergreifend als sinnvoll erwiesen, um dieser Problemstellung zu begegnen. Doch es wird festgestellt, dass durch die FAIR-Prinzipien allein keine Annahmen über die fachliche-inhaltliche Qualität eines Datensatzes getroffen werden können. Man ist sich nicht einig, ob weitere Prinzipien, wie die TRUST-Prinzipien, notwendig sind. Es gibt derzeit auch noch kaum fachspezifische Vorgaben, wie die FAIR-Prinzipien anzuwenden sind. Klar scheint, dass nicht alle vier Prinzipien immer gleichwertig behandelt werden können. Ist ein Datensatz nicht auffindbar, spielt es beispielsweise keine Rolle, ob die anderen Prinzipien erfüllt sind oder nicht. Ingrid Dillo formuliert zusammenfassend:

„So on all letters we still have a lot to do, but the F and the A of course are simpler than the I and the R.”

Man wundert sich, dass die Auffindbarkeit und die Zugänglichkeit von Daten immer noch ein größeres Problem darstellen. Zudem wird festgestellt, dass die beiden Prinzipien, also die Auffindbarkeit und die Zugänglichkeit von Daten, nicht allein durch das Internet gegeben sind. Daten müssen zumindest maschinenlesbar sein, um gefunden zu werden, was beispielsweise bei handschriftlichen Notizen zu Datensätzen nicht der Fall ist.

Bei den Prinzipien Interoperabilität und Nachnutzbarkeit wird erkannt, dass sie sich nur sehr schwer auflösen lassen. Wenn Interoperabilität letztendlich den Zweck erfüllen soll, Daten von verschiedenen Orten aus gleichzeitig und in Echtzeit in einen gemeinsamen Graphen einspeisen zu können, gibt es nicht nur technische Herausforderungen. So ein Unterfangen würde standardisierte, einheitliche Metadaten-schemata voraussetzen, die in der Wissenschaft so kaum zu finden sind, schon gar nicht fachübergreifend. Selbst wenn man das Internet als gemeinsames Metadaten-schemata erkennen würde, sind nicht alle Daten digitalisierbar. Fehlende und unvollständige Metadaten verhindern es, dass offene Datensätze weiterverwendet werden können. Es lässt sich nicht bestimmen, wann alle Metadaten zu einem Forschungsprojekt vollständig offengelegt wurden, weil man nicht *alle* Metadaten aufzeichnen kann. Im Voraus ist es schwer abzuschätzen, wofür die Daten nach Projektende, in drei oder in

50 Jahren verwendet werden könnten. Der ursprüngliche Erhebungszweck bestimmt den Korridor für die Weiterverwendung der Daten in dem Sinn, dass spezifische Daten in einem bestimmten Kontext erhoben wurden.

Datenkuration, wie die Überprüfung von Inhalten und Qualität oder die Pflege von Datensätzen, wird getrennt von den FAIR-Prinzipien gedacht. Hochwertige Datensätze sollten kuratiert sein, wobei dafür Zeit, Geld und Personal zur Verfügung stehen sollten. Die Data-Kuratoren-Konzepte sind noch nicht voll ausgeschöpft, sodass sich das Berufsbild erst finden muss.

5.1.7 Vernetzung

Kooperation, Zusammenarbeit, Wissenstransfer, Partnerschaft, Austausch – all diese Begriffe beschreiben aktive Vernetzung. Sie ist Kernaufgabe des wissenschaftlichen Betriebs, da neues Wissen nur Geltung erlangen kann, wenn es von anderen erkannt und geteilt wird. Das Unterfangen der Vernetzung scheint ebenso schwierig wie lohnenswert. In den Interviews konnten drei große Bereiche identifiziert werden, in denen mehr Vernetzung angestrebt wird: inter- und transdisziplinärer Austausch innerhalb der Wissenschaft, die Vernetzung mit Wirtschaft sowie die Vernetzung mit der Verwaltung. Die Möglichkeit der Nachnutzung von Daten für Forschungskontexte ist dabei ein wichtiges Motiv für Vernetzung.

Die Nationale Forschungsdateninfrastruktur (NFDI) ist, neben diversen anderen Kooperationsprojekten, das umfangreichste trans- und interdisziplinäre Vernetzungsprojekt. Es werden gemeinsame Prozesse und Heuristiken zum Forschungsdatenmanagement entwickelt. Auch die European Open Science Cloud (EOSC) wird als Metadaten-Repository schließlich der Vernetzung von Forschungsdaten dienen. Die soziale Komponente ist bei der Gestaltung und Durchsetzung neuer Standards, wie der FAIR-Prinzipien, vielschichtig und multidimensional. Es braucht Zeit und Ressourcen, um gemeinsame Ziele zu erreichen. Nicht alle Disziplinen sind im selben Maß miteinander vernetzt. In verwandten Fächern fällt die Zusammenarbeit leichter, weil es schon ähnliches Vokabular und Methoden gibt, auf die man sich berufen kann. Ähnlich sind Kooperationen zwischen wissenschaftlichen Institutionen, die sich räumlich oder thematisch nahe stehen. Manche Disziplinen stehen auch im direkten Austausch mit der Wirtschaft.

Die Zusammenarbeit zwischen *Wissenschaft* und *Wirtschaft* wird angestrebt, insbesondere weil man sich monetäre Vorteile erhofft. Die NFDI soll wirtschaftlich anschlussfähig sein und man sieht gemeinsame Potentiale mit dem europäischen Wirtschafts-Projekt Gaia-X. Allgemein wird betont, dass die Pflege von Kooperationen eine aufwendige, zeitintensive

Tätigkeit ist, die aber notwendig scheint, um neue Datensätze erschließen zu können. Der Wissenstransfer soll in beide Richtungen stattfinden: von der Wissenschaft in die Wirtschaft und von der Wirtschaft in die Wissenschaft. Die zweite Richtung trifft vor allem Firmen, die Datensätze besitzen, mit denen Forschung betrieben werden kann. Derzeit sind viele Datensätze nur zugänglich, wenn man individuell Kooperationen mit den Firmen eingeht. Der freie Zugang zu Forschungsdaten soll für alle Forschenden verbessert werden.

Im *Public Sector* besteht Interesse an einer engeren Zusammenarbeit mit der Wissenschaft. Das geschieht über gemeinsame Projekte oder wenn sich Akteur*innen aus Verwaltung in der NFDI engagieren, die als Verein organisiert ist. Vom *Public Sector* werden Daten offen zur Verfügung gestellt. Forschende, die öffentliche Daten nutzen, stehen oft im direkten Kontakt mit den Behörden, um Fragen und Wünsche abzuklären. Es besteht die Möglichkeit, Forschende schon vor der Generierung neuer Datensätze aktiv mit einzubeziehen, um die Qualität, insbesondere bezogen auf die Nachnutzbarkeit, zu erhöhen. Die Vernetzung zwischen den Bereichen kann nie vollständig abgeschlossen sein. Es ist kontinuierliche Zusammenarbeit nötig, damit möglichst hochwertige Datensätze generiert werden.

5.1.8 Infrastruktur

Als Infrastruktur werden hier alle sozialen und technischen Komponenten verstanden, die zur Sicherung des Wissenschaftssystems betrieben werden. In den Interviews wurde vor allem über zwei Projekte, NFDI und EOSC, gesprochen. Es bestehen Unsicherheiten, insbesondere, wenn es um die technische und finanzielle Resilienz der Infrastruktur geht. Die niederländische Forschungsdateninfrastruktur wird, als Best-Practice-Beispiel, näher beschrieben.

Die deutsche NFDI und die EOSC werden gemeinsam besprochen. Auffällig ist, dass den meisten Expert*innen nicht klar ist, was die Projekte NFDI und EOSC für sie bedeuten. Es gibt unterschiedliche Vorstellungen und Wünsche an die Projekte, die sich nicht unbedingt mit den tatsächlichen Aufbauprozessen decken. Um ein klares Bild von den Projekten zu zeichnen, werden insbesondere die Aussagen der Interviewpartner*innen berücksichtigt, die im nahen Kontakt mit den Projekt-Initiativen stehen.

Die NFDI ist ein Verein, der sich zum Ziel gesetzt hat, das Forschungsdatenmanagement in Deutschland zu verbessern. Die NFDI fördert und besteht aus Menschen und Kompetenzen. Der Mensch in der wissenschaftlichen Community steht im Vordergrund. Von 2019 bis 2028 ist die Initiative mit 90 Millionen Euro pro Jahr finanziert. Nach Ablauf der Projektzeit soll

sich das Vorhaben selbst tragen, wobei Kooperationen angestrebt werden. Der Verein befindet sich im Aufbau. Es wurden vier Querschnittsthemen identifiziert, die von Sektionen bearbeitet werden. In einem möglichst konsensualen Prozess werden Empfehlungen erarbeitet, mit denen sich der wissenschaftliche Senat im Herbst 2021 befasst hat. Man kann heute noch nicht sagen, wie sich die NFDI entwickelt und ob es zusätzlich technische Infrastruktur braucht. Eventuell können Basisdienste angeboten werden, wobei noch nicht klar ist, was unter solchen verstanden werden kann. Der Wissenschaftsrat hat 2025 eine Evaluierung angesetzt, bei der die Strukturen überprüft und verbessert werden sollen. Als mandatiertes Mitglied repräsentiert die NFDI Deutschland in der EOSC. Die Arbeiten der Projekte sollen sich ergänzen.

Die EOSC ist ein Metadaten-Repository für offene Daten. Unklar ist, ob und inwiefern geplant ist, sie mit den verschiedenen anderen europäischen Datenräumen zu vernetzen. Es besteht die Befürchtung, dass der Wissenstransfer nur in eine Richtung geplant wird: von der Wissenschaft in die Wirtschaft. Man könnte das Miteinander größer in den Vordergrund stellen, sodass es der Wissenschaft ermöglicht wird, auch an Daten aus der Wirtschaft zu forschen. Es mag einen Markt für Daten geben, doch die Datenverarbeitung durch die Wissenschaft sollte getrennt davon betrachtet werden, da sie eben geteiltes Wissen anstatt finanzielle Gewinne erzielen will.

Im Diskurs ist auffällig, dass die Projekte für außenstehende Wissenschaftler*innen sehr intransparent wirken, während jene, die im Prozess eingebunden sind, über zu viel Information berichten. Es ist unklar, wie sich die Projekte entwickeln werden und ob eine langfristige Finanzierung erreicht wird. Es könnte sein, dass technische Faktoren in den Vordergrund rücken, oder aber soziale Themenfelder wie Beratungsdienste und Schulungen. Einzelne Forschungseinrichtungen wünschen sich durch die stärkere Zentralisierung der Forschungsinfrastruktur Kostenersparnisse im täglichen Betrieb, weil Funktionen kostenlos zur Verfügung stehen könnten. Es ist unklar, ob das geleistet werden kann. Derzeit werden Infrastrukturprojekte ähnlich befristet finanziert wie Forschungsprojekte, wodurch es nicht sinnvoll scheint, langfristig auf derartige Dienste zu bauen. Wünschenswert ist, dass die NFDI die Lücken in der Infrastruktur schließt, die eine einzelne Universität nicht schließen kann: beispielsweise verbesserte Suchmöglichkeiten für Forschungsdaten, erhöhte Datensouveränität, fachspezifische Beratungsmöglichkeiten für FDM oder die Bereitstellung von nationalen bzw. europäischen Datengraphen, die Daten verschiedener Herkunft in

Echtzeit abbilden. Die Möglichkeiten sind vielfältig, wobei es eben nicht darum geht, mit schon bestehenden Diensten zu konkurrieren. Der Aspekt der Nachhaltigkeit, im Sinn des Klimaschutzes und der Resilienz, ist essentiell. Die Wissenschaft ist derzeit von Hyperscalern und anderen wirtschaftlich ausgerichteten Diensten abhängig. Das muss nicht zwangsläufig so bleiben.

Die Dateninfrastruktur in den Niederlanden, die hier als Best Practice dienen soll, konnte durch Frau Dillo näher beleuchtet werden. Das niederländische Datenverwaltungssystem kann man sich in Kreisen vorstellen: Als Kern gibt es eine nationale Basisinfrastruktur, bereitgestellt von DANS² und SURF³, in dem es insbesondere um die Archivierung von Daten geht. Darum herum gibt es drei Kompetenzzentren zu Sozialwissenschaften, Geisteswissenschaften und Naturwissenschaften, die wiederum einen ganzen Kreis an lokalen Digital-Kompetenzzentren betreiben. Es handelt sich um ein „front office back office model“ auf nationaler Ebene. Die basale, kritische, technische Infrastruktur (back office) sowie Trainings und Beratung (front office) werden staatlich bereitgestellt. Datentreuhänder sind im niederländischen System lokale Datenverwalter*innen, die an einzelnen Forschungseinrichtungen beschäftigt sind. Sie sind speziell geschult und kümmern sich um das FDM der Einrichtung und Fragen dazu.

5.2 Public Sector

5.2.1 Einstieg

Für den Untersuchungsbereich *Public Sector* werden hier ergänzend zur Literaturoswertung die Ergebnisse der fünf interviewten Expert*innen dargelegt. Da der Untersuchungsbereich des *Public Sectors* sehr heterogen ist, wurde sich hier für eine granulare Herangehensweise entschieden. Es wurden Interviewpartner*innen mit verschiedenen Rollen und Funktionsebenen aus dem *Public Sector* sowie aus unterschiedlichen Aktivitätsspektren, wie sie auch im Data-Life-Cycle und im FAIRification-Prozess eine Rolle spielen, eingeladen. Bei der Auswahl der Interviewpartner*innen wurde versucht, die unterschiedlichen Erfahrungen und Interessen aus der Verwaltung, den kommunalen Unternehmen sowie von Koordinierungs- und Beratungsstellen abzubilden. Das Aktivitätsspektrum umfasst Produzent*innen, Bereitsteller*innen, Anwender*innen sowie Regulationsorgane des *Public Sectors*. Die ausgewählten Stakeholder sind Vertreter*innen aus den Bereichen Datenportale,

² <https://dans.knaw.nl/en/>

³ <https://www.surf.nl/en>

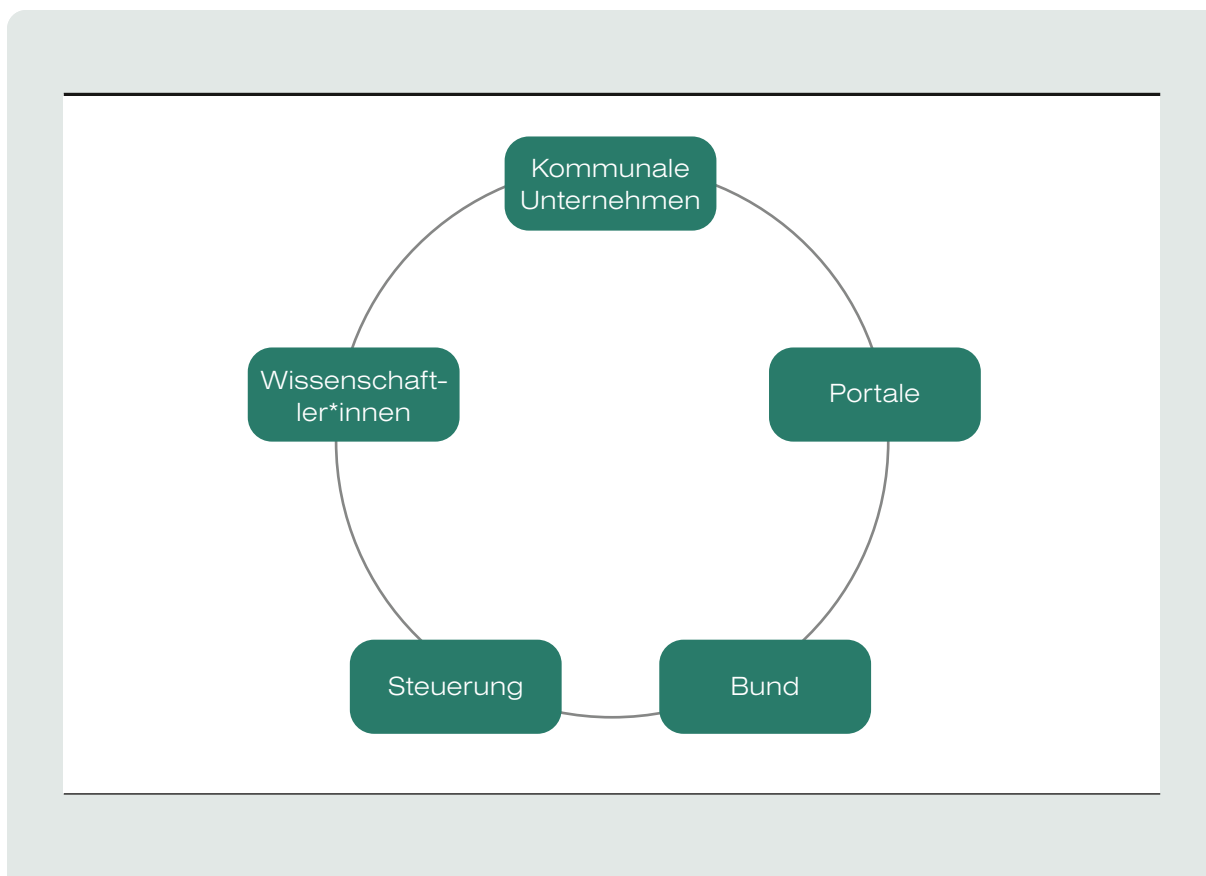


Abbildung 22: Auswahl von Stakeholdern für die Interview-basierte Analyse als zweite Untersuchungsphase (Quelle: EcoDM)

kommunale Unternehmen, Bundesbehörden und Steuerungsinitiativen sowie Wissenschaftler*innen (Abbildung 22). Alle Interviewpartner*innen zeichnen sich durch ihre Rollen und Tätigkeiten als Expert*innen in ihren jeweiligen Bereichen aus. Die Einschätzungen aus den Interviews werden in diesem Kapitel anonymisiert und thematisch zusammengefasst dargestellt.

5.2.2 Geschäftsmodell

Das Teilen von Daten erfolgt aufgrund von Vorgaben. Anreize oder Vorteile hingegen fehlen bzw. werden nicht aufgezeigt. Akteur*innen, die bereits in Projekten einen Mehrwert durch das Teilen von Daten hatten und technologisch aufgeschlossen sind, sind teilweise auch ohne Vorgaben bereit, Ressourcen für das Teilen von Daten einzusetzen. Gleichzeitig wurde die Pflege und Beschreibung von Daten über Personen, Organisationen oder Objekte als zentrale Aufgabe des *Public Sectors* benannt.

Die Frage nach der Finanzierung der Aufbereitung und Bereitstellung von Daten muss geklärt werden. Einige Behörden, wie beispielsweise Vermessungsämter, finanzieren sich zum Teil über den Verkauf ihrer Daten. Ein kostenloser Zugang zu diesen Daten würde für eine Lücke in der Finanzierung sorgen. Eine Finanzierung über andere Gebühren ist nicht möglich,

da Gebühren zweckgebunden verwendet werden müssen.

Die Gründe für den Mangel an Geschäftsmodellen im Bereich Open Data sind vielfältig: Sie reichen von unzureichender Absicherung und möglicherweise mangelnder Verfügbarkeit von Open Data – was wiederum eine mögliche Finanzierung einer Geschäftsidee basierend auf Open Data erschwert – über die Unmöglichkeit, Open Data mit vertraglichen Vereinbarungen wie Exklusivrechten zu kombinieren bis hin zu Wettbewerbsnachteilen für Unternehmen mit öffentlicher Beteiligung, die durch eine Datenteilungspflicht entstehen würde. Das Konzept der Datentreuhänder*innenschaft könnte zumindest eine Lösung für die Entlastung der Verwaltungen im Hinblick auf mangelnde personelle Ressourcen sein, doch bleiben sowohl die Frage der Finanzierung dieser Datentreuhänder*innen-Modelle ungeklärt als auch die Diskussion nach einer möglichen Gewinnerzielungsabsicht der Datentreuhänder*innen und damit einem möglichen Ausverkauf von Daten.

Es sind nicht alle öffentlichen Daten unabhängig von kommerziellen Aspekten, so stehen beispielsweise Verkehrsunternehmen in Konkurrenz zu anderen Unternehmen und nutzen ihre Daten auch kommerziell. Wenn Vorgaben Kooperationen verhindern, kann sich dies zugleich hemmend auf die Digitalisierung auswirken.

Der Bereich der Daseinsvorsorge kann im Sinne des Gemeinwohls als eines der wichtigsten Geschäftsmodelle für Open Government Data angesehen werden. Deswegen sollte die Pflicht für die gesicherte Herausgabe von Daten der Daseinsvorsorge mit Priorität behandelt werden. Die Regelungen sollten dabei für alle Unternehmen im Bereich der Energie-, Abfall- und Wasserwirtschaft gleichermaßen gelten und nicht nur für Unternehmen mit öffentlicher Beteiligung, um gleiche (Wettbewerbs-)Bedingungen zu schaffen. Hier steht nicht der wirtschaftliche Profit im Vordergrund, sondern vielmehr den Lebensstandard zu sichern und das Leben für alle Bürger*innen nachhaltiger und sicherer zu gestalten. Zusätzlich können Kooperationen mit wissenschaftlichen Einrichtungen oder Projekten die Daseinsvorsorge effizienter und umweltverträglicher machen, indem beispielsweise Künstliche Intelligenz (KI) im Bereich der Abfallwirtschaft zur Sortierung eingesetzt wird. Diese KI kann wiederum mit Open Data trainiert werden, denn der Datenbedarf für die Programmierung von KI ist immens und der Bedarf an frei zugänglichen offenen Daten ist durch KI deutlich gestiegen.

Es ist davon auszugehen, dass es durch die Rahmenbedingungen im öffentlichen Bereich zu Einschränkungen für Angebote zur Bereitstellung und Nutzung von Daten kommt. Dies kann sich beispielsweise auf Services auswirken, die auf Daten basieren, welche außerhalb der Dienstzeiten nur begrenzt zur Verfügung stehen.

Der Aufbau eines eigenen Datenbestandes wurde als Grundlage für eine bessere Verwaltung des eigenen Bereichs und als mögliches Hilfsmittel für Entscheidungsprozesse eingeschätzt. Im Hinblick auf die Grundversorgung sollte dem Datenmanagement und dem Datenteilen eine bevorzugte Rolle zukommen, welche sich in Vorgaben und Priorisierungen niederschlagen sollte.

5.2.3 Vorgaben und rechtliche Rahmenbedingungen

Die rechtlichen Rahmenbedingungen und Vorgaben sind in Deutschland aufgrund des Föderalismus sehr unterschiedlich. Grundsätzlich kann festgehalten werden, dass es in den letzten Jahren zu einer höheren Sensibilisierung beispielsweise im Bereich der Meldedaten oder anderen personenbezogenen Daten kam. So wird in Deutschland die Datenschutz-Grundverordnung (DSGVO) als EU-Umsetzung sehr restriktiv ausgelegt oder gesetzliche Grundlagen werden unter Berufung auf die DSGVO ausgehebelt. Dabei sollten gerade im Bereich der Gesundheitsdaten im Sinne des Gemeinwohls und der Innovation pseudonymisierte Daten für die Forschung verwendet werden dürfen. Hier fehlt es an klaren Regelungen.

EU-Richtlinien können in Deutschland teils aufgrund mangelnder Digitalisierung und einer nach wie vor papierbasierten Verwaltung nicht richtig umgesetzt werden. Eine Harmonisierung auf Bundesebene fehlt. Während manche Regionen mit gutem Beispiel vorangehen, liegen andere weit zurück. Dies ist auch eine Folge von fehlenden Konsequenzen bei Nicht-Erfüllung von Vorgaben, denn im Gegensatz zu zentralistisch organisierten Staaten kann der Bund hier gegenüber den Landesregierungen nur ein „ihr sollt“ einfordern. Die Zuständigkeiten für klare rechtliche Regelungen oder Gesetzesvorgaben werden zwischen Bund und Ländern hin und her geschoben. Es müssen hier eindeutige Lösungen geschaffen werden. Zudem regelt die PSI-Richtlinie zwar den öffentlichen Sektor, nicht aber die verschiedenen Domänen oder Datenräume. Hier sollte es gesonderte Absprachen geben.

Auch für vertragliche Regelungen fehlen größtenteils die rechtlichen Rahmenbedingungen. Das betrifft vor allem abgeleitete Datensätze oder auch wiederverwendete Daten, die dann für einen anderen Zweck genutzt werden sollen. Bei unterschiedlichen Rahmenverträgen kann eine Kombination von Daten nicht möglich sein, da sich vertragliche Absprachen widersprechen können. Der Gesetzgeber könnte beispielsweise über Ausschreibungen die Konsortial- oder Nutzungsverträge vorschreiben und damit Kooperationen in gemeinsame rechtliche Richtlinien zwingen.

Immerhin kann festgehalten werden, dass die Einführung des Datennutzungsgesetzes (DNG) den Druck auf die Verwaltungen erhöht, High-Value Data-Sets bereitzustellen. Auch die Verwendung empfohlener Lizenzen, wie der Datenlizenz Deutschland, setzt sich flächendeckend durch.

5.2.4 Datenbewusstsein

Im Gegensatz zu großen Unternehmen ist das Bewusstsein für die Relevanz von Daten und deren Auswertung in der Bevölkerung und auch in großen Teilen der Verwaltung noch nicht hinreichend geschärft, auch wenn das Thema innerhalb des *Public Sectors* an Bedeutung gewonnen hat. Das Teilen von Daten erfolgt jedoch meistens unsystematisch oder aufgrund von Vorgaben. Wurden bereits positive Erfahrungen im Datenteilen gemacht und darin Vorteile gesehen, steigert dies die Bereitschaft zu weiteren Kooperationen und zur Bereitstellung von Ressourcen für die Datenaufbereitung und -bereitstellung. Ein Übertragungseffekt sorgt in manchen Regionen für einen Anstieg in der Auseinandersetzung mit Daten. Wichtig ist in diesem Zusammenhang, wie Daten in einer Kommune gesammelt oder geteilt werden, da das Open-Data-Gesetz nicht hinreichend gelebt wird. Die Akteur*innen müssen aufgeklärt und motiviert werden. Jede*r

Einzelne muss verstehen, dass Data Driven Government Arbeit erspart und die Qualität verbessert. Das fehlende Selbstverständnis im Umgang mit Daten wurde in Verbindung mit mangelnden Kenntnissen als eine Ursache für eine Verzögerung der Digitalisierung innerhalb der Verwaltung gesehen. Ein Argument gegen eine zentralisierte Erfassung von Daten findet sich in der jüngeren deutschen Geschichte. Die Angst vor dem Versagen, in vielfältiger Hinsicht, ist einer der Hauptpunkte, warum Daten in Verwaltungen nicht hinreichend geteilt werden. Es gilt die Devise: Lieber nichts als etwas falsch machen. Auch die Angst vor Kontrollverlust oder mangelnder Expertise zur Interpretation der Daten hemmen die Bereitschaft zum Teilen. Diese Besitzstandswahrung verhindert potenzielle Umsetzungsprojekte, da die Daten nicht für Dritte verfügbar gemacht werden. Die Vielfältigkeit von Daten wird nicht erkannt, es wird vergessen, dass Daten in verschiedenen Kombinationen andere Potenziale entfalten können. Um den Mehrwert von Daten verständlich zu erklären, sind intensive Gespräche und praktische Umsetzungsprojekte notwendig. Eine gezielte Förderung von Kommunen in Form von zusätzlicher finanzieller Ausstattung, als Anreiz über Bundes- und Landesprojekte, kann für die nötige Motivation sorgen. Für frische Impulse sollte zudem der Austausch mit externen Expert*innen ermöglicht werden.

Der Kulturwandel muss in den nächsten Jahren als die zentrale Aufgabe innerhalb der Verwaltung gesehen werden. Dieser Prozess sollte engmaschig und strategisch begleitet werden. Die Umsetzung des Open-Data-Gesetzes kann durch ein zentral gesteuertes Change-Management unterstützt werden, denn Daten sind ein zentraler Treiber für die Verwaltungsdigitalisierung. Das alles kann allerdings nur gelingen, wenn unter den Verantwortlichen innerhalb der Kommunen das Vertrauen in die eigenen Daten gestärkt und die Verantwortung für das Sammeln und Verwalten der eigenen Daten übernommen wird. Dies bedeutet auch, klare Ansprechpartner*innen und Zuständigkeiten zu benennen. Hierzu sollen neben Verpflichtungen auch Anreize geschaffen und Vorteile aufgezeigt werden.

5.2.5 Kompetenzaufbau

Eine größere Kompetenz im Umgang mit Daten und deren Auswertung ist wichtig, um evidenzbasierte Entscheidungen treffen zu können. Aufgrund der mangelnden Expertise zur Auswertung und Simulation von Daten müssen sich die Verwaltungen aktuell noch auf externe Fachleute stützen, wodurch eine Überprüfung der Ergebnisse kaum möglich ist. Insgesamt wird übereinstimmend ein großer Mangel an Data Literacy und ein fehlendes Verständnis für die Granularität von Datensätzen sowie deren

Mehrwert attestiert. Der Kompetenzaufbau ist in den Kommunen unterschiedlich weit vorangeschritten, der Bedarf wird aber zunehmend von den Mitarbeiter*innen selbst formuliert und auch von den übergeordneten Behörden erkannt. Als Vorteil beim Kompetenzaufbau wird die Vorbildfunktion von Regionen und Kommunen gesehen, die bereits Daten teilen und andere Akteur*innen als Best-Practice-Beispiele zum Datenteilen animieren. Ein Grundwissen über den Umgang mit Daten wird für alle als unumgänglich angesehen, deshalb muss dem Aufbau dieser Kompetenzen in Zukunft eine noch höhere Priorität eingeräumt werden. Auch auf Bundesebene sollte bis auf Referatsebene eine größere Sensibilisierung im Umgang mit Daten vorhanden sein. Das Erlangen des Basiswissens sollte als Teil der Ausbildung oder des Studiums institutionalisiert werden. Zusätzliche Schulungen sorgen für eine größere Selbstsicherheit im Umgang mit Daten. Neben Grundwissen bei allen Mitarbeitenden soll es zukünftig spezialisierte Ansprechpartner*innen für Open Data in jedem Fachbereich geben. Durch Data Literacy werde die Grundlage gelegt, zu erkennen, welche Daten vorhanden sind und welchen Wert diese haben. Auch mögliche Risiken, die mit Daten in Verbindung stehen, können durch einen Kompetenzaufbau erkannt werden. Zugleich werde der Aufwand für die Pflege von Daten durch Fachkenntnisse reduziert.

Auf Bundesebene wird die Einrichtung von Data-Science-Teams in den Ministerien als Notwendigkeit angesehen. Diese sollen durch Datenrechtsexpert*innen unterstützt werden, um intern Ergebnisse und Rechtsfragen überprüfen zu können. Diese Data-Science-Teams sollen Themen von Datenöffnung, Datenbereitstellung bis hin zu technischen und rechtlichen Aspekten behandeln. Der Austausch zwischen Bund, Ländern und Kommunen soll stärker gefördert werden, beispielsweise durch eine zentrale Anlaufstelle wie das Kompetenzzentrum Open Data. Dieser Austausch soll auch eine bessere Information über Absprachen und Konventionen ermöglichen. In Bezug auf die Auswertung von Daten wurde der Föderalismus als Vorteil gesehen, da es immer Verzerrungen und verschiedene Interpretationen bei der Auswertung von Daten gebe.

Der Kompetenzaufbau wird maßgeblich durch den Kulturwandel unterstützt und kann in Form von Vorträgen und Workshops in die Verwaltung hineingetragen werden. Dies geschieht außerdem über eine individualisierte Beratung auf Basis einer regelmäßigen Zusammenarbeit. Wichtig ist es, das Wissen praxisbezogen durch Use- und Best-Cases in verschiedenen Ebenen einzubringen. Die Institutionalisierung der Durchsetzung von Open Data würde das Netzwerk verbessern und die Beratung erleichtern.

5.2.6 Datenqualität und Standards

Deutschland hat grundsätzlich eine gut funktionierende Verwaltung und gute Datenqualität, nur liegen die Daten meist noch ausschließlich in Papierform oder proprietären Formaten vor. Die Bereitstellerlandschaft ist allerdings durchaus heterogen. Während einige Behörden qualitativ gute Daten liefern, hängen andere in der Datenqualität noch etwas zurück. Open Data bietet insgesamt den Vorteil, dass dadurch nachnutzbare offene Daten bereitgestellt werden, die eine gewisse Qualität mitbringen und beispielsweise für KI verwendet werden können. Die Datenqualität wird als wichtiges Kriterium für die Nachnutzung von Daten benannt. Eine hohe Datenqualität ist beispielsweise Grundvoraussetzung für Auswertungen, die auf Abhängigkeiten von Daten untereinander basieren.

Neben der Datenqualität wird auch die Garantie einer langfristigen Verfügbarkeit der Daten als eine Grundlage für die Nachnutzung von Daten genannt. Die Entwicklung von datenbasierten Services und Kooperationen erfordert eine Garantie der Datenqualität und eine langfristige Verfügbarkeit. Um dies zu gewährleisten seien regelmäßige Aktualisierungen sowie die Pflege der Daten erforderlich.

Die bereichsübergreifende Vorgabe von Dateiformaten, Schnittstellen und die Schaffung von Standards wird als zentral für den Erfolg für das Teilen und Nachnutzen von Daten identifiziert. In der Schaffung von Standards wird das Potenzial zur Regulierung innerhalb eines Datenraumes, aber auch zur Verknüpfung mit anderen Datenräumen/Domänen gesehen. Daten sollen im Idealfall einem europaweiten Standard unterworfen werden, der im nächsten Schritt eine Verständigung auf globale Standards ermöglicht. Ein zentralisierter Metadatensatz wird dabei als dringend erforderlich angesehen.

Die fehlenden strukturellen Standards werden als Innovationsbremse genannt, da der Mehraufwand für die Aufbereitung der Daten beispielsweise durch Startups nicht geleistet werden kann. Auch innerhalb der Verwaltung erschweren das mangelnde Verständnis über die Vorteile maschinenlesbarer Daten und fehlende Schnittstellen behörden- und bereichsübergreifende Kooperationen. Gleichzeitig müssen auch Datenproduzent*innen, die nicht der staatlichen Kontrolle unterliegen, klaren Regeln unterworfen werden. Als Vorbild können europäische Richtlinien aus Datenräumen, wie beispielsweise die INSPIRE-Richtlinie⁴ aus dem Bereich der Geodaten dienen. So sollen Standards nicht nur in verschiedenen Bereichen gelten, sondern auch verschiedene Datenräume umfassen, um eine Vernetzung dieser Datenräume

problemlos zu ermöglichen. Die FAIR-Prinzipien sind innerhalb der Datenräume bereits Standard neben einigen zusätzlichen Ausprägungen wie FRAND-Lizenzbedingungen⁵.

Der Aspekt, dass die Festlegung von Standards durch einzelne große Unternehmen und nicht durch die Einbeziehung von Verbänden erfolge, wurde kritisch gesehen. Die Vorgabe von Standards in Förderprogrammen wird als mögliche Lösung genannt. Ein zentraler Dienst, der die Daten in einheitliche Strukturen und Formate bringt, könnte einerseits als Angebot genutzt werden und andererseits den Druck auf andere Datenproduzent*innen erhöhen.

5.2.7 Vernetzung

Während in der Wirtschaft das Interesse zur Zusammenarbeit als groß angesehen wird, sind Wissenschaft und Verwaltung eher kritisch gegenüber Kooperationen mit der Wirtschaft eingestellt. Ein Silodenken ist vorherrschend. Dieses löse sich zum Teil durch positive Erfahrungen mit bereichsübergreifenden Kooperationen auf. Dies erfolge jedoch nicht systematisch, sondern vielmehr, weil die Wirtschaft das Potenzial der Daten aus dem *Public Sector* für sich erkannt habe. Diese Kooperationen steigerten bei den Verwaltungen die Bereitschaft, Daten zu teilen und Lösungen anderer Akteur*innen nachzunutzen. Als Beispiel einer gelungenen Kooperation wurden die International Data Space Initiative aus dem Bereich Industrie 4.0 und GovTech aus Israel genannt.

Eine größere Transparenz zwischen den einzelnen Bereichen erleichtere das Teilen und Nachnutzen von Daten und ermögliche neue Lösungswege. Erfahrungen aus anderen Bereichen können genutzt und dadurch das Silodenken überwunden werden.

Das Teilen von Daten erfolge normalerweise lediglich in eine Richtung: So würden Daten aus dem *Public Sector* innerhalb der Wissenschaft genutzt, jedoch nutze der *Public Sector* selbst keine Daten aus der Wissenschaft. Auch müsse bei Projekten aktiv auf die Kommunen zugegangen werden, was durch die Struktur des Föderalismus zusätzlich erschwert werde. Das nötige Vertrauen kann über Nutzungsvereinbarungen und technische Lösungen sukzessive aufgebaut werden. Die Möglichkeiten von Angriffserkennungssystemen (IDS-Technologie, Intrusion Detection System) oder auch Apps, die innerhalb von Konnektoren installiert werden, garantieren beschränkte Zugänge zu beispielsweise sensiblen Daten. Dies bedeutet, die Rohdaten bleiben dem Datenbereiter vorbehalten. Als weitere technische Grundlagen für die Beförderung bereichsübergreifender Kooperationen werden

⁴ INSPIRE: Infrastructure for spatial information in Europe/ Europäische Geodateninfrastruktur für eine gemeinschaftliche Umweltpolitik, <https://inspire.ec.europa.eu/>

⁵ FRAND: FRAND: Fair, reasonable and non-discriminatory

die Schaffung eines zentralen Metadaten-Verzeichnisses, die Einigung auf gemeinsame Vokabulare, Standards und Schnittstellen genannt.

Um einen Mehrwert aus dem Teilen von Daten zu generieren, sollten diese nicht nur zur Verfügung gestellt werden, sondern gemeinsam mit den Projektpartner*innen eine Lösung oder eine Anwendung entwickelt werden. Hierbei spielt neben Mehrwert aus einer Kooperation auch die Finanzierung der Mehraufwände eine Rolle.

Kritisch bemängelt wird die fehlende Möglichkeit zum Austausch und zur Vernetzung in Open-Data-Portalen. Dadurch ist es nicht möglich, auf Fehler hinzuweisen oder mit der Community ins Gespräch zu kommen. Im Austausch mit Datennutzer*innen wäre es einfacher, die Datenbedarfe zu verstehen und daraufhin entsprechenden Einfluss auf die bereitgestellten Daten zu nehmen. Die Open-Data-Prinzipien sollen in diesem Zusammenhang nicht einfach übernommen, sondern im Hinblick auf ihre Wirksamkeit auch kritisch hinterfragt werden.

Ein größeres Netzwerk aus Stellen des Landes und des Bundes kann die Ausgestaltung von Open Data in der Verwaltung voranbringen. Durch regelmäßige Treffen über Fachforen können Bedarfe schneller erkannt werden, die Einbindung von Stakeholdern aus der Community bringt zudem die Fachexpertise aus Nutzer*innensicht mit ein.

Als eine Chance wird die Regulierung von Daten in Domänen verstanden. Die europäische Datenstrategie greift diesen Gedanken auf und wird daher positiv erwähnt.

5.2.8 Infrastruktur

Aufgrund des föderalen Systems sind die Daten in Deutschland auf viele verschiedene Akteur*innen verteilt. Dies erschwert die Kommunikation über und die Errichtung einer funktionierenden Infrastruktur. Übereinstimmend werden die infrastrukturellen Voraussetzungen für ein funktionierendes Datenmanagement in allen Ebenen der Verwaltung als ausbaufähig beschrieben. Als Hauptgrund dafür werden vor allem fehlende finanzielle und personelle Ressourcen benannt. Auch die fehlende Qualifizierung der Verwaltungsmitarbeiter*innen zum verantwortungsvollen Umgang mit Daten wird als ein wichtiger Grund angesehen. Das technische Know-how ist kaum vorhanden oder veraltet und erschwert die Durchsetzung von Open Data. Als zentraler Punkt zur Beförderung des Datenteilens und der Datennachnutzung wird die langfristige Finanzierung von technischer Ausstattung, Personal und Infrastrukturprojekten angesehen. Die Rekrutierung von geeignetem Personal erweist sich im Rahmen der TVöD-Richtlinien als schwieriges Unterfangen, da Data Scientists in der freien Wirtschaft

ein Vielfaches im Vergleich zur Verwaltung verdienen können. Hier wäre eine Anpassung der Lohnrichtlinien wünschenswert oder die Schaffung anderer Anreizsysteme.

Auf Ebene der Landkreise wird die Einrichtung von Data Lakes empfohlen. Diese sollten interoperabel aufgebaut werden, sodass bei Bedarf Daten zwischen verschiedenen Verwaltungen ausgetauscht werden können, wie es beispielsweise im Transparenzportal der Stadt Hamburg möglich ist. Diese dezentrale Verwaltung kann durch ein zentrales Metadaten-Portal durchsucht werden. Unterstützt werden sollte diese Art der Datenverwaltung durch eine zentrale Anlaufstelle, die organisatorische und technische Beratungsarbeit leistet. Um dieses System zu implementieren, sollten Kommunen über finanzielle Anreize sowohl durch Landes- als auch durch Bundesprojekte gefördert werden. Zudem wird es als wichtig angesehen, den Hauptaufwand nicht unbedingt bei der Datenbereitstellung zu verorten, sondern bei der Beantwortung von Anfragen. Dies müsse stärker personell berücksichtigt werden oder könne über eine*n Datentreuhänder*in abgefangen werden. Generell sollen sich die Landkreise ihrer Verantwortung beim Sammeln, Aufbereiten und Verwalten ihrer Daten (verstärkt) bewusst werden.

Innerhalb der Landesebene wurde die Vielzahl an Initiativen und Programmen positiv angemerkt. Als Best-Practice-Beispiele wurden Open.NRW sowie die Landesverwaltung von Schleswig-Holstein genannt, die gezielt Daten von den Kommunen einsammeln und an das Portal GovData liefern. Solche Landesdaten-Portale können die Daten aus den Kommunen aggregieren und in einheitliche Formate und Standards bringen.

Für eine Daten-Cloud und ein Metadatenprofil werden zentralisierte Lösungen eindeutig bevorzugt. Gerade im Bereich eines Metadaten-Portals sollen dabei keine neuen Angebote geschaffen, sondern der Ausbau der bestehenden Lösungen finanziell unterstützt werden. Durch die Einrichtung einer zentralen Vergabeplattform können Konventionen festgelegt und Datenheterogenität vermieden werden. Ein festes Regelwerk zur weiteren Nutzung der Daten schafft außerdem Vertrauen.

Zur Förderung von kleinen Unternehmen und Startups wäre es wichtig, die Daten des *Public Sectors* in Europa zu behalten und für diese bereitzustellen. Für Bereiche der Daseinsvorsorge wird die Schaffung von interoperablen Datenräumen präferiert. Gaia-X wird als mögliche Brückenlösung für den Austausch von Daten zwischen verschiedenen Datenräumen gesehen. Die Verwaltung dieser Datenräume von öffentlichen Stellen wird gewünscht.

Das Angebot einer europäischen IT-Infrastruktur mit Gaia-X wird als sinnvoll erachtet, da dadurch die

Idee einer europäischen Daten-Souveränität verfolgt wird. Gaia-X wird nicht nur als Bereitsteller freier Daten diskutiert, sondern könnte zugleich als ein Daten-Marktplatz gewerbliche Ansätze abdecken. Eine stärkere Beteiligung der EU bei Gaia-X wird positiv bewertet, außerdem könne durch einen einheitlichen Metadatenstandard bei Gaia-X eine Öffnung der verschiedenen Bereiche zueinander erreicht werden.

Die Souveränität über die Daten muss unbedingt gewährleistet werden. Das sei durch die Vereinbarung von Nutzungsbedingungen und durch technische Lösungen möglich.

5.3 Wirtschaft

5.3.1 Einstieg

Dieses Kapitel enthält die Zusammenfassung der Auswertungen der Expert*innen-Interviews für den Untersuchungsbereich *Wirtschaft*. In die Auswahl für die fünf Expert*innen im Untersuchungsbereich *Wirtschaft* sind Perspektiven von Unternehmen verschiedener Branchen und Größen, vom großen internationalen Konzern mit historisch gewachsenen Strukturen über mittelständische Unternehmen bis hin zu kleinen Start-ups eingeflossen. Darüber hinaus wurden über die Auswahl der Interviewpartner*innen spezifische Kompetenzen wie z. B. Erfahrungen im digitalen Datentransfer oder juristische Kompetenz eingebracht. Aufgrund dieser breit gefächerten Auswahl der Interviewpartner*innen und dem damit bewusst verknüpften breiten Spektrum an Blickwinkeln auf das Thema Nachnutzung von Daten, sind auch die hier zusammengefassten Inhalte der Interviewauswertungen sehr divers und reflektieren unterschiedliche Kompetenzen, Interessen und damit auch Unterschiede in Hinsicht auf die Benennung von Herausforderungen. Zusätzlich wurden für die Zusammenfassung in diesem Kapitel auch wirtschaftsrelevante Aussagen von den Interviewpartner*innen aus den anderen Themenbereichen, die in diesem Projekt unterschieden werden, berücksichtigt.

In den folgenden Abschnitten werden die Aussagen aus den Interviews nach den Themen *Geschäftsmodell*, *Vorgaben und rechtliche Rahmenbedingungen*, *Datenbewusstsein und Datenkompetenz*, *Datenqualität und Standards*, *Vernetzung* und *Infrastruktur* sortiert sowie auf Ähnlichkeiten und Dissonanzen analysiert.

5.3.2 Geschäftsmodelle

Grundsätzlich zeigte sich innerhalb des Untersuchungsbereichs *Wirtschaft* eine starke Erwartungshaltung gegenüber Daten: Sie müssen gewinnbringend, unternehmensfördernd und wettbewerbsorientiert

einsetzbar sein. Dies beinhaltet, dass dem Teilen von Daten ein Konzept zur Monetarisierung bzw. eine unternehmerische Effizienzsteigerung zugrunde liegen sollte. Eine Freigabe von Daten aus dem wirtschaftlichen Tätigkeitsfeld, wie es in der Wissenschaft im Rahmen von *Open Science* erfolgt, ist aus unternehmerischer Perspektive kaum denkbar. Sollen Unternehmen zur Freigabe eigener Daten zur Nachnutzung durch Dritte motiviert werden, kann dies am ehesten durch Angebote, welche die Profitorientierung von Unternehmen ansprechen incentiviert werden.

Dem entspricht eine erhöhte Bereitschaft zum Teilen von Daten, sofern damit Geschäftsmodelle bzw. die damit verbundenen Aussichten auf Gewinne verbunden werden können. Ein weiterer Ansatz zur Förderung des Datenteilens ist demnach die Unterstützung der Entwicklung von entsprechenden Geschäftsmodellen. Demgegenüber wurde von einzelnen Interviewpartner*innen vor einer pauschalen Verordnung der öffentlichen Zugänglichmachung von Daten gewarnt, weil damit das Risiko verbunden ist, etablierten Geschäftsmodellen bzw. Wertschöpfungsketten den Boden zu entziehen. Diese Problematik wird vermehrt in der Zusammenarbeit mit öffentlich finanzierten Projekten beobachtet, wo durch Data-Sharing-Modelle exklusive Zugriffsmöglichkeiten auf projektintern generierte Daten verloren gehen können. Aufgrund dieser Problematik sollten gesetzliche Regelungen, die Verpflichtungen zur Zugänglichmachung von Daten enthalten, so differenziert gestaltet werden, dass nicht intendierte Auswirkungen möglichst vermieden werden. Für Akteur*innen, die Daten erzeugen und diese für eine Nachnutzung durch Dritte zugänglich machen oder ihrerseits Daten von Dritten nutzen wollen, gewinnt das korrespondierende Rechtmanagement (Wem gehören die Daten? Wer darf sie wann nutzen? etc.) stark an Bedeutung. Bei Verbundprojekten, innerhalb derer Daten erzeugt werden, sollten deshalb Nachnutzungsszenarien und Rechtmanagement möglichst frühzeitig bedacht und projektintern klar geregelt werden. Es wird beobachtet, dass Geschäftsmodelle dann erfolgreicher sind, wenn sie Finanzierungskonzepte, die nicht allein die Förderung des öffentlich Zugänglichmachens von Daten beinhalten, sondern vielmehr auch deren proaktive Vermarktung und Qualität in Bezug auf Nachnutzung mit einbeziehen.

Die Förderung der technischen Voraussetzungen und regulatorischen Rahmenbedingungen zum Datenteilen gehen mit den Entwicklungen von Geschäftsmodellen einher. Neben der Entwicklung von eben jenen Geschäftsmodellen steckt ein wahrscheinlich größeres Wertschöpfungspotential in der Prozessoptimierung, die durch das Teilen bzw. Nachnutzen von Daten möglich wird. Dies wird bei der Förderung des Datenteilens oft nicht berücksichtigt und sollte zukünftig mehr in den Fokus rücken.

5.3.3 Vorgaben und rechtliche Rahmenbedingungen

Übergreifend wurde deutlich, dass bei der Förderung der deutschen und europäischen Digitalwirtschaft, speziell in der Entwicklung der rechtlichen Rahmenbedingungen, Handlungsbedarf besteht. Insbesondere bei der Aggregation großer Datenmengen durch eine kleine Anzahl von Großunternehmen braucht es die Entwicklung eines rechtlichen Instrumentariums, damit der Zugriff auf Daten von öffentlichem Interesse ermöglicht wird. Auch bei der Formulierung von Policies zur Förderung des Datenteilens sollte stärker berücksichtigt werden, dass Unternehmen im Regelfall darauf fokussiert sind, Intellectual Properties zu schaffen und zu verwerten. Der Zugriff auf jene Daten kann jedoch ermöglicht werden, wenn diese Zugriffserlaubnis Teil eines aus der Perspektive des Unternehmens funktionierenden Geschäftsmodells ist.

Es gibt viele Akteur*innen mit Interesse an personenbezogenen Daten. Hierbei ist die Datenanonymisierung in Zusammenhang mit Datentransfer ein vielschichtiges Thema, das als Moving Target zu behandeln bleibt. Obwohl durch die Datenschutz-Grundverordnung (DSGVO) in der Europäischen Union die Nachnutzung personenbezogener Daten unter Wahrung bestimmter Schutzniveaus ermöglicht wird, bedarf es einer Überarbeitung der aktuellen DSGVO, wobei hier der multi-relationale Charakter von Daten verstärkt mit aufgegriffen werden muss. Derzeit wird fast ausschließlich das bilaterale Verhältnis zwischen Datengeber*innen und Datenempfänger*innen geregelt, nicht aber die horizontalen Beziehungen von personenbezogenen Daten. Darüber hinaus braucht es in der Datenschutzgesetzgebung in Deutschland und Europa verbesserte Regularien an objektiv bewertbaren Kriterien, gerade im digitalen Bereich des Datenteilens, um diese handhabbarer zu gestalten bzw. Aufsichtsbehörden in die Lage zu versetzen, geltendes Recht auf handhabbare Weise zu implementieren. Des Weiteren gibt es einen Bedarf an der Entwicklung von differenzierteren, rechtlich-technischen Voraussetzungen für eine automatisierte Weitergabe datenschutzrechtlicher Präferenzen. Die derzeit verbreitete Praxis, Einwilligungen einzeln beim Aufruf von Webseiten abzufragen, führt vielfach zur Überforderung der Personen, die zu Einwilligungen aufgefordert werden. Die Umstände der Aufforderungen sind zudem vielfach so gestaltet, dass Zustimmungen unbeabsichtigt bzw. nur bedingt freiwillig gegeben werden.

Die wichtigsten Herausforderungen für die Förderung der europäischen Digitalwirtschaft und damit auch der Nachnutzung von Daten sind Systemsicherheit, *Identity Management* und Datenschutz. Daher sollte es, über deutsche rechtliche Regularien hinaus, das europäische Ziel sein, bestimmte Qualitätsanforderungen im

Bereich Datenschutz und im Bereich Datenverwendung in der EU als Regelwerk zu etablieren. Kurzfristig sollten dann auch die großen Hightech-Konzerne auf eben diesen europäischen Standard beim Datenteilen gebracht werden, mit der Hoffnung, dass dieser auch als weltweite Spielregel durchgesetzt werden kann.

Das Fehlen von klaren Strukturen bei Datentreuhandmodellen führt weiterhin dazu, dass große Akteur*innen der Datenwirtschaft sich teilweise einer Kontrolle der Einhaltung der Datenschutzgesetzgebung entziehen können. Parallel zu den technischen Entwicklungen von Systemen der Datentreuhänderschaft braucht es die Festlegung rechtlicher Rahmenbedingungen, die eine Verpflichtung der großen Konzerne mit umfangreichen Datenbeständen zur Kooperation enthalten. Diese Datentreuhandsysteme (z. B. E-Commerce, das Verkäufern und Käufern eine sichere Vertragsabwicklung ohne die Notwendigkeit der Preisgabe personenbezogener Daten ermöglicht) sollten öffentlich und nicht durch die Unternehmen, deren Praxen reguliert werden sollen, kontrolliert werden. Eine entsprechende Aufsicht durch unabhängige Datentreuhänder*innen, egal ob diese privatwirtschaftlich, gemeinnützig, genossenschaftlich oder auch staatlich organisiert sind, kann dies ermöglichen. Allerdings müssen diesbezüglich verschiedene Lösungsansätze in den unterschiedlichen Bereichen verfolgt werden.

5.3.4 Datenbewusstsein und Datenkompetenz

Ob bei der Generierung, Nutzung oder Vermarktung: Daten nehmen inzwischen einen hohen Stellenwert im wirtschaftlichen Tätigkeitsfeld ein. Das Teilen bzw. Nachnutzen von Daten hat durch das große Wertschöpfungspotential beim Datenteilen bereits stark an Bedeutung gewonnen. Dem steht gegenüber, dass in vielen großen Unternehmen großes wirtschaftliches Potential durch das Nicht-Teilen von Daten brach liegt. Dies steht nicht zwangsweise in Zusammenhang mit generellen Aversionen gegen das Datenteilen, welche beispielsweise aus Wettbewerbserwartungen und Monopolansprüchen gegenüber generierten Daten resultieren. Vielmehr weist es auf einen erhöhten Bedarf an Kompetenzen und Strategien zur Erweiterung der Beurteilungsfähigkeiten von Möglichkeiten für die Nutzung eben jener vorliegenden Daten hin. Dies beinhaltet vorweg das Entgegenwirken der oft vorherrschenden Orientierungslosigkeit zwischen dem schlichten öffentlich Zugänglichmachen von Daten sowie der öffentlichen Bereitstellung von verwertbaren Daten bzw. deren Verwertung. Zusätzlich gilt es, Kompetenzlücken zwischen der Forschung und Wirtschaft zu füllen. Die Nachnutzung von Know-how und Daten aus der Forschung scheidert oft am Mangel von vermittelnden Stellen für die Datenaufbereitung bzw.

-auswertung. Der Kompetenzaufbau eben jener vermittelnden Stellen (vgl. Kapitel 4.3.2; Datenintermediär) kann beispielsweise durch Ausgründungen aus der Wissenschaft entstehen, bedarf allerdings der Entwicklung neuer Geschäftsmodelle und deren gezielter Förderung.

Hinsichtlich der rechtlichen Rahmenbedingungen sowie dem Verhältnis von Offenheit und Transfer gibt es im Kontext der Förderung der Nachnutzung von Daten bei allen beteiligten Akteur*innen erhebliche Unsicherheiten. Darauf sollte ebenfalls mit Kompetenzaufbau reagiert werden. Auch innerhalb der Strukturen von industriellen Unternehmen braucht es zukünftig vermehrt gezieltes Change-Management, um bessere Kompetenzen beim Thema Datenteilen zu erlangen. Bisher stehen vor allem Droh-Szenarien wie Wertschätzungsverluste einzelner Personen sowie wirtschaftliche Verluste oder gar Kund*innenabwanderung im Vordergrund.

Ein erhöhter Bedarf zum Wandel und zu mehr Datenkompetenz besteht in der Form der Datenbereitstellung: Probleme mit Datenzugängen, schlechte Auffindbarkeit, unzureichende Formate und (fehlende) rechtliche Vorgaben erschweren das Nachnutzen von öffentlich zugänglichen Daten. Das betrifft gleichermaßen öffentliche sowie wissenschaftliche Tätigkeitsfelder. Zudem ist zu beobachten, dass der Datentransfer von großen Wirtschaftsunternehmen hin zu kleineren wirtschaftlichen Akteur*innen wie Start-ups, die für die Entwicklung ihrer Produkte auf die Nachnutzung von Daten Dritter angewiesen sind, nach wie vor nur stockend zunimmt. Oft scheitert es hier an fehlenden Geschäftsmodellen oder an der Monopolisierung und Marktmacht einzelner Unternehmen (z. B. Amazon und Google). Mit einer Open Data Policy könnten Start-ups gezielter gefördert werden.

In der Wirtschaft bedarf es in allen Unternehmen (von Start-ups bis zu Großkonzernen) eines Kulturwandels beim Datenteilen. Die Incentivierung des Datenteilens durch passende Geschäftsmodelle ist oft zu einseitig gedacht. Die Optimierung eigener Verfahren durch das Teilen und Nachnutzen von Daten ist perspektivisch. Dies kann die teilweise bestehenden starken Abhängigkeiten der Weltwirtschaft von wenigen oder in einzelnen Staaten angesiedelten Konzernen reduzieren und die Fähigkeit erhöhen, wirtschaftlich steuerungs-fähig zu bleiben. Wirtschaftliche Full-Stop-Situationen können durch mangelnde Datenqualität und fehlende Informationsketten entstehen (z. B. Suezkanal Blockade). Demnach ist nicht die Frage, ob Daten geteilt und nachgenutzt werden sollten, sondern vielmehr mit wem und in welchem Umfang. Zudem gibt es interessante Modelle, mit denen versucht wird, dies zu quantifizieren und zu ökonomisieren. Was soviel heißt wie: Wer seine/ihre Daten teilt, darf auch die Daten der Anderen nutzen.

Bei der Förderung der Nachnutzung von Daten müssen ethische Aspekte stärker mitberücksichtigt werden. Das öffentlich Zugänglichmachen von Daten kann keiner generalisierten Strategie folgen, sondern fordert kontextabhängige Entscheidungen. In diesem Zusammenhang bedarf das Mantra „As open as possible. As closed as necessary.“ dringend einer Elaboration.

5.3.5 Datenqualität und Standards

Insgesamt wurde sehr deutlich, dass die Realisierung einer potentiell möglichen Nachnutzung von Daten durch Unternehmen stark davon abhängt, ob diese Nachnutzung leicht realisierbar ist. Anders ausgedrückt: Für viele Unternehmen ist es wenig attraktiv oder unmöglich, Daten nachzunutzen, die vor ihrer Nutzung einer umfangreichen Aufbereitung bedürfen. Dies bedeutet, dass Daten aus Forschung und öffentlicher Verwaltung von Unternehmen dann nachgenutzt werden, wenn sie in für diese Unternehmen leicht verwertbarer Form vorliegen. Derzeit greifen viele Policies zur Förderung der Nachnutzung von Daten diesbezüglich zu kurz. Daher sollten Maßnahmen zur Förderung der Nachnutzung von Daten eine Förderung kund*innenorientierter Datenkuratierung einschließen. Bei der Bewertung wissenschaftlicher Leistung sollte die Erzeugung entsprechend aufgewerteter Datensätze stärker honoriert werden. Dies würde eine Verstärkung der Nachnutzung der Daten bewirken und die Entwicklung neuer Geschäftsfelder beispielsweise durch Ausgründungen fördern.

Datenstandards zur Nachnutzung von Daten sind im wirtschaftlichen Tätigkeitsfeld individuell definiert und abhängig von dem zugrunde liegenden Geschäftsmodell. Die FAIR-Prinzipien, die bei der Generierung, Aufbereitung, Veröffentlichung und Nachnutzung von Daten als Standard in der Wissenschaft fungieren sollten, spielen hier eine untergeordnete Rolle. Es steht vielmehr das gewinnbringende Konzept im Vordergrund, das sowohl die Form der Datengenerierung als auch die Art der Datenbearbeitung bestimmt und somit die Daten gezielt gewinnbringend einsetzbar macht. Der Anreiz zur Datengenerierung im wissenschaftlichen Tätigkeitsfeld wird vorwiegend gesteuert durch Publikationsfaktoren (h-Index). Dies resultiert zumeist in der Erhöhung der Quantität der Publikationen und nicht in der Qualität bzw. Interoperabilität der Daten. Eine Etablierung eines Wiederverwendungs-Index für wissenschaftliche Daten und Forschungsorganisationen würde die Nachnutzbarkeit mehr in den Fokus rücken. In diesem Zusammenhang sollte die öffentliche Zugänglichmachung von interoperablen Daten durch Forschungsorganisationen finanziell belohnt werden.

Die Förderung der Nachnutzung von Daten sollte sich aber nicht nur auf die Zugänglichmachung von Daten

konzentrieren, sondern vielmehr auf die Identifikation benötigter Daten und konkrete Szenarien, in welchen Datenanbieter*innen und Datennutzer*innen mit ihren jeweiligen Interessen bekannt sind. Zudem gilt es, den tatsächlichen Bedarf zur Förderung des Teilens von Forschungsdaten in die Wirtschaft zu prüfen. Es gibt bereits Branchen, die Modellprojekte entwickeln (Catena-X im Automobilbereich, Datenraum Mobilität, Agregaia in der Landwirtschaft, Financial Big Data Cluster (vgl. Europäische Kommission, 2020a)). Die Hoffnungen, die in den Nutzen unspezifischer Daten für die Industrie gesetzt werden, sind überzogen. Um die Nachnutzung von Daten zu erhöhen bedarf es insbesondere Standards, die dabei helfen, Daten interoperabel und nachnutzbar zu machen. Sie müssen jedoch so flexibel gehalten werden, dass verschiedene Variationen des Datenteilens möglich sind. Zusammen mit rechtlichen Rahmenbedingungen sollten diese Standards als Basis für den Aufbau digitaler Datenräume dienen.

5.3.6 Vernetzung

Das Interesse an der Vernetzung im Bereich des digitalen Datenaustauschs ist generell gegeben. Allgemein wird festgestellt, dass Ansprüche an den Umgang mit Daten, die im wirtschaftlichen Umfeld ganz anders einzuschätzen sind als im wissenschaftlichen oder administrativen Umfeld, zumeist eine effektive Kooperation behindern. Grund dafür ist oftmals, dass den Daten unterschiedliche Erwartungshaltungen zu Grunde liegen: Die Incentivierung des Datenteilens in Form gewinnbringender Geschäftsmodelle in wirtschaftlichen Tätigkeitsfeldern steht hierbei dem Open-Data-Ansatz gegenüber. Auch die unterschiedlichen Anforderungen an die Qualität der Daten bzw. deren Interoperabilität erschweren die Zusammenarbeit.

Insbesondere bei der Kooperation zwischen privater Digitalwirtschaft und öffentlicher Verwaltung sollten Anforderungen und Gesetze, die eine Kooperation erschweren oder verhindern, geprüft und überarbeitet werden. Damit ist auch gemeint, dass Konzeptionen entwickelt werden müssen, in denen eigene Gestaltungsräume möglich bleiben. Hierbei sollten vielmehr kontextabhängige Entscheidungen getroffen und generalisierte Regularien abgeschafft werden. Dies bedeutet beispielsweise für den öffentlichen Bereich, dass Daten, die schützenswert sind, nicht geteilt werden dürfen; das Teilen der übrigen Daten aber gefördert werden sollte. Estland hat hier mit X-Road sehr schnell ein sehr effizientes System zum Datenteilen zwischen behördlichen Ebenen ermöglicht. In Deutschland fehlt es in den öffentlichen Bereichen jedoch meist an technischen Voraussetzungen und Kompetenzen für eine innerbehördliche Vernetzung.

Im Allgemeinen gestaltet sich die Zusammenarbeit mit der Wissenschaft einfacher. Es gibt Projekte, in denen Daten in der Wissenschaft generiert wurden, um in der Industrie Anwendung zu finden. Wenn frühzeitig die Anforderungen an die Daten definiert werden, resultieren sie meist in gewinnbringenden Ergebnissen. Defizite in der interdisziplinären Kommunikation, die steigende Tendenz in der Forschung, eine immer größere Hürde bei dem Thema Intellectual Property zu schaffen und das Zurückhalten von Know-how in der Wissenschaft erschweren aber auch mit diesem Bereich die Zusammenarbeit und den Datentransfer. Vor allem Know-how aus der Wissenschaft sollte der Privatwirtschaft früher zugänglich gemacht werden. Zu oft streben Wissenschaftsorganisationen Produktentwicklungen an, ohne sie abschließen zu können. Daher sollten Wissenschaftsorganisationen ihre Technologietransfer-Strategien überprüfen. Das betont auch Andreas Weiss: „Was mich wirklich ärgert: Wir haben exzellente Forschung und andere verwerten das. Weil sie einfach flinker sind, weil sie weniger reguliert sind, weil sie nicht stundenlang oder tagelang oder wochenlang über IPs streiten, sondern sagen: „Na ja, das habt ihr ja soweit klar gemacht, das ist ja da öffentlich, dann nutzen wir das mal.“

Direkte Intellectual Property Rights fehlen im Datenbereich und erschweren die Zusammenarbeit. Die Lizenzierung von Daten mithilfe eines standardisierten und global verständlichen Sets von die Daten selbst betreffenden „Erlaubnis- und Verbots-Modulen“, aber gegebenenfalls ebenfalls betroffener Kontexte wie beispielsweise Datenbankrechte, Software-Lizenzen, Persönlichkeitsrechte etc. bedürfen gezielter Förderung. In diesem Zuge sollte die Freigabe zur Nachnutzung von Daten klar lizenziert werden, damit eindeutig kommuniziert wird, was bzw. wer etwas mit den Daten machen darf.

5.3.7 Infrastruktur

Deutschland und Europa benötigen vertrauenswürdige Datenräume. Diese werden gebraucht, weil viele Nutzungsszenarien für Daten nicht von einzelnen Datenblättern autonom realisiert werden können. Das gilt insbesondere für die Nutzung sensibler Daten. Hier bedarf es der Förderung von Projekten, in denen die europäischen Werte technisch umgesetzt werden. Die Eckpfeiler für eben jene vertrauenswürdigen digitalen Verfahren sind somit Datenschutz, Systemicherheit und Identitäten.

Die Datentransfer-Plattform Gaia-X ist weitgehend als eine innovative, konzeptionell richtige und visionäre Basisantwort akzeptiert. Anders als im Gesamtkonzept gedacht, nämlich dass in Gaia-X eine zentralisierte Verarbeitung von Daten zugrunde liegt, entstehen die Daten in Gaia-X immer noch primär dezentral

in der Edge. Als Gründe werden unter anderem Kosten und Datenlogistik angegeben. Die Herausforderung liegt hier darin, eine Datenverarbeitung zu schaffen, die über Provider- und Entitäten-Grenzen hinaus geht, und dafür braucht es klar definierte Nutzungsbedingungen. Diese fehlen und sollen verstärkt gefördert werden, um den Aufbau von Datenräumen nach gleichen Regeln und interoperablen Standards zu ermöglichen. Die Schwierigkeit besteht darin, eine Konzeption zu entwickeln, in der eigene, selbstverantwortliche Gestaltungsräume umsetzbar sind. Ein einheitliches Verfahren für alle Partikularinteressen kann es nicht geben, doch die Formulierung einheitlicher Regularien sollte bei der Art der Entstehung der Daten (Cloud vs. Edge) beginnen und muss den Umfang von Daten wie auch die Weitergabe von Domainwissen miteinschließen.

Auch andere Datentransferplattformen (z. B. Cate-na-X im Automobilbereich und Datenraum Mobilität, Agregai in der Landwirtschaft, Financial Big Data Cluster) haben sich als interessenbezogene Gruppen nach Gaia-X-Konzeptionen voran bewegt. Genauso wie bei Gaia-X liegt hier die Herausforderung in der Feinjustierung, um passende Lösungsmuster zu finden, höchstwahrscheinlich auch durch das Einbeziehen von Know-how aus anderen Domänen.

Bei der Gestaltung der in Europa angestrebten digitalen Souveränität ist darauf zu achten, der internationalen Kooperationsfähigkeit, die Voraussetzung für den internationalen Handel ist, nicht zu schaden. Hierbei sollte die Realitätstauglichkeit dieser Zielvorstellung überprüft bzw. die Vision auf realistische Ziele beschränkt werden. Zudem muss Europa eine an den eigenen Werten orientierte Wirtschaftsordnung und damit korrespondierende Wertschöpfungsketten entwickeln, statt eine Kopie der amerikanisch oder chinesisches vorherrschenden Plattformökonomie zu fördern. Daher sollte bei der Förderung auch eine Kommunikationsebene zwischen Datenerzeugern*innen und Nachutzern*innen eingeschlossen sein.

5.4 Qualifizierung

5.4.1 Einstieg

Für den Bereich Qualifizierung wurden fünf Expert*innen interviewt, um einen Blick auf aktuelle Tendenzen und Wahrnehmungen des Standes der Qualifizierung in der Praxis zu werfen. Da es sich um ein Querschnittsthema handelt, wurden Vertreter*innen für die Bereiche *Wissenschaft*, *Wirtschaft* und dem *Public Sector* ausgewählt. Die befragten Personen bringen neben ihren fachlich breitgefächerten Hintergründen (Informatik, Politik- und Sozialwissenschaften, Wirtschaft, Verwaltungswissenschaften) die

gewünschte Qualifizierungsperspektive mit, da sie im Alltag an der Hochschule, in der Beratung zum Umgang mit Daten, der Verbandsarbeit und Weiterbildung tätig sind. Für die individuellen Interviews wurden vom Interviewleitfaden vor allem die Fragen zu Qualifizierungsaspekten ausgewählt. Darüber hinaus flossen die Aussagen zum Thema Qualifizierung von weiteren 13 Expert*innen-Interviews aus den anderen Untersuchungsbereichen *Wissenschaft*, *Public Sector* und *Wirtschaft* in die Auswertung ein.

Die Darstellung der Interviewergebnisse erfolgt zusammenfassend und in anonymisierter Form. Sofern eine Zuordnung von speziellen Aussagen zu einem der Bereiche *Wissenschaft*, *Public Sector* und *Wirtschaft* möglich war, wurde dies gemacht, um wahrgenommene Unterschiede für die Bereiche deutlich zu machen. Die inhaltliche Zuordnung der Themen orientiert sich in der Abfolge der Themen an Kapitel 4.4 Qualifizierung.

5.4.2 Rahmenbedingungen und Infrastruktur

Auch für den Untersuchungsbereich *Qualifizierung* sind Rahmenbedingungen und Infrastruktur ein wichtiges Thema. Sie beeinflussen maßgeblich, wo und wie der Umgang mit Daten in der Praxis aussieht und damit mögliche Anknüpfungspunkte für das Thema Qualifizierung.

Von wissenschaftlich tätigen Expert*innen wurde deutlich gemacht, dass sie eine Änderung des Wissenschaftssystems für erforderlich halten, um den Umgang mit Daten zu verbessern. So sollte die Anerkennung von Tätigkeiten im Zusammenhang von FDM, ob in der Ausbildung, der Anwendung oder der Veröffentlichung von Daten eine Rolle spielen. Als mögliche Beispiele wurde die Berücksichtigung von FDM bei Berufungen auf Professuren, ein „Impact Factor“ für Datenveröffentlichung oder auch die Vereinbarung von FDM-Umsetzungen in Zielvereinbarungen genannt. Eine andere Möglichkeit wurde in verpflichtenden Vorgaben, zum Beispiel bei öffentlich finanzierter Forschung in Form von Projekten zur Erstellung eines Datenmanagementplans (DMP), der Datendokumentation, Datentransparenz und der Datenbereitstellung für Dritte gesehen.

Die Schaffung von zentralen Anlaufstellen mit Datenkompetenz und die verpflichtende Benennung von Beauftragten für Daten wurde vor allem für den *Public Sector* als wichtige und hilfreiche Maßnahme gesehen. Dies sollte mit der erforderlichen, dauerhaften Finanzierung und dem notwendigen zeitlichen Umfang der eingerichteten Stellen einhergehen und Mindestanforderungen für die Datenkompetenzen des entsprechenden Personals beinhalten.

Aufgrund der schnellen Entwicklungen im Bereich Forschungsdaten verändern sich die Infrastrukturen

und die Anforderungen an entsprechende Qualifizierungen in diesem Bereich sehr schnell. Es wurde vorgeschlagen, Bottom-up-Initiativen mit neuen Ideen, die sich etwa im Bereich von (qualifizierenden) Supportstrukturen bedarfsgerecht entwickeln, nicht durch organisatorische Vorgaben zu sehr einzuschränken oder gar zu verhindern, sondern großzügig finanziell und organisatorisch zu unterstützen. Als wichtig in diesem Zusammenhang wurde auch die Förderung der Vernetzung solcher Initiativen genannt.

5.4.3 Datenbewusstsein und Kulturwandel

Aus den Interviews wurde deutlich, dass der digitale Wandel und der damit verbundene Kulturwandel in der Gesellschaft noch nicht angekommen ist. Das betrifft insbesondere das Thema Datenbewusstsein.

Als wichtig für eine neue „Datenkultur“ wird das Verständnis von Daten und der informierte Umgang mit Daten gesehen. Um diesen Kulturwandel zu erreichen, wurde vorgeschlagen, Data Literacy, d. h. Kompetenz im Umgang mit Daten, als gesellschaftliches Ziel zu definieren. In diesem Zusammenhang wurde auf die Data-Literacy-Charta⁶ von Vertreter*innen der Wissenschaft, der Wirtschaft und des *Public Sectors* hingewiesen. Dabei wurde Data Literacy als eine Grundkompetenz für die gesellschaftliche und berufliche Teilhabe aller gesehen. Mehrere Interviewpartner*innen schlugen vor, Data Literacy möglichst breit im Kontext der Qualifizierung zu vermitteln.

In allen Untersuchungsbereichen wurde ein hoher Bedarf für ein neues Datenbewusstsein festgestellt. Dabei wurden jeweils unterschiedliche Beispiele genannt, die einer Nachnutzung von Daten entgegenstehen. So ist die Haltung „das sind meine Daten“ in der Wissenschaft noch sehr verbreitet. Im *Public Sector* sind die Angst vor „Fehlinterpretation“ durch Nutzer*innen bzw. vor „rechtlichen Fehlern bei der Daten-Veröffentlichung“ häufig Hinderungsgründe für eine Veröffentlichung und damit mögliche Nachnutzung. Im Bereich Wirtschaft wird oft fehlende „Wertschätzung von Datenkompetenz“ erkannt. Die Erkenntnis, dass Datenteile „eine Aufgabe der öffentlichen Verwaltung ist“, ist noch vergleichsweise wenig verbreitet. Von Vertreter*innen des *Public Sectors* wurde in diesem Zusammenhang empfohlen, mögliche Mehrwerte einer Datenveröffentlichung anhand von Beispielen etwa für eine bessere Auffindbarkeit oder eine mögliche Nachnutzung der Daten durch die erfolgreiche Verknüpfung von einzelnen Datensätzen aufzuzeigen. In der Wissenschaft werden Daten teilweise noch fragmentiert in Fachdisziplinen gedacht und „Silodenken“ bzw. die fehlende Sicht von

Zusammenhängen stehen einer (interdisziplinären) Zusammenarbeit entgegen.

Für den Kulturwandel ist ein Grundverständnis von Daten erforderlich. Daher sollte in allen Bereichen ein Basiswissen über Daten vermittelt werden. Für den *Public Sector* gibt es mehrfach den Vorschlag, dies bereits bei der Ausbildung der Verwaltungsmitarbeiter*innen zu vermitteln und zwar unabhängig davon, ob sie aktiv mit Daten arbeiten oder nicht. Hierbei sollen Ängste vor Fehlern genommen, Mehrwerte aufgezeigt und so Sicherheit im Umgang mit Daten geschaffen werden. Ein wichtiger Beitrag für den Kulturwandel ist laut Expert*innen, bereits (früh) in Ausbildung und Studium ein Grundverständnis für den Umgang mit Daten zu vermitteln.

Eine hohe Bedeutung für den angestrebten Kulturwandel wird Vorbildern beigemessen. Als Beispiel für den Bereich Wissenschaft wurden Lehrende genannt, die den Umgang mit Daten nicht nur theoretisch vermitteln, sondern in der eigenen Forschung für alle Studierenden erkennbar aktiv umsetzen. In der Wirtschaft und in der öffentlichen Verwaltung werden besonders die Vorgesetzten in der Verantwortung gesehen, die angestrebte Kultur vorzuleben.

Von den Expert*innen wurde die Wichtigkeit eines Datenbewusstseins bei der Vermittlung von technischem und fachlichem Grundwissen über Daten angeführt und als Blick für Zusammenhänge, Mehrwerte und mögliche Verknüpfungen charakterisiert.

5.4.4 Kompetenzprofile, Fähigkeiten und Skills

Trotz der verbreiteten Forderung nach Data Literacy, d. h. nach Kompetenz im Umgang mit Daten und der Forderung nach der Vermittlung von entsprechenden Grundkenntnissen, wurde in den Interviews deutlich, dass es aktuell nur Ansätze für Kompetenzprofile gibt. Es fehlt an Zusammenstellungen von konkret erforderlichen Fähigkeiten und Kenntnissen für die einzelnen Bereiche. So wurde auf entsprechende Fragen auf der einen Seite sehr allgemein mit „alle erforderlichen Kenntnisse im Umgang mit Daten, um beraten und unterstützen zu können“ oder die „FAIR-Prinzipien ergänzt um die CARE-Prinzipien decken das gut ab“ geantwortet. Die CARE-Prinzipien⁷ wurden von der Global Indigenous Data Alliance (GIDA) 2019 als Ergänzung der FAIR-Prinzipien veröffentlicht, mit dem Ziel, die Rechte und Interessen der indigenen Völker bei gemeinsamer Datennutzung stärker zu berücksichtigen. Auf der anderen Seite wurden in den Antworten sehr detaillierte Einzelaspekte angeführt. Die Auflistung reicht von „entscheiden können, welche

⁶ <https://www.stifterverband.org/charta-data-literacy>

⁷ CARE steht für Collective Benefit, Authority to Control, Responsibility und Ethics. Vgl. Carroll et al., 2019

Daten gelöscht werden können“ über „wissen was Metadaten“ sind und welche „fachspezifischen Repositorien“ es gibt bis hin zu „rechtliche Aspekte“ kennen, „Qualität von Datensätzen beurteilen“ können und vielen anderen Dingen.

Sehr deutlich wurde, dass es sehr unterschiedliche Zielgruppen gibt und sich die Anforderungen, d.h. mögliche Kompetenzprofile dementsprechend unterscheiden. Von mehreren Expert*innen wurde eine gewisse „Data Literacy“ der gesamten Bevölkerung als wichtiges Element für die gesellschaftliche Teilhabe gefordert. Die Data-Literacy-Charta (Quelle) wurde als wichtiger Anhaltspunkt genannt ohne die konkreten Inhalte zu benennen. Data Literacy soll helfen „die Bedeutung von Daten im Alltag“ zu erkennen. Sie muss ein Grundwissen beinhalten, um zu erkennen, was Daten sind, wie sie aussehen und was für Risiken (Thema Datenschutz) damit verbunden sein können. Außerdem muss bekannt sein, wie die eigenen (persönlichen) Daten verwendet werden, um mit diesem Grundverständnis an (gesellschaftlichen) Diskussionen teilnehmen zu können.

Für die Praxis werden Kompetenzen und Kenntnisse gefordert, die zum souveränen Umgang mit Daten an der jeweiligen Stelle befähigen. Für alle Verwaltungsmitarbeiter*innen, die mit Daten umgehen müssen, sind das beispielsweise Grundkenntnisse dazu, was ein Datensatz ist und dass er mit anderen Datensätzen kombiniert werden kann, um so Mehrwerte zu schaffen. Darüber hinaus sollte bekannt sein, dass Excel-Tabellen nicht maschinenlesbar sind, dass rechtliche Aspekte beachtet werden müssen und dass eine Veröffentlichung unter Beachtung von klaren Vorgaben möglich ist. Der grundlegende Gedanke ist hier ein Bewusstsein für die Daten zu schaffen. Für Spezialist*innen, die sich um die technische Umsetzung vor Ort kümmern, werden Kenntnisse zu Schnittstellen, zum Datenexport und zur Datenbereitstellung, d.h. alle technischen, aber auch rechtlichen und organisatorischen Aspekte gefordert.

Vertreter*innen der Wirtschaft sehen die Anforderung an die Kompetenzen von Mitarbeiter*innen eher bei Aspekten der Nutzung von Daten als in der reinen Bereitstellung. Bei allen müssen die notwendigen Kompetenzen für die Anwendung vorhanden sein, um die für ihren Bereich wichtigen Daten aktiv verwenden zu können. Unterstützt werden sie von technisch versierten Spezialist*innen, welche die Einbindung und Verfügbarkeit von Daten in Programmen und Anwendungen sicherstellen.

Bei Forschenden im Bereich Wissenschaft müssen grundlegende Kenntnisse zum Umgang mit Daten mit den jeweiligen fachspezifischen Ausprägungen kombiniert sein. Sie können durch Datenspezialist*innen unterstützt werden, die z. B. zu rechtlichen und ethischen Aspekten, zur Auffindbarkeit und Recherche

von externen Daten beraten oder sich um die technischen Umsetzungen kümmern.

Für die Qualifizierung bedeutet das, dass die Anforderungen von Expert*innen aus der jeweiligen Praxis kommen müssen. Für die Wissenschaft wurde beispielsweise für den Hochschulbereich vorgeschlagen, dass Empfehlungen von Kompetenzkatalogen für die Inhalte eines FDM-Kerncurriculum von den entsprechenden wissenschaftlichen Fachgesellschaften kommen sollen. Darüber hinaus wurde empfohlen, sich bei den Inhalten daran zu orientieren, was erforderlich ist, um Daten FAIR zu machen und zugleich ethische Aspekte der CARE-Prinzipien zu berücksichtigen.

Sowohl von Vertreter*innen der Wirtschaft als auch des *Public Sectors* wurde neben der reinen Datenkompetenz auch die Berücksichtigung von Kompetenzen zu Organisation und Change-Management in Kompetenzprofilen und bei der Vermittlung als besonders wichtig hervorgehoben. Ergänzend wurden Kenntnisse in den Bereichen Ethik, Recht und IT genannt.

Mehrfach wurde in den Interviews deutlich, dass eine Übersicht über alle Angebote zur Förderung der Datenkompetenz fehlt. Gewünscht wurde hier eine am Bedarf orientierte Detaillierung nach Fach- und Anwendungsgebieten und die Unterteilung in Aus- und Weiterbildungsangebote. Hinweise zur Einordnung der Qualität der Angebote wurden ebenfalls gewünscht, aber kurzfristig nicht als realistisch eingeschätzt.

5.4.5 Medienkompetenzvermittlung und Qualifikationswege

Von allen Interviewten wird Datenkompetenz, d.h. der kompetente Umgang mit Daten in den einzelnen Bereichen als wichtig erachtet und einer entsprechenden Qualifizierung der Mitarbeitenden ein hoher Stellenwert beigemessen. Sehr deutlich wurde in den Interviews geäußert, dass diese Qualifizierung möglichst früh erfolgen sollte und die gesamte Bevölkerung Grundkenntnisse zum Umgang mit Daten – auch für den Alltag – in unserer mehr und mehr digitalisierten Welt braucht.

Die Vermittlung von digitalen Kompetenzen bereits in der Grundschule wurde von einem Vertreter aus der Wirtschaft explizit gefordert. Alle Expert*innen sind sich einig, dass eine erste, grundlegende Medienkompetenzvermittlung bereits in der Schule erfolgen muss. Darauf aufbauend sollten diese Datenkompetenzen in der Ausbildung für die öffentliche Verwaltung – so die Forderung aus der Praxis – ergänzt werden. Dies gilt auch für Ausbildungen im Bereich Wirtschaft, hier insbesondere mit dem Fokus auf der Nutzung von Daten. Hierfür ist die Ergänzung von entsprechenden Inhalten in den Ausbildungen erforderlich.

Der Erwerb von FDM-Grundkenntnissen bereits im Grundstudium in allen Studienfächern ist Konsens

und sollte nicht erst auf dem Promovierenden-Level erfolgen. Curricula sind entsprechend anzupassen. Das könnte über die Integration in bestehende Kurse, neue Angebote oder als Bestandteil von grundlegenden „Schlüsselqualifikationen“ geschehen, so eine Vertreterin aus dem Hochschulbereich. Generische Angebote werden als wichtig erachtet, aber auch mehrfach auf eine große Lücke im Bereich der Vermittlung von fachdisziplinspezifischem FDM hingewiesen.

Nicht erst die Corona-Pandemie habe nochmals deutliche Defizite bei der Digitalisierung in der öffentlichen Verwaltung aufgezeigt, so ein Vertreter des *Public Sectors*. Expert*innen für diesen Bereich sind sich einig, dass es auch Schulungen und Weiterbildungen zu Grundlagen der Datenkompetenz für alle Verwaltungsmitarbeiter*innen geben muss, die diese Kenntnisse nicht bereits in ihrer Ausbildung erworben haben. Das wird für die Bereiche Wissenschaft und Wirtschaft ähnlich gesehen. Einigkeit herrscht auch darüber, dass entsprechende Schulungen bedarfsorientiert für die jeweilige Position erfolgen müssen. Als wichtig wird hierbei die Sicherstellung der Qualität der Fortbildungsanbieter*innen und ihrer Angebote, etwa über eine Zertifizierung von geeigneter Stelle, erachtet. Von mehreren Expert*innen wurde der aktuelle Wildwuchs und die Unübersichtlichkeit der Angebote bemängelt und mehrfach der Wunsch nach einer qualitätsgeprüften Übersicht von zentraler Stelle über Weiterqualifizierungsangebote zu FDM geäußert.

Die Expert*innen berichteten über den erfolgreichen Erfahrungsaustausch und die gute Zusammenarbeit in unterschiedlichen Kooperationen, Netzwerken und „Runden Tischen“ besonders in den Bereichen Wissenschaft und *Public Sector*. Konkret ging es hier um die Identifikation von Inhalten für mögliche Workshops und die Erfahrungen mit Online- und Präsenzkursen sowie die Nachnutzung von Erfahrungen und Angeboten. Darüber hinaus wurde von guten Erfahrungen mit zentralen Beratungsstellen zum Thema Umgang mit Daten bzw. Open Data berichtet, die in einzelnen Einrichtungen bereits zu einem stärkeren Bewusstsein für und professionelleren Umgang mit Daten geführt haben.

Als wichtiges Thema wird auch die Weiterbildung von Ausbildungs- und Lehrkräften gesehen. Gerade in der professionellen Ausbildung an Hochschulen sollte sichergestellt werden, dass auch langjährige Lehrkräfte stets auf dem aktuellen Stand im Umgang mit Daten sind, um diesen an ihre Studierenden vermitteln und auch als Vorbild fungieren zu können. Gefördert werden könnte dies beispielsweise über Zielvereinbarungen oder die Aufnahme von Angeboten zu Datenkompetenz in die Standardkataloge für Weiterqualifizierungsmaßnahmen.

5.4.6 Berufsbilder und Karrierewege

Der Bedarf an professioneller Datenkompetenz steigt in den Bereichen Wissenschaft, *Public Sector* und Wirtschaft im Zuge der fortschreitenden Digitalisierung und mit den damit verbundenen Anforderungen stetig. Neben der Weiterqualifizierung des vorhandenen Personals werden Expert*innen zu allen Aspekten im Umgang mit Daten für die Beratung, Organisation und Koordination einerseits sowie die technisch-organisatorische Umsetzung und den Support andererseits gesucht. Im Zuge des verstärkten Aufbaus einer technisch-organisatorischen Infrastruktur werden ursprünglich zeitlich eingeschränkte – zunächst nur als Zusatzaufgaben vergebene – Tätigkeiten zu vollzeitlichen Daueraufgaben, die von entsprechend qualifiziertem Personal mit neuen Tätigkeitsbeschreibungen erledigt werden müssen.

In öffentlichen Verwaltungen geht es vorrangig darum, der gesetzlich verankerten Pflicht zur Veröffentlichung von Daten (= open data) nachzukommen, in der Wirtschaft hauptsächlich um die Verwendung eigener oder fremder Daten verbunden mit ökonomischen Interessen und weniger um die Weitergabe eigener Daten an Dritte. In der Wissenschaft werden fremde Daten nachgenutzt und eigene Daten zur Ermöglichung der Prüfung von Forschungsergebnissen, dem Austausch mit Fachkolleg*innen bzw. aufgrund von Förderauflagen bereitgestellt. Dementsprechend divergieren die Anforderungen an die gesuchten Expert*innen hinsichtlich der Kompetenzen, die über die reine Datenkompetenz hinausgehen. In der öffentlichen Verwaltung geht es um die technisch-organisatorische, rechtssichere Datenbereitstellung. Dementsprechend wurde von den Interviewten mehrfach die Notwendigkeit der Benennung eines „Open Data Beauftragten“, bzw. eines „Data Stewards“ als Berater*in und Ansprechpartner*in, welche die verantwortliche Person bei der Entscheidungsfindung unterstützen, wenn es um die Veröffentlichung von Daten geht, gewünscht. Aus der Praxis einzelner Behörden wurde ergänzend der „Chief Digital Officer“ als Ansprechpartner für die Koordination von Open Data auf Ressortebene erwähnt. Für die Wirtschaft steht die Nutzung der Daten im Vordergrund und dementsprechend wurden als Berufsbezeichnungen mehrfach „Data Analyst“ und „Data Manager“ genannt. In der Wissenschaft wurden als aufkommende Berufsbezeichnungen „Data Steward“ – als Unterstützung im Umgang mit Daten – und „Data Scientist“ – mit fachwissenschaftlichem Hintergrund und Datenzusatzqualifikation – identifiziert. Der „Data Curator“ wurde in den verschiedenen Bereichen jedoch jeweils mit anderen Aufgabenschwerpunkten genannt.

Es herrschte Konsens, dass neue Berufsfelder im Bereich Daten noch in der Entwicklung sind. So seien

aufgrund ihrer unscharfen Profile, etwa bei Stellenanzeigen, nach wie vor die geforderten Datenkompetenzen von höherer Bedeutung als eine angeführte „neue“ Berufsbezeichnung.

Ergänzend und alternativ sind – sowohl in der öffentlichen Verwaltung als auch in der Wissenschaft – Kompetenzteams verbreitet, die beim Thema Daten

und Datenmanagement unterstützen und beraten. Die einzelnen Mitglieder eines „Data Science Teams“ oder „Team Datenmanagement“ bringen jeweils ihre spezielle Expertise ein. Skeptisch wurde gesehen, wenn den Mitarbeitenden das Thema Daten und Datenmanagement als Zusatzaufgabe on-top ohne entsprechendes Zeitbudget auferlegt wird.

BEREICHS- SPEZIFISCHE EMPFEHLUNGEN



6. Bereichsspezifische EcoDM-Empfehlungen

Dieses Kapitel beinhaltet die abgeleiteten bereichsspezifischen (Handlungs-)Empfehlungen zur Steigerung einer übergreifenden FAIRen Datenproduktion und -nachnutzung sowie Vernetzung der identifizierten Untersuchungsbereiche. Die bereichsspezifischen Empfehlungen richten sich an Akteur*innen und -gruppen der jeweiligen Bereiche sowie an verschiedene Stakeholder*innen. Sie leiten sich von der Auswertung der literatur-basierten Landscape- und Gap-Analysen (vgl. Kapitel 4) sowie den Ergebnissen aus den Expert*innen-Interviews (vgl. Kapitel 5) ab. In den vier Untersuchungsbereichen – *Wissenschaft*, *Public Sector*, *Wirtschaft*, *Qualifizierung* – gibt es unterschiedlich viele Empfehlungen und es liegen jeweils unterschiedlich große Überschneidungen bei den abgeleiteten Empfehlungen aus den LGA und den Expert*innen-Interviews vor, die im nächsten Kapitel weiter zu bereichsübergreifenden Empfehlungen von EcoDM kondensiert werden (vgl. Kapitel 7). Im Folgenden sind die Empfehlungen nach Kategorien, d.h. nach thematischen Zusammenhängen, aufgeführt.

6.1 Wissenschaft

Es wurden insgesamt 21 Empfehlungen für die Wissenschaft abgeleitet, die sich unter fünf Kategorien zusammenfassen lassen: (A) Vernetzung fördern, (B) Resiliente Infrastrukturen schaffen, (C) Forschungsförderung nachhaltig gestalten, (D) Expertise aufbauen und vermitteln und (E) Kompetenzräume, die Forschende entlasten, realisieren.

A. Vernetzung fördern

1. Forschungsdatenmanagement als gemeinsame Aufgabe erkennen

Die FAIR-Prinzipien sollen sich inter- und transdisziplinär etablieren. Die NFDI versucht, einen Teil zu leisten. Datenmanagement wird als disziplinäre und transdisziplinäre Herausforderung begriffen und gemeinsam im Sinne der Resilienz gelöst.

2. Forschungsk Kooperationen zwischen wissenschaftlichen Institutionen fördern

Forschungsdatenmanagement wird in wissenschaftlichen Einrichtungen unterschiedlich praktiziert. Gibt es Kooperationen zwischen Einrichtungen, kann man gemeinsam die FAIR-Prinzipien anwenden und voneinander lernen. Gerade im Bereich der (technischen) Infrastruktur sind gemeinsame Interessen zu erwarten.

3. Kooperationen mit dem Public Sector nutzen und stärken

Im *Public Sector* werden Daten öffentlich zur Verfügung gestellt, aber auch darüber hinaus gibt es diverse Forschungsk Kooperationen mit diesem Bereich. Diese Beziehungen gilt es zu stärken und proaktiv zu gestalten. Die Expertise zum Datenmanagement im *Public Sector* kommt auch aus der Wissenschaft.

4. Kooperationen mit der Wirtschaft auf Augenhöhe gestalten

Diverse Disziplinen pflegen enge Verbindungen mit der Wirtschaft. Kooperationen mit diesem Bereich sind lehrreich und notwendig. Die Offenlegung der Forschungsdaten aus derartigen Projekten erscheint notwendig, damit die Wissenschaft davon profitieren kann und nicht nur die Wirtschaft.

5. Bürger*innen-Wissenschaften anwenden

Methoden der Citizen Science können als wissenschaftliche Innovation betrachtet werden, die besonders durch die Digitalisierung und der damit verbundenen Vertrauenskrise relevant wird. Die heterogenen Möglichkeiten der Bürger*innen-Wissenschaften können dafür genutzt werden, einen engeren Austausch mit interessierten Bürger*innen zu schaffen.

B. Resiliente Infrastrukturen schaffen

1. System- und Datensicherheit ermöglichen

Das Wissenschaftssystem braucht Infrastrukturen, die Daten- und Systemsicherheit multidimensional gewährleisten können. Vertrauen ist keine Grundlage für wissenschaftliche Erkenntnisse, deshalb braucht es offene Formate und regelmäßige Pen-Tests für Forschungsinfrastrukturen.

2. FAIR-Prinzipien fördern und anwenden

Die FAIR-Prinzipien erleichtern u. a. das Suchen, Teilen und Finden von Forschungsdaten. Diese Prinzipien gilt es in der Praxis anzuwenden, wobei entsprechende infrastrukturelle Anreize gesetzt sein sollten.

3. Online-Kollaboration erleichtern

Forschende arbeiten oft gleichzeitig und online an kollaborativen Projekten. Infrastrukturen sollen so gestaltet sein, dass sie das digitale Arbeiten im Team erleichtern, ohne dabei an Resilienz zu verlieren, sodass Daten- und Systemsicherheit gewährleistet sind.

4. Umweltschonendes Arbeiten fördern

Datenverarbeitung braucht Energie und trägt damit aktiv zur Klimakrise bei. Die ökologischen Aspekte des wissenschaftlichen Arbeitens sollten in den Vordergrund rücken, insbesondere wenn es um die Verarbeitung großer Datenmengen geht. Es braucht klimaschonende Prozesse und Datenminimierung, soweit möglich und sinnvoll.

5. Langfristige Archivierung von Forschungsdaten absichern

Publizierte Forschungsdaten müssen so verwahrt werden, dass ihr Bestand langfristig vor Manipulation, Diebstahl, Löschung (...) gesichert ist, obwohl sie öffentlich zugänglich sind. Durch Versionierungen können Datenbestände oder Erkenntnisse aktualisiert werden. Es braucht resiliente technische Infrastrukturen, um das Wissen langfristig abzusichern.

C. Forschungsförderung nachhaltig gestalten

1. Öffentlich finanzierte Daten sind in der Regel öffentlich zugänglich

Datenveröffentlichungen müssen in Projekten vorgesehen und gewünscht sein. Studien sollen, sofern ethisch möglich, für alle transparent und replizierbar sein. Diese Herausforderung betrifft alle Disziplinen und Förderstellen.

2. In den Vergaberichtlinien werden offenes Arbeiten und Datenveröffentlichungen berücksichtigt

Um die Bereitstellung wissenschaftlicher Daten zu fördern, sollen Datenveröffentlichungen explizit in den Vergaberichtlinien genannt werden. Dabei soll auf offenes Arbeiten, beispielsweise mit Datenmanagementplänen, fokussiert werden.

3. Datenmanagement und Transparenz werden im Budget berücksichtigt

Gutes wissenschaftliches Datenmanagement soll nicht nur gefordert, sondern auch gefördert werden. Kurationsarbeiten für Datenveröffentlichungen, Aktualisierungen oder die Dokumentation von einzelnen Forschungs-Prozessschritten inkl. Metadaten (...) sind zeitaufwendige Unterfangen, die entsprechende Expertise voraussetzen. Es müssen genug budgetäre Mittel vorhanden sein, um entsprechende Arbeit leisten zu können.

4. Nachvollziehbarkeit und Transparenz einfordern

Es ist nicht nur notwendig, dass Nachvollziehbarkeit und Transparenz von Forschungsdaten gefordert und gefördert wird, sondern auch, dass die Daten daraufhin überprüft werden. Es braucht Kontrolle, ob Forschungsdaten nach den FAIR-Prinzipien veröffentlicht wurden.

5. Klare Trennung von Forschungs- und Infrastrukturprojekten

Forschungs- und Infrastrukturprojekte sind in ihrer Funktion grundsätzlich zu unterscheiden. In Forschungsprojekten wird Wissen in Form von Daten generiert. Das erfolgt über einen gewissen Zeitraum oder in Phasen. Infrastrukturprojekte hingegen wollen die Forschungstätigkeit kontinuierlich unterstützen und sind auf Resilienz ausgelegt.

D. Expertise aufbauen und vermitteln

1. Datenmanagement im Curriculum verankern

Die Digitalisierung gestaltet neue Möglichkeiten im Datenmanagement, die auch mit Risiken, Pflichten und Herausforderungen verbunden sind. Forschenden soll schon möglichst früh Wissen zum Datenmanagement vermittelt werden.

2. Fachspezifische Beratungsmöglichkeiten schaffen

Die Methodologie des Datenmanagements ist eine Wissenschaft für sich und es ist Forschenden nicht zumutbar, immer auf dem neuesten Stand zu sein. Es braucht fachspezifische Anlaufstellen, die Forschende beim Thema Datenmanagement disziplinspezifisch unterstützen können, um bessere Datenqualität zu gewährleisten.

3. Weiterbildungsmöglichkeiten etablieren

Es fehlt an Expertise zum Thema Datenmanagement in jeder Qualifikationsstufe. Etablierte Weiterbildungsmöglichkeiten können ein Anreiz sein, sich mehr mit diesem Thema zu befassen.

4. Fachexpertise institutionell nutzen und ausbilden

Nicht jeder Forschende muss Datenexpert*in sein. Es entstehen neue Berufsbilder: Datentreuhänder*innen, Daten-Kurator*innen, ... Dazu braucht es technische und juristische Expertise, um eine resiliente institutionelle Forschungsinfrastruktur zu schaffen. Fachwissen soll umfassend genutzt und ausgebildet werden.

E. Kompetenzräume, die Forschende entlasten, realisieren

1. Daten-Kompetenzräume schaffen und fördern

Es braucht übergreifende Daten-Kompetenzräume, um Wissen und Erfahrungen auszutauschen, Expertise zu sammeln und um Best Practices bereitzustellen. Alle Forschenden sollen die Angebote ohne Zwang nutzen können.

2. Intuitive Datenräume gestalten

Forschende wollen die Forschungsdaten – aus dem gesamten Forschungskreislauf – nach den FAIR-Prinzipien bereitstellen und abrufen. Das

sollte über die technische Infrastruktur intuitiv möglich sein. Also ohne, dass man dafür zusätzliche Expertise benötigt.

3. Klare Anleitungen und Regeln schaffen

Datenräume funktionieren nach klaren, expliziten und transparenten Regeln. Dazu gibt es Anleitungen und fachspezifische Ausführungen, wie der spezifische Raum genutzt werden kann.

6.2 Public Sector

Im *Public Sector* ließen sich 26 Empfehlungen ausarbeiten, die sich in fünf Kategorien unterteilen lassen: (A) Infrastruktur & Ressourcen, (B) Rechtliche Rahmenbedingungen, (C) Qualifizierung, Karrierewege und Kompetenzaufbau, (D) Kulturwandel, (E) Zusammenarbeit & Vernetzung.

A. Infrastrukturen & Ressourcen

1. Einrichtung von Infrastrukturen auf allen föderalen Ebenen

Auf kommunaler Ebene kann die Einrichtung von Data Lakes die dezentrale Verwaltung von Daten vereinfachen. Die Einrichtung von Landesdatenportalen kann die Vereinheitlichung von Formaten und eine Aggregation der Daten aus den verschiedenen Kommunen gewährleisten. Ein zentrales Metadaten-Portal ermöglicht außerdem die Einbindung von dezentralen Datenflüssen für Datenräume.

2. Bestehende Ressourcen nutzen und sichern

Um die Auffindbarkeit der Metadaten zu den dezentral verwalteten Daten zu gewährleisten, sollten diese Metadaten auf GovData eingestellt werden. Dieses nationale Datenportal sollte vorrangig bedient werden, bevor neue Lösungen aufgesetzt werden. Zudem müssen dafür langfristige Finanzierungen eingerichtet und die personellen wie auch technischen Ressourcen aufgestockt werden, um eine schnelle Anbindung von Datenlieferanten zu gewährleisten.

3. Verpflichtung zur Nutzung standardisierter Metadaten und Schnittstellen

Um die Interoperabilität von Daten zu gewährleisten, sollen einheitliche Metadatenstandards für Open Data über alle föderalen Ebenen hinweg etabliert werden. Dies schließt auch die Festlegung von standardisierten Schnittstellen mit ein.

B. Rechtliche Rahmenbedingungen

1. Harmonisierung von Regeln

EU-weite Verordnungen und Bundesgesetze sollen nicht durch Länderverordnungen ausgehebelt wer-

den können. Die mangelnde Verbindlichkeit bei der Bereitstellung von Daten in den Verwaltungen muss durch klare Konsequenzen bei Nichterfüllung ersetzt werden.

2. Zugangsrechte verankern

Zur Gewährleistung fairer (Wettbewerbs-)Bedingungen sowie im gesellschaftlichen Interesse an Problemlösungen und Innovation sollen gesetzliche Regelungen geschaffen werden, die gegenüber privaten und öffentlichen Datenhalter*innen Zugangsrechte verankern (= Verpflichtung zur Lizenzierung). Dabei sollte die Zugänglichmachung von Daten der Daseinsvorsorge priorisiert werden. Mögliche Zugangsrechte sollten nicht ausschließen, dass die konkrete Gewährung des Zugangs bepreist wird. Die Zulassung der Bepreisung macht eine Preiskontrolle notwendig, sodass die gesetzlich verankerten Zugangsrechte nicht über unangemessene Preisgestaltung ausgehebelt werden können.

3. Gebührenerhebung diskutieren und regeln

Die Frage nach der Möglichkeit zur Gebührenerhebung bzw. nach der Erlaubnis zur Monetarisierung von Datenbeständen bei öffentlichen Stellen bedarf einer grundsätzlichen gesellschaftspolitischen Klärung. Da es sowohl gute Argumente für als auch gegen eine Gebührenerhebung gibt, ist eine elementare öffentliche Debatte zu diesem Thema notwendig, um zu einer von der Gesellschaft getragenen gesetzlichen Regelung zu kommen. Neben einem grundsätzlich freien Zugang zu Daten, deren Erstellung bereits durch Steuergelder finanziert wurde, ist beispielsweise auch ein Stufenmodell denkbar, das den Aufwand für die Datenaufbereitung und -bereitstellung abbildet.

4. Open Source ins Vergaberecht

Bei öffentlichen Ausschreibungen sollte der Einsatz von Open-Source-Lösungen, offenen Standards und offenen Schnittstellen als Vergabekriterium eingeführt werden. Dadurch kann deren Einsatz in öffentlichen Einrichtungen gefördert werden.

5. Weiterverwendung von Daten

Um Daten, die im Rahmen von Forschungen erhoben wurden, für einen anderen als den ursprünglichen Erhebungszweck (wieder-)verwenden zu können, sind klare rechtliche Regelungen notwendig.

C. Qualifizierung, Karrierewege & Kompetenzaufbau

1. Qualifizierung des Personals sicherstellen

Die Vermittlung von Grundwissen über Daten soll innerhalb der Ausbildung oder des Studiums im

öffentlichen Sektor institutionalisiert werden. Die Kernbelegschaft sollte ein Basiswissen über Datenmanagement durch Schulungen und Beratungen erhalten.

2. Tatsächlichen Bedarf evaluieren

Der Schulungs- und Fortbildungsbedarf muss immer wieder an strategischen Zielen und tatsächlichen Bedarfen der Einrichtung ausgerichtet werden, um den wirklichen Erfordernissen begegnen zu können.

3. Anteilige Vollzeitäquivalente einrichten

Dem realen Aufwand der Digitalisierung und zunehmender Datenverwaltung sollte mit angemessener personeller und finanzieller Ausstattung begegnet werden.

4. Ausschreibungen in Zusammenarbeit entwickeln

Die Kompetenzprofile für Ausschreibungen sollten in Kooperationen, beispielsweise mit der NFDI, gemeinsam entwickelt werden. Es ist zu überlegen, an welcher Stelle die Positionen und Kompetenzen benötigt werden, sodass sich die Erfordernisse auch in der Verwaltungsstruktur abbilden. Man wird neue organisatorische Ebenen zwischen IT und Verwaltung einrichten müssen.

5. Neue Gehaltsoptionen einführen

Die TVL-/TVÖD-Richtlinien müssen an die geänderten Anforderungen, die sich aus den Digitalisierungsprozessen ergeben und neue Kompetenzen erfordern, angepasst werden, ansonsten wird die Rekrutierung von qualifiziertem Personal im Datenmanagement scheitern. Im Wettbewerb um die besten Köpfe muss der öffentliche Dienst auch im Gehaltsbereich konkurrenzfähig sein.

6. Datenkompetenz in Behörden etablieren

Für die interne Qualitätsprüfung sowie die Auswertung und Simulation von Daten sind Data-Science-Teams innerhalb der Behörden dringend notwendig. Diese internen Teams sollten ihre Ergebnisse gegenseitig begutachten und validieren und könnten darüber hinaus auch externe Experten prüfen. Ergänzt werden sollten die Teams außerdem um Datenrechtsexpert*innen, die sich mit den rechtlichen Fragestellungen und Möglichkeiten zur (Wieder-)Verwendung von Daten auseinandersetzen. Bisher ist ein Aufbau solcher Strukturen nur in den Bundesministerien geplant.

7. Open-Data-Kompetenzen in den Behörden ausbauen

Die Kompetenz der Datenschutz- und Informationsfreiheitsbeauftragten in den Behörden aller Verwaltungsebenen muss ausgebaut bzw. Stellen eingerichtet werden, wo sie noch nicht bestehen. Bisher

ist die Einstellung von Open-Data-Koordinator*innen nur für Bundesbehörden verpflichtend.

D. Kulturwandel

1. Data Management in each policy

Um die Bedeutung der Digitalisierung und des Handelns mit Daten zu unterstreichen und sichtbar zu machen, sollte Datenmanagement in jedem Leitfaden, Regelwerk oder Strategiepapier als Handlungsfeld Beachtung finden.

2. Den Nutzen der Digitalisierung vermitteln

Die Einrichtung einer benutzer*innenfreundlichen Dateninfrastruktur zum gesamtgesellschaftlichen Nutzen muss als Grundlage verstanden werden. Daten können als Treiber für die Verwaltungsdigitalisierung genutzt werden. Dieser Kulturwandel sollte mit einer Informationskampagne, die den Nutzen verbesserten Datenmanagements illustriert, unterstützt werden.

3. Data Driven Government aufbauen

Der Ansatz von Data Driven Government sollte flächendeckend in den Behörden initiiert werden. Dadurch können Verwaltungshandlungen effizienter gestaltet und stärker an tatsächlichen Gegebenheiten und den Bedürfnissen der Bürger*innen orientiert werden. Dazu braucht es vertrauenswürdige Prozesse, die sich nur durch aufbereitete Daten in einem vorgegebenen, demokratischen und transparenten Rechtsrahmen verwirklichen lassen.

4. Datenbewusstsein etablieren

Abteilungsdenken innerhalb von Verwaltungen sollte durch mehr Transparenz und die Übernahme von Lösungsansätzen aus anderen Bereichen überwunden werden. Dazu ist es notwendig, Vorteile aufzuzeigen, die sich durch das Teilen von Daten ergeben und somit ein Bewusstsein für Datenhandeln zu schaffen. Als Best-Practice-Beispiel kann das Transparenzportal der Stadt Hamburg dienen. Das behördenübergreifende Portal entlastet die Mitarbeiter*innen, da Daten von anderen Stellen ohne separate Anfrage selbsttätig abgerufen werden können.

5. Transparenz schaffen

Die Souveränität der Datengeber*innen muss durch Vereinbarungen über Nutzungsbedingungen oder technische Lösungen gewährleistet sein. Dies baut das nötige Vertrauen zwischen Datenbereitsteller*innen und Datennehmer*innen auf.

6. Daten-Souveränität fokussieren

Die Idee der Daten-Souveränität für Bürger*innen soll von der Europäischen Union noch stärker vorangetrieben werden. Insbesondere kann GAIA-X

hier einen Beitrag leisten und als doppelte Chance verstanden werden: einmal als Bereitsteller von freien Daten und zugleich als eine Art „Datenmarktplatz“, der auch eine gewerbliche Nachnutzung mitdenkt und ermöglicht.

E. Zusammenarbeit und Vernetzung

1. Gezielte Förderung etablieren

Die Kommunen, die bislang bei der Bereitstellung von Open Data zurückliegen, sollen gezielt durch Landes- und Bundesprojekte finanziell gefördert sowie durch intensive individualisierte Beratungen motiviert werden. Dies kann zu Übertragungseffekten führen, durch die umliegende Kommunen mitgezogen werden.

2. Mehr Dialog ermöglichen

Nicht nur bereichsübergreifend, sondern auch innerhalb der Verwaltung soll über die Ebenen von Bund, Land und Kommunen hinweg mehr Dialog ermöglicht werden. Dies ist durch den Aufbau von Netzwerken und regelmäßige Treffen in Fachforen zu bewerkstelligen. Diese Fachforen und der Austausch in den Netzwerken können auch dazu dienen, besser über Absprachen und Standards zu informieren. Die Einbindung von Stakeholdern aus anderen Communities bringen eine Fachexpertise aus Nutzer*innensicht mit ein.

3. Einbindung aller Akteur*innen

In Aushandlungsprozessen mit allen Gruppen kann der tatsächliche Bedarf ausgelotet und die Veröffentlichung von stärker nachgefragten Daten priorisiert werden.

4. Kooperationen mit Unternehmen aufbauen

Die Zusammenarbeit mit Unternehmen kann eine moderne Verwaltung befördern. Aus Kooperationen können Produkte entstehen, die den Datenaustausch einfacher und sicherer machen. Aufbereitete bzw. qualitätsgesicherte Daten aus der Verwaltung können eine Grundlage für Unternehmen bilden.

5. Forschungsprojekte einbinden

Die Zusammenarbeit mit Forschungsprojekten oder die Teilnahme an konkreten Umsetzungsprojekten mit Open Data können die Akzeptanz zum Teilen von Daten bei Mitarbeiter*innen der Verwaltung deutlich erhöhen. Zudem lässt sich so verdeutlichen, wie der Aufwand für das Aufbereiten von Daten unmittelbar wirksam wird.

sind: (A) Teilen von Daten steigern, (B) Nachnutzen von Daten steigern und (C) Bereichsübergreifenden Datenaustausch steigern.

A. Teilen von Daten steigern

1. Policies differenziert gestalten

Bei der Formulierung von Policies ist zu berücksichtigen, dass Unternehmen Daten als Firmeneigentum betrachten, das nur im Rahmen der Realisierung von Wertschöpfungsketten oder Vermarktungsstrategien geteilt wird. Daraus folgt, dass gesetzlich durchgesetzte Open-Data-Regeln die Kooperationsfähigkeit mit der Wirtschaft beeinträchtigen können.

2. Wertschöpfung in den Fokus stellen

Policies zur Förderung der Nachnutzung von Daten sollten sich primär an der Notwendigkeit der Schaffung neuer oder Verbesserung bestehender Geschäftsmodelle oder Wertschöpfungsketten orientieren. Es braucht eine gezielte Förderung bei der Entwicklung von funktionierenden Geschäftsmodellen sowie bei der Vermarktung zugänglich gemachter Daten.

3. Technische Voraussetzungen verbessern

Durch Verbesserung der technischen Voraussetzungen kann man den Handlungsspielraum für die Entwicklung neuer oder die Verbesserung bestehender Geschäftsmodelle und Wertschöpfungsketten erweitern. Einigen sich Vertragsparteien auf eine bestimmte Nutzung von Daten für einen festgelegten Zeitraum, ist es wichtig, ob die Einhaltung dieser Vereinbarung auch technisch abgesichert werden kann. Wenn die Kooperation technisch möglich ist, werden die Vertragsparteien vielfach nicht in der Lage sein, die technischen Voraussetzung selbst zu schaffen. Der Vertragsabschluss bzw. die Realisierung des damit korrespondierenden Geschäftes wird deshalb vielfach von Dienstleister*innen abhängen, die dieses technische Umfeld bieten können. Online-Handel auf Plattformen ist inzwischen verbreitet. Für die Nachnutzung von Daten fehlen solche Infrastrukturen vielfach.

4. Rechtliche Rahmenbedingungen weiterentwickeln

Die Realisierung von Nachnutzungsszenarien ist stark von den rechtlichen und technischen Rahmenbedingungen abhängig. Insbesondere im Kontext des IoT bestehen konkurrierende Interessen am Zugriff auf automatisch erzeugte Daten. Die Interessen von Wirtschaft, *Public Sector*, Wissenschaft und die der Konsument*innen müssen in Ausgleich gebracht werden. Benötigte Regelungen sollten international möglichst breite Geltung entfalten. Entsprechende Gesetzgebung sollte deshalb

6.3 Wirtschaft

Für die Wirtschaft wurden insgesamt 20 Empfehlungen abgeleitet, die unter drei Kategorien aufgeführt

ausschließlich auf EU-Ebene und als direkt geltende Verordnung realisiert werden, um eine größtmögliche Harmonisierung zu erreichen. Gleichzeitig sollte sich die Europäische Union, unterstützt durch ihre Mitgliedstaaten, in multilateralen Projekten zur Entwicklung und Harmonisierung des „Datenrechtes“ engagieren.

5. Nutzung organisationsinterner vorhandener Daten für Prozessoptimierungen

Eine eindimensionale Fokussierung von Fördermaßnahmen auf Szenarien der Nachnutzung von organisationsinterner Daten durch Dritte unterschätzt das große Potential, das vielfach in der internen Nutzung der Daten für Prozessoptimierung liegt. Einige Akteur*innen sehen in dieser Nutzung aktuell das größte Wertschöpfungspotential.

6. Fördergelder gezielt einsetzen

Förderungsgelder für den Aufbau von Infrastrukturen zum Teilen von Daten in der Wirtschaft müssen gezielter eingesetzt werden. Hierbei ist der tatsächliche Bedarf zu prüfen, da es bereits eine Reihe von Branchen gibt, die Modellprojekte auf europäischer Ebene entwickeln (Catena-X im Automobilbereich und Datenraum Mobilität, Agregai in der Landwirtschaft, Financial Big Data Cluster).

7. Klare rechtliche Rahmenbedingungen für Datentreuhänder*innen

Die technische Entwicklung von Systemen der Datentreuhänder*innenschaft bedarf einer parallelen Entwicklung rechtlicher Rahmenbedingungen, die auch eine Verpflichtung der großen Datennutzer bzw. -halter (z. B. Google, Amazon) zur Kooperation beinhaltet. Für Datentreuhandssysteme werden klare rechtliche Rahmenbedingungen benötigt, damit diesen Akteur*innen das Vertrauen entgegen gebracht wird, das für ihr funktionieren unerlässlich ist. Diese müssen öffentlich und nicht durch die Unternehmen, deren Praxen reguliert werden sollen, kontrolliert werden.

Auf europäischer Ebene

8. Vertrauenswürdige Datenräume schaffen

In Europa müssen vertrauenswürdige Datenräume geschaffen werden, um das Teilen von Daten sicher organisieren zu können. Vertrauenswürdige Datenräume werden benötigt, weil individuelle Datenhalter*innen nicht in der Lage sind, autonom eine sichere Nutzung geteilter Daten sicherzustellen. Dies gilt insbesondere für die Nachnutzung sensibler Daten.

9. Digitale Souveränität realistisch denken

Bei der Ausgestaltung der in Europa angestrebten digitalen Souveränität ist darauf zu achten, die in-

ternationale Kooperationsfähigkeit, welche die Voraussetzung für den internationalen Handel bildet, nicht zu behindern.

10. Rechtliche Rahmenbedingungen

Die europäische Datenschutzgesetzgebung muss handhabbar gestaltet bzw. Aufsichtsbehörden in die Lage versetzt werden, geltendes Recht auf handhabbare Weise zu implementieren. Bei der Förderung der europäischen Digitalwirtschaft besteht insbesondere bei der Entwicklung der rechtlichen Rahmenbedingungen starker Handlungsbedarf. Die wichtigsten Herausforderungen hierbei sind der Datenschutz, Identity Management und Systemsicherheit.

11. Asymmetrische Regelungen für Plattformen schaffen

Marktmächtige Plattformen sollten zum Datenteilen verpflichtet werden. Zudem braucht es eine asymmetrische Regelung der Datenteilpflicht für kleinere Plattformen und Unternehmen. Soziale Plattformen (soziale Netzwerke) müssten zu offenen Standards gezwungen werden (d. h. Interoperabilität und Portabilität), was auch ein klares Verbot von Exklusivbindungen mit sich bringt. All dies würde zur Stärkung des Wettbewerbs führen, zur Entwicklung neuer Innovationen und eine Monopolisierung erschweren sowie das Datenteilen und -nachnutzen durch kleinere wirtschaftliche Akteur*innen fördern.

B. Nachnutzung von Daten steigern

1. Nachnutzung von Daten zielgerichtet fördern

Bei der Formulierung von Policies zur Nachnutzung von Daten sollte mitberücksichtigt werden, dass Daten aus Forschung und öffentlicher Verwaltung durch Unternehmen oftmals nur dann nachgenutzt werden können, wenn sie in für die Unternehmen verwertbarer Form vorliegen. Die Förderungen zum Nachnutzen von Daten sollten daher auf konkrete Szenarien konzentriert werden, in welchen Datenanbieter*innen und Datennutzer*innen mit ihren jeweiligen Interessen bekannt sind. Der Kuratierung der Daten muss mehr Aufmerksamkeit geschenkt werden, weil mit der Aufbereitung von Daten die Wahrscheinlichkeit ihrer Nachnutzung gesteigert werden kann. Das Kuratieren von Daten kann sich zu einem eigenen Geschäftsfeld entwickeln.

2. Nachnutzungsszenarien entwickeln

Damit sich aus potentieller tatsächliche Nachnutzung entwickelt, sollten Plattformen zur Vermittlung zwischen Datenerzeuger*innen und Nachnutzer*innen aufgebaut werden. Diese sollte auch den Austausch von Ideen zur Nutzung der Daten

ermöglichen und die Kommunikation zwischen Datenanbieter*innen und Nachnutzer*innen unterstützen.

3. Die Beachtung ethischer Aspekte konkretisieren

Bei der Förderung der Nachnutzung von Daten muss die Beachtung ethischer Aspekte konkreter verankert werden. Das Mantra „As open as possible. As closed as necessary.“ ist zu unkonkret für die empirische Anwendung.

4. Rechtliche Klarheit schaffen

Beim Nachnutzen von Daten besteht bei allen Beteiligten regelmäßig Rechtsunsicherheit bzw. Beratungsbedarf. Auf dieses Desiderat sollte bei der Förderung der Nachnutzung von Daten durch die Förderung entsprechender Weiterbildungs- und Beratungsangebote reagiert werden. Im Zuge dessen bedarf es einer Lizenzierung der Daten, damit klar kommuniziert wird, wer was mit den Daten machen darf.

5. Den Wettbewerb stärken

Es braucht die Entwicklung eines rechtlichen Instrumentariums, um im öffentlichen Interesse und zur Realisierung von Wettbewerb Zugriffe auf die riesigen aggregierten Datenmengen weniger einzelner Großunternehmen zu ermöglichen. Start-ups, die für die Entwicklung ihrer Produkte auf die Nachnutzung von Daten Dritter angewiesen sind, sollten gezielt gefördert werden.

6. Datenverständnis und -kompetenzen aufbauen

Insbesondere große Unternehmen müssen ihre Fähigkeiten zur Beurteilung von Möglichkeiten zur Nachnutzung der bei ihnen vorliegenden Daten erweitern und Geschäftsmodelle zum Datenteilen entwickeln. Hier liegt großes wirtschaftliche Potential brach.

7. Technologietransferstrategien anpassen

Die Wissenschaftsorganisationen sollten ihre Technologietransferstrategien überprüfen. Zu oft streben Wissenschaftsorganisationen Produktentwicklungen an, ohne sie abschließen zu können. Es wäre zielführender, das Know-how früher für die Privatwirtschaft zugänglich zu machen.

C. Bereichsübergreifenden Datenaustausch steigern

1. Aufbau vermittelnder Stellen fördern

Es bedarf einer gezielten Förderung bei dem Aufbau vermittelnder Stellen, da die Nachnutzung von Know-how und Daten aus der Forschung oft an Kompetenzlücken zwischen Forschung und Wirtschaft scheitert. Für die Finanzierung dieser Mittler fehlen Finanzierungs- bzw. Geschäftsmodelle.

2. Anreize für die Produktion nachnutzbarer Daten schaffen

Eine Etablierung eines Wiederverwendungs-Index für wissenschaftliche Daten und Forschungsorganisationen würde die Nachnutzbarkeit mehr in den Fokus rücken. Darüber hinaus sollte die Erzeugung hochwertiger Datensets bzw. deren Aufwertung durch Kuratierung stärker honoriert werden. Die Veröffentlichung von Forschungsdaten sollte finanziell belohnt werden. Dies würde auch eine Verstärkung der Nachnutzung der Daten bewirken.

6.4 Qualifizierung

Es wurden insgesamt 30 Empfehlungen für den Untersuchungsbereich Qualifizierung abgeleitet. Der Gliederung der vorangegangenen Kapitel zu Qualifizierung folgend, sind sie nach fünf Kategorien geordnet: (A) Allgemeine Rahmenbedingungen, (B) Datenbewusstsein und Kulturwandel, (C) Kompetenzprofile, Fähigkeiten und Skills, (D) Medienkompetenzvermittlung und Qualifikationswege und (E) Berufsbilder und Karrierewege.

A. Allgemeine Rahmenbedingungen

Die folgenden Empfehlungen zur Schaffung und Verbesserung von Rahmenbedingungen richten sich vornehmlich an die Politik, den Gesetzgeber, aber auch an Hochschulen.

1. Anerkennung von Datenkompetenzen im Wissenschaftssystem verbessern

Qualifizierung im Umgang mit Daten muss durch das System gefördert werden. Aktuell sind Anerkennung von Datenkompetenzen und Anreize zu ihrem Erwerb noch wenig etabliert.

Die Berücksichtigung von Datenkompetenzen bei Berufungen auf Professuren könnte ein Anreiz sein. Ein „Impact Factor“ für Datenveröffentlichungen trüge dazu bei, den Erwerb und die Anwendung von Datenkompetenzen in der Forschung und bei Karrieren voranzubringen. Auch könnten überprüfbare Zielvereinbarungen Beschäftigte motivieren, sich im Umgang mit Daten weiter zu qualifizieren.

2. Eine gute Praxis im Umgang mit Daten durch verpflichtende Vorgaben etablieren

Klare Vorgaben und Pflichten beim Umgang mit Daten verändern die gelebte Praxis. So könnten z. B. bei öffentlich finanzierter Forschung durch Projekte an Bildungseinrichtungen wie Hochschulen neben Datenmanagementplänen (DMPs) auch Dokumentation, Transparenz und Nachnutzbar-

keit von Daten zur Pflicht und zur Voraussetzung für eine Förderung gemacht werden.

3. Zentrale Anlaufstellen fördern

Der Kompetenzaufbau in Form von (de-)zentralen Anlaufstellen (wie Kompetenzzentren) sollte durch die Schaffung der erforderlichen Rahmenbedingungen und Finanzierung entsprechender Einrichtungen und Stellen verbessert werden. Insbesondere im Bereich *Public Sector* könnten diese auch als Anlaufstellen für eine systematische Weiterqualifizierung (Zertifikate/Schulungen) dienen.

4. Datenbeauftragte benennen

Eine verantwortliche Person sollte als Beauftragte*r für Daten in allen Einrichtungen der untersuchten Bereiche benannt werden. Zumindest für den *Public Sector* kann dies durch gesetzliche Vorgaben verpflichtend vorgeschrieben werden. Dabei ist darauf zu achten, dass die Beauftragten mit einem passenden Zeitbudget ausgestattet werden und es Mindestvorgaben für die Anforderung an Kompetenzen im Datenbereich gibt, die ggf. über (Weiter-)Qualifizierungsmaßnahmen erfüllt werden können.

5. (Bottom-up-)Initiativen zu Qualifizierung in Supportstrukturen unterstützen

Supportstrukturen für Daten entwickeln sich in der Praxis am Bedarf orientiert weiter. Dies umfasst auch Ideen und Maßnahmen für die Qualifizierung der Beschäftigten im direkten Umgang mit diesen Strukturen bzw. ihrer Nutzer*innen. Für solche Initiativen sollten die Rahmenbedingungen für den erforderlichen Freiraum ohne (massive) Top-down-Einschränkungen geschaffen werden. Der Aufbau dieser Supportstrukturen und ihre Vernetzung sollte organisatorisch und finanziell gefördert werden.

B. Datenbewusstsein und Kulturwandel

Die Empfehlungen zur Förderung eines Kulturwandels haben zum Ziel, ein möglichst breites Bewusstsein und Verständnis für einen informierten Umgang mit Daten zu schaffen und so zu einer „neuen Datenkultur“ beizutragen. Sie richten sich an Politik und „Gesellschaft“ im Allgemeinen, aber auch an Verwaltung, Wirtschaft und Wissenschaft (Hochschulen).

1. Data Literacy als gesellschaftliches Ziel definieren

Data Literacy, verstanden als informierter Umgang mit Daten, ist eine Voraussetzung für die gesellschaftliche und berufliche Teilhabe in einer zunehmend digitalisierten Welt. Daher sollte es ein gesellschaftliches Ziel sein, allen Bürger*innen die notwendige Grundkompetenz im Bereich Daten zu vermitteln.

2. Data Literacy in allen Bildungsetappen vermitteln

Zu einer nachhaltigen Verbesserung des Datenbewusstseins gehört die Vermittlung von Kenntnissen im Umgang mit Daten. Daher sollte Data Literacy als wichtiges Element in allen Bildungsetappen und jeder Bildungsbiographie vorkommen.

3. Vorgaben und Umsetzungsrichtlinien für den Umgang mit Daten machen

Beschäftigte im *Public Sector* sind darin geschult, Vorgaben zu folgen und sich an Umsetzungsrichtlinien zu orientieren. Explizite Vorgaben und Umsetzungsrichtlinien, etwa um Datenteilen als Aufgabe zu etablieren, wären deshalb für diesen Bereich hilfreich.

4. Verständnis für das Potenzial von Daten vermitteln und Mehrwerte verdeutlichen

Datenbewusstsein beruht auf dem Verständnis für Daten, mit denen gearbeitet wird. Deshalb sollte das Potenzial von Daten durch das Aufzeigen von möglichen Verknüpfungen einzelner Datensätze sichtbar und nutzbar gemacht werden, um so die Forschung voranzutreiben. Mehrwerte durch Nachnutzung und Kombination von Daten(sätzen) könnten durch Erfolgsbeispiele illustriert werden.

5. Datenkompetenz als wichtige Qualifikation in der Wirtschaft stärken

In der Wirtschaft fehlt an einigen Stellen das Bewusstsein für die große Bedeutung von Datenkompetenz für den wirtschaftlichen Erfolg. Daher sollte diese durch geeignete Qualifizierungsmaßnahmen gerade auch beim vorhandenen Personal verbessert werden.

6. Vorurteile und Ängste durch Aufklärung und Information abbauen

Nach wie vor gibt es große Vorurteile und Ängste, dass beispielsweise Daten nur etwas für Expert*innen sind oder dass Datenteilen rechtlich bedenklich ist. Durch gezielte Aufklärung und Informationen können diese Bedenken, die insbesondere im *Public Sector* noch verbreitet sind, abgebaut werden.

7. Vorbildfunktion bereits in der Ausbildung nutzen

Neben der reinen Vermittlung von Kenntnissen ist der aktiv „vorgelebte“ Umgang mit Daten von großer Bedeutung. Deshalb sollte diese Vorbildfunktion bereits in der wissenschaftlichen Lehre oder Ausbildung genutzt werden. Um den Studierenden ein Bewusstsein für Daten zu vermitteln, könnten Dozent*innen beispielsweise mit jungen Forscher*innen Nachnutzungsfragen erörtern und durch ihre eigene Praxis im Umgang mit Daten zum Vorbild werden.

C. Kompetenzprofile, Fähigkeiten und Skills

Für eine erfolgreiche Medienkompetenzvermittlung sind die zu vermittelnden Inhalte von essentieller Bedeutung. Eine bereichs- und anwendungsbezogene Definition von grundlegenden Kenntnissen zum Umgang mit Daten für die Vermittlung an spezifische Zielgruppen in den Bereichen Wissenschaft (für Ausbildung und Forschung), *Public Sector* und Wirtschaft (hinsichtlich der Bereitstellung und Verwendung) ist daher erforderlich. Die folgenden Empfehlungen richten sich u. a. an Fachgesellschaften, Datenkompetenz-Vermittler*innen (Ausbildung/Hochschule) und Informationsstellen zu Weiterbildungsangeboten.

1. Kompetenzprofile und Kompetenzkataloge für FDM für Hochschulen erstellen

Im Bereich der Hochschulausbildung sind Kompetenzprofile und Kompetenzkataloge für den Umgang mit Daten eine wichtige Orientierungshilfe für die Entwicklung von Studieninhalten. Empfehlungen für ein fachspezifisches Curriculum könnten beispielsweise wissenschaftliche Fachgesellschaften machen.

2. Datenkompetenz als „Fähigkeit Daten FAIR zu machen“ verstehen

Eine Definition, was Datenkompetenz beinhaltet, ist für ihre Vermittlung von großer Bedeutung und meist nur im spezifischen Kontext möglich. Eine für die Praxis taugliche und zur Orientierung für ihre Vermittlung dienende Definition ist, Datenkompetenz als die „Fähigkeit Daten FAIR zu machen“ zu verstehen, d. h. Daten auffindbar, zugänglich, austauschbar/verknüpfbar und nachnutzbar zu machen.

3. Ethische Aspekte besonders berücksichtigen

Neben technisch-anwendungsbezogenen Kompetenzen, wie sie für die Umsetzung der FAIR-Prinzipien erforderlich sind, sollten auch Kompetenzen zu ethischen Aspekten für den Umgang mit Daten vermittelt werden. Eine gute ergänzende Orientierung geben hier u. a. die CARE-Prinzipien.

4. Ergänzende Inhalte zu Datenkompetenzen mit in Kompetenzprofile aufnehmen

In der Praxis reicht reine Datenkompetenz für einen guten Umgang mit Daten allein nicht aus. Gerade wenn es um die Etablierung von neuen Prozessen in Einrichtungen geht, sind auch Kompetenzen aus den Bereichen Organisation, Change-Management, IT, Recht und Kenntnisse zu ethischen Aspekten wichtig und sollten dementsprechend in Qualifizierungsmaßnahmen vermittelt werden.

5. Qualitätsgeprüfte Übersichten über Qualifizierungsangebote zu Daten erstellen

Der Markt für (Weiter-)Qualifizierungsangebote im Bereich Daten entwickelt sich schnell und es ist nahezu unmöglich, einen Überblick über alle Angebote zu gewinnen, insbesondere fehlt es an Gegenüberstellungen, die eine Orientierung über Zielgruppen und Vergleichbarkeit von Inhalten und Qualität geben könnten. Daher wäre eine systematische Zusammenstellung von Angeboten in den Bereichen Ausbildung und Weiterbildung, möglichst nach Fachbereichen bzw. Anwendungsgebieten sortiert und mit Hinweisen zur Einordnung der Qualität (durch Zertifikate etc.) eine große Hilfe.

D. Medienkompetenzvermittlung und Qualifikationswege

Die Vermittlung von Datenkompetenzen sollte möglichst früh beginnen und auf breiter Basis durch unterschiedliche Stellen, z. B. durch Schulen, Hochschulen, Ausbildungen und Weiterqualifizierungen erfolgen. Dabei gilt, viele Aspekte zu berücksichtigen. Die folgenden Empfehlungen richten sich u. a. an Lehrplan- und Curriculums-Entwickler*innen, Politik, Förder*innen, Arbeitgeber*innen, Arbeitnehmer*innen, Verantwortliche für Ausbildungsinhalte und Anbieter*innen von Fortbildungen.

1. Datenkompetenzen bereits in der grundlegenden Ausbildung vermitteln

Der Umgang mit Daten sollte integraler Bestandteil der Ausbildung an Schulen und Hochschulen sein. Dabei könnte die Vermittlung von Grundlagen bereits in den Schulen und die Vertiefung in allen Studienfächern erfolgen. Ergänzend sollten spezielle FDM-(Zusatz-)Qualifikationen und eigenständige Abschlüsse angeboten werden. Die Vermittlung von Datenkompetenzen in den jeweiligen Stationen von Schüler*innen, Studierenden, wissenschaftlichem Nachwuchs, Promovierenden sollte aufeinander aufbauend erfolgen.

2. Datenkompetenzen bereits in der Berufsausbildung vermitteln

Die Bedeutung von Datenkompetenzen steigt auch in den Ausbildungsberufen stetig. Daher sollte die Vermittlung von Grundkenntnissen im Umgang mit Daten in allen infrage kommenden Ausbildungen, insbesondere für Berufe in den Bereichen öffentliche Verwaltung (als Datenbereitsteller*innen), Wirtschaft (als Datennutzer*innen) und unterstützenden Berufen für die Wissenschaft (Daten-Support) erfolgen.

3. Medienkompetenzvermittlung durch Weiterqualifizierung im Beruf fördern

In vielen Berufen gibt es durch die zunehmende Digitalisierung einen stetig steigenden Bedarf an Kompetenzen im Umgang mit Daten. Dieser kann am besten durch beständige Weiterqualifizierung, beispielsweise durch die Teilnahme an anerkannten Zertifikatskursen von qualitätsgeprüften Fortbildungsanbietern, gedeckt werden. Bei der Auswahl müssen die speziellen Anforderungen für den spezifischen Bereich berücksichtigt werden. Schulungen können beispielsweise direkt am Arbeitsplatz erfolgen.

4. (Fachspezifische) Netzwerke fördern und zum Erfahrungsaustausch nutzen

Da die Anforderungen zum Umgang mit Daten in vielen Bereichen ähnlich sind, bietet sich die Zusammenarbeit und Vernetzung mit anderen an. In diesem Netzwerk können Erfahrungen und Informationen ausgetauscht und Kompetenzen vermittelt werden. Innerhalb der Netzwerke können Angebote zur Medienkompetenzvermittlung in Form von internen oder externen Workshops, Online- und Präsenzkursen sowie -Seminaren genutzt oder kooperative Beratungsstellen geschaffen werden.

5. Kontinuierliche Weiterbildung von Ausbildungs- und Lehrkräften fördern

Die Anforderungen an Datenkompetenzen im Berufsalltag verändern sich aufgrund der schnellen Veränderungen durch die Digitalisierung im Berufsalltag beständig. Dies muss auch in der Vermittlung berücksichtigt werden, d. h. dass Ausbildungs- und Lehrkräfte möglichst auf dem aktuellen Stand der Entwicklungen sind, um diesen vermitteln zu können. Hierfür müssen sie kontinuierlich weiterqualifiziert werden, um stets up to date zu sein. Eine Umsetzung in der Praxis kann z. B. durch Zielvereinbarungen oder durch die Aufnahme von Angeboten zu Datenkompetenz in Standardkataloge für Weiterqualifizierungsmaßnahmen gefördert werden.

E. Berufsbilder und Karrierewege

Eigene Berufsbilder im Bereich Daten sind noch in der Entwicklung. Noch eher selten kann qualifiziertes

Personal über spezifische Berufsbezeichnungen gezielt gesucht werden, da die Berufsbilder zumeist noch unklar und unscharf abgegrenzt sind. Es lassen sich jedoch drei Tendenzen erkennen. Erstens die Entwicklung von unterstützenden, zuarbeitenden Berufen und Berufsbildern (z. B. Data Steward, Data Manager, Datenkurator, ...), zweitens eine Zusatzqualifikation aufbauend und ergänzend zur eigentlichen beruflichen Qualifikation (z. B. Data Scientist) und drittens die Bildung von Kompetenzteams, in denen Spezialkenntnisse einzelner Beschäftigter gebündelt werden. Die folgenden Empfehlungen richten sich an Politik, Unternehmen und Institutionen.

1. (Neue) Berufe im Datenbereich klar definieren und staatlich anerkennen

Zur Etablierung von neuen Berufen gehört die klare Definition und möglichst auch eine Anerkennung von staatlicher Seite. Dadurch werden die neuen Berufsbezeichnungen aussagekräftig und „praxistauglich“, d. h. Arbeitsanforderungen, notwendige Qualifizierungen (Ausbildung, Studium, Weiterqualifizierungen) bis hin zu potenziellen Karrierewegen können beschrieben und genutzt werden.

2. Eigene Karrierewege im Supportbereich für Daten ermöglichen

Vielfach entwickeln sich neue Aufgaben und letztlich neue Berufe direkt aus den Anforderungen im Umgang mit Daten. Im Supportbereich zu Daten sind zwar oft die Infrastrukturen, aber eher selten die zugehörigen Stellen dauerhaft finanziert. Daher ist die Einrichtung von Dauerstellen und damit die Schaffung von Perspektiven für eine berufliche Karriere von hoher Bedeutung für die Gewinnung von qualifiziertem Personal für den Support.

3. Tätigkeiten innerhalb von Kompetenzteams für Daten fördern

In der Praxis verteilen sich einzelne Kompetenzen (z. B. technisch, rechtlich, organisatorisch) für neue berufliche Anforderungen im Bereich Daten auf mehrere Personen. Daher empfiehlt es sich diese innerhalb von Kompetenzteams zu bündeln, Dabei ist darauf zu achten, dass den Mitgliedern entsprechende Zeitbudgets gegeben werden und das Thema Daten nicht als reine „on-top“-Aufgabe und damit als Zusatzbelastung vergeben wird.

EcoDM- EMPFEHLUNGEN



7. Bereichsübergreifende EcoDM-Empfehlungen

Nachdem für die vier Untersuchungsbereiche *Wissenschaft, Public Sector, Wirtschaft* und *Qualifizierung* sowohl die jeweiligen Empfehlungen als Ergebnis aus den Landscape- und Gap-Analysen als auch aus den Leitfaden-gestützten Interviews analysiert wurden (vgl. Kapitel 6), haben sich übergreifende EcoDM-Empfehlungen herauskristallisiert.¹ Diese wurden vorab für eine breite Diskussion (vgl. Kapitel 8) unter dem Titel „*Data At Your Fingertips: Wegmarken einer exzellenten und vertrauenswürdigen Datenlandschaft*“ veröffentlicht (Bruch et al., 2022).

Die Umsetzung der im Folgenden gelisteten Empfehlungen kann wesentlich dazu beitragen, die Vision eines Ökosystems Datenmanagement zu realisieren. Diese Vision lässt sich bildlich wie folgt noch einmal skizzieren:

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 Sectors, der Wirtschaft, der Gesellschaft und der Privat-Interessierten.

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 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 entgegenzuwirken.
9. Bei der **Nutzung von Produkten von Hyperscalern**² 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.

¹ Diese Empfehlungen sind thematisch geclustert, aber innerhalb eines Clusters unterliegen sie keiner Priorisierung, d. h. die Reihenfolge bzw. Nummerierung unterliegt keiner Wertung.

² Z. B. Amazon Webservices, Microsoft Azure, Google Cloud Plattformen

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 wissenschaftsunterstützenden 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 Zugriffs 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 – „**lizenzierter**“ 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 Verbotsmodulen“ 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 Know-how 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 nutzer*innenzentrierte, 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.

FAZIT UND AUSBLICK



8. Fazit und Ausblick

Die bereichsübergreifenden EcoDM-Empfehlungen waren Thema einer Podiumsdiskussion im Februar 2022 mit vier Expert*innen, die als Abschlussveranstaltung des Projektes im Rahmen der RDA DE Konferenz 2022 durchgeführt wurde. Die Expert*innen wurden so ausgewählt, dass die Perspektiven der Stakeholder der vier Untersuchungsbereiche des Projektes auf dem Podium repräsentiert waren:

- für **Wissenschaft**: Dr.-Ing. Jörn Krupa, Leitung, Transfer & Innovation, Helmholtz-Zentrum Potsdam – Deutsches GeoForschungsZentrum GFZ
- für **Public Sector**: Marie Jansen, Senior Business Analyst, Public Sector, Business Technology Solutions, Capgemini Deutschland
- für **Wirtschaft**: Dr. Alexander Hirschfeld, Leiter Research, Bundesverband Deutsche Startups e.V.
- für **Qualifizierung**: Prof. Dr.-Ing. Habil. Ulrike Lücke, Professorin für Komplexe Multimediale Anwendungsarchitekturen, Universität Potsdam

Moderiert wurde die Diskussion von Dr. Jan-Martin Wiarda, einem freien Journalisten für Bildung und Wissenschaft. Die Podiumsteilnehmer*innen hatten die 31 Empfehlungen (vgl. Kapitel 7) vorab erhalten, um sich auf die Diskussion vorbereiten zu können. Über ein Etherpad (vgl. zum methodischen Vorgehen Kapitel 2.4.3) konnten die Empfehlungen am Tag der Podiumsdiskussion und während der folgenden Tage kommentiert werden. Das Feedback der Podiumsteilnehmer*innen, der weiteren Teilnehmer*innen der Veranstaltung sowie der Kommentator*innen, die das Etherpad als Forum nutzten, ist hier zusammenfassend dokumentiert.

Die Initiator*innen des Forschungsprojektes EcoDM führten zu Beginn der Podiumsdiskussion kurz in die Ziele des Projekts ein und stellten schlaglichtartig und abstrahierend drei zusammenfassende Erkenntnisse vor:

- Das Bewusstsein für die Chancen, die sich aus der Nutzung von Daten ergeben können, ist noch unzureichend ausgebildet.
EcoDM empfiehlt, den benötigten Bewusstseinswandel durch praxisnahe Projekte zu fördern, die den Mehrwert der Nutzung von Daten illustrieren.
- Die Fähigkeit zum Sammeln und Auswerten großer Datenmengen birgt auch Risiken.
EcoDM empfiehlt, begründeten Vorbehalten konstruktiv zu begegnen.

- Es fehlt an finanziellen und personellen Ressourcen für die Ausgestaltung eines leistungsfähigen Datenökosystems.

EcoDM empfiehlt, bei der Förderung des Teilens von Daten korrespondierenden Geschäftsmodellen bzw. deren Entwicklung mehr Aufmerksamkeit zu schenken, damit aus Kosten Investitionen werden sowie Qualifizierung in den Fokus zu nehmen.

Sowohl den ca. 90 Teilnehmenden der Podiumsdiskussion als auch einer breiteren Öffentlichkeit wurde die Gelegenheit gegeben, die in einer Draft-Veröffentlichung publizierten Empfehlungen¹ in einem Etherpad zu kommentieren, zu ergänzen etc.

Zentrales Ergebnis der Podiumsdiskussion war die übereinstimmende Einschätzung, dass mit den Empfehlungen die Sichtweisen der Akteursgruppen aus den Untersuchungsbereichen angemessen repräsentiert und die richtigen Impulse gesetzt werden. Unabhängig davon betonten alle Diskutant*innen die mit ihren Arbeitskontexten zusammenhängenden Aspekte bezüglich der Empfehlungen. Im Folgenden werden die Diskussionsrunden, die von dem Moderator zu bestimmten Fragestellungen initiiert wurden, thematisch zusammengefasst.

In der Podiumsdiskussion wurden zahlreiche Aspekte, die mit **Kulturwandel** im engeren und weiteren Sinne beschrieben werden können, wiederholt angesprochen. In zahlreichen Diskussionsbeiträgen wurde der Kulturwandel in Bezug gesetzt zu der Generierung, Weiterprozessierung, Auswertung sowie der Nachnutzung von Daten. Darüber hinaus wurden auch Themenbereiche des Kulturwandels sichtbar, die eher als Querschnittsthema zu betrachten sind. Im Folgenden werden die den Diskussionsteilnehmer*innen wichtigen Anliegen kurz vorgestellt:

- Datenkompetenz (Data Literacy) wurde wiederholt als ein zentrales Anliegen thematisiert. Dies reichte von der gewünschten Förderung – im Sinne der Data-Literacy-Charta (vgl. hierzu Kapitel 4.4.4.2), in der bereits die Elementarbildung berücksichtigen werden sollte – über die Integration von Datenkompetenz in allen Bildungsangeboten bis hin zu der festen Verankerung in den Curricula von der Primärbildung bis zum tertiären Bildungsbereich. Insgesamt müsse das Bewusstsein für die Notwendigkeit der Förderung von Datenkompetenz in allen Bereichen, beruflichen Stationen und Ausbildungswegen erhöht werden. Zertifizierte Schulungs- und Weiterbildungsangebote müssten kurz- und mittelfristig etabliert werden, bis Data Literacy und ihre

¹ Die Draft-Version wurde auf der EcoDM-Homepage unter den Projektpublikationen hochgeladen: <https://www.ecodm.de/informationsportal/publikationen/>

Vermittlung alle ausbildungs- und berufsbezogenen Bereiche vollständig erfasst.

- Weitestgehend Unverständnis, wenn nicht sogar phasenweise und vereinzelt Frust, kam darüber zum Ausdruck, dass bereits bestehende gesetzliche Vorgaben immer noch nicht selbstverständlich umgesetzt werden, wie zum Beispiel in Behörden. Die erste Version der EU-Richtlinie (2003/98/EG) über die Weiterverwendung von Informationen des öffentlichen Sektors wurde im Jahr 2003 beschlossen. Es gäbe seit vielen Jahren gute Praxisbeispiele für das Management und die Zugänglichmachung von Daten und die einschlägigen Gesetze wurden mehrfach novelliert und ergänzt. Es sei enttäuschend, dass das Datenmanagement sowie das Teilen und Nachnutzen von Daten sich nicht dynamischer entwickelt habe.
- Auch wenn viele Einrichtungen, Unternehmen, öffentliche Verwaltungen etc. grundsätzlich dem Datenmanagement nicht abgeneigt seien, mangle es offenbar immer noch an einer Vorstellungskraft für Innovationspotenziale, die in Daten liegen. Diese Potenziale würden nicht gesehen oder zumindest deren Nutzung nicht priorisiert, die Entwicklung entsprechender Geschäftsmodelle werde nicht ambitioniert genug betrieben. Dies wurde z. T. auch als Ursache dafür gesehen, dass Deutschland im internationalen Vergleich abgehängt sei. Dies gelte insbesondere für das Forschungs- bzw. Geschäftsfeld Künstliche Intelligenz. In den USA lägen die diesbezüglichen Investitionen um den Faktor 10 über den Investitionen in Deutschland.
- Es gebe zahlreiche Vorbehalte gegen die Erhebung und das Teilen bestimmter Daten. Diese Vorbehalte seien teilweise berechtigt, beispielsweise bei personenbezogenen Daten. Angst vor Missbrauch und Fehlnutzung von Daten sei begründet. Diese Vorbehalte sollten ernst genommen werden. Es gebe zahlreiche Möglichkeiten, diesen begründeten Sorgen adäquat zu begegnen. Hierzu zähle insbesondere die Schaffung von Foren, in welchen Konflikte rund um die Erhebung und Verarbeitung von Daten sachlich diskutiert werden können.
- Führungskräfte haben nach Ansicht der Diskussteilnehmer*innen eine entscheidende Rolle im Kulturwandel. Sie hätten es, ohne dass sie selbst Expert*innen für Digitalisierung und Open Data sein müssen, in der Hand, ihre Mitarbeitenden schulen zu lassen, diese selber aktiv weiterzuentwickeln, deren Ängste und Sorgen abzubauen, Möglichkeiten für Quereinsteiger*innen zu schaffen und sie als Multiplikator*innen zu etablieren, auch wenn es bislang noch keine definierten Karrierepfade oder konkrete Berufsbezeichnungen für diese gibt.
- Datenmanagement und damit der dokumentierte Umgang mit Daten müsse bereits vor Beginn von

Datenverarbeitungen als integraler Bestandteil von Forschungsdesigns oder anders gearteten Vorhaben eingeplant werden. Nur so sei zu gewährleisten, dass für Datenmanagement gleich zu Beginn ein Bewusstsein geschaffen werde. Dafür sei viel Aufklärungsarbeit notwendig und die für die Organisation dieser Aufgabe verantwortlichen Personen müssten explizit benannt werden.

Weiterhin war den Podiumsgästen das Thema **Anreiz- und Anerkennungsinstrumente** wichtig. Dies betreffe insbesondere Personen, die sich bereits maßgeblich im Bereich Datenmanagement engagieren. Für diese müssten geeignete Systeme entwickelt werden, z. B. durch die Schaffung definierter Karrierepfade, verlässlicher Berufsbezeichnungen und klar damit verbundener Kompetenzprofile. Nur wenn die Reputation und die Einkommen dieser Datenexpert*innen gesteigert werden, würde die Attraktivität dieses Berufsfeldes so erhöht, dass die aktuellen und zukünftigen personellen Bedarfe abgedeckt werden können.

Insgesamt wurden die **Berufs- und Karriereaussichten** für Datenexpert*innen (z. B. Datenmanager*innen, Datenanalyst*innen) als hervorragend angesehen. Dies habe sich vor wenigen Jahren noch anders dargestellt. In diesem Zusammenhang wurde mehrfach darauf hingewiesen, dass bereits jetzt zahlreiche Stellen für Datenexpert*innen aller Kompetenzstufen nicht besetzt werden können und diese Situation sich in den nächsten Jahren noch deutlich verschärfen werde.

Eng verknüpft damit sei ein Themenkomplex, der hier als **Motivation und Nutzen** bezeichnet wird. Nach Einschätzung der Diskutant*innen scheint der Nutzen von Datenteilung und -nachnutzung für viele Personen – unabhängig davon, ob sie aus Wirtschaft, Wissenschaft oder dem Public Sector kommen – nicht klar genug erkennbar. Die konkrete Frage laute hier: „Was habe ich davon?“. Unsicherheiten über den Nutzen des Teilens oder Nachnutzens von Daten wirkten sich insbesondere hemmend aus, weil Führungskräfte gezwungen sind, mit begrenzten Ressourcen hauszuhalten. Wenn ein direkter Nutzen nicht abgeleitet werden kann, sinkt naturgemäß die Motivation, sich zusätzliche Aufgaben und Verantwortungsbereiche aufzubürden. Ein Beispiel, welches in der Diskussion zirkulierte, betraf den zusätzlichen Aufwand für die adäquate Aufbereitung von Daten für die Nachnutzung durch Dritte. Dieser Aufwand und die Aussicht auf eine tatsächliche Nachnutzung werde vielfach als schwer kalkulierbar eingeschätzt. Gerade in Zeiten, in welchen dem effizienten Einsatz von Mitteln hohe Priorität eingeräumt wird, bleibe wenig Handlungsspielraum für Risikoinvestitionen.

Die Nachnutzung von Daten hat ein Mindestmaß an **Harmonisierung und Standardisierung** als Voraussetzung. Bezüglich des urheberrechtlichen, einschließlich des leistungsschutzrechtlichen Schutzes von Daten

sprachen sich die Diskutant*innen für die Nutzung der global verständlichen Lizenzen CC 0 (zero) und CC BY aus.² Die Zustimmung zur Nutzung dieser Lizenzen wurde mit der Warnung verbunden, das Entstehen eines „Lizenz-Dschungels“, wie es im Softwarebereich zu beklagen sei, zu vermeiden. Als wichtig wurde auch die (technische) Standardisierung von Schnittstellen zum Datenaustausch genannt. Grundsätzlich sollten sich Akteur*innen aus allen Bereichen auf nationaler und internationaler Ebene für die Entwicklung und Nutzung von Standards einsetzen.

Einigkeit bestand auch hinsichtlich der Bedeutung der **Vernetzung** von Anbietenden und Nutzenden von Daten. Daten-Anbietende und -Nachnutzende sollten dazu motiviert bzw. dabei unterstützt werden, ihre Möglichkeiten und Erwartungen stärker abzustimmen, um so das Teilen von Daten zielgerichteter befördern zu können. Als Motivationsbremse für das Zugänglichmachen von Daten wurde die Unsicherheit darüber, ob und von wem für welche Zwecke bestimmte Daten genutzt werden bzw. genutzt werden könnten eingeschätzt. Hier durch Vernetzung mehr konkrete Perspektiven zu eröffnen, wäre nach Einschätzung der Podiumsgäste eine geeignete Maßnahme zur Förderung der Nachnutzung von Daten.

Die Entwicklung von **Geschäftsmodellen** für die Nachnutzung von Daten ist offensichtlich ein Anliegen von Unternehmen. In der Diskussion wurde jedoch deutlich, dass im Kontext einer Konkretisierung von Geschäftsmodellen schnell auch Risiken, wie beispielsweise eine Verletzung von Datenschutzvorschriften, gesehen werden. Daher wurde vorgeschlagen, der Gesetzgeber solle seine Handlungsspielräume bei der Regulation dazu nutzen, spezifische Freiräume für kleinere und mittlere Unternehmen inklusive Start-ups zu schaffen. Daneben wurde in der Diskussion auch deutlich, dass die Entwicklung von Geschäftsmodellen im Kontext der Nachnutzung von Daten in einem engen Zusammenhang mit Vernetzung steht. Die Bereitschaft in Unternehmen zum Teilen von Daten und zu Investitionen zu vorgelagerter Kuratierung von dieser Daten sei eng mit der Identifikation konkreter Vermarktungschancen verknüpft. Um diese Chancen zu erkennen oder zu schaffen, sollten Unternehmen sich stärker mit potenziellen Kund*innen vernetzen.

Das Funktionieren eines Geschäftsmodells hänge auch davon ab, dass die Waren oder Dienstleistungen, für die ein Markt gesehen wird, zu einem für die Kund*innen akzeptablen Preis angeboten werden können. In der Diskussion wurde deshalb angesprochen, dass Unternehmen, die das Teilen oder die Nachnutzung von Daten erwägen, deshalb ein Interesse an der Existenz

von Standards haben, denn diese könnten Kostensenkungen sowohl bei der Kuratierung als auch bei der Nachnutzung von Daten ermöglichen.

Die Podiumsteilnehmer*innen forderten einmütig eine auskömmliche und dauerhafte **Finanzierung** für ein nachhaltiges Datenmanagement einschließlich der hierfür benötigten Infrastrukturen. In diesem Zusammenhang wurde auch angesprochen, dass teilweise durch eine zweckdienlichere Verwendung vorhandener Mittel eine Verbesserung der Situation erreicht werden könne. Die europäische Open-Data-Gesetzgebung verpflichtet die Mitgliedstaaten der Union dazu, Daten, die von ihren öffentlichen Verwaltungen gehalten werden, zu Grenzkosten oder kostenfrei zugänglich zu machen (zu öffentlichen Daten allgemein siehe: Richtlinie (EU) 2019/1024, 2019, (4), (36), (37); zu Forschungsdaten siehe: ebd., (27), (28)). Für Kulturinstitutionen in öffentlicher Trägerschaft gelten weniger strikte Regeln (ebd., (38)). Beschränkung hinsichtlich der Möglichkeit, über die Zugänglichmachung von Daten Einnahmen zu erzielen, kombiniert mit der gesetzlichen Verpflichtung zum Aufbau von Infrastrukturen zur Archivierung und Zugänglichmachung von Daten führen im Public Sector einschließlich der öffentlich finanzierten Wissenschaft seit vielen Jahren zur Forderung nach einer Steigerung der hierfür benötigten finanziellen Ressourcen. Diese Forderung wurde auch in der Podiumsdiskussion mehrfach erhoben. Dabei gehe es insbesondere in der Wissenschaft nicht nur um den Umfang gewährter finanzieller Mittel, sondern auch um die Dauer, für die diese Mittel gewährt werden und darum, bei der Vergabe der Mittel das Schaffen von „Parallel-Infrastrukturen“ zu vermeiden. Die in der Wissenschaft verbreitete Finanzierung über Projektmittel stünde in einem deutlichen Missverhältnis zur auf Dauer angelegten Aufgabe der Bewahrung und Zugänglichmachung von Daten. In der Podiumsdiskussion wurde darüber hinaus auch deutlich, dass die öffentlichen Dateninfrastrukturen nicht nur unterfinanziert sind, ihnen fehle vielfach auch qualifiziertes Personal. Dieser Personalmangel betreffe ebenfalls Unternehmen.

Ein größerer Raum zur Diskussion wurde auch dem Themenkomplex **Operationalisierung** eingeräumt. Große Einigkeit bestand darin, dass trotz angesprochener Hemmnisse nennenswerter Handlungsspielraum zur Steigerung des Umfangs geteilter bzw. nachgenutzter Daten bestehe. Zur Nutzung dieses Handlungsspielraumes empfahlen die Podiumsgäste zunächst Investitionen in die Erhöhung der Datenkompetenz von Mitarbeiter*innen. Im Folgenden bejahten die Podiumsteilnehmer*innen auch die Frage des Moderators, ob es zielführend sei, für eine

² Nicht angesprochen wurden die Datenlizenzen Deutschland. Hier ist der Hinweis angebracht, dass die Datenlizenzen Deutschland ein Beispiel für mehrere Ansätze aus Deutschland zur Etablierung liberaler Lizenzen sind, die sich international nicht durchgesetzt haben.

Nutzung bestehender Handlungsspielräume eine „Koalition von Willigen“ zu bilden. Die Aussage, es bestünde weniger Bedarf an neuen Regeln zur Zugänglichmachung von Daten als an Entschlossenheit zur Umsetzung bestehender Regeln traf auf große Zustimmung.

Auf der praktischen Ebene benannten die Diskutant*innen auch die komplexen Rechtsfragen, die sich im Kontext des Teilens von Daten stellen, als ständige Herausforderung. Der Public Sector und die öffentlich finanzierte Wissenschaft seien am ehesten in der Lage, durch eine durchgehende Lizenzierung der von ihnen zugänglich gemachten Daten eine nachhaltige Verbesserung der Situation zu bewirken. Diese Einschätzung sollte Unternehmen jedoch nicht davon entlasten, ihrerseits Möglichkeiten zur Zugänglichmachung von Daten beherzter zu nutzen, im Interesse der Generierung von Gewinnen, aber auch im öffentlichen Interesse. Gesetzliche Verpflichtungen von Unternehmen zum Teilen von Daten, die aktuell auch im Rahmen von Gesetzgebungsverfahren diskutiert werden, wurden von den Diskutant*innen nicht angesprochen.

Auf die Frage des Moderators, welche Aspekte bisher noch nicht genannt wurden und wie es konkret weitergehen könnte, wurde eine Reihe von Entwicklungsmöglichkeiten genannt. Als Desiderat wurde empfunden, dass es immer noch keinen definierten **Kompetenzkatalog** für bestimmte Tätigkeitsbereiche im Lebenszyklus von Daten in Deutschland gibt. Eng verknüpft damit war auch die Forderung nach der Entwicklung entsprechender **Referenz-Curricula**, die zumindest den Kanon der verschiedenen Kompetenzbereiche definieren sollten.

Thematisiert wurde auch die Bedeutung von **Standards**. Diese seien zwar bereits zahlreich vorhanden, würden aber nicht immer im wünschenswerten Umfang genutzt. Unter Umständen seien sie nicht bekannt (genug), es könnte auch Berührungängste geben, Standards aus anderen Communities zu nutzen oder konkrete Anwendungsbeispiele fehlen bzw. sind nicht bekannt. Hier könne durch die Schaffung eines Katalogs und Nachweissystems mit den wichtigen Standards sowie korrespondierender Anwendungs- und Community-spezifischer Praxisbeispiele schnell und unkompliziert Abhilfe geschaffen werden. Länger wurde darüber diskutiert, dass „wir endlich mal ins Doing kommen sollten“. Es müssten nicht immer alle zu 100 Prozent abgeholt werden. In der Diskussion zeigte sich auch, dass dringend motivierende Praxisbeispiele benötigt werden. Als ein konkretes Desiderat bzw. als Beispiel für ein jetzt anzupackendes Projekt wurde die Bereitstellung von **Trainingsdaten**, gerade

im Umfeld der Künstlichen Intelligenz genannt. Datenmanagementsysteme könnten hier (neu) entstehen, die dann Eingang in den praktischen Alltag vieler finden können. Wichtig hier wäre dann natürlich auch die rechtssichere Nutzung dieser Trainingsdaten. Dies würde auch dazu beitragen, **Entwicklungsfreiräume** zu schaffen, um einfach „auch mal etwas auszuprobieren“, was dann unter Umständen als Best Practice Eingang in andere Communitys findet und zur Nachahmung animiert.

Angeregt wurde, dem Thema **Lizenzierung**, das in den Empfehlungen nur einmal angesprochen wird, mehr Aufmerksamkeit zu schenken. Ohne dass dies in der Diskussion ausgesprochen wurde, kam hier zum Ausdruck, dass sich im Kontext der rechtssicheren Nutzung von Daten bzw. einer klaren Erläuterung ihrer jeweiligen rechtlichen Situation regelmäßig nicht nur urheberrechtliche Fragen stellen. Letztere können durch die Nutzung von CC-Lizenzen gut geklärt werden. Für die Kommunikation weiterer Einschränkungen der Nutzungsmöglichkeiten aufgrund anderer gesetzlicher oder vertraglicher Regelungen fehlen standardisierte Kennzeichnungen.

Insgesamt wurden die 31 EcoDM-Empfehlungen von den Diskussionsteilnehmer*innen als sehr umfassend und relevant beschrieben. An verschiedenen Punkten innerhalb der Diskussion wurde angeregt, die Empfehlungen pointierter zu formulieren. Diese Anregungen verdichteten sich jedoch nicht zu einer übergreifenden Einschätzung, die Empfehlungen seien zu weich formuliert. Einig waren sich dagegen alle Teilnehmer*innen, dass es äußerst hilfreich wäre, zu jeder dieser Empfehlungen konkrete Beispiele aus den verschiedenen Untersuchungsbereichen genannt zu bekommen. Auch wenn es nicht immer möglich sei, aus allen vier Untersuchungsbereichen – Wirtschaft, Wissenschaft, Public Sector und Qualifizierung – gemeinsame Best Practices, Standards, Umsetzungs-szenarien etc. zu listen, wäre es doch wünschenswert, einen **Katalog des (potenziell) Machbaren** pro Empfehlung und – wo sinnvoll – nach Community einsehen zu können.

Insgesamt fanden die Empfehlungen breite Unterstützung und Zustimmung. Angeregt wurde eine Anreicherung jeder Empfehlung durch ein praxisnahes Beispiel, das, wo möglich, die Spezifika der jeweiligen Community adressiert. Dies ließ sich allerdings im Rahmen des aktuellen Projektes nicht mehr realisieren.

Die Gedanken aus der Podiumsdiskussion aufgreifend, lautet die Schlussempfehlung aus dem EcoDM-Projekt: Endlich ins Machen kommen, auch wenn nicht gleich alles perfekt ist!

LITERATUR- UND QUELLEN- VERZEICHNIS



9. Literatur- und Quellenverzeichnis

- Ambrasat, J., & Heger, C. (2020). *Barometer für die Wissenschaft. Ergebnisse der Wissenschaftsbefragung 2019/20*. Deutsches Zentrum für Hochschul- und Wissenschaftsforschung GmbH. https://www.wb.dzhw.eu/downloads/wibef_barometer2020.pdf
- Anhalt-Depies, C., Stenglein, J. L., Zuckerberg, B., Townsend, P. A., & Rissman, A. R. (2019). Tradeoffs and tools for data quality, privacy, transparency, and trust in citizen science. *Biological Conservation*, 238, 108195. <https://doi.org/10.1016/j.biocon.2019.108195>
- Ausbildung und Qualifikation – Forschungsdaten.org. (o.J.). Abgerufen 21. Februar 2022, von https://www.forschungsdaten.org/index.php/Ausbildung_und_Qualifikation
- Bandtel, M., Kauz, L., & Weißker, N. (2021). Data Literacy Education für Studierende aller Fächer. Kompetenzziele, curriculare Integration und didaktische Ausgestaltung interdisziplinärer Lehr-Lern-Angebote. In Hochschulforum Digitalisierung (Hrsg.), *Digitalisierung in Studium und Lehre gemeinsam gestalten: Innovative Formate, Strategien und Netzwerke* (S. 395 – 412). Springer Fachmedien. https://doi.org/10.1007/978-3-658-32849-8_23
- Baums, A. (2015). Analyse – Was sind digitale Plattformen. In A. Baums, M. Schössler, & B. Scott (Hrsg.), *Industrie 4.0: Wie digitale Plattformen die Wirtschaft verändern – und wie die Politik gestalten kann* (S. 14 – 25). <http://plattform-maerkte.de/wp-content/uploads/2015/11/Kompendium-High.pdf>
- Biernacka, K., Buchholz, P., Danker, S. A., Dolzycka, D., Endelhardt, C., Helbig, K., Jacob, J., Neumann, J., Odebrecht, C., Petersen, B., Slowig, B., Trautwein-Bruns, U., Wiljes, C., & Wuttke, U. (2021). *Train-the-Trainer-Konzept zum Thema Forschungsdatenmanagement*. <https://doi.org/10.5281/ZENODO.5773203>
- Bierwirth, M., Glöckner, F. O., Grimm, C., Schimmler, S., Boehm, F., Busse, C., Degkwitz, A., Koepler, O., & Neuroth, H. (2020). *Leipzig-Berlin-Erklärung zu NFDI-Querschnittsthemen der Infrastrukturentwicklung*. <https://doi.org/10.5281/zenodo.3895209>
- Blank, M. (2019). Open Data Maturity Report 2019. European Commission, Directorate-General of Communications Networks, Content and Technology. <https://doi.org/10.2830/073835>
- Borges, G., & Hilber, M. (2021). *BeckOK IT-Recht*.
- Botta, J. (2021). Die Datenverkehrsfreiheit – Ein Beitrag zur Schutzgutdebatte im Datenschutzrecht. *Das Deutsche Verwaltungsblatt (DVBl.)*, 136(5), 290 – 296.
- Brenner, T. (2021). Kapitel 1: Wettbewerb und Digitalwirtschaft sowie weitere materiellrechtliche Änderungen. In F. Bien, T. Käseberg, G. Klumpe, T. Körber, & K. Ost (Hrsg.), *Die 10. GWB-Novelle: das neue Kartellrecht* (S. 11 – 110). C. H. Beck.
- Brink, S., & Wolff, H. A. (2021). *BeckOK Datenschutzrecht* (38. Edition, Stand: 01.11.2021). C. H. Beck. https://beck-online.beck.de/?vpath=bibdata/komm/BeckOKDatenS_38/cont/BeckOKDatenS.htm
- Bröckling, U. (2007). *Das unternehmerische Selbst: Soziologie einer Subjektivierungsform* (1. Aufl.). Suhrkamp.
- Bruch, C., Neuroth, H., & Senst, H. (2022). *Data At Your Fingertips: Wegmarken einer exzellenten und vertrauenswürdigen Datenlandschaft. Empfehlungen des EcoDM-Projekts*. <https://doi.org/10.5281/zenodo.5837408>
- Bundesministerium des Innern, für Bau und Heimat. (2021). *Cybersicherheitsstrategie für Deutschland 2021*. <https://www.bmi.bund.de/SharedDocs/downloads/DE/veroeffentlichungen/2021/09/cybersicherheitsstrategie-2021.pdf>
- Bundesministerium des Innern, für Bau und Heimat. (2021). *Open-Data-Strategie der Bundesregierung*. <https://www.bundesregierung.de/resource/blob/992814/1940386/1d269a2ad1b6346fcf60663bdea9c9f8/2021-07-07-open-data-strategie-data.pdf>
- Bundesministerium für Bildung und Forschung. (2021, Juni 16). *Karliczek: Mit FAIR-Data Spaces schaffen wir einen gemeinsamen Datenraum für Wissenschaft und Wirtschaft*. PRESSEMITTEILUNG: 129/2021. Abgerufen 01. März 2022, <https://www.bmbf.de/bmbf/shareddocs/pressemitteilungen/de/karliczek-mit-fair-data-spaces-er-wissenschaft-und-wirtschaft.html>
- Bundesministerium für Bildung und Forschung. (2018). *Forschung und Innovation für die Menschen. Die Hightech-Strategie 2025*. https://www.bmbf.de/SharedDocs/Publikationen/de/bmbf/1/31431_Forschung_und_Innovation_fuer_die_Menschen.pdf?__blob=publicationFile&v=6
- Bundesministerium für Wirtschaft und Energie. (2020). *Digitale Souveränität und Resilienz: Voraussetzungen, Treiber und Maßnahmen für mehr Nachhaltigkeit. Digital-Gipfel-Plattform 2 „Innovative Digitalisierung der Wirtschaft“*. Fokusgruppe Digitale Souveränität (Digital Gipfel 2020). <https://www.de.digital/DIGITAL/Redaktion/DE/Digital-Gipfel/Download/2020/digitale-souveraenitaet-und-resilienz.pdf>
- Bundesministerium für Wirtschaft und Energie. (2019). *Das Projekt GAIA-X. Eine vernetzte Dateninfrastruktur als Wiege eines vitalen, europäischen Ökosystems*. <https://www.bmwi.de/Redaktion/DE/Publikationen/Digitale-Welt/das-projekt-gaia-x.pdf>

- Bundesministerium für Wirtschaft und Klimaschutz, & Ministère de l'Économie, des Finances et de la Relance. (2021). *Gemeinsame Pressemitteilung: IPCEI Cloud geht in die nächste Phase – Interessenbekundungsverfahren in Deutschland gestartet und Beginn der Vorbereitungen zum europäischen Matchmaking*. Abgerufen am 28. Februar 2022, von <https://www.bmwi.de/Redaktion/DE/Pressemitteilungen/2021/07/20210709-ipcei-cloud-geht-in-die-naechste-phase.html>
- Bundesministerium für Wirtschaft und Klimaschutz. (2021). *IPCEI Nächste Generation Cloud Infrastrukturen und Services*. Abgerufen am 28. Februar 2022, von <https://www.bmwi.de/Redaktion/DE/Artikel/Industrie/ipcei-cis.html>
- Bundesministerium für Wirtschaft und Klimaschutz. (o.J.). *Smart Data – Innovationen aus Daten*. Abgerufen 21. Februar 2022, von <https://www.bmwi.de/Redaktion/DE/Artikel/Digitale-Welt/smart-data.html>
- Bundesregierung. (2019). *Eckpunkte einer Datenstrategie der Bundesregierung*. <https://www.bundesregierung.de/resource/blob/974430/1693626/60b196d5861f71cdefb9e254f5382a62/2019-11-18-pdf-datenstrategie-data.pdf>
- Bundesregierung. (2021a). *Datenstrategie der Bundesregierung. Eine Innovationsstrategie für gesellschaftlichen Fortschritt und nachhaltiges Wachstum* (Bundeskanzleramt, Hrsg.). <https://www.bundesregierung.de/resource/blob/992814/1845634/45aee6da9554115398cc6a722aba08cb/datenstrategie-der-bundesregierung-download-bpa-data.pdf>
- Bundesregierung. (2021b). *Digitalisierung gestalten. Umsetzungsstrategie der Bundesregierung* (Presse- und Informationsamt der Bundesregierung, Hrsg.). <https://www.bundesregierung.de/resource/blob/992814/1605036/ad8d8a0079e287f694f04cbc93f591/digitalisierung-gestalten-download-bpa-data.pdf>
- Burgelman, J.-C., Pascu, C., Szkuta, K., Von Schomberg, R., Karalopoulos, A., Repanas, K., & Schoupe, M. (2020). *Open Science, Open Data and Open Scholarship: European Policies to Make Science Fit for the 21st Century* (SSRN Scholarly Paper ID 3522398). Social Science Research Network. <https://dx.doi.org/10.2139/ssrn.3522398>
- Busch, C. (2019). *Der Mittelstand in der Plattformökonomie: mehr Fairness für KMU auf digitalen Märkten*. Friedrich-Ebert-Stiftung, Abteilung Wirtschafts- und Sozialpolitik. <https://library.fes.de/pdf-files/wiso/15493.pdf>
- Carroll, S. R., Hudson, M., Chapman, J., Figueroa-Rodríguez, O. L., Holbrook, J., Lovett, R., Materechera, S., Parsons, M., Raseroka, K., Rodriguez-Lonebear, D., Rowe, R., Sara, R., & Walker, J. (2019). *Die CARE-Prinzipien für indigene Data Governance*. <https://zenodo.org/record/5995059>
- Chen, X., Képuska, E., Mauw, S., & Ramírez-Cruz, Y. (2020). Active Re-identification Attacks on Periodically Released Dynamic Social Graphs. In L. Chen, N. Li, K. Liang, & S. Schneider (Hrsg.), *Computer Security – ESORICS 2020* (S. 185 – 205). Springer International Publishing. https://doi.org/10.1007/978-3-030-59013-0_10
- Country Fact-Sheet 2021 Germany. (2021). https://data.europa.eu/sites/default/files/country-factsheet_germany_2021.pdf
- Country Fact-Sheet 2020 Germany. (2020). https://data.europa.eu/sites/default/files/country-factsheet_germany_2020.pdf
- Country Fact-Sheet 2019 Germany. (2019). https://data.europa.eu/sites/default/files/country-factsheet_germany_2019.pdf
- Deiningner, E. K. (2018). Dezentrale Ökosysteme und Plattformen als Chance für Deutschland: Wir müssen aufwachen, größere Offenheit zeigen und gemeinsam handeln. In Begleitforschung Smart Data (Hrsg.), *Big Data, Smart Data, next?* (S. 110–113). https://www.digitale-technologien.de/DT/Redaktion/DE/Downloads/Publikation/Smart-Data/2018_08_19_Big_Data_Smart_Data_next.pdf
- Deuring, S. (2021). *Datenmacht: Zugang als Herausforderung* (1. Auflage). Nomos.
- Deutsche Forschungsgemeinschaft. (2021). *Checkliste zum Umgang mit Forschungsdaten*. https://www.dfg.de/download/pdf/foerderung/grundlagen_dfg_foerderung/forschungsdaten/forschungsdaten_checkliste_de.pdf
- Deutsche Forschungsgemeinschaft. (2020). *Digitaler Wandel in den Wissenschaften. Impulspapier*. <https://doi.org/10.5281/zenodo.4191345>
- Deutsche Forschungsgemeinschaft. (2019). *Leitlinien zur Sicherung guter wissenschaftlicher Praxis – Kodex*. https://www.dfg.de/download/pdf/foerderung/rechtliche_rahmenbedingungen/gute-wissenschaftliche_praxis/kodex_gwp.pdf
- Diaz-Bone, R. (2015). *Die „Economie des conventions“: Grundlagen und Entwicklungen der neuen französischen Wirtschaftssoziologie*. Springer VS.
- Diaz-Bone, R., & Salais, R. (2012). The économie des conventions: transdisciplinary discussions and perspectives ; introduction to the HSR focus. *Historical Social Research*, 37(4), 9 – 14. <https://doi.org/https://doi.org/10.12759/hsr.37.2012.4.9-14>
- Directorate-General for Research and Innovation (European Commission), & PwC EU Services. (2018). *Cost-Benefit Analysis for FAIR Research Data: Cost of Not Having FAIR Research Data*. Publications Office of the European Union. <https://data.europa.eu/doi/10.2777/02999>
- Directorate-General for Research and Innovation (European Commission), Lawrence, R. & Mendez, E. (2020). *Progress on open science: towards a*

- shared research knowledge system: final report of the open science policy platform*. Publications Office of the European Union. <https://data.europa.eu/doi/10.2777/00139>
- Djeffal, C. (2021). § 25 Öffentliche Verwaltung und Smart Cities. In P. Bräutigam & T. Kraul (Hrsg.), *Internet of Things*. (S. 925 – 944) C. H. Beck.
- Djeffal, C. (2019). „Privatheit 4.0“ im Spiegel von Recht und künstlicher Intelligenz. Das Recht als (Re)aktion und der status activus technicus. In H. Behrendt, W. Loh, T. Matzner, & C. Misselhorn (Hrsg.), *Privatsphäre 4.0* (S. 177 – 197). J.B. Metzler. https://doi.org/10.1007/978-3-476-04860-8_11
- Djeffal, C. (2018). Normative Leitlinien für künstliche Intelligenz in Regierung und öffentlicher Verwaltung. In R. Mohabbat Kar, B. E. P. Thapa, & P. Parycek (Hrsg.), *(Un)berechenbar? Algorithmen und Automatisierung in Staat und Gesellschaft* (S. 493 – 515). Fraunhofer-Institut für Offene Kommunikationssysteme FOKUS, Kompetenzzentrum Öffentliche IT (ÖFIT). <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-57618-7>
- Dreier, T., & Schulze, G. (2018). *Urheberrechtsgesetz: Verwertungsgesellschaftengesetz, Kunsturhebergesetz: Kommentar / von Dr. Thomas Dreier, Dr. Gernot Schulze unter Mitwirkung von Dr. Louisa Specht* (6. Auflage). C. H. Beck.
- Engelhardt, C., & Kusch, H. (2021). 5.3 Kollaboratives Arbeiten mit Daten. In M. Putnings, H. Neuroth, & J. Neumann (Hrsg.), *Praxishandbuch Forschungsdatenmanagement* (S. 451 – 476). De Gruyter Saur. <https://doi.org/10.1515/9783110657807-025>
- Engels, B. (2021). *Cybersicherheit: 52,5 Mrd. Euro Schaden durch Angriffe im Homeoffice* (IW-Kurzbericht Nr. 54/2021). Institut der deutschen Wirtschaft. <https://www.iwkoeln.de/studien/barbara-engels-525-mrd-euro-schaden-durch-angriffe-im-homeoffice-518890.html>
- Europäische Kommission. (2021). *ANHANG der Mitteilung der Kommission an das Europäische Parlament, den Rat, den Europäischen Wirtschafts- und Sozialausschuss und den Ausschuss der Regionen Digitaler Kompass 2030: der europäische Weg in die digitale Dekade*. https://eur-lex.europa.eu/resource.html?uri=cellar:12e835e2-81af-11eb-9ac9-01aa75ed71a1.0016.02/DOC_2&format=PDF
- Europäische Kommission. (2020b). *Vorschlag für eine Verordnung des europäischen Parlaments und des Rates über europäische Daten-Governance (Daten-Governance-Gesetz)*. <https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:52020PC0767>
- Europäische Kommission. (2020a). *Eine europäische Datenstrategie* (MITTEILUNG DER KOMMISSION AN DAS EUROPÄISCHE PARLAMENT, DEN RAT, DEN EUROPÄISCHEN WIRTSCHAFTS-UND SOZIALAUSSCHUSS UND DEN AUSSCHUSS DER REGIONEN). Europäische Kommission. <https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:52020DC0066>
- Europäische Kommission. (2010). *Eine Digitale Agenda für Europa* (MITTEILUNG DER KOMMISSION AN DAS EUROPÄISCHE PARLAMENT, DEN RAT, DEN EUROPÄISCHEN WIRTSCHAFTS- UND SOZIALAUSSCHUSS UND DEN AUSSCHUSS DER REGIONEN). [https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX:52010DC0245R\(01\)](https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX:52010DC0245R(01))
- Europäische Kommission. (2005). *i2010 – Eine europäische Informationsgesellschaft für Wachstum und Beschäftigung*. <https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=LEGISSUM%3Ac11328>
- European Commission. (2021). *Ethics-and-Data-Protection*. https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/ethics-and-data-protection_he_en.pdf
- European Commission. Directorate General for Communications Networks, Content and Technology. (2020). *Towards a European strategy on business-to-government data sharing for the public interest: final report prepared by the High Level Expert Group on Business to Government Data Sharing*. Publications Office. <https://data.europa.eu/doi/10.2759/731415>
- European Commission. Directorate General for Communications Networks, Content and Technology, IDC & The Lisbon Council. (2020). *The European data market monitoring tool: key facts & figures, first policy conclusions, data landscape and quantified stories: d2.9 final study report*. Publications Office. <https://doi.org/10.2759/72084>
- European Commission. Directorate General for Research and Innovation. (2018). *Mission-oriented research & innovation in the European Union: a problem solving approach to fuel innovation led growth*. Publications Office. <https://data.europa.eu/doi/10.2777/360325>
- European Data Protection Board. (2021). *EDPB-EDPS Joint Opinion 03/2021 on the Proposal for a regulation of the European Parliament and of the Council on European data governance (Data Governance Act) Version 1.1*.
- European Open Science Cloud (EOSC). (o. J.). European Commission. Abgerufen 9. Februar 2022, von https://ec.europa.eu/info/research-and-innovation/strategy/strategy-2020-2024/our-digital-future/open-science/european-open-science-cloud-eosc_en
- Falck, O., & Koenen, J. (2020). *Rohstoff „Daten“: volkswirtschaftlicher Nutzen von Datenbereitstellung – eine Bestandsaufnahme: Studie mit finanzieller Unterstützung der IHK für München und Oberbayern im Rahmen des Vertrags zur Erstellung volkswirt-*

- schaftlicher Studien*. ifo Institut, Leibniz-Institut für Wirtschaftsforschung an der Universität München e.V. <https://www.ifo.de/publikationen/2020/monographie-autorenschaft/rohstoff-daten-volkswirtschaftlicher-nutzen-von>
- Fraunhofer-Institut für Software- und Systemtechnik ISST. (2019). *Data Economy. Status Quo der deutschen Wirtschaft & Handlungsfelder in der Data Economy* (DEMAND. Data Economics and Management of Data driven Business) [White Paper]. [https://www.demand-projekt.de/paper/DEMAND-DataEconomicsAndManagementOfDataDrivenBusiness\(WhitePaper\).pdf](https://www.demand-projekt.de/paper/DEMAND-DataEconomicsAndManagementOfDataDrivenBusiness(WhitePaper).pdf)
- Freeman, R. E. (1984). *Strategic management : a stakeholder approach*. Pitman.
- Gebietskörperschaft: Definition, Begriff und Erklärung. (o.J.). In *JuraForum.de*. Abgerufen 9. Februar 2022, von <https://www.juraforum.de/lexikon/gebietskoerperschaft>
- Gegenhuber, T. (2020). *Eine Vision für das digitale Europa: von der widerspenstigen Zähmung der Plattformen zu einem digitalen Humanismus* (Friedrich-Ebert-Stiftung, Hrsg.). <http://library.fes.de/pdf-files/fes/16146.pdf>
- Gehring, P., & Lange, S. (2020, Oktober 2). Bitte nicht auf Autopilot schalten! *Jan-Martin Wiarda*. Abgerufen 01. März 2022, <http://www.jmwiarda.de/2020/10/02/bitte-nicht-auf-autopilot-schalten/>
- Gentemann, L., & Heidkamp, P. (2021). *Cloud-Monitor 2021. Eine Studie von Bitkom Research im Auftrag von KPMG* Pressekonferenz. https://www.bitkom-research.de/system/files/document/Bitkom_KPMG_Charts_Cloud%20Monitor%202021_final.pdf
- Gesetz über die Weiterverwendung von Informationen öffentlicher Stellen (Informationsweiterverwendungsgesetz – IWG)*, (2006). http://www.bgbl.de/xaver/bgbl/start.xav?startbk=Bundesanzeiger_BGBl&jumpTo=bgbl121s2941.pdf
- Gesetz zur Änderung des E-Government-Gesetzes und zur Einführung des Gesetzes für die Nutzung von Daten des öffentlichen Sektors*, (2021). https://www.bgbl.de/xaver/bgbl/text.xav?SID=&tf=xaver.component.Text_0&toctf=&qmf=&hlf=xaver.component.Hitlist_0&bk=bgbl&start=%2F%2F%5B%40node_id%3D%27940341%27%5D&skin=pdf&tlevel=2&nohist=1&sinst=E2A27D6D
- Gesetz zur Regelung des Zugangs zu Informationen des Bundes (Informationsfreiheitsgesetz – IFG)*, (2009). <https://www.gesetze-im-internet.de/ifg/IFG.pdf>
- Glaser, B. G., & Strauss, A. L. (1998). *The discovery of grounded theory: strategies for qualitative research*. Aldine Transaction.
- Gläser, J., & Laudel, G. (2010). *Experteninterviews und qualitative Inhaltsanalyse als Instrumente rekonstruierender Untersuchungen* (4. Aufl.). VS Verlag für Sozialwissenschaften.
- Gong, M., Wang, S., Wang, L., Liu, C., Wang, J., Guo, Q., Zheng, H., Xie, K., Wang, C., & Hui, Z. (2020). Evaluation of Privacy Risks of Patients' Data in China: Case Study. *JMIR Medical Informatics*, 8(2), e13046. <https://doi.org/10.2196/13046>
- Hartl, A., & Ludin, A. (2021). Was die Datenstrategien der EU sowie der Bundesregierung für die Gesetzgebung erwarten lassen., *MultiMedia und Recht*, 24(7), 534 – 538.
- Heckmann, D. (2021). Digitale Verwaltung. In D. Heckmann & A. Paschke (Hrsg.), *Juris-PraxisKommentar Internetrecht*. Juris.
- Heidrich, D. J., Bauer, P., & Krupka, D. (2018). *FUTURE SKILLS: ANSÄTZE ZUR VERMITTLUNG VON DATA LITERACY IN DER HOCHSCHULBILDUNG*. Zenodo. <https://doi.org/10.5281/zenodo.1413119>
- Helfferrich, C. (2019). Leitfaden- und Experteninterviews. In N. Baur & J. Blasius (Hrsg.), *Handbuch Methoden der empirischen Sozialforschung* (S. 669–686). Springer Fachmedien. https://doi.org/10.1007/978-3-658-21308-4_44
- Hessel, S., & Leffer, L. (2020). Rechtlicher Schutz maschinengenerierter Daten. *MultiMedia und Recht*, 23(10), 647–650.
- Hightech-Forum. (2020a). *Offene Wissenschaft und Innovation. Ein Impulspapier aus dem Hightech-Forum**. https://www.hightech-forum.de/wp-content/uploads/htf_impulspapier_offene_wissenschaft.pdf
- Hightech-Forum. (2020b). *Impulspapier: Zukunft der Wertschöpfung*. https://www.hightech-forum.de/wp-content/uploads/hightech-forum_impulspapier_wertsch%C3%B6pfung.pdf
- Hodgkinson, I. R., Mousavi, S., & Hughes, P. (2021). New Development: Citizen Science—Discovering (New) Solutions to Wicked Problems. *Public Money & Management*, 42(2), 133–136. <http://dx.doi.org/10.1080/09540962.2021.1967630>
- IDC Italia srl & The Lisbon Council. (2020). *D2.7 FINAL REPORT ON FACTS AND FIGURES. UPDATING THE EUROPEAN DATA MARKET MONITORING TOOL*. (SMART 2016/0063; UPDATE OF THE EUROPEAN DATA MARKET STUDY). https://datalandscape.eu/sites/default/files/report/EDM_D2.7_FinalReportFactsFigures_02.04.20.pdf
- International Science Council. (2020). *Open Science for the 21st Century. Draft ISC Working Paper June 2020*. https://council.science/wp-content/uploads/2020/06/International-Science-Council_Open-Science-for-the-21st-Century_Working-Paper-2020_compressed.pdf
- ISO/IEC 2382:2015. (2015). Abgerufen 01. März 2022, <https://www.iso.org/standard/63598.html>
- Jülicher, T., Röttgen, C., & v. Schönfeld, M. (2016). Das Recht auf Datenübertragbarkeit. Ein datenschutzrechtliches Novum. *ZD*, 6(8), 358–362.

- Junker, J. (2018). Risiken und Alternativen in der Plattformökonomie. In Begleitforschung Smart Data (Hrsg.), *Big Data, Smart Data, next?* (S. 96–99). https://www.digitale-technologien.de/DT/Redaktion/DE/Downloads/Publikation/Smart-Data/2018_08_19_Big_Data_Smart_Data_next.pdf
- Käseberg, T., Brenner, T., & Fülling, D. (2021). Das GWB-Digitalisierungsgesetz im Überblick. *WuW*, 71(5), 269 – 275.
- Kelle, U., & Kluge, S. (2010). *Vom Einzelfall zum Typus* (2. überarbeitete Auflage). VS Verlag für Sozialwissenschaften. <https://doi.org/10.1007/978-3-531-92366-6>
- Key Skills – EOSC FAIR4S. (2019). Abgerufen 01. März 2022, <https://eosc-fair4s.github.io/keyskills.html>
- Koch, H. (2021, Februar 11). Die Data-Literacy-Charta: Gemeinsam Bildungsprozesse für das 21. Jahrhundert gestalten. *Hochschulforum Digitalisierung – Hochschulbildung im digitalen Zeitalter*. Abgerufen 01. März 2022, <https://hochschulforumdigitalisierung.de/de/blog/data-literacy-charta>
- Kommission der europäischen Gemeinschaften. (2002). *eEurope 2005: Eine Informationsgesellschaft für Alle – Aktionsplan zur Vorlage in Hinblick auf den Europäischen Rat von Sevilla am 21./22. Juni 2002*. <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2002:0263:FIN:DE:PDF>
- Kommission der europäischen Gemeinschaften. (2000). *eEurope 2002: Eine Informationsgesellschaft für alle – Entwurf eines Aktionsplans der Europäischen Kommission zur Vorlage auf der Tagung des Europäischen Rates am 19./20. Juni 2000 in Feira*. <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2000:0330:FIN:DE:PDF>
- Kommission Wettbewerbsrecht 4.0. (2019). *Ein neuer Wettbewerbsrahmen für die Digitalwirtschaft. Bericht der Kommission Wettbewerbsrecht 4.0*. Bundesministerium für Wirtschaft und Energie. <https://www.bmwi.de/Redaktion/DE/Publikationen/Wirtschaft/bericht-der-kommission-wettbewerbsrecht-4-0.pdf>
- Kopp, F. O., & Ramsauer, U. (2021). *Verwaltungsverfahrensgesetz: Kommentar* (22., vollständig überarbeitete Auflage). C.H. Beck.
- Kraft, S., Schmalen, A., Seitz-Moskaliuk, H., Sure-Vetter, Y., Knebes, J., Lübke, E., & Wössner, E. (2021). Nationale Forschungsdateninfrastruktur (NFDI) e. V.: Aufbau und Ziele. *Bausteine Forschungsdatenmanagement*, 2, 1–9. <https://doi.org/10.17192/bfdm.2021.2.8332>
- Krempel, S. (2021, November 7). Gaia-X-Bruder: EU-Cloud-Großprojekt IPCEI-CIS geht in die Dating-Phase. *heise online*. <https://www.heise.de/news/Gaia-X-Bruder-EU-Cloud-Grossprojekt-IPCEI-CIS-geht-in-die-Dating-Phase-6134101.html>
- Krüger, S., Wiencke, J., & Koch, A. (2020). Der Datenpool als Geschäftsgeheimnis. *GRUR*, 122(6), 578–584.
- Kuhlmann, S., Franzke, J., Dumas, B., & Heine, M. (2021). *Daten als Grundlage für wissenschaftliche Politikberatung*. <https://www.uni-potsdam.de/fileadmin/projects/ls-kuhlmann/Politikberatung/Daten-Politikberatung-Kuhlmann-et-al-13-9-21.pdf>
- Lanier, J. (2021, Januar 25). STANDPUNKT: Daten-Genossenschaften statt Daten-Eigentum. *Tagesspiegel Background*. Abgerufen 01. März 2022, <https://background.tagesspiegel.de/digitalisierung/daten-genossenschaften-statt-daten-eigentum>
- Leimüller, G., Blümel, C., & Fecher, B. (2018). *INNOVATION DURCH STRATEGISCHE OFFENHEIT*. <https://doi.org/10.5281/ZENODO.1879801>
- Leitlinien zum Recht auf Datenübertragbarkeit, 16/EN WP 242 (2016). https://www.datenschutzgrundverordnung.eu/wp-content/uploads/2017/07/WP242de_Art_29-Gruppe_Datenubertragbarkeit.pdf
- Lernende Systeme – Die Plattform für künstliche Intelligenz. (2020). *Von Daten zu Wertschöpfung – Potenziale von daten- und KI-basierten Wertschöpfungsnetzwerken*. Lernende Systeme – Die Plattform für Künstliche Intelligenz. https://www.plattform-lernende-systeme.de/files/Downloads/Publikationen/PLS_Booklet_Datenoekosysteme.pdf
- Lis, D., Tagalidou, N., Lingelbach, K., & Spiekermann, M. (2019). *Positionspapier: Ökosysteme für Daten und künstliche Intelligenz*. Fraunhofer-Gesellschaft. <https://doi.org/10.24406/ISST-N-543753>
- Loewenheim, U., Leistner, M., & Ohly, A. (Hrsg.). (2020). *Urheberrecht: UrhG, KUG, VGG: Kommentar* (6., neu bearbeitete Auflage des von Prof. Dr. Dr. h.c. mult. Gerhard Schricker bis zur 3. Auflage herausgegebenen Werkes). C.H. Beck.
- Mazzucato, M. (2021). *Mission: auf dem Weg zu einer neuen Wirtschaft* (B. Schmid, Übers.). Campus Verlag.
- Mazzucato, M. (2019). *Wie kommt der Wert in die Welt? von Schöpfern und Abschöpfern* (B. Schmid, Übers.). Campus Verlag.
- Mazzucato, M. (2011). *The entrepreneurial state*. Demos. https://www.demos.co.uk/files/Entrepreneurial_State_-_web.pdf
- McKiernan, E. C., Bourne, P. E., Brown, C. T., Buck, S., Kenall, A., Lin, J., McDougall, D., Nosek, B. A., Ram, K., Soderberg, C. K., Spies, J. R., Thaney, K., Updegrave, A., Woo, K. H., & Yarkoni, T. (2016). How open science helps researchers succeed. *eLife*, 5, e16800. <https://doi.org/10.7554/eLife.16800>
- Mohabbat Kar, R., Tiemann, J., & Welzel, C. (2020). *DER STAAT AUF DEM WEG ZUR PLATTFORM Nutzungspotenziale für den öffentlichen Sektor* (FORSCHUNG FÜR DEN DIGITALEN STAAT). Kompetenzzentrum Öffentliche IT Fraunhofer-Institut für

- Offene Kommunikationssysteme FOKUS. <https://www.oeffentliche-it.de/documents/10181/14412/De+r+Staat+auf+dem+Weg+zur+Plattform>
- Nemitz, P., & Pfeffer, M. (2020). *Prinzip Mensch: Macht, Freiheit und Demokratie im Zeitalter der künstlichen Intelligenz*. Dietz.
- Ohly, A. (2019). Das neue Geschäftsgeheimnisgesetz im Überblick. *GRUR*, 121(5), 441 – 451.
- OSB Alliance entwickelt Sovereign Cloud Stack (SCS) für das GAIA-X Projekt. (o. J.). OSBA – Open Source Business Alliance. Abgerufen 28. Januar 2022, von <https://osb-alliance.de/sovereign-cloud-stack-scs>
- Oßwald, A. (2021). 3.5 Barrieren, Hemmschwellen und Gatekeeper. In M. Putnings, H. Neuroth, & J. Neumann (Hrsg.), *Praxishandbuch Forschungsdatenmanagement*. De Gruyter Saur. <https://doi.org/10.1515/9783110657807-016>
- Otto, B., Jürjens, J., Schon, J., Auer, S., Menz, N., Wenz, S., & Cirullies, J. (2016). *INDUSTRIAL DATA SPACE DIGITALE SOUVERÄNITÄT ÜBER DATEN* (White Paper). Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V., Industrial Data Space e.V. https://www.fraunhofer.de/content/dam/zv/de/Forschungsfelder/industrial-data-space/Industrial-Data-Space_whitepaper.pdf
- Otto, B., Steinbuß, S., Teuscher, A., & Lohmann, S. (2019). *Reference Architecture Model – Version 3.0*. International Data Spaces Association. <https://internationaldataspaces.org/wp-content/uploads/IDS-Reference-Architecture-Model-3.0-2019.pdf>
- Paal, B. P., & Pauly, D. A. (Hrsg.). (2021). *Datenschutz-Grundverordnung: Bundesdatenschutzgesetz* (3. Auflage). C.H. Beck.
- PARTHENOS, Hollander, H., Morselli, F., Uiterwaal, F., Admiraal, F., Trippel, T., & Di Giorgio, S. (2019). *PARTHENOS Guidelines to FAIRify data management and make data reusable*. <https://doi.org/10.5281/zenodo.3368858>
- Pawelke, A. (2020). *Daten teilen, aber wie?: Ein Panorama der Datenteilungsmodelle* (Bertelsmann Stiftung, Hrsg.). <https://doi.org/10.11586/2020079>
- Prüfer, J. (2020). *Die Datenteilungspflicht: Innovation und fairer Wettbewerb auf datengetriebenen Märkten* (Friedrich Ebert Stiftung, Hrsg.). <http://library.fes.de/pdf-files/fes/15990.pdf>
- Putnings, M. (2021). 1 Datenökosystem. In M. Putnings, H. Neuroth, & J. Neumann (Hrsg.), *Praxishandbuch Forschungsdatenmanagement* (S. 7 – 10). De Gruyter Saur. <https://doi.org/10.1515/9783110657807-001>
- Rat für Sozial- und Wirtschaftsdaten. (2020). *Handreichung Datenschutz* (2. Auflage). <https://www.konsortswd.de/aktuelles/publikation/handreichung-datenschutz-2/>
- Reinsel, D., Gantz, J., & Rydning, J. (2018). *The Digitization of the World. From Edge to Core* (IDC White Paper #US44413318; Data Age 2025). IDC. <https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf>
- RfII – Rat für Informationsinfrastrukturen. (2021). *Nutzung und Verwertung von Daten im wissenschaftlichen Raum – Empfehlungen zur Ausgestaltung von Datendiensten an der Schnittstelle zwischen Wissenschaft und Wirtschaft*. <https://rfii.de/?p=6961>
- RfII – Rat für Informationsinfrastrukturen. (2016). *Leistung aus Vielfalt. Empfehlungen zu Strukturen, Prozessen und Finanzierung des Forschungsdatenmanagements in Deutschland*. <https://rfii.de/?p=1998>
- RfII – Rat für Informationsinfrastrukturen. (2020b). *Datendienste nachhaltig gestalten. Ein Diskussionsimpuls zur Weiterentwicklung von Forschungsdateninfrastrukturen*. <http://www.rfii.de/?wpdmdl=4313>
- RfII – Rat für Informationsinfrastrukturen. (2020a). *Stellungnahme des Rates für Informationsinfrastrukturen (RfII) Datentreuhandstellen gestalten. Zu Erfahrungen der Wissenschaft*. <http://www.rfii.de/download/rfii-stellungnahme-zu-datentreuhandstellen/>
- RfII – Rat für Informationsinfrastrukturen. (2019b). *Digitale Kompetenzen – dringend gesucht! Empfehlungen zu Berufs- und Ausbildungsperspektiven für den Arbeitsmarkt* Wissenschaft. <https://rfii.de/?p=3883>
- RfII – Rat für Informationsinfrastrukturen. (2019a). *Herausforderung Datenqualität – Empfehlungen zur Zukunftsfähigkeit von Forschung im digitalen Wandel (zweite Auflage)*. <https://rfii.de/?p=4043>
- Richter, H. (2020). Zugang des Staates zu Daten der Privatwirtschaft. *Zeitschrift für Rechtspolitik*, 53(8), 245 – 248.
- Richtlinie (EU) 2019/1024 des Europäischen Parlaments und des Rates vom 20. Juni 2019 über offene Daten und die Weiterverwendung von Informationen des öffentlichen Sektors, 172 OJ L (2019). <http://data.europa.eu/eli/dir/2019/1024/oj/deu>
- Richtlinie 2003/98/EG des europäischen Parlaments und des Rates vom 17. November 2003 über die Weiterverwendung von Informationen des öffentlichen Sektors, 345 OJ L (2003). <https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:32003L0098>
- Richtlinie 2013/37/EU des Europäischen Parlaments und des Rates vom 26. Juni 2013 zur Änderung der Richtlinie 2003/98/EG über die Weiterverwendung von Informationen des öffentlichen Sektors Text von Bedeutung für den EWR, 175 OJ L (2013). <http://data.europa.eu/eli/dir/2013/37/oj/deu>
- Ridsdale, C., Rothwell, J., Smit, M., Bliemel, M., Irvine, D., Kelley, D., Matwin, S., Wuetherick, B., & Ali-Hassan, H. (2015). *Strategies and Best Practices for Data Literacy Education Knowledge Synthesis Report*. <https://doi.org/10.13140/RG.2.1.1922.5044>

- Roles of professional groups – EOSC FAIR4S.* (o. J.). Abgerufen 21. Februar 2022, von <https://eosc-fair4s.github.io/role-profiles>
- Röttger, U., Kobusch, J., & Preusse, J. (2018). *Grundlagen der Public Relations. Eine kommunikationswissenschaftliche Einführung* (3. Auflage). Springer Fachmedien. <https://doi.org/10.1007/978-3-658-17503-0>
- Rüegg-Stürm, J. (2004). Das neue St. Galler Management-Modell. In R. Dubs, D. Euler, J. Rüegg-Stürm, & C. E. Wyss (Hrsg.), *Einführung in die Managementlehre* (Bd. 1, S. 65 – 141). Haupt.
- Salais, R. (1989). L'analyse économique des conventions du travail. *Revue économique*, 40(2), 199 – 240.
- San Francisco Erklärung zur Forschungsbewertung. (o. J.). DORA. Abgerufen 19. Februar 2022, von <https://sfedora.org/read/read-the-declaration-deutsch/>
- Sattler, A. (2020). § 2 Schutz von maschinengenerierten Daten. In T. Sassenberg & T. Faber (Hrsg.), *Rechtshandbuch Industrie 4.0 und Internet of Things: Praxisfragen und Perspektiven der digitalen Zukunft* (2. Auflage). C. H. Beck ; Vahlen.
- Schmidt, H. (2018). Plattformökonomie: Chancen für Deutschland. In Begleitforschung Smart Data (Hrsg.), *Big Data, Smart Data, next?* (S. 106–109). https://www.digitale-technologien.de/DT/Redaktion/DE/Downloads/Publikation/Smart-Data/2018_08_19_Big_Data_Smart_Data_next.pdf
- Schmidt, H. (o. J.). Plattform-Ökonomie. Dr. Holger Schmidt. Abgerufen 28. Januar 2022, von <https://www.netzoekonom.de/plattform-oekonomie/>
- Schmidt, H., & Hosseini, H. (2021). *Top- 100 Plattformen der Welt.* unveröffentlichte Darstellung.
- Schmidt, S. A. (2020). *Zugang zu Daten nach europäischem Kartellrecht.* Mohr Siebeck.
- Schnurer, G. (2021, März). Gaia-X: Kentern unausweichlich. *ct*. Abgerufen 01. März 2022, <https://www.heise.de/select/ct/2021/3/2028614113239018944>
- Scholze, F., Ulrich, R., & Goebelbecker, H.-J. (2021). 2.2 Wissenschaftlicher Datenmarkt. In M. Putnings, H. Neuroth, & J. Neumann (Hrsg.), *Praxishandbuch Forschungsdatenmanagement* (S. 165 – 174). De Gruyter Saur. <https://doi.org/10.1515/9783110657807-009>
- Schössler, M. (2018). *Plattformökonomie als Organisationsform zukünftiger Wertschöpfung: Chancen und Herausforderungen für den Standort Deutschland.* Friedrich-Ebert-Stiftung, Abteilung Wirtschafts- und Sozialpolitik. <http://library.fes.de/pdf-files/wiso/14756.pdf>
- Schrögel, P., Rühland, S., Fischer, C., Göbel, C., Heimstädt, M., Humm, C., Rösener, R., Rössig, W., Röwert, R., Siegers, M., & Thieleis, I. (2020). *Von Offener Wissenschaft zu Zugänglicher Wissenschaft / Diskussionspapier.* Zenodo. <https://doi.org/10.5281/zenodo.3911600>
- Schüller, K., Koch, H., & Rampelt, F. (2021). *DATA-LITERACY-CHARTA* (Version 1.2). Stifterverband. https://www.stifterverband.org/sites/default/files/data-literacy-charta_v1_2.pdf
- Schur, N. (2021). Teil 6.8 Know-how und Geheimnisschutz von Daten. In A. Leupold, A. Wiebe, & S. Glossner (Hrsg.), *IT-Recht: Recht, Wirtschaft und Technik der digitalen Transformation* (4., überarbeitete und erweiterte Auflage). C. H. Beck.
- Schweda, M., & von Schreiter, F. (2021). Ran an die Datenschätze? Datenzugangsansprüche nach der 10. GWB-Novelle. *WuW*, 71(3), 145 – 154.
- Schweigel, H., Lindner, M., Mock, J., Straub, S., & Zinke, G. (2020). *OPEN PUBLIC DATA IN DEUTSCHLAND Rahmenbedingungen und Potenziale der Bereitstellung und Nutzung von Daten des öffentlichen Sektors.* Begleitforschung Smart Service Welt II Institut für Innovation und Technik (iit) in der VDI/VDE Innovation + Technik GmbH. https://www.digitale-technologien.de/DT/Redaktion/DE/Downloads/Publikation/SSW/2020/SSW_Open_Public_Data_in_Deutschland.pdf
- Sovereign Cloud Stack.* (o. J.). Sovereign Cloud Stack. Abgerufen 2. Februar 2022, von <https://scs.community/>
- Spiekermann, S. (2019). *Digitale Ethik: ein Wertesystem für das 21. Jahrhundert.* Droemer.
- Staab, P. (2019). *Digitaler Kapitalismus – Markt und Herrschaft in der Ökonomie der Unknappheit.* Suhrkamp.
- Stelkens, P., Bonk, H. J., & Sachs, M. (Hrsg.). (2018). *Verwaltungsverfahrensgesetz: Kommentar* (9. Auflage). C. H. Beck.
- Stiftung Datenschutz (Hrsg.). (2019). *Dateneigentum und Datenhandel.* Erich Schmidt Verlag.
- Strauss, A. L. (1987). *Qualitative Analysis for Social Scientists* (1. Aufl.). Cambridge University Press. <https://doi.org/10.1017/CBO9780511557842>
- Strübing, J. (2014). *Grounded Theory* (3., überarbeitete und erweiterte Aufl. 2014). VS Verlag für Sozialwissenschaften. <https://doi.org/10.1007/978-3-531-19897-2>
- Studieren in Deutschland – Studium in Deutschland.* (o. J.). Hochschulkompass. Abgerufen 14. Februar 2022, von <https://www.hochschulkompass.de/studium>
- Synergy Research Group | *Strategic Market Intelligence for Emerging IT & Cloud.* (o. J.). Abgerufen 12. Februar 2022, von <https://www.srgresearch.com/>
- The world's most valuable resource is no longer oil, but data. (2017, Mai 6). *The Economist.* Abgerufen 01. März 2022, <https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data>
- Thomas, J., Wendehorst, C., Duller, Y., & Schwamberger, S. (2021). *Response of the European Law Institute. Public Consultation on the Data Act.* European Law Institute. <https://europeanlawinstitute.eu/fileadmin/>

- user_upload/p_eli/Projects/Data_Economy/ELI_Response_Public_Consultation_on_a_Data_Act.pdf
- Thouvenin, F., Weber, R. H., & Früh, A. (2017). Data ownership: Taking stock and mapping the issues. In M. Dehmer & F. Emmert-Streib (Hrsg.), *Frontiers in Data Science*. CRC Press.
- Ulrich, P. (2001). *Integrative Wirtschaftsethik : Grundlagen einer lebensdienlichen Ökonomie* (3., revidierte Auflage). Haupt.
- Ungerer, B. (2021, April 8). Data Scientists dringend gesucht. *iX Magazin, heise online*. Abgerufen 01. März 2022, <https://www.heise.de/news/Data-Scientists-dringend-gesucht-6009166.html>
- van den Eynden, V., & Bishop, L. (2014). *Incentives and motivations for sharing research data, a researcher's perspective*. <https://repository.jisc.ac.uk/5662/1/KE-report-incentives-for-sharing-researchdata.pdf>
- van Hesteren, D., van Knippenberg, L., Weyzen, R., Huyer, E., & Cecconi, G. (2022). *Open Data Maturity Report 2021*. European Commission, Directorate-General of Communications Networks, Content and Technology. https://data.europa.eu/sites/default/files/landscaping_insight_report_n7_2021_0.pdf
- van Knippenberg, L. (2020). Open Data Maturity Report 2020. European Commission, Directorate-General of Communications Networks, Content and Technology. <https://doi.org/10.2830/619187>
- Verordnung (EU) 2016/679 des Europäischen Parlaments und des Rates vom 27. April 2016 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten, zum freien Datenverkehr und zur Aufhebung der Richtlinie 95/46/EG (Datenschutz-Grundverordnung) (Text von Bedeutung für den EWR), OJ L (2016). <http://data.europa.eu/eli/reg/2016/679/oj/deu>
- Volkert, J. (2021, Dezember 7). EUCLIDIA: Europäische Cloud-Infrastruktur gegen die Cloud aus China und den USA. *heise online*. Abgerufen 01. März 2022, <https://www.heise.de/news/EUCLIDIA-Gegen-die-Cloud-aus-China-und-den-USA-6135136.html>
- Was ist Datenaustausch? (o.J.). Support Centre for Data Sharing. Abgerufen 14. Februar 2022, von <https://eudatasharing.eu/de/what-data-sharing>
- Weinhardt, M. (2021). Big Data: Some Ethical Concerns for the Social Sciences. *Social Sciences*, 10(2), 36. <https://doi.org/10.3390/socsci10020036>
- Whyte, A., & Ashley, K. (2017). *D7.1: Skills landscape analysis and competence model*. <https://eoscpilot.eu/sites/default/files/eoscpilot-d7.1.pdf>
- Whyte, A., Leenarts, E., de Vries, J., Huigen, F., Kuehn, E., Sipos, G., Kalaitzi, V., Dijk, E., & Jones, S. (2019). *D7.5: Strategy for Sustainable Development of Skills and Capabilities*. <https://eoscpilot.eu/sites/default/files/eoscpilot-d7.5-v1.1.pdf>
- Wiebe, A. (2021). Teil 3. Schutz von Datenbanken, Computerspielen und Webangeboten. In A. Leupold, A. Wiebe, & S. Glossner (Hrsg.), *IT-Recht: Recht, Wirtschaft und Technik der digitalen Transformation* (4., überarbeitete und erweiterte Auflage). C. H. Beck.
- Wiebe, A., & Schur, N. (2019). Protection of trade secrets in a data-driven, networked environment – Is the update already out-dated? *GRUR Int*, 68(8–9), 746 – 751.
- Wiederkehr, S. (2021). Citizen Science. *o-bib. Das offene Bibliotheksjournal / Herausgeber VDB*, 1 – 13. <https://doi.org/10.5282/O-BIB/5727>
- Wilkins, A. (2021, Juni 30). Europa-Cloud Gaia-X: 16 Leuchtturmprojekte für datengetriebene Geschäftsmodelle. *heise online*. Abgerufen 01. März 2022, <https://www.heise.de/news/Gaia-X-16-Leuchtturmprojekte-fuer-datengetriebene-Geschaeftsmodelle-6124261.html>
- Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., ... Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3(1), 160018. <https://doi.org/10.1038/sdata.2016.18>
- Wissenschaftlicher Beirat der Bundesregierung Globale Umweltveränderungen (Hrsg.). (2011). *Welt im Wandel: Gesellschaftsvertrag für eine Große Transformation; [Hauptgutachten]* (2., veränd. Aufl). Wiss. Beirat der Bundesregierung Globale Umweltveränderungen (WBGU). https://www.wbgu.de/fileadmin/user_upload/wbgu/publikationen/hauptgutachten/hg2011/pdf/wbgu_jg2011.pdf
- Wissenschaftsrat. (2021). *Impulse aus der COVID-19-Krise für die Weiterentwicklung des Wissenschaftssystems in Deutschland | Positionspapier (Drs. 8834-21)*. Wissenschaftsrat. <https://www.wissenschaftsrat.de/download/2021/8834-21.pdf>
- Wölbart, C. (2020, Oktober 16). Europa-Cloud Gaia-X tritt gegen übermächtige Konkurrenz an. *heise online*. Abgerufen 01. März 2022, <https://www.heise.de/news/Europa-Cloud-Gaia-X-tritt-gegen-uebermaechtige-Konkurrenz-an-4920108.html>
- Zeunert, M., & Schneemann, C. (2021). *Forschungsdatenmanagement: Arbeitsschwerpunkte für Informationswissenschaftler*innen* [Universität Potsdam]. <https://doi.org/10.25932/publishup-50790>

ANHANG

10

10. Anhang

10.1 Forschungsdatenmanagementplan

Stand: 03.03.2022

Forschungsprojekt

Projektname: Ökosystem Datenmanagement Analysen – Empfehlungen – FAIRification [Ecosystem Data Management – EcoDM]

Projektbeschreibung: Mit dem BMBF geförderten Verbundprojekt EcoDM (<https://www.ecodm.de/>) wurde erforscht, welche Herausforderungen, Chancen und Hindernisse sich im Bereich des rasanten Wachstums an digitalen Daten ergeben und wie Rahmenbedingungen aussehen könnten, Daten systematisch und FAIR (Findable, Accessible, Interoperable, Reusable) nutzen und teilen zu können. Einen zentralen Teil der Untersuchung bildeten neben Landscape- und Gap-Analysen Leitfaden-gestützte Interviews mit Expert*innen aus den Bereichen Wissenschaft, Wirtschaft, Public Sector und Qualifikation. Basierend auf den Projektergebnissen wurden Empfehlungen zur Förderung des Teilens und Nachnutzens von Daten entwickelt. Daraus abgeleitete EcoDM-Empfehlungen wurden im Rahmen der RDA DE Konferenz 2022 auf einer Podiumsdiskussion vorgestellt und diskutiert.

Projektlaufzeit: 01.10.2019 – 31.03.2022 (30 Monate)

Projektkoordination: Fachhochschule Potsdam (FHP)

Projektpartner: Deutsches Archäologisches Institut (DAI), Fachhochschule Potsdam (FHP), Helmholtz-Zentrum Potsdam – Deutsches GeoForschungs-Zentrum (GFZ)

Projektverantwortliche: Henriette Senst (DAI), Heike Neuroth (FHP), Christoph Bruch (GFZ)

Kontaktperson Datenmanagement: Heike Neuroth

Anmerkungen: (ehemaliger Projektpartner MPCDF)

Förderung

Forschungsförderer: Bundesministerium für Bildung und Forschung

Förderprogramm: (keines)

Förderkennzeichen: 16DWWQP

Anforderungen

Förderer

Vorgaben des Forschungsförderers:

„Das BMBF erwartet einen „Verwertungsplan“, in dem unter anderem wissenschaftliche und/oder technische Nachnutzungsmöglichkeiten darzustellen sind: „Unabhängig von den wirtschaftlichen Erfolgsaussichten sollen die wissenschaftlichen und/oder technischen Erfolgsaussichten dargestellt werden (mit Zeithorizont) – u.a., wie die geplanten Ergebnisse in anderer Weise (z.B. für öffentliche Aufgaben, Datenbanken, Netzwerke, Transferstellen etc.) genutzt werden können. An dieser Stelle ist auch eine etwaige Zusammenarbeit mit anderen Einrichtungen, Firmen, Netzwerken, Forschungsstellen u.a. einzubeziehen.“ Weitere Informationen finden Sie in den [Richtlinien für Zuwendungsanträge auf Ausgabenbasis \(AZA\)](https://www.cms.hu-berlin.de/de/dl/dataman/arbeiten/dmp_erstellen/foerderer) des BMBF“ (vgl. https://www.cms.hu-berlin.de/de/dl/dataman/arbeiten/dmp_erstellen/foerderer)

Community-Vorgaben: DFG-Kodex Version 2019, https://www.dfg.de/foerderung/grundlagen_rahmenbedingungen/gwp/kodex/index.html

Datenrepositorium

Vorgaben der Einrichtungen der Projektpartner: DAI (IANUS): <https://www.ianus-fdz.de/it-empfehlungen/dateiformate>

1. Forschungsdaten

Art	Format	Beschreibung	Ursprung	Reproduzierbar
1. Videos	mp4	Videoaufzeichnung der Expert*inneninterviews (teilweise zusätzlich gekürzte Version)	projektintern erstellt	nein
2. Tonaufnahmen	m4a	separate Tonaufnahme der Expert*inneninterviews (nicht für alle Interviews vorhanden), als Backup	projektintern erstellt	nein
3. Transkripte	docx	Verschriftlichung der Videoaufnahmen	extern vergeben, projektintern QM-geprüft	nur solange Videos vorhanden
4. Auswertungen Rohdaten	mx20; xlsx; docx	Bearbeitung mit MAXQDA (mx20), Export der kodierten Texte in einem „Standardformat“ (Excel, Word)	projektintern erstellt	ja, solange Transkripte vorhanden (max. Aufbewahrung 10 Jahre)
5. Dokumentation zu den Interviews	pdf	Unterschiedene Einverständniserklärungen Schriftwechsel mit dem Datenschutzbeauftragten	projektintern erstellt	nein
6. Interviewleitfaden	pdf	Beschreibung des Aufbaus, Ablaufs des Interviews und Liste mit Fragen zur Auswahl	projektintern erstellt	nein
7. individuelle Fragenkataloge	pdf	Liste der ausgewählten Fragen und Ablauf für jedes einzelne Interview	projektintern erstellt	nein

2. Datenerhebung

Art	Erhebung	Bearbeitung	Verwendete Programme
1. Videos	26.03.2021 bis 29.07.2021	ggf. Kürzung für Transkription	Zoom
2. Tonaufnahmen	26.03.2021 bis 29.07.2021	keine	Zoom
3. Transkripte	12.05.2021 bis 15.09.2021	Überarbeitung durch das Projektteam: Korrektur von Lücken und Schreibfehlern	extern vergeben (MS Word)
4. Auswertungen Rohdaten	20.05.2021 bis 13.12.2021	Entwicklung von Kategorien mit bereichsspezifischer Kodierung aller Interviews	MAXQDA
5. Dokumentation zu den Interviews	15.01.2021 bis 29.07.2021	Erstellung der Einverständniserklärung	MS Word
6. Interviewleitfaden	20.01.2021 bis 15.04.2021	Erste Fassung bis 15.03., Anpassungen bis 15.04.	MS Word
7. individuelle Fragenkataloge	16.03.2021 bis 29.07.2021	Zuschnitt Fragenkatalog auf jeweilige*n Expert*in	MS Word

3. Datenspeicherung und Dokumentation

Art	Speicherort	Benötigter Speicherplatz
1. Videos	Zugriffsgesicherte Server der Fachhochschule Potsdam (FHP) inkl. Backup alle 24 h durch zentrale IT der FHP Alle Dateien bei externen Dienstleistern werden nach Vertragsende gelöscht (Auftragsverarbeitungsvertrag)	4,83 GB
2. Tonaufnahmen		280 MB
3. Transkripte		1,5 MB
4. Auswertungen Rohdaten	Zugriffsgesicherte Server der Fachhochschule Potsdam (FHP) inkl. Backup alle 24 h durch zentrale IT der FHP	80 MB
5. Dokumentation zu den Interviews		
6. Interviewleitfaden		
7. individuelle Fragenkataloge		

4. Rechtliches

Art	Urheberrecht	Persönlichkeitsrechte
1. Videos	Miturheberschaft → Zustimmung zur geplanten Verwertung/ Nutzung liegt von den Interviewten vor. Urheberrechte der Interviewer*innen sind als Pflichtwerke einzustufen, was eine automatische ausschließliche Übertragung der Nutzungsrechte an den Arbeitgeber zur Folge hat.	ja
3. Transkripte	Sind eine Vervielfältigung der Vorlage, die der Erlaubnis der Rechteinhaber*innen bedarf, wenn sie urheberrechtlich geschützt ist. Bei den EcoDM-Interviews ist von einer Miturheberschaft von Interviewer*innen und Interviewten auszugehen. Die Interviewten wurden über die geplante Nutzung informiert und haben zugestimmt. Damit ist auch die Vervielfältigung abgedeckt.	ja
4. Auswertungen Rohdaten	Bearbeitung von Dokumenten: urheberrechtlich nicht relevant, außer es werden Vervielfältigungen erstellt. Bei EcoDM liegt die Zustimmung der Urheber*innen vor.	ja
5. Dokumentation zu den Interviews	n/a	ja
6. Interviewleitfaden	n/a	nein
7. individuelle Fragenkataloge	n/a	ja

5. Datenarchivierung

Art	Archivierung geplant	Dauer	Speicherort	Format	Speicherplatz	Kosten
1. Videos	nein	-	-	-	-	-
2. Tonaufnahmen	nein	-	-	-	-	-
3. Transkripte	ja	10 a	IANUS: https://www.ianus-fdz.de/	PDF/A	3,24 MB	
4. Auswertungen Rohdaten	ja	10 a	IANUS: https://www.ianus-fdz.de/	xlsx, docx (csv)		
5. Dokumentation zu den Interviews	ja	10 a	IANUS: https://www.ianus-fdz.de/	PDF/A		
6. Interviewleitfaden	ja	10 a	IANUS: https://www.ianus-fdz.de/	PDF/A		
7. individuelle Fragenkataloge	ja	10 a	IANUS: https://www.ianus-fdz.de/	PDF/A		

6. Data Sharing

Art	Teilen möglich?	Vorgaben/Hinweise	Lizenzen
1. Videos	nein (nach Projektende gelöscht)	„Die Aufzeichnungen liegen während des Projekts auf zugriffsgesicherten Servern der Fachhochschule Potsdam (FHP) und werden bei Projektende gelöscht.“ (vgl. Einverständniserklärung)	-
2. Tonaufnahmen	nein (nach Projektende gelöscht)	„Die Aufzeichnungen liegen während des Projekts auf zugriffsgesicherten Servern der Fachhochschule Potsdam (FHP) und werden bei Projektende gelöscht.“ (vgl. Einverständniserklärung)	-
3. Transkriptionen	nur bei berechtigtem Interesse auf Nachfrage	„Im Sinne guter wissenschaftlicher Praxis werden die Transkriptionen bei Projektabschluss für einen Zeitraum von zehn Jahren geschützt beim Deutschen Archäologischen Institut aufbewahrt und nur berechtigten Forschenden auf Anfrage mit erneuter Zustimmung der Interviewpartner zugänglich gemacht.“ (vgl. Einverständniserklärung)	-
4. Auswertungen Rohdaten	nur bei berechtigtem Interesse auf Nachfrage	personenbezogene Daten	-
5. Dokumentation zu den Interviews	nein	personenbezogene Daten	-
6. Interviewleitfaden	ja	öffentlich (Bestandteil des Reports)	CC BY 4.0
7. individuelle Fragenkataloge	nein	nur archiviert	-

Veröffentlichungen im Rahmen von EcoDM

EcoDM-Empfehlungen:

Bruch, C., Neuroth, H., Senst, H. (2022). *Data At Your Fingertips: Wegmarken einer exzellenten und vertrauenswürdigen Datenlandschaft. Empfehlungen des EcoDM-Projekts.* <https://doi.org/10.5281/zenodo.5837408>

EcoDM-Report:

Bruch, C., Neuroth, H., Senst, H. (Hrsg.). (2022). *EcoDM – Ökosystem Datenmanagement: Analysen – Empfehlungen – FAIRifizierung.* <https://doi.org/10.5281/zenodo.6256398>

Beauftragte Übersetzungen

- Boserup Thestrup, J., Braskova, M., Krogh Kruuse, K., & Lembinen, L. (2021). *Die sechs Säulen für die Einbindung von Forschenden in das Forschungsdatenmanagement (FDM).* <https://doi.org/10.5281/zenodo.5795858>
- Carroll, S. R., Hudson, M., Chapman, J., Figueroa-Rodríguez, O. L., Holbrook, J., Lovett, R., Materechera, S., Parsons, M., Raseroka, K., Rodriguez-Lonebear, D., Rowe, R., Sara, R., & Walker, J. (2019). *Die CARE-Prinzipien für indigene Data Governance.* <https://doi.org/10.5281/zenodo.5995059>
- Carroll, S.R., Garba, I., Figueroa-Rodríguez, O.L., Holbrook, J., Lovett, R., Materechera, S., Parsons, M., Raseroka, K., Rodriguez-Lonebear, D., Rowe, R., Sara, R., Walker, J.D., Anderson, J., & Hudson, M. (2020). *Praxisbericht: Die CARE-Prinzipien für indigene Data Governance.* <https://doi.org/10.5281/zenodo.6007531>
- Devaraju, A., Mokrane, M., Cepinskas, L., Huber, R., Herterich, P., de Vries, J., Akerman, V., L'Hours, H., Davidson, J., & Diepenbroek, M. (2021). *Von der Konzeptualisierung zur Implementierung: FAIR-Evaluation von Forschungsdatenobjekten.* <https://doi.org/10.5281/zenodo.5877746>
- Europäische Kommission. (2021). *Ethik und Datenschutz.* <https://doi.org/10.5281/zenodo.6259754>
- Jacobsen, A., Azevedo, R. de M., Juty, N., Batista, D., Coles, S., Cornet, R., Courtot, M., Crosas, M., Dumontier, M., Evelo, C. T., Goble, C., Guizzardi, G., Hansen, K. K., Hasnain, A., Hettne, K., Heringa, J., Hooft, R. W.W., Imming, M., Jeffery, K. G., ... Schultes, E. (2020). *FAIR-Prinzipien: Interpretationen und Überlegungen zur Umsetzung.* <https://doi.org/10.5281/zenodo.6255623>
- Lin, D., Crabtree, J., Dillo, I., Downs, R. R., Edmunds, R., Giaretta, D., De Giusti, M., L'Hours, H., Hugo, W., Jenkyns, R., Khodiyar, V., Martone, M. E., Mokrane, M., Navale, V., Petters, J., Sierman, B., Sokolova, D. V., Stockhouse M., & Westbrook, J. (2020). *Die TRUST-Prinzipien für digitale Repositorien.* <https://doi.org/10.5281/zenodo.6256222>
- RDA FAIR Data Maturity Model Working Group. (2020). *Das FAIR Data Maturity Model. Spezifikation und Leitlinien.* <https://doi.org/10.5281/zenodo.5834115>
- Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., ... Mons, B. (2016). *Die FAIR-Prinzipien für das wissenschaftliche Datenmanagement und Data Stewardship.* <https://doi.org/10.5281/zenodo.6247015>

10.2 Vorlage der Einverständniserklärung



Ökosystem Datenmanagement: Analysen - Empfehlungen - FAIRification

Informationsblatt zur Datenerhebung und Datenverarbeitung

Projekt: EcoDM - Ökosystem Datenmanagement
E-Mail: ecodm-kontakt@listserv.dfn.de
Beteiligte Institutionen: Deutsches Archäologisches Institut, Fachhochschule Potsdam, Helmholtz-Zentrum Potsdam - Deutsches GeoForschungsZentrum
Ansprechpartner*in:
Kontakt:

Im Projekt EcoDM – Ökosystem Datenmanagement wird untersucht, wie die Potenziale des rasanten digitalen Datenwachstums in den Bereichen Wissenschaft, Wirtschaft, Public Sector und Qualifizierung systematisch und FAIR (Findable, Accessible, Interoperable, Reusable) genutzt werden können. Für die Erarbeitung von Handlungsempfehlungen für ein Ökosystem nachhaltigem Datenmanagements werden u. a. qualitative Expert*inneninterviews durchgeführt. Dabei arbeiten wir nach geltenden datenschutzrechtlichen Bestimmungen. Im Rahmen dieses Interviews werden folgende Daten erhoben:

Erhebung und Verarbeitung personenbezogener Daten

Ihr Name sowie Angaben über Ihre berufliche Tätigkeit. Diese Daten möchten wir wie im Folgenden dargelegt verwenden:

Wir zeichnen das Interview mithilfe der Software Zoom bzw. einer gemeinsam vereinbarten Alternative auf.¹

Diese Aufzeichnung wird anschließend durch einen externen (deutschen) Dienstleister transkribiert und von unserem Team in der Software MaxQDA (der Berliner Firma Verbi GmbH) ausgewertet. Bei der Transkription werden (Teil-)Abschriften der Aufzeichnungen erstellt. Die Aufzeichnungen und Transkripte werden nur für Analysen und darauf basierende Publikationen im Rahmen des Projekts genutzt. Die Aufzeichnungen liegen während des Projekts auf zugriffsgesicherten Servern der Fachhochschule Potsdam (FHP) und werden bei Projektende gelöscht. Im Sinne guter wissenschaftlicher Praxis werden die Transkriptionen bei Projektabschluss für einen Zeitraum von zehn Jahren geschützt beim Deutschen Archäologischen Institut aufbewahrt und nur berechtigten Forschenden auf Anfrage mit erneuter Zustimmung der Interviewpartner zugänglich gemacht. Alle anderen personenbezogenen Daten werden nach dem Abschluss des Projekts gelöscht.

Veröffentlichung

Die Veröffentlichung von Forschungsergebnissen in Publikationen oder auf Tagungen erfolgt grundsätzlich in anonymisierter Form und lässt zu keinem Zeitpunkt Rückschlüsse auf Sie als Person zu. Es ist geplant, Aussagen in Form von wörtlichen Zitaten bei der Öffentlichmachung von Ergebnissen (z. B. in Publikationen, Berichten, Webseiten) zu belegen.

¹ Weitere Informationen finden Sie unter <https://explore.zoom.us/de-de/privacy-and-security.html>. Da Zoom-Nutzerdaten in den USA gespeichert werden, kann nicht ausgeschlossen werden, dass im Rahmen der dort geltenden Gesetze auch staatliche Behörden – bei eingeschränkten Rechtsschutzmöglichkeiten – Zugriff auf diese Daten erhalten. Als mögliche Alternative steht beispielsweise BigBlueButton zur Verfügung.



Ökosystem Datenmanagement: Analysen - Empfehlungen - FAIRification

Informationsblatt zur Datenerhebung und Datenverarbeitung

Bei der Verwendung eines wörtlichen Zitats unter Nennung Ihres Namens holen wir vorab Ihre Genehmigung ein. Sollten Sie mit einer namentlichen Nennung nicht einverstanden sein, können Sie dieser Verwendung auf dem Formular ausdrücklich widersprechen, einer anonymisierten Form zustimmen oder die Form Ihrer Wahl für die Referenz angeben.

Rechte

Sie haben jederzeit die Möglichkeit folgende Rechte geltend zu machen:

- **Recht auf Widerruf** der Einwilligung: Sie haben das Recht, Ihre Einwilligung zum Erheben ihrer personenbezogenen Daten jederzeit mit Wirkung für die Zukunft zu widerrufen. (Art. 7 Abs. 3 DSGVO).
- **Auskunftsrecht**: Sie haben uns gegenüber das Recht, Auskunft darüber zu erhalten, welche Ihrer personenbezogenen Daten wir verarbeiten. (Art. 15 DSGVO).
- **Recht auf Berichtigung**: Sollten die Sie betreffenden personenbezogenen Daten nicht richtig oder unvollständig sein, so können Sie die Berichtigung unrichtiger oder die Vervollständigung unvollständiger Angaben verlangen. (Art. 16 DSGVO).
- **Recht auf Löschung**: Sie können jederzeit die Löschung ihrer personenbezogenen Daten verlangen. (Art. 17 DSGVO).
- **Recht auf Einschränkung der Verarbeitung**: Sie können die Einschränkung der Verarbeitung der Sie betreffenden personenbezogenen Daten verlangen. (Art. 18 DSGVO).
- **Widerspruchsrecht**: Sie können jederzeit gegen die Speicherung und Verarbeitung der Sie betreffenden personenbezogenen Daten Widerspruch einlegen. (Art. 21 DSGVO).

In jedem Fall gilt: Ihre Teilnahme an unserer Studie ist freiwillig. Lehnen Sie die Teilnahme ab, widerrufen oder beschränken Sie Ihre Einwilligung, entstehen Ihnen hieraus keine Nachteile.

Ihre Erklärungen zur Geltendmachung Ihrer Rechte sind grundsätzlich schriftlich per Post oder E-Mail an die Verantwortlichen (Projektleitung) zu richten.



Ökosystem Datenmanagement: Analysen - Empfehlungen - FAIRification

Einverständniserklärung zur Teilnahme am Interview und zur Verwendung Ihrer personenbezogenen Daten im Projekt EcoDM

Interviewdatum:

Interviewende*r:

Software:

Vorname:

Nachname:

Titel:

Institution:

Zur oben bezeichneten Befragung habe ich das Informationsblatt erhalten und dieses gelesen. Auch hatte ich Gelegenheit, Fragen zu stellen. Eventuelle Fragen wurden vollständig beantwortet.

Mir ist bewusst, dass meine Teilnahme vollkommen freiwillig ist und ich bei einer Verweigerung meiner Einwilligung keinerlei Nachteile erleide. Meine Einwilligung kann ich jederzeit mit Wirkung für die Zukunft widerrufen, ohne dass dies einer Begründung bedarf und ohne, dass mir daraus irgendwelche Nachteile entstehen. Im Fall eines Widerrufs werden meine personenbezogenen Daten anonymisiert oder gelöscht beziehungsweise gesperrt.

Eine Kopie der Informationsschrift und dieser Einwilligungserklärung habe ich erhalten. Das Original verbleibt bei der Fachhochschule Potsdam (FHP).

Ich erkläre meine Einwilligung, dass meine personenbezogenen Daten (einschließlich Klarname und Name meines Arbeitgebers) wie in der Informationsschrift beschrieben verwendet werden.

Der Aufzeichnung des Interviews stimme ich zu.

Im Rahmen von Zitaten in Publikationen, Berichten, im Web oder anderen Ergebnisdarstellungen stimme ich der Nennung in folgender Form zu (Zutreffendes bitte ankreuzen):

anonymisiert

namentliche Nennung

Nennung meines/r Arbeitgebers/Institution

Ort, Datum

Unterschrift

10.3 Informationsblatt zu den Interviews



Ökosystem Datenmanagement: Analysen - Empfehlungen - FAIRification

Informationsblatt

Wer sind wir?

Im BMBF-geförderten Verbundprojekt EcoDM – Ökosystem Datenmanagement erforschen wir Stand und Potenziale des digitalen Datenmanagements in Deutschland. Wir untersuchen, wie die stetig wachsenden digitalen Datenmengen aus unterschiedlichen Bereichen systematisch und entsprechend der FAIR-Prinzipien (Findable, Accessible, Interoperable, and Reusable) nachhaltig genutzt werden können. Die Projektpartner sind die Fachhochschule Potsdam, das Deutsche Archäologische Institut und das Helmholtz Open Science Office.

Was machen wir?

In einer Landscape- und Gap-Analyse zu Public Sector, Wirtschaft, Wissenschaft sowie dem Querschnittsthema Qualifizierung erkunden wir bestehende Strukturen, Inhalte, Akteure, Best Practices und ermitteln Desiderate. Die Ergebnisse werden durch bereichsspezifische Expert*inneninterviews und eine übergreifende Fokusgruppendifkussion erweitert. Auf dieser Basis formulieren wir Handlungsempfehlungen, um die Entwicklung eines nachhaltigen und FAIRen Datenmanagement-Ökosystems in Deutschland zu unterstützen.

Wie wird das Interview ablaufen?

Das Interview wird je nach Absprache über Zoom oder auf Wunsch alternativ z. B. mit BigBlueButton in digitaler Form durchgeführt. Die Dauer beträgt zwischen 30–45 Minuten. Sie sollten jedoch insgesamt etwa eine Stunde einplanen, damit vorab ein Technikcheck erfolgen kann. Gerne kann auch ein persönliches Vorgespräch erfolgen.

Nach einem allgemeineren Einstieg sollen maximal fünf Fragenkomplexe im Mittelpunkt des Interviews stehen. Dabei sind Ihre spezifische Perspektive und der Kontext, in dem Sie sich bewegen, von besonderem Interesse für uns. Das Ziel ist es, Bedarfe, Chancen und Hindernisse auf dem Weg zu einem FAIRen Datenmanagement zu ermitteln.

Was geschieht mit dem Interview?

Alle Interviews werden von einem externen Dienstleister transkribiert und dann vom Projektteam mit Hilfe der Annotationssoftware MaxQDA systematisch ausgewertet. Die Ergebnisse aus den Analysen der Expert*inneninterviews fließen in die kontinuierlich weitergeführte Landscape- und Gap-Analyse ein und dienen als Grundlage für die Fokusgruppendifkussion.

Alle weiteren Details zum Umgang mit den Interviews entnehmen Sie bitte der Einverständniserklärung.

www.ecodm.de | ecodm-kontakt@listserv.dfn.de | [@ecodm_projekt](https://twitter.com/ecodm_projekt)



10.4 Bereichsübergreifender Interview-Leitfaden

I. Aufbau und (zeitlicher) Ablauf des Interviews

1. Einstieg (Abholer)

[5 min]

Ziel:

Unterschiedliche Blickwinkel der Expert*innen; jeweils spezifischen Kontext identifizieren

Vorabüberlegungen:

- Erkenntnisinteresse für die spezifischen Interviews überlegen; in Stichworten formulieren, evtl. für den Einstieg nutzen

Beachten:

- Begrüßung der/des Interviewpartners/in. Dank für Bereitschaft, die Fragen zu beantworten.
- Als Einstieg zunächst die Bitte um kurze Vorstellung, mit Angaben zum Aufgaben- und Verantwortungsbereich.
- Ansage und Start der Aufzeichnung
- Vorstellung der Kolleg*innen, die am Interview beteiligt sind.
- Vorstellung des Projekts und seiner Ziele, ggf. angepasst an den spezifischen Kontext, Hintergrund und Kenntnisstand der/des Interviewpartner/in

Beispiel:

- EcoDM möchte Ansatzpunkte zur Unterstützung des Datenteilens liefern.
- Dabei sollen Hindernisse beim Teilen von Daten, aber auch Lösungen identifiziert werden.
- Unser Ziel ist es, Akteur*innen zum Teilen eigener Daten und zur Nachnutzung der Daten von Dritten zu motivieren.
- Zu diesem Zweck möchten wir bereichsübergreifend Handlungsempfehlungen entwickeln und fragen daher u. a. nach Best-Practices-Erfahrungen.
- Nach den FAIR-Prinzipien sollten Daten auffindbar, zugänglich, interoperabel und nachnutzbar sein.
- Ebenso sind weitere Rahmenbedingungen zu gestalten (etwa juristisch, politisch oder kulturell).

Beziehung zum Tätigkeitsfeld des/der Interviewten herstellen:

Wie sehen Sie die Themen in Ihrem Tätigkeitsfeld?

- Datenteilen
- Rahmenbedingungen
- FAIR-Prinzipien

Sollte der/die Interviewpartner*in näher auf FAIR eingehen:

- Sind die FAIR-Prinzipien als solche für Sie verständlich und praxistauglich?
- Konnten Sie bereits Daten von Dritten aufgrund mangelnder FAIRness schlecht oder gar nicht nutzen?
- Ist ein Defizit bei der FAIRifizierung von Daten ein wesentliches Hindernis bei deren Nachnutzung?

Andernfalls allgemeiner Einstieg:

- Was verstehen Sie unter einer „Kultur des Datenteilens“?
- Welche Rahmenbedingungen müssten sich dafür ändern? (politisch, juristisch, gesellschaftlich)
- Haben wir im Bereich Infrastruktur und Plattformen die Technik, die wir brauchen – und brauchen wir die Technik, die wir haben? (alte Grundsatzfrage der Technikphilosophie)

2. Assoziationen / Reaktion auf Antworten

[5 min]

... abhängig von Antworten mit 3. oder 4. weiterfragen

3. Ausgangslage / Interessen / Handlungslogik im Kontext des/der Interviewpartners/in

[15 – 20 min]

Mögliche Leitfragen:

- Wie kann Datenaustausch überhaupt funktionieren? (und zwischen wem?)

- Welche Voraussetzungen und Rahmenbedingungen sind nötig?
- Welche konkreten Schritte / Mittel sollten zur kontextadäquaten Umsetzung angewandt werden?

4. Chancen, Herausforderungen, Hürden, Bedenken, Lösungsstrategien [15 – 20 min]

Mögliche Leitfragen:

- Wenn nicht bereits Fokussierung auf einen Aspekt erfolgt ist: Welche Chancen, Herausforderungen, Bedenken und Hürden sehen Sie?
- Welche Strategien erscheinen Ihnen erfolgsversprechend, um den genannten Herausforderungen & Hürden zu begegnen?
- Wer sind Schlüsselakteur*innen, aus welchen Stakeholderkreisen kommen diese und welche Aufgaben kommen diesen zu?

5. Abschluss mit Blick auf Empfehlungen / Ergänzungen (offene Frage) [5 – 10 min]

Mögliche Fragen:

- Welche zusätzlichen, konkreten Handlungsempfehlungen würden Sie geben?
- Haben Sie noch Ergänzungen oder Beispiele für Good / Best Practices bzw. gescheiterte Wege?
- Vision: Was wünschen Sie sich in 10 Jahren, drei wichtigste Punkte?
- Alternativ: Utopie / Dystopie: Wie sieht unsere durch Daten geprägte Welt 2030 aus?

Abschluss / Abmoderation:

- Ggf. kurze Zusammenfassung
- Dank für die Teilnahme
- Erläuterungen, was mit den Aussagen weiter geschieht, spätere Rückmeldung
- Verabschiedung

II. Fragenkatalog

Die folgenden nach Themenfeldern geordneten Fragen dienen als Grundlage für die Vorbereitung der individuellen Interviews in den einzelnen Untersuchungsbereichen. Wie vorab mit den Interviewpartner*innen abgestimmt, werden Fragen aus bis zu fünf Themenblöcken für ein individuelles Interview ausgewählt und angepasst. Die Auswahl orientiert sich an der Expertise und Vorerfahrung der Interviewpartner*innen. Einige der Fragen sind dabei evt. nur bereichsspezifisch relevant.

Themenfelder:

1. Kultur des Datenteilens/Nachnutzung

a) Verständnis Kultur des Datenteilens

- Was verstehen Sie unter einer „Kultur des Datenteilens“?

b) Stand und Bedarfe bezüglich „Datenteilen/Nachnutzung“

- Wie sehen Sie den Stand bezüglich des Datenteilens und der Nachnutzung von Daten in ihrem Bereich?
- Wie können/werden Datenproduzent*innen zum Datenteilen motiviert (werden)?
- Funktioniert der Datenaustausch zwischen unterschiedlichen Bereichen?
- Was müsste getan werden, um ein systematisches Datenmanagement nach den FAIR-Prinzipien stärker zu verankern?
- Inwieweit ist aus Ihrer Sicht ein Kulturwandel nötig, was wären dessen zentrale Elemente und welche Rolle spielt Regulierung dabei?
- Stehen Teilen und Nachnutzen in einem ausgeglichenen Verhältnis?

Falls nicht:

- Warum nicht?
- Was sollte sich ändern?
- Gibt es bereichsspezifische Unterschiede?
- Bis zu welchem Grad werden bereits sektor- und bereichsübergreifend Daten geteilt / genutzt?
- Welche Rolle spielen Open Data bzw. Open Source Software in Diskurs und Praxis?
- Gibt es besondere, bereichsspezifische Bedürfnisse oder Grenzen, was Teilung und Nachnutzung von Daten angeht?

c) *Potenziale und Problemfelder bzgl. Datenteilens bereichsspezifisch/-übergreifend*

- Welche Potenziale im Bereich des Datenteilens sehen Sie?
- Welche Probleme ergeben sich beim Datenteilen zwischen unterschiedlichen Bereichen?
- Wo liegen aus Ihrer Sicht die Grenzen für Datenteilen und -nachnutzung?
- Was verhindert in der Praxis das Datenteilen?
- Welche Rahmenbedingungen müssten sich dafür ändern? (politisch, juristisch, gesellschaftlich)
- Wie lässt sich den unterschiedlichen Ansprüchen an Datenqualität Rechnung tragen?
- Was sind aus Ihrer Erfahrung die aktuell größten Hindernisse und Herausforderungen für die Bereitschaft zum Teilen eigener und Nachnutzen fremder Daten?
- Ist der Aufwand für die Aufbereitung und Pflege von Daten zum Zwecke des Austauschs zu groß?

2. FAIR-Prinzipien

a) *Kenntnisse der FAIR-Prinzipien*

- Sind Sie mit den FAIR-Prinzipien: Auffindbarkeit, Zugänglichkeit, Interoperabilität und Nachnutzung vertraut?
- Wenn ja, wie können sie zur Verbesserung im Umgang mit Daten beitragen?

b) *Bedeutung und Anwendung der FAIR-Prinzipien*

- Welche Bedeutung würden Sie den FAIR-Prinzipien beimessen?
- Welche Rolle spielen die FAIR-Prinzipien für eine Kultur des Datenteilens?
- Welche Rolle spielen die FAIR-Prinzipien bei Ihnen im täglichen Betrieb?
- Inwieweit folgt Ihre alltägliche Praxis dem Slogan „as open as possible – as closed as necessary“?
- Inwieweit ist das Teilen eigener und Nachnutzen anderer Daten in Ihrem Unternehmen aktuell bereits von den FAIR-Prinzipien (findable, accessible, interoperable, reusable) geleitet?
- Spielen die FAIR-Prinzipien beim Austausch zwischen unterschiedlichen Bereichen eine Rolle?

c) *Nachholbedarfe und angestoßene Veränderungen*

- Wo besteht bei der Implementierung der FAIR-Prinzipien Nachholbedarf?
- Welche Fortschritte wurden dadurch angestoßen?

d) *Abdeckung der Bedarfe durch FAIR-Prinzipien*

- Sind mit „auffindbar“, „zugänglich“, „interoperabel“ und „nachnutzbar“ die wesentlichen Aspekte abgedeckt?
- Welche Ergänzungen zu den FAIR-Prinzipien wären wünschenswert?

3. Wertebasis/normativer Anspruch/Wertvorstellung [falls passend]

Welche ethischen Dimensionen verbinden sich mit dem Teilen von Daten?

4. Infrastruktur

a) *Technologien und Plattformen*

- Klassische Grundsatzfrage der Technikphilosophie (leicht abgewandelt): Haben wir im Bereich Infrastruktur und Plattformen die Technik, die wir brauchen – und inwieweit erfüllt die Technik, die wir haben, die Bedürfnisse unterschiedlicher Akteur*innen aus Wissenschaft, Wirtschaft und Public Sector?
- Wie kann es gelingen, unterschiedlichen Bedarfen unterschiedlicher Disziplinen etc. Rechnung zu tragen?
- Wie stehen Sie zu Datenräumen für geschlossene Stakeholderkreise zum Austausch von Daten unter klaren Bedingungen?

b) *Infrastrukturprojekte (übergreifend / allgemeiner)*

- Momentan ist viel Dynamik im Infrastrukturbereich – Gehen künftige Ökosysteme mit NFDI und EOSC in die richtige Richtung?
- Welches Potenzial bzw. welchen Mehrwert sehen Sie in den im Aufbau befindlichen Datenökosystemen NFDI, EOSC und Gaia-X bzw. sovereign cloud stack?
- Inwiefern sind gegenwärtig im Aufbau befindliche Plattformökosysteme (Gaia-X, NFDI, EOSC) für Ihren Bereich relevant? Sind sie überhaupt ausreichend bekannt?

c) *GAIA-X*

- GAIA-X ist ökonomisch getrieben, hinsichtlich eines europäischen Binnenmarktes. Wie problematisch sehen Sie hier die Beteiligung von Google, Microsoft, Amazon, Huawei und Palantir („Schlüsselfirma in der Überwachungsindustrie“)?

d) *NFDI*

- Inwieweit trägt die NFDI wissenschaftlicher Vielfalt Rechnung? (Methoden, Ansprüche und Qualitätsstandards)
- Welche Mehrwerte über Disziplinengrenzen hinweg kann die NFDI schaffen?

5. Chancen und Risiken

a) allgemein / Einstieg

- Welche Chancen, Herausforderungen, Bedenken und Hürden sehen Sie in Ihrem Bereich
- beim Datenteilen?
- beim Datennachnutzen?

b) Lösungsmöglichkeiten

- Welche Strategien erscheinen Ihnen erfolgsversprechend, um den genannten Herausforderungen & Hürden zu begegnen?

c) Schlüsselakteur*innen bei der Überwindung von Hürden & Hindernissen

- Wer sind Schlüsselakteur*innen, aus welchen Stakeholderkreisen kommen diese und welche Aufgaben kommen diesen zu?

6. Interoperabilität

a) Probleme bei Interoperabilität

- Welche Probleme ergeben sich bezüglich Interoperabilität?

b) Maßnahmen zur Verbesserung der Nachnutzung

Welche Maßnahmen können neben Standardisierung die Interoperabilität und damit die Nachnutzung verbessern?

7. Standards

a) Vorhandene Standards und Formate

Welche Standards und strukturierten Vorgaben der Datenaufbereitung und -bereitstellung werden verwendet?

Welche nachnutzbaren, maschinenlesbaren Formate werden genutzt?

b) Kostenreduktion durch Standardisierung

Wie können Kosten der Aufbereitung minimiert werden?

c) FAIR-Prinzipien im Kontext von Standards

Spielen die FAIR-Prinzipien eine Rolle?

8. Anreize

a) Bestehende Anreize

- Welche Anreize bestehen gegenwärtig und wo liegen die größten Potenziale und Herausforderungen oder Hindernisse?
- zum Datenteilen
- zum verbesserten Umgang mit Daten

b) Wünschenswerte Anreize

- Welche Anreize und Belohnungen sollten geschaffen werden?
- Welche Anreize oder Belohnungssysteme können geschaffen werden, um den Datenaustausch zu fördern?

c) Interessen und Zielkonflikte in Bezug auf Anreizsysteme

- Inwieweit lassen sich die unterschiedlichen Anreizsysteme und Interessen verknüpfen?
- Wo sehen Sie Zielkonflikte und wie lässt sich diesen begegnen?

9. Empfehlungen

a) Best Practices

- Haben Sie Beispiele für Best Practices?

b) Weitergehende Empfehlungen

- Gibt es noch Empfehlungen, die Sie unserem Projekt mit auf den Weg geben wollen?

10. Vision/Zukunftsperspektive

- Welche Vision haben Sie mit Blick auf das Thema kurz-, mittel- und langfristig?
- Was wünschen Sie sich in 10 Jahren? Was wären die drei wichtigsten Punkte?
- Wie sieht unsere durch Daten geprägte Welt 2030 aus?

11. Datentreuhänder*innen

a) Rolle von bestehenden Datentreuhändern

Wie würden Sie die Rolle von Datentreuhändern in Ihrem Bereich beschreiben?

b) Datentreuhänder perspektivisch

- Inwieweit sind Datentreuhänder eine Option, um unterschiedlichen Interessen der beteiligten Akteur*innen Rechnung zu tragen, den Datenschutz zu gewährleisten und eine möglichst breite – auch ökonomische – Nachnutzung zu ermöglichen?
 - Welche Chancen oder Bedenken sehen Sie beim Einsatz von Datentreuhändern?
- c) *Organisationsform Datentreuhänder*
- Wie sollten Datentreuhänder organisiert sein? Privatwirtschaftlich, gemeinnützig, genossenschaftlich oder staatlich organisiert?

12. Rechtliche Aspekte

a) *Rechtliche Aspekte des Datenteilens*

- Welche Rechte müssen beim Teilen von Daten potenziell beachtet und mit Blick auf Interessen von Individuen, Allgemeinheit, Unternehmen und Machtasymmetrien abgewogen werden?

b) *Urheberrechtliche Aspekte*

- Inwieweit unterliegen Daten (in verschiedenen Stadien) dem Urheberrecht, aus dem sich Verwertungsrechte ableiten lassen können?
- Müsste sich etwas im Urheber- oder Datenschutzrecht ändern, um ein sektorenübergreifendes Datenökosystem – also zwischen öffentlichem Bereich, freier Wirtschaft und Forschung – zu ermöglichen?

c) *Rechtliche Aspekte der Nachnutzung*

- Wie kann die (ökonomische) Nachnutzung rechtlich geregelt werden?
- Wie sieht es mit Persönlichkeitsrechten (auch von Dritten) und damit mit dem Datenschutz aus? Mit Blick auf Diskriminierung ist ja auch group privacy ein Begriff, inwieweit ist diese bereits juristisch abgebildet?

d) *Weitere Rechtsvorgaben*

- Gibt es weitere Rechtsvorgaben, die zu beachten sind (etwa Recht auf digitale Selbstbestimmung für Individuen, Unternehmen und juristische Personen)?

Speziell für das Querschnittsthema Qualifizierung

13. Querschnittsthema Qualifizierung

ergänzende Fragen in den anderen drei Untersuchungsbereichen:

- Was können Sie in ihrem Bereich tun um gemeinsam mit anderen Akteur*innen das Thema Qualifizierung voranzubringen?
- Könnten „Data Stewards“, „Data Curator“, „Data Manager“ oder externe Datentreuhänder helfen?

14. Data Literacy (Verständnis)

- Was beinhaltet für Sie Data Literacy?
- Wie würden Sie Data Literacy kurz charakterisieren und was sind die wesentlichen Elemente, die vermittelt werden müssten?
- Bedeutet Data Literacy für Sie, dass ein Kulturwandel erforderlich ist?

15. Datenkompetenz in der Lehre

a) *Vermittlung von Datenkompetenz*

- Welche Möglichkeiten zur Vermittlung von Datenkompetenz sehen Sie in und außerhalb der Lehre?
- Sollten spezielle Module innerhalb des Studiums implementiert werden?
- Müssen Dozentinnen und Dozenten selbst in Fragen der Datenkompetenz geschult werden?

b) *Bewusstsein für Daten*

- Für wie wichtig halten Sie ein Bewusstsein für die Bedeutung von Daten?

c) *Vermittlung Datenkompetenz in der Lehre allgemein und spezifisch*

- Was sollten (alle) Studierenden unabhängig von der Disziplin lernen?
- Sehen Sie da disziplinspezifische Unterschiede, z. B. in den Geisteswissenschaften im Vergleich zu Naturwissenschaften, was gilt für alle, was wäre spezifisch?

16. Qualifikationswege

a) *Möglichkeiten und Orte der Qualifikation in Datenkompetenz*

- Sollten die Kenntnisse im Umgang mit Daten als Bestandteil von Ausbildungen und Studien vermittelt werden oder spricht mehr dafür, diese in speziellen Kursen bedarfsgerecht zu vermitteln?

- Welche Möglichkeiten zur Vermittlung von grundlegenden Kompetenzen im Umgang mit (Forschungs-)daten sehen Sie?

b) Zielgruppen

- Welche Zielgruppen sollten Ihrer Meinung nach angesprochen werden?

c) Entwicklungen

- Sehen Sie eine Weiterentwicklung hin zu neuen Angeboten?

d) Qualifikation von Lehrenden und Vermittelnden

- Wie sieht es aus mit der Weiterbildung von Dozierenden und Vermittelnden?

e) Allianzen und Netzwerke zur Qualifizierung

- Was halten Sie in diesem Zusammenhang von „Qualifizierungs-Allianzen“ und Netzwerken? Wer sollte da dabei sein?

f) Kompetenzaufbau zur Bereitstellung öffentlicher Daten

- Wie sehen Sie das Thema „Kompetenzen bzw. Kompetenzaufbau für die Bereitstellung öffentlicher Daten“?

17. (Neue) Berufsfelder – Karrierewege

Einstieg:

- Während sich in angelsächsischen Ländern, aber auch in den Niederlanden, neue Berufsfelder etablieren, scheint das in Deutschland noch nicht der Fall zu sein. Wie schätzen Sie die Lage ein?

Mögliche Nachfragen:

- Was halten Sie von den Berufsbezeichnungen „data steward“, „data scientist“, „data manager“?
- Welche Inhalte (Kompetenzen) verbinden Sie damit?
- Gibt es die Berufsbilder bereits oder wären sie wünschenswert?
- Welche weiteren würden Sie noch ergänzen wollen?

18. Kompetenzprofile/Fähigkeiten

- Welche Kompetenzen sind aus Ihrer Sicht bei den am Bereitstellungsprozess von Daten beteiligten Personen erforderlich?
- Besondere IT-Kenntnisse?
- Fachwissenschaftliche Kenntnisse in Abhängigkeit von den bereitzustellenden Daten?
- Rechtliche Kenntnisse?
- Rechtliche Kenntnisse welcher Art: Datenschutz, Persönlichkeitsrechte, Verwertungsrechte, Urheberrecht?
- Welche Kompetenzen sollten Ihrer Meinung nach vermittelt werden?

10.5 Verzeichnis der Projektbeteiligten

Herausgeber*innen

Dr. Christoph Bruch ist Referent des Helmholtz Open Science Office und im Projekt EcoDM – Ökosystem Datenmanagement Projektinitiator für den Bereich Wirtschaft. Er studierte Politikwissenschaft an der Johann Wolfgang Goethe-Universität Frankfurt und an der Freien Universität Berlin. Vor seiner Tätigkeit in der Helmholtz-Gemeinschaft hatte er berufliche Stationen an der Freien Universität Berlin, dem Deutschen Institut für Urbanistik und der Max-Planck-Gesellschaft. Ehrenamtlich setzt er sich für den Zugang zu Wissen ein, indem er sich in mehreren Nichtregierungsorganisationen engagiert, z. B. im Aktionsbündnis „Urheberrecht für Bildung in der Forschung“ und im Europäischen Netzwerk für Urheberrecht zur Unterstützung von Bildung und Wissenschaft. Christoph Bruch vertritt die Helmholtz-Gemeinschaft im Organizational Board der Research Data Alliance und gehört dem Vorstand des Vereins RDA DE an. Aktuelle Arbeitsschwerpunkte von ihm sind die Formulierung von Policies auf nationaler und internationaler Ebene, die den Übergang zum Open-Access-Publizieren bzw. zu Open Science unterstützen sowie die Begleitung damit zusammenhängender politischer Entscheidungsprozesse. Zu Letzteren zählt auch das Werben für ein wissenschaftsfreundliches Urheberrecht.

Prof. Heike Neuroth ist im Projekt EcoDM – Ökosystem Datenmanagement Projektinitiatorin für den Bereich Qualifizierung. Seit April 2015 ist sie Professorin an der Fachhochschule Potsdam im Fachbereich Informationswissenschaften und seit 2020 Forschungsprofessorin. Ihre Forschungsschwerpunkte liegen in den Bereichen Digital Humanities und Forschungsdaten(-management). Heike Neuroth hat 1997 in Geologie/Paläontologie an der Georg-August Universität Göttingen promoviert und 2002 ihren MA-Abschluss in Bibliothekswissenschaft an der Humboldt-Universität zu Berlin gemacht. Sie war von 1997 bis 2015 an der Niedersächsischen Staats- und Universitätsbibliothek (SUB) und hat dort zuletzt als Leiterin die Abteilung Forschung und Entwicklung verantwortet. Darüber hinaus war sie als eHumanities-Referentin von 2008 bis 2012 bei der Max Planck Digital Library (MPDL) tätig. Heike Neuroth ist in zahlreichen nationalen und internationalen Gremien engagiert, u. a. ist sie stellvertretende Vorsitzende im RDA DE Verein und Mitglied im Steuerungsgremien der GO FAIR Initiative. Heike Neuroth hat zahlreiche Forschungs- und Entwicklungsprojekte geleitet (z. B.

TextGrid als Virtuelle Forschungsumgebung, DARIAH als Forschungsinfrastruktur für die Geisteswissenschaften) und ist Mitherausgeberin einiger Sammelwerke. Seit 2020 verantwortet sie gemeinsam mit Prof. Vivien Petras von der HU Berlin den Weiterbildenden Master Digitales Datenmanagement.

Henriette Senst ist im Projekt EcoDM – Ökosystem Datenmanagement Projektinitiatorin für den Bereich Public Sector. Sie hat Bibliotheks- und Informationswissenschaft an der Humboldt-Universität zu Berlin studiert. Seit Juli 2019 ist Henriette Senst Bibliotheksdirektorin im Deutschen Archäologischen Institut und leitet dort gemeinsam mit dem IT-Direktor die „Zentralen Wissenschaftlichen Dienste“. Ihre Forschungsschwerpunkte liegen in den Bereichen Open Science und Forschungsdatenmanagement. Sie ist Co-Spokesperson der Task Area „Storage, Access and Dissemination“ im Rahmen der Antragstellung für die NFDI4Objects. Außerdem ist Henriette Senst im Vorstand des RDA Deutschland e.V. tätig.

Autor*innen

Christine Burkart ist Projektadministratorin und -koordinatorin im Projekt EcoDM – Ökosystem Datenmanagement und verantwortet für die Fachhochschule Potsdam (FHP) die organisatorische und koordinative Abwicklung des Projekts. Sie studierte Neuere und Neueste Geschichte, Allgemeine Rhetorik und Kunstgeschichte an der Universität Tübingen und schloss ihre Studien mit einem Magister ab. Danach war sie im Kulturbereich mehrere Jahre selbstständig und wirkte eigene fotografische und filmische Projekte und Auftragsarbeiten ab, schrieb Vermittlungskonzepte und gestaltete Führungen und Workshops in verschiedenen Museen in Berlin. Zu ihren Aufgaben gehört die Betreuung und Organisation des Projektmanagements, die Öffentlichkeitsarbeit (Webseite und Twitter) und die Erledigung des administrativen Tagesgeschäfts sowie die Erstellung von Texten zur Öffentlichkeitsarbeit und für Teile des Public Sectors. Ihre Expertise im fotografischen und künstlerischen Bereich fließt in die grafische Umsetzung in verschiedenen Gebieten mit ein.

Maria Chlastak ist akademische Mitarbeiterin an der FH Potsdam. Seit November 2021 ist sie Projektbereichskoordinatorin und -referentin für den Bereich Wissenschaft im Projekt EcoDM – Ökosystem Datenmanagement. Zuvor hat sie die Literaturrecherche für verschiedene Bereiche im Projekt unterstützt. Die gebürtige Wienerin studiert Sozialwissenschaften im Master an der Humboldt-Universität zu Berlin. Bis 2020 studierte Maria Politikwissenschaft an der Universität Wien und Innsbruck. Ihre Studienschwerpunkte liegen im Bereich Digitalisierung, Daten und

Demokratie. In diesem Kontext war sie beruflich in verschiedenen demokratie-fördernden Projekten tätig. Außerdem hatte sie ein kleines Unternehmen zur Datenerfassung, das Aufträge von Greenpeace Österreich und von Markt- und Meinungsforschungsunternehmen bearbeitete.

Christian Djeffal (Kapitel 3) ist Assistant Professor für Law, Science and Technology an der Technischen Universität München. Am Munich Center for Technology in Society und an der TUM School of Governance beschäftigt er sich mit dem Verhältnis von Recht und Technologie und arbeitet schwerpunktmäßig zu neuen Technologien wie künstlicher Intelligenz und dem Internet der Dinge. Dabei liegt sein Augenmerk auf integrativen und transdisziplinären rechtswissenschaftlichen Ansätzen. Er befasst sich aktiv mit der Digitalisierung von Recht und Rechtswissenschaft.

Kirsten Endres betreute im Projekt EcoDM – Ökosystem Datenmanagement als Projektreferentin und Projektbereichskoordinatorin den Aufgabenbereich Public Sector für das Deutsche Archäologische Institut. Zuvor unterstützte sie als wissenschaftliche Hilfskraft an der Fachhochschule Potsdam das Team des Projekts EcoDM im Bereich Öffentlichkeitsarbeit und durch die Erstellung des Datenmanagementplans. Sie war außerdem für den Public Sector mit der Kodierung und Auswertung der Expert*innen-Interviews sowie der Literaturauswertung beschäftigt. Hierbei kann sie auf ihre langjährigen Erfahrungen in der Beratung von Wissenschaftler*innen beim wissenschaftlichen Arbeiten und Publizieren in der Abteilung Research Services & Open Science des Fraunhofer IRB zurückgreifen. Parallel zur Projektarbeit und darüber hinaus erweitert Kirsten Endres ihre Kenntnisse im Bereich des Datenmanagements durch den Masterstudiengang Digitales Datenmanagement.

Dr. Reinhard Messerschmidt war Referent am Helmholtz Open Science Office und Projektbereichskoordinator „Wirtschaft“. Als interdisziplinärer Sozialwissenschaftler mit Promotion in Philosophie war er zuvor als wissenschaftlicher Referent für Digitalisierung beim Wissenschaftlichen Beirat der Bundesregierung Globale Umweltveränderungen (WBGU), als Referent in der Geschäftsstelle Hightech-Forum bei der Fraunhofer Gesellschaft sowie als wissenschaftlicher Mitarbeiter bei GESIS – Leibniz-Institut für Sozialwissenschaft und am Cologne Center for eHumanities tätig. Seit September 2021 arbeitet er bei der Zukunft – Umwelt – Gesellschaft (ZUG) gGmbH als fachlicher Projektmanager für Künstliche Intelligenz und befasst sich im dortigen Team primär mit der Umsetzung der Initiative „Civic Coding – Innovationsnetz KI für das Gemeinwohl“ (von BMAS, BMFSFJ und BMUV) und

dem Aufbau von KI-Ideenwerkstätten für Umweltschutz.

Eva Nürnberger (Kapitel 3) ist studentische Mitarbeiterin am Lehrstuhl für Recht, Wissenschaft und Technik an der TU München (Prof. Christian Djeffal). Sie studiert Rechtswissenschaften an der Ludwig-Maximilians-Universität München.

Dr. Janina Richter ist Projektreferentin und Projektbereichskoordinatorin im Projekt EcoDM – Ökosystem Datenmanagement und betreut dort den Aufgabenbereich Wirtschaft für das Helmholtz Open Science Office. Zu ihren Aufgaben gehört die Auswertung von Stakeholder-Interviews sowie die Analyse der Akteurslandschaft, Handlungsfelder und Infrastrukturen bei der Produktion und Nachnutzung von Daten im Bereich der Wirtschaft in Deutschland. Sie ist studierte Umweltgeologin und hat einen PhD in organischer Geochemie. Vor ihrer Tätigkeit im Projekt EcoDM war sie als Projektkoordinatorin und stellvertretende Geschäftsführerin in einem mittelständigen Bauunternehmen tätig.

Esther Schneidenbach war Projektreferentin und Projektbereichskoordinatorin im Projekt EcoDM – Ökosystem Datenmanagement und betreute dort den Aufgabenbereich Public Sector für das Deutsche Archäologische Institut (DAI). Zu ihren Aufgaben gehörte die Analyse des Standes des digitalen Datenmanagements im Public Sector in Deutschland. Gemeinsam mit den anderen Projektpartnern ist sie zusätzlich für den Bereich Wissenschaft im Projekt zur gleichen Themenstellung zuständig. Neben einer Landscape-Analyse widmete sie sich der Befragung von Stakeholdern in Form von Interviews und das Community Management. Vor ihrer Tätigkeit im Projekt EcoDM betreute sie die Abschlussphase der Informationsplattform forschungslizenzen.de, auf der Fragen zum Umgang mit Lizenzen in der Wissenschaft exemplarisch beantwortet werden. Zuvor war sie am Aufbau des Forschungsdatenzentrums der Archäologie und Altertumswissenschaften – IANUS beteiligt. Seit nunmehr über 10 Jahren beschäftigt sie sich mit digitalem Datenmanagement.

Claus Spiecker ist Projektreferent und Projektbereichskoordinator im Projekt EcoDM – Ökosystem Datenmanagement und betreut dort den Aufgabenbereich Qualifizierung für die Fachhochschule Potsdam (FHP). Er hat einen Magister in Geschichte der Naturwissenschaften und Technik (GNT) und Kunstgeschichte (1999) der Universität Stuttgart und einen Master in Library and Information Science der Humboldt-Universität zu Berlin (2011). Seit März 2016 arbeitet er am Fachbereich Informationswissenschaften

der FHP in verschiedenen Projekten (PARTHENOS, MaMoDaR) mit den Schwerpunkten Forschungsdatenmanagement, Open Science, Kommunikation und Qualifizierung. Von 2011-2016 war er als wissenschaftlicher Mitarbeiter am Fraunhofer IRB in Stuttgart in der Abteilung Research Services & Open Science des Fraunhofer IRB und hat – neben der Beratung zu Themen des wissenschaftlichen Publizierens und der Entwicklung von Qualifizierungsangeboten – das Projekt FORDATIS zum Aufbau des Forschungsdaten-Repository der Fraunhofer-Gesellschaft mitentworfen. Als Repräsentant der Fraunhofer-Gesellschaft war er in diverse Gremien zu den Themen Open Access und Forschungsdaten entsandt. Er lehrt seit 2003 an Hochschulen (Uni Stuttgart, FHP), konzipiert und führt Qualifizierungsangebote auch außerhalb von Hochschulen durch und ist über 10 Jahre im Bereich Forschungsdaten(management) aktiv.

Paul Walter ist Projektreferent und Projektbereichskoordinator im Projekt EcoDM – Ökosystem Datenmanagement und betreut dort den Aufgabenbereich Public Sector für das Deutsche Archäologische Institut (DAI). Zu seinen Aufgaben gehört die abschließende Auswertung der Untersuchungen zum Public Sector in Deutschland. Er hat Altertumswissenschaften und Ägyptologie an der Freien Universität Berlin sowie Denkmalpflege an der Martin-Luther-Universität Halle-Wittenberg studiert.

Beteiligte

Jasper Bothe unterstützt EcoDM seit August 2021 als studentische Hilfskraft. Er studiert Bibliothekswissenschaft an der Fachhochschule Potsdam, nachdem er zuvor Theater- und Filmwissenschaft an der Freien Universität Berlin studierte.

Dr. Tanja Friedrich ist Informationswissenschaftlerin und Leiterin der Organisationseinheit „Wissenschaftliche Information“ beim Deutschen Zentrum für Luft- und Raumfahrt. Zuvor war sie unter anderem 10 Jahre wissenschaftliche Mitarbeiterin beim GESIS Leibniz-Institut für Sozialwissenschaften.

Reiner Mauer ist stellvertretender Leiter der Abteilung Data Services for the Social Sciences bei GESIS – Leibniz-Institut für Sozialwissenschaften und leitet dort das Team Archiving. Er studierte Volkswirtschaft und Soziologie an der Universität Köln. Seit 1998 ist er in verschiedenen Projekten und Positionen bei GESIS mit der Kuratierung, Langzeitarchivierung und Bereitstellung von Forschungsdaten befasst. Seit 2006 ist er verantwortlich für die zentralen Prozesse der Datenarchivierung. Neben der operativen Steuerung, gehört die

konzeptionelle und praktische Weiterentwicklung der Datenarchivierung und damit verbundener Services zu seinen Aufgaben. Er ist u. a. Projektleiter des von DIPF, GESIS und IQB getragenen Verbund Forschungsdaten Bildung (VerbundFDB) und Mitglied des SprecherInnenkreises des AK Forschungsdaten der Leibniz-Gemeinschaft.

Florian Müller ist wissenschaftlicher Mitarbeiter und Doktorand am Fachgebiet „Soziologische Theorie“ an der Universität Kassel. Als Mitglied des DFG-Graduiertenkollegs „Privacy and Trust for Mobile Users“ beschäftigt er sich in seiner Forschung mit der Gestaltung, Bewertung und Regulierung von algorithmischen Koordinations- und Regelungsverfahren auf Online-Plattformen. Zuvor studierte er „Soziologie technikkundlicher Richtung“ an der Technischen Universität Berlin und hat als studentischer Mitarbeiter am Wissenschaftszentrum Berlin, am Weizenbaum-Institut und im Projekt EcoDM als wissenschaftliche Hilfskraft an der FHP gearbeitet.

Dr. Raphael Ritz studierte Physik an der TU München und hat dort in Theoretischer Physik promoviert. Heute leitet er den Bereich Datendienste der Max Planck Computing and Data Facility (MPCDF) und ist im nationalen (NFDI), europäischen (EOSC) sowie internationalen (RDA) Umfeld des Forschungsdaten Managements engagiert.

Laura Rothfritz ist wissenschaftliche Mitarbeiterin und Doktorandin am Lehrstuhl Information Management des Instituts für Bibliotheks- und Informationswissenschaft an der Humboldt-Universität Berlin. Sie hat ihr Masterstudium der Informationswissenschaften 2019 an der Fachhochschule Potsdam abgeschlossen und war vor ihrem Wechsel an die HU Berlin als Mitarbeiterin in unterschiedlichen Projekten im Bereich Forschungsdatenmanagement, u. a. EcoDM, am Fachbereich Informationswissenschaften der FH Potsdam angestellt. Ihre Forschungs- und Lehrtätigkeiten liegen im Bereich Forschungsdatenmanagement, digitale Forschungsinfrastrukturen und Open Science.

Dr. Wolfgang Schmidle ist seit über zehn Jahren wissenschaftlicher Mitarbeiter in der IT und in der Abteilung „Zentrale Wissenschaftliche Dienste“ des Deutschen Archäologischen Instituts (DAI). Er hat in Bonn Mathematik studiert und in Sunderland in Informatik promoviert. Zurzeit vertritt er das DAI in den beiden EU-Infrastruktur-Projekten SSHOC und IPERION HS. Er betreut außerdem den zur iDAI welt gehörenden Zeitbegriffe-Gazetteer ChronOntology. Ein Schwerpunkt seiner Arbeit ist die semantische Modellierung von Daten.

Weitere Informationen zum EcoDM-Projekt

www.ecodm.de