

Forschungsdatenkooperationen zwischen Wissenschaft und Wirtschaft ermöglichen

Bericht vom Workshop am 20.02.2025 im Quadriga Forum in Berlin

Hintergrund

Die fortschreitende Digitalisierung führt zu einem stetigen Datenwachstum, das sowohl für die Wirtschaft als auch für die Wissenschaft vielversprechende Möglichkeiten eröffnet. Unternehmen können datenbasierte Geschäftsmodelle entwickeln, während die Wissenschaft diese Datenbestände für innovative Grundlagenforschung nutzen kann. Allerdings sind diese Daten oft nicht frei zugänglich und in ihrer Verwertbarkeit oft schwer zu beurteilen. Kooperationen zwischen Wissenschaft und Wirtschaft, die den Austausch von Daten ermöglichen und realisieren, sind daher von zentraler Bedeutung, um wissenschaftlichen, wirtschaftlichen und damit gesamtgesellschaftlichen Nutzen aus dem Datenwachstum zu ziehen. Denn erfolgreiche Forschungsdatenkooperationen können öffentlichen und privaten Forschenden den Zugang zu qualitativ hochwertigen Daten ermöglichen, die für evidenzbasierte Forschung und damit für die Analyse zentraler Fragestellungen, die Vorhersage zukünftiger Entwicklungen oder die Entwicklung innovativer Produkte und Dienstleistungen benötigt werden.

Die Deutsche Forschungsgemeinschaft (DFG) hat sich in diesem Zusammenhang im Jahr 2023 dafür ausgesprochen, das freiwillige Teilen von Daten zwischen Wissenschaft und Wirtschaft in der Praxis zu intensivieren und Modelle zu entwickeln, um geeignete Rahmenbedingungen für Forschungsdatenkooperationen zu erarbeiten. Aus diesem Grund veranstaltete die DFG in Kooperation mit dem Stifterverband und dem Bundesverband der Deutschen Industrie (BDI) im Mai 2024 einen Workshop, bei dem insgesamt 48 Teilnehmer*innen aus Wirtschaft und Wissenschaft diskutierten, welche Schlüsselfaktoren für erfolgreiche Forschungsdatenkooperationen wichtig sind und wie Forschungsdatenkooperationen zukünftig erfolgreich initiiert und gestaltet werden können. Aus dem Workshop ergaben sich insgesamt fünf Handlungsempfehlungen für die Initiierung von Forschungsdatenkooperationen: 1. gegenseitiges Verständnis für die jeweilige Position schaffen und Vertrauen aufbauen, 2. Win-Win-Situationen schaffen, 3. rechtliche Rahmenbedingungen klären, 4. gemeinsame Standards für die Datenkuration etablieren, 5. Konflikt- und Erwartungsmanagement sowie offene Kommunikation der Kooperationspartner unterstützen.

Aufbauend auf diesen Ergebnissen veranstaltete die DFG in Kooperation mit dem Stifterverband im Februar 2025 einen zweiten Workshop. Ziel der Veranstaltung war es, ein besseres Verständnis dafür zu erlangen, welche Voraussetzungen für Forschungsprojekte mit Forschungsdatenkooperationen erfüllt sein müssen. Insgesamt 55 Teilnehmer*innen aus Wirtschaft, Wissenschaft und Verbänden diskutierten Fragen zu technischen und rechtlichen Rahmenbedingungen, Datenaustauschmodellen, Vernetzung von Initiativen, Datensouveränität und Wertschöpfung.

Die im Folgenden zusammengefassten Ergebnisse dieses Workshops sollen als Grundlage für eine mögliche Ausschreibung der DFG dienen, um Forschungsdatenkooperationen zu fördern und damit die Vernetzung von Wissenschaftler*innen und Unternehmen zu ermöglichen.

Workshopzusammenfassung

Der Workshop gliederte sich in zwei Blöcke, die sich sowohl in der inhaltlichen Schwerpunktsetzung als auch in der methodischen Herangehensweise unterschieden.

In **Block 1** wurde die Bedeutung von Forschungsdatenkooperationen aus unterschiedlichen Perspektiven in zwei Impulsvorträgen beleuchtet. Dr. Svetoslava Antonova-Baumann vom Stifterverband hob die Bedeutung von Forschungsdatenkooperationen als Innovationstreiber für Wissenschaft, Wirtschaft und Gesellschaft hervor und präsentierte Ergebnisse aus dem letzten Hochschulbarometer zur Lage und Entwicklung der Hochschulen aus Sicht ihrer Leitungen. Im Vortrag wurde einerseits deutlich, dass gegenseitiges Vertrauen zwischen den Kooperationspartnern sowie eine gewisse räumliche Nähe entscheidende Vorteile für den Erfolg von Kooperationen sind, und andererseits, dass Kooperationen zwischen Wissenschaft und Wirtschaft zwar als sehr wichtig, gleichzeitig aber auch als schwierig angesehen werden, insbesondere aufgrund von Hürden wie Datenschutzbestimmungen. Im Anschluss beleuchtete Prof. Dr. Tobias Kuhnimhof von der RWTH Aachen Potenziale, wie den Zugang zu qualitativ hochwertigen Daten, und Herausforderungen, wie den selektiven Charakter von Unternehmensdaten, bei Forschungsdatenkooperationen für die Wissenschaft. Zudem zeigte er drei Wege auf, wie Wissenschaftler*innen Zugang zu Daten aus Unternehmen erhalten können: 1. der Eins-zu-eins-Kontakt zwischen Wissenschaftler*innen und Unternehmen, 2. der Kauf von Daten über Datenaggregatoren, 3. die Bereitstellung und der Abruf von Daten über Datenmarktplätze (wie bspw. der Mobility Data Space) und weitere Infrastrukturen wie Datenvermittlungsdienste und Datentreuhänder. Aufgrund großer rechtlicher (z.B. IP-Rechte, Datenschutz, langwierige Verhandlungen über Non-Disclosure-Agreements (NDAs)) aber auch finanzieller Hürden (v.a. hohe Kosten und Gebühren der Datenaggregatoren) bei den ersten beiden Wegen warb Herr Kuhnimhof für kooperativ aufgebaute Datenmarktplätze als einen gangbaren Weg, um Forschungsdatenkooperationen zu realisieren.

Mit Blick auf das Leitthema des Workshops – Forschungsdatenkooperationen ermöglichen – konnten drei Schlussfolgerungen aus den Impulsvorträgen gezogen werden.

1. Die Kooperationspartner müssen Vertrauen aufbauen und ein **gegenseitiges Verständnis** für die Bedürfnisse und Anforderungen der jeweils anderen Seite entwickeln.
2. Die **Klärung der rechtlichen Rahmenbedingungen** ist wichtig, um die Interessen aller Beteiligten zu wahren und Kooperationen rechtssicher zu gestalten. Derzeit sind jedoch vor allem die finanziellen und zeitlichen Kosten für rechtssichere Daten- und Kooperationsverträge zu hoch und müssen prospektiv gesenkt werden.
3. **Infrastrukturen** wie Datenräume, Datentreuhänder oder Datenvermittlungsdienste sind notwendig, um Daten rechtssicher, effizient und kostengünstig über Sektorengrenzen hinweg teilen und austauschen zu können.

Block 2 zielte darauf ab, im Rahmen von themenzentrierten Gruppendiskussionen Anforderungen und Herausforderungen sowie Bedarfe und potenzielle Ziele einer möglichen Ausschreibung zu identifizieren. Diskutiert wurde in drei Themengruppen: 1. Gesundheits- und Medizindaten, 2. Verkehrs- und Wirtschaftsdaten, 3. Infrastrukturen für den Datentransfer.

Erkenntnisse aus der Gruppendiskussion „Gesundheitsdaten“: Zunächst grenzten die Teilnehmenden die unterschiedlichen Ziele und Erwartungen von Wissenschaft und Wirtschaft ab: Während die Wissenschaft an Erkenntnisgewinn und Reputation durch Grundlagenforschung interessiert ist, was sich durch Publikationen und Translation auch in ihrer Leistungsbewertung abbildet, sind Wirtschaftsunternehmen im Public Health-Bereich an einer wirksamen und rentablen Patientenversorgung interessiert, wobei sie neben dem gesellschaftlichen Nutzen auch ökonomische Dimensionen wie Umsatz, Gewinn und Shareholder Value berücksichtigen müssen. Eine gemeinsame Schnittstelle wurde darin gesehen, dass auf der einen Seite die wissenschaftlichen Akteure am Transfer ihrer Erkenntnisse in die medizinische Versorgung interessiert sind, während auf der anderen Seite die wirtschaftlichen Akteure ein großes Interesse an der Produktisierung neuester wissenschaftlicher Erkenntnisse haben. Dies könnte zu gesamtgesellschaftlichem Nutzen durch zielgerichtetes und strategisches Co-Development erreicht werden, setzt jedoch gegenseitiges Vertrauen und das Navigieren durch die teils unübersichtlichen und fragmentierten geltenden Rechtsvorschriften voraus.

Darüber hinaus sahen die Teilnehmenden einen besonderen Bedarf an digitalen Technologien und Infrastrukturen, die den Datentransfer, die Datennutzung und -nachnutzung sowie die Rekombinierbarkeit von Daten transparent und rechtssicher ermöglichen. Modelle der partizipativen Data Governance scheinen hier besonders erstrebenswert. Als weitere Herausforderungen für Forschungsdatenkooperationen speziell im Bereich der Gesundheitsdaten wurden die Notwendigkeit der Entwicklung von Unique Patient Identifiers sowie die Harmonisierung des nationalen und europäischen Rechtsrahmens genannt. Für Daten- und Kooperationsverträge einschließlich der Regelung von geistigem Eigentum und Datenschutzklauseln sei die Etablierung standardisierter Vertragsverfahren wünschenswert. Diese Verträge sollten zudem die mögliche Sekundärnutzung der Forschungsdaten regeln. Abschließend wurden bestehende Best-Practice-Modelle für IP-Rechte, Lizenzierung, Daten als Wirtschaftsgut, Kostenteilung im Gesundheits- und Medizinbereich, Data-Linkage, Datentreuhandmodelle sowie Repositorien gesammelt und diskutiert. Insgesamt wurde festgestellt, dass die Situation unübersichtlich und weitere Informations- und Austauschplattformen und -veranstaltungen wünschenswert seien.

Erkenntnisse aus der Gruppendiskussion „Verkehrs- und Wirtschaftsdaten“: Die Teilnehmenden in dieser Gruppe konzentrierten sich in ihrer Diskussion insbesondere auf Wünsche für und Ziele von zukünftigen Forschungsdatenkooperationen. Verschiedene Wünsche wurden geäußert, darunter insbesondere die Klärung von Haftungs- und Versicherungsfragen, die Überwindung von Datenschutzauflagen durch standardisierte und automatisierte Verfahren und die Etablierung rechtssicherer und standardisierter Data-Sharing-Templates. Die Notwendigkeit von (Muster-)Kooperationsverträgen bzw. einer standardisierten Vertragsgestaltung und -verhandlung wurde deutlich gemacht, um insbesondere die bisherige Praxis langwieriger individueller Vertragsverhandlungen zu Themen wie Non-Disclosure-Agreements, Lizenzierung und den richtigen Plattformen für den Datenaustausch deutlich zu vereinfachen und zu beschleunigen.

Darüber hinaus würden Forschungsdatenkooperationen von offenen, sektoren- und disziplinenübergreifenden Datenstandards im Sinne der FAIR-Prinzipien und von maschinenlesbaren Metadaten sowie von offenen und rechtssicheren Infrastrukturen für Datenaustausch und Datenverarbeitung profitieren. Als Ziele für Forschungsdatenkooperationen nannten die Teilnehmenden die Übertragbarkeit und Skalierbarkeit von Forschungsergebnissen in die wirtschaftliche/unternehmerische Praxis, Möglichkeiten für Co-Development-Ansätze, das Sichtbarmachen erfolgreicher Kooperationen, z.B. durch eine Datenbank, die Use Cases aufzeigt und damit einhergehend die Systematisierung dieser Use Cases.

Erkenntnisse der Gruppendiskussion „Infrastrukturen für das Datenteilen“: Die Diskussion zu Infrastrukturen konzentrierte sich auf die Themenkomplexe Data Sharing & Use sowie Datenqualität & Datenaufbereitung. Im Kontext des kooperativen Teilens und Nutzens von Daten wurden insbesondere Datentreuhandmodelle als vielversprechende infrastrukturelle Ansätze gesehen, um die Bereitstellung, Verfügbarkeit und Auffindbarkeit von Daten aus der Wirtschaft, aber auch von Forschungsdaten und Forschungsergebnissen zu gewährleisten. Datentreuhänder können als technischer, organisatorischer und rechtlicher Rahmen dienen, um das Teilen, das Nutzen und das Nachnutzen von Daten rechtssicher und standardisiert zu realisieren. Die Teilnehmenden sahen hier insbesondere das Datenpooling und abgestufte Zugänge für Nutzerkreise als großen Mehrwert von Datentreuhandmodellen, während besondere Unsicherheiten in der zeitlich begrenzten Projektfinanzierungslogik gesehen wurden, die im Widerspruch zur Notwendigkeit langfristiger und nachhaltiger Infrastrukturmodelle steht.

Zum Themenkomplex Datenqualität und Datenaufbereitung wurden von den Teilnehmenden verschiedene Hemmnisse identifiziert. Während für die Wissenschaft die hohe Personalfluktuation und ein noch notwendiger Kulturwandel im Hinblick auf den Aufbau langfristig verfügbarer digitaler Expertise in den wissenschaftsunterstützenden Diensten und Services genannt wurden, wurden für die Wirtschaft fehlende Standards und die damit verbundene Interoperabilität von Daten sowie - zumindest bei KMU - fehlendes Know-how angeführt. Des Weiteren wurde diskutiert, wie Infrastrukturen dazu beitragen können, den Aufwand für die gemeinsame Nutzung und das Teilen von Daten für alle Beteiligten zu reduzieren. Dies kann einerseits durch Datenpooling bei Datentreuhändern und andererseits durch die Vereinheitlichung von Metadatenstandards bzw. die Implementierung des FAIR Digital Objects (FDO) Konzepts erfolgen, das die Anschlussfähigkeit von Daten unabhängig von ihrer zugrundeliegenden Metadatenstruktur gewährleisten kann. In Bezug auf die Anschlussfähigkeit von Daten sollte die Nutzbarmachung/*Readiness* dieser Daten für Large Language Models und andere AI-basierte Werkzeuge intensiv verfolgt werden. Schließlich wurde auch in dieser Diskussionsrunde die große Bedeutung von standardisierten Daten- und Kooperationsverträgen betont.

Zusammenfassend lassen sich mehrere Gemeinsamkeiten aus den Ergebnissen der Gruppendiskussionen identifizieren:

1. Die **Notwendigkeit standardisierter Daten- und Kooperationsverträge**. Rechtssicherheit muss bei Forschungsdatenkooperationen gewährleistet sein, während langwierige individuelle Verhandlungen zwischen potenziellen Kooperationspartnern vermieden werden sollten.

2. Die **Bedeutung von Infrastrukturen** zur Ermöglichung von Datenteilen und Datennutzung. In allen drei Gruppen wurden Datentreuhandmodelle als ein vielversprechender infrastruktureller Ansatz hervorgehoben, um die Bereitstellung, Verfügbarkeit und Auffindbarkeit von Daten zu gewährleisten. Weiterhin wurde die Bedeutung von Technologien und digitalen Informationsinfrastrukturen, die Datentransfer, -nutzung, -weiterverwendung und -rekombinierbarkeit transparent und rechtssicher ermöglichen betont.
3. Die Rolle von **Datenstandards und Metadaten** für erfolgreiche Forschungsdatenkooperationen. Die Teilnehmenden wünschen sich offene, sektor- und disziplinübergreifende Datenstandards im Sinne der FAIR-Prinzipien und maschinenlesbare Metadaten. Das Konzept der FAIR Digital Objects (FDO) wird als zentraler infrastruktureller Ansatzpunkt für die cross-sektorale Nutzung von Daten genannt. In Kombination mit Datentreuhandstrukturen oder anderen geeigneten Infrastrukturen für die gemeinsame Nutzung und den Austausch von Daten würde dies auch die Anschlussfähigkeit von Daten aus unterschiedlichen Quellen und mit unterschiedlichen Typen verbessern.
4. Das Potenzial von **Co-Development und dem Transfer von Forschungsergebnissen** in die wirtschaftliche Praxis, um Anreize für Forschungsdatenkooperationen zu schaffen. Während die Wissenschaft am Erkenntnisgewinn durch Grundlagenforschung interessiert ist, hat die Wirtschaft ein Interesse an der Umsetzung wissenschaftlicher Erkenntnisse in neue Produkte und Dienstleistungen. Dabei ist allen Beteiligten bewusst, dass ein Interessensaustausch und eine vertrauensbasierte und vor allem rechtssichere Kooperation auf Augenhöhe zwischen den beteiligten Partnern geschaffen werden muss, um einen Mehrwert für beide Seiten zu generieren.

Ideen und Bedarfe für eine mögliche DFG-Ausschreibung zur Förderung von Forschungsdatenkooperationen

Ein zentrales Ziel des Workshops war es, Bedarfe, Anforderungen und Ziele für eine mögliche Ausschreibung der DFG zu sammeln, mit der die DFG Impulse setzen und Forschungsdatenkooperationen ermöglichen will. Bei der plenaren Vorstellung der Resultate aus den Gruppendiskussionen traten zahlreiche Übereinstimmungen bzw. ähnlich formulierte Vorschläge für eine potenzielle Ausschreibung zutage:

Ziele

- Vorrangiges Ziel der DFG ist die **Förderung des Erkenntnisgewinns durch die Nutzung miteinander geteilter Daten**. An dieser Stelle wurde die Bedeutung von **Co-Development** zwischen Wissenschaft und Wirtschaft hervorgehoben und als mögliches zweites Ziel einer Ausschreibung betont.
- **Förderung der Zusammenarbeit zwischen Wissenschaft und Wirtschaft**. Die Gruppen betonten die Notwendigkeit und den Wunsch nach intensiveren Kooperationen zwischen Wissenschaft und Wirtschaft. Ziel einer Ausschreibung sollte es daher sein, den Austausch und das gemeinsame Entwickeln von Kooperationsmodellen zu ermöglichen, um wissenschaftliche

Erkenntnisse und die Erschließung neuer Anwendungspotenziale und damit einen Mehrwert für beide Seiten zu generieren.

- **Schaffung von Rechtssicherheit und standardisierten Rahmenbedingungen.** Forschungsdatenkooperationen benötigen einen rechtlich abgesicherten Rahmen durch Verträge und die Vereinfachung von Verhandlungsprozessen durch Musterverträge. Die Ausschreibung sollte sich nicht nur selbst durch die Setzung von Rahmenbedingungen auszeichnen, sondern zusätzlich Anreize für die Etablierung von Standards bieten.
- Die **Nutzung von Infrastrukturen und Technologien**, die einen rechtssicheren, technischen und organisatorischen Rahmen gewährleisten, sollte bei einer Ausschreibung vorrangig berücksichtigt werden. Dadurch soll die **intersektorale und domänenunabhängige Datennutzung und -qualität** verbessert werden. Die Ausschreibung sollte Anreize schaffen, über die fach- und projektspezifische Ebene hinaus das weitergehende Teilen, Nutzen und Nachnutzen von Daten sowie die Aufbereitung und das Kuratieren von Daten fördern. Anträge sollten hierfür insbesondere die FAIR Data Prinzipien und etablierte Metadatenstandards berücksichtigen und darüber hinaus die *AI-readiness* der Daten sicherstellen.

Kriterien für die Förderung und Begutachtung

- Die **wissenschaftliche Qualität** und der **Beitrag zum Erkenntnisgewinn** sind zentrale Kriterien. Dennoch sollte eine Ausschreibung verstärkt auf kreative und disruptive Ideen abzielen, um so auch den Wissenstransfer und den Transfer in anwendungsähnere Kontexte gezielt zu fördern.
- Die **Machbarkeit und Skalierbarkeit** von Projekten sollte von Anfang an bedacht werden. Projektanträge sollten realistisch durchführbar sein und idealerweise das Potenzial zur Skalierung oder Übertragung auf andere Anwendungsbereiche haben und dies entsprechend darstellen.
- Die aktive **Vernetzung von Wissenschaft, Wirtschaft und Initiativen, Verbänden und Daten-Konsortien** (z.B. NFDI, GAIA-X, Catena-X, EHDA, Mobility Data Space, Datenraum Kultur etc.) sollte im Rahmen einer möglichen Ausschreibung von Anfang an eingeplant werden. Insbesondere die Einbindung und Nutzung bestehender Technologie, Datentreuhandstrukturen und Datenräume sollte Bedingung für die Förderung werden.
- Die gegenseitigen **Mehrwerte und Interessen der Kooperationspartner** müssen in den Blick genommen werden und Projekte sollten darauf eingehen, inwiefern ein Mehrwert sowohl für Wissenschaft als auch Wirtschaft geschaffen werden kann.
- **Umgang mit rechtlichen und ethischen Aspekten beachten.** Die Einhaltung von Datenschutzbestimmungen, die Klärung von Urheberrechten und der Umgang mit geistigem Eigentum müssen berücksichtigt werden.
- Bei der Begutachtung und Bewertung möglicher Projektanträge sollte der **unterschiedliche Erfahrungsstand** von Wissenschaft und Wirtschaft mit der DFG-Förderung berücksichtigt und die domänenspezifische Expertise der Wirtschaft eingeholt werden.

Rahmen und Ausgestaltung

- Die Teilnehmenden waren sich einig darüber, dass eine Ausschreibung ein möglichst **niedrigschwelliges Angebot** machen muss. Dementsprechend schwiebte den Teilnehmenden der Gruppe „Gesundheitsdaten“ ein Ideenwettbewerb vor, während die Teilnehmenden der

Gruppe „Verkehrs- und Wirtschaftsdaten“ ein **zweiphasiges Modell** mit Evaluation vorgeschlagen haben:

- **Explore-Phase:** Für einfache Skizzen/Konzepte mit geringem Budget zur Anschubförderung.
- **Exploit-Phase:** Für vollwertige Forschungsanträge, die auf erfolgreichen Explore-Projekten aufbauen.
- Wichtig erscheint insbesondere, den Zeitraum zwischen der Entwicklung einer Projektidee und einem Projektbeginn möglichst klein zu halten, um der Agilität gerade von Projekten unter Nutzung aktueller Daten Rechnung zu tragen.
- Als besonders wichtig wird die **aktive Förderung der Vernetzung, Vertrauensbildung und Begleitung** durch DFG und den Stifterverband angesehen. Dies könnte durch Matchmaking-Veranstaltungen vor der Explore-Phase (inkl. DFG-Workshops für unerfahrene Antragsteller*innen), Vernetzungsworkshops nach der Skizzeneinreichung geschehen.