

## DISKUSSIONSPAPIER

# Forschungsdaten

Der Begriff Forschungsdaten wird in der aktuellen öffentlichen und politischen Diskussion nicht einheitlich verwendet. Bei dem vorliegenden Papier handelt sich um erste generelle Überlegungen, die einen Beitrag zur Diskussion und Klärung weitergehender Fragen leisten sollen. Diese stellen keine abschließende Positionierung des VCI dar. Eine VCI-Kurzposition zum Forschungsdatengesetz findet sich hier ([VCI-Kurzposition Forschungsdatengesetz](#)).

## Was sind Forschungsdaten? (Definition der DFG)

Zu Forschungsdaten zählen, nach der Definition der Deutschen Forschungsgemeinschaft, u.a. Messdaten, Laborwerte, audiovisuelle Informationen, Texte, Survey-Daten, Objekte aus Sammlungen oder Proben, die in der wissenschaftlichen Arbeit entstehen, entwickelt oder ausgewertet werden. Methodische Testverfahren, wie Fragebögen, Software und Simulationen können ebenfalls zentrale Ergebnisse wissenschaftlicher Forschung darstellen und sollten daher ebenfalls unter den Begriff Forschungsdaten gefasst werden.<sup>1</sup>

Diese Daten liegen derzeit oft analog (Laborbücher etc.) aber auch vermehrt digital vor. Unter digitalen Forschungsdaten verstehen Kindling und Schirmbacher alle digital vorliegenden Daten, die während des Forschungsprozesses entstehen oder ihr Ergebnis sind.<sup>2</sup> Diese Daten umfassen somit nach unserem Verständnis sowohl unstrukturierte (u.a. in Dokumenten, Grafiken), semistrukturierte (u.a. XML und JSON) und strukturierte (u.a. Datenbanken) Datensätze. Zudem ist je nach Aggregationsgrad die Notwendigkeit von Metadaten zur „sinnvollen“ Verwertung der jeweiligen Datensätze (u.a. bei Rohdaten und „derived“-Data) zu beachten.

Für alle weiteren Betrachtungen im Rahmen dieses Diskussionspapiers werden wir diese bestehende Definition zu Grunde legen.

## Was ist ein Forschungsvorhaben?

Generell verstehen wir unter einem Forschungsvorhaben ein befristetes Projekt, dessen Ziel es ist, neue Erkenntnisse auf einem wichtigen oder besonders aktuellen Forschungsfeld zu erlangen. Das Vorhaben umfasst dabei alle Tätigkeiten der FuE-Abteilungen eines Unternehmens bzw. einer wissenschaftlichen Forschungseinrichtung zur Bearbeitung und Durchführung des Projekts. Hierbei können unterschiedliche Formen der Zusammenarbeit einzelner Projektpartner angewendet werden. Diese sind: unternehmenseigene FuE-Leistungen, Auftragsforschung/Forschungsdienstleistung und Kooperation.

---

<sup>1</sup> DFG:

[https://www.dfg.de/download/pdf/foerderung/grundlagen\\_dfg\\_foerderung/forschungsdaten/leitlinien\\_forschungsdaten.pdf](https://www.dfg.de/download/pdf/foerderung/grundlagen_dfg_foerderung/forschungsdaten/leitlinien_forschungsdaten.pdf)

<sup>2</sup> Kindling, Maxi und Schirmbacher, Peter: „Die digitale Forschungswelt“ als Gegenstand der Forschung. Information – Wissenschaft – Praxis 64 (2013): S. 130. [doi.org/10.1515/iwp-2013-0017](https://doi.org/10.1515/iwp-2013-0017)

Der Unionsrahmen für staatliche Beihilfen für Forschung, Entwicklung und Innovation (2014/C 198/01) legt für die letzten beiden Formen folgende Definitionen zu Grunde:

- ◆ **Auftragsforschung oder Forschungsdienstleistung** einer Forschungseinrichtung oder einer Forschungsinfrastruktur: Bei der Auftragsforschung legt das Unternehmen in der Regel die Vertragsbedingungen fest, ist Eigentümer der Ergebnisse der Forschungstätigkeit und trägt das Risiko des Scheiterns.
- ◆ **Zusammenarbeit** (Forschungskooperation) einer Forschungseinrichtung oder einer Forschungsinfrastruktur mit Unternehmen: Bei der Forschungskooperation verfolgen die Partner arbeitsteilig ein gemeinsames Ziel, legen gemeinsam den Gegenstand des Vorhabens fest, wirken an seiner Gestaltung mit, tragen zu seiner Durchführung bei und teilen die mit ihm verbundenen finanziellen, technischen, wissenschaftlichen und sonstigen Risiken sowie die erzielten Ergebnisse.

Je nach Form der Zusammenarbeit ergeben sich in Hinblick auf Data-Ownership, das Teilen von Daten, IP- sowie CBI-Fragen unterschiedliche Anforderungen.

Hinzu kommen Fragen der Abgrenzung zur „Produktentwicklung“. Forschung und damit auch verwendete Daten können auch direkt in ein Produkt einfließen, d.h. es handelt sich bei einem Forschungsvorhaben nicht ausschließlich um sogenannte Grundlagenforschung.

## Welche konkreten Anforderungen bestehen an die jeweiligen Daten?

Die Anwendung der **FAIR-Data Prinzipien**<sup>3</sup> ist eine Grundvoraussetzung für das Teilen von Daten und explizit von Forschungsdaten. Eine vollumfängliche Umsetzung dieser Prinzipien bedarf nach Möglichkeit u.a. globaler Standards von Dateiformaten sowie eine globale Governance von Infrastrukturen. Ergänzend müssen Aspekte des Datenschutzes und des Geschäftsgeheimnisschutzes sichergestellt werden. Letzteres betrifft vor allem den Aufbau der für die Datenteilung vorgesehenen nationalen und europäischen Infrastrukturen (z.B. NFDI, EOSC).

Privatwirtschaftliche Akteure, die ebenfalls in hohem Maße wissenschaftliche Daten erzeugen und nutzen werden in der aktuellen öffentlichen und politischen Diskussion primär als Datenlieferanten gesehen. Ein Interessensausgleich muss im Vordergrund stehen.

---

<sup>3</sup> findable, accessible, interoperable, reusable – siehe auch [https://www.forschungsdaten.org/index.php/FAIR\\_data\\_principles](https://www.forschungsdaten.org/index.php/FAIR_data_principles)

## **IP und Geschäftsgeheimnisse**

Die deutsche pharmazeutische und chemischen Industrie gehört zu den innovationsintensivsten Branchen in Deutschland. Sie investierte im Jahr 2022 rund 14 Milliarden in Forschung und Entwicklung (FuE), die wesentlich datenbasiert durchgeführt werden. Diese Anstrengungen sichern die Innovationsfähigkeit der Branche am hiesigen Standort. Gleichzeitig steht die Branche im internationalen Wettbewerb. Daher ist ein adäquater Schutz von IP und Geschäftsgeheimnissen zentral, um das Datenteil und datenbasierte Forschung & Entwicklung zu incentivieren.

Für kleine und mittlere Unternehmen (KMU) könnten nicht verpflichtende Standardverträge (z.B. zum Schutz des geistigen Eigentums sowie von Geschäfts- und Betriebsgeheimnissen oder für Kooperationen zwischen Wissenschaft und Wirtschaft) hilfreich sein. Die Flexibilität und die Vertragsfreiheit dürfen nicht verloren gehen.

## **Cybersicherheit**

Eine zunehmende Digitalisierung und Vernetzung innerhalb der (Daten-)Wertschöpfungskette führt zwangsläufig auch zu höheren Anforderungen im Bereich der Cybersicherheit. Diese Aspekte sind bei der Konzeption von Datenteilungsinfrastrukturen, wie der u.a. der NFDI, zu berücksichtigen. Darüber hinaus müssen Berechtigungskonzepte für den Zugriff auf Daten innerhalb dieser Infrastrukturen ausgearbeitet werden, damit sichergestellt ist, dass nur autorisierte Nutzer Zugriff erhalten. Hierbei ergeben sich weitere Fragen des weiteren Datenzugriffs. Unter anderem ist hierbei der Wert von aggregierten Daten<sup>45</sup> u.a. in Hinblick auf das Wettbewerbsrecht zu beachten. Clearingstellen (ggf. Datentreuhänder) als neutrale Player könnten entsprechende Aufgaben übernehmen.

## **Datenschutz**

Gerade im Bereich der Gesundheitsforschung stellen sich oft Fragen des Schutzes von personenbezogenen Daten. In der Europäischen Union gilt die Datenschutzgrundverordnung, die die Nutzung von personenbezogenen Daten einschränkt. Eine rechtssichere Anonymisierung bzw. Pseudonymisierung wie auch Opt-In Lösungen zur Nutzung bestimmter Daten könnten Möglichkeiten darstellen.

Zu beachten ist weiterhin, dass aus sogenannten Gesundheitsdaten im Laufe des Forschungsprozesses möglicherweise Forschungsdaten werden. In der Folge sind diese Daten möglicherweise von mehreren Gesetzen betroffen (z.B. Gesundheitsdatennutzungsgesetz und das geplante Forschungsdatengesetz). Eine entsprechende enge Verzahnung der Gesetzesinitiativen ist daher notwendig. Gleichzeitig gilt aber auch, dass nicht automatisch jedes Gesundheitsdatum ein Forschungsdatum ist.

---

<sup>4</sup> Kombination einzelner, zunächst unabhängiger Datensätze und/oder große Mengen an zusammenhängenden Daten.

<sup>5</sup> Weiter stellen sich Fragen der Wertigkeit von sogenannten Rohdaten („raw data“) und abgeleiteten / verarbeiteten Daten („derived data“).

Bei den gemachten Aussagen handelt es sich um generelle Bedürfnisse, die alle Datensätze in gleicher Form betreffen und für alle Formen der Zusammenarbeit in gleicher Weise gelten.

## Was würde die Bereitschaft zum Datenteilen auf Seiten von Unternehmen bzw. der Wirtschaft erhöhen?

Die digitale und zirkuläre Transformation in der chemisch-pharmazeutischen Industrie schreitet rasch voran. Bereits jetzt nutzt die Industrie die umfangreichen Daten, die im Zuge von Prozessen oder im Kontext der Forschung & Entwicklung anfallen. Perspektivisch ist davon auszugehen, dass einerseits die Menge an Daten und andererseits die Datennutzung auf Grund der Verfügbarkeit von Technologien stark zunehmen wird (z.B. Künstliche Intelligenz). Das Vorantreiben einer Datenökonomie ist daher von zentraler Bedeutung.

Um eine Datenökonomie zu incentivieren, muss der Austausch mit Daten so einfach wie möglich und rechtssicher gestaltet werden. Hierzu gehören auch Fragen einer europaweiten und – im Idealfall internationalen – Harmonisierung.

Die Erfahrung zeigt, dass bereits in dezentralen Organisationen der Austausch von Daten nicht immer optimal funktioniert. Dass Daten häufig ungenutzt bleiben, liegt jedoch nicht ausschließlich daran, dass Unternehmen diese Daten nicht nutzen wollen. Oft ist die geringe Datennutzung auch in der enorm aufwändigen Datenerschließung begründet. Es ist daher wichtig, zunächst Strukturvoraussetzungen (u.a. Qualitätssicherung von Daten) für einen Datenaustausch zu schaffen.

### Christian Bünger

Bereich Wirtschaft, Finanzen, Digitalisierung  
Abteilung Volkswirtschaft  
T +49 (69) 2556-1715 | E [christian.buenger@vci.de](mailto:christian.buenger@vci.de)

### Dr. Denise Schütz-Kurz

Bereich Wissenschaft, Technik und Umwelt  
Abteilung Wissenschaft und Forschung  
T +49 (69) 2556-1482 | E [schuetz@vci.de](mailto:schuetz@vci.de)

### Verband der Chemischen Industrie e.V. – VCI

Mainzer Landstraße 55  
60329 Frankfurt

[www.vci.de](http://www.vci.de) | [www.ihre-chemie.de](http://www.ihre-chemie.de) | [www.chemiehoch3.de](http://www.chemiehoch3.de)

[LinkedIn](#) | [X](#) | [YouTube](#) | [Facebook](#)

[Datenschutzhinweis](#) | [Compliance-Leitfaden](#) | [Transparenz](#)

- Registernummer des EU-Transparenzregisters: 15423437054-40
- Der VCI ist unter der Registernummer R000476 im Lobbyregister, für die Interessenvertretung gegenüber dem Deutschen Bundestag und gegenüber der Bundesregierung, registriert.

*Der VCI und seine Fachverbände vertreten die Interessen von rund 1.900 Unternehmen aus der chemisch-pharmazeutischen Industrie und chemienaher Wirtschaftszweige gegenüber Politik, Behörden, anderen Bereichen der Wirtschaft, der Wissenschaft und den Medien. 2022 setzten die Mitgliedsunternehmen des VCI rund 260 Milliarden Euro um und beschäftigten knapp 550.000 Mitarbeiterinnen und Mitarbeiter.*