

VCI-Beitrag zur BMBF-Konsultation zum Forschungsdatengesetz

VCI-Antworten in blau.

1. Bedarfe

Bitte erläutern Sie, welche Bedarfe an Daten im Kontext kommerzieller und nichtkommerzieller Forschung bestehen und gehen Sie hierbei auf folgende Aspekte ein:

a. Daten für die Forschung:

Zu welchen Daten besteht aus Ihrer Sicht ein Bedarf an besseren Zugangs-, Nutzungs- und Verknüpfungsmöglichkeiten z.B. lange Datenreihen, geographisch referenzierte Daten, gelabelte Daten, (nicht) personenbezogene Daten, kuratierte Daten, Rohdaten ohne Kontext, Spartendaten wie Klima, Wetter?

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). Dies betrifft unter anderem – aber nicht abschließend – folgende Bereiche:

- Logistik: Optimierung der Supply-Chain, Nutzung von Daten für Smart Contracts
- Produktion: digital vernetzte Produktionsprozesse (sowohl intern wie auch extern mit Partnern)
- Wartung: Predictive Maintenance
- Forschung und Entwicklung:
 - Einsatz von automatisierten Laborgeräten
 - Modellierung für die Versuchsplanung
 - Simulationsmethoden für das Design von Molekülen, Materialien sowie für die Vorhersage von Prozess- und Produkteigenschaften (on premise sowie in der Cloud)
 - Nutzung von Daten zur Abschätzung der Sicherheit chemischer Substanzen und somit Vermeidung / Reduktion von Tierversuchen
- Entwicklung neuer datengetriebener Geschäftsmodelle, z.B. Digital bzw. Smart Farming
- Kopplung interner Daten mit Daten von Kunden für z.B. digitale Produktpässe
- Gesundheitsdaten im Bereich der Pharmaforschung

Diese Liste zeigt die Fülle der in der chemisch-pharmazeutischen Industrie genutzten Daten. Eine abschließende Bewertung ist zum heutigen Stand auf Grund der Komplexität nicht möglich. Eine „one-size-fits-all“-Lösung wird es aller Voraussicht nach nicht geben.

Zwingend notwendig ist aus unserer Sicht zunächst eine Definition des Begriffs „Forschungsdaten“. Derzeit verstehen verschiedene Stakeholder unterschiedliche Dinge unter diesem Begriff.

Ebenso müssen „Forschungsdaten“ von „Gesundheitsdaten“ differenziert werden.

In welchen wissenschaftlichen Disziplinen bzw. interdisziplinären Kontexten für welche Forschungszwecke sind Datenzugänge besonders relevant (z.B. empirische Wirtschaftsforschung, Bildungsforschung, medizinische Forschung, interdisziplinäre Forschung)?

In der **Chemieforschung** sind u.a. in folgenden Bereichen Datenzugänge besonders relevant:

- Materialforschung (z.B. in der Katalysator- und Batterieforschung) sowie in der Synthesepaltung, Formulierung (molekulare und prädiktive Modellierung) oder Analytik (Auswertung von Messergebnissen). Hier sind insbesondere Meta-Daten der Analytik, wie beispielsweise Mess- und Auswerteparameter oder Standards zu betrachten.
- Wirkstoffforschung z.B. für Pflanzenschutzprodukte oder bei der Analyse und Vorhersage neuer biologischer Eigenschaften von Nutzpflanzen und bei biotechnologisch hergestellten Produkten
- Nutzung von Künstlicher Intelligenz z.B. zur Erkennung von Anomalien im Produktionsprozess
- Ein Datenaustausch entlang von Wertschöpfungsketten ist von hohem Interesse, um Entwicklungsprozesse zu beschleunigen und weitere Optimierungspotenziale zu heben.

Auch im Bereich der **Pharmaforschung** finden große Datenmengen (z.B. Real World Data oder klinische Studien) bereits heute Anwendung.

Die Relevanz von Daten und Anreizen, Daten zu teilen, ist je nach Industriebranche unterschiedlich ausgeprägt. Auch die Praxis des Datenteilens ist unterschiedlich weit entwickelt. Ebenso ist die Art der Daten unterschiedlich.

Diese Unterschiede und die Komplexität sind bei der Konzeption eines Forschungsdatengesetzes zwingend zu berücksichtigen: Beispielsweise hat die Gesundheitsindustrie einen großen Bedarf an Patientendaten. Die Chemieindustrie hat hingegen ein großes Interesse an technischen Daten aus Wertschöpfungsketten sowie Daten für die Nachhaltigkeitsbewertung.

Welche Herkunft/Quellen von Daten sind hier aus Ihrer Sicht besonders relevant (z.B. Daten aus der öffentlichen Verwaltung, Daten aus der Wirtschaft bzw. von Unternehmen, Forschungseinrichtungen oder Einzelpersonen, Maschinendaten, Daten der amtlichen Statistik, etc.)?

Von hoher Relevanz sind wissenschaftliche Daten, die nach „Open Science“-Prinzipien publiziert werden. Auch Daten von fehlgeschlagenen Experimenten können einen Mehrwert liefern. Hier sollte über Anreizsysteme nachgedacht werden, die das Publizieren dieser Art von Daten fördern.

Auch andere, nicht wissenschaftliche Publikationen enthalten Daten, die digitalisiert und geteilt werden könnten und so einen Mehrwert bieten.

Die amtliche Statistik bietet ebenfalls einen hohen Mehrwert. Wichtig ist hier allerdings, dass die Daten zeitnah vorliegen. Daten der öffentlichen Hand sollten kostenfrei zur Verfügung gestellt werden (z.B. Umwelt- und Klimadaten).

Im Bereich der Pharmaforschung sind u.a. folgende Daten von Interesse

- ePA-Daten
- Abrechnungsdaten der GKV und der PKV
- Registerdaten
- Gesundheitsdaten im Sinne der DSGVO
- Daten aus bildgebenden Verfahren

Damit ein breites Datenteilen allerdings auch Realität werden kann, müssen zunächst die notwendigen Strukturen in Form von Datenräumen und technischen Standards aufgebaut werden. Hierzu gehören auch Fragen des automatisierten Austauschs via API.

b. Daten aus der Forschung:

Zu welchen Daten aus der Forschung besteht aus Ihrer Sicht Bedarf an Zugang-, Nutzungs- und Verknüpfungsmöglichkeiten für öffentliche Stellen, Unternehmen und Zivilgesellschaft?

Bevor weitere Zwänge zum Datenteilen auferlegt werden, sollten zunächst Daten aufbereitet werden, die schon heute öffentlich zur Verfügung stehen. Wissenschaftliche Publikationen werden bereitwillig geteilt und enthalten intrinsisch einen großen – bisher in der Regel nicht gehobenen – Schatz an Daten.

Hier besteht noch Verbesserungspotenzial: Zu überlegen wäre, die Daten, die zu wissenschaftlichen Publikationen gehören, als Teil der Veröffentlichung in standardisierter Form zugänglich zu machen (ontologieunterstützte Vernetzung).

Der Austausch von Daten aus der Forschung sollte gleichberechtigt und dem Zweck und Ziel eines Forschungsvorhabens angemessen geregelt werden.

Zu welchen Zwecken sind Datenzugänge für Unternehmen, Zivilgesellschaft oder öffentlichen Stellen besonders relevant?

In Forschungsk Kooperationen ist das Teilen von Daten von hoher Relevanz. Hier können die Kooperationspartner in einem Forschungsprojekt immer gegenseitig vereinbaren, wie Daten geteilt werden. Entwicklungen in Richtung „Open Science“ sind vielversprechend und sollten vorangetrieben werden. Idealerweise ist ein Standard so definiert, dass er auch von Dritten ohne Offenlegung der Daten genutzt werden kann. So ergibt sich ein einheitlicher Datenraum für die Partner und die Möglichkeit für spätere Bereitstellung.

In nationalen und internationalen Notsituationen (z.B. Pandemien) kann ein Datenzugang ebenfalls sehr relevant werden. Möglicherweise kann auch ein Zwang angemessen sein. Notsituationen müssen aber sehr klar definiert werden. Außerdem müssen sie zeitlich begrenzt werden. Ebenso sollte deutlich werden, um welche Art von Daten es sich handelt. Patentdaten dürfen aus unserer Sicht nicht dazu gehören. Dies würde den Forschungsstandort Deutschland (bzw. Europa) für die Generierung von IP deutlich benachteiligen.

Daten, die öffentlich zur Verfügung stehen und einen potenziellen Nutzen versprechen, sollten in digitaler Form und über standardisierte Schnittstellen bereitgestellt werden. Das Prinzip „as open as possible, as closed as necessary“ sollte dabei die Grundlage sein.

Von einer generellen gesetzlichen Verpflichtung zum Datenteilen sollte derzeit abgesehen werden.

Noch wichtiger ist allerdings die Schaffung von Anreizen zum freiwilligen Datenteilen, um die Potenziale von nahtlosen Datenzugängen z.B. entlang von Wertschöpfungsketten zu heben.

Welche Herkunft/Quellen von Forschungsdaten sind hier aus Ihrer Sicht besonders relevant?

Jede wissenschaftliche Publikation ist eine potenzielle Datenquelle. Oft gibt es hier aber Probleme mit der strukturierten Bereitstellung von Daten (siehe oben).

Ein Forschungsdatengesetz sollte so ausgestaltet werden, dass flexibel und offen auf weitere, durch die digitale und zirkuläre Transformation entstehende Datenquellen im Forschungsumfeld sowohl der Wirtschaft als auch der Wissenschaft reagiert werden kann. Ein ausreichender Schutz von Geschäftsgeheimnissen (IP) muss auch bei Erweiterungen des Scopes sichergestellt werden.

c. Was würde die Bereitschaft zum Datenteilen auf Seiten von

- *Unternehmen bzw. der Wirtschaft*
- *Wissenschaft*
- *Zivilgesellschaft*
- *öffentlichen Stellen*

erhöhen? Welche Anreize könnten gesetzt werden?

Bei wissenschaftlichen Publikationen besteht heute schon eine große Bereitschaft zum Teilen von Ergebnissen. Die Möglichkeit, Datenquellen zu zitieren, wäre eine Maßnahme, die Herkunft der Daten wertzuschätzen und damit die Bereitschaft zum Teilen von Daten in Wirtschaft und Wissenschaft zu erhöhen. Es besteht die Chance, einen weiteren Schritt in Richtung einer neuen Datenkultur in der Wissenschaft zu tun.

Die Bereitschaft zum Datenteilen wird steigen, wenn sowohl Unternehmen wie auch Akademia positive Erfahrungen gemacht haben, die zu einer win-win Situation geführt haben. Für das Sammeln dieser Erfahrung wäre es sinnvoll, öffentlich geförderte Programme oder Kooperationen (PPP) aufzusetzen, die Experimentierräume schaffen, in denen die Beteiligten in einem geschützten Raum Erfahrungen sammeln können und ein gemeinsames Lernen erfolgen kann.

Kommunikation von best-practice Beispielen ist aus unserer Sicht ein weiterer Weg, eine neue Datenkultur zu fördern. Hierbei könnte das Dateninstitut eine Rolle spielen.

Entscheidend wird sein, dass Unternehmen das Vertrauen in die Prozesse gewinnen und die Rechtssicherheit im europäischen und internationalen Rechtsraum bekommen, dass durch das Datenteilen IP-Rechte (Intellectual Property) nicht eingeschränkt werden.

d. Wie schätzen sie die Bedeutung von Forschungsk Kooperationen (z.B. mit Partnern aus der Wissenschaft und der Wirtschaft) in Bezug auf das sektorübergreifende Teilen von Daten ein?

Forschungsk Kooperation sind bereits heute ein wichtiges Element und finden in der Chemie- und Pharmaindustrie breite Anwendung. Forschungsk Kooperationen bieten die Möglichkeit, in einem geschützten Raum positive Erfahrungen zu sammeln und funktionierende Datenteilungs-, Kooperations- und Geschäftsmodelle mit Modellcharakter zu entwickeln.

Von besonderer Bedeutung für die Chemieindustrie wird vor allem das Teilen von Daten über Sektor-Grenzen hinweg sein. Die Chemie steht am Anfang fast jeder industriellen Wertschöpfungskette. Datenteilen über Branchen hinweg kann Innovationen deutlich beschleunigen. Auch im Zuge der zirkulären Transformation werden diese sektorübergreifenden Daten noch wichtiger werden.

e. Welche Rolle und welche Aufgaben sollten Datenmittlerstrukturen beim Datenteilen einnehmen?

Ein Datenintermediär kann den Austausch der Daten unter den Beteiligten fördern. In der EU-Digitalgesetzgebung spielen sie eine wichtige Rolle. Die konkrete Ausgestaltung ist allerdings noch offen. Die praktische Umsetzung wird zeigen, wie sich die Kultur des Datenteilens entwickelt und wie gut die Rollen ausgeübt werden.

Datenintermediäre sollten klare rechtliche Rahmenbedingungen schaffen. Eine wichtige Frage ist u.a., wie Inhalte verwendet werden dürfen. International gibt es hier bereits sehr gute und leicht umsetzbare Regeln, die etwa über die non-Profit Organisation Creative Commons festgelegt werden.

Datentreuhänder könnten als Betreiber von Plattformen für den Datenaustausch (auf der Basis von verbindlichen und einfachen Regeln) in Frage kommen. Die Datensouveränität des

Dateninhabers muss dabei gewährleistet werden, z.B. indem die genutzten Daten nur in einer sicheren Verarbeitungsumgebung zur Verfügung gestellt werden.

2. Hindernisse

Bitte erläutern Sie, welche Hindernisse beim Zugang und im Umgang (Verknüpfung, Nutzung etc.) mit Daten für und aus der Forschung derzeit bestehen. Bitte gliedern Sie der Übersichtlichkeit halber gern in z.B. rechtliche, technische, organisatorisch Hindernisse und erläutern Sie (wenn möglich) jeweils anhand von Beispielen

Es besteht die Gefahr, dass durch ein zu breites und verpflichtendes Teilen von nicht öffentlichen Daten die nationale und internationale Wettbewerbsposition der Chemie- und Pharmabranche geschwächt wird.

Um eine Datenökonomie zu incentivieren, muss der Austausch mit Daten so einfach wie möglich gestaltet werden. So fallen derzeit z.B. Prüfungen an, ob personenbezogene Daten enthalten sind oder ob länderspezifische Anforderungen erfüllt sind. Um den Aufwand so gering wie möglich und die Hürde so niedrig wie möglich zu gestalten, ist zumindest eine europaweite Harmonisierung anzustreben.

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. Kuratieren von Daten) für einen Datenaustausch zu schaffen.

Für kleine und mittlere Unternehmen (KMU) könnten Standardverträge (z.B. zum Schutz des geistigen Eigentums sowie von Geschäfts- und Betriebsgeheimnissen oder für Kooperationen zwischen Wissenschaft und Wirtschaft) hilfreich sein. Derzeit gibt es diese nicht. Hier wäre eine Entwicklung hilfreich (z.B. durch das Dateninstitut).

3. Lösungsvorschläge

Mit Blick auf die beschriebenen Bedarfe und Hindernisse, welche legislativen Regelungen sollten aus Ihrer Sicht für ein Forschungsdatengesetz geprüft werden (konkrete Formulierungsvorschläge erwünscht)?

- a. *(zu den nicht-legislativen Lösungen s. 3 c)*
- b. *In welchen Gesetzen (deutschen/ggf. auch europäischen) besteht aus Ihrer Sicht Änderungs-/Ergänzungsbedarf mit Blick auf die oben beschriebenen Bedarfe und Hindernisse (Formulierungsvorschläge erwünscht)?*
- c. *Welche weiteren, nicht-legislativen Lösungsvorschläge sehen Sie mit Blick auf die beschriebenen Bedarfe und Hindernisse?*

Aus Sicht der deutschen chemisch-pharmazeutischen Industrie sollte zum jetzigen Zeitpunkt kein allgemeiner Zwang zum Teilen von Daten in einem Forschungsdatengesetz geschaffen werden. Vielmehr sollte auf Anreize gesetzt werden.

Ebenso sind aus unserer Sicht die Regulierungsbedarfe noch nicht ausreichend identifiziert. Es besteht die Gefahr mit einem breiten und horizontalen Ansatz über das eigentliche Ziel einer gut funktionierenden Datenökonomie hinauszuschießen. Vielmehr sehen wir Experimentierräume und Pilotprojekte als guten Ansatzpunkt, um regulatorischen Handlungsbedarf zu ermitteln. In diesen Projekten sollten dann auch praxisrelevante Themen wie Datenbankstrukturen, Zugriffsrechte, Datenstandards etc. adressiert werden.

Das neu gegründete Dateninstitut sollte hier eine wesentliche Rolle spielen und Pilotprojekte initiieren, die Durchführung begleiten und am Ende den Regulierungsbedarf identifizieren (siehe hierzu auch unsere Eingabe zu möglichen Use-Cases des Dateninstitutes).

- „Forschungsdaten“ sind ein komplexes Thema für das es keine „one-size-fits-all“-Lösung geben wird. Dennoch sind wir überzeugt, dass es Grundprinzipien (Referenzmodell) gibt, die für alle Anwendungsfälle sinnvoll sein könnten (z.B. Cool-Off-Periods, Fragen der Kennzeichnung von Geschäftsgeheimnissen, Standards zur Kategorisierung von Geschäftsgeheimnissen etc.).
- Das Dateninstitut wäre ein neutraler Player, der alle betroffenen Stakeholder zusammenbringen könnte, um so ein Referenzmodell zu entwickeln. Dabei sollte neben den verschiedenen Industriebranchen (möglicherweise gibt es hier weiteren Differenzierungsbedarf) auch die NFDI eingebunden werden. Das Dateninstitut wäre Moderator und bringt zeitgleich Expertise in die Diskussion ein. Auch würde so verhindert, dass parallel eine Vielzahl von Einzellösungen entwickelt werden, die am Ende möglicherweise nicht kompatibel sind.

4. Sonstiges

Was möchten Sie uns sonst noch zum Thema Forschungsdatengesetz/Forschungsklauseln mitgeben?

Da die Arbeit der NFDI mit öffentlichen Mitteln finanziert wurde, sollte eine stärkere Einbindung der Wirtschaft in die NFDI diskutiert werden. Der bisherige Dialog mit der Wirtschaft ist aus unserer Sicht noch nicht optimal. Die neu gegründete Industriesektion ist ein Schritt in die richtige Richtung, eine direkte Beteiligung der Wirtschaft in den Konsortien sollte allerdings ebenfalls in Erwägung gezogen werden.

Die Bundesregierung sollte sich für ein globales „open data level-playing field“ einsetzen. „Open Data“ kann in einem globalen System nur dann funktionieren, wenn in anderen Teilen der Welt Daten gleichermaßen geteilt werden. Ansonsten führt „open data“ zu einem einseitigen Wissensabfluss aus Deutschland und schwächt die Innovationskraft des Standorts.

Angesichts der geopolitischen Situation und des Strebens einiger Länder nach Technologieüberlegenheit kann nicht damit gerechnet werden, dass innovationsstarke „not like minded countries“ gleichermaßen bereit sind, ihre Daten zu teilen. Zum Schutz des deutschen Innovationssystems sind daher geeignete Strategien zu entwickeln. Das Grundprinzip sollte hier ebenfalls „as open as possible, as closed as necessary“ lauten.

Eine Harmonisierung der EU-Digitalgesetzgebung ist dringend geboten. Hierzu gehören auch nationale Forschungsdatengesetze.

Kontakt

Christian Bünger

Bereich Wirtschaft, Finanzen, Digitalisierung
Abteilung Volkswirtschaft
T +49 69 2556-1715 | E 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

Verband der Chemischen Industrie e.V. – VCI

Mainzer Landstraße 55
60329 Frankfurt

www.vci.de | www.ihre-chemie.de | www.chemiehoch3.de

[LinkedIn](#) | [Twitter](#) | [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 Verband der Chemischen Industrie (VCI) vertritt 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. 2021 setzten die Mitgliedsunternehmen des VCI rund 220 Milliarden Euro um und beschäftigten mehr als 530.000 Mitarbeiterinnen und Mitarbeiter.