

# Stellungnahme zum geplanten Forschungsdatengesetz

Anna Akimova, Lidya Benli, Johanna Fick, Florian Hoedt, Darius Kühler, Katharina  
Prothmann, Vivien Steinel, Britta Timmermann, Lena-Katharina Wrede

Braunschweig April 2023



Anna Akimova, Lidya Benli, Johanna Fick, Florian Hoedt, Darius Kühler, Katharina Prothmann, Vivien Steinel,  
Britta Timmermann, Lena-Katharina Wrede

Ansprechpartner: Florian Hoedt, [florian.hoedt@thuenen.de](mailto:florian.hoedt@thuenen.de)

Braunschweig / Germany April 2023

## Inhalt

<b>Interessensgruppe</b>	<b>2</b>
<b>1 Bedarfe</b>	<b>2</b>
<b>2 Hindernisse</b>	<b>3</b>
2.1 Rechtliche / Organisatorische Hindernisse	3
2.2 Technische Hindernisse	6
<b>3 Lösungsmöglichkeiten</b>	<b>6</b>
3.1 Zu prüfende legislative Regelungen	7
3.2 Nicht-legislative Lösungsvorschläge	7
<b>4 Sonstiges</b>	<b>9</b>

## Interessensgruppe

- Öffentliche Stelle
- Wissenschaft

## 1 Bedarfe

Das Thünen-Institut befasst sich mit komplexen in der Regel interdisziplinären Forschungsfragen aus den Bereichen „ländliche Räume“, „Agrar“, „Wald“ und „Fisch“, wobei in den drei zu Letzt genannten Bereichen jeweils die Forschungsdisziplinen Sozioökonomie, Ökologie und Technologie abgedeckt werden. Für die Beantwortung derartig komplexer Fragestellungen sind die Wissenschaftlerinnen und Wissenschaftler des Thünen-Instituts auf eine umfangreiche und verlässliche Datengrundlage angewiesen. Diese Stellungnahme stellt sowohl grundsätzliche Bedarfe, Hindernisse und Lösungsmöglichkeiten dar, ist zur Veranschaulichung jedoch auch mit Beispielen aus den einzelnen Disziplinen ergänzt.

### Datenzugänge und -quellen

Für das Thünen-Institut sind Datenzugänge mit Bezug zu folgenden wissenschaftlichen Disziplinen bzw. interdisziplinären Forschungskontexten, in quantitativer und qualitativer Form, hervorzuheben:

- agrar(ökonomisch)-ökologische-Forschung und längerfristige Monitoringvorhaben,
- ex-ante und ex-post Evaluierungsvorhaben von politischen Interventionen,
- wirtschafts- und humangeografische sowie soziologische Forschung und längerfristige Monitoringvorhaben sowie Panelstudien,
- naturwissenschaftliche und lebenswissenschaftliche, angewandte wie nicht angewandte Forschung,
- ingenieurwissenschaftliche Forschung und Informatik, hier besonders der Einsatz von künstlicher Intelligenz und Automatisierung durch Robotereinsatz.

Wichtige Datenquellen sind anderen Forschungseinrichtungen (national) und Daten aus Monitoringprogrammen des Bundes und Länder (in Meeresforschung BSH, BAW, NLVP). Ebenso werden

- allgemeine Verwaltungsdaten (z. B. Meldeämter, Finanzbehörden),
- spezifische Verwaltungsdaten der Landwirtschafts- und Umweltverwaltung,
- spezifische Daten der Statistischen Landesämter, des Statistischen Bundesamtes und weiterer Bundesbehörden (z. B. Umweltbundesamt, Bundesamt für Naturschutz, Bundesamt für Migration und Flüchtlinge, Bundesagentur für Arbeit),
- spezifische Daten einzelner Branchen,
- georeferenzierte Standortdaten und
- georeferenzierte Daten von *Social Media*-Anbietern und Daten der klassischen Massenmedien

benötigt, um unsere Forschungsarbeiten durchzuführen.

### Bedeutung von Forschungsk Kooperationen für das sektorübergreifende Teilen von Daten

Forschungsk Kooperationen sind meist essentiell für das sektorübergreifende Arbeiten und den Austausch von Daten über Forschungseinrichtungen hinweg. Sie ermöglichen interdisziplinäres Arbeiten, erlauben ein tieferes Verständnis der beteiligten Partner und die Verwendung konsistenter Datengrundlagen. Zugleich steigern sie die Effizienz der Datenbereitstellung, -nutzung- und -haltung, da mehrfache Anfragen, Freigaben, Übermittlungen, Speicherung, Aufbereitung etc. weitgehend vermieden werden können und führen somit tendenziell zu einer höheren Arbeitseffizienz auf Seiten der Datenbereinsteller und -nutzer. Sie erfordern zusätzliche Ressourcen für

das Management dieser Kooperation inkl. einem professionellen Datenmanagement inkl. Rechtsberatung u.a. hinsichtlich Eigentums- und Rechtgestaltung.

*Beispiel: Arbeitsschwerpunkte im Fachinstitut für Lebensverhältnisse in ländlichen Räumen*

Schwerpunkte der Arbeit im Fachinstitut für Lebensverhältnisse in ländlichen Räumen sind

- die räumliche differenzierte Abschätzung und Bewertung von Umwelteffekten der landwirtschaftlichen Landnutzung auf Grund von Änderungen des Marktumfeldes, der eingesetzten Technologie sowie des förder- und ordnungspolitischen Rahmens. Ferner beschäftigen wir uns mit den sozio-ökonomischen Implikationen der aufgeführten Veränderungen,
- Evaluation von Förderprogrammen der ländlichen Entwicklung,
- die Betrachtung der Lebensverhältnisse von Personen in ländlichen Räumen mit ihren sozio-ökonomischen Verflechtungen (z. B. wirtschaftlicher, gesellschaftlicher und demografischer Wandel).

Hierzu benötigen wir deutschlandweit verknüpfbare lokalbezogene, geographisch möglichst präzise direkt oder indirekt georeferenzierte, kleinräumige Mikrodaten in langen Zeitreihen. Als wichtige Daten sind zu nennen:

- allgemeine Verwaltungsdaten (u.a. Daten aus dem Melderegister, Steuerinformationsdaten, Ausländerzentralregister, Daten aus den Liegenschaftskatastern),
- spezifische Verwaltungsdaten der Landwirtschafts- und Umweltverwaltung (u.a. Integrierten Verwaltungs- und Kontrollsystem (InVeKoS), dem Herkunftssicherungs- und Informationssystem für Tiere (HIT), Aufzeichnungspflichtigen Düngeverordnung, Zahlstellendaten, Förderdaten)
- spezifische Daten der Statistischen Landesämter und des Statistischen Bundesamt (u.a. Daten der Agrarstatistik z. B. Agrarstrukturerhebung, Ertragsstatistik, u. Ä. aus Gesamt- oder Stichprobenerhebungen, Zensus)
- spezifische Daten anderer Bundes- oder Facheinrichtungen (z. B. Testbetriebsnetzes Landwirtschaft, andere Testbetriebsnetze, Daten des Deutschen Biomasse Forschungszentrums (DBFZ))
- spezifische Daten einzelner Branchen (u. a. Kassenärztliche Vereinigungen, Energieversorgern, Breitbandversorgung, Industrie- und Handelskammern sowie Handwerkskammer, landwirtschaftliche Beratungsunternehmen, Hersteller von Landtechnik) sowie
- georeferenzierte Standortdaten (u.a. Openstreetmap, Satelliten-Daten).

Zwingend für die Verarbeitung ist die Bereitstellung in maschinenlesbaren Formaten. Gegenwärtig sind diese Datensätze oftmals nicht (ohne weiteres) landes- und systemübergreifend insbesondere auf der Mikroebene verknüpfbar. Gründe hierfür sind zum einen technische Restriktionen, da die jeweiligen Erfassungssysteme nicht für Verknüpfungszwecke ausgelegt sind, und zum anderen rechtliche Restriktionen, die die Verknüpfung von Mikrodaten aus den Verwaltungs- und statistischen Quellen untersagen. Eine Verknüpfbarkeit dieser Daten verbessert die Qualität der Forschungsvorhaben, der Umweltberichterstattung als auch der Evaluation von Förderprogrammen und der Politikberatung erheblich und kann in der Bearbeitung zu einer nennenswerten Vereinfachung der Prozesse führen. Letztlich könnten Ergebnisse schneller bereitgestellt werden und die Empfehlungen zielgenauer formuliert, so dass die angebotenen Interventionen der öffentlichen Hand zur Erreichung der Nachhaltigkeitsziele möglichst passgenau ausgestaltet werden können.

## 2 Hindernisse

### 2.1 Rechtliche / Organisatorische Hindernisse

Aus unserer Sicht besteht in den Bereichen der Datenbeschaffung und -bereitstellung Optimierungsbedarf. Das Thünen-Institut als Ressortforschungseinrichtung im Geschäftsbereich des BMEL forscht zu den Bereichen

Ländliche Räume, Wald und Fischerei (15 Fachinstitute) sowie mittels einer Stabstelle zu den Bereichen Klima und Boden. Die Forschung erfolgt dabei durch eine Vielzahl an verschiedenen (Drittmittel-)Projekten in den Forschungsgebieten der Einrichtung. Die nachfolgend aufgeführten Hindernisse erstrecken sich dabei über alle Forschungsgebiete des Thünen-Instituts, sodass eine Verallgemeinerung der Problematiken anzunehmen ist.

Die Problemfelder liegen dabei insbesondere bei den Themen Langzeitforschung bzw. Monitoringvorhaben, Datenaufbereitung, Datensicherung (im Sinne der guten wissenschaftlichen Praxis) sowie der Übermittlung von Daten, welche bspw. bei den Bundesländern oder Forschungsdatenzentren vorhanden sind und an den Bund weitergegeben werden sollen. Hierbei werden Daten bspw. auch aus Befragungen generiert und in Form von Fragebögen erhoben (Marktforschung und -analyse). D. h. es können personenbezogene Daten sowie Betriebsdaten betroffen sein. Hier sollte das zu verabschiedende neue Forschungsdatengesetz ansetzen und Lösungen für die Problemfelder finden. Ziel muss dabei sein, eine vereinfachte Datenbeschaffung und Datenübermittlung für die praktische Forschung zu bewirken.

Zu Beginn der Ausführungen möchten wir auf die derzeit ungeklärte Situation bei Begriffsbestimmungen im Kontext der Forschungsdaten hinweisen. Diese Lücke besteht maßgeblich bei den Forschungsdaten selbst, die bislang nicht definiert sind, sowie die Begriffe der Forschung zu kommerziellen bzw. nicht-kommerziellen Zwecken bislang nicht höchstrichterlich geklärt sind, sodass keine Legaldefinition existiert. Für die Rechtsanwendung ergeben sich dadurch erhebliche Auslegungsfragen. Hinzuweisen ist für derartige Abgrenzungsfragen bspw. auf § 60c UrhG. Eine praxisnahe Definition sollte das Ziel der Stärkung und Erweiterung der Nutzung der Forschungsdaten für nicht-kommerzielle Zwecke verfolgen.

Insbesondere für Geodaten sind die Nutzungszwecke für Forschung nicht eindeutig (genug) spezifiziert. Dies führt z. B. bei Länderverwaltungen zu unterschiedlichen Auslegungen und Anforderungen für die Datennutzer. Beispielsweise kommt es bei einigen Ländern zu hohen Kostenforderungen, wenn die Daten für Auswertungen herangezogen werden sollen.

Aufgrund der Ausgestaltung des Thünen-Instituts als Forschungsinstitut und Bundesoberbehörde unterliegen wir neben unserer Forschungstätigkeit zusätzlichen rechtlichen Verpflichtungen aufgrund des Status als öffentliche Stelle. Bspw. unterliegt das Thünen-Institut den Informationsfreiheitsgesetzen des Bundes (IFG, UIG). Gerade im Hinblick auf unsere Forschungsbereiche ist es unerlässlich, mit Partnern aus der Wirtschaft zusammenzuarbeiten. Hier ergibt sich des Öfteren ein Spannungsfeld zwischen dem Anspruch auf Spitzenforschung und dem Gewinn von Partnern aus der Wirtschaft, da auf Seiten der Wirtschaftspartner Hemmnisse bei der Zusammenarbeit bestehen. Auf Seiten der Wirtschaftspartner besteht die Befürchtung, dass die zu Forschungszwecken übermittelten Daten in Folge der Anwendbarkeit dieser Gesetze durch Informationsbegehren Dritten gegenüber veröffentlicht werden. Dieser Befürchtung wird derzeit mit dem Abschluss von Geheimhaltungsvereinbarungen probiert, entgegenzuwirken. Der Abschluss solcher Vereinbarungen hindert die Anwendbarkeit der Informationsfreiheitsrechte des Bundes (IFG, UIG) jedoch nicht. Das Forschungsdatengesetz bietet an dieser Stelle die Möglichkeit, den Forschungsdaten einen hohen Schutzwert zuzusprechen. Im Kontext der Informationsfreiheitsgesetze des Bundes (IFG, UIG) bedeutet dies, dass zum einen die Möglichkeit besteht, den Anwendungsbereich dieser Gesetze für die noch näher zu definierenden Forschungsdaten nicht zu eröffnen. Andererseits können die aktuell bestehenden Ausnahmetatbestände um den Schutz von Forschungsdaten ergänzt bzw. dahingehend konkretisiert werden.

Bei Kooperationen mit der Wirtschaft stellt sich zudem die Frage, inwieweit hier mit öffentlichen Geldern bestimmten Akteuren in der Wertschöpfungskette ein Vorteil geschaffen wird, bzw. in wieweit sich die Forschung für bestimmte Fragestellungen von diesen abhängig macht. Darüber hinaus sind Fragen hinsichtlich Ergebnisverwertung, Eigentum- und Rechtgestaltung und Freiheit der Forschung zu klären sowie eine Verbindlichkeit durch Wirtschaftspartner zu gewährleisten (z. B. feste Zusagen zu Ressourcen über die Dauer der Kooperation ohne, dass der Wirtschaftspartner nach 1,5 Jahren eigenständig sagen kann, aufgrund firmeninterner Entscheidungen steigt er aus der Kooperation aus).

Weitere organisatorische Hindernisse bestehen bei einigen Datensammlungen (wie beispielsweise in der Agrarstatistik) bei denen oftmals eine Kontinuität bei der Merkmalerfassung nicht gegeben ist, d. h. Merkmale werden nur für einen begrenzten Erhebungszeitraum erfasst. Dadurch wird das Erstellen von Zeitreihen für gewisse Merkmale unmöglich.

Monitoringvorhaben benötigen, beispielsweise im Bereich „Landnutzung und natürliche Ressourcen“ oder im Bereich „Ländliche Räume“ lange Zeitreihen, da die Effekte einer Intervention sich z. T. erst mit einer Verzögerung von mehreren Jahren manifestieren. Hier ist eine verbindliche Ressourcenausstattung über einen langen Zeitraum erforderlich, um wissenschaftliche Erkenntnisse ableiten zu können.

*Beispiel: Verwendung von InVeKoS-Daten, HIT-Daten und weitere*

Das Zurverfügungstellen der beispielhaft aufgeführten Daten ist aufgrund der föderalistischen Ausprägung heterogen. So gibt es bereits auf Seiten der Bundesländer einen aufwendigen und komplizierten Weg die einzelnen Daten zusammenzutragen und an den Bund zu übermitteln. Dies setzt sich auf Bundesebene fort. Insbesondere auch aufgrund nicht eindeutiger Rechtsgrundlagen für die Datenübermittlung. Beispielsweise ist der Wortlaut des § 16 Abs. 3 GAPInVeKoSG nicht zweifelsfrei als Rechtsgrundlage für die Datenübermittlung zu identifizieren, weil er nur auf die Tätigkeit des Übermittels abstellt, aber nicht die Rechtsfolge klar herausstellt. Eine Änderung des Wortlautes in „Die Zahlstellen haben auf Anforderung ... zu übermitteln...“ käme diesem Bedürfnis nach Eindeutigkeit nach. Denn damit käme klar heraus, dass zum Zweck der wissenschaftlichen Forschung ein Anspruch auf die Daten besteht. Wünschenswert wäre es auch, wenn bei den Bundesländern zentrale Stellen für die Datensammlung und -übermittlung geschaffen werden, womit eine Zentralisierung der Daten einherginge. Zudem ist es von Relevanz eindeutige Rechtsgrundlagen für die Datenübermittlung an Forschungseinrichtungen zu schaffen, die einen Datentransfer vereinfachen. Hier könnte ferner auch auf die Öffnungsklauseln des Art. 89 DSGVO Bezug genommen werden.

Gerade im Hinblick auf Langzeitforschung und Monitoringvorhaben bestehen in den folgenden Gebieten Hindernisse in der praktischen Forschung: Einerseits wird teilweise durch gesetzliche Regelungen eine konkrete Löschpflicht für die Daten vorgeschrieben (bspw. § 7 InVeKoSDG). Bei dieser Art von Forschungsvorhaben kommt es aber grds. auf eine langfristige Verfügbarkeit der Daten an, um die Entwicklung analysieren zu können. Zudem ist meist zu Beginn der Vorhaben nicht absehbar, für welchen Zeitraum die erhobenen Daten benötigt bzw. gespeichert werden. An dieser Stelle könnte das neue Forschungsdatengesetz Abhilfe schaffen und eine längerfristige Aufbewahrung von Forschungsdaten regeln, um die Langzeitforschung zu fördern. Auch die Sekundärdatennutzung/Sekundäranalysen bei (Langzeit)forschung gibt Raum für Verbesserungspotential. So ist in diesem Forschungssegment nicht von vornherein absehbar, welche Daten für anderweitige Projekte oder Forschungsfragen verwendet werden können oder sollen. Die Sekundärdatennutzung hängt, wie auch die Primärnutzung, von der ausdrücklichen Einwilligung der betroffenen Personen ab, was gerade bei einer langfristig angelegten Forschung zu Problemen führen kann, wenn weitere Forschungsfragen erst nach langjähriger Forschung auftreten. An dieser Stelle könnte das neue Forschungsdatengesetz eine Rechtsgrundlage etablieren, die es Forschungseinrichtungen u. a. ermöglicht (Roh)daten für die Sekundärdatennutzung zu wissenschaftlichen Zwecken zu nutzen, ggf. mit Einschränkungen in den Veröffentlichungsrechten.

*Beispiel: Evidenzbasierte Politikberatung anhand der Agrarstatistik*

Am Beispiel der Agrarstatistik möchten wir die oben beschriebenen Hindernisse veranschaulichen. In der Agrarstatistik werden verschiedene Merkmale zu betrieblichen Strukturen, der Landnutzung und der landwirtschaftlichen Produktion erhoben. Aufgrund restriktiver Geheimhaltungsvorgaben werden oftmals Daten gesperrt (z. B. wenn in einer Region weniger als drei Beobachtungen des Merkmals vorliegen). Dieser Umstand erschwert eine Bereitstellung statistischer Informationen auf regionaler / lokaler Ebene erheblich bzw. macht diese unmöglich. Eine evidenzbasierte Politikberatung und Forschung zu agrarpolitischen Maßnahmen, die zunehmend auf kleinräumig abgegrenzte Gebiete (z. B. „rote“ Gebiete) zugeschnitten werden und ein

differenziertes Spektrum an politischen Interventionen nutzen, wird dadurch erheblich erschwert und ist nur sehr eingeschränkt möglich. Ferner ist die Verknüpfung der Daten der Agrarstatistik mit denen der Verwaltung auf Mikroebene (=Betriebsebene) rechtlich untersagt. Darüber hinaus verhindern gesetzlich geforderte Löschfristen für Verwaltungsdaten langfristige Analysen und Auswertungen beispielweise von umweltpolitischen Fragestellungen im Kontext der Politikberatung, da erst nachträglich benannt werden kann, welche Daten zur Analyse einer Problemstellung benötigen werden.

## 2.2 Technische Hindernisse

Datenaufbereitungen, insbesondere Datenumformatierungen, sind nach wie vor nötig, da die bereitgestellten Daten nicht interoperabel sind. Existierende Datenbanken sind oft instabil und nicht voll funktionsfähig. Oft sind die Nutzerinterfaces nicht oder nur für bestimmte Nutzergruppen, intuitiv bedienbar. Dadurch können die Datenbanken meist erst nach einer Einarbeitung oder Schulung bedient werden. In der Regel muss eine Vielzahl von fachspezifischen Datenbanken abgefragt werden, die in der Regel auf unterschiedlichen technischen Lösungen basieren, sodass eine Nutzung auch hier erst nach einer Einarbeitung möglich ist. Diese Einarbeitung in Datensuche, Datendownload, usw. sowie die Nutzungsbedingungen der einzelnen Datenbanken binden Arbeitskapazitäten, die nicht mehr für die eigentliche Forschungsleistung zur Verfügung stehen. Dadurch ist die Forschungsarbeit insbesondere bei interdisziplinären Forschungsvorhaben erschwert.

Öffentliche Einrichtungen (z. B. die statistischen Landesämter, aber auch spezifische Bundesbehörden) verfügen offenbar nicht über hinreichende Kapazitäten, aggregierte Daten in maschinenlesbarer Form bereitzustellen. Das bindet Ressourcen zur Datenaufbereitung und verlangsamt die oben beschriebenen Arbeiten. Dies gilt selbst für die aktuellen Erhebungsjahre; für Daten vor dem Jahr 2000 sind PDFs und physische Ausdrücke in der Regel die einzigen Bereitstellungsformen.

Wenn eine technische Infrastruktur gegeben ist, fehlen mittelfristig die finanziellen Ressourcen um die technische Infrastruktur zu erneuern, z. B. wenn Teile aus dem Support fallen, sie technisch entsprechend der (veränderten) Nutzungsanforderungen und Nutzerbedürfnisse, weiter zu entwickeln sowie den allgemeinen technischen Fortschritt einzupflegen. Hier gilt es konzeptionell Ressourcen für die Etablierung und kontinuierlich für den Betrieb und die Weiterentwicklung der technischen Infrastruktur vorzuhalten und bestehende Infrastrukturen auf Bundes- und Länderebene entsprechend des aktuellen State-of-the-Art-Standards zu ertüchtigen.

Als positiv Beispiel sei die Plattform govdata.de genannt, welche Datensätze gut auffindbar darbietet, aber je nach veröffentlichender Stelle die Nutzung des eigentlichen Datensatzes nicht intuitiv ermöglicht. Zudem ist diese mit klarem Fokus auf Verwaltungsdaten konzipiert, welche andere Anforderungen an die Metadaten und Speicherung als Forschungsdaten aufweisen.

## 3 Lösungsmöglichkeiten

Art. 5 Abs. 3 S. 1 GG stellt bereits den Grundsatz auf, dass Wissenschaft, Forschung und Lehre frei sind. Die aufgezeigten Hindernisse zeigen dennoch, dass in Teilen Beschränkungen bestehen. Das neue Forschungsdatengesetz bietet daher die Möglichkeit, als Ausgestaltung der verfassungsrechtlich garantierten Freiheit, bei den beschriebenen Problemfeldern Abhilfe zu schaffen und damit der Wissenschaft, Forschung und Lehre einen neuen Stellenwert in der Gesetzeslage zu verschaffen. Insbesondere sollte die Chance genutzt werden, ein einfachgesetzliches Forschungsprivileg in dem neuen Gesetz zu verankern, welches bei Auslegungs- und Abwägungsfragen herangezogen werden kann, um die beschriebenen Spannungsfelder zugunsten der Wissenschaft, Forschung und Lehre aufzulösen, ohne die Schutzbedürftigkeit der Daten in Frage zu stellen. Zu berücksichtigen sind dabei ferner anderweitige Bundesgesetze/Verordnungen sowie europarechtlich determinierte Grundsätze, sodass es nicht zu einer widersprüchlichen Gesetzeslage kommt. Dabei geht es bspw. um die Schranken der DSGVO und des UrhG, welche auf Beeinträchtigungen der praktischen Forschung hin



überprüft werden sollten. Bei Auftreten solcher Kollisionen sollten diese zugunsten der Wissenschaft, Forschung und Lehre aufgelöst werden. Dies betrifft zuvorderst das Spannungsverhältnis zwischen der Freiheit der Forschung und dem Schutz der Rechte Dritter, insbesondere wenn es um den Schutz personenbezogener Daten geht.

Für die praktische Anwendbarkeit des neuen Gesetzes ist es daher entscheidend, dass es klare und praxisnahe Definitionen und Begriffsbestimmungen enthält. Zu Beginn des Gesetzes sollten daher der persönliche und sachliche Anwendungsbereich eindeutig festgelegt werden. Hierbei sollte darauf Wert gelegt werden, dass bei der Definition der Forschungsdaten ein möglichst weiter Begriff - ergänzend durch einen nicht abschließenden Katalog - verwendet wird, um einen breiten Anwendungsfall des Gesetzes zu erzielen. Auch der persönliche Anwendungsbereich sollte sich an einem weiten Verständnis orientieren, um keinen Akteur bei der Datenerhebung, -übermittlung und -sicherung herauszunehmen. Nur so kann den aktuell bestehenden Unsicherheiten in der Rechtsanwendung bzw. in der Risikoabwägung im Umgang mit Forschungsdaten begegnet werden.

Das neue Forschungsdatengesetz sollte den Fokus ferner auf Digitalisierung legen. Dabei sollten digitale Lösungen für die Erhebung von und den Zugang zu Daten, sowie deren Übermittlungsprozess zwischen Datengeber und Datennutzer gleich zu Beginn mitgedacht werden.

### 3.1 Zu prüfende legislative Regelungen

Aus unserer Perspektive sollten die Statistikgesetze zukünftig eine Verknüpfung von Daten und eine Datensammlung zu Forschungszwecken zulassen. In den Gesetzen zur Datenerhebung sollte die Forschung, Politikberatung, Evaluation und Monitoringaufgaben (z.B. Umweltmonitoring, Monitoring ländlicher Räume) explizit als Nutzungszweck vorgesehen werden. Eine regelmäßige Prüfung der Nutzungsmöglichkeiten bei älteren Datenerhebungen sollte vorgesehen werden sowie eine Anonymisierung von Datensätzen für Forschungszwecke ermöglicht werden. Die Statistikgesetze sollten im Hinblick auf Geheimhaltungsvorgaben bei der Datennutzung durch Forschung und Politikberatung überprüft werden und diese verringert werden.

### 3.2 Nicht-legislative Lösungsvorschläge

Neben gesetzlichen Regelungen können auch technisch und organisatorische Maßnahmen zur Verbesserung des Status Quo führen. Diese können gegebenenfalls bei der Erstellung des neuen Forschungsdatengesetzes mitgedacht werden.

#### Organisatorische Vorschläge

Bei den Statistischen Ämtern der Länder und dem Statistischen Bundesamt sowie den Geodatenverwaltungen der Länder wünschen wir uns eine stärkere Orientierung auf Datennutzer außerhalb der jeweiligen Häuser inkl. einfachere und kostengünstigere Datennutzungsmöglichkeiten für Forschungszwecke auf der Basis aktueller technische Standards (z. B. Bereitstellung von standardisierten, maschinenlesbaren Schnittstellen).

Bei personenbezieharen bzw. personenbezogenen Daten ist es wünschenswert, für Forschungszwecke Anonymisierungs-Routinen zu etablieren, die eine einfachere Nutzung der Daten ermöglichen.

Darüber hinaus sind fachliche und technische Qualifikationen des benötigten Personals relevant, die im Kontext der Digitalisierung zunehmend nachgefragt werden, so dass eine hohe Konkurrenz um geeignete Fachkräfte entsteht. Im Bereich Forschung/Wissenschaft z. T. stärker als im Bereich der Wirtschaft, da neben IT-technischer Qualifikationen zudem ein grundlegendes Verständnis der jeweiligen Fachdisziplin notwendig ist. Hier ist ferner anzumerken, dass Personen mit geeigneter Qualifikation zwar essentielle Aufgaben in den Einrichtungen übernehmen (z. B. Datenmanagement, Betreuung der technischen Infrastruktur), aber nach TVÖD als technische

Kräfte deutlich schlechter als in der Wirtschaft entlohnt werden. Dies führt dazu, dass entsprechend qualifiziertes Personal kaum gewonnen werden kann bzw. schnell abwandert. Hier ist mehr Flexibilisierung erforderlich, wenn der öffentliche Dienst im Bereich IT-Infrastruktur/ Datenmanagement konkurrenzfähig werden soll.

### *Anreize für das Datenteilen*

Die Bereitschaft zum Datenteilen kann auf Seiten der Wissenschaft durch vereinfachte ggf. automatisierte Prozesse zum Daten veröffentlichen und stabilen Dateninfrastrukturen, die mit einem intuitiven Interface ausgestattet sind, erhöht werden. Ein Hemmnis für die Bereitstellung und den Austausch von Daten aus der Perspektive der Wissenschaft sind die rechtlichen Rahmenbedingungen, die dies häufig ausschließen oder erschweren. Hier ist es wünschenswert, wenn Forschungsergebnisse bereitgestellt werden können, auch wenn diese potentiell personenbeziehbar sind (z. B. bei Aggregation auf Gemeindeebene / 1km x 1km Raster o. Ä.). In Hinblick auf die Verwaltung besteht bei der Herausgabe von Mikrodaten der landwirtschaftlichen Landnutzung die latente Gefahr, dass die aufbereiteten Ergebnisse im Vollzug (und zum Nachteil der Landwirte) genutzt werden, dies kann mittelfristig die Qualität der Eingangsdaten deutlich beeinträchtigen und damit die Aussagekraft späterer Auswertungen negativ beeinflussen.

### *Rolle und Aufgaben von Datenmittlerstrukturen beim Datenteilen*

Datenmittler ermöglichen es, fachliches und technisches Know-how zu bündeln und effizient einzusetzen. Dabei kann die Datenakquise, das Datenmanagement und die Datenaufbereitung, -bereitstellung und -veröffentlichung hinsichtlich personeller und IT-struktureller Ausstattung und Auslastung effizient gestaltet und zwischen den Akteuren abgestimmt sowie weiterentwickelt werden. Datenmittler können daher bei entsprechender Konzeption und Ausstattung ein effizientes Vorgehen bei der Akquise fördern, Fehler bei der Datenaufbereitung minimieren und die Konsistenz sowie Auswertbarkeit der Daten erheblich verbessern. Agieren sie zudem als zentrale Strukturen für die Veröffentlichung der resultierenden Ergebnisse sowie den Austausch von Wissen zwischen den Datennutzern, kann die Qualität und Öffentlichkeitswirkung der Forschungsergebnisse weiter verbessert werden. Darüber hinaus sollte eine Einbindung der nationalen Datenmittler in europäische und weltweite Datennetzwerke angestrebt werden. Aus unserer Perspektive ist beispielsweise das Statistische Bundesamt ein Datenmittler.

### **Technische Vorschläge**

Die technischen Infrastrukturen auf Landes- und Bundesebene müssen auf den aktuellen Standard gebracht werden. Hier sind Finanzierungen (Erstbeschaffung und Folgekosten für den laufenden Betrieb) vorzuhalten und die Beschaffungsvorgänge zwingend zu beschleunigen.

Ferner ist zu gewährleisten, dass Forschungsdaten der Einrichtungen und deren Berechnungsgrundlagen zumindest aus den Bundesämtern (DWD, Statistisches Bundesamt, ...) entsprechend der wissenschaftlichen Vorgaben langfristig nach dem Abschluss der Forschungen abrufbar bleiben, um die wissenschaftliche Qualitätssicherung zu ermöglichen. Hierfür werden Datenarchive benötigt, die den technischen Anforderungen entsprechen und kontinuierlich aktuell gehalten werden (z. B. technischer Support, technischer Fortschritt).

Neben der rechtlich mögliche Datenbereitstellung sind auch personelle Ressourcen und technische Infrastrukturen, die meist nicht in dem dafür notwendigen Umfang vorhanden bzw. die gegenwärtig vorhandenen Systeme insbesondere in den öffentlichen Stellen nicht dafür konzipiert sind, benötigt. Eine zentrale Datenaustauschplattform mit hoher Nutzerfreundlichkeit hinsichtlich Up- und Downloads sowie Datensuche, guter Performanz und Datensicherheit hinsichtlich Zugriffsoptionen und Datenspeicherung/Backup existiert unserer Kenntnis nach in Deutschland nicht. Hier könnte eine zentrale und einheitliche Austauschplattformen, zum Beispiel im Rahmen des Research-Data-Commons Konzeptes der nationalen Forschungsdateninfrastruktur, als Ergänzung zu govdata.de geschaffen werden.

## 4 Sonstiges

Urheberrechte an einem Datenbankwerk können mit einem Leistungsschutzrecht des Datenbankherstellers kumulieren. Für das Leistungsschutzrecht des Datenbankherstellers sind die Schranken in § 87c UrhG einschlägig (Schulze, in: Dreier/Schulze, Urheberrechtsgesetz, 7. Aufl. 2022, § 23 UrhG, Rn. 43) und abschließend (Vohwinkel, in: BeckOK Urheberrecht, 37. Edition Stand 15.07.2022, § 87c UrhG, Rn. 1). Daher ist Teil 1 des UrhG nur anwendbar, wenn darauf verwiesen wird. Auf § 23 UrhG wird nicht verwiesen. Wahrt ein neues Werk einen hinreichenden Abstand nach § 23 Abs. 1 S. 2 UrhG bedarf es keiner Zustimmung des Urhebers der verwendeten Werke zu Veröffentlichung und Verwertung des neu geschaffenen Werkes (Ahlberg/Lauber-Rönsberg, BeckOK Urheberrecht, 37. Edition Stand 15.10.2022, § 23 UrhG, Rn. 27). Dadurch kann es aber bei kumulativ bestehendem Urheberrecht als Datenbankwerk und einem Leistungsschutzrecht des Datenbankherstellers vorkommen, dass das Leistungsschutzrecht des Datenbankherstellers immer noch verletzt ist, weil man sich außerhalb der Schranken des § 87c UrhG bewegt. Das Urheberrecht des Urhebers am Datenbankwerk wäre dagegen nicht verletzt. Auch wenn „nur“ ein Leistungsschutzrecht des Datenbankherstellers besteht, ist dies problematisch. Denn dadurch wird die Benutzung von Datenbanken für die wissenschaftliche Forschung beispielsweise durch Kombination und Umsetzung in andere Darstellungsformen (z. B. von mehreren Datensätzen in Tabellenform in eine Karte mit aggregierten Daten) erschwert und damit insbesondere auch die Wissenschaftskommunikation und Aufbereitung für den Bürger. Dies wirft zudem Haftungsfragen auf. Da die Forschung häufig auf aktuelle Daten angewiesen ist, kann auch die Schutzrecht-Dauer nicht abgewartet werden. Wünschenswert wäre ein Verweis in den §§ 87a ff. UrhG auf den § 23 UrhG oder aber eine vergleichbare Regelung, die es der Forschung ermöglicht, aus Datenbanken neue Werke zu schaffen. Da in der Praxis häufig nicht abseh- oder beurteilbar ist, ob (jedenfalls auch) ein Leistungsschutzrecht des Datenbankherstellers besteht und dieses im Zweifel immer mit bedacht werden muss, würde dies die Prozesse deutlich erleichtern. Die Einschränkung des Leistungsschutzrechts des Datenbankherstellers ist auch insofern gerechtfertigt, da es nicht einsehbar ist, weshalb bei demselben Sachverhalt der Urheber eines Datenbankwerkes (wenn z. B. die Schutzrechte kumulieren) „schlechter“ gestellt sein sollte, da er im Gegensatz zum Datenbankhersteller nicht nur eine Investition, sondern eine Schöpfung erbracht hat.

Nach einem Zeitraum von 5 Jahren nach Verabschiedung des neuen Forschungsdatengesetzes sollte eine Evaluation des Gesetzes unter Einbindung der Akteure aus Wissenschaft und Forschung erfolgen.