

Personenbezogene Forschungsdaten in unverdächtigen Disziplinen: Das Beispiel der Erd-, Umwelt- und Agrarwissenschaften

Niklas K. Hartmann

Am Beispiel der Erd- und Umweltwissenschaften (einschließlich der landschafts- und standortbezogenen Teilgebiete der Agrarwissenschaften) zeigt dieser Beitrag, dass auch in scheinbar „unverdächtigen“ Disziplinen personenbezogene Forschungsdaten vorkommen. Eine Auswertung der Literatur zeigt, dass allgemeine Handreichungen zum Datenschutz in der Forschung kaum Unterstützung bei der Arbeit mit den für diese Disziplinen besonders relevanten Fällen bieten. Für die in den Erd- und Umweltwissenschaften besonders relevanten raumbezogenen Daten kommt hinzu, dass selbst unter Fachjuristinnen Uneinigkeit über die datenschutzrechtliche Bewertung herrscht. Die Ergebnisse einer empirischen Vorstudie zeigen eine ganze Reihe verschiedener Arten personenbezogener Forschungsdaten auf, die in der Forschungspraxis der Erd- und Umweltwissenschaften eine Rolle spielen. Sie legen außerdem nahe, dass der Umgang mit personenbezogenen Daten in der Forschungspraxis der Erd- und Umweltwissenschaften auf Grund der mangelnden Vertrautheit mit dem Datenschutz nicht immer den rechtlichen Anforderungen entspricht. Auch Unterstützung durch Fachgesellschaften und Infrastruktureinrichtungen – etwa in Form disziplinspezifischer Handreichungen, qualifizierter Beratung oder institutionalisierten Möglichkeiten, Daten sicher zu archivieren und gegebenenfalls zugangsbeschränkt zu publizieren – bestehen kaum. Aus dieser Situation ergeben sich Herausforderungen an die Weiterentwicklung der disziplinären Datenkultur und Dateninfrastruktur, beispielsweise im Rahmen des Prozesses zum Aufbau einer Nationalen Forschungsdateninfrastruktur (NFDI). Zu den Möglichkeiten für Infrastruktureinrichtungen, diese Weiterentwicklung zu unterstützen, zeigt dieser Beitrag Handlungsoptionen auf.

Einleitung

Der Begriff der Nachhaltigkeit bleibt problematisch. Gerade in seiner nahezu beliebigen Ausweitung als Synonym für „gut gemacht, und zwar mehr als nur kurzfristig“ kann das Reden über Nachhaltigkeit maßgeblich zum Erhalt offensichtlich nicht dauerhaft tragfähiger Konfigurationen beitragen, wie Blühdorn (2007) für die umwelt- und entwicklungspolitische Herkunftsdomäne des Begriffs zeigt. Das ist insbesondere dann der Fall, wenn die Konzentration

auf die Effizienz einzelner Elemente und Abläufe den Blick auf schwierige und konfliktträchtige Aspekte des Ganzen verstellt: Auch die Steuerzentrale eines Braunkohlekraftwerks kann mit zertifizierter Green IT ausgestattet werden. Davon mag die Betreiberin auch einfacher zu überzeugen sein als von einer Stilllegung. Die CO₂-Bilanz wird das jedoch kaum beeinflussen.

Zugang zu Forschungsdaten ist nur dann „nachhaltig“ (im Sinne von längerfristig sicherzustellen), wenn er nicht ausschließlich informell gewährt wird. Daten, die nicht an eine Forschungsdateninfrastruktur übergeben werden, überleben selten die Emeritierung der Arbeitsgruppenleitung (vergleiche Vines *et al.* 2013). Damit Forschungsdaten an eine solche Infrastruktur übergeben werden können, müssen die Daten – neben vielen anderen Kriterien – entweder juristisch und forschungsethisch unbedenklich sein oder die Infrastruktur muss in angemessener Art und Weise beschränkten Zugang gewähren können. Für Forschungsgebiete, in denen regelmäßig „problematische“ Daten verarbeitet werden, müssen die relevanten Forschungsdateninfrastrukturen also in Abstimmung mit den Forschenden die nötigen rechtlichen und ethischen Standards, Kompetenzen sowie die organisatorischen und technischen Voraussetzungen entwickeln. Geschieht dies nicht, können die Infrastrukturen – egal wie „nachhaltig“ sie im technischen und im ökologischen Sinne betrieben werden – für das jeweilige Gebiet relevante Daten nicht „nachhaltig“ bereitstellen.

Ist von personenbezogenen Forschungsdaten die Rede, geht es meist um die Sozialwissenschaften, die Humanmedizin oder um andere Wissenschaften, für die der Mensch im Mittelpunkt des Erkenntnisinteresses steht, wie Psychologie und Pädagogik. In diesen Fächern ist ein Grundverständnis für den Datenschutz meist ein Teil der Methodenausbildung. Oft haben Fachgesellschaften bereits Leitlinien herausgegeben und disziplinspezifische Forschungsdateninfrastrukturen bieten spezialisierte Beratung an sowie die Möglichkeit, Daten zu archivieren und zugangsbeschränkt zur Nachnutzung bereitzustellen. Auch mit forschungsethischen Erwägungen mit Bezug auf Forschungsdaten über den rechtlich erforderlichen Datenschutz hinaus gibt es meist eine gewisse Vertrautheit.

In meiner Beratungstätigkeit im Rahmen meiner Aufgabe als Koordinator Forschungsdaten an der Universität Potsdam begegnen mir jedoch immer wieder Fälle, in denen in scheinbar „unverdächtigen“ Disziplinen Forschungsdaten verarbeitet werden, die Personenbezug haben, deren Personenbezug unklar ist oder die aus Gründen der Forschungsethik nicht allgemein verfügbar werden dürfen. Oft fällt dies erst auf, wenn Daten mit hohem Nachnutzungspotenzial veröffentlicht werden sollen. Einzelfälle kann ich natürlich immer von dem Forschungsdaten-Kontaktpunkt an Stellen wie die Datenschutzbeauftragte oder die Ethikkommission weiterverweisen. Häufen sich jedoch Fälle aus einzelnen Bereichen, liegt eine Bestandsaufnahme nahe, um festzustellen, ob Handlungsbedarf besteht und wenn ja, für wen (zum Beispiel Strukturen des Faches wie Fachgesellschaften oder disziplinspezifische Infrastruktureinrichtungen oder generische Strukturen wie Infrastruktureinrichtungen der Hochschulen) und in welchem Rahmen. Aktivitäten dieser Art können auch zusätzliche Handlungsoptionen eröffnen, die zu einem früheren Zeitpunkt im Forschungsprojekt-Managementzyklus beziehungsweise Forschungsdaten-Lebenszyklus ansetzen als erst zur Veröffentlichung.

In diesem Beitrag erfolgt eine solche Bestandsaufnahme für das Gebiet der Erd- und Umweltwissenschaften, einschließlich der landschafts- und standortbezogenen Teilgebiete der Agrarwissenschaften. Aus diesem Bereich haben mich in der Vergangenheit Anfragen erreicht und meine Ausbildung als Diplom-Geoökologe gibt mir den nötigen Einblick in die Eigenheiten der

betroffenen Fächer. Im Folgenden stelle ich zunächst die Forschungsdaten der Erd- und Umweltwissenschaften vor (Abschnitt 2), diskutiere dann ihren möglichen Personenbezug (Abschnitt 3), präsentiere die Ergebnisse einer empirischen Vorstudie (Abschnitt 4), leite schließlich Handlungsempfehlungen insbesondere für Infrastruktureinrichtungen ab (Abschnitt 5) und schließe mit einem Fazit samt Ausblick (Abschnitt 6).

Forschungsdaten in den Erd- und Umweltwissenschaften

Ein Kennzeichen von Forschungsdaten in den Erd- und Umweltwissenschaften ist ihre Heterogenität. Dabei gehen nicht nur verschiedene Teildisziplinen mit sehr unterschiedlichen Daten um, vielmehr wird selbst innerhalb einzelner Forschungsvorhaben oft mit heterogenen Daten gearbeitet. Die Daten unterscheiden sich nicht nur in ihrem Inhalt (welche Aspekte welcher Umweltkompartimente sie wie repräsentieren), sondern auch in ihrer Art und ihrem Entstehungskontext.

In vielen Teildisziplinen und Forschungsfeldern spielen nicht replizierbare und damit besonders wertvolle Beobachtungsdaten eine ebenso große Rolle wie experimentelle Daten. Sowohl experimentelle als auch Beobachtungsdaten können je nach Fragestellung im Labor oder im Freiland gewonnen werden. Die Daten können als qualitative Beschreibungen in verschiedenen Graden der Standardisierung ebenso vorliegen wie als quantitative Zähl- und Messdaten. Wo Land unter menschlicher Nutzung betroffen ist, werden im Freiland gewonnene Daten teils mit Einverständnis von Eigentümerinnen oder anderen Verantwortlichen erhoben (zum Beispiel beim Nehmen von Bodenproben oder bei der Arbeit über Wildtiere in Schutzgebieten), teils gemeinsam mit für die Kooperation entlohten Landmanagerinnen generiert (zum Beispiel in Feldversuchen), teils aber auch ganz ohne Zutun der Landnutzerinnen verarbeitet (etwa bei der Nutzung von Fernerkundungsdaten oder Daten aus Kartierungen und Zählungen). Viele dieser Daten haben einen expliziten Raumbezug, auch Zeitreihen werden regelmäßig erzeugt. Zur Verarbeitung der Daten werden oft Geographische Informationssysteme oder spezialisierte Forschungssoftware genutzt. Das Ergebnis sind oftmals komplexe Datenprodukte, teils in Form (digitaler) Karten. Die Entstehungskontexte der Daten reichen von Individualforschung über Arbeitsgruppen und kleinere Verbundprojekte bis zu großen Forschungsinfrastrukturen wie etwa Satelliten der Erdbeobachtung oder Dauerfeldversuchen der Agrar- und Forstwissenschaften beziehungsweise Dauerbeobachtungsflächen der terrestrischen Umweltforschung. Auf Grund ihrer Heterogenität können die meisten Daten der Erd- und Umweltwissenschaften, die nicht aus Forschungsinfrastrukturen stammen, aus Sicht der Dateninfrastrukturen als *long tail*-Daten bezeichnet werden (vergleiche Klump 2012; Bertelmann *et al.* 2014).

Während insbesondere in den Geowissenschaften *sensu strictu* Daten aus Großprojekten bereits seit langem von Dateninfrastrukturen vorgehalten und international verfügbar gemacht werden, gibt es für *long tail*-Daten in vielen Teildisziplinen eine verbreitete Kultur des informellen Teilens. Standardisierte Datenbanken (wie etwa für Sequenz- und Strukturdaten in den Biowissenschaften) entwickelten sich auf Grund der Heterogenität der Daten daher zunächst nicht. In den letzten etwa zehn Jahren waren die Erd- und Umweltwissenschaften jedoch eines der Wissenschaftsgebiete, in denen das digitale Publizieren von Forschungsdaten große Fortschritte gemacht hat (vergleiche Diepenbroek 2011, Klump 2012, Bertelmann *et al.* 2014). Die Recherchierbarkeit und Zitierbarkeit von Daten aus Großprojekten wurde ebenso verbessert wie die

Publikation von Daten aus Einzelprojekten, die verstärkt nicht nur von Fördermittelgeberinnen, sondern auch von Fachzeitschriften gefordert wird (vergleiche COPDESS 2015; Hanson, Lehnert und Cutcher-Gershenfeld 2015; McNutt *et al.* 2016; Bertelmann 2017). Einen aktuellen Überblick über die relevanten disziplinspezifischen Forschungsdateninfrastrukturen geben die Extended Abstracts und Absichtserklärungen der geplanten NFDI-Konsortien, NFDI4Agri, NFDI4Biodiversity und NFDI4Earth.¹

Eine große Relevanz für die Erd- und Umweltwissenschaften haben auch allgemein verfügbare oder auf Anfrage erhältliche Daten und Datenprodukte von Behörden, insbesondere Karten und Geodaten sowie Daten der EU-Agrarförderung. Diese wichtigen Datengrundlagen sind innerhalb der EU mittlerweile oft über die Online-Portale staatlicher Geodateninfrastrukturen verfügbar (vergleiche Bernard *et al.* 2016).

Personenbezug erd- und umweltwissenschaftlicher Forschungsdaten

„Personenbezogene Daten‘ [sind] alle Informationen, die sich auf eine identifizierte oder identifizierbare Person beziehen“ (Art. 4, Abs. 1 Datenschutzgrundverordnung der Europäischen Union, DS-GVO). Um Informationen zu sein, die sich auf eine Person beziehen, muss nicht nur eine Person identifizierbar sein, sondern die Daten müssen auch tatsächlich Angaben über diese Person enthalten (Gusy und Eichendorfer 2018, Rn. 47). Ansonsten wäre, um ein Beispiel von Forgó, Krügel und Reiners (2008, S. 19) aufzunehmen, jede Sachinformation über die Stadt Hannover personenbezogen, da wir sehr einfach Personen bestimmen können, die in Hannover wohnen.

Für die Verarbeitung personenbezogener Daten gilt seit dem 25. Mai 2018 die DS-GVO als unmittelbar geltendes Recht. Sie wird ergänzt durch das jeweils anwendbare Datenschutzgesetz (Landesdatenschutzgesetze für öffentliche Stelle in Trägerschaft der Länder, Bundesdatenschutzgesetz für Private und öffentliche Stellen in Trägerschaft des Bundes) sowie fachgesetzliche Regelungen, zum Beispiel der Statistik-, Sozial-, Archiv-, Geodaten- und Umweltinformationsgesetze. Dabei haben im Sinne des Subsidiaritätsprinzips spezialgesetzliche Regelungen Vorrang vor den allgemeinen Datenschutzgesetzen. Grundsätzlich gilt, dass personenbezogene Daten ohne Einwilligung der betroffenen Person nur zu bestimmten Zwecken verarbeitet werden dürfen (Art. 6, Abs. 1 DS-GVO). Die Verarbeitung bestimmter „besonderer Kategorien“ personenbezogener Daten ist grundsätzlich verboten und nur in Ausnahmefällen erlaubt (Art. 9, Abs. 1-2 DS-GVO). Bei der Verarbeitung personenbezogener Daten sind die Grundsätze des Art. 5 DS-GVO zu beachten, die betroffenen Personen haben weitgehende Rechte (Kapitel III DS-GVO), während den verarbeitenden Stellen diverse Pflichten auferlegt werden (Kapitel IV DS-GVO). Eine kurze Übersicht zu den Grundlagen des Datenschutzrechts nach In-Kraft-Treten der DS-GVO im Allgemeinen bietet zum Beispiel Hoeren (2018, S. 405ff).

Für den Umgang mit personenbezogenen Daten in der Forschung sind sowohl allgemeine als auch forschungsspezifische Regelungen der jeweils anwendbaren Rechtsgrundlagen relevant. Dabei geht es zum einen um die Frage, ob die Verarbeitung rechtmäßig ist, zum anderen um Grundsätze der Verarbeitung (zum Beispiel frühestmögliche Anonymisierung beziehungsweise

¹ Abrufbar unter <https://www.dfg.de/foerderung/programme/nfdi/absichtserklaerungen/index.html>.

Pseudonymisierung von Daten, die nicht anonymisiert werden können). Eine kompakte Übersicht über Anforderungen an den Datenschutz in der Forschung in Deutschland bietet RatSWD (2017). Allerdings werden dort die Auswirkungen der DS-GVO noch nicht berücksichtigt. Für ausführliche Informationen zum Datenschutz in der Forschung wird noch immer regelmäßig auf Metschke und Wellbrock (2002 [1994]) verwiesen. Die Auswirkungen der DS-GVO auf die Forschung thematisiert Schaar (2016). Bevor Forschende jedoch konkrete gesetzliche Regelungen in Anspruch nehmen, welche die Verarbeitung personenbezogener Daten erlauben, müssen sie zunächst selbst prüfen, ob die von ihnen angestrebte Verarbeitung für Forschungszwecke einer Abwägung der betroffenen Grundrechte auf informationelle Selbstbestimmung und Freiheit der Forschung standhält. Die Anforderungen dieser Abwägung stellen RatSWD (2017, S. 6f) sowie Metschke und Wellbrock (2002, S. 10–12) dar.

Für Gruppen natürlicher Personen oder juristische Personen gilt das Datenschutzrecht nicht. Falls sich jedoch aus der Erhebung, Verarbeitung oder dem Bekanntwerden von Daten ein Nachteil (zum Beispiel Stigmatisierung oder ein wirtschaftlicher Schaden) für eine Gruppe natürlicher Personen oder für juristische Personen ergeben kann, ist aus Gründen der Forschungsethik eine entsprechende Abwägung sowie die sichere Verarbeitung der Daten ebenfalls geboten. Gegebenenfalls ist eine Ethikkommission zu befragen.

Die Literatur zum Datenschutz in der Forschung und zur Forschungsethik hat oft die sozialwissenschaftliche, zeithistorische, humanwissenschaftliche oder medizinische Forschung im Blick. Dementsprechend liegt der Fokus oft auf den strengen Anforderungen für die Verarbeitung der „besonderen Kategorien“ von personenbezogenen Daten. Hinweise zu einem datenschutzrechtlich und forschungsethisch adäquaten Umgang mit in den Erd- und Umweltwissenschaften typischen Arten von Daten, die zum Beispiel über den Zustand von Flächen Rückschlüsse auf Vermögensverhältnisse von Eigentümerinnen oder die Bewirtschaftungspraxis von Landnutzerinnen zulassen, enthält diese Literatur nicht. Auch über die spezialgesetzlichen Regelungen im Bereich Zugang zu Geodaten und Umweltinformationen gibt die Literatur nur selten Auskunft, im Gegensatz zu den Regelungen in den Bereichen Statistik, Sozialstaat, Gesundheit und Archivwesen.

Personenbezug von raumbezogenen Daten, insbesondere Geodaten - Grundsätzliches

Für die Erd- und Umweltwissenschaften sind raumbezogene Daten von besonderer Bedeutung. Wird in der Literatur die Verarbeitung von Daten mit gleichzeitigem Personen- und Raumbezug in der Forschung behandelt, geht es in der Regel um die Verwendung anfallender Nutzungs- und Standortdaten mobiler Geräte und Anwendungen insbesondere in den Sozial- und Technikwissenschaften (etwa bei Hilty *et al.* 2012) oder um die Explikation und Nutzung des Raumbezugs klassischer wirtschafts- und sozialstatistischer Daten (wie in RatSWD 2012). Solche Daten sind unbestreitbar personenbezogen und können gegebenenfalls auch besondere Kategorien personenbezogener Daten enthalten. Dagegen sind für die Erd- und Umweltwissenschaften eher raumbezogene Daten zur natürlichen und bebauten Umwelt (im folgenden: Geodaten) relevant. Diese sind seit längerem Gegenstand von Auseinandersetzungen um die Reichweite des Datenschutzrechts.

Die rechtsdogmatische Diskussion wurde zunächst im Vorfeld der Einführung von Informationsrechten im Umweltbereich geführt (zum Beispiel Taeger 1991; Raum 1993) und setzte sich dann im Zuge des durch die INSPIRE-Richtlinie mandatierten Aufbaus von öffentlich zugänglichen Datenplattformen für behördliche Geodaten fort (Karg und Weichert 2007, S. 5-11). Auch in der Praxis legen Gerichte und Aufsichtsbehörden teils sehr unterschiedliche Maßstäbe an (Karg und Weichert 2007, S. 11–19). Anlass zu weiteren Diskussionen gab das Aufkommen von Internetdiensten wie Google Street View (vergleiche Gusy und Eichendorfer 2018, Rn. 47). So haben sich Geodaten zu einem wichtigen Beispiel entwickelt, an Hand dessen Abgrenzungsprobleme zwischen personenbezogenen und nicht-personenbezogenen Daten grundsätzlich verhandelt werden (Ende der 2000er Jahre zum Beispiel durch Weichert 2009; Forgó und Krügel 2010; aktuell zum Beispiel bei Krügel 2017; Schild 2018). Dies macht eine Rezeption der Rechtslage für juristische Laien nicht einfach und kann den Blick auf teilweise bestehende schulenübergreifende Übereinstimmungen in der praktischen Beurteilung verstellen.

Weitgehende Einigkeit scheint jedoch über die folgenden Grundsätze zu herrschen: Bei Geodaten handelt es sich in erster Linie um Sachdaten. Diese „sind zunächst keine personenbezogenen Daten. Sie beziehen sich auf eine Sache und beschreiben diese“ (Schild 2018, Rn. 22). Gerade „Geodaten können aber auch Personenbezug aufweisen“, was bei Katastern und katasterartigen Informationssystemen, über die sich zu Grundstücksflächen gehörige Adressen und Eigentümerinnen ermitteln lassen, offensichtlich wird (Schild 2018, Rn. 23). In weniger offensichtlichen Fällen kommt es neben der Genauigkeit des Raumbezugs auch auf den Informationsgehalt der Daten an, da dieser bestimmt, ob die Sachdaten Auswirkungen „auf rechtliche, wirtschaftliche oder soziale Positionen einer Person haben oder sich zur Beschreibung ihrer individuellen Verhältnisse“ haben (Schild 2018, Rn. 24; vergleiche auch Artikel-29-Datenschutzgruppe 2007, S. 13). Ein typischer Fall dafür aus der kommerziellen Verwendung von Geodaten ist das Geoscoring, also die Bewertung der Kreditwürdigkeit einer Person über auf ihre Wohnadresse bezogene Sachdaten (Schild 2018, Rn. 25–27).

Damit widerspricht der aktuelle Konsens der früher verbreiteten Einschätzung, dass grundsätzlich alle lagegenauen Geodaten personenbezogen seien, weil durch sie eine sachbezogene Information einer identifizierbaren Person, nämlich der Eigentümerin oder Nutzerin der Fläche, zugeordnet werden kann (so etwa Taeger 1991). Eine solche Überdehnung des Schutzbereichs des Datenschutzes ist weder praktikabel, noch unterstützt sie einen konsequenten Vollzug, sondern leistet im Gegenteil Argumenten für eine Aushöhlung eines vollkommen überdehnten Datenschutzes Vorschub (Weichert 2007, S. 23). Dass dies nicht nur theoretische Überlegungen sind, zeigen beispielsweise Beratungsergebnisse der Kommission Recht und Geodaten der Deutschen Gesellschaft für Kartographie (DGfK), die nicht nur für einen deutlich engeren Schutzbereich bei Fernerkundungsdaten plädieren, sondern selbst für Katasterdaten jeglichen Personenbezug bestreiten (Diez *et al.* 2009).

Personenbezug von raumbezogenen Daten, insbesondere Geodaten – bisher vorgeschlagene Prüfverfahren

Konkrete Prüfverfahren für den Personenbezug von raumbezogenen Daten wurden im Kontext der Umsetzung der INSPIRE-Richtlinie zunächst von zwei einschlägigen Gutachten etabliert

(Karg 2008; Forgó, Krügel und Reiners 2008), deren Ergebnisse in begleitenden Zeitschriftenartikeln weiter kontextualisiert wurden (Weichert 2009; Forgó und Krügel 2010). In beiden Verfahren von steht zu Anfang die Prüfung, ob die Daten sich auf eine identifizierbare Person beziehen lassen (Karg 2008, S. 8–16; Forgó, Krügel und Reiners 2008, S. 10–17, 22–23). Falls ja, ist zu prüfen, ob identifizierbare Geodaten tatsächlich Informationen über eine Person sind. Dazu wird im Anschluss an die Artikel-29-Datenschutzgruppe (2007) ein so genannter kontextbezogener Ansatz verfolgt (Karg 2008, S. 16–22; Forgó, Krügel und Reiners 2008, S. 18–21). Das heißt, es werden Arten und Weisen unterschieden, wie primär sachbezogene Geodaten Auskunft über Personen geben und sich „auf die Rechte und Interessen einer natürlichen Person auswirken“ können (Artikel-29-Datenschutzgruppe 2007, S. 13). Dabei muss ein Datum nur in eine der genannten Kategorien beziehungsweise Kontexte fallen, um personenbezogen zu sein und damit dem Datenschutz zu unterfallen.

Es werden drei so genannte Kontexte oder Elemente identifiziert (Karg 2008, S. 21–22; Forgó, Krügel und Reiners 2008, S. 19):

- **Inhaltskontext beziehungsweise Inhaltselement**

Es handelt sich um Daten, die direkt Auskunft über Einzelpersonen geben, zum Beispiel über Identität, Individualität, Intimsphäre, aber auch über Zustände oder tatsächliches Verhalten der Betroffenen und damit offensichtlich im klassischen Sinne um personenbezogene Daten. Bei diesen Daten kann es sich gegebenenfalls sogar um besondere Kategorien personenbezogener Daten handeln. Typisches Beispiel sind Ortungsdaten aus dem Mobilfunk.

- **Zweckkontext beziehungsweise Zweckelement**

Die Daten sind geeignet, das räumliche Umfeld einer Person zu bewerten und können so dazu führen, dass diese Person in einer bestimmten Weise beurteilt, behandelt oder beeinflusst wird. Dazu müssen die Daten nicht flächenscharf sein, wohl aber kleinräumig. Ein typisches Beispiel ist das bereits genannte Geoscoreing.

- **Ergebniskontext beziehungsweise Ergebniselement**

Das Bekanntwerden dieser Daten kann sich auf Rechte und Interessen einer bestimmten Person auswirken. Dabei geht es meist um wirtschaftliche Interessen und Verfügungsrechte über Eigentum. Die Person muss dabei spezifisch, d. h. im Unterschied zu anderen betroffen sein: Die Information, dass Spitzbergen abgelegt ist und ein rauhes Klima hat, ist sicherlich relevant für dortige Grundstückspreise, jedoch offensichtlich nicht personenbezogen. Ein typisches Beispiel für den Ergebniskontext ist bei weiter Auslegung bereits die Information, ob einem Gebäude Denkmalwert zukommt, und zwar auf Grund der damit verbundenen Beschränkung der Verfügungsbefugnis über Eigentum (Karg 2008, S. 21); bei enger Auslegung dagegen in etwa erst die Information, ob die Verursachung einer Altlast auf einem Grundstück dem aktuellen Eigentümer zuzurechnen ist, nicht aber die reine Information über das Vorhandensein einer Umweltbelastung an einem bestimmten Ort (Weichert 2007, S. 22).

Auch wenn Geodaten in all diesen Fällen dem Datenschutz unterliegen, ist der entstehende Eingriff in Grundrechte unterschiedlich schwerwiegend. Er ist bei vorliegenden Inhaltselementen größer als bei Zweckelementen und bei Ergebniselementen am kleinsten (Karg 2008, S. 21–22).

Diese Risikoabschätzung in zweifacher Hinsicht relevant: Zum einen betrifft sie die grundrechtliche Abwägung, ob die Erhebung und Verarbeitung der Daten nicht übermäßig und verhältnismäßig ist. Zum anderen hat sie Einfluss darauf, was angemessene Schutzmaßnahmen sind und welche gesetzlichen Ausnahmen zur Erhebung, Verarbeitung und Weitergabe gegebenenfalls auch ohne Einverständnis der Betroffenen in Betracht kommen. Eine solche abgestufte Beurteilung wird auch durch Erwägungsgrund 26 DS-GVO gestützt, der einen „graduellen“ Ansatz bei der Feststellung der Identifizierbarkeit unterstützt (vergleiche Gussy und Eichendorfer 2018, Rn. 47; Hofmann und Johannes 2017, S. 222–225).

Für die Bereitstellung von amtlichen Geodaten für die Wirtschaft nimmt das Gutachten von Karg (2008, S. 55–69) zudem für zahlreiche Geodaten eine allgemeine Einschätzung des Risikos nach einem Ampelsystem vor. Dabei fasst Karg (2008) die Reichweite des Ergebniskontexts deutlich weiter als beispielsweise Weichert (2007, S. 21–22) – und das, obwohl es sich um ein kooperierendes Autorenteam handelt (vergleiche Karg und Weichert 2007). Gerade in diesem für die Erd- und Umweltwissenschaften besonders relevanten Bereich besteht also wenig Einigkeit.

Personenbezug von Geodaten - Spezialgesetzliche Regelungen und Verwaltungsleitfäden

Für den Zugang zu bei Behörden vorliegenden personenbezogenen Geo- und Umweltdaten bestehen spezialgesetzliche Regelungen. Das Umweltinformationsgesetz (UIG) und die analogen Regelungen der Länder begründen ein Jedermannsrecht auf Zugang zu bei Behörden vorliegenden Umweltinformationen. Dabei wird der Datenschutz gegenüber den allgemeineren Regelungen des Informationsfreiheitsgesetzes (IFG) und den analogen Gesetzen der Länder deutlich eingeschränkt (Karg und Weichert 2007, S. 23–24). Obwohl bei der massenhaften öffentlichen Bereitstellung von Geodaten für automatisierte Abrufverfahren ein ganz anderes Gefährdungspotenzial vorliegt als beim Einblick in einzelne Akten auf Antrag (BFDI 2009, S. 89–90) und Einzelfallentscheidungen kaum möglich sind (Karg 2008, S. 42), verweist die Umsetzung der INSPIRE-Richtlinie in Deutschland, das Geodaten-Zugangsgesetz (GeoZG), auf die Regelungen des UIG.

Als Grundlage für die notwendigen Entscheidungen in den Verwaltungen wurde durch eine Bund-Länder-Arbeitsgruppe ein mit dem Bundesbeauftragten für Datenschutz und Informationsfreiheit abgestimmter Leitfaden entwickelt (IMAGI 2014). Dieser stellt vorsichtshalber unter Vernachlässigung des Kriteriums des inhaltlichen Bezugs ausschließlich auf die Identifizierbarkeit ab und geht daher davon aus, dass mit Ausnahme reiner Geländemodelle (ohne Abbildung von Gebäuden, Grundstücksgrenzen, etc.) jedes hinreichend lagegenaue Geodatum Personenbezug habe. Da aber auch schwierig zu beurteilen ist, ab welcher Genauigkeit die Bestimmbarkeit gegeben ist, sei vorsichtshalber bei allen Geodaten ein Personenbezug zu vermuten (IMAGI 2014, S. 8, A5). Zugleich wird aber empfohlen davon auszugehen, dass bei Daten mit größerem Raumbezug das öffentliche Interesse an den Daten die schützenswerten Interessen der Betroffenen im Regelfall überwiegt (und zwar unabhängig davon, ob die Maßstäbe des IFG oder des UIG/GeoZG anzulegen sind), so dass einer Weitergabe beziehungsweise öffentlichen Bereitstellung der Daten nichts entgegen stehe (IMAGI 2014, S. 11–13, A5). Die Verwaltungspraxis vermeidet also mit Hilfe einer argumentativen Krücke die Bewertung im Einzelfall und kann daher für den Umgang mit Geodaten, die im Rahmen von Forschungsprojekten entstehen keine Orientierung bieten.

Diskussion und Zwischenfazit

Die praxisbezogene Literatur zum Datenschutz in der Forschung behandelt schwerpunktmäßig Themen, die in den Erd- und Umweltwissenschaften kaum eine Rolle spielen. Dem Umgang mit „regulären“ (nicht den „besonderen Kategorien“ zuzurechnenden) personenbezogenen Daten wird vergleichsweise wenig Raum gewidmet. Zudem sind die für die Erd- und Umweltwissenschaften besonders wichtigen Geodaten in der juristischen Fachliteratur zum Beispiel geworden, an dem rechtsdogmatische Differenzen ausgetragen werden. Dies erschwert Laiinnen die Rezeption der Rechtslage. Die praxisorientierte Literatur sowie die Rechts- und Verwaltungspraxis beim Thema Datenschutz und Geodaten hat sich bisher vor allem mit systematisch und oft duldungspflichtig erhobenen Verwaltungsdaten, ihrer massenhaften Bereitstellung für die Öffentlichkeit und ihrer Nachnutzung in der Wirtschaft einerseits sowie mit von kommerziellen Anbietern öffentlich bereitgestellten Satellitenbildern und Lichtbildern von Wohnhausfassaden andererseits befasst. Aus dieser Konstellation abgeleitete implizite Annahmen prägen die Argumentationsweise ebenso wie Vermutungen über den datenschutzrechtlichen Status einzelner Arten von Datensätzen. Am deutlichsten wird dies in den pauschalen und sehr weitgehenden Vermutungen über den Personenbezug bestimmter Arten von behördlichen Geodaten, die der IMAGI-Leitfaden anstellt (IMAGI 2014). Aber auch in den Gutachten von Karg (2008) sowie Forgó, Krügel und Reiners (2008) werden einige für die Forschung relevante Arten von Geodaten gar nicht, andere übermäßig pauschal behandelt. Letztes gilt insbesondere für so genannte Punktdaten, denen unabhängig von ihrem Inhalt grundsätzlich Identifizierbarkeit und eine hohe Aussagekraft zugeschrieben wird.

Für den Umgang mit im Rahmen von Forschung erhobenen Geodaten und ihre Bereitstellung zur Nachnutzung bietet die Literatur daher kaum Anknüpfungspunkte. Nichtsdestotrotz liegen Vorschläge für Prüfverfahren zum Personenbezug von Geodaten von Karg (2008) sowie Forgó, Krügel und Reiners (2008) vor, die trotz Verankerung in verschiedenen rechtsdogmatischen Schulen zu praktisch sehr ähnlichen Prüfabläufen kommen. An diese könnten zu entwickelnde praxisorientierte Handreichungen zum Umgang mit Geodaten aus der Forschung anschließen, mit Hilfe derer offensichtlich unproblematische Fälle, offensichtlich problematische Fälle und Zweifelsfälle vergleichsweise einfach unterschieden werden könnten. Dies könnte zur Verbesserung der Praxis im Umgang mit potenziell personenbezogenen Geodaten in der Forschung beitragen, ohne dass für allzu viele Fälle der Aufwand einer vollständigen juristischen Prüfung nötig wird.

Vertrauliche Daten in der Forschungspraxis der Erd- und Umweltwissenschaften

Ziel des empirischen Teils dieser Untersuchung war es, relevante Fallgruppen von potenziell personenbezogenen Daten in den Erd- und Umweltwissenschaften und von offenen Fragen im Umgang mit diesen zu ermitteln. Dazu wurden leitfadengestützte, auf die Erhebung von Prozesswissen zielende Gespräche mit Expertinnen geführt (fundierende, informatorische Experteninterviews im Sinne Bogners 2014, S. 22-25). Deren relevante Inhalte wurden im Sinne einer qualitativen Inhaltsanalyse zusammengefasst. Wie für Vorstudien dieser Art von Bogner (2014,

S. 39-41) empfohlen, wurde direkt aus der Tonaufzeichnung exzerpiert. Dabei wurde jedes Vorkommen von als problematisch empfundenen Daten als ein Fall behandelt. Zur Gruppierung der Fälle wurde das Verfahren der qualitativen Typenbildung (Keller und Kluge 2010) angewendet.

Die Auswahl der Gesprächspartnerinnen erfolgte kriterienorientiert nach maximaler Variation in den drei Merkmalen „Art der Organisation“ (Hochschule oder außeruniversitäre Forschungseinrichtung), „Rolle der befragten Person“ (Forschung oder Forschungsunterstützung) und „Wissenschaftsgebiet“ (Landschafts- und standortbezogene Forschung in den DFG-Fachgebieten 34 Geowissenschaften und 23 Agrar- und Forstwissenschaften; andere Arten von Forschung im DFG-Fachgebiet 34; oder ökologische und naturschutzbezogene Forschung im DFG-Fachgebiet 21 Biologie). Potenzielle Interviewpartnerinnen wurden über eine systematische Suche auf den Websites relevanter universitärer Institute und außeruniversitärer Forschungseinrichtungen in Nord- und Ostdeutschland gefunden. Die Rekrutierung von Arbeitsgruppenleiterinnen aus der Forschung als Gesprächspartnerinnen gestaltete sich schwieriger, während Kolleginnen mit direktem Bezug zu Infrastruktur-Einheiten meist gesprächsbereit waren und zwar auch dann, wenn sie selbst (auch) Forschung betreiben.

Im Rahmen dieser Vorstudie konnten Gespräche mit insgesamt sechs Expertinnen geführt werden, die teils aus mehreren Positionen berichten konnten, so dass die Bereiche universitäre Arbeitsgruppe und außeruniversitäre Arbeitsgruppe je zwei Mal, der Bereich außeruniversitäre Infrastruktur vier Mal abgedeckt war. Dabei war die landschaftsbezogene Forschung drei Mal, die Geowissenschaften im engeren Sinne zwei Mal sowie Ökologie und Naturschutz ein Mal vertreten.

Die Interviews zeigen, dass von Erd- und Umweltwissenschaftlerinnen als (zumindest potenziell) vertraulich bewertete Daten sehr divers sind. In den Interviews lassen sich insgesamt 20 Fallgruppen vertraulicher Daten in den Erd- und Umweltwissenschaften identifizieren. Von diesen haben elf Bezug zur Fragestellung dieser Untersuchung, da personenbezogene Daten betroffen sein können. Die überwiegende Zahl dieser in die Typologie aufgenommenen Fallgruppen kommen aus der landschaftsbezogenen Forschung, unabhängig von ihrer genauen disziplinären Verortung. Aus dem Bereich der Geowissenschaften trägt vor allem die sich ebenfalls auf die menschliche Nutzung gerichtete Forschung zu Georisiken und Naturgefahren zur Konstituierung der Typen bei. Ökologie und Naturschutz sind ebenfalls nur dann betroffen, wenn es um den Zusammenhang mit Landnutzung geht.

Aus diesen elf relevanten Fallgruppen ließen sich analytisch sechs (Interview-übergreifende) Typen ableiten. Dabei teilen jeweils zwei Typen viele gemeinsame Eigenschaften, so dass sich eine Darstellung in drei Typgruppen (hier identifiziert durch Großbuchstaben A, B, C) mit je zwei Typen anbietet. Da die Typologie empirisch basiert ist (die Daten also so einteilt, wie eine idealisierte Interviewpartnerin es hätte tun können), sind die Kriterien zur Ausgrenzung der Typen nicht über die gesamte Typologie einheitlich. So wird als Kriterium teils die Art der Identifizierbarkeit, teils der Informationsgehalt der Daten und teils die methodische Provenienz der Daten herangezogen. Im Folgenden werden zunächst Ergebnisse dargestellt, die über die einzelnen konstruierten Typen hinweg für alle der Typologie zugeordneten Fallgruppen Gültigkeit besitzen. Anschließend werden die einzelnen Typen vorgestellt.

In den verbleibenden neun Fallgruppen geht es um andere Arten von Vertraulichkeit von Daten als (potenziellen) Personenbezug beziehungsweise berechtigte Interesse Betroffener. Die aus

ihnen gebildeten Typen werden im Anschluss an die Vorstellung der Typologie nur mit kurzer Beschreibung gelistet, eine ausführliche Analyse liegt außerhalb des Untersuchungsbereichs dieses Beitrags.

Typen-übergreifende Ergebnisse

Über alle sechs konstruierten Typen hinweg lässt sich feststellen, dass einige Befragte große Unsicherheit darüber äußern, welche Daten schutzwürdig sind, während andere sich in ihren Einschätzungen weitgehend sicher fühlen. Verschiedene Befragte, die sich in ihren Einschätzungen relativ sicher sind, kommen aber zu durchaus unterschiedlichen Beurteilungen des Status vergleichbarer Daten.

Die Wahrnehmung von Daten als vertraulich wird von den Befragten allerdings meist nicht mit einer Argumentation begründet, die sich im engeren Sinne auf das Persönlichkeitsrecht bezieht. Dies zeigt sich auch daran, dass Betriebe und teils auch Personengruppen nicht notwendigerweise als weniger schutzwürdig angesehen werden als Einzelpersonen. Auch die Frage der Identifizierbarkeit spielt für die Befragten nicht die Hauptrolle. Vielmehr halten die Befragten Daten vor allem dann für vertraulich, wenn eine nicht-intendierte außerwissenschaftliche Nutzung das Potenzial hat, betroffenen Dritten (insbesondere Eigentümerinnen beziehungsweise Landnutzerinnen) zu schaden. Dafür werden zwei mögliche Arten von Szenarien angegeben:

- Die Daten könnten zeigen, dass die Nutzungs- und Bewirtschaftungspraxen von Landnutzerinnen nicht der guten Praxis folgen oder sogar unzulässig sind.
- Die Daten könnten, auch ohne dass ein Bezug zum Verhalten der aktuellen Nutzerinnen besteht, negative Auswirkungen auf wirtschaftliche Interessen der Betroffenen haben (zum Beispiel auf Grund von Naturgefahren oder bestehenden Umweltbelastungen).

Die Wahrscheinlichkeit einer Rezeption dieser Daten außerhalb der Wissenschaft wird jedoch als in den meisten Fällen sehr gering eingeschätzt. In die unterschiedlichen Einschätzungen der Vertraulichkeit von Daten scheint auch einzufließen, inwieweit die Forschenden etwaige Landnutzerinnen als Forschungssubjekte sehen und sich ihnen daher forschungsethisch verpflichtet fühlen oder ob diese gar Kooperationspartner sind, denen sich die Forschenden nicht nur verpflichtet fühlen, sondern von denen sie auch abhängig sind.

Daten, bei denen ein solches vornehmlich wirtschaftliches Schadenspotenzial nicht gesehen wird, scheinen oft als grundsätzlich harmlos angesehen zu werden. Die – aus Perspektive der Sozialwissenschaften selbstverständlich erscheinenden – persönlichkeits- und datenschutzrechtlichen Herausforderungen bei der Verarbeitung von Daten, die aus anderen Gründen personenbezogen sind, werden daher teils gar nicht erkannt. Gute Beispiele sind hier die Veröffentlichung von Transkripten qualitativer Interviews oder der Umgang mit Daten, die Auskunft über das Wohnumfeld bestimmbarer Personen geben.

In den Antworten auf die Frage, wie mit potenziell problematischen Daten dieser Art umgegangen wird, fallen folgende Aspekte auf:

- Die Befragten äußern sich vor allem zur Veröffentlichung von Daten. Ob die Daten überhaupt hätten erhoben und verarbeitet werden dürfen, spielt dagegen seltener eine Rolle –

wohl auch, weil das Thema erst im Umfeld von Datenveröffentlichungen virulent geworden ist.

- Es bestehen Unsicherheiten und Defizite im Bereich der Einwilligung. Eine solche wird in der Regel eingeholt, allerdings erfolgt diese oft mündlich, die Qualität schriftlicher Einverständniserklärungen ist uneinheitlich und eine Nachnutzung meist nicht mit abgedeckt.
- Auch wo Daten als vertraulich eingeschätzt werden, wird von technischen und organisatorischen Maßnahmen der Datensicherheit (wie Zugangsbeschränkung, Pseudonymisierung beziehungsweise formale Anonymisierung oder Verschlüsselung) bei der Datenverarbeitung in der Arbeitsgruppe während des Forschungsprozesses nicht berichtet.
- Die Befragten an Universitäten berichten, keine Möglichkeit zur institutionellen Archivierung vertraulicher Daten zu haben beziehungsweise zu kennen. An den außeruniversitären Forschungseinrichtungen gibt es zusätzlich zum öffentlichen Repositorium teils interne Datenbanken, in denen auch vertrauliche Forschungsdaten abgelegt werden. Die in solchen Fällen übliche Zugangsbeschränkung ist, dass die ursprüngliche Projektverantwortliche beziehungsweise AG Leiterin darüber entscheidet, wem auf Anfrage Zugang gewährt wird. Regelungen für den Fall, dass diese nicht mehr verfügbar ist, scheinen nicht üblich zu sein. Gelegentlich werden gerade seit längerem vorhandene („legacy“) Datensammlungen in solchen internen Repositorien als der Einrichtung gehörend verstanden und die Zugangsgewährung erfolgt durch die Leitung des Bereichs.
- In den Aussagen zur Datenveröffentlichung werden meist nur die Optionen Open-Access-Veröffentlichung versus komplettes Zurückhalten der Daten thematisiert. Als mögliche Zugangsbeschränkung für veröffentlichte Daten werden nur Embargo-Perioden und Lizenzverträge für kommerzielle Daten genannt, wissenschaftsfreundliche Modelle für den beschränkten Zugang zu Daten (wie sie etwa an den RatSWD-zertifizierten Datenzentren üblich sind) sind weitgehend unbekannt.
- Befragte berichten von der Erfahrung, dass im Bereich Open Science engagierte Editorinnen und Gutachterinnen sich in Einzelfällen mit Verweis auf die sich entwickelnde disziplinäre *best practice* teils auch dann schwer tun, Fachartikel ohne Open-Access-Veröffentlichung der zugehörigen Daten zu akzeptieren, wenn diese Daten aus Sicht der Autorinnen begründbar vertraulich sind. Gleichzeitig sind aber sowohl Journals als auch Förderorganisationen in der Regel auch mit der Publikation von Daten zufrieden, die so stark aufbereitet sind, dass sie gar nicht mehr sinnvoll nachnutzbar sind.

Typgruppe A: Abschätzbare Risiken bei Geo- und Umweltdaten aus naturwissenschaftlicher Feldforschung

Ein Großteil der potenziell zum Nachteil der Betroffenen identifizierbaren Daten, von denen die Befragten berichten, stammt aus naturwissenschaftlicher Feldforschung. Es geht also um Daten, die entweder direkt im Gelände erhoben wurden oder aus Analyse im Feld genommener Proben stammen. Dabei kann es sich um sehr verschiedene Daten handeln, fast das gesamte Methodenspektrum der Erd- und Umweltwissenschaften ist betroffen: Es geht sowohl um experimentelle Daten (aus Feldversuchen) als auch um Beobachtungsdaten. Es kann sich um qualitative Daten (Zuordnungen zu Klassen wie Bodentypen oder pflanzensoziologische Einheiten), ebenso

handeln wie um quantitative Daten (von Zählungen von Wildtiersichtungen über Messungen physikalischer und chemischer Parameter in einzelnen Umweltkompartimenten bis hin zu Erkundungen von Oberfläche und Untergrund mit geophysikalischen Methoden). Zeitlich kann es sich um Einzelerhebungen (Ergebnisse von Einzelaufnahmen beziehungsweise von einmaligen Mess- beziehungsweise Beobachtungskampagnen) oder um Zeitreihen (zeitlich diskrete longitudinale Erhebungen oder quasi-kontinuierliche Daten aus automatisierten Messstationen, so genannten Datenloggern) handeln.

Die Befragten bestätigen, dass es für viele dieser Daten plausible Nachnutzungsszenarien gibt. Der zunehmende Druck seitens der Fördermittelgeber, möglichst alle in Projekten beziehungsweise über geförderte Infrastrukturen erhobene Daten zu veröffentlichen, sowie die Forderung renommierter Zeitschriften, zumindest die direkt zur Textpublikation gehörenden Daten zu veröffentlichen, wird daher insgesamt positiv beurteilt. Im Zuge der weiteren Verbreitung, des Ablegens von Daten in Repositorien und des Aufbaus neuer fachbezogener Dateninfrastrukturen, wird aber durchaus auch wahrgenommen, dass sich diese Daten auf verschiedene Art und Weise in Bezug zur Eigentümerin beziehungsweise Nutzerin der betroffenen Fläche setzen lassen und dass dies in bestimmten Fällen eine gewisse Vertraulichkeit der Daten begründen sollte. Da es sich um Daten aus Feldforschung handelt, die eine begrenzte Zahl von Parametern aufnehmen, erscheint eine Abschätzung des Risikos in der Regel dabei in erster Näherung möglich.

Potenziell identifizierbare und gegebenenfalls vertrauliche Geo- und Umweltdaten aus naturwissenschaftlicher Feldforschung unterscheiden sich wesentlich in der Art und Weise der möglichen (Re-)Identifikation und den jeweiligen Möglichkeiten der Anonymisierung. **Typ A1** fasst durch genauen Raumbezug (Koordinaten oder bestimmte Fläche) identifizierende Daten zusammen. Es handelt es sich im Wesentlichen um Daten, die in Geoinformatik und Kartographie als thematische Daten, im geodatenbezogenen Behördendeutsch als Fachdaten bezeichnet werden. **Typ A2** beschreibt durch Inhalt und Kontext des Datensatzes identifizierbare Daten. Dies ist insbesondere dann der Fall, wenn die untersuchten Standorte Teil einer kleinen Grundgesamtheit sind. Eine Identifizierbarkeit kann in diesem Fall nicht nur aus veröffentlichten Daten erwachsen, sondern sich bereits aus der (Text-)Publikation von Ergebnissen ergeben. Wo es nötig und möglich ist, eine (Re-)Identifikation durch Anonymisierungsmaßnahmen zu verhindern, schränken diese für beide Typen die Nachnutzbarkeit der Daten in der Regel massiv ein.

Typgruppe B: Daten mit schwer abzuschätzenden Risiken

Bestimmte Daten wurden nur von einem Teil der Befragten als problematisch identifiziert, bereiten aus deren Sicht jedoch besondere Probleme. Dabei geht es um Daten, bei denen nur schwer einzuschätzen ist, welche Informationen sich aus ihnen ableiten lassen und das Risiko sich daher kaum bestimmen lässt.

In **Typ B1** geht es um zusätzliche Risiken, die sich durch möglich werdende Verknüpfungen von vielen, für sich genommen zunächst unproblematisch erscheinenden Datensätzen in einem zukünftigen Open Data-Regime in den Erd- und Umweltwissenschaften ergeben können. In **Typ B2** geht es dagegen um Daten, die auf Grund der Erhebungsmethode einen extrem hohen Informationsgehalt haben, so dass vorab kaum geklärt werden kann, welche Angaben sich aus den Daten ableiten lassen, aber vermutet werden muss, dass sich evtl. problematische

Rückschlüsse ziehen lassen. Dies betrifft insbesondere selbst erhobene kleinräumige Fernerkundungsdaten, gegebenenfalls aber auch hochauflösende beobachtende Zeitreihen von terrestrisch erhobenen Messwerten. Gleichzeitig besteht in beiden Fällen auf Grund des hohen Informationsgehalts besonderes Interesse an einer Verfügbarkeit der Daten für die Nachnutzung.

Typgruppe C: Daten aus nicht-naturwissenschaftlichen Methoden

In allen Bereichen der Erd- und Umweltwissenschaften kommen nach Auskunft der Befragten immer wieder auch nicht-naturwissenschaftliche Methoden zum Einsatz. Bei den so generierten Daten handelt es sich teilweise um „klassische“ personenbezogene Daten, wie sie aus den Wirtschafts- und Sozial- und Kommunikationswissenschaften bekannt sind. In den Erd- und Umweltwissenschaften werden auch innerhalb hauptsächlich naturwissenschaftlich ausgerichteter Arbeitsgruppen Befragungen durchgeführt, Dokumente analysiert, konventionelle und soziale Medien ausgewertet sowie statistische Wirtschafts- und Sozialdaten nachgenutzt oder auch betriebs-, flächen- oder gebäudebezogene quantitative Daten selbst erhoben.

Dabei umfasst **Typ C1** Wirtschafts- und Sozialdaten als Paradata, die zur Begleitung und Ergänzung naturwissenschaftlich ausgerichteter Studien erhoben werden. Entscheidend ist, dass diese Daten in Verbindung mit den oben diskutierten (raumbezogenen) naturwissenschaftlichen Daten stehen, die ihrerseits schlecht zu anonymisieren sind und daher identifizierend wirken können. Diese zur Kontextualisierung naturwissenschaftlicher Untersuchungen erhobenen Daten haben oft kein eigenständiges Nachnutzungspotenzial. Dagegen steht **Typ C2** für eigenständige wirtschafts-, sozial- und kommunikationswissenschaftliche Daten. Auch die Fragestellungen sind dann, trotz des naturwissenschaftlichen Kontexts, im weiteren Sinn wirtschafts-, sozial- und kommunikationswissenschaftlich geprägt. Oft wird für solche Untersuchungen entweder mit Gruppen aus den betreffenden Disziplinen kooperiert oder Personen mit entsprechender Expertise eingestellt. Hier hängt das Nachnutzungspotenzial von der Art der erhobenen Daten ab.

Weitere Problemfelder

Neben den in die Typologie aufgenommenen Fallgruppen, bei denen Rechte und Interessen Dritter betroffen sind, wurden auch andere Konstellationen genannt, in denen Daten als vertraulich eingeschätzt werden. Da diese Punkte die Fragestellung dieser Arbeit nur am Rande berühren, werden Sie im folgenden nur kurz genannt und in der weiteren Diskussion nicht behandelt.

Unerwünschte außerwissenschaftliche Nachnutzung raumbezogener Geo- und Umweltdaten kann in Einzelfällen nicht nur Personen betreffen, sondern auch *Gefahr für Umwelt und Natur* bergen. Dies gilt insbesondere für Daten aus der Naturschutzökologie, zum Beispiel zu Vorkommen geschützter Arten. Manche Daten, insbesondere aus den Geowissenschaften, haben *unmittelbar außerwissenschaftliche, insbesondere wirtschaftliche Bedeutung*. Dies kann die erhebenden Einrichtungen vor Zielkonflikte (Verwertung vs. offene Daten) stellen oder ein Risiko für die Reputation der Disziplin beziehungsweise der Einrichtung darstellen (zum Beispiel in Konflikten um Ressourcen oder Landnutzung). Werden Daten veröffentlicht, geht dies zwangsläufig

mit dem *Verlust der Kontrolle* über ihre Interpretation einher. Während dies im wissenschaftlichen Diskurs zumindest für Messdaten explizit gewünscht ist, wird dies bei außerwissenschaftlicher Nachnutzung gerade für komplexe Datenprodukte teils als Problem gesehen. Insbesondere für räumlich explizite Modelle und bei der Erstellung räumlich expliziter Datenprodukte wie digitalen Karten werden oft Daten aus diversen Quellen zusammengeführt und verarbeitet, die sehr unterschiedlichen Lizenzbedingungen unterliegen. Dies kann die *Lizenzierung von Datenprodukten zur Nachnutzung* erschweren. Bei *Daten aus Industriekooperation* sperren sich die privaten Partnerinnen oft dagegen, sich vorab zu verpflichten, alle im Rahmen eines Projekts erhobenen Daten zu veröffentlichen und zwar auch dann, wenn diese offensichtlich keine Betriebsgeheimnisse enthalten.

Diskussion und Zwischenfazit

Die Ergebnisse des empirischen Teils bestätigen, dass auch in den – im Diskurs um personenbezogene Forschungsdaten weithin als unverdächtig geltenden – Erd- und Umweltwissenschaften tatsächlich Daten anfallen, die Personen zugeordnet werden können, zum Beispiel Eigentümerinnen und Nutzerinnen von Flächen. Insofern solche identifizierbaren Daten auch im datenschutzrechtlichen Sinne Personenbezug haben, wird es sich in der Regel um gewöhnliche personenbezogene Daten handeln. In Einzelfällen können aber durchaus „besondere Kategorien“ personenbezogener Daten nach Art. 9 DS-GVO anfallen und /oder weitere Rechtsgebiete betroffen sein, sei es bei Ortungsdaten von Landmaschinen in Forschung zur Präzisionslandwirtschaft (Arbeitsrecht) oder bei Fotos aus Citizen Science-Projekten (Urheberrecht).

Besonders betroffen sind die Teildisziplinen, die landschafts- und landnutzungsbezogene Forschung betreiben. In anderen Teilbereichen der Erd- und Umweltwissenschaften sind Daten eher aus anderen Gründen vertraulich. Diese anderen Fälle sind für die fachspezifische Unterstützung im Bereich Forschungsdatenmanagement durchaus relevant und sollten in eventuellen fachspezifischen Handreichungen berücksichtigt werden, können hier jedoch nicht weiter diskutiert werden.

Die Typen-übergreifenden Ergebnisse zeigen, dass die Forschenden den Personenbezug ihrer Daten selten in Isolation in Betracht ziehen. Vielmehr trennen sie den Personenbezug in ihren impliziten forschungsethischen Abwägungen und in ihren durch das Interview veranlassten Explikationen nicht von anderen berechtigten Interessen Dritter. Die Interessen von Personengruppen und juristischen Personen von den Befragten nicht als grundsätzlich weniger schützenswert angesehen als die Interessen von Einzelpersonen. Zudem machen die Befragten den Umfang des als schützenswert angesehenen Bereichs auch vom Verhältnis zwischen Forschenden und Betroffenen abhängig. Beides ist forschungsethisch einleuchtend, geht aber über eine rein datenschutzrechtliche Betrachtung hinaus. Zugleich werden jedoch andere Dimensionen von Persönlichkeitsrelevanz, etwa dass auch erd- und umweltwissenschaftliche Daten in Einzelfällen zum Beispiel Auskunft über das private Wohnumfeld geben könnten, von den Befragten oft nicht gesehen. Die Überlegungen der Befragten zur Vertraulichkeit von Daten erfolgen vor allem im Kontext von Datenpublikationen, die Frage nach den Anforderungen an ihre Erhebung und Bedingungen ihrer Verarbeitung stellt sich ihnen seltener. Wo Daten tatsächlich personenbezogen sind, wird die Forschungspraxis den strengen Anforderungen an Information der Betroffenen, Einverständniserklärungen und Grundsätze der Verarbeitung wohl in der Regel nicht gerecht.

Es besteht also einerseits über den engen Geltungsbereich des Datenschutzrechtes hinaus Klärungsbedarf im Bereich von Forschungs-, Publikations- und Datenethik in den Erd- und Umweltwissenschaften. Die Frage, wie hier eine gute Praxis im Umgang mit vertraulichen Daten entwickelt werden kann, lässt sich nicht auf das Datenschutzrecht und technisch-organisatorische Maßnahmen der Datensicherheit reduzieren. Andererseits besteht im Hinblick auf die Anforderungen des Datenschutzes sowohl ein Sensibilitäts- und Informationsdefizit auf Seiten der Forschenden als auch ein Defizit an Beratungskompetenz und Infrastrukturen bei den disziplinären Datenzentren.

Für verschiedene Typen von Daten sind weitere typspezifische Ergebnisse relevant. Zunächst ist festzuhalten, dass teils auch im Rahmen vornehmlich naturwissenschaftlich ausgerichteter Studien «ganz klassische» personenbezogene Daten beziehungsweise Daten, welche auch ohne Ansehen ihres Raumbezugs die berechtigten Interessen Dritter betreffen, erhoben werden (Typgruppe C). Während betriebswirtschaftliche Daten klar als vertraulich eingeschätzt werden, scheint der Umgang mit den mittels sozialwissenschaftlicher Methoden gewonnenen Daten uneinheitlich und insbesondere bei den nicht einfach zu anonymisierenden qualitativen Daten nicht immer angemessen zu sein.

Ein weiteres Cluster von Antworten bezieht sich auf Daten, deren Risiko besonders schlecht einzuschätzen ist (Typgruppe B). Hier äußern Forschende Unverständnis, dass sie alleine für die Unbedenklichkeit der von ihnen bei Repositorien abgelieferten Daten verantwortlich sein sollen. Gleichzeitig ist auch bei den Infrastrukturen beschäftigten Personen unklar, wie sie bei einer solchen Beurteilung unterstützen sollten. Es wird einerseits zugestanden, dass diese Daten nicht im Open Access veröffentlicht werden sollten und andererseits beklagt, dass anscheinend nur in Ausnahmefällen eine nachnutzbare Archivierung und Zugänglichmachung erfolgt.

Der größte Komplex von genannten Fällen bezieht sich jedoch auf die bei einzelnen Datensätzen gegebenenfalls in Kombination mit einfach zu ermittelnden Adress- und Katasterdaten vorliegenden Risiken bei Geo- und Umweltdaten aus naturwissenschaftlicher Feldforschung (Typgruppe A). Hier unterscheiden die Befragten klar zwei Typen: Daten, bei denen formale Anonymisierung – also die Entfernung direkter Identifikatoren einschließlich exaktem Raumbezug – ausreicht, um faktische Anonymität sicherzustellen (Typ A1) und Daten, die auch bei formaler Anonymisierung identifizierbar bleiben (Typ A2). Dabei wird jedoch betont, dass viele Daten in anonymisierter Form kaum noch nutzbar sind. Dies gilt auch für rein formale Anonymisierung, da Daten in der Regel nur im Kontext interpretierbar sind. Darüber, welche dieser Daten überhaupt schützenswert sind, besteht jedoch keine Einigkeit zwischen den Befragten und bei manchen Befragten auch deutliche individuelle Unsicherheit. Im Wesentlichen ist strittig, in welchem Umfang ein Datum Rückschlüsse auf die wirtschaftliche Lage oder Details des Verhaltens einer Nutzerin oder der Eigentümerin zulassen muss beziehungsweise wie viel Potenzial, den Wert eines Grundstücks zu beeinflussen, das Datum haben muss, um personenbezogen zu sein oder – falls Personenmehrheiten oder juristische Personen betroffen sind – aus forschungsethischen Gründen als vertraulich zu gelten. Damit besteht in der forschungsethischen Bewertung durch die Akteure Unklarheit an demselben Punkt, der auch in der juristischen Fachliteratur strittig ist.

Einige Befragte beschreiben die begrenzte Nachnutzbarkeit weit aufbereiteter Daten, wie sie oft in Supplementen enthalten sind. Dies gilt verstärkt für Daten, die anonymisiert wurden. Archivierung bei Infrastrukturen mit Recherchierbarkeit der Metadaten und *scientific use* (begrenzter

Zugriff auf den vollständigen Datensatz zu geprüften Forschungszwecken nur nach Abschließen eines Lizenzvertrags) gibt es im Bereich der Erd- und Umweltwissenschaften bisher aber fast ausschließlich für (in Bezug auf Rohstoffe) wirtschaftlich relevante Daten aus den Geowissenschaften der festen Erde.

Insbesondere an Hochschulen scheint es oft keine Möglichkeit zum institutionellen Archivieren vertraulicher Daten zu geben, da in der Regel die Unbedenklichkeit der Daten bei Übergabe an einen Archivierungsdienst bescheinigt werden muss. Werden Daten zur Sicherung der guten wissenschaftlichen Praxis oder zur Nachnutzung innerhalb einer Arbeitsgruppe aufgehoben, aktuell aber nicht benötigt, ist jedoch gerade für vertrauliche Daten eine sichere institutionelle Datenhaltung statt Datenhaltung im Workspace der Arbeitsgruppe angezeigt.

Handlungsbedarf

Aus den Ergebnissen der Literaturlauswertung zum Datenschutz in der Forschung, insbesondere bei Geodaten, und den Ergebnissen der Vorstudie zum Umgang mit potenziell personenbezogenen Daten in den Erd- und Umweltwissenschaften ergibt sich ein Bedarf, die Datenpraxis im Rahmen der Disziplinlkultur zu verbessern. Vorrangig ist dies natürlich eine Aufgabe der Fächer, zum Beispiel im Rahmen wissenschaftsgeleiteter Weiterentwicklungsprozesse wie der Nationalen Forschungsdateninfrastruktur (NFDI), in der Datenschutz und Datenethik als „cross-cutting issue“ behandelt werden sollen.² Die Ergebnisse zeigen jedoch auch mehrere Handlungsfelder für Infrastruktureinrichtungen auf und legen bestimmte Handlungsoptionen nahe.

In folgenden drei Themenbereichen ist **Beratungsbedarf der Forschenden** sichtbar geworden, den Infrastruktureinrichtungen adressieren sollten:

- Einschätzung des Personenbezugs beziehungsweise der Vertraulichkeit auf Grund berechtigter Interessen Dritter bei im Rahmen von Forschungsvorhaben erhobenen Geodaten.
- Umgang mit wirtschafts- und sozialwissenschaftlichen Paradata zu naturwissenschaftlichen Untersuchungen.
- Einhaltung der Anforderungen an die Erhebung und Verarbeitung personenbezogener Daten in Forschungsprojekten (nicht erst an Weitergabe und Veröffentlichung).

Der Umfang der relevanten Informationen legt nahe, hier eine eigenständige Handreichung zu erarbeiten. Der geeignete Moment im Datenlebenszyklus, um diese Themen frühzeitig zu berücksichtigen, ist die Projektplanung. Daher sollte diese Handreichung mit vorliegenden und entstehenden disziplinspezifischen Informationen zu Datenmanagementplänen abgestimmt werden. Zur Einschätzung des Personenbezugs bei Geodaten aus der Forschung bietet es sich an, eine Anleitung zur Risikoabschätzung in Anlehnung an die existierenden, mit Blick auf Behördenaten entwickelten Prüfverfahren zu entwickeln.

Im Bereich der Herstellung von **Archivierbarkeit personenbezogener und anderweitig vertraulicher Daten in institutionellen Datenrepositorien** käme jede entsprechende Verbesserung nicht nur den Erd- und Umweltwissenschaften zu Gute, sondern allen Disziplinen, die nicht

²Vergleiche die unter https://www.dfg.de/en/research_funding/programmes/nfdi/ zur Verfügung gestellten Unterlagen und Glöckner *et al.* (2019).

alle ihre vertraulichen Daten bei fachbezogenen Infrastrukturen abgeben können. Hier sind zunächst die institutionellen Datenarchive beziehungsweise Repositorien insbesondere der Hochschulen angesprochen. Sollte sich herausstellen, dass gute Praxis im Umgang mit vertraulichen Daten in institutionellen Datenarchiven und Repositorien (zumindest an Hochschulen) tatsächlich um ein bundesweites Desiderat und kein Problem einzelner Einrichtungen handelt, würde sich ein entsprechende Forschungs- und Entwicklungsprojekt an einer gut aufgestellten Einrichtung anbieten. Wünschenswert wäre es dabei, die besonderen Bedingungen bei kooperativen und gehosteten Diensten wie RADAR mit zu betrachten.³

Wenn es in Zukunft möglich sein soll, auch **personenbezogene und anderweitig vertrauliche Daten aus den Erd- und Umweltwissenschaften zu publizieren**, statt nur institutionell zu archivieren, das heißt die Daten auffindbar und für *scientific use* begrenzt zugänglich zu machen, wäre es notwendig, bei disziplinären Datenzentren entsprechende Modi des Zugangs als Standard-Dienstleistung mit definierten Zugangsbedingungen zu etablieren. Dies setzt jedoch nicht nur die Entwicklung von Workflows, sondern auch den Aufbau von Expertise und die Entwicklung von disziplinspezifischen Richtlinien zu vertraulichen Daten und Datenschutz, effektiver Anonymisierung und vertraglichen Bedingungen der Zugangsgewährung voraus.

Die **Risiken der Kombination von Datensätzen** aus den Erd-, Umwelt- und Agrarwissenschaften sind zurzeit auch für Datenexpertinnen bezüglich der Infrastrukturen kaum überschaubar. Es besteht Klärungsbedarf dazu, welche inhaltlichen Rückschlüsse über den Informationsgehalt der Einzeldatensätze hinausgehenden und aus der Kombination vieler identifizierbarer, aber als Einzeldatensatz harmloser Forschungsdaten mit den offenen Verwaltungsdaten im Geo-, Umwelt- und Agrarförderungsbereich sowie anderen leicht verfügbaren Daten (Adressverzeichnisse, Kataster) tatsächlich getroffen werden können. Es liegt nahe, dieses Thema im Zusammenhang mit Projekten zur Schaffung von Strukturen für *scientific use* von Forschungsdaten der Erd- und Umweltwissenschaften zu bearbeiten.

Fazit und Ausblick

Am Beispiel der Erd- und Umweltwissenschaften habe ich aufgezeigt, dass auch in scheinbar „unverdächtigen“ Disziplinen personenbezogene Forschungsdaten vorkommen. Die Ergebnisse der empirischen Vorstudie zeigen eine ganze Reihe verschiedener Arten personenbezogener Forschungsdaten auf, die in der Forschungspraxis der Erd- und Umweltwissenschaften eine Rolle spielen. Sie legen außerdem nahe, dass der Umgang mit ihnen in der Forschungspraxis auf Grund der mangelnden Vertrautheit mit dem Datenschutz nicht immer den rechtlichen Anforderungen entspricht. Auch Unterstützung in Form disziplinspezifischer Handreichungen, qualifizierter Beratung oder institutionalisierte Möglichkeiten, Daten sicher zu archivieren und gegebenenfalls zugangsbeschränkt zu publizieren, bestehen kaum. Zudem ergibt sich aus der Auswertung der Literatur, dass bisherige Antworten auf die Frage, wann Geodaten unter den Datenschutz fallen, die spezifischen Erfordernisse erd- und umweltwissenschaftlicher Forschungsdaten nicht berücksichtigen. Allgemeine Handreichungen zum Datenschutz in der Forschung geben keine Auskunft, Handreichungen zum Datenschutz bei Geodaten der Verwaltung sind nicht auf die Forschung anwendbar und selbst unter Fachjuristinnen herrscht Uneinigkeit über die Bewertung. Aus dieser Situation ergeben sich Herausforderungen an die Weiterentwicklung

³Zu RADAR vgl. <https://www.radar-service.eu/de>.

der disziplinären Datenkultur und Dateninfrastrukturen. Dabei können die Infrastruktureinrichtungen wertvolle Unterstützung leisten, wie die dargestellten Handlungsoptionen zeigen. Insbesondere die Kooperation zwischen Forschenden und Infrastruktureinrichtungen im Rahmen des NFDI-Prozesses bietet hier gute Möglichkeiten, diese Herausforderungen anzugehen.

Anmerkungen

Diese Veröffentlichung basiert auf einer Masterarbeit im weiterbildenden Masterstudiengang Bibliotheks- und Informationswissenschaft (Library and Information Science, M. A. (LIS)) im Fernstudium am Institut für Bibliotheks- und Informationswissenschaft der Humboldt-Universität zu Berlin aus dem Sommersemester 2018. Ich danke Prof. Dr. Wolfram Horstmann (Staats- und Universitätsbibliothek Göttingen und Humboldt-Universität zu Berlin) und Maxi Kindling (damals Humboldt-Universität zu Berlin) herzlich für die Betreuung der Arbeit. Angewandte Forschung dieser Art ist unmöglich ohne die „kindness of strangers“ der Interviewpartnerinnen, die sich auf ein Gespräch über potenziell problematische Aspekte ihrer Forschungspraxis beziehungsweise ihrer forschungsunterstützenden Arbeitspraxis in der Infrastruktur einließen. Auch ihnen sei an dieser Stelle herzlich gedankt.

Der Autor dieses Beitrags ist kein Jurist, dementsprechend enthält der Text *keine* verbindlichen Auskünfte zu Rechtsthemen. Diese Arbeit nutzt das generische Femininum, wo sich genderneutrale Partizipialkonstruktionen nicht anbieten. Alle Links wurden am 23.09.2019 zuletzt geprüft, soweit nicht anders vermerkt.

Angaben zur Datenverfügbarkeit: Die Aufzeichnungen der Gespräche wurden nach Abschluss der Untersuchung gelöscht. Die der Auswertung zu Grunde liegenden Exzerpte sind für Dritte nicht nachnutzbar und lassen sich nicht verlässlich anonymisieren. Sie werden daher nicht zugänglich gemacht.

Bibliographie

Artikel-29-Datenschutzgruppe (2007). *Stellungnahme 4/2007 zum Begriff „personenbezogene Daten“*. WP 136. Brüssel: Datenschutzgruppe nach Artikel 29 der Richtlinie 95/46/EG.

Bernard, Lars *et al.* (2016). „Geodateninfrastrukturen“. In: *Handbuch der Geodäsie*. Hrsg. von Willi Freeden und Reiner Rummel. Springer Reference Naturwissenschaften. Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-662-46900-2_66-1.

Bertelmann, Roland (2017). „Geowissenschaften“. In: *Praxishandbuch Open Access*. Hrsg. von Konstanze Söller und Bernhard Mittermeier. Berlin, Boston: De Gruyter. Kap. 5c, S. 261–265.

Bertelmann, Roland *et al.* (2014). *Einstieg ins Forschungsdatenmanagement in den Geowissenschaften*. Handreichung. Potsdam, Berlin: DFG-Projekt EWIG. <https://doi.org/10.2312/lis.14.01>.

BFDI (2009). *Tätigkeitsbericht zum Datenschutz für die Jahre 2007 und 2008*. 22. Tätigkeitsbericht. Berlin: Bundesbeauftragter für den Datenschutz und die Informationsfreiheit.

Blühdorn, Ingolfur (2007) Sustaining the unsustainable: Symbolic politics and the politics of simulation. *Environmental Politics* 16. 2, S. 251-275, <https://doi.org/10.1080/09644010701211759>

Bogner, Alexander (2014). *Experteninterviews: Theorien, Methoden, Anwendungsfelder*. Wiesbaden: VS Verlag für Sozialwissenschaften.

COPDESS (2015). *COPDESS Statement of Commitment: Statement of Commitment from Earth and Space Science Publishers and Data Facilities*. Coalition on Publishing Data in the Earth und Space Sciences. <http://www.copdess.org/statement-of-commitment/>.

Diepenbroek, Michael und Gerold Wefer (2011). „Forschungsdateninfrastrukturen in den Bio- und Geowissenschaften“. In: *Zeitschrift für Bibliothekswesen und Bibliographie* 58.3-4, S. 167-171. <https://doi.org/10.3196/18642950115834125>.

Diez, Dietrich *et al.* (2009). „Schutz des Persönlichkeitsrechts bei der Verarbeitung von Geodaten“. In: *zfv* 134.6, S. 357-362.

Forgó, Nikolaus und Tina Krügel (2010). „Der Personenbezug von Geodaten: Cui bono, wenn alles bestimmbar ist?“ In: *MultiMedia und Recht* 12.1, S. 17-23. <https://beck-online.beck.de/Bcid/Y-300-Z-MMR-B-2010-S-17-N-1>.

Forgó, Nikolaus, Tina Krügel und Nico Reiners (2008). *Forschungs- und Entwicklungsauftrag zum Thema Geoinformation und Datenschutz*. Rechtsgutachten. Hannover: Leibniz Universität Hannover, Institut für Rechtsinformatik im Auftrag der X GmbH. https://iri.uni-hannover.de/tl_files/pdf/Geodaten_Datenschutz_Gutachten.pdf.

Glöckner, Frank Oliver *et al.* (2019). *Berlin Declaration on NFDI Cross-Cutting Topics*. Zenodo. <http://doi.org/10.5281/zenodo.3457213>

Gusy und Eichendorfer (2018). „BDSG 2018 §1“. In: *BeckOK Datenschutzrecht*. Hrsg. von Wolff und Brink. 23. Aufl. Stand: 25.07.2017. Beck-Online, Rn. 27-50.

Hanson, Brooks, Kerstin Lehnert und Joel Cutcher-Gershenfeld (2015). „Committing to Publishing Data in the Earth and Space Sciences“. In: *EOS Earth & Space Science News* 96. <https://doi.org/10.1029/2015EO022207>.

Hilty, Lorenz *et al.* (2012). *Lokalisiert und identifiziert: Wie Ortungstechnologien unser Leben verändern*. TA-SWISS. Zürich: vdf Hochschulverlag der ETH. <https://doi.org/10.3218/3477-6>.

Hoeren, Thomas (2018). *Internetrecht*. Vorlesungsskript. Münster: Universität Münster, Institut für Informations-, Telekommunikations- und Medienrecht. <https://www.itm.nrw/lehre/materialien/>.

Hofmann, Johanna M und Paul C Johannes (2017). „DS-GVO: Anleitung zur autonomen Auslegung des Personenbezugs“. In: *Zeitschrift für Datenschutz* 7.5, S. 221-226. <https://beck-online.beck.de/Bcid/Y-300-Z-ZD-B-2017-S-221-N-1>.

IMAGI (2014). *Behördenleitfaden zum Datenschutz bei Geodaten und -diensten*. Leitfaden. Interministerieller Ausschuss für Geoinformationswesen. https://www.imagi.de/IMAGI/DE/Arbeit_sgruppen/AG-Geodatschutz/ag-geodatschutz_node.html.

Karg, Moritz (2008). *Datenschutzrechtliche Rahmenbedingungen für die Bereitstellung von Geodaten für die Wirtschaft*. Rechtsgutachten. Kiel: Unabhängiges Landeszentrum für Datenschutz Schleswig-Holstein im Auftrag der GIW-Kommission. url: <https://www.bmwi.de/Redaktion/DE/Publikationen/Geobusiness/datenschutzstudie-2-die-ampelstudie.html>.

Karg, Moritz und Thilo Weichert (2007). *Datenschutz und Geoinformation*. Rechtsgutachten. Kiel: Unabhängiges Landeszentrum für Datenschutz Schleswig-Holstein im Auftrag des Bundesministeriums für Wirtschaft und Technologie. <https://www.bmwi.de/Redaktion/DE/Publikationen/Geobusiness/datenschutzstudie-1-datenschutz-und-geoinformation.html>.

Kelle, Udo und Susann Kluge (2010 [1999]). *Vom Einzelfall zum Typus: Fallvergleich und Fallkontrastierung in der qualitativen Sozialforschung*. 2., überarb. Aufl. Bd. 15. Qualitative Sozialforschung. Wiesbaden: VS Verlag für Sozialwissenschaften.

Klump, Jens (2012). „Geowissenschaften“. In: *Langzeitarchivierung von Forschungsdaten: Eine Bestandsaufnahme*. Hrsg. von Heike Neuroth et al. Boizenburg: Verlag Werner Hülsbusch. <https://nestor.sub.uni-goettingen.de/bestandsaufnahme/>.

Krügel, Tina (2017). „Das personenbezogene Datum nach der DS-GVO: Mehr Klarheit und Rechtssicherheit?“ In: *Zeitschrift für Datenschutz* 7.10, S. 455–460. <https://beck-online.beck.de/Bcid/Y-300-Z-ZD-B-2017-S-455-N-1>.

McNutt, Marcia et al. (2016). „Liberating field science samples and data“. In: *Science* 351.6277, S. 1024–1026. <https://doi.org/10.1126/science.aad7048>.

Metschke, Rainer und Rita Wellbrock (2002 [1994]). *Datenschutz in Wissenschaft und Forschung*. 3. überarb. Aufl. Berlin: Berliner Beauftragter für Datenschutz und Informationsfreiheit, Hessischer Datenschutzbeauftragter.

RatSWD (2012). *Georeferenzierung von Daten: Situation und Zukunft der Geodatenlandschaft in Deutschland*. Abschlussbericht der AG „Georeferenzierung von Daten“. Berlin: Scivero Verlag für den Rat für Wirtschafts- und Sozialdaten. <https://www.ratswd.de/publikationen/georeferenzierung-von-daten>.

– (2017). *Handreichung Datenschutz*. Output Series 5. Berlin: Rat für Wirtschafts- und Sozialdaten. <https://doi.org/10.17620/02671.6>.

Raum, Bertram (1993). „Umweltschutz und Schutz personenbezogener Daten“. In: *Computer und Recht* 9, S. 162–170.

Schaar, Katrin (2016). *Was hat die Wissenschaft beim Datenschutz künftig zu beachten?* RatSWD Working Paper 257. Aktualisierte Fassung, Erstpublikation in *Zeitschrift für Datenschutz*, 6.9 (2016), 224ff. Berlin: Rat für Wirtschafts- und Sozialdaten. <https://doi.org/10.17620/02671.19>.

Schild (2018). „DS-GVO Artikel 4: Begriffsbestimmungen“. In: *BeckOK Datenschutzrecht*. Hrsg. von Wolff und Brink. 23. Aufl. Stand: 01.02.2018. Beck-Online, Rn. 22–27.

Taege, Jürgen (1991). „Umweltschutz und Datenschutz“. In: *Computer und Recht* 7, S. 681–688.

Vines et al. (2013). The Availability of Research Data Declines Rapidly with Article Age. *Current Biology* 24, S. 94–97. <https://doi.org/10.1016/j.cub.2013.11.014>

Weichert, Thilo (2007). „Der Personenbezug von Geodaten“. In: *Datenschutz und Datensicherheit* 31.1, S. 17–23. <https://doi.org/10.1007/s11623-007-0006-3>.

– (2009). „Geodaten: Datenschutzrechtliche Erfahrungen, Erwartungen und Empfehlungen“. In: *Datenschutz und Datensicherheit* 33.6, 347ff. <https://doi.org/10.1007/s11623-009-0071-x>.

Niklas Hartmann studierte Umweltwissenschaften in Potsdam und Wageningen, Niederlande sowie Wissenschaftssoziologie in Lancaster, Großbritannien. Seit 2016 ist er Koordinator Forschungsdaten an der Universität Potsdam und Fachreferent in der Universitätsbibliothek.

<https://orcid.org/0000-0002-7328-3625>