

April 2020

Stellungnahme des Rates für Informationsinfrastrukturen (RfII) Datentreuhandstellen gestalten – Zu Erfahrungen der Wissenschaft

Die Bundesregierung befasst sich in der Ausarbeitung ihrer Datenstrategie mit dem Thema Datentreuhänderschaft. Sie verweist in einem Eckpunktepapier darauf, dass Anstrengungen unternommen werden, um die „verantwortungsvolle Bereitstellung und Nutzung von Daten durch Personen und Institutionen (...) signifikant zu steigern.“¹ Dabei sollen neue Datenmonopole verhindert und eine „gerechte Teilhabe“ gesichert werden. Ebenso hat die Europäische Kommission in ihrer Datenstrategie befürwortet, Datenpools „in strategischen Sektoren und Bereichen von öffentlichem Interesse“ einzurichten, die – auch unter dem Gemeinwohl-Aspekt – den Datenzugang für verschiedene Akteure erleichtern sollen.²

Der Diskurs über Datentreuhänder will also auf das Problem, wie Daten auf eine dem Wettbewerb ein Stück weit entzogene Form von neutraler Seite bereitgehalten werden können, eine Antwort finden. Die Debatte hierzu ist eine primär außerwissenschaftliche, sie weist allerdings einige Analogien zu Aushandlungsprozessen in der Wissenschaft auf. So geht es um Austausch und um eine faire Nutzbarkeit der Daten, was beispielsweise bedeuten kann, dass unter mehreren Akteuren ein gleichberechtigter Zugriff auf Daten gewährleistet ist. Datentreuhänder können dazu motivieren, dass mehr als bislang Datenbestände zusammengeführt werden. Dies kann innovationspolitischen Interessen dienen, birgt aber auch erhebliches Potenzial zur Lösung großer gesellschaftlicher Herausforderungen in Bereichen wie Gesundheit, Klima, Mobilität oder Bekämpfung von Armutsursachen. Datentreuhänder können auch zum Einsatz kommen, wo ein Datenaustausch unter Akteuren in Konstellationen starker Machtasymmetrie oder Konkurrenz organisiert werden soll – zum Beispiel zwischen global agierenden Anbietern von Internetplattformen, Start-ups, Forschenden oder individuellen Verbraucherinnen und Verbrauchern.

Derzeit wird über das Konzept der Datentreuhänderschaft ergebnisoffen, aber auch vage diskutiert. Es besteht noch kein öffentliches oder politisches Einverständnis darüber, wie Datentreuhänder beziehungsweise Datentreuhandstellen aufgebaut sein sollen.³

¹ Die Bundesregierung (2019): Eckpunkte einer Datenstrategie, S. 1.

² Europäische Kommission (2020): Eine europäische Datenstrategie. Mitteilung der Kommission an das Europäische Parlament, den Rat, den Europäischen Wirtschafts- und Sozialausschuss und den Ausschuss der Regionen, Brüssel, S. 25.

³ Auch die vom Bundesministerium für Wirtschaft und Energie eingesetzte Kommission „Wettbewerbsrecht 4.0“ hat in ihrem Bericht 2019 zunächst lediglich befürwortet, die Einrichtung von Datentreuhändern und die verschiedenen Modelle zu prüfen. Siehe BMWi - Bundesministerium für Wirtschaft und Energie (2019): Ein neuer Wettbewerbsrahmen für die Digitalwirtschaft. Bericht der Kommission Wettbewerbsrecht 4.0.

Der Rfll nimmt diese Überlegungen mit Aufmerksamkeit zur Kenntnis. Für die Erschließung neuer Datensätze und für deren Weiternutzung ist die Grundidee der nachhaltigen Bereitstellung von datenbasierten Dienstleistungen seitens neutraler Stellen, die Nutzerinnen und Nutzern faire Bedingungen anbieten, von hoher Relevanz. Datentreuhänder könnten Teil einer gesamtgesellschaftlichen Dateninfrastruktur werden.

Aus dieser Grundidee ergeben sich allerdings erhebliche Anforderungen an die konkrete Ausgestaltung einer solchen Treuhänderrolle, an den zu garantierenden vertrauensvollen Umgang mit personenbezogenen und anderweitig sensiblen Daten sowie an die Voraussetzungen ihrer sicheren, langfristigen Aufbewahrung. Der Rfll regt, um hier von existierenden Vorbildern zu lernen, einen Erfahrungsaustausch mit der Wissenschaft an: Wissenschaftliche Fachgemeinschaften/Communities verfügen bereits über eine längere Tradition des fairen Teilens qualitäts gesicherter Daten und haben Regularien entwickelt, um die Rechte von Datengebern zu sichern. Zugleich hat die Wissenschaft auch eigene Erwartungen an Datentreuhandstellen, die außerhalb der Wissenschaft entstehen.⁴

Der Rfll ist ein Beratungsgremium, das sich mit dem Aufbau und der Weiterentwicklung von Informationsinfrastrukturen in der Wissenschaft befasst. Vor diesem Hintergrund ordnet er das Konzept der Datentreuhänderschaft begrifflich ein und formuliert Empfehlungen hinsichtlich der Ausgestaltung von Datentreuhandstellen in der Zivilgesellschaft und in der Wirtschaft, vor allem – aber nicht nur – mit einem besonderen Blick auf deren Nutzen für Forschung und Entwicklung.

BEGRIFFSKLÄRUNG

Allgemeines zum Treuhandbegriff

Die Vorstellung einer treuhänderischen Übernahme von Verantwortung oder Pflichten hat eine vielfältige Tradition. Sie wird hier nur kurz umrissen – auch um deutlich zu machen, dass Gestaltungsspielräume bestehen.

Im juristischen Verständnis ist die **Treuhanderschaft** ein Rechtsverhältnis zwischen einem **Treuhand** (*fiduciary/trustee*) und einem **Treugeber**: Dem Treuhand werden seitens des Treugebers bestimmte Rechte im Hinblick auf das **Treugut** übertragen oder eingeräumt, die der Treuhand im Außenverhältnis im eigenen oder auch fremden Namen wahrnehmen darf. Im Verhältnis zum Treugeber bleibt er dennoch an vereinbarte Vorgaben, Bedingungen und Grenzen gebunden, wobei der Treuhandvertrag gesetzlich nicht im Detail geregelt ist. Je nach Treuhandverhältnis sind daher verschiedene rechtliche Regelungen anwendbar. Ein Treuhandverhältnis kann zur Sicherung der Interessen des Treuhänders („eigennützige Treuhand“) oder der Interessen des Treugebers („fremdnützige Treuhand“) begründet werden. Ein

⁴ Vgl. hierzu RatSWD - Rat für Sozial- und Wirtschaftsdaten (2019): Big Data in den Sozial-, Verhaltens- und Wirtschaftswissenschaften: Datenzugang und Forschungsdatenmanagement, Berlin, S. 21ff. Der RatSWD formuliert hier auch bereits einen Aufgaben- und Kriterienkatalog für die Ausgestaltung einer Datentreuhandstelle in den Wirtschafts- und Sozialwissenschaften.

Treuhänder kann aber insbesondere auch im Falle widerstreitender Interessen verschiedener Personen tätig werden („doppelseitige Treuhand“). Im letzteren Fall dient das Treuhandverhältnis dazu, den Interessen der Beteiligten durch die Einschaltung eines **in der Sache uneigennützig** handelnden und vertrauenswürdigen Dritten gerecht zu werden.

Der **wirtschaftliche Treuhandbegriff** ist demgegenüber weiter gefasst, indem hier bereits die Wahrnehmung fremder Interessen als Treuhandschaft verstanden werden kann. Auch im **englischsprachigen Kontext** kann *trustee* ganz allgemein einen vertrauenswürdigen Verwalter meinen. An solche offenen, eher pragmatischen Vorstellungen von der Treuhänderrolle schließt auch die Diskussion über Datentreuhandstellen an.⁵

Datentreuhänder

Der Begriff des „Datentreuhänders“ hat sich im Digitalisierungsdiskurs als Metapher oder Analogie für Stellen herausgebildet, die in besonderer Weise als **Intermediäre** zwischen einem oder mehreren Datengebern und einer Nachfrageseite agieren. Es gibt bereits eine Reihe von Beschreibungen und Anwendungsfeldern. Für Anwendungsfälle im Medizinbereich wird ein Datentreuhänder beispielsweise definiert als eine „rechtlich, räumlich und personell selbstständige und unabhängige Stelle, die idealerweise einer besonderen Geheimhaltungspflicht unterliegt“.⁶ Der Begriff wird zum Teil auch weiter gefasst. So definiert die Bundesdruckerei einen Datentreuhänder als „eine unabhängige Vertrauensinstanz, die Daten zwischen Datengeber und Datennutzer sicher und gesetzeskonform vermittelt.“⁷ Dies kann insbesondere bedeuten, dass der Datentreuhänder im Falle personenbezogener Daten Aufgaben der Pseudonymisierung oder der Anonymisierung wahrnimmt und die Daten nur pseudonymisiert, anonymisiert oder aggregiert zur Verfügung gestellt werden.

Innerhalb der Wissenschaft bestehen keine Treuhandverhältnisse im engeren juristischen Sinn. Dennoch nehmen wissenschaftliche Institutionen die Rolle neutraler, das heißt nicht-kommerzieller und inhaltlich unabhängiger Datenarchive wahr. Ein Beispiel sind hier Biobanken oder auch Forschungsdatenzentren in den Sozial- und Wirtschaftswissenschaften.

Rechts- und digitalpolitische Forderungen nach der Schaffung sowie gegebenenfalls auch Kodifizierung der Rolle eines oder mehrerer Datentreuhänder sind vergleichsweise neu. Sie werden beispielsweise im Zusammenhang mit besonders umfangreichen (und häufig zugleich in ihrer Integrität schutzbedürftigen) Datensätzen artikuliert, die durch den Einsatz neuer IT-gestützter Datenverarbeitung anfallen. So besteht ein erhebliches Interesse daran, **Medizin-, Mobilitäts- oder Nutzungsdaten**, die im wirtschaftlichen Kontext oder in Alltagsanwendungen

⁵ Insbesondere wird so auf die Idee von Dateneigentum verzichtet. Vgl. hierzu die Empfehlung der Datenethikkommission; Datenethikkommission der Bundesregierung (2019): Gutachten der Datenethikkommission, Berlin, S. 18.

⁶ Pommerening, Klaus et al. (2014): Leitfaden zum Datenschutz in medizinischen Forschungsprojekten. Generische Lösungen der TMF 2.0 (Schriftenreihe der TMF, 11), Berlin, S. 209.

⁷ https://www.bundesdruckerei.de/system/files/dokumente/pdf/BDR.de_Datentreuhaender.pdf (zuletzt geprüft am: 03.04.2020). Vgl. zur Definition des Datentreuhänders u.a. auch RatSWD - Rat für Sozial- und Wirtschaftsdaten (2017): Handreichung Datenschutz (RatSWD Output, 5 (5)), Berlin, S. 16.

des Internets der Dinge produziert werden, besser für die Weiterverwendung zu erschließen, das heißt für Dritte, die Gesellschaft oder die Forschung nutzbar zu machen. Im zivilgesellschaftlichen Bereich werden Datentreuhandssysteme auch als Option für einen faireren und stärker **selbstbestimmten Umgang mit Verbraucherdaten** diskutiert (*Personal Data Trusts, Personal Information Management Systems*). Hier steht die individuelle Datensouveränität im Vordergrund. Die Datenethikkommission der Bundesregierung hat in diesem Kontext allerdings ausdrücklich darauf aufmerksam gemacht, dass der Einzelne vor „vermeintlichen Interessenverwaltern“, die vor allem wirtschaftliche Eigeninteressen vertreten, zu schützen sei.⁸

EMPFEHLUNGEN

Aus Sicht des Rfll können etablierte Praktiken des Datenzugangs und der Datennutzung in der Wissenschaft auch als Rollenmodell für den Aufbau von Dateninfrastrukturen in anderen gesellschaftlichen Teilbereichen fungieren. An den Fernerkundungsdaten der großen Satellitenmissionen lässt sich beispielsweise zeigen, wie ein **fares Teilen von Daten** im Rahmen **größerer Fachgemeinschaften/Communities** gut funktionieren kann und für diese eine langfristige Nutzbarkeit sicherstellt. Erst dadurch wird die Erforschung komplexer Systemzusammenhänge überhaupt ermöglicht. Die Großforschung, etwa am Large Hadron Collider des CERN, ist mit ihrer **hochkomplexen, arbeitsteiligen Datenarchivierung** ein Beispiel dafür, wie Forschenden weltweit eine Nutzung großer Datenmengen unter Gewährleistung verbindlicher Standards für die Datenqualität ermöglicht werden kann. Das räumlich dezentrale, aber sachlich zentral koordinierte deutsche **Bibliothekssystem** verfügt über seit Jahrhunderten sich evolutionär entwickelnde Formen der Metadatenvergabe. Die **Forschungsdatenzentren in den Sozial- und Wirtschaftswissenschaften** liefern Modelle für den rechtskonformen Zugang zu und die Nachnutzung von Daten, die aus Gründen des unternehmerischen Eigentums oder des Datenschutzes sensibel sind. Diese Beispiele aus der Wissenschaft zeigen, wie ein mit einheitlichen Standards qualitätsgesicherter und gleichzeitig unter mehreren Akteuren fair organisierter Datenaustausch aufgebaut werden kann.

Ausgehend von diesen Erfahrungen versteht der Rfll Datentreuhandstellen nachfolgend in einem institutionellen Sinn, nämlich als eine Dateninfrastruktur besonderen Typs.

Dateninfrastrukturen für Wissenschaft, Wirtschaft und Gesellschaft gut ausbalanciert gestalten

Die aktuellen Diskussionen zur Intensivierung der Nutzung von Daten und zur Etablierung von Modellen der Datentreuhänderschaft zeigen der Politik Wege auf, in welche Richtung die zahlreichen Projekte und Initiativen im Bereich der Dateninfrastrukturen strategisch weiterentwickelt werden sollten. In der *International Data Spaces Association* entsteht ein Referenzmodell für Plattformen, die Unternehmen den Zugang, die Nutzung und den sicheren Austausch von Daten erleichtern wollen. Mit dem Projekt GAIA-X wird für Unternehmen und Organisationen ein europaweit offenes „Ökosystem“ vertrauenswürdiger Werkzeuge, Dienst-

⁸ Datenethikkommission (2019): Gutachten der Datenethikkommission, S. 21.

leistungen sowie Speicher- und Rechenkapazitäten angestrebt. Mit Blick auf Gemeinwohl und innovationspolitische Ziele will die EU-Kommission sogenannte „Datenpools“ für spezifische Sektoren einrichten, darunter Industrie, Umwelt und Gesundheit.⁹ Weltweit erschließen Behörden ihre Datenbestände und stellen sie über Portale wie Transparenzregister, GovData oder das *European Data Portal* zur Verfügung. Ganz ähnliche Anstrengungen werden unter den besonderen Vorzeichen von Wissenschaft seit einigen Jahren mit der Einrichtung von wissenschaftlichen Datenzentren für die Forschung unternommen.

Der Charakter von Unternehmens- und Verbraucherdaten kann sich von denen der wissenschaftlichen Datenbestände unterscheiden. Beiden gemeinsam ist allerdings, dass sie bestimmte Qualitätsmerkmale erfüllen müssen. Auch Kriterien wie Vertrauenswürdigkeit, Rechtskonformität sowie Transparenz des Zugangs haben in beiden Welten den gleichen Wert. In diesen Punkten können sich die Interessen der Marktteilnehmer wie auch von Wissenschaft, Politik und Gesellschaft treffen.

Die **Idee eines Systems von „Treuhandstellen“**, die einem Grundkonzept der gleichermaßen effektiven wie rechtskonformen und im Sinne der Datensouveränität nachvollziehbaren Nutzung von qualitätsgesicherten Datenbeständen folgt, erscheint dem Rfll daher überaus sinnvoll für die Entwicklung einer Datenstrategie. Ein Ausbau solcher Dateninfrastrukturen sollte **mit gesamtgesellschaftlichem Auftrag** vorangetrieben werden, gegebenenfalls mit öffentlicher Förderung. Allerdings sollte auch genau geprüft werden, in welchen Bereichen Treuhandmodelle funktionieren und ob sie – ihrer Zielsetzung entsprechend qualitätsgesicherte – Datenbestände zusammenführen oder einen Datenaustausch zwischen verteilten Ressourcen koordinieren sollen.

Zugang für Forschung und Entwicklung sichern

Der Rfll sieht in Datentreuhandstellen ein großes Potenzial, sowohl einen vertrauensvollen **Austausch von Daten** nicht nur innerhalb eines Sektors (u. a. Kooperation unter Wettbewerbern in Innovationsprozessen der Industrie, Verfügbarmachung und Zusammenführung von gesundheitsrelevanten Daten, Verwaltungsdaten, etc.) zu organisieren, als auch im Sinne von Schnittstellen zwischen Wissenschaft, Wirtschaft und Gesellschaft zu fungieren. Hierdurch ließe sich das wirtschaftliche und gemeinwohlorientierte **Innovationspotenzial** ausschöpfen, das mit *data sharing* verbunden ist.

In Fällen, in denen die öffentliche Hand als Treuhandstelle fungiert, ist es aus Sicht des Rfll wichtig, dass ein möglichst weitreichender Datenzugang für Forschung und Entwicklung gewährleistet wird. Wenn unter wettbewerbsrechtlichen Aspekten Regulierungsmaßnahmen ergriffen werden, um die Daten großer Plattformen auch für andere Marktteilnehmer zu öff-

⁹ Europäische Kommission (2020): Eine europäische Datenstrategie.

nen,¹⁰ dann sollte ein solcher **Zugang** auch und erst recht öffentlich geförderter Forschung **in sachgerechtem Umfang** eingeräumt werden.

Der Rfll möchte die Bedarfe der Wissenschaft in diesem Prozess deutlich artikulieren: Sektorale Datenmonopole oder andere sachwidrige Beschränkungen des Zugangs für die Wissenschaft sind zu vermeiden. Demnach wäre in Regelungsvorhaben jeweils eine **Forschungsklausel** zu etablieren, in der die Belange der Forschung in sachgerechter Art und Weise berücksichtigt sind.

Darüber hinaus gibt es zahlreiche Anwendungsbeispiele für Datenplattformen, in denen sich ein Datenzugang durch Forschungsverträge, Kooperationsvereinbarungen oder das Bereitstellen von Forschungsdatensätzen erreichen lässt. Treuhandstellen können hier je nach Schutzstufe der Daten eine Vielzahl an Aufgaben wahrnehmen, bis hin zur Entwicklung von Datenprodukten für spezifische Nachnutzungsfälle und unterschiedliche Zielgruppen. Derartige Szenarien sind bereits in der Forschung erprobt und werden in zahlreichen Anwendungsbereichen schon mit Erfolg eingesetzt.

Qualitätssicherung für Treuhandstellen etablieren

Datentreuhänder sollen als Vertrauensinstanz transparent und fair eine Nutzung von Daten für möglicherweise auch miteinander konkurrierende Akteure organisieren. Eine weitergehende Festlegung der funktionalen Dimensionen, die solche Infrastrukturen zu erfüllen haben, steht allerdings noch aus. Um einer künftig inflationären Verwendung des Datentreuhänder-Begriffs vorzubeugen, schlägt der Rfll vor, für solch neutrale anerkannte Vertrauensinstanzen die **Bezeichnung „Datentreuhandstelle“** zu etablieren und diese Bezeichnung durch ein Siegel oder eine Zertifizierung zu **schützen**. Aus Sicht des Rfll sollte mit dem Begriff ein Qualitätsanspruch verbunden sein, der durch entsprechende Qualitätssicherungsmaßnahmen untermauert wird und die gewünschte Neutralität sichert. Bei der Festlegung der hierfür erforderlichen Kriterien sollte auch die Datenqualität Berücksichtigung finden.¹¹ Auch wenn Datentreuhandstellen nicht explizit für die Qualitätssicherung der Daten zuständig sind beziehungsweise sein sollten, so müssten sie zumindest eine Beurteilung der bereitgestellten Daten ermöglichen und Informationen über die Qualität der Daten an die Nutzerinnen und Nutzer weitergeben.

Da sowohl die Daten in ihrer Sensibilität als auch die geplanten Nutzungen und die jeweils damit verbundenen Bedingungen große Unterschiede aufweisen können, bietet es sich an, verschiedene Stufen von Qualitätsgarantien für die Vertrauensinstanzen vorzusehen. So werden bei hoher Kritikalität und großem Risikopotenzial Überprüfungen durch Dritte, etwa un-

¹⁰ Dies sieht zurzeit beispielsweise ein Konzept der Projektgruppe „KI und Wirtschaft“ der Enquete Kommission „Künstliche Intelligenz – Gesellschaftliche Verantwortung und wirtschaftliche, soziale und ökologische Potenziale“ vor: Projektgruppe „KI und Wirtschaft“ (2019): Zusammenfassung der vorläufigen Ergebnisse, Stand: 18. Dezember 2019.

¹¹ Vgl. hierzu Rfll - Rat für Informationsinfrastrukturen (2019): Herausforderung Datenqualität – Empfehlungen zur Zukunftsfähigkeit von Forschung im digitalen Wandel, 2. Aufl., Göttingen.

abhängige Zertifizierungsstellen oder staatliche Überwachungsstellen, notwendig sein. Bei geringerer Kritikalität können auch Selbstverpflichtungsmechanismen zum Beispiel über *Codes of Conduct* erwogen werden. Die Wissenschaft kennt solche Stufenmodelle beispielsweise in der Zertifizierung „vertrauenswürdiger Langzeitarchive“.¹² Der Erwerb von Siegeln oder Akkreditierungen beinhaltet zudem oftmals die Aufnahme in eine *Community of Practice*, in der die Datenzentren Erfahrungsaustausch und laufende Weiterentwicklung der Verfahren organisieren.

Ausgestaltung und Regulierung von Datentreuhandstellen sektorenübergreifend voranbringen

Der Rfll regt an, zeitnah Gespräche zwischen den Akteuren in Politik, Wirtschaft, Verwaltung und Wissenschaft anzustoßen, um konkrete Ansätze für ein System zertifizierter Datentreuhandstellen zu entwickeln – letztlich auch mit Blick auf die europäische Ebene. In diesen Prozess sollte die Wissenschaft nicht nur ihre Erfahrungen mit der Regelung des Zugangs und der Nutzung von Daten einbringen, sondern auch ihre eigenen spezifischen Bedarfe artikulieren. Gestaltungsmerkmale und Regulierungen für Datentreuhandstellen werden insbesondere in Abhängigkeit von Kritikalität der jeweils bereitgestellten Daten und der Nutzungszwecke festzulegen sein. Eine **Förderung von Vorhaben**, die Anforderungen und Regelungsbedarfe detaillierter ausarbeiten, wäre ein hilfreicher Impuls, um den Aufbau solcher Dateninfrastrukturen kontrolliert voranzubringen.

¹² Konformität kann hier in unterschiedlich detaillierter Form vom peer-review-gestützten Selbstevaluierungsverfahren des sogenannten Core Trust Seal bis hin zu einem formalen externen Audit nachgewiesen werden. Grundlage ist jeweils das OAIS-Referenzmodell (ISO-Standard 14721).

Impressum

Rat für Informationsinfrastrukturen (RfII) - Geschäftsstelle

Papendiek 16, 37073 Göttingen

Fon 0551-392 70 50

E-Mail info@rfii.de

Web www.rfii.de

Dieses Werk ist lizenziert unter einer Creative Commons Namensnennung –
keine Bearbeitung 4.0 Lizenz (CC BY-ND).

