

„Innovative Datenpolitik: Potenziale und Herausforderungen“

“Innovative data policy: potential and challenges”

Öffentliche Anhörung / Public Hearing

Ausschuss für Digitales/ Committee on Digital Affairs,

Stellungnahme von Dr. Stefaan G. Verhulst

Deutscher Bundestag

26. Juni 2024, 14:30 Uhr bis 16:30 Uhr MESZ

Erschließung des Potenzials von Daten: Innovative Maßnahmen für eine verantwortungsbewusste Wiederverwendung von Daten und den Umgang mit Datenasymmetrien

Sehr geehrte Frau Vorsitzende, sehr geehrte Mitglieder des Ausschusses für Digitales,

ich freue mich, diese Stellungnahme dem Ausschuss für Digitales im Rahmen der öffentlichen Anhörung zum Thema „Innovative Datenpolitik: Potenziale und Herausforderungen“ vorstellen zu dürfen.

Mein Name ist Stefaan Verhulst, ich bin Mitbegründer von The Governance Lab (GovLab) in New York City und The DataTank in Brüssel. Darüber hinaus habe ich eine Forschungsprofessur an der Tandon School of Engineering der New York University inne.

Das Potenzial: Verantwortungsbewusste Wiederverwendung von Daten

Lassen Sie mich zunächst erläutern, welches Potenzial Daten haben, wenn sie verantwortungsbewusst genutzt und wiederverwendet werden. Obwohl wir viel über die Risiken der Nutzung von Daten hören – und viele der Befürchtungen sind sicherlich gerechtfertigt –, ist es auch wichtig, die sehr realen Möglichkeiten zu berücksichtigen, die Daten im Hinblick auf die Förderung des Gemeinwohls bieten.

Wir leben in einer datengesättigten Welt, die durch ein beispielloses Angebot – ja geradezu einer Schwemme – an Daten gekennzeichnet ist. In dieser Welt sind Daten zu einer sehr wichtigen Ressource geworden, die in politische Prozesse und Entscheidungsprozesse einfließt. Wenn sie richtig analysiert und genutzt werden, können Daten eine entscheidende Rolle dabei spielen, politische Entscheidungsträger – und andere Interessenträger – bei der Lösung einer Reihe kritischer Probleme zu unterstützen, und zwar in so unterschiedlichen Sektoren wie öffentliche Gesundheit, Klima, Innovation und wirtschaftliche Entwicklung, Bekämpfung des Verfalls von Städten – und vieles mehr.

Manchmal sind diese Daten leicht verfügbar. Meistens aber nicht. Einer der Bereiche mit dem größten Potenzial – aber mit auch erheblichen Herausforderungen – ist die Wiederverwendung von Daten, d.h. Daten, die bereits für einen bestimmten Zweck erhoben wurden, werden für einen anderen Zweck wiederverwendet. Die Wiederverwendung von Daten kann unschätzbare Einblicke in aktuelle Phänomene liefern, die uns helfen, die Ursachen für aufkommende Trends zu verstehen und uns bei der Entwicklung von effektiven Lösungen für dringende Probleme zu unterstützen. Darüber hinaus kann die Analyse aus der Wiederverwendung von Daten als leistungsfähiges Instrument dienen, um zukünftige Entwicklungen zu antizipieren und gezielte Interventionen vorzugeben.

In den letzten Jahren wurden immer mehr Belege gefunden, die das Potenzial für die Wiederverwendung von Daten aufzeigen und uns den Weg weisen, wie man sicherstellen kann, dass dies zum Gemeinwohl beiträgt und gleichzeitig Risiken minimiert. Um dieses Potenzial zu verdeutlichen, betrachten wir das Beispiel des Klimawandels.

- Die Wiederverwendung von Satellitendaten kann beispielsweise genaue Messungen von Treibhausgasemissionen wie Methan und deren Quellen liefern.
- Die Wiederverwendung von Social-Media-Daten kann Einblicke in die Ansichten der Menschen zum Thema Klimawandel geben.
- Die Aggregation traditioneller und nicht-traditioneller Daten kann die Modellierung der langfristigen Auswirkungen des Klimawandels auf verschiedene Ökosysteme und gefährdete Menschen und Gruppen – wie etwa Kinder – ermöglichen.

- Die Wiederverwendung von Daten der Wasserversorgungsunternehmen kann beispielsweise dazu beitragen, die Wirksamkeit verschiedener Abschwächungs- und Anpassungsstrategien zu bewerten.
- Die Wiederverwendung von Handydaten kann dazu beitragen, die Ressourcenzuweisung für Projekte zu optimieren, die im Zusammenhang mit Klimaresilienz stehen.

Die Herausforderungen: Datenasymmetrien und verpasste Chancen

Ungeachtet des eigentlichen Potenzials von Daten und deren Wiederverwendung stehen wir unbestreitbar vor großen Herausforderungen in Bezug auf die Aufgabe, den gesellschaftlichen Wert von Daten in vollem Umfang auszuschöpfen.

Eines der Haupthindernisse ist der fehlende Zugang zu qualitativ hochwertigen, aktuellen Daten für den öffentlichen Sektor, die Zivilgesellschaft und andere Gruppen, die sich für das Gemeinwohl einsetzen.

Wir leben heute in einer paradoxen Situation, die einerseits durch die Verfügbarkeit einer noch nie dagewesenen Menge von Daten und andererseits durch eine noch nie dagewesene Asymmetrie in Bezug auf den Zugang zu diesen Daten für die Weiterverwendung im öffentlichen Interesse gekennzeichnet ist.

Ich glaube, dass die zunehmende Asymmetrie zwischen denjenigen, die über Daten verfügen (oft aus dem privatwirtschaftlichen Sektor), und denjenigen, die am besten in der Lage sind, sie für das Gemeinwohl zu nutzen, eine der größten Herausforderungen unserer Zeit darstellt.

Die Datenpolitik konzentriert sich bisher in erster Linie auf die Verhinderung des Missbrauchs von Daten, oft aus triftigen Gründen, wie bereits erwähnt. Bei diesem Ansatz wurden jedoch versehentlich die verpassten Verwendungsmöglichkeiten von Daten übersehen – also die Möglichkeiten, die wir aufgrund einer zu restriktiven Politik oder eines Mangels an innovativen Rahmenbedingungen für den Datenaustausch und die Datennutzung ungenutzt lassen.

Innovative Ansätze in der Datenpolitik

Was kann die Politik angesichts dieser Herausforderungen tun? Welche Schritte können politische Entscheidungsträger wie Sie – und andere Interessenträger aus dem Privatsektor, der Wissenschaft und der Zivilgesellschaft – unternehmen, um das Potenzial unserer datengestützten Gesellschaft und Wirtschaft zu maximieren und sicherzustellen, dass die Vorteile unseres Datenzeitalters auf möglichst gerechte und inklusive Weise maximiert werden?

Im Folgenden möchte ich einige der Schlüsselkomponenten eines innovativeren, flexibleren und verantwortungsbewussteren Rahmens für die Datenpolitik skizzieren. Obwohl die Elemente, die ich hier skizziere, nicht endgültig oder umfassend sind, basieren sie auf umfangreicher Forschung in Europa und weltweit sowie auf jahrelanger Zusammenarbeit mit politischen Entscheidungsträgern, dem Privatsektor, der Zivilgesellschaft, der Wissenschaft und anderen.

Lassen Sie mich also zunächst drei konkrete Wege skizzieren, wie wir zu einem ausgewogeneren Datenökosystem beitragen können. Die drei politischen Empfehlungen, die ich skizzieren werde, konzentrieren sich auf drei spezifische Formen der Asymmetrie, die ich wie folgt bezeichne:

- Asymmetrie der Agentur,
- Datenasymmetrie und
- Technische und rechnerische Asymmetrie.

1. Bekämpfung der Asymmetrie der Agentur: Digitale Selbstbestimmung vorantreiben

Die „Asymmetrie der Agentur“ bezieht sich auf die ungleichen Bedingungen, unter denen Einzelpersonen und Gruppen derzeit über die Verwendung ihrer Daten bestimmen können. Diese Asymmetrien können bestehende sozioökonomische Ungleichheiten widerspiegeln; oft verschärfen sie aber auch diese Spaltungen oder schaffen neue.

Traditionelle Datenrichtlinien, beispielsweise die DSGVO, haben sich stark auf das Konzept der Einwilligung als primären Mechanismus für Einzelpersonen zur Kontrolle ihrer Daten gestützt. Die meisten Menschen denken bei „Einwilligung“ an die Kontrollkästchen, die aufploppen und uns auffordern, einer Datenverwendungsrichtlinie zuzustimmen, die wir wahrscheinlich nie gelesen haben. Die Idee hinter der Einwilligung ist sehr wichtig (dem Einzelnen mehr Kontrolle und Auswahl über seine Daten zu geben), aber es ist heute allgemein anerkannt, dass der Mechanismus selbst höchst unzureichend ist. Dafür gibt es mindestens drei Gründe:

- Erstens erfolgt die Einwilligung aufgrund zu komplexer Bedingungen selten tatsächlich auf informierter Basis.
- Zweitens wird die Einwilligung in der Regel als eine binäre Entscheidung dargestellt, wodurch übersehen wird, dass bei der Entscheidung über die Verwendung und Wiederverwendung von Daten mehr Nuancen erforderlich sind.
- Und schließlich konzentrieren sich die vorhandenen Zustimmungsmechanismen auf individuelle Präferenzen, wobei Gemeinschafts- oder Gruppeninteressen vernachlässigt werden.

Um diese Grenzen zu überwinden, habe ich an anderer Stelle argumentiert, dass wir die Einwilligung ergänzen und zu einem Modell der digitalen Selbstbestimmung übergehen müssen. Digitale Selbstbestimmung ist eine Ergänzung zur Einwilligung, die darauf abzielt, dass Einzelpersonen und Gemeinschaften selbst bestimmen können, wie ihre Daten auf leistungsfähigere Weise verwendet und wiederverwendet werden.

Sie umfasst drei Schlüsselkomponenten:

- **Erstens: Prozessinnovation.** Wir müssen neue Wege entwickeln, um uns mit Menschen und Gemeinschaften in Bezug auf die Nutzung von Daten auszutauschen und über bestehende Einwilligungsmethoden hinauszugehen. Unsere Forschung hat beispielsweise gezeigt, dass „Datenversammlungen“ – Bürgerversammlungen, die sich mit der Wiederverwendung von Daten befassen – verschiedene Interessengruppen an einen Tisch bringen können, um gemeinsam Fragen der Datenverwaltung (Data Governance) zu erörtern und darüber zu entscheiden. Wir haben solche „Datenversammlungen“ durchgeführt, um die öffentliche Meinung über die Wiederverwendung von Daten während der COVID-19-Pandemie zu verstehen und herauszufinden, wie junge Menschen die Wiederverwendung von Daten für oder über sie in Bezug auf ihre Gesundheit und ihr Wohlbefinden empfinden.
- **Zweitens erfordert die digitale Selbstbestimmung auch Innovationen in Politik und Politikgestaltung:** Es werden neue Mechanismen benötigt, um die Präferenzen und Erwartungen der Gemeinschaft zum Ausdruck zu bringen. Ein mögliches solches Instrument, das wir in unserer Forschungsarbeit untersucht haben, ist die Entwicklung einer „sozialen Lizenz“ für die Datennutzung, in der vereinbarte Grundsätze und Richtlinien für verantwortungsvolle Datenpraktiken innerhalb einer Gemeinschaft oder eines Sektors formuliert werden.
- **Und schließlich erfordert digitale Selbstbestimmung produktbezogene und institutionelle Innovationen:** Wir sollten neue Strukturen wie vertrauenswürdige Datenräume erkunden, die sichere Umgebungen für Datenaustausch und Analyse bieten und dabei individuelle und gemeinschaftliche Präferenzen berücksichtigen.

Diese drei Komponenten werden aktiv vom Netzwerk Digitale Selbstbestimmung entwickelt, das von der Schweizer Regierung geleitet wird. Das Netzwerk steht auch anderen Akteuren und Regierungen offen, darunter auch Akteuren aus Deutschland.

2. Bekämpfung der Datenasymmetrie: Open Data und Datenkooperationen

Eine zweite wichtige - und problematische - Form der Asymmetrie ist die Datenasymmetrie. Datenasymmetrie bezieht sich auf den begrenzten und ungleichen Zugang für diejenigen, die am besten in der Lage sind, Daten für das Gemeinwohl zu nutzen, insbesondere wenn dieser Zugang sektorübergreifend ist - z.B. der Zugang von politischen Entscheidungsträgern zu Daten des privatwirtschaftlichen Sektors.

Insbesondere haben wir das Potenzial – und die Grenzen – von sektorübergreifenden Partnerschaften untersucht, die den Zugang zu nicht traditionellen Datenquellen zugunsten des öffentlichen Interesses ermöglichen. So haben wir weltweit Beispiele gesehen, bei denen Mobiltelefonaten von unschätzbarem Wert für Stadtplanung, Katastrophenschutz oder Maßnahmen im Bereich der öffentlichen Gesundheit sein können.

In Europa wurde mit mehreren Instrumenten wie der Datenverordnung und dem Data-Governance-Rechtsakt eine politische Grundlage für die Bekämpfung von Datenasymmetrien geschaffen. Um diese Ansätze jedoch systematischer, nachhaltiger und verantwortungsbewusster zu gestalten, brauchen wir politische Innovationen in verschiedenen Bereichen:

- Zunächst müssen wir daran arbeiten, die Datennachfrage zu stärken und zu priorisieren: Allzu oft konzentrieren sich Dateninitiativen auf Infrastruktur und Datenverfügbarkeit; sie suchen nach Lösungen, ohne vorher die wichtigsten Fragen zu verstehen, zu deren Beantwortung die Daten genutzt werden könnten. Wir brauchen eine innovative Datenpolitik und -strategie, die der Identifizierung und Formulierung der entscheidenden Fragen und Anwendungsfälle, die von einem verbesserten Datenzugang profitieren würden, Priorität einräumt. Dies ist einer der Gründe, warum wir im GovLab unsere „100-Fragen-Initiative“ entwickelt haben, mit der wir einen kulturellen Wandel herbeiführen wollen, indem wir neue Fragen entwickeln und den Prozess der Priorisierung von Fragen in verschiedenen Bereichen neugestalten. Einfach ausgedrückt: Wir wollen die Nachfrage nach Daten stärken, damit sie sich stärker auf das Gemeinwohl richtet. Im europäischen Kontext gab es Bemühungen, hochwertige Datensätze zu identifizieren, aber ohne hochwertige Fragen ist die Identifizierung solcher Datensätze im privaten Sektor schwer zu bewerkstelligen oder zu begründen.
- Zweitens brauchen wir neue Berufe, die helfen können, verschiedene Formen der Asymmetrie zu beseitigen. Wir sollten beispielsweise die Etablierung von „Chief Data Stewards“ fördern: Fachleute, die den verantwortungsbewussten Datenzugang und die Wiederverwendung von Daten sektorübergreifend erleichtern können. Diese Personen würden die menschliche Infrastruktur bereitstellen, die für effektive Datenkooperationen erforderlich ist, und ganz allgemein eine verantwortungsvolle, gesunde Ökologie für die Wiederverwendung und den Austausch von Daten fördern.
- Drittens müssen wir neue institutionelle Ansätze entwickeln: Zum Beispiel sollten wir neue Arten von Organisationen unterstützen, die als vertrauenswürdige Vermittler für die gemeinsame Nutzung von Daten fungieren können. Die Bestimmungen des Europäischen Data-Governance-Rechtsakt für Datenintermediäre sind ein Schritt in die richtige Richtung, aber wir sollten auch andere Modelle wie Data-Commons-Ansätze untersuchen, die Datenbestände auf eine kollektivere Weise verwalten könnten.
- Ein weiterer Ansatz zur Reduzierung von Datenasymmetrien ist die Einrichtung eines Dateninnovationsfonds: ein spezieller Finanzierungsmechanismus, der die Erprobung neuer

Datenaustauschmodelle und die Ausweitung erfolgreicher Initiativen unterstützt. Dies könnte ähnlich funktionieren wie ein Fonds für den universellen Dienst, der im Telekommunikationssektor üblicherweise zur Überbrückung der letzten Meile bei der Bereitstellung von Telekommunikationsdiensten eingesetzt wird, aber auf die Zusammenführung von Daten für Zwecke des öffentlichen Interesses ausgerichtet ist.

3. Überwindung der technischen und rechnerischen Asymmetrie

Um schließlich eine verantwortungsbewusstere, effektivere und gerechtere Datenökologie zu fördern, müssen wir uns mit der ungleichmäßigen Verteilung der technischen Fähigkeiten und Rechenressourcen befassen – was ich als „technische Asymmetrien“ bezeichne.

Technische Kapazitäten sind eine Voraussetzung (wenn auch für sich genommen nicht ausreichend), um das Potenzial unserer Datenökologie voll auszuschöpfen und deren Schäden zu minimieren. Gleichwohl ist erwiesen, dass die globale und nationale Landschaft durch tiefgreifende Ungleichheiten bei diesen Kapazitäten gekennzeichnet ist, insbesondere im Bereich der künstlichen Intelligenz. Allerdings gibt es einige Schritte, die wir unternehmen können, um diese Ungleichheiten zu beseitigen. Dazu gehören:

- **Demokratisierung des Zugangs zu Rechenkapazitäten:** Wir brauchen Strategien und Initiativen, die den Zugang zu Hochleistungsrechenressourcen für nicht-kommerzielle und öffentliche Interessen ermöglichen – genau die Organisationen, die in der Regel am stärksten an der Nutzung von Daten für das Gemeinwohl beteiligt sind. Solche Initiativen könnten beispielsweise Partnerschaften mit Cloud-Computing-Anbietern oder die Schaffung einer Public-Computing-Infrastruktur umfassen.
- **Entwicklung zweckgebundener Modelle:** Wir sollten auch die Entwicklung von KI-Modellen und Methoden des maschinellen Lernens fördern, die speziell für Anwendungen im öffentlichen Interesse konzipiert wurden. Dieser Ansatz stellt sicher, dass fortgeschrittene analytische Fähigkeiten zur Bewältigung gesellschaftlicher Herausforderungen und nicht nur kommerzieller Interessen zur Verfügung stehen. Ein kürzlich veröffentlichter Bericht des Open Data Policy Lab des GovLab beispielsweise enthält eine Reihe von Anwendungsfällen, in denen sich generative KI und offene Daten aus offiziellen Quellen für öffentliche Zwecke überlappen.
- **Aufbau von Daten und KI-Kompetenzen:** Schließlich ist es wichtig, dass wir in Bildungs- und Ausbildungsprogramme investieren, um die Datenkompetenz in der gesamten Gesellschaft zu erhöhen. Fragenkompetenz ist unerlässlich, um definieren zu können, welche Fragen Priorität haben sollten, um Fortschritte bei der Bewältigung gesellschaftlicher Herausforderungen zu erzielen (oder die SDG-Ziele für 2030 zu erreichen). Technische Kompetenz ist entscheidend, um technische Asymmetrien (und im weiteren Sinne alle Formen der Asymmetrie) anzugehen. Wenn mehr Menschen und Organisationen in der Lage sind, Daten effektiv und verantwortungsbewusst zu nutzen und zu interpretieren, ist dies der Schlüssel für die Erschließung des Potenzials des Datenzeitalters. Wir brauchen Kompetenzen, um Datenintelligenz in Entscheidungsintelligenz umzumünzen.

Empfehlungen

Zusammenfassend lässt sich sagen, dass das Potenzial von Daten, einen positiven gesellschaftlichen Wandel voranzutreiben, immens ist, aber das Ausschöpfen dieses Potenzials erfordert innovative Ansätze in der Datenpolitik. Indem wir uns mit agentur- und datenbezogenen sowie technischen Asymmetrien befassen, können wir ein gerechteres und effektiveres Datenökosystem schaffen, das

dem öffentlichen Interesse dient und gleichzeitig die Rechte des Einzelnen und der Gemeinschaft schützt.

Aufbauend auf diesen Beobachtungen möchte ich dem Ausschuss einige spezifische politische Empfehlungen zur Prüfung geben. Viele dieser Empfehlungen können durch das geplante Dateninstitut operationalisiert und umgesetzt werden – je nachdem, wie es strukturiert und verwaltet werden soll:

- Schließen Sie sich dem Internationalen Netzwerk für digitale Selbstbestimmung an, um einen Rahmen für die digitale Selbstbestimmung zu schaffen, der über traditionelle Einwilligungsmuster hinausgeht und Entscheidungsprozesse auf Gemeinschaftsebene für die Datenverwaltung integriert.
- Führen Sie eine oder mehrere Datenversammlungen durch – Bürgerversammlungen rund um die Prioritäten und Bedingungen für die Wiederverwendung von Daten (d.h. nicht traditionellen Daten) im öffentlichen Interesse.
- Entwickeln Sie eine 100-Fragen-Initiative für Deutschland, um vorrangige gesellschaftliche Fragen zu identifizieren, bei denen die Weiterverwendung von Daten, einschließlich Daten des Privatsektors, wertvoll sein könnte.
- Investieren Sie zu diesem Zweck in Fragekompetenzen als eigentliche Grundlage für Datenwissenschaft und KI.
- Richten Sie einen nationalen oder EU-weiten Dateninnovationsfonds zur Unterstützung des Zugangs zu nicht traditionellen Daten und experimentellen Ansätzen für die Datenkooperation zum gesellschaftlichen Nutzen ein.
- Investieren Sie in die Entwicklung von „Data Stewardship“ als Berufsfeld, was die Erarbeitung von Bildungsprogrammen und professionellen Zertifizierungen beinhaltet.
- Richten Sie die Stelle des „Chief Data Steward“ im öffentlichen und privaten Sektor ein, um Datenkooperationen zu beschleunigen.
- Bauen Sie Partnerschaften zwischen öffentlichem und privatem Sektor auf, um Rechenressourcen und technische Unterstützung für gemeinnützige Organisationen und Forschende, die an Projekten von öffentlichem Interesse arbeiten, zur Verfügung zu stellen.
- Sorgen Sie dafür, dass so genannte „Datenfolgenabschätzungen“ in wichtige politische Initiativen verpflichtend aufgenommen werden, um sicherzustellen, dass datengesteuerte Erkenntnisse in die Entscheidungsprozesse einfließen.
- Entwickeln Sie Labors zur Beschleunigung von Entscheidungen, die die Übertragung von Datenintelligenz in Entscheidungsintelligenz vorantreiben können.

Zusammen betrachtet stellen diese politischen Innovationen einen Weg zu einer gerechteren und produktiveren Datenlandschaft dar. Sie können einige der bestehenden – und wachsenden – Ungleichgewichte in unserer technischen Ökologie korrigieren und dazu beitragen, dass die Früchte des Datenzeitalters gerechter verteilt werden.

Um wirksam zu sein, erfordern diese Ansätze jedoch nachhaltiges Engagement, Investitionen und sektorübergreifende Zusammenarbeit.

Ich fordere diesen Ausschuss auf, diese Empfehlungen zu berücksichtigen, wenn Sie die Zukunft der Datenpolitik in Deutschland gestalten und möglicherweise breitere EU-weite Initiativen beeinflussen. Durch die Förderung von Innovationen in der Art und Weise, wie wir Daten verwalten, teilen und nutzen, können wir ihr transformatives Potenzial erschließen, um unsere drängendsten gesellschaftlichen Herausforderungen zu bewältigen.

Ich danke Ihnen für Ihre Aufmerksamkeit. Ich freue mich darauf, all Ihre Fragen zu beantworten.