

Positionspapier „Transparenz 2.0: Open Data gegen Korruption“

09/2023

Zusammenfassung

Open Data kann einen großen Beitrag zur Bekämpfung und Prävention von Korruption leisten und stellt einen Paradigmenwechsel in der Kontrolle staatlicher und wirtschaftliche Akteure dar, denn erst Open Data ermöglicht eine effektive, unkomplizierte und umfassende Analyse und Nachnutzung von Regierungsdaten. Auch die [Open-Data-Strategie der Bundesregierung](#) hebt den Nutzen offener Daten für ein offenes Regierungs- und Verwaltungshandeln hervor. Allerdings ist zurzeit die Verfügbarkeit und damit auch die Nutzung von offenen Daten in Deutschland aufgrund der verbesserungswürdigen Rechtslage und der zurückhaltenden Verwaltungspraxis noch sehr dürftig ausgeprägt, gerade in Bezug auf Daten, die für eine wirkungsvolle zivilgesellschaftliche Kontrolle notwendig wären. Daher formuliert Transparency Deutschland die folgenden Kernforderungen:

- Es müssen mehr korruptionsrelevante Daten, wie beispielsweise Registerinformationen, Vergabedaten, Lobbyaktivitäten und Parteienfinanzierung in einem offenen, maschinenlesbaren Format leicht zugänglich und zentral per API und Direktdownload zur Verfügung gestellt werden.
- Dabei sind bundeseinheitliche, gut dokumentierte Datenstandards und eindeutige Kennungen zu verwenden, die eine einfache Zusammenführung von Daten erlauben.
- Bei der Bereitstellung der Daten sollte auf einen niedrighwelligen Zugang geachtet werden, durch zugängliche Visualisierungen und interaktive Plattformen.
- “Open Data by Default” sollte auch in einem Bundestransparenzgesetz festgeschrieben werden, wobei Abweichungen von diesem Grundsatz zu begründen und nur unter eng gefassten Umständen erlaubt sein sollten.
- Die Einbeziehung der Zivilgesellschaft bei der Bereitstellung von Open Data sollte institutionalisiert werden und die Zuständigkeit der Informationsfreiheitsbeauftragten als Ansprechstelle für Open Government Data erweitert werden. Kurzfristig sollte unter Einbeziehung der Zivilgesellschaft die Veröffentlichung bestimmter, besonders relevanter Datensätze priorisiert werden.

Was ist Open (Government) Data?

Unter *Open Data* beziehungsweise *Open Government Data* versteht Transparency Deutschland Daten, die von staatlichen Stellen stammen und die der [Open Definition](#) entsprechend von jedermann frei verwendet, nachgenutzt und verbreitet werden können. Auch Daten, die von privaten und zivilgesellschaftlichen Organisationen stammen, sobald sie einen Beitrag zum Gemeinwohl beitragen können oder ihre Veröffentlichung im Interesse der Allgemeinheit liegt, werden von Transparency Deutschland als Open Data verstanden. Insbesondere müssen dafür die folgenden Kriterien erfüllt sein:

- **Verfügbarkeit und Zugänglichkeit:** Die Daten sollen vollständig und gebührenfrei im Internet verfügbar gemacht werden, wobei auf eine zweckmäßige und modifizierbare Form zu achten ist.
- **Wiederverwendung und Nachnutzung:** Die Daten müssen unter denjenigen Bedingungen bereitgestellt werden, die die Wiederverwendung, Nachnutzung und Verbindung mit anderen Datensätzen in einfacher Weise (z.B. durch die Nutzung von eindeutigen Kennungen) erlauben. Die Daten müssen maschinenlesbar sein und sollten standardisierte Metadaten verwenden.
- **Universelle Beteiligung:** Jede Person muss grundsätzlich in der Lage sein, die Daten zu nutzen und zu verwenden. Es darf keine Diskriminierung gegen beziehungsweise Einschränkung für bestimmte Handlungsfelder, Einzelpersonen oder Gruppen vorliegen.

Der deutsche Status Quo

Auch wenn es in Deutschland in den letzten Jahren Fortschritte bei der Veröffentlichung von korruptionsrelevanten Informationen wie Lobbydaten und Nebeneinkünften gegeben hat, bestehen weiterhin große Defizite bei der Art und Weise, wie diese Informationen zur Verfügung gestellt werden. Zwar legt §7 des Datennutzungsgesetzes die Bereitstellung von Daten "soweit möglich" in einem offenen Format fest. Dies lässt jedoch einen sehr großen Spielraum für Datenbereitsteller diesen Grundsatz in der Praxis zu unterlaufen. §12a des E-Government-Gesetzes schränkt diesen schwachen Veröffentlichungsgrundsatz um einen breiten Katalog an Ausschlussstatbeständen weiter ein. So sind beispielsweise alle Daten ausgenommen, die nicht nur "ausschließlich Tatsachen enthalten, die außerhalb der Behörde liegende Verhältnisse betreffen." Allerdings sind gerade behördeninterne Daten für die Vermeidung und Bekämpfung von Korruption von entscheidender Bedeutung. Des Weiteren bietet die derzeitige Rechtslage keine Pflichten oder Anreize auf die Möglichkeit der Bereitstellung von Open Data hinzuwirken. Auch werden explizit keine Rechtsansprüche begründet. Im Ergebnis bleiben Bürger:innen auf Auskunftsbegehren im Rahmen

des Informationsfreiheitsgesetzes angewiesen und eine Pflicht zur proaktiven Veröffentlichung besteht nur in wenigen Fällen. So trägt die aktuelle Rechtslage in einem nur unzufriedenstellenden Umfang zur Verbreiterung des Angebots von Open Data bei. Insgesamt wird die Nutzung dieser Daten im Sinne demokratischer Kontrolle erschwert. Doch auch in den wenigen Fällen, wo korruptionsrelevante Daten bereits veröffentlicht werden, werden diese in der Praxis noch zu häufig

- nicht proaktiv veröffentlicht, sondern müssen per Informationsfreiheitsantrag beantragt werden
- in suboptimalen Formaten veröffentlicht, wie beispielsweise gescannte Dokumente als nicht-maschinenlesbare und nicht durchsuchbare PDFs, als nicht disaggregierbare Zusammenfassungen oder als unstrukturierte, dezentral verteilte Informationen
- in sogenannten “Datenfriedhöfen” (also schlecht strukturierten Transparenzportalen) versenkt werden, wo sie praktisch unauffindbar bleiben.
- nur nominell und unvollständig veröffentlicht, wobei wichtige korruptionsrelevante Eckdaten weggelassen werden
- nicht in einem einheitlichen Datenstandard veröffentlicht, der für eine unkomplizierte Analyse nötig wäre
- ohne eindeutige Kennung (“unique identifier”) veröffentlicht, die eine eindeutige Zuordnung oder die Zusammenführung relevanter Daten aus verschiedenen Quellen zulassen würde
- in minderer Datenqualität mit unzuverlässigen und sogar falschen Informationen veröffentlicht, sodass eine effektive Nutzung kaum möglich ist.

Von wenigen Good Practice-Fällen abgesehen stellt dies zurzeit noch die Regel bei der Veröffentlichung korruptionsrelevanter Daten dar. Dies ist ein Mitgrund, warum die Nutzung von Daten in Deutschland bisher weit hinter anderen europäischen Ländern zurückliegt. Im internationalen Vergleich liegt Deutschland laut dem [Global Data Barometer](#) lediglich im globalen Mittelfeld. Zwar ist zu begrüßen, dass die Bundesregierung in ihrer Open-Data-Strategie von 2021 den grundsätzlichen Nutzen von Open Data anerkennt und die Bereitstellung von offenen Regierungsdaten ausweiten möchte, jedoch wurde das Potential für die Bekämpfung und -prävention von Korruption auch hier nicht substantiell berücksichtigt.

Nutzen für die Bekämpfung von Korruption

Transparentes Regierungshandeln ist ein grundlegendes Mittel im Kampf gegen Korruption. Informationen müssen jedoch in einem geeigneten Format zur Verfügung gestellt werden, damit

Transparenz seine Wirkung als Instrument der Korruptionsbekämpfung und -prävention entfalten kann. Unzweckmäßige technische Einschränkungen oder die Verwendung suboptimaler Formate machen den erwarteten Mehrwert von Transparenz häufig zunichte. Bei der Herstellung von Transparenz muss daher darauf geachtet werden, Informationen in einem geeigneten – den oben aufgeführten Kriterien entsprechenden – “Open Data”-Format zu veröffentlichen.¹

Zivilgesellschaftliche Organisationen, Medien und insbesondere einzelne Bürger:innen können mit begrenzten Ressourcen aufwendige Recherchen häufig nicht leisten – so bleibt die angestrebte Wirkung von Transparenz hinter ihrem Potential zurück. Durch eine grundsätzliche Bereitstellung korruptionsrelevanter Informationen als Open Data können Transparenzangebote der Politik auch effektiv genutzt werden – von zivilgesellschaftlichen Organisationen, aber auch durch Strafverfolgungsbehörden und anderen staatlichen Stellen, was viele langwierige Amtshilfeersuchen überflüssig machen kann.

Wie sollte Open Data zur Verfügung gestellt werden?

Ob Regierungsdaten eine breite Nutzung durch Bürger:innen, Journalist:innen und der Zivilgesellschaft erfahren, bleibt in hohem Maße abhängig von der Art und Weise, wie diese zur Verfügung gestellt werden. So sollte Zugang und Nutzung von Daten grundsätzlich gebührenfrei, ohne Einschränkungen und gemeinfrei oder unter einer offenen Lizenz erfolgen können. Der Informationszugang muss grundsätzlich anonym und ohne Registrierungs- oder Identifizierungserfordernis möglich sein. Informationen sollten aktuell zu halten sein. Entsprechend sollten Aktualisierungen als Nebenprodukt digitalisierten Verwaltungshandelns automatisch und in Echtzeit erfolgen. Wo es sinnvoll ist, sollte im Sinne der Nachvollziehbarkeit eine Versionierung der Datensätze erfolgen und ältere Versionen weiterhin zugänglich bleiben. Auch synoptische Darstellungen bieten sich an. Für die Veröffentlichung der Daten sollten vorzugsweise offene und häufig genutzte Dateiformate wie zum Beispiel CSV, XML und JSON genutzt werden. Dabei ist auch auf eine adäquate Dokumentation veröffentlichter Datensätze sowie die Bereitstellung beschreibender Metadaten zu achten, da ansonsten Informationen häufig weder nachvollziehbar noch interpretierbar sind. Dies umfasst auch die Dokumentation der Vollständigkeit und Validität der Daten.

Um ihren vollen Nutzen entfalten zu können, muss bei der Veröffentlichung der Daten auf eine hohe Datenqualität geachtet werden. Dafür braucht es einerseits einheitliche Datenstandards, von denen bereits für etliche Bereiche praxiserprobte Vorschläge existieren, sowie andererseits eindeutige Kennungen (“unique identifiers” oder IDs), beispielsweise für beteiligte Dienststellen, Parteigliederungen, Abgeordnete oder Lobbyverbände. Eine einheitliche Anwendung der IDs in

allen veröffentlichten Datensätzen erlaubt darüber hinaus eine einfache und unkomplizierte Zusammenführung der Informationen und macht aus einfachen Open Data Formaten sogenannte "Linked Open Government Data" (LOGD), die eine sektorübergreifende Analyse von Korruptionsrisiken zulassen. Schließlich sollten gleichartige, aber dezentral erhobene Daten wie zum Beispiel kommunale Auftragsvergaben zentral auf einer (möglichst bundesweiten) Plattform gebündelt und veröffentlicht werden, anstatt dass sie nur dezentral auf den Webauftritten der Kommunen auffindbar zu veröffentlichen.

Ausgehend von diesen so zu schaffenden technischen Rahmenbedingungen sollte bei der Bereitstellung der Daten auf die Bedürfnisse und technischen Kompetenzen verschiedener Nutzer:innengruppen geachtet werden. Während grundsätzlich alle offenen Daten über offene und gut dokumentierte Schnittstellen (sogenannte "Application Programming Interfaces", kurz APIs) und als Direktdownload abrufbar sein sollten, um eine Nachnutzung durch die Zivilgesellschaft zu ermöglichen, dürfen auch niedrighwellige Angebote wie Visualisierungen und interaktive Formate nicht fehlen. Während APIs den Aufbau eigener Plattformen und anderer Anwendungen sowie die Zusammenführung und Analyse verschiedener Datensätze ermöglichen, helfen Visualisierungen durch beispielsweise Dashboards oder auch Suchfunktionen bei der Recherche konkreter Einzelfälle, deren Verwendung kein technisches Wissen der Benutzer:innen voraussetzt.

Institutioneller Rahmen

Um eine bundesweite und möglichst einheitliche Praxis bei der Bereitstellung von Open Data sicherzustellen und einen Rechtsanspruch auf Open Data zu begründen, ist die Verabschiedung eines Bundestransparenzgesetzes mit einer starken Verankerung von Open Data unerlässlich. Daher sollte ein Bundestransparenzgesetz das Prinzip "Open Data by Default" festschreiben und möglichst hohe Anforderungen an eine Abweichung von diesem Grundsatz stellen. So sollten Datensätze, die personenbezogene Daten enthalten, nicht pauschal zurückgehalten werden können, sondern ggf. anonymisiert oder pseudonomisiert zur Verfügung gestellt werden. Außerdem müssen verbindliche, bundeseinheitliche Vorgaben gemacht werden, insbesondere zur Nutzung von Standards, IDs, Schnittstellen, Formaten und Dokumentation, statt diese Ausgestaltung den einzelnen potentiell transparenzpflichtigen öffentlichen Stellen in Deutschland zu überlassen. Ebenfalls sollte geregelt werden, dass alle Veröffentlichungen gemeinfrei oder zumindest offen zu lizenzieren sind. Darüber hinaus sollten institutionalisierte Konsultationsformate bei der Bereitstellung neuer sowie der Verbesserung der Qualität bestehender Datensätze mit zivilgesellschaftlichen und privaten Organisationen geschaffen werden, um eine Orientierung an den tatsächlichen Bedürfnissen von Nutzenden sicherzustellen. In diesem Zusammenhang ist auch die Einführung einer institutionellen Ombudsperson empfehlenswert, die die Interessen und Rechte von

Transparenzbegehrenden gegenüber transparenzpflichtigen Stellen vertritt und nötigenfalls per Anordnungsrecht durchsetzt. So könnte die Zuständigkeit der Datenschutz- und Informationsfreiheitsbeauftragten um diese Funktion erweitert werden. Analog zu der Verabschiedung eines Transparenzgesetzes auf Bundesebene sollten die Länder entsprechende Landestransparenzgesetze einführen. Dort wo noch keine ausreichenden einheitlichen Bundes- und Landesregelungen getroffen wurden, sollte durch Austauschformate, Best Practices und Kapazitätsentwicklungsinitiativen wie das Kompetenzzentrum Open Data die Verbreitung von Open Data gefördert werden. Einen wichtigen Beitrag kann auch die anstehende Registermodernisierung leisten, bei welcher die Bereitstellung von Open Data mitberücksichtigt werden sollte. Darüber hinaus sind eine zeitnahe Etablierung und ausreichende finanzielle und personelle Ausstattung des Dateninstituts geboten, um den Datenaustausch über Sektorengrenzen hinweg zu fördern.

Welche Daten sollten als Open Data zur Verfügung gestellt werden?

Für die Stärkung der Demokratie sind Daten, die zur Prävention und Verfolgung von Korruption relevant sind, von höchster Bedeutung. Dazu zählen beispielsweise Daten zur Rechtsstaatlichkeit, Leistungsfähigkeit der Verwaltung, Rechenschaftspflicht und Transparenz der öffentlichen Hand. Die Erfahrungen mit Open Data und Transparenzgesetzen in verschiedenen Bundesländern sowie anderen Mitgliedsstaaten der Europäischen Union haben gezeigt, dass es Sinn ergibt, zunächst besonders relevante Datensätze zu priorisieren, statt alle Daten der Verwaltung auf einen Schlag als Open Data zur Verfügung stellen zu wollen. So definiert die Europäische Kommission im Zuge der "Open Data"-Richtlinie 2019/1024 in ihrer Durchführungsverordnung 2023/138 bestimmte sogenannte hochwertige Datensätze, die die Mitgliedsstaaten innerhalb von 16 Monaten zu veröffentlichen haben. Leider wurde hier die Gelegenheit versäumt, korruptionsrelevante Daten in die Liste hochwertiger Datensätze aufzunehmen. Nichtsdestotrotz wäre ein vergleichbares Vorgehen auch in Deutschland sinnvoll, wobei unter Einbeziehung der Zivilgesellschaft auch auf die Aufnahme von korruptionsrelevanten Datensätzen geachtet werden sollte.

Im Bereich der Korruptionsbekämpfung und -prävention bietet es sich an, zunächst die Qualität und Veröffentlichungspraxis von bereits in offenen Formaten vorliegenden Informationen wie die Daten des Lobbyregisters oder für Bergbaukonzessionen weiter aufzuwerten. Zudem sollten bereits veröffentlichte, aber noch nicht als offene Daten vorliegende Informationen wie Parteispenden und die Nebeneinkünfte von Bundestagabgeordneten in ein offenes Format überführt werden. Weitere Beispiele für die Veröffentlichung offener Verwaltungsdaten betreffen das öffentliche Vergabewesen, einen legislativen Fußabdruck, das Handelsregister, das Wettbewerbsregister,

Subventions- und Förderdaten, Prüfergebnisse der Rechnungshöfe oder auch Erwerbstätigkeiten im Sinne von §6a Bundesministergesetz. Außerdem sollte die Bundesregierung zügig ihrer Verpflichtung nachkommen, die Public Integrity-Indikatoren der OECD zu bestücken. Darüber hinaus sind noch viele weitere Datensätze von potentieller Bedeutung für eine effektive demokratische Kontrolle und den Kampf gegen Korruption durch die Zivilgesellschaft, die es in einem kollaborativen Prozess zu identifizieren gilt.

In den letzten Jahren ist es besonders deutlich geworden, dass die Bekämpfung von Geldwäsche – als einer besonderen Ausprägung von Korruption – von Open Data profitiert. Zentrale Rolle spielen dabei die nationalen Transparenzregister über sogenannte wirtschaftliche Berechtigte (*“Beneficial Ownership”*). Neben dem Nutzen für staatliche Stellen wie die Financial Intelligence Unit (FIU) oder Strafverfolgungsbehörden und Finanzämter können zivilgesellschaftliche Organisationen und Journalist:innen zu der Aufdeckung von Geldwäschenetzwerken und Korruption beitragen. Offene Zugänge können darüber hinaus bei der Verbesserung der Datenqualität helfen, durch das Aufzeigen von Unstimmigkeiten mit anderen Quellen.

Juni 2023

Federführend Julian Brummer, Leiter der Arbeitsgruppe Vergabe