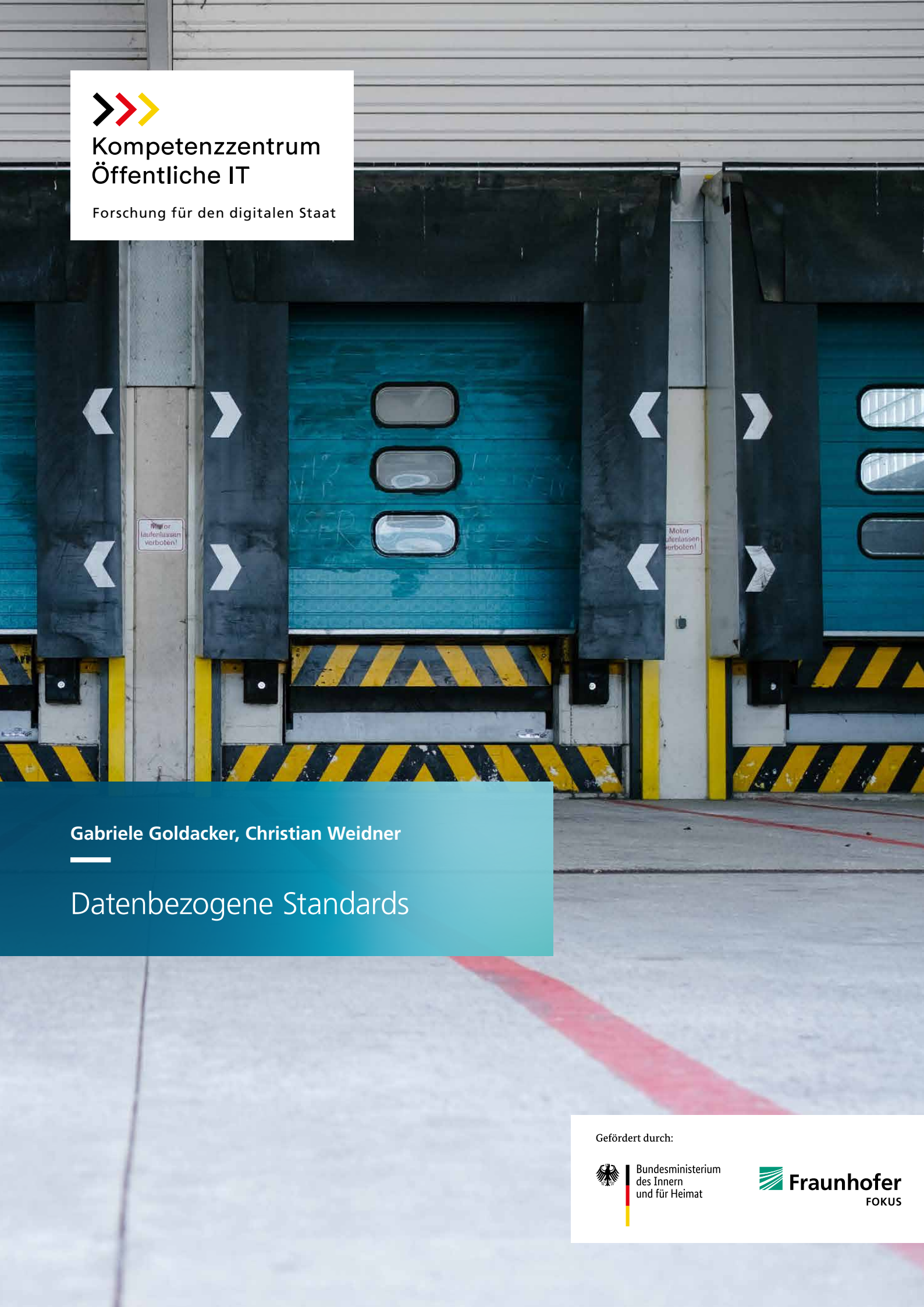


Kompetenzzentrum Öffentliche IT

Forschung für den digitalen Staat

Gabriele Goldacker, Christian Weidner

Datenbezogene Standards

Gefördert durch:

Bundesministerium des Innern und für Heimat

Fraunhofer FOKUS

Impressum

Autor:innen:
Gabriele Goldacker, Christian Weidner

Gestaltung:
Reiko Kammer

Herausgeber:
Kompetenzzentrum Öffentliche IT
Fraunhofer-Institut für Offene Kommunikationssysteme FOKUS
Kaiserin-Augusta-Allee 31, 10589 Berlin
Telefon: +49-30-3463-7173
info@oeffentliche-it.de
www.oeffentliche-it.de
www.fokus.fraunhofer.de

ISBN: 978-3-948582-22-7

1. Auflage März 2024

Icons für Infografik: https://fontawesome.com/

Das letzte Abrufdatum der Onlinequellen in den Fußnoten ist der 04.03.2024.

Bildnachweis

Seite	Autoren	Quelle
1	Hannes Egler	unsplash.com
6	Jason Thompson	unsplash.com
11	Chuttersnap	unsplash.com
14	Chuttersnap	unsplash.com
20	Mika Baumeister	unsplash.com
32	Ruchindra Gunasekara	unsplash.com
36	Venti Views	unsplash.com
39	Timelab	unsplash.com
41	Marcin Jozwiak	unsplash.com

Vorwort

Daten – auch solche, die von öffentlichen Einrichtungen erhoben, gespeichert, verarbeitet oder (nicht) bereitgestellt werden – stehen im Fokus unterschiedlicher Debatten. Nur einige Beispiele: Die weltweit aufgebauten Datenberge wachsen rasant und exponentiell. Datensilos und Datenfriedhöfe werfen Fragen nach Sinn und Effizienz von Erhebung und Speicherung auf, während Datenschutzaspekte, -bedenken und -verstöße die oft noch nebulöse Diskussion um Datentreuhänderschaft befördern.

Voraussetzung für jegliche – nicht nur elektronische – sinnvolle Nutzung von Daten sind datenbezogene Standards: In welcher Form werden Daten gespeichert oder übermittelt, wie werden sie durch Metadaten angereichert, sodass sie von beliebigen zukünftigen Nutzenden »verstanden« und möglichst unkompliziert weiterverarbeitet werden können? Wie kann der gezielte automatisierte Zugriff auf offene Daten durch a priori unbekannte Dritte erfolgen? Wie werden Daten erfolgreich gegen Nichtverfügbarkeit, Manipulation und unzulässige Offenlegung geschützt? Wie wird sichergestellt, dass Datenerhebung, -speicherung und -nutzung anerkannten ethischen Grundsätzen entsprechen? Die Antwort auf alle diese Fragen lautet: Durch die Anwendung datenbezogener Standards.

In einem gegebenen Kontext ist allerdings nicht jeder formal »passende« Standard auch gleichzeitig ein guter Standard: Zu häufig werden neue Standards mit nur geringem oder sogar nur vermeintlichem Mehrwert geschaffen und isoliert zur Vorgabe gemacht!

Standards für die elektronische Abbildung von Daten für Speicherungs- oder Übermittlungszwecke gibt es bereits viele. Auch für bestimmte Arten des automatisierten Datenabrufes gibt es insbesondere für Daten des öffentlichen Sektors relativ verbreitete Standards. Trotzdem werden in Datenverarbeitungsketten über Organisationsgrenzen hinweg noch nicht durchgängig einheitliche Daten- und Metadatenstandards benutzt, selbst wenn es geeignete Standards gibt. Noch seltener finden an sich geeignete (Meta-)Datenstandards ihren Weg von einer in eine andere Anwendungsdomäne.

Verantwortungsbewusster Umgang mit Daten erfordert sinnvolle, über unterschiedliche Akteure hinweg abgestimmte Maßnahmen zur Informationssicherheit. In anderen Fällen muss das oft schwierig zu bestimmende richtige Maß zwischen Abschottung und Offenlegung gefunden werden, damit Datenschätze nicht weiterhin ihrer Befreiung aus den Depots harren müssen. Die für diese Zwecke zahlreich vorhandenen Standards sind aber teilweise zu wenig bekannt.

Der Begriff Datenethik ist noch relativ jung, Fragen zum ethisch angemessenen Umgang mit Daten hingegen beschäftigen viele Gesellschaften bereits seit Zeiten, in denen Computer noch nicht in großem Umfang eingesetzt wurden. Auch wenn bei vielen derartigen Fragen noch um detaillierte »Standards« gerungen wird, ist es notwendig, die relevanten Akteure und deren Zuständigkeit sowie die eigenen Ansprechpartner zu kennen.

Gründe für die unterschiedliche Nutzung datenbezogener Standards und für »Standardbrüche« sind nicht nur unabhängig gewachsene Strukturen, die aktiv überwunden werden müssen, sondern auch Unkenntnis und Unsicherheit – darüber, wo man Standards findet, wie man deren Einsatzreife, Akzeptanz und Perspektive bewertet und wie man feststellt, ob sie für den eigenen Zweck nutzbar sind.

Dieses White Paper möchte Anfänger:innen wie Fortgeschrittene dazu ermutigen, (weitere) Schritte in die Welt der datenbezogenen Standards zu gehen, dabei auch ungewohnte Wege zu beschreiten und – bei wirklichem Bedarf – auch neue Wege aktiv zu erschließen.

Ihr Kompetenzzentrum Öffentliche IT

Inhalt

Vorwort **3**

1. Thesen **5**

2. Weshalb werden datenbezogene Standards benötigt – insbesondere im öffentlichen Sektor? **7**
2.1 Technische (Meta-)Datenstandards 8
2.2 Technischer und organisatorischer Schutz von Daten 9
2.3 Ethischer Umgang mit Daten 10

3. Der Weg zum geeigneten datenbezogenen Standard und seiner Umsetzung . . **12**
3.1 Expert:innen für IT-Standards 13

4. Reicht nicht die Vereinbarung eines einzigen datenbezogenen Standards? **15**
4.1 Technische (Meta-)Datenstandards 17
4.2 Technischer und organisatorischer Schutz von Daten 19
4.3 Ethischer Umgang mit Daten 19

5. Kategorien datenbezogener Standards im öffentlichen Sektor **21**

6. Was macht einen guten (datenbezogenen) Standard aus? **33**

7. Vom guten Umgang mit (datenbezogenen) Standards **37**
7.1 Generelle Gesichtspunkte 37
7.2 Eigene Standards 37

8. Übergeordnete Handlungsempfehlungen **40**

1. Thesen

Aufgrund fehlender Informationen über existierende Standards wird zu oft »das Rad neu erfunden«.
Selbst dann, wenn ein Standard in Teilen des öffentlichen Sektors etabliert ist und erfolgreich genutzt wird, fehlt in anderen Teilen die Möglichkeit, sich darüber zu informieren. Darüber hinaus sind außerhalb des öffentlichen Sektors nicht nur Normen, sondern vielfältige offengelegte datenbezogene Standards unterschiedlicher Gremien für Nutzungszwecke verbreitet, die auch in öffentlichen Stellen auftreten. Gerade angesichts der Vielfalt der datenbezogenen Standards und ihrer Quellen mangelt es an einem übergreifenden diesbezüglichen Wissensmanagement des öffentlichen Sektors.

Datenbezogene Standards müssen angemessen und für alle Betroffenen praxistauglich sein.
Sowohl bei den datenbezogenen Standards selbst als auch bei der Verpflichtung zu deren Umsetzung muss »das richtige Maß« gefunden werden. Ein Standard, der absehbar nicht ausreichend zukunftsfest ist, ist ebenso unangemessen wie ein Standard, der weit mehr verpflichtend regelt, als in der Praxis notwendig ist. Insbesondere Standards, die einseitig festgelegt werden und nur den Bedarf und den Kontext der festlegenden Seite berücksichtigen, können unangemessenen Anpassungsaufwand bei den weiteren Betroffenen verursachen.

Mehrere aufeinander aufbauende Standards sind besser als ein Alleskönner.
Ein datenbezogener Standard muss und sollte nicht als ein Monolith alle Aspekte einer Datennutzung – z. B. Verarbeitung, Übertragung und Speicherung – originär abdecken. Besser ist es, eine geeignete Modularisierung vorzunehmen und dabei insbesondere bereits existierende Module – auch anderer Domänen – wiederzuverwenden. Dies erleichtert die Erstellung in sich konsistenter Standards, deren spätere Anpassung an neue Anforderungen und technischen Fortschritt sowie den Austausch einzelner Komponenten.

Ohne standardisierte Metadaten keine Nachnutzung.
Zumindest die – im öffentlichen Sektor immer wichtiger werdende – Nachnutzung von Daten erfordert nicht nur, dass die Daten selbst in einem standardisierten Format vorliegen, sondern auch, dass bestimmte spezifische Metadaten vorhanden sind und ebenfalls einem standardisierten Format folgen. Nicht umsonst setzt auch die Datenstrategie der Bundesregierung auf standardisierte Datenbeschreibungen.[1]

Die Nutzung von datenbezogenen Standards erfordert Kenntnisse des funktionalorganisatorischen Rahmens.
Für die Spezifikation von (datenbezogenen) technischen Standards werden häufig Spezifikationsstandards (z. B. UML – Unified Markup Language – oder bestimmte Teile der XÖV-Rahmenvorgaben) benutzt, die man kennen und verstehen muss, um die interessierenden Standards verstehen und eigene Standards gut (wieder)verwendbar spezifizieren zu können. Für viele (Meta-)Datenstandards gibt es wiederverwendbaren Referenzcode in gängigen Programmiersprachen oder Testumgebungen, in denen die Korrektheit eigener Implementierungen geprüft werden kann. Profile und Conformance Statements erleichtern bei Standards mit viele Optionen das Realisieren von Interoperabilität und Konformität (Conformance) unterhalb des kompletten Standards.

Auch für Daten gilt: Nicht alles, was möglich ist, ist auch zulässig und vertretbar.
Neue technische Möglichkeiten erwecken seit einiger Zeit den Eindruck, hartnäckige Probleme (nur) durch die Erhebung, Erfassung und Verarbeitung noch größerer und vielfältigerer Datenmengen in den Griff bekommen zu können. Dabei müssen jedoch stets die Angemessenheit und die »Risiken und Nebenwirkungen« der Lösung, z. B. im Hinblick auf die informationelle Selbstbestimmung der Bürger:innen, im Blick behalten werden. So ist beispielsweise noch unbekannt, welche besonderen Kategorien personenbezogener Daten zukünftig, z. B. durch den Einsatz künstlicher Intelligenz, aus welchen Kombinationen von jeweils für sich genommen unkritischen Einzeldaten bestimmt werden können. Deshalb sind auch explizite Standards für den ethischen Umgang mit Daten erforderlich.

1 Die Bundesregierung: »Fortschritt durch Datennutzung – Strategie für mehr und bessere Daten für neue, effektive und zukunftsweisende Datennutzung«, August 2023; https://www.bmi.bund.de/SharedDocs/pressemitteilungen/DE/2023/08/nationale-datenstrategie.html


MARITIME
tex
Lloyd
SOD

2. Weshalb sind datenbezogene Standards notwendig – besonders im öffentlichen Sektor?

Die Datenstrategie der Bundesregierung bringt es so auf den Punkt: »Normen und Standards sind eine Grundlage für ein langlebig nutzbares Datenökosystem«.[2]

Standards sind immer dann hilfreich, wenn Dinge oder Prozesse zusammenpassen müssen und dabei häufig gleichartige Paarungen auftreten oder viele Dinge bzw. Prozesse gleichzeitig zueinander in Beziehung treten. Entsprechende Standards beschreiben die Anforderungen an die jeweils betroffenen Dinge oder Prozesse bzw. deren notwendige Eigenschaften. Andere Standards dienen dazu, allgemeinere Eigenschaften oder Leistungen nach einheitlichen Kriterien beschreiben bzw. kategorisieren zu können und so den Vergleich mit Erwartungen und Anforderungen zu erleichtern. Standards vermeiden in diesen Fällen, dass jedes Mal und zwischen allen Beteiligten zunächst Absprachen getroffen werden müssen.

Im engeren Sinne und in digitalem Kontext wird oft erst dann von einem (datenbezogenen) Standard gesprochen, wenn damit auf eine Vereinbarung eines Gremiums mit Vertretern mehrerer, rechtlich voneinander unabhängiger Organisationen verwiesen wird. Ebenso wie für Sprachen gilt: Von je mehr Unternehmen, Computerprogrammen, Personen usw. ein (datenbezogener) Standard verstanden und benutzt wird, umso nützlicher ist er. Ein Gremienstandard hat in der Regel mindestens zwei wesentliche Vorteile: Die Unterstützung durch das Gremium sichert ihm von vornherein eine größere Grundakzeptanz und Verbreitung. Und die Erarbeitung durch das Gremium, das typischerweise mit Experten aus (zumindest leicht) unterschiedlichen Kontexten besetzt ist, lässt eine breitere Anwendbarkeit und eine höhere Zukunftsfestigkeit des Standards erwarten, als wenn er von einem einzigen Unternehmen oder einer einzelnen Behörde für die eigenen Bedürfnisse entwickelt worden wäre. Eine besondere Form von Standards sind Normen. Sie werden in Gremien nationaler (DIN, DKE) oder internationaler Normungsorganisationen (z. B. CEN, CENELEC, ETSI, ISO und IEC) entwickelt und gepflegt.[3] Die Normungsorganisationen sind staatlich bzw. überstaatlich anerkannt, in ihnen haben Staaten bzw. ihre Vertreter:innen besondere Abstimmungsrechte und teilweise können ihnen Normungsaufträge der öffentlichen Hand erteilt werden.

Zur Festlegung eines Standards gehören zudem in der Regel zwei Vorgänge: die Spezifikation, also die inhaltliche Festlegung des Standards, und die Verpflichtung (oder Empfehlung), den Standard zu nutzen. Für ersteres sind meist Standardisierungsgremien (gegebenenfalls auch lediglich interne der öffentlichen Hand), für letzteres ein oder mehrere Beteiligte aufseiten der Daten Bereitstellenden oder Nutzenden zuständig. In technischen Kontexten gibt es im Hinblick auf den öffentlichen Sektor eine langjährig erfolgreiche »Arbeitsteilung«[4]: Die Rechtsetzung gibt in Rechtsnormen Ziele, Rechte und Pflichten (das »Was«) vor – z. B. zum barrierefreien Zugang zu den Websites und mobilen Anwendungen öffentlicher Stellen[5] –, identifiziert

2 Die Bundesregierung: »Fortschritt durch Datennutzung – Strategie für mehr und bessere Daten für neue, effektive und zukunftsweisende Datennutzung«, August 2023, S. 24; https://www.bmi.bund.de/SharedDocs/pressemitteilungen/DE/2023/08/nationale-datenstrategie.html

3 DIN: Deutsches Institut für Normung.
DKE: Deutsche Kommission für Elektrotechnik Elektronik Informationstechnik.
CEN: Comité Européen de Normalisation, Europäisches Komitee für Normung.
CENELEC: Comité Européen de Normalisation Electrotechnique, Europäisches Komitee für elektrotechnische Normung.
ETSI: European Telecommunications Standards Institute, siehe Europäisches Institut für Telekommunikationsnormen.
ISO: International Organization for Standardization, Internationale Organisation für Normung.
IEC: International Electrotechnical Commission, Internationale Elektrotechnische Kommission.

4 Eine detaillierte Beschreibung des Prozesses zum hier genutzten Beispiel findet sich unter https://digital-strategy.ec.europa.eu/de/policies/web-accessibility-directive-standards-and-harmonisation

5 »Richtlinie (EU) 2016/2102 des Europäischen Parlaments und des Rates vom 26. Oktober 2016 über den barrierefreien Zugang zu den Websites und mobilen Anwendungen öffentlicher Stellen (Text von Bedeutung für den EWR)«; https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX%3A32016L2102

gegebenenfalls entsprechende technische Standardisierungsbedarfe (für das »Wie«) und vergibt möglicherweise auch Normungsaufträge[6]. Standardisierungsgremien erarbeiten die entsprechenden Standards[7] – die Beschreibungen des »Wie« –, die eventuell wiederum von der Rechtsetzung offiziell anerkannt oder sogar für bestimmte Kontexte als verbindlich festgelegt werden[8]. In informations- und kommunikationstechnischem Kontext sind neben den staatlich und überstaatlich anerkannten Normungsorganisationen viele weitere Standardisierungsgremien aktiv und mit der Etablierung ihrer Standards erfolgreich. In diesem White Paper wird daher in der Regel allgemein von Standardisierungsgremien gesprochen, womit auch, aber nicht nur Gremien der Normungsorganisationen gemeint sind.

Vor allem in anderen, teilweise aber auch in technischen Kontexten ist die Arbeitsteilung nicht so klar. Oder die Rechtsetzung wird erst durch die technische Entwicklung oder die Ergebnisse von Standardisierungsaktivitäten angestoßen. Deshalb wird in diesem White Paper keine absolute Abgrenzung zwischen der Erarbeitung und der Verbindlichmachung von Standards vorgenommen, sondern der Blick auf alle datenbezogenen Regelungsbedarfe gerichtet, die für öffentliche oder privatwirtschaftliche Einrichtungen entstehen und von diesen selbstständig umgesetzt werden können.

Der öffentliche Sektor erhebt, verarbeitet und produziert große Mengen unterschiedlichster Daten und interagiert dabei mit vielfältigen internen und vor allem auch externen Akteuren, z. B. Leistungsempfänger:innen. Effizienter Umgang mit Daten und faire Ergebnisse der Datenverarbeitung sind daher nur mit dem Einsatz entsprechender Standards erreichbar. Dies gilt zum Beispiel, wenn der Allgemeinheit nach einem E-Government- oder einem Informationsfreiheitsgesetz offene Daten bereitgestellt, wenn elektronische Rechnungen bearbeitet oder wenn Daten mit anderen Behörden ausgetauscht werden müssen.

In diesem White Paper sind technische Datenstandards, technische Metadatenstandards, Standards für den technischen und organisatorischen Schutz von Daten sowie Standards (»Leitlinien«) für den ethischen Umgang mit Daten unter dem Begriff »datenbezogene Standards« zusammengefasst und damit gleichzeitig gegen die große Menge anderer Standards abgegrenzt, die sich nicht mit der Verarbeitung elektronischer Daten befassen, sondern z. B. mit Anforderungen an den Arbeitsschutz oder die Kompatibilität technischer Bauteile.

2.1 Technische (Meta-) Datenstandards

Technische Datenstandards sind Standards zur elektronischen Datenrepräsentation – vor allem für den Austausch von Daten zwischen im Einzelnen nicht vordefinierten Partnern, aber auch für die Datenspeicherung – sowie Standards für den (automatisierten) Datenabruf. Im weiteren Sinne umfasst der Begriff »technischer Datenstandard« auch die Spezialisierung »technischer Metadatenstandard«. Diese Spezialform ist daher in der Regel gedanklich mitzulesen, auch wenn sie zwecks besserer Lesbarkeit weggelassen wurde.

Wir sind ständig von den Ergebnissen der Anwendung mannigfaltiger Datenstandards umgeben, meist ohne uns dessen bewusst zu sein: Gesprochene Sprache, Schrift, Zahlen, Verkehrszeichen und andere Piktogramme – all dies und vieles mehr ist für uns nur verständlich, weil unsere Interpretation auf denselben Standards beruht, die die Urheber für die Erstellung benutzt haben.

Sehr frühe »Datenstandards« benutzten bereits die Steinzeitmenschen, indem sie möglichst naturgetreue Zeichnungen anfertigten, die wir oft auch heute noch sicher bekannten Naturphänomenen, Tierfamilien usw. zuordnen können. Allerdings sind dabei uns ebenfalls interessierende Metadaten – zusätzliche (Hintergrund-)Daten, die die relevanten Daten beschreiben – meist nicht aufgezeichnet worden, oder wir sind nicht in der Lage, diese aus den Zeichnungen zu extrahieren, weil wir nicht wissen, wie sie »kodiert« wurden. Metadaten wären beispielsweise der Entstehungszeitpunkt oder der Zweck der Zeichnungen (z. B. künstlerische Tätigkeit, Nachrichtenübermittlung, Datenspeicherung, rituelle Nutzung).

Ein Datenstandard wird bereits immer dann benötigt, wenn Daten – also (abstrakte) Abbildungen realer oder originär abstrakter Dinge, Umstände, Aussagen usw. – so aufgezeichnet, gespeichert oder übermittelt werden sollen, dass die Daten beliebigen Datennutzenden einen möglichst eindeutigen Rückschluss auf die Originale ermöglichen. Ein derartiger Datenstandard beschreibt, wie die Abbildung der Daten (»Kodierung«) zu erfolgen hat, und impliziert, wie sie (gedanklich) rückgängig gemacht werden kann. Gleichzeitig oder alternativ kann auch

6 Hier »C(2017)2585 – Normungsauftrag M/554: Durchführungsbeschluss der Kommission vom 27.4.2017 über einen Normungsauftrag an die europäischen Normungsorganisationen zur Unterstützung der Richtlinie (EU) 2016/2102 des Europäischen Parlaments und des Rates über den barrierefreien Zugang zu den Websites und mobilen Anwendungen öffentlicher Stellen«; https://ec.europa.eu/growth/tools-databases/enorm/mandate/554_de

7 Hier DIN EN 301549:2022-06: »Barrierefreiheitsanforderungen für IKT-Produkte und -Dienstleistungen; Englische Fassung EN 301 549 V3.2.1 (2021-03); Text Deutsch«; https://www.beuth.de/de/norm/din-en-301549/353869627

8 Hier »Durchführungsbeschluss (EU) 2021/1339 der Kommission vom 11. August 2021 zur Änderung des Durchführungsbeschlusses (EU) 2018/2048 über die harmonisierte Norm für Websites und mobile Anwendungen (Text von Bedeutung für den EWR)«; https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:32021D1339

ein anderer Datenstandard zum Einsatz kommen, mittels dessen Teile der Daten entfernt oder so verändert – z. B. verschlüsselt oder aggregiert – werden, dass damit gezielt die Rückschlussmöglichkeit für unberechtigte oder alle Nutzungsinteressent:innen eingeschränkt wird. Ein »technischer« Datenstandard ist erforderlich, wenn die Aufzeichnung, Speicherung, Änderung oder Übermittlung der Daten mit technischen Hilfsmitteln erfolgt, die den Standard einhalten (»implementieren«) müssen, damit die (verbleibenden) Daten für die Nutzungsberechtigten verständlich bleiben.

Neben den grundlegenden Abbildungsstandards bedarf es für viele Nutzungszwecke der Daten weiterer Standards, z. B. zu den (Mindest-)Anforderungen an die Genauigkeit der Abbildung, wenn Daten sensorisch erfasst werden, oder zu den für eine sinnvolle Nutzung der Daten erforderlichen Metadaten.

Besonders im Bereich des öffentlichen Sektors können zu den relevanten Metadaten nicht nur solche zum »Wann«, »Wo«, »Wie«/ »Womit«, »Wer« oder »Warum« gehören, sondern auch solche zur Korrektheit bzw. der Glaubwürdigkeit der Daten selbst sowie der übrigen Metadaten. Auf die konkreten entsprechenden Standards wird hier angesichts der Komplexität dieses Themas aber bewusst nicht speziell eingegangen.

Wird der Datenzugang über Mensch-Maschine-Schnittstellen – in der Regel von Portalen – bereitgestellt, verbirgt sich dahinter häufig die Nutzung Webbrowser-typischer Standards für den Datenabruf. Für den automatisierten Abruf – egal ob über einen derartigen oder einen anderen Zugang – sind zusätzliche Standards erforderlich, mittels derer die angeforderten Daten beschrieben werden.

(Gemeinsame,) explizierte Datenstandards sind umso wichtiger, je unterschiedlicher einerseits die aufzeichnende, erzeugende bzw. speichernde oder ändernde und andererseits die weiterverarbeitende Seite der Daten sind, bzw. wenn die konkrete weiterverarbeitende Seite gar nicht vorbestimmt ist. Fehlende Vorbestimmtheit ist zudem nicht nur bei sogenannten »offenen« (Verwaltungs-)Daten gegeben, auch innerhalb einer Organisation können sich Nutzungsbedarfe erst im Laufe der Zeit, z. B. durch neue gesetzliche Regelungen oder aus Praktikabilitätserwägungen, ergeben.

Unterschiedlichkeit zwischen Datenbereitstellenden und den jeweiligen Datennutzenden kann in vielen verschiedenen Dimensionen vorliegen. Ein Beispiel aus der Schriftsprache: Wer die zugrunde liegende Sprache nicht kennt und versteht, in der ein bestimmter Sachverhalt detailliert beschrieben ist, kann ihn sich trotz elektronischer Übersetzungshilfen oft nur schwer oder gar nicht erschließen. Wer die Schriftzeichen nicht kennt, ist oft nicht einmal in der Lage, diese in eine elektronische Übersetzungshilfe einzugeben. Bei unterschiedlichen Fachsprachen, die mit demselben Begriff verschiedene Dinge bezeichnen, sind sogar Muttersprachler ohne Kenntnis des Kontextes eventuell verloren. Letzteres gilt auch für die Fachsprachen der verschiedenen Bundesministerien, die zu unterschiedlichen Definitionen z. B. der Begriffe »Kind«[9] und »Einkommen«[10] geführt haben. Dies hat zur Folge, dass für ein gemeinsames, eindeutiges Verständnis des jeweils Gemeinten der definierende gesetzliche Kontext als Metadatum vorhanden sein muss.

Für die öffentliche Hand ist die gemeinsame Benutzung eines oder mehrerer, wohldefinierter Datenstandards vor allem dann notwendig, wenn elektronische Daten zwischen (Fach-)Anwendungen oder mit anderen Behörden, Bürger:innen oder Unternehmen ausgetauscht werden, da nur so die Grundlage für eine einheitliche Interpretation gegeben ist.

Neben den praktischen Vorteilen, die für eine Nutzung von Datenstandards sprechen, kann die Nutzung sogar verpflichtend sein, z. B. wenn die öffentliche Hand unter die sogenannte PSI-Richtlinie der EU[11] fallende hochwertige Datensätze (gemäß der Definition der Richtlinie) bereitstellt.

2.2 Technischer und organisatorischer Schutz von Daten

Daten- bzw. Informationssicherheit umfasst nach breit akzeptierter Definition die Bereiche Datenverfügbarkeit, Datenintegrität und Datenvertraulichkeit.[12] Informationssicherheit macht dabei keinen prinzipiellen Unterschied zwischen »natürlichen«

9 Siehe z. B. Mohabbat Kar, Resa; Thapa, Basanta E. P.; Hunt, Simon Sebastian; Parycek, Peter: »Recht digital: maschinenverständlich und automatisierbar – Impuls zur digitalen Vollzugstauglichkeit von Gesetzen«. Kompetenzzentrum Öffentliche IT. 1. Auflage, September 2019; https://www.oeffentliche-it.de/documents/10181/14412/Recht+Digital+-+Maschinenverst%C3%A4ndlich+und+automatisierbar

10 Siehe z. B. Nationaler Normenkontrollrat (Auftraggeber): »Digitale Verwaltung braucht digitaltaugliches Recht – Der modulare Einkommensbegriff«. 1. Auflage, Juni 2021; https://www.normenkontrollrat.bund.de/Webs/NKR/SharedDocs/Downloads/DE/Gutachten/2021-digitale-verwaltung-braucht-digitaltaugliches.pdf?__blob=publicationFile&v=7

11 Siehe »Datennutzungsgesetz vom 16. Juli 2021 (BGBl. I S. 2941, 2942; 4114)«; https://www.gesetze-im-internet.de/dng/, das die »Richtlinie (EU) 2019/1024 des Europäischen Parlaments und des Rates vom 20. Juni 2019 über offene Daten und die Weiterverwendung von Informationen des öffentlichen Sektors (Neufassung)«, https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:32019L1024&rid=1, in deutsches Recht umsetzt, sowie die »Durchführungsverordnung (EU) 2023/138 der Kommission vom 21. Dezember 2022 zur Festlegung bestimmter hochwertiger Datensätze und der Modalitäten ihrer Veröffentlichung und Weiterverwendung«, https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:32023R0138&qid=1691491183280

12 Siehe z. B. Bundesamt für Sicherheit in der Informationstechnik (BSI): »BSI-Standard 200-1 – Managementsysteme für Informationssicherheit (ISMS)«, Version 1, Seite 8; https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Grundschutz/BSI_Standards/standard_200_1.pdf?__blob=publicationFile&v=2

Bedrohungen, wie technischem oder menschlichem Versagen, und vorsätzlichen Angriffen. Dies ist gerechtfertigt, da Schutzmaßnahmen oft gegen mehrere Arten von Bedrohungen wirken.

Unter der Annahme, dass nur relevante Daten vorgehalten und verarbeitet werden, besteht für alle Daten die Notwendigkeit, deren Integrität und angemessene Verfügbarkeit sicherzustellen, da sie ansonsten im Bedarfsfall nutzlos sind oder sogar – im Fall verfälschter, nicht oder nicht rechtzeitig verfügbarer Daten – ein Schaden eintreten kann. Für vielfältige Daten muss darüber hinaus Vertraulichkeit gewährleistet sein: Daten können unter Datenschutzvorschriften fallen oder es kann sich um andere schutzpflichtige Daten über Dritte (z. B. Unternehmensdaten) handeln; Daten bezüglich der eigenen Organisation könnten Dritten unerwünschte Wettbewerbsvorteile verschaffen und manche Daten müssen zum Schutz des Staates oder seiner Organe geheim gehalten werden. Während die jeweiligen Schutzvorschriften in Gesetzen und Ähnlichem typischerweise nur festlegen, was in welchem Umfang und wogegen zu schützen ist, dienen technische und organisatorische Schutzstandards der Beschreibung (und, durch Referenzierung solcher Standards, auch der Festlegung) der konkreten Maßnahmen und Mechanismen, um den erforderlichen Schutz sicherzustellen. Aus der Perspektive von Datenhalter:innen oder -verarbeiter:innen können alle derartigen Maßnahmen und Mechanismen unter dem Begriff der Informationssicherheit zusammengefasst werden.

Daten sind in allen drei Bereichen der Informationssicherheit stets nur so sicher wie das schwächste Glied in einer Erfassungs-, Verarbeitungs-, Übertragungs- und Speicherungskette. Um ein einheitliches minimales Datensicherheitsniveau über eine derartige Kette hinweg etablieren, aufrechterhalten und nachweisen zu können, ist die gemeinsame Nutzung geeigneter Schutzstandards unabdingbar.

Standards für Informationssicherheit können nicht nur technische Aspekte – wie beispielsweise konkrete Verschlüsselungsmechanismen für die Erreichung von Vertraulichkeit – umfassen, sondern z. B. auch Regeln für organisatorisches Vorgehen sowie für personen- bzw. rollenspezifische und bauliche Maßnahmen enthalten. Manche Aspekte liegen auch im Schnittbereich zwischen technisch und organisatorisch, z. B. Datensparsamkeit und das »Need-to-know«-Prinzip, bei denen zunächst – organisatorisch – festgelegt wird, welche Daten erfasst werden bzw. wer auf welche Daten Zugriff haben soll, und dies dann technisch unterstützt umgesetzt wird.

Das Thema Informationssicherheit umfasst eine große Menge an zu berücksichtigenden Gesichtspunkten, die hier nicht im Detail dargestellt werden. Als möglichen Einstieg verweisen wir auf die umfangreiche IT-Grundschutz-Dokumentation des Bundesamtes für Sicherheit in der Informationstechnik (BSI)[13], namentlich das IT-Grundschutz-Kompendium in der aktuellen Version[14].

2.3 Ethischer Umgang mit Daten

Nicht erst seit dem Wirken einer Datenethikkommission auf Bundesebene zeigt sich mit zunehmender Dringlichkeit, dass ethisch veranlasste Datenstandards notwendig sind. Dabei geht es nicht nur um besonders sensible Daten wie z. B. im Gesundheitsbereich, sondern mehr und mehr um Fragen, die das alltägliche Leben aller Bürger:innen betreffen, wie »soziales Wohlverhalten« (»Welche Daten dürfen erfasst und wie dürfen sie kombiniert werden?«) oder kritische Entscheidungen autonomer Fahrzeugsteuerungen (»Welche Daten dürfen für Entscheidungen herangezogen werden?«).

Bei vielen datenethischen Fragestellungen befinden sich Gesellschaft und Politik noch in der Phase, das »Was« von Grund auf zu bestimmen und zu detaillieren. Die öffentliche Hand muss daher, wenn sie auf entsprechenden datenethischen Erwägungen basierende Maßnahmen durchführen will, häufig zunächst auch das »Was« – gegebenenfalls vorläufig – eingrenzen und auf konkrete Situationen abbilden.

In den öffentlichen Debatten zum ethischen Umgang mit Daten stehen Fragestellungen rund um personenbezogene Daten im Vordergrund. Daneben gilt es, einen weiteren Themenkreis zu berücksichtigen, der eher selten explizit in diesen Kontext gestellt wird: die angemessene Präsentation von Information für und die Interaktion mit Menschen mit Beeinträchtigungen und Sprachschwierigkeiten.

Darüber hinaus werden vorwiegend in Fachkreisen Fragen zur ethischen Angemessenheit bestimmter Formen der Datenerzeugung und -erfassung bei Tieren diskutiert, z. B. wenn Daten durch Schmerzreize erzeugt werden oder Implantate zum Einsatz kommen, die nur der Datenerfassung dienen.

13 https://www.bsi.bund.de/DE/Home/home_node.html

14 BSI: »IT-Grundschutz-Kompendium«, Stand Februar 2023. https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Grundschutz/Kompendium/IT_Grundschutz_Kompendium_Edition2023.pdf?__blob=publicationFile&v=4.

80
P7
P7
P6
P6
STOP
STOP
90
STOP

3. Der Weg zum geeigneten datenbezogenen Standard und seiner Umsetzung

Datenbezogene Standards betreffen immer sowohl die Bereitstellenden als auch die Nutzenden von Daten, auf die die jeweiligen Standards angewendet werden. Oft müssen beide Seiten aktiv werden, indem sie beispielsweise zu technischen Standards konforme Kommunikationsschnittstellen für die Bereitstellung bzw. die Entgegennahme der Daten verfügbar machen und nutzen. Besonderes Augenmerk beider Seiten erfordern Standards für die Erhebung, Speicherung und Verarbeitung Menschen betreffender Daten, wenn die Daten nicht von diesen Menschen aktiv bereitgestellt, sondern beispielsweise von technischen Systemen über sie erhoben werden.

Bedarfsträger für Standards – nicht nur datenbezogener Art – sind in der Regel Unternehmen und öffentliche Stellen, aber auch zivilgesellschaftliche Organisationen. Bedarfe für technische Standards entstehen vor allem, wenn den Bedarfsträger:innen wirtschaftliche Nachteile bzw. hohe Kosten durch Inkompatibilitäten entstehen (würden). Wenn Einzelpersonen sich für bestimmte Aspekte technische Standards wünschen, werden sie trotzdem nur äußerst selten zu unmittelbaren Akteur:innen, da ihnen die (organisatorischen und finanziellen) Mittel fehlen, allein einen Standard zu etablieren. Für technische und organisatorische Schutzstandards sind wirtschaftliche oder rechtliche Risiken die treibenden Faktoren. Die öffentliche Hand kann zudem Bedarfe z. B. zum Schutz grundlegender Menschenrechte, von Verbrauchern oder der Umwelt identifizieren, was häufig durch entsprechende Schutzorganisationen initiiert wird.

Im optimalen Fall folgt auf die Feststellung des Bedarfes die Prüfung, ob ein fachlich geeigneter Standard bereits existiert. Dazu müssen zumindest die Eckpunkte des erforderlichen Standards feststehen, also der Zweck und die betroffenen Daten sowie wesentliche technische Rahmenbedingungen (beispielsweise, ob ein abfrageinduzierter Einzeldaten(satz)abruf oder eine vom Übermittler selbst getriggerte Massendatenübertragung erfolgen soll). Die Festlegung der Eckpunkte muss unter fachlichen Gesichtspunkten erfolgen, während für die Verfügbarkeitsprüfung Personal eingesetzt werden sollte, das mit datenbezogenen Standards vertraut ist.

Eine angemessene Verfügbarkeitsprüfung kann sich besonders im Bereich der technischen Datenstandards weder auf Standards des eigenen organisationalen Umfeldes noch auf Normen beschränken. Datenbezogene Standards des öffentlichen Sektors z. B. sind nicht unbedingt ausreichend relevant für die in den – breiter aufgestellten – zuständigen Normungsgremien aktiven Unternehmen und Interessengruppen, sodass oft kein entsprechendes Normungsvorhaben zustande kommt. Viele in der Wirtschaft verbreitete datenbezogene Standards stammen von Standardisierungsgremien, die keine Normungsorganisationen sind, z. B. der XML[15]-Standard des W3C. Für die Vergabe öffentlicher Aufträge hat die EU-Kommission sogar mittels einer Reihe von Durchführungsbeschlüssen[16] die Bezugnahme auf bestimmte Standards, die keine Normen sind, der Referenzierung von Normen gleichgestellt. Existieren für einen Anwendungsbereich kein verbindlicher Standard, aber geeignete Normen, gleichgestellte oder sonstige Gremienstandards, berücksichtigt die Verfügbarkeitsprüfung bevorzugt zunächst diese (in der angegebenen Reihenfolge).

15 World Wide Web Consortium (W3C): »Extensible Markup Language (XML) 1.0 (Fifth Edition)«, W3C Recommendation 26 November 2008 (Note: On 7 February 2013, this specification was modified in place to replace broken links to RFC4646 and RFC4647.), https://www.w3.org/TR/xml/, und »Extensible Markup Language (XML) 1.1 (Second Edition)«, W3C Recommendation 16 August 2006, edited in place 29 September 2006, https://www.w3.org/TR/xml/

16 Gestützt auf die »Verordnung (EU) Nr. 1025/2012 des Europäischen Parlaments und des Rates vom 25. Oktober 2012 zur europäischen Normung, zur Änderung der Richtlinien 89/686/EWG und 93/15/EWG des Rates sowie der Richtlinien 94/9/EG, 94/25/EG, 95/16/EG, 97/23/EG, 98/34/EG, 2004/22/EG, 2007/23/EG, 2009/23/EG und 2009/105/EG des Europäischen Parlaments und des Rates und zur Aufhebung des Beschlusses 87/95/EWG des Rates und des Beschlusses Nr. 1673/2006/EG des Europäischen Parlaments und des Rates«. Die Durchführungsbeschlüsse im Einzelnen: https://eur-lex.europa.eu/search.html?scope=EURLEX&text=durchf%C3%BChrungsbeschluss+%221025%2F2012%22+%22IKT-Bereich%22+%22technischen+spezifikationen%22&lang=de&type=quick&qid=1689671233476&FM_CODED=DEC_IMPL

Stellt sich die Notwendigkeit eines neuen Standards heraus, ist eine enge Kooperation zwischen fachlicher Seite und Expert:innen für IT-Standards erforderlich, um eine sachgerechte, stabile und in das bereits existierende Gefüge von Standards passende Lösung erarbeiten zu können. In der Regel ist es vorteilhaft, bereits im Vorfeld der Erarbeitung möglichst alle Bedarfsträger:innen und Betroffenen(gruppen) sowie ihre jeweiligen Expert:innen zusammenzubringen, um so zu einem konkreten und zukunftssicheren Bedarf zu gelangen. Bürger:inneninteressen z. B. können dabei bevorzugt von demokratisch legitimierten, demokratisch verfassten oder zumindest teilnahmeoffenen Organisationen eingebracht werden. Durch die Zusammenführung wird organisationsübergreifend entsprechendes Wissen zielgerichtet gebündelt und etwaigen Parallel- und Fehlentwicklungen frühzeitig entgegengewirkt. In dieser Phase ist es stets sinnvoll, zu prüfen, ob die Erarbeitung des Standards im Rahmen eines Standardisierungsgremiums durchgeführt werden kann.

Ist ein geeigneter Standard (oder ein Bündel geeigneter Standards) identifiziert oder geschaffen, dann wird wiederum eigenes oder beauftragtes IT-technisches Fachpersonal benötigt, das diesen Standard in eine konforme Implementierung überführt. Dies gilt sowohl für technische und technisch umgesetzte ethische Datenstandards als auch für viele Standards zum technischen Schutz von Daten sowie für Standards zum organisatorischen Schutz von Daten, die auf informationstechnische Unterstützung angewiesen sind.

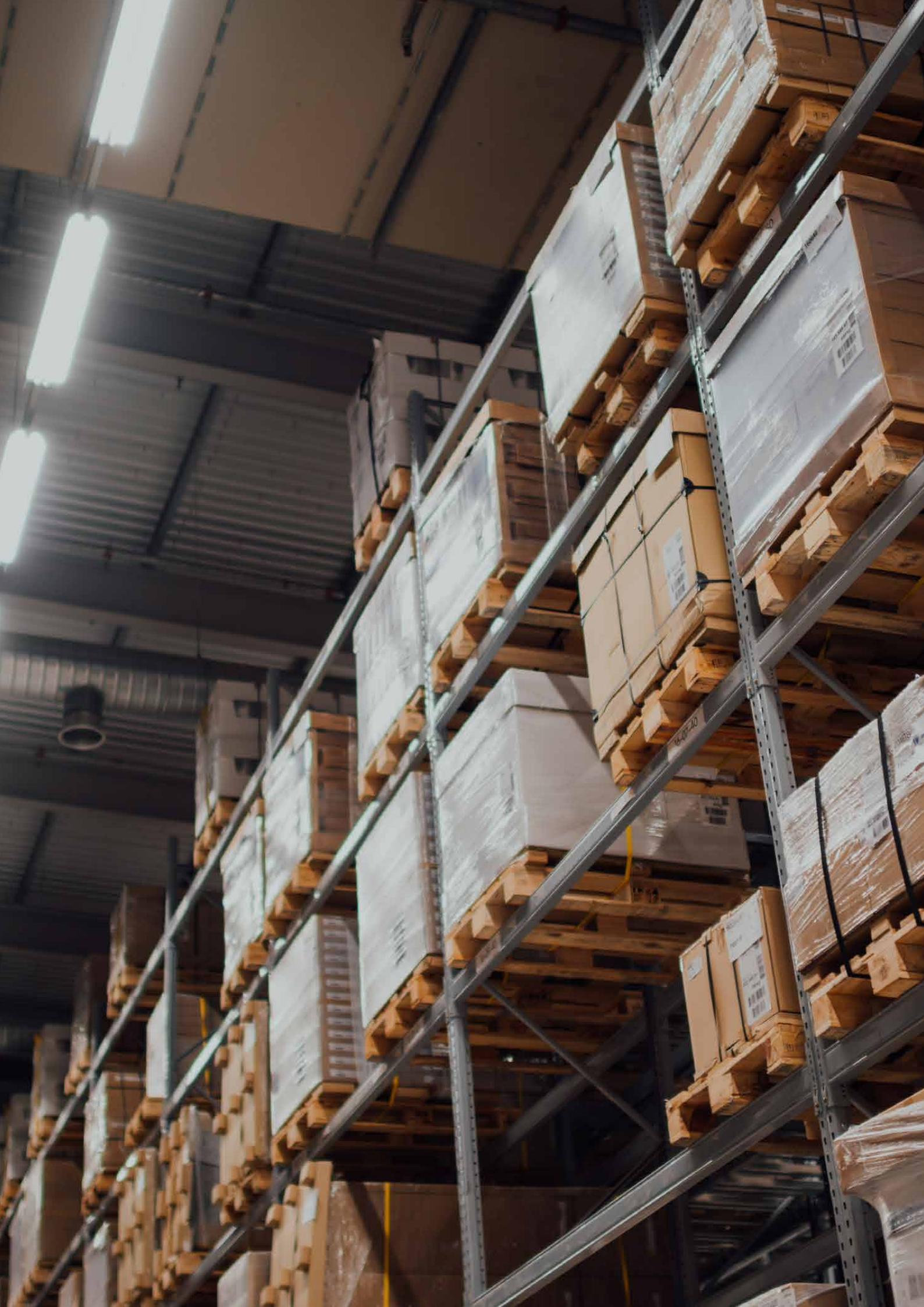
Schließlich muss die implementierte Lösung in Eigenregie oder durch einen Dienstleister produktiv betrieben werden. Wird der Betrieb einer Lösung von einer Bedarfsträger:in selbst übernommen, muss dafür dauerhaft (eigenes oder fremdes) Personal bereitstehen. Nutzendenunterstützung und die Beseitigung technischer Probleme – verursacht durch die Lösung selbst oder durch ihr technisches Umfeld – können jederzeit und auch immer wieder erforderlich sein. IT-typische Änderungen am technischen Umfeld, wie andere Hardware, neue Softwareversionen und zusätzliche Koexistenzanforderungen, führen meist sogar dazu, dass nach einiger Zeit des Betriebes die Zahl der technischen Probleme steigt. Änderungen des technischen Umfeldes führen zudem oft früher oder später dazu, dass Teile einer Software angepasst bzw. neu implementiert werden müssen, obwohl sich für die Nutzenden nichts ändert.

Hinzu kommt, dass auch datenbezogene Standards normalerweise nicht »in Stein gemeißelt« werden können. Erst nach der Finalisierung erkannte Schwachstellen, aber vor allem veränderte Anforderungen – im öffentlichen Sektor häufig durch neue gesetzliche Regelungen verursacht – führen dazu, dass Standards angepasst werden müssen. Es ist also auch nach der Fertigstellung eines Standards eine Stelle erforderlich, die Änderungsbedarfe zusammenführt und für die Aktualisierung des Standards sorgt. Wurde der Standard von einem Standardisierungsgremium entwickelt, dann ist dieses in der Regel auch für die Pflege zuständig. Um den Standard aktuell zu halten, ist allerdings die kontinuierliche Einbeziehung aller Betroffenen und auch deren aktive Beteiligung erforderlich. Daneben kann – speziell für sehr verbreitete Standards – für einzelne Gruppen von Bedarfsträger:innen, z. B. des öffentlichen Sektors, jeweils ein Gremium mit klar definierter Zuständigkeit und mit Ressourcen notwendig sein, das Verbesserungsvorschläge (z. B. für Bedienung oder Abläufe) sammelt und die entsprechende (Vor-)Abstimmung innerhalb der jeweiligen Gruppe herbeiführt (siehe auch den Abschnitt »Den Standard nach einem transparenten Verfahren pflegen und weiterentwickeln.«).

3.1 Expert:innen für IT-Standards

Größere Organisationen – auch z. B. große Kommunen und (Bundes-)Ministerien – mit zumindest deutlicher IT-Abhängigkeit besitzen in der Regel bereits seit einiger Zeit Expert:innen für IT-Standards, die die Entwicklung (möglicherweise) organisationsrelevanter Standards und die Relevanz entsprechender Standardisierungsgremien beobachten und bewerten sowie häufig auch aktiv in der Standardisierung mitarbeiten.

Auch kleinere Organisationen benötigen Zugang zu derartigem Expert:innenwissen. Es bietet sich an, dass dieses von gemeinsamen Strukturen bereitgestellt wird. Für Kommunen kommen dafür z. B. die Landes- oder die Landkreisebene sowie interkommunale Zusammenarbeit, für KMUs die zuständigen Kammern und Innungen infrage.



4. Reicht nicht ein einziger datenbezogener Standard?

Die Verwendung technischer (Meta-)Datenstandards orientiert sich in erster Linie an den Fachanforderungen an die jeweiligen Daten. Die benutzten Standards können demnach für dieselben Daten auch von Anwendung zu Anwendung bzw. von Übertragungsform zu Übertragungsform (z. B. Echtzeitzugriff auf einzelne Datensätze, Massendatenübertragung, Transport via Speichermedium) variieren. Ein Datensatz ist dabei eine Datenmenge, die ein einzelnes Objekt durch einen oder mehrere Werte beschreibt, z. B. einen Menschen durch Vorname(n), Nachname und Geburtsdatum. Die Gesamtheit einer Vielzahl von Datensätzen, z. B. in einer Datenbank, wird ebenfalls oft als Datensatz bezeichnet. Um diese Doppeldeutigkeit zu vermeiden, wird in diesem White Paper im letzteren Fall allgemein von einer strukturierten Datenmenge gesprochen.

Standards für Informationssicherheit sind an den anwendungsübergreifenden, inhärenten Sicherheitsanforderungen für die Daten orientiert und können anforderungsbedingt weitere Komponenten – wie z. B. Programmcode für den Zugriff auf oder die Verarbeitung der Daten oder physische Objekte (Speichermedien, Übertragungswege …) – sowie organisatorische Maßnahmen betreffen. Ethisch veranlasste Datenstandards wiederum bestimmen sowohl die Sicherheitsanforderungen für die Daten als auch die für sie zulässigen Anwendungen und Übertragungsformen mit. Dies zeigt bereits, dass voneinander unabhängige technische Datenstandards, Standards für Informationssicherheit und ethisch veranlasste Datenstandards sinnvoll sind und notwendig sein können, weil sie unterschiedlichen Zielen dienen und daher in Abhängigkeit von den betroffenen Daten in unterschiedlichen Kombinationen zum Einsatz kommen.

Hinzu kommt: Im Bereich der elektronischen Datenverarbeitung und -kommunikation sind die meisten Standards ein Kompromiss zwischen der Umfänglichkeit (»Breite«) dessen, was durch den Standard abgedeckt ist, und dem Aufwand, der für Umsetzung (Implementierung) und Betrieb (Speicherbedarf, Verarbeitungsaufwand, Kommunikationsvolumen usw.) erforderlich ist. Zu »breite« Standards werden oft nicht vollständig umgesetzt. Dies führt zu Problemen, sobald eine Anwendungsseite Funktionen voraussetzt, die die andere nicht unterstützt. Deshalb kann es eine bewusste und angemessene Entscheidung sein, absehbar nur äußerst selten benötigte Funktionen, Formate, Abbildungen und Ähnliches (zunächst) nicht zu standardisieren. Ein solches, oftmals pragmatisches oder auch von zeitlichen Rahmenbedingungen getriebenes Vorgehen führt in der Folge oft zur Notwendigkeit von Ergänzungen oder separater, komplementärer Standards. Ähnlich verhält es sich, wenn ein datenbezogener Standardisierungsbedarf von vornherein »horizontal« in verschiedene funktional in sich abgeschlossene Teile zerlegt wird, z. B. weil bestimmte Teile eines Zeichensatzes oder bestimmte Schutzmechanismen nur in speziellen Situationen benötigt werden.

Und last but not least: Erfahrungen mit der Erstellung von Standards haben gezeigt, dass die Chancen überproportional steigen, einen entsprechend seinem Zweck vollständigen, in sich widerspruchsfreien und auch ansonsten fehlerfreien Standard zu produzieren, wenn man umfangreiche Standardisierungsbedarfe in »vertikal« aufeinander aufbauende Abstraktionsebenen (»Layer«) mit klaren Schnittstellen zwischen den Ebenen zerlegt (s. a. Infokasten Datenabstraktion). Ein weiterer Vorteil eines solchen Vorgehens ist, dass elementarere, niedrigere Layer häufig für verschiedene übergeordnete Ausprägungen wiederverwendet werden können (vergleiche die Benutzung der lateinischen Schriftzeichen für viele Sprachen), was bei einem »All-in-One«-Ansatz vermutlich weder erkannt würde noch technisch umsetzbar wäre, weil für Letzteres dann die notwendigen Schnittstellen fehlen würden.

Exkurs: Datenabstraktion

Vertikale, in mehreren Schritten durchgeführte Datenabstraktion dient dazu, Daten aus einer Form, die zwischenmenschlich üblich ist, in eine effizient maschinenverarbeitbare Form (und umgekehrt) abzubilden. Dabei wird in jedem Schritt ein abgegrenzter, überschaubarer Abstraktionsaspekt umgesetzt und die Kette der Schritte realisiert das gewünschte Ergebnis.

Sollen beispielsweise Namen ausschließlich in **lateinischer Schreibweise** bearbeitet werden (technischer Datenstandard 1), ist für alle Namen, die originär (z. B. »auf Papier«) in anderer Schreibweise – z. B. griechisch – existieren, die Benutzung einer einheitlichen **Transliterationstabelle** (hier gem. ISO 843[1], technischer Datenstandard 2) erforderlich. Für einen elektronischen Datenaustausch oder eine elektronische Speicherung müssen alle (ggf. transliterierten) Schriftzeichen auf eine einheitliche numerische Darstellung abgebildet werden, dafür bedarf es eines Standards, wie ein Zeichen – z. B. »ü« – auf einen numerischen Wert (»Codepoint«) – hier »$00FC_{(16)}$« (Abbildung gem. String.Latin+ 1.2/ DIN 91379[2], technischer Datenstandard 3) – abzubilden ist. Je nach der Anzahl der Zeichen im zu unterstützenden Zeichensatz bzw. je nach dem Wertebereich der Codepoints werden die Codepoints eventuell mittels eines weiteren Standards auf eine **Bytefolge** – hier »C3 $BC_{(16)}$« (Abbildung auf UTF-8[3] gem. Unicode 15.0[4]), technischer Datenstandard 4) – abgebildet. Die Codepoints von String.Latin+ 1.2 reichen bspw. von »00 $09_{(16)}$« bis »22 $65_{(16)}$« und definieren damit 930 Schriftzeichenabbildungen, die jeweils 2 Bytes erfordern (würden). Die Abbildung auf UTF-8 ermöglicht eine effizientere Speicherung bzw. schnellere Übermittlung und vermeidet Codeüberschneidungen mit betriebssystemspezifischen Steuerzeichen.

1 ISO 843:1997-01: »Information and documentation – Conversion of Greek characters into Latin characters«; https://www.beuth.de/de/norm/iso-843/2410095

2 DIN 91379:2022-08: »Zeichen und definierte Zeichensequenzen in Unicode für die elektronische Verarbeitung von Namen und den Datenaustausch in Europa«; https://www.beuth.de/de/norm/din-91379/353496133

3 UTF-8: 8-Bit UCS Transformation Format, UCS: Universal Coded Character Set; 8-Bit-Transformationsformat für UNICODE.

4 ISO/IEC 10646:2020-12: »Informations Technologie - Universeller codierter Zeichensatz (UCS)« in Verbindung mit ISO/IEC 10646 AMD 1:2023-07: »Informationstechnik – Universeller codierter Zeichensatz (UCS) – Änderung 1: CJK Unified Ideographs Extension H, Vithkuqi, Old Uyghur, Cypro-Minoan und andere Zeichen«; https://www.beuth.de/de/norm/iso-iec-10646/334145221 bzw. https://www.beuth.de/de/norm/iso-iec-10646-amd-1/371146442

Abb. 1: Abstraktionsschritte bei der allgemeinen Textkodierung

Geht man gedanklich noch einen Schritt weiter und betrachtet beispielsweise den Austausch von Staatsnamen, fällt schnell auf, dass eine weitere Ebene der Einheitlichkeit vorteilhaft ist: Während sich Deutschland selbst als »(Bundesrepublik) Deutschland« bezeichnet, lautet die französische Bezeichnung »(République fédérale de l')Allemagne«, die finnische »Saksa(n liittotasavalta)«. Für die international einheitliche **(Kurz-)Bezeichnung von Staaten** gibt es zweibuchstabige Abkürzungen – »DE« (gem. DIN EN ISO 3166-1[5], Alpha-2-Kodierung, technischer Datenstandard 5a) – und dreibuchstabige – »DEU« (gem. DIN EN ISO 3166-1, Alpha-3-Kodierung, technischer Datenstandard 5b) sowie »GER« (gem. Ländercodes des Internationalen Olympischen Komitees, technischer Datenstandard 6). Es ist also notwendig, sich auf den Standard (bzw. Substandard) zu einigen, nach dem die Staatsbezeichnung menschenlesbar kodiert wird. Die einzelnen Buchstaben können dann, wie oben, iterativ weiter abgebildet werden. Alternativ zur zeichenweisen Abbildung ist auch die Nutzung einer Codeliste möglich, die die menschenlesbare Form jeweils direkt auf einen Codepoint abbildet. Für den eher maschineninternen Gebrauch, der auf »sprechende« Bezeichnungen verzichten kann, gibt es zudem drei- und vierstellige numerische Codes, die unmittelbar als Codepoints genutzt werden können – im Beispiel »$276_{(10)}$« bzw. »$1195_{(10)}$« (ISO 3166-1, Num-3- bzw. Num-4-Kodierung, technischer Datenstandard 5c bzw. 5d) sowie »$000_{(10)}$« (DESTATIS-Kodierung[6], technischer Datenstandard 7).

5 DIN EN ISO 3166-1:2020-12: »Codes für die Namen von Ländern und deren Untereinheiten – Teil 1: Codes für Ländernamen (ISO 3166-1:2020); Englische Fassung EN ISO 3166-1:2020«; https://www.beuth.de/de/norm/din-en-iso-3166-1/324815714

6 Statistisches Bundesamt: »Staats- und Gebietssystematik«, Stand 01. Januar 2023. https://www.destatis.de/DE/Methoden/Klassifikationen/Staat-Gebietsystematik/Staatsangehoerigkeitsgebietsschluessel_pdf.pdf?__blob=publicationFile

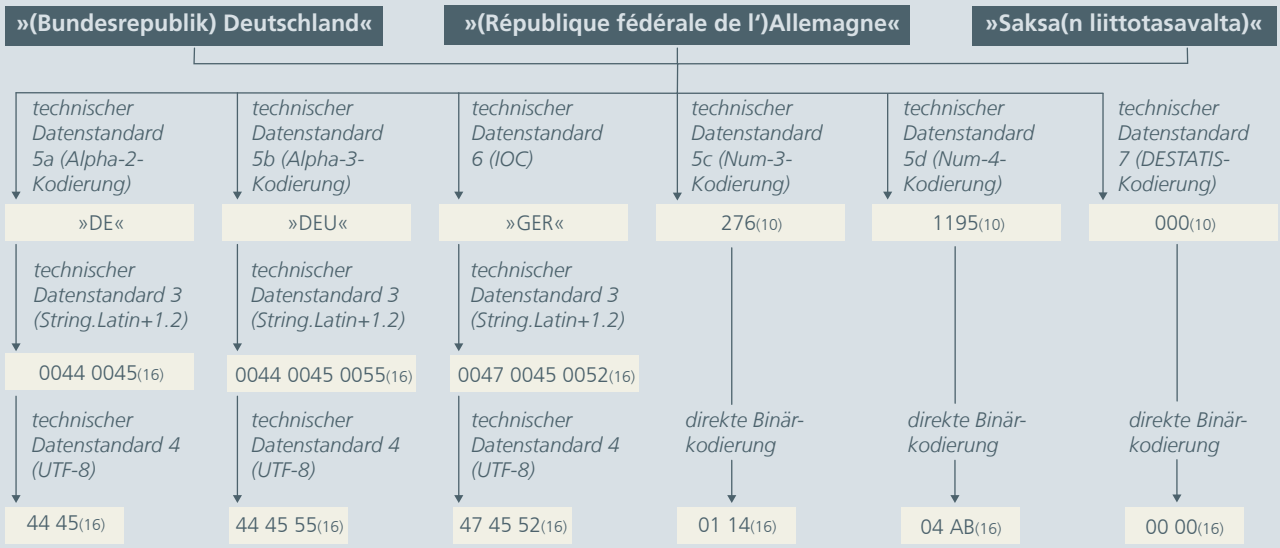


Abb. 2: Abstraktionsschritte bei der Kodierung von Staatsbezeichnungen

4.1 Technische (Meta-) Datenstandards

Bei Geräten oder Objekten, die sogar gegenüber Smartphones deutlich eingeschränkte Speicher- oder Verarbeitungskapazitäten besitzen, z. B. RFID-Tags oder optisch gelesene Karten, besteht auch heute noch oft die Notwendigkeit, spezifische, besonders »sparsame« technische Datenstandards einzusetzen, die beispielsweise nur einen sehr eingeschränkten Zeichensatz für Kommunikation und Darstellung unterstützen (s. a. Infokasten XhD). Die Festlegung eines bzw. Einigung auf einen einzigen technischen Datenstandard ist somit oft nicht ausreichend, denn ein solcher Standard könnte nur das umfassen, was auf allen Geräten und Objekten umsetzbar wäre. Dies könnte für die breite Masse der ausreichend leistungsfähigen Geräte als störende Einschränkung empfunden werden, z. B. wenn keine Unterscheidung zwischen Groß- und Kleinschreibung möglich wäre. Gravierender wäre, dass mit derartigen Beschränkungen ein Präzisionsverlust verbunden sein kann.

XhD

Vollständiger Name	BSI TR-03123-1 bis -3: »Technische Richtlinie – XML-Datenaustauschformat für hoheitliche Dokumente (TR XhD)«
IT-Planungsratsbeschlüsse	keine
Entwickelt von	Bundesamt für Sicherheit in der Informationstechnik (BSI)
Aktuelle Version	1.6.1 vom 25.01.2021
Erstveröffentlichung	30.06.2009
Art	Datenaustauschstandard
Basiert auf	u. a. DIN SPEC 91379[1] bzw. XÖV-Bibliothek[2],BSI TR-03121[3], ICAO 9303[4], XML Schema[5]
Lizenz	
Webseite	https://www.bsi.bund.de/DE/Themen/Unternehmen-und-Organisationen/Standards-und-Zertifizierung/Technische-Richtlinien/TR-nach-Thema-sortiert/tr03123/TR-03123_node.html

1 DIN SPEC 91379:2019-03: »Zeichen und definierte Zeichensequenzen in Unicode für die elektronische Verarbeitung von Namen und den Datenaustausch in Europa«, zwischenzeitlich ersetzt durch DIN 91379:2022-08.

2 In der Fassung vom 31.08.2020.

3 BSI TR-03121: »Biometrie in hoheitlichen Anwendungen, XML-Schema; Version 4.6«, https://www.bsi.bund.de/DE/Themen/Unternehmen-und-Organisationen/Standards-und-Zertifizierung/Technische-Richtlinien/TR-nach-Thema-sortiert/tr03121/TR-03121_node.html

4 ICAO (International Civil Aviation Organization) 9303: »Machine Readable Travel Documents«, hier insbes. Parts 3, 4 und 5, https://www.icao.int/publications/pages/publication.aspx?docnum=9303

5 W3C: »XML Schema Part 0: Primer Second Edition«, 28.10.2004; http://www.w3.org/TR/xmlschema-0/. W3C: »XML Schema Part 1: Structures Second Edition«, 28.10.2004; http://www.w3.org/TR/xmlschema-1/. W3C: »XML Schema Part 2: Datatypes Second Edition«, 28.10.2004; http://www.w3.org/TR/xmlschema-2/.

XhD ist ein dreiteiliger Daten(austausch)standard, der primär das Format der zur Erstellung eines hoheitlichen Dokumentes – Personalausweis, Reisepass, Reiseausweis, Aufenthaltstitel, ID-Karte – erforderlichen Daten festlegt. Er ist gegliedert in ein Rahmendokument (BSI TR-03123-1), Dokumentenprofile (BSI TR-03123-2) und Funktionsmodule (BSI TR-03123-3). Die Datenspezifikationen sind auf die BSI TR-03123-2 (beschreibend) und die BSI TR-03123-3 (UML-Notation) verteilt: Zulässige Feldlängen findet man beispielsweise in der BSI TR-03123-2, während Codelisten in der BSI TR-03123-3 festgelegt sind.

Hintergrund

Für die inhaltliche und formale Korrektheit hoheitlicher Dokumente gelten besonders hohe Anforderungen. Deutsche hoheitliche Dokumente werden zentral hergestellt, die Bestellungen erfolgen dezentral, vor allem durch die zuständigen Meldebehörden. Hoheitliche Dokumente müssen deutschen und Vorgaben der Europäischen Union sowie internationalen Regeln entsprechen, damit sie z. B. auch von Berechtigten ohne Deutschkenntnisse interpretiert oder elektronisch gelesen werden können. Beispielsweise ist der im maschinenlesbaren Teil der Dokumente zugelassene Zeichensatz auf Großbuchstaben (ohne Umlaute), Ziffern und »<« beschränkt, und die Kürzung von Angaben, die nicht vollständig in die maschinenlesbaren Zeilen passen, ist genau geregelt.

Bei der Bestellung eines hoheitlichen Dokumentes muss die zuständige Behörde alle für das Ausfüllen des Dokumentes erforderlichen Daten in der Form liefern, wie sie im Dokument dargestellt werden sollen. Soll beispielsweise bei notwendigen Kürzungen von Namensbestandteilen im maschinenlesbaren Teil von den allgemeinen Kürzungsregeln abgewichen werden, um ehrverletzende Ergebnisse zu vermeiden, dann muss die gekürzte Form explizit angegeben werden.

Daneben spezifiziert der Datenaustauschstandard die Formate und die Kodierungen der während des Bestellvorganges auftretenden Nachrichten und der darin übermittelten organisatorischen Daten.

Jede:r kann die in XhD spezifizierten inhaltlichen Daten anhand eines üblicherweise vorhandenen persönlichen Personaldokumentes nachvollziehen.

4.2 Technischer und organisatorischer Schutz von Daten

Unterschiedliche Anforderungen an die Sicherheit von Daten in unterschiedlichen Anwendungskontexten und Schutzmaßnahmen, die nur für bestimmte Gegebenheiten geeignet sind, können dazu führen, dass es vorteilhafter oder sogar notwendig ist, situationsspezifische Schutzstandards zu entwickeln und einzusetzen. Neben dieser Differenzierung in der Breite ist bei informationstechnisch umgesetzten Schutzstandards oft auch eine vertikale Differenzierung sinnvoll, z. B. indem logisch »unten« ein verschlüsselter Datenübertragungskanal zwischen einander vertrauenden Organisationen die Vertraulichkeit gegenüber Dritten sicherstellt und logisch »darüber« eine weitere Verschlüsselung den Zugriff auf sensible Daten zusätzlich auf ausgewählte Mitarbeitende der kooperierenden Organisationen beschränkt.

Für viele Daten des öffentlichen Sektors und von Unternehmen existieren hohe Integritätsanforderungen, da sie die Basis für folgenreiche Entscheidungen der Datenhalter:innen oder Dritter, die die Daten nutzen, bilden. Die Verfügbarkeitsanforderungen an Daten hingegen haben eine deutliche Spannbreite: Während bestimmte, auch personenbezogene Daten z. B. für Kontrollen praktisch sofort und immer verfügbar sein müssen, reichen für andere Daten Zugriffsmöglichkeiten innerhalb einiger Tage oder erst nach vorheriger Absprache. Ähnliches gilt für die Vertraulichkeit: Personenbezogene und Firmendaten müssen meist vertraulich behandelt werden, was unter anderem auch organisationsintern beschränkte Zugriffsrechte bedeuten kann (s. a. Beispiel des vorigen Absatzes). Andererseits bestehen für explizit als offen kategorisierte Daten keinerlei Vertraulichkeitsanforderungen.

4.3 Ethischer Umgang mit Daten

In Bezug auf Daten gibt es viele ethische Themen und Gesichtspunkte und deren angemessene Berücksichtigung ist in der Regel komplex. Deshalb erfolgt die Betrachtung der Themen und Gesichtspunkte üblicherweise (zumindest teilweise) separat und auf mehreren Detailebenen. Allgemeinere, aber dadurch universellere ethisch begründete Ziele, Anforderungen, Rechte und Pflichten lassen sich beispielsweise aus dem Grundgesetz, der Allgemeinen Erklärung der Menschenrechte der Vereinten Nationen (englisch: United Nations, UN)[17] und der Europäischen Menschenrechtskonvention des Europarates[18] ableiten. »Die Charta der Menschenrechte und Prinzipien für das Internet«[19] basiert auf der Allgemeinen Erklärung der Menschenrechte der UN sowie auf den beiden darauf aufbauenden Internationalen Pakten über wirtschaftliche, soziale und kulturelle Rechte[20] bzw. über bürgerliche und politische Rechte[21]. Diese Charta wurde durch die Internet Rights and Principles Coalition des UN Internet Governance Forum entwickelt. Die im Gutachten der von der Bundesregierung eingesetzten Datenethikkommission vom Oktober 2019[22] enthaltenen ethischen Grundsätze und Prinzipien sind vom Grundgesetz und den Menschenrechten abgeleitet, die im Gutachten dargestellten Anforderungen, Rechte und Pflichten sind eine Detaillierungsstufe im Hinblick auf Daten. Ethisch veranlasste Standards im engeren Sinne, also für das »Wie«, sind z. B. erforderlich, wenn es darum geht, Grenzwerte (beispielsweise bei Anonymisierungsverfahren) oder konkrete Mechanismen und Algorithmen zur Umsetzung ethischer Anforderungen, Rechte oder Pflichte festzulegen. Solche Standards sind eventuell gleichzeitig technische Datenstandards, Standards zum technischen Schutz von Daten oder untrennbarer Teil solcher Standards.

17 Vereinte Nationen: »Allgemeine Erklärung der Menschenrechte« (»The Universal Declaration of Human Rights« (UDHR)); https://unric.org/de/allgemeine-erklaerung-menschenrechte/

18 Europarat: »Konvention zum Schutze der Menschenrechte und Grundfreiheiten in der Fassung der Protokolle Nr. 11, 14 und 15«; https://rm.coe.int/1680a6eaba

19 Internet Rights and Principles Coalition: »Die Charta der Menschenrechte und Prinzipien für das Internet«; https://drive.google.com/file/d/11Q3J7CsgzuDnzbivbXyWBlLu6zdKK6Zt/view?usp=sharing

20 Vereinte Nationen: »Internationaler Pakt über wirtschaftliche, soziale und kulturelle Rechte vom 19. Dezember 1966«; https://www.institut-fuer-menschenrechte.de/fileadmin/Redaktion/PDF/DB_Menschenrechtsschutz/ICESCR/ICESCR_Pakt.pdf

21 »Gesetz zu dem Internationalen Pakt vom 19. Dezember 1966 über bürgerliche und politische Rechte« vom 15. November 1973; http://www.bgbl.de/xaver/bgbl/start.xav?startbk=Bundesanzeiger_BGBl&jumpTo=bgbl273s1533.pdf

22 Datenethikkommission: »Gutachten der Datenethikkommission«, Stand: 10. Oktober 2019; https://www.bmi.bund.de/SharedDocs/downloads/DE/publikationen/themen/it-digitalpolitik/gutachten-datenethikkommission.pdf?__blob=publicationFile&v=6


PL-16-156
HT

5. Kategorien datenbezogener Standards im öffentlichen Sektor

Für den öffentlichen Sektor sind die folgenden, teilweise sehr unterschiedlichen Kategorien von datenbezogenen Standards besonders relevant. Die Kategorien sind nicht zwangsläufig disjunkt – so sind z. B. Wettermessdaten für einen längeren Zeitraum, die gesammelt einem Dritten zur Verfügung gestellt werden, sowohl »Vermessungs«-Daten als auch Teil einer strukturierten Datenmenge – und beziehen sich teilweise auf unterschiedliche Aspekte der Daten bzw. des Datenzugangs. Sie wurden vor allem unter pragmatischen Gesichtspunkten bestimmt. Das bedeutet: Wer seinen Bedarf klar einer einzigen Kategorie zuordnen kann, sollte bei der Suche nach geeigneten Standards vornehmlich jene in den Blick nehmen, die für diese Kategorie geschaffen wurden oder verbreitet eingesetzt werden. Nichtsdestotrotz können auch oder sogar nur Standards geeignet sein, die einer anderen Kategorie zuzuordnen sind. Fällt ein Bedarf in mehrere Kategorien, sollte zunächst gezielt nach Standards gesucht werden, die alle betroffenen Kategorien abdecken oder in allen betroffenen Kategorien zum Einsatz kommen.

Darüber hinaus sind für jede Kategorie auch Querschnittsfragen zu beantworten, z. B.: Wird ein Standard für die (dauerhafte) Speicherung von Daten (z. B. in Registern), für den Austausch von Daten oder beides benötigt? Dabei ist auch relevant, ob die Daten eher statisch oder dynamisch bzw. sogar volatil, also kurzlebig sind. Aus rechtlicher, ethischer und Informationssicherheits-Perspektive ist beispielsweise zu überprüfen, ob es sich bei den Datensätzen um personenbezogene Daten bzw. vertrauliche Unternehmensdaten handelt.

Unmittelbare Erfüllung öffentlicher Aufgaben

Die betroffenen Daten sind beispielsweise im Großen Daten für das Meldewesen und den Personenstand, für Unternehmens- und Vereinsregister oder für das Grundbuch, im Kleineren für die Schülerverzeichnisse der einzelnen Schulen oder die tierspezifischen Daten im Hunderegister. Dabei handelt es sich überwiegend um Daten zur (oft möglichst eindeutigen) Beschreibung natürlicher oder juristischer Personen, bedeutender Wertgegenstände oder – für das Zusammenleben im öffentlichen Raum relevanter – privater Objekte (z. B. Fahrzeuge, Hunde, Jagd- und Sportwaffen).

Die Daten haben oft eine sehr lange andauernde Relevanz, an ihre Korrektheit und Vollständigkeit werden hohe Anforderungen gestellt und nur spezifisches Personal der öffentlichen Hand erhält die Berechtigung, derartige Daten zu erfassen oder zu ändern. Vielfach gelten Vertraulichkeitsanforderungen. Wenn Daten (in der Regel einzelne bzw. wenige Datensätze in von Menschen interpretierbarer Form) aus solchen Registern/Verzeichnissen berechtigten Einzelpersonen (z. B. Unfallgeschädigten) bzw. -firmen oder der Öffentlichkeit zur Verfügung gestellt werden, erwarten die Empfänger:innen ein hohes Maß an Verlässlichkeit und Aktualität der Daten, weil solche Daten häufig Grundlage für Entscheidungen mit großer Tragweite sind. Soweit zur Verfügung gestellte Daten sich auf identifizierbare Personen, Unternehmen oder Gruppen solcher Objekte beziehen, haben zudem die betroffenen Objekte das Recht auf Datenkorrektheit.

Diese Datenkategorie ist gleichzeitig die Kernkategorie, wenn es um die Offenlegung im Sinne von Open Government Data geht. Sofern es sich um mit ethischen Fragen behaftete, z. B. personenbezogene Daten handelt, kann eine unmittelbare Offenlegung ausgeschlossen sein und die Offenlegung bestimmter Aspekte eine Vorverarbeitung, z. B. Aggregierung oder Anonymisierung erforderlich machen.

Die in diesem Bereich zu verwendenden Standards legt in der Regel die öffentliche Hand fest, vorteilhaft dabei ist angesichts der Forderungen der EU, vielfältigen grenzüberschreitenden elektronischen Datenaustausch zu ermöglichen, und angesichts des tatsächlich steigenden Bedarfs für derartige Möglichkeiten sich an vorhandenen, insbesondere internationalen Standards zu orientieren und frühzeitig die Abstimmung mit den anderen EU-Staaten zu suchen und zu fördern, um die Forderungen und Erwartungen ohne zusätzlichen Aufwand erfüllen zu können.

Internes Funktionieren öffentlicher Stellen

Bei den betroffenen Daten handelt es sich z. B. um solche für das Beschaffungs-, Inventarisierungs- oder Rechnungswesen, für das Personalwesen, für Projektmanagement usw. Für die in diesem Zusammenhang bearbeiteten Daten gelten zu Unternehmen vergleichbare Anforderungen. Bearbeitet werden können sie prinzipiell von allgemein für den Einsatzzweck der Daten geschultes Personal, auch wenn eventuell spezifische Gesetze oder Verordnungen für den öffentlichen Sektor zu beachten sind – ähnliche Regeln sind häufig auch in Unternehmenscodizes vorhanden.

Die zunehmend übergreifend automatisierte Verarbeitung entsprechender Daten durch den öffentlichen Sektor und Unternehmen – z. B. Rechnungen oder Bestellungen – erfordert den Einsatz gemeinsamer technischer Datenstandards, damit die Daten einheitlich interpretiert werden. In diesem Bereich ist die Wirtschaft dem öffentlichen Sektor oft voraus und es gibt bereits breit etablierte Standards, die der öffentliche Sektor übernehmen sollte, weil der Erfüllungsaufwand für die Wirtschaft bei neu durch die öffentliche Hand gesetzten Standards nicht vertretbar wäre.

Die EU-Richtlinie zur elektronischen Rechnungsstellung[23] beispielsweise beinhaltet einen Auftrag (an das CEN) für die Normung eines semantischen Datenmodells einer Rechnung und für die Benennung geeigneter Syntaxen zur Umsetzung des Modells sowie die Verpflichtung für die öffentliche Hand der Mitgliedstaaten, dieses Modell und die Syntaxen zu unterstützen. Der Auftrag wurde durch die vom DIN national übernommene europäische Norm DIN EN 16931-1[24] und den ebenfalls vom DIN übernommenen Technischen Standard des CEN DIN CEN/TS 16931-2[25] erfüllt. Auch die fortlaufende Pflege dieser Dokumente obliegt dem CEN. Das semantische Datenmodell nach DIN EN 16931-1 bietet Spielräume für nationale Ausgestaltungen und Ergänzungen. Der XÖV-Standard XRechnung[26] wurde entwickelt, um spezielle Bedarfe der deutschen öffentlichen Verwaltung abzubilden – z. B. die Leitweg-ID – sowie Kernelemente und nationale Ergänzungen in einem Dokument zu vereinen (s. a. Infokasten XRechnung).

Daten des Beschaffungs-, Inventarisierungs- und Rechnungswesens müssen vertrauenswürdig und verfügbar sein, beispielsweise um Anforderungen der Innenrevision oder der zuständigen Rechnungshöfe zu erfüllen. Für die Verfügbarkeit der Daten wird meist ausschließlich die öffentliche Seite verantwortlich sein – und somit auch die dafür anzuwendenden Standards bestimmen können. Die Vertrauenswürdigkeit vieler Daten erfordert hingegen das Zusammenwirken der öffentlichen Seite und ihrer externen Geschäftspartner: Angebote und Rechnungen beispielsweise müssen gegebenenfalls vom Anbieter digital signiert werden, wobei der öffentliche Sektor in der Regel die einzusetzenden Verfahren und Mittel festlegt.

Anforderungen der öffentlichen Seite oder solche der jeweiligen (externen) Geschäftspartner:in können eine vertrauliche Behandlung zumindest eines Teils der Daten erfordern. Für die interne, gegenüber Dritten vertrauliche Behandlung von Daten in seinem Besitz ist der öffentliche Sektor verantwortlich. Er kann die Standards dafür – unter Beachtung des Standes der Technik – allein festlegen. Jedoch kann es sinnvoll sein, bestimmte Maßnahmen und Mittel, wie z. B. Verschlüsselung, übergreifend einzusetzen und so Gefährdungen an Verfahrensübergängen zu vermeiden. Werden derartige Daten – im Rahmen der Zulässigkeit – mit Externen ausgetauscht, dann kann die öffentliche Seite zwar Anforderungen festlegen, wird aber nur solche Maßnahmen und Mittel fordern können, die den Externen zur Verfügung stehen oder von der öffentlichen Seite zur Verfügung gestellt werden.

23 »Richtlinie 2014/55/EU des Europäischen Parlaments und des Rates vom 16. April 2014 über die elektronische Rechnungsstellung bei öffentlichen Aufträgen (Text von Bedeutung für den EWR)«; https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=uriserv%3AOJ.L_.2014.133.01.0001.01.DEU.

24 Aktuelle Fassung: DIN EN 16931-1 (Dezember 2020): »Elektronische Rechnungsstellung – Teil 1: Semantisches Datenmodell der Kernelemente einer elektronischen Rechnung; Deutsche Fassung EN 16931-1:2017+A1:2019 + AC:2020«; https://www.beuth.de/de/norm/din-en-16931-1/327729047

25 Aktuelle Fassung: DIN CEN/TS 16931-2 (November 2017): »Elektronische Rechnungsstellung – Teil 2: Liste der Syntaxen, die EN 16931-1 erfüllen; Deutsche Fassung CEN/TS 16931-2:2017«; https://www.beuth.de/de/technische-regel/din-cen-ts-16931-2/274991011

26 Aktuelle Fassung: KoSIT: »Spezifikation XRechnung – Standard und Extension – Version XRechnung 2.3.1«, veröffentlicht 03.02.2023, gültig ab 01.08.2023; https://xeinkauf.de/app/uploads/2023/02/231-XRechnung-2023-02-03.pdf

XRechnung

Vollständiger Name	XRechnung
IT-Planungsratsbeschlüsse	2022/49, 2017/22, 2017/18, 2015/34
Entwickelt von	KoSIT
Aktuelle Version	3.0.1 (in Kraft getreten: 01.02.2024)
Erstveröffentlichung	1.0 (10.05.2017)
Art	Datenaustauschstandard
Basiert auf	EN 16931-1[1], EN ISO 3166-1[2], ISO 4217[3], ISO 8601:2004[4], ISO 15000-5:2014 Anhang B[5], Leitweg-ID[6], XÖV[7], CII[8], UBL[9], XML[10], UTF-8[11]
Lizenz	Lizenzabkommen der Europäischen Kommission (EK) und des Europäischen Komitees für Normung (CEN)
Webseite	https://xeinkauf.de/xrechnung/versionen-und-bundles/

1 DIN EN 16931-1:2020-12: »Elektronische Rechnungsstellung – Teil 1: Semantisches Datenmodell der Kernelemente einer elektronischen Rechnung«; https://www.beuth.de/de/norm/din-en-16931-1/327729047

2 DIN EN ISO 3166-1:2020-12: »Codes für die Namen von Ländern und deren Untereinheiten – Teil 1: Codes für Ländernamen«; https://www.beuth.de/de/norm/din-en-iso-3166-1/324815714

3 ISO 4217:2015-08: »Codes für die Darstellung von Währungen«; https://www.beuth.de/de/norm/iso-4217/240783291

4 ISO 8601:2004: »Data elements and interchange formats«; https://www.iso.org/standard/40874.html

5 DIN ISO 15000-5:2014-10 »Erweiterbare Auszeichnungssprache für das elektronische Geschäftswesen (ebXML) – Teil 5: Spezifikation der Kernkomponenten (CCS) (ISO 15000-5:2014)«; https://www.beuth.de/de/norm/din-iso-15000-5/209146415

6 KoSIT: »Leitweg-ID. Format-Spezifikation Version 2.0.2«, Juli 2021; https://xeinkauf.de/app/uploads/2022/11/Leitweg-ID-Formatspezifikation-v2-0-2-1.pdf

7 Standard gemäß XÖV-Handbuch; https://www.xoev.de/xoev/xoev-produkte/xoev-handbuch-5060

8 UNECE: »Cross Industry Invoice (CII)«; https://tfig.unece.org/contents/cross-industry-invoice-cii.htm

9 OASIS: »Universal Business Language (UBL)«; https://www.oasis-open.org/committees/ubl/

10 W3C Recommendation: »Extensible Markup Language (XML) 1.0 (Fifth Edition)«, 2008; https://www.w3.org/TR/xml/

11 ISO/IEC 10646:2020-12: »Informations Technologie – Universeller codierter Zeichensatz (UCS)«; https://www.beuth.de/de/norm/iso-iec-10646/334145221

XRechnung ist ein Datenstandard zum Austausch elektronischer Rechnungen mit öffentlichen Auftragnehmern. Der Standard setzt die Vorgabe innerhalb Deutschlands um, EU-weit interoperabel elektronische Rechnungen erstellen, entgegennehmen und verarbeiten zu können. Dem Standard entsprechende Rechnungen sind sprach- und bezeichnungsunabhängig, d. h., eine z. B. in Spanien erstellte Rechnung kann in Deutschland ohne individuelle Übersetzung verarbeitet werden und synonym verwendete Begriffe – wie »Bestellungsnummer« und »Bestellreferenz« – bzw. mehrdeutige Begriffe – wie »Bestellnummer« – führen nicht mehr zu erhöhtem Interpretationsaufwand. Der Standard deckt das breite Feld der in DIN EN ISO 3166-1 spezifizierten typischen Elemente einer Rechnung (der sogenannten Kernelemente) – wie »Rechnungsnummer« oder »Lieferanschrift« – sowie deutschlandspezifische Zusatzanforderungen – wie die etwaige Angabe der »Leitweg-ID« im Kernelement »Referenz des Käufers« – ab.

Hintergrund

Grundlage für die Standardisierung des Austauschs elektronischer Rechnungen ist die 2014 in Kraft getretene EU-Richtlinie über die elektronische Rechnungsstellung bei öffentlichen Aufträgen (2014/55/EU)[12]. Anlass der Richtlinie war eine bereits existierende und zunehmende Vielfalt nicht interoperabler Standards für elektronische Rechnungen in den Mitgliedstaaten, die unerwünschte Marktzutrittsschranken und Handelshemmnisse darstellte. Gegenstände der Richtlinie sind die EU-weite Verpflichtung, bei öffentlichen Aufträgen elektronische Rechnungen, die einem einheitlichen semantischen Datenmodell und einer Syntax aus einer begrenzten Anzahl zugelassener Syntaxen

12 »Richtlinie 2014/55/EU des Europäischen Parlaments und des Rates vom 16. April 2014 über die elektronische Rechnungsstellung bei öffentlichen Aufträgen«; https://eur-lex.europa.eu/eli/dir/2014/55/oj

entsprechen, entgegennehmen und verarbeiten zu können, sowie der Normungsauftrag für das Datenmodell.

Zur Umsetzung der Richtlinie wurde an das Europäische Komitee für Normung (CEN, französisch: Comité Européen de Normalisation) der Auftrag zur Erstellung des später als EN 16931 verabschiedeten, mehrteiligen Normenwerkes gegeben. Gemeinsam mit Vertretern des DIN nahmen Beauftragte der Koordinierungsstelle für IT-Standards (KoSIT) als Interessenvertretung der deutschen öffentlichen Auftraggeber am europäischen Normungsverfahren teil.[13] Die DIN EN 16931 verweist auf mehrere international anerkannte Kodierungen: z. B. zur vereinheitlichten Deklaration von Währungen, für Datum- und Zeitangaben oder für die Bezeichnung von Staaten und deren Untereinheiten. Darüber hinaus wird mit der »erweiterbaren Auszeichnungssprache für das elektronische Geschäftswesen« (ebXML)[14] ein ebenfalls internationaler Standard als Grundlage des Datenformates festgelegt.

Die Erarbeitung des EN-16931-konformen Standards XRechnung wurde parallel vom IT-Planungsrat auf nationaler Ebene koordiniert und 2017 in der ersten Version von der KoSIT herausgegeben. XRechnung stellt dabei eine CIUS (»Core Invoice Usage Specification«) der europäischen Norm dar, die das europäische Format übernimmt und durch nationale und branchenspezifische Elemente, wie die deutsche Leitweg-ID, erweitert.[15] Seit dem 27.11.2019 ist die Entgegennahme und Verarbeitung elektronischer Rechnungen im gesamten öffentlichen Auftragswesen des Bundes verpflichtend, seit dem 27.11.2020 auch die Ausstellung derartiger Rechnungen.[16]

XRechnung ist inzwischen Bestandteil der neu geschaffenen Gremien- und Organisationsstruktur »XStandards Einkauf« (XSE). Das Vorhaben bündelt mehrere fachrelevante Standards des öffentlichen Beschaffungswesens und soll u. a. die Harmonisierung und Interoperabilität zwischen öffentlichen Beschaffungsprozessen gewährleisten. Mitglieder der Gremienstruktur sind KoSIT, IT-PLR und FITKO. Ferner soll XSE gebündelt deutsche Interessen in europäischen Gremien hinsichtlich des öffentlichen Beschaffungswesens, z. B. durch die Mitarbeit im CEN, vertreten.[17]

Diese Beschreibung enthält kein Anwendungsbeispiel, da sinnvolle Beispiele meist sehr umfangreich sind und ihre Interpretation detaillierte Kenntnisse des Standards erfordert.

13 IT-Planungsrat: »Spezifikation XRechnung. Standard und Extension«, Version 2.3.1 vom 03.02.2023; https://xeinkauf.de/xrechnung/versionen-und-bundles/

14 ebXML: Electronic Business using eXtensible Markup Language.

15 »Bei der Übermittlung und Verarbeitung elektronischer Rechnungen müssen die geltenden Anforderungen an Datenschutz und Datensicherheit erfüllt sein«. Wirtsch Inform Manag 11, 342–346 (2019). https://doi.org/10.1365/s35764-019-00201-w.

16 § 11 der Verordnung über die elektronische Rechnungsstellung im öffentlichen Auftragswesen des Bundes (E-Rechnungsverordnung – ERechV); https://www.gesetze-im-internet.de/erechv/__11.html

17 IT-Planungsrat: »Betriebskonzept XStandards Einkauf (XSE). Die Standardfamilie des öffentlichen Einkaufs«, Version 1.0 vom 10.11.2022. https://www.it-planungsrat.de/beschluss/beschluss-2022-49

Automatisierte »Vermessung« der physischen Welt

Die automatisierte »Vermessung« der physischen Welt durch von der öffentlichen Hand oder auch von Dritten betriebene Sensoren, die Verarbeitung solcher Daten und die Steuerung physischer Objekte (Aktuatoren) durch vom öffentlichen Sektor übermittelte Daten nehmen weiter zu und benötigen entsprechende datenbezogenen Standards. In diese Kategorie fallen beispielsweise Standards für vielfältige Daten, die im Rahmen von Wetter-, Verkehrs- oder Umweltbeobachtungen gewonnen oder zur Steuerung von Verkehr, Warn- oder Schutzeinrichtungen eingesetzt werden.

Zum Einsatz kommende technische Standards müssen vor allem sicherstellen, dass die gewonnenen oder zur Steuerung eingesetzten Daten korrekt interpretierbar sind. In der Regel werden die für die Kodierung von Rohdaten (Daten und Metadaten, die unmittelbar von Sensoren stammen) verwendeten Standards von den Herstellern der Sensoren festgelegt. Aus Kompatibilitätsgründen mit anderen vom öffentlichen Sektor genutzten oder von ihm zu Verfügung gestellten Daten kann es sinnvoll sein, Rohdaten vor bestimmten Schritten der Weiterverarbeitung oder Speicherung in ein übergreifend genutztes Format zu konvertieren. Als Zielformat sollte bevorzugt ein etabliertes, breit eingesetztes Format gewählt werden. Zur Steuerung physischer Objekte kann es allerdings notwendig sein, (Meta-)Daten (wiederum) in ein herstellerspezifisches Format umzukodieren.

Viele Sensoren und Aktuatoren unterstützen nur einen sehr eingeschränkten Satz von Informationssicherheitsmechanismen

und -mitteln. In diesem Zusammenhang ist es Aufgabe der öffentlichen Hand, bereits vor Beschaffung bzw. Einsatz von Sensoren und Aktuatoren zu prüfen, ob und wie das erforderliche Sicherheitsniveau mit den ins Auge gefassten konkreten Produkten erreichbar ist. Gegebenenfalls können dazu auch organisatorische Maßnahmen erforderlich sein, z. B., indem bestimmte Sensoren/Aktuatoren nur in separaten Netzen mit zusätzlicher Zugangskontrolle betrieben werden.

Zumindest, wenn Daten über identifizierbare Personen oder Personengruppen erfasst werden (sollen), dürfte eine ethische Betrachtung erforderlich sein.

Gemeinsames Erbringen von Leistungen für Bürger:innen oder Unternehmen durch die öffentliche Hand und Dritte

Beispielsweise im Smart-City-Kontext werden öffentliche und private Leistungsangebote zunehmend enger verflochten, z. B. im Rahmen flexibler Mobilitätskonzepte, die sowohl den ÖPNV als auch private Angebote der »Sharing-Economy« – wie E-Scooter-, Bike- und Carsharing – umfassen. Für eine nutzendenfreundliche, anbieterübergreifende Echtzeitpräsentation der verfügbaren Verkehrsmittel, ihrer Nutzungsvoraussetzungen usw. bedarf es zumindest gemeinsamer technischer (Meta-) Datenstandards, u. a. zur Beschreibung von Positionsdaten.

Die öffentliche Hand wird in einer solchen Konstellation stets gemeinsame technische (Meta-)Datenstandards unterstützen müssen, unabhängig davon, ob sie als Datenlieferantin, als Datenkonsumentin oder lediglich als (vertrauenswürdige) Datenvermittlerin tätig ist. Je nach der Rolle und dem Anteil der öffentlichen Hand bei der gemeinsamen Erbringung von Leistungen werden bzw. sind diese Standards häufig nicht von der öffentlichen Seite (allein) festgelegt.

Zumindest dann, wenn die Leistungen auch Buchungsmöglichkeiten oder Ähnliches umfassen, sind auch Standards für Informationssicherheit (aller drei Kategorien) erforderlich, auf die die Nutzenden derartiger Leistungen vertrauen können.

Die Akzeptanz gemeinsam erbrachter Leistungen ist stark abhängig von einem einheitlichen, einfachen und erfolgreichen Nutzungserlebnis. Daher ist es vorteilhaft, wenn auch in den Bereichen, in denen jeweils nur eine einzelne Leistungserbringer:in betroffen ist, beispielsweise für Busverbindungen ein ÖPNV-Unternehmen der öffentlichen Hand, weitestgehend übergreifende Standards für die Interaktion mit den Nutzenden eingesetzt werden, also z. B. für alle Buchungsvorgänge mindestens ein gemeinsames Bezahlverfahren mit einheitlicher Authentifizierung der Nutzenden.

Wenn in einer solchen Konstellation ethische Fragen zu beantworten sind, z. B. die Frage der Zulässigkeit, individuelle Nutzendenprofile zu erstellen und auszuwerten, muss die öffentliche Hand die Einhaltung der eigenen ethisch veranlassten Datenstandards auch von den Dritten einfordern, da die Bürger:innen unangemessenes Verhalten der Dritten auch der öffentlichen Seite anlasten werden.

Strukturierte und maschinenverarbeitbare Datenmengen, die Dritten zur (Nach-)Nutzung bereitgestellt werden

Die öffentliche Hand besitzt, erfasst und erhebt vielfältige, oft nicht geheimhaltungsbedürftige Daten, die auch über die Grenzen des öffentlichen Sektors hinweg relevant sind. Dabei handelt es sich vorwiegend um Daten für die Erfüllung öffentlicher Aufgaben (siehe auch den Abschnitt »Unmittelbare Erfüllung öffentlicher Aufgaben«). Der Kreis der Interessierten erstreckt sich von einzelnen Bürger:innen über die gemeinwohlorientierte Zivilgesellschaft (Civic Tech) bis hin zur (regionalen) Privatwirtschaft. Darüber hinaus profitieren auch Forschung und Wissenschaft, aber auch andere Behörden und Einrichtungen der öffentlichen Hand von derartigen Daten. Deren Bereitstellung – z. B. als Open (Government) Data – wird zunehmend erwartet und gefordert.[27] Wenn die Nachnutzung nicht ausschließlich im Rahmen eines gemeinsamen (Fach-)Verfahrens stattfindet und insbesondere, wenn eine nicht vorab festgelegte Art der Nachnutzung möglich sein soll, ist es erforderlich, die Daten durch möglichst vielfältige Metadaten zu beschreiben. Oft können nur damit mögliche Nachnutzende entscheiden, ob die Daten für ihre Zwecke geeignet sind (z. B. ob sie sich auf einen relevanten Zeitraum beziehen). Und: Wenn dem Urheberrecht unterliegende Daten bereitgestellt werden, ist beispielsweise die Angabe der Nutzungslizenz in den Metadaten Voraussetzung für eine rechtssichere Nachnutzung.

Häufig sind solche Datenmengen sehr groß und bestehen aus vielen Datensätzen. Auch in den heutigen Zeiten hoher Übertragungsraten und niedriger Übertragungskosten bedarf es daher oft spezieller technischer Datenstandards für den gezielten Abruf (s. a. Infokasten OParl) und die effiziente Übertragung, um das erforderliche Datenvolumen zu reduzieren. In dieselbe Kategorie fallen vergleichbare Datenmengen Dritter, die zukünftig eventuell auch vermehrt von der öffentlichen Hand genutzt werden. Ethische Anforderungen können eine Vorverarbeitung

27 Siehe z. B. Tagesspiegel Background: »Bitkom fordert mehr Open Government Data«, 01.04.2020; https://background.tagesspiegel.de/digitalisierung/bitkom-fordert-mehr-open-government-data

der Daten, z. B. Aggregation oder Anonymisierung notwendig machen. Einen guten Einstieg in das Thema Open Government Data bieten die auf der Webseite »Open Data – Handbuch für offene Verwaltungsdaten«[28] vom Bundesverwaltungsamt bereitgestellten Dokumente, insbesondere das diesbezügliche Handbuch.

Je nach Relevanz und Nutzungsrahmen der Daten können sehr unterschiedliche Informationssicherheitsanforderungen existieren: Werden beispielsweise umfang- und detailreiche Wirtschafts- oder Gesundheitsdaten – z. B. für Forschungszwecke – vom öffentlichen Sektor zur Verfügung gestellt, kann es notwendig sein, dass Datennutzende auch nach längerer Zeit noch die Herkunft und die Integrität der Daten belegen können, um ihre eigenen Verarbeitungsergebnisse zu rechtfertigen. Werden Daten hingegen für die Echtzeitverarbeitung herangezogen, ist deren Verfügbarkeit in der Regel ausschlaggebend. Gelten in einem solchen Kontext Vertraulichkeitsanforderungen für die Daten, bedarf es insbesondere standardisierter, klarer und rechtssicherer Vereinbarungen mit den Nutzenden.[29]

28 Bundesverwaltungsamt: »Open Data – Handbuch für offene Verwaltungsdaten«; https://www.bva.bund.de/DE/Services/Behoerden/Beratung/Beratungszentrum/Methoden/_documents/stda_open_data.html

29 Siehe z. B. International Data Spaces Association: »Position Paper: Usage Control in the International Data Spaces, Version 3.0«, März 2021. https://internationaldataspaces.org/wp-content/uploads/dlm_uploads/IDSA-Position-Paper-Usage-Control-in-the-IDS-V3..pdf

OParl

Vollständiger Name	OParl – Spezifikation einer einheitlichen Schnittstelle zum Abruf von maschinenlesbaren Informationen aus Ratsinformationssystemen
IT-Planungsratsbeschlüsse	keine
Entwickelt von	Open Knowledge Foundation Deutschland e. V.
Erstveröffentlichung	1.0 (11.07.2016)
Aktuelle Version	1.1 (20.06.2018)
Art	Webserviceschnittstelle für anonymen, lesenden Zugriff auf öffentliche Inhalte parlamentarischer Informationssysteme
Basiert auf	-
Implementierung verwendet	- JSON[1] zur unmittelbaren Kodierung von Daten und zur Beschreibung konkreter Datenobjekte (z. B. Dokumente) - RFC 3986[2] zur Identifizierung/Adressierung von Daten und Datenobjekten - ISO 8601[3] zur Kodierung von Datum/Uhrzeit
Spezifikation verwendet	JSON[1]-Notation/-Format und -Datentypen
Lizenz	CC BY-SA 4.0

OParl ist ein Schnittstellenstandard für die strukturierte Suche nach offenen Verwaltungsdaten in Ratsinformationssystemen sowie für die unmittelbare strukturierte Kodierung und die Beschreibung derartiger Daten.

1 Internet Engineering Task Force (IETF): »RFC 7159: The JavaScript Object Notation (JSON) Data Interchange Format«, März 2014; https://datatracker.ietf.org/doc/html/rfc7159/

2 Internet Engineering Task Force (IETF): »RFC 3986 / STD 66: Uniform Resource Identifier (URI): Generic Syntax«, Januar 2005; https://datatracker.ietf.org/doc/rfc3986/

3 DIN ISO 8601:2006-09: »Datenelemente und Austauschformate – Informationsaustausch – Darstellung von Datum und Uhrzeit«; https://www.beuth.de/de/norm/din-iso-8601/90754621

Hintergrund
Seinen Ursprung hat OParl in einer Initiative von Verbänden, zivilgesellschaftlichen Organisationen und Initiativen, Softwareanbietern sowie interessierten Einzelpersonen. Diese hat sich Anfang der 2010er Jahre das Ziel gesetzt, die Bereitstellung strukturierter offener Verwaltungsdaten durch die Spezifikation einer an bestehenden Ratsinformationssystemen orientierten Webserviceschnittstelle und damit die Nutzung solcher Daten – auch durch engagierte Bürger:innen – zu fördern. Erklärte Unterziele sind möglichst geringer Aufwand für die Implementierung aufseiten der Softwareanbieter und für die Migration und Datenbereitstellung aufseiten der Betreiber.

Der Einsatz von OParl wurde beispielsweise im Rahmen des Open.NRW-Pilotprojektes »Kommunales Open Government« 2017 in 27 nordrhein-westfälischen Kommunen umgesetzt. Inzwischen unterstützen mehr als 120 Kommunen, Landkreise und Stadtbezirke die OParl-Schnittstelle.

OParl macht sich in hohem Maße das Linked-Data-Prinzip zunutze, indem z. B. für voneinander abhängige bzw. aufeinander bezogene Daten anstelle der unmittelbaren Daten URLs angegeben werden, unter denen diese Daten zu finden sind.

Schwachstelle von OParl-Implementierungen
Schwachstelle vieler mittels OParl beschriebener Datenbestände sind jedoch sehr eingeschränkte Suchmöglichkeiten. So ist es beispielsweise oft nicht möglich, anhand des Namens einer Person deren Eintrag zu suchen. Dies ist weder ein Mangel der Spezifikation noch dem Datenschutz geschuldet. Häufig können die Daten durch eine Websuche auf einer offiziellen Seite der Kommune gefunden werden, während eine OParl-Abfrage bestenfalls eine – oft sehr große – Menge von Personenbeschreibungen liefert, von allen Personen, die jemals eine erfasste kommunale Rolle innehatten (siehe das Anwendungsbeispiel unten). Diese Liste muss dann lokal nach dem Namen der Person durchsucht werden, um deren Beschreibung extrahieren zu können. Damit werden diese Implementierungen gerade nicht dem Anspruch von OParl gerecht, übertragene Datenmengen und Antwortzeiten zu minimieren.

Anwendungsbeispiel

https://oparl.oparlstadt.de/bodies.asp

liefert Basisinformationen über alle im OParl-Server von Oparlstadt repräsentierten Körperschaften, unter anderem die Information, dass die Körperschaft »Bezirksverordnetenversammlung von Oparlstadt« die »Body-ID« (Körperschaftskennnummer auf diesem OParl-Server) »1« hat:

```
»id«: »https://oparl.oparlstadt.de/bodies.asp?id=1«,
»name«: »Bezirksverordnetenversammlung von Oparlstadt«
```

Folgerichtig liefert

https://oparl.oparlstadt.de/persons.asp?body=1

die Beschreibungen aller Personen, die jemals in der Bezirksverordnetenversammlung eine Rolle innehatten.

Erstellen und Bereitstellen von Webseiten und -oberflächen

Die Bereitstellung von Informationen für Bürger:innen oder Unternehmen mittels visueller, audiovisueller oder akustischer Darstellung auf oder über Webseiten sowie die Interaktion mit Bürger:innen und Unternehmen mithilfe webbasierter Anwendungen werden immer wichtiger. Für alle genannten Darstellungsformen gibt es etablierte Standards, von denen die öffentliche Hand solche anwenden muss, die für den jeweiligen Zweck und die verwendeten Daten geeignet sind, da nur dann sichergestellt ist, dass die Informationen den Adressaten überhaupt und in der gewünschten Art und Weise präsentiert bzw. eingegebene oder hochgeladene Daten und Dokumente von öffentlichen Einrichtungen korrekt interpretiert werden können.

Inhaltlich sind auch für die Erstellung von Webseiten ethisch veranlasste Standards relevant, z. B., wenn es darum geht, welche Informationen zu Funktionsträger:innen veröffentlicht werden (dürfen) oder wie eine inklusive Gestaltung für Menschen mit und ohne Beeinträchtigungen aussehen sollte.

Metadaten

Wie gut die Auffindbarkeit und Nachnutzbarkeit einer Datensammlung ist, hängt vor allem von der Qualität der darauf bezogenen Metadaten ab. Metadaten sollten grundsätzlich strukturiert und einheitlichen Standards folgend Aufschluss über den Inhalt der beschriebenen Daten geben. Je nach Art der beschriebenen Daten kann auch die Quantität der Datenbeschreibung (z. B. Redundanz bei Schlüsselworten) nützlich sein, insbesondere bei nicht von vornherein festgelegter (Nach-) Nutzung und bei vielfältiger Nutzung durch unterschiedliche Disziplinen. Redundanz und Breite der Metadaten können eine einfache zweckbezogene Auffindbarkeit deutlich erleichtern. Relevante Metadaten können z. B. Aussagen zu Entstehungszeitpunkt und -ort der beschriebenen Daten, zur Erzeugungsmethode, zur Genauigkeit der Daten oder zu Lizenzbestimmungen sein.

Metadatenstandards, z. B. für die Katalogisierung offener Verwaltungsdaten (s. a. Infokasten DCAT-AP.de[30]), legen die Bezeichnungen für die einzelnen Metadatenpunkte, deren Datenformate und ggf. deren Codelisten[31] fest. Dabei können – insbesondere für die Formate der Metadaten – auch technische Standards Wiederverwendung finden, die originär für Produktivdaten spezifiziert wurden.

30 Insbesondere innerhalb der EU hat sich der Metadatenstandard DCAT-AP als »de facto«-Standard für die Katalogisierung offener Verwaltungsdaten durchgesetzt.

31 Codelisten sind Abbildungen zwischen menschenverständlichen Wertebezeichnungen (z. B. »Flensburg, Stadt«) und in Datensätzen stattdessen gespeicherten Werten (z. B. »01001000«).

DCAT-AP.de

Vollständiger Name	Deutsche Adaption des »›Data Catalogue Application Profile‹ (DCAT-AP) für Datenportale in Europa«
IT-Planungsrats-beschlüsse	2012/43 (erste Erwähnung, Metadaten zu standardisieren) 2013/29 (DCAT-AP indirekt erwähnt) 2013/33 (Ernennung eines Ansprechpartners) 2016/37 (Konkretisierung) 2018/11 (Fertigstellung) 2018/30 (verbindliche Nutzung)
Entwickelt von	]init[AG für GovData
Aktuelle Version	2.0 vom 01.03.2022
Erstveröffent-lichung	28.09.2017
Art	Metadatenstandard
Basiert auf	DCAT-AP[1], DCAT[2], ISO 15836-1[3], ADMS[4], FOAF[5], RDF[6]
Lizenz	CC BY 4.0[7]

1 Joinup: »DCAT Application Profile for data portals in Europe«; https://joinup.ec.europa.eu/collection/semic-support-centre/solution/dcat-application-profile-data-portals-europe/about

2 W3C Recommendation: »Data Catalog Vocabulary (DCAT) - Version 2«; https://www.w3.org/TR/vocab-dcat-2/

3 ISO 15836-1: »Information und Dokumentation – Das Dublin Core Metadaten Elemente Set – Teil 1: Core Elemente«, Mai 2017; https://www.beuth.de/de/norm/iso-15836-1/274483729

4 SEMIC: »Asset Description Metadata Schema (ADMS)«; https://semiceu.github.io/ADMS/releases/2.00/

5 Brickley, D. & Miller, L.: »FOAF Vocabulary Specification«; http://xmlns.com/foaf/0.1/

6 W3C: »Resource Description Framework (RDF)«; https://www.w3.org/RDF/

7 Creative Commons: »CC BY 4.0 Lizenzvertrag. Namensnennung 4.0 International«; https://creativecommons.org/licenses/by/4.0/legalcode.de

DCAT-AP.de ist ein Metadatenstandard für die Beschreibung und Katalogisierung offener Verwaltungsdaten in Deutschland. Der Standard soll die Auffindbarkeit und Wiederverwendbarkeit von Verwaltungsdaten, insbesondere über das Datenportal GovData[8], gewährleisten.[9]

Hintergrund

Die ersten Bestrebungen der deutschen Verwaltung, Metadaten zu standardisieren, gehen mindestens auf das Jahr 2012 zurück. In einem Zwischenbericht der Bund-Länder-Arbeitsgruppe »Open Government« des IT-Planungsrates wurde damals bereits der Bedarf für eine einheitliche Strukturierung der Metadaten im Rahmen eines Datenportal-Prototypen festgestellt. Die verbindliche Einführung eines entsprechenden Standards erfolgte 2018 mit Beschluss des IT-Planungsrates.

Bei dem DCAT-AP.de-Standard handelt es sich um die deutsche Adaption des vom W3C[10] herausgegebenen »Data Catalog Vocabulary« (DCAT) bzw. der von der europäischen SEMIC[11] gepflegten Spezifikation »DCAT Application Profile for data portals in Europe« (DCAT-AP). DCAT und davon abgeleitete Spezifikationen verwenden dabei Klassen und Eigenschaften bereits etablierter Standards, wie des ISO-Standards »Dublin Core«[12] sowie des Vokabulars »Asset Description Metadata Schema« (ADMS) und des RDF-Schemas »Friend of a Friend« (FOAF).[13] DCAT-AP und die spezifizierten nationalen Derivate haben sich in der EU als Metadatenstandards für offene Verwaltungsdaten etabliert. Das Datenportal der EU bietet hierfür auch das »Metadata Quality Assessment«[14] an, das als Online-Tool beschreibende Metadaten prüft und anhand einer Vielzahl von Indikatoren, wie der DCAT-AP-Konformität, bewertet. Im EU-weiten Vergleich liegt Deutschland in der Standardkonformität im oberen Viertel.[15] Die Auffindbarkeit offener Daten aufgrund standardkonformer Metadaten kann dabei auch einen Einfluss auf die Verwaltungseffizienz oder die Effektivität der Politikgestaltung haben. Laut dem Open Data Maturity Report 2022[16] ist dieser potenzielle »Impact« durch auffindbare offene Daten in Deutschland allerdings bislang nur durchschnittlich.

Die Offenheit der DCAT-Standardisierung ermöglichte früh die Einbindung des Standards insbesondere in entsprechende Open-Source-Projekte wie CKAN[17], die wiederum die Einrichtung etwaiger Datenportale vereinfachen.[18]

Anwendungsbeispiel

Das reale Anwendungsbeispiel zeigt die Verwendung von DCAT-AP.de in der Auszeichnungssprache XML. Die Metadaten beschreiben eine Liste von Trinkbrunnen in Steglitz-Zehlendorf (dct:title), die vom Bezirksamt Steglitz-Zehlendorf (dct:publisher) als CSV-Datei (dct:mediaType) bereitgestellt wird. Die Metadaten enthalten unter anderem den Speicherort (dcat:accessURL) und die Lizenz (dct:license) der Daten, beschreibende Schlagwörter (dcat:keyword), z. B. Wasser, sowie eine E-Mail-Adresse, um Kontakt aufzunehmen (dcat:contactPoint).

Das GovData-GitHub-Repository bietet darüber hinaus ausgewählte fiktive Anwendungsbeispiele in den Syntaxen von XML, JSON und Turtle an.

8 https://www.govdata.de/

9 Siehe]init[AG im Auftrag der GKSt GovData (2017): »DCAT-AP.de Pflegehandbuch«, S. 16; https://www.it-planungsrat.de/fileadmin/beschluesse/2018/Beschluss2018-30_TOP12_Anlage1_DCAD_AP.pdf

10 World Wide Web Consortium; https://www.w3.org/

11 Semantic Interoperability Community (SEMIC); https://joinup.ec.europa.eu/collection/semic-support-centre/about

12 ISO 15836-1: »Information und Dokumentation – Das Dublin Core Metadaten Elemente Set – Teil 1: Core Elemente«, Mai 2017; https://www.beuth.de/de/norm/iso-15836-1/274483729

13 Siehe]init[AG im Auftrag der GKSt GovData (2017): »DCAT-AP.de Pflegehandbuch«, S. 10; https://www.it-planungsrat.de/fileadmin/beschluesse/2018/Beschluss2018-30_TOP12_Anlage1_DCAD_AP.pdf

14 https://data.europa.eu/mqa/methodology?locale=de

15 Publications Office of the European Union. (2022): »Open Data Maturity Report 2022«, S. 84; https://doi.org/10.2830/70973

16 Publications Office of the European Union. (2022): »Open Data Maturity Report 2022«, S. 51; https://doi.org/10.2830/70973

17 https://ckan.org/

18 Nogueras-Iso, J., Lacasta, J., Ureña-Cámara, M. A., & Ariza-López, F. J. (2021): »Quality of Metadata in Open Data Portals«, IEEE Access, 9, S. 60364 – 60365; https://doi.org/10.1109/ACCESS.2021.3073455

```
<rdf:RDF>
<dcat:Dataset rdf:about="https://datenregister.berlin.de/dataset/e36a2425-8dae-4c96-8b23-7f06471ddf93">
  <dct:title>Trinkbrunnen in Steglitz-Zehlendorf</dct:title>
  <dct:description>Liste der öffentlichen Trinkbrunnen im Bezirk Steglitz-Zehlendorf.
  </dct:description>
  <dct:identifier>e36a2425-8dae-4c96-8b23-7f06471ddf93</dct:identifier>
  <dcat:keyword>gesundheit</dcat:keyword>
  <dcat:keyword>hitze</dcat:keyword>
  <dcat:keyword>trinkbrunnen</dcat:keyword>
  <dcat:keyword>wasser</dcat:keyword>
  <dct:issued rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2023-07-21T12:57:01.042600
  </dct:issued>
  <dct:modified rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2023-07-25T07:26:04.572245
  </dct:modified>
  <dct:language rdf:resource="http://publications.europa.eu/resource/authority/language/DEU"/>
  <dcat:theme rdf:resource="http://publications.europa.eu/resource/authority/data-theme/HEAL"/>
  <dct:conformsTo>http://dcat-ap.de/def/dcatde/1.0.1/</dct:conformsTo>
  <dcat:contactPoint>
    <vcard:Organization rdf:nodeID="N4e7b62c89dea4351b6e65e18e97ba3a2">
      <vcard:fn>Bezirksamt Steglitz-Zehlendorf</vcard:fn>
      <vcard:hasEmail rdf:resource="mailto:internet@ba-sz.berlin.de"/>
    </vcard:Organization>
  </dcat:contactPoint>
  <dct:publisher>
    <foaf:Organization rdf:nodeID="N139421419f9c44d8afe5dd51f07192de">
      <foaf:name>Bezirksamt Steglitz-Zehlendorf</foaf:name>
    </foaf:Organization>
  </dct:publisher>
  <dcat:distribution>
    <dcat:Distribution rdf:about="https://datenregister.berlin.de/dataset/e36a2425-8dae-
    4c96-8b23-7f06471ddf93/resource/de71020f-34c6-47e5-b179-a99c56e45832">
      <dct:title>Trinkbrunnen in Steglitz-Zehlendorf</dct:title>
      <dct:license rdf:resource="http://dcat-ap.de/def/licenses/cc-by"/>
      <dcat:accessURL rdf:resource="https://www.berlin.de/ba-steglitz-zehlendorf/service/
      daten/trinkbrunnen-in-steglitz-zehlendorf.csv"/>
      <dcat:mediaType>text/csv</dcat:mediaType>
      <dct:format rdf:resource="http://publications.europa.eu/resource/authority/file-type/CSV"/>
      <dct:issued rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2023-07-25T07:26:04.605753
      </dct:issued>
      <dct:modified rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2023-07-25T07:26:04.578846
      </dct:modified>
      <dcatde:licenseAttributionByText>Bezirksamt Steglitz-Zehlendorf</dcatde:licenseAttributionByText>
    </dcat:Distribution>
  </dcat:distribution>
  <dcatde:politicalGeocodingLevelURI rdf:resource="http://dcat-ap.de/def/politicalGeocoding/Level/
    administrativeDistrict"/>
  <dcatde:contributorID rdf:resource="http://dcat-ap.de/def/contributors/berlinOpenData"/>
  <dcatde:politicalGeocodingURI rdf:resource="http://dcat-ap.de/def/politicalGeocoding/regionalKey/
    110010001006"/>
  <dct:spatial>
    <dct:Location rdf:nodeID="N9219d6ec325c4e7093e83f937e448fe9">
      <locn:adminUnitL2 rdf:resource="http://dcat-ap.de/def/politicalGeocoding/regionalKey/110010001006"/>
    </dct:Location>
  </dct:spatial>
</dcat:Dataset>
</rdf:RDF>
```

Technischer und organisatorischer Schutz von Daten

Neben den spezifischen Standards für die Vielzahl von Datenarten, die bei der öffentlichen Hand relevant sind, bedarf es im öffentlichen Sektor wie in allen modernen Unternehmen angemessener allgemeiner Standards für Informationssicherheit und Datenschutz. Dies ist notwendig, weil viele datengetriebene Prozesse der öffentlichen Hand auf aktuelle, korrekte und verfügbare Daten angewiesen sind und in großem Umfang personenbezogene Daten (von Bürger:innen sowie eigenen und fremden Mitarbeitenden) sowie schutzpflichtige Unternehmensdaten elektronisch verarbeitet werden müssen. Daneben hält die öffentliche Hand naturgemäß in großem Umfang Daten, die unter Staatsschutzgesichtspunkten geheim zu halten sind.

Für einige Datenarten, die in den in diesem Abschnitt genannten Kategorien verwendet werden – z. B. personenbezogene Daten oder Geschäftsgeheimnisse – sind beispielsweise in Gesetzen spezifische Schutzanforderungen und teilweise auch abstrakte Schutzmaßnahmen festgelegt. Darüber hinaus enthalten manche technischen Datenstandards solche Anforderungen und legen in einigen Fällen – z. B. durch Referenzierung der entsprechenden Standards – die konkret anzuwendenden Maßnahmen fest. Daneben werden stets auch generelle Regelungen und entsprechende Standards benötigt, die auf die Gegebenheiten der jeweiligen Behörde oder Einrichtung als Ganzes abgestimmt sind.

Ethischer Umgang mit Daten

Öffentliche Einrichtungen sind in unterschiedlichem Maß mit Fragen zum ethischen Umgang mit Daten konfrontiert, sowohl was die Häufigkeit des Auftretens solcher Fragen als auch was die Tragweite der jeweiligen Entscheidung betrifft. Für einige Datenarten und Datenzusammenstellungen existieren bereits übergeordnete, ethisch veranlasste Rechtsnormen. Diese enthalten z. B. spezifische Schutzanforderungen – deren Umsetzung durch Standards für den technischen und organisatorischen Schutz von Daten konkretisiert werden muss – oder Erfassungs-, Speicherungs- oder Verarbeitungsverbote. Für den Bereich inklusiver Datenpräsentation und Interaktion gibt es sowohl für öffentliche Stellen verbindliche Rechtsnormen[32] als auch Normen und weitere Standards, die konkrete Barrierefreiheitsanforderungen festlegen[33].

Öffentliche Stellen sollten daher derartige Fragen – auch proaktiv – ermitteln und bewerten sowie verpflichtende bzw. geeignete Standards identifizieren und diese bei Notwendigkeit an die Gegebenheiten der Stelle anpassen bzw. eigene Standards entwickeln (lassen), sofern noch nichts Passendes existiert. Ein Beispiel, das unter anderem Aspekte verschiedener existierender Regelwerke vereint, ist das Datenethikkonzept der Stadt Ulm[34].

Die immer noch rasant wachsenden Möglichkeiten Daten zu erfassen, zusammenzustellen und zu verarbeiten, sowie neu erkannte Formen des Datenmissbrauchs erfordern eine regelmäßige Überprüfung der Angemessenheit der in der jeweiligen Stelle etablierten Regelwerke und Standards.

---

32 Z. B. die »Verordnung zur Schaffung barrierefreier Informationstechnik nach dem Behindertengleichstellungsgesetz (Barrierefreie-Informationstechnik-Verordnung – BITV 2.0)« (Barrierefreie-Informationstechnik-Verordnung vom 12. September 2011 (BGBl. I S. 1843), die zuletzt durch Artikel 1 der Verordnung vom 21. Mai 2019 (BGBl. I S. 738) geändert worden ist); https://www.gesetze-im-internet.de/bitv_2_0/index.html

33 Z. B. DIN EN 301549: »Barrierefreiheitsanforderungen für IKT-Produkte und -Dienstleistungen«; Juni 2022; https://www.beuth.de/de/norm/din-en-301549/353869627

34 https://www.ulm.de/aktuelle-meldungen/z%C3%B6a/oktober-2020/datenethikkonzept-2020_10

6. Was macht einen guten (datenbezogenen) Standard aus?

Nur das ist festgelegt, was tatsächlich notwendig ist.

Überflüssige Festlegungen können den Blick auf das Wesentliche verstellen, vermeidbare Einschränkungen verursachen und schlimmstenfalls zu (vermeintlicher) Nicht-Konformität zum Standard oder zu (vermeintlicher) Nicht-Interoperabilität zwischen Systemen führen. In der Regel ist es beispielsweise nicht erforderlich, in einem technischen Datenaustauschstandard (Kommunikationsstandard) festzulegen, wie die auszutauschenden Daten lokal in den am Austausch beteiligten Systemen gespeichert werden.

Referenzieren geht vor Neuspezifizieren, Vereinheitlichen geht vor Optimieren jedes Einzelfalles.

Die explizite inhaltliche Einbeziehung (Referenzierung) bereits etablierter, stabiler Standards (bzw. konkreter Teile solcher Standards) ist stets die bessere Alternative zu einer erneuten Festlegung gleichartiger Sachverhalte.[35] Die inhaltliche Einbeziehung existierender Standards ermöglicht nicht nur die mengenmäßige Reduzierung eines Standards, sondern vor allem die Nutzung des in die existierenden Standards eingeflossenen Know-hows und der bei ihrer Entwicklung gemachten Erfahrungen. Zudem erleichtert gemeinsame Referenzierung die Interoperabilität bzw. Abbildbarkeit zwischen Standards.

Vor einer von etablierten Standards abweichenden Festlegung empfiehlt es sich, die Notwendigkeit eines solchen Vorgehens sorgfältig zu prüfen und nachvollziehbar zu begründen. In der Regel behindern voneinander abweichende Festlegungen den Datenaustausch zwischen Systemen, für den Austausch erfordern sie oft Umkodierungen der Daten und führen zu Präzisions- oder Kontextverlusten. Zudem ist die technische Realisierung solcher Umkodierungen eine häufige, meist von vornherein vermeidbare Fehlerquelle.

Insbesondere Unternehmen übertragen Daten bereits in erheblichem Maß elektronisch und oft in großen Mengen. Dabei werden dieselben Daten vielfach mehrmals genutzt. Gerade dabei gilt es, bereits verbreitet eingesetzte Standards auch für neue Anwendungen zu benutzen, um den Anpassungsaufwand aller Beteiligten gering zu halten.

Der Zweck ist angegeben.

Technische (Meta-)Datenstandards können unter verschiedenen Gesichtspunkten geschaffen werden. Beispiele sind: eindeutige Abbildung zwischen Datendarstellungen (z. B. Transliteration), effiziente Datenspeicherung, -verarbeitung oder -übertragung, Verarbeitung von Daten in verschlüsseltem Zustand (homomorphe Verschlüsselung), Bitfehler-resistente Datenspeicherung oder -übertragung und vieles mehr. Ein guter Standard besitzt vor allem eines: einen klaren, wohldefinierten und überschaubaren Zweck.

In der Kürze liegt die Würze.

In der Regel ist es vorteilhaft, wenn der Standard einen eher geringen Umfang besitzt.[36] Je komplexer ein Standard, desto größer das Risiko, dass er unentdeckte Lücken, Widersprüche oder andere Fehler enthält. (Für eigene Standardisierungsprojekte siehe dazu auch Abschnitt 7.)

35 Ein Beispiel für intensive Referenzierung: »DCAT also makes extensive use of terms from other vocabularies, in particular Dublin Core […]. DCAT defines a minimal set of classes and properties of its own.« (Data Catalog Vocabulary (DCAT) – Version 3; https://www.w3.org/TR/vocab-dcat-3/).

36 Prominentes Gegenbeispiel ist der Unicode-Standard (https://unicode.org/main.html), der inzwischen rund 1700 Druckseiten umfasst. Dies allerdings ist seinem besonderen Zweck geschuldet, möglichst alle gängigen und historischen (Schrift-)Zeichen auf einen eindeutigen numerischen Code abzubilden.

Erweiterungsperspektiven sind berücksichtigt.

Nur selten ist vorab absolut klar, dass der Standard alle – auch zukünftig – möglichen Fälle abdeckt. Er sollte daher technisch einfach erweiterbar sein, ohne dass ein Bruch gegenüber Regeln, Algorithmen usw. des aktuellen Standards erforderlich ist.

Vorangehende und folgende Transport- und Verarbeitungsschritte der Daten sind mitbedacht.

Dazu gehört insbesondere, im vor- oder nachgelagerten Teil der Verarbeitung der Daten bereits etablierte Standards nach Möglichkeit in neuen Verarbeitungsschritten wiederzuverwenden, um Umkodierungen und neue Fehlerquellen weitestgehend zu vermeiden. Ein guter datenbezogener Standard muss und kann dabei nicht unbedingt für jeden Schritt und jede Art der Speicherung und Verarbeitung eines bestimmten Datums gleichermaßen gut geeignet sein.[37] Es kann beispielsweise aus Effizienz- und Geschwindigkeitsgründen vorteilhaft sein, dass Daten in dem Format gespeichert sind, in dem sie – z. B. im Rahmen eines Fachverfahrens – an andere Stellen übermittelt werden. Dies ist jedoch nicht unbedingt notwendig. Voraussetzung ist allerdings stets, dass sich beide Formate eindeutig ineinander überführen lassen und bei dieser Überführung keine notwendigen Datenbestandteile und keine notwendigen Datendetails verloren gehen.

Wie in Abschnitt 2.2 beschrieben, muss bei Erfassung, Speicherung, Verarbeitung und Transport von Daten in allen vorangegangenen Einzelschritten ein einheitliches Mindestniveau an Informationssicherheit eingehalten worden sein, damit dieses für einen beliebigen Zeitpunkt der Datenexistenz angenommen werden kann. Ein isolierter und nur für bestimmte Einzelschritte geeigneter oder ein nicht durchgängig verpflichtender Informationssicherheitsstandard können daher ihren Zweck verfehlen.

Ethische Erwägungen können die Offenlegung einer bestimmten anonymisierten strukturierten Datenmenge, die auf personenbezogenen Datensätzen beruht, zulassen. Wenn jedoch das Risiko besteht, dass Daten unzulässigerweise durch ebenfalls verfügbare Daten de-anonymisiert werden können, müssen gegebenenfalls in ethisch veranlassten Standards entsprechende Verarbeitungsbeschränkungen festgelegt oder die Offenlegung der Daten verboten werden.

Vielfältige und breite Nutzbarkeit sind inhaltlich und organisatorisch möglich.

Zumindest mittel- bis langfristig betrachtet ist ein guter Standard einer, der von möglichst vielen Nutzenden für seinen Zweck verwendet wird und sich dabei gegebenenfalls erfolgreich gegen konkurrierende Standards durchgesetzt hat. Die hier betrachteten technischen (Meta-)Datenstandards dienen in erster Linie der Verständlichkeit und Eindeutigkeit der Daten über verschiedene Hürden hinweg, wobei diese Hürden beispielsweise unterschiedliche Sprachen, Schriftsysteme, kulturelle Hintergründe oder Zeitepochen, aber auch differierende technische Voraussetzungen bzw. funktionale oder qualitative Anforderungen der Beteiligten sein können. In einem solchen Zusammenhang kann durchaus ein Standard (bzw. ein Standardbündel), der einen etwas breiteren Zweck erfüllt oder nicht die für die aktuelle Technik effizienteste Lösung darstellt, der »bessere« Standard sein, wenn er z. B. bereits etabliert ist oder sich besser in Transport- und Verarbeitungsketten einordnet.

Zur vielfältigen und breiten Nutzbarkeit von Standards trägt auch bei, wenn diese zwar präzise, aber möglichst allgemeinverständlich formuliert sind und dabei z. B. – soweit möglich – auf sektorspezifische Terminologie verzichtet wird. Auch die Formulierung eines Standards in englischer Sprache kann selbst dann vorteilhaft sein, wenn zunächst nur von einer Nutzung in Deutschland ausgegangen wird. Durch das wirtschaftliche und gesellschaftliche Zusammenwachsen der Staaten des Europäischen Wirtschaftsraumes ergeben sich mehr und mehr transnationale Datenflüsse und gemeinschaftliche Standardisierungsbedarfe. Ein in englischer Sprache vorliegender Standard kann in einem solchen Fall unmittelbar als Diskussionsgrundlage eingebracht oder sogar bereits vorab freiwillig in anderen Staaten genutzt werden, wohingegen die Verbreitung eines ausschließlich in deutscher Sprache verfügbaren Standards bereits an Sprachbarrieren scheitern kann.

37 Ein technischer Datenstandard für die Speicherung von Daten in einem Computer und den effizienten Zugriff auf die gespeicherten Daten muss berücksichtigen, welche Werte die Daten annehmen können, wie Daten in Speichermedien organisiert sind, wie die Daten zwischen Speichermedium und Verarbeitungseinheit des Computers übertragen werden und welche Operationen auf den Daten möglich sind. Nehmen wir an, es ginge um ein einfaches Bestellsystem, bei dem die maximale Bestellnummer kleiner als 1.000.000 ist. Angesichts der heutzutage niedrigen Anschaffungskosten für Speichermedien könnte ein entsprechender Standard beispielsweise vorsehen, dass die Bestellnummer, für die nur 3 Bytes (Wertebereich für natürliche Zahlen: 0 – 16 777 215) notwendig wären, in 4 Bytes gespeichert wird, weil dadurch der Zugriff in modernen Rechnern effizienter ist. Dieser fiktive Standard wäre beispielsweise offensichtlich ungeeignet, wenn ein solches Datum auf einem (ebenfalls fiktiven) RFID-Chip mit nur 3 Bytes Kapazität gespeichert werden soll.

Ein guter Standard ist öffentlich und über eine gut auffindbare, zentrale Stelle verfügbar.[38] Bei Normungsorganisationen und Standardisierungsgremien kann dies z. B. deren Website oder ein dedizierter Webshop (https://www.beuth.de) sein, andere gute Beispiele sind Themenportale wie das XRepository (https://www.xrepository.de) der Koordinierungsstelle für IT-Standards (KoSIT)[39]. Auch andere Organisationen wie beispielsweise das BSI und die Bundesnetzagentur für Elektrizität, Gas, Telekommunikation, Post und Eisenbahnen[40] entwickeln und veröffentlichen Standards (von diesen »Technische Richtlinie (TR)« genannt). Dass auf den Websites solcher vergleichsweise spezifischen und oft national fokussierten Organisationen allerdings potenzielle Nutzende – auch der öffentlichen Hand – überhaupt nachschauen, dürfte bereits weniger wahrscheinlich sein.

Normungsorganisationen und privatwirtschaftliche Standardisierungsgremien finanzieren sich teilweise durch den Verkauf ihrer Standards, weshalb derartige Standards oft nicht kostenfrei sind. Dies steht zwar der Nutzbarkeit nicht entgegen, kann aber das Ziel einer verbreiteten Nutzung beeinträchtigen.

Er wurde gemeinsam mit allen Betroffenengruppen entwickelt.

Nicht nur, aber auch unter dem Gesichtspunkt der anzustrebenden Nutzungsmaximierung muss ein guter Standard die Einsatzumgebung des Standards bei allen angestrebten oder verpflichteten Nutzenden berücksichtigen. Auch wenn z. B. die öffentliche Hand die zu verwendenden datenbezogenen Standards für eine gesetzlich vorgeschriebene Datenlieferung festlegt, sind auch die Belange der Datenlieferant:innen angemessen zu berücksichtigen. Optimale Praxistauglichkeit und Nutzungsmaximierung lassen sich in der Regel nur erreichen, wenn von Dritten zu befolgende datenbezogene Standards gemeinsam mit diesen spezifiziert werden. Dabei kann es vorteilhaft sein, die Entwicklung im Rahmen eines offenen Standardisierungsgremiums durchzuführen, das das Standardisierungsvorhaben rechtzeitig publiziert und interessierten Betroffenengruppen die Möglichkeit der Teilnahme bietet.[41]

Das erfolgreiche Ergebnis einer solchen Zusammenarbeit von u. a. Softwareanbietern, Wirtschaftsprüfern und öffentlichen Stellen sind die nationalen Taxonomien für die Finanzberichterstattung auf Basis von XBRL (eXtensible Business Reporting Language).[42] Die Verwendung der entsprechenden XBRL-Taxonomie ist beispielsweise für die Abgabe der Bilanz gegenüber den Finanzbehörden vorgeschrieben.

Er wird nach einem transparenten Verfahren gepflegt und weiterentwickelt.

Kein Standard ist dauerhaft vollkommen. Oft gibt es kleine Ungenauigkeiten und Fehler, die Expert:innen als solche erkennen und mit denen sie intuitiv korrekt im Sinne des Zweckes bzw. der Ersteller umgehen. Für nicht fachkundige potenzielle Nutzende oder Personen, die sich zunächst über den Inhalt eines Standards informieren möchten, können derartige Mängel jedoch eine gravierende Hürde für das Verständnis darstellen.

Technischer Fortschritt, sich weiterentwickelnde Nutzungsanforderungen und Änderungen des regulatorischen Rahmens können ebenso wie Erkenntnisse aus dem Betrieb standardkonformer Lösungen die Anpassung von Standards erfordern.

Deshalb muss ein guter Standard auch nach seiner erstmaligen Fertigstellung von einem fachlich qualifizierten Gremium weiterhin betreut, gepflegt und angepasst werden.

Die Entwicklung ist nachvollziehbar.

In den Entwicklungsphasen eines Standards sind viele Entscheidungen verschiedenster Art zu fällen, z. B. zu den Grenzen des Standards oder zu prioritären bzw. später zu bearbeitenden Inhalten, aber auch zur konkreten Kodierung von Objekten oder der Gestaltung von Kommunikationsabläufen. Für manche Entscheidungen gibt es gute Gründe, bei anderen gilt es lediglich, eine von vielen geeigneten Alternativen auszuwählen. Für das Verständnis eines Standards und um Erkenntnisse des Entwicklungsprozesses für weitere Standards nutzbar zu machen, ist es hilfreich, die wesentlichen Entscheidungen zu dokumentieren und offen zu legen. Ebenso sollten ältere, nicht mehr gültige Versionen eines Standards für Interessierte leicht zugänglich sein, da auch aus dem Vergleich nützliche Erkenntnisse gewonnen werden können.

38 Bzgl. des öffentlichen Sektors siehe dazu auch: Databund: »Top10 – Standards für die Digitalisierung der öffentlichen Verwaltung«, 31.07.2023; https://databund.de/wp-content/uploads/sites/20/2023/08/DATABUND-Top-10-der-notwendigen-Standards.pdf

39 https://www.xoev.de/

40 https://www.bundesnetzagentur.de/cln_112/DE/Home/home_node.html

41 Siehe dazu auch: Deutsches Institut für Normung (DIN): »Whitepaper Normung und Standardisierung bei der Digitalisierung er öffentlichen Verwaltung«, Januar 2023, Abschnitt 3.1: »Grundsatz der Transparenz«, https://www.din.de/resource/blob/892574/33686a700a1d35e688b5e88c06a6a49c/whitepaper-normung-standardisierung-digitalisierung-oeffentliche-verwaltung-data.pdf

42 https://de.xbrl.org/ueber-uns/ueber-uns-mitglieder/

7. Vom guten Umgang mit (datenbezogenen) Standards

7.1 Generelle Gesichtspunkte

Profile festlegen.

Ein Profil spezifiziert, welcher Teil bzw. welche Teile eines Standards im konkreten Fall zum Einsatz kommen. Bei einem umfangreichen Standard, der nicht immer komplett benötigt wird, ist es hilfreich, ein oder mehrere Profile festzulegen. Dadurch reicht es zur Herstellung von Interoperabilität und Konformität, die im vereinbarten Profil festgelegten Teile zu unterstützen. Dieses Vorgehen kann auch vorteilhaft sein, wenn ein Standard Teile beinhaltet, die noch stark in Entwicklung sind, aber nur bereits stabile Teile benötigt werden.

Profile können helfen, Fehler zu vermeiden und Implementierungen übersichtlich und klein zu halten. Die Größe einer Implementierung spielt auch heute noch oft eine Rolle, z. B. bei speicherbeschränkten Geräten, wie es Sensoren und Aktuatoren vielfach noch sind.

Bezug festlegen.

Bei Standards kann auf die stets aktuellste oder auf eine ganz bestimmte, dauerhaft feststehende Version Bezug genommen werden.

Die Bezugnahme auf die stets aktuellste Version erfordert in der Regel die Festlegung einer Übergangszeit, nach der nach einem – genau definierten – Aktualisierungsereignis diese Version tatsächlich unterstützt werden muss, da ja erst nach der formalen Aktualisierung etwaige Änderungen an Prozessen, Software und Ähnlichem (abschließend) umgesetzt werden können.

Vorteil dieser Variante ist, dass alle konformen Implementierungen sich gemeinsam mit dem Standard weiterentwickeln und Änderungen und Ergänzungen nach kurzer Zeit überall zur Verfügung stehen und ohne weitere Absprachen genutzt werden können. Nachteilig ist, dass alle (potenziell) interoperierenden Partner einen dauerhaften Aktualisierungsprozess benötigen, der von jeder neuen Version des Standards getriggert wird und Anpassungskosten verursacht.

Vorteil einer Bezugnahme auf eine bestimmte Version eines Standards ist die Stabilität einer entsprechenden Implementierung. Sollen allerdings spätere Änderungen des Standards doch zum Einsatz kommen, ist ein Prozess erforderlich, dies mit den (potenziellen) Interoperationspartner:innen abzustimmen und – wie bei der kontinuierlichen Anpassung – den Verfügbarkeitszeitpunkt festzulegen.

In beiden Fällen kann es erforderlich sein, für einen Übergangszeitraum oder sogar dauerhaft auch ältere Versionen des Standards weiterhin zu unterstützen. Auch dies muss festgelegt werden. Der Bezug kann auch Teil eines Profils (s. o.) sein.

7.2 Eigene Standards

Bei eigenen Standards sind diese weiteren Punkte hilfreich:

Die horizontale und die vertikale Modularisierbarkeit im Blick behalten.

Es sollte geprüft werden, ob sich der Zweck horizontal oder vertikal sinnvoll zerlegen lässt und ein Standardbündel (horizontale Zerlegung) oder einen Standardstack (vertikale Zerlegung) geschaffen werden kann.

Horizontale Modularisierung: Daten, die zwar in einem (Fach-) Verfahren gemeinsam genutzt werden, aber weder in starker Wechselbeziehung zueinander stehen, noch gemeinsam ein- oder ausgeliefert werden, müssen nicht unbedingt vom selben Standard abgedeckt sein. Oft ist es vorteilhafter, sie in kleineren

separaten Standards zu behandeln, die für das konkrete Verfahren gemeinsam referenziert werden.

Es kann beispielsweise sinnvoll sein, in der Leistungsverwaltung für die (personenbezogenen) Daten der Leistungsempfänger:innen und die (personenunabhängigen) leistungsbezogenen Daten separate technische Datenstandards zu verwenden. Dies erleichtert die verfahrensübergreifende Erkennung, Spezifikation und Wiederverwendung von gleichartigen Elementen.

Vertikale Modularisierung: Wie bereits in Abschnitt 4 ausführlich dargestellt, besitzen Daten häufig mehrere, sozusagen vertikal aufeinander aufbauende syntaktische und semantische Ebenen. Ähnliches gilt für die Datenkommunikation. Wenn man diese Ebenen sorgfältig und in sinnvoller Weise identifiziert und separiert, ist es einfacher, die Anforderungen einer konkreten Ebene zu erkennen, sich darauf zu konzentrieren und für diese eine gute Lösung zu finden. Anforderungen der Ebenen untereinander müssen dabei getrennt von der Betrachtung genau einer Ebene gesammelt werden, damit sie in die Spezifikation aller betroffenen Ebene(n) einfließen können.

Wenn beispielsweise die Betrachtung, welche Daten über einen Leistungsempfänger:innen vorliegen müssen, um über die Leistungsgewährung entscheiden zu können, davon entkoppelt wird, wie diese Daten kodiert und wie sie in einem Datenspeicher abgelegt oder über einen elektronischen Kommunikationsweg versandt werden, erleichtert dies den Überblick über die Daten der »Entscheidungs«ebene erheblich. Natürlich müssen die benutzten Ebenen geeignet spezifiziert sein bzw. werden, um die an sie gerichteten Anforderungen – z. B. zu den erforderlichen Wertebereichen der Daten – zu erfüllen.

Beides – horizontale wie vertikale Modularisierung – erleichtert zudem das Erkennen der Existenz geeigneter Standards und damit deren Wiederverwendung.

Offene Standardisierung bevorzugen.

Insbesondere im Open-Data-Kontext sowie beim Datenaustausch mit Dritten außerhalb des öffentlichen Sektors sind internationale Standards vorteilhaft. Im Zuge europäischer Integrationsbestrebungen und einer Internationalisierung des Verwaltungsangebotes sind entsprechende Schnittstellen und abgestimmte Standards notwendig. Der deutsche öffentliche Sektor sollte hierfür aktiv an der (Weiter-)Entwicklung und Ausgestaltung der Standards in entsprechenden Gremien und Komitees mitwirken, um neben der inhaltlichen Weiterentwicklung der Standards auch eigene Interessen im jeweiligen Standard durchsetzen zu können. Je größer die Nutzung und der Kreis der Nutzenden eines Standards ist, desto größer ist die Wahrscheinlichkeit, dass er weitere Nutzende hinzugewinnt und dass er aus Eigeninteresse der Nutzenden weiterentwickelt und aktuell gehalten wird. Der Erfolg des datenbezogenen Standards wäre so – ähnlich wie bei (Open-Source-)Software-Produkten[43] – durch transparente und offene Organisationsstrukturen selbstverstärkend. Dies kommt wiederum auch den nutzenden öffentlichen Stellen zugute.

Existieren fachspezifische Standards nicht oder nur in unzureichender Weise, sollten sie gemeinsam mit entsprechenden Interessengruppen entwickelt werden. So könnten für potenzielle Betroffene eines Standards offene bzw. zugängliche Arbeitsgruppen gegründet werden, die die Entwicklung und Pflege des jeweiligen Standards sicherstellen. Dabei sollte über die Arbeitsweise und Teilnahmemöglichkeiten möglichst detailliert informiert werden. Der europäische Metadatenstandard DCAT-AP (s. a. Infokasten S. 28) beispielsweise nutzt zur Bekanntmachung und Informationsbereitstellung die Kollaborationsplattform Joinup[44], die dem Austausch über Digitalisierungsprojekte der EU und ihrer Mitgliedstaaten dient. Darüber hinaus werden etablierte Versionsverwaltungssysteme genutzt, um die niedrigschwellige Beteiligung externer Fachleute zu erleichtern. Das Projektteam von DCAT-AP betreibt ein GitHub-Repository[45] als zentrale Anlaufstelle für konkrete inhaltliche Erweiterungsvorschläge, Korrekturbedarfe und Verbesserungsvorschläge für die Dokumentation sowie generelle Kommentare.

Interessant für (internationale) Normung und Standardisierung werden.

Es kann sinnvoll sein, zu versuchen, einen Bedarf über eine (internationale) Normungs- bzw. Standardisierungsaktivität abzudecken. Dazu bedarf es in der Regel weiterer Mitstreiter:innen außerhalb des öffentlichen Sektors bzw. außerhalb Deutschlands, z. B. die Unterstützung und Mitarbeit von Fachverfahrensherstellern, von Fachdienstleistern für betroffene Unternehmen oder von öffentlichen Stellen anderer europäischer Staaten.

Um das Interesse und die Mitarbeit von Unternehmen und weiterer Organisationen zu erreichen, wird es häufig notwendig sein, den Fokus des zu schaffenden Standards weiter zu fassen als für den unmittelbaren Bedarf des deutschen öffentlichen Sektors notwendig. Eine derartige Vorgehensweise kann trotzdem vorteilhaft sein, wenn dadurch z. B. weiterer Fachverstand eingebunden werden kann, umfassendere Prozesse effizienter gestaltet werden können oder der Standard zukunftssicherer aufgestellt werden kann.

---

43 Vgl. Thapa, Basanta E. P.; Weidner, Christian; Grosch, Dorian: »Ein Open-Source-Ökosystem für die öffentliche Verwaltung«, Kompetenzzentrum Öffentliche IT, 1. Auflage, August 2022; https://www.oeffentliche-it.de/publikationen/?doc=255694

44 https://joinup.ec.europa.eu/collection/semic-support-centre/solution/dcat-application-profile-data-portals-europe

45 https://github.com/SEMICeu/DCAT-AP/issues

8. Übergeordnete Handlungsempfehlungen

Neben den zahlreichen Hinweisen und Empfehlungen zu guten Standards (siehe Abschnitt 6) und dem guten Umgang mit Standards (siehe Abschnitt 7), die sich in erster Linie an die Nutzenden und die Entwickelnden von Standards richten, gibt es im Folgenden einige Empfehlungen für umfassendere Maßnahmen, die der Unterstützung durch Entscheider:innen bzw. Strateg:innen bedürfen.

»Verbündete« für (datenbezogene) Standardisierungsbedarfe suchen.

Egal, ob ein Standard in Eigenregie oder unter dem Dach einer Standardisierungsorganisation erstellt werden soll: Es ist stets vorteilhaft, Standardisierungsbedarfe öffentlich zu machen und andere Bedarfsträger:innen sowie weitere Betroffene zur Mitarbeit zu ermuntern. So können Know-how zusammengeführt, unterschiedliche Perspektiven eingebunden, breite Nutzbarkeit und Unterstützung erreicht sowie Parallel- und Fehlentwicklungen vermieden werden.

Verbindliche und empfohlene (datenbezogene) Standards katalogisieren und die Informationen dem gesamten öffentlichen Sektor über ein zentrales Portal zugänglich machen.

Viele datenbezogene Standards, die für Teile des öffentlichen Sektors unmittelbar oder mittelbar verbindlich oder empfohlen sind, können auch von anderen öffentlichen Stellen oder in weiteren Anwendungsfeldern eingesetzt werden, vorausgesetzt, dass dort das notwendige Wissen über die Existenz und Eignung der Standards existiert. Dazu ist es auch hilfreich, unmittelbar zu den jeweils zuständigen Pflegegremien und zu bereits nutzenden Stellen Kontakt aufnehmen zu können.

Etablierte (datenbezogene) Standards für den gesamten öffentlichen Sektor leicht, zentral und pauschal zugänglich machen.

Damit eine öffentliche Stelle die konkrete Eignung eines datenbezogenen Standards prüfen kann, um ihn z. B. in einer Ausschreibung für einen neuen Anwendungsbereich verbindlich vorgeben zu können, muss sie Zugang zu den infrage kommenden Standards haben. Hilfreich ist dazu ein auf die Belange des öffentlichen Sektors abgestimmtes zentrales Portal mit zeitgemäßen Suchmöglichkeiten, sodass die Interessierten keine Vorkenntnisse über Standardisierungsgremien und Fachtermini benötigen.

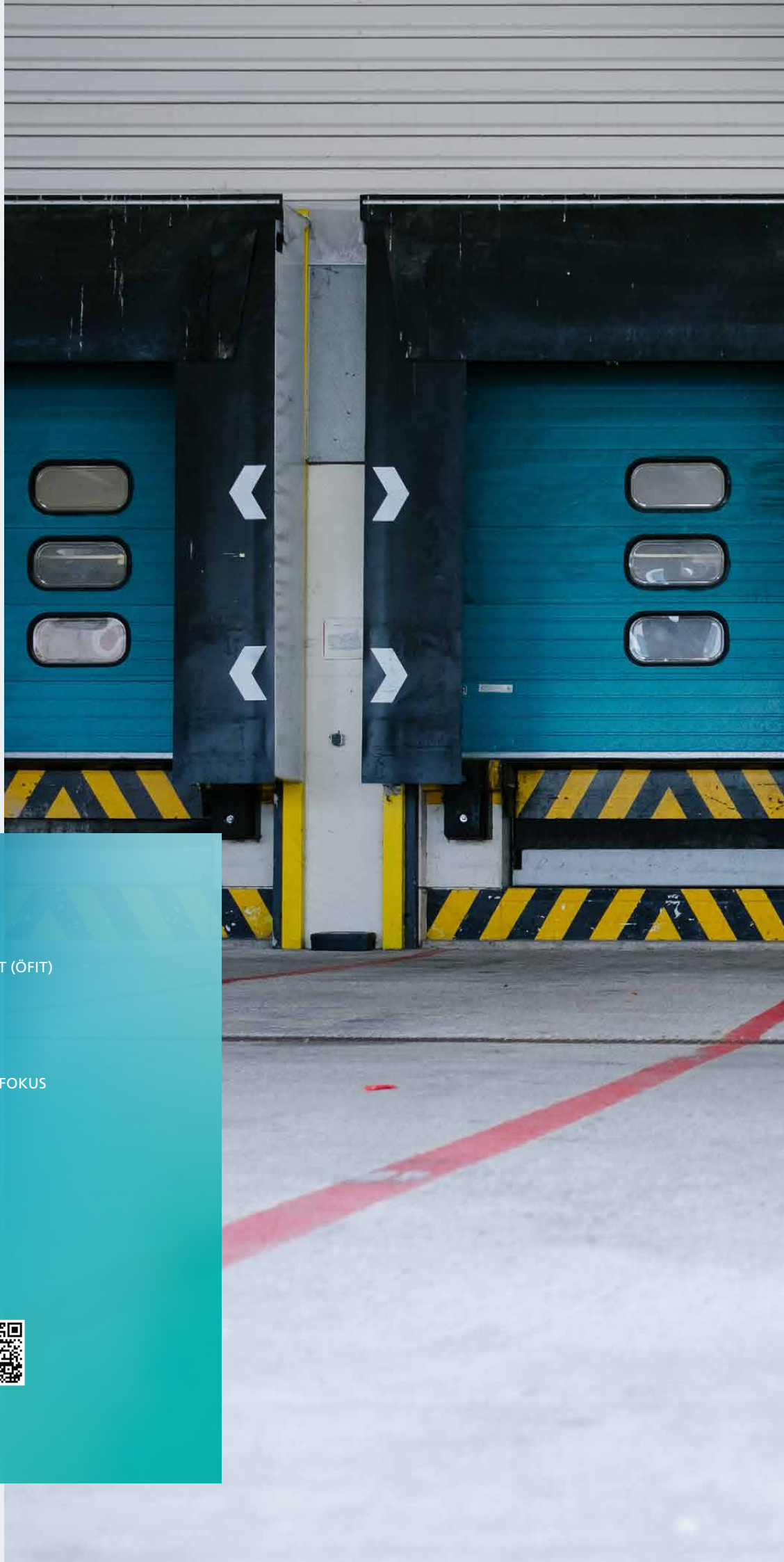
Abrechnungsmodelle, die der abrufenden Stelle jeden Zugriff auf eine konkrete Norm separat in Rechnung stellen, behindern die breite Berücksichtigung von Normen, obwohl gerade für diese die Annahme der Eignung für den definierten Zweck gilt.

Ansprechpartner:innen für (datenbezogene) Standards etablieren.

Kleinere Organisationen, egal ob Unternehmen oder öffentliche Stellen, können das erforderliche Wissen zu datenbezogenen Standards und eventuell notwendigen Standardisierungsprozessen meist weder effektiv noch wirtschaftlich vertretbar allein vorhalten. Deshalb sind für diese Organisationen übergreifend zuständige Ansprechpartner:innen erforderlich, um zügig und flächendeckend zu auf koordinierten Standards basierenden Lösungen zu gelangen.

Verfolgung von Standardisierungsprojekten für alle Betroffenen ermöglichen.

Beteiligung an einem Standardisierungsprojekt kann auf vielfältige Weise geschehen, z. B. durch kontinuierliche aktive Mitarbeit, durch gezielte Mitarbeit in bestimmten Phasen oder auch durch bloße Verfolgung des Fortschritts, um zeitnah einen neu geschaffenen Standard umsetzen oder dessen Einhaltung fordern zu können. Zumindest die Eckdaten von Standardisierungsprojekten – wie z. B. Zweck des Standards, Zeitrahmen, Beteiligte, Status und Kontaktmöglichkeiten – sollten öffentlich und gut auffindbar sein. Hilfreich sind auch Beschlussprotokolle, aus denen ersichtlich ist, unter welchen Annahmen Entscheidungen getroffen wurden und inwieweit dabei jeweils Vorteilsbetrachtungen eine Rolle spielten. Aktive Beteiligungsmöglichkeiten für öffentliche Stellen und Dritte sowie das öffentliche Zugänglichmachen von Entwurfsfassungen sollten der Normalfall sein.



Kontakt

Gabriele Goldacker
Kompetenzzentrum Öffentliche IT (ÖFIT)
Tel.: +49 30 3463-7173
info@oeffentliche-it.de

Fraunhofer-Institut für
Offene Kommunikationssysteme FOKUS
Kaiserin-Augusta-Allee 31
10589 Berlin

www.fokus.fraunhofer.de
www.oeffentliche-it.de
X: @OeffentlicheIT

ISBN: 978-3-948582-22-7

3I24