DATA DRIVEN GOVERNMENT – WIE DER STAAT ZUM VORREITER WIRD

Capgeminis Ansatz zu den Rahmenbedingungen von Data Driven Government

INHALTSVERZEICHNIS

1. Einleitung **04**

2. Rahmenbedingungen von DDG **07**

2.1 Gesellschaft und Stakeholder **08**

2.2 Fachliche Herausforderungen als Treiber für DDG **13**

2.3 Technologie und Architektur **17**

2.4 Rechtlicher Rahmen und Gesetzeslage **19**

2.5 Governance und Personal **19**

3. Ausblick und Fazit **24**

EXECUTIVE SUMMARY

Data Driven Government beschreibt eine datengestützte Verwaltung – ein Thema, welches den deutschen Staat intensiv beschäftigt und in Zukunft eine entscheidende Rolle bei der Weiterentwicklung der Verwaltung spielt. Capgemini positioniert sich hier mit weitreichender Expertise:

In der Wirtschaft sind datengetriebene Geschäftsmodelle seit Jahren Garant für Wachstum und hohe Börsenbewertungen. Auch Regierungen haben mittlerweile verstanden, dass Daten der Schlüssel zum Erfolg von Nationen und Gesellschaftsmodellen sein können und auf europäischer und nationaler Ebener entsprechende Datenstrategien verabschiedet. Zumindest in unserer europäischen Werteordnung sollten aber die Logiken und Regeln der Datennutzung immer die Balance zwischen den Chancen und den Risiken halten und die Datensouveränität der Bürgerinnen und Bürger wahren. Initiativen wie Gaia-X versuchen diese Balance herzustellen, und während Gaia-X meist als Cloud-Initiative zitiert wird, so ist es die Datennutzung für Analysen und KI, die nach europäischen Werten und Regeln gefördert werden soll und ein Gleichgewicht zu den Datenschätzen der GAFAM bzw. des chinesischen Staatsapparats herstellen soll.“

Marc Reinhardt
Leitung Global Public Sector

Von staatlicher Seite verifizierte, aktuelle Daten sind der Schlüssel für ein schnelleres und besseresVerwaltungshandeln der Zukunft. Die Daten helfen, in Krisensituationen schnell zu reagieren und langsam ablaufende, gesellschaftliche Trends zu identifizieren. Data Driven Government erlaubt, in staatlichem Handeln von Vermutungen zu bestätigtem Wissen über komplexe, gesellschaftliche Wirkzusammenhänge zu kommen.”

Frank Jacobsen
Leitung Public Sector Deutschland

Die Ampelkoalition hat die Veröffentlichung und Nutzung von Daten im Koalitionsvertrag verankert. Neben dem großen Innovationspotential für Wirtschaft und Gesellschaft ermöglichen hochwertige Daten auch für Politik und Verwaltung ein enormes Potenzial und bereits jetzt vielzählige Lösungen, um das Verwaltungshandeln effektiver zu gestalten. Durch das hohe Innovationstempo werden in den nächsten Jahren noch zahlreiche weitere Lösungen hinzukommen.”

Carlos Ferrero Calle
Leitung Strategische Initiativen Public Sector

Data Driven Government kann die Handlungsspielräume für politische und administrative Entscheidungsträger:innen erweitern und unterstützt damit Entscheider in der operativen, taktischen und strategischen Entscheidungsfindung. . Über Steuerungs -und Informationsdashboards kann zusätzlich die notwendige Transparenz für Politik, Industrie sowie die Bevölkerung im Kontext “ geschaffen werden.”

Peter Steinleitner
Leitung Data Driven Government Initiative

1 EINLEITUNG

Stellen Sie sich eine Welt vor, in der politisch-gesellschaftliche Entscheidungen unter Zuhilfenahme von vorliegenden Daten getroffen werden; eine Welt, in der Daten geteilt, gemeinsam erhoben und verbessert werden. Entscheidungen werden so transparenter und die Teilhabe der Bürger:innen leichter. Die Idee, dass der Staat Informationen und Wissen nutzen sollte, um auf zukünftige Ereignisse besser vorbereit zu sein, hatte bereits der Staatsphilosoph Machiavelli. Die unzähligen Daten und Informationen in Kombination mit der Digitalisierung bieten dem Staat die Möglichkeit, eine durch Daten unterstützte Basis für Entscheidungen aufzubauen.

Corona-Pandemie und Extremwetterereignisse zeigen, wie wichtig die Analyse, Vorhersage und Bereitstellung von Daten ist. In künftigen Krisensituationen, wie Flutkatastrophen, Waldbränden, Dürren oder medizinischen Notständen, muss die Reaktion unseres politisch-administrativen Systems schneller, präziser und nachvollziehbarer werden.

Das politisch-administrative System in Deutschland hat die Bedeutung und Potenziale der Nutzung von Daten erkannt. Im Jahr 2021 hat die Bundesregierung die Datenstrategie und die Open-Data-Strategie veröffentlicht. Beide Strategien heben die Bedeutung von Daten hervor und benennen viele Maßnahmen, die eine intensive Nutzung von Daten in der näheren Zukunft vorbereiten und anstoßen sollen. Der Grundstein für eine datenbasierte Entscheidungsfindung auf Bundesebene wurde damit gelegt. Auch während und nach der Bundestagswahl 2021 war die Digitalisierung der Verwaltung ein Thema aller Parteien, die den Einzug in den Bundestag geschafft haben. Darauf aufbauend zeigt der Koalitionsvertrag der neuen Bundesregierung, dass die Digitalisierung der Verwaltung und die vielfältigere Nutzung von Daten in der nächsten Legislaturperiode von hoher Bedeutung sein wird (Rechtsanspruch auf Open Data, Datengesetz, Dateninstitut und weitere Punkte). Neben den Initiativen auf Bundesebene lassen sich auch auf Landes- sowie Kommunalebene Vorhaben im Bereich Daten-Aufbereitung und -Nutzung identifizieren.

Die intelligente Datennutzung kann eine enorme Hilfe bei der Lösung gesellschaftlicher Probleme oder beim Krisenmanagement sein. Mit neuen datenbasierten Technologien sind Herausforderungen besser zu begreifen und zu bewältigen. Auf der Grundlage von Daten lassen sich Unsicherheiten und Komplexität reduzieren, sowie zielgenaue Lösungen für zentrale gesellschaftliche Herausforderungen von heute und morgen finden.

Politisch-administrative Entscheidungsträger:innen stehen vor der Herausforderung, Lösungen für kurz- mittel- und langfristige Probleme zu finden. Sie müssen aber auch vorhandene, täglich anfallende Daten als Ausgangspunkt für den politischen Austausch über Ziele nutzen und grundsätzlich neue administrative Handlungsoptionen erschließen. Dies gilt für alle, die für Bund, Land und Kommunen Entscheidungen tragen. Wie kann der Staat hier seiner Vorbildfunktion gerecht werden?

Data-Driven-Government-Ansatz

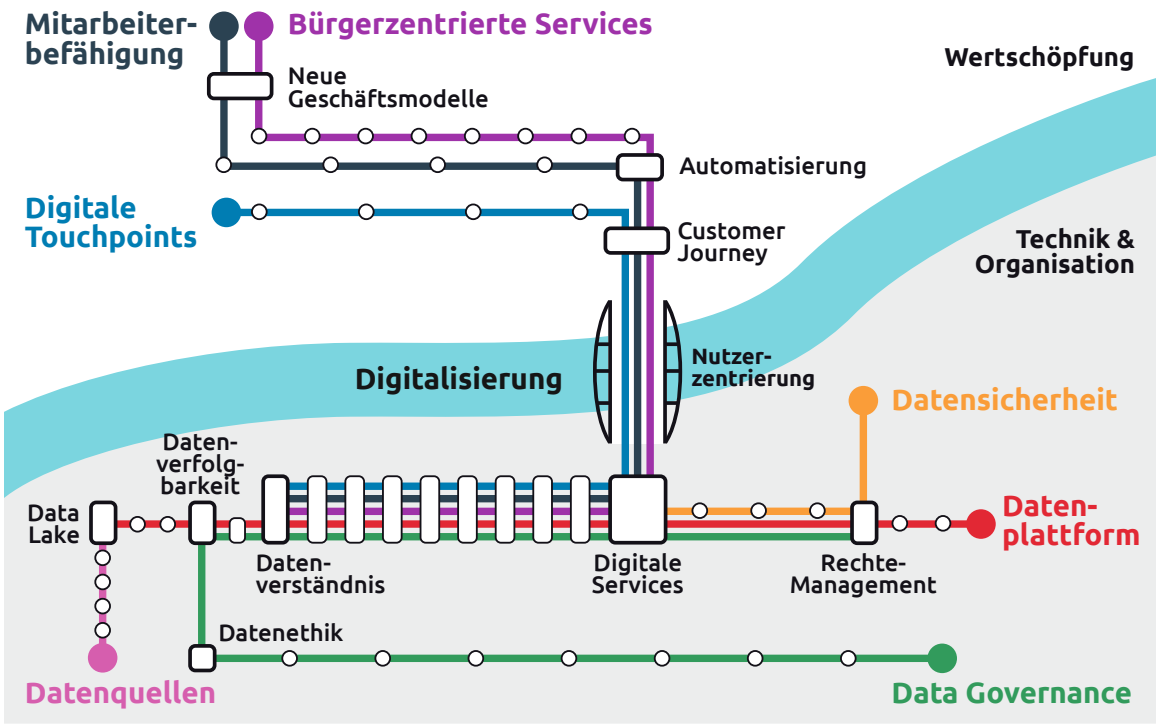
Um ebendiese genannten Herausforderungen in Zukunft noch besser bewältigen zu können, spielt ein daten- und evidenzbasiertes Verwaltungshandeln eine entscheidende Rolle. Der Begriff Data Driven Government (DDG) bedeutet eine datenbasierte Entscheidungsfindung auf Grundlage von Datenanalysen in der Verwaltung und politisch-administrative Führung. Eine digitale Verwaltung von morgen wird nicht nur von der Einführung neuer Technologien und Vereinfachung von Prozessen geprägt, sondern auch von der intelligenten Nutzung der entstehenden und vorhandenen Daten. Der gesellschaftliche Mehrwert: Daten ermöglichen Transparenz und die Teilhabe von Bürger:innen; ein öffentlicher Diskurs über Herausforderungen der Zukunft wird möglich.

„Data Driven" bedeutet hier, vorhandene Daten und Informationen zu erkennen, bereitzustellen, nutzbar zu machen und auszuschöpfen, um für eine bestimmte Aufgabe verwertbare Erkenntnisse zu gewinnen. Im Sinne von DDG müssen relevante Daten und Informationen für die Entscheidungsträger:innen zu deren Unterstützung bei Bedarf bereitstehen. DDG geht damit über einzelne Digitalisierungsprojekte hinaus und nutzt Daten, Infrastruktur, Prozesse und Anliegen der verschiedenen Interessensgruppen, um datengestützt politische und administrative Gestaltungen vorzunehmen.
„Data Driven" ist dabei nicht als „von Daten getrieben" zu verstehen. Denn die Konsequenz eines solchen Verständnisses wäre, dass die Daten einen Entscheider vor sich hertreiben, Entscheidungen durch Algorithmen vorgeben oder gar Entscheidungen aufzwingen. Vielmehr soll der Ansatz die Handlungsspielräume für politische und administrative Entscheidungsträger:innen mit Fakten anreichern und erweitern - nicht vermindern. DDG bezieht sich folglich auf die unterstützende Funktion der Daten im Entscheidungsfindungsprozess und in der strategischen Steuerung.

Im sozio-technologischem System eines digitalen Staates bilden unter anderem Daten, wie zum Beispiel amtliche Daten, Forschungsdaten, Wirtschaftsdaten, Infrastrukturdaten oder Geodaten, die Basis für Digitalisierungsansätze. Bei der Digitalisierung der Verwaltung spielen Daten eine zentrale Rolle in Projekten, wie Digitalisierung der Aktenführung, Prozessoptimierung oder E-Government. Aber auch die Digitalisierung des Lebensraumes der Bürger:innen basiert auf Daten. Lösungen für den Infrastrukturausbau, die Daseinsvorsorge, die Steigerung der Lebensqualität oder Smart Cities lassen sich nur unter Einbeziehung verschiedener Daten umsetzen. Um politisch im Sinne von DDG zu gestalten, müssen die Voraussetzungen eines digitalen Staats im Bereich Daten, Infrastruktur und Prozesse auf- und ausgebaut werden.

Bei der Betrachtung von DDG wird schnell bewusst, dass es je nach Reifegrad und Herausforderung einer Organisation verschiedene Einstiegsmöglichkeiten gibt. Wir haben diese identifiziert und im Capgemini Netzplan für Data Driven Government veranschaulicht (Abb. 1).

Abbildung 1: Capgemini Netzplan für Data Driven Government



Die Einstiegsmöglichkeiten sind alle miteinander verknüpft und ergeben ein ganzheitliches Bild im Sinne von DDG. Einige Behörden oder Organisationen werden sich zunächst mit potentiellen Datenquellen befassen, andere wiederum wählen den Einstieg über den Aufbau von Datenkompetenzen innerhalb bestimmter Organisationseinheiten. Unabhängig des Einstiegs wird deutlich, dass eine Betrachtung der verschiedenen verknüpften Themenfelder essenziell ist, wenn man ein vollumfängliches Verständnis und eine Ende-zu-Ende-Umsetzung in Richtung Data Driven Government erreichen möchte.

Weiterhin stellt sich jedoch auch die Frage, welche Faktoren die Einstiegsmöglichkeiten beeinflussen und welche Rahmenbedingungen gegeben sind. Wenn man einen Einstiegspunkt wählt, sollte man die jeweiligen Rahmenfelder in Betrachtung ziehen, um ein vollumfängliches Verständnis und eine Ende-zu-Ende- Umsetzung in Richtung Data Driven Government zu erreichen.

Ziel dieses Whitepapers ist es, in diese Rahmenbedingungen einzuführen und einen Einblick die relevanten Felder zu geben.

2 RAHMENBEDINGUNGEN VON DDG

Die Einführung und Definition zu Data Driven Government hat dargestellt, dass es sich hier um ein weitreichendes Konstrukt handelt, welches Potenziale für eine bessere datengestützte Entscheidungsfindung bietet. Eine sich stetig verändernde Gesellschaft, unterschiedlichste Bedürfnisse von Stakeholdern und eine rasante Technologieentwicklung sind nur einige der Umfeldfaktoren, welche auf dieses Konstrukt Einfluss nehmen. Capgemini identifiziert hier fünf übergreifende Rahmenbedingungen, dargestellt in Abbildung 2. In den folgenden Kapiteln werden die einzelnen Rahmenbedingungen näher beleuchtet und anhand verschiedener Beispiele dargelegt. Zusätzlich wird aufgezeigt wie Capgemini innerhalb dieser verschiedenen Bereichen unterstützen kann und Expertise aufgebaut hat. Ziel ist, ein Grundverständnis zu schaffen, um am Weg in Richtung DDG Leitplanken zur Orientierung zu stellen.

Abbildung 2: Rahmenbedingungen von DDG

2.1 GESELLSCHAFT UND STAKEHOLDER

Unsere Gesellschaft wandelt sich stetig; komplexe, weitgreifende Herausforderungen gilt es zu bewältigen; Trends und Entwicklungen wie Digitalisierung, Globalisierung, demografischer Wandel und Klimawandel fungieren als Treiber. Diese Veränderungen beeinflussen das Leben der Gesellschaft. So verändern Digitalisierung und Klimaschutz unsere Art zu arbeiten und uns fortzubewegen. Insbesondere eine Extremsituation wie die Corona-Pandemie hat gezeigt, wie schnell sich ein Virus in einer globalisierten, vernetzten Welt verbreiten kann und Einfluss auf das Gesundheitssystem, die Arbeit und das gesellschaftliche Leben nimmt. Diese, teilweise globalen, Entwicklungen machen die Entscheidungsfindungen hochgradig komplex. In der gleichen Zeit steigt die globale Datenmenge täglich. Die erhobenen und punktuell bereits bereitgestellten Daten bieten die Chance, in einer immer komplexer werdenden Welt Überblick zu schaffen, sich auszutauschen, und evidenzbasierte Entscheidungen zu treffen.

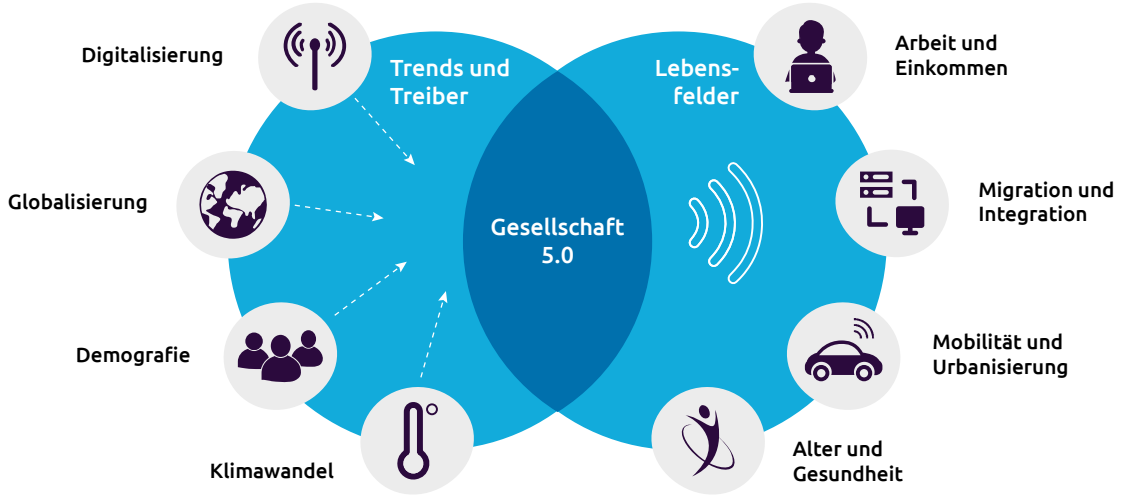
Abbildung 3 skizziert eine Auswahl dieser Herausforderungen unter dem Begriff Gesellschaft 5.0.

Die Gesellschaft 5.0 steht für eine nachhaltige, intelligente und vernetzte Gesellschaft, welche durch die Nutzung von modernen Technologien eine Lebensqualitätssteigerung und die Verbesserung gesellschaftlicher Probleme erreichen möchte. Die vier Einflussfaktoren Digitalisierung, Globalisierung, Demografie und Klimawandel wirken als Trend und Treiber auf die Lebensfelder der Gesellschaft 5.0.

Die digitale Transformation bietet viele Ansatzpunkte und Möglichkeiten, die vielschichtigen Herausforderungen zu bewältigen; zugleich stellt sie aufgrund der vielen gesellschaftlichen Veränderungen die öffentliche Verwaltung als Steuerungsakteurin selbst vor neue Herausforderungen:

- Anzahl und Komplexität von Aufgaben der Verwaltung steigen rasant und damit auch die Geschwindigkeit, mit der sie bearbeitet werden müssen. Zeitgleich stehen diese Faktoren im Ungleichgewicht zu der Fähigkeit, die Strukturen und Organisation der Verwaltung an die neuen Anforderungen anzupassen.
- Die durch den demographischen Wandel ausgelösten Lücken in Wissen und Erfahrung lassen sich mit heutigem Vorgehen nicht allein durch bessere Fachverfahren oder technische Unterstützung schließen.
- Die Geschwindigkeit, mit der die Digitalisierung in privaten Lebensfeldern und Unternehmen voranschreitet, vergrößert den Abstand zwischen Verwaltung und Bürger:innen. Dies erhöht den Handlungsdruck auf die Politik.
- Der enorme Nachholbedarf in elementaren Digitalisierungsaufgaben (vgl. Umsetzung des Online-Zugangsgesetzes) hindert die Verwaltung daran, die großen Chancen, die in den verfügbaren Daten liegen, zum Wohle der Gesellschaft zu nutzen.

Abbildung 3: Die sich verändernde Gesellschaft und ihre aktuellen Herausforderungen



Die öffentliche Verwaltung muss sich zu Nutze machen, dass sie auf riesige Daten- und Informationsmengen zugreifen kann, um die gesellschaftlichen und ihre eigenen Herausforderungen zu bewältigen.

Für Entscheidungsträger:innen können Daten beispielsweise in Form von Dashboards aufbereitet werden. Warum sollte eine Bürgermeisterin ohne ausreichende Faktengrundlage über die in ihrer Stadt zu tätigenden Investitionen entscheiden? Ein Datencockpit kann ihr genau zeigen, welche Investitionen heute sinnvoll sind und extrapolieren, welche Auswirkungen Investitionen in der Zukunft haben, wenn keine Steuerungs- und Gestaltungsmaßnahmen ergriffen werden.

Mit der stetig besseren Verfügbarkeit kontinuierlich wachsender Datenmengen und den erweiterten Auswertungsmöglichkeiten setzen weltweit immer mehr Regierungen und Verwaltungen verstärkt auf datengestützte Entscheidungsprozesse. Mit Daten entstehen für die politisch-administrativen Entscheidungsträger:innen erweiterte Handlungsoptionen und neue Instrumente strategischer Steuerung. Die Daten lassen sich dabei vielfältig einsetzen: als Quelle neuer Erkenntnisse, als Unterstützung bei strategischen Entscheidungen, als Hilfe bei Fehleridentifikationen und Prozessoptimierungen sowie als Grundlage für Vorhersagen und Analysen. Dies bietet viele Chancen und Möglichkeiten:

- Komplexität ist reduzierbar.
- Verantwortliche werden bei Entscheidungen jeglicher Art unterstützt.
- Prozesse lassen sich vereinfachen.
- Herausforderungen verschiedener Dimensionen lassen sich bewältigen.
- Zukunft wird gestaltbarer.

Diese vielfältigen Datenanwendungen erfordern ein Datenökosystem, welches den Austausch und die Weiterverarbeitung von Daten zwischen Politik, Verwaltung, Wirtschaft und Gesellschaft ermöglicht und unterstützt.

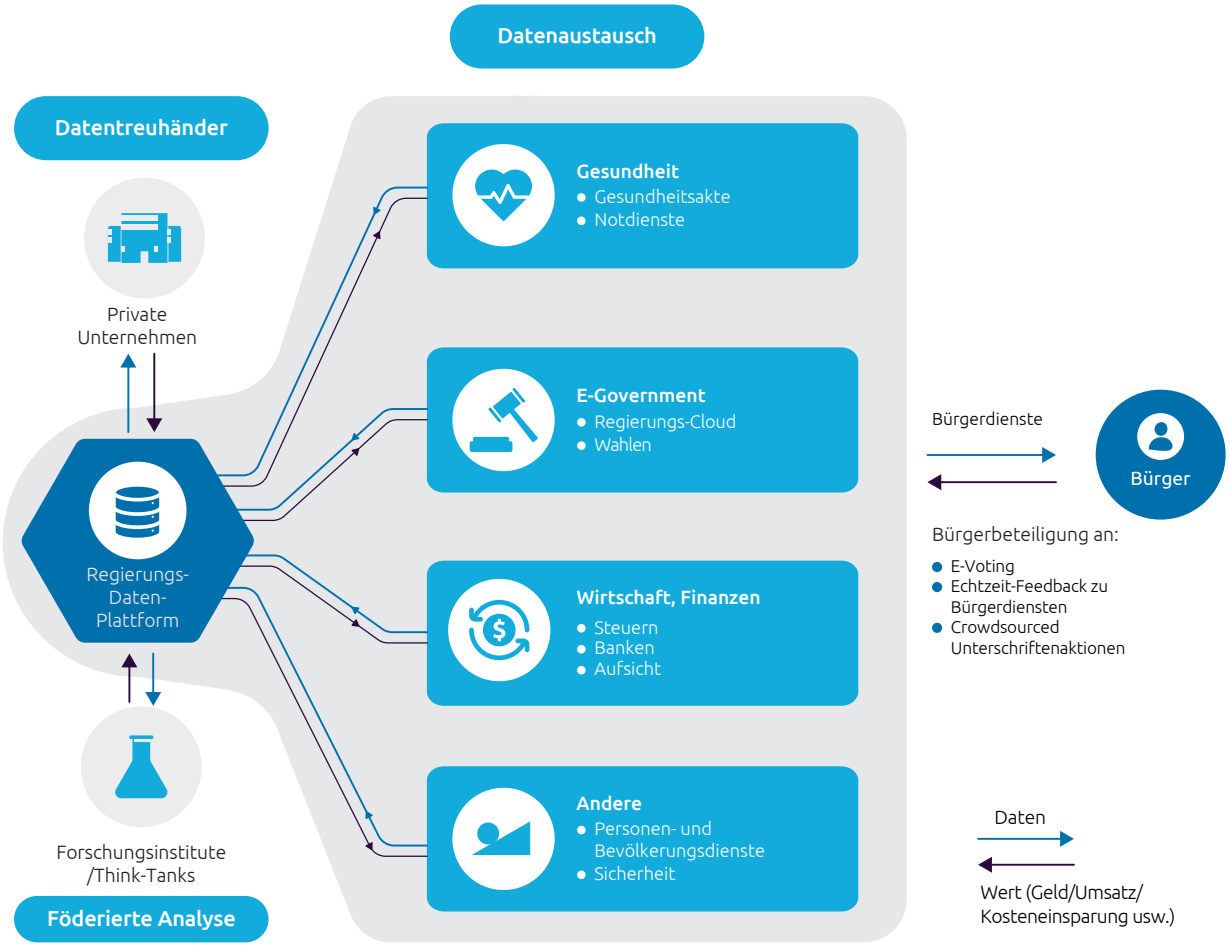
Datenökosysteme

Immer mehr Organisationen und Unternehmen wollen Daten von Dritten in ihre Analysen einbeziehen, um höherwertige Erkenntnisse zu gewinnen und Entscheidungen treffen zu können. Die Verwendung vieler Quellen erhöht die Diversität und Vollständigkeit von Informationen. Jedoch haben zahlreiche Organisationen dabei einige Hindernisse erfahren, beginnend mit dem (fehlenden) Vertrauen in externe Daten: Wie ist die Qualität des Datenlieferanten? Welche Bedeutung haben die einzelnen Datenfelder genau? Außerdem spielen vertragliche Aspekte der Datenbereitstellung eine Rolle.

All diese Punkte adressiert das Konzept Datenökosystem. Organisationen bzw. Unternehmen nehmen an einem Datenökosystem teil, indem sie Daten über eine gemeinsame Plattform bereitstellen oder konsumieren. Der Datenaustausch ist darauf ausgerichtet, gegenseitigen Mehrwert zu erzielen. Datenökosysteme sind häufig branchenübergreifend.

Eine aktuelle Studie des Capgemini Research Institute beleuchtet den aktuellen Stand und die weitere Entwicklung von Datenökosystemen (Abb. 4). Die Studie zeigt, dass Datenökosysteme eine effizientere Verbreitung der Dienstleistungen der öffentlichen Hand ermöglichen. Behörden sammeln unzählige Informationen aus der Privatwirtschaft. Hiervon können verschiedenste Organisationen, wie beispielsweise aus der Forschung signifikant profitieren. Daher gelten Behörden als wesentliche Akteure vieler Datenökosysteme.

Abbildung 4: Datenökosysteme ermöglichen eine effizientere Verbreitung öffentlicher Dienste[1]



Der nahtlose Datenaustausch zwischen öffentlichen Einrichtungen und Städten kann helfen, den Übergang zu „Smart Cities" zu vollziehen. Er gibt Bürger:innen die Möglichkeit, sich an der Entwicklung ihrer Städte zu beteiligen. Beteiligte Privatunternehmen können über diese Ökosysteme ihre Dienstleistungen für die Bevölkerung anbieten und erweitern.

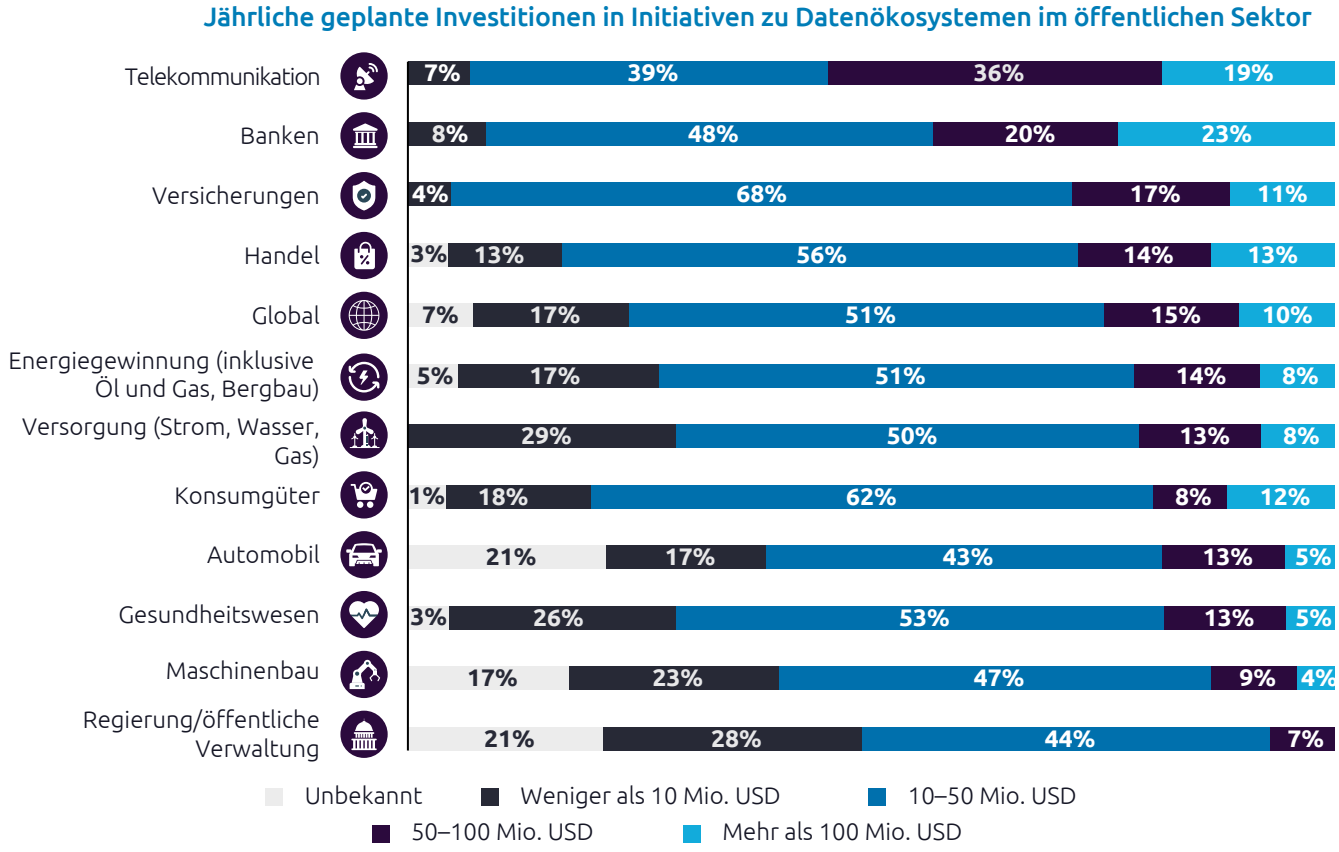
Öffentliche Dienste können in Bezug auf Governance, Sicherheit, Geschäft und Finanzen an eine Datenaustauschplattform angebunden werden und so Daten gemeinsam nutzen und den Menschen E-Services nahtlos bereitstellen. Bürger:innen haben die Möglichkeit, darüber ihre Meinung zu äußern, Fragen und Anliegen jederzeit und von überall einzubringen oder mitzuteilen, dass ein Dienst ausgefallen ist. Hinweise aus dem Crowdsourcing[2] bringen Transparenz und decken Probleme auf, die einen Großteil der Bevölkerung betreffen.

Datenpartnerschaften, heißt ein organisierter offener Datenaustausch zwischen mehreren Parteien, beschleunigen die Mobilität der Bevölkerung; sie beziehen alle Arten von Verkehrsmitteln ein – vom Bus bis zum E-Scooter. So lassen sich die besten Mobilitätsdienste anbieten.

In dem Capgemini-Report „Data Sharing Masters" wird das Beispiel Estland aufgegriffen: Estland hat e-Estonia implementiert und bietet somit 99 Prozent der öffentlichen Dienste online mit 24/7-Verfügbarkeit an. Die Bausteine dieses Ökosystems umfassen E-Identity, Gesundheitswesen, Sicherheit und E-Governance. Dieser Datenaustausch stiftet großen Nutzen: Estland spart jährlich 844 Jahre Arbeitszeit und Kosten in Höhe von zwei Prozent des Bruttosozialproduktes. Mehr als 99 Prozent der Banktransaktionen und Steuererklärungen werden online ausgeführt.

[1] Capgemini Research Institute: "Data Sharing Masters" https://www.capgemini.com/research/data-sharing-masters/ (Abruf am 15.02.2022)
[2] Crowdsourcing: Aufteilung von Aufgaben und Prozessen und Einholen von Ideen oder Informationen von einer undefinierten Masse von Nutzern

Abbildung 5: Capgemini Research: Geplante Investitionen in den einzelnen Sektoren in Datenökosysteme[3]



Der Report legt allerdings auch dar, dass Datenökosysteme im öffentlichen Sektor, im Vergleich zum privaten Sektor, wenig etabliert sind. In Abbildung 5 wird deutlich, dass dem öffentlichen Sektor der Nachholbedarf bewusst ist und Investitionen im zweistelligen Millionen-Bereich plant, um u.a. auch die Erwartungen der Privatwirtschaft und Bürger:innen zu erfüllen

Aus dem Report geht ebenfalls hervor, dass jedes zehnte Unternehmen der Regierung/des öffentlichen Sektors plant, in den nächsten drei Jahren mehr als 50 Millionen US-Dollar in Initiativen zum Datenökosystem zu investieren.

[3] Capgemini Research Institute: "Data Sharing Masters" https://www.capgemini.com/research/data-sharing-masters/ (Abruf am 15.02.2022)

Um Datenökosysteme zu etablieren und zu verbessern, ist es hilfreich, die Aufgaben und Beziehungen zwischen den Akteuren im Datenökosystem zu verstehen:

- **Datentreuhänder:** Vermittler:innen sammeln und aggregieren Daten verschiedener Datenlieferanten, die sie dann in aggregierter Form an Daten-Nutzer:innen bereitstellen. Das ist zum Beispiel die typische Arbeitsweise von Marktforschungsinstituten.
- **Informationskette:** Informationen werden von einem Unternehmen zum nächsten bereitgestellt, von diesem
- weiter aufgewertet und angereichert und letztlich den Informationsnutzer:innen angeboten.
- **Föderierte Analyse:** Dieses Vorgehen adressiert regulatorische Hindernisse bei der Weitergabe sensibler Daten (z. B. personenbezogene Daten). Hierbei sammelt ein vertrauenswürdiges Unternehmen die Daten von den Lieferanten ein und analysiert sie. Nutzer:innen erhalten nur die Erkenntnisse dieser Analysen – nicht die zugrundeliegenden Daten.
- **Kollaborative Datenlieferketten:** Hier stellen unterschiedliche Datenlieferanten wie Bürger:innen oder die Kundschaft den Organisationen Daten bereit. Diese sammeln und reichern die Daten an, um anschließend Analysen und Erkenntnisgewinne zu generieren. Darauf aufbauend können den Datenlieferanten verbesserte oder neue Services angeboten werden.

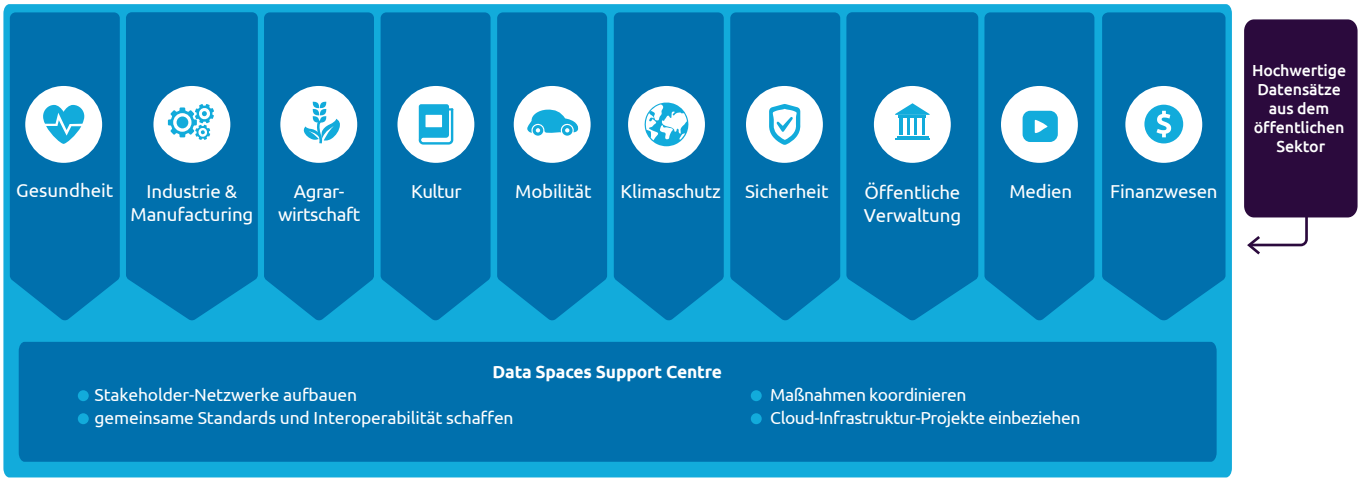
Datenräume:
Die Datenstrategie der EU-Kommission Februar 2020[4] skizziert europäische Datenräume (Abb. 6). Das sind fachlich abgegrenzte Datenökosysteme wie z. B. der Datenraum für Mobilität, in dem sich Autohersteller, -lieferanten, Mobilitätsdienstleister u. a. Stakeholder austauschen.

Innerhalb eines abgeschlossenen Datenraums wird das sichere, souveräne Teilen von Daten ermöglicht. Somit werden Teilnehmende in die Lage versetzt, grenzübergreifend europaweit Daten zu teilen und zu konsumieren; gleichzeitig bewahren sie die eigene Datensouveränität. Um weitere Daten aus anderen Bereichen nutzen zu können, wird ein Austausch mit anderen Datenräumen ermöglicht. Gegenwärtig wird an Datenräumen für den öffentlichen Sektor gearbeitet.

Abbildung 6 verdeutlicht, wie verschiedene Akteure in einem Datenraum partizipieren und Daten sowie Erkenntnisse austauschen. Die Teilnahme an Datenräumen ist nicht nur Organisationen oder Unternehmen vorbehalten. Für die Betrachtung verschiedener Aspekte und Datenquellen in einem Datenraum ist auch eine Beteiligung der Bürger:innen möglich. Ein Datenraum im öffentlichen Sektor kann beispielsweise für bürgernahe Dienste Mehrwerte schaffen. Beispiele sind E-Voting, Feedback der Bürger:innen in Echtzeit und Petitionen.

Datenökosysteme bieten vielfältige Möglichkeiten, Daten bereitzustellen und auszutauschen. Dabei bestehen hohe Anforderungen an die Daten hinsichtlich Vertrauen und Qualität. Sind diese gewährleistet, können von Behörden betriebene Datenökosysteme die Qualität und Effizienz staatlicher Dienstleistungen verbessern. Die konkreten Anforderungen an die Veröffentlichung von Daten der öffentlichen Verwaltung behandelt der Abschnitt zu Open Data.

Abbildung 6: Gemeinsame europäische Datenräume – Vorschlag aus der Europäischen Datenstrategie[5]



[4] European Commision: "European State Strategy" (19.02.2020) https://ec.europa.eu/commission/presscorner/detail/en/fs_20_283 (Abruf am 15.02.2022)
[5] Capgemini Research Institute: "Data Sharing Masters" https://www.capgemini.com/research/data-sharing-masters/ (Abruf am 15.02.2022)

2.2 FACHLICHE HERAUSFORDERUNGEN ALS TREIBER FÜR DDG

Eine weitere Rahmenbedingung ist die fachliche Orientierung der jeweiligen Behörde und Organisation. Beim Bearbeiten behördlicher Aufgaben und Erbringen von Leistungen ergeben sich kontinuierlich neue Herausforderungen. die es zu bewältigen gilt. Doch wie können Daten, deren Auswertung und Kombination eine bessere Basis zur Entscheidungs- und Lösungsfindung bilden? Das Entwickeln eines fachlichen Anwendungsfalls setzt zunächst ein grundlegendes Verständnis verschiedener Datenttypen voraus. Darauf aufbauend geht es an das Lösen der komplexen und fachlichen Herausforderungen durch einen datengestützten Ansatz. Um dabei Gesellschaft, Politik, Wissenschaft und Zivilgesellschaft einzubeziehen, können Behörden ihre im Rahmen der öffentlich-rechtlichen Aufgaben erhobenen Daten frei maschinenlesbar über Datenportale zur Verfügung stellen. Im Folgenden erläutern wir den Hintergrund und die Vorteile von Open Data, um anschließend einen von Capgemini aufgebauten Use Case vorzustellen – den Proof of Concept Waldvitalität.

Datenquelle Open Data

Ein umfassender Datenbestand ist die zwingend erforderliche Grundlage, damit die Umsetzung des DDG-Ansatzes funktioniert. Daten lassen sich in öffentlich zugängliche Daten (Open Data), gemeinsam verwendete Daten (Shared Data) und nicht öffentlich zugängliche Daten (Closed Data) unterteilen. Abbildung 7 veranschaulicht dieses Spektrum von Daten.

Open (Government) Data stehen in einer großen Vielfalt zur weiteren freien Nutzung zur Verfügung. Dazu zählen statistische Daten, Wetterdaten und Satellitendaten; sie lassen sich beispielweise über das Datenportal für Deutschland (govdata.de) finden und nutzen.

Abbildung 7: Datenspektrum[6]

[6] Open Data Institute London: „The Data Spectrum" https://theodi.org/about-the-odi/the-data-spectrum/ (Abruf am 15.02.2022)

Open Government Data[7] sind

- von Behörden oder im Auftrag Dritter erhobene Daten,
- frei über öffentlich zugängliche Netze verfügbar,
- von Maschinen lesbar,
- über freie Lizenzen frei verwendbar, d. h. die Rechte Dritter werden nicht verletzt,
- ohne Personenbezug und
- frei von sicherheitsrelevanten Informationen.

Durch diese Grundsätze von offenen Daten sind sie für jedermann gleich zugänglich und wiederverwendbar. So können durch die bereitgestellten offenen Daten einfache Analysen und Anwendungen entstehen, Entscheidungen können nachvollzogen und neues Wissen generiert werden.

Die Open-Data-Gesetzgebung auf Bundes- und Landesebene sowie weitere veröffentlichte Open-Data-Strategien werden die Veröffentlichung und Verfügbarkeit von Open Data weiter vorantreiben. Bisher veröffentlicht die Verwaltung die eigenen Daten jedoch eher zurückhaltend. Die Open-Data-Landschaft in Deutschland stellt sich in Quantität und Qualität der bereitgestellten Daten heterogen dar. Eine bloße rechtliche Verpflichtung der Bundes- und Landesbehörden hat bisher nicht zu einer umfassenden Datenbereitstellung geführt (vgl. 1. Open-Data-Fortschrittsbericht der Bundesregierung). Abzuwarten bleibt, welche Potenziale sich durch konkrete Maßnahmen zur Förderung von Open Data entfalten. Wie wird sich damit die Verfügbarkeit von Daten des öffentlichen Sektors, wie sie beispielsweise in der Datenstrategie sowie der Open-Data-Strategie der Bundesregierung erfasst sind, entwickeln?

Entwickelt die öffentliche Verwaltung die Bereitstellung von Verwaltungsdaten als Open Data umfassend weiter, schafft sie selbst eine solide (Daten-)Basis für den Einsatz von DDG. Die bereitgestellten Daten lassen sich beispielsweise dazu nutzen, die Herausforderungen des Klimawandels zu meistern. Wie es ganz konkret möglich ist, mit DDG die Folgen des Klimawandels zu bekämpfen, skizzieren wir nachfolgend am Beispiel Wald.

Capgeminis Proof of Concept Waldvitalität

Capgemini hat in der Vergangenheit verschiedene Proof of Concepts (PoC) und konkrete Lösungen erstellt, die die Vorteile der Datennutzung und-verwertung belegen, zum Beispiel:

- Plastikmüllerkennung an Küstendurch Satellitenbilder[8]
- Vorhersage der verfügbarer Betten in Krankenhäusern[9]
- Überwachung der Mojave-Wüste in Nevada zum Begrenzen von Schäden am Ökosystem[10]

An einem weiteren Proof of Concept zur Waldvitälität erläutern wir nun die verschiedenen Phasen von der Konzeption des fachlichen Use Cases und den fachlichen Anforderungen über Datenaufbereitung bis hin zu Analyse und KI und Visualisierung in einem Dashboard.

Beweggründe für diesen PoC gibt es genug: Das Waldsterben verändert seit einigen Jahren nicht nur das äußere Erscheinungsbild des Waldes, sondern auch das hochkomplexe und für die Menschheit wichtige Ökosystem Wald. Der Waldbericht der Bundesregierung[11] zeigt: Stürme, extreme Dürren und Borkenkäferbefall haben in deutschen Wäldern in den letzten Jahren immense Schäden verursacht. Dabei wurden in einigen Regionen wichtige Waldfunktionen, wie der Erhalt von Wasser- und Bodenschutz oder die Klimaschutzwirkung, erheblich beeinträchtigt. Um den Zustand und die Veränderungen von Wäldern frühzeitig zu erkennen, wurde ein Informationsdashboard für die öffentliche Hand und interessierte Personen entwickelt. Das Dashboard bildet den Gesundheitszustand von Wäldern flächendeckend ab und überwacht diesen. Dabei kommen Satellitenbilder und KI zum Einsatz.

[7] Da Open Government Data eine Unterkategorie von Open Data darstellen, die im Übrigen dieselben Eigenschaften teilen, wird im Folgenden der Begriff Open Data als Synonym für Open Government Data verwendet.
[8] The European Space Agency: "Implemented OSIP ideas — June 2020" https://www.esa.int/Enabling_Support/Preparing_for_the_Future/Discovery_and_PreparationImplemented_OSIP_ideas_June_2020 (Abruf am 16.02.2022)
[9] Capgemini: „AI4Health" https://www.capgemini.com/wp-content/uploads/2021/03/AI4Health_Brochure.pdf (Abruf am 16.02.2022)
[10] Capgemini: „Sustaining a national wonder with AI" https://www.capgemini.com/featured-article/sustaining-a-national-wonder-with-ai/ (Abruf am 04.03.2022)
[11] Bundesministerium für Ernährung und Landwirtschaft: "Waldbericht der Bundesregierung 2021" (14.07.2021) https://www.bmel.de/DE/themen/wald/wald-in-deutschland/waldbericht2021.html (Abruf am 04.03.2022)

Die Vorteile dieses Vorgehens: Satellitenbilder bieten deutlich umfassendere, detailreichere Bilder als herkömmliche Luftaufnahmen. Die gewonnenen Informationen und Satellitendaten sind bereits digitalisiert und somit einfacher zu verarbeiten.

Im Gegensatz zu einer manuellen Untersuchung von Satellitendaten, kann die entwickelte KI unter anderem Tausende von Bildern gleichzeitig automatisiert untersuchen und so Muster und Anomalien deutlich schneller erkennen. Unterschiedliche Indikatoren wie Biomasse oder Luftqualität lassen sich ebenfalls aus den Daten ableiten und im Informationsdashboard abbilden; und auch ein Vergleich der Indikatoren aus verschiedenen Zeiträumen ist möglich.

Dazu lässt sich der PoC in folgende vier Aufgaben gliedern:

- Boden- und Baumklassifizierung,
- Gesundheitszustand des Waldes,
- Biomasse des Waldes und
- Luftqualität.

Für die **Bodenklassifizierung** wurden Sentinel-2- und Corine-Land-Cover-Satellitenbilder ausgewertet. Corine Land Cover sind Daten über die Bodenklassifizierung von 2018 mit einer geringen räumlichen Auflösung. Der PoC hat das Ziel, die Bodenklassifizierung in guter zeitlicher Auflösung und räumlicher Genauigkeit anzuzeigen. Dazu wurde ein Klassifizierungsmodell mit Machine Learning (Random Forest) trainiert, indem historische Bodenklassifizierungsdaten von Corine Land Cover (als Labels) in Beziehung zu Sentinel-2-Daten (als Features) gesetzt wurden. Das ermöglicht bei Vorliegen von neuen Sentinel-2-Aufnahmen, auf eine aktuelle Bodenklassifizierung mit hoher räumlicher Auflösung zu schließen.

Das Dashboard (Abb. 8) zeigt dann an, zu wieviel Prozent beispielsweise das gekennzeichnete Gebiet aus Feld, Straße, Bäumen, Wiese usw. besteht. Bei der Baumklassifizierung werden die Flächen der Bäume nochmals in Laub- und Nadelbäume und in Mischwald aufgeteilt.

Für das Berechnungsmodell der **Gesundheits- bzw. Vegetationsindizes** des Waldes wurden anhand der Sentinel-2-Satellitenbilder bestimmte Spektralbänder in Relation gesetzt. Der PoC zeigt dafür die Indizes NDVI (Normalized Difference Vegetation Index), SAVI (Soil-Adjusted Vegetation Index) und NDWI (Normalized Difference Water Index). Diese Indizes sind die am häufigsten genutzten Indizes, um die Vitalität von Vegetation (bzw. das Chlorophyllgehalt) zu messen.

Die berechneten Indizes sind auch Basis für die Vorhersage der **Biomasse**; als Trainingsdaten werden sie gemeinsam mit Daten zum Holzvorrat (Biomassedaten) aus dem Gebiet in Lübeck in einen Machine-Learning-Algorithmus (Random Forest) gegeben. Um die Biomasse für das gekennzeichnete Gebiet zu erhalten, erhält die KI die Indizes (NDVI, SAVI, NDI45, SR, DVI) dieses Gebietes als Input und kann dann die Vorhersage durchführen. Das Dashboard (Abb. 9) präsentiert anschließend die Daten zur Biomasse, zum Kohlenstoffbestand und zum CO_2-Bestand jeweils oberirdisch und unterirdisch.

Den Bereich rund um die **Luftqualität** decken beispielsweise Daten zu Feinstaubpartikel der Größe $PM_{2,5}$ und PM_{10}, Kohlenstoffmonoxid und weitere ab. Hierfür erfolgte keine eigene Berechnung; die Daten wurden nur auf der openweathermap.org-API abgerufen.

Abbildung 8: Digital-Forest-Bodenklassifizierung

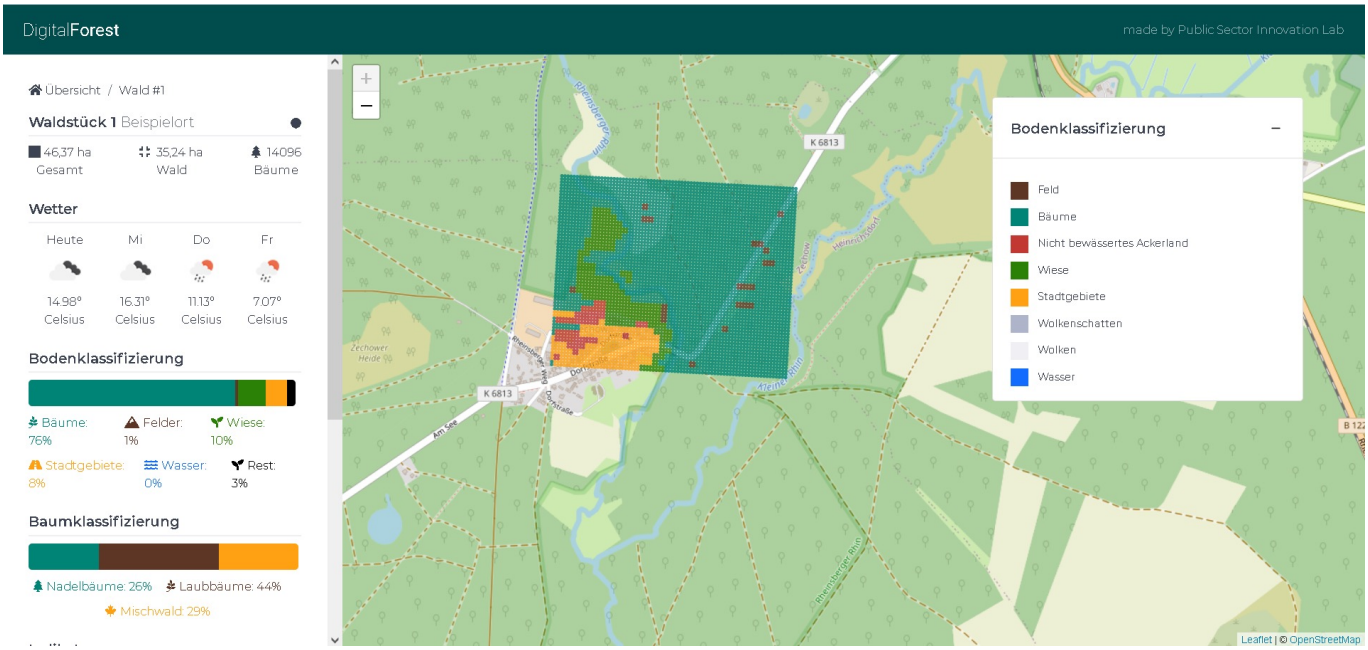
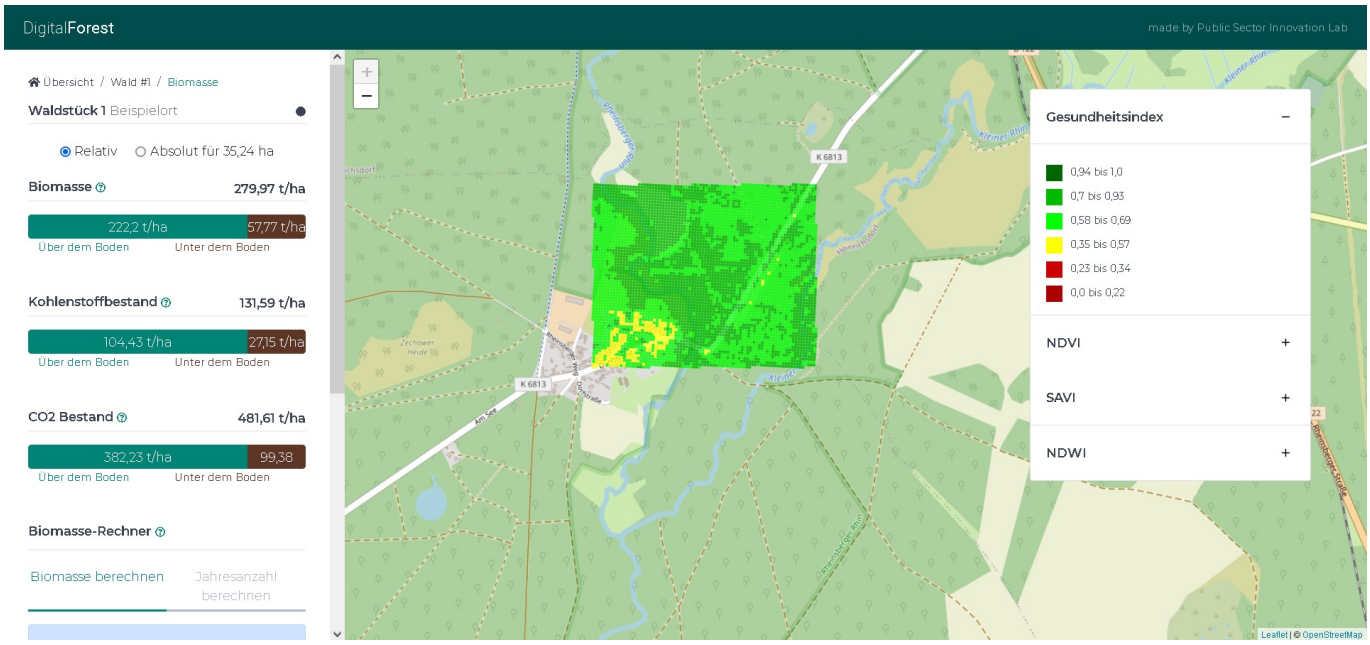


Abbildung 9: Digital-Forest-Gesundheitsindex



Das Beispiel POC Waldvitalität zeigt, dass Capgemini für unterschiedliche fachliche Anforderungen Dashboards erstellen kann. Dabei gehen wir auf die spezifischen Anforderungen der Behörden ein, identifizieren Probleme in der Datenaggregation und entwickeln ein auf die individuellen Bedürfnisse zugeschnittenes DDG-Dashboard.

Für so gut wie jedes Problem kann eine Datengrundlage geschaffen werden. Eine ausreichende Menge verfügbarer Daten ermöglicht sodann einen Entscheidungsfindungsprozess, der von hochwertigen Daten begleitet wird.

2.3 TECHNOLOGIE UND ARCHITEKTUR

Um datengestützt Entscheidungen treffen zu können, bedarf es nicht nur Daten oder fachliche Anwendungsfälle, sondern auch eine technologische Grundlage, auf der Daten verarbeitet und abgebildet werden können. Die rasante technologische Entwicklung bietet den Stakeholdern hier verschiedene Möglichkeiten, Datenplattformen aufzubauen. In diesem Kapitel erläutern wir, welche funktionalen Fähigkeiten eine Plattform mitbringen muss, auf der sich Data-Analytics-Anwendungen der nächsten Generation erstellen lassen.

Die Fortentwicklung bei modernen Datenanalyse-plattformen hat im Wesentlichen vier Treiber, die jeweils unterschiedliche Konzepte unterstützen:

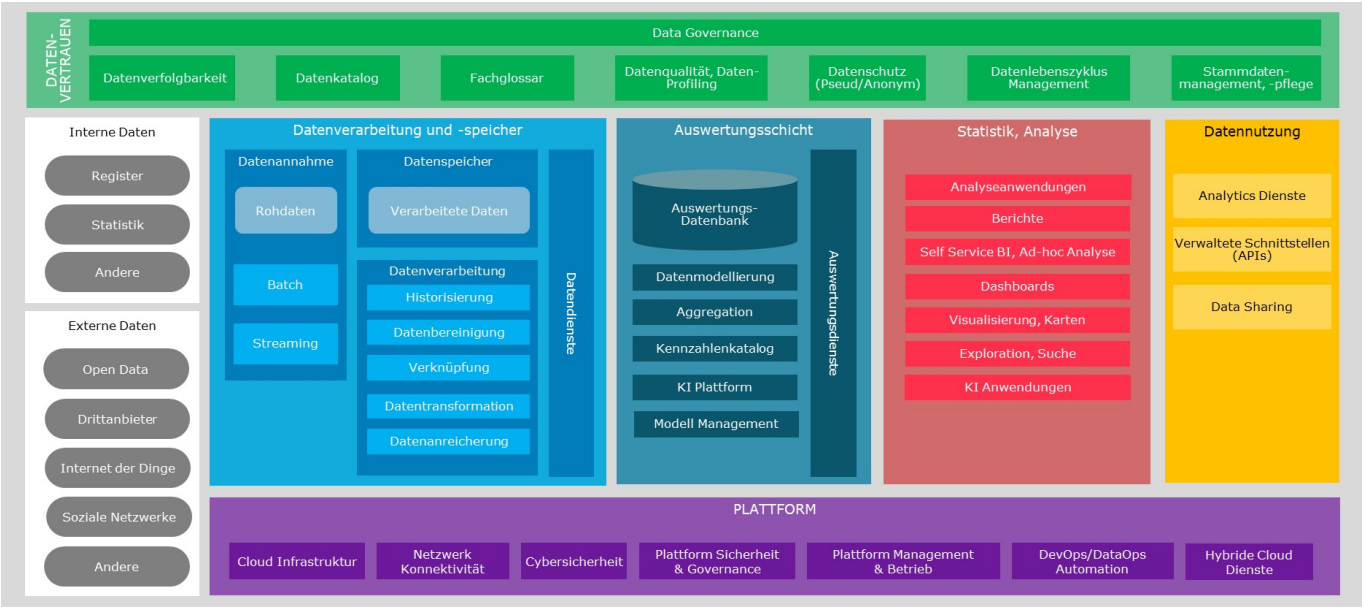
- **Agilität:** Cloud, Managed Services, CI/CD, DataOps, MLOps
- **Skalierbarkeit:** Big-Data-Konzepte, Cloud
- **Datenaustausch:** Managed APIs, Data Sharing
- **Regulatorische Compliance:** Verschlüsselung, Anonymisierung, Aggregation

Die Referenzarchitektur von Capgemini (Abb. 10) bietet eine moderne Datenanalyseplattform, welche diese vier Treiber integriert hat.

Referenzarchitektur für die moderne Datenanalyseplattform

Capgeminis Referenzarchitektur beschreibt die Fähigkeiten einer modernen Datenanalyseplattform. Sie ist herstellerneutral und lässt sich daher auf verschiedenen technologischen Plattformen on-premise oder in einer Cloud umsetzen.
Der Datenfluss verläuft auf dem Architekturbild von links nach rechts und beginnt mit den internen und externen Datenquellen. Die Daten werden aus den unterschiedlichen Quellen gewonnen, verarbeitet und gespeichert. Anschließend werden sie für Auswertungen vorbereitet, modelliert, aggregiert und in Kennzahlen für die Analyse angeboten. Fachanwender:innen ist es dadurch möglich, verschiedenartige datenbasierte Analysen und Statistiken durchzuführen. Die gefundenen Erkenntnisse werden dann zur Nutzung bereitgestellt. Dies ermöglicht faktenbasierte fachliche Entscheidungen und macht fortlaufende Geschäftsprozesse intelligenter.
Neben dem Datenfluss gibt es zwei übergreifende Fähigkeiten der Referenzarchitektur: Zum einen ist eine Plattforminfrastruktur erforderlich, auf der die Prozesse der Datenverarbeitung, -speicherung und Analyse ablaufen. Zum anderen sollte sie für die Speicherung und Verarbeitung dieser Vielzahl an Daten einen Cloud-Ansatz bieten, der einen schnellen, kostengünstigen und ortsunabhängigen Zugriff ermöglicht.

Abbildung 10: Capgeminis Referenzarchitektur für die moderne Datenanalyseplattform



Wege in die Cloud

Die Nutzung von Cloud-Lösungen ist in deutschen Behörden und Organisationen nur wenig verbreitet. Dabei haben Unternehmen anderer Branchen oder Länder bewiesen, dass die Mittel der Cloud die Datenanalyse auf eine neue Stufe heben. Im Wesentlichen gibt es drei Arten der Cloud, deren Vor- und Nachteile in Abbildung 11 dargestellt sind. Bei einer rein technischen Betrachtung der drei Cloud-Arten wäre die **Public Cloud** prädestiniert für die Verarbeitung und Speicherung von Daten. Der US Cloud Act verunsichert allerdings viele Interessierte. Die sogenannten Hyperscaler können durch den US Cloud Act zur Herausgabe von Daten gezwungen werden. Das hebelt das europäische Datenschutzrecht aus, obwohl die Daten in europäischen Datenzentren liegen. Die Hyperscaler versuchen der europäischen Kundschaft so weit wie möglich entgegenzukommen. Insbesondere Datensouveränität und das Misstrauen, sensible Daten an ausländische Serviceangebote zu geben, ist bei Verantwortung tragenden Personen der Verwaltung präsent. Der aktuelle Stand der Technik ist, dass Unternehmen ihre sensiblen Daten verschlüsseln und selbst den Schlüssel managen (sog. „Bring your own key").

Die **Private Cloud** ermöglicht es Anwendern, ihre Daten einem lokalen Betreiber mitzugeben und die volle Kontrolle zu behalten. In diesem Szenario muss der Betreiber seinerseits die Private Cloud vor Angriffen schützen. Das Thema Cybersecurity rückt demnach auch bei den Hyperscalern noch mehr in den Fokus. Zudem ist die Skalierbarkeit der Private Cloud ab einem gewissen Punkt begrenzt. Mehr Skalierbarkeit geht meist mit höheren Kosten einher. Die Virtualisierung der Infrastruktur ist in einer Private Cloud gut zu bewerkstelligen. Jedoch ist das funktionale Angebot zur Datenanalyse kaum vergleichbar mit dem, was die Hyperscaler samt Drittanbieter auf ihren Plattformen anbieten.

Die Data Acceleration Platform (DAP) von Capgemini bietet einen Weg, eine moderne Datenanalyseplattform in die Private Cloud zu integrieren. Hierbei wird eine neue Lösung mit Testdaten in der Public Cloud entwickelt. Dabei nutzen wir die Möglichkeiten der agilen Entwicklung für einen PoC wie der im vorherigen Kapitel beschriebene PoC Waldvitalität oder ein Minimal Viable Product (MVP) sowie die unkomplizierte Provisionierung von Infrastruktur und Skalierbarkeit der Public Cloud. Sobald die Lösung erfolgreich fertiggestellt ist, wird sie in die Private Cloud umgesiedelt und dort mit echten, sensiblen Daten produktiv betrieben. Der Umzug wird von verschiedenen Plattformanbietern unterstützt. Ist der PoC nicht erfolgreich, lässt er sich in der Infrastruktur der Public Cloud einfach abschalten. Ideen sind damit schnell und ohne große Investitionen testbar.

In Zukunft möchten verschiedene Anbieter die Option der souveränen Cloud herstellen. Die **souveräne Cloud** versucht die Vorteile von Public und Private Cloud zu kombinieren und die Nachteile zu verringern. Damit ist sie eine sehr interessante Option für deutsche Behörden und Organisationen, die ihre Datenanalyse auf moderne, rechtskonforme Füße stellen wollen. Capgemini und Orange bauen eine ähnlich gelagerte souveräne Cloud Namens Bleu[12] für französische Behörden auf, mit dem Ziel, dieses Konzept auf globaler Ebene auszurollen.

Abbildung 11: Wege in die Cloud

Kriterien	Cloud-Optionen		
	Public	Sovereign	Private
Skalierbarkeit	hoch	hoch	begrenzt
Plattformdienste	marktführend	begrenzt je nach Betreiber	vergleichsweise begrenzte Lösungen
Datenschutz, -lokation	wg. US Cloud Act kann Zugriff nicht gänzlich ausgeschlossen werden	volle Kontrolle des deutschen Cloud-Anbieter allein nach deutschem Gesetz	kontrolliert und verantwortet vom Betreiber
Preis	bestes Preis-Leistungs-Verhältnis	etwas teurer	stark abhängig vom lokalen Betreiber

2.4 RECHTLICHER RAHMEN UND GESETZESLAGE

Aus rechtlicher Perspektive berührt das Thema Data-Driven-Government unterschiedliche Regelungsbereiche. Bedeutsam sind etwa die E-Government-Gesetze von Bund und Ländern, das Onlinezugangsgesetz, die Open-Data-Gesetze auf Bundes- und Landesebene und das neue Bundes-Datennutzungsgesetz, um nur eine Auswahl zu nennen. Diese fügen sich in den Rahmen des Datenschutz- sowie IT-Sicherheitsrechts ein.

Daneben hat auch das Verwaltungsverfahrensrecht Einfluss auf DDG. Das Verwaltungsverfahrensrecht bildet die Strukturen und wesentlichen Entscheidungen von Verwaltungsverfahren ab und regelt die Interaktion zwischen Verwaltung und den Bürger:innen. Durch eine Entwicklung hin zu datenbasierten Entscheidungen können diese Verfahren und Interaktionen an unterschiedlichen Stellen verändert werden. Ein Aspekt ist beispielsweise das vielfach in Fachgesetzen verankerte Antragserfordernis, das sich künftig, je nach Daten- und Informationslage, möglicherweise automatisieren ließe. Eine Grundlage für weitere Entwicklungen hat § 35a Verwaltungs-verfahrensgesetz geschaffen, welcher das Erlassen von Verwaltungsakten vollständig durch automatische Einrichtungen zulässt.

Diese Regelungen haben eine normierende Funktion, dienen mitunter aber auch als DDG-Treiber. Die Schaffung eines entsprechenden Rechtsrahmens zur Weiterverwendung von Daten auf EU-Ebene, die Novellierung der Open-Data- und Public-Sector-Information-Richtlinie, wurde hierzulande im 2. Open-Data-Gesetz einschließlich Datennutzungsgesetz umgesetzt. Insgesamt zeigen die Regelungen ein klares Bekenntnis der EU-Mitgliedstaaten, die Bereitstellung und Nutzung „hochwertiger" Daten in allen EU-Mitgliedstaaten zu fördern.

Für Data-Driven-Government sind offene (Verwaltungs-) Daten im Zuge datenintensiver Anwendungen wie KI und Trends wie die vernetzte Stadt (Smart City) besonders wichtig. Hier geht es vor allem darum, dass der Staat seinen Teil zu einer funktionierenden Dateninfrastruktur beiträgt.

Die neuen Regelungen zu Open Data und zur Datennutzung schaffen dabei günstige Rahmenbedingungen für die Umsetzung von DDG.

Sind personenbezogene Daten im Rahmen von DDG von Interesse, dürfte die strenge datenschutzrechtliche Zweckbindung thematisiert werden, da eine Anonymisierung nicht immer mit vertretbarem Aufwand realisierbar oder zielführend ist.

Grundsätzlich bilden die gesetzlichen Grundlagen in Richtung Daten in Deutschland den Rahmen hinsichtlich der Verwend- und Verarbeitbarkeit verschiedener Daten. Durch innovative Gesetzesentwicklungen konnte die deutsche Bundesregierung hier viele Hindernisse aus dem Weg räumen. Beim Aufbau einer datengestützten Verwaltung sollte eine grundlegende Prüfung der rechtlichen Rahmenbedingungen durchgeführt werden.

2.5 GOVERNANCE UND PERSONAL

Um datengestützte Entscheidungsmechanismen technisch zu implementieren, ist die IT-Infrastruktur zu erneuern. Für eine erfolgreiche Einführung von DDG in Behörden sind darüber hinaus Kultur und Verhaltensweisen der Mitarbeitenden sowie die organisatorische Struktur innerhalb der Behörden anzupassen.

Data Governance befasst sich mit der organisationsweiten Kontrolle und Betreuung von Daten innerhalb einer Organisation durch einheitliche Begriffe, Standards und Richtlinien sowie eine kontinuierliche Verbesserung der Datenqualität. Der Umgang mit Daten und deren Verwaltung muss aktiv gehandhabt werden. Hierzu empfiehlt sich, eine Organisationsstruktur für eine Data Governance mit dedizierten Rollen und Verantwortlich-keiten einzuführen, die in der gesamten Organisation verankert ist. So werden in der Organisation Daten verfügbar gemacht, die sicher, zugreifbar, vertrauenswürdig und dem beabsichtigten Zweck dienlich sind.

Eine erfolgreiche Data Governance, die Mehrwerte für die Organisation generiert, findet immer im Dreiklang der drei Perspektiven Mensch, Technologie und Prozesse statt.

[12] Capgemini: "Capgemini and Orange announce plan to create 'Bleu', a company to provide a 'Cloud de Confiance' in France"(27.05.2021) https://www.capgemini.com/news/capgemini-and-orange-announce-plan-to-create-bleu-a-company-to-provide-a-cloud-de-confiance-in-france/ (Abruf am 16.02.2022)

Mensch

Bei der Einführung und Umsetzung einer datengetriebenen Organisation spielt die Kultur der jeweiligen Organisation eine entscheidende Rolle. Eine Studie von Capgemini bestätigt: Unternehmen, die ihre Kultur an die Digitalisierung anpassen, sind erfolgreicher und haben zufriedeneres Personal. In der Studie bestätigten 62 Prozent der Unternehmen die Kultur als größtes Hindernis für die digitale Transformation13. Ein Bewusstsein für eine datengetriebene Kultur in der Organisation zu schaffen, ist daher eine Voraussetzung, um die Akzeptanz und das Einbinden von Daten in den Alltagsprozessen zu erhöhen. Mitarbeiter:innen sollten Daten als hohes Gut betrachten, das einen hohen Mehrwert bei der Entscheidungsfindung bietet und auch die alltäglichen Arbeitsprozesse idealerweise qualitiativ verbessert. Change Management stellt die Weichen; beispielsweise mit der Vision und Roadmap einer datengesteuerten Kultur, die den positiven Zugang zu Daten erhöht, sodass diese erfolgreich eingesetzt werden. Ein solcher Kulturwandel ist selbstverständlich nur dann erfolgreich, wenn ein Commitment aller - und insbesondere der obersten Führungsebene - sichtbar ist und der kulturelle Wandel hin zu einer Data-Driven-Organisation unterstützt und gelebt wird.

Adäquate Kompetenzen zum richtigen Umgang und zum Einsatz der Daten in der Organisation sind zu verankern, um den Schritt zur Data Driven Culture zu gehen und das Potential der Daten vollumfänglich in der Organisation auszuschöpfen. Um aus Daten wertvolle Erkenntnisse zu gewinnen und die Entscheidungsfindung zu verbessern, muss das positive Mindset der Mitarbeiter:innen gegenüber Daten geschärft werden. Die geeigneten Kompetenzen rund um Datenverarbeitung, Datensammlung und Datennutzung sind ebenso zu schulen wie die richtige Anwendung der Tools, um sie gekonnt im Arbeitskontext einzusetzen. Nur so entsteht der gewünschte Mehrwert aus den Daten. Um diese Kompetenzen gezielt aufzubauen, bedarf es eines strategischen Rahmens, welcher die Kompetenzen nach den strategischen Zielen der Organisation ausrichtet. Durch Maßnahmen wie Umschulung (Reskilling) oder Weiterbildung (Upskilling) lassen sich die benötigen Kompetenzen aufbauen und intelligent nutzen. Zu den erforderlichen Kompetenzen in diesem Zusammenhang zählen analytisches und kritisches Denken, Problemlösung und die Fähigkeit zum Selbstmanagement. Mit diesen Fähigkeiten sind die Beschäftigten auch auf künftige Veränderungen vorbereitet und können sich schneller einer sich wandelnden Umgebung anpassen. Letztlich lässt sich in der Organisation über die Förderung dieser Kompetenzen und ein unterstützendes Change Management eine Kultur etablieren, die Daten als Chance anstatt als Herausforderung begreift.

Der Wandel hin zur Data-Driven-Governance greift tief in die Organisation. Strukturen, Technologie, Prozesse und Kultur sind neu zu denken und aufzustellen. Die Umstrukturierung stellt hohe Anforderungen an die Organisation und ihre Mitarbeiter:innen. Der tendenziell als veränderungsresistent angesehene Sektor muss Veränderungen initiieren und gleichzeitig mit branchenspezifischen Herausforderungen kämpfen. Eine öffentliche Verwaltung ist weniger als die Privatwirtschaft mit Innovationsdruck konfrontiert, muss sich weniger den Anforderungen der Kundschaft stellen und hat kaum Konkurrenz. Daher fehlt die Erfahrung, mit Veränderungsprozessen umzugehen, sodass die notwendige Geschwindigkeit und die Nachhaltigkeit bei Veränderungsmaßnahmen fehlen. Erst der erwähnte gesellschaftliche, soziale und technologische Wandel setzt auch den öffentlichen Sektor unter Handlungsdruck; nun müssen auch die öffentlichen Organisationen ihre Kultur, Struktur und Prozesse überdenken.

Die Impulse aus dem Umfeld können der Anlass sein, die organisatorische Realität auf den Prüfstand und den derzeitigen Reifegrad der datengetriebenen Organisation und insbesondere der Organisationkultur zu betrachten. Wichtige Schritte auf dem Weg zur DDG sind, eine Vision, Roadmap und eine konsistente Strategie über alle Ebenen hinweg aufzustellen. Zu Recht gilt der Kulturwandel oft als Marathon, der eine gute Vorbereitung und Geduld braucht, um die gewünschten Effekte zu erreichen.

Technologie

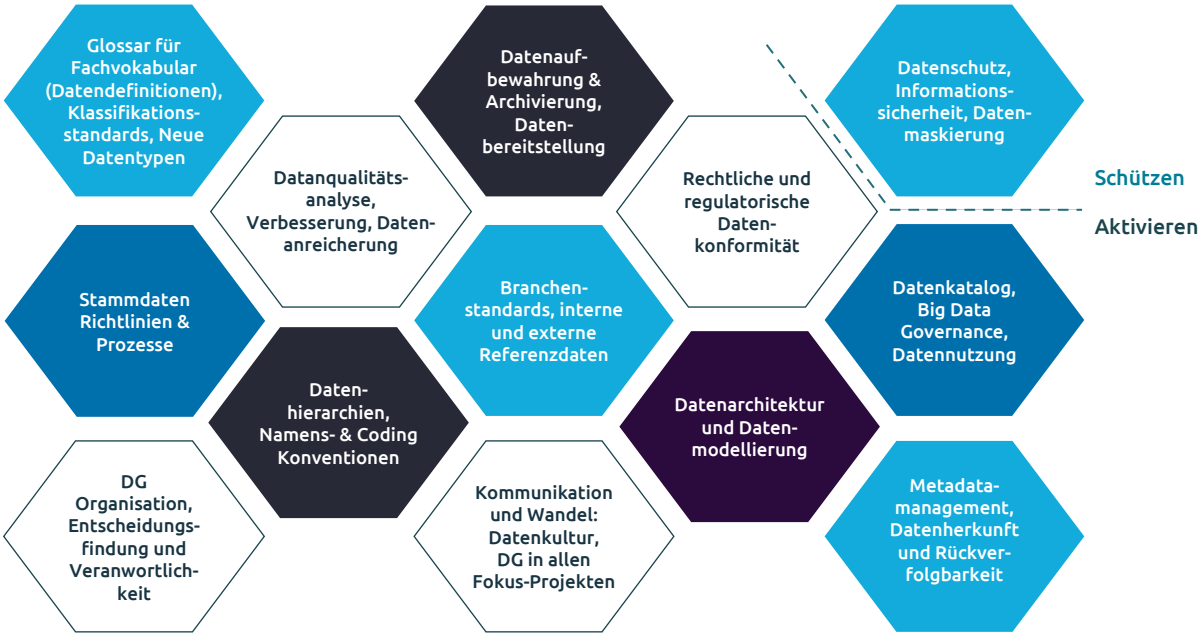
Der Data-Governance-Organisation bieten sich viele Technologien und Tools. Um eine funktionierende Data Governance aufzubauen, bedarf es einem ganzheitlichen Blick auf notwendige Fähigkeiten, welche in der Organisation implementiert werden sollten. Die folgende Grafik (Abb. 12) zeigt die Komponenten des Data-Governance-Gerüsts auf. Mit einem Datenkatalog lässt sich beispielsweise eine Übersicht über die verfügbaren Daten, deren Auffindbarkeit und Zugriff generieren. Glossare helfen, Begriffe einheitlich zu definieren; Tools ermöglichen die Nachvollziehbarkeit der Datenherkunft und die Rückverfolgbarkeit und sorgen so für Transparenz und Rechtssicherheit.

Eine mehrwertstiftende Nutzung dieser Tools erfordert fachliche Vorgaben und Prozesse. Datenkataloge sollten kontinuierlich befüllt und gepflegt werden, um ein gesamtheitliches, aktuelles Bild über die vorhandenen Daten zu erhalten. Einheitliche Datenhierarchien sowie Namenskonventionen schaffen ebenfalls Transparenz und Nachvollziehbarkeit; sie sind fachlich zu definieren und fortlaufend zu pflegen.

Organisatorisch ist zu entscheiden und zu gewährleisten, dass die Einhaltung von Datenschutz, Informationssicherheit und Datenmaskierung sichergestellt ist; auch hierfür gibt es dedizierte Tools.

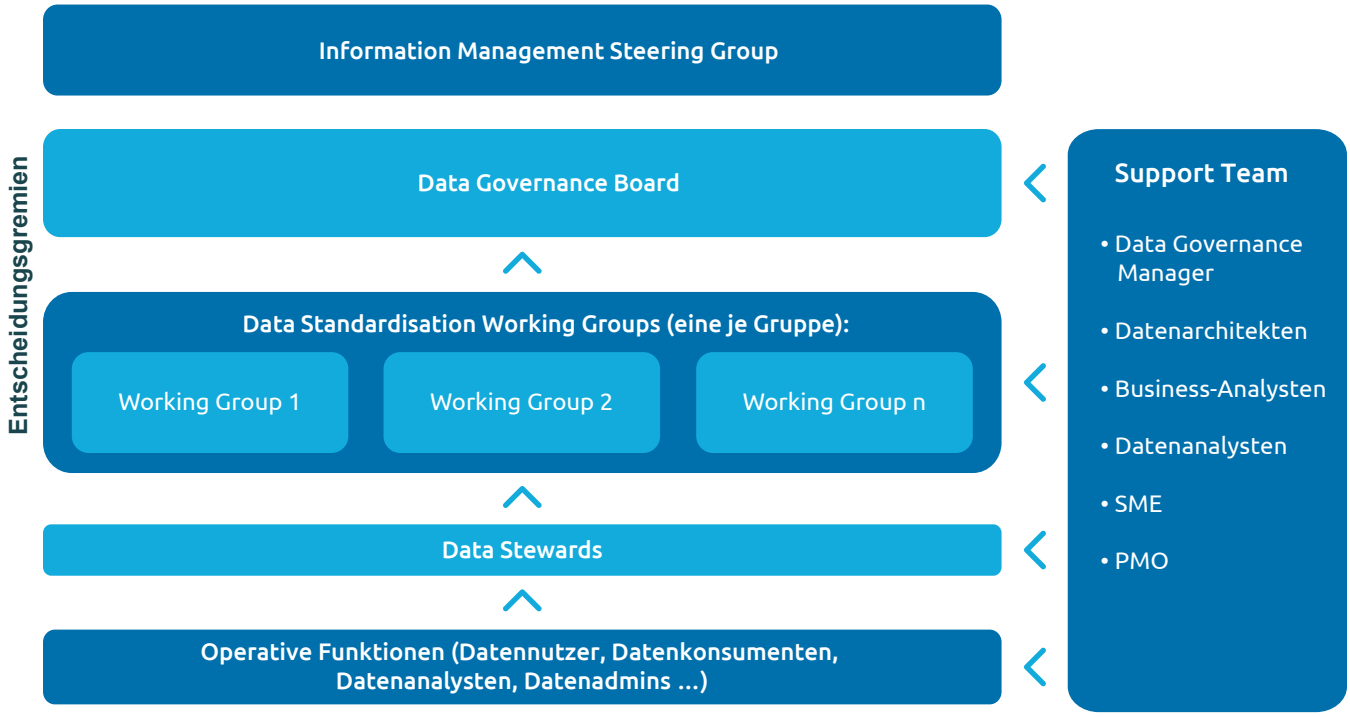
Als Teil eines Gesamtkonstrukts bauen die einzelnen Komponenten aufeinander auf und schließen sich nicht gegenseitig aus. Das Data-Governance-Gerüst ist hier ein Teil des **Capgemini QuickStart Framework für Information Governance**, welches unsere Kunden in 3 Schritten zu einer eigenständigen und funktionierend Governance verhilft.

Abbildung 12: Bausteine des Data-Governance-Gerüsts (Framework)



Prozesse

Abbildung 13: Capgeminis Referenz-Organisationsstruktur



Capgemini hat aus der jahrelangen Erfahrung mit Data Governance-Projekten im privaten Sektor verschiedene Standards und Vorgehensmodelle entwickelt. Im Folgenden gehen wir auf die Data-Governance-Referenz-Organisationsstruktur ein, die wir für den öffentlichen Sektor adaptiert haben.

Es gibt drei Entscheidungsgremien, welche vornehmlich mit fachlichen Stakeholdern besetzt sind, die durch IT-Rollen ergänzt werden.

Die **Information Management Steering Group** ist das höchste Entscheidungsgremium; es besteht aus Abteilungsleiter:innen und dem CIO. Das Gremium gibt die Rahmenbedingungen vor und prägt das Bewusstsein und die Kultur rund um Daten. Es verändert Organisation, Berufsbilder und Rollen und definiert strategische Richtungen und Ziele für Daten. Hierzu gehört auch, Ressourcen, Ausstattung und Finanzmittel bereitzustellen. Das Gremium genehmigt und überwacht Verbesserungsprojekte; es setzt Prioritäten zur Durchführung von Initiativen und prüft und genehmigt Vorschläge.

Die Information Management Steering Group steuert auch das **Data Governance Board**. Dieses setzt sich aus Gruppenleiter:innen, Chief Data Officer(n), Datenschutzbeauftragten und den Data Ownern zusammen. Diesen kommt eine Schlüsselrolle zu, welche im weiteren Verlauf detailliert dargelegt wird. Das Data Governance Board stellt sicher, dass sich das Data-Governance-Programm an den Behördenzielen und den Anforderungen der Transformationsprogramme ausrichtet. Auch prüft es den Fortschritt des Data-Governance-Programms; genehmigte Verbesserungsmaßnahmen verfolgt es nach und schließt sie ab. Das Board ernennt die Data Owner. Letztlich prüft das Data Governance Board die Ergebnisse der Working Groups und nimmt sie ab.

Die einzelnen Referatsgruppen bilden eine **Working Group**. Diese bestehen aus erfahrenem operativem Personal mit fundiertem Wissen über Datenbestände und Probleme. Je Referat wird ein Repräsentant entsandt. Zusätzlich hat jede Working Group einen Data Owner, einen Data Scientist und System Manager. In den Working Groups finden wöchentliche Arbeitsgruppensitzungen statt, um referatsspezifisches Wissen in Diskussionen einzubringen und diese ggf. im Anschluss an die Arbeitsgruppen im Detail zu diskutieren. Die Working Groups sind an der Erstellung neuer Richtlinien, Standards und Definitionen beteiligt. Sie stellen relevante Unterlagen bereit und erstellen abgestimmte Dokumente.

Die Entscheidungsgremien unterstützt ein **Central Data Governance Support Team**, welches aus erfahrenen Data Governance Managern, Business-Analysten, Datenanalysten und Datenarchitekten besteht und die Entscheidungsgremien mit seiner jeweiligen Expertise unterstützt.

Data Owner haben in den Entscheidungsgremien eine zentrale Rolle. Als Führungskräfte auf mittlerer Ebene der jeweiligen Gruppe haben sie eine gewisse Wissenstiefe in ihrem Datendomänenbereich, jedoch keine detaillierten Kenntnisse der Datenbestände.

Sie definieren Datenstandards, Richtlinien, Zugriffsrechte, Datenqualitäts-KPIs und relevante Qualitätskriterien, die für ihren Datensatz gelten. Sie geben regulatorischer Berichte frei und eskalieren funktionsübergreifende Datenprobleme an das Data Governance Board. In ihrer Verantwortung liegt es, Gelegenheiten zur Datenstandardisierung zu identifizieren, vorhandene Datenstandards und die Datenqualität zu verbessern und Datenqualitätsaspekte in der Organisation durch Wissenstransfer aktiv zu fördern. Data Owner stellen sicher, dass gemeldete Probleme im entsprechendem Funktionsdatenbereich behoben werden. Außerdem nominieren sie die Data Stewards für ihre Zuständigkeitsbereich.

Data Stewards sind nicht Teil eines Entscheidungsgremiums und haben eine eher operative Rolle. Sie besteht darin, Änderungen zu implementieren und alltägliche Probleme, Fragen und Probleme im Zusammenhang mit Datenvorgängen zu managen, welche die operativen Prozesse Ende-zu-Ende abdecken. Die Rolle eines Data Stewards können alle Mitarbeiter:innen der Behörde ausführen, die eine Beziehung zu Daten haben. Es ist keine bestimmte Position in der Organisation. Ein Data Steward ist eine Person, die Daten als Teil ihrer Rolle definiert, produziert oder verwendet. Sie fungieren als Bindeglied zwischen Data Ownern und den operativen Funktionen, die mit Daten arbeiten. Ihnen kommt daher eine Schlüsselrolle bei der erfolgreichen Implementierung von Data Governance zu: Sie treiben einerseits die Umsetzung der beschlossenen Standards und Richtlinien voran und unterstützen die operativen Funktionen bei deren Einhaltung. Andererseits tragen sie über den Data Owner

Das breite Aufgabenspektrum der Data Stewards zeigt die folgende Tabelle.

Abbildung 14: Aufgaben eines Data Stewards

Unterstützung Data Owner	Unterstützung Operative Funktionen
Unterstützung der Data Quality/Data Governance Teams bei der Definition von Datenstandards, Abschätzung der Auswirkungen von Änderungen und Management der Änderungen	Information über neue Standards, Richtlinien und Verfahren
Zusammenarbeit mit den Data Quality/Data Governance Teams bei der Nachverfolgung der Compliance von Datenstandards sowie Kommunikation von Compliance-Problemen und Lösungsmöglichkeiten	Zusammenarbeit, um Standards, Richtlinien und Verfahren zu implementieren und alltägliche Prozess- und Datenprobleme zu managen
Unterstützung des Data Owners bei der Präsentation neuer oder geänderter Standards zur Abnahme im Data Governance Board	Unterstützung von Projektteams durch Überprüfung neuer Projektanforderungen, um neue Standards, Änderungsbedarfe bestehender Standards und die Einhaltung dieser zu identifizieren
Ursachenanalyse und Korrektur von Datenqualitätsproblemen	Eskalation datenbezogener Probleme an den Data Owner
Dokumentation und Pflege von Datenstandards	
Abnahme neuer Datenobjekte und Datenwerte	

datenbezogene Probleme, die sich aus dem operativen Umgang mit Daten ergeben, in die Entscheidungsgremien. Um die Potentiale von Data-Driven-Government vollends für eine öffentliche Organisation zu nutzen, ist sicherzustellen, dass eine gesamtheitliche Data Governance über die organisationseigenen Daten eingesetzt wird. Diese soll, wie beschrieben, den Kulturwandel fördern, Standards setzen, die Veröffentlichung und Nutzung von Daten beaufsichtigen und einen Überblick über alle Datenaktivitäten erzielen.

Hierzu ist es erforderlich, dass Date Governance bereits in initialen DDG-Anwendungen berücksichtigt wird. Dabei sind Data Governance und die Data-Governance-Organisation jedoch kein Selbstzweck. Vielmehr richten sich die benötigte Komplexität und Tiefe nach dem Reifegrad der Organisation in Bezug auf DDG-Anwendungen und den Erfordernissen der Anwendungen im Fokus. Sie wächst mit der Daten-Nutzung und der Anzahl der Use Cases.

So reicht zunächst eine schlanke Organisation, die sich auf priorisierte Abteilungen und Referate beschränkten. Erfolgskritisch für die Data Governance ist jedoch, dass die oberste Führungsebene sie anstößt, einfordert und vorlebt. Auch lassen sich die einzusetzenden Tools zunächst mithilfe einfacher, in der Organisation etablierter Technologien umsetzen und mit zunehmendem Reifegrad durch Spezialsoftware ersetzen.

3 AUSBLICK UND FAZIT

> Was für Krisen gilt, gilt in gleicher Weise auch für das Alltagsgeschäft von Politik und Verwaltung (…) Evidenzbasiertes Handeln setzt aussagekräftige, aktuelle Daten voraus. Datengetriebenes Regieren ('Data Driven Government') sollte daher systematisch gefördert und zu einer Kernfähigkeit ministerieller und behördlicher Arbeit werden – gerade auch bei der Vorbereitung und Wirkungskontrolle von Gesetzen."

Nationaler Normenkontrollrat[13]

Es ist Teil des Koalitionsvertrags, Bundeskanzler Olaf Scholz spricht davon in seiner Regierungserklärung und viele Akteure im behördlichen Umfeld, wie beispielsweise der Nationale Normenkontrollrat als unabhängige Beratungsgremium der Bundesregierung, beschäftigen sich mit dem Thema Data Driven Government. Die Gründe hierfür sind verschieden: Digitale Trends und die Treiber Digitalisierung, Klimawandel, Globalisierung und demografischer Wandel beeinflussen die Lebensfelder der Gesellschaft im engeren Sinne; sie bestimmen aber auch die Anforderungen an die Verwaltung. Dieses Whitepaper zeigt die Potenziale von DDG als wichtigen Part für die Zukunftsfähigkeit einer stabilen, verlässlichen Verwaltung, die im Dienst der Gesellschaft, Wirtschaft und Forschung arbeitet.

Um eine datengestützte Verwaltung aufzubauen, gibt es nicht einen richtigen Weg. In diesem Whitepaper wurden die fünf von Capgemini identifizierten **Rahmenbedingungen** erkennbar, welche als Leitplanken Orientierung in Richtung datengestützte Verwaltung geben: Gesellschaft und Stakeholder, fachliche Herausforderungen, Technologie und Architektur, Gesetzeslage sowie Governance und Personal. Diese Rahmenbedingungen formen die Anforderungen an DDG und clustern alle Umfeldfaktoren in überschaubare und thematische Felder. Die Capgemini DDG Map zeigt, dass es die unterschiedlichsten Einstiegspunkte gibt und es nie den einen richtigen Weg geben wird. Dazu ist die Ausgangslage – geprägt durch die demokratischen und föderalen Strukturen – viel zu komplex. Ungeachtet des individuellen Einstiegs ist es jedoch wichtig, die grundsätzlichen Rahmenbedingungen zu berücksichtigen.

Um DDG in den Strukturen, Prozessen und Organisationen von Verwaltungen erfolgreich zu implementieren, muss DDG von Beginn an bei Transformationsprojekten mitgedacht werden. Damit dies gelingt, ist der Implementierungsprozess im Dreiklang der Perspektiven **Mensch, Technologie und Prozesse** zu denken.

[13] Nationaler Normenkontrollrat: "Initiative Leistungsfähige Verwaltung – Zukunftsfester Staat." (Juni 2021) https://www.normenkontrollrat.bund.de/resource/blob/72494/1936832/503cc3c5b6d4393803f9b39ef4a0cd99/120625-initiative-leistungsfaehige-verwaltung-zukunftsfester-staat-data.pdf (Abruf am 05.03.2022)

Auf der menschlichen Ebene bedarf es eines von oben vorgelebten Kulturwandels und der Anpassung an neue Abläufe und Prozesse. Die Verwaltung muss ein Bewusstsein für den vorhandenen, enorm wertvollen Datenschatz schaffen, über den sie verfügt. Verfügbare Forschungs- und Geodaten liefern bereits jetzt wertvolle Entscheidungshilfen im Bereich Klimaschutz. Dies zeigt der Wald-PoC von Capgemini.

Wünschenswert für die Zukunft ist, dass **Datendashboards** mit relevanten Daten und Informationen beliefert werden und Entscheidungsträger:innen bei komplexen Fragen in den unterschiedlichen Dimensionen und Lebensfeldern unterstützen können. Dafür bedarf es neben Kulturwandel auch eines kräftigen Investitions- und Innovationsschubs im technischen Bereich. Ein zentraler Bestandteil für die Etablierung von DDG sind **Datenökosysteme**, in denen alle relevanten Stakeholder integriert sind. Daten zu finden und auszuwerten, ist für den DDG-Ansatz essentiell. Es ist erkennbar, dass insbesondere der Datenaustausch einen beliebig skalierbaren Mehrwert bringen kann. Durch den Aufbau von Datenökosystemen lässt sich ein Netzwerk schaffen, welches es ermöglicht, eigene Daten anzureichern, und das dadurch aufgebaute Wissen zu erweitern. Verwaltungen müssen das Prinzip der Datenökosysteme mitdenken und umsetzen, um die bestmögliche Entscheidungsgrundlage aufzubauen. Dies bedeutet auch: Das Gerüst des Datenhoheitsprinzips ist zu überdenken.

Technologisch bieten sich verschiedene Optionen je nach Anforderungen an, um eine datengestützte Verwaltung aufzubauen. Capgemini hat eine moderne, herstellerneutrale Referenzarchitektur für moderne Datenanalyseplattformen entwickelt, die für unterschiedliche technologische Plattformen individuell angepasst und eingeführt werden kann. Im Zuge der Betrachtung einer Referenzarchitektur, der vorhandenen Daten sowie des angekündigten exponentiellen Wachstums an Daten liegt die Frage nach der zugrunde liegenden Technologie nicht fern. Die deutsche Verwaltung wird sich mit verschiedenen **Cloudansätzen** beschäftigen müssen: Public, Sovereign oder Private Cloud. Allerdings bieten momentan nur die Hyperscaler aus den USA technische Top-Leistung. Eine **Data Acceleration Platform** wie die von Capgemini ermöglicht die Vorteile der Public Cloud im Entwicklungsprozess zu nutzen und anschließend zu entscheiden, auf welche Cloud-Art die entwickelte Lösung aufgebaut werden soll.

Aus der prozessualen Perspektive spielt die Data-Governance-Referenz-Organisationsstruktur eine wichtige Rolle. Mit einer strukturierten, auf DDG angepassten **Organisationskultur** können Behörden und Verwaltungen die Kultur einer datenunterstützenden Arbeitsweise erfolgreich adaptieren.

Fünf Rahmenbedingungen und viele Ansatzpunkte- DDG erweist sich als komplexes Vorhaben. Es gilt zu erkennen, dass Daten und deren Auswertung und Nutzung ein essentieller Baustein der digitalen und zukunftsorientierten Verwaltung sind. Mit der Nutzbarmachung der Daten und der Einführung starker daten- und evidenzgestützter Entscheidungshilfen trägt DDG dazu bei, das Ungleichgewicht zwischen den wachsenden Anforderungen an die Verwaltung und den Anpassungsfähigkeiten der gegebenen Verwaltungsstruktur und -organisation auszugleichen. Durch die aktuellen politischen Entwicklungen, Entscheidungen und Pläne in Deutschland und in Europa wird der Weg in Richtung Data Driven Government geebnet. Capgemini positioniert sich hier klar und möchte die breit aufgestellte und gesammelte Expertise dafür nutzen, den Staat zu unterstützen, als Data-Pionierin voranzuschreiten und in Zukunft Daten in wertvolles Wissen umzuwandeln.

UNSERE AUTORINNEN UND AUTOREN:

• Melanie Adelt • Till Busche • Yanik Elixmann • Nathalie Miesch • Alexandra Nitzsche

• Axel Schräer • Peter Steinleitner • Kai Broek • Rüdiger Eberlein • Marie Jansen

• Kim Anh Nguyen • Lena Nowak • Hendrik Seidel

HABEN SIE FRAGEN ZU DATA DRIVEN GOVERNMENT? SPRECHEN SIE UNS GERNE AN!

Peter Steinleitner
Data Driven Government

Email : peter.steinleitner@capgemini.com

Alexandra Nitzsche
Data Driven Government

Email : alexandra.nitzsche@capgemini.com

ABBILDUNGSVERZEICHNIS

Abbildung 1: Capgemini Netzplan für Data Driven Government **5**
Abbildung 2: Rahmenbedingungen von DDG **7**
Abbildung 3: Die sich verändernde Gesellschaft und ihre aktuellen Herausforderungen **8**
Abbildung 4: Datenökosysteme ermöglichen eine effizientere Verbreitung öffentlicher Dienste **10**
Abbildung 5: Capgemini Research: Geplante Investitionen in den einzelnen Sektoren in Datenökosysteme **11**
Abbildung 6: Gemeinsame europäische Datenräume - Vorschlag aus der Europäischen Datenstrategie **12**
Abbildung 7: Datenspektrum **13**
Abbildung 8: Digital-Forest-Bodenklassifizierung **16**
Abbildung 9: Digital-Forest-Gesundheitsindex **16**
Abbildung 10: Capgeminis Referenzarchitektur für die moderne Datenanalyseplattform **17**
Abbildung 11: Wege in die Cloud **18**
Abbildung 12: Bausteine des Capgemini Data-Governance-Gerüsts (Framework) **21**
Abbildung 13: Capgeminis Referenz-Organisationsstruktur **22**
Abbildung 14: Aufgaben eines Data Stewards **23**

Über Capgemini

Capgemini ist einer der weltweit führenden Partner für Unternehmen bei der Steuerung und Transformation ihres Geschäfts durch den Einsatz von Technologie. Die Gruppe ist jeden Tag durch ihren Purpose angetrieben, die Entfaltung des menschlichen Potenzials durch Technologie zu fördern – für eine integrative und nachhaltige Zukunft. Capgemini ist eine verantwortungsbewusste und diverse Organisation mit einem Team von über 325.000 Mitarbeiterinnen und Mitarbeitern in mehr als 50 Ländern. Eine 55-jährige Unternehmensgeschichte und tiefgehendes Branchen-Know-how sind ausschlaggebend dafür, dass Kunden Capgemini das gesamte Spektrum ihrer Geschäftsanforderungen anvertrauen – von Strategie und Design bis hin zum Geschäftsbetrieb. Dabei setzt das Unternehmen auf die sich schnell weiterentwickelnden Innovationen in den Bereichen Cloud, Data, KI, Konnektivität, Software, Digital Engineering und Plattformen. Der Umsatz der Gruppe lag im Jahr 2021 bei 18 Milliarden Euro.

Get the Future You Want | www.capgemini.com/de

Die in diesem Dokument enthaltenen Informationen sind rechtlich geschützt. ©2022 Capgemini. Alle Rechte vorbehalten.