



Bern, 8. November 2023

Konkrete Anwendungsfälle von Datenwissenschaft zum Gemeinwohl entlang des Prozesses der Politikgestaltung

Im Rahmen der Umsetzung der Datenwissenschaftsstrategie des Bundes

Inhaltsverzeichnis

Zusammenfassung	3
1 Einleitung	6
2 Konsolidierte Erkenntnisse	8
3 Dokumentation der Anwendungsfälle	12
3.1 Übersicht	12
3.2 Nutzung von Open Source Intelligence (OSINT).....	13
3.3 Ganzheitliche Lagebilder in Krisensituationen.....	19
3.4 Datenbasierte Unterstützung von Aufsicht	27
3.5 Ganzheitliche Modellsysteme zur Entscheidungsunterstützung	33
3.6 Qualitätssicherung von Daten durch Plausibilitätsprüfungen	40
3.7 Digitale Zwillinge	46
4 Anhang	54
4.1 Datenwissenschaft im Prozess der Politikgestaltung	54
4.2 Ansatz zur Erarbeitung des Berichts	55
4.2.1 Vorgehen.....	55
4.2.2 Charakteristika eines Anwendungsfalls	56
4.2.3 Mitglieder der Begleitgruppe und der Arbeitsgruppen	57

Zusammenfassung

Dem Eidgenössischen Departement des Innern (EDI) wurde der Auftrag erteilt, in Zusammenarbeit mit den anderen Departementen und dem ETH-Bereich einen Bericht über konkrete Anwendungsfälle von Datenwissenschaft in der Bundesverwaltung zum Gemeinwohl entlang des Prozesses der Politikgestaltung zu erstellen. Ziel des Berichts ist zum einen, das Potenzial von Datenwissenschaft in der Bundesverwaltung aufzuzeigen, und zum anderen, erforderliche Rahmenbedingungen für ihren breiten, transversalen Einsatz zu identifizieren.

In mehreren Arbeitssitzungen wurden mit Vertretungen der Bundesverwaltung und des ETH-Bereichs sechs Anwendungsfälle ausgearbeitet und analysiert. Die Anwendungsfälle sind: (1) «Nutzung von Open Source Intelligence (OSINT)», (2) «Ganzheitliche Lagebilder in Krisensituationen», (3) «Datenbasierte Unterstützung von Aufsicht», (4) «Ganzheitliche Modellsysteme zur Entscheidungsunterstützung», (5) «Qualitätssicherung von Daten durch Plausibilitätsprüfungen» und (6) «Digitale Zwillinge». Die Auswahl der Anwendungsfälle erhebt keinen Anspruch auf Vollständigkeit und stellt keine Priorisierung dar. Sie sind exemplarisch und sollen eine möglichst grosse Bandbreite an datenwissenschaftlichen Methoden und Tätigkeitsfelder der Bundesverwaltung abbilden.

Die Erarbeitung der Anwendungsfälle hat sowohl zu neuen Erkenntnissen geführt als auch den bereits in der Datenwissenschaftsstrategie des Bundes (DSStB) identifizierten Handlungsbedarf untermauert. So akzentuieren die Anwendungsfälle das grosse Potenzial der Datenwissenschaft für die Bundesverwaltung. Bereits heute setzen einige Verwaltungseinheiten Datenwissenschaft ein, doch in der Breite bleibt viel Potenzial in der Bundesverwaltung ungenutzt. Für einen breiteren und transversalen Einsatz der Datenwissenschaft müssen die Rahmenbedingungen in verschiedenen Handlungsfeldern geschaffen beziehungsweise verbessert werden. Die wichtigsten Rahmenbedingungen für den Einsatz von Datenwissenschaft umfassen unter anderem den Führungswillen und die Führungsunterstützung, die Aus- und Weiterbildung im Bereich Datenwissenschaft sowie die Gewinnung von Fachkräften, die Datenverfügbarkeit und Interoperabilität (insb. durch die Förderung vertrauenswürdiger und interoperabler Datenräume), eine Rechtssetzung, die datenbasierte Entscheidungsunterstützung erleichtert, und die Berücksichtigung von Methoden aus dem Bereich der künstlichen Intelligenz (KI) als wichtige Werkzeuge in den datenwissenschaftlichen Anwendungen. Der Steuerung und Koordination («Governance») der Datenwissenschaft in der Bundesverwaltung kommt eine besondere Bedeutung zu.

Die Schaffung der identifizierten Rahmenbedingungen für den Einsatz von Datenwissenschaft wird in der Bundesverwaltung teilweise bereits in initiierten Vorhaben wie beispielsweise den Folgeaufträgen zur Umsetzung der DSStB adressiert. Die Ergebnisse des vorliegenden Berichts fliessen entsprechend in deren Bearbeitung ein. Zu diesen Vorhaben zählen insbesondere die Evaluation des Kompetenzzentrums für Datenwissenschaft (DSCC) und die Evaluation des Kompetenznetzwerks für KI (CNAI), die Entwicklung eines Verhaltenskodex für menschenzentrierte und vertrauenswürdige

Datenwissenschaft, die Schaffung von vertrauenswürdigen und interoperablen Datenräumen und die Konzeptionierung einer «kollaborativen Datenwissenschaftsplattform».

In der folgenden Abbildung sind die wichtigsten Erkenntnisse auf einen Blick dargestellt.

Das Potenzial für Datenwissenschaft ist gross, doch bleibt unausgeschöpft

Einzelne Verwaltungseinheiten setzen bereits heute sehr kompetent Datenwissenschaft ein, doch in der Breite bleibt viel Potenzial in der Bundesverwaltung ungenutzt.

Um das Potenzial auszuschöpfen, sind Rahmenbedingungen zu schaffen



Führungswille und -unterstützung



Aus- und Weiterbildung sowie Rekrutierung



Werkzeuge, Plattformen und Rechenleistung



Datenverfügbarkeit und Interoperabilität



Transversale, interdisziplinäre Zusammenarbeit



Nachhaltige Betriebs- und Finanzierungsmodelle



Etablierung der Grundprinzipien beim Einsatz von Datenwissenschaft



Rechtssetzung, die datenbasierte Entscheidungsunterstützung erlaubt



Künstliche Intelligenz als wichtiges Hilfsmittel



Steuerung in manchen, Koordination in vielen Bereichen

Vieles ist bereits auf den Weg gebracht, doch es bleibt Handlungsbedarf

Die aufgeführten Handlungsbedarfe betreffen in vielen Teilen bereits initiierte Vorhaben oder Evaluationen. Sie untermauern deren Bedeutung und müssen konsequent umgesetzt werden. Punktuell sind weitere Massnahmen nötig.

Abbildung 1: Übersicht der konsolidierten Erkenntnisse.

Eine Rahmenbedingung wird nicht explizit durch andere Aufträge adressiert, ist jedoch von herausragender Bedeutung: die strategische Verankerung von Datenwissenschaft in den Departementen und Ämtern durch die Führungskräfte in Politik und Verwaltung. Des Weiteren sind die Zuständigkeiten und nächsten Schritte im Bereich der Aus- und Weiterbildung von Datenwissenschaft in der Bundesverwaltung sowie der Rekrutierung von externen Talenten zu definieren und konsequent im Rahmen der Digitalisierungsstrategie des Bundes (SI-6) voranzubringen.

1 Einleitung

In Folge der Gutheissung der Datenwissenschaftsstrategie des Bundes (DSSStB)¹ hat der Bundesrat am 2. Dezember 2022 verschiedene Aufträge vergeben, um die Umsetzung der Strategie zu fördern.

Dem Eidgenössischen Departement des Innern (EDI), speziell dem Bundesamt für Statistik (BFS), wurde der Auftrag erteilt, in Zusammenarbeit mit den anderen Departementen und dem ETH-Bereich einen Bericht über konkrete Anwendungsfälle von Datenwissenschaft in der Bundesverwaltung zum Gemeinwohl entlang des Prozesses der Politikgestaltung zu erstellen. Der Bericht hat zweierlei Ziele.

Einerseits soll der Bericht einzelne Verwaltungseinheiten dabei unterstützen, das Potenzial der Datenwissenschaft zu erkennen und ihren Einsatz zu fördern. Darüber hinaus ermöglicht er die Identifikation von beispielhaften, übergreifenden Anwendungsfällen in der gesamten Bundesverwaltung. Zu diesem Zweck erläutert der Bericht den Einsatz datenwissenschaftlicher Anwendungen im Arbeitsalltag verschiedener Verwaltungseinheiten des Bundes. Dadurch wird einfach verständlich und praxisnah aufgezeigt, wie Führungskräfte aus Politik und Verwaltung sowie Fachleute auch in ihrem Arbeitsalltag Datenwissenschaft einsetzen können, um die Wirksamkeit und Effizienz der staatlichen Tätigkeit zu erhöhen.² Entsprechend liegt der Fokus nicht auf technischen Aspekten der Umsetzung, sondern auf den Aktivitäten der Bundesverwaltung. Andererseits sollen die notwendigen Rahmenbedingungen und aktuellen Herausforderungen hinsichtlich des Einsatzes von Datenwissenschaft in der Bundesverwaltung aufgearbeitet werden. Auf diese Weise sollen Handlungsfelder abgeleitet werden, um den weiteren Einsatz der Datenwissenschaft in der Bundesverwaltung zu fördern.

Vor dem Hintergrund dieser Ziele erhebt die Auswahl der Anwendungsfälle keinen Anspruch auf Vollständigkeit und stellt keine Priorisierung dar. Die Beschreibungen der Anwendungsfälle sind nicht im Sinne von Projektanträgen oder -planungen zu verstehen. Sie sind exemplarisch und sollen eine möglichst grosse Bandbreite der möglichen datenwissenschaftlichen Methoden und Tätigkeitsfelder der Bundesverwaltung abbilden. Die Anwendungsfälle wurden von einer interdepartementalen Begleitgruppe ausgewählt (siehe Kapitel 4.2).

Der vorliegende Bericht ist wie folgt aufgebaut:

- In Kapitel 2 werden die wesentlichen Erkenntnisse aus der Erarbeitung der Anwendungsfälle erläutert. Zudem werden die nächsten Schritte aufgezeigt.
- Kapitel 3 beschreibt die sechs ausgewählten Anwendungsfälle und die wichtigsten Anforderungen je Anwendungsfall.

¹ <https://www.news.admin.ch/news/message/attachments/74233.pdf> und <https://www.bfs.admin.ch/bfs/de/home/dscc/dscc.gnpdetail.2022-0738.html>

² Vor diesem Hintergrund ist die Ambition der DSSStB formuliert: «Datenbasierte Entscheidungsunterstützung in der Bundesverwaltung: verstärkte Wirksamkeit und Effizienz im staatlichen Handeln durch Datenwissenschaft». 6/58

Anwendungsfälle der Datenwissenschaft entlang des Prozesses der Politikgestaltung

- Im Anhang werden die strategischen und konzeptionellen Grundlagen aus der DSStB zusammengefasst. Zudem wird das Projektvorgehen aufgezeigt.

2 Konsolidierte Erkenntnisse

Die Erarbeitung sowie die zugrundeliegenden Gespräche mit Vertretungen der Bundesverwaltung und des ETH-Bereichs haben sowohl zu neuen Erkenntnissen geführt als auch den bereits in der DSStB identifizierten Handlungsbedarf untermauert.

Das Potenzial der Datenwissenschaft für die Bundesverwaltung ist gross. Über umfassende Lagebilder in Krisensituationen, die Planung von humanitären Einsätzen, die (Teil-)Automatisierung von Aufsichtstätigkeiten bis hin zur langfristigen Simulation der Schweizer Ernährungsnachhaltigkeit – datenwissenschaftliche Anwendungen können die Arbeit der Bundesverwaltung in zahlreichen Politikfeldern und Tätigkeitsbereichen unterstützen. Zugleich ist ihr Einsatz schon heute möglich und kein Thema für die «nächste Geländekammer».

Einige Verwaltungseinheiten setzen bereits heute Datenwissenschaft ein, doch in der Breite bleibt viel Potenzial in der Bundesverwaltung ungenutzt. Viele der vorgestellten transversalen Anwendungsfälle basieren auf bereits laufenden Vorhaben oder Überlegungen der jeweiligen Verwaltungseinheiten. Die involvierten Mitarbeitenden sind motiviert und verfügen über die notwendigen Fähigkeiten. Doch es ist eine Nische der Fachexpertinnen und -experten. Das Wissen über die Einsatzmöglichkeiten durchdringt weder die Amtsgrenzen, die Disziplinen noch die Hierarchiestufen.

Für einen breiten Einsatz der Datenwissenschaft müssen die Rahmenbedingungen in diversen Bereichen geschaffensweise verbessert werden. In der DSStB wurden Handlungsbereiche zur Förderung der Datenwissenschaft in der Bundesverwaltung identifiziert. Diese Handlungsbereiche wurden erhärtet und präzisiert. Die wichtigsten Handlungsbereiche sind:

- **Führungswille und -unterstützung:** Führungskräfte in Politik und Verwaltung müssen den Einsatz von Datenwissenschaft in ihren jeweiligen Departementen oder Ämtern fördern und einfordern. Sie müssen deren Potenzial und die Herausforderungen für die jeweiligen Politik- und Tätigkeitfelder kennen und dabei unterstützt werden, politische und organisatorische (Informations-)Bedarfe in konkrete datenwissenschaftliche Anforderungen zu übersetzen («Requirements Engineering»). Es braucht insbesondere ein klares Verständnis über den Einsatzzweck und die Bereitschaft, Ressourcen zu investieren.
- **Aus- und Weiterbildung sowie Rekrutierung:** Die Bundesverwaltung muss in ihre Mitarbeitenden investieren und klar definierte Berufsbilder und -profile im Bereich der Datenwissenschaft etablieren. Neben dedizierten Fachkräften sowie Datenwissenschaftsspezialistinnen und -spezialisten sind Führungskräfte und Fachmitarbei-

tende stufengerecht auszubilden. Im Rahmen der Digitalisierungsstrategie des Bundes (SI-6)³ wird dieser Handlungsbereich bereits adressiert. Die Zuständigkeiten und nächsten Schritte sind zu definieren und konsequent voranzubringen.

- **Werkzeuge, Plattformen und Rechenleistung:** Die Verfügbarkeit von Datenwissenschaftswerkzeugen und -plattformen ist in der Bundesverwaltung aktuell nur punktuell gewährleistet und in aller Regel unzureichend. Aus diesem Grund wurde bereits als Folgeauftrag zur DSSStB die Erarbeitung eines Konzepts für eine «kollaborative Datenwissenschaftsplattform» lanciert. Der laufende Einsatz datenwissenschaftlicher Anwendungen erfordert eine entsprechende Rechenleistung. Entsprechend erfordert es eine stärkere Koordination der Cloud-Beschaffung in der Bundesverwaltung mit Blick auf Rechenleistung («compute») und Speicherung («storage») und eine Abwägung der ökologischen Nachhaltigkeit.
- **Datenverfügbarkeit und Interoperabilität:** Der Nutzen datenwissenschaftlicher Anwendungen beruht in hohem Masse auf der Verfügbarkeit relevanter Daten und der Möglichkeit, diese einfach auszutauschen. Dies betrifft insbesondere transversale Anwendungsfälle, in welchen beispielsweise verschiedene Bundesämter und Staatsebenen involviert sind und Daten rasch zusammengezogen werden müssen (z.B. Lagebilder für Naturkatastrophen). Verschiedene Vorhaben des Bundes zielen bereits auf eine Verbesserung der Datenverfügbarkeit und Interoperabilität ab. Die Verfügbarkeit von Geodaten ist mit der Bundes Geodaten-Infrastruktur (BGDI)⁴ schon heute in weiten Teilen auf Bundesebene gewährleistet, während auf föderaler Ebene das Konzept für eine Nationale Geodaten-Infrastruktur (NGDI)⁵ in Erarbeitung ist. Auf Basis des Berichts des UVEK (BAKOM) und des EDA (DV) zur «Schaffung von vertrauenswürdigen Datenräumen basierend auf der digitalen Selbstbestimmung» wurden Folgeaufträge erteilt, um den Betrieb von vertrauenswürdigen und interoperablen Datenräumen zu fördern⁶. Dadurch soll unter anderem das Datennutzungspotenzial künftig besser ausgeschöpft und gewährleistet werden, dass Individuen und Unternehmen selbstbestimmt über ihre Daten verfügen können. Im Programm Nationale Datenbewirtschaftung (NaDB) wird die Harmonisierung der generellen Datenbestände gefördert und gemeinsame Standards auf der «I14Y-Interoperabilitätsplattform» publiziert.⁷ Der Datenschutz und das Datensparsamkeitsprinzip sind dabei zu wahren. Aus diesem Grund erarbeitet das BFS momentan ein Konzept über die Anwendung von Datenwissenschaft zur Wahrung der Privatsphäre bei der gemeinsamen Nutzung von Daten in der Bundesverwaltung.
- **Transversale, interdisziplinäre Zusammenarbeit:** Der Einsatz von Datenwissenschaft erfordert das Zusammenspiel von verschiedenen Fachbereichen, Disziplinen und Organisationen innerhalb der Bundesverwaltung sowie mit Dritten (u.a. Wis-

³ https://www.bk.admin.ch/bk/de/home/digitale-transformation-ikt-lenkung/strategie-planung/ikt-strategie_bund_2020-2023.html

⁴ [Bundes Geodaten-Infrastruktur \(admin.ch\)](https://www.bk.admin.ch/bk/de/home/digitale-transformation-ikt-lenkung/strategie-planung/ikt-strategie_bund_2020-2023.html)

⁵ [Nationale Geodateninfrastruktur \(NGDI\): e-geo.ch - Kanton Luzern](https://www.ngeo.ch/)

⁶ <https://www.admin.ch/gov/de/start/dokumentation/medienmitteilungen.msg-id-87780.html>

⁷ <https://www.bfs.admin.ch/bfs/de/home/nadb/nadb.html>

senschaft und Forschung, Kantone, Wirtschaft). Dies erfordert zum einen eine offene Zusammenbeitskultur und zum anderen Strukturen und Prozesse zur Zusammenarbeit.

- **Nachhaltige Betriebs- und Finanzierungsmodelle:** Datenwissenschaftliche Anwendungen müssen langfristig finanziert, betrieben, weiterentwickelt und evaluiert werden. Ein Business-Case ist zu erarbeiten. Mit dem Abschluss des Projekts zur Entwicklung einer Anwendung sind die entsprechenden Ressourcen bereitzustellen sowie Strukturen, Prozesse und Rollen festzulegen. Auch nach dem «Go-Live» braucht es «Management Attention». Insbesondere für Anwendungen, welche in Krisen zum Einsatz kommen, wie beispielsweise Lagebilder, müssen die erforderlichen Rahmenbedingungen geschaffen werden.
- **Etablierung der Grundprinzipien beim Einsatz von Datenwissenschaft:** Um die Qualität und die Ethik der datenwissenschaftlichen Anwendungen zu gewährleisten, müssen die in der DSStB definierten Grundprinzipien konsequent umgesetzt werden und das Rechtsstaatsprinzip einhalten. Zudem ist der Stand der Umsetzung der Öffentlichkeit transparent zu machen. Nur so kann Vertrauen aufgebaut werden. Die Erarbeitung des Verhaltenskodex für menschenzentrierte und vertrauenswürdige Datenwissenschaft ist ein erster Schritt. Weitere Schritte müssen folgen, um die Grundprinzipien kulturell und institutionell zu verankern.
- **Rechtssetzung, die datenbasierte Entscheidungsunterstützung erleichtert:** Gesetze, Verordnungen und Weisungen müssen die zunehmend digitale und datenbasierte Arbeit der öffentlichen Verwaltung berücksichtigen und antizipieren. Dazu gehören beispielsweise Vorgaben an Dritte hinsichtlich der Datenproduktion, -qualität und -übermittlung (z.B. Anstalten). Ein Beispiel hierfür ist das Informationssystemgesetz im Bereich Sozialversicherung, das aktuell in Vorbereitung ist.
- **Künstliche Intelligenz als wichtiges Hilfsmittel:** Der Einsatz künstlicher Intelligenz (KI) ist ein wesentliches Werkzeug, um das volle Potenzial der Datenwissenschaft auszuschöpfen. Die Bundesverwaltung muss die weitere Entwicklung ihres Einsatzes über das gesamte Methodenspektrum hinweg (z.B. Methoden aus der erweiterten Statistik, dem maschinellen Lernen und dem Bereich der KI) ganzheitlich steuern und koordinieren.
- **Steuerung in manchen, Koordination in vielen Bereichen («Governance»):** Es besteht ein Steuerungs- und Koordinationsbedarf in verschiedenen Bereichen. Vergleichbar mit der digitalen Transformation und Informatik ist es erforderlich, durch einen gemeinsamen, abgestimmten Ansatz Synergien zu schaffen, die Interoperabilität innerhalb der Bundesverwaltung und mit Dritten sicherzustellen und das Vertrauen der Öffentlichkeit in den Einsatz der Datenwissenschaft zu stärken. Dies betrifft verschiedene der hier aufgeführten Rahmenbedingungen für den Einsatz von Datenwissenschaft (z.B. Plattform und Werkzeuge, Standardisierung, Ethik, Rechtssetzung, Aus- und Weiterbildung). Im Rahmen der Evaluation des Kompetenzzentrums für Datenwissenschaft (DSCC) und der Evaluation des Kompetenznetzwerks für KI (CNAI) werden diese Fragestellungen aufgegriffen und vertieft.

Wie jeweils angedeutet betreffen die aufgeführten Handlungsbedarfe in vielen Teilen bereits initiierte Vorhaben oder laufende Evaluationen. Nur eine Rahmenbedingung wird nicht adressiert: die **strategische Verankerung von Datenwissenschaft in den Departementen und Ämtern durch die Führungskräfte in Politik und Verwaltung**. Die Entscheidung zum Einsatz von Datenwissenschaft ist eine Führungsaufgabe in den Fachämtern. Verwaltungseinheiten wie das CNAI und das DSCC können verbinden, beraten und unterstützen, doch die Führungskräfte in den Departementen und Ämtern müssen Verantwortung für eine systematische Berücksichtigung und Anwendung von Datenwissenschaft in ihren Fachprozessen und den zugrundeliegenden Rechtsgrundlagen übernehmen. Zudem sind die Zuständigkeiten und nächsten Schritte im Bereich der Aus- und Weiterbildung von Datenwissenschaft in der Bundesverwaltung sowie der Rekrutierung von externen Talenten zu definieren.

Eng damit verbunden sind die erforderlichen Ressourcen. Sollen die in der DSSStB gesteckten Ziele erreicht werden, sind durch die einzelnen Verwaltungseinheiten Investitionen in konkrete datenwissenschaftliche Vorhaben zu tätigen, vergleichbar mit den beschriebenen Anwendungsfällen. Zudem sind die Mittel zur Umsetzung der oben aufgeführten übergeordneten bereits initiierten Vorhaben erforderlich.

3 Dokumentation der Anwendungsfälle

3.1 Übersicht

In diesem Kapitel sind insgesamt sechs ausgewählte Anwendungsfälle beschrieben.⁸ Die Anwendungsfälle sind: (1) «Nutzung von Open Source Intelligence (OSINT)», (2) «Ganzheitliche Lagebilder in Krisensituationen», (3) «Datenbasierte Unterstützung von Aufsicht», (4) «Ganzheitliche Modellsysteme zur Entscheidungsunterstützung», (5) «Qualitätssicherung von Daten durch Plausibilitätsprüfungen» und (6) «Digitale Zwillinge».

Die Beschreibung der Anwendungsfälle besteht aus je vier Abschnitten. Zuerst werden die dem Anwendungsfall zugrundeliegenden datenwissenschaftlichen Methoden sowie deren Mehrwert und Relevanz im Allgemeinen erläutert. Im zweiten Abschnitt folgt ein praktisches Beispiel zur Anwendung von Datenwissenschaft in einer Verwaltungseinheit des Bundes im Rahmen eines konkreten Tätigkeits- oder Politikbereichs. Dieses Praxisbeispiel basiert einerseits auf bereits laufenden Vorhaben oder Überlegungen der jeweiligen Verwaltungseinheit und andererseits auf hypothetischen Entwicklungen, welche in den Gesprächen mit Expertinnen und Experten der Bundesverwaltung und des ETH-Bereichs erarbeitet wurden. Im dritten Abschnitt wird jeweils darauf eingegangen, in welchen anderen Bereichen der Bundesverwaltung der Anwendungsfall eingesetzt werden könnte. Abschliessend wird aufgeführt, welche Anforderungen erfüllt sein müssen, um den Anwendungsfall in der Bundesverwaltung transversal und in der Breite einzusetzen. Dabei wurden sieben Dimensionen betrachtet (siehe Abbildung 2).



Abbildung 2: Betrachtete Anforderungsdimensionen.

Die Dokumentation der Anwendungsfälle wurde jeweils von einer dedizierten Arbeitsgruppe bestehend aus Expertinnen und Experten der Bundesverwaltung und des ETH-

⁸ Eine übergeordnete Gruppierung von Anwendungsfällen zur strategischen Steuerung der Anwendung von Datenwissenschaft in der Bundesverwaltung kann im Rahmen der Evaluation CNAI stattfinden.

Bereichs erarbeitet. Das Vorgehen zur Erarbeitung der Anwendungsfälle und die Zusammensetzung der Arbeitsgruppen sind in Kapitel 4.2 beschrieben.

3.2 Nutzung von Open Source Intelligence (OSINT)

Am Beispiel von humanitärer Hilfe in Krisensituationen

Allgemeiner Teil

Beschreibung

Die Nutzung von Open Source Intelligence (OSINT) basiert auf der systematischen Sammlung, Analyse und Auswertung von öffentlich zugänglichen Informationen. Diese Informationen können aus vielfältigen Quellen stammen, einschliesslich Nachrichtenmedien, Social-Media-Beiträgen, Blogs, Foren und publizierten Berichten, Dokumenten oder Daten von Regierungen, privaten Unternehmen oder Nichtregierungsorganisationen (NGOs). Der Schlüssel zur Verarbeitung und Nutzung dieser grossen Menge an Informationen liegt in der Anwendung von datenwissenschaftlichen Methoden und Algorithmen.

Durch den Einsatz von KI, insbesondere Methoden des maschinellen Lernens, können grosse Mengen an Daten, einschliesslich unstrukturierter Daten, automatisiert analysiert und verarbeitet werden. Dies beinhaltet u.a. Textdaten und Fernerkundungsdaten. Mit Hilfe von datenwissenschaftlichen Methoden und Algorithmen können entscheidungstragende dabei unterstützt werden, z.B. Stimmungen in der Bevölkerung zu erfassen, die Auswirkungen von Politikmassnahmen zu bewerten und neue Themen zu identifizieren, die möglicherweise politische Aufmerksamkeit erfordern.

Nachdem die Daten analysiert wurden, stehen verschiedene Ansätze zur Verfügung, um die aggregierten Werte und Ergebnisse für die Nutzenden verständlich und zugänglich zu machen. Eine dieser Möglichkeiten ist die Entwicklung von KI-gestützten Informationssystemen, welche dabei helfen, relevante Dokumente in einer vorhandenen internen Datenbank mit Hilfe natürlicher Sprache intuitiv aufzufinden.

Mehrwert und Relevanz

Die Nutzung von OSINT bietet bedeutende Mehrwerte für den Prozess der Politikgestaltung. Im Kern geht es darum, das volle Potenzial öffentlich zugänglicher Informationen zu nutzen, um fundiertere und effizientere Entscheidungen treffen zu können.

Bislang werden Daten aus öffentlich verfügbaren Quellen meist in mühsamer Handarbeit analysiert und ausgewertet, was vor allem bei unstrukturierten Textdaten einen erheblichen Aufwand bedeutet. Mit der Anwendung von datenwissenschaftlichen Methoden und Algorithmen, insbesondere aus dem maschinellen Lernen und dem Bereich der KI, z.B. natürliche Sprachverarbeitung (Natural Language Processing, NLP), kann dieser Prozess jedoch wesentlich effizienter gestaltet werden. Die automatisierte Analyse, Verarbeitung und

Strukturierung dieser Daten ermöglichen eine zielgerichtete und systematische Analyse grosser Datenmengen. Zudem entlastet sie die Mitarbeitenden. Die Folge sind schnellere und effizientere Entscheidungsprozesse in der Politikgestaltung, ohne die menschliche Analyse- und Entscheidungskompetenz zu ersetzen.

Ein weiterer Mehrwert liegt in der Entwicklung von benutzerfreundlichen Oberflächen als Zugangsstelle für OSINT. Diese ermöglichen es den Nutzenden, schnell und gezielt auf für sie relevante Informationen zuzugreifen, ohne lange danach zu suchen.

Mit dem Einsatz von datenwissenschaftlichen Methoden und Algorithmen können zudem Muster und Trends in den Daten identifiziert werden. Auf dieser Basis lassen sich Prognosen erstellen. Insbesondere bei grossen Datenmengen ist dies für die Mitarbeitenden eine erfolgskritische Unterstützung.

Im Kontext des politischen Entscheidungsprozesses – von der Identifikation und dem Verständnis des Problems über die Entwicklung von Optionen bis hin zur Evaluation – bietet OSINT und der damit verbundene Einsatz von Datenwissenschaft das Potenzial, den gesamten Prozess erheblich zu verbessern.

Anwendungsbeispiel aus der Praxis

OSINT-basierte Bedürfniserhebung in humanitären Krisen

Ausgangslage

Krisensituationen, insbesondere nach Naturkatastrophen wie Erdbeben oder Überschwemmungen, erfordern eine schnelle und effektive humanitäre Reaktion. In solchen Fällen ist die initiale, möglichst zeitnahe Bewertung der Bedürfnisse der betroffenen Bevölkerung von entscheidender Bedeutung, um die notwendige Hilfe zu leisten und das Ausmass und den zeitlichen Rahmen für weitergehende Unterstützung zu ermitteln. Allerdings sind die Bedingungen in den ersten Stunden und Tagen nach einer Krise oft chaotisch und unklar. Es gilt, eine Fülle von Informationen aus verschiedenen Quellen auszuwerten, während der physische Zugang teilweise oder komplett verunmöglicht ist. Solche Informationen liegen in der Regel als unstrukturierte und verteilte Daten vor. Sie reichen von Social-Media-Beiträgen, die direkte Zeugenaussagen der betroffenen Bevölkerung liefern können, bis hin zu offiziellen Pressemitteilungen oder Berichten von Regierungen und NGOs, Nachrichtenartikeln der lokalen Presse sowie Daten aus Fernerkundung (u.a. Messdaten von Satelliten-, Flugzeug- und Drohnensensoren), soweit diese öffentlich zugänglich sind.

Herausforderung

Die Analyse der vorliegenden Informationen und die Extraktion relevanter Daten zur Beurteilung der Bedürfnisse stellen eine bedeutende Herausforderung dar. Die manuelle Durchsicht und Analyse dieser umfangreichen, unstrukturierten Datenmenge ist sehr zeitaufwendig und kann insbesondere aufgrund der schnellen Taktung während einer humanitären Krise unvollständig oder bereits wieder obsolet sein. Darüber hinaus ist es oft schwierig, aus den gewonnenen Informationen konkrete Schlussfolgerungen zu ziehen und die spezifischen Bedürfnisse zu bestimmen. Zudem ist die aktuelle Praxis grösstenteils reaktiv. Für vorbereitende Interventionen, die auf wahrscheinlichen Szenarien oder Trends

basieren, fehlen in der Regel die erforderlichen fundierten Informationsgrundlagen.

Lösungsansatz

Eine OSINT-Anwendung unterstützt Fachexperten und Fachexpertinnen des Eidgenössischen Departements für auswärtige Angelegenheiten (EDA), die im Ausland in Koordination mit bundesinternen und dritten Stellen humanitäre Hilfe leisten, dabei, krisenspezifische Bedürfnisse zügig und umfassend zu ermitteln, indem auf Basis von datenwissenschaftlichen Methoden und Algorithmen (insbesondere aus dem maschinellen Lernen und dem Bereich der KI, z.B. «Computer Vision» und «NLP») relevante OSINT-Daten laufend identifiziert und analysiert werden. Damit sollen Fachexperten und Fachexpertinnen und Entscheidungstragende befähigt werden, datenbasiert und zeitnah eine Lageeinschätzung anzugeben und basierend darauf Entscheidungen zu treffen.

Die Mitarbeitenden des EDA können im Fall einer humanitären Krise auf diese aufbereiteten Daten zugreifen. In einem Dashboard werden die wichtigsten Kennzahlen (sowie Trends) für die Planung der humanitären Hilfe auf Basis der OSINT-Daten dargestellt (z.B. Versorgung mit Trinkwasser, Lokalisierung der Zivilbevölkerung auf der Flucht). Die Mitarbeitenden können je nach Krise und Region die Kennzahlen anpassen. Die berücksichtigten OSINT-Daten umfassen dabei unter anderem die lokalen Medien, Social-Media-Beiträge und Berichte nationaler und internationaler Hilfsorganisationen in den relevanten Sprachen. Zur Lokalisierung von flüchtenden Menschengruppen können z.B. georeferenzierte Informationen zu Social-Media-Logins und aktuelle Satellitenbilder verwendet werden.

Die Nachvollziehbarkeit der dargestellten Informationen wird durch die Angabe der Datenhauptquelle (z.B. Bericht einer NGO) bzw. der Methode (z.B. Social-Media-Auswertung) gewährleistet. Zudem wird die Unsicherheit einer Information transparent gemacht und auf widersprüchliche Informationen hingewiesen (z.B. stark abweichende Angaben zu Flüchtenden in Berichten unterschiedlicher NGOs).

Da gerade in Krisensituationen Falschinformationen (beispielsweise aufgrund politischer Propaganda) ein grosses Risiko für die betroffenen Menschen darstellen, sind datenwissenschaftliche Mechanismen integriert (z.B. Analysen von Netzwerktopografien), die Muster erkennen, welche auf derartige Kampagnen hinweisen.

Ein wesentlicher Aspekt dieses Dashboards ist die Möglichkeit, dass der zugrundeliegende Datensatz rasch und wenn möglich in Echtzeit mit neuen Datenpunkten erweitert werden kann. Dadurch wird eine schnelle Aktualisierung der Bedürfnisanalyse gewährleistet, was insbesondere in sich schnell entwickelnden Krisensituationen von entscheidender Bedeutung ist.

Die Qualität der Datenanalyse ist von grosser Bedeutung. Hierzu stellt eine Kombination von einem KI-System und menschlicher Expertise sicher, dass die gewonnenen Erkenntnisse zuverlässig und präzise sind. Mithilfe des

menschlichen Inputs kann das System auch lernen, problematische Daten und Datenquellen, z.B. aufgrund tiefer Qualität, besser zu identifizieren.

Auf Basis der zur Verfügung gestellten Informationen können die Mitarbeitenden Szenarien und idealerweise Prognosen zu den unmittelbaren Auswirkungen der humanitären Krise entwickeln. Dieses sogenannte «nowcasting» ermöglicht beispielsweise, auf Basis von Informationen über Fluchtbewegungen und vorhandene Wasserreserven in einer Region den Wasserbedarf der kommenden Tage zu schätzen. Auf dieser Grundlage können die erforderlichen Massnahmen definiert werden. Dies unterstützt den Übergang von einer rein reaktiven zu einer proaktiven Herangehensweise bei der Bewältigung von Krisen.

Übertragbarkeit

Die Unterstützung von Fachexperten und Fachexpertinnen sowie Entscheidungstragenden mithilfe der Aufbereitung und Analyse von OSINT-Daten ist für jedes Politikfeld relevant, in dem Meinungen zu inländischen und ausländischen Gegebenheiten in grösserem Umfang eingeholt werden sollen. So könnten ähnliche Auswertungen auch als Grundlage für Inputs von Briefings politischer Entscheidungstragender dienen, z.B. zu sicherheitspolitischen Fragestellungen. Darüber hinaus könnten auch andere Politikfelder adressiert werden, bei denen es darum geht, internationale Trends zu identifizieren, z.B. in Bezug auf die Entwicklung der Klimapolitik in anderen Ländern und auf globaler Ebene.

Die Schweizer Vertretungen im Ausland könnten ebenfalls Nutzende des Dashboards sein, wobei das Konzept spezifischer ausgerichtet wäre. Es könnte auch die humanitäre Gemeinschaft, insbesondere das internationale Genf, davon profitieren.

Ein weiterer Aspekt der Übertragbarkeit besteht in der Möglichkeit, eine Verknüpfung mit internen Daten zu implementieren, möglicherweise sogar mit vertraulichen Daten. Hier sollte dann aber der Datenzugang für bestimmte Dokumente je nach nutzender Person eingeschränkt werden, um die Geheimhaltung bestimmter Daten und Informationen sicherzustellen.

Anforderungen

Daten

- Zugang zu den relevanten öffentlich verfügbaren Datenquellen.
- Berücksichtigung unterschiedlicher Sprachen, Schreibweisen, Terminologien und Abkürzungen in den Dokumenten in der Datenanalyse und bezüglich Benutzeroberfläche.
- Kontinuierliche Aktualisierung und Erweiterung der Datensätze.
- Gewährleistung der Datenqualität und Datenqualitätsmanagement.
- Effiziente Verarbeitung und Speicherung grosser Datenmengen.
- Sicherstellung der Aktualität der gesammelten Daten.
- Nachvollziehbarkeit von Meta-Daten (z.B. Geolokalisierung, Zeitstempel).

Regulierung und Ethik (in Anlehnung an DSStB-Grundprinzipien)

- Gewährleistung von Daten- und Informationsschutz und Informationssicherheit, z.B. durch Zugriffsbeschränkungen, insbesondere bei der Verwendung von Personendaten.
- Klärung der Zugangsrechte für die Anwendung.
- Sicherstellung der Nachvollziehbarkeit der Datenverarbeitung.
- Gewährleistung von Transparenz der eingesetzten Algorithmen sowie der Erklärbarkeit etwaiger Analyseergebnisse.

Technik und Infrastruktur

- Skalierbare, flexible und leistungsfähige Architektur, die mit der wachsenden Menge an Daten umgehen kann.
- Technisch und regulatorisch konforme Integration der Anwendung in die bestehende IT-Infrastruktur.
- Ausreichende, skalierbare Rechenleistung zum Trainieren von Modellen und für den Einsatz von z.B. «Natural Language Processing» (NLP).
- Bereitstellung einer verwaltungsinternen Cloud-Plattform für die Anwendung («storage» und «computing»).
- Gewährleistung des technischen Zugangs zur Anwendung für alle designierten Nutzenden, auch ausserhalb der Bundesverwaltung (z.B. Schweizerisches Korps für Humanitäre Hilfe).

Management

- Etablierung von Projektmanagement-Strukturen für die Implementierung der Anwendung, inkl. Monitoring und Evaluierung des Projekterfolgs.
- Sicherstellung der langfristigen Wartung und Weiterentwicklung der Anwendung.
- Förderung von interdisziplinärer Zusammenarbeit zwischen Nutzenden, Fachexperten und Fachexpertinnen, IT-Fachleuten und Datenwissenschaftlern und Datenwissenschaftlerinnen bei der Weiterentwicklung der Anwendung.
- Langfristige Gewährleistung von finanziellen Ressourcen für den Aufbau und Wartung der Anwendung.
- Gewährleistung der notwendigen Unterstützung und Akzeptanz des Projekts in der Organisation.

Wissen und Fähigkeiten

- Aufbau der erforderlichen Fähigkeiten im Bereich der KI und des maschinellen Lernens.
- Förderung der Aus- und Weiterbildung von Datenwissenschaftlern und Datenwissenschaftlerinnen sowie Fachexperten und Fachexpertinnen, um

die erforderlichen Fähigkeiten für diese konkrete Anwendung (weiter) zu entwickeln.

- Organisation von Schulungen und Workshops, um das Bewusstsein für die Möglichkeiten und Herausforderungen dieser konkreten Anwendung in der eigenen Organisation zu erhöhen.
- Profundes Fachwissen zur Einschätzung der Datenqualität.
- Fähigkeit, technische und analytische Möglichkeiten bzw. Konsequenzen aus der Datenqualität abzuschätzen.
- Einbindung des notwendigen Wissens aus der Forschung, insbesondere bei Projekten mit starker Relevanz für Grundlagenforschung.
- Fähigkeit der Mitarbeitenden und Entscheidungstragenden zum kritischen Umgang mit Informationen, gerade bezüglich der Güte von Daten und Datenquellen, Desinformationen sowie mangelnden oder spärlichen Informationen.
- Ausbildung der Mitarbeitenden (im Feld und in der Zentrale), um den korrekten Umgang mit dem Dashboard auch unter grossem Stress sicherzustellen.

Stakeholder

- Einbindung von Stakeholdern, Experten und Expertinnen aus verschiedenen Bereichen (humanitärer Bereich, Forschung, etc.), Nutzenden und Entscheidungstragenden in allen Prozessen der Entwicklung der Anwendung, z.B. zur Identifizierung der wichtigsten Informationsbedürfnisse / Produktanforderungen.
- Kommunikation mit der Öffentlichkeit und den betroffenen Parteien, um Bedenken und Anliegen zu adressieren und Transparenz zu gewährleisten.
- Austausch mit relevanten Stakeholdern über die Vorteile und Nutzungsmöglichkeiten der Anwendung, sowie über potenzielle Risiken.

Ökologische Nachhaltigkeit

- Festlegung und Überwachung der notwendigen Leistung des Modells und des Ressourcenbudgets.
- Definition von Massnahmen zur Kompensation von anfallenden CO₂-Emissionen.
- Identifikation, Analyse und gegebenenfalls Quantifizierung der positiven und negativen Nachhaltigkeitspotenziale der Anwendung.
- Bestimmung der Kriterien, um sicherzugehen, dass die verwendete Hardware und die genutzten Rechenzentren nachhaltig sind.
- Verwendung nachhaltiger Datensammlungs-, Speicher- und Modelltrainingspraktiken, einschliesslich Minimierung von Datenredundanz und Einsatz von Modellkompressionstechniken.

3.3 Ganzheitliche Lagebilder in Krisensituationen

Am Beispiel eines grossflächigen Schadenbebens in der Schweiz

Allgemeiner Teil

Beschreibung

In Krisensituationen, etwa bei Naturkatastrophen oder Energieengpässen, bedarf es einer schnellen und fundierten Entscheidungsfindung. Ganzheitliche Lagebilder sind in solchen Szenarien von entscheidender Bedeutung, um einen umfassenden Zugang zu allen relevanten Informationen zu gewährleisten.

Ganzheitliche Lagebilder integrieren und analysieren verschiedene Datenquellen, die normalerweise isoliert sind. Ziel ist es, ein umfassendes Verständnis der vorliegenden Situation zu erreichen und so die Entscheidungsfindung in Krisensituationen zu unterstützen. Hierbei werden unterschiedliche Lagebilder, beispielsweise Informationen über das Wetter, den Zustand der Strassen, die Strom- und Wasserversorgung, den Katastrophenschutz und die Verfügbarkeit von Personalressourcen, zusammengeführt.

Zur Verarbeitung und Analyse dieser Daten kommen Methoden aus der Datenwissenschaft zum Einsatz, einschliesslich Methoden aus der erweiterten Statistik, dem maschinellen Lernen und der KI. Mit ihrer Hilfe werden die Daten erfasst, gesammelt, aggregiert und analysiert, damit Entscheidungstragende gezielte Massnahmen ableiten können.

Für die Nutzenden stellt sich das Lagebild in Form eines Dashboards dar, das mit einem zeitlich und räumlich organisierten Modell verknüpft ist.

Um ganzheitliche Lagebilder effektiv zu erstellen, ist es notwendig, einen Plan zu haben, der bestimmt, welche Daten benötigt werden, wo sie zu finden sind, welche Formate sie haben und wie sie zusammengeführt werden können. Dabei muss oft über organisatorische Grenzen hinweg koordiniert und zusammengearbeitet werden, um eine möglichst vollständige und präzise Darstellung der Lage zu gewährleisten.

Mehrwert und Relevanz

Durch den Einsatz ganzheitlicher Lagebilder können Entscheidungstragende in Krisensituationen besser informiert agieren. Diese Form der datenbasierten Entscheidungsunterstützung trägt zu einer effizienteren Bewältigung von Krisen bei, erhöht die Sicherheit der Bevölkerung und minimiert mögliche Schäden.

Datenwissenschaft kann dabei helfen, das Potenzial von ganzheitlichen Lagebildern in Krisensituationen auszuschöpfen, gerade bei solchen, die sich nicht auf bestimmte Themenbereiche beschränken lassen und die alle Staatsebenen betreffen.

Die Datenwissenschaft erleichtert es, grosse Datenmengen schnell zu erfassen und zu sammeln, in die notwendigen Formate und Qualität zu überführen,

zusammenzuführen und schliesslich einfach verständlich darzustellen. Insbesondere unstrukturierte Daten, die oft schwer zu analysieren sind, wie Freitexte und Bilder, können z.B. mit «NLP» und «Computer Vision» schnell aufbereitet und für Entscheidungsfindungen zugänglich gemacht werden. Dieses Vorgehen ist besonders wertvoll, da es oft nicht vollständig vorhersehbar ist, welche Informationen in einer Krisensituation benötigt werden.

Weiteres Potenzial liegt in der Möglichkeit, zentrale «top-down» Informationen beispielsweise aus Geoinformationssystemen mit dezentralen «bottom-up» Informationen beispielsweise von zivilen Führungsorganen und Einsatzkräften vor Ort zu kombinieren.

Darüber hinaus können auf Basis solcher Lagebilder Szenarien und idealerweise Prognosen für unmittelbare Folgen der Krise entwickelt werden (sogenanntes «nowcasting»).

Ansätze basierend auf maschinellem Lernen, welche selbst Handlungsempfehlungen entwickeln, sind in Krisensituationen weniger geeignet. Einerseits ist gerade in Krisen eine Nachvollziehbarkeit der Entscheidungsfindung erforderlich, was in Systemen des maschinellen Lernens nur beschränkt möglich ist.

Andererseits sind solche Systeme nicht in der Lage, mit für sie neuen und einzigartigen Situationen umzugehen. Krisensituationen zeichnen sich in der Regel jedoch dadurch aus (sogenannte Stationarität).

Gerade in einer zunehmend volatilen und von Krisen geprägten Welt, in der Krisenereignisse immer häufiger auftreten, ist es von grosser Bedeutung, dass Entscheidungstragende über die besten vorhandenen Informationen verfügen, um schnell und effektiv handeln zu können.

Anwendungsbeispiel aus der Praxis

Ganzheitliches Lagebild für ein grosses Schadenbeben in der Schweiz

Ausgangslage

Ein grosses Schadenbeben (ein Erdbeben, das zu Schäden an Gebäuden und Infrastruktur führt), das grosse Teile der Schweiz getroffen hat, führte zu grossflächigen Zerstörungen mit umfangreichen Sach- und Personenschäden. Die Gemeinden im Hauptschadensraum sind weitgehend von der Aussenwelt abgeschnitten, mit blockierten Zufahrtsstrassen und unterbrochenen Kommunikationsleitungen, was die Situation noch komplizierter macht. Anhaltende Niederschläge verschärfen die Lage weiter, da sie Überschwemmungen und Erdbeben auslösen und die bereits prekäre Situation der Einwohnerinnen und Einwohner verschlimmern können. Zudem ist zum aktuellen Zeitpunkt unklar, inwieweit kritische Infrastrukturen (z.B. Staudämme, Kraftwerke, Rechenzentren) beschädigt sind.

Die sofortige Priorität ist die Sicherheit der betroffenen Bevölkerung und die Minimierung weiterer Schäden. Dafür ist es entscheidend, ein detailliertes und

aktuelles Lagebild der Situation zu erstellen, um die Notfallmassnahmen zu koordinieren und die verfügbaren Ressourcen optimal zu nutzen.

Darüber hinaus besteht ein dringender Bedarf an Szenarien und idealerweise Prognosen zu den unmittelbaren Auswirkungen eines Krisenereignisses (z.B. Beschädigung von kritischer Infrastruktur, Rutschungen). Diese Situation stellt ein komplexes und dynamisches Szenario dar, das eine ganzheitliche und datengetriebene Herangehensweise erfordert, um effektive Lösungen zu liefern.

Herausforderung

Das Erstellen ganzheitlicher Lagebilder in Krisensituationen wie einem schweren Erdbeben stellt die Behörden vor eine Vielzahl von Herausforderungen.

Erstens variieren die benötigten Informationen je nach Krisenphase und können sich im Verlauf der Krise verändern. Zu Beginn sind beispielsweise detaillierte georeferenzierte Informationen über das Gesundheitssystem, die Energieversorgung oder die Infrastruktur (vor und nach dem Ereignis) von zentraler Bedeutung. In späteren Phasen können andere Bereiche in den Vordergrund rücken, wie etwa der Wiederaufbau von Infrastrukturen oder die Wiederherstellung des Güter- und Personenverkehrs.

Zweitens sind bei der Bewältigung von Krisen viele Akteure auf allen Staatsebenen beteiligt. Diese verfügen über unterschiedliche Ressourcen, Kapazitäten und Bedürfnisse, was die Koordination und Kommunikation erschweren. Zudem sind die Daten, die für die Erstellung eines umfassenden Lagebildes benötigt werden, oft über verschiedene Institutionen und Einrichtungen verteilt und in verschiedenen Qualitäten und Formaten, auch zum Teil unstrukturiert, vorhanden.

Drittens machen die Komplexität und Dynamik der Krisensituation eine prospektive Vorbereitung und Planung oft schwierig. Die Situation kann sich schnell und unvorhersehbar verändern, und es kann sein, dass viele der benötigten Massnahmen und Lösungen erst während der Krise entwickelt und implementiert werden können.

Lösungsansatz

Als Reaktion auf die identifizierten Herausforderungen wird eine «Lagebild-Plattform» entwickelt, welche im Fall eines grossen Erdbebens in der Schweiz das notwendige Lagebild zügig bereitstellen kann. Diese Plattform definiert ausgewählte Daten sowie die technischen und organisatorischen Grundlagen.

Da jede Krise und auch jedes Erdbeben und seine Auswirkungen einzigartig sind, lässt sich der Informationsbedarf nicht abschliessend prognostizieren. Aus diesem Grund steht im Zentrum der «Lagebild-Plattform» die Modularität. Wo möglich werden bereits prospektiv der Austausch und die Aufbereitung von Daten vorbereitet, so dass diese im Fall einer Krise je nach Bedarf genutzt werden können (z.B. allgemeine Geoinformations- oder Wetterdaten). Je nach Krisenphase und -typ oder Anwender bzw. Anwenderin (z.B. Gemeindepräsi-

dentin; Feuerwehrkommandant) können die verschiedenen «Informationsschichten» aktiviert werden. Dies erfordert eine Unterscheidung zwischen akuten und mittel- bis langfristig benötigten Informationen und die Definition der relevanten Datenquellen für jede Phase.

Neben zentralen «top-down» bereitgestellten Informationsschichten wie Geoinformationssysteme können Informationen von Einsatzkräften und Privatpersonen vor Ort und frei verfügbaren Informationen aus dem Internet beispielsweise aus Social-Media (sogenannte Open Source Intelligence; OSINT) systematisch zugeschaltet werden.

Um in einer Krisensituation andere Informationsbedarfe schnell zu integrieren, die entweder nicht vorhersehbar, sehr krisenspezifisch oder für die keine allgemeinen Datengrundlagen verfügbar sind, sind die erforderlichen Ressourcen und Infrastrukturen bereitgestellt. Datenwissenschaftler und Datenwissenschaftlerinnen begleiten das Krisenmanagement. Sie helfen, die notwendigen Datenquellen zu identifizieren sowie die Daten aufzubereiten, zu plausibilisieren und in das Lagebild zu integrieren. Sie werden durch datenwissenschaftliche Anwendungen unterstützt (z.B. für die Strukturierung von Daten oder deren Plausibilisierung). Ein mögliches Beispiel für einen solchen krisenspezifischen, akuten Informationsbedarf kann die Auswirkung einer durch das Erdbeben zerstörten Verkehrsinfrastruktur in einem Nachbarland auf die Gesundheitsversorgung in einem grenznahen Kanton sein. Zudem können, dort wo «echte Daten» fehlen, datenwissenschaftliche Methoden (einschliesslich Methoden aus der erweiterten Statistik, dem maschinellen Lernen und dem Bereich der KI) helfen, diese Lücken durch Simulationen (z.B. von komplexen Infrastrukturen mit kaskadierenden Effekten) und/oder die Generierung synthetischer Daten zu schliessen. Beim Einsatz synthetischer Daten ist die Nachvollziehbarkeit und Erklärbarkeit der Datengenerierung zu gewährleisten.

Für die Nutzenden bietet die «Lagebild-Plattform» ein benutzerfreundliches Dashboard, das individuell angepasst werden kann. Damit können Anwender und Anwenderinnen das Lagebild an ihre spezifischen Bedürfnisse und die spezifische Krisensituation oder -phase anpassen. Auf diese Weise kann eine lokal agierende Feuerwehrkommandantin wie auch ein national aufgestellter Krisenstab die jeweils für sich relevanten Informationsschichten zuschalten und die Darstellung adjustieren (z.B. mithilfe von «Widgets»). Die Visualisierung der eingespeisten Daten als Dashboard verknüpft mit einem zeitlich und räumlich organisierten Modell ermöglicht es, komplexe Informationen auf einfache und verständliche Weise darzustellen.

Ein wichtiger Aspekt in Krisensituation ist die Nachvollziehbarkeit der dargestellten Informationen und deren Verlässlichkeit. Sie wird durch die Angabe der Datenhauptquelle (z.B. öffentliche Statistik, Polizeibericht) bzw. der Methode (z.B. Social-Media-Auswertung) gewährleistet. Zudem wird die Unsicherheit einer Information transparent gemacht und auf widersprüchliche Informationen hingewiesen (z.B. stark abweichende Angaben zum Ausmass von Gebäudeschäden in unterschiedlichen Quellen).

Neben der Darstellung der aktuellen IST-Situation erlaubt die Plattform die Entwicklung von Szenarien und idealerweise Prognosen zu den unmittelbaren Auswirkungen eines Krisenereignisses. Dieses sogenannte «nowcasting» ermöglicht beispielsweise auf Basis von Informationen über die Stärke und Verteilung des Erdbebens und georeferenzierten Modellen von Gebäuden und Infrastrukturen ein detailliertes Bild über wahrscheinliche Schäden. Darauf aufbauend können die Lagebilder zeitnah Hinweise für Entscheidungstragende liefern (z.B. zur Abschaltung von Ventilen für Gasleitungen in bestimmten Regionen).

Übertragbarkeit

Die Übertragbarkeit dieses Ansatzes auf andere Krisensituationen ist weitreichend, vorausgesetzt, die rechtlichen Rahmenbedingungen für die Zusammenführung und Verwendung der notwendigen Daten sind gegeben.

Eine weitere Voraussetzung für die Übertragbarkeit dieses Ansatzes auf andere Situationen ist die Verfügbarkeit einer generischen, universellen Lösung. Diese sollte kompatibel mit verschiedenen Schnittstellen sein, da in Krisensituationen eine Vielzahl von Akteuren, wie Kantone, Gemeinden oder Hilfsorganisationen, sowohl Daten liefern als auch nutzen. Dabei sollte sie flexibel genug sein, um an spezifische Anforderungen angepasst zu werden, ohne dass für jede Krisensituation eine komplett neue Lösung entwickelt werden muss. Ausgangsbasis könnte die digitale Landkarte der Schweiz sein (<https://test.map.geo.admin.ch>), bei welcher beliebig viele zusätzliche Schichten georeferenziert aufgeschaltet werden könnten.

Dieser Ansatz ist besonders relevant für Situationen, in denen umfassende Lagebilder rasch zur Verfügung stehen müssen, wie etwa bei Pandemieausbrüchen, Kriegsfällen, terroristischen Anschlägen oder Notfallsituationen aufgrund extremer Wetterereignisse. Diese Ereignisse können aufgrund geopolitischer oder klimatischer Veränderungen in Zukunft sowohl in ihrer Häufigkeit als auch in ihrer Intensität zunehmen. In all diesen Fällen können die Integration und Analyse verschiedener Datenquellen zu fundierteren Entscheidungen und effektiverem Handeln führen.

Anforderungen

Daten

- Zusammenführung verschiedener Datenquellen, wie Bilddaten (Situation), Wetterinformationen (Gefährdung, «hazard»), Verkehrsdaten, Infrastruktur (Exponierung, «exposure»), Katastrophenschutz und Personalressourcen (Massnahmen) sowie Informationen zur spezifischen Verletzlichkeit («vulnerability»).
- Sicherstellung der Aktualität, Genauigkeit, Vollständigkeit und Rückverfolgung der Daten, um präzise, verlässliche und überprüfbare Entscheidungsgrundlagen zu schaffen.

- Vorbereitung vorhersehbar notwendiger Daten (Schaffung der Datengrundlage als wichtiger Teil der Vorbereitung für Krisensituationen, z.B. als Aufgabenbereich des Krisenstabs).
- Bereitstellung umfassender Metadaten.
- Übersicht über vorhandene Daten (d.h. Metadatenkatalog).

Regulierung und Ethik (in Anlehnung an DSStB-Grundprinzipien)

- Festlegung klarer Regeln und Zuständigkeiten für den Zugriff auf und die Verwendung von Daten in Krisensituationen.
- Gewährleistung der Privatsphäre und Sicherheit der Daten und Informationen, insbesondere im Hinblick auf persönliche Informationen und sensible Infrastrukturdaten.
- Sicherstellung einer fairen und diskriminierungsfreien Datenverarbeitung in Krisensituationen.
- Gewährleistung von Transparenz und Nachvollziehbarkeit der Algorithmen und Ergebnisse.

Technik und Infrastruktur

- Entwicklung einer zentralen, benutzerfreundlichen und sicheren Plattform, um Daten und Datenanalysen (inkl. Modellierungen) bereitzustellen und zu visualisieren.
- Schaffung einer robusten und flexiblen Infrastruktur, die in Krisensituationen schnell an veränderte Bedingungen angepasst werden kann.
- Implementierung starker Sicherheitsmassnahmen, um die Plattform und die darin enthaltenen Daten vor unbefugtem Zugriff und Cyber-Angriffen zu schützen.
- Schaffung von Kanälen im IT-System, welche die rasche Datenzusammenführung in Krisenzeiten, aber auch die Isolationen der Daten in Normalzeiten sicherstellt.
- Adressatengerechte Zugänglichkeit für verschiedene Akteure mit jeweils relevanten Informationen (z.B. Rettungsdienste vor Ort, Regierungsrat oder Bundesrat) beispielsweise über Schnittstellen in bestehende Systeme oder dedizierte Applikationen.
- Gewährleistung computerlesbarer Datenschnittstellen (APIs, «Application Programming Interfaces»). Dies erlaubt auch die Einbindung in andere/ergänzende Systeme und ermöglicht und fördert die Interoperabilität.
- Bereitstellung der datenwissenschaftlichen Instrumente, um in Krisensituationen rasch die notwendigen Daten (oftmals in unterschiedlichen Formaten) in das Lagebild integrieren zu können.

Management

- Etablierung klarer Prozesse (inkl. Aufgaben, Kompetenzen und Verantwortlichkeiten) für die konkrete Krisensituation (z.B. «AKV» oder «RACI»)

Matrizen⁹⁾ sowie allfälliger Nutzungs- und Verrechnungsmodelle, so dass die verschiedenen Akteure und Datenquellen effektiv gesteuert und koordiniert werden.

- Sicherstellung einer klaren, zeitnahen und zielgruppengerechten Kommunikation von Informationen und Entscheidungen.
- Implementierung eines Feedback- und Lernsystems, um aus Erfahrungen in Krisensituationen zu lernen und den Ansatz laufend zu optimieren und Unsicherheiten im Lagebild zu reduzieren.
- Enge Abstimmung und Integration in die Strukturen und Prozesse zum Krisenmanagement von Bund, Kantonen und Gemeinden sowie weiteren relevanten Akteuren.
- Mechanismus, welcher in Krisensituationen die richtigen Personen an den Tisch bringt, gerade im Bereich Datenwissenschaft, um flexibel die notwendigen Modelle Schicht für Schicht zu bauen und weiterzuentwickeln sowie Schnittstellen zu schaffen und Datenanalysen durchzuführen.
- Bereitstellung der notwendigen personellen und finanziellen Ressourcen in den Bereichen Datenwissenschaft, Problemanalyse und Krisenmanagement, sowohl in der Erstellung des Systems wie auch für die operative Nutzung in Krisensituationen.
- Ein Denkansatz, der anerkennt, dass aktuelle Herausforderungen grundsätzlich auf einer interdisziplinären Ebene, die traditionelle Politikfelder sowie Fachbereiche und Datenwissenschaften überschreitet, bewältigt werden sollten.
- Prozesse und Mitarbeitende zur raschen Übersetzung von politischen Informationsbedarfen in konkrete datenwissenschaftliche Anforderungen / Aufgaben («Requirements Engineering»).

Wissen und Fähigkeiten

- Verbesserung der Kompetenzen im Umgang mit unvorhersehbaren Situationen, in denen nicht sofort klar ist, welche Daten benötigt werden, wie z.B. die Bewältigung von unvollständigen Datensätzen durch den Einsatz von «Proxies».
- Fähigkeit der Mitarbeitenden zum Umgang mit unsicheren Informationen, gerade bezüglich der Güte von Daten und Datenquellen (z.B. Social-Media-Auswertungen), inkl. Abwägung zwischen schnellem Handeln und notwendiger Güte für Entscheidungen.
- Förderung von interdisziplinärem Wissen und Fähigkeiten (bezüglich z.B. sowohl Datenwissenschaft wie auch Umweltwissenschaft), um fachbereichsübergreifende Herausforderungen (wie sie oft in Krisensituationen auftauchen) effektiver anzugehen.
- Regelmässige Schulungen und Übungen für Entscheidungstragende und Fachpersonal, um den Umgang mit der Plattform und den Daten zu schulen (inkl. generelles digitales Upskilling).

⁹⁾ AKV: Aufgaben, Kompetenzen und Verantwortlichkeiten; RACI: Responsible, Accountable, Consulted and Informed

- Wissen und Fähigkeiten, um in Krisensituationen rasch Daten mit unterschiedlichen Formaten in das Lagebild zu integrieren.

Stakeholder

- Aktive Beteiligung von Stakeholdern, wie Behörden, Energieversorgern, Verkehrsbetrieben und Rettungsdiensten und etwaigen Betroffenen in Bevölkerung und Wirtschaft, bei der Entwicklung und Umsetzung des Ansatzes («Co-Design and -Development»). Insbesondere Übungen (siehe unter «Wissen & Fähigkeiten») bieten auch Gelegenheit, Verbesserungen und Prototypen unter «Quasi-Echtbedingungen» zu testen.
- Schaffung von Bewusstsein und Verständnis für die Bedeutung von Datenintegration und -analyse in Krisensituationen bei allen relevanten Akteuren.

Ökologische Nachhaltigkeit

- Auswahl von energieeffizienten Servern und Rechenzentren, um den Energieverbrauch und CO₂-Emissionen zu minimieren.
- Verwendung nachhaltiger Datensammlungs- und Speicherpraktiken, einschliesslich der Minimierung von Datenredundanz und dem Einsatz von Datenkompressionstechniken.

3.4 Datenbasierte Unterstützung von Aufsicht

Am Beispiel der Luftfahrtbranche

Allgemeiner Teil

Beschreibung

Die Bundesverwaltung führt im Rahmen ihrer Aufgaben und Aufsichtszuständigkeit eine Reihe von Tätigkeiten durch, sei es die Entwicklung, Überwachung und Durchsetzung von Gesetzen und Vorschriften wie auch die Ausstellung von Genehmigungen und Lizenzen. Zur Erfüllung dieser Aufsichtsaufgaben werden unterschiedlichste Daten und Informationen aus vielfältigen Quellen generiert, erhoben, gesammelt, verarbeitet und analysiert. Datenwissenschaftliche Methoden können in verschiedenen Schritten dieses Prozesses eingesetzt werden, um die Bewältigung der umfangreichen, heterogenen Datenmengen zu erleichtern und dadurch die Aufsicht effizienter und effektiver zu gestalten.

Mehrwert und Relevanz

Die Datenwissenschaft kann einen wesentlichen Beitrag zur Unterstützung des Aufsichtspersonals der Bundesverwaltung leisten. Die zunehmend dynamische technologische und gesellschaftliche Entwicklung (mit wachsenden Datenmengen als Folge) akzentuiert ihr Potenzial.

Erstens können datenwissenschaftliche Methoden durch die Unterstützung der Datensammlung und -validierung dazu beitragen, die Aufsicht auf eine qualitativ hochwertige Datenbasis zu stellen und das Aufsichtspersonal wesentlich zu entlasten. Z.B. können solche Methoden eingesetzt werden, um unstrukturierte elektronische Dokumente wie PDF-, XLS- oder Word-Dokumente in strukturierte Datenformate zu überführen oder die Datensammlung selbst weitgehend zu automatisieren. Durch die Validierung der generierten bzw. gesammelten Daten mithilfe datenwissenschaftlicher Methoden können z.B. Anomalien zügig erkannt werden, die entweder auf Fehler in den Daten oder neue Entwicklungen im Aufsichtsbereich hinweisen.

Zweitens unterstützt die Datenwissenschaft die Durchführung der Aufsicht selbst. Durch die Analyse und Kombination einer Vielzahl von Indikatoren ermöglicht sie eine fundierte Beurteilung der Risiken und Leistungsqualität beaufsichtigter Organisationen bzw. Bereiche. Auf dieser Basis kann die Aufsichtsplanung risiko- oder leistungsorientiert erfolgen. So können z.B. besonders risikoreiche Bereiche priorisiert werden, wodurch die Aufsicht und der Ressourceneinsatz optimiert werden. In weniger komplexen, risikoarmen Fällen können Systeme basierend auf KI bestimmte Schritte der Aufsicht sogar automatisch durchführen (z.B. Verlängerung von Bewilligungen). Auch bei der Aufsichtsplanung können Prozessschritte mithilfe solcher Systeme automatisiert und stetig verbessert werden.

Schliesslich dient die Datenwissenschaft auch der Verbesserung der regulatorischen Grundlagen der Behörden. So kann maschinelles Lernen durch die

Auswertung grosser Datenmengen und Erfahrungswerten dazu beitragen, Regulationsvorgaben kontinuierlich zu verbessern und an veränderte und zusammenhängende Rahmenbedingungen anzupassen.

Anwendungsbeispiel aus der Praxis

Aufsicht der Luftfahrtsicherheit

Ausgangslage

Das Bundesamt für Zivilluftfahrt (BAZL) ist für die Aufsicht über die zivile Luftfahrt in der Schweiz zuständig und trägt die Verantwortung für eine nachhaltige Entwicklung der Luftfahrt. Damit leistet es einen wichtigen Beitrag für die Sicherheit im Bereich der Zivilluftfahrt und stimmt sich eng mit der militärischen Luftfahrtbehörde (MAA) ab.

Zur Erfüllung seiner Aufsichtspflicht ist das BAZL auf qualitativ hochwertige und zuverlässige Daten zahlreicher Akteure in der Luftfahrt angewiesen (u.a. Anbieter von Luftfahrtinformations-, Flugnavigations- und Flugverkehrsdiensten (Skyguide, MeteoSchweiz, Engadin Airport), Flugplätzen, Airlines, Piloten, ausländische Behörden, internationale Organisationen). Diese Daten bilden unter anderem die Grundlage für sichere Flugoperationen oder die Anweisungen an Piloten und sind erfolgskritisch für eine sichere und effiziente Luftfahrt.

Um die hohen Standards der Luftfahrtsicherheit zu gewährleisten, muss das BAZL seine Ressourcen gezielt und effizient einsetzen, insbesondere in Bereichen, in denen das Potenzial zur Verbesserung der Flugsicherheit besteht.

Herausforderung

Die Aufsicht im Bereich der Luftfahrtsicherheit steht vor verschiedenen Herausforderungen.

Eine der Schwierigkeiten liegt in der Heterogenität und Qualität der erhaltenen Daten. Die Datensätze, die aus verschiedenen Quellen und in unterschiedlichen Formaten gesammelt werden, weisen bezüglich ihrer Qualität mitunter erhebliche Unterschiede auf. Derartige Unstimmigkeiten und Mängel müssen nachträglich behoben werden, um eine gleichbleibend hohe Datenqualität sicherzustellen. Diese Aufgabe bindet jedoch wertvolle Ressourcen. Die starke Heterogenität kann auch dazu führen, dass Datenfehler in den Aufsichtsprozess einfließen und diesen sogar negativ beeinträchtigen können.

Heutzutage erfolgt die Sammlung von Indikatoren manuell in Excel-Tabellen und Access-Datenbanken, welche ebenfalls manuell als Prioritätenlisten zur Bestimmung der Aufsichtstätigkeit aktualisiert werden. Diese Vorgehensweise weist mehrere Schwachstellen auf. Zum einen ist sie äusserst zeitaufwendig, da jede Information manuell eingegeben und zusammengeführt werden muss. Zum anderen ist sie anfällig für menschliche Fehler, wie etwa falsche Datenzusammenführungen oder Fehleingaben. Diese Mängel können sowohl die Qualität der gesammelten Daten als auch die darauf basierenden Entscheidungen beeinträchtigen. Darüber hinaus erschwert die manuelle Datenverarbeitung das Erkennen von Mustern und Zusammenhängen, beispielsweise wiederkeh-

rende Wartungsprobleme oder systematische Verfehlungen bei der Durchführung vorgeschriebener Verfahren. Für die kontinuierliche Verbesserung der Flugsicherheit können geeignete Werkzeuge und Methoden zur Datenanalyse entscheidend sein, damit solche wertvollen Erkenntnisse nicht unentdeckt und ungenutzt bleiben.

Lösungsansatz

Die Anwendung der Datenwissenschaft kann in jedem Aspekt der Datenverwendung im Rahmen der Luftfahrtsicherheitsaufsicht eine wesentliche Verbesserung der derzeitigen Aufsichtsprozesse bewirken. Voraussetzung dafür ist eine ausreichend hohe Güte der Daten, Datenquellen und datenwissenschaftlichen Methoden, welche es zu gewährleisten gilt.

Das BAZL hat das Potenzial bereits erkannt und entwickelt aktuell den sogenannten «Data Collection Service» (DCS). Dies ist ein typisches Beispiel einer digitalen Transformation mit dem Ziel, die Datensammlung in der geforderten Datenqualität digital sicherzustellen und die Luftfahrt Daten im Kontext von weiteren Daten, Informationen und regulatorischen Anforderungen zu validieren. Durch automatische Prüfungen bezüglich regulatorischer und technischer Anforderungen sollen Daten medienbruchfrei übermittelt, effizient aufbereitet und Beurteilungen aus verschiedenen Luftfahrtkreisen zusammengeführt werden. Das System soll schliesslich die Grundlage für regulatorische Entscheidungen zur Erteilung von Genehmigungen für die anschliessende Datenpublikation bilden. Ein besonderes Augenmerk wird dabei auf die Vereinfachung und Automatisierung der Prozesse gelegt, damit die Arbeitsprozesse aller Akteure medienbruchfrei durchgeführt und die Datenqualitätsanforderungen in jedem Fall garantiert werden können.

Auf Basis des DCS können weitere datenwissenschaftliche Anwendungen zur Verbesserung der Aufsicht und der Qualität der Luftfahrt Daten lanciert werden. Erstens kann eine Anomalie-Erkennung das Aufsichtspersonal bei der Freigabe der Daten unterstützen. Darüber hinaus können KI-Systeme dazu beitragen, Bereiche zu identifizieren, in denen die Wahrscheinlichkeit einer niedrigeren Datenqualität höher ist. Dies ermöglicht es, den Fokus auf diese Bereiche zu legen und die Qualität und Zuverlässigkeit der Daten zu erhöhen. Die Datengenerierenden erhalten anschliessend Feedback aus dem System zu Fehlern und Fehlerquellen, um so schrittweise die Leistungsqualität zu verbessern.

Der zweite Bereich ist die Analyse der Risiken und die Bewertung der Leistungsqualität der datengenerierenden Stellen. Hintergrund sind zwei Aufsichtsprinzipien, die in der Luftfahrtbranche üblicherweise angewendet werden: die «risikobasierte Aufsicht» (Risk Based Oversight, RBO) und die «leistungsbaasierte Aufsicht» (Performance Based Oversight, PBO). Die Aufsichtstätigkeiten werden demnach proportional zum Umfang des Risikos, das als Folge einer unzureichenden Einhaltung der Vorschriften entstehen könnte, ausgestaltet. Bei der risikobasierten Aufsicht wird der Fokus der Aufsichtstätigkeiten auf Bereiche mit hohem Risiko gelegt. Die leistungsorientierte Aufsicht basiert auf Leistungsindikatoren, wie z.B. die Bewertung und das Benchmarking der Leis-

tung verschiedener Organisationen bezüglich der Datenlieferungen, dem Umgang mit Sicherheitsrisiken oder der Reaktionszeiten bei der Umsetzung von Abhilfemassnahmen bis zur Auswertung. Um die beiden Aufsichtsansätze zu ergänzen und um etwaige Routinen, die sich einstellen könnten, zu durchbrechen, werden zudem zufällige Untersuchungen durchgeführt. Datenwissenschaftliche Anwendungen unterstützen durch die Analyse einer Vielzahl von Indikatoren eine fundierte, komplexe Beurteilung der Risiken und Leistungsqualität von beaufsichtigten Organisationen bzw. Bereichen. Auf dieser Basis kann die Aufsichtsplanung zielgenauer erfolgen und der Ressourceneinsatz optimiert werden.

Drittens kann die Datenwissenschaft auch zur Entwicklung eines Aufsichtsplans genutzt werden. Dabei wird eine Vielzahl von Kriterien berücksichtigt, wie die Verfügbarkeit und Kompetenz des Aufsichtspersonals, die Komplexität der Aufsichtsmaßnahme und die geografische Lage des betroffenen Objektes, um die Aufsicht wirksam und effizient zu planen. Es ist auch möglich, aufgrund von Vorfällen oder unvorhergesehenen Ereignissen, Last-Minute-Anpassungen im Planungssystem vorzunehmen. Durch den Einsatz von lernbasierten Systemen kann aus vergangenen Aufsichtstätigkeiten gelernt und die Planung entsprechend den gewonnenen Erkenntnissen angepasst werden.

Schliesslich kann die Datenwissenschaft dazu beitragen, Vorschriften kontinuierlich weiterzuentwickeln, und unter anderem Grenzwerte in Vorschriften optimal festzulegen. Beispielsweise können durch die Auswertung grosser Datenmengen und Erfahrungswerten des Aufsichtspersonals auf Basis von maschinellem Lernen die Ausgestaltung und Wirkung von Regulierungen evaluiert werden (z.B. die Anpassung der maximal zulässigen Höhe von Luftfahrthindernissen oder die Reduktion der Bewilligungspflicht (damit auch der Einzelprüfung) von Objekten einer tieferen Gefährdungsstufe für die Luftfahrt).

Übertragbarkeit

Die Übertragbarkeit dieses datenwissenschaftlichen Ansatzes auf andere Politikfelder ist vielfältig und vielversprechend. In nahezu allen staatlichen Handlungsbereichen besteht ein Bedarf, die Entscheidungsfindung und den Vollzug im Bereich der Aufsicht zu verbessern und Rechtsvorschriften auf eine rationale, evidenzbasierte Weise zu optimieren. Besonders in Bereichen, in denen komplexe und umfassende Datenauswertungen erforderlich sind, um regulatorischen Anforderungen zu entsprechen, bevor sie genehmigt werden können, sind datenwissenschaftliche Methoden von grossem Wert für eine rationale Informationsaufbereitung. Ein ähnlicher datenwissenschaftlicher Ansatz könnte beispielsweise im Energiebereich eingesetzt werden, um neue Technologien und Verfahren zu prüfen und zu genehmigen, bevor sie implementiert werden.

Anforderungen

Daten

- Zugang zu relevanten und qualitativ hochwertigen Datenquellen.
- Etablierung von Datenstandards und -protokollen für eine konsistente und wo möglich automatisierte Datensammlung in einem einfachen, maschinenlesbaren Format.

Regulierung und Ethik (in Anlehnung an DSStB-Grundprinzipien)

- Gewährleistung von Daten- und Informationsschutz und Informationssicherheit.
- Sicherstellung der notwendigen gesetzlichen Grundlage für autonome Entschiede basierend auf datenwissenschaftlichen Algorithmen und Methoden aus dem maschinellen Lernen und dem Bereich der KI.
- Sicherstellung der Transparenz von datenwissenschaftlichen Algorithmen und Methoden, die in der Analyse verwendet werden, und deren Prüfung auf mögliches diskriminierendes Verhalten, um Vertrauen bei den beteiligten Stakeholdern aufzubauen und Diskriminierung zu vermeiden.
- Entwicklung von Mechanismen, welche nachvollziehbar und erklärbar machen, wie die Outputs des Analysesystems entstehen (z.B. Transparenz in Prozessen).
- Gesetzliche Grundlage für Weisungen an Datengenerierende / -liefernde bezüglich Daten, Datenformat, Übermittlungsfrequenz und -kanal, technischen Standards, Prozessen der Qualitätssicherung und Plausibilisierung etc. (entweder in den jeweiligen Fachgesetzen oder dedizierten Gesetzen).
- Etablierung klarer Verantwortlichkeiten für den gesamten Prozess der Plausibilitätsprüfung.

Technik und Infrastruktur

- Skalierbare, flexible und leistungsfähige Architektur, die mit der wachsenden Menge an Daten umgehen kann.
- Bereitstellung der notwendigen Cloud-Infrastruktur.
- Bereitstellung einer zentralisierten Datenplattform für die nutzerfreundliche Eingabe, Speicherung und Verarbeitung von Daten, inkl. einer notwendigen Datenaustauschinfrastruktur mit definierten, nutzerfreundlichen Schnittstellen.
- Ausreichende, skalierbare Rechenleistung zum Trainieren von datenwissenschaftlichen Algorithmen und Methoden.
- Gestaltung neuer Prozesse im Rahmen der Aufsichtstätigkeiten unter Berücksichtigung bestehender und potenzieller Datengrundlagen und relevanter Kapazitäten (z.B. Rechenleistung), um das Potenzial von Daten und Datenwissenschaft optimal auszuschöpfen.
- Tools und Schnittstellen zur Unterstützung der Datengenerierenden und -liefernden in der Datengenerierung und -Lieferung

Management

- Exakte Beschreibung des Zwecks und der Aufgaben im Rahmen der Aufsicht, für welche Daten gesammelt und verarbeitet werden (inkl. Kommunikation an Beaufsichtigte).
- Etablierung einer effektiven Gouvernanz-Struktur für die Implementierung und Nutzung von Datenwissenschaft in der Aufsicht.
- Förderung des Verständnisses und der Akzeptanz von datenwissenschaftlichen Methoden zugunsten eines moderneren Verwaltungsmanagements, das sich an den zunehmend dynamischeren Bedürfnissen der Gesellschaft orientiert.
- Schaffung von Synergien und Kooperationen zwischen verschiedenen Aufsichtsbehörden und -bereichen.
- Entwicklung von Strategien zur kontinuierlichen Verbesserung und Anpassung der datenwissenschaftlichen Algorithmen und Methoden.

Wissen und Fähigkeiten

- Aufbau von Kompetenzen und Fachwissen im Bereich Datenwissenschaft und KI innerhalb der Aufsichtsbehörden, z.B. durch Implementierung von Schulungs- und Weiterbildungsmaßnahmen für Mitarbeitende im Umgang mit Datenwissenschaft und KI.
- Förderung des Wissensaustauschs und der Zusammenarbeit zwischen Datenwissenschaftlern und Datenwissenschaftlerinnen sowie Fachexperten und Fachexpertinnen.
- Aufbau und Stärkung der Kapazitäten, digitale Transformationen zu begleiten, u.a. Fähigkeiten in den Bereichen Innovation und Change-Management.

Stakeholder

- Einbindung von Stakeholdern aus der Industrie, der Wissenschaft und militärischen Luftfahrt (Luftwaffe, LW und Military Aviation Authority, MAA) in die Entwicklung und Anwendung von datenwissenschaftlichen Aufsichtsansätzen für die zivile Luftfahrt (Gesellschaft).
- Förderung des Verständnisses und der Akzeptanz von datenwissenschaftlichen Methoden der Aufsicht bei den betroffenen Stakeholdern.

Ökologische Nachhaltigkeit

- Verwendung von energie- und ressourceneffizienten Technologien und Infrastrukturen, um den Energieverbrauch und ökologischen Fussabdruck zu minimieren.
- Einsatz von umweltfreundlichen Speicherlösungen, wie Cloud-Speicherdiensten auf Basis erneuerbarer Energien oder umweltschonenden physischen Speichermedien.

3.5 Ganzheitliche Modellsysteme zur Entscheidungsunterstützung

Am Beispiel von Ernährungsnachhaltigkeit

Allgemeiner Teil

Beschreibung

Computerbasierte, ganzheitliche Modellsysteme sind integrierte Softwaresysteme, die eine Fülle von Informationen erfassen, zusammenführen, analysieren und abbilden. Sie kombinieren Daten aus Sensoren, historischen Datenbanken und Expertenwissen, um ein umfassendes Verständnis eines untersuchten Systems zu erlangen. Sie berücksichtigen ökonomische, ökologische, soziale und technische Aspekte sowie deren Wechselwirkungen, um die Komplexität des untersuchten Systems angemessen abzubilden.

Als solche liefern sie Prognosen, Simulationen und Analysen, die auf verschiedenen Datenebenen und -quellen basieren. Durch das Bewerten einer Vielzahl von Szenarien und Handlungsoptionen unterstützen sie Entscheidungsstragende, die möglichen Auswirkungen auf unterschiedliche Aspekte eines Systems abzuschätzen.

Es gibt zahlreiche Anwendungsfälle für diese Systeme, darunter das Umweltmanagement und die Risikobewertung. Wichtig ist dabei zu betonen, dass ganzheitliche Modellsysteme sich von digitalen Zwillingen und ganzheitlichen Lagebildern unterscheiden. Sie beschränken sich nicht nur auf die virtuelle Repräsentation von realen Objekten oder die Darstellung der aktuellen Situation. Vielmehr beziehen sie auch abstrakte Elemente wie theoretische mathematische Modelle, sozioökonomische Faktoren und kulturelle Aspekte ein, um ein umfassenderes Verständnis für komplexe Systeme zu erlangen und dynamische Simulationen durchzuführen. Weitere mögliche Bestandteile solcher Systeme sind Algorithmen und Methoden aus dem Bereich des maschinellen Lernens und der KI. Durchaus kann beispielsweise auch das Verhalten eines digitalen Zwillings einer von mehreren Faktoren in einem ganzheitlichen Modellsystem sein.

Mehrwert und Relevanz

Der Einsatz von ganzheitlichen Modellsystemen kann in der Politikgestaltung erhebliche Vorteile bieten, gerade im Kontext komplexer Herausforderungen in aussergewöhnlichen Situationen oder für lange Zeithorizonte.

Die erste Stärke dieser Systeme liegt in ihrer Fähigkeit, Entscheidungsprozesse auf eine solide Datenbasis zu stellen. Anstatt isolierte Datenpunkte zu betrachten, integrieren ganzheitliche Modellsysteme eine Fülle relevanter Informationen aus verschiedenen Quellen. Sie extrahieren und analysieren diese Daten und ermöglichen so eine holistische Betrachtung der Problemsituation.

Zweitens ermöglichen ganzheitliche Modellsysteme ein besseres Verständnis komplexer Zusammenhänge und Wechselwirkungen. Sie veranschaulichen die Interaktionen zwischen verschiedenen Faktoren innerhalb des betrachteten Systems. Mit einem detaillierten Verständnis dieser Zusammenhänge können

Entscheidungstragende geeignete Massnahmen ergreifen und unerwünschte Nebenwirkungen minimieren.

Der dritte Mehrwert von ganzheitlichen Modellsystemen liegt in ihrer Fähigkeit, verschiedene Szenarien zu simulieren und somit Entscheidungstragende zu unterstützen, die Wirksamkeit von Massnahmen im Voraus einzuschätzen. Anstatt lediglich historische Daten und Erfahrungen darzustellen und mit extrapolierten Prognosen zu erweitern, können diese Systeme mögliche zukünftige Szenarien abbilden und die Auswirkungen verschiedener politischer Optionen bewerten. Sie bieten so eine wichtige Entscheidungsgrundlage, indem sie Risiken und potenzielle sowie unerwartete Auswirkungen von Politikmassnahmen frühzeitig identifizieren.

Schliesslich fördern ganzheitliche Modellsysteme eine kontinuierliche Anpassung und Verbesserung von Entscheidungsprozessen. Sie können lernfähig sein und Rückmeldungen, u.a. im Rahmen von menschlichen Feedback-Loops, und Erfahrungen aus früheren Entscheidungen aufnehmen und in zukünftige Simulationen und Analysen integrieren. Dadurch tragen sie zur kontinuierlichen Verbesserung der Entscheidungsqualität bei und helfen Entscheidungstragenden, den sich ständig ändernden Anforderungen und Rahmenbedingungen gerecht zu werden.

Anwendungsbeispiel aus der Praxis

Ganzheitliches Modellsystem für Ernährungsnachhaltigkeit

Ausgangslage

Die Gewährleistung der politischen Rahmenbedingungen für eine sichere Versorgung der Bevölkerung mit Nahrungsmitteln ist eine der grundlegendsten Aufgaben jedes Staates. In der heutigen Zeit sind dazu unterschiedliche Anforderungen zu erfüllen. Einerseits muss die Nahrungsmittelversorgung der Bevölkerung zu jedem Zeitpunkt gewährleistet werden können, auch in Situationen, in welchen es zu Ertragsausfällen oder eingeschränkten Importmöglichkeiten gekommen ist. Andererseits ist es entscheidend, den Agrar- und Ernährungssektor langfristig nachhaltig zu gestalten, um nicht nur den aktuellen, sondern auch den zukünftigen Bedarf an Nahrungsmitteln zu decken, und dabei gleichzeitig negative Auswirkungen auf die Umwelt oder die Gesellschaft zu minimieren und eine gesunde Ernährung zu ermöglichen. Diese Kombination aus kurz- und mittelfristiger Sicherstellung der Nahrungsmittelversorgung und langfristiger ökologischer, ökonomischer und sozialer Nachhaltigkeit des Agrar- und Ernährungssektors bildet den Kern dessen, was wir hier als «Ernährungsnachhaltigkeit» verstehen.

Zur Unterstützung der Gewährleistung der Versorgungssicherheit hat das Bundesamt für wirtschaftliche Landesversorgung (BWL) in Zusammenarbeit mit Agroscope, dem Schweizer Kompetenzzentrum für landwirtschaftliche Forschung, ein Modellsystem der Land- und Ernährungswirtschaft entwickelt. Dieses Modellsystem bildet die produktionstechnischen Zusammenhänge zwischen Agrarflächen, Nutztierbeständen, Lebensmittelverarbeitung, Lagerhal-

tung, Aussenhandel und Ernährungsbedarf ab. Ziel des Systems ist es, für spezifische Szenarien schwerer Mangellagen zu ermitteln, welche Art von Nahrungsmittelproduktion in der Schweiz den Energie- und Nährstoffbedarf der Bevölkerung am besten decken würde.

Das Modellsystem wird derzeit weiterentwickelt. Es wird zu einem umfassenden Nachhaltigkeitsmodell erweitert, das ökonomische, ökologische und soziale Indikatoren integriert. Dieses erweiterte Modell ist Teil des Agroscope-Projekts «Agri-food economy» und zielt darauf ab, eine noch ganzheitlichere Betrachtung des Agrar- und Ernährungssektors zu ermöglichen, um die Resilienz und Nachhaltigkeit zukünftiger schweizerischer Ernährungssysteme beurteilen zu können.

Herausforderung

Die Komplexität, Dynamik und Verwundbarkeit der Agrar- und Ernährungswirtschaft haben zugenommen, getrieben durch Faktoren wie veränderte Essgewohnheiten, ökologische Bedenken, Handelsbeschränkungen, volatile Lieferketten (z.B. Dünger aus Russland) und die Effekte der klimatischen Veränderungen.

In einem solch dynamischen und komplexen System kann die Gewährleistung von Ernährungsnachhaltigkeit mit den aktuellen Modellen nur ungenügend unterstützt werden. Es braucht eine umfassendere Betrachtung der Agrar- und Ernährungswirtschaft, welche alle Elemente, die einen wesentlichen Einfluss auf die Nachhaltigkeitsaspekte haben können, mitsamt ihren Wechselwirkungen im System berücksichtigt.

Zum einen erhöht die wachsende Komplexität der landwirtschaftlichen und ernährungswirtschaftlichen Prozesse, z.B. hinsichtlich der internationalen Lieferketten von Nahrungs- und Produktionsmitteln, den Bedarf an unterschiedlichen Datenquellen und somit den Aufwand für die Datensammlung und die Integration der Daten in das Modellsystem. Aufgrund der starken Heterogenität der Daten und deren Qualität kann dies einen erheblichen Zeitaufwand bedeuten und birgt zudem das Risiko von Fehlern, die die Ergebnisse der inhärenten Modellsimulationen verzerren könnten, und erhöht die Datenunsicherheit. Dazu kommt, dass je nach Nutzungszweck (kurzfristig vs. langfristig) unterschiedliche Daten für Entscheidungstragenden relevant sind. Das Modell muss die unterschiedlichen Anforderungen der Nutzenden erfüllen können und ausreichend flexibel an sich ändernde Bedürfnisse anpassbar sein.

Zum anderen erhöht die steigende Komplexität und Dynamik auch die Notwendigkeit, die Logiken des Modellsystems und die darunterliegenden Daten regelmässig zu aktualisieren und auf seine Korrektheit zu prüfen. Nur so kann sichergestellt werden, dass das Modellsystem weiterhin valide Aussagen und Prognosen erlaubt und seine Funktion als Entscheidungsunterstützungsinstrument erfüllen kann.

Aufgrund der hohen Komplexität des Modellsystems ist es auch von zentraler Bedeutung, dass Entscheidungstragende die Nutzung des Systems und die

Interpretation der Ergebnisse einfach verstehen und nachvollziehen können, ohne dass dies einen grossen zeitlichen und kognitiven Aufwand erfordert. Ansonsten besteht die Gefahr, dass die Erkenntnisse des Modellsystems nur wenig oder gar nicht in die Entscheidungsfindung einfließen, was dessen Effektivität als Instrument zur Entscheidungsunterstützung erheblich einschränken könnte.

Lösungsansatz

Um der zunehmenden Komplexität des Ernährungssystems gerecht zu werden und eine datenbasierte Entscheidungsunterstützung für Ernährungsnachhaltigkeit zu liefern, wird das bestehende Modellsystem weiter ausgebaut und verfeinert. Das Modellsystem bildet das Ernährungssystem in seiner ganzen Komplexität ab und bietet somit eine zuverlässige Grundlage für strategische und operative Entscheidungen sowohl hinsichtlich der Optimierung der kurz- bis mittelfristigen Versorgung von Nahrungsmitteln als auch der langfristigen Versorgungssicherheit und Nachhaltigkeit.

Das Modellsystem integriert zum einen eine Vielzahl von Informationen, um einen umfassenden Überblick über das Ernährungssystem zu ermöglichen. Dies beinhaltet den Bedarf an Produktionsmitteln in der Agrar- und Ernährungswirtschaft, wie z.B. Maschinen, Mineraldünger und Transportkapazitäten. Auch Produktionsfunktionen und -alternativen (bei mangelnder Verfügbarkeit von Produktionsmitteln), Bodeneignungsdaten, Informationen zur aktuellen Versorgungslage und zu möglichen Gefährdungen sowie die Kosten und Wirkungen von betrieblichen und politischen Massnahmen sollten berücksichtigt werden.

Zum anderen ist das Modellsystem in der Lage, die Versorgungssicherheit und Nachhaltigkeit zukünftiger Ernährungssysteme zu beurteilen. Hierzu wird es um Informationen zu zukünftigen Rahmenbedingungen (z.B. Produktionsverfahren, Erträge), Ernährungstrends und -verhalten, gesundheitlichen Auswirkungen von Ernährungsweisen sowie Indikatoren zur Nachhaltigkeit der Ernährung erweitert.

Damit das bestehende Modellsystem trotz dieser Erweiterungen handhabbar bleibt, ist es erforderlich, dass die beabsichtigten Funktionen zu Beginn der Entwicklung bzw. des Ausbaus definiert werden. So sind z.B. für die Simulation von kurzfristigen Versorgungsengpässen andere Datengrundlagen und Algorithmen nötig als für die Prognose und Beurteilung zukünftiger Ernährungssysteme. Es muss von Beginn an klar definiert werden, welche Fragestellungen das Modellsystem beantworten kann und welche nicht. Gleichzeitig sollte das Modellsystem aber auch modular aufgebaut werden, um eine optionale Skalierung und Anpassung des Systems zu ermöglichen, wenn weitere Aspekte wie z.B. Lebensmittelsicherheit in den Entscheidungsfindungsprozess einfließen sollen.

Die Herausforderung besteht darin, all diese Informationen effizient und fehlerfrei zu sammeln und zu integrieren. Hier helfen datenwissenschaftliche Methoden aus der erweiterten Statistik, dem maschinellen Lernen und dem Bereich

der KI. Sie werden beispielsweise dazu genutzt, die Prozesse der Datensammlung und -plausibilisierung zu automatisieren und zu verbessern. So dienen Anomalie-Erkennungsverfahren auf Basis datenwissenschaftlicher Methoden dazu, abweichendes Verhalten in Datensätzen aufzuspüren. Sie ermöglichen es, Ausreisser in den Daten (teil-)automatisch zu identifizieren und an Fachexperten und Fachexpertinnen oder den Datenlieferanten rasch weiterzuleiten, um zu überprüfen, ob es sich tatsächlich um fehlerhafte Werte handelt.

Neben der Datenverarbeitung spielt auch die Darstellung der Ergebnisse eine entscheidende Rolle. Das Modellsystem präsentiert die Ergebnisse der Datenanalysen, Prognosen und Simulationen in einer leicht verständlichen und ansprechenden Form, um Entscheidungstragenden einen schnellen Überblick über die aktuelle Lage und die möglichen Handlungsoptionen zu bieten. Nutzerfreundliche Schnittstellen und Funktionen, wie z.B. (teil-)automatisch generierte Empfehlungen, sind hierbei hilfreich. Dabei wird den Entscheidungstragenden die Güte der Zahlen, Aussagen und Empfehlungen des Modells transparent kommuniziert. Wenn beispielsweise die Unsicherheit der Modelloutputs hoch ist (gemessen mit datenwissenschaftlichen Methoden, z.B. durch Sensitivitätsanalysen oder Monte-Carlo-Simulationen), d.h. dass es wahrscheinlich ist, dass der tatsächliche Wert signifikant von dem dargestellten Wert abweicht, wird dieser entsprechend gekennzeichnet.

Durch den Einsatz der Datenwissenschaft ist das Modellsystem zudem lernfähig. So kann es auf Basis von Erkenntnissen vergangener Ereignisse und Feedback von Experten und Expertinnen sowie Nutzenden z.B. besser vorher-sagen oder simulieren, wie sich bestimmte politische Massnahmen auf die Ernährungsnachhaltigkeit in der Schweiz über einen längeren Zeithorizont auswirken würden.

Übertragbarkeit

Ein solches ganzheitliches Modellsystem, welches modular aufgebaut ist, kann in vielen anderen Politikfeldern eingesetzt werden, in denen komplexe Zusammenhänge und Interaktionen eine Rolle spielen. Beispielsweise könnte es in der Gesundheitspolitik angewendet werden, um Massnahmen zur Bekämpfung von Epidemien zu bewerten, oder in der Energiepolitik, um Ressourcenplanung und die Umsetzung von nachhaltigen Technologien zu unterstützen.

Anforderungen

Daten

- Sicherstellung der Datenqualität, u.a. durch Entwicklung von Standards und Protokollen zur Vereinheitlichung unterschiedlicher Datenformate und -strukturen aus diversen Quellen.
- Sicherstellung der Aktualität der Datengrundlage, u.a. durch die Bereitstellung der notwendigen Schnittstellen.
- Etablierung validierter Prozesse zur Erfassung, Aufbereitung, Verarbeitung und Speicherung von Daten, um Datenintegrität und -zugänglichkeit zu gewährleisten.

- Sicherstellung grundlegender Datenverfügbarkeit (z.B. Daten zur Anbaueignung der Landwirtschaftsflächen oder zu alternativen Produktionsprozessen).
- Quantifizierung der Datenvariabilität als Basis der Simulationen.

Regulierung und Ethik (in Anlehnung an DSStB-Grundprinzipien)

- Gewährleistung von Daten- und Informationsschutz und Informationssicherheit, z.B. durch Zugriffsbeschränkungen oder Implementierung von Mechanismen zur Anonymisierung personenbezogener Daten, um die Privatsphäre der betroffenen Personen zu wahren.
- Sicherstellung der Transparenz von datenwissenschaftlichen Algorithmen und Methoden, die verwendet werden, um Vertrauen bei den beteiligten Stakeholdern aufzubauen und Diskriminierung zu vermeiden.
- Wahrung der Geschäftsgeheimnisse privater Akteure und der Persönlichkeitsrechte von Individuen (z.B. Bewegungsflüsse).
- Entwicklung von Mechanismen, welche nachvollziehbar und erklärbar machen, wie die Outputs des Datenanalysesystems entstehen.

Technik und Infrastruktur

- Bereitstellung einer leistungsfähigen, flexiblen und skalierbaren IT-Infrastruktur, um mit der wachsenden Menge an Daten umgehen zu können und rechenintensive Simulationen und Datenanalysen zu ermöglichen.
- Effiziente, langfristige Speicherung der Input- und Output-Daten zum Zweck von Versionenkontrollen und der Reproduzierbarkeit der Modellrechnungen.
- Gewährleistung der Kompatibilität der Systeme mit bestehenden Technologien und Standards und Antizipierung zukünftiger Technologiewechsel.
- Ausreichende, skalierbare Rechenleistung zum Trainieren von datenwissenschaftlichen Algorithmen und Methoden (gerade, wenn Elemente aus maschinellem Lernen und aus dem Bereich der KI eingesetzt werden).
- Bereitstellung einer verwaltungsinternen Cloud-Plattform für die Anwendung («storage» und «computing»).
- Implementierung robuster Sicherheitsmassnahmen, um die Systeme vor Cyberangriffen und Datenlecks zu schützen.

Management

- Etablierung eines effektiven Projektmanagementansatzes, um die Implementierung und Weiterentwicklung der ganzheitlichen Modellsysteme zu steuern.
- Regelmässige Überprüfung der Systemleistung und Ergebnisse zur Validierung der datenwissenschaftlichen Algorithmen und Methoden und allfälliges Korrigieren.
- Identifikation und Mitigation potenzieller Risiken, die mit der Implementierung und Nutzung der ganzheitlichen Modellsysteme einhergehen, u.a.

durch permanente Kontrolle der Qualität des ganzen Modellsystems mit Hilfe eines Monitoringsystems.

Wissen und Fähigkeiten

- Sicherstellung der Verfügbarkeit von Expertinnen und Experten aus verschiedenen Disziplinen, um die Komplexität des untersuchten Systems/Prozesses angemessen abzubilden.
- Regelmässige, stufengerechte Schulungen und Weiterbildungsmassnahmen für Systemnutzende als auch Entscheidungstragende, um den optimalen Einsatz der Modellsysteme zu gewährleisten, z.B. Schulung im Umgang mit Unsicherheit im Modellsystem, z.B. die Interpretation von Ergebnissen mit Wahrscheinlichkeiten und Vertrauensintervallen, und die Bereitstellung der notwendigen Informationsgrundlagen zur Güte der Daten und Analysen im Modell.
- Implementierung von Prozessen zur Sammlung, Speicherung und Weitergabe von Wissen und Erfahrungen, um kontinuierliche Verbesserungen und Innovationen zu ermöglichen.

Stakeholder

- Aktive Einbindung der betroffenen Stakeholder in den Entwicklungs- und Entscheidungsprozess, um Transparenz zu gewährleisten und das Vertrauen in ganzheitliche Modellsysteme zu stärken.
- Etablierung von Kanälen und Verfahren, um Rückmeldungen und Anregungen von Stakeholdern zu sammeln, die in die Weiterentwicklung der ganzheitlichen Modellsysteme einfließen.

Ökologische Nachhaltigkeit

- Auswahl von energieeffizienten Servern und Rechenzentren, um Energieverbrauch und CO₂-Emissionen zu minimieren.
- Verwendung nachhaltiger Datensammlungs-, Speicher- und Modelltrainingspraktiken, einschliesslich Minimierung von Datenredundanz auf den für die Datensicherheit erforderlichen Umfang und Einsatz von Daten- und Modellkompressionstechniken.

3.6 Qualitätssicherung von Daten durch Plausibilitätsprüfungen

Am Beispiel von Sozialversicherungen

Allgemeiner Teil

Beschreibung

Die Qualitätssicherung von Daten bildet eine wichtige Grundlage für fundierte politische Entscheidungen sowie für den ordnungsgemässen Vollzug und eine ordnungsgemässe Aufsicht. Bei der ständig wachsenden Menge an zur Verfügung stehenden (und zu überwachenden) Daten können datenwissenschaftliche Methoden und Algorithmen zur «Anomalie-Erkennung» helfen, ungewöhnliche Muster in grossen Datenmengen zu identifizieren. Diese Muster können auf unbeabsichtigte Fehler (oder in Einzelfällen auch auf bewusst falsche Angaben) hindeuten, etwa bei der Datengenerierung oder -übertragung, wie auch in der Verarbeitung der Daten.

Zur Identifikation dieser Fehler analysieren datenwissenschaftliche Methoden und Algorithmen (insbesondere aus dem maschinellen Lernen und dem Bereich von KI) automatisch eingegangene Daten. Eindeutige Fehler in den Daten können durch das System automatisiert gekennzeichnet und korrigiert werden. Unsichere Fälle werden zur menschlichen Überprüfung an Fachexperten und Fachexpertinnen weitergeleitet bzw. zurückgemeldet. Z.B. ist die Meldung eines Mutterschaftsurlaubs für eine männliche Person definitiv ein Fehler und kann automatisch korrigiert werden. Die Immatrikulation einer grossen Zahl fünfzigjähriger Studierenden in einem Semester ist hingegen zwar unwahrscheinlich, aber nicht zwangsläufig ein Fehler und sollte dem Datenlieferanten, in diesem Fall einer Hochschule, zur Plausibilisierung zurückgemeldet werden.

Um solche «Anomalien» in Daten zu erkennen, werden die Modelle auf das «normale Verhalten» von Daten und die Identifikation eines davon abweichenden Verhaltens trainiert. Eine kontinuierliche Überprüfung und Aktualisierung des Modells ermöglicht es, Feedback zu integrieren und die Genauigkeit zu verbessern.

Gerade bei Daten, die in der Durchführung generiert werden und im Anschluss eine zentrale Rolle für Aufsicht- und Steuerungszwecke der Bundesverwaltung spielen, ist eine effiziente und effektive Qualitätssicherung der Daten durch Plausibilitätsprüfungen von grosser Bedeutung. Werden fehlerhafte Daten nicht als solche erkannt, kann dies unter anderem finanzielle Auswirkungen für Betroffene haben (z.B. im Subventionsbereich) oder zu falschen Schlüssen hinsichtlich der Regulierung führen (z.B. in der Spitalplanung).

Mehrwert und Relevanz

Der Einsatz von datenwissenschaftlichen Methoden und Algorithmen zur «Anomalie-Erkennung» (insbesondere aus dem maschinellen Lernen und dem Bereich von KI) für die Qualitätssicherung von Daten eröffnet vielfältige Vorteile. Ein entscheidender Vorteil ist die Effizienzsteigerung der Qualitätssicherung. Diese Modelle sind in der Lage, automatisch Muster in grossen Datenmengen

zu erkennen, die manuell möglicherweise übersehen würden bzw. deren manuelle Identifikation einen grossen Aufwand erfordert hätte. Dadurch ermöglichen sie eine systematische, vollständige und genaue Überprüfung der Daten. Darüber hinaus liegt ein Mehrwert in der Identifikation noch unbekannter Fehler durch datenwissenschaftliche Methoden und Algorithmen (insbesondere aus dem maschinellen Lernen und dem Bereich der KI). Während herkömmliche Plausibilisierungsmethoden, ob manuell oder automatisiert, nur bekannte, vorab definierte Fehler erkennen, decken Methoden und Algorithmen zur «Anomalie-Erkennung» auch bisher unbekannte Fehler auf. Sie basieren auf der Wahrscheinlichkeit von «Anomalien» und nicht nur auf vordefinierten Fehlern bzw. Regeln (z.B. «Mutterschaftsurlaub nur für weibliche Personen»).

Ein weiterer Vorteil ist die Optimierung des Prozesses durch Automatisierung routinemässiger Aufgaben, wobei Fachexperten und Fachexpertinnen nur noch bei komplexen Fragestellungen eingebunden werden. Dies führt zu einer effizienteren Nutzung der vorhandenen Expertise und erhöht die Arbeitszufriedenheit der Fachleute, indem sie sich auf komplexere und herausfordernde Aufgaben konzentrieren können.

Zuletzt schaffen plausibilisierte Daten die Grundlage dafür, weitere Analysen im Bereich der «Anomalie-Erkennung» unverzerrt durchzuführen, um Fachmitarbeitende und Entscheidungstragende durch die Identifizierung von möglichen Betrugs- oder Missbrauchsfällen (beispielsweise bei Steuererklärungen oder beim Beantragen von Sozialleistungen) zu unterstützen.

«Anomalie-Erkennung» zur Qualitätssicherung von Daten hat eine erhebliche Bedeutung in allen Phasen des politischen Entscheidungsprozesses. Durch Qualitätssicherung mittels Plausibilitätsprüfungen wird sichergestellt, dass politische Entscheidungen sowie der Vollzug und die Aufsicht in allen Phasen auf soliden und qualitativ hochwertigen Daten beruhen.

Anwendungsbeispiel aus der Praxis

Politische Reformen bei Sozialversicherungen

Ausgangslage

Das Bundesamt für Sozialversicherungen (BSV) ist mit der Verwaltung umfangreicher Mengen an Personendaten betraut, welche für Aufsicht, Steuerung, statistische Berichterstattung und Analyse sowie die Weiterentwicklung der sozialen Sicherheit genutzt werden. Datenströme werden kontinuierlich von verschiedenen Durchführungsstellen, darunter Ausgleichskassen und IV-Stellen, an das BSV übermittelt. Obwohl die Frequenz der übermittelten Informationen unterschiedlich ist, bekommt das BSV normalerweise schnell und regelmässig die notwendigen Informationen.

Der Datenflussprozess sieht vor, dass die Informationen von den Durchführungsstellen zur Zentralen Ausgleichsstelle (ZAS) gelangen. Hier werden die Daten geprüft, plausibilisiert und aufbereitet, bevor sie dem BSV zur Verfügung gestellt werden.

Bislang erfolgt die Prüfung der Daten zum einen bei den Durchführungsstellen durch Stichprobenkontrollen und zum andern durch automatisierte Plausibilitätstests bei der Datenübermittlung, die in den entsprechenden Weisungen geregelt sind. Diese Plausibilitätsprüfungen sind sowohl technischer (d.h. entspricht die Datenübermittlung dem vorgegebenen Format) wie auch fachlicher Natur. Der Fokus der fachlichen Plausibilitätsprüfungen liegt hauptsächlich auf der Ebene der versicherten Person. So wird beispielsweise kontrolliert, ob eine Person Zahlungen von mehreren Durchführungsstellen erhalten hat, ob der Anspruchsbeginn vor dem Anspruchsende liegt oder ob die Anzahl der entschädigten Tage das gesetzliche Maximum übersteigt.

Herausforderung

Nach politischen Reformen ist das Interesse an der unmittelbaren Abschätzung der Auswirkungen dieser Veränderungen besonders gross. Allerdings dauert es naturgemäss bis erste Zahlen vorliegen, da Leistungsansprüche, die unter das neue Recht fallen, erst von den Versicherten angemeldet, dann von den Durchführungsstellen geprüft, im Anschluss bei einer Gutheissung ausbezahlt und dann über die Übermittlungskanäle gemeldet werden müssen. Zusätzlich bergen solche Umbrüche, insbesondere in Übergangsphasen und zu Beginn eines neuen Systems, eine erhöhte Anfälligkeit für Fehler in der Datensammlung und -übermittlung. Diese Unstetigkeiten können vorübergehend zu einer Verringerung der Datenqualität führen, was wiederum die Einschätzung der Reformauswirkungen weiter verzögert oder verfälscht.

In diesen sensiblen Phasen ist eine effiziente und umfassende Datenüberprüfung von grosser Bedeutung. Die bisherige (teil-)automatisierte Prüfung beschränkt sich hauptsächlich auf die Plausibilität individueller Dateneinträge, wohingegen eine systematische Überwachung von Entwicklungen oft vernachlässigt wird. So könnten Unregelmässigkeiten auf Systemebene, wie z.B. eine plötzliche Abnahme der gemeldeten Versicherten durch eine Durchführungsstelle oder ein unerwarteter Anstieg oder Abfall des Gesamtauszahlungsbetrags, übersehen werden. Solche «Anomalien» können gerade in Phasen des Systemwandels auf Datenfehler hinweisen und sollten daher nicht unbeachtet bleiben.

Lösungsansatz

Eine umfassende Plausibilitätsprüfung der eingehenden Daten im Bereich der Sozialversicherungen adressiert die aufgeführten Herausforderungen sowohl im operativen Geschäft als auch im Kontext grösserer Reformen. Die Plausibilitätsprüfung erfolgt mithilfe eines Modells zur «Anomalie-Erkennung», das auf datenwissenschaftlichen Methoden und Algorithmen beruht - einschliesslich Methoden traditioneller und erweiterter Statistik, des maschinellen Lernens und aus dem Bereich der KI. Durch das Modell können sowohl bereits geläufige Fehler erkannt als auch neu auftretende Fehler identifiziert werden (beispielsweise solche, die erstmals aufgrund einer Reform auftreten). Insbesondere erlaubt das Modell eine Plausibilitätsprüfung auf System und nicht nur Einzelfallebene. Wenn es z.B. nach einer Reform zu einer Zunahme der AHV-Aufschübe kommt, verringert sich die Wahrscheinlichkeit, dass ein Neurentner oder eine

Neurentnerin 65-jährig ist. Das Modell würde erkennen, dass eine Ausgleichskasse aus diesem Schema ausschärft, was ein Hinweis auf einen Programmierfehler sein kann und entsprechend von einer Fachperson geprüft werden muss.

Grundsätzlich finden die Prüfungen automatisch nach Eingang der Daten von den Durchführungsstellen statt. Abhängig von der Schwere der aufgedeckten «Anomalien» erfolgt eine automatische Rückmeldung an die entsprechenden Stellen mit Aufforderung zur Korrektur bzw. Plausibilisierung der Daten. Auf diese Weise werden Lernzyklen etabliert, welche zum einen die Eingabe der Daten seitens der Durchführungsstellen und zum anderen das Modell zur «Anomalie-Erkennung» verbessern.

Auf einer plausibilisierten Datenbasis können anschliessend inhaltliche «Anomalien» sowie Auswirkungen von Reformen oder anderen Massnahmen untersucht werden.

Um die Effektivität und Zuverlässigkeit der «Anomalie-Erkennung» zu gewährleisten, ist es unerlässlich, die Algorithmen kontinuierlich zu überwachen und zu aktualisieren. Dabei ist es wichtig, dass mögliche «Anomalien» von Fachleuten interpretiert werden, um zu gewährleisten, dass diese tatsächlich auf Fehler hindeuten.

Eine zusätzliche Massnahme zur Verbesserung der Datenqualität neben dem systematischen Einsatz von Modellen zur «Anomalie-Erkennung» könnte in der Einrichtung einer IT-Plattform für Datengenerierende bestehen. Diese Plattform könnte die Generierenden dabei unterstützen, Daten korrekt zu übermitteln. Methoden und Algorithmen aus dem Bereich der KI könnten dabei helfen, die Daten im korrekten Format und ohne Fehler einzureichen.

Übertragbarkeit

Dieser Anwendungsfall ist besonders relevant für Bundesstellen, die etwelche Daten erhalten, aufbereiten, analysieren und darauf basierend Entscheidungen treffen. In diesen Situationen haben die Verwaltungseinheiten oft keine Kontrolle über die Datensammlung und -übertragung, da sie die Daten von externen Datengenerierenden erhalten und so die Datenqualität nicht selbst sicherstellen können.

Beispielsweise betrifft dies die meisten Aufsichts- und Statistikbereiche, die auf Daten von Drittanbietern angewiesen sind, um ihre Aufgaben zu erfüllen. Weitere Anwendungsfälle sind in Bereichen mit kantonalem Vollzug zu finden (z.B. Abfallmeldungen). Hier sind die Daten von externen Stellen abhängig und es ist von entscheidender Bedeutung, dass diese Daten präzise und verlässlich sind.

Anforderungen

Daten

- Integration der notwendigen Datenquellen zum Training der Methoden und Algorithmen, unter Berücksichtigung von Datenschutz- und Datensicherheitsstandards.
- Sicherstellung einer standardisierten Übermittlung von Daten an die Bundesverwaltung in einem einfach maschinenlesbaren Format.

Regulierung und Ethik (in Anlehnung an DStB-Grundprinzipien)

- Gesetzliche Grundlage für Weisungen an Datengenerierende / -liefernde bezüglich Datenformat, Übermittlungsfrequenz und -kanal, technischen Standards, Prozessen der Qualitätssicherung und Plausibilisierung etc. (entweder in den jeweiligen Fachgesetzen oder dedizierten Gesetzen wie dem Informationssystemgesetz im Bereich Sozialversicherung (in Arbeit)).
- Gewährleistung von Daten- und Informationsschutz sowie Informationssicherheit, z.B. durch Zugriffsbeschränkungen auf Methoden und Algorithmen.
- Offenlegung der verwendeten datenwissenschaftlichen Methoden und Algorithmen, um Vertrauen bei den Stakeholdern aufzubauen und verantwortungsbewusstes Handeln sicherzustellen.
- Vermeidung von Diskriminierung und Verzerrungen in den datenwissenschaftlichen Methoden und Algorithmen, z.B. Sicherstellung von akkuratem «Labeling» der Daten.
- Etablierung klarer Verantwortlichkeiten für den gesamten Prozess der Plausibilitätsprüfung.

Technik und Infrastruktur

- Bereitstellung der notwendigen, benutzerfreundlichen Schnittstellen zur Datenübermittlung.
- Bereitstellung von ausreichenden Rechen- und Speicherressourcen für die Verarbeitung und Analyse grosser Datenmengen.
- Implementierung automatisierter Verfahren zur Datenaufbereitung, Modellbildung und Aktualisierung, um die Effizienz und Genauigkeit zu erhöhen.
- Regelmässige Aktualisierung der Modellgewichte.
- Infrastrukturen zur Datenhaltung und zu erforderlichen Datenzugängen.
- Vorhandensein geeigneter datenwissenschaftlicher Werkzeuge und Plattformen.
- Sicherstellung der Interoperabilität auf Stufe Daten und datenwissenschaftlicher Methoden, Werkzeuge und Plattformen.

Management

- Bereitstellung der notwendigen finanziellen und personellen Ressourcen zur Umsetzung.

- Einführung effektiver Projektmanagementmethoden, um eine strukturierte Planung, Umsetzung und Überwachung des Anomalie-Erkennungsprojekts sicherzustellen.
- Förderung der Zusammenarbeit zwischen Datenwissenschaftlerinnen und Datenwissenschaftlern sowie Fachexpertinnen und Fachexperten, um den Wissenstransfer in der (Weiter-)Entwicklung des Systems zu ermöglichen.
- Implementierung der Prozesse zur Pflege und kontinuierlichen Verbesserung durch regelmässige Überprüfung, Aktualisierung und Feedback-Integration.
- Bereitstellung von Strukturen zur Messung des Effekts der Anwendung.
- Bereitschaft bei Entscheidungstragenden in der ganzen Bundesverwaltung und über Staatsebenen hinweg, gemeinsame digitale Lösungen und Strukturen zu unterstützen, und die Verfügbarkeit der notwendigen Ressourcen, um diese auch nutzen zu können.

Wissen und Fähigkeiten

- Sicherstellung, dass Datenwissenschaftlerinnen und Datenwissenschaftler sowie Fachexpertinnen und Fachexperten über die erforderlichen Kenntnisse und Fähigkeiten verfügen, um Projekte der «Anomalie-Erkennung» durchzuführen.
- Bereitstellung von Schulungen und Weiterbildungsmöglichkeiten für Mitarbeitende, um ihre Kenntnisse im Bereich der Datenwissenschaft und im Fachbereich zu vertiefen und auf dem neuesten Stand zu halten. Dies beinhaltet insbesondere datenwissenschaftliche Grundkenntnisse für Fachexperten und Fachexpertinnen sowie Personen in Führungspositionen, um das Verständnis zum Prozess der Datenwissenschaft in der ganzen Organisation zu stärken.

Stakeholder

- Permanenter Dialog mit Datengenerierenden zur Verbesserung der Qualität des Dateninputs.
- Sicherstellung der gleichen Basis von plausibilisierten Daten bei Kollaborationen.
- Einbindung relevanter Stakeholder in den Prozess der Anwendungsentwicklung (wo möglich), um deren Akzeptanz und Unterstützung zu gewährleisten.

Ökologische Nachhaltigkeit

- Nutzung energieeffizienter Technologien und Infrastrukturen, um den Energieverbrauch und den CO₂-Fussabdruck zu reduzieren.

3.7 Digitale Zwillinge

Am Beispiel eines digitalen Zwillings für die Schweizer Landwirtschaft

Allgemeiner Teil

Beschreibung

Ein digitaler Zwilling ist ein virtuelles Modell eines physischen, dynamischen Objekts oder Systems. Er wird vor allem bei hochkomplexen Objekten und Systemen verwendet, bei denen es schwierig ist, Zusammenhänge zu verstehen und Auswirkungen abzuleiten oder bei denen die Datenerhebung mit anderen Mitteln (z.B. Sensoren) zu teuer oder zeitaufwändig ist. Ein Beispiel dafür sind landwirtschaftliche Gebäude zur Tierhaltung mit umfassenden Lüftungs- und Heizsystemen. Bei der Steuerung dieser Systeme müssen auch die komplexen biologischen Bedürfnisse der Tiere berücksichtigt werden.

Der digitale Zwilling sammelt und verarbeitet Daten und Informationen über seinen physischen Zwilling und entwickelt auf dieser Basis ein digitales Abbild. Diese Datenübertragung kann dabei sowohl in Echtzeit als auch zeitversetzt erfolgen. Der digitale Zwilling spiegelt alle relevanten Aspekte des realen Systems wider, von Form und Bewegungen über Materialien und Umgebungsbedingungen bis hin zu Leistungskennzahlen.

Für den Nutzenden präsentiert sich der digitale Zwilling als ein Dashboard, das mit einem zeitlich und räumlich organisierten Modell verknüpft ist. Eine solche Benutzeroberfläche verfügt zudem oftmals über Eingabefelder zur Veränderung bestimmter Parameter und Integration neuer Inputs, mit deren Hilfe Nutzende die Auswirkungen von Veränderungen simulieren und analysieren können. Dies erlaubt es flexibel und schnell verschiedene Optionen auszuprobieren bzw. Szenarien durchzuspielen.

Das Konzept des digitalen Zwillings wird bereits in verschiedenen Bereichen der industriellen Produktion, der Logistik oder Medizin eingesetzt. Ein Beispiel dafür ist die Modellierung von Zügen und Zugsystemen, einschliesslich der zugehörigen Infrastruktur wie Schienen, Signalanlagen und Bahnhöfen. Durch Simulationen können Verkehrsflüsse analysiert und Fahrpläne optimiert werden. Ebenso ermöglichen sie das Testen und Analysieren des Verhaltens des Zugsystems unter verschiedenen Bedingungen, wie z.B. Wetteränderungen, technischen Ausfällen oder verändertem Passagieraufkommen.

Ein weiteres Beispiel ist die Modellierung von Lebensmitteln und Impfstoffen in Lieferketten, einschliesslich der zugehörigen Infrastruktur wie Kühlketten, Lagerung und Transportmittel. Die digitalen Zwillinge erleichtern es hierbei, die Haltbarkeit einzelner Produkte unter verschiedenen Bedingungen zu analysieren, beispielsweise bei Temperaturschwankungen, logistischen Problemen oder sich ändernden Nachfragebedingungen. Dadurch lassen sich Lieferketten und Verkaufsprozesse optimieren.

Mehrwert und Relevanz

Digitale Zwillinge bringen auch in der Politikgestaltung einen bedeutenden Mehrwert. Zum einen modellieren sie komplexe Systeme und Objekte und schaffen durch Visualisierung und Vereinfachung ein tiefgreifendes Verständnis für deren Elemente und deren dynamische Interaktionen.

Zum anderen kann die Überwachung solcher hochkomplexen Objekte oder Systeme mithilfe digitaler Zwillinge umfassend, in einigen Fällen sogar in Echtzeit, erfolgen. Dies ermöglicht Erkenntnisse zur Verbesserung von Kontrollen und Wartung von Objekten oder Systemen und fördert eine positive Auswirkung auf die Qualität der Outputs.

Digitale Zwillinge können zudem zur Unterstützung der Berichterstattung detaillierte und aktuelle Daten über den Zustand und die Effizienz von Infrastrukturen, Ressourcen und Abläufen bereitstellen. Ein Beispiel hierfür wäre die genaue Erfassung des Pestizid- und Wasserverbrauchs in landwirtschaftlichen Betrieben.

Ein weiterer zentraler Aspekt ist die Möglichkeit, verschiedene Szenarien und politische Massnahmen vor ihrer Implementierung zu simulieren. Dadurch können die Auswirkungen politischer Entscheidungen, beispielsweise der Einführung eines Road-Pricing-Systems, auf verschiedene Bereiche wie Umwelt, Wirtschaft oder Verkehr abgeschätzt werden. Die so generierten Daten sind besonders wertvoll, wenn ihre Erhebung durch andere Methoden grossen menschlichen bzw. finanziellen Aufwand oder natürliche Ressourcen erfordert (z.B. Experimente).

In einer Zeit steigender Komplexität wirtschaftlicher und gesellschaftlicher Systeme und zunehmender Anforderungen an die Nachvollziehbarkeit politischer Entscheidungen wird der Einsatz von digitalen Zwillingen wichtiger. Angesichts der heute vorhandenen Datengrundlage und Rechenkapazität besteht die Möglichkeit, das Potenzial digitaler Zwillinge zur Bewältigung dieser Herausforderungen zu nutzen.

Anwendungsbeispiel aus der Praxis

Digitaler Zwilling des Schweizer Landwirtschaftssystems

Ausgangslage

Digitale Zwillinge werden in der Schweiz bereits erfolgreich in Teilbereichen der Landwirtschaft und der Agrarpolitikgestaltung eingesetzt. Sie unterstützen bei der Entwicklung und Umsetzung nationaler Aktionspläne, beispielsweise bei der Verringerung der Verwendung von Pflanzenschutzmitteln und bei der Kontrolle von Nährstoffflüssen. Darüber hinaus können sie mit ihren Ergebnissen zur Erstellung von Umweltpolitikmodellen und bei der Durchführung von Ökobilanzen beitragen. Diese Anwendungen ermöglichen eine präzisere und effizientere Planung und Optimierung von Produktionsprozessen und tragen dazu bei, den Ressourceneinsatz und die Umweltauswirkungen landwirtschaftlicher Tätigkeiten zu minimieren und den Ertrag zu optimieren. Diese digitalen Zwillinge betrachten meistens individuelle landwirtschaftliche Betriebe. Sie sind auf

die spezifischen Gegebenheiten und Anforderungen des jeweiligen Betriebs zugeschnitten und ermöglichen so eine zielgerichtete Planung und Optimierung der Produktionsprozesse. Der Grad der Automatisierung und der Anbindung an andere landwirtschaftliche Modelle, Modellsysteme und Datensätze aus dem Bereich Agrarwirtschaft ist bislang aber nur in einem begrenzten Umfang umgesetzt.

Herausforderung

Die Herausforderungen im Bereich der landwirtschaftlichen Produktionsprozesse und der Arbeitswissenschaft in der Schweiz im Speziellen liegen vor allem in der Bewältigung der systemischen Komplexität. Aktuelle Tools zur Arbeitsplanung und -optimierung decken die Komplexität einzelner Betriebe ab und konzentrieren sich vorwiegend auf die Verbesserung der Effizienz auf Einzelbetriebsebene. Im Zentrum steht die Optimierung des Zusammenspiels von Menschen und Maschine.

Zum einen fehlt ein Ansatz, der die Auswirkungen und Zusammenhänge auf regionaler und nationaler Ebene abbildet. Zum anderen ist die Einbindung weiterer Faktoren erforderlich, welche es ermöglichen über den bisherigen Effizienzfokus hinaus das Landwirtschaftssystem zu verstehen. Im Fokus stehen dabei ökologische und soziale Faktoren.

So erfordern beispielsweise die Planung und Umsetzung von Massnahmen zur Verringerung der Umweltauswirkungen und der Berücksichtigung der Bedürfnisse der Menschen (z.B. der Arbeitskräfte) im landwirtschaftlichen Prozess eine detaillierte Analyse auf übergeordneter Ebene. Es ist essenziell, die Wechselwirkungen zwischen den verschiedenen Elementen des Systems - wie Pestizidrückstände, Wasserverbrauch oder Wohlbefinden der Mitarbeitenden - zu verstehen, auch solche, welche bisher in traditionellen wirtschaftlichen Logiken wenig Eingang gefunden haben, wie eben ökologische Nachhaltigkeit und die Bedürfnisse der Mitarbeitenden der landwirtschaftlichen Betriebe.

Eine solche umfassende Sichtweise hilft beispielsweise dabei, die Konsequenzen politischer Massnahmen, wie einer Erhöhung des Anteils ökologischer Landwirtschaft, Subventionen für bestimmte Anbauarten oder Auflagen für die Tierhaltung, genau abzuschätzen. Damit können die Auswirkungen solcher Massnahmen auf das Gesamtsystem daher genau und ohne Verzögerung erhoben und analysiert werden.

Werden politische Entscheidungen auf Basis von Modellen getroffen, die nur Teilaspekte abbilden, können ungewünschte Effekte die Folge sein. Für eine effektive Politikgestaltung ist daher eine ganzheitlichere und datenbasierende Herangehensweise erforderlich, die das Zusammenspiel aller Elemente in der Landwirtschaft berücksichtigt.

Lösungsansatz

Der digitale Zwilling des gesamten schweizerischen Landwirtschaftssystems ist eine komplexe, aber übersichtliche digitale Darstellung des realen landwirtschaftlichen Systems mit seinen verschiedenen Produktionsprozessen. Über

eine benutzerfreundliche Oberfläche - wie ein Dashboard - erhalten Entscheidungstragende Zugang zu den wichtigsten Parametern und Informationen des digitalen Zwillings. Hier können nicht nur aktuelle Informationen zum System abgeholt, sondern auch die Auswirkungen der simulierten Szenarien, wie der zukünftig benötigte Bedarf an Arbeitskräften, Wasser oder Düngemittel oder die Veränderungen in der Flächennutzung aufgrund politischer Entscheidungen, einfach verständlich präsentiert werden. Veränderungen in aktuellen, aber auch zukünftigen Arbeitsabläufen, beispielsweise durch fortschreitende Digitalisierung und Mechanisierung, aber auch Veränderungen im Anbaumanagement (z.B. Fruchtfolge, veränderte Kultursorten, Pestizidreduktion) können somit analysiert und auch mit anderen Datensätzen und Modellen verknüpft werden.

Der digitale Zwilling nutzt dazu Daten aus bestehenden Statistiken zur Landwirtschaft und Umwelt sowie weiteren Quellen wie der Geoinformation. Der Datensatz des digitalen Zwillings kann zusätzlich mit datenwissenschaftlichen Modellen und physikalischen Kalkulationen ergänzt werden. So können datenwissenschaftliche Methoden für Simulationen oder die Erzeugung von synthetischen Daten angewendet werden, wenn bestimmte Informationen durch Datenschutzbestimmungen eingeschränkt sind oder keine Messdaten aus realen Systemen zur Verfügung stehen. Beispielsweise sind in der Präzisionslandwirtschaft detaillierte Messdaten zu Feldgrößen, Bepflanzung und Bewässerung nötig. Ein mit Sensoren ausgestattetes Landwirtschaftsfahrzeug könnte diese Daten liefern. Sind diese Sensoren jedoch nicht installiert, können die Werte durch die Anwendung datenwissenschaftlicher Methoden synthetisch erzeugt werden.

Der digitale Zwilling verbessert das Verständnis des Landwirtschaftssystems, indem eine übersichtliche Abbildung aller relevanten Merkmale wie Produktion, Arbeitsmittel, Lieferketten und -wege, Ressourcenverbrauch (z.B. Personal, Düngemittel), Bodenbeschaffenheit, Pflanzenwachstum, Nährstoff- und Pflanzenschutzmitteleinsatz, Wetterbedingungen, Wasserverbrauch und Energieeffizienz und soziale Aspekte wie Arbeitsbedingungen sowie deren Interaktionen geschaffen werden.

Überwachung und Steuerung des Systems werden durch den digitalen Zwilling erheblich erleichtert. Politische Entscheidungstragende haben die Möglichkeit, das ganze Landwirtschaftssystem in der Schweiz, aber auch nur spezifische Regionen oder Anbauflächen zu überwachen und effektive Massnahmen zu implementieren. Dies könnte beispielsweise die Anpassung des Pflanzenschutzmitteleinsatzes in einem bestimmten Anbauggebiet beinhalten. Ob solche Massnahmen den gewünschten Effekt erzielen, kann mit dem digitalen Zwilling simuliert werden.

Die im digitalen Zwilling zusammengetragenen und generierten Daten können für die Berichterstattung im Landwirtschaftsbereich und zur Förderung der Transparenz im System genutzt werden. So kann nachvollzogen werden, wie sie z.B. der Einsatz von Pestiziden nach Einführen einer politischen Mass-

nahme verändert hat. Ein wesentlicher Vorteil des digitalen Zwillings ist in einem solchen Kontext, dass Parameter nicht nur zentral bzw. «top-down» (beispielsweise durch eine politische Entscheidung) gesteuert und gemessen werden können, sondern auch dezentral bzw. «bottom-up» (z.B. durch einzelne Landwirte) Veränderungen der Parameter angestossen bzw. gemessen werden können.

Der digitale Zwilling des gesamten schweizerischen Landwirtschaftssystems ermöglicht es zudem, verschiedene Szenarien zu simulieren und so optimierte Lösungen für das gesamte landwirtschaftliche Produktionssystem in der Schweiz zu entwickeln. Beispielsweise könnten Auswirkungen eines heisseren, trockeneren Sommers auf den Wasserverbrauch oder die Wahl der Landwirte bezüglich Anbautechniken und Frucht- und Getreidefolgen simuliert werden, oder wie eine Änderung der Düngemittelverwendung das Nährstoffgleichgewicht in den Böden oder die landwirtschaftliche Produktion beeinflusst. Die Auswirkungen dieser Veränderungen können damit für längere Zeiträume prognostiziert werden, so dass die langfristigen Auswirkungen simuliert und die Folgen möglicher negativer Auswirkungen quantifiziert werden können.

Gleichzeitig erlaubt der digitale Zwilling auch Auswirkungen von Massnahmen im landwirtschaftlichen System auf andere Bereiche abzuschätzen, wenn der Zwilling in die entsprechenden Modelle integriert wird, z.B. den Einfluss der Nutzung resilienter Anbaupflanzen auf landwirtschaftliche Versicherungsunternehmen oder Konsumentenschutz/-information (z.B. Labels), Anpassungen der Landwirtschaftsflächen auf Geoinformationssystem oder Veränderungen der Subventionspolitik auf AHV und IV.

Der digitale Zwilling des Schweizer Landwirtschaftssystem ermöglicht es auch, die landwirtschaftlichen Systeme anderer Länder besser zu verstehen und mit der Schweizer Landwirtschaft zu vergleichen. Dies geschieht in dem das Gerüst des digitalen Zwillings für die Schweiz auf das System in einem anderen Land angewandt wird. Diese Vergleichbarkeit liefert wertvolle Erkenntnisse für unsere eigene Agrarpolitik und fördert das Verständnis für die Landwirtschaftssysteme anderer Länder. Diese Informationen können wiederum in der Politikgestaltung oder für die operativen Tätigkeiten von anderen Einheiten der Bundesverwaltung hilfreich sein, z.B. bei der Konzeption und Umsetzung von Entwicklungsmassnahmen in anderen Ländern.

Die Möglichkeit, den digitalen Zwilling durch die Verknüpfung mit weiteren Datenquellen und anderen digitalen Zwillingen zu erweitern, z.B. von weiteren Produktionsstätten entlang der Wertschöpfungskette wie Schlachthöfe und Milchherstellungsanlagen, bietet Potenzial für eine noch präzisere Modellierung und Analyse.

Übertragbarkeit

Digitale Zwillinge sind in allen Bereichen von Relevanz, wo ein ganzheitlicher Blick eines physischen Systems und/oder die Simulation hypothetischer Entwicklungen auf physischer Systemebene benötigt werden, und die erforderlichen Daten vorhanden sind. Digitale Zwillinge sind in verschiedenen weiteren

Politikfeldern denkbar, zum Beispiel in der Stadt- und Verkehrs- und Raumplanung, dem Umweltschutz, der Energie- sowie Lebensmittel- und Gesundheitsversorgung. Im Falle der Energieversorgung können dynamische Prozessmodelle basierend auf datenwissenschaftlichen Methoden zum Beispiel zur Optimierung von Stromnetzen und zur Erkennung von Schwachstellen in der Versorgung beitragen.

Anforderungen

Daten

- Regelmässige und wenn möglich automatisierte Aktualisierung der Datenquellen, um sicherzustellen, dass die digitalen Zwillinge die aktuellen Informationen widerspiegeln und somit Datenanalysen (z.B. genaue Prognosen) ermöglichen. Dazu zählen spezifische Informationen zu Feldfrüchten und Tieren sowie Entfernungsdaten.
- Sicherstellung der Datenqualität durch beispielsweise regelmässige Überprüfung der Genauigkeit, Vollständigkeit und Konsistenz der verwendeten Daten.
- Verfügbarkeit detaillierter und präziser Daten (d.h. nicht nur aggregierte Daten wie Durchschnittswerte), um die unterschiedlichen Aspekte des realen Objekts oder Systems genau darstellen zu können.
- Schaffung einer Datenerfassungspipeline innerhalb eines zentralen oder dezentralen Datenbankmanagementsystems, um die verschiedenen Datenquellen zu konsolidieren und den Zugriff auf die Daten für alle beteiligten Stakeholder zu erleichtern.
- Gewährleistung semantischer Interoperabilität zwischen den konsolidierten Datensätzen.
- Erfüllung weiterer notwendiger Datenanforderungen wie die Verfügbarkeit der absoluten Werte (z.B. absolute Veränderungen über Zeit) und mehrjähriger Zeitreihendatensätze.

Regulierung und Ethik (in Anlehnung an DSStB-Grundprinzipien)

- Entwicklung von Datenschutzrichtlinien, um die Vertraulichkeit der gesammelten Daten und den Schutz der Privatsphäre der betroffenen Personen zu gewährleisten.
- Implementierung von ethischen Leitlinien und Risiko-Assessments für die Anwendung von Datenwissenschaft in der Politikgestaltung, um potenzielle Missbräuche zu verhindern und faire Entscheidungen zu gewährleisten sowie ungewollte negative Auswirkungen von Datenanalysen und Entscheidungen zu verhindern.
- Sicherstellung der Transparenz von datenwissenschaftlichen Algorithmen und Methoden, die in digitalen Zwillingen verwendet werden, und deren Prüfung auf mögliches diskriminierendes Verhalten, um Vertrauen bei den beteiligten Stakeholdern aufzubauen und Diskriminierung zu vermeiden.
- Wahrung der Geschäftsgeheimnisse privater Akteure und Persönlichkeitsrechte von Individuen (z.B. Bewegungsflüsse)

- Ausrichtung der Modellgestaltung auf die Gewährleistung der Reproduzierbarkeit von Ergebnissen, z.B. durch den Aufbau eines «Provenance Systems», zur Nachverfolgung, vergangener Einstellungen und Inputwerte des Modells
- Entwicklung von Mechanismen, welche nachvollziehbar und erklärbar machen, wie die Outputs der digitalen Zwillinge entstehen (z.B. Transparenz in den Prozessen und den einzelnen Prozessschritten)
- Entwicklung und Betrieb des digitalen Zwillings nach Open-Source- und Open-Data-Standards

Technik und Infrastruktur

- Etablierung einer leistungsfähigen IT-Infrastruktur, um die Verarbeitung und Speicherung der umfangreichen Datenmengen zu ermöglichen, die für die Erstellung digitaler Zwillinge erforderlich sind.
- Ausreichende, skalierbare Rechenleistung für die Modellsimulationen und die Anwendung von datenwissenschaftlichen Algorithmen und Methoden.
- Vorhandensein einer adäquaten Cloud-Infrastruktur, primär für Rechenleistungen (Computing), aber auch Datenhaltung.
- Bereitstellung von fortschrittlichen Analyse- und Simulationswerkzeugen, um komplexe Szenarien effizient modellieren und optimieren zu können.
- Schaffung einer sicheren und benutzerfreundlichen Plattform mit Dashboards für den Zugriff auf digitale Zwillinge und deren Analyse- und Simulationsfunktionen.
- Gewährleistung einer ausreichenden IoT-Systemlandschaft als Datenquelle.
- Bereitstellung eines «Institutional Repositories», um den Austausch von Meta-Daten zu erleichtern

Management

- Einführung eines interdisziplinären Projektmanagements, um die Zusammenarbeit zwischen Datenwissenschaftlerinnen und Datenwissenschaftlern, Dateningenieurinnen und Dateningenieuren, Datenanalytistinnen und Datenanalysten, Fachexpertinnen und Fachexperten, Vertreterinnen und Vertretern von politischen Entscheidungsprozessen (insbesondere aus dem Generalsekretariat) und anderen beteiligten Stakeholdern zu fördern.
- Schaffung von Strukturen und Ressourcen, um alle relevanten Expertisen in die Entwicklung und den Betrieb des Modells einzubinden und regelmäßigen Austausch im Rahmen der Weiteentwicklung des Modells zu fördern
- Festlegung klarer Ziele und Leistungskennzahlen für die Anwendung des digitalen Zwillings
- Implementierung von Monitoring- und Evaluierungsprozessen, um die Effektivität und den Mehrwert der digitalen Zwillinge kontinuierlich zu überprüfen und Verbesserungen vorzunehmen.
- Langfristige Gewährleistung von finanziellen und personellen Ressourcen für den Aufbau und Wartung des digitalen Zwillings.

Wissen und Fähigkeiten

- Förderung der Aus- und Weiterbildung von Datenwissenschaftlerinnen und Fachexpertinnen, um die erforderlichen Fähigkeiten für diese Anwendung (weiter) zu entwickeln.
- Schaffung eines internen Wissensmanagementsystems, um den Informationsaustausch und die Zusammenarbeit zwischen den beteiligten Akteuren zu fördern.
- Organisation von Schulungen und Workshops, um das Bewusstsein für die Möglichkeiten und Herausforderungen dieser Anwendung in der eigenen Organisation zu erhöhen.

Stakeholder

- Einbeziehung der betroffenen Stakeholder in den Entwicklungsprozess der digitalen Zwillinge, um deren Bedürfnisse und Anliegen zu berücksichtigen.
- Schaffung eines Dialogforums, um den Austausch von Erfahrungen, Ideen und «Best Practices» zwischen Stakeholdern zu fördern und die Akzeptanz der digitalen Zwillinge und der daraus resultierenden politischen Entscheidungen zu erhöhen.
- Regelmässige Kommunikation und Berichterstattung über den Fortschritt und die Erfolge dieser Anwendung von Datenwissenschaft in der Politikgestaltung, um Transparenz zu gewährleisten und das Vertrauen der Stakeholder zu stärken.

Ökologische Nachhaltigkeit

- Einsatz energieeffizienter Hardware und Cloud-Lösungen, um den Energieverbrauch bei der Implementierung von digitalen Zwillingen zu minimieren
- Verwendung nachhaltiger Datensammlungs-, Speicher- & Modelltrainingspraktiken, einschliesslich Minimierung von Datenredundanz und Einsatz von Daten- und Modellkompressionstechniken (insbesondere für Daten relevant bei umfassender Datenspeicherung, z.B. bei starker Nutzung von Sensoren)

4 Anhang

4.1 Datenwissenschaft im Prozess der Politikgestaltung

Der Prozess der Politikgestaltung («Policy Making») ist ein zyklischer Ablauf, der die Stadien der Gestaltung, Umsetzung und Evaluierung von Politik umfasst. In diesem Prozess spielt die Datenwissenschaft eine entscheidende Rolle.

Der Prozess beginnt mit der Identifikation eines Problems und des entsprechenden Handlungsbedarfs. In dieser Phase kann die Datenwissenschaft bei der Analyse und Auswertung relevanter Daten eingesetzt werden, um ein besseres Verständnis der Problematik sowie der relevanten Zusammenhänge und Abhängigkeiten zu erlangen.

Mit einem festen Verständnis der Problematik bewegt sich der Prozess weiter zur Identifikation von möglichen Lösungen, wie beispielsweise einer Gesetzesnovelle oder einem Verkehrsplanungskonzept. Datenbasierte Simulationen können in diesem Stadium eingesetzt werden, um die Auswahl der Optionen zu verfeinern und die Entscheidungsfindung zu unterstützen.

Die aus der Datenanalyse gewonnenen Erkenntnisse unterstützen im weiteren Verlauf des Zyklus die politische Diskussion und die Weiterentwicklung oder Präzisierung der identifizierten Lösungen. Sie ermöglichen fundierte Entscheidungen auf Basis von zuverlässigen Daten und Erkenntnissen.

Nach der Entscheidung kann die Datenwissenschaft verwendet werden, um die notwendige Ressourcenplanung zu analysieren und die potenziellen Auswirkungen der Entscheidung zu simulieren. Dies trägt zur Effizienz und Effektivität der Politikumsetzung bei.

In der Umsetzungsphase kann die Datenwissenschaft zur Aufbereitung von Informationen und zur Automatisierung von Prozessen beitragen, was zu erheblichen Verbesserungen führen kann.

Schliesslich, in der Phase der Evaluierung und möglichen Revision, können die gesammelten Daten analysiert werden, um Rückschlüsse auf den Erfolg der umgesetzten Politik zu ziehen. Dies trägt zur kontinuierlichen Verbesserung bei und hilft, das Verständnis des ursprünglichen Problems zu schärfen.

Zusammenfassend lässt sich sagen, dass die Datenwissenschaft das Potenzial hat, den gesamten Prozess der Politikgestaltung zu unterstützen und zu verbessern, indem sie vorhandene Instrumente der Politikgestaltung und -umsetzung ergänzt und erweitert. Dies wird mit der Abbildung 3 des Prozesses der Datenwissenschaft in jedem Prozessschritt verdeutlicht.

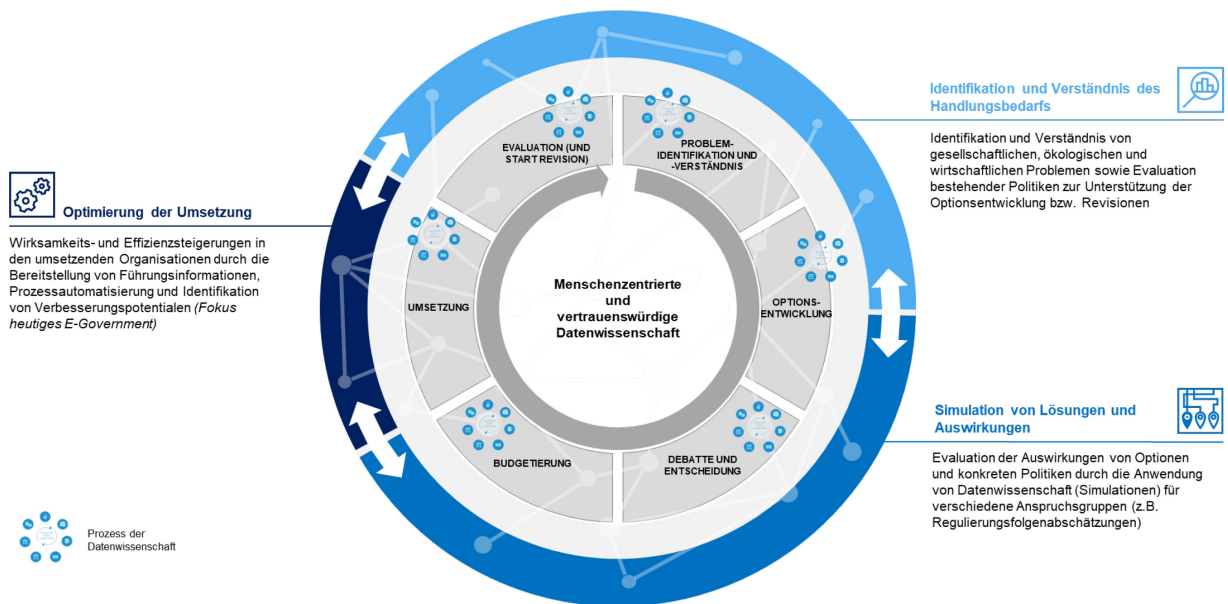


Abbildung 3: Datenwissenschaft entlang des Prozesses der Politikgestaltung («Policy Making»). (Quelle: DSStB).

4.2 Ansatz zur Erarbeitung des Berichts

4.2.1 Vorgehen

Die Entwicklung dieses Berichts basiert auf einer engen Kollaboration von Expertinnen und Experten der Bundesverwaltung und des ETH-Bereichs. Der Bericht wurde zwischen März und August 2023 erarbeitet. Er wurde in den folgenden Schritten erarbeitet.

Zuerst hat die interdepartementale Begleitgruppe «Datenwissenschaft» zusammen mit Vertretungen des ETH-Bereichs in einer Sitzung ein gemeinsames Verständnis für die Definition eines Anwendungsfalls, Kriterien für dessen Auswahl und eine grobe Struktur für dessen Beschreibung erarbeitet. In einem zweiten Schritt hat die Begleitgruppe gemeinsam mit Vertretungen des ETH-Bereichs erste Ideen für Anwendungsfälle gesammelt und sechs davon in einer gemeinsamen Sitzung priorisiert und präzisiert. Danach wurden Spezialistinnen und Spezialisten für die ausgewählten Anwendungsfälle aus der Bundesverwaltung und dem ETH-Bereich identifiziert, die jeweils eine Arbeitsgruppe gebildet haben. In drei Sitzungen haben die Arbeitsgruppen ihre Anwendungsfälle iterativ ausgearbeitet. Zur Orientierung hat das BFS einen hypothetischen Anwendungsfall vorgelegt.

Auf Basis der erarbeiteten Anwendungsfälle und der intensiven Diskussionen in den Arbeitsgruppen wurde der vorliegende Bericht verfasst. Er konsolidiert die wichtigsten Erkenntnisse sowie Anforderungen und verknüpft sie mit den laufenden Vorhaben und Evaluationen. Der Bericht wurde der Begleitgruppe zur Validierung und Präzisierung vorgelegt.

4.2.2 Charakteristika eines Anwendungsfalls

Ein Anwendungsfall in der Datenwissenschaft ist definiert als ein spezifisches Szenario, in dem Datenwissenschaft angewendet wird, um eine bestimmte Aufgabe zu erfüllen oder ein bestimmtes Ziel zu erreichen. Für den Kontext der Bundesverwaltung und im Rahmen dieses Berichts sollen die identifizierten Anwendungsfälle bestimmte Charakteristika aufweisen. Bei der Beschreibung eines Anwendungsfalls sollen folgende Aspekte laufend in Betracht gezogen werden:

1. **Relevanz:** Ein Anwendungsfall sollte für die Arbeit der Bundesverwaltung und, falls zutreffend, auch für die anderen beiden Staatsebenen relevant sein. Er sollte auf eine reale Herausforderung oder ein reales Problem eingehen, das im Rahmen der öffentlichen Verwaltung gelöst werden muss.
2. **Transferierbarkeit:** Ein Anwendungsfall sollte transferierbar sein und für verschiedene Politikfelder eingesetzt werden können. Dies erhöht seine Anwendbarkeit und ermöglicht es, den Nutzen der Datenwissenschaft in einer Vielzahl von Kontexten zu demonstrieren.
3. **Identifizierung notwendiger Grundlagen:** Der Anwendungsfall sollte dazu beitragen, die notwendigen Grundlagen zu identifizieren und deren Schaffung zu begründen. Dies kann die Identifizierung und Begründung von Rechtsgrundlagen, technischen Ressourcen oder organisatorischen Strukturen beinhalten.
4. **Menschenzentrierte und vertrauenswürdige Datenwissenschaft:** Der Anwendungsfall sollte auf der Anwendung vertrauenswürdiger Datenwissenschaft basieren und (direkt oder indirekt) einen Nutzen für den Menschen schaffen («menschenzentriert»). Dies bedeutet, dass die angewandten datenwissenschaftlichen Methoden, Techniken und Praktiken ethischen und rechtlichen Standards entsprechen und die Datenqualität und -sicherheit gewährleistet sind.
5. **Konkrete Beschreibung:** Der Anwendungsfall sollte ausreichend konkret beschrieben und mit praktischen Beispielen erläutert sein, damit Politik und Verwaltungseinheiten ihn auf ihre Arbeit übertragen können. Dies erleichtert das Verständnis und die Anwendung der Datenwissenschaft in der praktischen Politikgestaltung.

Die Berücksichtigung dieser Charakteristiken bei der Identifizierung und Beschreibung von Anwendungsfällen gewährleistet, dass diese sowohl relevant als auch nutzbringend für die Bundesverwaltung sind und zur Förderung der Anwendung von Datenwissenschaft in der Politikgestaltung beitragen.

4.2.3 Mitglieder der Begleitgruppe und der Arbeitsgruppen

Begleitgruppe:

Name des Mitglieds (Organisation)

Baeriswyl Dario (GS-EFD)	Fraefel Marianne (BK)	Mielewczik Michael (Agroscope)
Blarer Albert (armasuisse)	Hagi Ivan (ISC-EFPD)	Moret Lionel (MeteoSuisse)
Bremer Walter (BK)	Johansson Baker Kerstin (BFS)	Roth Anja (BSV)
<i>Brüning Jakob (BIT)*</i>	<i>Kaydul Alan (EDA)*</i>	Schütz Samuel (BFS)
Chatagny Florian (EFV)	Kellenberger Tobias (swisstopo)	Siegenthaler Claudia (GS-VBS)
Choirat Christine (BFS)	Keller Eva (EDA)	Singler Sebastian (PwC)
Cochard Jean-Luc (BAR)	Klopfenstein Hans-Peter (BIT)	Stephan Thomas (ESTV)
Droz-Georget Serge (EDA)	Krebs Daniel (GS-UVEK)	Vancauwenberghe Laure (SDSC)
Dumont Georges (BAZG)	Kuonen Diego (BFS)	Willi Lukas (PwC)
Egloff Florian (GS-VBS)	Loison Bertrand (BFS; Leitung der Begleitgruppe)	
<i>Flury Marco (BIT)**</i>	Michlig Roger (GS-VBS)	

* Im Laufe des Projekts aus der Begleitgruppe ausgeschieden

** Ersatz für Hans-Peter Klopfenstein an 2. Koordinationssitzung (14. Juni 2023)

Übersicht Arbeitsgruppe:

Anwendungsfall	Bundesverwaltung	ETH-Bereich
Nutzung von Open Source Intelligence (OSINT)	Eva Keller, Alan Kaydul, Mary Brown, Sibylle Obrist, Serge Droz-Georget (EDA); Albert Blarer (armasuisse); Diego Kuonen (BFS)	Andreas Wenger, Sascha Langenbach, Corinne Bara (Center for Security Studies, ETH Zürich); Elliott Ash (Center for Law & Economics, ETH Zürich)
Ganzheitliche Lagebilder in Krisensituationen	Markus Aeschlimann, Markus Böhm, Markus Müller, Philipp Angehrn (BAFU); Samuel Schütz (BFS)	David Bresch (Dep. Umweltsystemwissenschaften, ETH Zürich); Gianfranco Guidati (Energy Science Center, ETH Zürich)
Datenbasierte Unterstützung von Aufsicht	Gianpaolo Cecchin, Markus Luginbühl, Pascal Imoberdorf, Caroline Gerber (BAZL); Bertrand Loison (BFS)	Giovanni Sansavini (Dep. Maschinenbau und Verfahrenstechnik, ETH Zürich)

Anwendungsfälle der Datenwissenschaft entlang des Prozesses der Politikgestaltung

Ganzheitliche Modellsysteme zur Entscheidungsunterstützung	Albert von Ow, Mélanie Douziech (Agroscope); Diego Kuonen (BFS)	Stephan Pfister (Gruppe für ökologisches Systemdesign, ETH Zürich)
Qualitätssicherung von Daten durch Plausibilitätsprüfungen	Anja Roth (BSV); Diego Kuonen (BFS)	Matthias Bannert, Samad Sarferaz (KOF, ETH Zürich); Sergio Galletta (Gruppe für Ökonomie, Recht und Datenwissenschaften ETH Zürich)
Digitale Zwillinge	Michael Mielewczik (Agroscope)	Thijs Defraeye (Empa, ETH Zürich)