



Kompetenzzentrum
Öffentliche IT

Forschung für den digitalen Staat



Luise Ruge, Maximilian Kupi, Nicole Opiela, Mike Weber

Auf dem Weg zur evidenzbasierten Politik: Eine Analyse der Datenlabore der Deutschen Bundesregierung

Gefördert durch:



Bundesministerium
des Innern
und für Heimat

 **Fraunhofer**
FOKUS

aufgrund eines Beschlusses
des Deutschen Bundestages

Impressum

Autor:innen:

Luise Ruge, Maximilian Kupi, Nicole Opiela, Mike Weber

Gestaltung:

Reiko Kammer

Herausgeber:

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

ISBN: 978-3-948582-29-6

1. Auflage Juni 2024

Dieses Werk steht unter einer Creative Commons Namensnennung 3.0 Deutschland (CC BY 3.0) Lizenz (sofern nicht anders gekennzeichnet). Es ist erlaubt, das Werk bzw. den Inhalt zu vervielfältigen, zu verbreiten und öffentlich zugänglich zu machen, Abwandlungen und Bearbeitungen des Werkes bzw. Inhaltes anzufertigen sowie das Werk kommerziell zu nutzen. Bedingung für die Nutzung ist die Angabe der Namen der Autor:innen sowie des Herausgebers.

Logos und vergleichbare Zeichen dürfen nur im Kontext des Werkes genutzt und nicht abgewandelt werden.

Von uns verwendete Zitate unterliegen den für die Quelle geltenden urheberrechtlichen Regelungen.

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

Das letzte Abrufdatum der Onlinequellen ist der 21.06.2024.

Bildnachweis

Seite	Autoren	Quelle
1	Louis Reed	unsplash.com
6	Logan Moreno Gutierrez	unsplash.com
11	Julia Koblitz	unsplash.com
22	Fulvio Ciccolo	unsplash.com
26	Brands & People	unsplash.com

Das Wichtigste in Kürze

Die Datenlabore sind das aktuell umfassendste Instrument zur Förderung evidenzbasierter Politikgestaltung auf Bundesebene.

Sie zeigen auf, welche Mehrwerte Digital- und Datenkompetenz bieten können und sind Beweis für die Innovationsoffenheit der öffentlichen Verwaltung sowie ein wichtiger Impuls für den digitalen Kulturwandel.

Die Datenlabore agieren agil und nutzer:innenzentriert als Katalysator und Schnittstellenmanager.

Sie unterstützen die ministerielle Arbeit mit Daten und bilden die Schnittstelle zwischen fachlichen, datenanalytischen und arbeitsprozessbezogenen Aspekten. Dabei agieren die Datenlabor-Teams in partizipativer, agiler und nutzer:innenzentrierter Weise und sind damit auch ein Treiber für neue Arbeitsweisen in der Verwaltung.

Ein Datenlabor in jedem Ministerium ermöglicht die Anpassung an die spezifischen Bedarfe des jeweiligen Hauses.

So unterscheiden sich die Labore teilweise wesentlich hinsichtlich ihrer Entstehung und institutionellen Einbettung, ihrer personellen und infrastrukturellen Ausstattung sowie ihrer inhaltlichen Ausrichtung und können so den Anforderungen unterschiedlicher Arbeitskulturen, (technischer) Standards und Datenhoheiten gerecht werden.

Die Datenlabore sind ein gutes Beispiel interministerieller Kooperation.

Ein wesentliches Mittel dieser intensiven Zusammenarbeit ist der kontinuierliche Austausch unter den Datenlaboren in einer interministeriellen Arbeitsgruppe.

Der Bewährungsdruck der Datenlabore führt zum Vorrang kurzfristiger, sichtbarer Datenprojekte gegenüber längerfristiger, struktureller Grundlagenarbeit.

Angesichts der kurzfristigen Finanzierung liegt der Fokus der Datenlabore oftmals auf Datenanwendungen, um den haus-eigenen Verwaltungsalltag zu erleichtern und damit die eigene Nützlichkeit herauszustellen. Essenzielle Grundlagenarbeit für eine nachhaltige Implementierung datengetriebener Politikgestaltung, wie bspw. der Aufbau einer verbesserten (ggf. häuser-übergreifenden) IT- und Dateninfrastruktur, gerät dadurch ins Hintertreffen.

Potenziale einer stärkeren Außenkommunikation und -kooperation werden noch kaum ausgeschöpft.

Die Datenlabore sind in der Öffentlichkeit nur punktuell sichtbar. Transparentere Kommunikation und Kooperation mit anderen Verwaltungsakteuren sowie der Zivilgesellschaft, der Wissenschaft und der Wirtschaft könnten den gemeinwohlorientierten Nutzen datengetriebener Innovationen weiter erhöhen.

Es ist zu früh, die positive Langfristwirkung der Datenlabore zu beurteilen.

Die Datenlabore sind notwendige Experimentier- und Lernräume für datengetriebene Verwaltungsinnovationen, denen vor ihrer vollständigen Entfaltung in vielen Ministerien das Ende droht. Es bedarf einer längerfristigen Perspektive, um das Momentum errungener Fortschritte aufrecht zu erhalten und aus gemachten Erfahrungen weiter gemeinsam zu lernen. Nur so kann datengetriebene Politikgestaltung in der Bundesverwaltung und darüber hinaus verstetigt werden.

Inhalt

Das Wichtigste in Kürze	3
Hintergrund und Methodik	5
1. Entstehung und Verortung der Datenlabore	7
2. Mitarbeitende der Datenlabore	9
3. Ziele der Datenlabore	12
3.1 Evidenzbasierung	12
3.2 Datenkompetenzaufbau	13
3.3 Weitere Ziele	13
3.4 Zielgruppen	13
4. Output der Datenlabore	14
4.1 Bereitstellung von Datenprodukten	15
4.1.1 Sprachmodell-basierte Anwendungen	15
4.1.2 Dashboards	16
4.2 Erhöhung der Datenkompetenz	17
4.3 Aufbau der Dateninfrastruktur	17
4.4 Etablierung neuer Arbeitsweisen und Kooperationen	18
4.4.1 Hausinterne Kooperation	19
4.4.2 Kooperation im Datenlaborökosystem	20
4.4.3 Kooperation mit Öffentlichkeit und Zivilgesellschaft	21
5. Herausfordernde Wirkungsmessung	23
6. Fazit: Wirkungspotenziale der Datenlabore	25
7. Handlungsempfehlungen	27
Literaturverzeichnis	28

Hintergrund und Methodik

Data Science und künstliche Intelligenz (KI) sind Schlüsselemente des digitalen Wandels im öffentlichen Sektor. Sie können dazu beitragen, den öffentlichen Sektor »effizienter, gerechter, reaktionsfähiger, vorausschauender und nachhaltiger« (Kupi et al. 2022, S. 1) zu gestalten, indem sie unter anderem das Erkennen und Vorhersagen wichtiger Trends, das Simulieren und Bewerten politischer Alternativen sowie die Personalisierung und Automatisierung von Politikmaßnahmen unterstützen (ebd.). Doch wenn es darum geht, dieses Potenzial aktiv zu nutzen, steht die Bundesregierung noch am Anfang. Um dies zu ändern, wurden entsprechend der Empfehlung der Datenstrategie 2021 die Datenlabore der Bundesregierung eingeführt. Diese sollen Expertise im Bereich Data Science bündeln, die Fachabteilungen bei konkreten Fragen und Projekten unterstützen und als Multiplikatoren allgemein Wissen zu Datenkompetenz in den Fachabteilungen vermitteln und aufbauen. Zugleich sollen sie proaktiv Data Science-Projekte und -Anwendungen zur Unterstützung und Vereinfachung der Arbeit in den Ministerien im Einzelnen sowie der Bundesregierung insgesamt umsetzen. Die Datenlabore haben demnach drei Aufträge: die Beratung und Unterstützung, die Wissensvermittlung und die Entwicklung und Umsetzung, wobei sich je nach Datenlabor unterschiedliche Schwerpunkte ergeben können.

Über die Datenlabore, ihre Arbeitsweisen und Aktivitäten ist bislang wenig bekannt. Die vorliegende Studie macht es sich zur Aufgabe, erste beobachtbare und geplante Wirkungen der Datenlabore der Bundesregierung aufzuzeigen und zu analysieren. Im Zentrum steht hierbei die Frage, auf welche Weise die Datenlabore die oben skizzierten Aufträge zu erfüllen versuchen und wie sie sich in ihrem Vorgehen unterscheiden. Kurz gefragt: **Inwiefern tragen die Datenlabore dazu bei, Data Science in die Arbeitsweise der Ministerien und der Bundesregierung insgesamt zu integrieren und dort Datenkompetenz in der Breite aufzubauen?**

Um diese Fragen zu beantworten, wurden eine Literaturrecherche sowie darauf aufbauend 16 qualitative, leitfadengestützte Expert:inneninterviews mit Leitenden von 14 Datenlaboren

sowie Mitarbeitenden der Ministerien Ende 2023 durchgeführt. Einige der befragten Studienteilnehmenden baten um Anonymisierung ihrer Häuser, weswegen Namen nur bei ausdrücklichem Einverständnis genannt werden. Die Untersuchung orientierte sich am Input-Output-Outcome-Impact-Framework (IOOI-Framework) (siehe Bühner et al. 2022; Agora Digitale Transformation 2024), einem mehrstufigen Modell, welches eine umfassende Wirkungsanalyse ermöglicht. Die Auswertung der Daten und Ableitung der Handlungsempfehlungen erfolgte unter Einbezug von Feedback aus der Vorstellung erster Ergebnisse in einem erweiterten Expert:innenkreis. Da die Labore erst 2021 initiiert wurden und teils erst seit Kurzem ihre operative Arbeit aufgenommen haben, ist die Frage nach ihrer mittel- bis langfristigen Wirkung noch schwer zu beantworten. Es gibt jedoch erste beobachtbare Aktivitäten und unmittelbare Wirkungen, die es ermöglichen, sich der Beantwortung dieser Frage zumindest anzunähern.

Die Studie gliedert sich wie folgt: Zuerst werden die Hintergründe der Datenlabore, ihre Entstehung und ihre Rahmenbedingungen sowie die personelle Zusammensetzung und die Zielsetzung dargestellt. Darauf aufbauend wird auf die Aktivitäten und den Output der Datenlabore sowie die Arbeitsweise und Kooperationen innerhalb und außerhalb der Verwaltung eingegangen. Nach einer Analyse der bisher absehbaren Wirkungen der Arbeit der Datenlabore werden die Ergebnisse im Fazit zusammengefasst und Handlungsempfehlungen für politische Entscheidungsträger:innen formuliert.



MEIJI
TECHNO
JAPAN

Plan-Xc
20x
1.89

ME

1. Entstehung und Verortung der Datenlabore

In der im Januar 2021 verabschiedeten **Datenstrategie der Bundesregierung** wurde – neben wesentlichen Beschlüssen zur Datenpolitik – die Einrichtung von Datenlaboren in allen Bundesministerien sowie dem Bundeskanzleramt vorgesehen: »Wir wollen, dass alle Bundesministerien und/oder ihre nachgeordneten Behörden in eigener Verantwortung interne Datenlabore etablieren« (Bundeskanzleramt 2021, S. 102). Auch die Einrichtung der Posten von Chief Data Officers oder Chief Data Scientists wurde im Rahmen dieser Strategie angestrebt (mehr dazu auf S. 9f.). Im Kontext der deutschen öffentlichen Verwaltung definieren wir den Begriff »Datenlabor« wie folgt: **eine spezialisierte Einheit innerhalb einer Regierungs- oder Verwaltungsorganisation, die sich mit der Analyse und Nutzung großer Datenmengen beschäftigt und dabei hilft, die entsprechenden Anwendungen und Kompetenzen in der Organisation zu entwickeln und zu verbreiten.**

Die Gleichzeitigkeit der Einrichtung der Datenlabore in allen Bundesressorts kann als historische Besonderheit angesehen werden und ermöglicht den direkten Vergleich der 16 Labore in Bezug auf ihren Output und ihre Wirkungen. Während sich die Datenlabore hinsichtlich des Zeitpunktes und der Begleitumstände ihrer Entstehung, ihres Auftrages, ihrer organisatorischen Verortung sowie ihrer finanziellen Mittel und personellen und technischen Ausstattung (Inputs) ähneln, so unterscheiden sie sich doch in Bezug auf die von ihnen entwickelten und bereitgestellten Produkte, die Art und Intensität ihrer Fortbildungsaktivitäten (Outputs) und ihre geplanten und beobachtbaren Wirkungen (Impact/Outcome). Auf wesentliche Gemeinsamkeiten und Unterschiede im Aufbau und in der Ausgestaltung der Datenlabore soll im Folgenden eingegangen werden. Abbildung 1 gibt dazu einen ersten Überblick.



Abbildung 1: Dimensionen der unterschiedlichen Ausgestaltung der Datenlabore

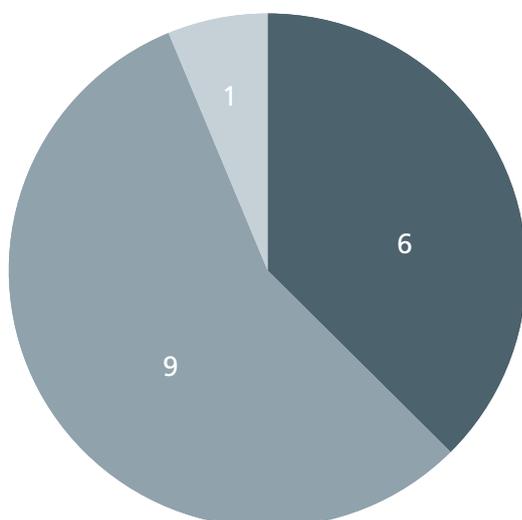
1. Entstehung und Verortung der Datenlabore

Die Finanzierung der Datenlabore stammt aktuell mit insgesamt 240 Mio. Euro aus dem **Deutschen Aufbau- und Resilienzplan (DARP)**, also EU-Geldern aus dem Aufbauinstrument **Next-GenerationEU**, das als Reaktion auf die Coronakrise aufgelegt wurde. Es war ein Novum, dass EU-Gelder direkt in deutsche Bundesministerien fließen. Die Bundesressorts und das Bundeskanzleramt haben für ihre Datenlabore jeweils 15 Mio. Euro bis Ende 2024 zur Verfügung gestellt bekommen, die jedoch wegen der Größe des Verwaltungsaufwandes nicht notwendigerweise auch überall vollständig ausgegeben werden können. Diese Gelder gaben jedoch in allen Bundesressorts sowie dem Bundeskanzleramt die Initialzündung für den Aufbau der Datenlabore.

Der Aufbau verlief hierbei in sehr unterschiedlichem Tempo: In einigen Fällen konnten die Ministerien auf umfassende Vorarbeiten im Bereich Data Science in ihren eigenen Häusern oder den ihnen nachgeordneten Behörden zurückgreifen, aber auch externe Expertise war im Aufbau der Datenlabore entscheidend. Innovationsinitiativen für den öffentlichen Sektor wie Work4Germany oder einschlägige Beratungshäuser wurden ebenfalls einbezogen.

Während einige Ministerien sofort mit der Etablierung ihres Datenlabors begannen oder bestehende Einheiten umorganisierten, erarbeiteten andere zunächst eine Vision und Mission im Rahmen partizipativer Prozesse (mehr zur partizipativen Arbeitsweise ab S. 12). Dies führte dazu, dass manche der Datenlabore bereits seit 2021 bestehen, während andere erst Ende 2023 ihre inhaltliche Arbeit aufnahmen.

Ein weiterer Unterschied zwischen den Ministerien ist die organisationale Aufhängung der Datenlabore. Es musste entschieden werden, ob die Datenlabore in der Zentralabteilung oder einer Grundsatzabteilung angesiedelt werden und ob eine neue Organisationseinheit geschaffen wird oder die Labore in bestehende Organisationseinheiten integriert werden.



Sechs der 16 Datenlabore, die im Zuge des DARP und der Datenstrategie gegründet wurden, sind der Zentralabteilung ihrer Ministerien zugeordnet, was oftmals zu einem verstärkten Fokus auf die Hausleitung als Zielgruppe führt (siehe dazu S. 13). Neun der Datenlabore sind in der Strategie- oder der Grundsatzabteilung ihres Hauses verortet. Für einige Datenlabore gibt es »*ein eigenes Kästchen im Organigramm*«. Teils werden organisatorisch aber auch andere Wege beschritten: Das Datenlabor des Bundesministeriums für Umwelt, Naturschutz, nukleare Sicherheit und Verbraucherschutz (BMUV) beispielsweise hat den Umgang mit Fach- und Sachdaten bei einer Geschäftsbereichsbehörde, dem Umweltbundesamt, angesiedelt.¹ Im Fall des Bundesministeriums des Innern und für Heimat (BMI) wurde gar die gesamte Datenlaborarbeit an eine Geschäftsbereichsbehörde, das Statistische Bundesamt, ausgelagert. Diese Entwicklung wird von Akteuren aus anderen Häusern teils durchaus kritisch gesehen:

»Das halte ich für nicht richtig oder für nicht zielgerichtet, weil wir einfach in einer Zeit leben, in der die Daten so viel zugenommen haben und so an Bedeutung gewonnen haben, dass vernünftige Regierungsarbeit Daten zu einem großen Teil enthält und dass wir nicht mehr sagen können, »wir sind juristische Experten, wir müssen diesen Teil nicht verstehen, das können wir uns erklären lassen.«

Auch im Hinblick auf ihre Zuständigkeiten unterscheiden sich die Datenlabore. Manche Labore erfüllen fast ausschließlich die in der Datenstrategie für sie vorgesehenen Aufgaben, andere sind darüber hinaus auch noch für weitere Themen zuständig, wie beispielsweise Open Data, Data Governance, Datenmanagement oder strategisches Controlling des gesamten Hauses. Alle Datenlabore folgen jedoch Prinzipien agiler, partizipativer und nutzer:innenzentrierter Arbeitsweisen oder verfolgen diese zumindest als idealtypisches Ziel.

¹ Siehe: <https://www.umweltbundesamt.de/themen/digitalisierung/anwendungslabor-fuer-kuenstliche-intelligenz-big#das-ki-lab-am-uba-fur-das-umweltressort>

- Zentralabteilung
- Grundsatzabteilung
- Nachgeordnete Behörde

Abbildung 2: Verortung der Datenlabore im Organigramm (N-16)

2. Mitarbeitende der Datenlabore

In Bezug auf Personal und Kompetenzen sind die Datenlabore nach eigenen Angaben – von einigen wenigen Ausnahmen abgesehen – zufriedenstellend ausgestattet. Die personellen Ressourcen und Kompetenzen variieren jedoch zwischen den einzelnen Häusern teilweise erheblich. Mit einer Spannweite von fünf bis 14 Mitgliedern bestehen beispielsweise deutliche Unterschiede in der Teamstärke. Zusätzlich wurden teilweise auch Stellen in der jeweiligen IT-Abteilung oder dem Geschäftsbereich geschaffen, die den Datenlaboren zuarbeiten.

Die Laborteams unterscheiden sich auch in Bezug auf ihre Kompetenzen. Die fachlichen Hintergründe der Leiter:innen sind divers und reichen von Astrophysik über Mathematik bis hin zu Luft- und Raumfahrttechnologie. Bevor sie die Leitung der Datenlabore übernahmen, arbeiteten sie zum Teil in Start-ups, Beratungsfirmen oder anderen Unternehmen. Die meisten von ihnen sind promoviert und haben Führungserfahrung. Die Diversität der Hintergründe sowie die Berufserfahrung außerhalb des Verwaltungskontextes verstehen viele der Leiter:innen als Vorteil, da so »alle auch frische Ideen mitbringen«.

»Wir sind auch insofern eine spezielle Gruppe. [...] Es sind extrem wenige Juristen, natürlich ganz anders als [...] insgesamt im Verwaltungsumfeld. Wir sind jünger, [...] als der Durchschnitt der Häuser und technikaffiner.«

Die Leiter:innen wirken dabei für ihre Teams sowohl nach innen wie auch nach außen. Nach innen besteht ihre Aufgabe darin, als klassische Führungskraft Mitarbeitende einzustellen, die Qualität der Arbeit zu sichern, Datenprojekte auf ihre Förderwürdigkeit hin zu prüfen und Aufgaben zu verteilen. Insgesamt ist die Arbeit in das eigene Team von verwaltungstechnischen Aufgaben bestimmt, die auch trotz dominierender Projektgruppenstruktur anfallen. Zusätzlich sehen die Leiter:innen es aber auch als ihre Aufgabe, ihre Teams politisch-strategisch auf die Ziele der Datenstrategie oder des Förderrahmens NextGenerationEU einzustimmen. Nach außen hin wirken die Teamchef:innen der Datenlabore als »Außenminister«, bzw. als Vertretung in Richtung der Leitungsebenen anderer (Unter-)Abteilungen sowie der Staatssekretär:innen. Die Rolle der Außenvertretung

erstreckt sich dabei auch auf die Repräsentation in der **Interministeriellen Arbeitsgruppe Datenlabore** (mehr zur IMAG ab S. 20).

Sowohl innerhalb ihrer eigenen Organisation als auch auf politischer Ebene streben die Leiter:innen der Datenlabore danach, für die Transformation hin zu mehr datengestützter Politikgestaltung und für verbesserte Datenkompetenz zu werben. Um alle Interessent:innen mitzunehmen und die Unterstützungsbasis auszubauen, gilt es »Marketing zu betreiben für dieses Thema« und zu kommunizieren, inwiefern Daten dabei helfen können, die Arbeit des Hauses »verlässlicher, schneller, effizienter und auch zielgerichteter« zu machen. Dieser Aufgabe kommen sie neben ihrer Funktion als Datenlaborleiter:in oft auch in Doppelfunktion als Chief Data Officer bzw. Chief Data Scientist ihres Hauses nach.

Aber nicht nur die Leiter:innen sind in ihren Profilen breit aufgestellt, auch die Teammitglieder bringen vielfältige fachliche Hintergründe und Kompetenzen mit. Darunter finden sich IT-Expert:innen, Mathematiker:innen, Ökonom:innen, Wirtschaftsethiker:innen, Coaches, Bürosachbearbeiter:innen, Datenschutzjurist:innen und IT-Projektmanager:innen. Ein Datenlabor beschäftigt Personen mit einem Hintergrund in Controlling, Haushalts- und Kassenrecht, da es als ein großes Hindernis wahrgenommen wird, das bewilligte Geld gezielt und rechtzeitig zu verausgaben:

»Und das ist auch eine wesentliche Hürde, warum wir nicht so viel Geld ausgeben, wie wir vielleicht wollten oder könnten, weil es einfach sehr viel Recherchearbeit ist. Aus welchem Titel muss das kommen? Kann das Geld innerhalb eines Jahres von Titel A nach Titel B geschoben werden? Wie wirkt sich jetzt die aktuelle Haushaltssperre darauf aus? Hinzu kommen Rechnungsvorgänge per Papier, die sehr aufwendig bearbeitet werden müssen. Bei der Auslegung des Haushaltsrechts scheinen sich einige Experten auch nicht immer einig zu sein. Also Haushaltsrecht ist ein Werk des Teufels. (lacht)«

2. Mitarbeitende der Datenlabore

Neben den Leiter:innen arbeiten in jedem der Datenlabore mindestens drei Data Scientists, teilweise mehr. Hinzu kommen Data Analysts und Data Engineers in unterschiedlicher Anzahl. Bei alledem ist das Verständnis davon, was ein Data Scientist im Rahmen der Datenlabore leisten kann und soll, aber weit gefasst und ein Abschluss in Data Science oder Informatik ist dafür nicht zwingend Voraussetzung. Auch Natur-, Wirtschafts- oder Politikwissenschaftler:innen mit Programmiererfahrung und starkem Bezug zu Data Science besetzen diese Positionen.

Data Scientists entwickeln fortgeschrittene Modelle und führen komplexe Analysen durch, um aus großen Datenmengen Sinn zu extrahieren.

Data Analysts konzentrieren sich auf das Sammeln und Interpretieren von Daten, um spezifische Fragen zu beantworten und bessere Entscheidungen zu treffen.

Data Engineers erstellen und warten die Infrastruktur, die für die Datenspeicherung und -verarbeitung benötigt wird.

Die Besetzung vakanter Stellen in den Datenlaboren wird als unterschiedlich herausfordernd wahrgenommen. So sei vor allem die Befristung der Stellen und die Dauer der Einstellungsverfahren eine Hürde. Über Mangel an passenden Bewerber:innen beklagten sich aber nur wenige. Vielmehr gibt es auch Fälle, in denen rund 900 Bewerbungen auf zehn Stellen zu verzeichnen waren. Die intrinsische Motivation vieler Bewerber:innen, die Verwaltung mitgestalten zu wollen, wird als einer der Gründe für die hohe Bewerber:innenzahl gesehen – trotz Befristung und mit der freien Wirtschaft oftmals nicht konkurrenzfähigem Gehalt:

»Das sind wirklich kompetente Menschen, die auch teilweise aus der Wirtschaft bereit sind, in die Verwaltung zu wechseln, obwohl die ganz anderen Regeln eine klare Herausforderung darstellen. Sie wollen dazu beitragen, diese Situation zu verändern, sie sind intrinsisch motiviert.«

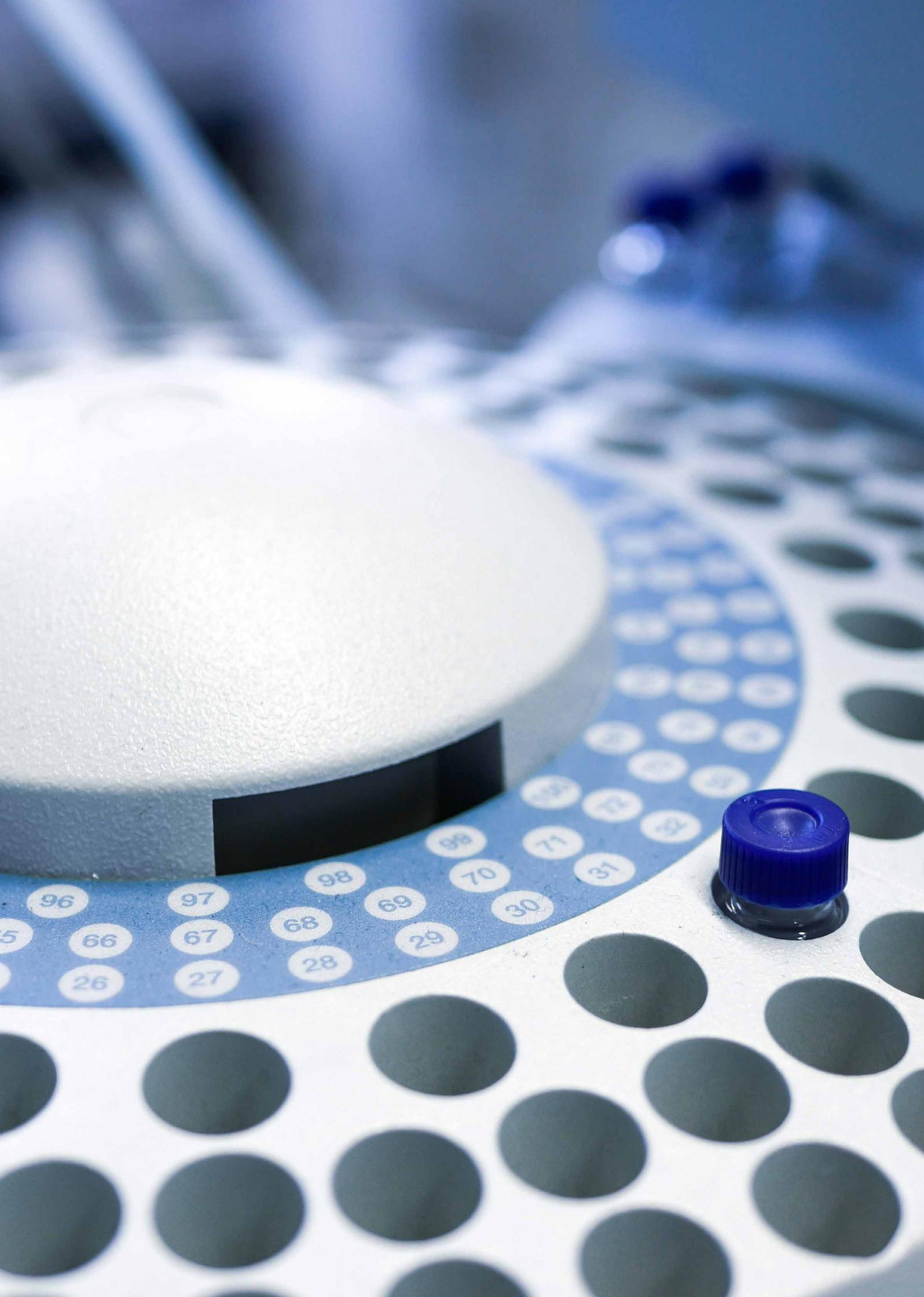
»Ich habe noch nie so intrinsisch motivierte Teams in der Bundesverwaltung gesehen, die insbesondere mit dieser Sonder-situation der Befristung so gut klarkommen. [...] Also ich finde, dafür sind das krasse Höchstleistungen, welche die Leute hier an den Tag legen.«

Auch andere Abteilungen der Ministerien unterstützen die Datenlabore teilweise in ihrer Arbeit:

»Wir haben gute Unterstützung bei uns aus dem Haus im Bereich Data Literacy, wo wir in einem großen Projekt unterwegs sind. Wir haben gute Unterstützung im Bereich IT und Dienstleister, einschließlich Strategieentwicklung.«

Die finanziellen Ressourcen der Datenlabore konnten auch dafür genutzt werden, Aufstockungen in den jeweiligen IT-Abteilungen oder Geschäftsbereichen zu erlangen, die dann den Datenlaboren zuarbeiten bzw. für diese in den Rest der Organisation wirken.

Auch Beratungsfirmen werden teils ergänzend zum Kernteam der Datenlabore hinzugezogen, um insbesondere Umsetzungsunterstützung bei konkreten Aufträgen zu leisten und bestehende Kompetenz- oder Kapazitätslücken zu kompensieren.



3. Ziele der Datenlabore

Die Zielsetzungen für die Datenlabore der Bundesregierung ergeben sich aus der Datenstrategie. Dies ist zum einen die Evidenzbasierung der Politik und zum anderen der Kompetenzaufbau im Bereich des Umgangs mit Daten (Data Literacy) bei den Mitarbeitenden der Bundesverwaltung. Die Teams der Datenlabore nehmen sich aber durchaus auch die Freiheit, diese Ziele auf ihre Weise zu interpretieren und zu erweitern.

Im Zuge des Laboraufbaus erarbeiteten und formulierten die Datenlabore auch eigene Zielvorstellungen und gingen dabei strategisch methodenbasiert vor: Ein Datenlabor entwickelte in einem agilen Prozess ein Umsetzungskonzept, welches als weiteres Ziel die Erarbeitung von Prototypen beinhaltet. Die Erarbeitung dieser Prototypen half beispielsweise zu verstehen, welche Akteure im Haus mit Daten arbeiten und deren Verständnis von Datenarbeit zu klären. Außerdem wurden auf diese Weise die Hürden und Schmerzpunkte erkennbar und es konnte nachvollzogen werden, warum bisherige Datenvorhaben gescheitert waren. Dies wurde als großer Lernpunkt wahrgenommen, auch im Hinblick auf die Aufstellung der hauseigenen IT.

In einem anderen Haus wurde eine Arbeitsgruppe aus drei Abteilungen gebildet, die gemeinsam Bedarfe, Ziele und Aufgaben erarbeiteten. So musste die Arbeit des Datenlabors zunächst von der Arbeit anderer Abteilungen abgegrenzt werden und man fragte sich: »Was macht ein Datenlabor überhaupt und was kann es für das Haus tun?«. Darauf aufbauend erarbeiteten die Beteiligten ein Konzept inklusive Zielstellung für das hauseigene Labor.

Im Auswärtigen Amt (AA) wiederum entwickelte das Team des Datenlabors aus der Zusammenschau von KI-Strategie, Datenstrategie und Open-Data-Strategie der Bundesregierung eine eigene Datenstrategie für sein Haus und ließ sich hierbei auch von internationalen Good-Practice-Beispielen, insbesondere aus den Arbeiten des US-Außenministeriums, inspirieren, wie beispielsweise der Enterprise Data Strategy (EDS)². In der hieraus

hervorgegangenen Datenstrategie sind auch die Ziele und Handlungsstränge für das Datenlabor des AA festgelegt.

Das BMUV wiederum erarbeitete eine rein strategische, nicht politische, interne Datenstrategie für das eigene Haus, in der auch eine Zielhierarchie für das Datenlabor enthalten ist.

Die genannten Beispiele verdeutlichen die Gemeinsamkeiten und Unterschiede bei der Entwicklung der Zielsetzungen der Datenlabore zwischen den einzelnen Häusern. Als Gemeinsamkeit lässt sich festhalten, dass in den meisten Fällen die Zielsetzung durch die Datenlabore selbst entwickelt wurde, die hierbei im Rahmen ihres politischen Auftrages große Freiheiten hatten. Ein Unterschied besteht beispielsweise dahingehend, wie stark die Mitarbeitenden des Hauses in die Missions- und Visionsentwicklung miteinbezogen wurden. Im Folgenden werden die zwei wesentlichen aus der Datenstrategie abgeleiteten Ziele der Evidenzbasierung und des Kompetenzaufbaus nochmals im Detail beschrieben, bevor auf die weiteren verfolgten Ziele und generellen Zielgruppen eingegangen wird.

3.1 Evidenzbasierung

»Wie bekomme ich die Evidenz in die Entscheidungsprozesse? Das ist ein langfristiger Prozess.«

Dem Sinn nach wird die Evidenzbasierung der Politik in allen Datenlaboren als wesentliches Ziel verfolgt, es bestehen jedoch unterschiedliche Schwerpunktsetzungen und Benennungen. So sieht das Datenlabor des BMUV einen wichtigen Schritt zur Evidenzbasierung der Politik darin, »die richtigen Informationen den richtigen Personen zur richtigen Zeit bedarfsgerecht aufbereitet zur Verfügung« zu stellen. In einem anderen Datenlabor wird hingegen nicht von Evidenzbasierung gesprochen, sondern vom Ziel des »wirksamen Regierens«, weil dieses Konzept im gesamten Politikzyklus auch die gesellschaftliche Wirkung mit in den Blick nimmt:

² Siehe: <https://www.state.gov/wp-content/uploads/2021/09/Reference-EDS-Accessible.pdf>

»Und wirksames Regieren heißt dann eben auch, die eigenen Strukturen, die eigenen Angebote so gestalten zu können, dass sie tatsächlich wirksam sind, also dort ankommen, wo sie ankommen müssen, und tatsächlich auch messbare und nachvollziehbare Wirkungen entfalten können.«

Das Verständnis von Evidenzbasierung erstreckt sich aber auch darauf, die Datenarbeit im eigenen Ministerium insgesamt »auf ein neues Niveau zu heben« und dem Ziel »Regieren in Echtzeit« näher zu kommen. Eine Modernisierung oder ein Neuaufbau der Arbeitsabläufe rund um die Datenarbeit ist damit ebenfalls eingeschlossen.

3.2 Datenkompetenzaufbau

»Und gerade gelingt es uns, in den Datenlaboren Kompetenzaufbau innerhalb der Häuser zu schaffen, da spannende Leute reinzubringen, voneinander zu lernen.«

Das zweite übergeordnete Ziel der Datenlabore, welches sich aus der Datenstrategie speist, bezieht sich auf den Kompetenzaufbau im Bereich der Data Literacy der Ministerialmitarbeitenden. Die Erhöhung der Datenkompetenz wird dabei als essenzielle Querschnittsaufgabe verstanden. Die Kolleg:innen im eigenen Haus zu befähigen, mit Daten und KI zu arbeiten, wird dabei auf einer Ebene mit Aufgaben gesehen, wie »eine Kantine [zu] betreiben oder dafür [zu] sorgen, dass jeder von uns einen vernünftigen Rechner hat oder dafür [zu] sorgen, dass die Post ankommt«. Bevor es die Datenlabore gab, war es gängige Praxis, für Datenfragen externe Expertise durch Dienstleister einzukaufen. Dank der Datenlabore besteht nun die Möglichkeit, diese grundlegende Expertise auch im eigenen Haus zu kultivieren.

In der alltäglichen Arbeit der Datenlabore ist damit das Ziel des Datenkompetenzaufbaus ständig präsent, beispielsweise im Kontakt mit den Fachreferaten. Dabei steht nicht nur der Kompetenzaufbau für die praktische Arbeit mit Daten im Fokus, sondern auch die Vermittlung von Kompetenzen im Sinne der Schaffung eines Verständnisses beispielsweise für statistische Kennzahlen und Analysen. Dies wird vornehmlich durch Fortbildungsmaßnahmen für die Kolleg:innen im Haus erreicht.

3.3 Weitere Ziele

Über die beiden Hauptziele der Evidenzbasierung und des Kompetenzaufbaus hinaus formulierten einige Datenlabore für sich noch spezifischere Zielstellungen. Zu diesen gehören:

- Stärkung der Data Governance
- Klare und verbindliche Rollen und Verantwortlichkeiten für den Umgang mit Daten und zukünftig möglicherweise auch

KI, zum Teil durch die Etablierung gänzlich neuer Rollen und Profile

- Zugewandter, offener Ansatz im Umgang mit Daten (datensensible Arbeitskultur) einschließlich der Schaffung von Räumen für Experimente und Innovation
- Erhöhung der Maschinenlesbarkeit eigener Daten

3.4 Zielgruppen

»Wir kümmern uns darum, wie wir die Mitarbeiter:innen aus unserem Haus bestmöglich unterstützen [können], mit digitalen Werkzeugen.«

Die Zielgruppen der Datenlaborarbeit variieren in den untersuchten Ministerien. Im Fokus stehen insbesondere die Mitarbeitenden des eigenen Hauses, mit Unterscheidungen dahingehend, inwiefern die Führungsebene des Hauses besonders adressiert wird. Damit einher geht auch die Frage, wie divers der Adressatenkreis in Bezug auf Anspruch und Kompetenzen ist. Ein Fokus auf alle Mitarbeitenden, also über unterschiedliche Qualifikationsebenen hinweg, bedeutet auch, ein großes Spektrum an Interessen bedienen zu müssen zwischen Sachbearbeitung und Hausleitung:

»Da hat eine Leitung andere Erwartungen daran, was man mit Daten anstellen kann, als das zum Beispiel eine Sachbearbeitung in einem Förderreferat oder ein:e Referent:in in einem gesetzgebenden Referat hat.«

Manche Datenprodukte, besonders jene aus dem Bereich der Datenvisualisierung, adressieren besonders die Führungsebene der Ministerien und erleichtern ihre Arbeit, beispielsweise bei Reisen in bestimmte Wahlkreise oder ins Ausland:

»Also klar, alle Abgeordneten im Ministerium, Staatssekretäre, reisen, und bisher war es ein recht großer Aufwand für die entsprechenden Referate, das vorzubereiten, weil man ja wissen muss, natürlich neben dem, wohin es geht, worüber sprechen die und welche Daten brauchen wir.«

Je nach Produkt kann die Arbeit des Datenlabors dem ganzen Haus dienen oder ein spezielles Problem bestimmter Mitarbeitender oder einer Abteilung lösen:

»Wenn wir dafür sorgen, dass das Haus beispielsweise auf Knopfdruck Wirtschaftssteckbriefe oder Wirtschaftsdatenblätter ziehen kann, dann ist das etwas, was den Arbeitsalltag von Vielen einfacher macht. Wenn wir sagen, wir kaufen Datenquellen ein und stellen diese zur Verfügung, dann kommt es auf die Datenquelle an. Aber es hilft in diesem Fall einer bestimmten Gruppe von Kolleginnen und Kollegen bei der Erledigung ihrer Fachaufgaben.«

4. Output der Datenlabore

Der geplante und beobachtbare Output der Datenlabore besteht aus viererlei: 1) Bereitstellung von Datenprodukten, 2) Erhöhung der Datenkompetenz, 3) Aufbau der Dateninfrastruktur und 4) Etablierung neuer Arbeitsweisen und Kooperationen.

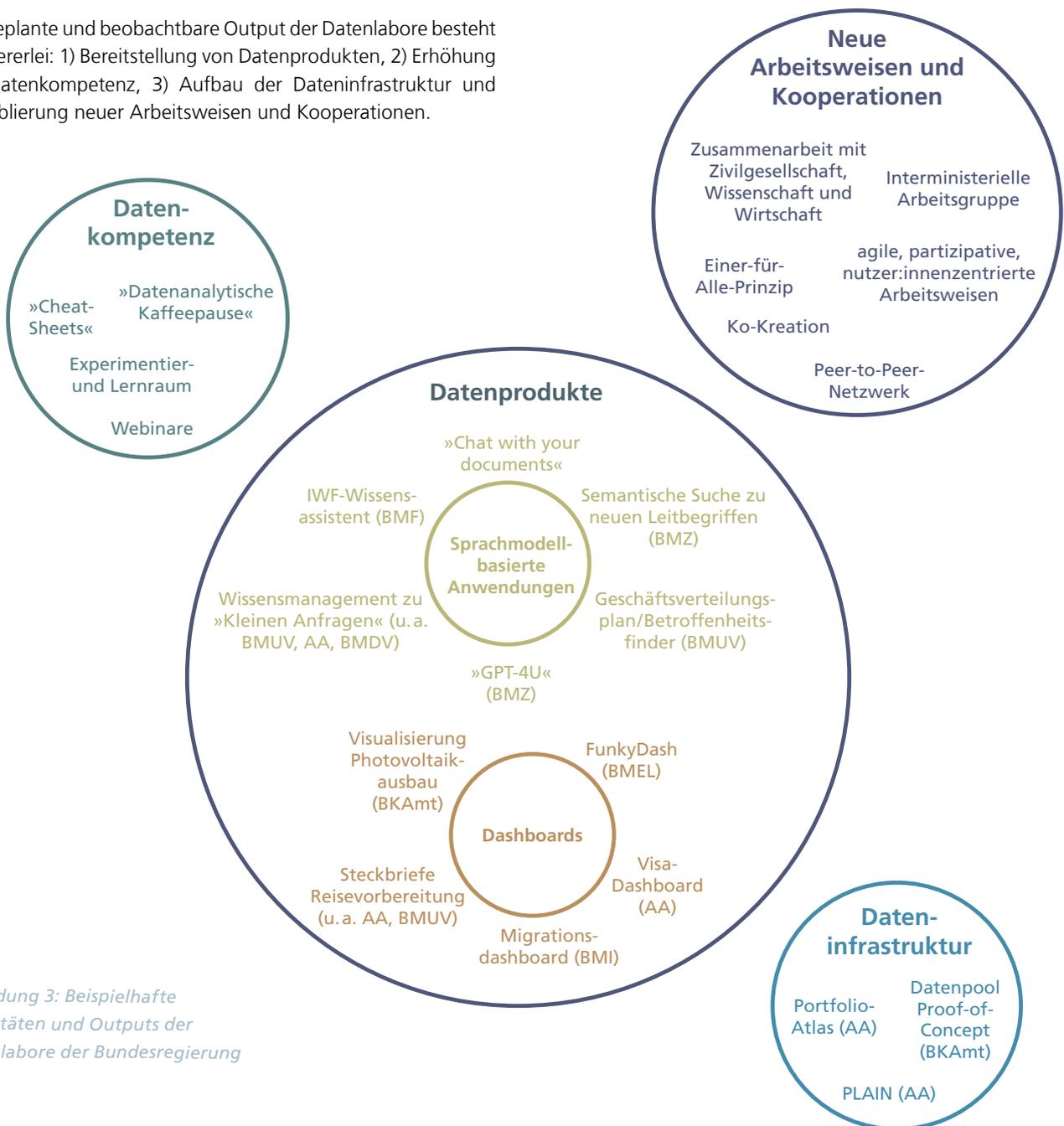


Abbildung 3: Beispielhafte Aktivitäten und Outputs der Datenlabore der Bundesregierung

Ein wesentlicher Zielkonflikt in der Arbeit der Datenlabore besteht darin, ob sie sich auf kurzfristige, außenwirksame Projekte oder auf langfristige, zunächst weniger sichtbare Vorhaben konzentrieren sollen. Gerade vor dem Hintergrund der Unsicherheiten bei der Verstetigung ist dies eine Herausforderung, bei der sich viele Datenlabore für ersteres entschieden, um ihre Nützlichkeit schnell sichtbar zu machen.

»Wir haben auch einen gewissen Zielkonflikt, weil wir auf der einen Seite das Datenlabor nachhaltig aufbauen wollen, auf der anderen Seite jedoch noch Projektcharakter haben und für eine Verstetigung mehr Sichtbarkeit und Akzeptanz in der Verwaltung benötigen. Der nachhaltige Weg bedeutet aufgrund der hohen Sicherheitsanforderungen und der mangelnden Data Governance den notwendigen Aufbau einer eigenen Infrastruktur, Entwicklungsumgebung, und Datenpipelines. Da die meisten Datenquellen der Verwaltung per Default als VS-NfD [Anmerkung: Verschlussache – Nur für den Dienstgebrauch] eingestuft sind, können diese beispielsweise nicht in einer Public Cloud verarbeitet werden. Diese Datenquellen als »nicht eingestuft« klassifizieren zu lassen, ist sehr langwierig und muss politisch gewollt sein und ist manchmal gar nicht möglich. Darüber hinaus bedarf es einer Analyse des Status quo inklusive aller vorhandenen Datenbanken und Systeme. Also die Vorarbeiten, bevor das Datenlabor mit der eigentlichen Datenarbeit beginnen kann. Aber das sind keine Arbeiten, die wir gut verkaufen können, die Sichtbarkeit in der Hierarchie bringen. Das hat sicher nicht den Sexyness-Faktor, der uns bei der Weiterfinanzierung und Verstetigung hilft. Aufgrund des Projektcharakters der Datenlabore sind wir noch stark im »Clicki-bunt«-Modus, also dem schnellen Aufbauen von kleinen Lösungen. Diese sind meist nicht nachhaltig und schnell veraltet, da die Daten händisch eingegeben werden und sich nicht automatisch aktualisieren. Hinzu kommt die eingeschränkte Datenverfügbarkeit, da viele Datenquellen in den Verwaltungen zwar über Schnittstellen verfügen, diese jedoch nicht freigegeben werden.«

Auch vom Reifegrad des Hauses im Umgang mit Daten hängt es ab, ob kurzfristig sichtbare oder langfristige, aber weniger sichtbare Vorhaben gewählt werden. So kann sich ein Datenlabor in einer Organisation mit höherem Reifegrad auf die Verfeinerung der Datenauswertung konzentrieren, während man anderswo darum bemüht sein muss, überhaupt *»schnell einen sichtbaren Erfolg«* zu erzeugen. Diese Konzentration auf schnelle Erfolge birgt jedoch die Gefahr, übergeordnete Ziele, für die man einen *»langen Atem«* benötigt, wie beispielsweise die Etablierung einer Data Governance, aus den Augen zu verlieren.

Als Beispiele für langfristige Ziele können auch alle ressortübergreifenden Infrastrukturvorhaben angesehen werden, wie die Erstellung von Datenkatalogen, die Erarbeitung von Schnittstellen oder die Etablierung gemeinsamer Datenpools. Die

Einführung eines ressortübergreifenden Datenpools und eines Datenatlas (zumindest) für interne Zwecke der Bundesregierung wird dabei von vielen Interviewten als besonders erstrebenswert erachtet, da diese Mittel im Idealfall eine einheitliche Arbeitsbasis für die verschiedenen Ministerien schaffen. Allerdings variieren die Rahmenbedingungen und Voraussetzungen in den verschiedenen Ressorts erheblich, was die Umsetzung solcher Vorhaben deutlich erschwert. Im Folgenden wird auf konkrete Outputs, also Produkte und Maßnahmen, der Datenlabore eingegangen.

4.1 Bereitstellung von Datenprodukten

Bis zum Abschluss der Expert:inneninterviews für diese Studie wurden in den Datenlaboren nahezu 100 Anwendungsfälle identifiziert, die in verschiedene Datenprodukte mündeten. Dabei handelt es sich beispielsweise um Sprachmodell-basierte Produkte oder visualisierende Dashboards.

4.1.1 Sprachmodell-basierte Anwendungen

Beispiele für einige in den Datenlaboren erarbeitete Anwendungsfälle, die mittels Sprachmodellen umgesetzt wurden, konzentrieren sich auf die Erleichterung des textbasierten Informationszugangs, wie die nachfolgenden Beispiele zeigen.

Bearbeitung »Kleiner Anfragen« mit KIRAS (KI-Rechercheassistent): Im Datenlabor des BMUV wird an einer KI-gestützten Rechercheanwendung gearbeitet, mittels derer die Beantwortung Kleiner Anfragen erleichtert werden soll. Die Anwendung bietet Interaktionsmöglichkeiten wahlweise per App oder über eine Web-Schnittstelle und war zum Befragungszeitpunkt für eine Kernnutzer:innengruppe von 40 Personen im Haus freigeschaltet. Der Anwendungsfall der Kleinen Anfragen beschäftigt auch einige andere Datenlabore, die daher an der Übertragung von KIRAS auf ihre Bedarfe arbeiten. Damit ist KIRAS ein Beispiel für eine gelungene interministerielle Kooperation.

Chat with your documents: Dank dieser KI-Applikation können mehrere hundert Seiten an Dokumenten schnell durchsucht und erfasst werden. Die Mitarbeitenden des Ministeriums können auf diese Weise auch rasch Handlungsanweisungen herausfiltern, die meist verständlicher sind als das Originaldokument. Es wurden Prototypen getestet zu Verordnungen und Handreichungen, z. B. zu spezifischen Arbeitsschritten bei der Schließung von Verträgen oder entsprechenden rechtlichen Vorgaben.

Betroffenheitsfinder: Der Betroffenheitsfinder ist ebenfalls eine Entwicklung des Datenlabors des BMUV. Die Suche nach den zuständigen Referaten im Geschäftsverteilungsplan im Rahmen neuer Vorgänge ist normalerweise ein manueller und, insbesondere bei querliegenden Zuständigkeiten, auch ein aufwendiger Prozess. Um diesen Prozess zu erleichtern, wurde der Betroffenheitsfinder entwickelt, der den Nutzenden nach Eingabe des Textes vorschlägt, welche Referate an einem bestimmten Vorgang zu beteiligen sind. Dazu vergleicht die Anwendung den Text semantisch mit dem Geschäftsverteilungsplan, um möglicherweise betroffene Referate zu identifizieren. Dieser Ansatz zielt darauf ab, den manuellen Suchprozess zu (teil-)automatisieren und effizienter zu gestalten. Zum Befragungszeitpunkt befand sich die Anwendung noch in der Testphase.

GPT-4U: Der auf ChatGPT basierende offene KI-Chatbot GPT-4U ist ein Chat-Interface, das – anders als die öffentlich zugängliche Version – keine Daten speichert. Der Chatverlauf wird nur im Browser der Anwender:innen gespeichert. Diese KI-Konversation des BMZ läuft auf einer Cloud-Instanz in der EU, also auf »DSGVO-Grund«. Gleichzeitig ist der Umgang mit GPT-4U auch ein Beispiel für die partizipative Arbeitsweise des Datenlabors des BMZ. So wurde die Applikation in einer Testphase in zehn Referaten erprobt. Die Ergebnisse aus der Testphase werden geteilt und sollen veröffentlicht werden, um auch für die Außenwelt die Erfahrungen und Möglichkeiten der GPT-Anwendung nachvollziehbar zu machen. Die Studie zur Verwendung von GPT-4U basiert auf Interviews und Umfragen mit den Mitarbeitenden, die aktiv an der Testphase teilgenommen haben. Nach dieser Studie kann die KI-Anwendung beliebig groß skaliert werden, ggf. gar auf die gesamte deutsche Verwaltung.

IWF-Wissensassistent: Eine Anwendung aus dem Datenlabor des Bundesministeriums der Finanzen (BMF) ist der ChatGPT-basierte IWF-Wissensassistent, der gemeinsam mit der Innovationseinheit der Bundesbank entwickelt wurde. Er dient zur Analyse der Länderberichte des Internationalen Währungsfonds (IWF). Diese Länderberichte sind textlich umfangreich und müssen für die Hausleitung stets zeitnah aufbereitet werden. Mithilfe des IWF-Wissensassistenten gelingen diese Vorarbeiten nun schneller und weniger personalintensiv. So werden gezielte Abfragen möglich, beispielsweise von Daten zur wirtschaftlichen Entwicklung oder von Abweichungen zwischen den in den Berichten enthaltenen Prognosen und den tatsächlichen Entwicklungen. Zum Befragungszeitpunkt befand sich der Wissensassistent noch in der Testphase.

Einbettung von Leitbegriffen: Auch um die politischen Leitbegriffe der Hausleitung in die Arbeit des Ministeriums einzubetten, hat das Datenlabor des BMZ ein Sprachmodell trainiert.

Das Besondere an den Sprachmodellen ist ihre vielfältige Einsetzbarkeit. So hat sich etwa in einer Testphase herausgestellt, dass die Anwendung GPT-4U von den Mitarbeitenden für verschiedene Arten der Textarbeit eingesetzt wurde, für die den Ministerialmitarbeitenden keine für die Aufgabe dedizierte Software zur Verfügung stand – quasi als funktionaler »Lückenfüller«. Dieses Beispiel verdeutlicht, dass oftmals erst in der Anwendung die Fülle an Verwendungsmöglichkeiten der entwickelten Lösungen offenbar wird.

Die von den Datenlaboren entwickelten Sprachmodell-basierten Anwendungen erhalten besonders viel Aufmerksamkeit innerhalb der Ministerien, was teils auch kritisch gesehen wird:

»Eines der wenigen Themen, die Sichtbarkeit und Akzeptanz für Datenanwendungen und die Arbeit der Datenlabore schaffen, sind KI und Sprachmodelle. Das basiert auf dem noch wenig ausgeprägten Verständnis für technische Themen in der Verwaltung und das aktuell in den Medien sehr präsente Thema KI durch die Weiterentwicklung der Sprachmodell-technologie. Jedoch können wir auch mit wesentlich einfacheren Technologien und Anwendungen Mehrwert schaffen; z. B. durch Dashboards, regelbasierte Systeme oder einfacher Digitalisierung von Prozessen. Das ist dann aber nicht mehr so sexy.«

4.1.2 Dashboards

Die identifizierten Anwendungsfälle der Datenlabore münden aber auch in diverse Dashboard-Entwicklungen.

Steckbriefe: Datenprodukte, die viele Häuser angepasst auf ihre Bedürfnisse entwickeln, sind Steckbriefe zu bestimmten Themen, die nun in Form von Dashboards dargestellt werden. Hierbei geht es um die bedarfsgerechte, anschauliche Aufbereitung von Informationen. Ein Beispiel sind Steckbriefe zu Landkreisen mit Informationen zur Regierung, Wahlergebnissen, Erhalt von Fördermitteln des Ministeriums sowie weiteren ressortspezifischen Daten, die unter anderem der Vorbereitung von Reisen dienen. Für die Ministeriumsmitarbeitenden bedeutet diese Anwendung eine deutliche Arbeitserleichterung, weil Informationen, die sie zuvor händisch zusammentragen mussten, nun automatisiert vorliegen.

Visa-Dashboard: Im Datenlabor des AA wurde ein Visa-Dashboard erarbeitet. Hierauf kann echtzeitnah abgelesen werden, wie sich die deutschen Visazahlen weltweit entwickeln, wie lang die Wartezeiten sind, ob man das eigene Auftragsvolumen erfüllen kann und ob die Ziele im Hinblick auf die Fachkräfteeinwanderung erreicht werden.

Migrationsdashboard: Das Migrationsdashboard des BMI beziehungsweise des Statistischen Bundesamtes bildet mit aggregierten und anonymisierten Daten die gegenwärtig Migrationslage ab und unterstützt das Fachpersonal in Behörden bei der Planung und Entscheidung zur Unterbringung und Integration von Geflüchteten. Alle 16 bundesdeutschen Länder und zwei Drittel der Landkreise nutzen das Tool bereits.

Veranschaulichung des Photovoltaikausbaus: Im Datenlabor des Bundeskanzleramtes (BKAMt) wurde basierend auf den Daten der Bundesnetzagentur ein Dashboard entwickelt, das den Photovoltaikausbau in Deutschland visualisiert.

Veranschaulichung der Mobil- und Breitbandversorgung: Im Datenlabor des Bundesministeriums für Ernährung und Landwirtschaft (BMEL) wurde die Funky-Dash-App entwickelt, welche die Mobil- und Breitbandversorgung in ländlichen Räumen auf Landkreisebene visualisiert.

Weiterbildungsformate: Die Entwicklung von Weiterbildungsformaten wurde ursprünglich vom Bundesministerium für Bildung und Forschung (BMBF) federführend begonnen, die Durchführung jedoch bald überwiegend vom Datenlabor des BMZ übernommen. Das BMZ bewirbt seine Fortbildungsangebote innerhalb der Datenlabor-Community. Die Angebote bestehen aus thematisch fokussierten Onlinesitzungen und vermitteln unter anderem grundlegende Kenntnisse zu Themen der Datenkompetenz, zur Nutzung von Programmiersprachen, zu Visualisierungstechniken und zu relevanten KI-Anwendungen.

»**Cheat-Sheets**«: Ein anschauliches Beispiel für niederschwellige Weiterbildungsformate sind auch die Veröffentlichungen von Handreichungen und Anleitungen, sogenannte Cheat-Sheets. In diesen Orientierungshilfen veröffentlicht ein Datenlabor Statistiksätze und Schritt-für-Schritt-Anleitungen, wie man beispielsweise bei der Suche bestimmter Daten vorgeht.

»**Datenanalytische Kaffeepause**«: Auch eine Datenanalytische Kaffeepause wurde als Format getestet, bei dem in zehnmütigen Minivorträgen vor Ort im Ministerium informiert und die Arbeit des Datenlabors vorgestellt wird.

Experimentier- und Lernraum: Ein Datenlabor etabliert mit einem Experimentier- und Lernraum einen physischen Raum für kollaboratives Arbeiten, für Seminare und Veranstaltungen. Der Raum wird für Videokonferenzen multimedial ausgestattet und wird über die erforderliche technische Ausstattung verfügen, um Datenanalysen durchführen zu können. Er soll ein Ort des gemeinsamen Kompetenzaufbaus werden.

4.2 Erhöhung der Datenkompetenz

Alle Datenlabore nehmen das in der Datenstrategie 2021 festgeschriebene Ziel der Datenkompetenzerhöhung ernst und setzen es auf ihre Weise um. Dabei wird der »*Lehrauftrag*« in den meisten Aktivitäten der Datenlabore von Anfang an mitgedacht. Das heißt, es gilt zu »*sensibilisieren*« und die »*Blackbox so klein wie möglich zu halten*«, wenn Kontakt mit den Fachreferaten besteht oder Datenprodukte entwickelt werden.

Darüber hinaus gibt es formalisierte Weiterbildungsangebote, wie thematisch fokussierte Videokonferenzsitzungen (Aufzeichnungen davon finden sich auch auf YouTube) oder Kurse an der **Bundesakademie für öffentliche Verwaltung**, die von den Datenlaboren angeboten werden. Auch niedrighschwellige Angebote wie Handreichungen, Schritt-für-Schritt-Anleitungen und Kurzvorträge sind im Angebot. Die Aktivitäten im Bereich des Wissensaufbaus sind überaus vielfältig und werden mitunter auch über die Ressortgrenzen hinweg genutzt.

4.3 Aufbau der Dateninfrastruktur

Der wohl bedeutendste Output der Datenlabore besteht darin, die Infrastruktur für die Datenarbeit der Ministerien im Speziellen und der Bundesregierung insgesamt zu schaffen oder auszubauen. Die entstehende Datenarbeitsinfrastruktur hat dabei die unterschiedlichen Ausgangsbedingungen der Häuser zu berücksichtigen. In Bezug auf die eingebundenen Daten (schutzbedürftige interne Daten oder offene Daten) und die Art der Bereitstellung (zentral oder dezentral) schlugen die verschiedenen Datenlabore unterschiedliche Wege ein. Maßgeblich bei diesen Entscheidungen waren für alle Datenlabore die Ausbalancierung von Informationssicherheit und Nutzen.

4. Output der Datenlabore

Einige Datenlabore entschieden sich für die Nutzung der hausinternen IT-Infrastruktur. Dieses Vorgehen bietet nach Ansicht der Befragten wesentliche Vorteile, insbesondere wenn man sich auf das leistungsfähige IT-Team des Hauses verlassen könne und die eigene IT-Infrastruktur hochperformant sei, Voraussetzungen, die nicht alle Befragten als gegeben ansehen. Die Verfasser:innen der Nutzung interner IT-Infrastruktur sahen zum Beispiel einen wesentlichen Vorteil darin, einfach mit hausinternen, also hoch schutzbedürftigen Daten arbeiten zu können und anderen Mitarbeitenden Zugriff auf die Datenanalyse-Infrastruktur ermöglichen zu können:

»Ja, absolut, wir haben den Vorteil, dass wir mit allen hausinternen Daten dadurch grundsätzlich arbeiten können und jeden [Internen] auf unsere Infrastruktur lassen können, aus dem regulären System heraus und nicht sozusagen einen Rechner hier haben, mit dem wir sozusagen ins E-Mail-Postfach kommen, und daneben einen zweiten Rechner stehen haben, auf dem die Datenlaborinfrastruktur läuft, sondern das ist alles in einem System.«

Andere Datenlabore arbeiten aus Überzeugung ausschließlich mit offenen Daten und vermeiden interne Datenquellen aus Zeit- und Kostengründen:

»Das heißt, wir arbeiten eigentlich außerhalb der Zone mit hohem Schutzbedarf. Warum? Nach meiner Erfahrung ist der Faktor von Geschwindigkeit und Kosten, den Sie von einer offenen Umgebung in eine geschützte und von einer geschützten in eine Behördenumgebung haben, jeweils ungefähr 5 bis 10. Das heißt, Sie brauchen für eine Entwicklung in einer geschützten Behördenumgebung ungefähr 25mal so lang. Und das ist 25mal so teuer in etwa.«

Die Vorhaben zur Schaffung einer ressortübergreifenden Infrastruktur für die Datenarbeit sind vielfältig und reichen von der Bereitstellung von Datenkatalogen und Schnittstellen bis hin zur Etablierung gemeinsamer Plattformen wie **PLAIN** (siehe Infokasten). Als Zielvision geht es dabei darum, dass man *»nicht auf Inseln sitzt und sich über den Fluss oder den See zuruft, sondern sagt ›okay, hier ist auch ein Steg‹ oder noch besser: Es gibt eine zentrale Insel, auf der man sich auch mal treffen kann zum Arbeiten und zum Feiern von Erfolgen«*.

Portfolio-Atlas: Der Portfolio-Atlas des AA ermöglicht es, das deutsche Förderengagement im Ausland darzustellen. Er wurde vom Haushaltsausschuss des Bundestages in Auftrag gegeben.

Datenpool Proof-of-Concept: Der Proof-of-Concept (Machbarkeitsnachweis) wird durch das Datenlabor des BKAm koordiniert und soll die Verfügbarkeit maschinenlesbarer Daten der Bundesverwaltung erhöhen. Hierdurch sollen datengetriebene Prozesse und Entscheidungen in den Bundesbehörden, ein effektiveres Wissensmanagement, das ressortübergreifende Teilen von Daten und die Aufbereitung und Bereitstellung von Open Data unterstützt werden.

PLAIN als Vorschlag zur gemeinsamen Infrastruktur: PLAIN ist eine Analytics Plattform für die Bundesregierung, ein *»Werkzeugkasten, der von allen genutzt werden kann«* und wurde durch das AA in Zusammenarbeit mit der Bundesdruckerei entwickelt. Sie ist Multi-Mandanten-fähig und erlaubt es der Bundesverwaltung, auf einer einheitlichen Plattform geschützt mit Daten und KI analytisch zu arbeiten. Dabei soll PLAIN keine Fachverfahren ersetzen. Im Frühsommer 2023 ging Version 1.0 von PLAIN live. Jedes Bundesressort kann PLAIN buchen und benutzen und dort eigene Data-Science-Cases umsetzen und seiner Organisation zur Verfügung stellen. Aber auch die Kooperation zwischen Ministerien wird dadurch möglich, sodass zwei (oder mehr) Häuser etwa Daten und Code miteinander austauschen können. Im Herbst 2024 soll die Version 2.0 zur Verfügung gestellt werden, die noch mehr Freiheitsgrade für die Mandanten bieten soll und auf performanterer Hardware eines neuen Rechenzentrums aufbaut. Bis zum Jahr 2025 soll dann der Rollout des »finalen« PLAIN Stable folgen.

Es gibt jedoch auch Bedenken hinsichtlich der Etablierung von PLAIN als gemeinsame Infrastruktur. Die Kritik konzentriert sich dabei auf drei Aspekte: Zum einen wird bemängelt, dass die spezifischen Bedürfnisse anderer Ministerien und Behörden nicht ausreichend berücksichtigt werden, was die Eignung als universelle Lösung für die gesamte Bundesverwaltung einschränkt. Zum anderen stehen die Geschwindigkeit der Anwendung und die als überhöht empfundenen Kosten im Fokus der Beanstandungen.

4.4 Etablierung neuer Arbeitsweisen und Kooperationen

Die Kooperation und Vernetzung der Datenlabore wird als wichtige Triebkraft der Transformation hin zu einer datenbasierten Politikgestaltung angesehen und kann damit als ein weiterer wesentlicher Output der Arbeit der Datenlabore nach dem IOOI-Framework betrachtet werden. Kooperationen umfassen dabei erstens die hausinterne Zusammenarbeit, zweitens die

Vernetzung im Datenlaborökosystem mit anderen Ministerien oder öffentlichen Institutionen, und drittens die Kooperation mit der Wissenschaft, Zivilgesellschaft, Wirtschaft und der Öffentlichkeit insgesamt. Auf diese drei Bereiche wird im Folgenden gesondert eingegangen.

4.4.1 Hausinterne Kooperation

Die hausinterne Kooperation ist der wichtigste Wirkungskreis der Datenlabore. So fungieren die Labore generell als »Katalysator« und Schnittstellenmanager zwischen fachlicher, datenanalytischer, arbeitsprozessbezogener und technischer Arbeit. Diese Rolle füllen die Laborteams insbesondere proaktiv aus, indem sie gezielt mit Informationen, Angeboten und Bedarfsanfragen auf andere Abteilungen und nachgeordnete Behörden zugehen. Die Schnittstellenfunktion nehmen sie zu allen Seiten im Haus hin wahr und sind so Ansprechpartner in Fragen der Datenarbeit für alle Mitarbeitenden des Hauses. Die Aufgabe der Datenlabore besteht auch darin, die Bedarfe der Mitarbeitenden in technische und prozessbezogene Anforderungen zu übersetzen und darauf aufbauend an die Mitarbeitenden zu spiegeln, was möglich ist und was nicht.

In Fragen der hausinternen Kooperation ist es auch von Bedeutung, welche Nähe zwischen den Datenlaboren und den jeweiligen Hausleitungen besteht. In manchen Ministerien hat die Hausleitung ein direktes Interesse und ist über die wesentlichen Entwicklungen ihres Datenlabors im Bilde. In anderen Häusern besteht ein Interesse an spezifischen Datenprodukten, die besonders hohe Aufmerksamkeit und Interesse wecken, wie den Sprachmodell-basierten Anwendungen. Auch die nachgeordneten Behörden der einzelnen Ministerien sind wichtige Kooperationspartner für die Datenlabore, da diese häufig auf eine längere Erfahrung mit der Datenarbeit zurückblicken. Beispiele hierfür sind die Bundesbank oder die Arbeitsagentur.

Die konkrete Zusammenarbeit der Datenlabore mit anderen Einheiten ihres Hauses ist unterschiedlich intensiv. Sie reicht von der einmaligen Zusammenarbeit, z. B. im Rahmen bestimmter Anwendungsfälle oder der Teilnahme an Webinaren, bis hin zur institutionalisierten Zusammenarbeit in fester Formation. Die Arbeit an den Datenprodukten ist dabei in allen Häusern eine Kooperationsleistung zwischen dem Datenlabor und der jeweiligen Fach- oder Zentralabteilung. Die Zusammenarbeit unterscheidet sich jedoch im Grad ihrer Formalisierung. Manchmal geschieht die Zusammenarbeit ad hoc und fallbezogen, in anderen Fällen ist sie strukturiert und nahezu institutionalisiert.

Ein Beispiel für eine institutionalisierte Zusammenarbeit im Haus findet sich im BMUV. Hier hat das Datenlabor eine »Kernnutzer:innen-Gruppe« benannt, die aus 40 Ministeriumsmitarbeitenden unterschiedlicher Abteilungen besteht und ein weitgehend repräsentatives Abbild potenzieller Nutzender darstellt.

Diese Mitarbeitenden fungieren als Tester:innen und Ko-Konstrukteur:innen der Datenprodukte. Sie sind diejenigen, die »erste Anwendungen testen, anschauen, ob die funktionieren und sinnvoll sind und wie wir sie noch verbessern können«.

Ein Beispiel eher fallbezogener Zusammenarbeit findet sich bei der Entwicklung von GPT-4U (siehe S. 16). Um die Anwendung innerhalb des Ministeriums zu erproben, wurden verschiedene Abteilungen innerhalb des BMZ und bis zu 100 Mitarbeitende miteinbezogen. Ein weiteres Beispiel ist der IWF-Wissensassistent des BMF (siehe S. 16). Er wurde in Kooperation zwischen IT-Entwickler:innen und der Fachabteilung für den Währungsfonds entwickelt. Im Rahmen gemeinsamer Workshops beschrieb die Fachabteilung den IT-Entwickler:innen, welche Zwecke das Programm erfüllen muss. Diese wurden von Entwickler:innen wiederum technisch in die Lösung übersetzt.

Auch entscheiden die Datenlabore, inwiefern sie den Anspruch verfolgen, alle Mitarbeitenden des Hauses einzubeziehen oder sich in ihrer Arbeit auf jene zu fokussieren, die aktives Interesse zeigen. Meist ist es eine Kombination: Teilweise suchen die Datenlaborteams einzelne Abteilungen auf und fragen sie nach ihren Bedürfnissen im Hinblick auf die Datenarbeit, in anderen Fällen tragen auch Interessierte Probleme und Anwendungsfälle an das Team heran:

»Nicht immer müssen wir proaktiv auf andere zugehen. Es gibt auch ein Beispiel, wo es erfreulicherweise auch so funktioniert: die kommen mit ihrer Idee zu uns – >sagt mal, was möglich ist« – und dann schauen wir mal, ob wir die Möglichkeit haben, etwas mit ihnen gemeinsam zu entwickeln.«

So ging etwa das Datenlabor des BMZ proaktiv auf verschiedene Abteilungen zu, um sich vorzustellen und die Kolleg:innen nach ihren Wünschen zu fragen, respektive wo die vorhandenen Arbeitsinstrumente zur Datenanalyse nicht mehr ausreichten. Diesen Input übersetzte das Team des Datenlabors dann in kürzester Zeit in Anwendungen, eine davon gar binnen drei Stunden Arbeitszeit.

Auch bei anderen Datenlaboren kontaktierte das Team alle Referate, die fachlich-inhaltlich mit Daten arbeiten, und erfragte die dabei auftretenden Hürden. Im Rahmen dieser Kontakte stellte sich dann heraus, in Zusammenarbeit mit welchen Referaten besonders schnell Verbesserungen erreicht werden konnten. Auf den anschließenden Aufruf zur Zusammenarbeit war die Resonanz groß:

»Und seitdem das so ein bisschen angerollt ist, ist das Interesse auch extrem groß geworden an diesen Unterstützungsleistungen. Und so groß, dass man eigentlich schon sagen muss, dass man sich kaum davor retten kann, was ja gut ist, was aber eben auch zeigt, dass der Bedarf einfach da ist.«

4. Output der Datenlabore

Um die Anfragen zu systematisieren, die von den Ministeriumsmitarbeitenden an die Datenlabore herangetragen werden, haben einige Datenlabore feste Prozessabläufe etabliert. Ein solcher Prozessablauf sieht beispielhaft wie folgt aus:

Anfrage	Über ein Funktionspostfach adressiert die Fachseite eine strukturierte Anfrage basierend auf vorgegebenen Fragen an das Datenlabor
Erste Einordnung	Das Datenlaborteam nimmt eine erste Einordnung vor. Wichtige Kriterien dabei sind: <ul style="list-style-type: none">■ Umsetzbarkeit■ Neuheitsgrad für das Haus■ Neuheitsgrad der Funktionalität■ Potenzial zum Einfluss auf politische Entscheidungen■ Möglichkeit zur Nachnutzbarkeit innerhalb der Bundesregierung → Entscheidung, ob Vorhaben prioritär, später oder gar nicht bearbeitet wird
Erstes Vorgespräch	Ein Treffen zwischen Datenlabor und der Arbeitseinheit mit Austausch über Vorstellungen zum Datenprodukt findet statt
Zweite Einordnung	Das Datenlaborteam nimmt eine zweite Einordnung vor und bewertet, ob es die Kapazitäten für die Umsetzung hat
Start der Bearbeitung	Start der Bearbeitung unter Einbeziehung der Fachseite

Neben der konkreten ko-kreativen Arbeit an Datenprodukten bedienen sich die Datenlabore verschiedener Formate der Kommunikation und Interaktion innerhalb ihres Hauses, darunter fallen:

- Multiplikatorenveranstaltungen
- Berichte im Intranet
- Hausticker
- Hausinterne Newsletter
- Lunch-Formate
- Sonderaktionen
- Vorstellung eines zufriedenen Kunden
- Besuch der Abteilungen
- Vorstellung des Datenlabors auf Führungsklausuren
- Datensprechstunde

4.4.2 Kooperation im Datenlaborökosystem

Die Datenlabore des Bundes nutzen im Wesentlichen zwei Foren für den ressortübergreifenden Austausch. Die **Interministerielle Arbeitsgruppe Datenlabore (IMAG)** sowie ein selbstorganisiertes **Peer-to-Peer-Netzwerk**.

Im Rahmen der **Interministeriellen Arbeitsgruppe** tauschen sich alle Leiter:innen der Datenlabore unter Federführung des Leiters des Datenlabors des Bundeskanzleramtes aus. Zu diesen Treffen werden regelmäßig auch externe Gäste eingeladen.

Im Rahmen des **Peer-to-Peer-Netzwerkes** treffen sich neben den Leiter:innen auch die Mitarbeitenden der Datenlabore, hier hat sich eine Form der Selbstorganisation etabliert. Ein Beispiel für die Kooperation innerhalb dieses Netzwerkes ist der Austausch zum Datenprodukt Steckbriefe (siehe S. 16). Hier hat ein Datenlaborteam seinen Piloten zur Diskussion gestellt, Anmerkungen eingeholt und einen Testzugang für andere Datenlabore bereitgestellt.

Sowohl die IMAG als auch das Peer-to-Peer-Netzwerk werden von den Datenlaboren als überaus produktiv wahrgenommen:

»Ich finde es beeindruckend, in was für einem anderen Geist die Datenlaborakteure zusammenarbeiten. Ich habe das noch nie erlebt in ressortübergreifender Zusammenarbeit, dass man sich so unter die Arme greift, so offen miteinander redet, auch so einfach abstimmt, Code teilt, Ideen teilt, Erfahrungen teilt. Das ist wirklich super. Ich habe das noch nie, wirklich noch nie auch nur ansatzweise so gesehen.«

»Es ist gut, es ist auch, glaube ich, einmalig gut zwischen den Ministerien, aber das geht natürlich immer noch weiter.«

»Und dass wir unabhängig auch von der Farbe des Hauses, unabhängig eben von der politischen Großwetterlage, uns verknüpfen können, ist einmalig, bin ich ziemlich überzeugt.«

Trotz dieser positiven Einschätzung und der oftmals praktischen Anwendung des **Einer-für-Alle-Prinzips** stellt sich gerade bei der Bearbeitung ähnlicher Problemstellungen die Frage, wie die Koordination, trotz beispielsweise Unterschieden in der Fachlichkeit, gelingen kann. Ein Beispiel ist die Bearbeitung Kleiner Anfragen, mit der sich neben dem Datenlabor des BMUV mit seiner KI-gestützten Rechercheanwendung (KIRAS, siehe S. 15) auch andere Datenlabore beschäftigen:

»Es gibt eine Entwicklung bezüglich Kleiner Anfragen [...]. Diese tauchen regelmäßig in jedem Ministerium auf und da ist es so offensichtlich gewesen, dass aus Datenlaborsicht etwas verbessert werden kann. Diese Idee gab es in vielen Ministerien gleichzeitig. Zu dem Thema setzten wir uns nun zusammen, bestimmen eine Federführung und entwickeln gemeinsam ein Produkt als Lösung. Inhaltlich gibt es zwar Differenzen bzgl. der Fachlichkeit in den jeweiligen Ressorts. Aber interessant für uns ist die Vorgehensweise, die technisch einheitlich sein kann, zum Beispiel eine Suchfunktion [...] ist natürlich absolut identisch, und das kann man ja dann auch nutzen.«

4.4.3 Kooperation mit Öffentlichkeit und Zivilgesellschaft

Neben der Kooperation in und zwischen den Ministerien pflegen die Datenlabore den Austausch mit Akteuren der Zivilgesellschaft, der Wissenschaft und der interessierten Öffentlichkeit. Dies tun sie allerdings in unterschiedlicher Intensität. Einige Datenlabore, wie die des AA und des BMZ, kommunizieren offen nach außen an das Fachpublikum, legen ihre Arbeit und ihre Anwendungsfälle dar und veröffentlichen, wo möglich, auch die eigenen Arbeitsergebnisse. Andere Datenlabore wiederum sehen sich nicht in der Lage, eigene Öffentlichkeitsarbeit zu betreiben, da sie dafür weder die Reife noch die personelle Ausstattung haben. Zudem sollte an vielen Stellen zunächst die hausinterne Kommunikation gestärkt werden, damit die Mitarbeitenden nicht zuerst über externe Quellen von den Fortschritten der Datenlabore erfahren.

Manche Datenlabore forcieren konkret die Zusammenarbeit mit der Zivilgesellschaft. Dies geschieht etwa in Form des Data Innovation Labs des AA, angesiedelt am GovTech Campus, sowie des Civic Data Labs des Bundesministeriums für Familien, Senioren, Frauen und Jugend (BMFSFJ). Von diesen Einrichtungen erhoffen sich die jeweiligen Datenlabore insbesondere, dass die zivilgesellschaftlichen oder wissenschaftlichen Akteure so eine Stimme in wichtigen Prozessen, wie beispielsweise bei Standardisierungsfragen oder bei der Erarbeitung der KI-Charta des AA, bekommen. Hier öffnet sich ein Raum, Diskussionen führen zu können, bei denen *»nicht gleich der Bundesadler draufsteht«*.

Für die Datenlabore übergreifend kommuniziert der Leiter des Datenlabors des Bundeskanzleramtes und Federführer der IMAG:

»Ich habe [...] in unserem Ressortkreis nachgefragt, wer möchte denn was [...] hier auch präsentiert sehen – auch mal in der Öffentlichkeit [...]. Weil natürlich ist es so: Diese Datenlabore, die tun Gutes, die tun manchmal nur inkrementell Gutes, aber was wir bisher noch nicht so richtig gemacht haben, [...] ist, viel darüber zu reden und das ist jetzt ein Moment, wo wir [...] so weit sind, dass wir sagen können: »Wir wollen darüber reden.«« (Frederik Blachetta im Rahmen der Smart Country Convention, 08.11.2023)



5. Herausfordernde Wirkungsmessung

Die wesentlichen Zielsetzungen der Datenlabore liegen in der Stärkung evidenzbasierter Politik, dem Aufbau von Datenkompetenz in den Bundesministerien und der Effizienzsteigerung in der ministeriellen Arbeit durch Datenanwendungen. Ein essenzieller Zusatznutzen liegt zudem in den in diesem Zusammenhang etablierten ressortübergreifenden produktiven Kooperationsformen. Die Wirkungsmessung der Datenlabore gewinnt gerade im Rahmen der angestrebten Professionalisierung und Stabilisierung an Bedeutung, sieht sich aber auch spezifischen Problemen gegenüber, die über die schwere Fassbarkeit mancher Kriterien, wie der Stärkung von Evidenzbasierung, hinausgehen. Die Datenlabore unternehmen unterschiedliche Anstrengungen zur Wirkungsmessung, die, insbesondere aufgrund des Projektcharakters der Datenlabore, auch immer unter dem Druck steht, Erfolge nachweisen zu müssen. So bestand angesichts der Ungewissheit über die Fortführung der Datenlabore zum Erhebungszeitpunkt bei vielen Datenlaboren Unbehagen beim Formulieren von Erfolgsindikatoren:

»Und ich traue mich im Moment nicht, Ziele für das nächste Jahr auszugeben, wenn ich dann im März im Zweifel allen meinen Mitarbeitern eigentlich raten muss, sich woanders zu bewerben und denen eher helfe, einen geilen neuen Job zu finden, was ich dann nämlich tun würde für jeden Einzelnen.«

Einige Datenlabore nehmen gänzlich Abstand von messbaren Indikatoren, weil dies als »Bürde« empfunden wird, der Reifegrad des Labors dies noch nicht zulasse und zunächst die Herstellung der Arbeitsfähigkeit des Labors in den Mittelpunkt gestellt werden müsse.

Eine generelle Schwierigkeit der Erfolgsmessung besteht zusätzlich darin, dass zuvor ein Status quo erhoben werden müsste, um im Outcome eine Veränderung zu diesem Ausgangszustand nachweisen zu können. Veranschaulichen lässt sich dies am Beispiel der Erhöhung der Datenkompetenz, zu der keine Daten ex ante vorliegen. Auch die Zielmarke sei hier schwer zu ermessen, denn »niemand weiß, wie viel Datenkompetenz denn eigentlich notwendig ist«.

»Ja, das ganze Evaluationsthema ist unfassbar schwierig. Oft ist man ja versucht, gerade dann, wenn es schwer wird, den Input zu messen. Bestenfalls misst man irgendwie den Output. Aber was man ja eigentlich braucht, ist der Outcome und wir haben an vielen Stellen dieses Problem, beispielsweise wenn es jetzt um die Datenkompetenz geht.«

Über die eigene Wirksamkeit reflektieren die 14 befragten Datenlabore ungeachtet dieser Schwierigkeiten und haben unterschiedliche Formen gefunden, die eigenen Ziele in Indikatoren zu fassen oder dies zumindest perspektivisch zu erarbeiten.

Quantitative Indikatoren: Das Datenlabor des BMZ zieht als Erfolgsmaßstab beispielsweise die Anzahl der Teilnehmenden an den eigenen Webinaren heran. Rund 100 bis 200 Teilnehmende werden für die Webinare verzeichnet, bis zu 80 Prozent von ihnen stammen aus verschiedenen Bundesverwaltungen jenseits des BMZ und aus bis zu 20 verschiedenen Bundesbehörden. Damit habe man die gesetzten Ziele übertroffen. Für das Datenlabor des BMF wiederum könnte die Vorgehensweise im Innovation Lab des BMF vorbildhaft sein. Im Innovation Lab wird bereits mit **Key Performance Indikatoren (KPIs)** gearbeitet, die aktuell noch stark auf In- und Output fokussieren, perspektivisch aber auch Aussagen darüber treffen können, wie erfolgreich die Arbeit des Labors war. Die KPIs sollen dann beispielsweise abbilden, wie viele Menschen sich in die Prozesse eingebracht oder an der Entwicklung von Prototypen oder Strategieworkshops teilgenommen haben.

Qualitative Indikatoren: Leichter fällt den Datenlaboren hingegen die Etablierung qualitativer Maßstäbe für die eigene Arbeit. Eines der Datenlabore will sich daran messen lassen, inwiefern die entwickelten Datenprodukte im Haus »wirklich gebraucht werden«. Es gehe aber auch darum, »dass genug Leute und vor allem auch die richtigen Leute [...] erkannt haben, dass es eine hohe Relevanz hat«. Das Datenlabor des BMUV erfragt die subjektiv wahrgenommene Nützlichkeit der eigenen Anwendungen anhand verschiedener Kriterien, etwa im Rahmen der Arbeit mit ihrer Kernnutzer:innengruppe. Ein

5. Herausfordernde Wirkungsmessung

anderes Erfolgskriterium bezieht sich auf die Datenkompetenz. Frage man Ministeriumsmitarbeitende nach Anwendungsfällen, hätten diese häufig keine Vorstellung davon, was mit Data-Science-Anwendungen möglich sei. Diesen Vorstellungsraum erweitert zu haben, könne ebenfalls ein Kriterium zur Erfolgsmessung sein.

Zusammenfassend lässt sich festhalten, dass, obzwar eine abschließende Wirkungsmessung über alle Datenlabore zu diesem Zeitpunkt weder zielführend noch umsetzbar ist, alle

Datenlabore kreative Mittel und Wege gefunden haben, um in ihrem jeweiligen organisationalen Kontext auf die Ziele der Datenstrategie einzuzahlen und die Transformation hin zu mehr Evidenzbasierung und Datenkompetenz in den Bundesministerien voranzutreiben. Die dahinterstehenden Wirkungsbausteine sind in Abbildung 4 dargestellt.

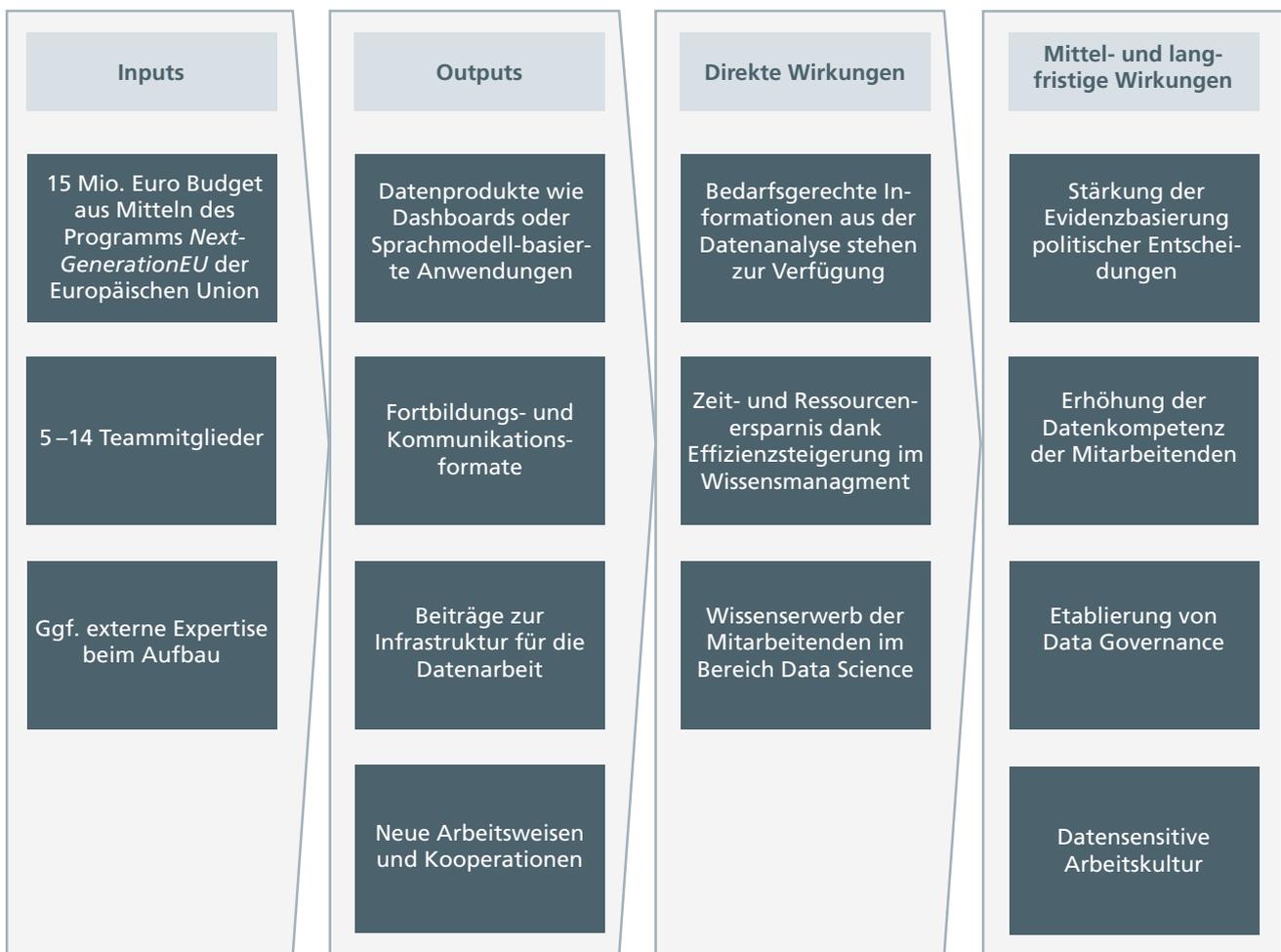


Abbildung 4: Schematische Darstellung der Wirkungsbausteine eines Datenlabors

6. Fazit: Wirkungspotenziale der Datenlabore

Die Datenlabore unterscheiden sich wesentlich in ihren Herangehensweisen und in ihrem Aufbau. Ausgehend von unterschiedlichen Rahmenbedingungen, Erfahrungen mit Datenarbeit, personeller Ausstattung und technischer Infrastruktur besitzen die Datenlabore auch unterschiedliche Wirkungspotenziale. Diese Potenziale zeigen sich bisher vor allem...

→ ... in der Etablierung essenzieller Infrastruktur für die Datenarbeit

Die Datenlabore arbeiten daran, die Grundlagen für die Datenarbeit in ihren Ministerien zu schaffen bzw. auszubauen. In dieser mühsamen und meist wenig außenwirksamen, aber unumgänglichen Grundlagenarbeit für datengetriebene Entscheidungsfindung, ist eines der größten Wirkungspotenziale der Datenlabore zu sehen. Diese Wirkung könnte mit einer längerfristigen gesicherten Finanzierung noch gestärkt werden.

→ ... in der Verbesserung des Wissensmanagements

Durch den Einsatz analytischer Verfahren und künstlicher Intelligenz wird das Wissensmanagement verbessert, was insbesondere vor dem Hintergrund des bevorstehenden Ruhestandes der Babyboomer-Generation und dem damit verbundenen Wissensverlust von großer Bedeutung ist. Die Datenlabore ermöglichen es beispielsweise, mithilfe Sprachmodell-basierter Anwendungen institutionelles Wissen niederschwellig zugänglich zu machen.

→ ...im Kompetenzaufbau

Alle Datenlabore zahlen auf verschiedene Weise auf den Datenkompetenzaufbau der Verwaltungsmitarbeitenden ein. Durch diesen internen Kompetenzaufbau können die Ministerien die gesteckten Ziele im Bereich der evidenzbasierten Politikgestaltung langfristig besser erreichen, eine zu große Abhängigkeit von externen Dienstleistern vermeiden und diese Dienstleister zielgenauer steuern.

→ ... in neuen Arbeitsweisen

Durch die Angebote und Anwendungen der Datenlabore werden die Mitarbeitenden an digitale, datenbasierte Arbeitsweisen gewöhnt und entsprechende Bildungsmaßnahmen in ihren Alltag integriert. Dies führt zu einer effizienteren und effektiveren Arbeitsweise, da insbesondere der Umgang mit unstrukturierten (Text-)Daten durch Datenanwendungen erheblich erleichtert werden kann. Für die Arbeit mit umfangreichen Dokumenten und Datenquellen bietet dies enorme Potenziale.

→ ... in der interministeriellen Zusammenarbeit

Die Datenlabore sehen sich selbst als ein Vorzeigebispiel im Hinblick auf interministerielle Kooperation und praktizieren eine ebenso offene wie zielführende Form der ressortübergreifenden Zusammenarbeit. Regelmäßig findet ein intensiver Austausch von Wissen, Ideen und Erfahrungen statt.

→ ... in der Schaffung der Grundlagen für evidenzbasierte Politikgestaltung

Durch ihre Arbeit tragen die Datenlabore dazu bei, die Dateninfrastrukturen der Ministerien technisch und methodisch zu verbessern und datenbasiertes Arbeiten in der Breite zu etablieren. Hiermit legen sie einen wesentlichen Grundstein für die Evidenzbasierung politischer Entscheidungen.



7. Handlungsempfehlungen

Aus der Betrachtung der Analyse der Arbeit der Datenlabore lassen sich die folgenden **Handlungsempfehlungen für politische Entscheidungsträger:innen** ableiten:

Datenbasierte Transformation der ministeriellen Arbeit weiter institutionalisieren.

Auf die grundsätzliche Frage, wie Datenpolitik in den Bundesministerien institutionalisiert werden sollte, sind die Datenlabore die aktuell einzige umfassende Antwort. Zwar mögen die Datenlabore noch der institutionellen Modifikation und Evaluation bedürfen, die Umsetzung eines vergleichbar umfassenden Ansatzes, Datenarbeit und Evidenzbasierung in den Ressorts zu verankern, ist gegenwärtig nicht absehbar. Durch die Datenlabore ist es geglückt, eine Infrastruktur für datenbasierte Arbeit zu schaffen. Kompetenzen, Personal und Wissen zu sichern, gelingt jedoch besser mit längerfristigen Planungshorizonten. Die wesentlichen Beiträge zur Schaffung einer Infrastruktur für die Datenarbeit in den jeweiligen Ministerien sollten durch eine kontinuierliche finanzielle Ausstattung langfristig angelegter Organisationseinheiten gesichert werden.

Datenlabore als Gemeinschaftsaufgabe aller Ressorts fördern.

Die Wirkung der Datenlabore entfaltet sich gerade auch durch ihr Selbstverständnis als Gemeinschaftsaufgabe aller Ressorts. Die Frage der Institutionalisierung von Datenarbeit in der Bundesverwaltung erfordert daher auch ein gemeinsames Vorgehen aller Ministerien. Eine weitere Fragmentierung in der Institutionalisierung der Datenlabore gilt es zu vermeiden, um den Potenzialen der interministeriellen Kooperation zur weiteren Entfaltung zu verhelfen.

Die Diversität der Bundesministerien auch bei der künftigen Ausgestaltung der Datenarbeit weiterhin berücksichtigen.

Den Datenlaboren gelingt es, die Besonderheiten ihrer Ministerien im Hinblick auf Arbeitskulturen, Datenaffinität oder Infrastruktur in ihrer Arbeit abzubilden und in maßgeschneiderte Anwendungen und Bildungsformate zu übersetzen. Diese Diversität sollte auch in der Institutionalisierung der Datenlabore

weiterhin Berücksichtigung finden. Gleichzeitig darf die Diversität nicht dem Erarbeiten einer gemeinsamer Dateninfrastruktur im Wege stehen.

Anstrengungen zur gemeinsamen Dateninfrastruktur verstärken.

Eine bessere Koordination bereits laufender Initiativen ist anzustreben und insbesondere dem Wunsch nach behördenübergreifendem Datenaustausch sollte noch stärker nachgekommen werden. Die nutzer:innen- und stakeholderzentrierte Durchführung eines solchen Vorhabens unter Berücksichtigung aktueller Interoperabilitätsstandards ist dabei essenziell.

Öffnung der Datenlabore nach außen vorantreiben.

Innovation wird durch das Zusammenwirken und voneinander Lernen vieler und heterogener Akteure begünstigt. Einige Datenlabore berücksichtigen dies bereits durch entsprechende Kooperationen. Mehrheitlich agieren die Datenlabore allerdings noch mit wenig Kontakt zu Akteur:innen außerhalb der Bundesverwaltung. Um von externem Wissen zu profitieren, sollten sich die Datenlabore langfristig stärker öffnen und auch die Expertise externer Akteur:innen aus Zivilgesellschaft, Wirtschaft und Wissenschaft stärker einbeziehen.

Den Raum für Experimente und gemeinsames Lernen beibehalten.

Auch wenn der Umgang mit Daten in allen Ressorts eine Daueraufgabe darstellt, ist es hilfreich, den Laborcharakter mit Raum für neue Experimente und weiteres gemeinsames Lernen zu wahren. So ist es möglich, vorab nicht im Detail planbare, aber dennoch wertvolle Entwicklungen zuzulassen, daraus zu lernen und sie gegebenenfalls nutzbar zu machen. Solche institutionalisierten Räume für organisationales Lernen außerhalb des üblichen Verwaltungsalltags sind insbesondere bei sich schnell entwickelnden Feldern wie den datenbasierten (KI-)Technologien und Anwendungen unabdingbar.

Literaturverzeichnis

Agora Digitale Transformation, 2024: *Forschungsprojekt E-Valuate. Wie wird das Regierungshandeln wirkungsvoller?* Unter Mitarbeit von Benedikt Göller und Vivien Benert. Online verfügbar unter <https://www.agoradigital.de/projekte/e-valuate>.

Bührer, Susanne; Feidenheimer, Alexander; Walz, Rainer; Lindner, Ralf; Beckert, Bernd; Wallwaey, Elisa, 2022: *Concepts and methods to measure societal impacts – an overview.* Discussion Papers Innovation Systems and Policy Analysis No. 74. Karlsruhe: Fraunhofer ISI. Online verfügbar unter <https://publica-rest.fraunhofer.de/server/api/core/bitstreams/06cee3b3-f386-4ef5-8026-26c9311e0923/content>.

Bundeskanzleramt, 2021: *Datenstrategie der Bundesregierung. Eine Innovationsstrategie für gesellschaftlichen Fortschritt und nachhaltiges Wachstum.* Kabinettsfassung, 27. Januar 2021. Die Bundesregierung. Online verfügbar unter <https://www.bundesregierung.de/resource/blob/974430/1960032/f073096a398e59573c7526feaadd43c4/2021-08-12-datenstrategie-deutsch-data.pdf?download=1>.

Bundesministerium für Digitales und Verkehr, 2023: *Nationale Datenstrategie der Bundesregierung.* Online verfügbar unter <https://bmdv.bund.de/SharedDocs/DE/Anlage/K/nationale-datenstrategie.pdf>.

Kupi, Maximilian; Jankin, Slava; Hammerschmid, Gerhard, 2022: *Data Science und KI in der Verwaltung. Warum Organisationen des öffentlichen Sektors eigene Expertise für Data Science und künstliche Intelligenz benötigen.* Policy Brief. Hg. v. Centre for Digital Governance & Data Science Lab. Hertie School. Online verfügbar unter https://hertieschool-f4e6.kxcdn.com/fileadmin/2_Research/2_Research_directory/Research_Centres/Centre_for_Digital_Governance/5_Papers/Other_papers/HS_Policy_Brief_German_Final_Version_Print.pdf.

Kontakt

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

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

www.fokus.fraunhofer.de
www.oeffentliche-it.de
X: @OeffentlicheIT
[mastodon.social/@OeffentlicheIT](https://mstdn.social/@OeffentlicheIT)

ISBN: 978-3-948582-29-6

