



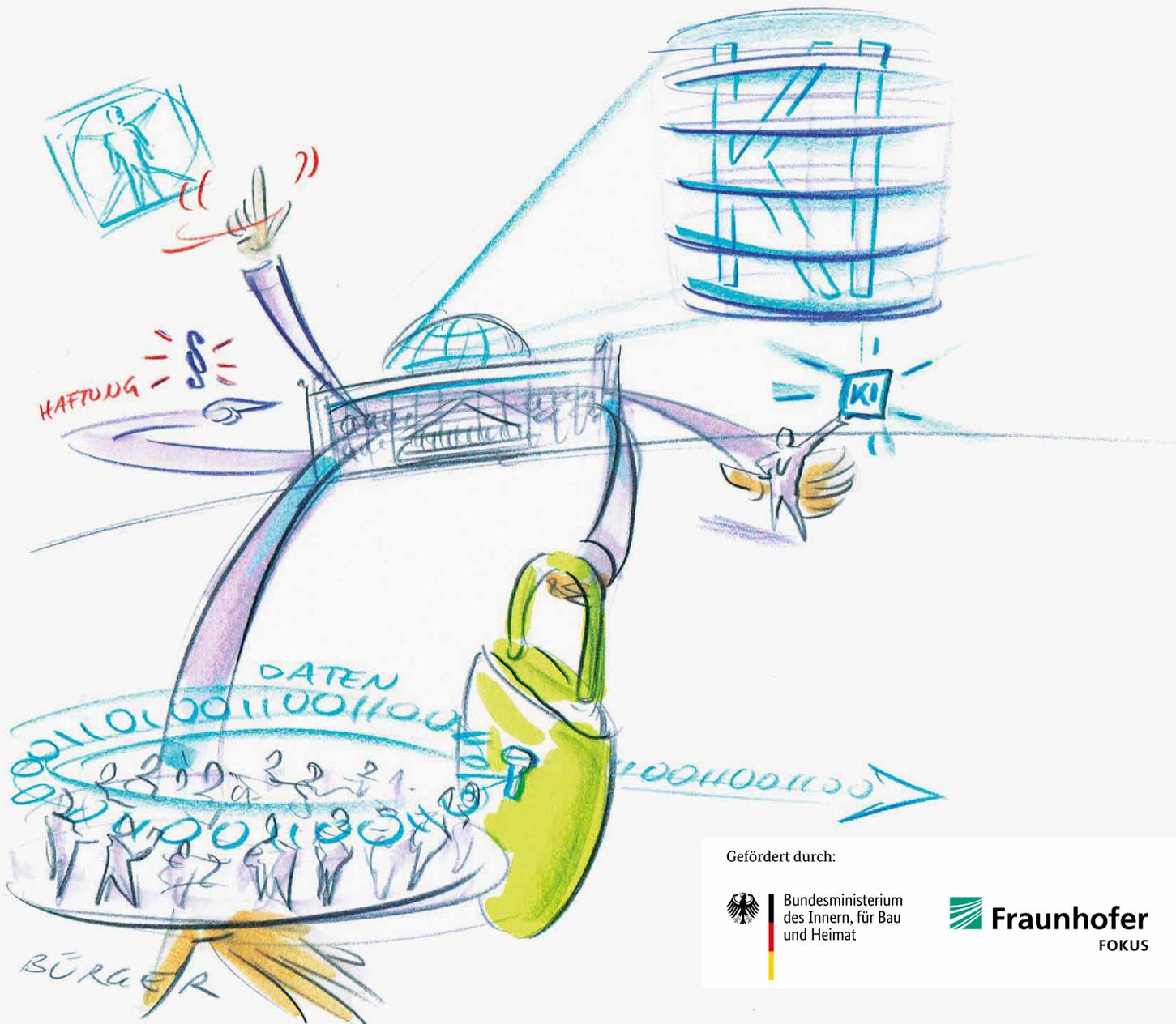
Kompetenzzentrum
Öffentliche IT

iRights.Lab 
Think Tank für die
digitale Welt

FORSCHUNG FÜR DEN DIGITALEN STAAT

KI IM BEHÖRDENEINSATZ: ERFAHRUNGEN UND EMPFEHLUNGEN

Jan Engelmann, Michael Puntschuh



Gefördert durch:



Bundesministerium
des Innern, für Bau
und Heimat

 **Fraunhofer**
FOKUS

IMPRESSUM

Autoren:

Jan Engelmann, Michael Puntschuh

Weitere Mitwirkende:

Annika Albert, Bastian Denker,
Dr. Wiebke Glässer, Anja Henschel,
Philipp Otto, Hannah Willing

Projektleitung:

Nicole Opiela, Dr. Mike Weber

Layout:

Reiko Kammer (Gestaltung),
Christoph Löffler (Grafiken), Heiko Stöber (Illustration)

Herausgeber:

Kompetenzzentrum Öffentliche IT
Fraunhofer-Institut für Offene Kommunikationssysteme FOKUS
Kaiserin-Augusta-Allee 31, 10589 Berlin
Telefon: +49-30-3463-7173
Telefax: +49-30-3463-99-7173
info@oeffentliche-it.de
www.oeffentliche-it.de
www.fokus.fraunhofer.de
Twitter: @OeffentlicheIT

Auftragnehmer:

iRights.Lab GmbH
Schützenstraße 8
10117 Berlin
Telefon: +49 30 40 36 77 230
Telefax: +49 30 40 36 77 260
kontakt@irights-lab.de
https://irights-lab.de
Twitter: @IRightsLab
Instagram: @iRights.Lab
Facebook: @iRights.Lab

ISBN: 978-3-948582-04-3

1. Auflage Dezember 2020

Dieses Werk steht unter einer Creative Commons Namensnennung 3.0 Deutschland (CC BY 3.0) Lizenz. 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.

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

ZUSAMMENFASSUNG

Die vorliegende Kurzstudie liefert auf Basis einer qualitativen Umfrage neue Erkenntnisse zum Einsatz von KI-Technologien in der öffentlichen Verwaltung in Deutschland. Dabei stehen neben Prozessen in der bürgernahen Kernverwaltung (etwa Dokumenten- und Antragsmanagement) auch Bereiche im Blickfeld, die z. B. die datengestützte Verkehrsplanung in Kommunen oder das Gebäudemanagement von Behörden erleichtern. Die Abfrage und Bewertung der 12 untersuchten Fallbeispiele stützte sich dabei auf neun zentrale Kriterien. Die vorliegende Analyse ermöglicht es anderen geplanten KI-Projekten, aus den bereits gemachten Erfahrungen in der Verwaltung zu lernen.

In der Gesamtschau der Beispiele lassen sich über die unterschiedlichen Ausgangsprobleme und Anwendungskontexte hinweg zentrale Erkenntnisse ableiten:

- Viele der untersuchten Projekte stellten an den Anfang eine klare Zielformulierung und eine vorherige Analyse möglicher Zielkonflikte. Einige KI-Anwendungen gehen darüber hinaus und beziehen die späteren Anwender:innen (Domänenwissen) bei der iterativen Entwicklung der Software-Architektur intensiv mit ein.
- Die Überlegungen zum Wirkungsmonitoring sind teilweise noch unausgereift oder unpräzise. Ebenso fehlen konkrete Maßnahmen zur Transparenz und Nachvollziehbarkeit der Funktionsweisen der Systeme.
- Darüber hinaus fehlen in bestimmten Bereichen gute Praktiken, die das Vertrauen der Beschäftigten und ggf. Klient:innen in die jeweiligen KI-Anwendungen erhöhen könnten. So gibt es zwar vereinzelt erste Ansätze wie Explainable AI oder Schulungen, erstere sind jedoch nicht immer für alle Beteiligten verständlich und letztere bilden noch die Ausnahme.
- Weiterhin ist eine nur sehr begrenzte Einbindung externer Expertise, insbesondere von zivilgesellschaftlichen Organisationen, zu beobachten. Nur eines der untersuchten Projekte arbeitet dezidiert und umfassend mit der organisierten Zivilgesellschaft zusammen.
- Insgesamt wünschen sich die Projektverantwortlichen mehr Experimentierräume und Offenheit aufseiten ihrer öffentlichen Auftraggeber:innen respektive anderer Abteilungen. Auch plädieren sie für einen stärkeren Austausch zwischen den Verwaltungseinheiten, um z. B. Synergien zwischen KI-Anwendungen des Bundes und der Länder zu schaffen und gelungene Beispiele von KI auf andere Anwendungsfelder zu übertragen.

- Als Vorzeige-Beispiele können jene Projekte fungieren, die von Anfang an ein Querschnittsdenken beherzigen und eine Anschlussfähigkeit und Skalierbarkeit des eigenen Projekts für andere Verfahren mitplanen. Insbesondere hieran zeigt sich, dass die Aufgabe des Silodenkens für digitale Innovationen in der Verwaltung gewinnbringend sein kann.

Die Handlungsempfehlungen am Schluss der Studie zielen auf mehr Wirkungsanalyse bei der Bewertung öffentlicher KI-Projekte, konsequentere Bemühungen um (technische) Nachvollziehbarkeit sowie (soziale) Akzeptanz. Desgleichen wird der Aufbau eigener datenanalytischer Expertise angemahnt, um sich von existierenden Abhängigkeiten schrittweise zu lösen. Hoffnung gibt der Aufbau eigener informeller *communities of practice* in der öffentlichen Verwaltung, denn nur praktizierte Offenheit und geeignete Strukturen für Wissenstransfer könnten den Sektor langfristig vom Getriebenen zu einem Treiber im Themenfeld KI machen.

DIE ANALYSE ZEIGT,
DASS DER KI-EINSATZ IN DER
DEUTSCHEN ÖFFENTLICHEN VERWALTUNG
UNTER BEACHTUNG NACHAHMENSWERTER
LEITLINIEN ERFOLGT.

INHALTSVERZEICHNIS

	Zusammenfassung	3
1.	Einleitung	5
2.	Neun Kriterien für KI-Best Practices	7
3.	Vorstellung der Fallbeispiele	11
4.	Bewertung der Fallbeispiele	15
5.	Erkenntnisse zum KI-Einsatz in der deutschen Verwaltung	19
6.	Fünf Handlungsempfehlungen	20
	Annex: Steckbriefe	21
	1 – ARC-D	21
	2 – C-19	22
	3 – Erkennung von Studienbescheinigungen für Kindergeld	23
	4 – Handwerksstatistiken	25
	5 – Indoor Robot	26
	6 – KI4LSA	27
	7 – Mit KI kinderleicht zum Kindergeld	28
	8 – KI gegen Kinderpornografie (informelle Bezeichnung)	30
	9 – KI zur Steuerung der Fallprüfung im Regressmeldeverfahren	31
	10 – Maschineller Übersetzungsservice der Justiz	32
	11 – PREVIEW	33
	12 – Profilanalyse von Gesprächsprotokollen	34

1. EINLEITUNG

Künstliche Intelligenz wird aktuell breit diskutiert: Dabei geht es bspw. um Chancen des Einsatzes von KI, Herausforderungen für die mögliche Regulierung oder Risiken, die mit der Nutzung von KI-Systemen einhergehen. Dieser Diskurs um einen unscharfen Begriff¹ ist längst auch in der öffentlichen Verwaltung angekommen.² Hier sind der digitale Anpassungsdruck und die Erwartungen der Bürger:innen an eine nachzuholende Digitalisierung hoch. Deshalb setzt sich die öffentliche Verwaltung zunehmend mit dem konkreten Nutzen von KI-Systemen z. B. für eine effektivere Gestaltung von Prozessen, personalisierte Angebote und eine gleichzeitige Entlastung des Personals auseinander.³ Dabei bilden nicht immer Effizienzgewinne die Begründung für den KI-Einsatz, sondern häufig auch eine Qualitätssteigerung bei der Erbringung staatlicher Leistungen.

Auch wenn eine solche KI-gestützte öffentliche Verwaltung flächendeckend noch keine Realität ist, zeigen erste Anwendungen schon heute, wohin der Weg führen kann. Sie demonstrieren, wie KI-Anwendungen entwickelt und eingesetzt werden, um Verwaltungsprozesse zu vereinfachen und zu beschleunigen,

Bürger:innen neue Informationskanäle zur Verfügung zu stellen sowie Entscheidungsprozesse und Planungen zu verbessern. Diese Beispiele bilden aber aktuell eher die Ausnahme: Ein Whitepaper des praxisnahen Austauschforums NExT-Netzwerk, das Bundes-, Landes- und Kommunalbeschäftigte zum Thema digitale Verwaltung versammelt, kam im April 2020 zu dem Ergebnis, dass im öffentlichen Sektor »entwickelte KI nur in einer überschaubaren Zahl von Fällen genutzt [wird]«⁴.

Um den aktuellen Stand des Einsatzes von KI-Systemen in der öffentlichen Verwaltung zu erfassen, ist ein strukturierter Überblick über deren unterschiedliche Einsatzbereiche erforderlich. Ein vollständiges Mapping kann die vorliegende Kurzstudie nicht leisten, dafür präsentiert sie jedoch stellvertretend 12 Beispiele von KI-Systemen, die bereits heute im öffentlichen Sektor eingesetzt werden oder sich im Forschungs- und Entwicklungsstadium befinden. Es geht ausdrücklich nicht um die Erstellung einer Rangliste, sondern um die Gewinnung von Bausteinen für eine gute Praxis.

¹ Mit Blick auf die Genese des Begriffs, der zum ersten Mal in einem Förderantrag an die Rockefeller Foundation im Jahr 1956 auftauchte, sprechen Paul Nemitz und Matthias Pfeffer gar von einem »Marketinginstrument erster Ordnung« in: dies., Prinzip Mensch. Macht, Freiheit und Demokratie im Zeitalter der Künstlichen Intelligenz, Berlin 2020, S. 37. Im Rahmen der Recherche wurde deutlich, dass das populäre Etikett »KI« auch für Anwendungen verwendet wird, die nichts mit KI zu tun haben – sicherlich kein singuläres Problem im Zuge sich ständig ablösender (medialer) Technologie-Hypes. Die Autor:innen der vorliegenden Studie arbeiteten mit einer pragmatischen Definition von KI, die darunter all jene informatischen Verfahren versteht, die sich von rein deterministisch arbeitenden Algorithmen (im Sinne von regelbasierten Abläufen) unterscheiden und als »selbstlernende«, mit Daten oder Feedback trainierte, Systeme begriffen werden können (überwachtes Lernen, unüberwachtes Lernen, bestärkendes Lernen). Diese Verfahren werden eingesetzt, um Muster und Anomalien zu erkennen, Prognosen zu erstellen, Empfehlungen zu geben und Entscheidungen vorzubereiten. Durch ihre Outputs wirken sie auf ihre virtuelle und reale Umwelt. Eine tiefere Einführung in die unterschiedlichen Methodiken von KI liefern Christian Welzel und Dorian Grosch im ÖFIT-Trendsonar Künstliche Intelligenz, Berlin 2018, <https://www.oeffentliche-it.de/documents/10181/14412/Das+%C3%96FIT-Trendsonar+K%C3%BCnstliche+Intelligenz> (letzter Aufruf: 21.10.2020).

² Christian Djeffal: Künstliche Intelligenz in der Verwaltung. Hrsg. vom Nationalen E-Government-Kompetenzzentrum e.V., Berlin 2018, S. 5.

³ Dies wird bspw. beim Blick auf die aktuellen Themen bei Konferenzen der öffentlichen Verwaltung deutlich. Ebenso wurde dieser Anspruch in der KI-Strategie der Bundesregierung aus dem Jahr 2018 festgestellt: »Der Einsatz von KI bietet im Bereich der öffentlichen Verwaltung die Chance, Informationen und Leistungen zielgerichteter, passgenauer und niedrigrschwelliger für Bürgerinnen und Bürger sowie verwaltungsintern bereitzustellen.«, in: Strategie Künstliche Intelligenz der Bundesregierung, 2018, S. 31, <https://www.bundesregierung.de/resourcelblob/975226/1550276/3f7d3c41c6e05695741273e78b8039f2/2018-11-15-ki-strategie-data.pdf?download=1> (letzter Aufruf: 21.10.2020)

⁴ Dr. Iliya Nickelt/NExT-Werkstatt »Digitale Projekte«, KI-Projekte in Behörden beginnen (Whitepaper), V01, Berlin 2020.

ZUM AUFBAU DER STUDIE:

Zunächst wird eine von uns entwickelte Kriterienmatrix⁵ mit neun Kriterien für Best Practices beim KI-Einsatz in der Verwaltung vorgestellt (Kapitel 2). Im Anschluss daran werden die untersuchten Fallbeispiele kurz skizziert (Kapitel 3). Anhand der zuvor entwickelten Kriterienmatrix werden die Fallbeispiele bewertet und einzelne Elemente nachahmenswerter Praxis hervorgehoben (Kapitel 4). Auf Grundlage der Bewertung werden übergreifende Erkenntnisse zum Einsatz von KI in der öffentlichen Verwaltung dargestellt (Kapitel 5). Abschließend werden politische Handlungsempfehlungen formuliert, wie die Entwicklung und der Einsatz solcher Best Practices unterstützt werden können (Kapitel 6).

Steckbriefe zu allen 12 Fallbeispielen befinden sich im Anhang dieser Studie. Die Steckbriefe wurden jeweils durch semistrukturierte Interviews mit mindestens einer projektverantwortlichen Person entwickelt und abgestimmt. Die Beispiele wurden hinsichtlich föderaler Ebene, Anwendungsfeld, Funktionsumfang und Kontextbedingungen möglichst vielfältig ausgewählt.

⁵ Zu diesem Zweck wurden in einem ersten Schritt existierende Prinzipienkataloge für »gute KI« dahingehend untersucht, ob sie auch im Hinblick auf digitalisiertes Verwaltungshandeln wichtige Orientierungspunkte bieten. Vgl. Jobin, A., Ienca, M. & Vayena, E., 2019: The global landscape of AI ethics guidelines, in: Nature Machine Intelligence 1, S. 389 – 399, <https://doi.org/10.1038/s42256-019-0088-2> (letzter Aufruf: 20.10.2020). Hagendorff, T., 2020: The Ethics of AI Ethics: An Evaluation of Guidelines, in: Minds and Machines, S. 99 – 120, <https://link.springer.com/article/10.1007%2Fs11023-020-09517-8> (letzter Aufruf: 20.10.2020). Konkrete Kriterienkataloge und Studien, die in die Entwicklung der Kriterienmatrix einfließen, waren zudem: Bertelsmann Stiftung / iRights.Lab 2019: Algo.Rules, https://algorules.org/fileadmin/files/alg/Algo.Rules_DE.pdf (letzter Aufruf: 27.11.2019); Hallensleben, S. et al., 2020: From Principles to Practice – An interdisciplinary framework to operationalise AI ethics, <https://www.ai-ethics-impact.org/resource/blob/1990526/c6db9894ee73aefa489d6249f5ee2b9f/aieig---report---download-hb---en-data.pdf> (letzter Aufruf: 21.10.2020); High-Level Expert Group on Artificial Intelligence, 2019: Ethics Guidelines for Trustworthy AI, https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60419 (letzter Aufruf: 21.10.2020); Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (OECD), 2019: OECD Council Recommendation on Artificial Intelligence, <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449> (letzter Aufruf: 21.10.2020); Datenethikkommission der Bundesregierung, 2019: Gutachten der Datenethikkommission, https://datenethikkommission.de/wp-content/uploads/191128_DEK_Gutachten_bf_b.pdf (letzter Aufruf: 21.10.2020); Diakopoulos, N. et al., 2016: Principles for Accountable Algorithms and a Social Impact Statement for Algorithms, <https://www.fatml.org/resources/principles-for-accountable-algorithms> (letzter Aufruf: 21.10.2020) sowie Future of Life Institute, 2017: Asilomar AI Principles, <https://futureoflife.org/ai-principles/> (letzter Aufruf: 21.10.2020). Diese allgemeinen Prinzipien mussten für den Bereich der öffentlichen Verwaltung konkretisiert und insbesondere um Kriterien für Innovation ergänzt werden. Hierbei standen Aspekte der Disruption, Skalierungsfähigkeit, Nachnutzung und Kooperationsbereitschaft im Vordergrund. Gestützt auf diese Vorüberlegungen wurden die Kriterien entlang einer Phasenbildung geordnet: der Prozess der Bedarfsermittlung, die Programmierung des Systems, der Übergang in den Produktivbetrieb und schließlich die Evaluation der Wirksamkeit. Die Anordnung der Kriterien ist damit nicht hierarchisch, sondern sequenziell.

2. NEUN KRITERIEN FÜR KI-BEST PRACTICES

Um die vorliegenden KI-Fallbeispiele bewerten zu können, haben sich die Autor:innen an internationalen Meta-Analysen von ethischen Kriteriensets im Bereich KI und algorithmische Entscheidungssysteme (ADM) orientiert. Die dabei am häufigsten genannten Kriterien wurden in der Folge daraufhin untersucht, ob sie auch hinreichend Plausibilität für die digitale Transformation der öffentlichen Verwaltung (z. B. neue Anforderungen an die staatliche Leistungserbringung, Reorganisation und Verbesserung von Arbeitsprozessen, Vorgaben zur Ausgabensenkung, Anforderungen an Servicequalität) aufweisen. War dies nicht der Fall, wurden andere Kriterien hinzugefügt. In einem dritten Schritt wurden die Kriterien auf neun reduziert, um eine Operationalisierung in Form semistrukturierter Interviews zu gewährleisten. Die Indikatorik innerhalb der Kriterienmatrix (s. folgende Abb.) bilden idealtypische Maßnahmen bzw. default settings, die zu unterschiedlichen Graden erfüllt sein können.

Zielorientierung: Das Kriterium ist angelehnt an die allgemeinen Grundsätze guten Projektmanagements und die Prinzipien der Wirkungsorientierung, siehe dazu: Kursbuch Wirkung. Bettina Kurz/Doreen Kubek: Das Praxishandbuch für alle, die Gutes noch besser tun wollen. Berlin: Phineo 2013, <https://www.phineo.org/kursbuch-wirkung> (letzter Aufruf: 12.11.20)

Veränderungsumfang: Interne Veränderung wird im Rahmen dieser Studie als disruptiver Effekt für die eigene Verwaltungseinheit bzw. deren Arbeitsorganisation verstanden. Das Kriterium liegt nahe an dem der Innovation, nimmt aber vor allem jene Standardprozesse in den Blick, die durch den Einsatz von KI-Technologien beeinflusst bzw. neu aufgesetzt werden.

Innovationsmarker: Das Kriterium ist angelehnt an Jürgen Hauschildts Fassung des Innovationsbegriffs, in: ders., Innovationsmanagement, 5. Auflage, München 2011. In diesem Zusammenhang sei auch auf den Innovationskompass verwiesen, ein kollaboratives Sammelprojekt unter Beteiligung von Politics for Tomorrow, Fraunhofer FOKUS und anderen: <https://www.innovationskompass.net/>

Projektmanagement: Eine sehr kompakte Einführung in das Thema bieten Hans-Dieter Litke, Ilonka Kunow, Heinz Schulz-Wimmer: Projektmanagement, Freiburg 2010.

Sicherheit: Das Kriterium findet sich in den meisten ethischen Kriteriensets zu Algorithmic Decision Making (ADM) und Künstlicher Intelligenz wieder, belegt in den Meta-Analysen von Jobin, A./Ienca, M./Vayena, E., 2019: The global landscape of

AI ethics guidelines, in: Nature Machine Intelligence 1, S. 389 – 399 sowie Hagendorff, T., 2020: The Ethics of AI Ethics: An Evaluation of Guidelines, in: Minds and Machines, S. 99 – 120.

Datenhaltung und -qualität: Die einzelnen Indikatoren finden sich bspw. in den Kriterien »Justice« und »Transparency« des KI-Ethik-Labels wieder, Hallensleben, S. et al., 2020: From Principles to Practice – An interdisciplinary framework to operationalise AI ethics. Einzelne Aspekte sind auch den Orientierungsfragen zur »Überprüfung der technischen Elemente« der Algo.Rule 3 »Ziele und erwartete Wirkung dokumentieren« entnommen, siehe Bertelsmann Stiftung/iRights.Lab, 2019: Praxisleitfaden zu den Algo.Rules.

Wirkungsmonitoring: Das Kriterium ist angelehnt an quantitative und qualitative Methodiken des Outcome-Measurements. Siehe dazu einführend: Christian Schober/Volker Then: Praxishandbuch Social Return on Investment: Wirkung sozialer Investitionen messen, Stuttgart 2015. Zur Anschlussfähigkeit der Wirkungsorientierung für Ansätze des New Public Management siehe Gerhard Hammerschmid/Alexander Grünwald: Einführung der wirkungsorientierten Verwaltungssteuerung. Erfolge – Potentiale – Perspektiven, Wien 2014.

Nachvollziehbarkeit: Das Kriterium findet sich in den meisten ethischen Kriteriensets zu Algorithmic Decision Making (ADM) und Künstlicher Intelligenz wieder, siehe Kriterium 5 (Sicherheit). Mit Explainable Artificial Intelligence (XAI) hat sich ein eigener Forschungszweig etabliert, der technische Verfahren der Transparenzbildung bei den Ergebnissen von KI-gestützten Verfahren untersucht, siehe dazu einführend Paul Voosen: How AI detectives are cracking open the black box of deep learning, in: Science, 6. Juli 2017, <https://www.sciencemag.org/news/2017/07/how-ai-detectives-are-cracking-open-black-box-deep-learning> (letzter Aufruf: 12.11.2020)

Akzeptanz: Das Kriterium bezieht sich vor allem auf die Akzeptanz innerhalb der Verwaltungseinheit, muss z. B. bei Chatbots oder neuen Services für Bürger:innen aber auch deren Akzeptanz bzw. Beschwerdemöglichkeiten einbeziehen. Das Kriterium ist inspiriert von Ansätzen des Change Managements, die dem Faktor »Interne Akzeptanz« eine sehr hohe Bedeutung zuschreiben. Die Herstellung externer Akzeptanz findet besondere Berücksichtigung im neunten Prinzip der Algo-Rules »Beschwerden ermöglichen«, siehe <https://algorules.org/de/startseite>.



1. Zielorientierung

Idealerweise erfasst durch:

- 1.1 Ziele wurden in der Frühphase des Projekts klar definiert.
- 1.2 Ziele werden regelmäßig reflektiert und ggf. angepasst.
- 1.3 Ziele wurden gemeinsam mit Stakeholdern (Anwender:innen, Betroffene) erarbeitet.
- 1.4 Zielkonflikte wurden identifiziert und beschrieben sowie Schlussfolgerungen gezogen.
- 1.5 Nicht-Ziele sowie Grenzen des Anwendungsbereichs des Systems wurden definiert.
- 1.6 Messindikatoren für das Erreichen der Ziele wurden entwickelt.



2. Interne Veränderung

Idealerweise erfasst durch:

- 2.1 Das System wird für die Lösung eines Problems eingesetzt, das vorher nicht befriedigend gelöst werden konnte, bzw. das System wird nicht nur ergänzend zu menschlichen Handlungen eingesetzt, sondern ist substituierend oder gänzlich neu.
- 2.2 Einzelne Prozesse in der Verwaltung sind maßgeblich verändert worden.
- 2.3 Der Gesamtprozess in der Verwaltung, in den das System eingebettet ist, wurde maßgeblich verändert.
- 2.4 Die Entwicklung des KI-Systems hat neue Kooperationen zu anderen öffentlichen Institutionen, Hochschulen, zivilgesellschaftlichen Organisationen und/oder privaten Unternehmen hervorgebracht, die die Perspektive der involvierten Verwaltungseinheit erweitert haben.
- 2.5 Es besteht ein Austausch mit anderen Verwaltungseinheiten, die ein ähnliches System einsetzen (möchten).



3. Innovationsmarker

Idealerweise erfasst durch:

- 3.1 Das KI-System ist (unabhängig vom Eigenentwicklungsanteil) neu für die öffentliche Verwaltung.
- 3.2 Der Entwicklungs- und Einsatzprozess wird öffentlich dokumentiert.
- 3.3 Erfahrungen werden regelmäßig initiativ, z. B. bei Konferenzen, geteilt.
- 3.4 Das KI-System ist unter einer freien Lizenz (z. B. GNU General Public Licence) oder als Public Domain verfügbar.
- 3.5 Das System ist so entwickelt worden, dass es auch in andere Projektkontexte ohne größere Modifikationen implementierbar wäre (Skalierbarkeit).



4. Projektmanagement

Idealerweise erfasst durch:

- 4.1 Es gibt eine eindeutig und öffentlich benannte, für das Projekt verantwortliche Person.
- 4.2 Diese verantwortliche Person verfügt über das notwendige Wissen, alle Gestaltungsprozesse nachvollziehen zu können.
- 4.3 Das Projekt wird agil entwickelt. Es gab/gibt mindestens eine Explorationsphase sowie eine Testphase.
- 4.4 Innerhalb der Behörde sind Fachwissen, verfügbare Daten und ggf. Berechnungsmodelle ausreichend vorhanden, um die Ergebnisse des Gesamtsystems im Zweifel ohne größeren Aufwand durch Menschen überprüfen zu können.
- 4.5 Das System kann durch definierte Entscheidungsprozesse für Einzelentscheidungen sowie insgesamt außer Betrieb gesetzt werden.



5. Sicherheit

Idealerweise erfasst durch:

- 5.1 Regelmäßige Funktionstests finden statt (sowohl in festen zeitlichen Abständen als auch nach größeren Veränderungen des Systems).
- 5.2 Redundante/voneinander abgrenzbare technische Systeme und Komponenten sind vorhanden.
- 5.3 Es besteht ein Notfallplan für das Vorgehen bei Systemausfall.
- 5.4 Eine automatisierte Angriffserkennung ist eingebaut.
- 5.5 Eine (technische und organisatorische) Notabschaltung des Systems ist vorhanden.



6. Datenhaltung und -qualität

Idealerweise erfasst durch:

- 6.1 Ursprung und Aufbau/Zusammensetzung von Trainingsdaten, Inputdesign, Inputdaten und Outputdesign sind dokumentiert.
- 6.2 Trainingsdaten, Inputdesign, Inputdaten und Outputdesign werden jeweils zu ihrer Einführung sowie regelmäßig (in festen Intervallen sowie nach signifikanten Änderungen) auf Verzerrungen überprüft.
- 6.3 Diese Überprüfung wird (mithilfe) von externen Expert:innen durchgeführt.
- 6.4 Es bestehen versionierte Repositorien für Trainings- und Betriebsdaten.
- 6.5 Die Voraussetzungen für ein Re-Training des Systems sind definiert.
- 6.6 Trainingsdaten werden regelmäßig aktualisiert und erweitert mit dem Ziel, das System laufend zu verbessern.



7. Wirkungsmonitoring

Idealerweise erfasst durch:

- 7.1 Die Einhaltung der Indikatoren der Zieldefinition wird laufend überprüft.
- 7.2 Das Analyse- oder Berechnungsmodell ist überprüfbar.
- 7.3 Es finden regelmäßige Überprüfungen der weiteren Auswirkungen des Einsatzes des Systems statt, z. B. auf Arbeitsprozesse, auf Betroffene, auf die Kultur der jeweiligen Behörde etc.
- 7.4 Es finden regelmäßige Überprüfungen auf nicht-intendierte Wirkungen statt.
- 7.5 Diese Überprüfungen finden unter Einbeziehung externer Expert:innen und Stakeholder statt.
- 7.6 Es bestehen Schnittstellen, die eine unabhängige Überprüfung des Systems durch Externe ermöglichen.



8. Nachvollziehbarkeit

Idealerweise erfasst durch:

- 8.1 Es ist für jede Entscheidung nachvollziehbar, welche Daten verwendet wurden.
- 8.2 Informationen, die die Funktionsweise des Systems insgesamt beschreiben, sind öffentlich verfügbar (globale bzw. allgemeine Nachvollziehbarkeit).
- 8.3 Der Einsatz des Systems ist nach außen hin, insbesondere gegenüber Klient:innen, gekennzeichnet.
- 8.4 Wenn Klient:innen mit Entscheidungen konfrontiert sind, bei denen das System zum Einsatz kam, werden sie ohne Aufforderung über dessen Einsatz, die zugrunde liegenden Daten und die in diesem konkreten Fall relevanten Faktoren aufgeklärt (lokale bzw. spezifische Nachvollziehbarkeit).
- 8.5 Diese Informationen sind für Klient:innen, Führungskräfte und Externe (z. B. Betroffene) jeweils zielgruppenspezifisch und verständlich aufbereitet, z. B. in Textform oder als Infografik.
- 8.6 Für die Ermittlung jener Einstellungen im KI-Design, die für die Nachvollziehbarkeit der Ergebnisse erforderlich sind, werden technische Maßnahmen zu deren Kenntlichmachung getroffen. Diese Maßnahmen wurden bereits im Entwicklungsprozess berücksichtigt.



9. Akzeptanz

Idealerweise erfasst durch:

- 9.1 Mitarbeitende werden gestaffelt und angepasst an die jeweiligen Erfordernisse zum KI-System geschult.
- 9.2 Alle Mitarbeitenden der Organisationen werden über den Einsatz des Systems informiert.
- 9.3 Es stehen Schulungen zur Verfügung, die es Mitarbeitenden ermöglichen, ihre Kompetenzen fachspezifisch zu vertiefen.
- 9.4 Die Zufriedenheit der Anwender:innen innerhalb der Verwaltungseinheit wird regelmäßig gemessen.
- 9.5 Diese Zufriedenheit ist im Zeitverlauf nachweisbar gestiegen.
- 9.6 Es bestehen einfach zugängliche und geschützte interne Hinweisgebersysteme bzw. ein effektiver Schutz für Whistleblower.
- 9.7 Es bestehen einfach zugängliche und verständliche externe Beschwerdemechanismen.
- 9.8 Die Beschwerdemechanismen sind an rechtlich garantierte Beschwerdebearbeitungsverfahren gebunden. Durch Beschwerden aufgedeckte Fehler werden behoben.
- 9.9 Es wurde definiert, ab welchem Grad an Fehlern bzw. Beschwerden das System abgeschaltet wird.



3. VORSTELLUNG DER FALLBEISPIELE

In der Mehrzahl der Fälle wird KI zur **Entlastung** der Mitarbeitenden im öffentlichen Sektor genutzt, um Prozesse effizienter zu gestalten und Kapazitäten für andere Aufgaben zu schaffen oder Kapazitäten abbauen zu können. Darunter fällt etwa das mFund-Projekt **Automatisches Zustandsmonitoring von Autobahnen** mit KI – ARC-D (Automatic Road Condition Detection), das Videodaten aller Bundesautobahnen ausliest, klassifiziert und dadurch die Dringlichkeit von Instandsetzungsmaßnahmen anzeigt. Durch die perspektivische Einbindung von Bürger:innen, die z. B. via Dashcam oder Smartphone Straßenabschnitte erfassen könnten, wird der finanzielle und personelle Entlastungsaspekt gesteigert. Auch die **automatisierte Erkennung von Studienbescheinigungen** für die Kindergeldstellen der Agentur für Arbeit nutzt maschinelle Lernverfahren zur Beschleunigung von Antragsverfahren. Mit der KI-gestützten Validierung der eingereichten Dokumente werden Zeitressourcen für den Prüfaufwand an anderer Stelle frei. Bei den **Handwerksstatistiken**, angefertigt durch das Statistische Bundesamt, wird der generelle Mehrwert von KI-gestützten Musterprozessen deutlich: Während die Zuordnung von Daten zu einem Register vorher händisch erfolgte, haben KI-Verfahren die korrekte Zuordnung derart beschleunigt, dass inzwischen darüber nachgedacht wird, die Methodik auch in anderen statistischen Arbeitsfeldern einzusetzen. Der gerade pilotierte **maschinelle Übersetzungsservice** des baden-württembergischen Ministeriums der Justiz und für Europa (JUM) nutzt neuronale Netze⁶, um fremdsprachliche juristische Dokumente schneller zu erfassen und auf die wesentlichen Punkte zu stoßen. Zwar gibt es dabei – wie bei allen untersuchten Systemen – immer eine menschliche Gegenprüfung, die Effizienzsteigerung ist aber bereits jetzt evident.

KI wird häufig auch an der direkten **Schnittstelle zwischen Verwaltung und Bürger:innen** eingesetzt. Ziel ist die einfachere und aufwandsärmere Gestaltung bestehender Prozesse (bspw. die Beschleunigung von Antragsverfahren) oder die Gestaltung neuer Kanäle (bspw. zum Informationsaustausch).

Beispiele hierfür sind etwa der ressortübergreifende **Chatbot C-19 zum Thema Corona**, den das Informationstechnikzentrum des Bundes in Rekordzeit koordiniert und umgesetzt hat, sowie der vom Bundesministerium für Verkehr und digitale Infrastruktur (BMVI) geförderte **Indoor Robot**, der in öffentlichen Bestandsgebäuden Bilddaten und Klassifikationen nutzt, um z. B. Rollstuhlfahrer:innen die Orientierung zu erleichtern und deren Feedback für erforderliche Umbaumaßnahmen aufzunehmen. Ebenfalls direkt auf die Zufriedenheit der Bürger:innen bezogen ist das Hamburger Projekt **»Mit KI kinderleicht zum Kindergeld«**, das die aufwändigen Behördengänge für Eltern unmittelbar nach der Geburt eines Kindes reduziert und zugleich die Antragsverfahren beschleunigt. Mit der angestrebten Integration einer Sprachassistentz wird nun auch konsequente Barrierefreiheit angestrebt.

Kritisch beäugt wird KI immer dann, wenn es um Entscheidungsprozesse geht. Dabei liegt ihr Mehrwert zumeist in der **Entscheidungsvorbereitung**, gerade in Bezug auf Sichtung und Mustererkennung bei großen Datenmengen. Dabei können auch neue Daten gesammelt und analysiert werden, die zuvor nicht Teil des Entscheidungsprozesses waren. Ziel ist es, eine effektivere, auf einem größeren Datenbestand basierende oder auch konsistentere und fairere Entscheidung zu treffen. Unter diese Kategorie fällt der KI-Einsatz beim **Kampf gegen Kinderpornografie und sexuellen Missbrauch** in Nordrhein-Westfalen. Bei der Beschlagnahmung entsprechenden Materials muss unmittelbar entschieden werden, ob alle Tatbestandsmerkmale erfüllt sind, und ggf. noch bestehende Missbrauchsverhältnisse müssen aufgedeckt werden. Ziel ist es also, auf sicherer Grundlage eine rasche Erstbeurteilung der Lage zu fällen. Die finale Entscheidung über das polizeiliche Vorgehen treffen Staatsanwält:innen und nicht die Maschinen. In einem anderen Beispiel nutzt die Berufsgenossenschaft Energie Textil Elektro Medienerzeugnisse (BG ETEM) KI zur Steuerung ihrer **Fallprüfung im Regressmeldeverfahren**, sodass sie Aufwand und Ertrag eines angestregten Verfahrens auf Basis der Daten vorab kalkulieren und somit informierter beurteilen kann. Das Bundesamt für Migration und Flüchtlinge (BAMF) nutzt neuro-linguistische Verfahren der Texterkennung, um mit der **Filterung von Gesprächsprotokollen** relevante Hinweise im Hinblick auf terroristische Vereinigungen u. Ä. zu erhalten.

⁶ Eine Form maschinellen Lernens. Ein neuronales Netz besteht aus mehreren Schichten von miteinander verbundenen künstlichen Neuronen, die die Funktionsweise ihrer biologischen Gegenstücke simulieren sollen. Über Eingangsneuronen werden die zu analysierenden Daten in das neuronale Netz übertragen, das Ergebnis lässt sich an den Ausgangsneuronen ablesen. Zwischen Eingabe- und Ausgabeschicht können unterschiedlich viele innere Schichten aus verknüpften Neuronen zum Einsatz kommen. Siehe dazu ausführlicher: <https://www.oeffentliche-it.de/-/neuronale-netze> (letzter Aufruf: 30.10.2020)

Vorstellung der Fallbeispiele

Auch bei Prozessen der **Planung** spielt KI eine wichtige Rolle. Hierbei werden Daten, die Grundlage für Entscheidungen in der öffentlichen Verwaltung sind, unterstützend durch KI aufbereitet. Im Gegensatz zur Entscheidungsvorbereitung geht es dabei nicht um fallspezifische Einzelentscheidungen, sondern um strategische planerische Entscheidungen. Ziele dieser Anwendungen können die effektivere Verteilung von bestehenden Ressourcen oder die Prognose zukünftiger Entwicklungen sein. Dazu gehört das KI-System **PREVIEW zur Verbesserung der internationalen Krisenfrüherkennung** im Auswärtigen Amt. Zunächst wurden Datenquellen diversifiziert, um dann auf Basis eines maschinellen Regressionsverfahrens validere Aussagen über künftige Konfliktherde, Kriege oder humanitäre Katastrophen zu treffen.

Bei Prozessen der **Überwachung** werden zu einem bestimmten Zustand oder einer Situation laufend Daten gesammelt und analysiert. Ziel ist zumeist die Erkennung von Anomalien oder Besonderheiten in einer größeren Datenmenge. Hierzu zählt das **Projekt KI4LSA**, das bereits im Testbetrieb in der Stadt Lemgo läuft und den dortigen Verkehrsfluss optimieren soll. Bild- und Lärmsensoren bilden jeweils die Verkehrssituation an einer Kreuzung ab. Das aus den Daten geschaffene digitale Abbild wird unter Nutzung von KI analysiert, um Ampelanlagen zu steuern. Durch die intelligente Anpassung der Ampelschaltungen an die akute Verkehrssituation können bestimmte verkehrspolitische Ziele (Vermeidung von Stehzeiten, Minderung der Lärmbelastung, Verringerung der Umweltbelastung durch Abgase) besser erreicht werden.

Wie in der Analyse der Fallbeispiele deutlich geworden ist, beschreiben die Prozesskategorien Idealtypen, die sich nicht gegenseitig ausschließen. Die Fallbeispiele sind häufig in Schnittbereichen dieser Typologie angesiedelt. Wie die KI-Anwendungsbereiche in den Fallbeispielen ineinandergreifen, zeigt beispielhaft die Funktionsweise des Projekts KI4LSA: Hier wird zunächst die Verkehrssituation mittels Sensorik überwacht (Überwachung), um dann optimierte Ampelschaltungen vorzubereiten (Entscheidungsvorbereitung). Die Daten sollen auch langfristig genutzt werden, um optimierte Verkehrsflüsse zu planen (Planung).

KI4LSA

Künstliche Intelligenz für Lichtsignalanlagen

1 Verkehrskamera

Die Verkehrskamera nimmt die Straße vor der Ampel auf und sendet die Daten an den Edge-Computer, der die empfangenen Daten lokal prozessiert.

2 Radarsensor

Der Radarsensor sendet Position und Geschwindigkeit von Fahrzeugen an den Edge-Computer.

3 Edge-Computer

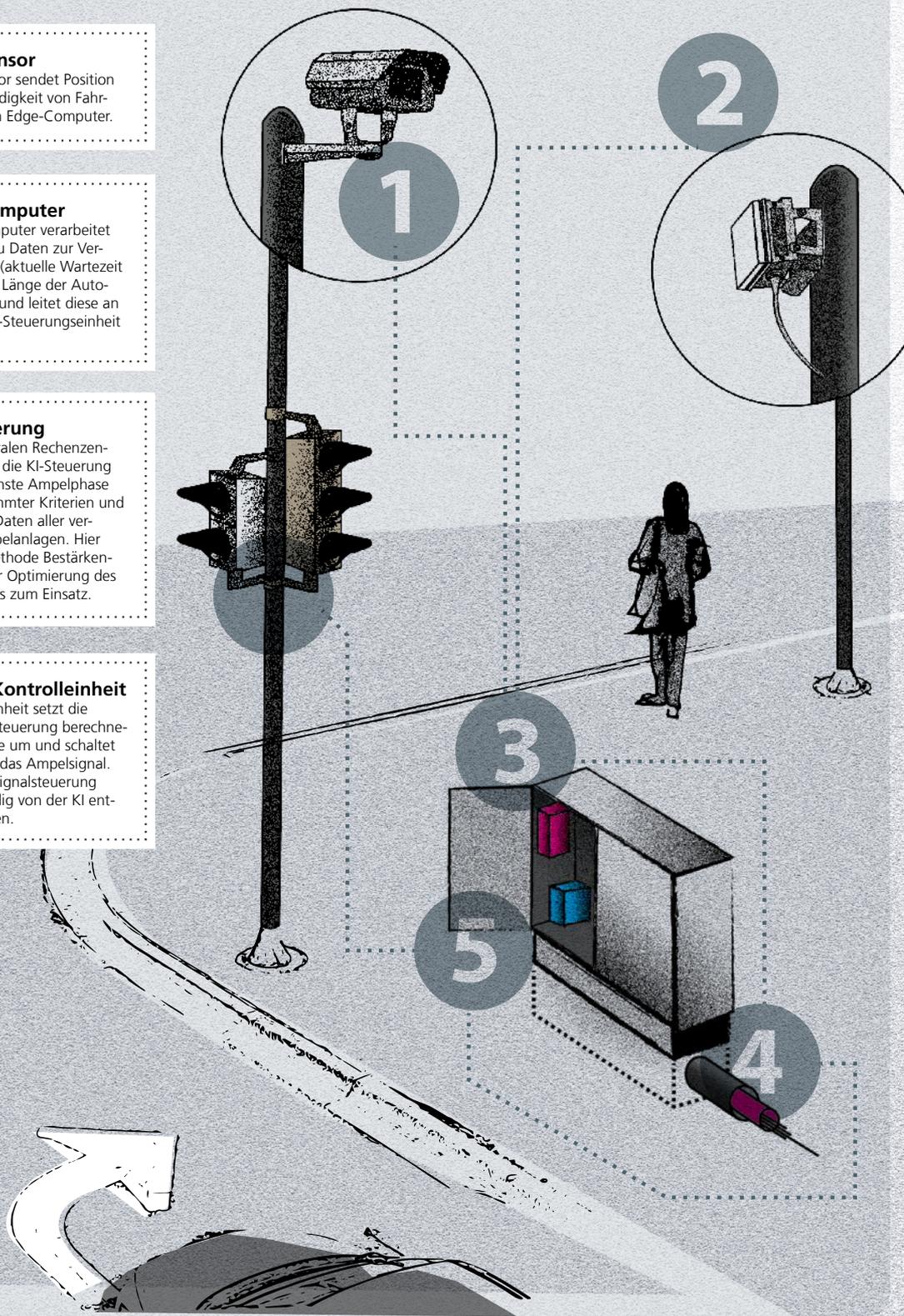
Der Edge-Computer verarbeitet Sensordaten zu Daten zur Verkehrssituation (aktuelle Wartezeit an der Ampel, Länge der Autoschlange etc.) und leitet diese an die zentrale KI-Steuerungseinheit weiter.

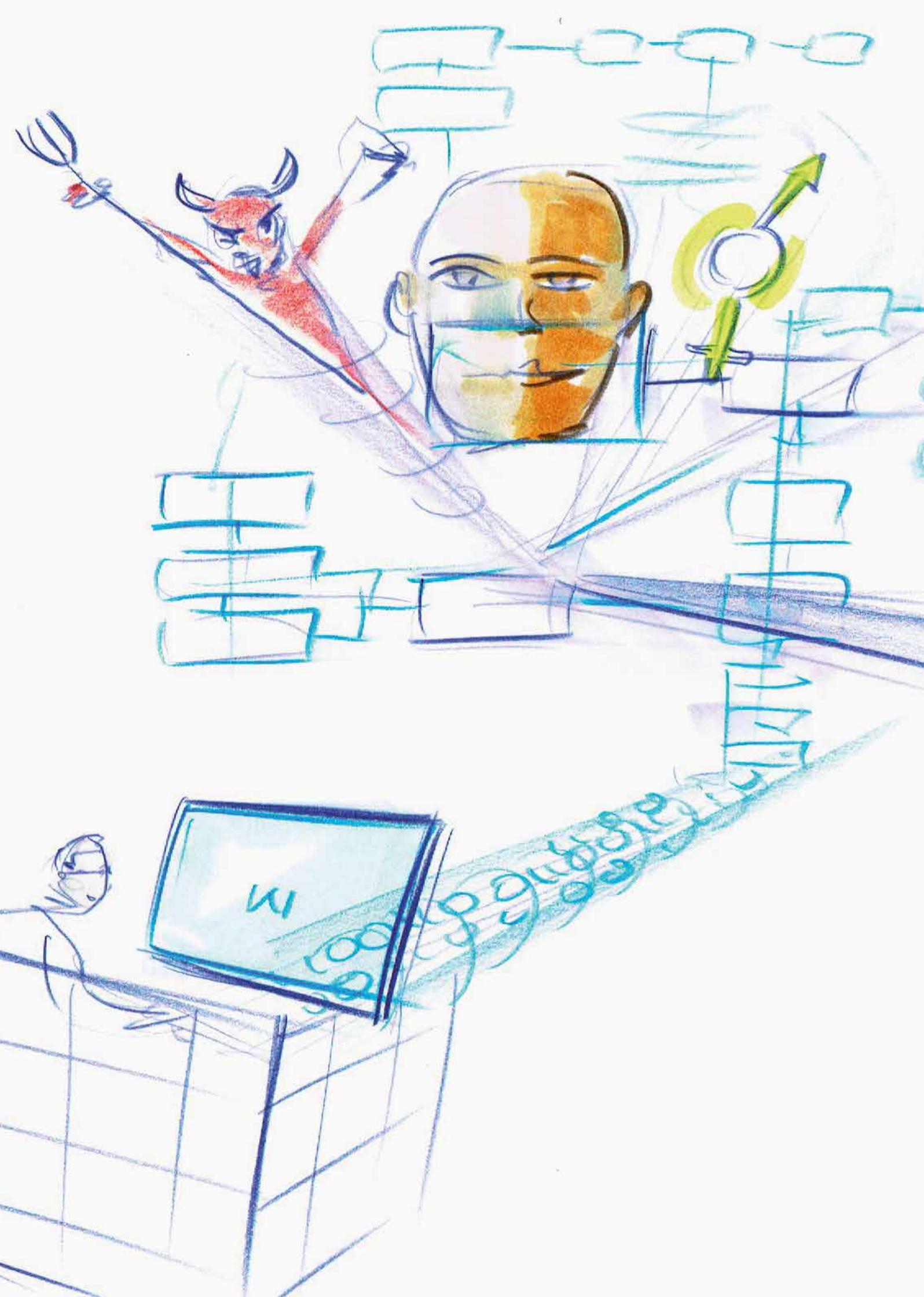
4 KI-Steuerung

In einem zentralen Rechenzentrum ermittelt die KI-Steuerung die ideale nächste Ampelphase entlang bestimmter Kriterien und auf Basis der Daten aller verknüpften Ampelanlagen. Hier kommt als Methode Bestärkendes Lernen zur Optimierung des Verkehrsflusses zum Einsatz.

5 Lokale Kontrolleinheit

Die Kontrolleinheit setzt die durch die KI-Steuerung berechnete Ampelphase um und schaltet entsprechend das Ampelsignal. Wichtig: Die Signalsteuerung kann vollständig von der KI entkoppelt werden.





4. BEWERTUNG DER FALLBEISPIELE

Die Fallbeispiele wurden anhand der in Kapitel 2 dargestellten Kriterienmatrix analysiert. Dabei wurden für jede Kategorie diejenigen Maßnahmen identifiziert, die als ein Beispiel guter Praxis gewertet werden können. Das Vorgehen bestand aus folgenden Schritten:

Erstens wurden die Fallbeispiele anhand der Kriterienmatrix analysiert und jeweils zutreffende ideale Antworten (Unterpunkte der Indikatoren) notiert.

Zweitens wurden diejenigen Antworten aussortiert, die von vielen oder allen Beispielen erfüllt werden. Dazu gehören insbesondere die Definition von Zielen (1.1), ein grundsätzlich iteratives und agiles Vorgehen (4.3), grundlegende Sicherheitsmaßnahmen wie zumindest einmalige Testphasen und das Vorhandensein einer technischen Notabschaltung (teilw. 5.5) sowie versionierte Repositorien von verwendeten Datensätzen in Training und Betrieb (6.4).

Drittens wurden zur besseren Übersicht und Verdeutlichung der Ergebnisse jene Indikatoren hervorgehoben, die im Vergleich zu den anderen Fallbeispielen eine besonders gute Praxis darstellen. Die identifizierten Indikatoren und Unterpunkte sind in der folgenden Tabelle dargestellt. Sie beziehen sich auf die Kriterienmatrix aus Kapitel 2. Im Anschluss an die Tabelle werden ausgewählte Beispiele guter Praxis beschrieben.

Zielorientierung

Um eine möglichst präzise und umfassende Zielorientierung vorzunehmen, führen alle Projekte grundsätzliche Überlegungen zu ihren Bedarfen und Zielen durch. Dabei kann es hilfreich sein, das KI-System so aufzubauen, dass es sich bei veränderten Zielen anpassen kann, wie bspw. bei **KI4LSA** (6). Zusätzlich ist es wichtig, Zielkonflikte zu identifizieren, wie es bei der automatisierten Unterstützung von **Handwerksstatistiken** (4) der Fall ist. Hierbei wurden besondere Herausforderungen wie Mathwashing⁷ und Feedback-Loops als Quellen für Verzerrungen diskutiert.

Interne Veränderung

Projekte werden zumeist eingesetzt, um Aufgaben besser zu lösen, als dies bisher geschieht. Nur in wenigen Fällen werden bisherige Arbeitsprozesse neu gedacht, um Aufgaben insgesamt besser zu lösen. So nahm das Projekt **Mit KI kinderleicht zum Kindergeld** (7) konsequent den Bedarf nach einer Integration vormals separater Geschäftsprozesse auf. Dies erhöhte signifikant die Zufriedenheit und die Effizienz für Bürger:innen sowie die Verwaltung. Maschinelle **Übersetzungen in der Justiz** (10) ermöglichen einen schnelleren Überblick über relevante Fachtexte, der vorher so nicht möglich war. Einige Projekte haben neue Kooperationen geschaffen, bspw. über Behörden- und Ministeriengrenzen hinweg (**C-19**, 2) oder mit Akteur:innen aus der Zivilgesellschaft wie dem Sehbehinderten-Verband beim **Indoor Robot** (5). Die Idee, Bürger:innen direkt in die Datensammlung des Projektes einzubinden, ist noch in keinem Projekt realisiert, aber bei **ARC-D** (1) als Zielvision angedacht.

Innovationsmarker

Da bereits die Recherche der Fallbeispiele auf innovative Projekte ausgerichtet war, ist dies auch die Kategorie, in der viele der untersuchten Fälle gute Praktiken etabliert haben. Dazu gehört die Verwendung von Open Source-Komponenten oder die Bereitstellung der eigenen Anwendung – zumindest teilweise – als Open Source (**ARC-D**, 1; **Handwerksstatistiken**, 4). Ausgesprochen viele Projekte bedenken eine mögliche Skalierung des eigenen Ansatzes und haben das KI-System so entwickelt, dass es auch als Blaupause für ähnliche Anwendungen in anderen Themen- oder Geschäftsbereichen dienen kann, wie der Covid-19-Chatbot **C-19** (2), die jährliche Abfrage von gültigen **Studienbescheinigungen** für die Auszahlung von Kindergeld (3), die Klassifikation von **Handwerksstatistiken** (4), die maschinellen **Übersetzungen in der Justiz** (10) sowie die Sprachassistenten bei **Mit KI kinderleicht zum Kindergeld** (7). Abschließend ist die Praxis des Austauschs mit anderen Behörden und der Öffentlichkeit hervorzuheben, bspw. im Rahmen des behördlichen Netzwerks NEXt beim KI-System zur Erkennung von **Studienbescheinigungen** (3) oder zur **Profilanalyse von Gesprächsprotokollen** (12).

⁷ Der Begriff bezeichnet ein unhinterfragtes Vertrauen in die Ergebnisse mathematischer Berechnungen und damit die Ausblendung möglicher inhärenter und menschengemachter Verzerrungen, Interessen oder Machtkonstellationen.

	Zielorientierung 	Veränderungsumfang 	Innovationsmarker 	Menschliche Aufsicht 	Sicherheit 	Technisches Monitoring 	Wirkungsmonitoring 	Nachvollziehbarkeit 	Akzeptanz 
1 ARC-D		2.1, 2.3	3.2					8.4	
2 C-19 Chatbot		2.4	3.5						
3 Erkennung von Studienbescheinigungen		2.4 2.5	3.5				7.5		
4 Handwerkstatistiken	1.4, 1.5		3.2 3.5						
5 Indoor Robot		2.4							9.1
6 KI4LSA	1.2				5.2 5.3			8.2 8.5	
7 Mit KI kinderleicht zum Kindergeld	1.2, 1.3	2.1 bis 2.5	3.1 3.5				7.4 bis 7.6		
8 KI gegen Kinderpornografie	1.5			4.3		6.2 6.6			
9 KI-Fallprüfung im Regressmeldeverfahren					5.1 5.2 5.5	6.2		8.4	
10 Übersetzungsservice der Justiz		2.1 2.3	3.3						9.4
11 PREVIEW						6.2		8.2 8.5	9.1
12 Profilanalyse von Gesprächsprotokollen		2.5		4.4					9.1 9.8

Projektmanagement

Ein gutes Projektmanagement mit der Benennung einer verantwortlichen Person ist in allen Projekten vorhanden, ebenso eine zumindest in Grundzügen agile Vorgehensweise. Explorations- und Testphasen sind nicht bei allen Projekten verbreitet. Bei **KI gegen Kinderpornografie** (8) wurde dies jedoch unter erschwerten Bedingungen (bereits der Besitz der notwendigen Trainingsdaten wäre außerhalb des geschützten Verwaltungsumfelds strafbar) durchgeführt, um die Fehlerquoten zu ermitteln und zu verringern. Besonders erwähnenswert ist das Projekt zur **Profilanalyse von Gesprächsprotokollen** (12), bei dem fachverantwortliche Personen der eigenen Behörde die Qualität des KI-Systems einschätzen und individuelle Anpassungen vornehmen können, ohne externe IT-Kräfte beauftragen zu müssen.

Sicherheit

Auch im Bereich Sicherheit sind viele gute Praktiken verbreitet – die in die Kriterienmatrix aufgenommenen Indikatoren aber noch nicht. Eine besondere Vorsichtsmaßnahme trifft das Projekt **KI4LSA** (6), indem ein Teil der Rechenleistung dezentral erfolgt und – im Falle eines Angriffs – die lokalen Steuerungseinheiten der einzelnen Ampelanlagen jederzeit wieder von der zentralen KI-Steuerung entkoppelt werden können. Ebenso ist die Maßnahme, wöchentlich ein Monitoring des Systems zur **Fallprüfung von Regressmeldeverfahren** (9) durchzuführen und Daten parallel abzuspeichern, für die Sicherheit hilfreich.

Datenhaltung und -qualität

Versionierte Repositorien für Trainings- und Betriebsdaten sind in den untersuchten Projekten üblich. Besonders hervorzuheben ist hier erneut das wöchentliche Monitoring bei der **KI-Fallprüfung von Regressmeldeverfahren** (9). Dabei wird überprüft, ob es z. B. Auffälligkeiten bei den Datenquellen gibt und ob sich der Input und Output verändert haben. Beim Projekt **KI gegen Kinderpornografie** (8) werden neue Daten einer gründlichen Prüfung unterzogen, die sich insbesondere nach den hohen rechtlichen Anforderungen für die Arbeit mit solchen Daten richtet.

Wirkungsmonitoring

Das System zur Erkennung von **Studienbescheinigungen** beim Kindergeld (3) wurde von einem externen Forschungsinstitut überprüft. Dies erlaubte eine unabhängige Einschätzung und brachte externe Expertise in das Projekt ein. Auch im Pro-

jekt **Mit KI kinderleicht zum Kindergeld** (7) wurde die Ergebnisanalyse in jeder Projektphase mitgedacht, indem der schrittweise Umbau der Leistungen konsequent an den Bedürfnissen und Praxen der Kund:innen ausgerichtet wurde.

Nachvollziehbarkeit

Nachvollziehbarkeit umfasst Maßnahmen, die die Funktionsweise des Systems bspw. mittels Simulationen (**KI4LSA**, 6) oder durch Erklärpapiere und Schulungen (**PREVIEW**, 11) erklären und näherbringen. Besonders der Einsatz von Erklärungen über reine Erklärtexthe hinaus ist lobenswert. Wenige Projekte, wie **ARC-D** (1) und die **Fallprüfung von Regressmeldeverfahren** (9), schaffen auch die Nachvollziehbarkeit von einzelnen Entscheidungen des KI-Systems, indem die jeweils für die Entscheidung relevanten Faktoren hervorgehoben und eine individuelle Wahrscheinlichkeit ausgegeben werden.

Akzeptanz

In Bezug auf die Akzeptanz ist es gute und in einigen Projekten etablierte Praxis, Schulungen für Mitarbeitende (**PREVIEW**, 11; **Profilanalyse von Gesprächsprotokollen**, 12) oder Klient:innen (**Indoor Robot**, 5) durchzuführen und sie mit der Arbeit mit dem KI-System vertraut zu machen. Erwähnenswert ist zudem die Begleitung des maschinellen **Übersetzungsservices der Justiz** (10). In diesem Fall werden seit der Pilotphase zweimal jährlich Anwender:innenbefragungen mit sieben bis acht Zufriedenheitsskalen durchgeführt. Diese Ergebnisse werden dann an das Entwicklerteam weitergeleitet.

Diese hervorgehobenen Beispiele zeigen einige gute Praktiken, die als Vorbild für andere Projekte dienen können. Die Darstellung zeigt jedoch auch, dass kein Projekt alle Indikatoren der Kriterienmatrix vollständig erfüllt und auch zwischen den Beispielen deutliche Unterschiede festzustellen sind. Ohne die Beispiele zu bewerten, ergeben sich erhebliche Potenziale und Defizite, die im folgenden Kapitel näher betrachtet werden.



5. ERKENNTNISSE ZUM KI-EINSATZ IN DER DEUTSCHEN VERWALTUNG

Anhand der Interviews und Steckbriefe wurde überprüft, welche Bausteine guter Praxis in den jeweiligen Fallbeispielen realisiert wurden. Die Schlussfolgerungen in diesem Kapitel gehen von diesen Einzelbewertungen aus und stellen die übergreifenden Erkenntnisse zum KI-Einsatz in der deutschen öffentlichen Verwaltung in den Mittelpunkt.

Insgesamt **fehlen gute Praktiken** in bestimmten Bereichen, die **das Vertrauen der Beschäftigten und ggf. Klient:innen** in die jeweiligen KI-Anwendungen erhöhen könnten. So gibt es in den Kategorien »Nachvollziehbarkeit« und »Akzeptanz« zwar erste Ansätze, bspw. Explainable AI und Schulungen, erstere sind jedoch nicht immer für alle Beteiligten verständlich und letztere bilden noch die Ausnahme. Teilweise wurden in den Interviews Maßnahmen für Nachvollziehbarkeit als unnötig abgetan. Dies zeigt, dass dieses Best-Practice-Kriterium in der öffentlichen Verwaltung bislang noch wenig berücksichtigt wird. Wie wichtig die Akzeptanz der Systeme ist, veranschaulicht die Aussage eines Interviewpartners: »Schwierig wird es, wenn ein KI-System im Arbeitsalltag implementiert werden soll. Da wird man sicherlich auf Widerstände der Beschäftigten stoßen.« Die Grenzen und die Funktionsweisen der Systeme besser zu erklären, kann solchen Widerständen vorbeugen. Auch grundlegende Informationen zu den Systemen fehlen vielen Beteiligten: »Es sollen keine Entscheidungsprozesse ersetzt werden, viele Leute denken das aber.«

Ein weiterer Bereich, der fast keine konkreten Praktiken aufzeigt, ist das **Wirkungsmonitoring**. So konnte keines der untersuchten Fallbeispiele eine stringente Methodik vorweisen, um die mit dem KI-Einsatz angestrebten Ziele fortlaufend zu messen. Ohne eine systematische fortlaufende Messung der Zielerreichung lässt sich der konkrete Nutzen des KI-Einsatzes jedoch nicht bestimmen. Ein Gesprächspartner kritisierte: »Das ist ein bisschen die Krux von Förderprojekten, da diese auf die Entwicklung der Technik und deren Pilotierung ausgelegt sind. Evaluierung müsste eigentlich nicht nur für zwei Jahre stattfinden.«

Weiterhin ist eine nur sehr **begrenzte Einbindung** externer Expertise insbesondere **von zivilgesellschaftlichen Organisationen** zu beobachten. Nur ein untersuchtes Projekt arbeitet ausdrücklich mit der Zivilgesellschaft zusammen. Teilweise ließe sich die mangelhafte Einbindung dadurch erklären, dass viele der Fallbeispiele für die Verwaltung selbst entwickelt und dort

eingesetzt werden. Dies steht einer Einbindung von einschlägigen Initiativen oder Vereinen jedoch nicht prinzipiell entgegen.

Besonders hervorzuheben ist hingegen das interne **Querschnittsdenken** einiger Projekte, die eine **Anschlussfähigkeit und Skalierbarkeit** des eigenen Projekts für andere Verfahren mitplanen. Dies kann auch durch interne Impulse entstehen: Das referatsgetriebene Projekt **Handwerksstatistiken** lieferte dem Statistischen Bundesamt eine Blaupause für neue Verfahren bei der Klassifikation, das Projekt **PREVIEW** begründete im Auswärtigen Amt den Einstieg in eine stärker datengetriebene Prognostik und der **C-19-Chatbot** ist Teil einer Blaupause für eine Chatbot-Infrastruktur in vielen Behörden. Das Projekt **Mit KI kinderleicht zum Kindergeld** formuliert sogar dezidiert den Anspruch, eine »nachbaubare« digitale Vorlage für einen integrierten Geschäftsprozess für sämtliche Verwaltungsebenen zu liefern.

Auch in puncto Sicherheit denken einige Projekte weiter und überlegen im Voraus, wie sie mit Angriffsfällen umgehen und wie eine Notabschaltung des Systems technisch und organisatorisch durchzuführen wäre. An dieser Stelle wird nicht nur die technische Sicherheit des KI-Systems betrachtet, sondern auch die **erweiterte Sicherstellung des (Entscheidungs-)Prozesses**, in den das KI-System eingebettet ist.

Abschließend sollen noch zwei Aspekte festgehalten werden, die sich nicht aus der Bewertung der Fallbeispiele anhand der Kriterienmatrix ergeben, aber in vielen Interviews thematisiert wurden. Erstens zeigten sich insbesondere Vertreter:innen von Forschungsprojekten mitunter erstaunt, wie stark die **Beharrungskräfte und Pfadabhängigkeiten** in der Verwaltung sind: Es bestehe eine »eingebaute Latenz« in den Behörden in Bezug auf die Umsetzung von KI-Projekten. »Die Begeisterung für Veränderungschancen fehlt an vielen Stellen. Die Leute sehen diese Themen als eine Gefahr von der Seite. Es ist im ersten Moment wie eine Provokation für das Bestehende.«

Zweitens zeigte sich im Prozess unserer Interviewanfragen, dass bei vielen KI-Projekten im öffentlichen Sektor das **Wissen** über deren Funktions- und Arbeitsweisen stark **auf einzelne Mitarbeitende konzentriert** ist. Abwesenheiten, Umbesetzungen oder auch die Pensionierungswelle im öffentlichen Dienst führen zu Wissenslücken oder dem Abfluss von fachlicher Expertise.

6. FÜNF HANDLUNGSEMPFEHLUNGEN

Aus der Analyse ergeben sich folgende Handlungsempfehlungen:

Wirkung messen!

Auf Projektebene braucht es eine mitlaufende Methodik, welche die mit dem KI-Einsatz avisierten Effekte analysiert und misst. Dabei sollten insbesondere Auswirkungen über den engen Projektkontext hinaus einbezogen werden. Dies scheint bisher nicht der Fall zu sein. Auf politischer Ebene bedeutet das, dass solche Wirkungsanalysen auch bei jeder Förderung mitgedacht und vor allem auch mitbudgetiert werden müssen. Nur so können die richtigen Anreize gesetzt werden.

Nachvollziehbarkeit und Akzeptanz einfordern!

Während manche Projekte innovative Ansätze wie Simulationen und laufende Zufriedenheitsbefragungen durchführen, wird bei anderen Projekten Nachvollziehbarkeit – auch als Mittel zur Stärkung der Akzeptanz – vernachlässigt. In den Projektinterviews wurde mehrfach berichtet, Nachvollziehbarkeit sei im konkreten Fall nicht relevant. Entscheidungsträger:innen in Politik und Verwaltung müssen hier klare Rahmenbedingungen setzen, sodass Maßnahmen zur Nachvollziehbarkeit sowie Beschwerdemanagement im Zweifel (rechtlich) eingefordert werden. Auch eine Anlehnung an die Empfehlungen der Dataethikkommission ist denkbar, die ein risikoadaptiertes Regulierungsregime für die algorithmische Verarbeitung von Daten skizziert hat.

KI-Entwicklung ergebnisoffen angehen!

Auf Projektebene bedeutet das einerseits ein freieres Planungsdenken: »Wünschenswert wäre ein offenes Mindset, was die Kommunen angeht. (...) An den Stellen, wo man Dinge positiv wie negativ beeinflussen könnte, sollte mehr Experimentierfreudigkeit und eine positive Einstellung gegenüber Innovationen herrschen, damit wir Wissenschaftler nicht schon im Vornherein beweisen müssen, dass alles funktionieren wird.« Auf der politischen Ebene muss andererseits der Einsatz von KI insgesamt ergebnisoffener angegangen werden. KI ist kein Selbstzweck, sondern ein bewusst einzusetzendes Werkzeug – nicht für jedes Problem stellt KI eine Lösung dar. Ein Interviewpartner sprach offen von seinem Gefühl, »dass KI in manchen Bereichen fehlt, in anderen wiederum angewandt wird, wo es kei-

nen Sinn ergibt«. Es gilt daher, diejenigen Problemlagen zu identifizieren, bei denen KI tatsächlich eine Lösung bieten kann, und für diese dann Förderlinien oder Finanzierung bereitzustellen.

Datenanalytische Expertise aufteilen!

Innerhalb der Verwaltung braucht es mehr Mitarbeitende und Teams, die über datenanalytische Expertise verfügen. Gerade im Projektmanagement haben einige Projekte behördenintern Fachexpertise und spezifisches Domänenwissen aufgebaut, das sich aber auf wenige Beteiligte oder nur eine Person konzentriert. Bei KI-Themen in der Verwaltung sollte man »sich nicht zu stark von externen Zulieferungen abhängig machen und eine gewisse Autonomie beibehalten«, wie es in einem Interview gefordert wurde. Leiter:innen von Behörden sollten daher Projektverantwortung auf mehrere Schultern verteilen und langfristig eigene KI-Kapazitäten aufbauen. Vorhandene Expertise kann dann in Projekte eingebracht werden, anstatt dass – wie bisher – die Expertise erst im konkreten Projektkontext aufgebaut wird.

Projekte nach außen öffnen!

Die Praxis einiger Projekte, das eigene System teilweise oder vollständig als Open-Source-Software nachnutzbar zu machen, sollte fortgeführt und verstärkt werden. Gleiches gilt für den Austausch mit anderen Verwaltungseinheiten, die vor ähnlichen Herausforderungen stehen, und die Entwicklung von skalierbaren, leicht auf neue Projektkontexte anpassbaren KI-Lösungen. Strukturell kann die Öffnung auch über die Verwaltung hinaus erfolgen: Expert:innen und andere Stakeholder wie die Zivilgesellschaft in die Coding- bzw. Reviewprozesse einzubinden, zeigte beim Projekt Indoor Robot, wie neue Potenziale für das Gemeinwohl freigesetzt werden können. Auf Entscheidungsebene bedarf es hier eines grundsätzlichen Einstellungswandels: »Entscheidend ist, dass man Teile seiner behördlichen Mentalität aufgibt.« Ein entsprechender Austausch kann strukturell durch den Aufbau und die Stärkung hilfreicher Foren wie der NExt- oder N3GZ-Netzwerke erfolgen, die verwaltungsinterne Herausforderungen und Erfahrungen aggregieren und anschlussfähig für andere Fachbereiche machen.

ANNEX: STECKBRIEFE

1 – ARC-D

Name des Systems	Automatisches Zustandsmonitoring von Autobahnen mit KI – ARC-D (Automatic Road Condition Detection)
Einsetzende Behörde	Forschungsprojekt im Rahmen des mFUND des BMVI, Autobahn GmbH des Bundes
Anwendungsbereiche	 Entlastung  Planung  Überwachung
Fachbereich	Verkehr
Themenbereich	Straßenbau / Zustandsüberwachung
Beschreibung	<p>Problem: Aktuell gibt es keine automatisierte Erfassung von Zustandsdaten aller Bundesautobahnen. Eine aktuelle Datenbasis und eine automatisierte Einschätzung der Qualität einzelner Streckenabschnitte wären aber wichtig, um bspw. frühzeitig geeignete Instandsetzungsmaßnahmen einzuleiten und so kostenintensive Langzeitschäden zu verhindern.</p> <p>Lösungsansatz: substituierend und ergänzend. Anhand von zuvor durch Expert:innen klassifizierte Bilddaten wird ein System entwickelt, das es ermöglicht, Videodaten (teil-)automatisiert auszuwerten, um Straßen bezüglich ihres Zustandes zu klassifizieren. Langfristig sollen auch andere Sensordaten einfließen.</p> <p>Outputs: Klassifizierte Straßenabschnitte sowie »Grenzfälle« zur weiteren Überprüfung durch Menschen. Die genaue Darstellungsweise steht noch nicht fest.</p>
Technische Umsetzung	Unüberwachtes Lernen
Verwendete KI- Komponenten	Deep Learning mit Computer Vision/Bilderkennung, Anomaly Detection
Unternehmen für Entwicklung und Betrieb	Stand-alone- und Neuentwicklung. Die Palaimon GmbH führt die Machbarkeitsstudie zusammen mit den assoziierten Partnern Hessen Mobil Straßen- und Verkehrsmanagement sowie Die Autobahn GmbH des Bundes durch.
Dauer des Einsatzes	-
Bisheriger Projektverlauf	Machbarkeitsstudie läuft seit Dezember 2019, Laufzeit bis Ende 2020
Was macht dieses Beispiel zu einer Best Practice?	<p>Interne Veränderung:</p> <ul style="list-style-type: none"> • Den Zustand von Autobahnen mit Videodaten mittels KI (teil-)automatisiert auszuwerten, ist neu. Perspektivisch ist auch eine Einbindung von Bürger:innen denkbar, die via Dashcam oder Smartphone Straßen erfassen, die sie befahren. <p>Innovationsmarker:</p> <ul style="list-style-type: none"> • Zumindest eine Basisversion soll als Open Source veröffentlicht werden. <p>Nachvollziehbarkeit:</p> <ul style="list-style-type: none"> • Die Klassifikation der KI wird durch Maßnahmen der Nachvollziehbarkeit unterstützt. So werden bei Bildschnipseln die tatsächliche und die von der KI ermittelte Klassifizierung gegenübergestellt. Zudem wird eine eigene (Un-)Sicherheitsschätzung ausgewiesen. Diese Einschätzung sehen auch die Anwender:innen.
Ansprechperson	Fabian Kraut fabian.kraut@palaimon.io

2 – C-19

Name des Systems	C-19 (Ressortübergreifender Chatbot zum Thema Corona)
Einsetzende Behörde	ITZBund, Bundesministerien und Behörden
Anwendungsbereiche	 Entlastung  Bürger*innen-Beziehung
Fachbereich	Kommunikation
Themenbereich	Regelungen, Vorschriften und Hilfen im Kontext der Covid-19-Pandemie
Beschreibung	<p>Problem: Bürger:innen wollen schnell und einfach verständliche Informationen erhalten, auch zu Covid-19-Themen. Dafür stehen ihnen diverse Kanäle zur Verfügung (u. a. Post, E-Mail, Telefon, persönliches Vorsprechen), die aber teilweise an Öffnungszeiten gebunden sind und in den Behörden Personalressourcen binden. Seit Beginn der Pandemie ist in den Servicebereichen des öffentlichen Dienstes ein bis zu zehnmals höheres Anfrageaufkommen festzustellen.</p> <p>Lösungsansatz: ergänzend. Ein Chatbot beantwortet automatisiert Fragen von Bürger:innen. Dabei werden bestehende FAQs als Bausteine für einen Dialog eingepflegt. Zukünftig sind für Chatbots Funktionen zu Antragserstellung und Statusklärung angedacht. Die KI-Komponenten werden vorrangig zur Spracherkennung genutzt, aber auch zur Aufarbeitung der geführten Dialoge.</p> <p>Output: Vorformulierte Antworten und Verweise auf weitere Informationsquellen.</p>
Technische Umsetzung	Überwachtes Lernen
Verwendete KI- Komponenten	Spracherkennung via NLP über Open-Source-Framework
Unternehmen für Entwicklung und Betrieb	Stand-alone-Entwicklung mit Open-Source-Framework
Dauer des Einsatzes	Seit Februar 2020 vorgeplant, im Mai umgesetzt
Bisheriger Projektverlauf	Die Beauftragung des Chatbots erfolgte am 30.04.2020 durch das BMI, ca. 20 andere Behörden waren involviert, die Entwicklung gelang innerhalb von ca. 40 Arbeitstagen.
Was macht dieses Beispiel zu einer Best Practice?	<p>Innovationsmarker</p> <ul style="list-style-type: none"> • Der Covid-19-Chatbot ist Teil der IT-Konsolidierung des Bundes. In diesem Rahmen dient er auch als Blaupause für weitere Chatbots, die zukünftig in verschiedenen Behörden zu verschiedenen Zwecken eingesetzt werden können. <p>Interne Veränderung</p> <ul style="list-style-type: none"> • Die Entwicklung des Chatbots erforderte und ermöglichte eine behördenübergreifende Zusammenarbeit, die in sehr kurzer Frist realisiert wurde.
Ansprechperson	René Gürth Rene.Guerth@itzbund.de

3 – Erkennung von Studienbescheinigungen für Kindergeld

Name des Systems	Klassifikation und Informationsextraktion von Studienbescheinigungen für die Familienkasse
Einsetzende Behörde	Bundesagentur für Arbeit
Anwendungsbereiche	 Entlastung
Fachbereich	Familienkasse
Themenbereich	Dokumentenerkennung
Beschreibung	<p>Problem: Zur Beantragung des Kindergeldes muss zweimal jährlich die aktuelle Studienbescheinigung des betreffenden Kindes vorgelegt werden. Dadurch entsteht ein hohes Aufkommen von händisch zu prüfenden Dokumenten, zumal sich die deutschen Hochschulen nicht auf ein einheitliches Formular verständigen konnten.</p> <p>Lösungsansatz: ergänzend. Der/die Klient:in lädt die Studienbescheinigung auf einer entsprechenden Webseite hoch. Der Text der Studienbescheinigung wird mit OCR (Optical Character Recognition) extrahiert und dem System zur Klassifikation übergeben. Vier Sachverhalte werden überprüft:</p> <ul style="list-style-type: none"> • Handelt es sich um eine Studienbescheinigung (ja/nein) und mit welcher Wahrscheinlichkeit (Score)? • Handelt es sich um das richtige Kind? • Handelt es sich um den richtigen Gültigkeitszeitraum/das korrekte Semester? • Handelt es sich um eine bekannte deutsche Hochschule? <p>Outputs: Das Ergebnis der Klassifikation wird den Mitarbeitenden im Fachverfahren der Familienkasse angezeigt. Zugleich wird das Ergebnis des Prüfverfahrens »Studienbescheinigung ja/nein« angegeben. Dieses muss dann aktiv von dem/der Mitarbeiter:in angenommen werden.</p>
Technische Umsetzung	Überwachtes Lernen
Verwendete KI- Komponenten	Die Trainingsdaten bestehen derzeit aus 60.000 Studienbescheinigungen, die semi-automatisiert annotiert und dann zum Training verwendet werden. Durch einen eigens entwickelten Algorithmus werden die personenbezogenen Daten vor ihrer Klassifikation anonymisiert und können daher auch als Trainingsdaten eingesetzt werden. Die Klassifikation der Studienbescheinigung wird vorgenommen durch: Naive Base/Support Vector Machine/Logistische Regression. Zur Datenextraktion mittels NLP (Natural Language Processing) wird u. a. Named Entity Recognition verwendet.
Unternehmen für Entwicklung und Betrieb	Stand-alone-Entwicklung (Eigenentwicklung)
Dauer des Einsatzes	Die Entwicklung begann Mitte 2018 und befindet sich seit Anfang 2019 im Produktivbetrieb.
Bisheriger Projektverlauf	2018: Entwicklung eines Prototyps 2019: Produktivsetzung und Re-Training 2020: Optimierung und Re-Training

Was macht dieses Beispiel zu einer Best Practice?	<p>Veränderungsumfang:</p> <ul style="list-style-type: none">• Das System ermöglicht eine große Arbeitserleichterung für die Mitarbeiter:innen, da sie von fehleranfälligen und monotonen Massenaufgaben einfacher Art entlastet werden.• Die Kund:innen-Zufriedenheit steigt, da eine schnellere Rückmeldung bzw. Bearbeitungszeit ermöglicht wird. <p>Innovationsmarker:</p> <ul style="list-style-type: none">• Die Prüfung für die Anwendung des Systems in anderen Fachbereichen läuft bereits. Die Übertragbarkeit auf andere Bereiche ist relativ einfach mit bedingten Anpassungen verschiedener Komponenten zu ermöglichen.• Die Erfahrungen des Projekts fließen bereits in andere Projekte ein.• Das Verfahren zur Anonymisierung der Bescheinigungen wurde an andere Fachbereiche weitergegeben.• Das System wurde auch von externer Stelle qualitätsgeprüft.• Es besteht ständig ein interner Austausch in der Data-Science-Community, an der jede/r aus der BA teilnehmen kann.• Regelmäßige Teilnahme an Konferenzen und Austausch mit Universitäten sowie in der NEXT-Community. <p>Datenhaltung und -qualität: Jährliches Re-Training des Systems Bei Problemen gibt es direkte Rückmeldungen von den Bearbeiter:innen.</p>
Ansprechperson	Dr. Frank Pelzel frank.pelzel@arbeitsagentur.de

4 – Handwerksstatistiken

Name des Systems	Handwerksstatistiken
Einsetzende Behörde	Statistisches Bundesamt
Anwendungsbereiche	 Entlastung  Entscheidungs- vorbereitung
Fachbereich	Statistik
Themenbereich	Klassifikation von Rohdaten
Beschreibung	<p>Problem: Ausgangspunkt ist ein Klassifikationsproblem: Um zu entscheiden, ob ein Datenpunkt (z. B. ein Handwerksunternehmen) in eine Statistik aufgenommen wird, muss dieser als relevant oder nicht relevant klassifiziert werden. Bisher wurde dies händisch erledigt. Dies bringt allerdings einen hohen Zeitaufwand mit sich und ist zudem fehleranfällig. Ferner gibt es dafür nur knappe Personalkapazitäten.</p> <p>Lösungsansatz: ergänzend. Anhand von zuvor bestimmten Klassifikationsmustern für Einheiten (Informationen aus Unternehmensregistern) wird ein Klassifizierungssystem trainiert, das neue Datenpunkte automatisch als relevant oder nicht relevant einstuft. Sachbearbeiter:innen bekommen dasselbe Zeitfenster wie zuvor, um die nun vorsortierten Fälle weiterzubearbeiten.</p> <p>Outputs: klassifizierte Datenpunkte (Handwerksunternehmen relevant/nicht relevant für die Statistik)</p>
Technische Umsetzung	Überwachtes Lernen
Verwendete KI- Komponenten	Support Vector Machine, Random Forest
Unternehmen für Entwicklung und Betrieb	Stand-alone- und Eigenentwicklung. Partielle Zusammenarbeit mit der Universität Bayreuth
Dauer des Einsatzes	Seit Mitte 2016
Bisheriger Projektverlauf	Das System ist seit Mitte 2016 im Einsatz, jedoch noch nicht vollautomatisch einsatzfähig.
Was macht dieses Beispiel zu einer Best Practice?	<p>Zielorientierung:</p> <ul style="list-style-type: none"> • Mögliche Zielkonflikte wurden offengelegt und diskutiert. Insbesondere die Herausforderungen von Mathwashing (s. Fußnote 7) und Feedback-Loops (Skalierung von Fehlern aufgrund ihrer Einspeisung in den Trainingsprozess) als mögliche Folgen des Einsatzes eines solchen halbautomatisierten Systems wurden diskutiert. Um dem vorzubeugen, werden die Ergebnisse durch Menschen aufbereitet, auch bevor sie für das weitere Training des Systems genutzt werden. <p>Innovationsmarker:</p> <ul style="list-style-type: none"> • Das System wurde spezifisch für Handwerksstatistiken entwickelt, kann als Musterprozess aber auf andere Klassifikationsverfahren übertragen werden. • Die KI-Komponenten sind keine von Grund auf neue Software. Vielmehr wurden Komponenten entwickelt, die je nach Bedarf kombiniert und für andere Klassifikationsprozesse angepasst werden können.
Ansprechperson	Joerg Feuerhake Joerg.Feuerhake@destatis.de

5 – Indoor Robot

Name des Systems	Indoor Robot
Einsetzende Behörde	Pilotprojekt gefördert vom Bundesministerium für Verkehr und digitale Infrastruktur (BMVI)
Anwendungsbereiche	 Entlastung  Bürger*innen-Beziehung  Planung
Fachbereich	Barrierefreiheit
Themenbereich	Building Information Modeling (BIM)
Beschreibung	<p>Problem: Personen müssen in öffentlichen Bestandsgebäuden (selbstständig) navigieren können. Insbesondere Menschen mit Beeinträchtigungen wie Blinde und Rollstuhlfahrer:innen brauchen Informationen über die Gegebenheiten des Gebäudes sowie eine einfache und barrierefreie Navigation. In der Praxis zeigt sich, dass aktuelle Gebäudepläne oft nicht die notwendigen Informationen bereitstellen.</p> <p>Lösungsansatz: neu. Ein Roboter vermisst mittels Laserscanner und Objektkamera das gesamte Gebäude. Aus den einzelnen Messpunkten wird unter Nutzung von KI ein Vektormodell des Gebäudes erstellt, das dann für Orientierungskarten und deren Ausspielung auf mobilen Devices genutzt werden kann.</p> <p>Outputs: 3D-Modell des öffentlichen Gebäudes</p>
Technische Umsetzung	Überwachtes Lernen
Verwendete KI- Komponenten	Deep Learning: Training und Verwendung von neuronalen Netzen zur Klassifizierung von Objekten, die für die Navigation wichtig sind, z. B. Bürotüren, Ausgänge und Wände.
Unternehmen für Entwicklung und Betrieb	Eigenentwicklung Fraunhofer, Neuentwicklung für den öffentlichen Sektor
Dauer des Einsatzes	-
Bisheriger Projektverlauf	Der Forschungszeitraum war September 2018 bis August 2020. Das Projekt wurde aufgrund der Covid-19-Pandemie um drei Monate verlängert.
Was macht dieses Beispiel zu einer Best Practice?	<p>Interne Veränderung:</p> <ul style="list-style-type: none"> Im Zuge der Entwicklung wurden neue Kooperationen mit privaten Unternehmen und zivilgesellschaftlichen Organisationen aufgebaut und Stakeholder eingebunden, bspw. der Sehbehinderten-Verband. Dies hat auch Auswirkungen auf die Gestaltung der Anwendung oder eine mögliche Integration in andere Dienste. So werden bei der Auswertung der Daten auch Navigationskarten für Rollstuhlfahrer:innen erstellt, z. B. in Kooperation mit wheelmap.org. <p>Akzeptanz:</p> <ul style="list-style-type: none"> Gebäudebetreiber bzw. Servicekräfte sollen durch eine Schulung befähigt werden, eigenständig mit dem Roboter Gebäude zu vermessen.
Ansprechperson	Konstantin Klipp konstantin.klipp@fokus.fraunhofer.de

6 – KI4LSA

Name des Systems	Künstliche Intelligenz für Lichtsignalanlagen – KI4LSA
Einsetzende Behörde	Stadt Lemgo (Testbetrieb)
Anwendungsbereiche	 Entscheidungs-vorbereitung  Planung  Überwachung
Fachbereich	Mobilität
Themenbereich	Verkehrssteuerung / Smart City
Beschreibung	<p>Problem: Ampeln steuern den Verkehr in Städten maßgeblich. Dabei sind die Ampelanlagen in ihrer Schaltung oft relativ statisch oder werden nur mit Verzögerung von Menschen gesteuert. Weiterhin sind viele Ampeln nicht miteinander synchronisiert. Eine mögliche Optimierung des Verkehrsflusses kann nicht vorgenommen werden.</p> <p>Lösungsansatz: substituierend. Eine noch zu entwickelnde Sensorik bildet jeweils die Verkehrssituation an einer Kreuzung ab. Das aus den Daten geschaffene digitale Abbild wird unter Nutzung von KI analysiert, um andere Ampelanlagen zu steuern. Dafür werden bestehende Ampelanlagen verkettet. Die Ampelanlage soll sich so an den Verkehrsfluss anpassen. Dadurch soll die Transitzeit um 25 Prozent reduziert und eine erhebliche Minderung der Lärmbelastung sowie der Umweltbelastung durch Abgase erreicht werden. Langfristig sollen anormale Verkehrssituationen erkannt und Verkehrsteilnehmende entsprechend informiert werden.</p> <p>Outputs: Steuerung der Ampelanlagen und darauf aufbauend eine Optimierung des Verkehrsflusses. Ferner sollen die gesammelten Daten als Open Data zur Verfügung gestellt werden.</p>
Technische Umsetzung	Verstärkendes Lernen
Verwendete KI- Komponenten	Neuronales Netz
Unternehmen für Entwicklung und Betrieb	Eigenentwicklung Fraunhofer, Betrieb durch Stadtwerke Lemgo GmbH Projektpartner Stührenberg GmbH (Hersteller Lichtsignalanlagen), Cichon Automatisierungstechnik GmbH (Sensorik)
Dauer des Einsatzes	-
Bisheriger Projektverlauf	Der Projektstart erfolgte am 05.03.2020. Die Laufzeit des Projekts beträgt 2,5 Jahre.
Was macht dieses Beispiel zu einer Best Practice?	<p>Zielorientierung:</p> <ul style="list-style-type: none"> • Da die Ziele für das Projekt von Stadt zu Stadt unterschiedlich sein können, kann das System auf verschiedene Ziele hin optimiert werden. Neben allgemeingültigen Zielen (Verringerung von Stop-and-Gos, der Durchfahrtszeit, des Lärms sowie der Abgase) wäre dies z. B. die Steuerung der Geschwindigkeit durchfahrender LKWs. <p>Sicherheit:</p> <ul style="list-style-type: none"> • Das System wurde im Rahmen eines »Cybersecurity by design«-Konzepts u. a. durch zwei zusätzliche Maßnahmen gesichert: Ein Teil der Rechenleistung erfolgt dezentral in Edge-Computern, sodass ein Angriff hier ggf. nur eine Ampel ausschaltet und nicht das gesamte Ampelnetz der Stadt. Außerdem können die lokalen Steuerungseinheiten der einzelnen Ampelanlagen jederzeit wieder von der zentralen KI-Steuerung entkoppelt werden, sodass sie auch bei Ausfall der Zentralsteuerung weiter funktionieren können. <p>Nachvollziehbarkeit:</p> <ul style="list-style-type: none"> • Im Projekt ist die Entwicklung einer Simulation vorgesehen, die die Effekte des Systems für Bürger:innen nachvollziehbar machen soll.
Ansprechperson	Prof. Jürgen Jasperneite juergen.jasperneite@iosb-ina.fraunhofer.de

7 – Mit KI kinderleicht zum Kindergeld

Name des Systems	Mit KI kinderleicht zum Kindergeld (Arbeitstitel)
Einsetzende Behörde	Amt für IT und Digitalisierung in der Senatskanzlei der Freien und Hansestadt Hamburg
Anwendungsbereiche	 Entlastung  Bürger*innen-Beziehung
Fachbereich	Familien und Kind im Rahmen des Onlinezugangsgesetzes (OZG)
Themenbereich	Antragswesen
Beschreibung	<p>Problem: Gesetzlich sind Eltern dazu verpflichtet, bereits in den ersten Tagen nach der Geburt ihres Kindes dem Standesamt Angaben zur Beurkundung zu machen, den Namen festzulegen und damit auch die erste formale Voraussetzung für den Bezug des Eltern- und Kindergeldes zu schaffen. Der bürokratische Aufwand erhöht sich dadurch, dass Dokumente an verschiedenen Stellen in redundanter Weise abgefragt werden und es zumeist (noch) keinen elektronischen Datenaustausch gibt. Direkt nach der Geburt entsteht so – neben der oft angespannten häuslichen Situation – ein zusätzlicher Aufwand bei vielen Eltern.</p> <p>Lösungsansatz: neu. Das Leitmotiv des Vorhabens lautet: »Die Daten laufen, nicht die Eltern«. Zunächst wurde 2018 ein bürgerfreundliches, nur noch dreiseitiges Kombiformular zum einmaligen Eingeben aller notwendigen Daten entwickelt. Dank eines institutionsübergreifenden Geschäftsprozesses können Eltern es direkt in der Geburtsklinik abgeben, die Kooperationspartner kümmern sich um alles Weitere. Die Eltern erhalten nun deutlich schneller auf dem Postweg die Geburtsurkunde ihres Kindes, die Steuer-ID, den Eintrag ins Meldewesen und den Kindergeld-Bescheid. Das Kindergeld wird direkt auf ihr Konto überwiesen. Auf dieser Basis wurde 2019 ein Online-Dienst entwickelt, der seit Juni 2020 als Pilot zunächst einer eingeschränkten Zielgruppe in der Asklepios-Klinik Altona angeboten wird. Der Pilot umfasst auch den Einsatz eines Sprachassistenzsystems, um den kombinierten Antrag einfacher und barriereärmer ausfüllen zu können. An stationären Terminals in Kliniken oder über Tablets oder Smartphones werden Eltern sprachlich durch den Prozess geleitet.</p> <p>Outputs: Ausgabe der Antragsdaten bei den zuständigen Standesämtern oder der Familienkasse. Ein nachgeschaltetes System überprüft, ob mögliche Nachfragen erforderlich sind.</p>
Technische Umsetzung	Überwachtes Lernen
Verwendete KI- Komponenten	Spracherkennung (Speech-to-Text), Deep Neural Networks; gekoppelt an weitere Systeme zur intelligenten Kommunikation und zur regelbasierten Verarbeitung der Bürger:innenanfragen mit Constraint Handling Rules sowie Sprachausgabe (Text-to-Speech)
Unternehmen für Entwicklung und Betrieb	Fraunhofer FOKUS, Fraunhofer IDMT (für die Konzeption und Softwareentwicklung), Dataport (für die Implementierung und den Betrieb)
Dauer des Einsatzes	Einjähriger Pilotbetrieb zum Sammeln von Erfahrungen, geplant ab Mitte 2021
Bisheriger Projektverlauf	Das analoge Kombiformular läuft seit Juni 2018 stabil und wurde in allen Geburtskliniken Hamburgs eingeführt. Im Juni 2020 wurde der Pilotstart des Online-Dienstes »Kinderleicht zum Kindergeld« gestartet, der ebenfalls erfolgreich von den Eltern in der Asklepios-Klinik Altona angenommen wurde. Neben der Erweiterung der Zielgruppen und dem Ausrollen auf weitere Geburtskliniken wird seit April 2020 die Integration einer KI-Sprachassistenz entwickelt, die Mitte 2021 für die Pilotierung vorgesehen ist.

Was macht dieses Beispiel zu einer Best Practice?**Interne Veränderung:**

- Ein vormalig getrennter und vielschichtiger Geschäftsprozess zwischen Geburtskliniken, Standesämtern und Familienkassen konnte erstmals erfolgreich institutionsübergreifend zum Wohle der Eltern etabliert werden. Das entlastet diese erheblich von Bürokratie und gestaltet Verwaltung modern und familienfreundlich. Die Integration einer Sprachassistenten im Online-Dienst schafft einen modernen Zugang für Eltern, der höhere Datenqualität und damit auch Entlastung für die behördlichen Prozesse erwarten lässt.

Innovationsmarker:

- Die Grundidee des Projekts, d. h. das Kombiformular und der institutionenübergreifende Geschäftsprozess, lassen sich auf andere Behördenvorgänge übertragen.
- Die Pilotierung in Hamburg bietet eine Grundlage für einen möglichen bundesweiten Einsatz.
- Gemeinsame Entwicklung des Services im regelmäßigen Austausch mit Nutzer:innen (Eltern) und Kooperationspartnern (Krankenhäuser, Standesämter, Familienkassen)

Wirkungsmonitoring:

- Evaluation und Monitoring erfolgen unter Einbindung der Nutzer:innen und Kooperationspartner.
- Messungen der Zufriedenheit und Effizienz ergaben eine sehr hohe Zufriedenheit und Entlastung der Eltern.
- Eltern erhalten die Verwaltungsdienstleistungen in der Hälfte der bisher benötigten Zeit.

Ansprechperson

Dr. Brigitte Klamroth
Brigitte.Klamroth@sk.hamburg.de

8 – KI gegen Kinderpornografie (informelle Bezeichnung)

Name des Systems	KI gegen Kinderpornografie (informell)
Einsetzende Behörde	Staatsanwaltschaft Köln
Anwendungsbereiche	 Entlastung  Entscheidungs- vorbereitung  Überwachung
Fachbereich	Strafverfolgung
Themenbereich	Kinderpornografie
Beschreibung	<p>Problem: Für die Strafverfolgung der Verbreitung, des Erwerbs und des Besitzes kinder- und jugendpornografischer Schriften muss regelmäßig eine große Menge gesicherter und möglicherweise beweisrelevanter Daten ausgewertet werden. Bspw. müssen nicht justiziable Urlaubsbilder von kinderpornografischen Bildern unterschieden werden.</p> <p>Lösungsansatz: ergänzend. Eine automatisierte Bilderkennung analysiert mögliche Bilder und sortiert solche aus, die am ehesten Kinder- oder Jugendpornografie abbilden. Diese werden dann den Sachbearbeiter:innen zur näheren Untersuchung weitergeleitet.</p> <p>Outputs: Mutmaßliche kinder- oder jugendpornografische Bilder, die zur Entscheidung der Staatsanwaltschaft vorgelegt werden.</p>
Technische Umsetzung	Überwachtes Lernen
Verwendete KI- Komponenten	Bilderkennung via neuronale Netze
Unternehmen für Entwicklung und Betrieb	Forschungskonsortium aus der Universität Saarland, Microsoft und dem Deutschen EDV Gerichtstag e.V.
Dauer des Einsatzes	In Erprobung, Produktivbetrieb noch nicht absehbar
Bisheriger Projektverlauf	<p>Ende 2016: Projektidee, Bedarfsmittelungen</p> <p>2017 bis 2018: Rechtliche Prüfungen zu Bildernutzung und Cloud</p> <p>April bis Juli 2019: Entwicklung eines Abstraktionslayers</p> <p>November 2019 bis Januar 2020: Auswahl und Training mit abstrahierten Echtdaten</p> <p>April 2020: Abschluss dieser Forschungsphase</p> <p>Mai bis November 2020: Implementierungsanalyse und Demonstrator</p>
Was macht dieses Beispiel zu einer Best Practice?	<p>Zielorientierung:</p> <ul style="list-style-type: none"> • Die Grenzen des Systems sind klar gesteckt, weil es sich um ein rechtlich sensibles Verfahren handelt. Das KI-System soll die menschliche Entscheidungsebene nicht ersetzen, sondern Inhalte als Kinderpornografie vorfiltern. <p>Projektmanagement:</p> <ul style="list-style-type: none"> • Es wurden Testphasen durchgeführt, in denen das System laufend mit menschlichen Prozessen verglichen wurde, um die Fehlerquote zu ermitteln. <p>Datenhaltung und -qualität</p> <ul style="list-style-type: none"> • Die Trainingsdaten werden regelmäßig von IT-Fachkräften und der Projektsteuerung überprüft. Dieses Monitoring wird unter hohen rechtlichen Zugangsbeschränkungen zu dem kinder- oder jugendpornografischen Material durch die Behörde selbst durchgeführt.
Ansprechperson	Markus Hartmann E-Mail: Markus.Hartmann@sta-koeln.nrw.de

9 – KI zur Steuerung der Fallprüfung im Regressmeldeverfahren

Name des Systems	KI zur Steuerung der Fallprüfung im Regressmeldeverfahren
Einsetzende Behörde	Berufsgenossenschaft Energie Textil Elektro Medienerzeugnisse (BG ETEM)
Anwendungsbereiche	 Entscheidungs- vorbereitung  Entlastung
Fachbereich	Versicherung
Themenbereich	Regressverfahren, Fallprüfung
Beschreibung	<p>Problem: Kommt es am Arbeitsplatz oder auf dem Arbeitsweg zu einem Unfall, verursacht durch betriebsfremde Dritte, prüft die Berufsgenossenschaft, ob diese dritte Person in Regress genommen werden kann. Im Falle eines Regressverdachts wurde ein standardisiertes Dokument an die zentrale Stelle für Regressverfahren weitergeleitet und händisch vorgeprüft.</p> <p>Lösungsansatz: substituierend. Anhand von etwa einer Million abgeschlossener (erfolgreicher und nicht erfolgreicher) Regressverfahren wird eine KI trainiert, zu bewerten, inwieweit in aktuellen und alten Fällen eine ausreichende Wahrscheinlichkeit für ein erfolgreiches Regressverfahren besteht. Mithilfe von Textmining werden die Daten zusätzlich nach kritischen Schlagwörtern durchsucht, um weitere unterstützende Informationen für die spätere Entscheidung durch den/die Sachbearbeiter:in bereitzustellen (Wurde die Polizei hinzugerufen? Kam es zu einer Tötlichkeit?). Die Prognosen und alle zur Entscheidung notwendigen Informationen werden der Regressabteilung über ein Web-Interface zur Verfügung gestellt.</p> <p>Outputs: Die KI berechnet die Plausibilität für ein mögliches Regressverfahren, die Höhe der Gesamtkosten und prognostiziert zugleich die Erfolgswahrscheinlichkeit.</p>
Technische Umsetzung	Überwachtes Lernen
Verwendete KI- Komponenten	Decision Tree, Textmining
Unternehmen für Entwicklung und Betrieb	Es handelt sich um eine Eigenentwicklung durch die anacision GmbH eigens für die BG ETEM. Partielle Nutzung von IBM-Frameworks.
Dauer des Einsatzes	Seit Dezember 2019 im laufenden Betrieb
Bisheriger Projektverlauf	Nach einer Machbarkeitsstudie 2017 wurde das System 2018 ausgebaut. Nach einer dreimonatigen Testphase ging das System Dezember 2019 in Betrieb.
Was macht dieses Beispiel zu einer Best Practice?	<p>Sicherheit:</p> <ul style="list-style-type: none"> • Es findet ein wöchentliches Monitoring inkl. Plausibilitätschecks statt. Bei der Datensicherheit gibt es zudem eine Parallelstruktur, indem die Daten zusätzlich bei den entsprechenden Organisationen (z. B. Polizei, Ärzt:innen) gespeichert werden. • Es besteht jederzeit mit moderatem Aufwand die Möglichkeit, zu einer manuellen Vorgehensweise zurückzukehren. <p>Datenhaltung und -qualität:</p> <ul style="list-style-type: none"> • Im Rahmen des wöchentlichen Monitorings wird zudem überprüft, ob es z. B. Auffälligkeiten bei den Datenquellen gibt und ob sich Input und Output verändert haben. Das ermöglicht gemeinsam mit dem ständigen Zuwachs an Daten eine wöchentliche Optimierung der Prognosegüte. • Ergänzend findet jährlich eine umfassende Evaluation der prognostizierten und eingeleiteten Regressverfahren statt. <p>Nachvollziehbarkeit:</p> <ul style="list-style-type: none"> • Die verwendeten Anwendungsmodelle werden transparent gemacht. So können z. B. die Regeln extrahiert werden, die hinter jeder Entscheidung stehen, und wichtige Entscheidungsmerkmale visualisiert werden (»predictor importance«).
Ansprechperson	Dr. Marcel Beining marcel.beining@anacision.de

10 – Maschinelles Übersetzungsservice der Justiz

Name des Systems	Maschinelles Übersetzungsservice der Justiz
Einsetzende Behörde	Ministerium der Justiz und für Europa (JUM), Baden-Württemberg
Anwendungsbereiche	 Entlastung
Fachbereich	Justiz
Themenbereich	z. B. Internationale Rechtshilfe, Justizverfahren
Beschreibung	<p>Problem: Im Justizumfeld müssen immer wieder kurzfristig Rechtstexte in Fremdsprachen als einschlägig identifiziert sowie anschließend übersetzt werden. Bestehende Angebote großer Plattformen sind aufgrund der mangelnden Qualität und des Datenschutzes nicht nutzbar. Gleichzeitig ist eine verlässliche händische Übersetzung vieler verschiedener Sprachen mit eigenen Kompetenzen nicht verfügbar.</p> <p>Lösungsansatz: substituierend. Unter Nutzung eines mit übersetzten Texten trainierten neuronalen Netzes werden maschinelle Übersetzungen erstellt. Fremdsprachige Dokumente können so gezielter durchsucht werden, um sich z. B. einen schnellen Überblick zu verschaffen. Ein Ziel des weiteren Ausbaus des Projekts ist, ein Translation Memory zu entwickeln: Das bedeutet, das alles, was schon einmal als Input bzw. Output prozessiert wurde, wiedererkannt wird. Im Regelbetrieb soll dies auf eine zentrale Übersetzungsplattform verlagert werden. Insbesondere zeitkritische Übersetzungen sollen damit in Zukunft strukturierter erfolgen. Zudem soll das System die Schwelle für Arbeitsübersetzungen senken, die in Sekundenschnelle selbst angefertigt werden können, und einen ersten Zugriff auf fremdsprachige Dokumente erlauben.</p> <p>Outputs: Übersetzte Texte bei beibehaltener Formatierung. Das eingespeiste Dokument wird ggf. händisch geprüfert.</p>
Technische Umsetzung	Überwachtes Lernen
Verwendete KI- Komponenten	Neuronale Netzwerke
Unternehmen für Entwicklung und Betrieb	SYSTRAN (Frankreich)
Dauer des Einsatzes	Noch in Pilotierungsphase
Bisheriger Projektverlauf	2017 wurde die Projektidee entwickelt. Im Oktober 2018 wurde das System installiert und eine Pilotphase begann. 2019 wurde das System evaluiert und von zehn auf 25 Sprachen erweitert. Der Produktivbetrieb ist für den Jahresbeginn 2021 vorgesehen.
Was macht dieses Beispiel zu einer Best Practice?	<p>Innovationsmarker:</p> <ul style="list-style-type: none"> • Die Übersetzungsplattform wird zunächst für die eigenen Zwecke entwickelt, soll aber skalierbar gemacht werden. Das Übersetzen kann dann behördenübergreifend strukturierter ablaufen. <p>Interne Veränderung:</p> <ul style="list-style-type: none"> • Die Identifikation relevanter Dokumente, die in der Folge für eine Übersetzung näher betrachtet werden, wurde vollständig automatisiert. Arbeitsübersetzungen können kurzfristig erstellt werden, Prozesse werden dadurch vereinfacht. <p>Wirkungsmonitoring:</p> <ul style="list-style-type: none"> • Auch in der Pilotphase werden zweimal jährlich Anwender:innenbefragungen mit sieben bis acht Zufriedenheitsskalen durchgeführt. Diese Ergebnisse werden dann an das Entwickler:innenteam weitergeleitet.
Ansprechperson	Björn Beck Bjoern.Beck@jum.bwl.de

11 – PREVIEW

Name des Systems	PREVIEW - Prediction Visualisation and Early Warning
Einsetzende Behörde	Auswärtiges Amt
Anwendungsbereiche	 Entscheidungs-vorbereitung  Planung  Überwachung
Fachbereich	Krisenprävention und Stabilisierung
Themenbereich	Instrumente und Prozesse in der Krisenfrüherkennung
Beschreibung	<p>Problem: Um politische Krisen zu verhindern bzw. einzudämmen, sollten diese im Voraus erkannt werden können. Die aktuelle Berichterstattung kann dabei wichtige Hinweise geben. Hierbei handelt es sich um eine große Menge an Daten, die bislang händisch auszuwerten war.</p> <p>Lösungsansatz: ergänzend. Ein multifaktorieller Ansatz bei der Erstellung von Lagebildern führt zu einer präziseren Früherkennung von Krisen. Dafür werden öffentlich verfügbare Daten genutzt, darunter aktuelle Berichterstattungen über Krisensituationen (allerdings noch ohne Einbeziehung von Social Media), NGO-Berichte, Konfliktereignisse, diverse sozioökonomische Indikatoren sowie Wetter- und Klimadaten. Anhand dieser Daten kann bspw. prognostiziert werden, wie viele Konflikttote in einer bestimmten Region zu erwarten sind oder welche Regionen als konfliktisikoreich eingestuft werden müssen. Diese werden dann ggf. qualitativ näher untersucht. Dabei besteht das System aus einer einheitlichen Struktur und Entwicklungsumgebung, während die jeweiligen Forecast-Modelle individuell erstellt werden.</p> <p>Outputs: unterschiedlich, bspw. Identifizierung von Regionen mit Konfliktisiko, Schätzung bzw. Hochrechnung von Konflikttoten für bestimmte Regionen/Länder und konkreten Zeitraum.</p>
Technische Umsetzung	Unüberwachtes Lernen
Verwendete KI- Komponenten	Maschinelles Lernen, Random Forecast
Unternehmen für Entwicklung und Betrieb	Stand-alone- und Eigenentwicklung, nur stellenweise Beteiligung von externen Dienstleistern
Dauer des Einsatzes	2 Jahre
Bisheriger Projektverlauf	2016-2018: Projektentwicklung 2018-2020: Aufbauphase mit erstem Betrieb seit Anfang 2020: Regelbetrieb
Was macht dieses Beispiel zu einer Best Practice?	<p>Datenhaltung und -qualität: Die Daten werden bei jedem Durchlauf neu geprüft und die Modelle auch immer wieder mit neuen Daten trainiert. Trainings- und Betriebsdaten werden in versionierten Repositorien festgehalten.</p> <p>Nachvollziehbarkeit: Es besteht regierungsintern eine vollständige Kontrolle und Offenheit über die zugrunde liegenden Modelle, Codes und Methoden. Im Rahmen eines Expert:innenworkshops wird für die regierungsinterne Nutzung erläutert, welche Daten herangezogen werden. Es gibt zudem ein Erklärpapier für Anwender:innen.</p>
Ansprechperson	Hans-Christian Mangelsdorf s05-9@auswaertiges-amt.de

12 – Profilanalyse von Gesprächsprotokollen

Name des Systems	Profilanalyse von Gesprächsprotokollen
Einsetzende Behörde	Bundesamt für Migration und Flüchtlinge (BAMF)
Anwendungsbereiche	 Entscheidungs- vorbereitung  Entlastung
Fachbereich	Asylverfahren/Sicherheit
Themenbereich	Profilanalyse
Beschreibung	<p>Problem: Das BAMF hat Kontakt mit Geflüchteten, die z. B. als Opfer von Terrorismus über Informationen verfügen können, die für die Sicherheitsbehörden des Bundes von Interesse sind. Um entsprechende Personen zu identifizieren, werden Anhörungen überprüft. Das Erkennen von melderelevanten Inhalten war bislang mit einem hohen Zeitaufwand für die Mitarbeitenden verbunden.</p> <p>Lösungsansatz: substituierend. Das System basiert auf diversen Teilkomponenten. Kernmodule sind Texterkennung sowie Module zur maschinellen Verarbeitung natürlicher Sprache. Dafür wurden bisherige Gesprächsprotokolle und die darin markierten relevanten Textstellen als Grundlage für ein maschinelles Lernverfahren verwendet. Mithilfe der KI werden Anhörungsprotokolle nun nach ähnlichen Stellen durchsucht, die den Sachbearbeiter:innen zur Prüfung für die Meldung an Nachrichtendienste vorgelegt werden.</p> <p>Outputs: melderelevante Anhörungen bzw. darin enthaltene Textstellen</p>
Technische Umsetzung	Unüberwachtes Lernen
Verwendete KI- Komponenten	Neuronale Netze und linguistische Modelle
Unternehmen für Entwicklung und Betrieb	Das System wurde durch IBM entwickelt und ist in den Meldeprozess integriert.
Dauer des Einsatzes	Pilotbetrieb seit Ende 2018
Bisheriger Projektverlauf	Das Pilotprojekt wurde 2017 initiiert. Eine umfassende Einführung ist, abhängig vom weiteren Verlauf der Covid-19-Pandemie, für Mitte 2021 vorgesehen.
Was macht dieses Beispiel zu einer Best Practice?	<p>Innovationsmarker:</p> <ul style="list-style-type: none"> • In diversen anderen Bereichen im BAMF gibt es Interesse daran, ähnliche Prozesse ebenfalls maschinell zu unterstützen. Zudem findet ein allgemeiner Austausch zu generellen Erfahrungen und Umsetzungsansätzen, u. a. im Rahmen des NeXT-Netzwerks, regelmäßig statt. <p>Projektmanagement:</p> <ul style="list-style-type: none"> • Die Entwicklungsergebnisse werden wöchentlich geprüft. Um den Prüfprozess zu unterstützen, wurden Systeme entwickelt, die es fachverantwortlichen Personen ermöglichen, die Qualität des KI-Systems einschätzen und individuelle Anpassungen vornehmen zu können, ohne externe IT-Kräfte beauftragen zu müssen. <p>Akzeptanz:</p> <ul style="list-style-type: none"> • In Mitarbeiter:innenschulungen werden der Umgang mit dem System und grundlegende Informationen zu den eingesetzten Techniken vermittelt. • Anwender:innen können Ausgaben des Systems einfach mit »Ja/Nein« bewerten und mit Hinweisen bzw. Kommentaren versehen. Diese Informationen stehen Entwickler:innen später zur Optimierung des Systems zur Verfügung.
Ansprechpersonen	<p>Christian Wolter Christian.Wolter@bamf.bund.de</p> <p>Johannes Wenisch Johannes.Wenisch@bamf.bund.de</p>

KONTAKT

Nicole Opiela
Kompetenzzentrum Öffentliche IT (ÖFIT)
Tel.: +49 30 3463-7173
Fax: +49 30 3463-99-7173
info@oeffentliche-it.de

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

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

ISBN: 978-3-948582-04-3

