

Von Einzelexperimenten zur Gesamtstrategie: KI in der Schweizer Verwaltung nutzen

Dr. Melanie Kolbe-Guyot

Berkeley

**TECH
POLICY
FELLOWSHIP**

EPFL

**Center for
Digital Trust**



Kanton Basel-Stadt

Einführung

- **Grosses Interesse**, aber bislang geringe, pilotgetriebene und rechtlich unzureichend geregelte KI-Nutzung in der Schweizer Verwaltung.^{***}
- Viele Pilotprojekte bleiben **isoliert und wenig skaliert**.
- KI-Einsatz konzentriert sich auf **interne Effizienz** und generative KI-Anwendungen; bürgerorientierte Anwendungen beschränken sich meist auf **(sichtbare) Chatbots**.



Wie kann die Schweizer Verwaltung den Übergang von isolierten Pilotprojekten zu einer breiten, verantwortungsvollen und wirkungsvollen KI-Nutzung schaffen?

^{***} Fokus auf KI-Einsatz in der öffentlichen Verwaltung. Ausgeschlossen: Justiz, Polizei, Grenzkontrolle und Privatsektor, ausser zur Illustration relevanter Entwicklungen und Risiken.

Hauptargumente

- **KI in der Verwaltung** beschränkt sich nicht auf Chatbots oder generativer KI, sondern **umfasst ein breites Spektrum an Anwendungen** zur Unterstützung staatlicher Aufgabenerfüllung.
- Der breite, sichere und wirksame Einsatz von KI in der Schweizer Verwaltung wird durch eine **Reihe struktureller Barrieren** eingeschränkt.
 - Diese erfordern ein **koordiniertes, strategisches Vorgehen** statt isolierter und technologiegetriebener Einzelinitiativen ...
 - und eine **gezielte Stärkung** der rechtlichen, institutionellen und technischen **Grundlagen**.
- Führungskräfte müssen KI konsequent am **öffentlichen Mehrwert/Nutzen** ausrichten und deren vertrauenswürdigen, skalierbaren und wirkungsvollen Einsatz aktiv steuern.

Fundament der präsentierten Ergebnisse

C4DT Insight #5 From Cautious Experimentation to Coherent Strategy: Harnessing AI's Potential in the Swiss Public Administration

- **Design:** Literatur- und Policy-Review/Analyse + 20 Interviews mit Experten und Expertinnen (Feb–Apr 2025) in Bundes-/Kantons-verwaltung, Wissenschaft, Zivilgesellschaft und Consulting.
- **Triangulation:** Verwandte Studien, Desk Research und internationalen Vergleiche.
- **Use-Case-Mapping:** Eigene Forschung und externen Repositorien.
- **Limitationen:** Nicht-randomisierte Expertenstichprobe; dynamische und begrenzte Transparenz bei KI-Einsatz im öffentlichen Sektor → Ergebnisse des Use-Case-Mapping sind indikativ, nicht final.



Definition von KI und algorithmischen Systemen

- **KI kein klar abgegrenzter Begriff**, oft uneinheitlich verwendet und reicht von einfachen Modellen bis zu komplexen, lernenden Systemen.

OECD Definition: Künstliche Intelligenz (KI) wird als ein maschinenbasiertes System verstanden, das für explizite oder implizite Ziele aus den erhaltenen Eingaben ableitet, wie Ausgaben wie etwa Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen erzeugt werden, welche physische oder virtuelle Umgebungen beeinflussen können. KI-Systeme unterscheiden sich dabei hinsichtlich ihres Autonomiegrads und ihrer Anpassungsfähigkeit nach der Inbetriebnahme.

- Wichtige Unterscheidung: algorithmische Systeme vs. KI
 - **Regelbasierte Systeme:** feste „wenn-dann“-Regeln, transparent, deterministisch, aber nicht selbstlernend → **technisch gesehen keine KI.**
 - **Machine Learning (KI):** lernt Muster aus Daten, adaptiv, leistungsfähig, aber weniger transparent und probabilistisch.
 - **„Algorithmische Systeme“:** Umfasst beide, da **ähnliche Risiken und Konsequenzen** (Bias, Diskrim.).

KI im öffentlichen Sektor (vs. im privaten)

- Ziel ist **Effizienz und Verbesserung bestehender Verwaltungsaufgaben**, nicht Gewinnmaximierung oder neue Geschäftsmodelle.
- **KI muss öffentlichen Mehrwert (Public Value) schaffen**, z. B. bessere Services, mehr Fairness und höhere Servicequalität für Einwohner und Einwohnerinnen.
- **Höhere Anforderungen an Transparenz, Fairness und Rechenschaftspflicht**, da potenzielle negative Auswirkungen persönliche Rechte und öffentliches Vertrauen direkt betreffen.
- **Grössere Risiken und institutionelle Herausforderungen**, z. B. durch rechtliche Vorgaben, vielfältige Stakeholder und hohe Risikoaversion.



KI in der Verwaltung: Motivationen und Use-Cases (international)

Frequenz	Motivation	Anwendung	Use-Case Beispiele
■■■■■	Effizienzsteigerung interner Abläufe	Automatisierung repetitiver Aufgaben und Entlastung von Mitarbeitenden	Automatisierung (Sortieren, Abfragen, Anonymisieren) und Assistenzsysteme
■■■■□	Verbesserte Aufsicht und Betrugsbekämpfung	Erkennung von Risiken und Unregelmässigkeiten zur Einhaltung staatlicher Aufsicht u. Rechenschaftspflicht	Interne Compliance oder Betrugserkennung in Steuern, Zoll oder Sozialleistungen
■■■□□	Bessere Servicequalität	Schnellere, zugänglichere und personalisierte öffentliche Dienstleistungen	Chatbots oder Assistenzsysteme für Informationszugang, Steuererklärungen etc.
■■■□□	Demografischer Wandel und Fachkräftemangel	Kompensation begrenzter personeller Ressourcen und Sicherung von Wissen	Virtuelle Assistenten, Wissensmanagement (Kategorisierung und Auffindbarkeit von Informationen), Entscheidungshilfen
■■□□□	Verbesserte Entscheidungsgrundlagen (Strategische Planung und Policy Evaluation)	Analyse grosser Datenmengen für Prognosen, Monitoring, Optimierung	Frühwarnsysteme, Smart Cities und Infrastrukturplanung, Resourcemanagement
■□□□□	Verbesserte Entscheidungsgrundlagen (Entscheidungssysteme)	Unterstützt Entscheidungen durch Datenanalyse und Empfehlungen sowie teilweise Automatisierung dieser	Automatisierte Entscheidungs- und Profiling-Systeme (Anspruchsprüfung, Risiko- bzw. Erfolgsbewertung)

Sensibler KI-Einsatz: Sozial- und Steuerverwaltung

- **Effizienzpotenzial:** Automatisierte und unterstützte Fallbearbeitung; verbesserte Betrugs- und Risikoerkennung.
- **Breites Anwendungsspektrum:** Von Assistenzsystemen (z. B. Chatbots, Antragsautomatisierung) bis zu Anspruchsprüfung, Profiling und Risikobewertung.
- **Zentrale Problemfelder und Risiken** (Analyse 9 internationaler Fallbeispiele):
 - Datenschutz- und Überwachungsprobleme, Diskriminierung und Bias, mangelnde Transparenz und Erklärbarkeit sowie fehleranfällige Bewertungen und unklare Verantwortlichkeiten.
 - Fehlerhafte Systeme können zu finanziellen, sozialen Schäden sowie staatlichem Vertrauensverlust führen.



Wegen direkter Auswirkungen auf Lebensgrundlagen sind Transparenz, Rechtsschutz und menschliche Aufsicht entscheidend.

Sensibler KI-Einsatz: Automatisierte Entscheidungen

- **KI unterstützt Entscheidungen vor allem halb-automatisiert**, z. B. durch Datenanalyse, Mustererkennung und Entscheidungsempfehlungen.
- **Vollautomatisierte Entscheidungen sind selten**, aufgrund rechtlicher, ethischer und Transparenzprobleme.
 - Vorwiegend **regelbasierte Systeme** auf Basis gesetzlicher Vorgaben.
- **Menschliche Aufsicht** bleibt zentral, aber: begrenzte Zeit, Ressourcen und Expertise.
 - Risiko von „**Automation Bias**“.
- **Offene Grundsatzfrage**: Konsistentere KI-Entscheidungen vs. Risiko systematischer Fehler und begrenzter Kontextberücksichtigung - akzeptabel Fehlerrisiko?



Spannung zwischen Effizienzgewinnen und Anforderungen an Fairness, Transparenz und Rechenschaftspflicht.

KI in der Schweizer Verwaltung: Governance

- Mangels horizontaler KI-spezifischer Gesetze, Orientierung an **bestehenden sektoralen Regulierungen** (z. B. Datenschutzbestimmungen), **Fachgesetzen und internen Leitlinien**.
- Geplante Regulierung im Zuge der Umsetzung der KI-Konvention des Europarats weiterhin, aber weiterhin **Unsicherheiten bei Definition, Governance, Aufsicht und Grundrechtsschutz**.
- Bisher **nur wenige umfassende KI-Strategien** für die Verwaltung (Bundesverwaltung, Kantone Aargau und Luzern).
- Gleichzeitig verzichten manche Kantone bewusst auf eine übergreifende Strategie.

KI in der Schweizer Verwaltung: Diskurs, Motivation, Bedenken

Wahrnehmung und Verständnis von KI

- ❖ **Ambivalente Wahrnehmung:** KI wird zugleich als Modernisierungschance und als Risiko für Vertrauen, Beschäftigte und staatliche Legitimität gesehen.
- ❖ **Uneinheitliches Begriffsverständnis:** KI umfasst verschiedenste Anwendungen; generative KI dominiert aber die öffentliche und interne Wahrnehmung.
- ❖ **Konzeptionelle Unklarheit verzerrt Prioritäten:** KI wird teils als Buzzword genutzt (für verpasste Digitalisierung), wodurch strategisch relevante Anwendungen übersehen und der Anwendungsfokus uneinheitlich gesetzt werden.

Motivationen

- ❖ **Pragmatische Treiber dominieren:** Hauptmotive sind Effizienzsteigerung, Zeitersparnis, Pensionswelle, internes Wissensmanagement sowie Modernisierungs- und Leistungsdruck; strategische Transformation oder bürgerzentrierte Innovation stehen weniger im Fokus.

KI in der Schweizer Verwaltung: Diskurs, Motivation, Bedenken

Zentrale Bedenken und Risiken

- ❖ **Recht und Governance:** Unklare gesetzliche Grundlagen, Verantwortlichkeiten und fehlende Aufsichtsstrukturen erschweren einen rechtssicheren Einsatz.
- ❖ **Transparenz und Vertrauen:** Begrenzte Nachvollziehbarkeit sowie Risiken von Verzerrungen und Diskriminierung gefährden Vertrauen in staatliche Entscheidungen.
- ❖ **Einsatz und Kontrolle:** Unsicherheit über geeignete Anwendungsbereiche unterstreicht die Notwendigkeit klarer Grenzen und wirksamer menschlicher Aufsicht.



KI in der Schweizer Verwaltung: Use-Case Mapping

Kategorie	Beschreibung	Interviews / Forschung (# / %)	Atlas of Automation (# / %)	CNAI DB (# / %)
Chatbots und virtuelle Assistenten	Bereitstellung von Informationen, Beantwortung von Anfragen und Unterstützung von Nutzenden auf Websites oder Portalen, häufig unter Verwendung von Retrieval-Augmented Generation (RAG).	8 (27%)	19 (38%)	13 (29%)
Analytische Entscheidungsunterstützung und Planung	Einsatz von Machine Learning für Prognosen, Ausreisser-Erkennung und Politikplanung.	6 (20%)	12 (24%)	13 (29%)
Automatisierte Entscheidungsfindung (KI- und regelbasierte Systeme)	Automatisierung von Entscheidungen, Genehmigungen und Zuweisungen in Bereichen mit geringem Ermessensspielraum.	5 (17%)	7 (14%)	0 (0%)
Textverarbeitung, Transkription und Informationsmanagement	Automatisierung der Textverarbeitung, Übersetzung/Transkription, Anonymisierung sowie Strukturierung von Dokumenten und Berichten.	7 (23%)	3 (6%)	4 (9%)
Prozessautomatisierung und Workflow-Unterstützung	Automatisierung repetitiver Aufgaben, Kommunikation und Ressourcenallokation.	1 (3%)	9 (18%)	7 (15%)
Wissensmanagement	Verbesserung des Zugangs zu institutionellem Wissen durch semantische Suche, Bild- und Dokumentenklassifikation sowie Datenabruf.	3 (10%)	0 (0%)	8 (18%)
Gesamtzahl Anwendungsfälle		30	50	45

Muster der KI-Nutzung in der Schweizer Verwaltung

Adoption: begrenzt und fragmentiert

- Viele Pilotprojekte, aber wenige produktive, skalierte Anwendungen
- Initiativen oft dezentral und bottom-up statt strategisch gesteuert
- Insgesamt vorsichtige und selektive Einführung

Use Cases: intern, unterstützend und risikoarm

- Fokus auf interne Effizienz (Zusammenfassungen, Suche, Transkription, Wissensmanagement)
- GenAI-Tools und Chatbots dominieren die Praxis
- Automatisierung v. a. in strukturierten Bereichen (z. B. Steuerwesen)

Governance & Constraints: hohe Hürden für transformative Nutzung

- Einsatz in sensiblen Entscheidungsbereichen bislang kaum verbreitet
- Regelbasierte Systeme häufiger als datengetriebene Entscheidungs-KI
- Vollautomatisierte Entscheidungen bleiben Ausnahme, da rechtliche, ethische und vertrauensbezogene Anforderungen oft nicht erfüllt sind.

Plädoyer für mehr Public Value-Orientierung

- Bisher Fokus auf operative KI, wenig direkter öffentlicher Mehrwert.
- Das ist schade, denn KI kann ... :
 - **staatliche Leistungen verbessern**, z. B. durch schnelleren Zugang, höhere Servicequalität und stärkere Personalisierung der Angebote.
 - die **Qualität und Fairness von Entscheidungen erhöhen**, etwa durch mehr Konsistenz, Analysestärke und bessere Unterstützung benachteiligter Gruppen.
 - Durch mehr **Zeit** für komplexe **Anliegen**/Fälle, **innovative Projekte**, etc. schaffen.



Öffentlicher Mehrwert entsteht nur, wenn Effizienzgewinne zu konkreten Vorteilen für Einwohner und Einwohnerinnen führen – nicht nur zu Kostensenkung oder sogar Stellenabbau.

Warum stockt die Umsetzung?



Fehlende Grundlagen

Fragmentierte Daten, unzureichende Infrastruktur und mangelndes Fachpersonal erschweren einen systematischen Einsatz.

Geringe organisatorische und strategische Reife

Mangelnde KI- sowie Prozesskompetenzen. Viele Pilotprojekte scheitern an Skalierung, fehlender Governance, Ressourcen und strategischer Einbettung.

Rechtliche Unsicherheit und fehlende Governance

Unzureichende gesetzliche Grundlagen, Verantwortlichkeiten und Aufsichtsstrukturen bremsen den Einsatz, insbesondere in sensiblen Bereichen.

Technische Grenzen und Vertrauensrisiken

Begrenzte Erklärbarkeit, Genauigkeit und Verzerrungen erschweren einen verlässlichen und fairen Einsatz.

Abhängigkeit von ausländischen Anbietern

Fehlende digitale Souveränität und eigene (Cloud-) Infrastruktur begrenzen Kontrolle über Daten und strategische KI-Entscheidungen.

Empfehlungen

Die Verwaltung muss von isolierten Pilotprojekten zu einer strategischen, rechtskonformen und kompetenzgestützten KI-Nutzung übergehen – mit klarem Fokus auf öffentlichen Mehrwert, Vertrauen und nachhaltige Umsetzung.

1. Zielgerichteten und strategischen Ansatz wählen

- KI gezielt dort einsetzen, wo sie messbaren Nutzen für Staat und Bevölkerung schafft („public value first“) – „KI um der KI willen“ vermeiden.
- Klare Ziele und Prioritäten definieren und Governance-Strukturen für KI-Projekte etablieren.
- Zentrale Koordinationsstelle für KI schaffen (kantonal oder departemental), um KI-Anwendungsfälle systematisch zu steuern.
- Klare Mindestanforderungen definieren (Datenreife, öffentlicher Mehrwert) und systematische Prozessanalysen und Evaluationen verlangen.



Empfehlungen

2. Regulierung und Governance stärken

- Klare gesetzliche Grundlagen schaffen und Transparenz, Erklärbarkeit, Rechenschaftspflicht sowie Einspruchsrechte sicherstellen.
- KI-Anwendungen nach Risiko bewerten und strenge Schutzmassnahmen, insbesondere in sensiblen Bereichen, verbindlich machen.
- Risiko- und Folgenabschätzungen (vor und während des Einsatzes) verbindlich machen.
- Unabhängige Aufsicht stärken, Audits etablieren und öffentliche KI-Register aufbauen.



Empfehlungen

3. Infrastruktur, Daten und digitale Souveränität ausbauen

- In interoperable Dateninfrastruktur und souveräne IT- und KI-Systeme investieren.
- Gemeinsame Plattformen und Zusammenarbeit über Kantonsgrenzen hinweg fördern, um Ressourcen effizient zu nutzen.
- Datenqualität und -verfügbarkeit, Integration und sichere Verarbeitung verbessern.

4. Kompetenzen und organisatorische Fähigkeiten entwickeln

- KI-Kompetenzen, Data Literacy und interdisziplinäre Fähigkeiten gezielt fördern.
- AI-Champions einsetzen, um verantwortungsvolle Nutzung zu fördern.
- Einheitliche Leitlinien, Checklisten und Tools bereitstellen, um Entwicklung, Einsatz und Governance von KI zu strukturieren.
- Systematische Pilotierung ermöglichen und erfolgreiche Lösungen skalieren.
- Mitarbeitende und Öffentlichkeit frühzeitig und transparent einbinden.



Fazit

- KI zentral für **Leistungsfähigkeit, Effizienz und Resilienz des Staates** – bei steigender Komplexität, Serviceerwartung und Fachkräftemangel.
- **Es fehlen aber zentrale Grundlagen**: Infrastruktur, Daten, rechtliche Klarheit, Governance und Kompetenzen.
- **Ohne strategischen Aufbau** bleibt KI-Potenzial ungenutzt und das Risiko von **Fehlinvestitionen** und **verpassten Innovationschancen** steigt.



Vielen Dank!
Merci beaucoup!
Grazie mille!