

KI verändert die Organisationslogik

Mögliche Handlungsfelder für die SEVAL

**Bericht von der gemeinsamen Frühjahrstagung der DeGEval-Arbeitskreise
«Methoden in der Evaluation» und «Verwaltung», 21.–22. Mai 2026, Hoch-
schule Campus Wien, Wien**

Inhalt

1. Rückblick auf die Frühjahrstagung	4
Die Ausgangslage: KI verändert die Organisation schneller, als Strukturen entstehen	5
KI ist keine weitere Technologie – sie verändert die Organisationslogik	5
Organisationale Lern- und Steuerungsfähigkeit entscheidet über Erfolg	6
Die Steuerungsinstrumente des DEval: Reifegradmodell und Integrationsprozessmodell.....	7
Bedeutung der KI-Transformation für die Evaluation	8
Organisationale Voraussetzungen für einen verantwortbaren KI-Einsatz	9
2. Mögliche Handlungsfelder für die SEVAL	9
Von den Tagungsbefunden zu möglichen Handlungsfeldern der SEVAL.....	9
Vertrauenswürdigkeit als Qualitätsdimension für KI-unterstützte Evaluationen klären.....	10
Qualitätsinfrastruktur und professionelle Kompetenz bereitstellen.....	10
Chancengleichheit im Evaluationsmarkt unterstützen.....	11
Rolle der SEVAL im KI-Transformationsprozess definieren	11
Kooperationen und realistische Umsetzung.....	11
Fazit	12
Literaturverzeichnis	13

Notizen abrufbar:

<https://www.degeval.org/arbeitskreise/methoden-in-der-evaluation/aktuelles/>

- FH-Prof.in Mag.a Dr.in Elisabeth Haslinger-Baumann, Vizerektorin für Forschung und Entwicklung der Hochschule Campus Wien: **»Grußwort«**
- Prof. Dr. Ulrich Lenz: **»Zwischen Heilsversprechen und systemischem Risiko: Ethische Leitlinien für die verantwortliche KI-Transformation«**
- Isabella E. Wagner & Thomas Palfinger: **»Generative KI als Faktor im Evaluationssystem: veränderte Beziehungen und Praktiken«**
- Dorian Wachsmann: **»KI & Evaluation: Lessons learned aus der öffentlichen Verwaltung«**
- Markus Tumeltshammer: **»Algorithmische Praktiken der Verwaltung als offener Evaluationsgegenstand und collaborative research«**
- Werkstattberichte
 - Dr. Sascha Kopczynski: **»Einsatz von KI-Unterstützung in der Hochschulevaluation – Erfahrungsbericht über Software-automatisierte Inhaltsanalysen von Freitextkommentaren in Studierendenbefragungen«**
 - Dr. Anke Schad-Spindler & Lilian Häge: **»Qualitative Inhaltsanalysen mit MAXQDA AI Assist: Mehr Effizienz, mehr Kontrollbedarf?«**
 - Tabea Krauter: **»KI zur Unterstützung qualitativer Analyse in Evaluationen«**
 - Laura Kunert & Kai Rompczyk: **»KI-Integration im Spannungsfeld zwischen Innovation und Regulierung«**
- Prof. Dr. Jan Hense: **»Wrap-up und Abschlussdiskussion zur Tagung«**

Bericht zur gemeinsamen Frühjahrstagung der DeGEval-Arbeitskreise «Methoden in der Evaluation» und «Verwaltung», 21.–22. Mai 2026, Hochschule Campus Wien, Wien

Impulsbeitrag: Künstliche Intelligenz verändert die Organisationslogik – mögliche Handlungsfelder für die SEVAL

1. Rückblick auf die Frühjahrstagung

Zwei Tage lang diskutierten rund 100 Teilnehmende aus Wissenschaft, Verwaltung und Evaluationspraxis aus dem D-A-CH-Raum auf der gemeinsamen Frühjahrstagung, wie generative Künstliche Intelligenz (genKI) Organisationen, Evaluation und Evaluationssysteme verändert und welche methodischen, organisatorischen, technischen und ethischen Konsequenzen sich daraus ergeben.

Nach dem Grusswort von FH-Prof.in Mag.a Dr.in Elisabeth Haslinger-Baumann (Hochschule Campus Wien) spannte die Keynote von Prof. Dr. Ulrich Lenz (Hochschule für angewandtes Management; LENZ Advisory Services) den Bogen von der Funktionsweise generativer KI über den EU AI Act bis zu Responsible AI und KI-Kompetenz. Isabella E. Wagner (fteval) und Thomas Palfinger (Ludwig Boltzmann Gesellschaft) stellten generative KI als Faktor im Evaluationssystem und als Veränderung von Beziehungen und Praktiken dar. Dorian Wachsmann (Fraunhofer FOKUS, Kompetenzzentrum Öffentliche IT) zeigte anhand der öffentlichen Verwaltung, wo GenKI-Systeme bereits entstehen und welche Stolpersteine bei Evaluationen mit und von GenKI auftreten. Markus Tumeltshammer (Austrian School of Government/Bundeskanzleramt) rückte algorithmische Praktiken der Verwaltung als Evaluationsgegenstand und als Feld kollaborativer Forschung in den Mittelpunkt. Die Werkstattberichte von Dr. Sascha Kopczynski (HSPV NRW), Dr. Anke Schad-Spindler und Lilian Häge (EDUCULT), Tabea Krauter (Fraunhofer ISI) sowie Laura Kunert und Kai Rompczyk (DEval) vertieften konkrete Einsatzfelder: Freitextanalysen in der Hochschulevaluation, qualitative Inhaltsanalyse mit MAXQDA AI Assist, KI-gestützte qualitative Analyse sowie die organisationale KI-Integration im Spannungsfeld zwischen Innovation und Regulierung. Prof. Dr. Jan Hense ordnete die Tagung im abschliessenden Wrap-up historisch und professionspolitisch ein und stellte Vertrauen, Verantwortung und die aktive Gestaltung der KI-Transformation in den Mittelpunkt.

Der vorliegende Beitrag greift aus der Breite der Tagungsbeiträge die organisationale Verankerung von KI heraus. Ausgangspunkt ist die Beobachtung, dass KI den Evaluationsprozess zunehmend durchdringt – von Recherche, Datenerhebung und Datenaufbereitung über Analyse und Berichterstattung bis zu Dissemination und Nutzungsförderung. Die Tagung behandelte zugleich die Evaluation von GenKI-Systemen selbst, algorithmische Verwaltungspraktiken als Evaluationsgegenstand, Bias- und Fairnessfragen, digitale Souveränität sowie neue Anforderungen an Auftraggebende, Evaluierende und Fachgesellschaften. Diese Perspektiven werden dort einbezogen, wo sie für Governance, Qualität und Vertrauen zentral sind.

Die SEVAL-Standards bestimmen Evaluationsqualität anhand von Nützlichkeit, Durchführbarkeit, Korrektheit und Genauigkeit; in den DeGEval-Standards tritt Fairness ausdrücklich hinzu. Unter den Bedingungen generativer KI lassen sich diese Qualitätsdimensionen nicht allein durch die Wahl geeigneter Werkzeuge sichern. Erforderlich sind ebenso klare Verantwortlichkeiten, belastbare Daten- und Kontrollprozesse sowie eine angemessene Governance. Denn generative KI arbeitet probabilistisch, ihre Ergebnisse sind nicht immer vollständig nachvollziehbar und ihre Wirkung entsteht erst im Zusammenspiel mit Daten, Rollen, Abläufen und Entscheidungsstrukturen. Ihr Einsatz ist deshalb stets auch eine organisationale und professionsethische Entscheidung.

Der Impulsbeitrag verbindet zentrale Ergebnisse der Frühjahrstagung mit dem Forschungsstand und leitet daraus mögliche Handlungsfelder für die SEVAL ab. Im Mittelpunkt steht die Frage, wie Evaluationsqualität, Fairness, Transparenz und Vertrauen gesichert werden können, wenn KI in Organisationen und Evaluationsprozessen zunehmend mitwirkt.

Die Ausgangslage: KI verändert die Organisation schneller, als Strukturen entstehen

In vielen Organisationen wird KI bereits eingesetzt, ohne dass abschliessend geklärt ist, nach welchen Regeln sie genutzt und gesteuert werden soll. Die Keynote von Ulrich Lenz, die Beiträge von Dorian Wachsmann und Markus Tumeltshammer sowie der Werkstattbericht des DEval beleuchteten diese Entwicklung aus transformationsbezogener, verwaltungswissenschaftlicher und praktischer Perspektive. Trotz unterschiedlicher Zugänge führten sie zu demselben Befund: KI lässt sich nicht «nebenbei» einführen, sondern verlangt eine bewusste organisationale Gestaltung.

KI ist keine weitere Technologie – sie verändert die Organisationslogik

Die zentrale These der Keynote von Ulrich Lenz¹ lautete: KI ist weder ein blosses IT-Projekt noch ein Beschaffungsvorhaben, sondern eine Organisations- und Transformationsaufgabe der strategischen Führung. Sie verlangt Menschen, die zwischen Fach- und Technikwelt übersetzen, Strukturen, die sich weiterentwickeln können, und Leitlinien, die Orientierung geben.

Der Grund liegt weniger in der Leistungsfähigkeit einzelner Anwendungen als in der Funktionsweise der Technologie. Während Organisationen traditionell auf abgegrenzten Zuständigkeiten, Hierarchien und Prozessen beruhen, verbindet generative KI Daten bereichsübergreifend, wirkt auf Entscheidungsprozesse ein und schafft neue Formen von Verantwortung und Einfluss.

Lenz begründete diese These mit sechs miteinander verbundenen Argumentationslinien:

1. KI greift tief in die Organisation ein: Generative KI geht weit über eine neue Softwareanwendung hinaus. Sie verändert die Informationsverarbeitung, wirkt auf Entscheidungen ein und stellt Teile der Aufbau- und Ablauforganisation infrage.
2. Durch generative KI entstehen neue Macht- und Verantwortungsstrukturen: Einfluss gewinnt, wer die Werkzeuge beherrscht, über die Dateninfrastruktur verfügt und definiert, wofür KI eingesetzt wird. Das gilt innerhalb von Organisationen ebenso wie zwischen ihnen.
3. Generative KI wird zu einem neuen Akteur in der Organisation: Sie ist nicht nur ein Werkzeug, sondern erzeugt eigenständig Vorschläge und wird in Entscheidungs- und Kommunikationsprozesse eingebunden. Gerade deshalb müssen menschliche Verantwortung und Letztentscheidung eindeutig geregelt bleiben.
4. Fehlhaltungen erzeugen systemische Risiken: Sowohl das Heilsversprechen, Effizienz- und Qualitätsgewinne als selbstverständlich anzunehmen, als auch die pauschale Ablehnung von KI umgehen die notwendige differenzierte Prüfung. Entscheidend ist, für welche Aufgaben KI einen belastbaren Mehrwert schafft, wo ihre Grenzen liegen und wie Fehlentscheidungen, Anbieterabhängigkeiten und Verantwortungsdiffusion begrenzt werden können. Lenz plädierte für Human-Centered AI, die menschliches Denken unterstützt, aber nicht ersetzt.
5. Regulierung allein reicht nicht aus: Den EU AI Act wertete Lenz als wichtigen Schritt für Rechtssicherheit und Risikobegrenzung. Weil technologische Entwicklungen schneller verlaufen als staatliche Regulierung, braucht es ergänzend technische Schutzmechanismen, organisationsinterne Governance, freiwillige Selbstverpflichtungen und systematisches KI-Kompetenzmanagement. Das systemische Risiko betrifft nicht nur einzelne Fehlklassifikationen, sondern auch Vertrauen, Methodenintegrität, Ressourcenverteilung und Verantwortung im gesamten Evaluationsfeld.
6. KI-Kompetenz wird zur strategischen Kernaufgabe: Verantwortungsvolle Nutzung setzt Kompetenzen voraus, die über die Bedienung einzelner Anwendungen hinausreichen und methodische, rechtliche, ethische, organisationale und technische Aspekte verbinden. Ihr Aufbau ist eine Führungsaufgabe.

¹ Ulrich Lenz lehrt an der Hochschule für angewandtes Management zu KI-Ethik und Digital Transformation und begleitet als Transformationsberater Führungsteams bis zur Vorstandsebene.

Es genügt daher nicht, KI an bestehende Strukturen anzudocken. Die Strukturen müssen sich mit den Lern- und Veränderungsprozessen rund um KI weiterentwickeln; daraus folgt die Notwendigkeit eines angepassten Organisationsdesigns.

Zwischen dem ersten produktiven Einsatz eines Sprachmodells und einer Organisation, die KI strategisch, sicher und qualitätsgesichert einsetzt, liegt ein längerer Transformationsweg. Wie er praktisch begonnen und gesteuert werden kann, zeigte der Werkstattbericht des DEval.

Organisationale Lern- und Steuerungsfähigkeit entscheidet über Erfolg

Während Ulrich Lenz die organisationale Verankerung von KI aus der Perspektive des Transformationsmanagements beschrieb, zeigten Kai Rompczyk und Laura Kunert² am Deutschen Evaluierungsinstitut für Entwicklungszusammenarbeit (DEval), wie sich dieser Transformationsprozess im Organisationsalltag beginnen lässt.

Ihre Kernbotschaft lautet: Zwischen dem ersten produktiven Einsatz eines KI-Systems und einer Organisation, die KI strategisch und qualitätsgesichert einsetzt, liegt ein längerfristiger organisationaler Entwicklungsprozess. Über seinen Erfolg entscheidet die Lern- und Steuerungsfähigkeit der Organisation.

Der Werkstattbericht bestätigte damit die theoretische Kernthese der Keynote: KI scheitert selten an der Technologie selbst, häufiger jedoch an der Art und Weise, wie Organisationen sie einführen und sich dabei weiterentwickeln. Vier Befunde sind besonders relevant:

1. **KI-Kompetenz entsteht durch Erfahrung und Übersetzung zwischen Fach- und Technikwelt:** Beim DEval erfolgte der Kompetenzaufbau nicht als Resultat einzelner Schulungen, sondern als kontinuierlicher Lernprozess. Am Beispiel einer Textklassifikation zur Portfolioidentifikation wurden Projektbeschreibungen aus CRS-Daten der OECD-DAC für die Jahre 2020 bis 2022 binär klassifiziert und mit Begründungen versehen. Entscheidend war nicht allein die Trefferquote, sondern das Zusammenspiel von Ergebnisqualität, Nachvollziehbarkeit, Halluzinationen, menschlichem Kontrollaufwand und Kontextwissen. Erst die enge Zusammenarbeit mit Data Scientists machte sichtbar, wie wichtig eine gemeinsame Sprache zwischen Evaluationsmethodik und Informatik ist.
2. **KI erzwingt organisationale, nicht nur technische Entscheidungen:** Das DEval untersagte frühzeitig die unkontrollierte Nutzung von ChatGPT aus Datenschutzgründen und entwickelte beziehungsweise prüfte kontrollierte Lösungen. Damit verlagerte sich die Herausforderung von der Werkzeugauswahl auf Verantwortlichkeiten, Governance, Ressourcen und tragfähige Infrastrukturen. Ein strukturiertes Verfahren ermöglichte es Teams, konkrete Anwendungsfälle in begrenzten Kontexten zu testen. Die Pilotprojekte dienten dem organisationsweiten Erfahrungsaufbau, machten Bedarfe – etwa für Transkription, Anonymisierung, Wissensmanagement oder die Extraktion von Ursache-Wirkungs-Beziehungen – sichtbar und zeigten zugleich Grenzen auf, wenn Datenqualität, Kontextwissen oder Wirkungsberichte für automatisierte Analysen nicht ausreichten.
3. **Datensouveränität ist eine strategische Schlüsselfrage:** Das DEval prüfte unter anderem kontrollierte Chat-Lösungen, europäische Anbieter und EU-konforme Konfigurationen. Dabei zeigte sich, dass Konformität weit über den Serverstandort hinausgeht. Entscheidend sind Datenflüsse, Sicherheitszertifizierungen, Speicherorte, Zugriffskontrollen, Modelltraining und Nachvollziehbarkeit. Sichere und souveräne Infrastrukturen erfordern erhebliche personelle und finanzielle Ressourcen; dadurch beeinflusst die organisationale KI-Verankerung auch die Chancengleichheit im Evaluationsmarkt.
4. **Der Wert der Evaluation verschiebt sich auf Leistungen, die sich nicht automatisieren lassen:** Die Werkstattberichte zu Freitextanalysen, MAXQDA AI Assist und dialogbasierter qualitativer Analyse zeigten Effizienzpotenziale, aber ebenso zusätzlichen Kontroll-, Dokumentations- und Validierungsbedarf. Je mehr standardisierbare Schritte unterstützt werden können, desto wichtiger werden hochwertige

² Kai Rompczyk ist seit 2021 Evaluator mit Schwerpunkt auf quantitativen Methoden am DEval (Deutsches Evaluierungsinstitut für Entwicklungszusammenarbeit) und befasst sich mit den Folgen Künstlicher Intelligenz. Laura Kunert ist seit 2019 Evaluatorin am DEval und arbeitet nach ihrem Master in Development Studies an der Universität Passau.

Interviews, die Moderation komplexer Prozesse, die Kontextualisierung von Befunden, das Verständnis organisationaler Machtstrukturen und die transparente Dokumentation des KI-Einsatzes.

Der erfolgreiche KI-Einsatz ist damit eine organisationale Entwicklungsaufgabe, die dauerhaftes Lernen, Steuerung und Qualitätssicherung erfordert.

Die Steuerungsinstrumente des DEval: Reifegradmodell und Integrationsprozessmodell

Das DEval hat seine praktischen Erfahrungen in zwei Steuerungsinstrumente überführt, die den organisationalen Wandel beobachtbar und steuerbar machen. Das Reifegradmodell dient der Standortbestimmung; das Integrationsprozessmodell beschreibt den Entwicklungsweg der Organisation.

Das Reifegradmodell bündelt acht Dimensionen in den drei Bereichen organisationale Rahmenbedingungen, Qualität und Effizienz sowie Praxis. Jede Dimension wird entlang der Reifegrade niedrig, mittel und hoch beschrieben und ermöglicht eine systematische Einschätzung des aktuellen Entwicklungsstands.

Grafik: Reifegradmodell des DEval mit drei Bereichen, acht Dimensionen und drei Reifegraden.

Dimension	Niedriger Reifegrad	Mittlerer Reifegrad	Hoher Reifegrad
ORGANISATIONALE RAHMENBEDINGUNGEN			
1.  Strategische Steuerung	Unkoordiniert Kein Zielbild, Verantwortlichkeiten sind nicht geregelt.	Strukturiert Das Thema KI ist in strategischen Leitlinien verankert, wird aber dezentral bearbeitet.	Institutionalisiert KI ist als eigenständiges Handlungsfeld verankert und wird zentral gesteuert.
2.  Institutionelle Ressourcen	Nebenläufig KI-Nutzung erfolgt ohne eingeplante Ressourcen und beruht auf Einzelinitiativen.	Eingeplant Ressourcen werden in der Planung von KI-Vorhaben berücksichtigt.	Dauerhaft ausgestattet Für die Integration von KI sind alle relevanten Arbeitsbereiche dauerhaft ausgestattet.
3.  Daten- und IT-Infrastruktur	Dezentral, unkontrolliert Unkontrollierte Nutzung von KI-Tools ohne gesteuerte KI-Infrastruktur. Externe Tool-Nutzung.	Kontrolliert, aber beschränkt Sichere IT-Infrastruktur ist vorhanden, aber keine Verarbeitung nichtöffentlicher Daten und eingeschränkte Performanz.	Sicher und souverän KI-Infrastruktur ist hochperformant, sicher und souverän.
QUALITÄT UND EFFIZIENZ			
1.  Qualität der Anwendung	Nicht ausreichend Ergebnisqualität ist nicht ausreichend oder schwankt stark.	Teilweise gesichert Die Ergebnisqualität ist in Teilbereichen ausreichend.	Verlässlich KI-Anwendungen liefern verlässlich hohe Ergebnisqualität.
2.  Effizienz der Anwendung	Anekdotisch Zeitersparnis ist zufällig, individuell.	Nachweisbar Effizienzziele sind definiert und werden teilweise erreicht.	Systematisch wirksam Effizienzgewinne werden organisationsweit realisiert.
3.  Reichweite der Anwendung	Punktuell KI wird für einzelne Aufgaben genutzt.	Ausgedehnt Mehrere Anwendungsgebiete sind erschlossen. Tätigkeiten sind nicht automatisiert.	Umfassend KI wird als generalistisches Tool in allen Tätigkeiten eingebunden und automatisiert geeignete Prozesse.
PRAXIS			
1.  Kompetenzen und Kultur	Individuell KI-Kompetenz liegt bei einzelnen interessierten Personen; Nutzung und Akzeptanz sind uneinheitlich.	Im Aufbau Mitarbeitende nutzen Lernformate zum Aufbau von KI-Kompetenzen.	Professionell verankert KI-Kompetenzen sind bei allen Beschäftigten vorhanden und fester Bestandteil der Arbeitskultur.
2.  Wissenschaftliche Integrität	Unreflektiert KI-Nutzung erfolgt ohne Berücksichtigung von geregelten Standards.	Geregelt Standards zur Nutzung von KI sind definiert.	Qualitätsgesichert Standardeinhaltung bei der Nutzung von KI wird systematisch dokumentiert und evaluiert.

(Quelle: Reifegradmodell des DEval)

Ergänzend beschreibt das Integrationsprozessmodell den Weg von ersten Einzelinitiativen bis zur organisationsweiten Verankerung von KI. Es unterscheidet vier aufeinander aufbauende Entwicklungsphasen: Exploration, geregelte Testphase, Skalierung und Institutionalisierung. KI-Implementierung wird damit nicht als einmalige Einführung, sondern als kontinuierlicher organisationaler Lern- und Veränderungsprozess verstanden.

Grafik: Integrationsprozessmodell des DEval.



(Quelle: Integrationsprozessmodell des DEval)

Bedeutung der KI-Transformation für die Evaluation

Eine zentrale Erkenntnis der Frühjahrstagung betrifft bereits den Gegenstand selbst: Der Sammelbegriff «Künstliche Intelligenz» ist zu unscharf, um daraus unmittelbar Regeln, Verantwortlichkeiten oder Qualitätsanforderungen abzuleiten. Er umfasst Systeme mit sehr unterschiedlichen Fähigkeiten, Risiken und Einsatzmöglichkeiten. Klassische Taxonomien reichen angesichts generativer Foundation Models nur begrenzt aus (Bommasani et al., 2021; Feuerriegel et al., 2024; OECD, 2022). Hinzu kommt die sogenannte jagged frontier: Generative KI kann bei ähnlichen Aufgaben sehr unterschiedliche Ergebnisse liefern (Dell’Acqua et al., 2023). Der Einsatz ist deshalb auf der Ebene konkreter Aufgaben zu beurteilen (Huang & Rust, 2018).

Für die organisationale Steuerung sind **zwei Differenzierungen** besonders hilfreich. **Technologisch** ist zu unterscheiden, welche Daten ein System benötigt, wie transparent und fehleranfällig es ist und welchem Regulierungsbedarf es unterliegt. **Funktional** ist entscheidend, ob KI eine klar begrenzte Teilaufgabe innerhalb eines menschlich gesteuerten Workflows übernimmt, Entscheidungen vorbereitet oder Probleme weitgehend eigenständig bearbeitet. Je weiter der Handlungsspielraum des Systems reicht, desto höher sind die Anforderungen an Kontrolle, Dokumentation und Rechenschaft.

Algorithmische Praktiken sind dabei nicht nur Werkzeuge der Evaluation, sondern selbst ein Evaluationsgegenstand. Daten und ihre algorithmische Verarbeitung sind in technische, organisationale, rechtliche, politische und kulturelle Strukturen eingebettet. Evaluationen müssen deshalb auch fragen, welche Akteure an Planung, Implementierung und Alltagspraxis beteiligt sind, wie Werte in Prozesse übersetzt werden, welche nicht beabsichtigten Wirkungen entstehen und wie sich Handlungsmacht, Zuständigkeiten und Machtverhältnisse verschieben.

Der Forschungsstand und die Tagungsbeiträge verweisen auf wiederkehrende Grenzen und Umsetzungsbarrieren: fehlende Übersetzungskompetenz zwischen Fach- und Datenexpertise, unzureichende Datensouveränität, begrenzte Nachvollziehbarkeit und Eigenkontrolle vieler Modelle sowie Halluzinationen, Bias, Black-Box-Probleme und Modellkollaps (Rudin, 2019; Feuerriegel et al., 2024; Raisch & Krakowski, 2021; Shumailov et al., 2024). Gerade bei qualitativen Analysen müssen Ergebnisse geprüft, Prompts dokumentiert, Zitate kontrolliert und mögliche Verzerrungen reflektiert werden.

Als General-Purpose-Technologie entfaltet KI ihren Nutzen häufig erst nach Investitionen in Prozesse, Kompetenzen und Rollen (Brynjolfsson et al., 2021). Gleichzeitig neigen Menschen dazu, KI entweder zu überschätzen und ihr zu viel Vertrauen entgegenzubringen oder sie nach einzelnen Fehlern vorschnell abzulehnen (Dell’Acqua et al., 2023; Dietvorst et al., 2015). Mit zunehmender Automatisierung steigen zudem die Anforderungen an die verbleibenden menschlichen Entscheidungen – die von Bainbridge (1983) beschriebene «Ironie der Automatisierung».

Für Evaluationen folgt daraus ein doppelter Auftrag: Der KI-Einsatz im eigenen Prozess muss methodisch und organisatorisch kontrolliert werden. Zugleich müssen KI-Systeme und algorithmische Praktiken als mögliche Evaluationsgegenstände verstanden werden. In beiden Fällen gewinnen Transparenz, menschliche Aufsicht, Rechenschaftspflicht und die nachvollziehbare Zuordnung von Verantwortung an Bedeutung (Potluka et al., 2025; Rompczyk, 2025).

Organisationale Voraussetzungen für einen verantwortbaren KI-Einsatz

Aus Forschung und Tagungsbeiträgen lassen sich die notwendigen Elemente einer verantwortbaren KI-Implementierung kompakt zusammenfassen: **Erfolgreiche KI-Integration beginnt nicht mit dem Werkzeug, sondern mit einer realistischen Einschätzung von Reifegrad und Bereitschaft der Organisation.** Dazu gehören Führung und Strategie, Ressourcen, Kompetenzen, Kultur sowie eine tragfähige Datenbasis (Dreyling et al., 2024; Jöhnk et al., 2021; Neumann et al., 2024).

Anschließend sind konkrete Anwendungsfälle nach ihrem erwarteten Mehrwert, ihren Risiken und ihrer Vereinbarkeit mit Auftrag und Werten der Organisation zu priorisieren. Kleine, klar begrenzte Pilotprojekte schaffen Erfahrungswissen und liefern eine Grundlage für spätere Investitions- und Skalierungsentscheidungen; sie dienen vor allem dem organisationalen Lernen (Fountaine et al., 2019; Sánchez et al., 2025).

Kompetenzentwicklung, Organisationskultur und Übersetzerrollen zwischen Fach- und Datenexpertise müssen gemeinsam entwickelt werden. Professionelles Change Management greift Unsicherheiten früh auf und bindet Mitarbeitende in die Veränderung ein (Kurup & Gupta, 2022; Murire, 2024).

Governance ist von Beginn an mitzudenken. Dazu gehören geregelte Datenverantwortlichkeiten, transparente Entscheidungsprozesse, menschliche Aufsicht, Dokumentationspflichten und ein interdisziplinäres Zusammenwirken von Fachbereichen, IT, Datenschutz, Recht und Leitung. Diese ermöglichenden Fähigkeiten sind Voraussetzung für vertrauenswürdige Anwendungen (Gama & Magistretti, 2025).

Erst wenn Pilotprojekte ausgewertet, Kompetenzen aufgebaut und Governance-Strukturen etabliert sind, sollte skaliert werden. Mit wachsendem Reifegrad verschiebt sich der Schwerpunkt von Einzelanwendungen zur kontinuierlichen Steuerung des KI-Portfolios, zum Wissensaustausch und zum organisationsweiten Lernen (Dreyling et al., 2024; Neumann et al., 2024).

Diese Voraussetzungen bilden die Brücke zu den Handlungsfeldern der SEVAL: Eine Fachgesellschaft kann die organisationale Verantwortung ihrer Mitglieder nicht übernehmen, sie kann aber gemeinsame Qualitätsanforderungen präzisieren und eine Infrastruktur bereitstellen, die verantwortbares Handeln erleichtert.

2. Mögliche Handlungsfelder für die SEVAL

Von den Tagungsbefunden zu möglichen Handlungsfeldern der SEVAL

Der Rückblick auf die Frühjahrstagung zeigt, dass generative Künstliche Intelligenz die Evaluation auf mehreren Ebenen verändert: als Werkzeug im Evaluationsprozess, als Gegenstand von Evaluationen und als Faktor, der Rollen, Verantwortlichkeiten und Organisationsstrukturen verschiebt. Damit geht es nicht nur um die Frage, welche Anwendungen eingesetzt werden können, sondern vor allem darum, unter welchen Voraussetzungen ihr Einsatz mit den Qualitätsansprüchen professioneller Evaluation vereinbar ist.

Die SEVAL-Standards bleiben dafür der zentrale Referenzrahmen. Nützlichkeit, Durchführbarkeit, Korrektheit und Genauigkeit verlieren durch KI nicht an Bedeutung; vielmehr müssen sie für neue Risikosituationen konkretisiert werden. Wo Datenflüsse, Modellverhalten oder Bearbeitungsschritte nur begrenzt nachvollziehbar

sind, gewinnen Transparenz, Dokumentation, menschliche Aufsicht, Sensibilität gegenüber Verzerrungen, Datensouveränität und eine eindeutig zugeordnete Letztverantwortung zusätzlich an Gewicht.

Diese Anforderungen lassen sich nicht allein durch die Wahl eines vermeintlich sicheren Werkzeugs erfüllen. Sie setzen organisationale Lern- und Steuerungsfähigkeit voraus: klare Regeln, definierte Verantwortlichkeiten, ausreichende Kompetenzen, kontrollierte Pilotierungen sowie Verfahren zur Prüfung und Validierung von KI-Ergebnissen. Zugleich unterscheiden sich die Handlungsmöglichkeiten der Akteure erheblich. Grössere Institutionen können eigene Fachstellen und technische Infrastrukturen aufbauen. Einzelpersonen und kleine Evaluationsbüros verfügen meist nicht über vergleichbare Ressourcen.

Daraus entsteht eine fachgesellschaftliche, nicht jedoch operative Verantwortung der SEVAL. Sie kann und soll die Verantwortung der einzelnen Evaluierenden, Auftraggebenden und Organisationen nicht übernehmen. Sie kann aber einen gemeinsamen Orientierungsrahmen schaffen, evaluationsspezifische Anforderungen in die berufliche Praxis übersetzen, gute Praxis bündeln und Unterstützungsangebote entwickeln, die auch für kleinere Akteure zugänglich sind.

Die nachfolgenden Handlungsfelder sind daher nicht als abschliessendes Programm, sondern als Anregungen für eine schrittweise fachliche Diskussion zu verstehen. Sie betreffen die Klärung und Operationalisierung von Qualitätsanforderungen, den Aufbau einer praxistauglichen Infrastruktur, die Stärkung professioneller Kompetenz und Chancengleichheit sowie die Rolle und mögliche Kooperationen der SEVAL im KI-Transformationsprozess.

Die KI-Transformation betrifft nicht nur einzelne Methoden, Werkzeuge oder Arbeitsschritte. Sie verändert die Bedingungen, unter denen Evaluationen geplant, durchgeführt, dokumentiert, beurteilt und genutzt werden. Daraus ergibt sich für die SEVAL eine berufspolitische Aufgabe: dazu beizutragen, dass Evaluation auch unter den Bedingungen generativer KI nützlich, durchführbar, korrekt, genau, fair, nachvollziehbar und vertrauenswürdig bleibt. Die Statuten schaffen hierfür eine formale Grundlage, indem sie den Informations- und Erfahrungsaustausch zwischen Politik, Verwaltung, Hochschulen und Beratung sowie die Qualität und Verbreitung von Evaluation fördern.

Vertrauenswürdigkeit als Qualitätsdimension für KI-unterstützte Evaluationen klären

Die Frühjahrstagung zeigte, dass KI nicht nur Evaluationsprozesse verändert, sondern auch die Voraussetzungen ihrer Glaubwürdigkeit. Im Wirkmodell von Wagner und Palfinger bildet Vertrauen eine eigene Systemdimension. Jan Hense bezeichnete Vertrauen als die «wichtigste Währung» im Umgang mit KI. Die Beiträge zu Bias, algorithmischen Praktiken und digitaler Souveränität machten zugleich deutlich, dass Vertrauen organisatorisch, methodisch und technisch hergestellt werden muss (Palfinger et al., 2025).

Die Qualitätsdimensionen der SEVAL und DeGEval bleiben das Fundament professioneller Evaluation. Fachlich zu klären ist jedoch, ob Vertrauenswürdigkeit als eigenständige Qualitätsdimension oder als übergreifendes Prinzip zu verstehen ist. In jedem Fall müssen Transparenz, Nachvollziehbarkeit, Erklärbarkeit, Sensibilität gegenüber Verzerrungen, Datenverantwortung, Dokumentation des KI-Einsatzes, menschliche Aufsicht und Letztverantwortung operationalisiert werden.

Folgendes Vorgehen bietet sich an: Zunächst sollte eine Handreichung die bestehenden Standards auf typische KI-bezogene Risikosituationen beziehen und darlegen, welche Vorkehrungen die angestrebte Evaluationsqualität sichern. Sollte sich zeigen, dass zentrale Anforderungen mit den geltenden Standards nicht ausreichend erfasst werden können, wäre eine Revision der Standards zu prüfen.

Qualitätsinfrastruktur und professionelle Kompetenz bereitstellen

Die SEVAL könnte ihren Mitgliedern praxistaugliche Instrumente anbieten: Muster für interne KI-Richtlinien, Checklisten zur Beurteilung von KI-Tools, Vorlagen für Datenschutz- und Vertraulichkeitsprüfungen, Entscheidungshilfen zur Zulässigkeit bestimmter Anwendungen, Musterklauseln für Offerten und Verträge,

Dokumentationsvorlagen sowie einfache Reifegrad- und Selbstprüfungsinstrumente. Sie sollten ausdrücklich auch für Einzelpersonen und kleine Evaluationsbüros mit begrenzten Ressourcen nutzbar sein.

Ein zweiter Baustein ist die Aus- und Weiterbildung. KI-Kompetenz wird zu einer professionellen Grundkompetenz der Evaluation. Schulungen sollten deshalb nicht nur Anwendungen vorstellen, sondern Urteilsfähigkeit stärken: Welche Aufgaben eignen sich für KI-Unterstützung? Welche Daten dürfen in welche Systeme eingegeben werden? Wie werden Ergebnisse validiert? Wann entstehen methodische, rechtliche oder ethische Risiken? Wie verändern sich die Rollen von Auftraggebenden, Evaluierenden und evaluierten Organisationen?

Weiterbildung allein genügt jedoch nicht. Datenschutz, Informationssicherheit, KI-Governance, Dokumentationspflichten, Auftragsbearbeitung, Cloud-Nutzung, Haftungsfragen und interne Weisungen von Auftraggebenden können kleine Anbieter rasch überfordern. Ein Fachverband erfüllt deshalb auch eine infrastrukturelle Funktion: Er bündelt Aufgaben und Wissen, die einzelne Mitglieder allein kaum effizient und fachlich sicher bewältigen können.

Chancengleichheit im Evaluationsmarkt unterstützen

Wenn nur grosse Institutionen über die Mittel verfügen, KI-Governance, Datenschutz und technische Infrastruktur professionell aufzubauen, entsteht eine strukturelle Ungleichheit. Kleine Anbieter können fachlich ausgezeichnet arbeiten und dennoch an regulatorischen und technischen Anforderungen scheitern. Die SEVAL kann ihnen die Verantwortung nicht abnehmen, aber die Voraussetzungen verbessern, damit sie diese realistisch wahrnehmen können.

Die SEVAL kann folgendes anbieten: Verbandsseitige Arbeitshilfen, Schulungen und gegebenenfalls vergünstigte Beratungsangebote sollten so konzipiert werden, dass sie nicht nur Organisationen mit eigenen Stabsstellen, sondern auch kleine Evaluationsanbieter erreichen.

Rolle der SEVAL im KI-Transformationsprozess definieren

Die SEVAL kann Orientierung geben, Mindestanforderungen klären, gute Praxis sichtbar machen, Vorlagen bereitstellen und Erfahrungswissen bündeln. Sie sollte jedoch nicht den Eindruck erwecken, individuelle Rechtsprüfung, Datenschutzberatung oder IT-Sicherheitsprüfung ersetzen zu können. Die Letztverantwortung bleibt bei den Vertragspartnern des jeweiligen Evaluationsvorhabens.

Naheliegend ist ein gestuftes Unterstützungsmodell:

- frei zugängliche Grundinformationen und Merkblätter,
- vertiefte Praxispakete für Mitglieder sowie
- bei Bedarf Beratungsformate mit qualifizierten Partnerinnen und Partnern, deren Vergütung beispielsweise an das Preismodell der Beratungsstelle anknüpfen könnte.

So kann die SEVAL niederschwellig unterstützen, ohne ihre Verantwortung zu überdehnen oder in Konkurrenz zu Mitgliedern zu treten, die selbst Beratungsleistungen anbieten.

Kooperationen und realistische Umsetzung

Eine solche Qualitätsinfrastruktur muss die SEVAL nicht allein entwickeln. Weil KI-Implementierung methodische, technische, rechtliche, ethische und organisationale Fragen verbindet, bieten sich Kooperationen mit der DeGEval sowie mit Fachgesellschaften und Netzwerken aus Organisationsentwicklung, Digitalisierung, Datenschutz, öffentlicher Verwaltung und KI-Governance an. Ziel ist nicht, allgemeine Leitfäden zu duplizieren, sondern vorhandenes Wissen in die spezifische Praxis der Evaluation zu übersetzen.

Mögliche Kooperationsformen sind gemeinsame Arbeitsgruppen, D-A-CH-Austauschformate, Peer Reviews, Schulungsreihen und gemeinsam entwickelte, für die Schweiz adaptierte Materialien. Die SEVAL sollte dabei zunächst die evaluationsspezifischen Fragen priorisieren: Transparenz des KI-Einsatzes, Validierung KI-generierter Analysen, Schutz vertraulicher Daten, Dokumentation methodischer Entscheidungen und Rollenklärung zwischen Auftraggebenden und Evaluierenden.

Realistisch zu berücksichtigen ist, dass die Arbeit der SEVAL weitgehend auf freiwilligem Engagement beruht und evaluative Expertise allein die technischen, rechtlichen und sicherheitsbezogenen Fragen nicht abdeckt. Tragfähig wäre ein Modell, in dem eine kleine SEVAL-Arbeitsgruppe den Bedarf definiert, Prioritäten setzt und die evaluative Anschlussfähigkeit sichert, während externe Fachpersonen punktuell für KI-Technologie, Datenschutz, Informationssicherheit, Recht, Organisationsentwicklung und die Umsetzung in Kleinstorganisationen beigezogen werden.

Ein schlankes Umsetzungsmodell könnte freiwillige fachliche Steuerung durch die SEVAL, Kooperation mit verwandten Fachgesellschaften, den punktuellen Einkauf externer Expertise und eine Finanzierung über Kurse, Praxispakete, Projektbeiträge, Beratungen oder Partnerschaften verbinden.

Damit würde die SEVAL ihre eigentliche Stärke nutzen: evaluative Erfahrung bündeln, komplexe Anforderungen in berufliche Praxis übersetzen und gemeinsame Qualitätsmassstäbe sichern. Sie würde zu einer Wissens-, Compliance- und Professionalisierungsinfrastruktur für ihre Mitglieder, ohne Aufgaben zu übernehmen, die ausserhalb ihres fachlichen Kerns liegen.

Fazit

Vertrauenswürdigkeit entsteht nicht allein in einzelnen Evaluationen; sie benötigt auch eine professionelle Infrastruktur. Wenn KI Evaluationen verändert, müssen sich Fachgesellschaften weiterentwickeln. Die SEVAL kann als Ort des Austauschs, Hüterin und Weiterentwicklerin von Qualitätsanforderungen, Weiterbildungsplattform und praktische Unterstützungsstruktur eine wichtige Rolle übernehmen. Entscheidend ist, Qualität nicht nur zu postulieren, sondern ihre Umsetzung unter neuen technischen, organisatorischen und regulatorischen Bedingungen konkret zu ermöglichen und Kontrollmöglichkeiten zu schaffen.

Literaturverzeichnis

- Bainbridge, L. (1983). Ironies of automation. *Automatica*, 19(6), 775–779. [https://doi.org/10.1016/0005-1098\(83\)90046-8](https://doi.org/10.1016/0005-1098(83)90046-8)
- Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M. S., Bohg, J., Bosselut, A., Brunskill, E., Brynjolfsson, E., Buch, S., Card, D., Castellon, R., Chatterji, N. S., Chen, A. S., Creel, K., Davis, J. Q., Demszky, D., ... Liang, P. (2021). *On the opportunities and risks of foundation models*. arXiv. <https://doi.org/10.48550/arXiv.2108.07258>
- Brynjolfsson, E., Rock, D., & Syverson, C. (2021). The productivity J-curve: How intangibles complement general purpose technologies. *American Economic Journal: Macroeconomics*, 13(1), 333–372. <https://doi.org/10.1257/mac.20180386>
- DeGEval – Gesellschaft für Evaluation e. V. (2026). Frühjahrstagung 2026 „KI in der Evaluation“, 21./22. Mai 2026, Wien: Aktuelles und Tagungsdokumentation. <https://www.degeval.org/arbeitskreise/methoden-in-der-evaluation/aktuelles/>
- Dell’Acqua, F., McFowland, E., III, Mollick, E. R., Lifshitz-Assaf, H., Kellogg, K., Rajendran, S., Kraymer, L., Candelon, F., & Lakhani, K. R. (2023). *Navigating the jagged technological frontier: Field experimental evidence of the effects of artificial intelligence on knowledge worker productivity and quality* (Harvard Business School Working Paper No. 24-013). <https://doi.org/10.2139/ssrn.4573321>
- Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, 144(1), 114–126. <https://doi.org/10.1037/xge0000033>
- Dreyling, R., Lemmik, J., Tammet, T., & Pappel, I. (2024). An artificial intelligence maturity model for the public sector: A design science approach. *TalTech Journal of European Studies*, 14(2), 219–243. <https://doi.org/10.2478/bjes-2024-0023>
- European Parliament & Council of the European Union. (2024). *Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*. Official Journal of the European Union, L 2024/1689. <http://data.europa.eu/eli/reg/2024/1689/oj>
- Feuerriegel, S., Hartmann, J., Janiesch, C., & Zschech, P. (2024). Generative AI. *Business & Information Systems Engineering*, 66(1), 111–126. <https://doi.org/10.1007/s12599-023-00834-7>
- Fountain, T., McCarthy, B., & Saleh, T. (2019). Building the AI-powered organization: Technology isn’t the biggest challenge, culture is. *Harvard Business Review*, 97(4), 62–73.
- Gama, F., & Magistretti, S. (2025). Artificial intelligence in innovation management: A review of innovation capabilities and a taxonomy of AI applications. *Journal of Product Innovation Management*, 42(1), 76–111. <https://doi.org/10.1111/jpim.12698>
- Haslinger-Baumann, E. (2026, Mai 21). Grußwort [Präsentation]. Gemeinsame Frühjahrstagung der DeGEval-Arbeitskreise „Methoden in der Evaluation“ und „Verwaltung“, Hochschule Campus Wien, Wien, Österreich.

- Hense, J. (2026, Mai 22). *Wrap-up und Abschlussdiskussion zur Tagung [Präsentation]. Gemeinsame Frühjahrstagung der DeGEval-Arbeitskreise „Methoden in der Evaluation“ und „Verwaltung“*, Hochschule Campus Wien, Wien, Österreich.
- Huang, M.-H., & Rust, R. T. (2018). Artificial intelligence in service. *Journal of Service Research*, 21(2), 155–172. <https://doi.org/10.1177/1094670517752459>
- Jöhnk, J., Weißert, M., & Wyrski, K. (2021). Ready or not, AI comes: An interview study of organizational AI readiness factors. *Business & Information Systems Engineering*, 63(1), 5–20. <https://doi.org/10.1007/s12599-020-00676-7>
- Kopczynski, S. (2026, Mai 22). *Einsatz von KI-Unterstützung in der Hochschulevaluation – Erfahrungsbericht über software-automatisierte Inhaltsanalysen von Freitextkommentaren in Studierendenbefragungen [Werkstattbericht]*. Gemeinsame Frühjahrstagung der DeGEval-Arbeitskreise „Methoden in der Evaluation“ und „Verwaltung“, Hochschule Campus Wien, Wien, Österreich.
- Krauter, T. (2026, Mai 22). *KI zur Unterstützung qualitativer Analyse in Evaluationen [Werkstattbericht]*. Gemeinsame Frühjahrstagung der DeGEval-Arbeitskreise „Methoden in der Evaluation“ und „Verwaltung“, Hochschule Campus Wien, Wien, Österreich.
- Kunert, L., & Rompczyk, K. (2026, Mai 21–22). *KI-Integration im Spannungsfeld zwischen Innovation und Regulierung [Werkstattbericht]*. Gemeinsame Frühjahrstagung der DeGEval-Arbeitskreise „Methoden in der Evaluation“ und „Verwaltung“, Hochschule Campus Wien, Wien, Österreich.
- Kurup, S., & Gupta, V. (2022). Factors influencing the AI adoption in organizations. *Metamorphosis: A Journal of Management Research*, 21(2), 129–139. <https://doi.org/10.1177/09726225221124035>
- Lenz, U. (2026, Mai 21–22). *Zwischen Heilsversprechen und systemischem Risiko: Ethische Leitlinien für die verantwortliche KI-Transformation [Keynote]*. Gemeinsame Frühjahrstagung der DeGEval-Arbeitskreise „Methoden in der Evaluation“ und „Verwaltung“, Hochschule Campus Wien, Wien, Österreich.
- Murire, O. T. (2024). Artificial intelligence and its role in shaping organizational work practices and culture. *Administrative Sciences*, 14(12), Article 316. <https://doi.org/10.3390/admsci14120316>
- Neumann, O., Guirguis, K., & Steiner, R. (2024). Exploring artificial intelligence adoption in public organizations: A comparative case study. *Public Management Review*, 26(1), 114–141. <https://doi.org/10.1080/14719037.2022.2048685>
- Organisation for Economic Co-operation and Development. (2022). *OECD framework for the classification of AI systems* (OECD Digital Economy Papers No. 323). OECD Publishing. <https://doi.org/10.1787/cb6d9eca-en>
- Palfinger, T., Gaisbauer, F., Wagner, I., & Beck, S. (2025). Wie wirkt künstliche Intelligenz im Evaluationssystem? Diskussionsanstöße für die Gestaltung des Evaluationssystems von morgen. *fteval Journal for Research and Technology Policy Evaluation*, 55, 1–35. <https://doi.org/10.22163/FTEVAL.2024.644>
- Potluka, O., Harten, S., Kocks, A., & Dvorak, J. (2025). Digitalization in evaluations and evaluations of digitalization: The changing landscape of evaluations. *Evaluation*, 31(3), 289–302. <https://doi.org/10.1177/13563890251357650>

- Raisch, S., & Krakowski, S. (2021). Artificial intelligence and management: The automation–augmentation paradox. *Academy of Management Review*, 46(1), 192–210.
<https://doi.org/10.5465/amr.2018.0072>
- Rompczyk, K. (2025). Technological revolution in evaluation: Artificial intelligence and the adherence to evaluation standards. *Evaluation*, 31(3), 331–351. <https://doi.org/10.1177/13563890251331066>
- Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1(5), 206–215.
<https://doi.org/10.1038/s42256-019-0048-x>
- Sánchez, E., Calderón, R., & Herrera, F. (2025). Artificial intelligence adoption in SMEs: Survey based on TOE–DOI framework, primary methodology and challenges. *Applied Sciences*, 15(12), Article 6465.
<https://doi.org/10.3390/app15126465>
- Schad-Spindler, A., & Häge, L. (2026, Mai 22). Qualitative Inhaltsanalysen mit MAXQDA AI Assist: Mehr Effizienz, mehr Kontrollbedarf? [Werkstattbericht]. Gemeinsame Frühjahrstagung der DeGEval-Arbeitskreise „Methoden in der Evaluation“ und „Verwaltung“, Hochschule Campus Wien, Wien, Österreich.
- Shumailov, I., Shumaylov, Z., Zhao, Y., Papernot, N., Anderson, R., & Gal, Y. (2024). AI models collapse when trained on recursively generated data. *Nature*, 631(8022), 755–759.
<https://doi.org/10.1038/s41586-024-07566-y>
- Tumeltshammer, M. (2026, Mai 22). Algorithmische Praktiken der Verwaltung als offener Evaluationsgegenstand und collaborative research [Präsentation]. Gemeinsame Frühjahrstagung der DeGEval-Arbeitskreise „Methoden in der Evaluation“ und „Verwaltung“, Hochschule Campus Wien, Wien, Österreich.
- Wachsmann, D. (2026, Mai 21). KI & Evaluation: Lessons learned aus der öffentlichen Verwaltung [Präsentation]. Gemeinsame Frühjahrstagung der DeGEval-Arbeitskreise „Methoden in der Evaluation“ und „Verwaltung“, Hochschule Campus Wien, Wien, Österreich.

Hinweis zur Methodik: Die Synthese der Tagungsbeiträge und der Literatur sowie die Ableitung der Handlungsfelder sind Eigenleistungen der Verfassenden.