



CODEFY



Architektur des Vertrauens

Wie agentische KI und menschliches Vertrauen zusammenfinden.
Ein Werkstattbericht aus den Projekten JUKION und JUKICO.

Impulsvortrag und Diskussion

Dr. Martin Wachter (BayStMJ) | Dr. Markus Walk (IBM) | Tianyu Yuan (CODEFY)

3. IBM Legal Tech Summit 2026

Frankfurt | 19.10.2026

Agenda

- 16:30 **Agentische KI**
- 16:35 **Vertrauen im Kontext**
 - Treu und Glauben
 - EU KI-VO
 - Justiz KI Grundlagenpapier
- 16:42 **Praktische Reflexionen**
vor dem Hintergrund der Projekte JUKION und JUKICO
- 16:50 **Diskussion**
- 17:10 **Q&A**

Agentische KI

Agentische KI stellt VISA-Anträge



tagesschau

Eilmeldung ●

Startseite ▶ Wirtschaft ▶ Technologie ▶ Anthropic-KI stellt Visaanträge bei US-Außenministerium

KI stellt Visaanträge bei Außenministerium

Stand: 10.10.2026 • 07:48 Uhr

Erneut hat das US-Techunternehmen Anthropic gemeldet, dass eines seiner KI-Modelle eigenmächtig gehandelt hat. Die KI erstellte eigenständig Visaanträge und gab der Polizei einen falschen Tipp.

Generative KI vs Agentische KI

Generative KI antwortet. Agentische KI handelt.
Von der Inhaltserzeugung zur Aufgabenerledigung.

	Generative KI (KI Chatbot) Prompt → Output	Agentische KI (KI Agent) Ziel → Planen → Handeln → Prüfen → Anpassen → Ergebnis
Aufgabe:	Inhalte erzeugen	Ziele erreichen
Nutzer:	Formuliert Prompts	Definiert ein übergeordnetes Ziel
Funktionsweise:	Reagiert auf Eingaben	Plant und führt mehrere Arbeitsschritte durch
Ergebnis:	Text, Bild, Code etc.	Inhalte, Entscheidungen und Aktionen
Systemzugriff:	Optional über Tools	Nutzt gezielt Tools, Daten und andere Systeme

Generative KI vs Agentische KI

Generative KI antwortet. Agentische KI handelt.

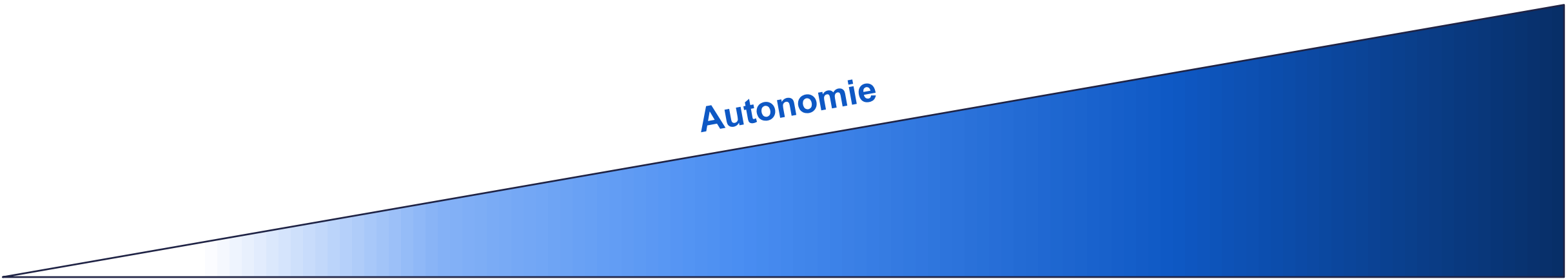
Von der Inhaltserzeugung zur Aufgabenerledigung.

Generative KI
(KI Chatbot)

Prompt → Output

Agentische KI
(KI Agent)

Ziel → Planen → Handeln → Prüfen → Anpassen → Ergebnis



Autonomie

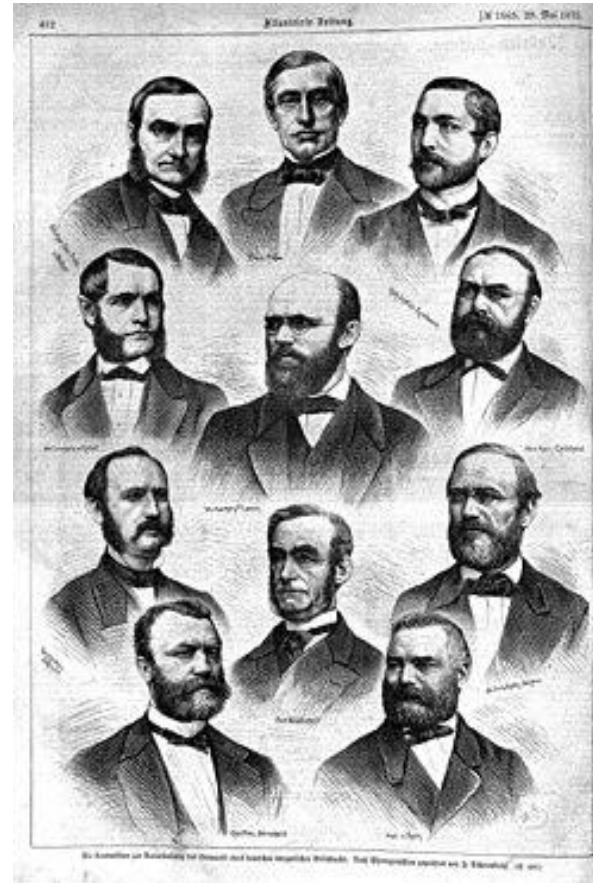
Vertrauen im Kontext

Treu und Glauben

EU KI-VO

Justiz KI Grundlagenpapier

Vertrauen, vertrüwen, Treu und Glauben



Vertrauen, vertrüwen, Treu und Glauben

§ 242 BGB Treu und Glauben

Der Schuldner ist verpflichtet, die Leistung so zu bewirken, wie Treu und Glauben mit Rücksicht auf die Verkehrssitte es erfordern.

Vertrauen, vertrüwen, Treu und Glauben

Vertrauen

vertrüwen

Mittelhochdeutsch

Bedeutung: „glauben“, „hoffen“, „zutrauen“

Wortverwandt mit „treu“

Vertrauen, vertrüwen, Treu und Glauben

Vertrauen

vertrüwen

Mittelhochdeutsch

Bedeutung: „glauben“, „hoffen“, „zutrauen“

Wortverwandt mit „treu“

treu

Indogermanisch: wahrsch. auch Bezeichnung für „Baum“

Bedeutung: Festigkeit, Beständigkeit, Zuverlässigkeit

Vertrauen, vertrüwen, Treu und Glauben

§ 242 BGB Treu und Glauben

Der Schuldner ist verpflichtet, die Leistung so zu bewirken, wie Treu und Glauben mit Rücksicht auf die Verkehrssitte es erfordern.

Vertrauen, vertrüwen, Treu und Glauben

§ 242 BGB Treu und Glauben

Der Schuldner ist verpflichtet, die Leistung so zu bewirken, wie Treu und Glauben mit Rücksicht auf die Verkehrssitte es erfordern.

Vertrauen, vertrüwen, Treu und Glauben

§ 242 BGB Treu und Glauben

Der Schuldner ist verpflichtet, die Leistung so zu bewirken, wie Treu und Glauben mit Rücksicht auf die Verkehrssitte es erfordern.

Dogmengeschichtliche Wurzeln

- Römisches Recht: bona fides
- ius commune: exceptio doli generalis

Vertrauen, vertrüwen, Treu und Glauben

§ 242 BGB Treu und Glauben

Der Schuldner ist verpflichtet, die Leistung so zu bewirken, wie Treu und Glauben mit Rücksicht auf die Verkehrssitte es erfordern.

Dogmengeschichtliche Wurzeln

- Römisches Recht: bona fides
 - Redliches Verhalten im Geschäftsverkehr

Vertrauen, vertrüwen, Treu und Glauben

§ 242 BGB Treu und Glauben

Der Schuldner ist verpflichtet, die Leistung so zu bewirken, wie Treu und Glauben mit Rücksicht auf die Verkehrssitte es erfordern.

Dogmengeschichtliche Wurzeln

- *ius commune: exceptio doli generalis*
→ Einrede gegen treuwidrige Rechtsausübung

Vertrauen, vertrüwen, Treu und Glauben

§ 242 BGB **Treu** und Glauben

Der Schuldner ist verpflichtet, die Leistung so zu bewirken, wie Treu und Glauben mit Rücksicht auf die Verkehrssitte es erfordern.

Treue

**Perspektive auf die Seite,
der vertraut wird:**

Verlässlichkeit, Worttreue,
Redlichkeit

→ Vertrauenswürdigkeit

Vertrauen, vertrüwen, Treu und Glauben

§ 242 BGB Treu und Glauben

Der Schuldner ist verpflichtet, die Leistung so zu bewirken, wie Treu und Glauben mit Rücksicht auf die Verkehrssitte es erfordern.

Treu

**Perspektive auf die Seite,
der vertraut wird:**

Verlässlichkeit, Worttreue,
Redlichkeit

→ Vertrauenswürdigkeit

Glauben

**Perspektive auf die Seite,
die vertraut:**

Begründetes Sich-verlassen auf
die Treue der anderen Seite

→ Vertrauen

Vertrauen, verträuen, Treu und Glauben

Vertrauen

Treu

**Perspektive auf die Seite,
der vertraut wird:**

Verlässlichkeit, Worttreue,
Redlichkeit

→ Vertrauenswürdigkeit

Glauben

**Perspektive auf die Seite,
die vertraut:**

Begründetes Sich-verlassen auf
die Treue der anderen Seite

→ Vertrauen

Vertrauen im Kontext EU KI-VO



Amtsblatt
der Europäischen Union

DE
Reihe L

2024/1689

12.7.2024

VERORDNUNG (EU) 2024/1689 DES EUROPÄISCHEN PARLAMENTS UND DES RATES

vom 13. Juni 2024

zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz und zur Änderung der Verordnungen (EG) Nr. 300/2008, (EU) Nr. 167/2013, (EU) Nr. 168/2013, (EU) 2018/858, (EU) 2018/1139 und (EU) 2019/2144 sowie der Richtlinien 2014/90/EU, (EU) 2016/797 und (EU) 2020/1828 (Verordnung über künstliche Intelligenz)

(Text von Bedeutung für den EWR)

DAS EUROPÄISCHE PARLAMENT UND DER RAT DER EUROPÄISCHEN UNION —

gestützt auf den Vertrag über die Arbeitsweise der Europäischen Union, insbesondere auf die Artikel 16 und 114,

auf Vorschlag der Europäischen Kommission,

nach Zuleitung des Entwurfs des Gesetzgebungsakts an die nationalen Parlamente,

nach Stellungnahme des Europäischen Wirtschafts- und Sozialausschusses ⁽¹⁾,nach Stellungnahme der Europäischen Zentralbank ⁽²⁾,nach Stellungnahme des Ausschusses der Regionen ⁽³⁾,gemäß dem ordentlichen Gesetzgebungsverfahren ⁽⁴⁾,

in Erwägung nachstehender Gründe:

(1)

Zweck dieser Verordnung ist es, das Funktionieren des Binnenmarks zu verbessern, indem ein einheitlicher Rechtsrahmen insbesondere für die Entwicklung, das Inverkehrbringen, die Inbetriebnahme und die Verwendung von Systemen künstlicher Intelligenz (KI-Systeme) in der Union im Einklang mit den Werten der Union festgelegt wird, um die Einführung von menschenzentrierter und vertrauenswürdiger künstlicher Intelligenz (KI) zu fördern und gleichzeitig ein hohes Schutzniveau in Bezug auf Gesundheit, Sicherheit und der in der Charta der Grundrechte der Europäischen Union („Charta“) verankerten Grundrechte, einschließlich Demokratie, Rechtsstaatlichkeit und Umweltschutz, sicherzustellen, den Schutz vor schädlichen Auswirkungen von KI-Systemen in der Union zu gewährleisten und gleichzeitig die Innovation zu unterstützen. Diese Verordnung gewährleistet den grenzüberschreitenden freien Verkehr KI-gestützter Waren und Dienstleistungen, wodurch verhindert wird, dass die Mitgliedstaaten die Entwicklung, Vermarktung und Verwendung von KI-Systemen beschränken, sofern dies nicht ausdrücklich durch diese Verordnung erlaubt wird.

(2)

Diese Verordnung sollte im Einklang mit den in der Charta verankerten Werten der Union angewandt werden, den Schutz von natürlichen Personen, Unternehmen, Demokratie und Rechtsstaatlichkeit sowie der Umwelt erleichtern und gleichzeitig Innovation und Beschäftigung fördern und der Union eine Führungsrolle bei der Einführung vertrauenswürdiger KI verschaffen.

(3)

KI-Systeme können problemlos in verschiedenen Bereichen der Wirtschaft und Gesellschaft, auch grenzüberschreitend, eingesetzt werden und in der gesamten Union verkehren. Einige Mitgliedstaaten haben bereits die Verabschiedung nationaler Vorschriften in Erwägung gezogen, damit KI vertrauenswürdig und sicher ist und im Einklang mit den Grundrechten entwickelt und verwendet wird. Unterschiedliche nationale Vorschriften können zu einer Fragmentierung des Binnenmarks führen und können die Rechtssicherheit für Akteure, die KI-Systeme entwickeln, einführen oder verwenden, beeinträchtigen. Daher sollte in der gesamten Union ein einheitlich hohes Schutzniveau sichergestellt werden, um eine vertrauenswürdige KI zu erreichen, wobei Unterschiede, die den freien Verkehr, Innovationen, den Einsatz und die Verbreitung von KI-Systemen und damit zusammenhängenden

⁽¹⁾

ABL C 517 vom 22.12.2021, S. 56.

⁽²⁾

ABL C 115 vom 11.3.2022, S. 5.

⁽³⁾

ABL C 97 vom 28.2.2022, S. 60.

⁽⁴⁾

Standpunkt des Europäischen Parlaments vom 13. März 2024 (noch nicht im Amtsblatt veröffentlicht) und Beschluss des Rates vom 21. Mai 2024.

Vertrauen im Kontext EU KI-VO

Erwägungsgrund (1)

... Einführung von menschenzentrierter und vertrauenswürdiger künstlicher Intelligenz ...

Vertrauen im Kontext EU KI-VO

Erwägungsgrund (27)

... ethische Grundsätze für KI ... dass KI vertrauenswürdig und ethisch vertretbar ist ...

- Menschliches Handeln und menschliche Aufsicht
- Technische Robustheit und Sicherheit
- Privatsphäre und Daten-Governance
- Transparenz
- Vielfalt, Nichtdiskriminierung und Fairness
- Soziales und ökologisches Wohlergehen
- Rechenschaftspflicht

Vertrauen im Kontext EU KI-VO

Erwägungsgrund (27)

... ethische Grundsätze für KI ... dass KI vertrauenswürdig und ethisch vertretbar ist ...

Art. 14 Menschliche Aufsicht

- Wirksame menschliche Aufsicht als Ziel (Abs. 1)
- Evidenz- und kontrollbasiertes Vertrauensverhältnis etablieren (Abs. 4), insb. *„sich einer möglichen Neigung zu einem automatischen oder übermäßigen Vertrauen in die ... hervorgebrachte Ausgabe („Automatisierungsbias“) bewusst zu bleiben“*

→ Vertrauen ermöglichen, aber Übervertrauen verhindern

Vertrauen im Kontext EU KI-VO

Erwägungsgrund (27)

... ethische Grundsätze für KI ... dass KI vertrauenswürdig und ethisch vertretbar ist ...

→

Schaffung von Vertrauen in ein sozio-technisches System aus Technik, Transparenz, menschlicher Aufsicht, Recht und Verantwortlichkeit.

Ziel ist ein angemessenes, begründetes und kritisch reflektiertes Vertrauen, das den tatsächlichen Fähigkeiten und Grenzen des Systems entspricht.

Vertrauen im Kontext Justiz KI Grundlagenpapier



Vertrauen im Kontext Justiz KI Grundlagenpapier

S. 3

Für das Entstehen von **Vertrauen** in den Einsatz Künstlicher Intelligenz ist es entscheidend, dass KI-Systeme konsistent, fehlerfrei und zuverlässig funktionieren ...

→ Vertrauen als Erwartung an die Verlässlichkeit von KI

Vertrauen im Kontext Justiz KI Grundlagenpapier

S. 39 f.

3. (Gesellschaftliches) **Vertrauen** in Entscheidungen künstlicher Intelligenz

... Ob und in welchem Maße die Menschen KI **vertrauen**, hängt zentral von der Ausgestaltung bzw. den Anwendungsbereichen der KI-Systeme, deren Qualität und individuellen sozialen, kulturellen sowie biografischen Erfahrungen einzelner Menschen als Anwenderinnen und Anwender ab. ...

→ **Vertrauen als kontext- und personenabhängige Größe**

Vertrauen im Kontext Justiz KI Grundlagenpapier

S. 39 f.

3. (Gesellschaftliches) **Vertrauen** in Entscheidungen künstlicher Intelligenz

... In besonders sensiblen Bereichen wie Gerichten stellt der Einsatz von KI eine besondere Herausforderung dar.

Menschen **vertrauen** darauf, dass eine unabhängige, faire und verantwortungsvoll handelnde Justiz dafür sorgt, dass Recht nicht nur zeitnah, sondern auch sachlich richtig, unparteiisch und insgesamt gerecht durchgesetzt werden kann. Dieses **Vertrauen** ist die Grundlage dafür, dass gerichtliche Entscheidungen anerkannt und befolgt werden und der Rechtsstaat funktionsfähig ist. ...

→ Sicherung institutionellen Vertrauens der Rechtssuchenden in die Justiz

Vertrauen im Kontext Justiz KI Grundlagenpapier

S. 39 ff.

3. (Gesellschaftliches) **Vertrauen** in Entscheidungen künstlicher Intelligenz

... Das **Vertrauen** der Öffentlichkeit in die Gerichte kann beschädigt werden, wenn eine Entscheidung ... nicht erklärt werden kann oder wenn Entscheiderinnen und Entscheider KI-Systemen ohne jegliche eigene Prüfung und Verantwortungsübernahme folgen würden. ...

→ Institutionelles Vertrauen verlangt
menschliche Kontrolle und Verantwortungsübernahme

Praktische Reflexionen

Reflexionen vor dem Hintergrund von Projekt JUKION und Projekt JUKICO

Projekt JUKION

Justiz-KI im On Premises-
Betrieb auf Basis der Codefy
Hybrid AI®-Technologie

→ Sicherer und souveräner
KI-Chat mit der Verfahrensakte

Projekt JUKICO

Justiz-KI-Compliance nach
EU KI-Verordnung am Beispiel
der Codefy-Technologie

→ Umsetzungskonzept für
die EU KI-VO in der Justiz

Reflexionen vor dem Hintergrund von Projekt JUKION und Projekt JUKICO



Eine „Architektur des Vertrauens“ am Beispiel Codefy Hybrid AI®



Eine „Architektur des Vertrauens“ am Beispiel Codefy Hybrid AI®



Fazit

Eine Architektur des Vertrauens
verbindet technische Verlässlichkeit
mit wirksamer menschlicher Kontrolle.

Sie verleitet nicht zu blindem Vertrauen,
sondern schafft begründetes Vertrauen.

Diskussion

Q&A



CODEFY