

02.12.2025 | Bankhaus Metzler

KI-Agent – bitte übernehmen Sie! Wertschöpfungspotentiale für die Industrie

Fraunhofer Allianz BIG DATA KI

Agenda

Teil 1 Technische Perspektive

Teil 2 Wirtschaftliche Perspektive

KI revolutioniert die Welt (immer noch) in Rekordzeit

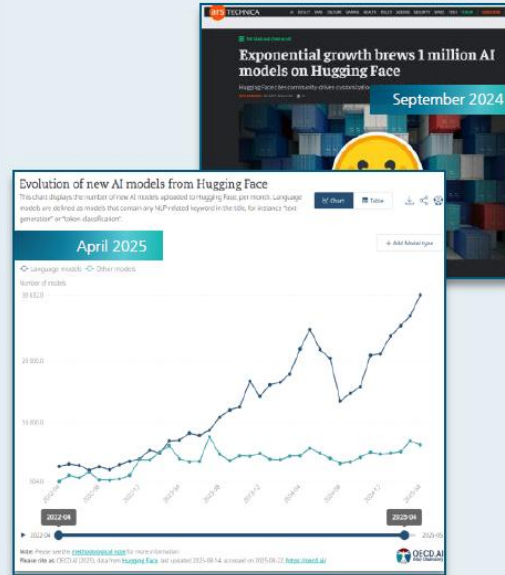
Meilensteine der letzten Monate

Oktober 2024
Zwei Nobelpreise für das Thema KI



Quelle: [zdf heute](#), [inside IT](#)

April 2025
Monatlich >30.000 Sprachmodelle auf Hugging Face



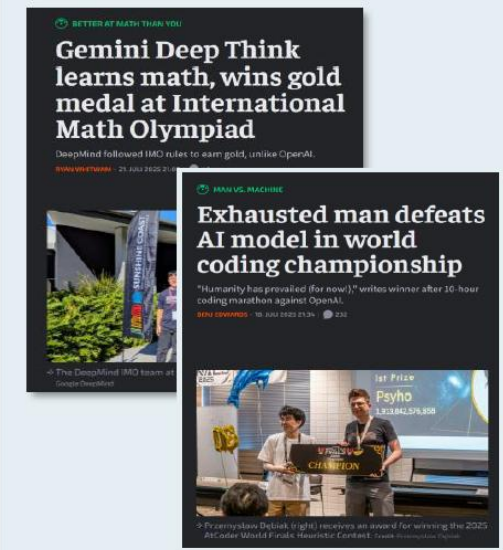
Quelle: [ars technica](#), [OECD AI \(2025\)](#)

Juli 2025
Chinas zweiter KI-Durchbruch bestätigt Aufholjagd



Quelle: [tagesschau](#), [FAZ](#)

Juli 2025
KI siegt in Mathe, Mensch (noch) im Coding



Quelle: [Ars Technica](#), [Ars Technica](#)

Trends mit Transformationspotential 2025+

Aktuelle Herausforderungen spiegeln sich auch teilweise in den Trends wider

AI Agents

Von einfachen Assistenten die in Unternehmensanwendungen eingebettet sind, über aufgabenspezifischen Agenten bis zum Multiagenten-Ökosystemen

AI Ready Data

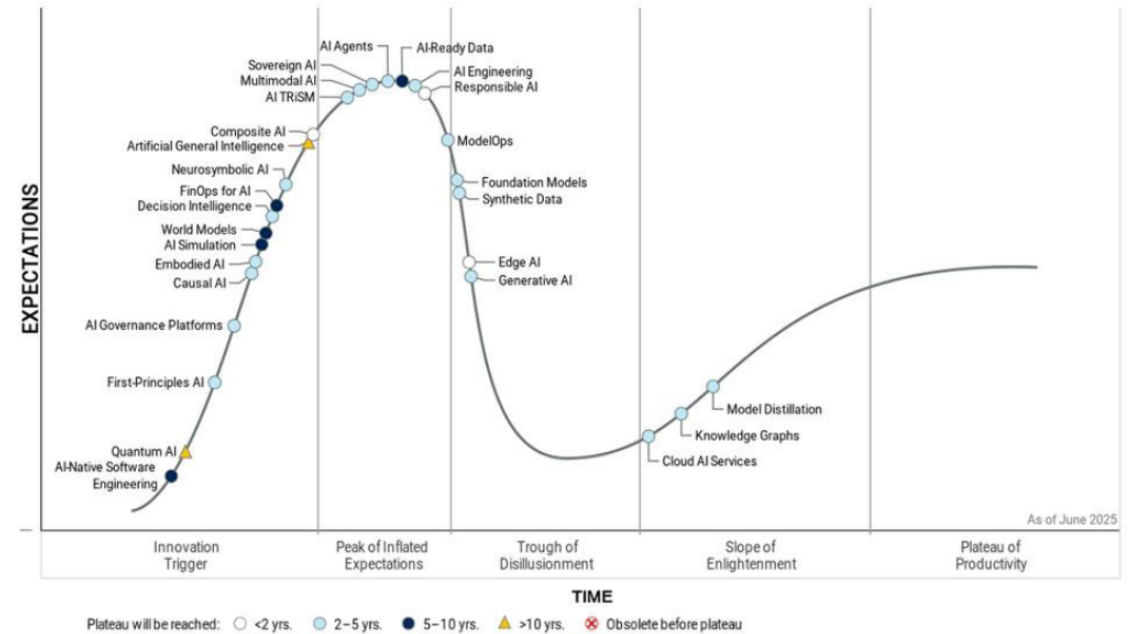
Datensätze, die für den Einsatz in KI-Anwendungen optimiert, kontextbezogene Genauigkeit und Effizienz, Vertrauen, Risiken und Compliance

Multimodal AI

Hilft, die Welt besser zu erfassen, eröffnet neue Möglichkeiten für KI in jeder Anwendung und Softwarelösung über alle Branchen hinweg

AI TRiSM

Technologien zur Sicherstellung von ethischer, sicherer und regelkonformer KI-Nutzung.



Gartner


Quelle: <https://www.gartner.com/en/newsroom/press-releases/2025-08-05-gartner-hype-cycle-identifies-top-ai-innovations-in-2025>

Überschneidung von Transformationspotenzialen und Herausforderungen. Ein Wendepunkt: Weg vom Hype hin zu skalierbaren, vertrauenswürdigen und nachhaltigen KI-Lösungen mit echtem Mehrwert

„Klassische“ KI vs. GenAI

Was kann die eine, was die andere?

Here is a Python script that prints the first 100 prime numbers:

```
PYTHON   
  
def is_prime(n):  
    if n <= 1:  
        return False  
    for i in range(2, int(n**0.5) + 1):  
        if n % i == 0:  
            return False  
    return True  
  
def first_n_primes(n):  
    primes = []  
    num = 2  
    while len(primes) < n:  
        if is_prime(num):  
            primes.append(num)  
        num += 1  
    return primes  
  
# Print the first 100 prime numbers  
print(first_n_primes(100))
```

You can run this script in a Python environment to see the first 100 prime numbers.

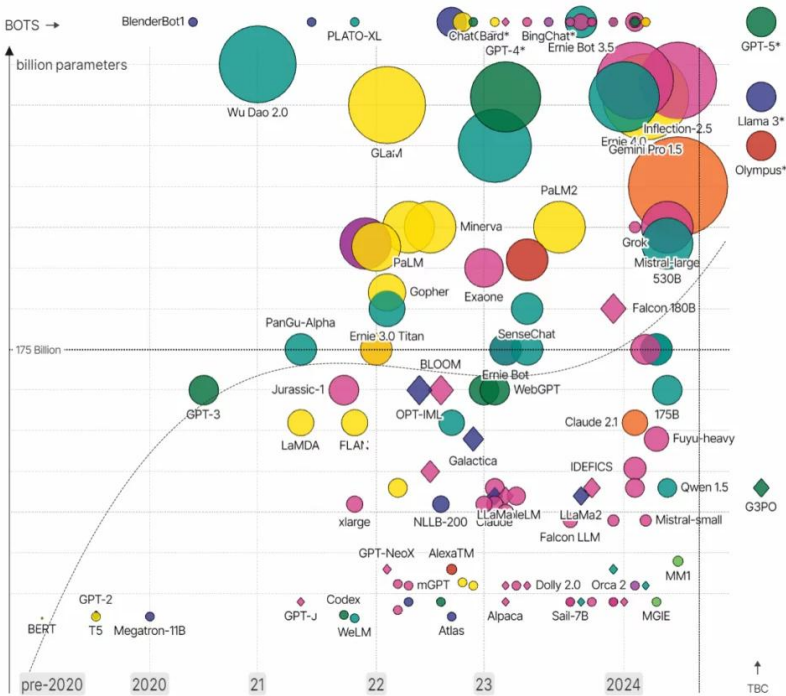
GenAI:

- Erste Anfänge von „echtem Verständnis“:
 - Schreibe ein Programm in Python, dass die ersten 100 Primzahlen ausgibt!
- Ergebnis:
 - Software entwickelt
 - Rechenoperationen korrekt ausgeführt
 - Liste korrekt ausgegeben

Was ist „groß“ bei Large Language Models?

Große Sprachmodelle der Generativen KI

Modell	Größe	Benötigter Rechner
openAI GPT 3	175B	HPC-Cluster
openAI GPT 4	1,76T	HPC-Cluster
Llama-2	70B	Kleiner HPC-Cluster
Teuken	7B	Großer PC



David McCandless, Tom Evans, Paul Barton
Information is Beautiful // UPDATED 20th Mar 24

source: news reports, [LifeArchitectureAI](#)
* = parameters undisclosed // see [the data](#)

MADE WITH [VIZsweat](#)

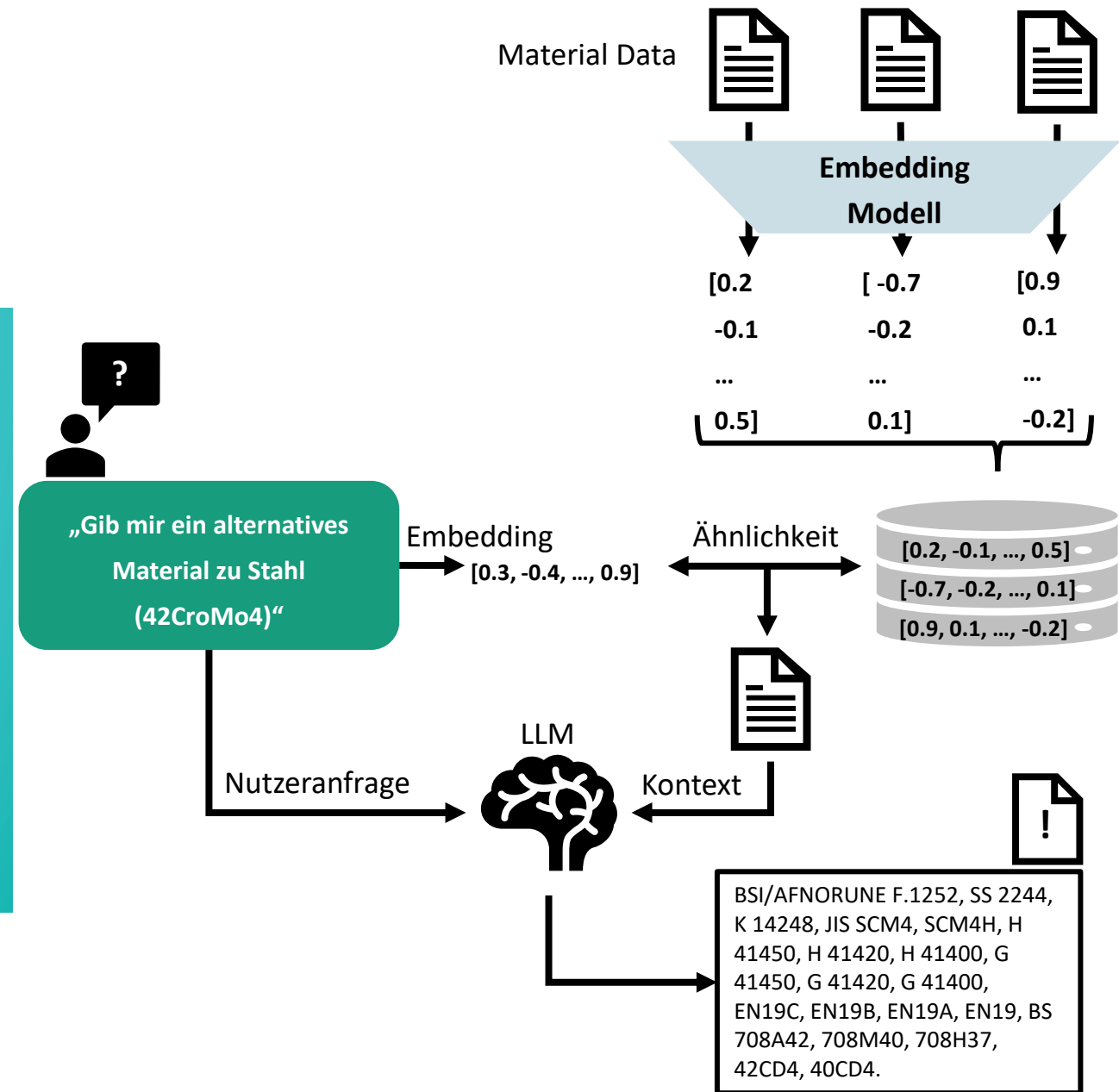
The rise and rise of AI-based Large Language Models (LLMs) like GPT4, LLaMA, LLaMa, PaLM and Jurassic-2.

Quelle: <https://nestify.io/blog/top-large-language-models/>

„Wissen“ in LLMs

Generalisten vs. Spezialisten

- Möglichkeit 1: Frage LLM direkt
„Wieviel ist 2+2?“
„Ist dieses NDA mit unseren Policies vereinbar?“
- Möglichkeit 2: Bitte LLM, die richtige Lösung aus geprüften Dokumenten herauszusuchen
„Sieh in der beigelegten Tabelle nach: Wieviel ist 2+2?“
„Vergleiche dieses NDA mit unseren Policies, die im angehängten Dokument beschrieben sind!“



Ausblick GenAI: Technische Perspektive

Densing Law (Xiao et al.)

Die Kapazitätsdichte verdoppelt sich alle 3,3 Monate.

- **Kapazitätsdichte = Größe Referenz LLM / Größe aktuelles LLM**
- **Kapazitätsdichte ist entscheidende Metrik für Verbesserung von Sprachmodellen**
- **Künftige KI-Fortschritte: Modelle effizienter und nicht nur größer**

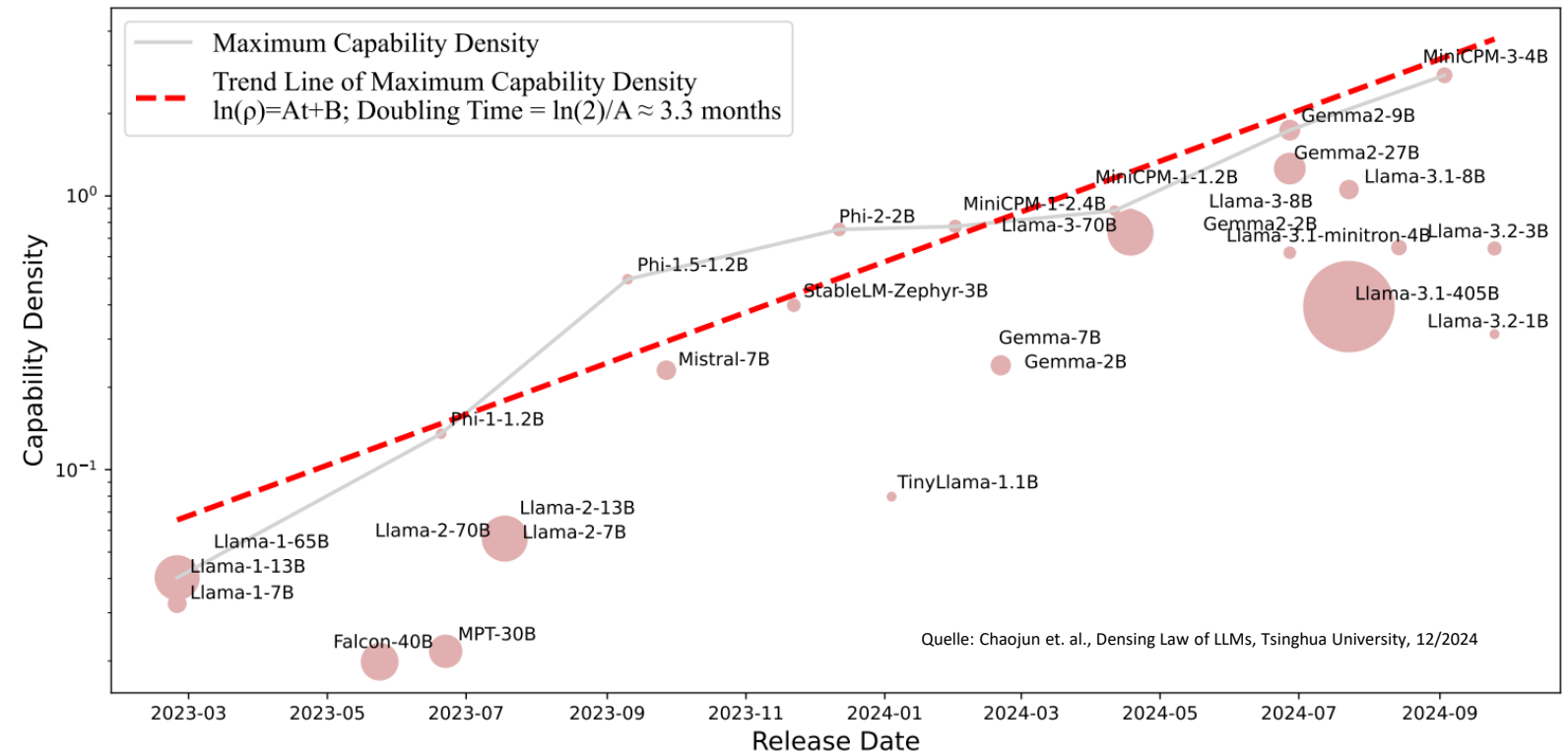


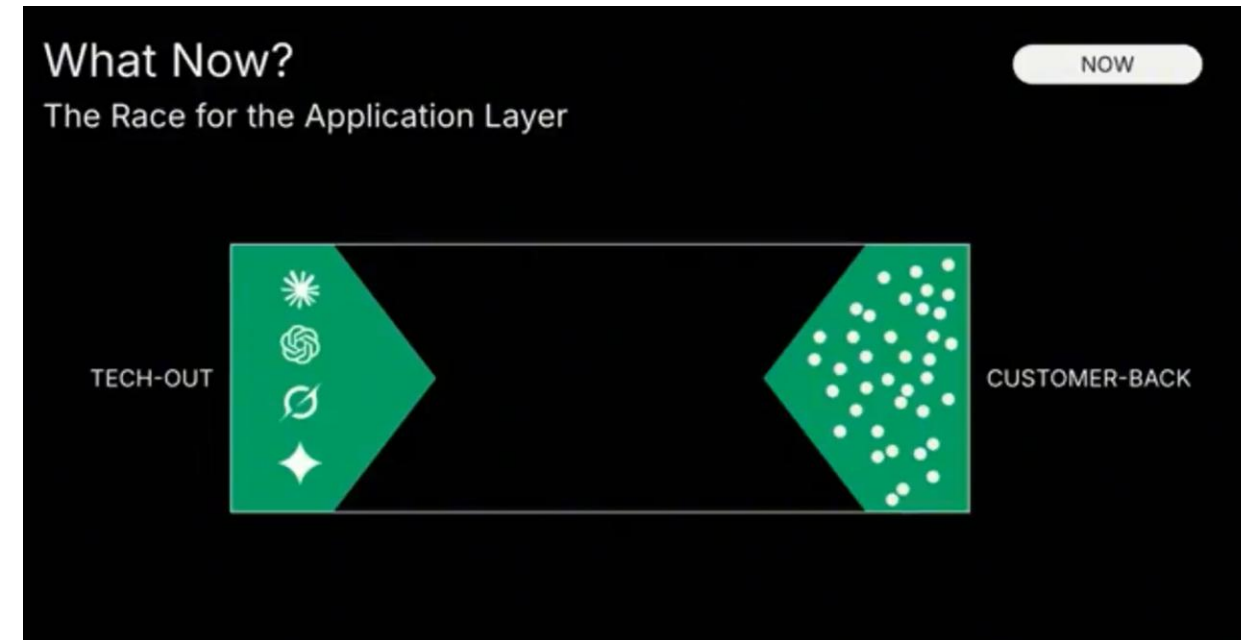
Figure 1: The estimated capability density of open-source base LLMs.

Ausblick GenAI: Technische Perspektive

Blickwinkel aus Investorensicht

KI beginnt 10-mal größer als der Cloud-Boom

- Start von ChatGPT „The rails were in place“
- Reale Wertschöpfung
 - Modelle sind mächtig, aber Wert liegt in den Anwendungen (Application layer)
 - Gleichzeitig scheren Basismodelle aus ihrer Spur, dringen in Anwendungsgebiete vor
 - Infrastruktur ist Grundlage, aber wichtiger ist von Menschen genutzte Anwendungsebene (Software)
 1. Bei Cloud
 2. Bei Mobile
 3. Nun selbes Muster bei KI
 - KI Applikationen vom Kunden aus aufbauen (bestimmte Nutzer, bestimmter Markt, echte Arbeitsabläufe)

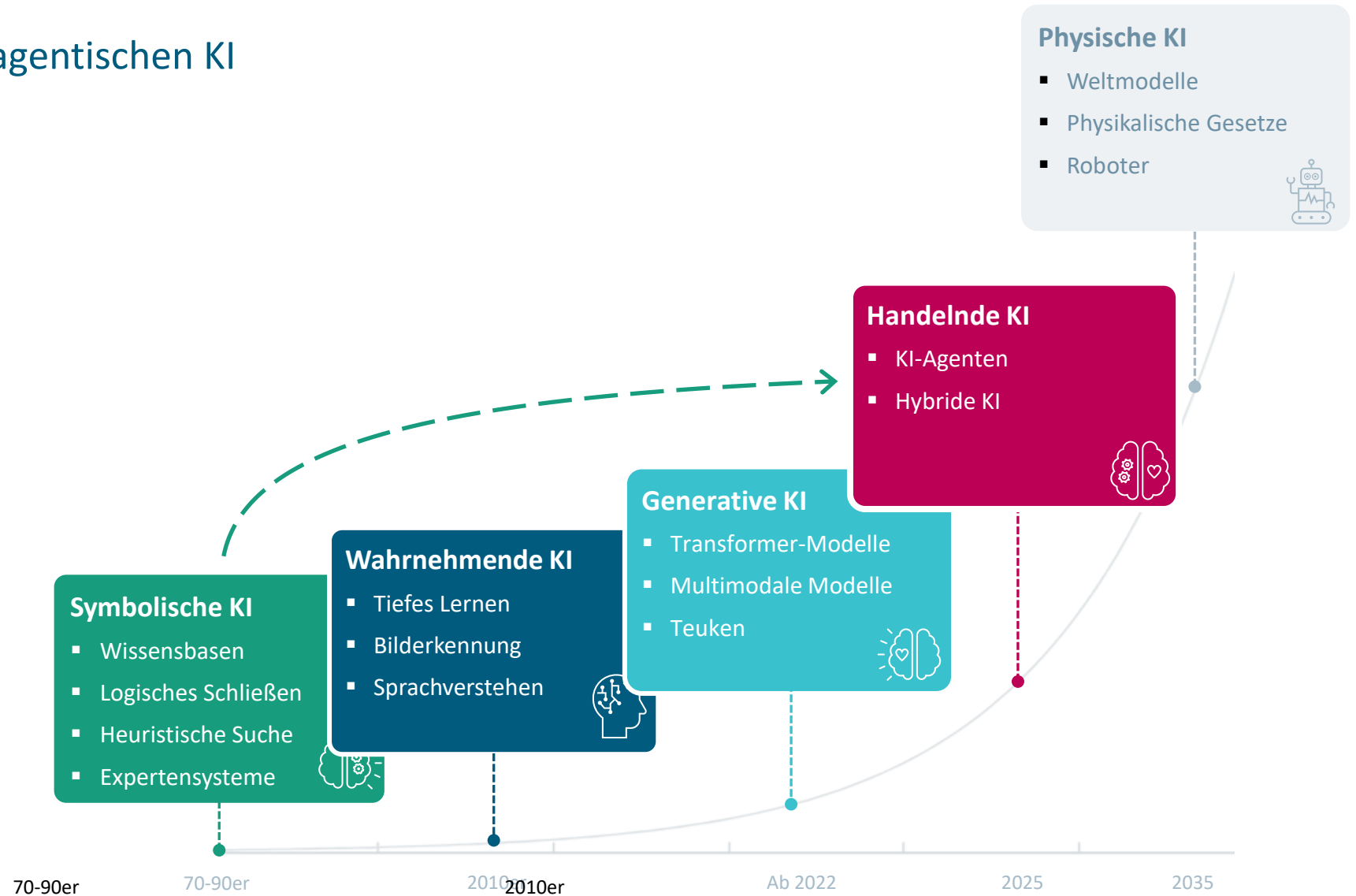


Quelle: Sequoia AI Ascent, 05/2025

KI-Evolution

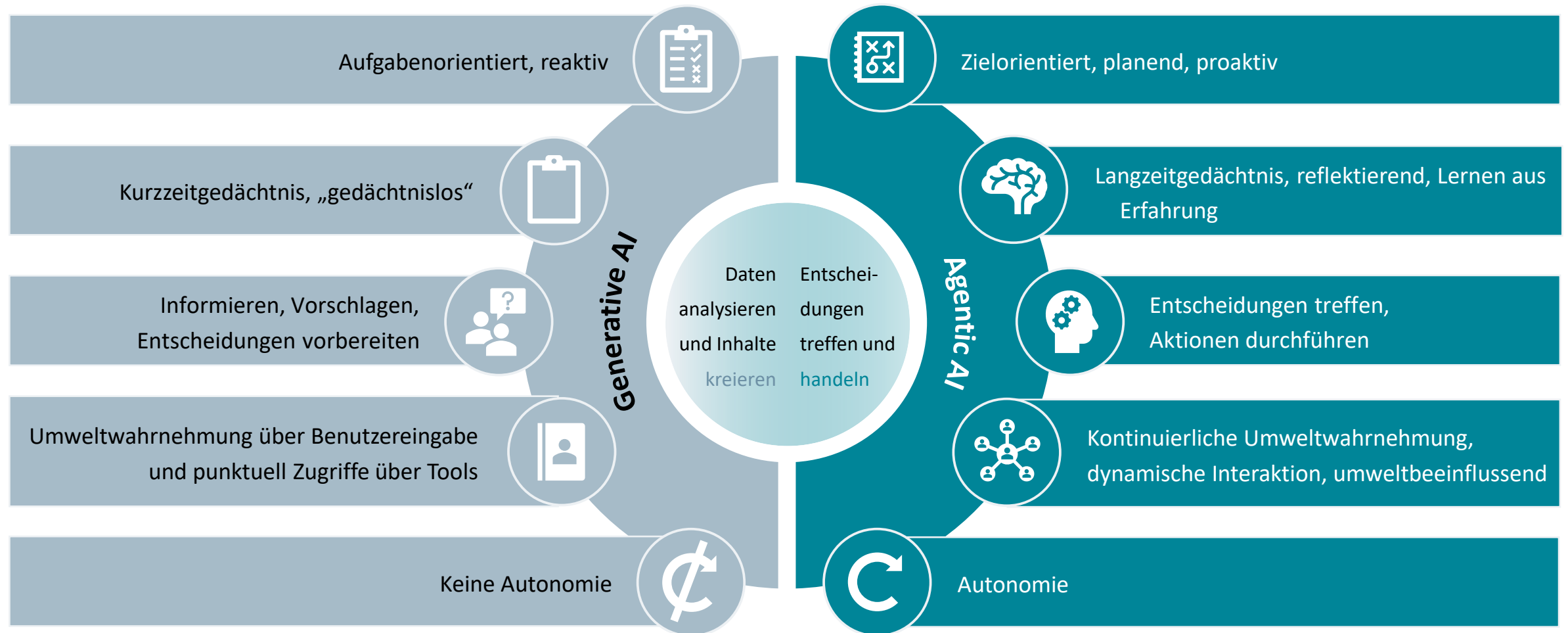
2025: Von der generativen zur agentischen KI

>>> Handelnde KI interagiert mit ihrer Umgebung und löst komplexe Aufgaben



Von Generativer AI zu Agentic AI

Vom Informieren und Kreieren zum Entscheiden und autonomen Handeln



Verschiedene Evolutionsstufen der KI und der Rolle des Menschen

Evolutionsstufe	Initiator	Produzent	Weiterverarbeiter	Beispiel
Ohne KI	Mensch	Mensch	Mensch	
Prompted AI	Mensch	KI	Mensch	ChatGPT
Assisting AI	KI	KI	Mensch	Copilot in MS-Office
Agentic AI	KI	KI	KI	Terminplaner, Antragsbearbeitung

GenAI and beyond -> das »Agentic Web«

Agenten werden Hauptnutzer des Internets

Menschliche Oberflächen verlieren an Bedeutung, Agenten prüfen Verfügbarkeiten, führen Preisverhandlungen und wickeln Serviceanfragen eigenständig ab

Ziel

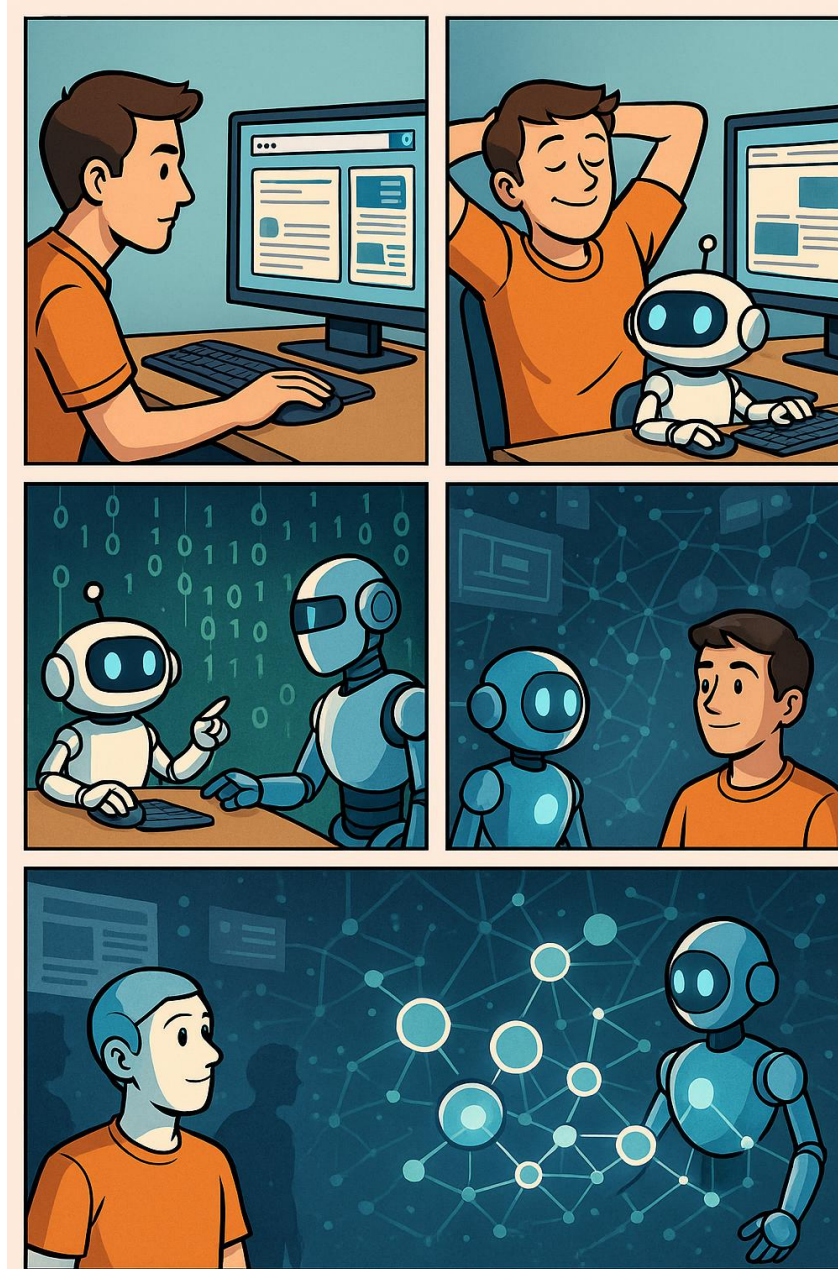
- Aufgaben Ende-zu-Ende zu erledigen

Herausforderung

- Faktenbasierte Entscheidungen treffen
- Vorsicht vor Fehlerketten – Fehler von Agent zu Agent verstärken sich

Konsequenz ->

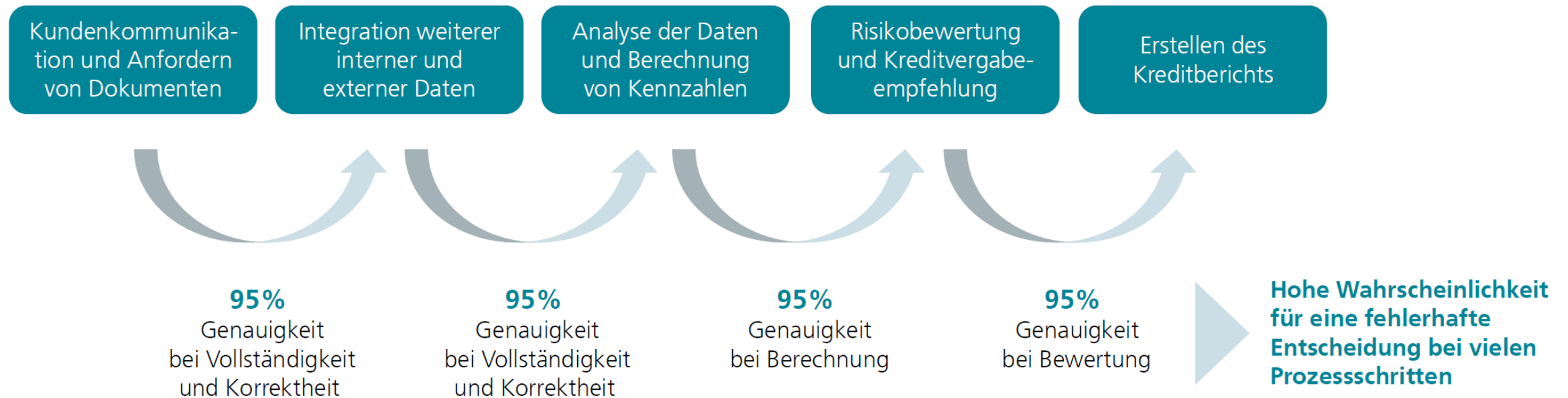
- Webseiten dienen nur noch als Informationsquellen für KI-Agenten und haben eine Oberfläche für die Handlungen der Agenten
- Webseiten werden zu KI-Schnittstellen optimiert
- Agenten müssen stetig evaluiert werden



Einsatz erfordert vertrauenswürdige Agentic AI in jeder Prozessstufe

Ungenauigkeiten verstärken sich über die Prozesskette

Beispiel der Kreditvergabeentscheidung



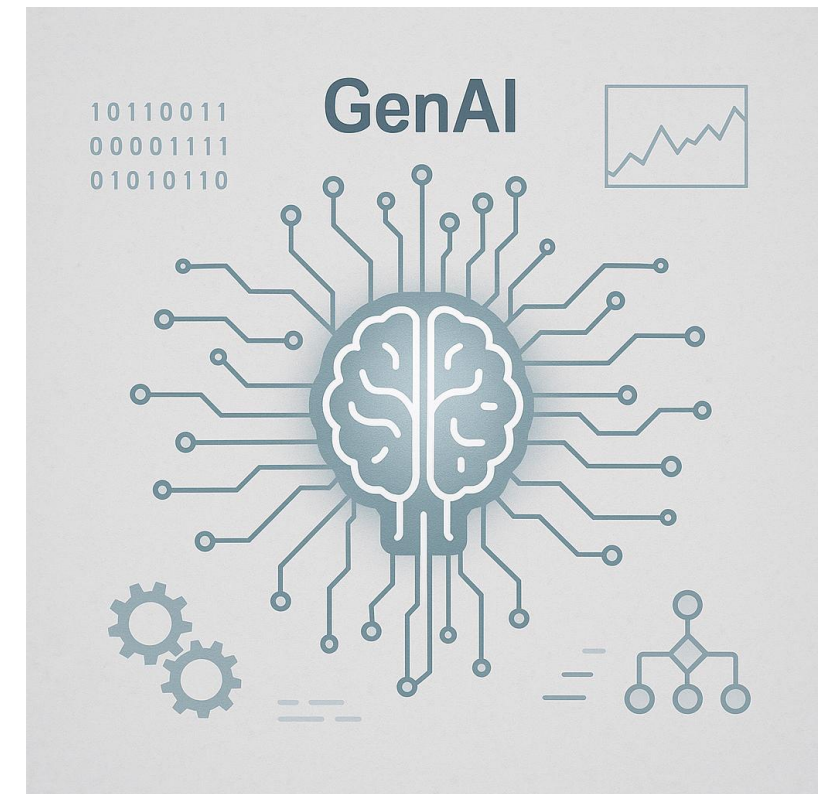
Take-home Message Teil 1

Generative/Agentische KI ...

- „versteht“ in einem gewissen, praktischen Sinn die Aufgaben und liefert kreative Lösungen.
- kann halluzinieren: Manchmal ist GenAI „zu kreativ“.
- Derzeit rasante Entwicklung: Mixture of experts, Agentic AI / MCP, RAG, ... macht GenAI
 - leistungsfähiger und gleichzeitig
 - Kleiner / billiger

Anleitung zum Glücklichein, technische Perspektive:

- KI ist ein Werkzeug. Generative KI ist ein Sparingspartner, der "mitdenken" kann!
- AgenticAI unbedingt auf dem Schirm behalten! Die Entwicklung verläuft stürmisch!
- Der nötige Aufwand sinkt ständig stark, die Leistungsfähigkeit steigt ständig stark.
=> Ständig über neue Einsatzfelder nachdenken!



Agenda

Teil 1 Technische Perspektive

Teil 2 Wirtschaftliche Perspektive

Wirtschaftliche Perspektive

Fraunhofer-Erfahrungen

3 verschiedene branchenunabhängige Geschäftsmodell-Varianten

Interne Prozessoptimierung
im Unternehmen

- erhöht Produktivität

Produkterweiterung & -
entwicklung

- steigert Umsatz

Eigenentwicklung – später
in Branche ausrollen

- Produktivität verbessern
- Ansatz des branchenweiten Ausrollens durch eigen entwickeltes Know-How

Interne Prozessoptimierung

Elektrotechnik: Angebotsanfrage

- Tägliche Bearbeitung einer großen Anzahl von Ausschreibungen und Angebotsanfragen (RFQs)
- **KI-Assistent “AI RFQ Processor”**
 - Abgleich der Anforderungen in Ausschreibungen und Angebotsanfragen mit den entsprechenden Produkten im Produktkatalog (über 50.000 Produkte)
 - Integration in Infrastruktur und Angebotsprozess
- Steigerung der Effizienz und Qualität der Dienstleistungen

Kosten- und Zeiteinsparungen für Zulieferer und Produzenten im Vertriebsprozess



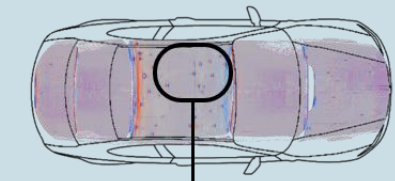
Interne Prozessoptimierung Prüfdienstleister: Qualitätskontrolle

- Das bildbasierte KI-System ermöglicht die automatische Erkennung, Klassifizierung und Messung von Schäden an Materialoberflächen (z. B. im Automobilbereich).
- Vielfältige Anwendungen sind die Qualitätssicherung in der industriellen Produktion und die Zustandsbewertung nach der Inbetriebnahme von Produkten (z.B. Überprüfung von Leasingfahrzeugen auf Schäden).
- Agenten-Workflow bereitet für Gutachten alle Vor-Informationen auf

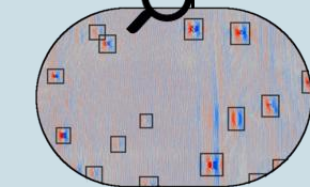
Object capture via video



2D reconstruction of the surface



Detection of damages using Deep Learning



Quelle: KI.NRW / <https://showroom.ki.nrw/bildbasierte-qualitaetskontrolle-damage-detection>

Echtzeit-Analyse sowie einfache
Systemintegration und Bedienung



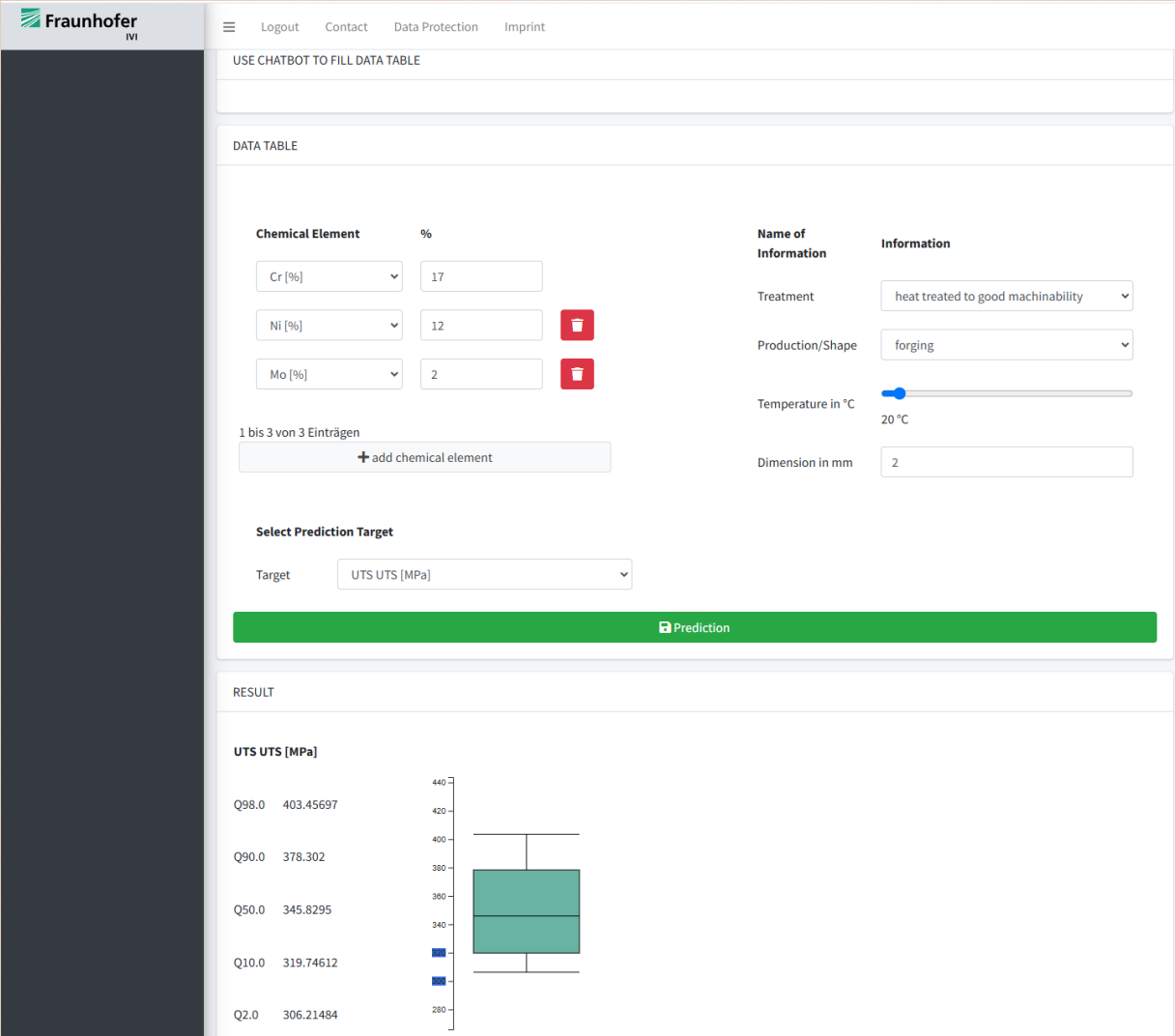
Produktweiterung & -entwicklung

Bau- & Chemiebranche

Effiziente Suche nach neuen Materialkombinationen

- Zusammenführung von chemischen Elementen in Form von Legierungen (350.000) in einem System
- Antrainiertes numerisches Generatives KI-Modell ermöglicht die automatische Suche neuer Material-Kombinationen und deren physikalischen Eigenschaften
- Nutzbar im Zuge von Rohstoffpreiserhöhung, Abschneiden von Lieferketten (De-Coupling), Handelsbeschränkungen und Multi-Sourcing Strategien

Analyse von Eigenschaften im materialwissenschaftlichen und prüfspezifischen Umfeld

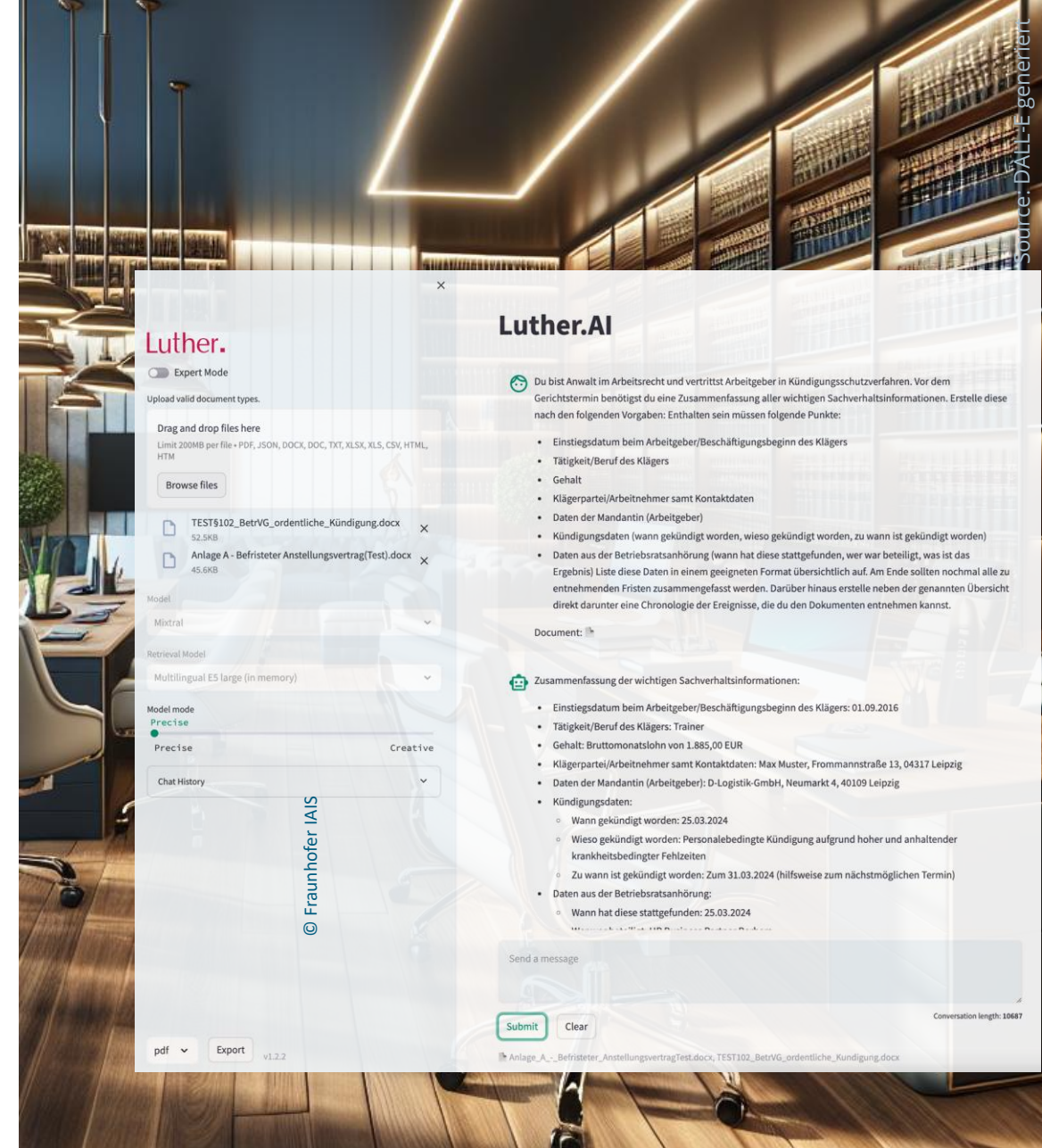


Eigenentwicklung – später in Branche ausrollen

Branche: Rechtsanwaltschaft - juristisches Analysetool

- Luther.AI ermöglicht Mitarbeitenden der Anwaltskanzlei Luther die Nutzung von **GenAI** auch im Kontext von **sensiblen Daten**
- Tool ermöglicht **sicheren Upload** von Dokumenten, auch mit sensiblen Kundeninformationen, und einfache und effiziente Interaktion mit Inhalten des Dokuments per Chat-Funktion
- Zwei Kernkomponenten:
 - Für **nicht-sensible Daten**: OpenAI's GPT-Modelle als Basis. Mitarbeitende lernen Umgang mit Large Language Models (LLMs) und es wird Bibliothek mit nützlichen Prompts erstellt
 - Für **sensible Daten**: Luther-Unternehmensmodell auf Grundlage von Open-Source-LLMs, die für juristischen Bereich optimiert wurden

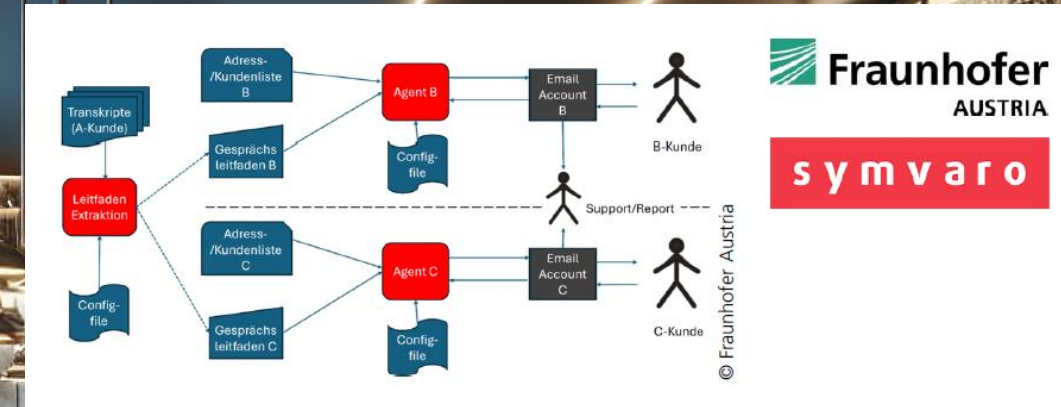
Analyse von sensiblen Rechtsdokumenten und Verträgen im unternehmensspezifischen LLM



Eigenentwicklung – später in Branche ausrollen

Branche: E-Mail Kundendialog

- Agenten-basierter Ansatz zur Automatisierung von E-Mail Kundendialog bei symvaro
- Automatisierung durch Agenten-basierte Klassifikation, Gesprächsstrukturierung und themenspezifische Reaktion
- Automatisch Mails Beantworten, Weiterleiten, Ignorieren
- Automatisierte Generierung und Versand personalisierter Startmails
- Zusammenfassung von Mailverläufen für internes Reporting



Analyse von sensiblen Rechtsdokumenten und Verträgen im unternehmensspezifischen LLM

Eigenentwicklung – später in Branche ausrollen

ALI – Das Automatisierungstool für die Wirtschaftsprüfung

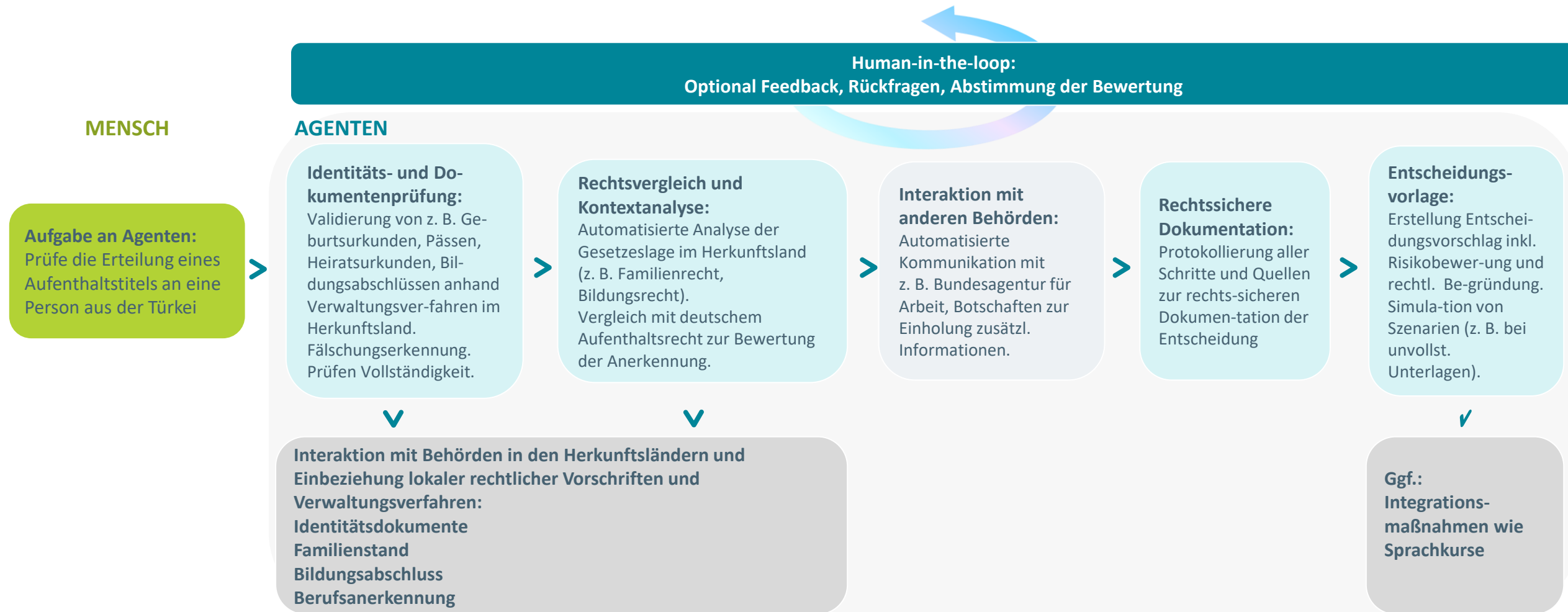
- KI-unterstützte Prüfung von Geschäftsberichtsanhängen
- ALI bietet KI-unterstützte Abarbeitung von Audit-Checklisten in Bezug auf Vollständigkeit und Voraussetzungenkonformität zu prüfender Dokumentation
- Anhangsprüfungsprozesse dauern zwischen 25 bis 12.000 Stunden; Checklisten von Wirtschaftsprüfern umfassen mehr als 1.000 Punkte
- ALI wird derzeit in 150 Ländergesellschaften ausgerollt.
- Weitere Anwendungsgebiete: (Teil-)automatische Berichterstellung, Vollständigkeits- und Konsistenzprüfung und Unterstützung in Ausschreibungsverfahren


Durch ALI kann jährlich **25-50 %** der Bearbeitungszeit für Anhangs-Checklisten von Geschäftsberichten eingespart werden



Automatisierte Prüfung und Erteilung von Aufenthaltstiteln

Agenten entlasten Ausländerbehörden bei komplexen und aufwändigen Aufgaben





Bundesamt
für Migration
und Flüchtlinge

Falls der Antrag über einen Kursträger gestellt wird,
bitte Namen und Kennziffer eintragen:

Antrag auf Zulassung zur Teilnahme am Integrationskurs gemäß § 5
Absatz 1 Integrationskursverordnung (IntV)

Name:

Ggf. Geburtsname:

Vorname:

Geburtsdatum:

Geschlecht: ☐ männlich ☐ weiblich ☐ divers

Straße/Hausnummer:

Postleitzahl/Ort:

Ggf. wohnhaft bei (c/o):

Telefonnummer:

Falls sich Ihre Anschrift ändert, teilen Sie dies dem Bundesamt bitte unverzüglich mit.

Jahr der Einreise:

Ich beantrage die Zulassung zu einem Integrationskurs (Sprach- und Orientierungskurs) gemäß § 44 Abs. 4
Satz 1 AufenthG.

Ich versichere, dass ich mich in keiner schulischen Ausbildung befinde und bisher noch an keinem durch das
Bundesamt geförderten Integrationskurs teilgenommen habe.

Dem Antrag ist eine Kopie des aktuellen Aufenthaltstitels beizufügen. Bei Staatsangehörigen eines
Mitgliedstaates der EU genügt eine Kopie des Ausweisdokuments (z. B. Reisepass).

Grund der Einreise nach Deutschland

☐ Familiennachzug zu einem deutschen Ehegatten (Nachweis der deutschen Staatsangehörigkeit beifügen)

☐ Familiennachzug zu einem ausländischen Ehegatten (Kopie des aktuellen Aufenthaltstitels des Ehegatten
beifügen)

☐ Sonstige Gründe:

Begründung des Antrages auf Zulassung zum Integrationskurs

☐ Nachweis ausreichender Kenntnisse der deutschen Sprache und Grundkenntnisse der Rechts- und
Gesellschaftsordnung und der Lebensverhältnisse im Bundesgebiet für die Erteilung einer
Niederlassungserlaubnis
Ein entsprechender Nachweis, dass ein Verfahren zum Erhalt einer Niederlassungserlaubnis bereits läuft, ist
beizufügen.

☐ Nachweis ausreichender Kenntnisse der deutschen Sprache für die Einbürgerung

☐ Sonstige Gründe:

NR: 418/0071 BASIF 01/2024

Seite 1 von 2

"Мене звати Олена Ковальчук. Я народилася 15 квітня 1988 року в Києві, Україна. На момент в'їзду до Німеччини, я була не одружена, тому мій дівочий прізвище та заразівше співпадають. Через війну в Україні, я вирішила шукати притулок в Німеччині і зараз проживаю за тимчасовою адресою: Ганзестрассе 5, 20099, Гамбург, Німеччина. Моя телефонний номер: +49 176 1234567. Я приїхала до Німеччини у 2023 році з метою знайти безпеку та захист. Моя мета в'їзду - пошук притулку від війни.

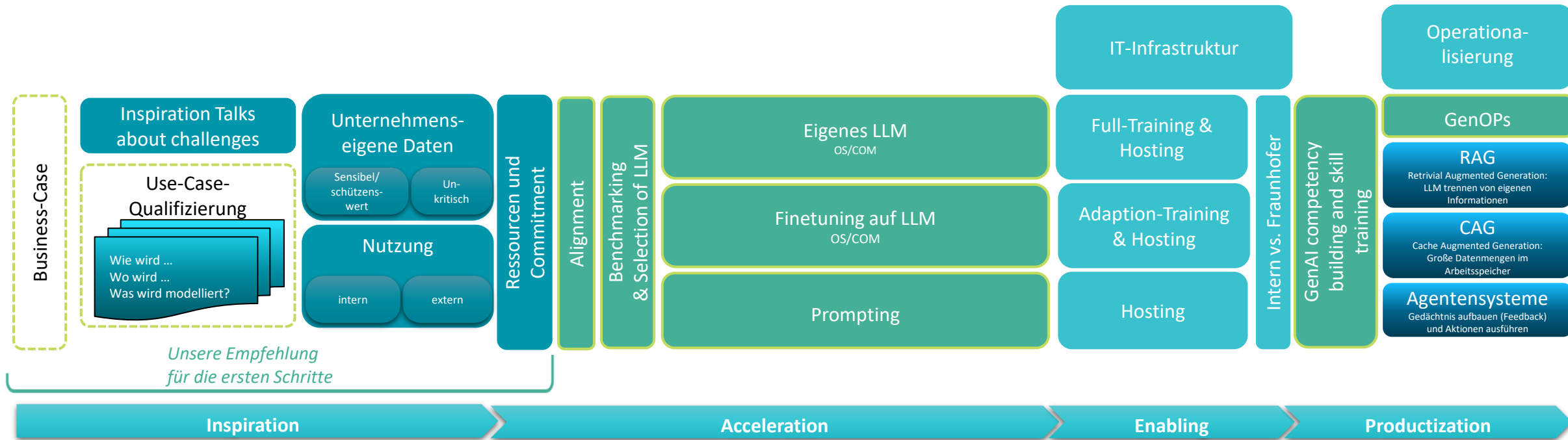
Я подаю заявку на участь в інтеграційному курсі, тому що хочу швидше адаптуватися до життя в Німеччині, вивчити мову та культуру цієї країни. Я не знаходжусь в шкільній освіті та не брала участі в інтеграційних курсах, фінансованих державою раніше.

Щодо фінансових питань, я подаю заявку на звільнення від сплати за курс, оскільки мої доходи обмежені, і я перебуваю в складній фінансовій ситуації. Я готова надати всі необхідні документи для підтвердження мого статусу.

Я повністю усвідомлю важливість надання правдивої інформації та зобов'язуюся негайно повідомляти про будь-які зміни моїх обставин. Я сподіваюсь на вашу підтримку та допомогу в цьому процесі."

Lösungsentwicklung basierend auf Generative AI & Foundation Models

Enabling GenAI



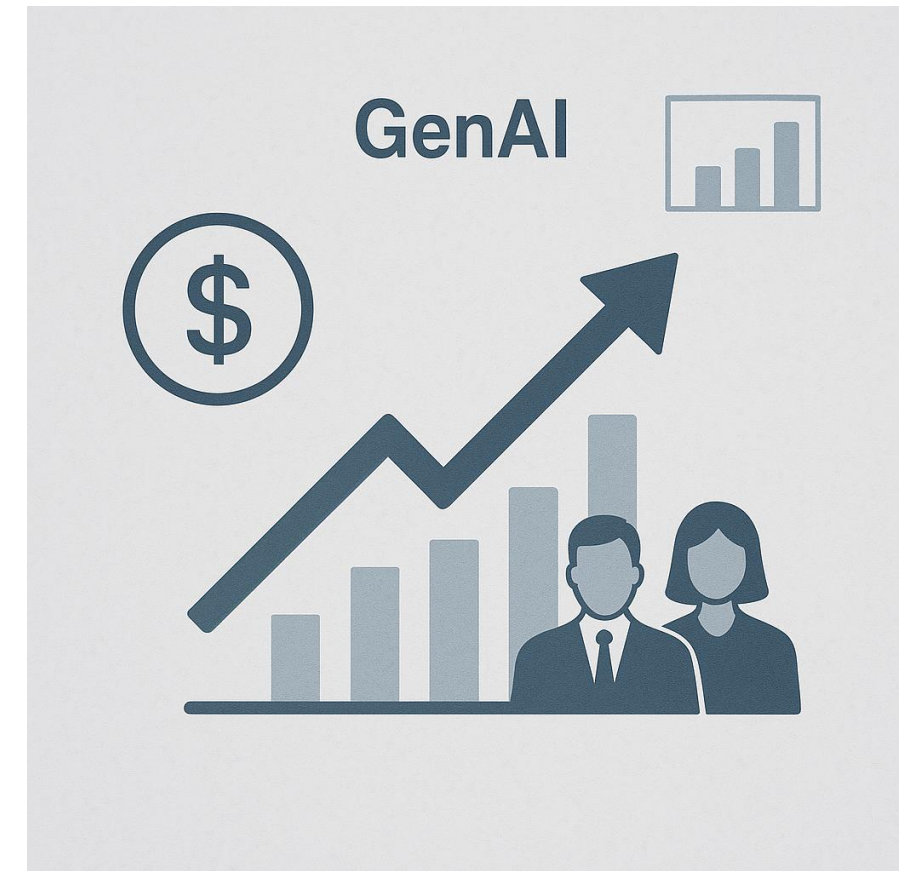
Take-home Message Teil 2

■ Mit Generativer/Agentischer KI ...

- eigene Produktivität erhöhen
- Umsatzpotentiale mit Produkten heben
- Branchenlösung entwickeln als neues Geschäftsmodell

■ Anleitung zum Glücklichen, wirtschaftliche Perspektive:

- Out-of-the-Box denken – über die eigenen Geschäftsbereiche hinaus blicken
- Leichtgewichtige Lösungen sind besser als keine



„Wer sich immer alle Türen offen hält, wird sein
Leben auf dem Flur verbringen.“

Kontakt



André Rauschert
Head of Division Computational AI
Tel. +49 351 4640-681
andre.rauschert@ivi.fraunhofer.de

Back up

Biowissenschaften &
Gesundheit



Logistik & Mobilität



Produktion &
Industrie 4.0



Energie &
Umwelt



Handel &
Finanzwesen



Sicherheit



Fraunhofer-Allianz Big Data und AI

Connecting People and Data

Mehr als 30 Fraunhofer Institute bündeln ihr branchen-übergreifendes Know-how für Big Data und KI und bieten Unternehmen:



Optimierung von **Prozessen und Produkten** mit Hilfe von Big Data und KI



Unterstützung auf dem Weg zur **Data Driven Company**



Ausbildung von Fach- und Führungskräften zu Data Scientists

Wie können Sie mit der Fraunhofer-Allianz zusammenarbeiten?



Beratung

- ✓ Strategieentwicklung
- ✓ Innovationsmanagement
- ✓ Prozess- und Strukturanalyse
- ✓ Anforderungsanalyse
- ✓ Technologieberatung
- ✓ Studien und Whitepaper



Technologieentwicklung und Implementierung

- ✓ Proof of Concept / Machbarkeitsstudien
- ✓ Prototypen
- ✓ Software
- ✓ Hardware
- ✓ Analysedienste
- ✓ Plattformen und Infrastrukturen



Qualifizierung

- ✓ Aus- und Weiterbildung
- ✓ Lernplattformen
- ✓ KI-Lab
- ✓ Wissenstransfer

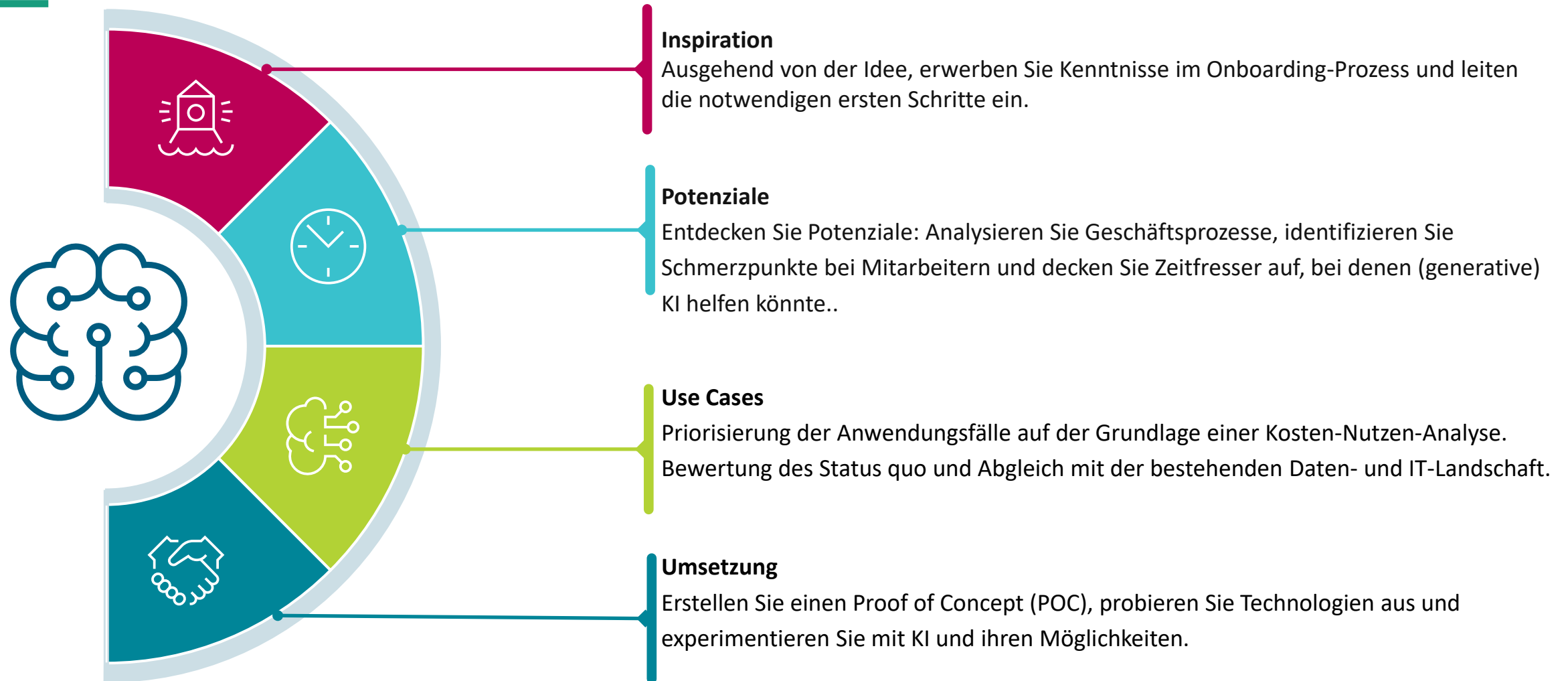


Qualitätssicherung

- ✓ KI-Absicherung und Zertifizierung
- ✓ KI-Readiness Check
- ✓ Schutz kritischer Infrastrukturen
- ✓ Risiko- und Vulnerabilitätsanalysen

Bilder: fizkes - stock.adobe.com (l.), Fraunhofer IAIS (l. m.), MIND AND I - stock.adobe.com (r. m), Gorodenkoff - stock.adobe.com (r.)

Möglichkeiten der Zusammenarbeit - Erste Schritte



Bildquelle: © jay - stock.adobe.com / Fraunhofer IAIS.

Einsatz von großen Sprachmodellen in Unternehmen

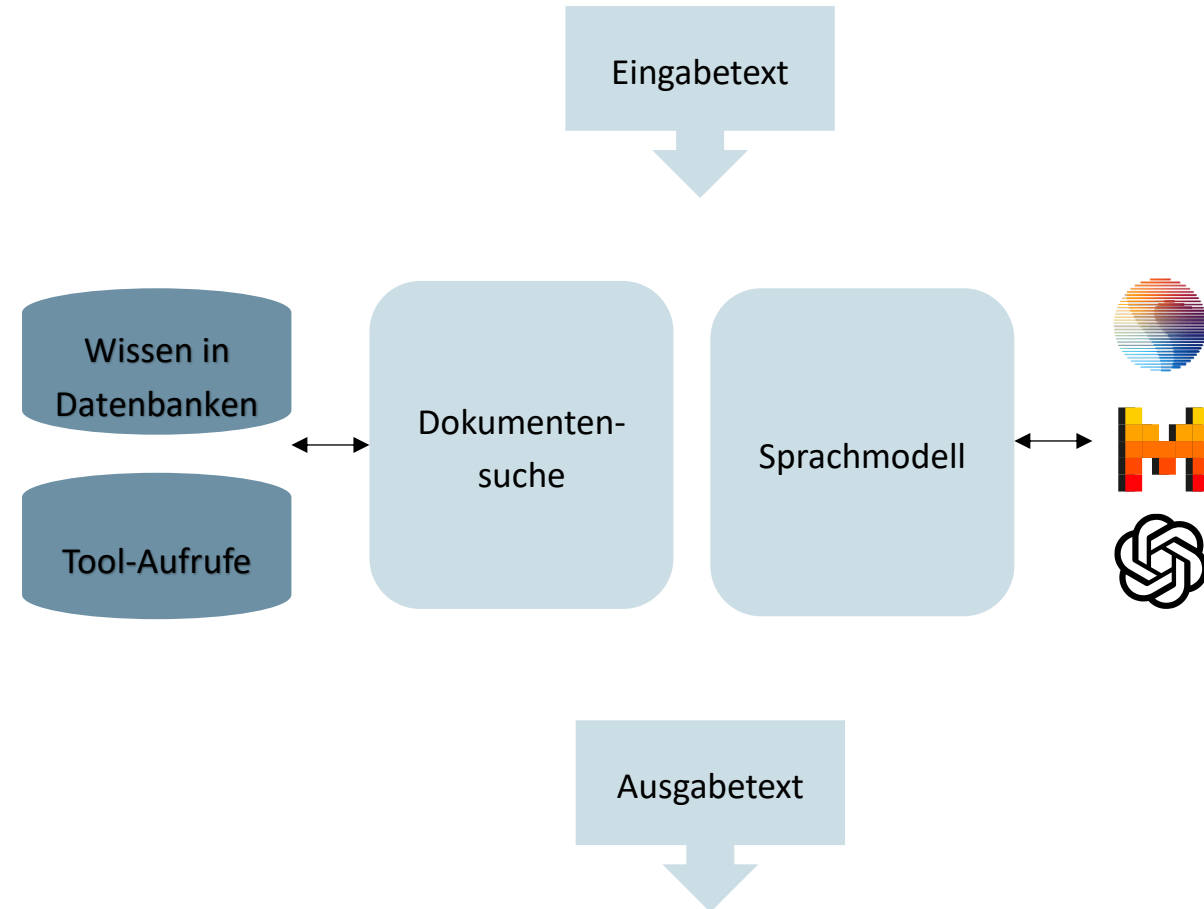
Dialoge zu kundenspezifischen Dokumenten – Industriemodelle

Industrialisierung von LLM-Technologie, **Ziele:**

- Natürlichsprachliche Dialoge zu internen und externen Dokumenten
- Fokus auf verlässliche Informationen, Produktsicherheit und Faktentreue
- unangemessene, falsche oder gefährliche Sprachausgabe verhindern
- Flexibilität bzgl. Hintergrundwissen und Sprache

Kombination aus mehreren Machine Learning Modellen:

- KI Suche: Modelle für das Finden von Antworten in Kundenwissen (Machine Reading Comprehension, RAG)
- KI Sprachmodell: Generative Language Model (GLM)
- Antwortselektion zur Auswahl der besten Antwort für den Nutzer



Established training program of the Fraunhofer-Alliance Big Data and AI

Certified (up to 5 days)
Our certification program qualifies you as a certified Data Scientist in three stages. Or become a data scientist specialized in one of the big data or AI fields.



Specific (individually)
In our industry-specific training courses, you can familiarize yourself with the most important concepts, methods and tools for your specific application area.



For Beginners or Experts (1-2 days)
Start as a Beginner, Manager or Expert with introduction courses and broaden your expertise. Use the method-specific training courses to deepen this knowledge. (Micro certificates soon available)

10 institutes, 29 modules and ca. 2.100 training days per year

Werden Sie fit für Big-Data- und KI-Projekte

Auswahl aus einer Vielzahl an methoden- und branchenspezifischen Schulungen

Kompakteinstiege			Vertiefungskurse (Foundation Level)		Zertifizierungskurse		
Prompting Grundlagen für Generative KI	Fortgeschrittene Prompting Techniken für Generative KI	Moderne KI-Tools im Führungsalltag	Time Series Analysis	Cognitive Robotics	Data Scientist Basic Level	Data Scientist specialized in MLOps	Data Scientist specialized in Production
In-house Bereitstellung und Betrieb von generativen KI-Modellen	Künstliche Intelligenz	Healthcare Prompting	Image and Video Understanding	Deep Learning for Text Mining	Data Scientist specialized in Big Data Analytics	Data Scientist specialized in Trustworthy AI	Data Scientist specialized in Process Mining
Large Language Models mit Hugging Face	Computer Vision	Machine Learning Operations (MLOps)	Machine Learning in Production Excite	Energy Data Analyst	Data Scientist specialized in Deep Learning and Generative AI	Data Scientist specialized in Assuring Safety	Data Scientist specialized in Quantum Machine Learning
Reinforcement Learning for Robotics and Automation					Data Scientist specialized in Machine Data Analytics	Data Scientist specialized in Data Management	Data Scientist specialized in Data Analytics
Maschinelles Lernen mit Python					Data Scientist specialized in Edge AI	Quantum Computing Professional - Basic	Certified Data Scientist Specialized in the EU AI Act
+ umfassendes Angebot zu AI Literacy Trainings							

www.bigdata-ai.fraunhofer.de/data-scientist