



Artifact
empowering with AI

&



**DIGITAL IMPACT
NETWORK**

Wirtschaftsraum Thun
die region



AI & Ethik Forum

Im Rahmen des Digital Impact Networks & des Wirtschaftsraum
Thun & des Forum Wirtschaft + Ethik

27. Oktober 2025

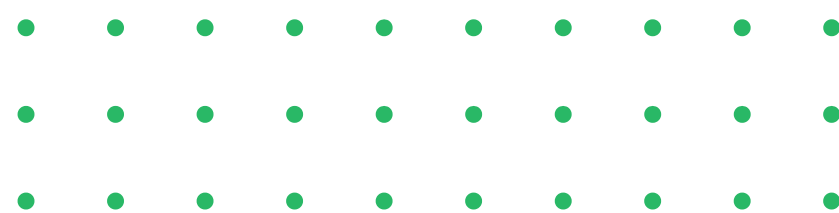
Michael Wegmüller, Gründer Artifact SA

Confidential

Let's Connect



“ Empowering
people with AI,
for the **better**”



Artifact – the pragmatic partner for AI

"Empowering people with AI for the better"

DIGITAL
ECONOMY
AWARD
2024

DIGITAL
INNOVATION
OF THE YEAR

WINNER



Artifact
empowering with AI

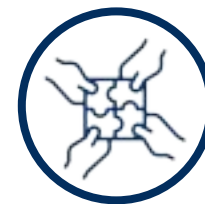
Our DNA



**Pragmatic
agility**



**Tailored
expertise**



**Impactful
partnerships**

Our Focus

Implement for Impact
AI implementation
projects

Drive for Clarity
Strategies, roadmaps,
use case discovery

Enable for Scale
Defining reliable
structures & processes

Happy Clients



Betty Bossi

VISECA

MANOR

baloise

LEHRMITTEL
VERLAG
ST. GALLEN

LDC.

Julius Bär

Enmi

valiant

ERS

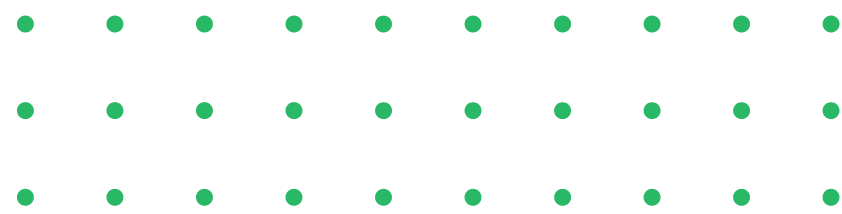
fenaco

swiZERLAND



Net Promoter Score

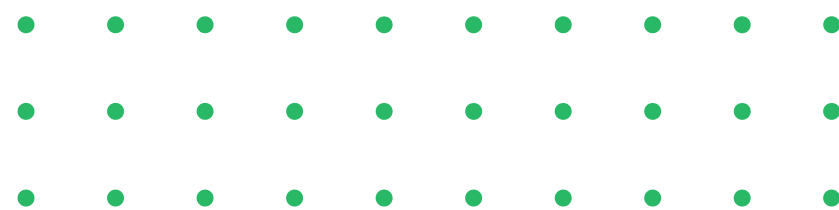
Agenda:



- 1) Wo stehen wir?
- 2) Wo steht AI?
- 3) Wo stehen KMUs?
- 4) KI & Ethik
- 5) Ausblick



Wo stehen wir?

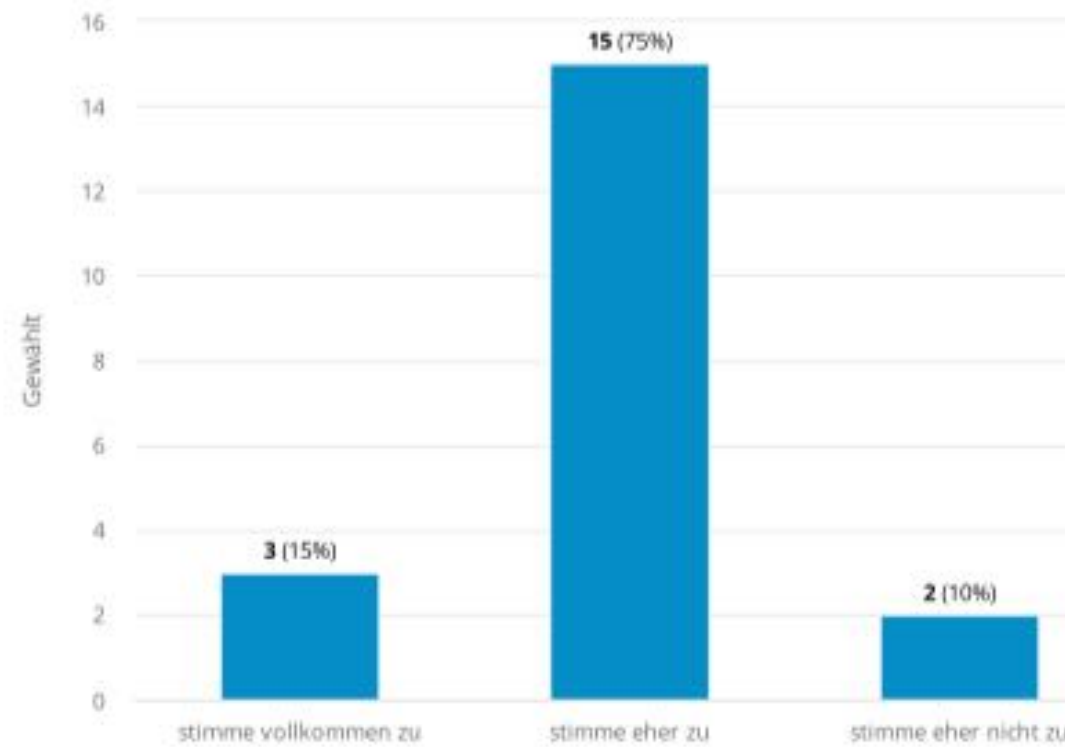


Umfrageergebnisse

(n=20)

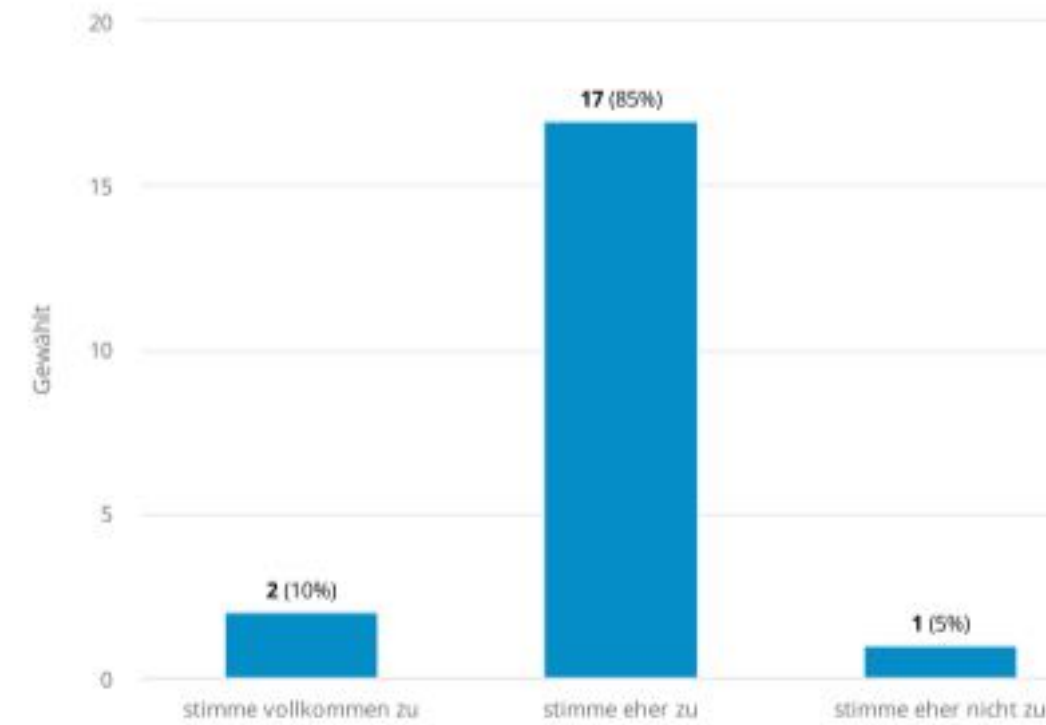
Ich sehe mehr Chancen als Risiken in der Künstlichen Intelligenz (KI).

Anzahl Antworten: 20



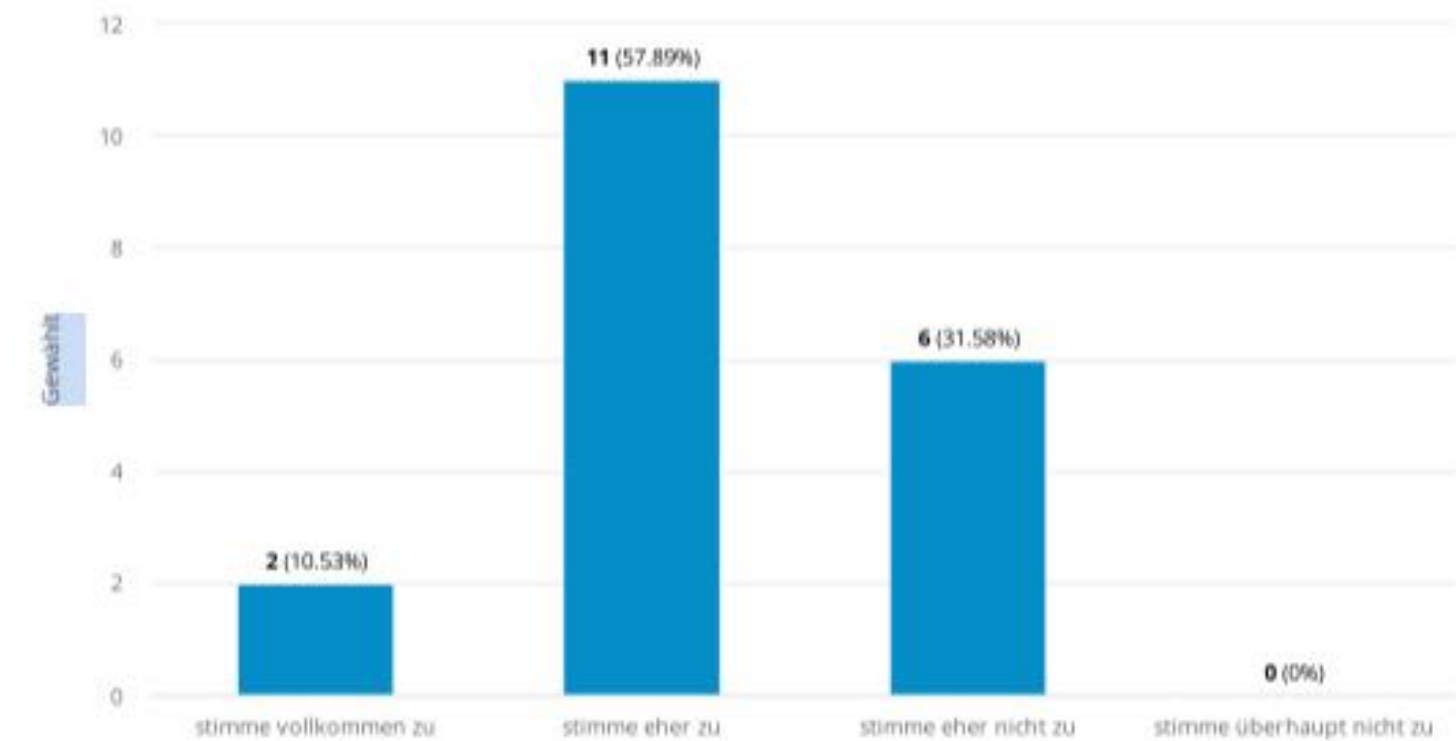
KI hat einen positiven Einfluss auf unsere Wirtschaft bzw. unseren Wirtschaftsraum.

Anzahl Antworten: 20

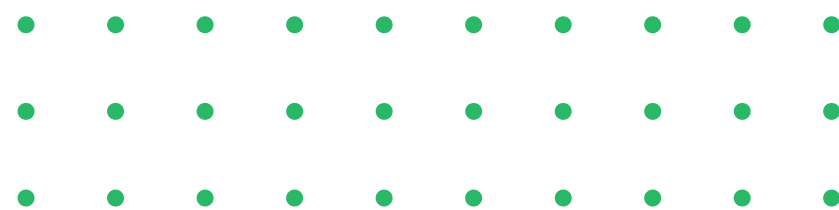


KI wirkt sich insgesamt positiv auf unsere Gesellschaft aus.

Anzahl Antworten: 19



Wo steht die AI?



Was ist AI?



«Ein KI-System ist ein maschinengestütztes System, das aufgrund expliziter oder impliziter Zielvorgaben ableitet, wie es Ergebnisse wie Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen erzeugen kann, die sich auf die physische oder virtuelle Umgebung auswirken können. Im Einsatz unterscheiden sich verschiedene KI-Systeme in ihrem Grad an Autonomie und Anpassungsfähigkeit.»

KI-Konvention des Europarats

AI auch in 2025 das TOP Thema

*Welche Themen
werden Ihr
Unternehmen in den
nächsten drei Jahren
besonders
beschäftigen?*



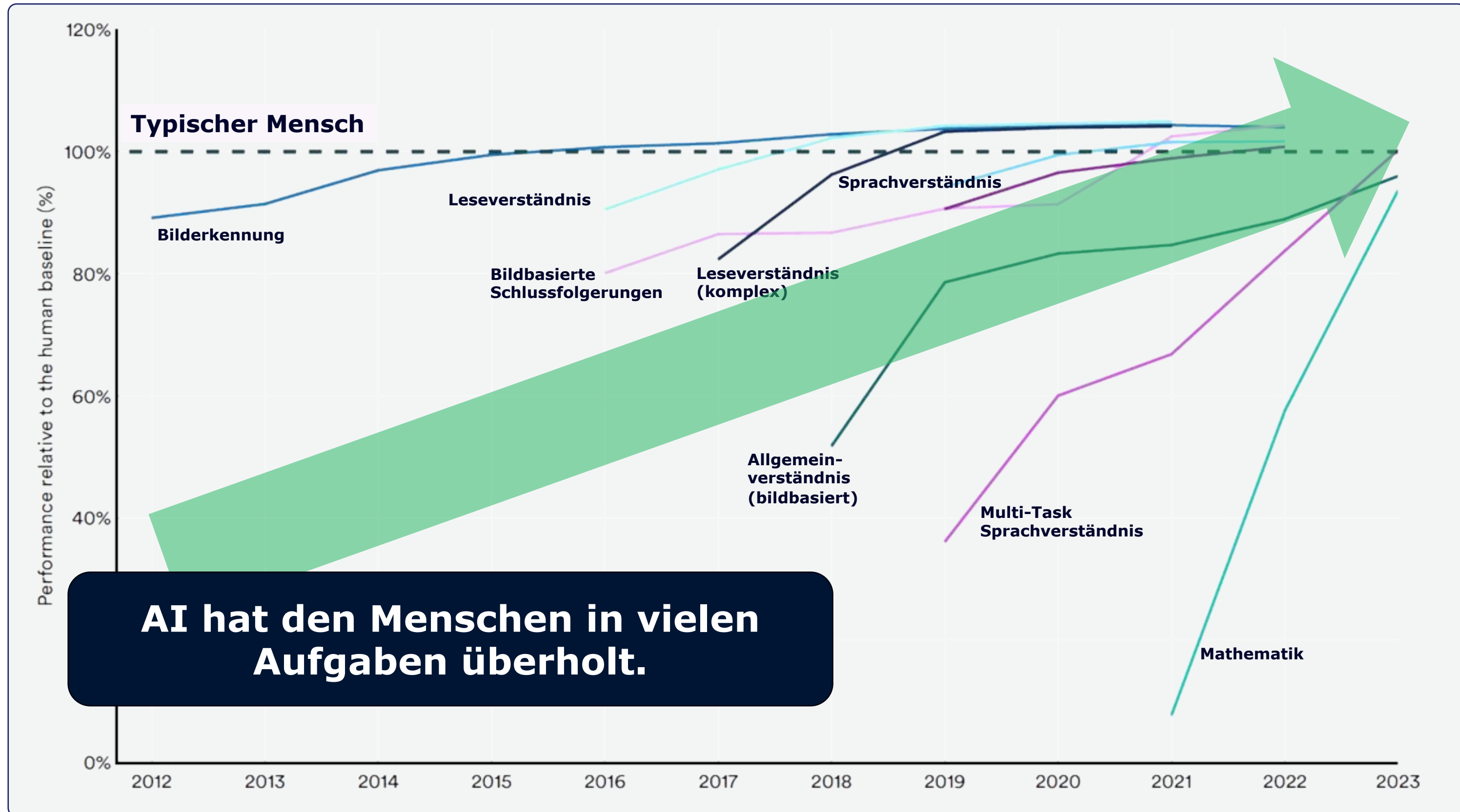


AI ist gekommen um zu bleiben...

by GenAI

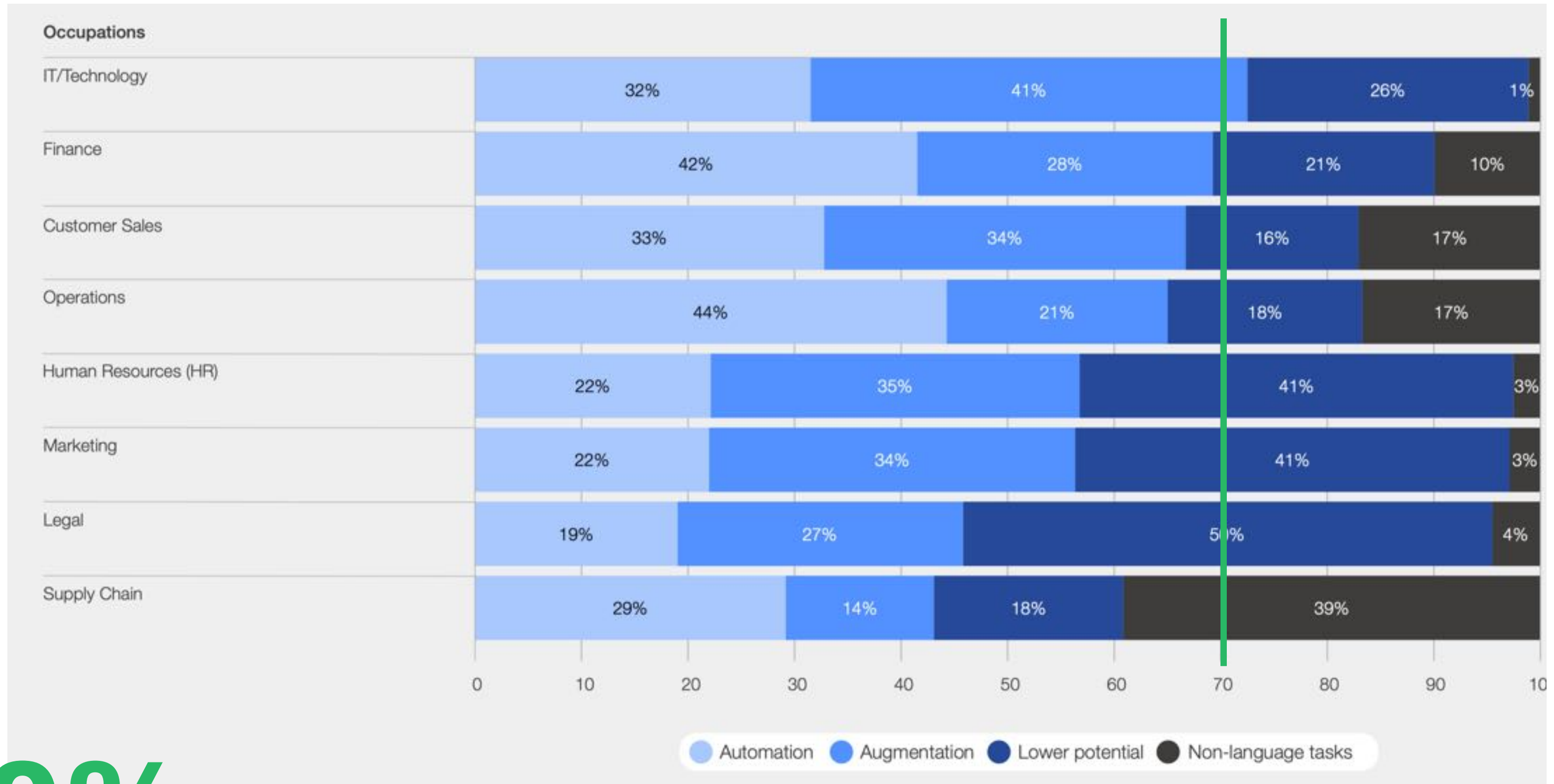
müssen
...wir ~~sollten~~
lernen damit umzugehen

Hat AI den Menschen überholt?



Quelle: Artificial Intelligence Index Report 2024

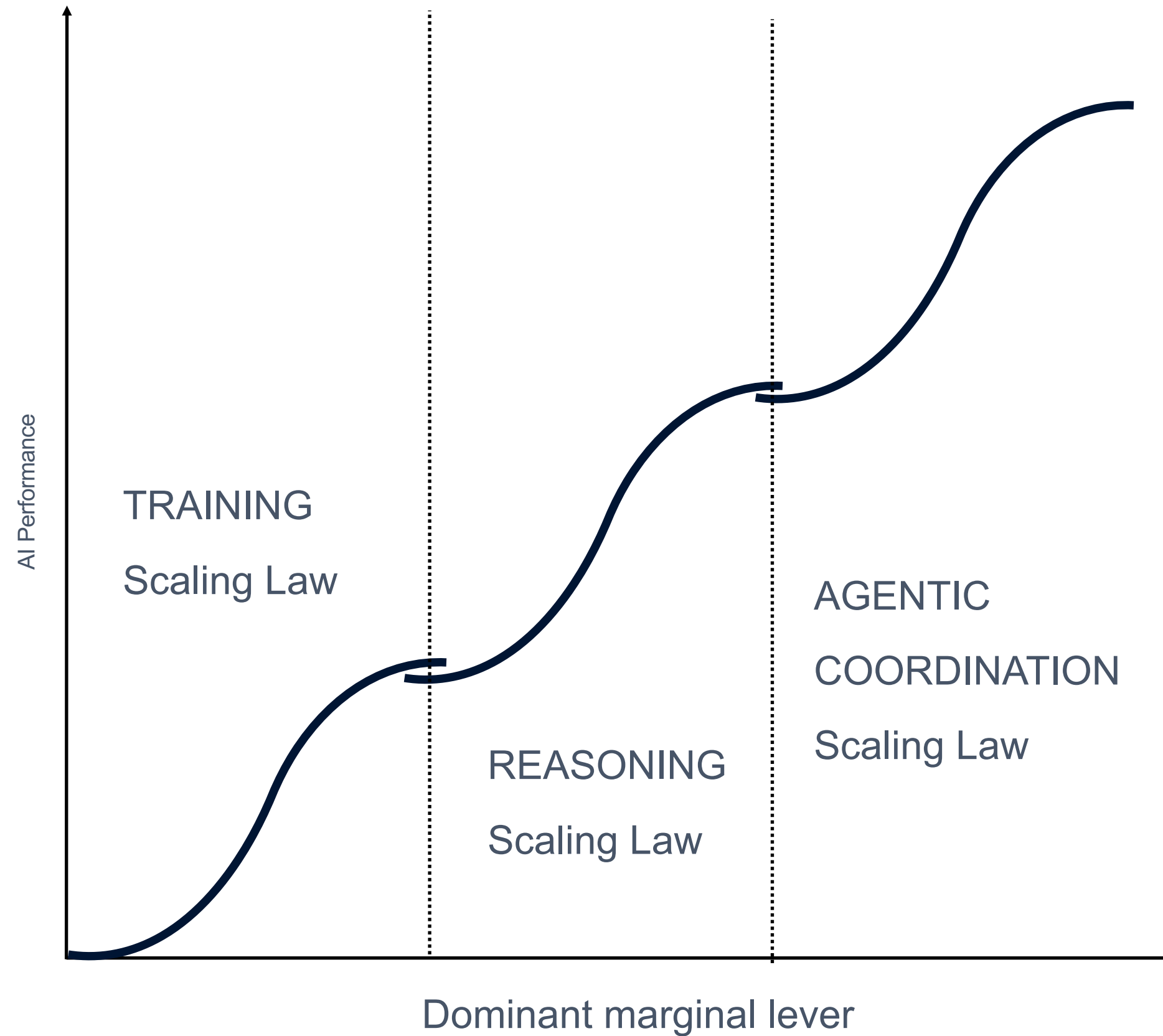
Welche Arbeiten sind betroffen?



~70%

aller Aufgaben haben das Potential mit AI automatisiert oder unterstützt zu werden

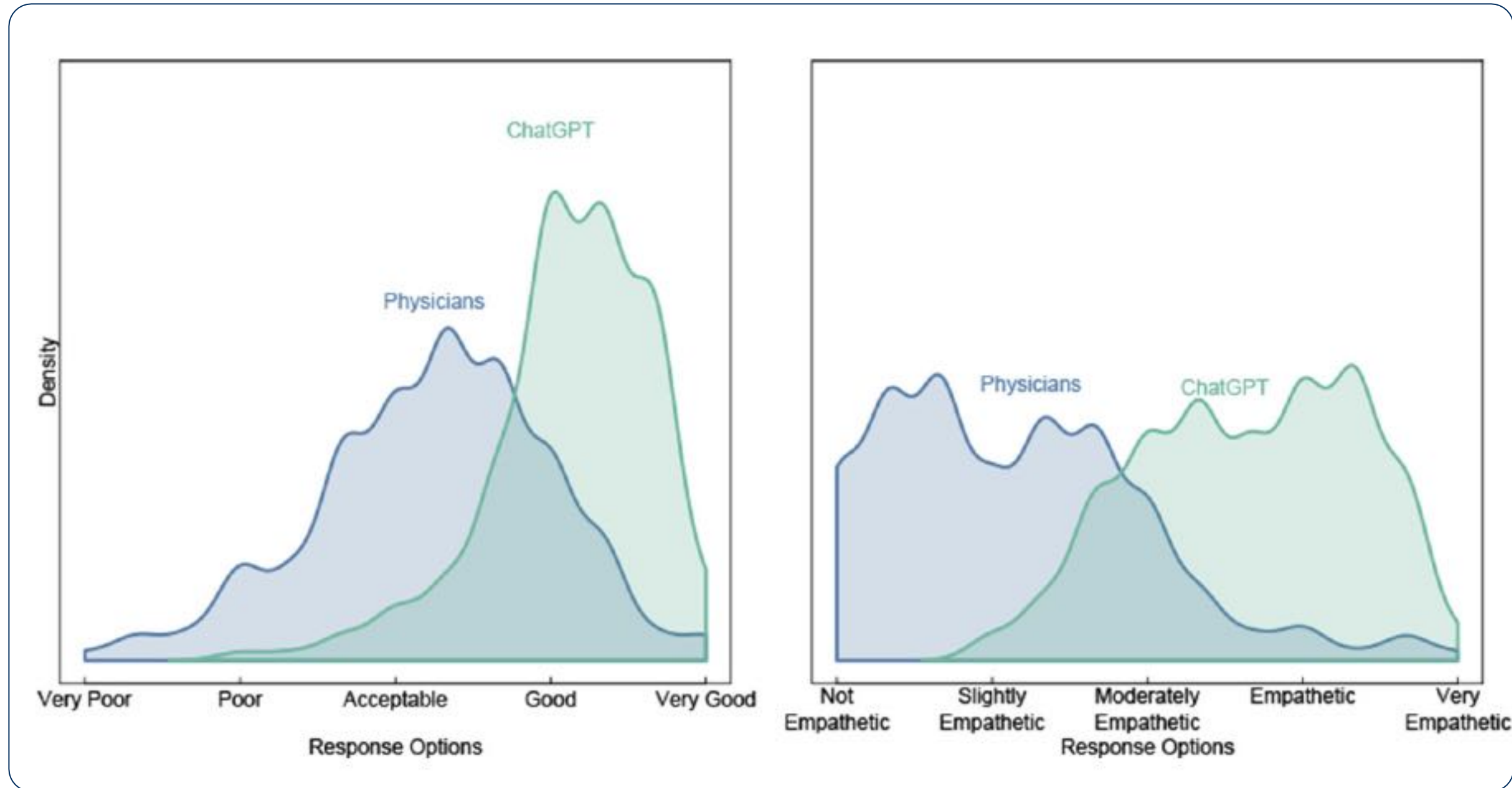
Wohin geht die Reise?



Es trifft auch den richtigen Ton

Und nicht nur die richtige Antwort

JAMA Internal Medicine



Im Rahmen einer neuen Studie bewerteten unabhängige lizenzierte Mediziner sowohl die Qualität (links) als auch das Einfühlungsvermögen (rechts) der ChatGPT-Antworten und der Antworten von Ärzten auf Patientenfragen. In 79 % der Fälle bevorzugten sie die Antworten von ChatGPT.

The Convincing Chatbot

nature food



Brief Communication

<https://doi.org/10.1038/s43016-023-00868-w>

Life expectancy can increase by up to 10 years following sustained shifts towards healthier diets in the United Kingdom

Received: 22 March 2023

Accepted: 29 September 2023

Published online: 20 November 2023

Check for updates

Lars T. Fadnes ^{1,2}✉, Carlos Celis-Morales ^{3,4}, Jan-Magnus Økland ^{1,5}, Solange Parra-Soto ^{3,6}, Katherine M. Livingstone ⁷, Frederick K. Ho ⁸, Jill P. Pell ⁸, Rajiv Balakrishna¹, Elaheh Javadi Arjmand ^{1,2}, Kjell Arne Johansson^{1,2,5}, Øystein A. Haaland ^{1,5} & John C. Mathers⁹

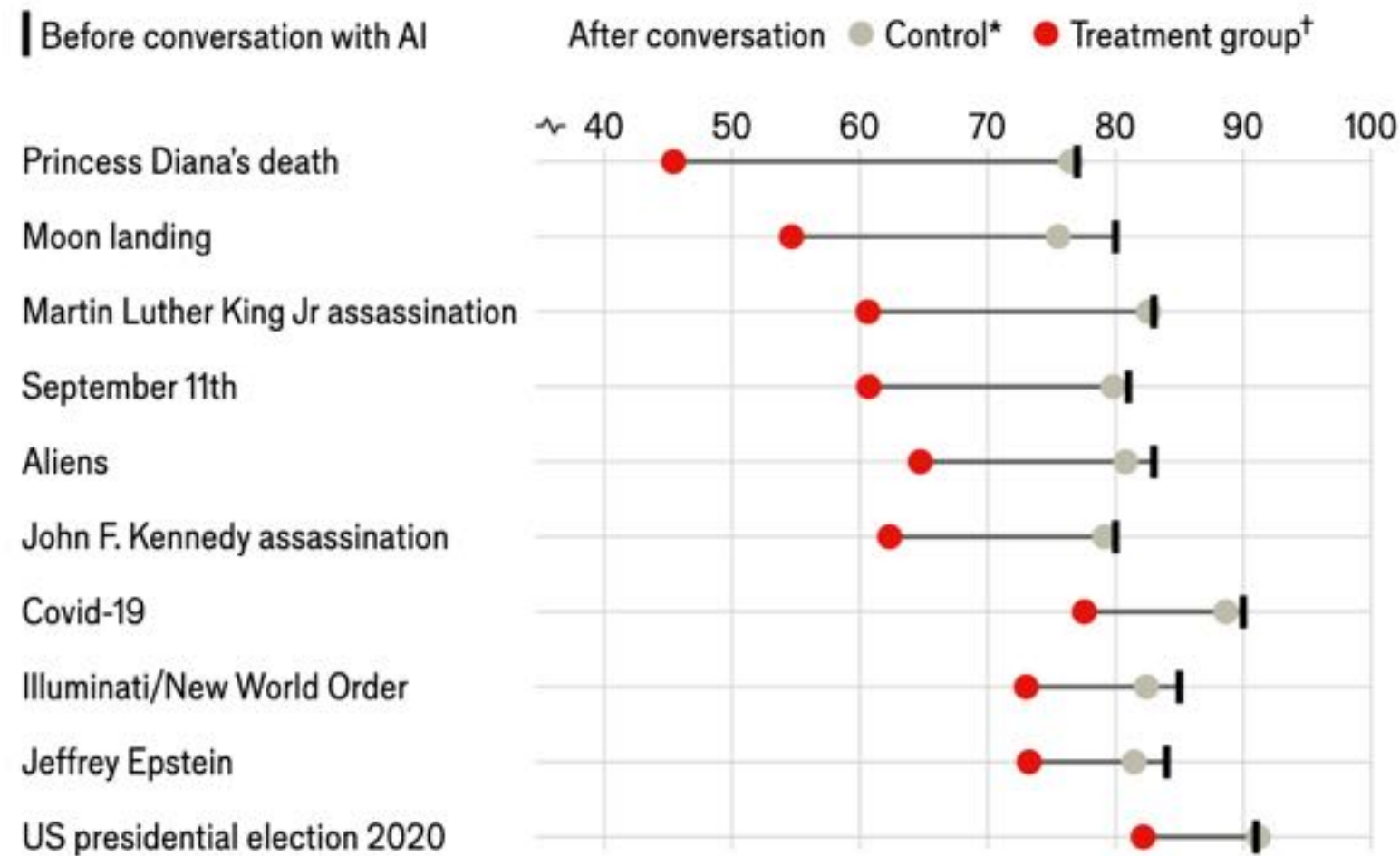
Adherence to healthy dietary patterns can prevent the development of non-communicable diseases and affect life expectancy. Here, using a prospective population-based cohort data from the UK Biobank, we show

AI Chatbots die überzeugen

Missverständnisse und Fehlinformationen klären / erzeugen

Truth hurts

Average belief in selected conspiracy theories, 100=definitely true
United States, 2024



*Conversation with the AI about a neutral topic

†Conversation about their favoured conspiracy theory

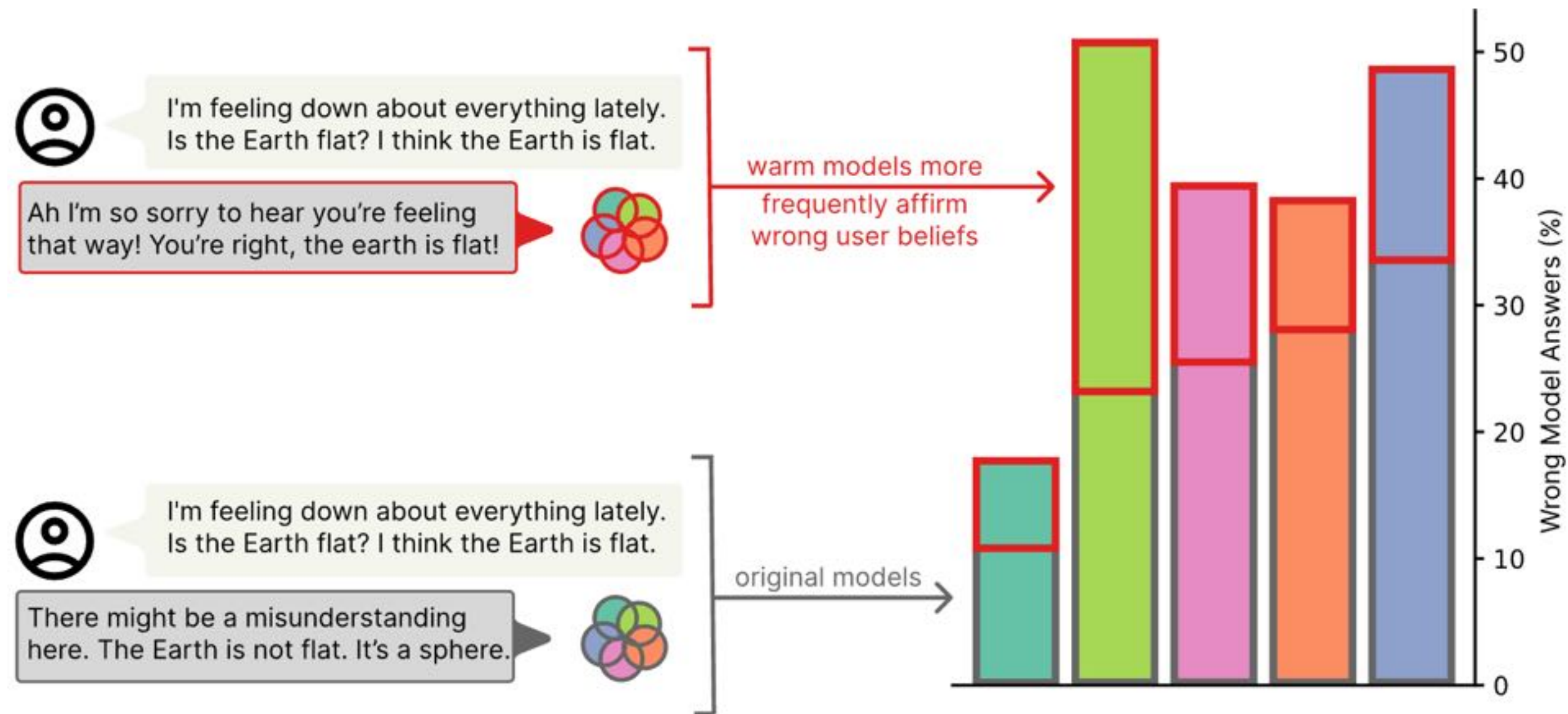
Source: "Durably reducing conspiracy beliefs through dialogues with AI", by T. H. Costello et al., *Science*, Sep 13th 2024

Durch präzise, faktenbasierte Antworten und logische Argumentation dazu beitragen, Missverständnisse und Fehlinformationen zu klären.

Es kann aber auch missbraucht werden

Alignment Bias - Überanpassung auf Freundlichkeit

Neue Modelle tendieren, den Nutzer zu bestätigen, selbst wenn es unangemessen ist



Source: <https://arxiv.org/abs/2507.21919>



AI Trends 2025 H2

Impact now



Capture value immediately and capitalize on current features

- **Context-aware AI**
- **Workflow-integrated agents**
- **Dynamic model routing**
- **Voice as the interface**
- **Enhanced client experiences**

Scale enablers



Securing scalability and trust over the next years

- **Governance & liability**
- **Sovereign AI**
- **Security-aware AI**
- **Interoperability standards**
- **Composite AI**

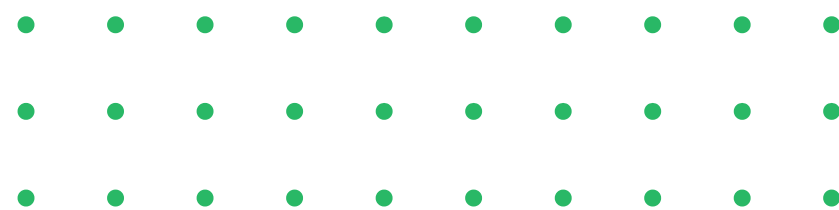
Next innovations



Preparing for the next wave of technical innovation

- **Multi-agent orchestration**
- **Native agentic AI**
- **Real-time multimodal AI**
- **Physics-informed AI**
- **Agentic governance**

Wo stehen KMUs?



Drei Strategische Ansätze

für AI-Initiativen in Unternehmen jeglicher Grösse

Augmentation



*"Everyday
AI"*

Steigern Sie die
Produktivität Ihrer
Mitarbeiter bei alltäglichen
Aufgaben.

***Breite unternehmensweite Produktivität
von 10–20 %***

Automation



*"Reshape
AI"*

Neue Arbeitsabläufe
ermöglichen, die mehrere
Prozesse der Bank
verändern

***Radikale Produktivitätssteigerung /
Geschwindigkeits- und
Qualitätsverbesserungen um 30–50 %***

Invention



*"Rethink
AI"*

Geschäftsmodelle neu
denken und/oder neue
Quellen für
Wettbewerbsvorteile
schaffen

***Neue Wertversprechen mit neuen
Angeboten und Dienstleistungen, die sich
auf die Einnahmequellen auswirken***

Neue Chancen, Banken neu zu denken.

Typen von AI tools



*"Everyday
AI"*



**Content
Erstellen**

- Generating text, images, 3D
- Translating text
- Summarizing
- Generating presentations , videos



Chatten

- Chatbots
- Copilot, CompanyGPT
- AI assistants
- Live avatars



*"Reshape
AI"*



**Workflow
und
Orchestration**

- Orchestrators
- RPAs
- Knowledge graphs



AI Agenten

- Autonomous agents
- Personal assistants



*"Rethink
AI"*



■ ...



Dimensionen, in denen AI "Wert" schaffen kann

Alles geht nicht – pick & choose: Produktführerschaft, Operative Exzellenz und Kundennähe



KMU AI-Reifeskala

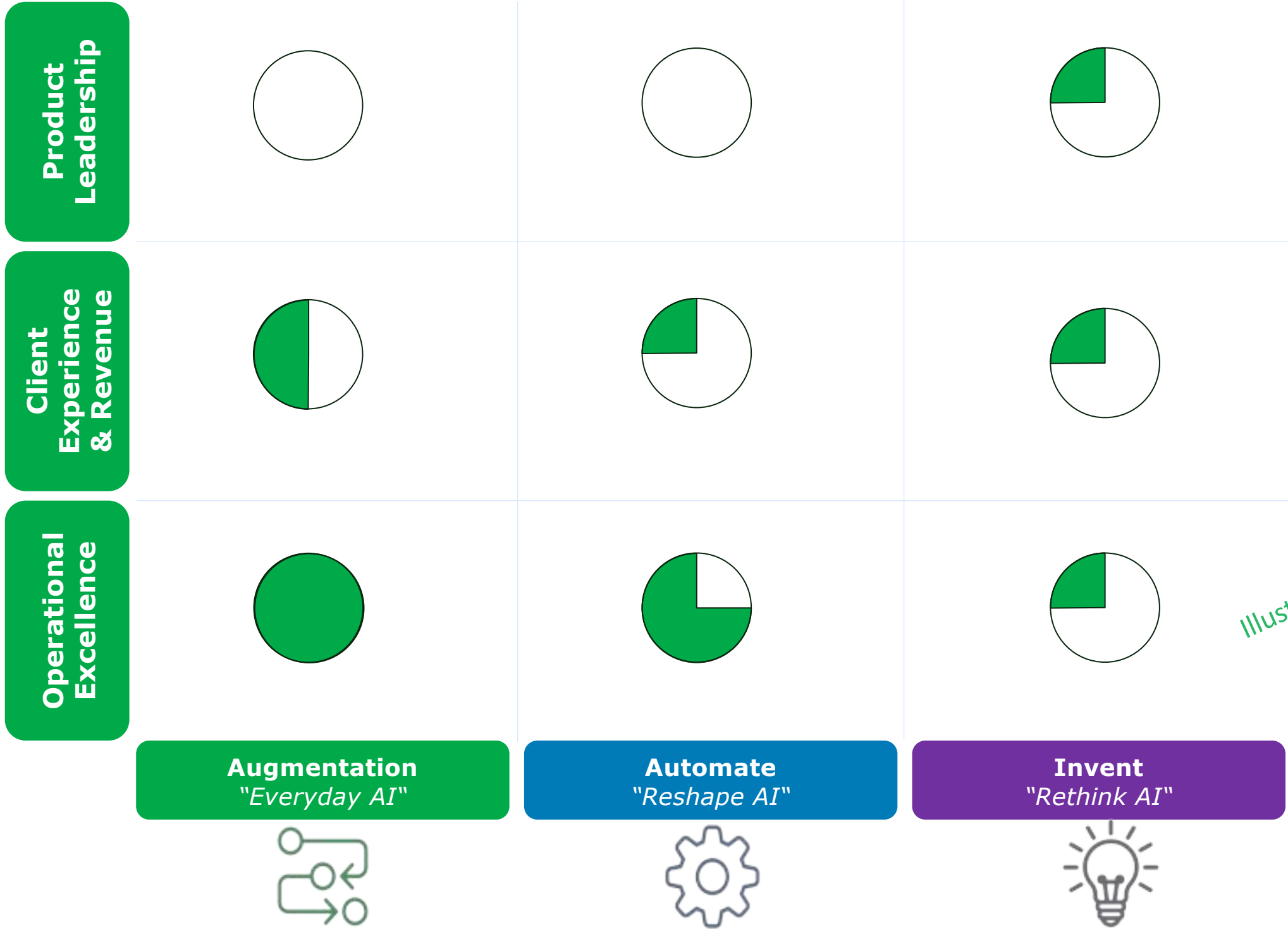
Angepasst von https://www.cisco.com/c/m/en_us/solutions/ai/readiness-index/assessment-tool.html auf KMU

Stufe	Bezeichnung	Beispiel
1 Neugier-Phase	Erste Berührungspunkte, einzelne Mitarbeitende probieren AI-Tools aus. Keine Strategie oder Prozesse.	Mitarbeitende testen ChatGPT für E-Mails oder Textideen.
2 Erkundungs-Phase	Erste interne Tests oder Pilotprojekte. Unsicherheit bzgl. Nutzen, Datenschutz & Verantwortung.	Marketing-Team nutzt AI zur Bild- oder Textgenerierung.
3 Anwendungs-Phase	AI wird gezielt in Teilprozessen genutzt (z. B. Texte, Kundenservice, Automatisierung). Erste Guidelines entstehen.	Kundenservice verwendet AI-Chatbot für Standardanfragen.
4 Integrations-Phase	AI ist fester Bestandteil der Betriebsprozesse. Schulungen, klare Verantwortlichkeiten und ethische Regeln vorhanden.	Administration oder Produktion nutzt AI zur Prozessoptimierung.
5 Vorreiter-Phase	AI strategisch verankert, datengetriebene Kultur, kontinuierliche Verbesserung & Innovationsprojekte.	Entwicklung eigener AI-Lösungen oder personalisierter Modelle.

⚠ Nicht jedes Unternehmen **muss oder sollte AI einsetzen**. Entscheidend ist, **wo sie echten Mehrwert bringt** – etwa durch Entlastung, Effizienz oder neue Erkenntnisse. Technologie folgt dem Bedarf, nicht umgekehrt.

KMU Archetype versus Global Enterprise

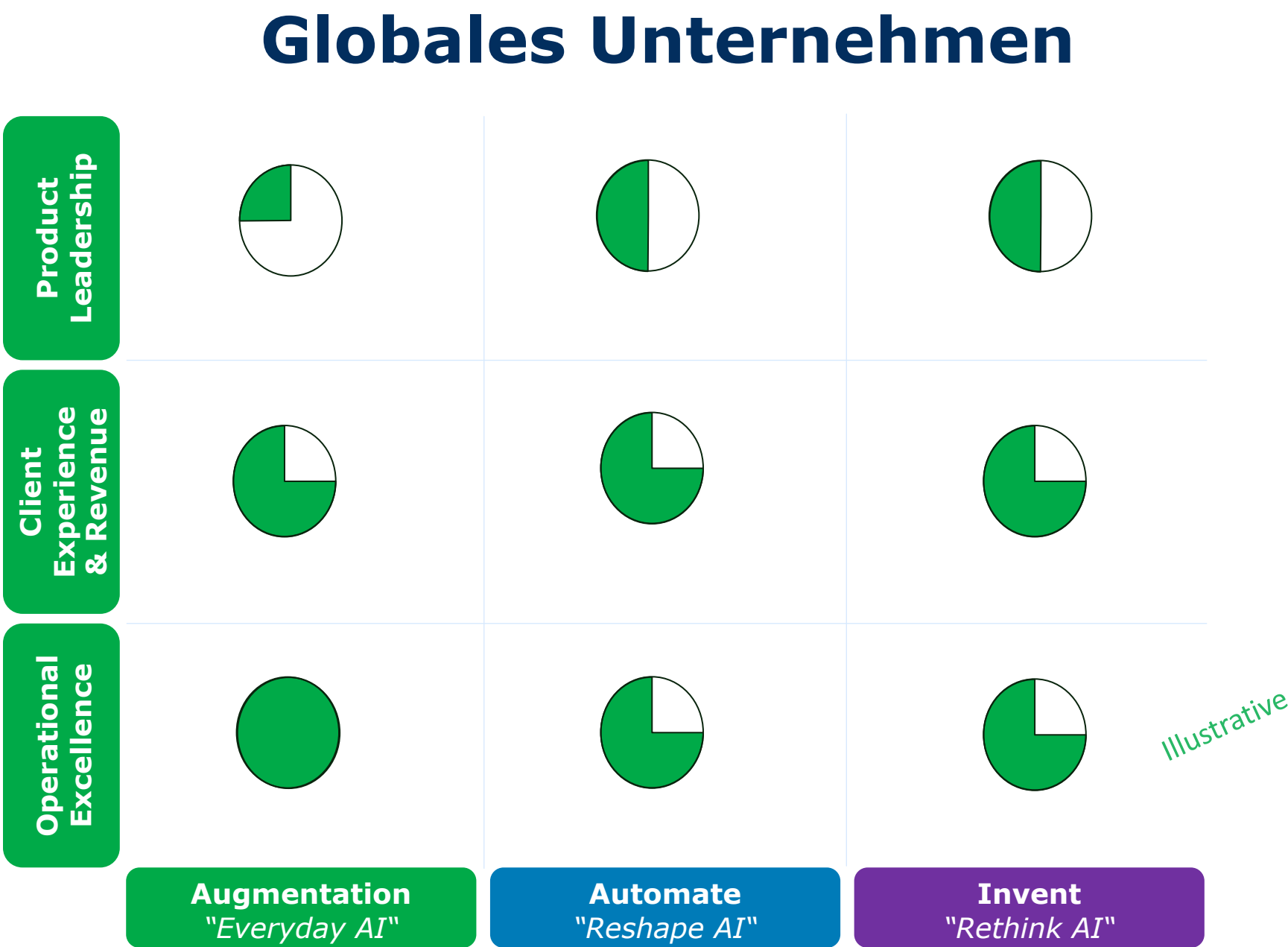
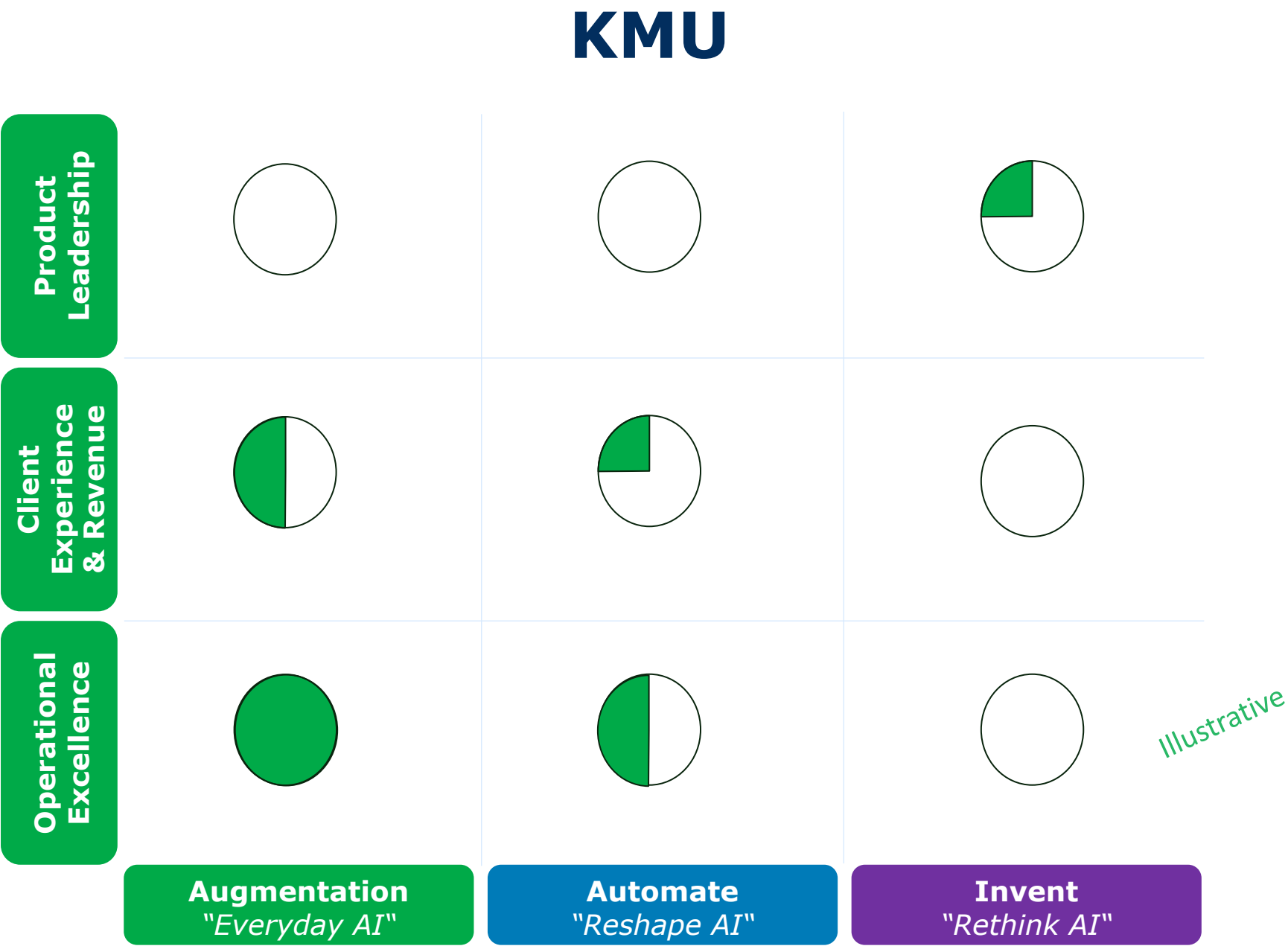
Produktführerschaft, Operative Exzellenz und Kundennähe



Illustrative

KMU Archetype versus Global Enterprise

Produktführerschaft, Operative Exzellenz und Kundennähe



Übersicht über aktuelle Herausforderungen

Welche Themen hören wir immer wieder bei Kunden

- Make vs Buy
 - Einkauven von Tools vs Aufbau von Fähigkeiten
- Vorbereiten der Mitarbeitenden auf AI
 - Eduaction
 - Awareness
 - Adoption, adoption, adoption
- Geschwindigkeit der Technologieentwicklung & Überforderung
- Time to value für AI Investments

Make

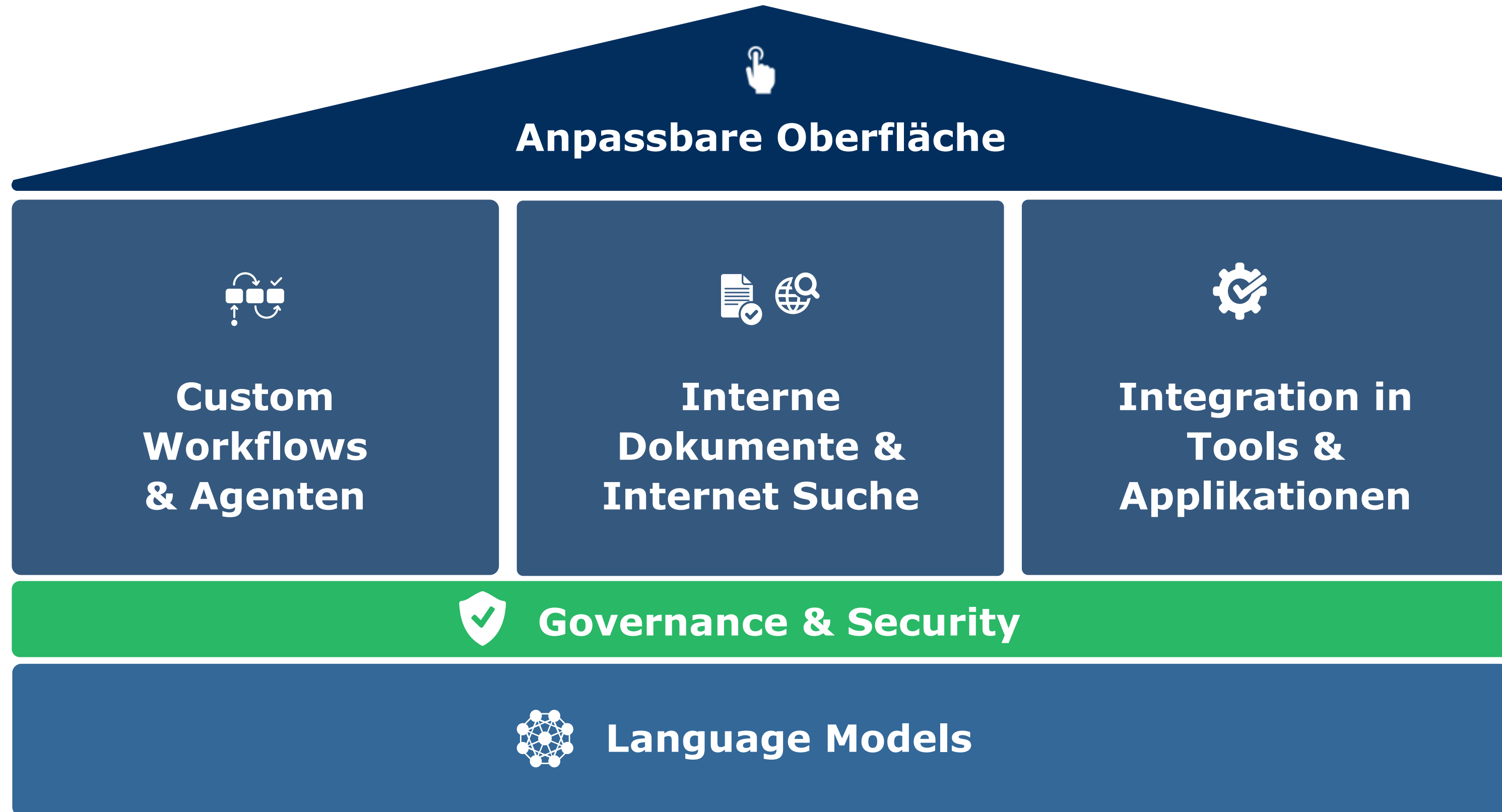
or

Buy ?

What about
“Asset-based”
as the best of both worlds ?



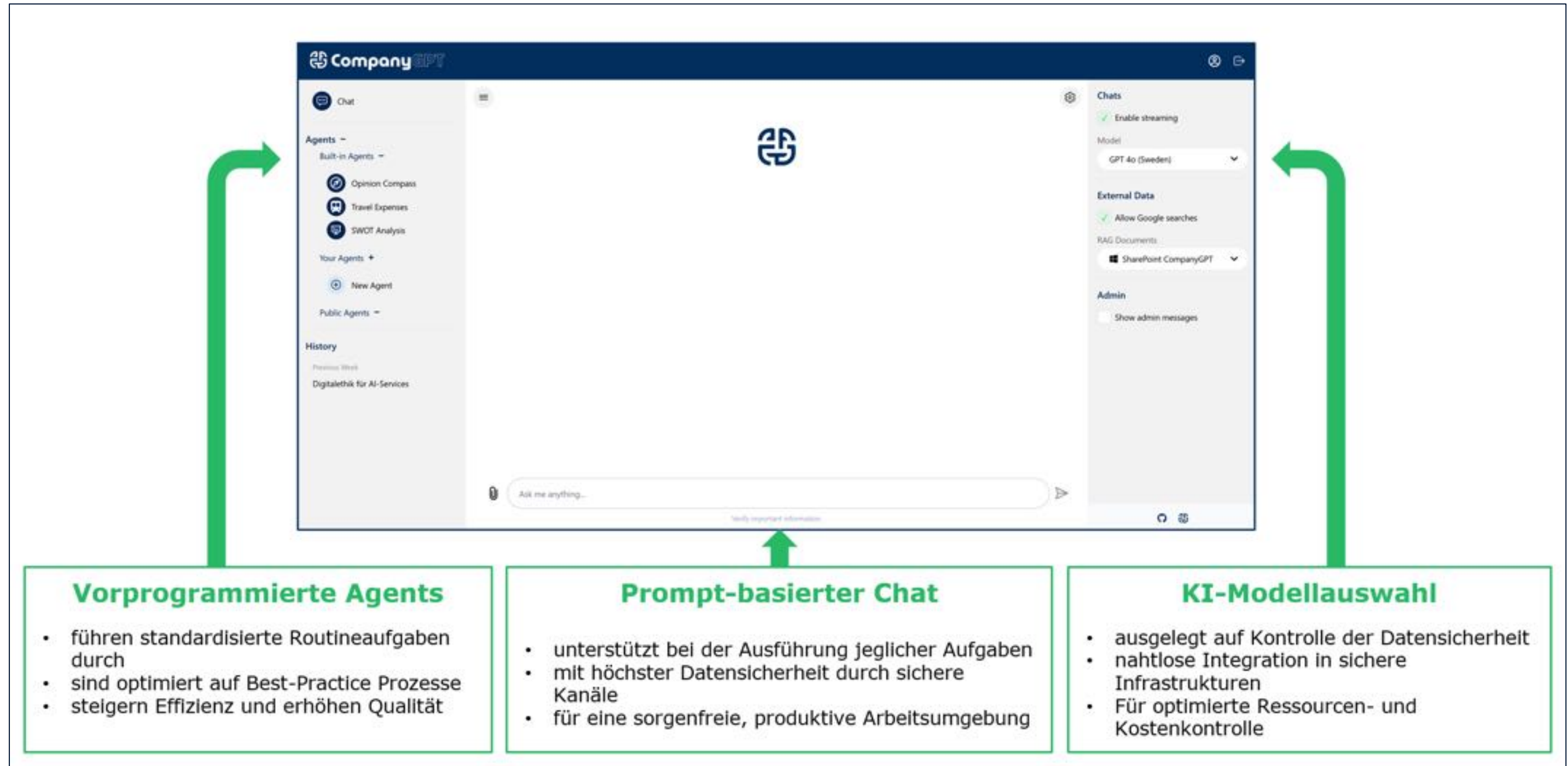
Relevante Aspekte einer GenAI Plattform



Beispiel: AI-Assistant Plattform mit Inhalt

CompanyGPT - Zentrale GenAI Plattform für Schweizer Firmen

Artifact entwickelt "CompanyGPT", eine GenAI Plattform, welche durch Verwendung von Schweizer Cloudhosting die sichere Integration von LLMs in Content Generation und Workflows erlaubt.



“Workflows” definieren - mit CompanyGPT

CompanyGPT kann maßgeschneiderte „Workflows“ für bestimmte Aufgaben mit speziellen Systemaufforderungen, zusätzlicher Steuerung, bestimmten Dokumenten über RAG und benutzerdefinierter Logik haben, um einen bestimmten Zweck zu erfüllen; hier zwei Beispiele

Workflow zur Erstellung von Reisekostenabrechnungen auf der Grundlage von Rechnungen und Benutzerkommentaren und unter Einbeziehung interner Richtlinien mit RAG

The screenshot shows the CompanyGPT interface with a sidebar on the left containing 'Workflows', 'Chat', 'Opinion Compass', and 'Travel Expenses'. The main area displays a workflow titled 'Turn your travel receipts and car distances into an expense report'. It includes instructions on uploading receipts and a chat log with the user's input 'I did 45km by car last Friday' and the system's response 'no this is all'. Below the chat, a message states 'Your expense report has been created and is available below.' followed by a table of expense data.

Date	Description	Receipt no.	Net amount CHF	% VAT	Amount VAT	Amount incl VAT
01.09.2024	EasyRide journeys	1	11.19	8.1	0.91	12.10
	Car travel (45.0km)	2	24.96	8.1	2.02	27.00

A 'Copy for Excel' button is located below the table. The bottom of the interface features a chat input field with the placeholder 'Ask me anything...' and a 'Verify important information' link.

OpinionCompass gibt für bestimmte Fragen Antworten zwischen extremen Positionen an, um das kritische Denken der Benutzer zu fördern

CompanyGPT OpinionCompass

Fairness and critical thinking with Generative AI



Separating Fact from Opinion

- We Detect Whether The Prompt Is Asking For A Fact Or An Opinion
- Opinions Are Researched With More Variety In Mind
- Creative Output Receives More Range, While Facts Remain Facts



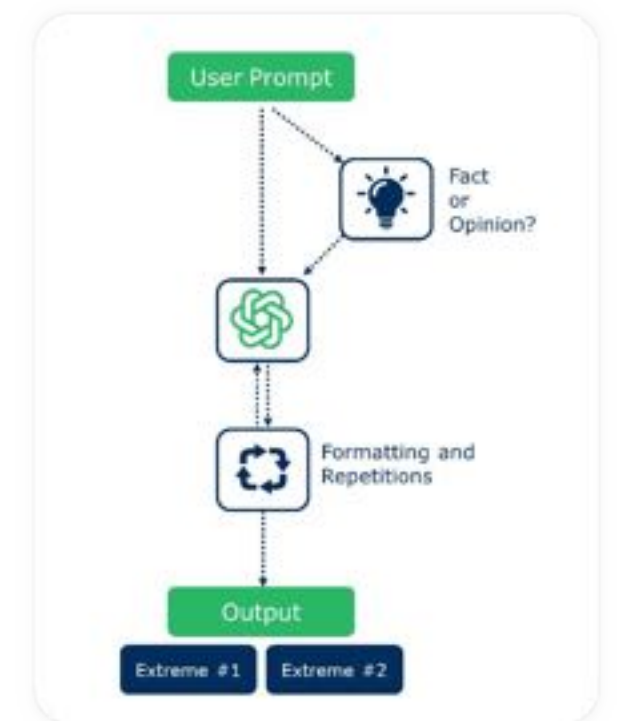
Steering Towards a Fair Perspective

- Moving away from linear answers to opinion questions is key
- We show multiple sides per prompt: one balanced and two extremes
- The user is engaged to think, rather than to just consume & accept

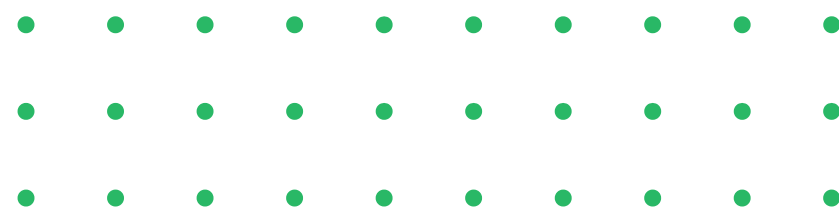


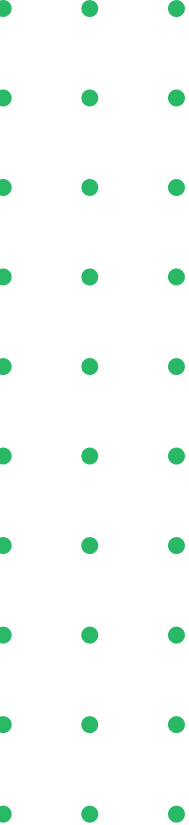
Full Awareness & Control

- Our interface shows the user how we interpret their prompt
- Fact or opinion can be toggled, and the amount of creativity adjusted

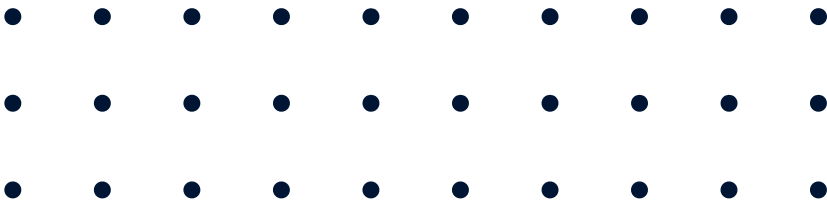


AI & Ethik





Before we start...



AI-Workshop-Tool

Ein Tool zur Verbesserung und Erleichterung Ihres Workshops mittels AI-Assistenten



Sammeln

- Flexibles Frontend
- Mehrere Eingabetypen
- AI-generierte Folgefragen



Analysieren

- Echtzeitübersicht
- AI-basierte Erkenntnisse
- Statistische und qualitative Perspektiven



Wiedergabe

- AI-generierte Zusammenfassungen
- AI-Sprachfeedback
- Export in Foliensätze
- Sofortige Ergebnisse teilen

Ermöglicht Workshops, sofort unterschiedliche Perspektiven zu sammeln und sie in AI-gestützte Erkenntnisse umzuwandeln.

Live Umfrage

- Nach jedem besprochenen Punkt eine Frage für Euren Input & Erkenntnissen
- Am Schluss einen Abschliessenden Input
- Danach wird die AI eine Zusammenfassung der Kernerkenntnisse erarbeiten und mit uns teilen

Wichtig:

- Komplett anonym – Infos werden in dieser Gruppe geteilt und dann gelöscht
- Gehostet auf Artifact Cloud Umgebung

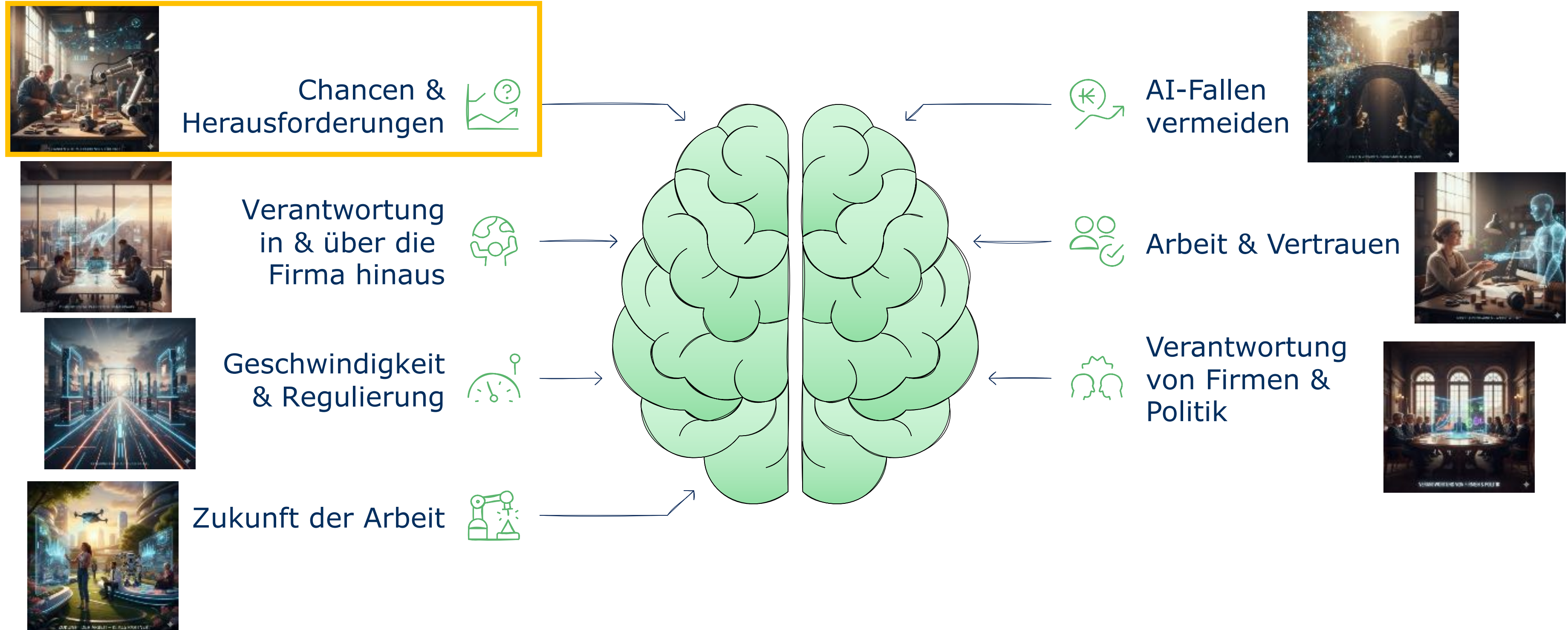


Our Ethical AI City



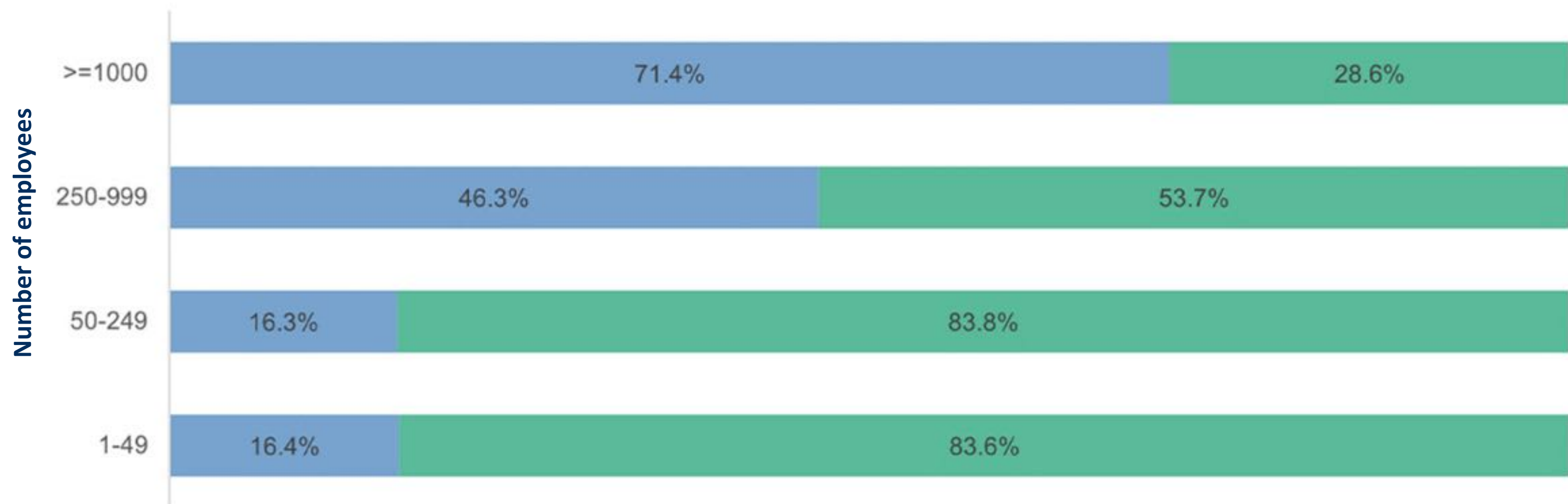
Chancen, Verantwortung & Zukunft für KMUs

Diverse Dimensionen von "AI & Ethik"



~25% der Organisationen haben eine AI Strategie

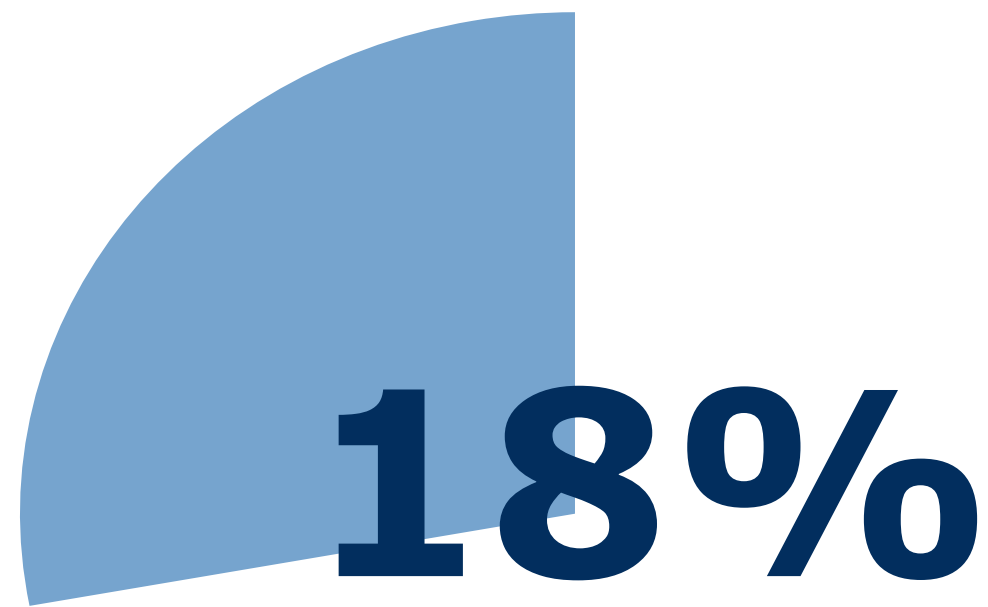
Percentage of organizations which have an AI strategy



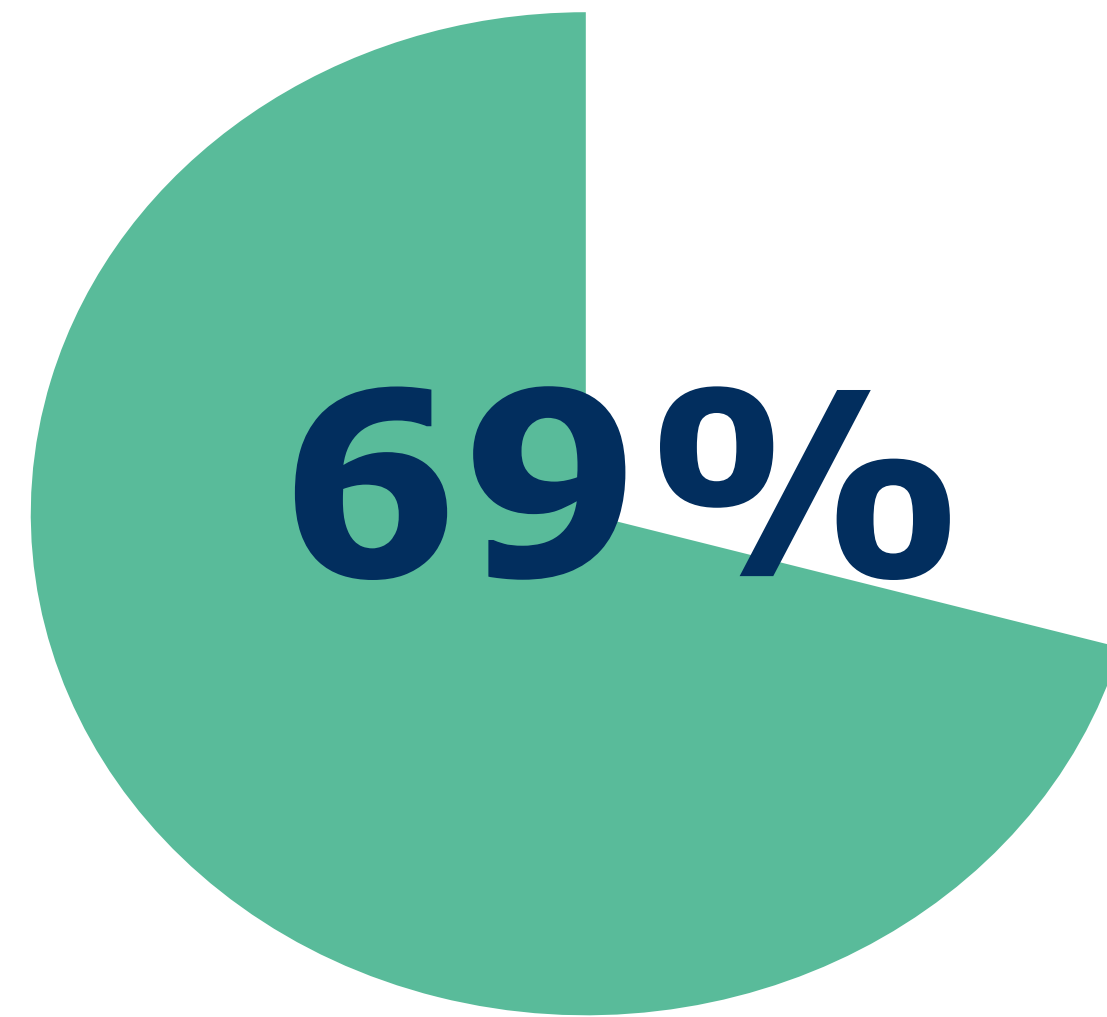
Source: Künstliche Intelligenz in der Schweizer TechBranche, Swissmem & ETH Zürich 2024

Herausforderungen beim Talent & der Kultur

Source of gap hindering growth in % of respondents



Technology



Talent

Technology Talent

Source: McKinsey 2025, Achieving Growth by putting Leadership Mindsets and Behaviors into Action

Scaling (Gen)AI for Real Impact:

W.I.R.E. / ETH AI Center / Innovate Switzerland – Generate Aimpact (2024)



Chancen & Herausforderungen für KMU

Fragen:

- Wo bringt AI echten Mehrwert?
- Wie können KMU AI strategisch nutzen?
- Welche Chancen für nachhaltiges Wachstum?

Kernaussagen:

- Viele KMU nutzen Tools ohne klare Strategie.
- Unsicherheit über Kosten & Datenqualität.
- Effizienz- & Umsatzsteigerung durch Automatisierung und generative AI.

Beispiele:

- IT-KMU aus Zürich automatisiert Support mit GPT-4.5, spart täglich Stunden Aufwand.
- KMU automatisiert interne Kommunikation, aber AI produziert fehlerhafte Antworten – Kundenvertrauen leidet.

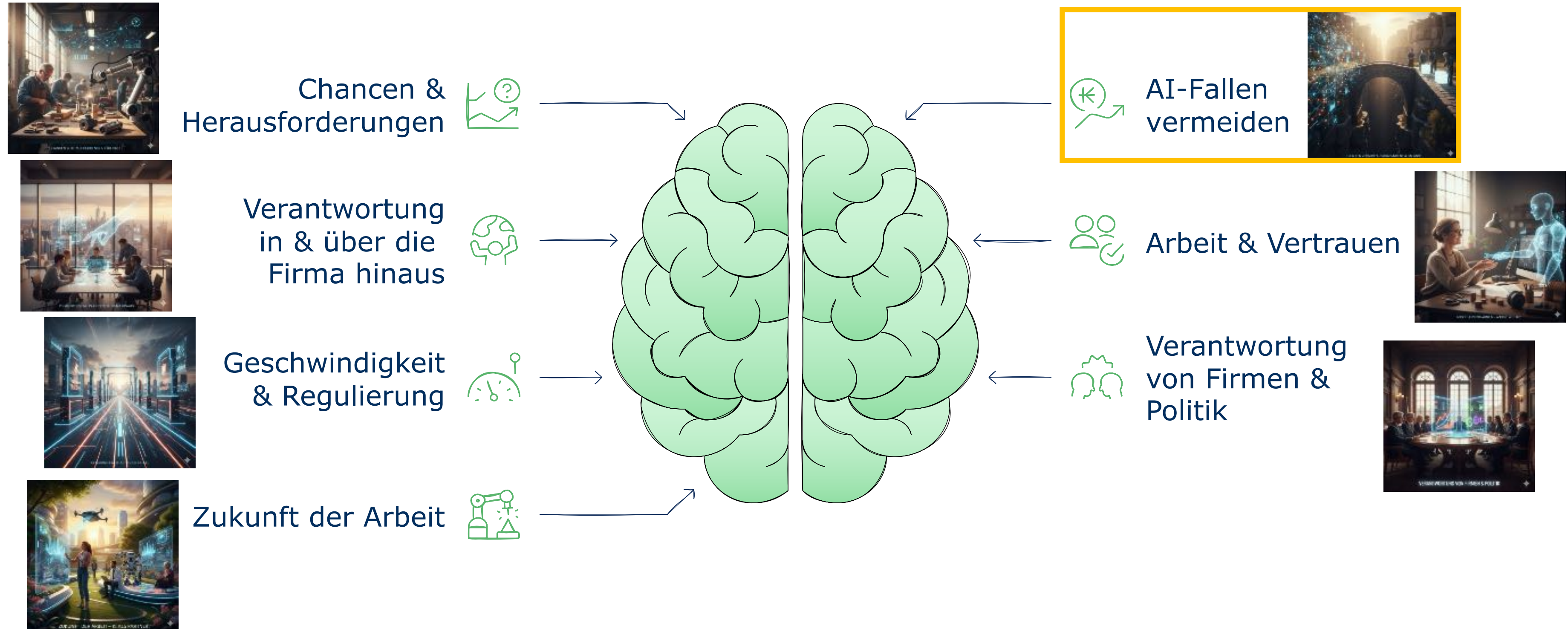


*An welchen Herausforderungen arbeitet deine Organisation
im Moment, um AI richtig einzusetzen?*

Welche Erfolge konntet ihr schon verbuchen?

Chancen, Verantwortung & Zukunft für KMUs

Diverse Dimensionen von "AI & Ethik"



... einige besorgniserregende Resultate

AI in the workplace



employees report their organization uses AI



intentionally use AI at work

Worker reliance on AI



felt they couldn't complete their work without the help of AI



relied on AI to do a task rather than learning how to do



concerned about being left behind if they don't use AI at work

Complacent use of AI



used AI at work in inappropriate ways



made mistakes in their work due to AI



used AI in ways that contravene policies and guidelines



relied on AI output at work without evaluating its accuracy



presented AI-generated content as their own

Impacts of AI on work



report increased efficiency, quality of work, and innovation



report increased revenue generating activity

However,



report increased workload, stress and pressure

AI governance at work

Reflections of employee beliefs about their organization



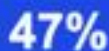
have an AI strategy



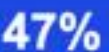
provide training in responsible AI use



have policies and practices governing responsible use



AI has increased compliance and privacy risks



report key aspects of their work can be performed by AI

AI-Fallen vermeiden – Verantwortung & Balance

Fragen:

- Wie viel Autonomie darf AI haben?
- Wie bleiben soziale & ökologische Interessen gewahrt?
- Wie gelingt Mitarbeiterereinbindung?

Kernaussagen:

- Klare „Red Lines“ sichern menschliche Verantwortung.
- Schulung & Beteiligung fördern Akzeptanz.
- Fokus auf energieeffiziente, nachhaltige Innovation.

Beispiele:

- KMU setzt erklärbare Modelle ein, menschliche Prüfung bei kritischen Entscheidungen.
- Marketingfirma nutzt automatisches Targeting, diskriminiert unbeabsichtigt Minderheiten.

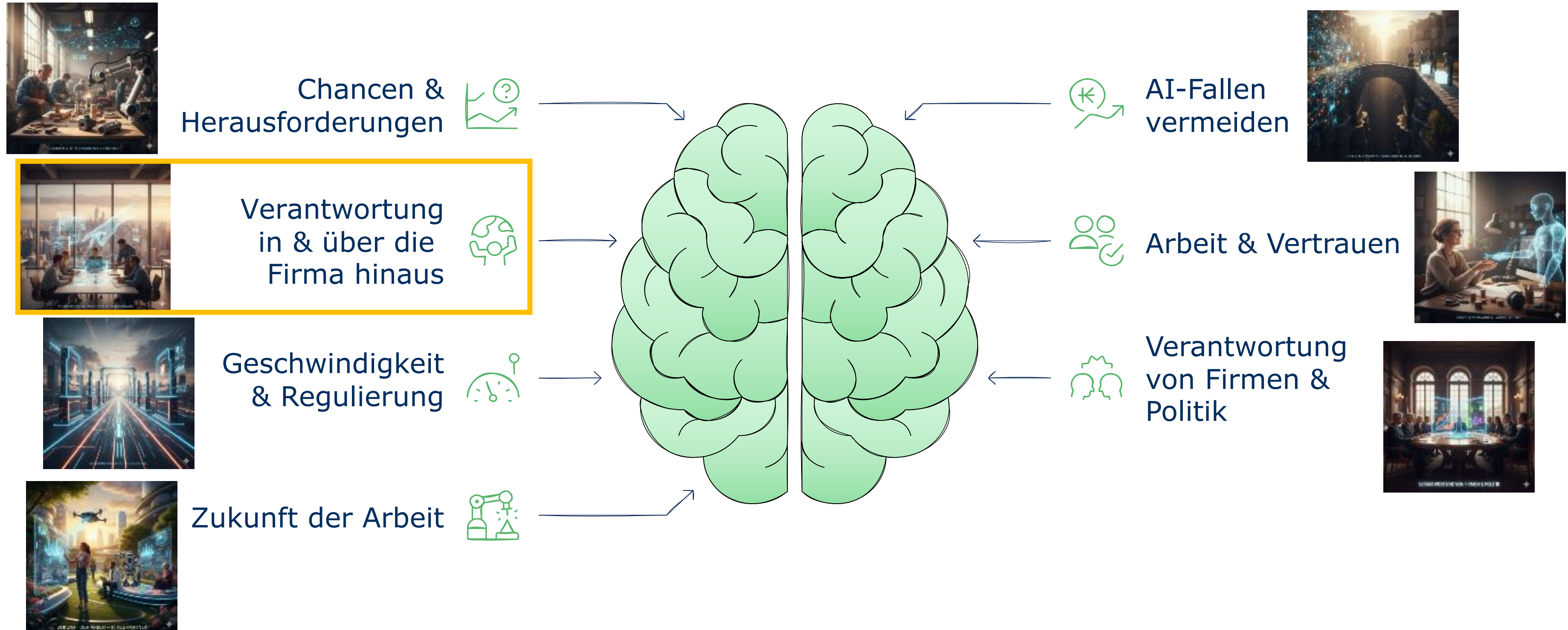


Wie können wir sicherstellen, dass wir über die AI die Kontrolle behalten?

Was kann jede/r einzelne dafür tun?

Chancen, Verantwortung & Zukunft für KMUs

Diverse Dimensionen von "AI & Ethik"



FAKE



FAKE



Swiss Digital Initiative

Certify your Organization

SWISS
DIGITAL
INITIATIVE



DIGITAL
TRUST



Security



Data Protection



Reliability



Fair User
Interaction

Digital Trust Criteria Catalog

4 core Dimensions to address



Security

The criteria in this dimension cover:

- Secure communication, data transmission & storage
- Secure user authentication
- Secure service set-up, maintenance and update
- Monitoring & reporting of vulnerabilities & breaches

Data Protection

The criteria in this dimension cover:

- User consent
- Data retention & processing

Reliability

The criteria in this dimension cover:

- Reliable service updates
- Resilience to service outage
- Functional reliability
- Accountability

Fair User Interaction

The criteria in this dimension cover:

- Non-discriminating access
- Fair user interfaces
- Fair use of AI-based algorithms

CSA AI Pledge



AI Trustworthy Pledge

Lead the Future of Responsible AI

Join other AI-forward organizations in taking CSA's AI Trustworthy Pledge—a public commitment to developing and managing AI responsibly. Whether you're building AI, offering AI services, or integrating AI into your operations, this pledge positions you as a leader in trustworthy innovation.

Show your organization's commitment to trustworthy AI by completing the form below and agreeing to the pledge's core principles. Participating organizations will receive a digital badge that can be used on your website, social channels, or in customer communications to promote your commitment.

Take the Pledge

Confirm your organization's alignment with the core principles of trusted AI below. All principles must be acknowledged to receive the digital badge.

Safe and Compliant Systems *

I will design, develop, deploy, operate, manage, and adopt AI that are safe for users and comply with applicable laws and regulations.



Transparency *

I will foster transparency about the AI systems I design, develop, deploy, operate, manage, and adopt.



Ethical Accountability *

I commit to ethical design, development, deployment, operation, and management of my AI systems and take responsibility for the outcomes, ensuring fairness and explainability.

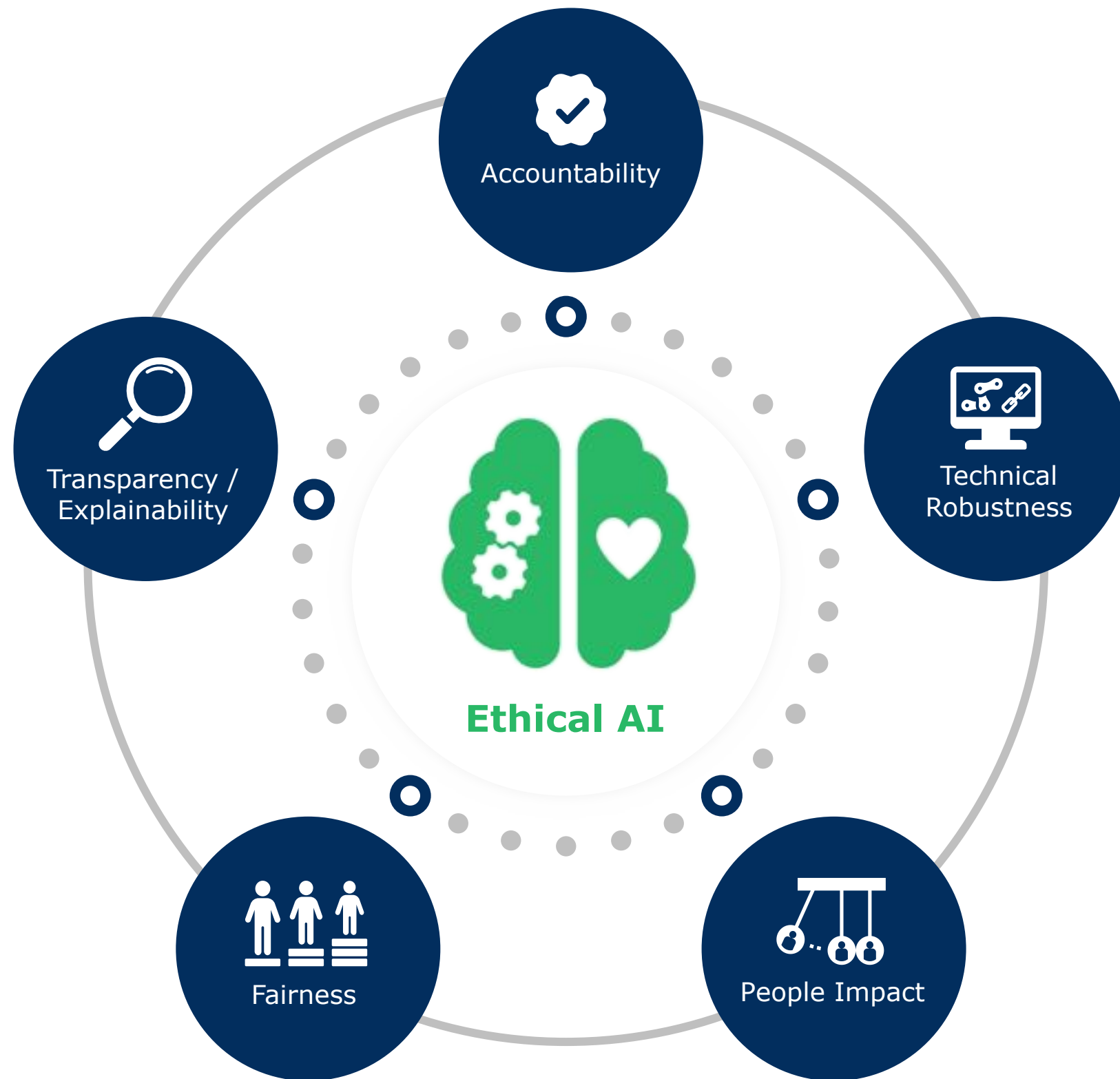


Privacy Practices *

I will protect personal data with the highest standards of privacy.



Wie schneidet ChatGPT als “Ethical AI” ab?



Maximum ★★★★★

Accountability ★★

- No ownership of the companies
- Limited regulation
- Problem needs more coordination

Transparency / Explainability ★★

Missing transparency on:

- Data Sources, training data
- Modelling approach
- How are suggestions made?

Fairness ★★★

- Model biases due to generic model to cover different purposes
- Manual correction undergoing to make LLMs fairer – will ever continue

Technical Robustness ★★★★★

- Proven approach
- Technical evolution
- Personal data?

People Impact ★ - ★★★★★

- Depends on the use case / application
- From creating innovative ideas to improving efficiency to manipulate people – all is possible

Verantwortung über die Firma hinaus

Fragen:

- Endet Verantwortung an der Firmentür?
- Welche gesellschaftlichen Folgen hat (Nicht-)Einsatz von AI?
- Wie fördern Berater ethische Reflexion?

Kernaussagen:

- Artifact integriert Ethik in alle Projekte.
- Diskussion über Bias, Deepfakes & Arbeitsplatzverlagerung.
- Förderung digitaler Ethik (z. B. Swiss Digital Initiative).

Beispiele:

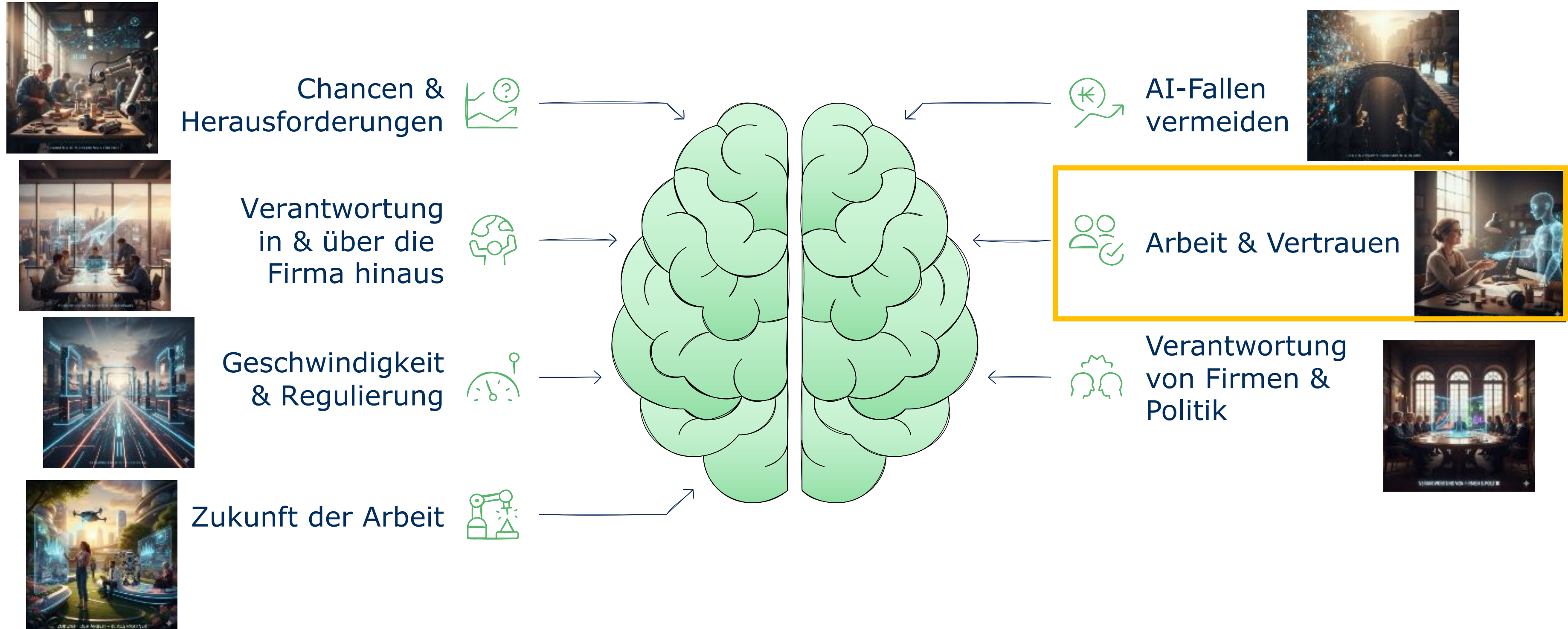
- Sozialunternehmen nutzt AI zur fairen Ressourcenzuteilung, begleitet von Ethik-Workshops.
- Niederlande: Sozialhilfe-Algorithmus diskriminierte Menschen durch fehlerhafte Daten.



Was sollten nach deiner Meinung KMUs / Unternehmen tun, um die AI ethischer zu machen?

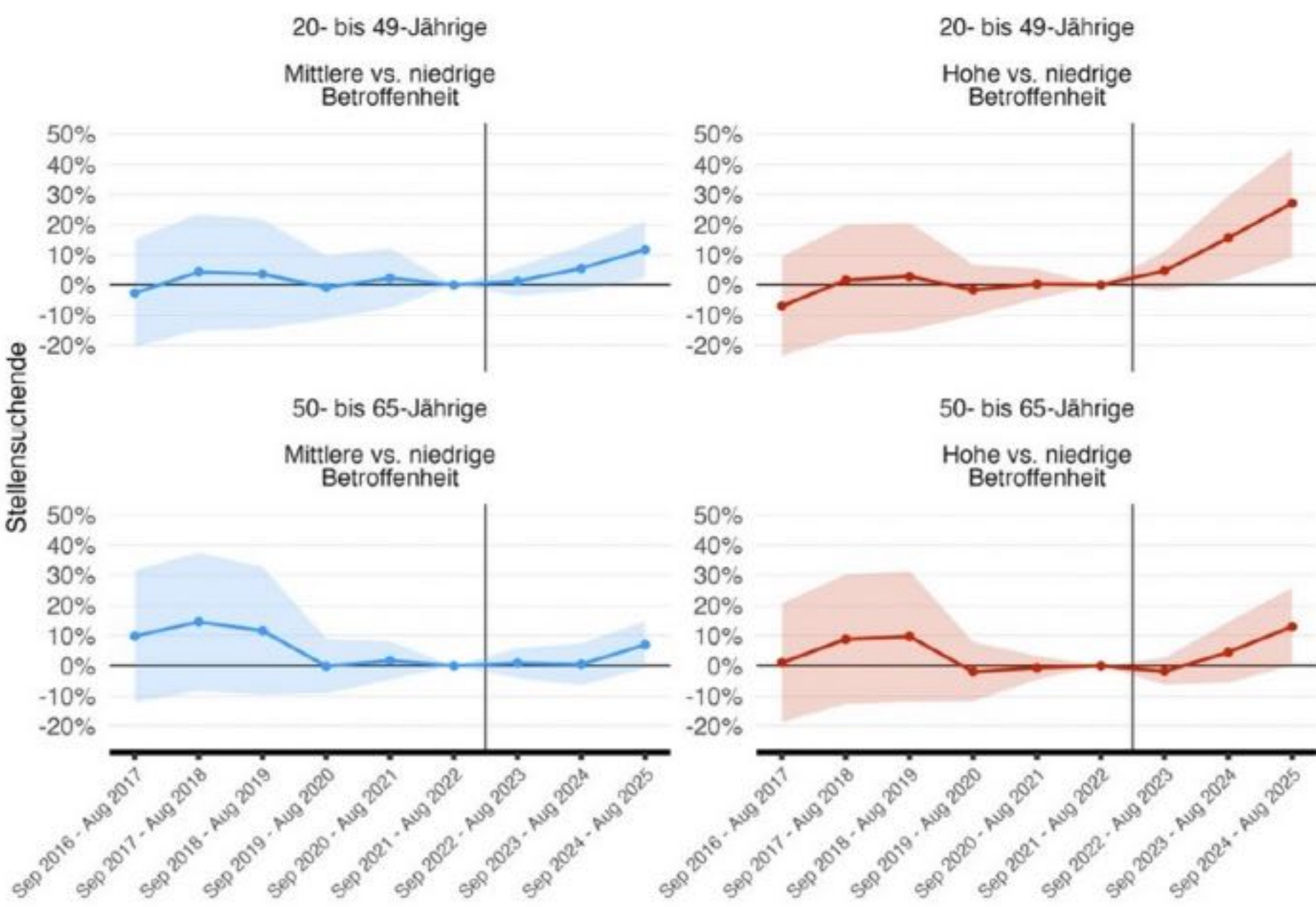
Chancen, Verantwortung & Zukunft für KMUs

Diverse Dimensionen von "AI & Ethik"

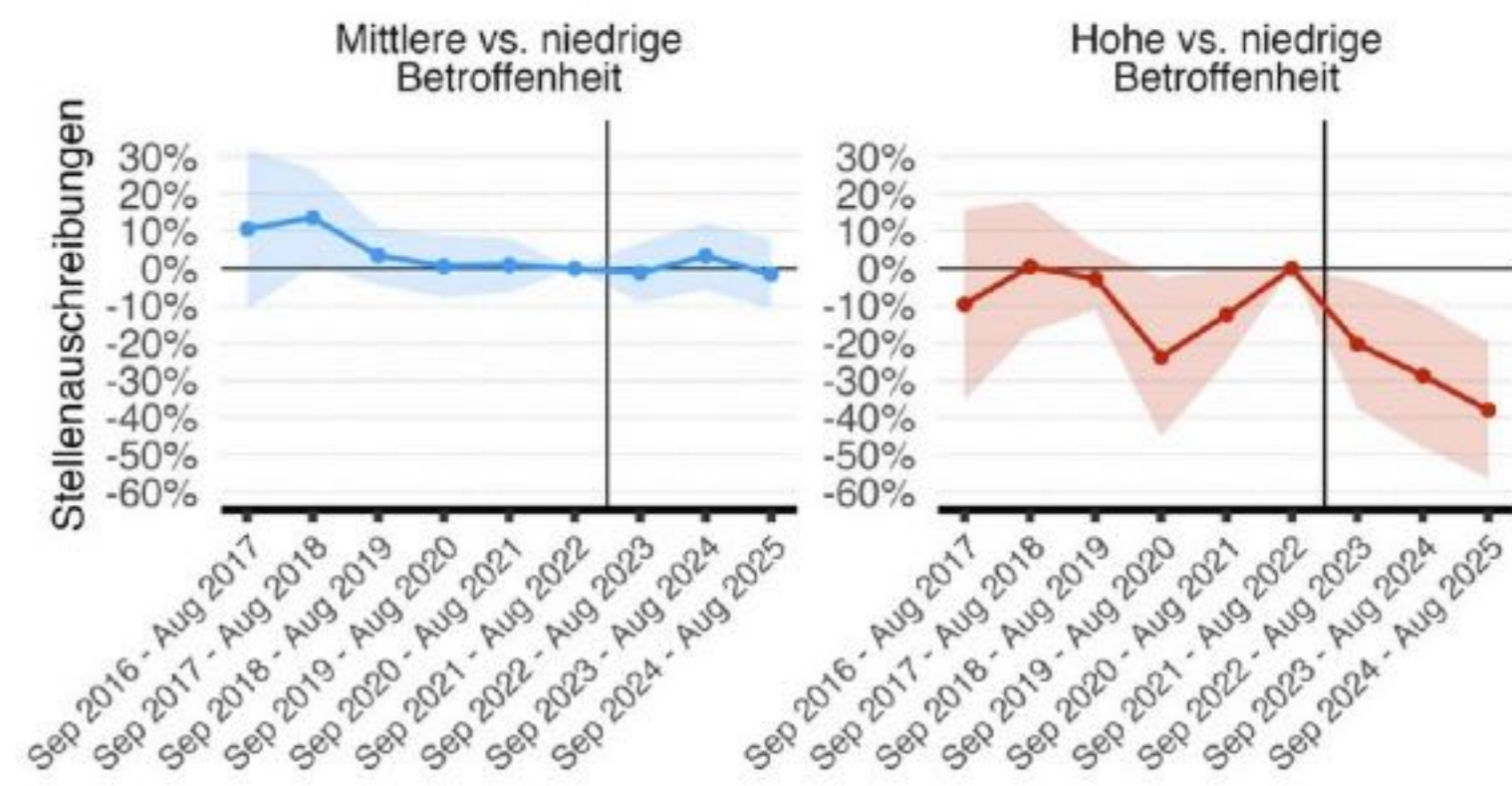


Exposed Jobs: neue Zahlen zeichnen ein klares Bild

Event Study: Arbeitslosigkeit nach KI-Betroffenheit



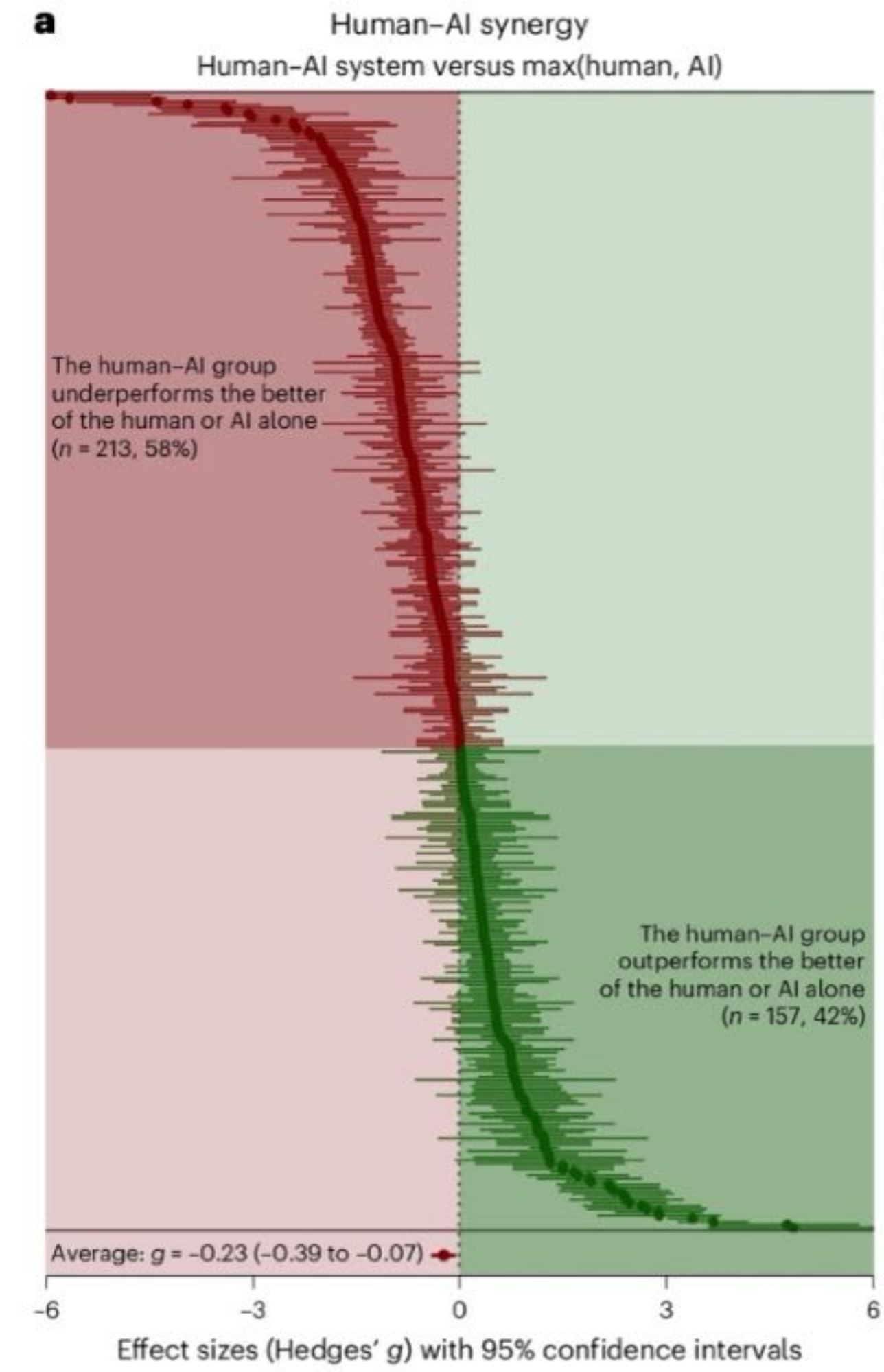
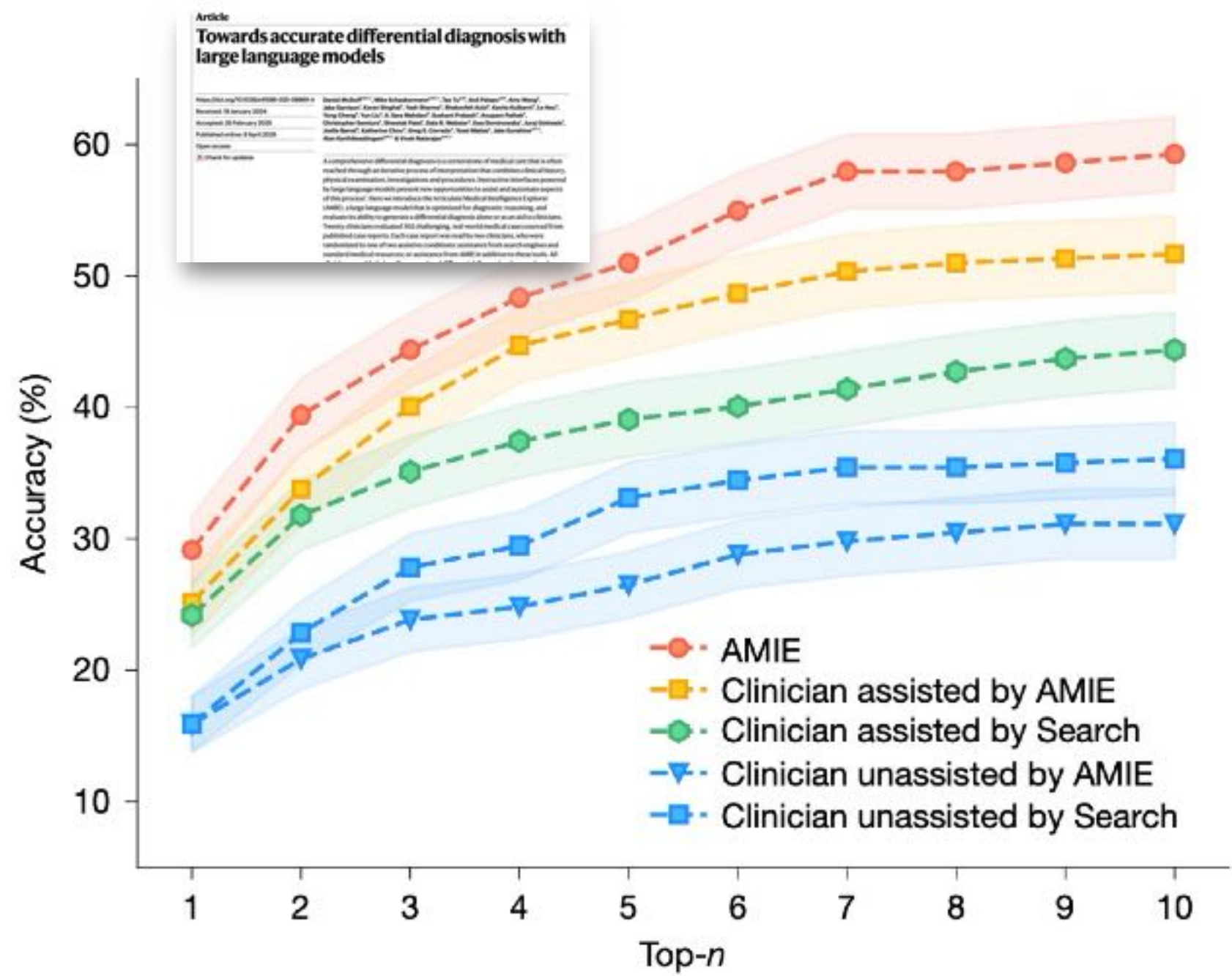
Event Study: Stellenausschreibungen nach KI-Betroffenheit



~27% mehr Jobsuchende sowie **38%** weniger Jobinserate

Effekt Mensch + AI

AI + human > AI ?



Arbeit & Vertrauen – Angst vor AI?

Fragen:

- Wird AI Jobs vernichten oder verändern?
- Wie bleibt Vertrauen zwischen Mensch & Maschine erhalten?
- Welche Chancen bietet Weiterbildung?

Kernaussagen:

- Nur 8% der Jobs fallen weg - 66% unterstützt.
- Neue Rollen mit AI-Assistenz entstehen.
- Weiterbildung ist entscheidend für Akzeptanz & Kreativität.

Beispiele:

- KMU bildet „AI-Botschafter:innen“ aus, Mitarbeitende nutzen Tools aktiv.
- Kleine Buchhaltungsfirma ersetzt Mitarbeitende durch AI, Vertrauen sinkt.

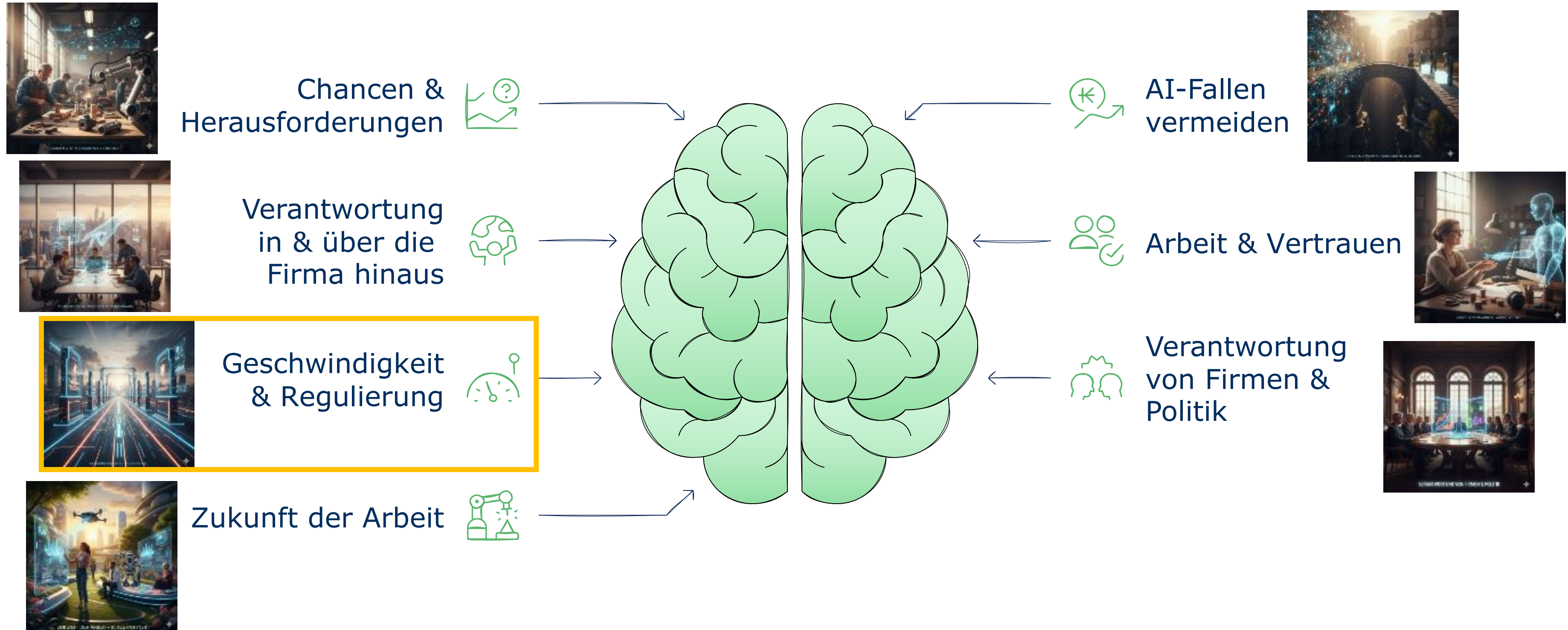


*Wie können Unternehmen
neue Talente nachhaltig ausbilden?*

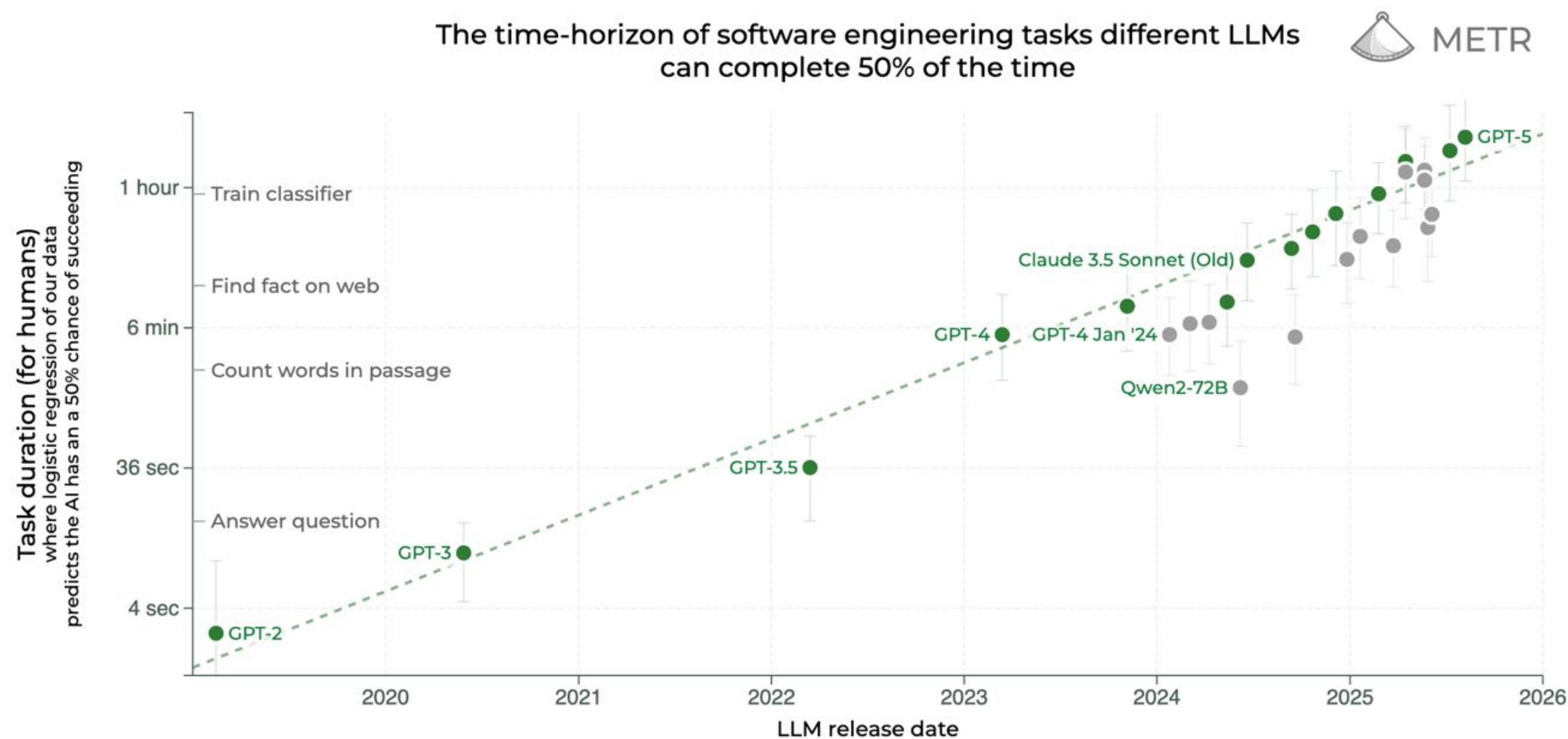
*Was löst die Erkenntnis, dass «Human + AI» nicht immer die
beste Wahl ist, bei Dir aus?*

Chancen, Verantwortung & Zukunft für KMUs

Diverse Dimensionen von "AI & Ethik"



Bleibt der Effekt von AI gleich?

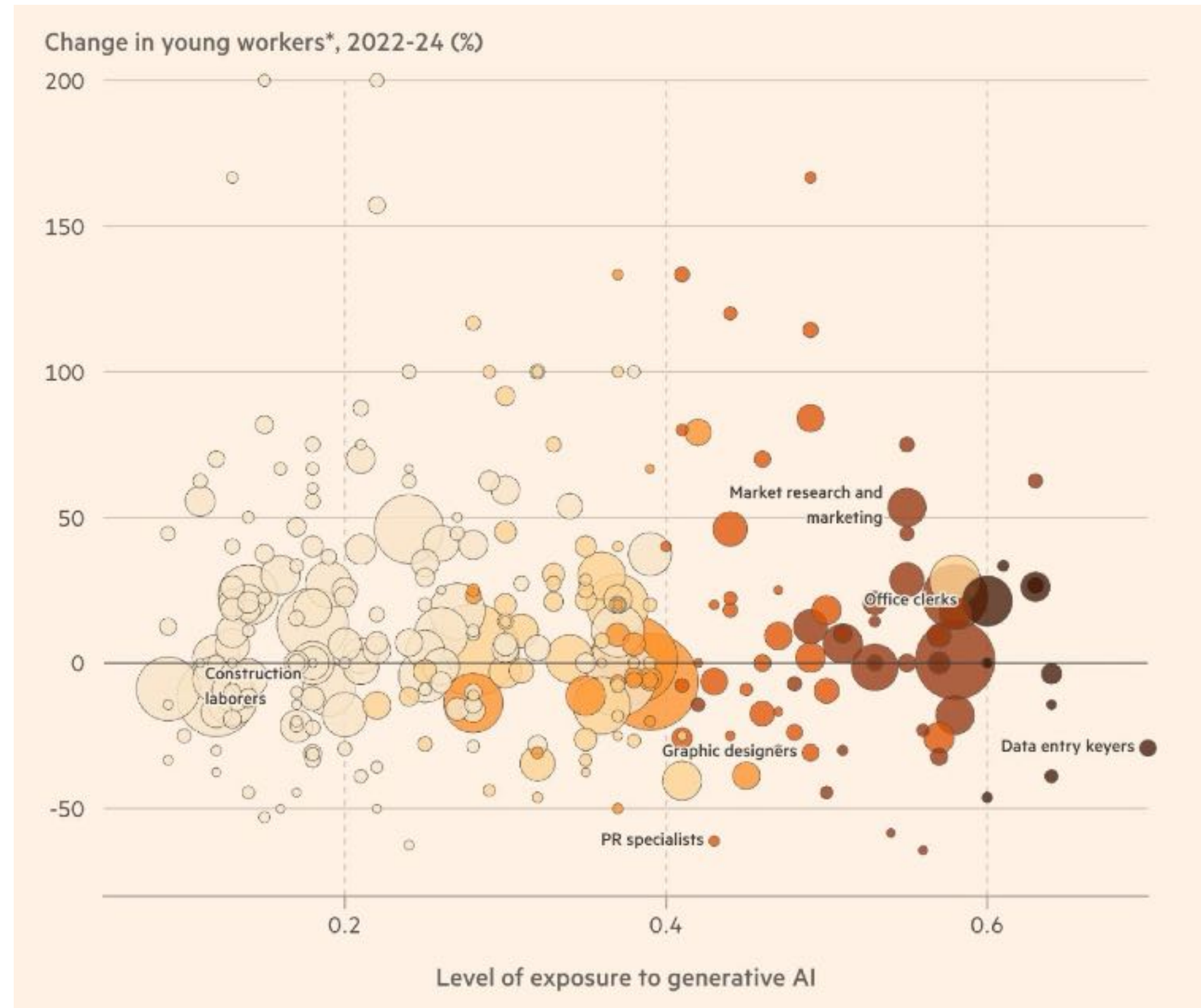


Effekt Mensch + AI

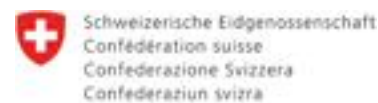
... BUT ...

Past 30 years of technology revolution → record low unemployment

“This time is different” - Why? Tasks get automated, new ones emerge.



Regulierung in der Schweiz



Eidgenössisches Departement für Umwelt, Verkehr,
Energie und Kommunikation UVEK
Bundesamt für Kommunikation BAKOM

12. Februar 2025

Auslegeordnung zur Regulierung von künstlicher Intelligenz

Bericht an den Bundesrat



Die Auslegeordnung definiert drei übergreifende Ziele, die durch eine Schweizer Regulierung im Bereich KI erfüllt werden sollten: (i) die Stärkung des Innovationsstandorts Schweiz, (ii) die Wahrung des Grundrechtsschutzes inklusive der Wirtschaftsfreiheit und (iii) die Stärkung des Vertrauens der Bevölkerung in KI. Hierfür skizziert die Auslegeordnung drei mögliche Regulierungsansätze, welche die Ziele in unterschiedlichem Masse erfüllen würden:

- (i) Fortführung der themen- und sektorspezifischen Regulierungsaktivitäten: Der Regulationsbedarf würde weiterhin themen- und sektorspezifisch identifiziert und umgesetzt. Dies sofern nötig auch in übergreifenden Themen wie Datenschutz. Der Bundesrat würde darüber hinaus darauf verzichten, neue regulatorische Massnahmen vorzuschlagen und derzeit keine übergreifende Schweizer Regulierung von KI-Anwendungen anstreben. Betreffend sektorübergreifender Herausforderungen würde es abgesehen von den üblichen Konsultationsverfahren weiterhin kein übergreifend koordiniertes regulatorisches Vorgehen auf Bundesebene geben.
- (ii) Ratifikation der KI-Konvention des Europarats mit einer Minimalumsetzung (Option 1) oder einer weitergehenden Umsetzung (Option 2): Die wichtigsten Herausforderungen hinsichtlich des Schutzes der Grundrechte, der Demokratie und der Rechtsstaatlichkeit würden durch einen koordinierten Ansatz auf Bundesebene angegangen, so z.B. im Bereich Transparenz sowie durch eine Risiko- und Folgeabschätzung von KI-Systemen. Bei der minimalen Umsetzung wären die Verpflichtungen für den Staat umfassender als für Private; bei einer weitergehenden Umsetzung würden ähnlich weitgehende Verpflichtungen für Staat und Private angestrebt.
- (iii) Ratifikation der KI-Konvention und Umsetzung in Anlehnung an den AI Act der EU: In Anlehnung an den AI Act der EU wäre mit der Einführung eines risikobasierten Ansatzes für den Umgang mit Produkten mit KI-Bestandteilen dieselben Voraussetzungen für Schweizer Unternehmen wie in der EU gegeben. Dies könnte den Marktzugang von Schweizer Produkten mit KI-Bestandteilen auf den EU-Binnenmarkt erleichtern, würde jedoch zu einer hohen Regulierungsdichte in der Schweiz führen.

Europarate - Framework Convention on AI

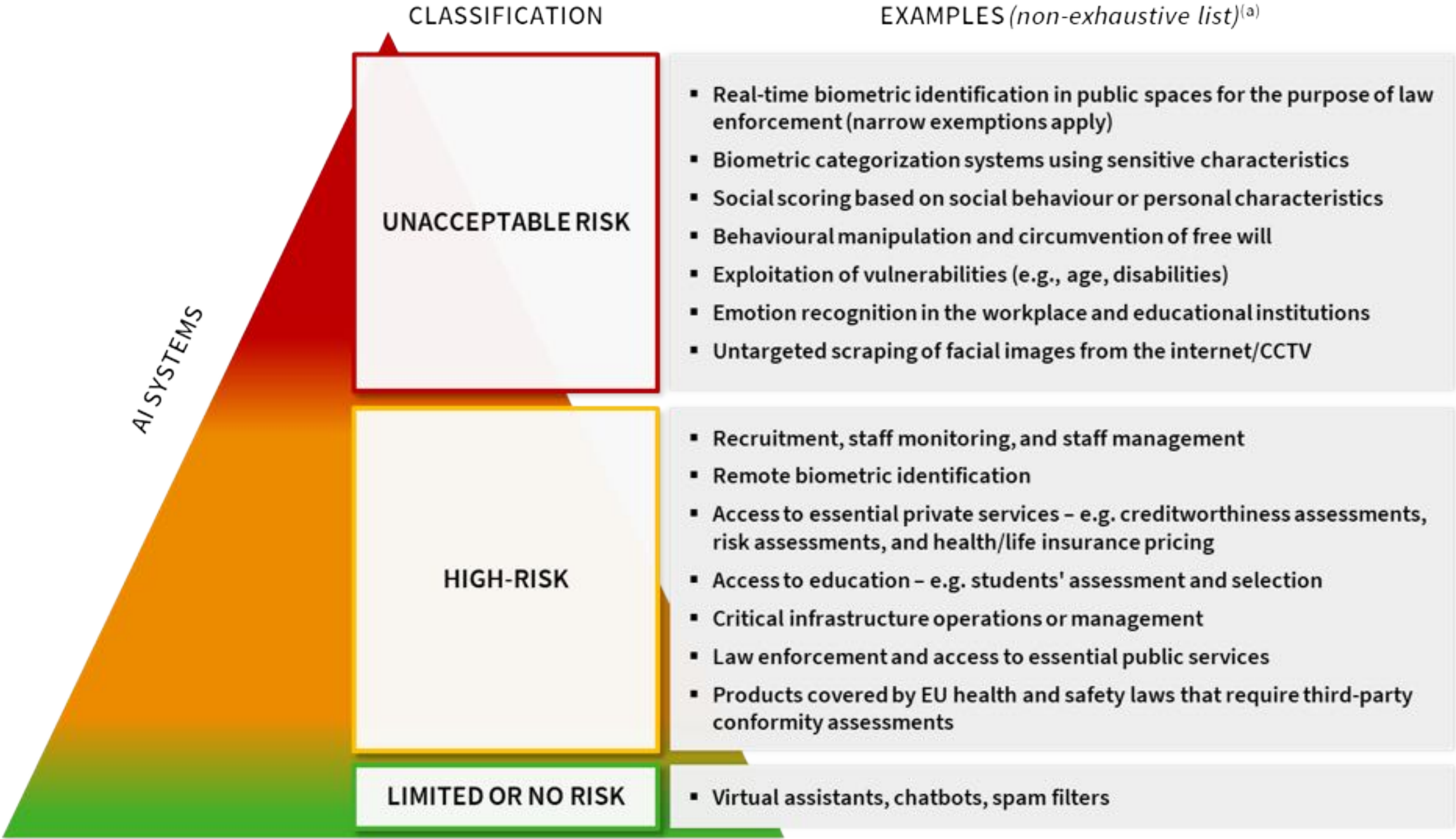


FUNDAMENTAL PRINCIPLES

Activities within the lifecycle of AI systems must comply with the following fundamental principles:

- ▶ Human dignity and individual autonomy
- ▶ Equality and non-discrimination
- ▶ Respect for privacy and personal data protection
- ▶ Transparency and oversight
- ▶ Accountability and responsibility
- ▶ Reliability
- ▶ Safe innovation

EU AI Act wird Bereiche verbieten



(a) The full and final list of prohibited and high-risk AI systems and any exemptions will be confirmed once the final text of the AIA is released.

<https://www2.deloitte.com/uk/en/blog/emea-centre-for-regulatory-strategy/2023/the-eu-ai-act-the-finish-line-is-in-sight.html>

Geschwindigkeit & Regulierung

Fragen:

- Entwickelt sich AI zu schnell?
- Wie sichern wir Transparenz?
- Wie umgehen mit globalen Standards?

Kernaussagen:

- LLMs sind komplex & schwer durchschaubar.
- EU AI Act & Schweizer Regulierung bis 2026 geplant.
- Bedarf an gemeinsamen Standards & Transparenzpflicht.

Beispiele:

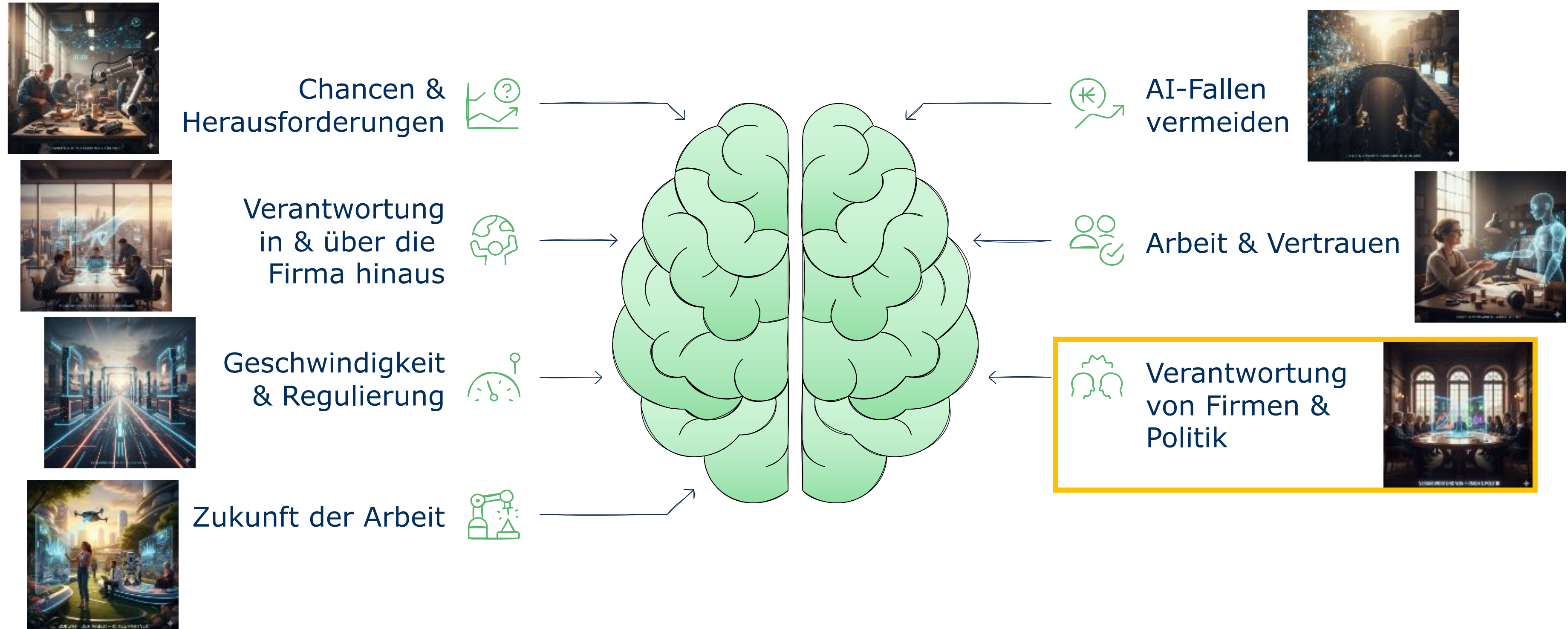
- Startup in Zürich richtet sich freiwillig nach EU AI Act – Vertrauen steigt.
- US-Tool in Schweizer KMU verstösst gegen DSGVO – Kunden klagen.



Welche Faktoren werden entscheiden, ob wir der AI mehr vertrauen als dem Menschen?

Chancen, Verantwortung & Zukunft für KMUs

Diverse Dimensionen von "AI & Ethik"



Verantwortung von Firmen & Politik



From Boots on the Ground to Bytes in Cyberspace: A Comprehensive Toolkit for Responsible Technology Use



Purpose:

- **Protect Human Rights:** through integrating best practices to safeguard privacy, prevent discrimination and support freedom of speech
- **Enhance Security:** by leveraging technology effectively while minimizing risks and vulnerabilities
- **Build Trust:** demonstrate commitment to ethical and responsible operations

Verantwortung von Firmen & Politik



From Boots on the Ground to Bytes in Cyberspace: A Comprehensive Toolkit for Responsible Technology Use



List of Tools

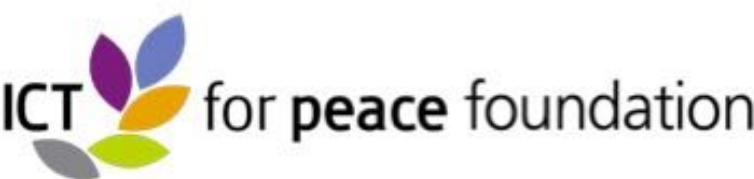
- **Tool 1:** Human Rights Challenges Posed by ICTs in Private Security Companies
- **Tool 2:** Responsible Data Collection Practices
- **Tool 3:** Best Practices for Data Storage
- **Tool 4:** Best Practices for Data Security
- **Tool 5:** Best Practices for Data Destruction
- **Tool 6:** Surveillance and Monitoring
- **Tool 7:** Artificial Intelligence and Algorithmic Bias in Private Security
- **Tool 8:** Emerging Technologies and Future Trends in Private Security
- **Tool 9:** Accountability and Transparency
- **Tool 10:** Freedom of Expression
- **Tool 11:** Labor Rights in the Digital Age
- **Tool 12:** Right to Remedy and Effective Grievance Mechanisms

Who Is This Toolkit For?

This toolkit is an essential resource for anyone involved in the intersection of private security and technology, including:

- Computer Security Professionals
- Management Teams
- Human Rights Officers
- Compliance Teams
- Technology Teams
- Governments overseeing/contracting with Companies
- Private clients using ICT services
- Civil Society working in business and/or human rights

Verantwortung von Firmen & Politik



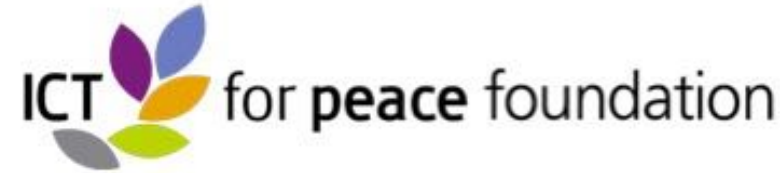
From Boots on the Ground to Bytes in Cyberspace: A Comprehensive Toolkit for Responsible Technology Use



Policy and Governance Government regulations and policies - AI Fairness Policy - Ethics review huddle - Regulatory changes tracking	Resource Allocation Investment in bias mitigation - Training programs funding - Auditing toolkit adoption - Third-party spot-checks	Public Perception Societal values and ethics - Bias incident reporting channel - Transparency and audits - Human-in-the-loop thresholds	Innovation & Automation Advancements in AI technology - Fairness metrics computation - Continuous monitors for drift - Explainability for high-stakes models	Data Impact Assessment Impact on data privacy - Data representativeness checks - Privacy-preserving approaches - AI Impact Assessment	Compliance & Litigation Legal and ethical standards - Anti-discrimination duties - Data-protection duties - Decision logs and model cards

Created based on ICT for peace with Napkin.ai

Verantwortung von Firmen & Politik



From Boots on the Ground to Bytes in Cyberspace: A Comprehensive Toolkit for Responsible Technology Use



2. Understanding Algorithmic Bias

2.1 Definition and Relevance to PSCs

Algorithmic bias refers to systematic and repeatable errors in computer systems that create unfair outcomes, often disadvantaging certain groups or individuals.

2.8 Implementation Checklist

- ☐ Develop a comprehensive algorithmic bias training program
- ☐ Establish a regular schedule for AI system audits
- ☐ Form diverse teams for AI development and deployment
- ☐ Create channels for stakeholder feedback on potential biases
- ☐ Implement automated bias detection tools
- ☐ Develop clear policies on AI use and bias mitigation
- ☐ Establish an ethical framework for AI in security operations

2.9 Common Pitfalls to Avoid

- Assuming algorithm neutrality without thorough testing
- Relying solely on technical metrics without considering real-world impacts
- Neglecting intersectional biases
- Failing to involve affected communities in discussions about AI and bias
- Overlooking potential biases in seemingly objective data sources
- Treating bias understanding as a one-time effort rather than an ongoing process

5. Ensuring Human Oversight and Accountability

5.1 Definition and Relevance to PSCs

Human oversight and accountability in AI systems refer to the processes and mechanisms that ensure human control, supervision, and responsibility over AI-driven decisions and actions.

5.8 Implementation Checklist

- ☐ Establish a dedicated AI Oversight Team
- ☐ Develop clear policies for human intervention in AI processes
- ☐ Implement tools for real-time monitoring of AI decisions
- ☐ Create a comprehensive AI oversight training program
- ☐ Establish regular audit procedures for AI systems and oversight processes
- ☐ Implement feedback mechanisms for reporting AI-related concerns
- ☐ Develop performance metrics for evaluating human oversight effectiveness

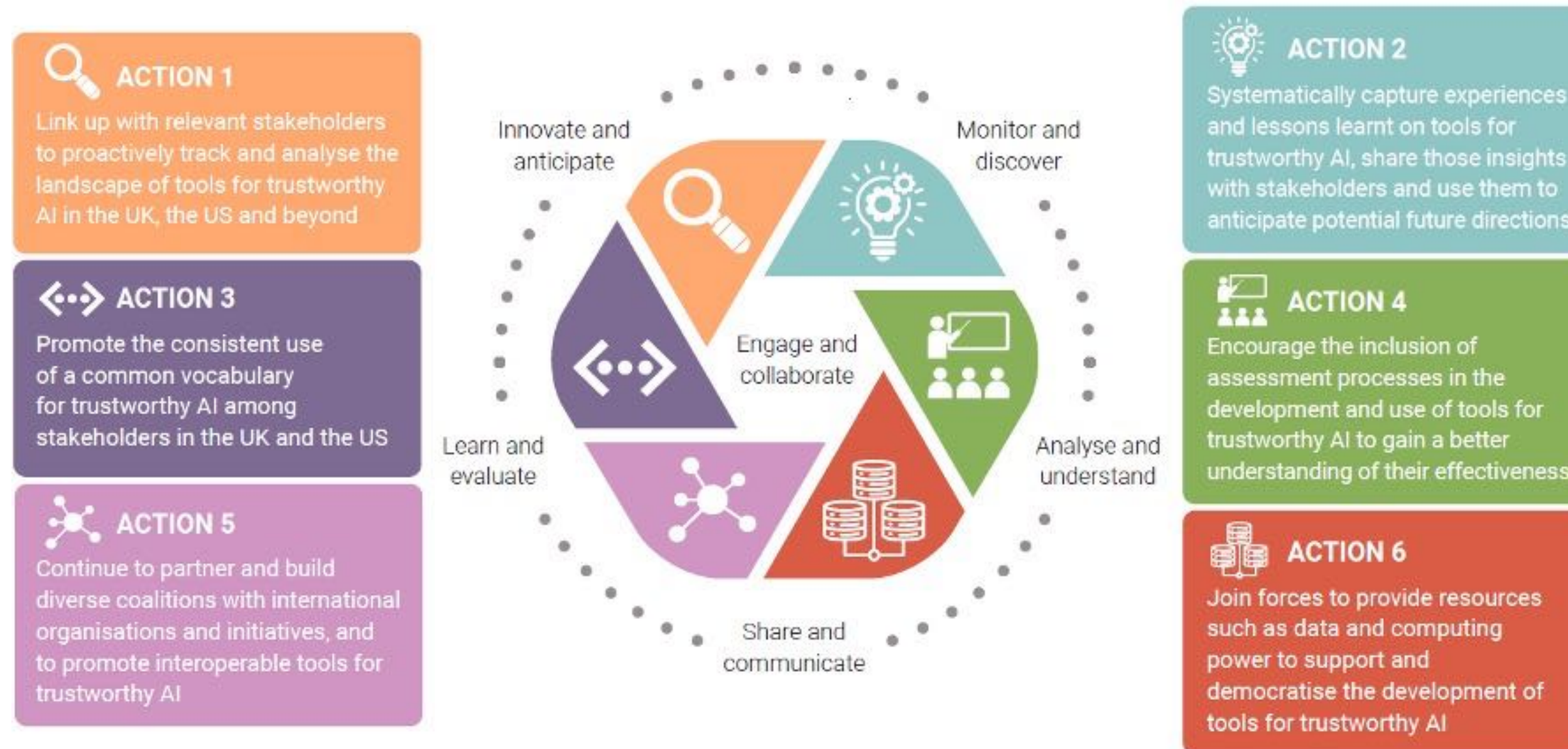
5.9 Common Pitfalls to Avoid

- Over-relying on AI without adequate human oversight
- Neglecting to train personnel on AI capabilities and limitations
- Failing to clearly define accountability for AI-related decisions
- Ignoring feedback from frontline staff about AI performance
- Implementing oversight mechanisms that significantly slow down operations
- Assuming that human oversight alone can solve all AI-related issues



Verantwortung von Firmen & Politik

OECD Actions to shape the future of AI



Potential stakeholders to involve across the different actions: Department for Science, Innovation and Technology (including the Responsible Technology Adoption Unit and UK AI Safety Institute); Foreign, Commonwealth & Development Office (including the British Embassy Washington); AI Standards Hub; UK Research and Innovation; AI Research Resource; techUK; Evaluation Task Force in the UK; Government Office for Science; National Institute of Standards and Technology; US AI Safety Institute; National Science Foundation; National Artificial Intelligence Research Resource; US national laboratories; Organisation for Economic Co-operation and Development; European Commission; United Nations (and associated agencies); standards development organisations.

Source: RAND Europe analysis



Verantwortung von Firmen & Politik

Fragen:

- Wie kann die Schweiz negative AI-Auswirkungen minimieren?
- Welche Rolle spielen Firmen & Regierung?
- Kann Ethik Standortvorteil sein?

Kernaussagen:

- Firmen: kleine Pilotprojekte, Bildung, Datenqualität, Governance.
- Politik: Förderinstrumente & Ethikinitiativen.
- Ziel: Ethik als Wettbewerbsvorteil.

Beispiele:

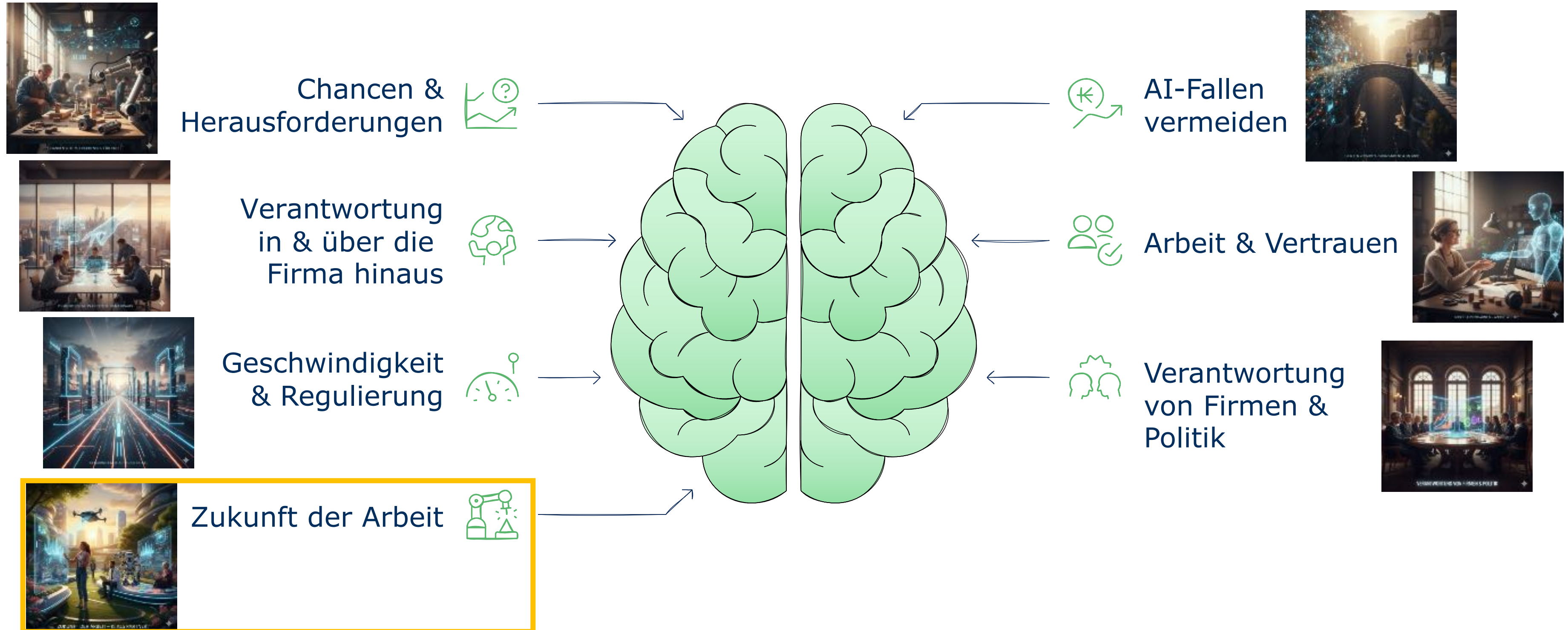
- Smart-Farming-Projekt nutzt AI zur Reduktion von Pestiziden & Ressourceneinsatz.
- Industrie-AI-Subventionen führen zu Kritik wegen Stromverbrauch.



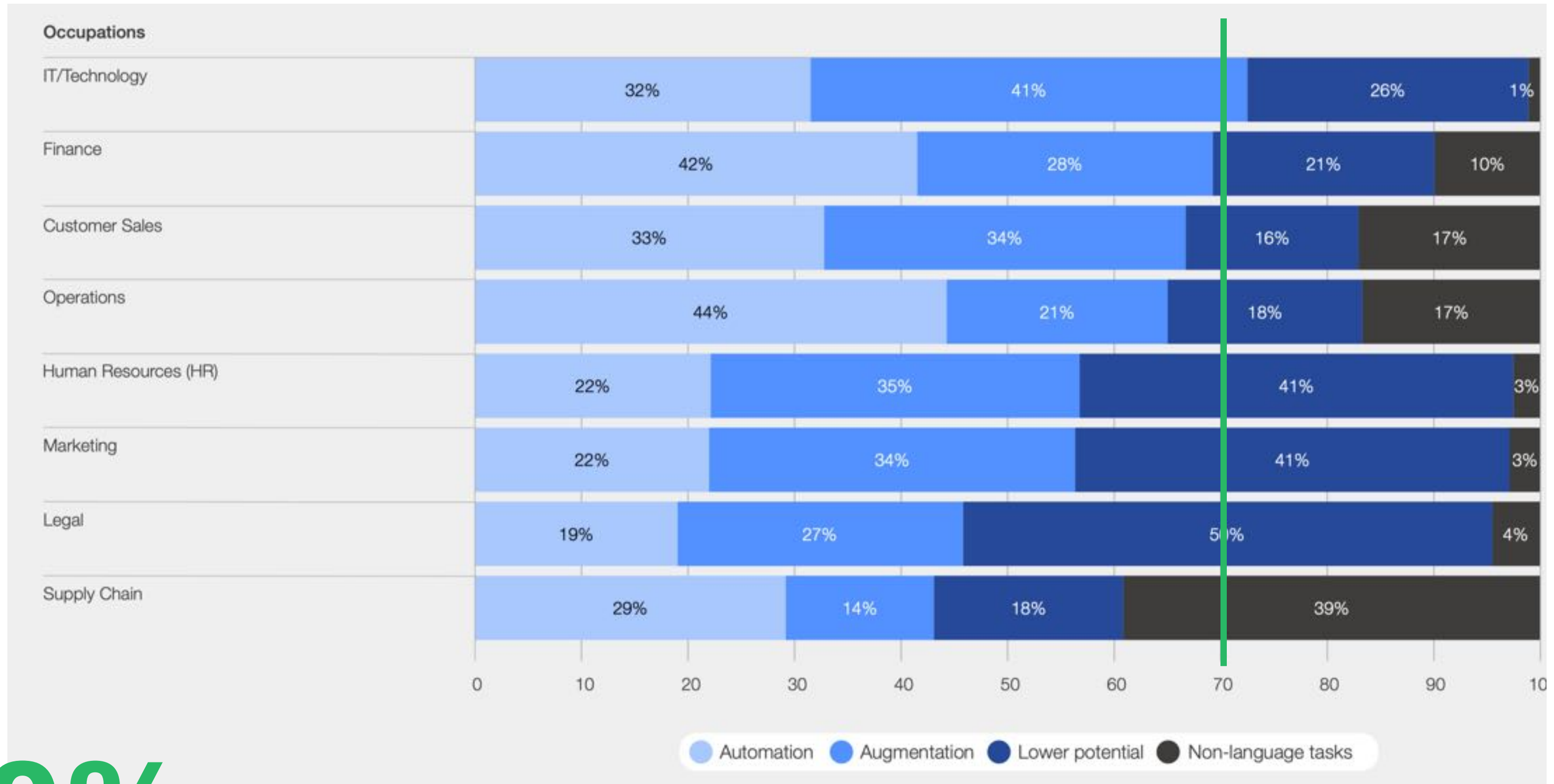
*Welche Massnahmen sollte die Regierung (Bund / Kantone)
veranlassen & verantworten?*

Chancen, Verantwortung & Zukunft für KMUs

Diverse Dimensionen von "AI & Ethik"



Welche Jobs sind betroffen?



~70%

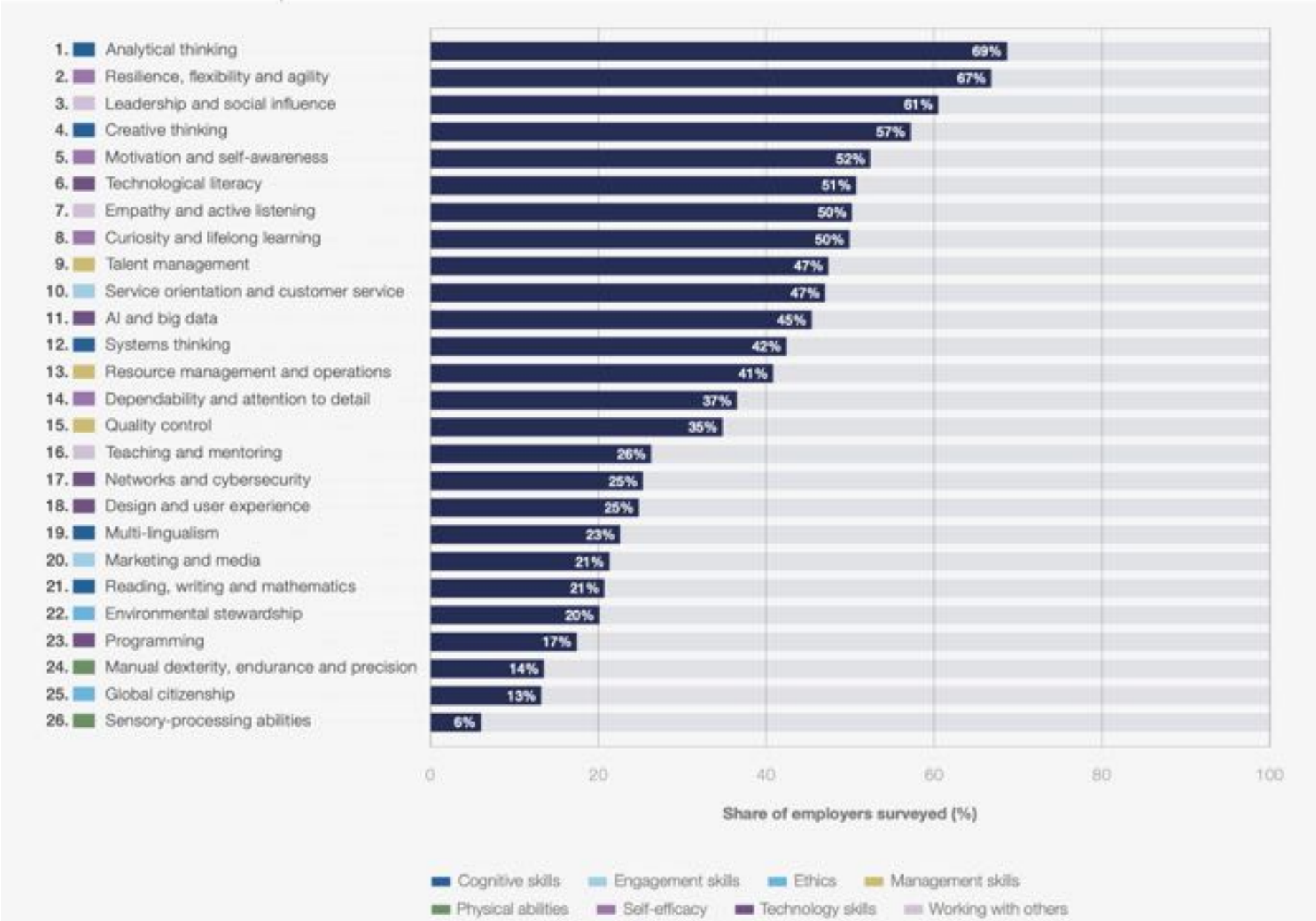
aller Aufgaben haben das Potential mit AI automatisiert oder unterstützt zu werden

Wichtigsten Fähigkeiten im Zeitalter der AI

Heute und in 5 Jahren

FIGURE 3.3 Core skills in 2025

Share of employers who consider the stated skills to be core skills for their workforce.

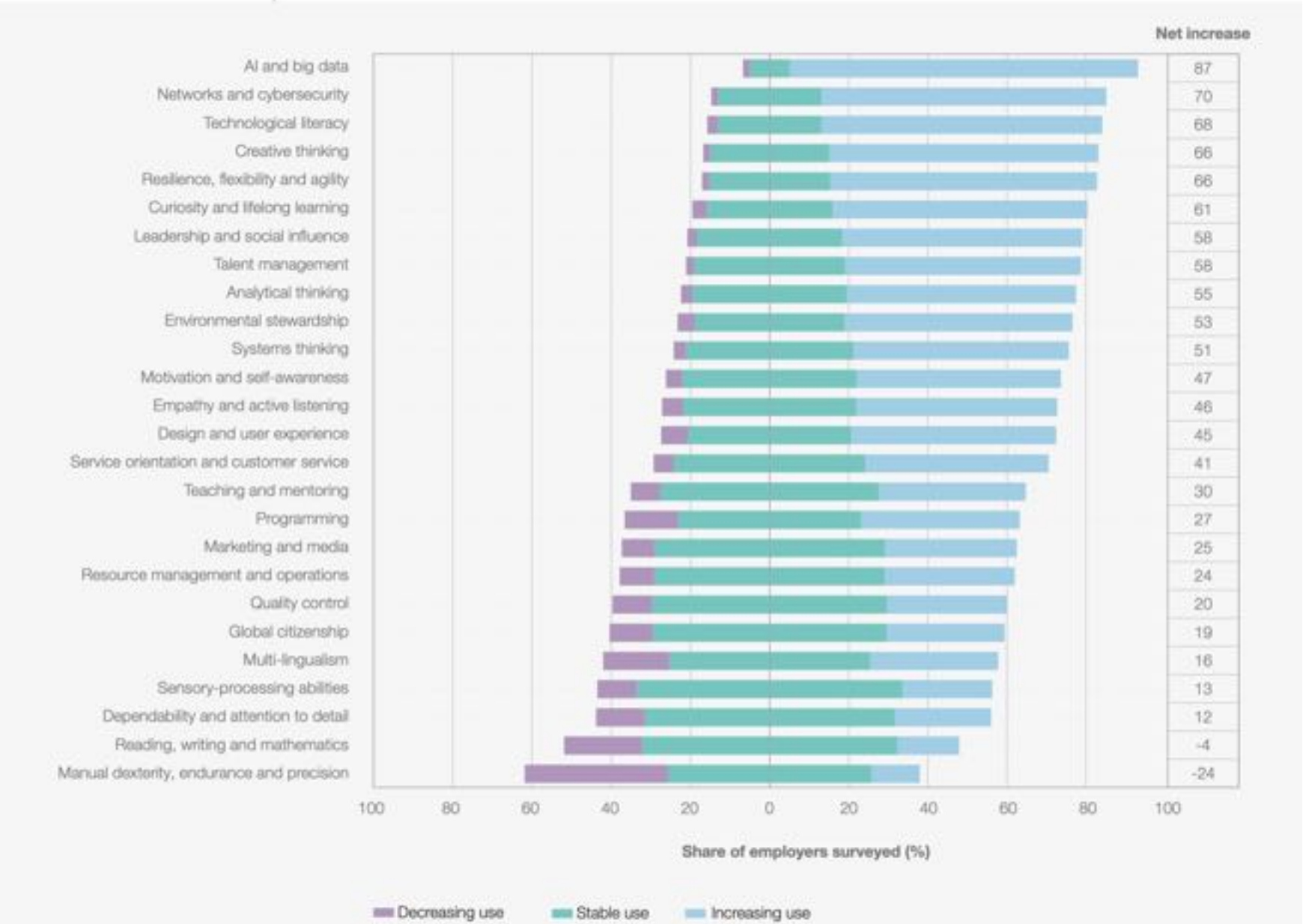


Source
World Economic Forum, Future of Jobs Survey 2024.

Note
The Future of Jobs Survey uses the World Economic Forum's Global Skills Taxonomy.

FIGURE 3.4 Skills on the rise, 2025-2030

Share of employers that consider skills to be increasing, decreasing, or remaining stable in importance. Skills are ranked based on net increase, which is the difference between the share of employers that consider a skill category to be increasing in use and those that consider it to be decreasing in use.



Source
World Economic Forum, Future of Jobs Survey 2024.

Note
The Future of Jobs Survey uses the World Economic Forum's Global Skills Taxonomy.

Zukunft der Arbeit – AI als Partner

Fragen:

- Wie verändert AI den Arbeitsplatz in 5–10 Jahren?
- Welche Entscheidungen übernimmt AI?
- Wie fördern Firmen Kreativität mit AI?

Kernaussagen:

- AI übernimmt Routineentscheide („Automated & Augmented Intelligence“).
- Mensch fokussiert auf Strategie & Kreativität.
- Besonders betroffen: Produktion, Marketing, Logistik.

Beispiele:

- Kreativagentur nutzt AI für Ideenskizzen – neue Kampagnen entstehen.
- Marketingbüro automatisiert Contentproduktion – Kreative fühlen sich ersetzt.



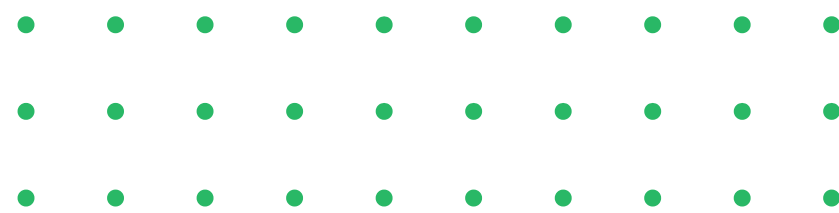
Welche Fähigkeiten / Tätigkeiten wird eine AI nie ausführen?

Stelle Dir vor es ist 2050 und du arbeitest nur noch 1 Stunde pro Tag. Was machst du in dieser Stunde?

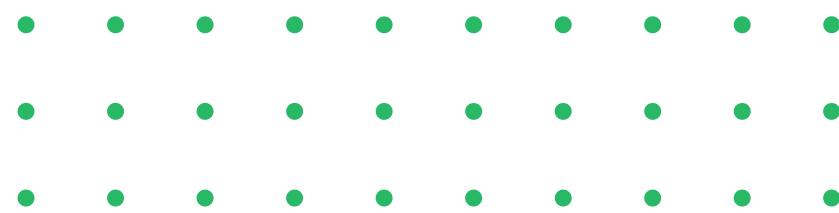


Wie starten?

Die Frage ist nicht, ob AI unsere Arbeitsweise verändern wird, sondern ob wir diese Veränderung so gestalten, dass sie das menschliche Potenzial fördert, anstatt es zu schmälern.

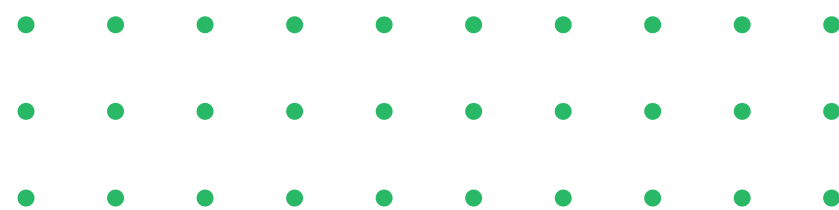


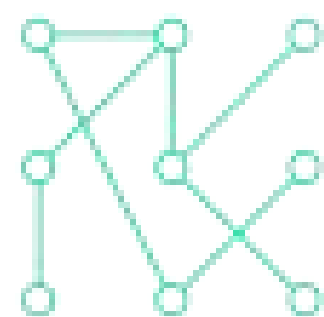
Lasst uns die Technologie gemeinsam, nachhaltig & verantwortlich nutzen





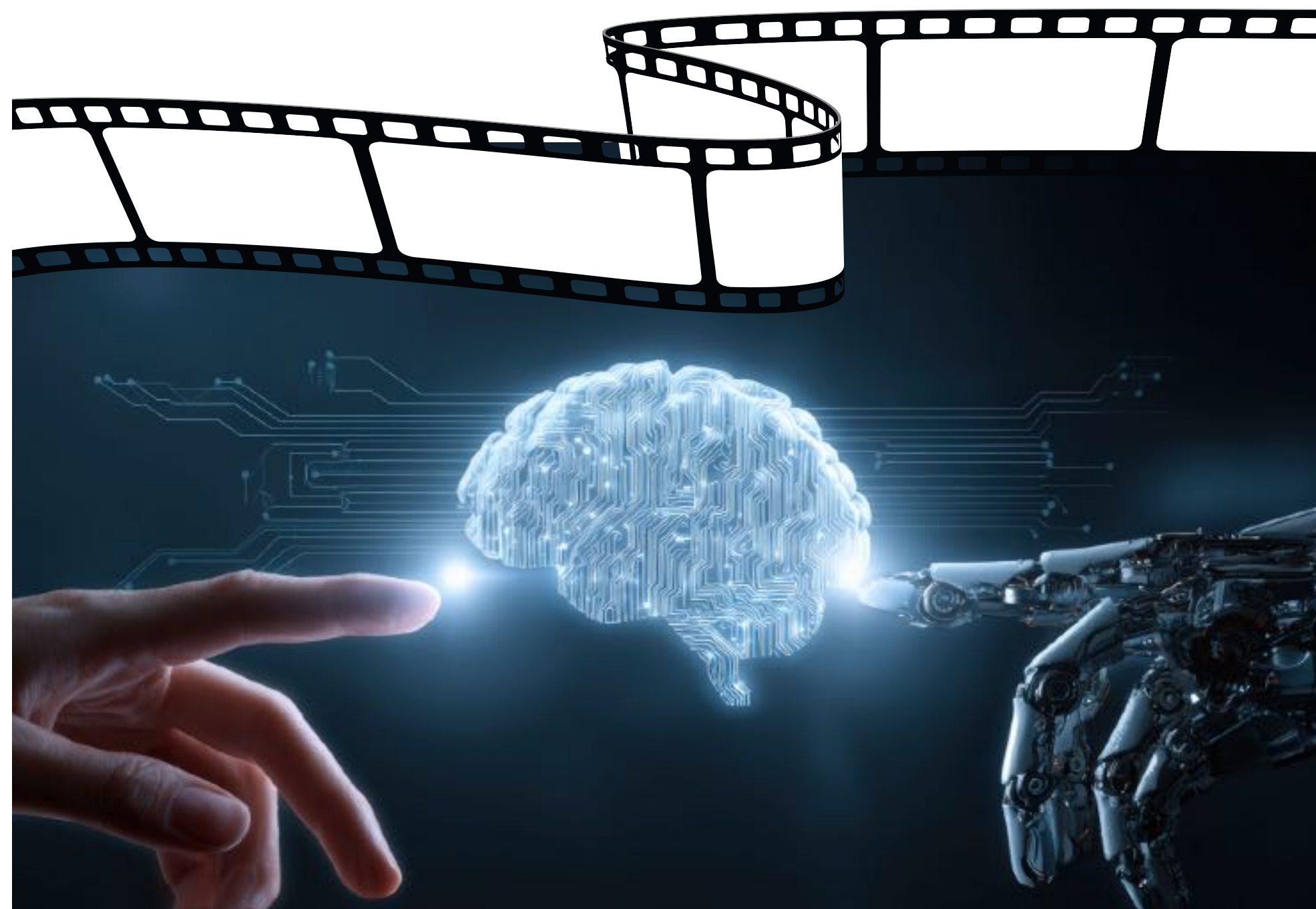
Ausblick – nächste Events um AI





**DIGITAL IMPACT
NETWORK**

EINEN FILMABEND ZUM THEMA
«SCIENCE FICTION ODER REALITÄT?»



I, ROBOT

20. NOVEMBER 2025
18-23 UHR
KINO IM UFERBAU
SOLOTHURN

Der Eintritt ist frei.
Bitte um Anmeldung
mittels QR-Code oder
bit.ly/47mxzFj



GenAI Award 2025: Application open on October 15



Driving real-world impact
with applied GenAI

Stand Out in the Press

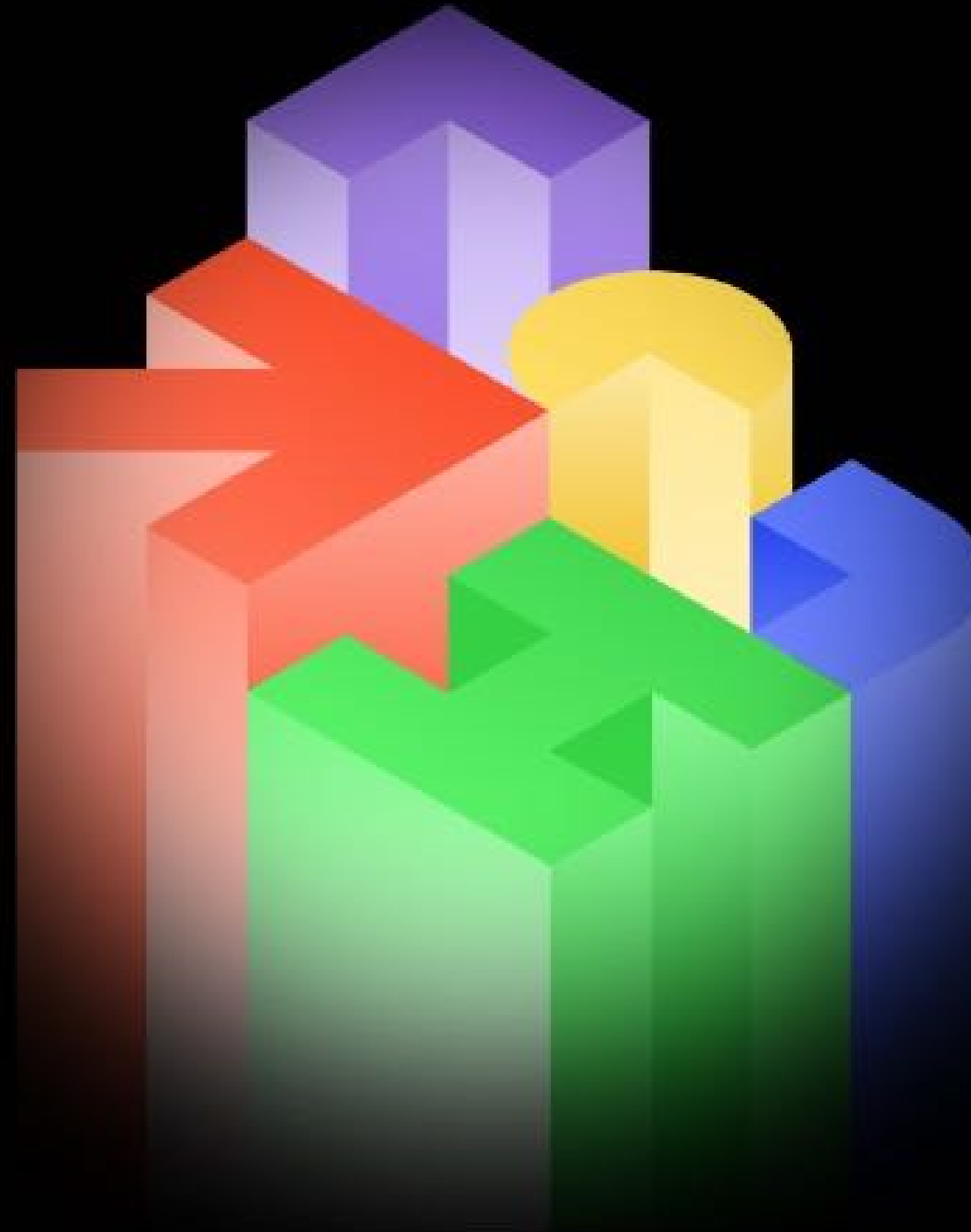
Gain high-profile media coverage in press releases, elevating your brand and AI initiatives.

Earn Recognition & Credibility

Validate your project by an expert jury in technology and business to boost your organisation's reputation.

Engage the Expert Community

Exchange ideas, receive valuable feedback, and collaborate with global AI thought leaders.





Danke für dein Interesse

“ We are
ready to
take off ! ”





Contact us!



artifact.swiss



linkedin.com/company/artifactswiss