



© Adobe Stock / 1757331612

Preparing for the Age of AI

A Living Outlook for Decision-Makers

Contents

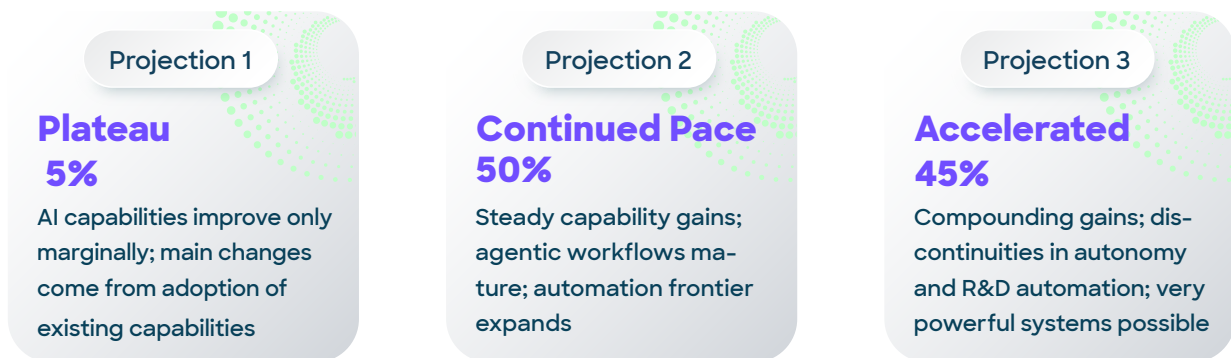
Executive summary	3
Chapter 1: AI advancement projections (3-Year Horizon)	5
1.1 Purpose and scope of this chapter	5
1.2 Key definitions (Working Terms)	6
1.3 Baseline conditions: constraints that shape all projections	6
1.4 The three projections: differentiated by capability speed	10
1.5 Projection comparison: one-page decision aid	15
1.6 What this chapter does not determine	16
1.7 How to read the rest of this document	16
Chapter 2: Speed drivers & projection probabilities	17
2.1 Purpose of this chapter	17
2.2 Driver selection criteria	17
2.3 The four core drivers	18
2.4 Driver 1: Architectures & training paradigms	19
2.5 Driver 2: Agentic autonomy & orchestration	24
2.6 Driver 3: Automation of AI R&D and software engineering	29
2.7 Driver 4: Robotics & embodied AI	32
2.8 Key benchmarks reference	36
2.9 Interactions between drivers	37
2.10 Consolidating drivers into projection probabilities	38
2.11 Breakthrough triggers: when to re-estimate	41
2.12 Summary: projection probabilities and key uncertainties	42
Chapter 3: Concrete impacts (opportunities & risks) by projection	43
3.1 Purpose and structure of this chapter	43
3.2 Impact categories	44
3.3 Impact category 1: Labour market & skills	45
3.4 Impact category 2: Public finance & social systems	48
3.5 Impact category 3: Industry & competitiveness	51
3.6 Impact category 4: Innovation & startups	56
3.7 Impact category 5: Science system	60
3.8 Impact category 6: Security & resilience	64
3.9 Impact category 7: Digital public sphere & democracy	68
3.10 Impact category 8: Health & care	71
3.11 Impact category 9: Education system	75
3.12 Impact category 10: Functioning local institutions & liveability	78
3.13 Cross-cutting observations	81
3.14 Summary: Impact Severity by Projection	83
3.15 What this chapter does not determine	84
Chapter 4: Measures & appliedAI portfolio logic	85
4.1 Purpose of this chapter	85
4.2 Measures design principles	85
4.3 Part A measures inventory (What should be addressed overall)	86
4.4 Part B - appliedAI Portfolio Selection (What We Will Do)	92
Chapter 5: Update system: Research agent + expert panel + versioning	99
5.1 Purpose of this chapter	99
5.2 Update triggers and cadence	100
5.3 Research agent system	100
5.4 Expert panel	101
5.5 Versioning and change management	102
5.6 Quality assurance	103
5.7 Contact and contribution	103
Reference	107

Executive summary

What this document provides

This AI Perspective Whitepaper provides European decision-makers with a structured orientation for navigating AI advancement over the next 36 months. This report does not present full scenarios. Instead, it offers structured projections of technological capability development as one central driver of future outcomes, providing a framework for preparing across plausible futures.

Our three perspectives (differentiated by capability speed)



Key drivers we monitor

1. **Architectures & Training Paradigms** – whether scaling continues, new paradigms emerge (Nested Learning, Causal AI), or both
2. **Agentic Autonomy & Orchestration** – reliability of multi-step autonomous operation (SWE-bench Verified now at 79%, OSWorld at 73%)
3. **Automation of AI R&D and Software Engineering** – feedback loops now materialising (Claude Code writes ~90% of Anthropic's code)
4. **Robotics & Embodied AI** – sim-to-real gap largely closed; Chinese manufacturers achieving 99%+ success rates in production

Top impacts requiring attention

1. **Labour market pressure** – entry-level knowledge workers already affected (6–20% employment decline in exposed roles)
2. **Information integrity crisis** – 52% of internet content now machine-generated; deepfakes doubling every 6 months
3. **Competitiveness gap** – EU controls only 4.8–5% of global high-end AI compute; 13.48% enterprise adoption vs. 41% for large firms
4. **Institutional adaptation** – only 6% of local governments prioritise AI; 77% of citizens distrust government AI use

Measures framework

- 19 no-regret measures applicable across all projections (probability = 1)
- 12 projection-conditional measures for P2/P3 preparation
- 5 appliedAI opportunity fields where we focus our contribution

How to use this document

- **Policy leaders:** Focus on no-regret measures and institutional capacity building
- **Company leaders:** Assess exposure across impact categories; build adoption capacity
- **Ecosystem partners:** Identify collaboration opportunities in measures inventory

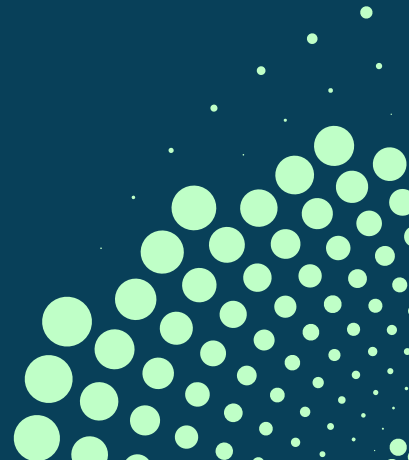
Give us your feedback

We welcome inputs across all aspects of this document:

- **Projections (Chapter 1):** Perspectives on projection framing, baseline conditions, and projection descriptions
- **Drivers & probabilities (Chapter 2):** Insights on main drivers, important trends, and probability assessments
- **Impacts (Chapter 3):** Observations on implications across impact fields; additional evidence or perspectives
- **Measures (Chapter 4):** Additions, ideas, and comments on proposed measures
- **Collaboration (Chapter 5):** If you would like to collaborate with us on our activities or if you want us to highlight your activities

Contact: [ai-projections@appliedai-institute.de]

Chapter 1: AI advancement projections (3-Year Horizon)



1.1 Purpose and scope of this chapter

This chapter presents three plausible projections for AI capability advancement over the next 36 months. The projections are distinguished solely by the speed of capability progress—not by the direction of societal response, regulatory choices, or adoption patterns. Those factors shape outcomes within each projection but do not define the projections themselves.

Why a 3-year horizon?

Beyond 36 months, uncertainty compounds to the point where projection planning loses decision-relevance. Feedback loops between capability gains, investment cycles, talent flows, and institutional responses become difficult to disentangle. We therefore constrain our outlook explicitly and recommend revisiting assumptions annually.

What this chapter does not do:

- It does not assign probabilities to projections (see Chapter 2).
- It does not identify the technical drivers that determine which projection materializes (see Chapter 2).
- It does not describe impacts on specific sectors or institutions (see Chapter 3).

Instead, Chapter 1 provides a decision-relevant framing: what each speed regime would mean in practical terms, and what baseline conditions constrain all projections equally.

1.2 Key definitions (Working Terms)

To ensure clarity across the document, we define the following terms as used in this whitepaper:

Term	Working Definition
AI advancement	Measurable increases in the capability and reliability of AI systems on economically and socially relevant tasks. This includes improvements in reasoning, planning, tool use, multimodal understanding, and agentic execution.
Agentic AI / Agents	AI systems that plan and execute multi-step tasks using tools, interfaces, and memory, operating under varying degrees of human supervision. Agents may coordinate with other agents or systems to accomplish goals.
Frontier models	State-of-the-art general-purpose AI models and the products built on them. These systems typically define the current capability ceiling for a given task category.
Automation frontier	The boundary of tasks that can be reliably performed by AI systems with acceptable quality, cost, and oversight. This frontier shifts as capabilities improve.
Capability speed	The rate at which the automation frontier expands and at which reliability on complex tasks improves. This is the single dimension that distinguishes our three projections.

1.3 Baseline conditions: constraints that shape all projections

Before describing the projections, we establish the baseline conditions that apply regardless of which projection materializes. These are structural realities that constrain the speed of change, the feasibility of responses, and Europe's room for manoeuvre.

1.3.1 Compute, energy, and infrastructure cycles

Current reality:

AI capability advancement, particularly for frontier models, depends on access to large-scale compute infrastructure. Training runs for leading models require tens of thousands of specialised accelerators (GPUs, TPUs, or custom ASICs) operating in purpose-built data centres with substantial power and cooling requirements. The top 4 US hyperscalers together are investing more than \$600 Billion in 2026 alone, more than the entire German federal budget! [1]

Constraint dynamics within 36 months:

- Data centre buildout is a multi-year process. Site acquisition, permitting, construction, and equipment installation typically require 2–4 years from decision to operation. Projects announced today will largely come online towards the end of our projection horizon or beyond.
- Grid capacity and energy supply are binding constraints in many regions. Even where capital is available, connecting large facilities to reliable power (especially low-carbon power) faces infrastructure and regulatory bottlenecks.
- Chip supply chains remain concentrated. Advanced semiconductor fabrication is dominated by a small number of facilities, primarily outside Europe. Lead times for new fabrication capacity are measured in years.
- Quantum computing will not materially affect AI capabilities within this timeframe. While quantum computing research continues to advance, practical quantum systems capable of accelerating AI training or inference remain years to decades away.

Practical implication:

Within our 36-month window, the physical infrastructure supporting AI advancement will be shaped primarily by decisions already made in unprecedented amounts. New announcements will matter more for the period after our horizon. In the near term, progress will depend heavily on efficiency gains: better algorithms, improved training recipes, optimised inference, and more effective use of existing hardware.

1.3.2 Europe’s dependency exposure

Current reality:

Europe occupies a structurally dependent position in several layers of the AI stack:

Layer	Dependency pattern
Chips	Advanced logic chips (sub-7nm) are fabricated almost exclusively outside Europe (and the US). European chip design capabilities exist but rely on non-European foundries.
Cloud infrastructure	Hyperscale cloud providers are predominantly US-based. European “alternatives” exist theoretically but lack comparable scale, geographic reach, and ecosystem integration. There is no single player properly suited for large scale AI use
Foundation models	The majority of frontier general-purpose models are developed by US-based (and increasingly Chinese) organisations. European model development is growing but remains significantly behind the capability frontier.
Talent	Europe produces strong AI research talent but faces competition for retention, particularly at the applied/engineering end of the spectrum.

Constraint dynamics within 36 months:

- Structural dependencies will not be resolved within this timeframe; building semiconductor fabs or hyperscale infrastructure requires longer cycles.
- However, Europe’s ability to deploy, govern, and adapt AI systems is not fully determined by supply-side dependencies. Competence in integration, responsible deployment, and institutional readiness are within European control.

Practical implication:

Europe’s strategic position over the next 36 months will be shaped less by whether it resolves structural dependencies and more by whether it builds operational competence: the ability to adopt AI effectively, govern it responsibly, and maintain agency over critical systems despite upstream dependencies.

1.3.3 Institutional and organisational capacity

Current reality:

The speed at which AI capabilities translate into real-world impact depends on institutional capacity to adopt, integrate, and govern these systems. This capacity varies widely:

- Public sector: Procurement cycles are slow (often 12–24 months for significant IT projects). Legal and compliance capacity is stretched. IT modernisation backlogs are substantial. Workforce AI literacy is low in most agencies.
- Large enterprises: Adoption is accelerating but unevenly. Many organisations are in pilot or early deployment phases. Integration with legacy systems, data governance, and change management remain significant barriers. [2][150]
- SMEs: Awareness is growing, but capacity to adopt is constrained by resources, expertise, and access to tailored solutions.

Constraint dynamics within 36 months:

- Institutional capacity does not scale quickly. Training programmes, procurement reform, and organisational change take time.
- Bottlenecks are not primarily technical; they are organisational, legal, and cultural.
- Even in an “accelerated” capability projection, adoption will be gated by how fast institutions can adapt.

Practical implication:

In all projections, the execution gap, the difference between what AI systems can do in principle and what organisations actually achieve with them, will be a central determinant of outcomes. Building institutional capacity is a no-regret priority regardless of which projection materializes.

1.3.4 Competitive pressure and forced adaptation

Current reality:

While institutional capacity constrains adaptation speed, competitive pressure can override organisational inertia. In globally competitive markets, the pace of adoption is increasingly set not by internal readiness but by external necessity.

Emerging dynamics:

- **Speed-to-market compression:** Organisations that integrate AI into product development and operational workflows can dramatically reduce cycle times. When a competitor releases software updates in weeks rather than months, or responds to customer requests with high-quality proposals in minutes rather than days, the market benchmark shifts for everyone.
- **Productivity arbitrage:** Firms that achieve higher output per employee through AI augmentation gain structural cost and quality advantages. Competitors that cannot match this productivity face margin pressure, talent drain, and market share loss.
- **Winner-take-more dynamics:** In markets where speed and quality advantages compound (e.g., software, professional services, design, engineering), early adopters may pull ahead faster than laggards can catch up.

The “San Francisco consensus”:

In July 2025, Eric Schmidt, former Google CEO, described a belief among Silicon Valley technology leaders that “in the next two to four years (the average is three years) the entire world will change.” This expectation of rapid, transformative change is driving investment decisions, talent allocation, and competitive behaviour in ways that affect European organisations regardless of their own beliefs about AI timelines. [3]

Constraint dynamics within 36 months:

- For organisations operating in globally competitive, exposed sectors, adaptation may not be optional. The choice becomes: adopt rapidly or face existential competitive risk.
- For sheltered sectors (parts of the public sector, regulated industries, local services), competitive pressure is less immediate but not absent. Indirect effects, such as talent expectations, citizen comparisons, and supply chain dependencies, will transmit pressure over time.

Practical implication:

Institutional readiness is a constraint, but market pressure is a forcing function. In projections with rapid capability gains (P2 and especially P3), organisations in competitive markets will be compelled to adapt faster than their internal change capacity would normally allow—with significant implications for workforce, governance, and organisational stability.

1.4 The three projections: differentiated by capability speed

We now describe three projections for AI capability advancement over the next 36 months. Each projection is defined by the speed at which AI systems improve in capability and reliability on economically and socially relevant tasks.

Important: These projections describe capability trajectories, not outcomes. Outcomes depend on how societies, institutions, and organisations respond—which is addressed in Chapters 3 and 4.

Projection 1 (P1): Plateau

AI capabilities improve only marginally compared to today's frontier models.

What this projection looks like:

In this projection, the rapid capability gains observed in recent years do not continue at the same pace. Progress becomes incremental rather than step-changing. The fundamental architecture and training paradigms that underpin current systems encounter diminishing returns or unforeseen bottlenecks.

Typical improvements in this projection:

- Incremental gains in reliability, latency, and cost efficiency
- Better user interfaces and integration tooling
- Modest improvements in specific domains through fine-tuning and specialisation
- Continued reduction in deployment friction (APIs, SDKs, enterprise connectors)
- What remains difficult:
- Robust long-horizon autonomy (multi-step tasks over extended timeframes with minimal supervision)
- Consistently correct reasoning under ambiguity and uncertainty
- Reliable operation in open-ended, high-stakes environments
- Verifiable alignment with complex human intentions
- Transfer of virtually trained models to real-world robotic applications (sim-to-real gap remains a significant barrier, limiting practical deployment of robots in unstructured environments)

Why this projection still matters:

While AI projection probabilities are highly contested among experts, P1 should be considered the generally accepted lower boundary for capability advancement. Thus, all impacts described for P1 should be understood as the minimum expected impacts, and all measures identified for P1 should be considered the absolute minimum response required.

A plateau in capability advancement does not mean a plateau in impact. Current AI systems are already capable enough to transform significant portions of knowledge work, administrative processes, and creative production. The question is whether organisations can adopt and integrate these capa-

bilities effectively. [4] [5]

In Projection 1, the primary driver of change is diffusion and adoption, not new capabilities. Organisations that invest in process redesign, workforce development, and systematic integration can realise substantial productivity gains. Those that do not will fall behind—not because of new AI breakthroughs, but because competitors extract more value from existing capabilities.

Projection 2 (P2): Continued pace

AI capabilities continue to improve at roughly the pace observed in recent years; agentic workflows become materially more useful.

What this projection looks like:

In this projection, the trajectory of recent years continues. Frontier models become meaningfully more capable through a combination of scaling, algorithmic improvements, better training data, and advances in post-training techniques (reinforcement learning, inference-time reasoning, tool use). Progress is steady but not explosive.

Typical improvements in this projection:

- Stronger planning and multi-step reasoning capabilities
- More reliable tool use and structured outputs (code, data manipulation, API calls)
- Improved grounding on enterprise data and context
- Better multimodal handling (text, image, audio, video in combination)
- Partially autonomous workflows for bounded domains (software development, analytics, customer service, document processing)
- Significant improvement in sim-to-real transfer for robotics, enabling practical deployment of robots in diverse fields of application (logistics, manufacturing, service environments)

New “normal” capabilities within 12–36 months:

- AI agents that can execute multi-step workflows with moderate supervision (e.g., draft → review → revise cycles in writing, coding, or analysis)
- Autonomous handling of well-defined operational tasks (scheduling, monitoring, routine decisions) under explicit controls
- Meaningful productivity gains in software engineering, back-office operations, and knowledge-intensive services
- Software development becomes substantially automated, with AI handling the majority of routine coding, testing, and debugging tasks
- Core organisational processes begin running with limited human supervision in leading organisations

Adoption pattern:

Diffusion remains uneven. Organisations with strong change capacity, data infrastructure, and AI-literate workforces pull ahead. Laggards, including many public sector bodies and SMEs, struggle to keep pace. The gap between leaders and followers widens.

Projection 3 (P3): Accelerated

AI capabilities compound rapidly; discontinuities in autonomy and R&D automation become plausible within 36 months. Some experts predict the emergence of very powerful AI, maybe even AGI (Artificial General Intelligence) systems within this timeframe.

What this projection looks like:

In this projection, multiple drivers of progress align and reinforce each other. Advances in architectures, training methods, or emergent capabilities produce step-changes that accelerate the rate of improvement itself. AI systems begin to contribute meaningfully to AI research and software engineering, creating feedback loops that compress timelines significantly.

The possibility of reaching highly capable, general-purpose AI systems (sometimes referred to in public discourse as approaching “AGI-level” capabilities) enters the plausible range within 36 months. While definitions vary and uncertainty remains high, the practical implication is that AI systems may achieve or approach human-level performance across a wide range of cognitive tasks, not just narrow domains.

Expert views on accelerated timelines:

Leading AI researchers and executives have offered increasingly specific predictions about AGI timelines during 2025–2026:

Expert	Organisation	Timeline view	Key statement	Source
Dario Amodei	Anthropic CEO	2026–2027	„A country of geniuses in a data centre“ matching Nobel Prize-level capabilities; 50% probability within 1–3 years	[6] [15]
Sam Altman	OpenAI CEO	Mid-2020s	„We are now confident we know how to build AGI“; superintelligence within 5 years	[7]
Demis Hassabis	DeepMind CEO	2028–2030	Revised from „5–10 years“ to „3–5 years away“; 50% by end of decade	[8] [9] [10] [15]
Eric Schmidt	Former Google CEO	2–6 years (avg. 3)	„San Francisco Consensus“ among tech leaders; recursive self-improvement within a decade	[3]

Shane Legg	DeepMind Co-founder	2028 (minimal AGI)	50% probability of minimal AGI by 2028; full AGI 3–6 years later	[11]
Ilya Sutskever	SSI CEO	Research-dependent	2025 onward as return to „age of research“ where scaling alone insufficient	[12]
Yoshua Bengio	AI Pioneer	Few years to decade	„Many researchers now consider human-level AI plausible within a few years to a decade“	[13] [15]
Yann LeCun	Former Meta Chief Scientist (left December 2025)	Not within 2 years, but „probably very soon“	Argues current LLM approaches alone insufficient; pursuing world models at AMI Labs	[14] [15]
Llion Jones	Sakana.ai Co-founder	Focused on recursive self-improvement	Estimates 1–2 breakthroughs needed for AGI; believes recursive self-improvement is the key path	[15]

The architecture debate:

A fundamental disagreement exists among experts about whether AGI requires new architectures beyond transformers: [

- Scaling advocates (Altman, Amodei): Current transformer paradigm with refinements (reasoning modules, tool use, inference-time compute) is likely sufficient. OpenAI's o-series models achieving 75.7% on ARC-AGI (vs. GPT-4o's 5%) suggest transformers with reasoning can adapt to novel tasks. [16]
- Architecture innovators (LeCun, Jones, Markram): Transformers face fundamental limitations. Yann LeCun left Meta in December 2025 to launch AMI Labs (€3B valuation) focused on «world models» rather than scaling LLMs. Llion Jones co-founded Sakana.ai to focus on recursive self-improvement, estimating 1–2 key breakthroughs are needed. Henry Markram's INAIT is pursuing causal learning as the foundation for understanding AI rather than pattern-matching AI. [14] [17] [18]
- Hybrid perspective (Hassabis, Bengio): Additional breakthroughs needed, particularly continual learning, world models, reasoning, hierarchical planning, and creative hypothesis generation, but these may build on rather than replace current paradigms. [19]

Typical improvements in this projection:

- Rapid cycles of capability increases, with shorter intervals between major advances
- Substantially more autonomous execution across broad task categories
- Meaningful AI contribution to research and engineering workflows (experiment design, code generation, debugging, optimisation, scientific discovery)
- Emergence of capabilities that were not explicitly trained or anticipated

- Robotic applications expand dramatically, including viable deployment of humanoid robots in commercial and industrial settings, enabled by breakthroughs in sim-to-real transfer, perception, and manipulation

Agentic capabilities:

- Autonomy in substantial end-to-end processes becomes viable, not just bounded subtasks
- AI systems can manage complex, multi-stage projects with minimal human intervention
- Coordination between multiple AI agents becomes routine for complex workflows
- Human oversight shifts from task-level supervision to goal-level governance

Stress points:

An accelerated projection creates significant stress across institutions and systems:

- Policy and regulation: Governance frameworks designed for slower change may become outdated before implementation is complete. The gap between capability and governance widens.
- Education and workforce: Curricula and training programmes cannot adapt quickly enough; cohorts of workers face rapid skill obsolescence across many professions simultaneously.
- Public services: Demand for support (unemployment, retraining, fraud prevention, social stability) may spike while institutional capacity lags severely.
- Information integrity: The scale and sophistication of synthetic content (text, image, audio, video) may fundamentally outpace detection and attribution capabilities, challenging the foundations of trust in digital communication.
- Economic structure: Fundamental questions about value distribution, employment, and economic organisation may arise faster than political systems can address them.

Competitive dynamics and forced adoption:

In this projection, competitive pressure becomes extreme. Organisations in globally competitive markets face existential pressure to adopt AI at a pace that may exceed their capacity for orderly change:

- Product and service cycles compress dramatically; organisations that cannot keep pace lose market position within months, not years.
- The productivity gap between AI-native organisations and traditional competitors becomes a chasm.
- Workforce transformation cannot wait for gradual reskilling; organisations must simultaneously deploy AI and manage significant workforce restructuring.
- Some organisations will thrive by embracing radical transformation; others will face severe disruption or failure despite best efforts, simply due to the speed of change.

1.5 Projection comparison: one-page decision aid

The following table summarises the three projections across key dimensions. It is intended as a quick reference for decision-makers, not a comprehensive analysis.

Dimension	P1: Plateau	P2: Continued Pace	P3: Accelerated
Capability speed	Marginal improvements; gains mostly in cost/efficiency	Steady capability gains; regular new model generations	Compounding gains; potential discontinuities; very powerful systems possible
Automation frontier	Broadly similar to today; expands via adoption	Expands into multi-step workflows and bounded autonomy	Expands rapidly; autonomy in substantial end-to-end processes
Reliability trajectory	Improves slowly; edge cases remain problematic	Improves materially; deployment confidence grows	Uneven but fast; reliability sufficient for autonomous operation in many domains
Agentic viability	Limited; supervision-heavy	Growing; practical for bounded enterprise tasks	Autonomy in substantial end-to-end processes; goal-level oversight replaces task-level supervision
Robotics	Sim-to-real gap persists; limited practical deployment	Diverse robotic applications become viable	Broad deployment including humanoids; robotics becomes economically transformative
Main bottleneck	Adoption and integration capacity	Adoption + governance + workforce readiness	Societal and institutional adaptation speed; governance gap
Diffusion pattern	Leaders gain from existing capabilities	Leaders widen gap over laggards	Extreme divergence; forced rapid adoption in exposed sectors; potential for systemic shocks
Europe's primary risk	Complacency and slow diffusion	Widening competitiveness gap	Loss of control/agency; dependency deepens; institutional overload
Urgency of response	Moderate (but not negligible)	High	Very high

1.6 What this chapter does not determine

To avoid misinterpretation, we explicitly note what this chapter leaves open:

1. Projection probabilities: We do not assign likelihoods in this chapter. Probability estimates require analysing the underlying drivers (Chapter 2).
2. Which projection is “good” or “bad”: Each projection contains both risks and opportunities. Outcomes depend on responses, not just capability trajectories.
3. Specific impacts: How each projection affects labour markets, public services, democracy, or competitiveness is addressed in Chapter 3.
4. What to do: Measures and priorities are derived in Chapter 4, after impacts are assessed.
5. Technological drivers: The technical factors that determine which projection materializes—architectures, agentic reliability, R&D automation, robotics—are analysed in Chapter 2.

1.7 How to read the rest of this document

With the projection frame established, the remainder of this whitepaper proceeds as follows:

- Chapter 2 identifies the key technical drivers that determine projection probabilities and provides our current best estimates.
- Chapter 3 maps concrete impacts (opportunities and risks) across sectors and systems for each projection.
- Chapter 4 derives measures from those impacts, distinguishes no-regret actions from projection-conditional ones, and specifies the appliedAI Institute's portfolio response.
- Chapter 5 describes how we keep this document current: continuous monitoring, expert input, and versioning protocols.

The goal is a decision-ready orientation: not a prediction of what will happen, but a structured basis for preparing across plausible futures.

Chapter 2: Speed drivers & projection probabilities

2.1 Purpose of this chapter

Chapter 1 established three projections differentiated solely by the speed of AI capability advancement. This chapter identifies the technical drivers that determine which projection materializes and provides our current best-estimate probability assessments.

The logic:

1. Identify a small set of drivers that are (a) highly uncertain within 36 months, (b) meaningfully affect capability speed, and (c) can be tracked via observable signals.
2. For each driver, describe what P1/P2/P3 would look like for that driver specifically.
3. Consolidate driver assessments into projection probabilities.
4. Specify breakthrough triggers that would prompt re-estimation.

Note on probability estimates: We publish point estimates as a decision aid, not as precise forecasts. These estimates reflect our current best judgment informed by available evidence. Probabilities are re-estimated when monitoring detects material breakthroughs or disruptions.

2.2 Driver selection criteria

A topic belongs in this chapter only if it meets all three criteria:

Criterion

Highly uncertain within 12-36 months

Rationale

If the outcome is predictable, it is a baseline condition (Chapter 1), not a driver.

Criterion

Meaningfully changes capability speed

Rationale

If progress on this dimension does not affect how fast the automation frontier expands, it belongs elsewhere (e.g., adoption, governance).

Criterion

Trackable via observable signals

Rationale

If we cannot monitor progress, we cannot update probabilities.

What is excluded from this chapter:

- Compute/energy buildout: Largely determined by decisions already made; affects all projections similarly within 36 months. (Covered in Chapter 1 as a baseline condition.)
- Open-source ecosystem health: Affects Europe's ability to respond, not capability speed itself. (Relevant to Chapter 4.)
- Regulation and standards: Affects deployment and governance, not fundamental capability advancement. (Relevant to Chapters 3 and 4.)
- Security and misuse: These are impacts of capability advancement, not drivers of it. (Covered in Chapter 3.)

2.3 The four core drivers

Based on the selection criteria, we identify four drivers that will determine which projection materializes:

1. Architectures & Training Paradigms
2. Agentic Autonomy & Orchestration
3. Automation of AI R&D and Software Engineering
4. Robotics & Embodied AI

Each driver is analysed below with:

- Definition and mechanism
- What P1/P2/P3 looks like for this driver
- Current evidence and expert views
- Signals and indicators to monitor
- Directional impact on projection probabilities

2.4 Driver 1: Architectures & training paradigms

2.4.1 What this driver means

This driver captures whether AI capability gains continue to come primarily from scaling and engineering improvements within current paradigms (primarily transformer-based architectures), or whether new architectural approaches emerge that unlock qualitatively different capabilities.

Key questions:

- Do we see continued gains from scaling (more compute, more data, larger models)?
- Do post-training techniques (RLHF, inference-time reasoning, chain-of-thought) continue to yield meaningful improvements?
- Does a genuinely new paradigm emerge (e.g., world models, hybrid neuro-symbolic systems, nested learning, causal AI, or fundamentally different architectures)?

Emerging paradigms: nested learning and causal ai

Two significant architectural innovations have emerged in 2025–2026 that may represent pathways beyond transformer limitations:

Google DeepMind's Nested Learning (NeurIPS 2025):

Nested Learning, introduced in the paper “Nested Learning: The Illusion of Deep Learning Architectures,” bridges the gap between model architecture and optimisation algorithms by treating them as a unified system of interconnected, multi-level optimisation problems. Key innovations include:

- Continuum Memory Systems (CMS): Memory is viewed as a spectrum of modules, each updating at different frequency rates, creating richer memory systems for continual learning—directly addressing the «catastrophic forgetting» problem that limits current LLMs.
- Deep Optimizers: By viewing optimizers as associative memory modules, the approach derives new formulations for concepts like momentum that are more resilient to imperfect data.
- Hope Architecture: A proof-of-concept self-modifying architecture that achieves superior performance in language modelling and demonstrates better long-context memory management than existing state-of-the-art models including Titans, TTT, and Mamba2.

The Hope architecture demonstrates that principled unification of architecture and optimisation can lead to more expressive, capable, and efficient learning algorithms—potentially closing the gap between current LLM limitations and human-like continual learning abilities. [20] [21]

EPFL/INAIT Causal AI:

Henry Markram, who led the Blue Brain Project and Human Brain Project, has launched INAIT to commercialise causal learning—a fundamentally different approach to AI that focuses on understanding cause and effect rather than pattern correlation:

- Causal vs. Correlational Learning: Current AI observes patterns; INAIT's approach enables AI to participate in and understand the causal dynamics of systems, a prerequisite for genuine understanding and autonomous action.
- Three Product Lines: Future Complete (causal forecasting), Cognition Complete (causal decision-making), and Behaviour Complete (full-scale robotics with continuous learning).
- Transparency Advantage: Causal AI provides inherent explainability because decision chains can be traced through causal relationships, addressing a key limitation of current «black box» neural networks.
- Commercial Deployment: INAIT's Future Complete product launches in 2026, with Microsoft as distribution partner, targeting financial services, supply chain, and robotics applications.

Markram describes this as moving from “knowledge AI” (pattern matching) to “understanding AI” (causal reasoning), a transition that could be as significant as the original deep learning revolution. [18] [22]

2.4.2 Expert views on architectures and timelines

Leading AI experts hold divergent views on the role of new architectures in reaching very powerful AI systems:

Expert	Organisation	AGI Timeline	Architecture View	Breakthroughs Needed	Source
Dario Amodei	Anthropic	2026–2027	Current paradigm + scaling + AI-assisted R&D likely sufficient	Very few; mostly continuation	[6]
Sam Altman	OpenAI	Mid-2020s	Transformer-style models extended with tools, agents, compute	Small number of insights + engineering	[7]
Demis Hassabis	DeepMind	2028–2030	Needs world models, continual learning, planning, robotics–hybrid systems	Several major but related advances	[8] [9] [10]
Yann LeCun	Former Meta Chief Scientist / AMI Labs	Not within 2 years, but „probably very soon“	LLMs alone insufficient; needs new architectures with world models	Multiple breakthroughs; new paradigm	[14]

Llion Jones	Sakana.ai Co-founder	Focused on recursive self-improvement	Recursive self-improvement is the key path; current paradigm may be local optimum	1-2 key breakthroughs needed for AGI	[15]
Shane Legg	DeepMind	2028 (minimal)	Distinguishes minimal vs. full AGI; scaling may suffice for minimal	Minimal: few; Full: several more	[11]
Henry Markram	EPFL / INAIT	Imminent for domain-specific	Causal learning required for understanding; current AI only correlates	Causal AI solves key limitation	[18]

The architecture debate in detail:

Transformer advocates argue that the current paradigm, with refinements, can reach AGI:

- OpenAI's o-series models demonstrate that inference-time reasoning can dramatically improve performance on novel tasks (o3 achieved 75.7% on ARC-AGI vs. GPT-4o's 5%)
- Continued improvements in training recipes, data curation, and post-training techniques suggest substantial headroom remains

Architecture innovators argue that fundamental limitations require new approaches:

- Yann LeCun stated at Davos 2026 that AGI will not arrive within 2 years but is "probably coming very soon," and that transformers cannot learn causality or understand physical dynamics. He launched AMI Labs (€3B valuation) in December 2025 specifically to pursue "world models."
- Llion Jones (transformer co-inventor) co-founded Sakana.ai to focus on recursive self-improvement, estimating that 1-2 key breakthroughs are needed for AGI. He warns of a "gravitational well" where transformer success prevents exploration of alternatives.
- Henry Markram argues that current AI captures knowledge but doesn't understand it—AI is an "observer" while brains are "participants" in the world. INAIT's causal learning enables AI to understand cause and effect, which he considers prerequisite for true intelligence.
- Missing capabilities identified by multiple experts: continual learning, world models, hierarchical planning, creative hypothesis generation, causal reasoning.

Hybrid perspectives suggest breakthroughs within or extending current paradigms:

- Hassabis identifies specific missing capabilities but frames them as extensions within deep learning
- Google's Nested Learning approach demonstrates that continual learning can be achieved through architectural innovation within neural network paradigms

- Joint Embedding Predictive Architecture (JEPA) variants from Meta represent attempts to add world-model capabilities to neural approaches
- Neuro-symbolic integration combining neural networks with structured reasoning

2.4.3 Driver-specific projection mapping

Projection 1

Plateau

What This Looks Like for Architectures.

No major paradigm shift. Scaling yields diminishing returns. Post-training techniques saturate. Gains are incremental (cost, latency, reliability) rather than capability expansion.

Projection 2

Continued Pace

What This Looks Like for Architectures.

Continued engineering gains within current paradigms. Post-training advances (inference-time reasoning, tool use, RLHF improvements) drive measurable capability increases. No single breakthrough, but steady compounding.

Projection 3

Accelerated

What This Looks Like for Architectures.

Either (a) scaling + engineering continues to yield surprising gains, or (b) a paradigm shift emerges—Nested Learning, Causal AI, world models, hybrid architectures, or other innovations that materially increase reasoning depth, planning horizon, or self-improvement capability.

2.4.4 Signals and indicators to monitor

Signal	Why It Matters	What We Track	Current Read	Confidence
ARC-AGI-2 and novel reasoning benchmarks	Tests genuine adaptation with both high adaptability and high efficiency;	ARC-AGI-2 scores, cost efficiency	Human Panel: 100% at \$17; Gemini 3 Deep Think: 84.6% at \$13.62	Medium-High
New architecture announcements with verified gains	Would signal paradigm shift	Major lab releases, benchmark results, research publications	AMI Labs, Nested Learning, INAIT, Sakana.ai launched; multiple approaches emerging	Medium

Architectural innovation research	Tests whether new paradigms emerge beyond transformers	Papers on continual learning, causal AI, recursive self-improvement, world models, neuro-symbolic integration	Active research across multiple organisations; growing publication volume	Medium-High
Compute infrastructure scaling	Tests continued validity of scaling laws	Data centre capacity announcements, training run scales	OpenAI: 1 GW capacity early 2026, additional 1 GW by end of year; major scaling investments continuing	High
World model demonstrations	Core capability gap identified by sceptics	DeepMind Genie, Meta JEPAs, AMI Labs	Research active; no production deployment	Low-Medium

2.4.5 Directional impact on projection probabilities

- Evidence of paradigm shift (Nested Learning, Causal AI, recursive self-improvement, or other approaches achieving production deployment with verified superior performance): shifts probability towards P3 [23]
- Sustained engineering gains without paradigm shift: supports P2
- Diminishing returns on scaling and post-training: shifts towards P1

Current Assessment: The emergence of multiple promising architectural innovations—including Nested Learning (DeepMind, NeurIPS 2025) demonstrating superior continual learning and memory management, INAI's commercial launch of Causal AI with Microsoft partnership, Richard Socher's work on recursive self-improvement, AMI Labs pursuing world models, and other organisations exploring neuro-symbolic integration—provides strong evidence that architectural innovation is accelerating across the field. These are not isolated efforts but represent a broader trend of multiple companies and research groups pursuing alternatives and extensions to current paradigms.

Furthermore, it is reasonable to assume that architectural innovations will continue to emerge. We observe how efficiently the biological brain operates compared to current AI models—the brain achieves remarkable capabilities with approximately 20 watts of power and learns from vastly less data than current models require. This efficiency gap suggests substantial room for improvement in AI architectures.

Additionally, all scaling laws for pre-training, mid-training, and post-training continue to hold. OpenAI's expansion to 1 GW of data centre capacity in early 2026, with an additional 1 GW planned by year-end, indicates continued confidence in scaling approaches. It is reasonable to assume that with significantly larger compute resources, substantially better models will emerge.

The combination of (a) multiple promising architectural paths reducing P1 probability, (b) strong continued investment in scaling supporting P2, and (c) potential for architectural breakthroughs or scaling surprises supporting P3, creates the current probability distribution where both P2 and P3 have substantial weight.

2.5 Driver 2: Agentic autonomy & orchestration

2.5.1 What this driver means

This driver captures the practical ability of AI systems to execute multi-step goals using tools, memory, and coordination patterns—with acceptable reliability, cost, and oversight requirements.

Key questions:

- Can AI agents reliably complete bounded but meaningful workflows (not just single tasks)?
- How fast is the “time horizon” of reliable autonomous operation expanding?
- Can multiple agents coordinate effectively on complex tasks?
- What supervision level is required for acceptable error rates?

2.5.2 Current benchmark evidence (February 2026)

The benchmark landscape for agentic AI has matured substantially, with dramatic improvements across all major benchmarks:

Real-world task performance:

Benchmark	What It Measures	Best Performance	Leading Model/System	Human Baseline	Gap	Source
SWE-bench (bash)	Real GitHub issue resolution	76.8%	Claude Opus 4.5	~70–80%	Closed/Exceeded	[24]
SWE-bench Verified	Human-verified solvable issues	79.2%	Claude Opus 4.5 (High Compute)	100% (by design)	21 points	[25]
GAIA	Multi-step research/reasoning tasks	74.55%	Claude Sonnet 4.5	92%	17 points	[26]
WebArena	Interactive web task completion	74.3%	DeepSeek v3.2	~85%	11 points	[27]
OSWorld	Real OS environment tasks	73%	Multimodal agent system	72%	Closed	[28]

Key observations:

- SWE-bench has reached human parity. The benchmark uses 2,294 real GitHub issues; success requires understanding complex codebases, identifying root causes, and implementing working patches. Claude Opus 4.5 at 79.2% now matches or exceeds estimated human performance.
- OSWorld gap has closed dramatically. The jump from 12.24% to 73% represents one of the most significant capability improvements in the past year, demonstrating that real computer environment interaction—previously the largest gap—is now approaching human-level performance.
- GAIA shows strong progress towards the 92% human baseline, with Claude Sonnet 4.5 achieving 74.55% on multi-step research and reasoning tasks requiring tool use and complex problem-solving.
- WebArena has seen major improvement with DeepSeek v3.2 achieving 74.3%, substantially closing the gap to human performance and demonstrating that web automation is becoming practically viable.

Time-horizon analysis:

Research suggests frontier agents can now reliably complete tasks equivalent to hours of expert human work for bounded domains, with effective time horizons expanding rapidly. Key benchmarks measuring task duration capability:

METR task horizon benchmark:

METR proposes measuring AI performance in terms of the length of tasks AI agents can complete. Their research shows that this metric has been consistently exponentially increasing over the past 6 years, with a doubling time of around 7 months. Extrapolating this trend predicts that, by the end of the 3-year timeframe of this document, we will reach 2 weeks of continuous work done by agents within a fraction of a time. [29]

- Current state of the art: GPT-5.2 achieves 6 hours 34 minutes task length for software tasks completed 50% of the time. [30]

Vending-bench 2:

Vending-Bench 2 is an advanced AI benchmark developed by Andon Labs (released late 2025/early 2026) that evaluates the long-term, agentic capabilities of Large Language Models. It simulates a 365-day, high-stakes projection where an AI agent must operate a vending machine business, focusing on maintaining coherence, managing inventory, setting prices, and maximising profit over a long time horizon (>20M tokens).

- Current state of the art: Claude Opus 4.6 achieves approximately \$7,500 profit over the simulated year.
- Progress indicator: Six months ago, frontier models achieved roughly \$0 profit on this benchmark, indicating dramatic improvement in long-horizon coherence and planning.

The SWE-bench results demonstrate that agents can now handle substantial software engineering tasks that previously required experienced developers working over extended periods. Combined with METR's exponential trend and Vending-Bench 2's demonstration of year-long planning coherence, the evidence indicates that agentic time horizons are expanding faster than many anticipated. [31]

2.5.3 OpenClaw: The emergence of system-integrated agents (updated)

OpenClaw (formerly Clawdbot or Moltbot) has emerged as one of the most significant frameworks for agentic autonomy, marking the transition from passive chatbots to proactive, system-integrated agents that can act independently. Its emergence surprised the entire AI community with how quickly system-level agent integration became practically viable.

Core capabilities:

- **System Integration (“AI with Hands”):** Unlike isolated web interfaces, OpenClaw runs locally (macOS, Linux, Windows) with root-level access to the file system, shell commands, and browser control.
- **Proactive Action:** Through «heartbeats» and cron jobs, agents can initiate tasks (e.g., portfolio monitoring, email triage) without direct user prompts.
- **Persistent Memory:** Stores interaction histories and user preferences locally (e.g., as Markdown), enabling adaptive, session-consistent personality.

Orchestration architecture:

- **Model-Agnostic Routing:** Functions as a gateway that dynamically routes tasks to different LLMs (Claude, GPT, DeepSeek), including automatic failover chains.
- **Declarative Skills System:** Capabilities defined through structured SKILL.md files, enabling controlled and auditable orchestration where agents access only context-relevant tools.
- **Multi-Agent Teams:** Users can configure specialised agent families that delegate tasks and track progress among themselves. [32]

Security concerns:

Despite its capabilities, OpenClaw is considered an “identity security nightmare” for enterprises:

- **Privilege Escalation:** Agents often inherit full user permissions, which can lead to critical system interventions with unpredictable behaviour.
- **Vulnerability Rate:** Approximately 26% of community-provided «skills» showed security vulnerabilities in analyses.
- **Decentralisation:** Because OpenClaw runs locally and distributed, there is no central «kill switch,» complicating regulatory control.

OpenClaw represents the leading tool for users seeking maximum functional independence for their agents, but requires high technical competence in security and governance. [33]

Implications for Projection Assessment:

The rapid emergence and adoption of OpenClaw demonstrates that system-integrated agentic AI is not a future possibility but a current reality. This has several implications:

1. Capability validation: The fact that users are successfully deploying agents with file system access, browser control, and proactive task initiation validates that agentic reliability has reached practical thresholds.
2. Security concerns become urgent: The 26% vulnerability rate in community skills indicates that agentic deployment is outpacing security governance—a pattern that will likely intensify.
3. Enterprise adoption pressure: Despite security concerns, competitive pressure will drive enterprise adoption of similar capabilities, forcing organisations to develop governance frameworks rapidly.

Outlook:

We expect in 2026 that this space will further develop along two tracks:

- Corporate versions: Enterprise-grade variants of OpenClaw with improved security controls, audit capabilities, and centralised governance will emerge from major vendors.
- Malicious use: Bad actors will adopt similar tools for automated fraud, social engineering at scale, and autonomous cyberattack orchestration, requiring defensive capabilities to advance in parallel.

2.5.4 Driver-specific projection mapping



2.5.5 Signals and Indicators to Monitor

Signal	Why It Matters	What We Track	Current Read	Confidence
SWE-bench progression	Direct measure of practical coding capability	Quarterly leaderboard updates	77.5-82% best; at/above human level	High
OSWorld progression	Real environment interaction	Leaderboard updates	73%; dramatic improvement	High
GAIA scores	Holistic assistant capability	Top scores, gap to human	74.55% vs. 92% human	High
Enterprise deployment announcements	Market validation	Major vendor releases, case studies	Growing rapidly; production deployments	High
OpenClaw adoption and security incidents	System-integrated agent maturity	Community metrics, incident reports	High adoption; security concerns documented	Medium
Multi-agent coordination results	Complex workflow viability	Research benchmarks, production deployments	Active development; production emerging	Medium

2.5.6 Directional Impact on Projection Probabilities

- Continued benchmark improvement + successful enterprise deployment: strongly supports P2/P3
- OSWorld gap widens: strong signal towards P3
- Major security incidents with agentic systems: could slow deployment but unlikely to reverse capability
- SWE-bench exceeding 90%: would indicate P3 trajectory

Current Assessment: The dramatic improvement in SWE-bench (to 77-80%, matching human performance), OSWorld (from 12% to 73%, matching human performance), and GAIA (to 74.55%) represents transformative progress. Combined with the emergence of system-integrated frameworks like OpenClaw, the evidence strongly supports P2 as the baseline with significant probability of P3. The agentic capability gap that previously distinguished projections has largely closed within the past year.

2.6 Driver 3: Automation of ai r&d and software engineering

2.6.1 What this driver means

This driver captures the extent to which AI systems can accelerate the production of AI capabilities themselves—creating potential feedback loops that compress timelines.

Key questions:

- How much of software engineering work is AI-assisted or AI-automated?
- Are AI systems contributing meaningfully to AI research (experiment design, hypothesis generation, code optimisation)?
- Is there evidence of “recursive self-improvement”—AI systems improving AI systems?

2.6.2 Current evidence (February 2026)

Code generation benchmark status:

Benchmark	Best Performance (Pass@1)	Leading Model	Details	Source
HumanEval	96.2%	o1-mini	Extremely high accuracy on standard Python tasks	[34]
MBPP	~90-92%	GPT-5 / Claude 4.5	Estimated values based on HumanEval correlation	[129]
LiveCodeBench	91.7%	Gemini 3 Pro	Contamination-free evaluation on fresh problems	[35]
SWE-bench Verified	79.2%	Claude Opus 4.5	Real-world software engineering	[24] [25]

Key insight: HumanEval is now saturated at 96%+, representing superhuman performance on standard coding benchmarks. The more significant metric is SWE-bench, where 79% on verified issues demonstrates that AI can now handle the majority of real-world software engineering tasks.

Claude Code: The Transformation of Software Development

Claude Code has become the definitive tool for agentic autonomy in software development and AI R&D. Unlike traditional code assistants that primarily offered writing assistance, Claude Code operates as an autonomous “AI Engineer.”

Impact on software engineering:

- **Autonomous Debugging:** Analyses error messages, searches for root causes across entire repositories, and independently implements fixes including tests.
- **Legacy Modernisation:** Excels at understanding outdated codebases and performing large-scale refactoring that exceeds the context limits of conventional tools.
- **Productivity Impact:** Anthropic reports that Claude Code now writes approximately 90% of the company's code and has increased engineer productivity by approximately 67%. [36] [37]

Impact on AI R&D automation:

- **Experiment Automation:** Can set up research environments, write hyperparameter scripts, and autonomously evaluate test run results.
- **Knowledge Integration:** Through the Model Context Protocol (MCP), Claude Code connects directly with scientific databases (e.g., PubMed, bioRxiv) or lab tools to automate research pipelines.
- **Agent Orchestration:** Developers use the tool to create agent teams that work in parallel on different research subtasks. [38] [39] [40]

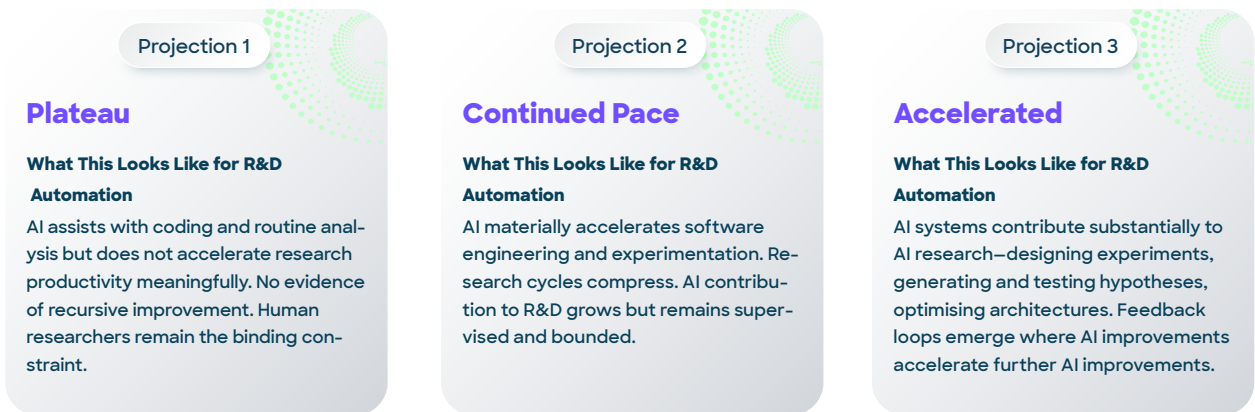
R&D automation evidence:

Domain	Current State	Evidence
Code assistance	Mainstream; dominant	90% of Anthropic code AI-written; ~67% productivity increase; majority of developers at AI labs use assistants daily
Experiment automation	Production-scale	AI systems running experiment pipelines autonomously; automated ML standard practice
Research contribution	Growing rapidly	„AI Scientist“ systems producing workshop-level papers; all major labs report using models for internal R&D
Recursive improvement	Materializing	AI-generated synthetic training data; AI-assisted evaluation and red-teaming; automated architecture search; AI writing AI code

Lab statements on internal AI use:

- Anthropic: ~90% of code written by Claude Code; 67% productivity increase reported
- OpenAI: Uses models for code review, test generation, experiment scripts, security review
- DeepMind: AutoML, NAS, automated experiment pipelines standard practice
- Meta: LLaMA-based tools for code generation, data labelling, documentation

2.6.3 Driver-Specific projection mapping



2.6.4 Signals and Indicators to Monitor

Signal	Why It Matters	What We Track	Current Read	Confidence
SWE-bench Verified scores	Real-world engineering capability	Quarterly updates	79% (Claude Opus 4.5)	High
LiveCodeBench scores	Contamination-free code capability	Quarterly updates	91.7% (Gemini 3 Pro)	High
AI code contribution metrics	R&D acceleration	Lab announcements, productivity data	90% at Anthropic; widespread internal use	High
AI-authored research	Indicates research contribution	Publications, „AI Scientist“ results	Growing; workshop-quality demonstrated	Medium
Closed-loop improvement evidence	Compounding dynamics	Research publications, capability jumps	Strong evidence emerging	Medium-High

2.6.5 Directional impact on projection probabilities

This is the key “swing factor” between P2 and P3. The evidence that AI systems are now contributing substantially to AI research and software development—with Anthropic reporting 90% of code AI-written—suggests feedback loops are materialising.

- Strong evidence of closed-loop R&D acceleration: shifts probability towards P3
- AI assists but does not fundamentally accelerate research: supports P2
- Limited impact on research productivity: shifts towards P1

Current Assessment: The Claude Code metrics (90% of code AI-written, 67% productivity increase) provide concrete evidence that R&D automation feedback loops are no longer theoretical—they are operational at frontier AI labs. This represents a qualitative shift from the previous assessment and significantly increases the probability of P3. The recursive improvement dynamic that was speculative 12 months ago now appears to be materialising.

2.7 Driver 4: Robotics & embodied AI

2.7.1 What this driver means

This driver captures the speed at which AI capabilities translate into robust physical-world autonomy that is economically deployable at scale.

Key questions:

- Is sim-to-real transfer becoming reliable enough for practical deployment?
- Are humanoid and other advanced robots approaching economic viability?
- How fast is deployment expanding beyond controlled environments?

2.7.2 Current evidence (February 2026)

China's robotics revolution:

China's humanoid robotics industry has undergone a remarkable transformation, transitioning from experimental prototypes to commercial-scale production at unprecedented pace. The convergence of government investment exceeding \$25 billion, advanced manufacturing infrastructure, and talent from leading tech companies has positioned Chinese manufacturers at the forefront of solving sim-to-real transfer challenges. [41] [42] [43]

Market leadership:

- AgiBot shipped 5,100 humanoid robots in 2025—the first manufacturer globally to achieve this milestone—capturing 39% of global market share (Omdia). [44] [45]
- UBTECH has begun mass production of Walker P2 with trajectory to 5,000 units in 2026, doubling to 10,000 by 2027. Orders exceed 800 million yuan (~\$112 million). [46] [47]
- Unitree deployed ~5,000 G1 units in H1 2025 to Amazon, Stanford, MIT, and UT Austin at \$13,500 per unit. [48]
- Xpeng targets mass production of IRON humanoid by end of 2026, with plans for 1 million units annually by 2030. [49] [50]

Sim-to-Real Transfer: Now Solved for Production Workloads:

The critical sim-to-real transfer challenge has been overcome for practical manufacturing applications:

- Figure AI at BMW: Over 1,250 operational hours spanning 11 months at BMW Plant Spartanburg. Figure 02 loaded 90,000+ parts, contributed to 30,000+ X3 vehicle production. Achieved >99% placement accuracy per shift with zero human interventions required. [51]
- CATL Battery Assembly: Spirit AI's Xiaomi robot deployed at world's largest EV battery manufacturer, plugging high-voltage battery connectors at 99% success rate while maintaining 3x human daily workload. Autonomously detects wiring anomalies in real-time. [52]

These deployments validate that simulation-trained policies achieve sufficient reliability for economic deployment in complex manufacturing environments.

Cost trajectory:

Time Frame	Unit Cost Range	Evidence	Source
2023-2024	\$150,000-\$500,000	Low-volume production, custom components	[53]
2026	\$30,000-\$150,000	Current production; varies by capability	[54]
2027-2028	\$20,000-\$30,000	Documented 20-30% annual cost reduction	[54]
2030	Sub-\$20,000	Tesla, Chinese manufacturers target	[55]

Key enabler: 90%+ of humanoid robot components now manufactured domestically in China (except certain computing chips), enabling vertical integration and economies of scale unavailable to geographically distributed supply chains. [55]

Economic viability:

ROI calculations now strongly favour deployment in high-labour-cost environments:

- US Manufacturing: At \$156,000 annual loaded labour cost, humanoid robots achieve payback in ~1.9 months (single shift) or ~1.0 month (dual shift replacing two workers).

- 5-Year ROI: 2,070% under baseline assumptions; 1,400% under conservative 70% task automation projections. [56]
- Healthcare: Rehabilitation robots achieve 3-6 month payback with 800%+ 5-year ROI due to \$150-400/session billing rates.
- Robot-as-a-Service: \$6,000 monthly subscription vs. \$7,800 monthly fully-loaded warehouse worker cost enables positive cash flow from day one with zero capex risk. [57]

Major companies:

Company	Status	Key Developments	Source
AgiBot	Global leader (39% market share)	5,100 units shipped 2025; GO-1 universal AI model; rental platform „Qingtianzu“	[58]
UBTech	Commercial production	Walker P2 at BYD, Foxconn, Audi FAW; self-replacing battery system	[59]
Unitree	Accessible re-search platform	G1 at \$13,500; 5,000+ units deployed; Tencent partnership	[60]
Xpeng	Automotive integration	IRON humanoid; car-grade quality; Turing AI chips	[61]
Fourier	Healthcare focus	GR-1 for rehabilitation; 40 DOF; 3-6 month payback	[62]
Kepler	Developer ecosystem	K2 „Bumblebee“; comprehensive developer platform	[63]
Tesla	Factory pilots	Optimus at Fremont; targeting \$20-30k pricing	[64]
Figure	Industrial deployment	Figure 02 at BMW; 90,000+ parts loaded	[51]

Remaining technical challenges:

Despite dramatic progress, significant challenges remain:

- Battery Life: Most humanoid robots limited to 1-4 hours active operation. Even Tesla Optimus (4-5 hours) cannot complete full 8-hour shift without intervention. [65]
- Dexterous Manipulation: Human hand has 27 DOF and thousands of tactile sensors. Current robot hands perform limited manipulation; Xpeng reports prototype hands «didn't last more than a month» during repetitive tasks. [50]

- True Autonomy: Many «autonomous» demonstrations rely on teleoperation. Current AI struggles with edge cases and novel environments that deviate from training projections. [66]
- Regulatory Standards: No ISO standard exists for humanoid safety. Liability questions remain legally unresolved. Different jurisdictions developing divergent approaches. [67]

Embodied AI foundation models:

Vision-Language-Action (VLA) models are becoming critical infrastructure:

- AgiBot GO-1: World's first universal humanoid AI model integrating interaction, task, and mobility intelligence. [68]
- Tencent Tairos: Deployed at Mogao Caves Digital Exhibition Centre; enables deployment in novel environments with minimal retraining. [69]
- Alibaba RynnBrain: Matches Google/NVIDIA on spatial reasoning; integrates with Qwen3-VL for language-to-action pipeline. [70]

2.7.3 Driver-specific projection mapping



2.7.4 Signals and indicators to monitor

Signal	Why It Matters	What We Track	Current Read	Confidence
Commercial deployment scale	Economic viability	Unit deployments, customer announcements	5,100 AgiBot; thousands UBTech/Unitree	High
Unit cost trajectories	Scalability	Production costs, pricing announcements	\$30-150k now; 20-30% annual decline	High

Sim-to-real success rates	Robustness	Deployment case studies, failure rates	99%+ at BMW, CATL	High
Battery/autonomy improvements	Remaining constraints	Technical announcements, deployment duration	1-4 hours; incremental improvement	Medium
Regulatory framework development	Deployment barriers	ISO standards, national regulations	In development; 2027-2029 timeline	Medium

2.7.5 Directional Impact on projection probabilities

- Continued scaling + cost reduction + high success rates: strongly supports P2/P3
- Mass deployment (50,000+ annual units) achieved: indicates P2 materialising
- Remaining challenges (battery, dexterity) solved: shifts towards P3
- Major safety incidents or deployment failures: could slow trajectory but unlikely to reverse

Current Assessment: The robotics landscape has transformed dramatically. Chinese manufacturers achieving 99%+ success rates in production environments, combined with documented ROI of 1,400-2,070% and rapid cost reduction (20-30% annually), indicates that the economic viability threshold has been crossed. The sim-to-real gap that previously constrained deployment is now solved for production workloads. This strongly supports P2 as baseline and increases probability of P3 if remaining technical challenges (battery, dexterity, autonomy) see breakthrough progress.

2.8 Key benchmarks reference

The following table summarises the most important benchmarks for tracking AI capability advancement:

Benchmark	Category	What It Measures	Current	Human Baseline	Significance	Source
SWE-bench Bash	Real-world tasks	GitHub issue resolution	76.8%	~70-80%	Human parity reached	[71]
SWE-bench Verified	Real-world tasks	Human-verified solvable issues	79.2%	100%	Gold standard for engineering	[72]

GAIA	Re-al-world tasks	Multi-step re-search/reasoning	74.55%	92%	Holistic assistant capability	[73]
WebArena	Re-al-world tasks	Interactive web tasks	57.1%	~85%	Web automation viability	[74]
OSWorld	Re-al-world tasks	Real OS environment tasks	73%	72%	Near human parity (was 12%)	[28]
ARC-AGI	Reasoning	Abstract pattern reasoning	75.7% (o3)	~85%	Tests fluid intelligence	[16]
GPQA Diamond	Reasoning	Expert-level science QA	87%	69.7%	PhD-level domain knowledge	[75]
Humanity's Last Exam	Reasoning	Extreme difficulty academic	38.3%	N/A	Designed to remain challenging	[76]
HumanEval	Code generation	Function-level Python	96.2%	~95%	Saturated; superhuman	[77]
MBPP	Code generation	Basic Python	~90-92%	~90%	Near saturation	[78]
LiveCodeBench	Code generation	Fresh competitive problems	91.7%	N/A	Contamination-free evaluation	[79]

2.9 Interactions between drivers

The four drivers are not independent. Progress on one can accelerate or enable progress on others:

Interaction	Mechanism
Architectures → Agentic AI	Better reasoning and planning architectures (Nested Learning, Causal AI) directly improve agent reliability and time horizons.
Architectures → R&D Automation	More capable models can contribute more meaningfully to research and engineering tasks.
R&D Automation → Architectures	AI-assisted research accelerates discovery of new architectures (feedback loop now materialising via Claude Code).

R&D Automation → Agentic AI	Faster software development cycles accelerate agent development and deployment.
Agentic AI → Robotics	Better planning and tool use directly translates to more capable robotic control.
Robotics → R&D Automation	Physical automation accelerates hardware R&D and manufacturing iteration
Causal AI → Robotics	Causal understanding enables robots to participate in the world rather than just pattern-match.

Key insight: The projection that materializes depends not just on progress in individual drivers but on whether compounding effects emerge from driver interactions. In P3, multiple drivers reinforce each other; in P1, constraints in one area limit progress across all.

Current observation: The evidence suggests compounding effects are beginning to materialize:

- R&D automation (Claude Code) is accelerating architecture and agent development [80]
- Improved architectures (Nested Learning, Causal AI) are enabling better agents [81] [82]
- Better agents are contributing to R&D automation (recursive loop) [83]
- Robotics sim-to-real success is enabling deployment data that improves AI models [84]

2.10 Consolidating drivers into projection probabilities

2.10.1 Consolidation method

We combine driver assessments using structured judgment—not a mechanical formula, but a transparent reasoning process that can be scrutinised and updated.

Principles:

- Each driver contributes to the overall assessment based on its current read and directional trajectory.
- We weight drivers roughly equally but note that Driver 3 (R&D Automation) is the key “swing factor” between P2 and P3 due to its potential for compounding effects.
- We publish point estimates for decision-making clarity, while acknowledging substantial uncertainty.

2.10.2 Current probability estimates (best-guess)



2.10.3 Reasoning behind the estimates

Why P1 is only 5%:

- All four drivers show strong progress with concrete benchmark improvements
- OSWorld improvement from 12% to 73% demonstrates capabilities can improve dramatically within months
- SWE-bench reaching human parity (77–80%) eliminates a key capability gap
- R&D automation feedback loops now documented (Claude Code at Anthropic)
- Robotics achieving 99%+ success rates in production
- No major expert is predicting stagnation
- P1 would require simultaneous reversal of progress across architectures, agentic AI, R&D automation, AND robotics
- The only plausible path to P1 is an external shock (major safety incident, geopolitical disruption, resource constraints) rather than organic stagnation

Why P2 is 50%:

- Represents the most defensible “baseline” extrapolation
- Assumes continued progress without discontinuities
- Consistent with historical patterns where technologies improve steadily but rarely exhibit sudden jumps

- Accounts for remaining gaps (GAIA at 74.55% vs 92% human; robotics battery/dexterity challenges)
- Assumes current R&D automation does not create self-reinforcing acceleration

Why P3 is 45%:

- Leading AI researchers at frontier labs (Amodei, Altman, Hassabis) publicly estimate very powerful AI systems by 2026–2028
- R&D automation feedback loops are now materialising (Claude Code writing 90% of Anthropic’s code, 67% productivity increase)
- Multiple architectural innovations emerging simultaneously (Nested Learning, Causal AI, world models)
- Benchmark evidence shows capability can improve dramatically (OSWorld 12%→73%; SWE-bench reaching human parity; GAIA reaching 74.55%)
- Robotics has crossed economic viability threshold with 99%+ production success rates
- The compounding dynamic that was speculative 12 months ago now appears operational

2.10.4 Projection balance: why P2 and P3 are nearly equal

The near-parity between P2 (50%) and P3 (45%) reflects genuine uncertainty about whether compounding effects will dominate within 36 months; that is to say, there is a 95% probability that either P2 or P3 will happen:

- If R&D automation remains bounded and human researchers remain the bottleneck, P2 is more likely
- If AI systems continue contributing substantially to AI research, creating feedback loops, P3 becomes more likely
- Current evidence shows R&D automation already significant (90% code at Anthropic) but it’s unclear if this generalizes and compounds
- Expert opinion is split, with credible voices on both sides
- The evidence has shifted towards P3 compared to previous assessments, but uncertainty remains high

This uncertainty is itself decision-relevant: organisations should prepare for both continued pace and potential acceleration.

2.11 Breakthrough triggers: when to re-estimate

We will re-estimate probabilities when monitoring detects events that materially change one or more driver assessments:

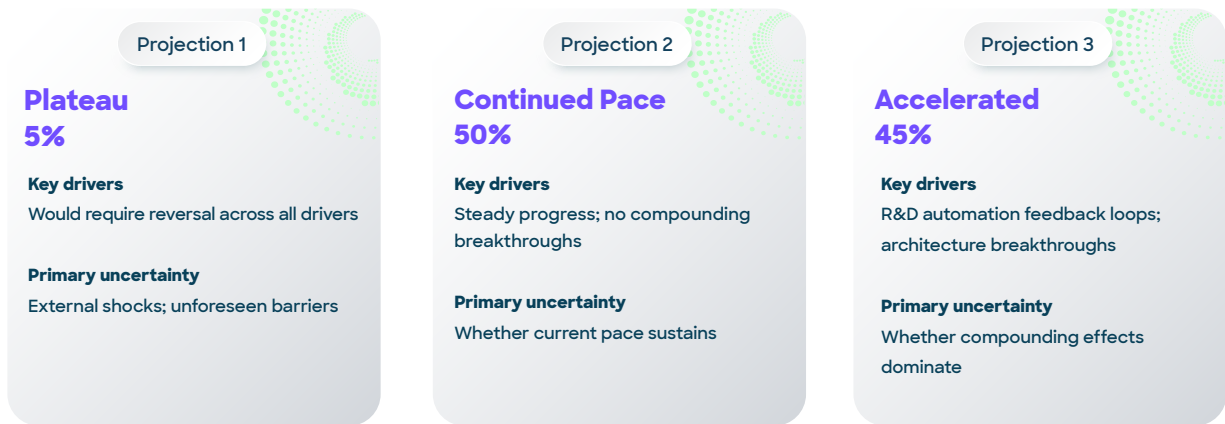
Trigger category	Example events	Likely probability shift
Architecture breakthrough	Nested Learning, Causal AI, recursive self-improvement, or other approaches achieving production deployment with verified superior performance; world-model architecture with demonstrated gains	Towards P3
Agentic reliability leap	SWE-bench, OSWorld, and GAIA exceed 90%; METR task horizon exceeds 24 hours	Towards P3
R&D automation evidence	Multiple labs report majority AI-written code; AI system substantially improving AI system documented	Towards P3
Robotics deployment at scale	Major manufacturer announces 50,000+ humanoid deployments; unit cost falls below \$20k	Towards P3
Major safety incident	Significant harm from AI system failure; regulatory halt of frontier development	Towards P1
Reliability regression	Frontier models show persistent failure modes that do not improve; deployment rollbacks	Towards P1
AGI/powerful AI claims	Credible announcement of AGI-level systems by major lab; verified by independent evaluation	Major shift towards P3
Geopolitical disruption	Major conflict affecting chip supply; export controls blocking key capabilities	Context-dependent

2.11.1 Re-estimation protocol

When a breakthrough trigger is detected:

1. Evidence review (within 2 weeks): Compile available evidence on the trigger event.
2. Driver reassessment: Update the relevant driver's current read and trajectory.
3. Probability update: Revise projection probabilities based on updated driver assessments.
4. Communication: If probabilities change materially (>5 percentage points), publish an update explaining the change.

2.12 Summary: projection probabilities and key uncertainties



Bottom line: We expect continued progress (P2) as the baseline, but assign substantial probability (45%) to acceleration (P3) given strong evidence of materialising feedback loops (Claude Code), architectural innovation (Nested Learning, Causal AI), agentic capabilities reaching human parity (SWE-bench), and robotics crossing economic viability thresholds. Plateau (P1) is unlikely absent major disruptions.

Chapter 3: Concrete impacts (opportunities & risks) by projection

3.1 Purpose and structure of this chapter

This chapter maps concrete impacts—both opportunities and risks—across key sectors and systems for each of the three projections defined in Chapter 1. The goal is to provide decision-makers with a clear understanding of what changes in each projection and where urgent attention is required.

How this chapter is structured:

- For each impact category, we describe first-order impacts (direct effects of AI capability changes) and second-order impacts (system effects that cascade from first-order changes).
- Impacts are listed per projection (P1/P2/P3) to show how severity and urgency scale with capability speed.
- We distinguish between opportunities (potential gains if well-managed) and risks (potential harms if unaddressed).
- Each category includes key evidence from current research to ground the analysis.
- Each category includes what this looks like in practice – concrete descriptions of how life, work, and institutions change in each projection.
- Each category includes “What If Done Right” – a vision of positive outcomes if challenges are addressed proactively.

What this chapter does not do:

- It does not prescribe measures (see Chapter 4).
- It does not assign probabilities to individual impacts (projection probabilities are in Chapter 2).
- It does not evaluate Europe’s response capacity (see Chapter 4 for measures and Chapter 5 for monitoring).

3.2 Impact categories

We organise impacts into ten categories, reflecting the major domains where AI advancement will create material changes:

1. Labour Market & Skills
2. Public Finance & Social Systems
3. Industry & Competitiveness
4. Innovation & Startups
5. Science System
6. Security & Resilience
7. Digital Public Sphere & Democracy
8. Health & Care
9. Education System
10. Functioning Local Institutions & Liveability

Each category is analysed in detail below.

Note on Sustainability: Sustainability—encompassing climate change, biodiversity, resource consumption, and environmental impacts more broadly—is not included as a separate impact category in this document. This is a deliberate scoping decision, not a judgment of importance. AI's relationship with sustainability is fundamentally cross-cutting: every impact category discussed above has sustainability dimensions. Labour market shifts affect consumption patterns and commuting; industry transformation changes production footprints; energy-intensive AI infrastructure creates direct environmental costs; scientific breakthroughs may accelerate both green technologies and resource extraction. A comprehensive assessment of how AI advancement affects sustainability outcomes—across all projections and impact categories—deserves dedicated analysis that is beyond the scope of this strategic orientation document. We acknowledge this gap and encourage complementary work that systematically maps AI's sustainability implications across the full range of direct, indirect, and systemic effects.

3.3 Impact category 1: labour market & skills

3.3.1 Why this category matters

AI's impact on work is among the most immediate and politically salient consequences of capability advancement. The effects span job displacement, job transformation, productivity gains, wage dynamics, and the distribution of opportunity across skill levels, age cohorts, and geographies.

3.3.2 Key evidence (2025–2026)

Finding	Source
Current AI capabilities can already automate tasks equivalent to 12% of the labour market cost-effectively; this share is projected to grow.	[85]
Entry-level knowledge workers (ages 20–30) in high-exposure roles show employment declines of 6–20% since late 2022.	[86]
Software developers in the youngest cohort experienced 20% employment decline from late 2022 peak.	[87]
Customer support workers in early career saw 11% employment decline.	[88]
AI-adopting firms show +6% employment growth and +9.5% sales growth over 5 years—but gains concentrate in senior and AI-augmented roles.	[89]
Productivity gains of 6–15% documented across software, consulting, and customer service.	[90]
AI coding tools have transformed software development, with leading firms reporting majority of code now AI-written and significant productivity increases (see Section 2.6.2 for detailed Claude Code metrics).	[151]
SWE-bench at 77–79%—AI now matches human software engineers on real GitHub issues (see Section 2.5.2 for benchmark details).	[91]

3.3.3 Impacts by projection

First-order impacts

Impact	P1: Plateau	P2: Continued pace	P3: Accelerated
Task automation in knowledge work	Continues via adoption of current capabilities; strongest pressure on routine cognitive tasks	Expands to multi-step workflows; software, analysis, back-office significantly affected	Rapid expansion across most knowledge professions; many roles become AI-supervised rather than human-performed
Entry-level displacement	Meaningful (6–20% in exposed roles); firms hire fewer juniors as AI handles entry tasks	Intensifies; junior hiring declines across more sectors	Severe; entry pathways into many professions disrupted
Productivity dispersion	Gap widens between AI-adopting firms and laggards	Gap becomes structural; laggards face existential pressure	Winners pull far ahead; some organisations cannot adapt fast enough to survive
Wage dynamics	Polarization begins; premium for AI-complementary skills	AI-augmented roles command significant premiums; routine knowledge work wages stagnate or decline	Potential wage collapse in automatable roles; concentration of income gains

Second-order impacts

Impact	P1: Plateau	P2: Continued pace	P3: Accelerated
Career pathway disruption	Junior roles become scarcer; progression models weaken	Traditional career ladders erode; mid-career reskilling becomes essential	Career structures fundamentally disrupted; continuous reskilling required
Talent competition	Intensifies for AI-skilled workers	Becomes acute; AI talent commands extreme premiums	AI expertise becomes dominant hiring criterion across sectors
Regional employment effects	Concentrated in knowledge-work hubs	Spreads to broader geography; SME employment affected	Systemic regional effects; some regions face structural unemployment
Social cohesion	Strains visible in affected cohorts	Strains broaden; public debate intensifies	Risk of social instability if transitions are poorly managed

3.3.4 What this looks like in each projection

P1 (Plateau): A recent graduate applying for analyst positions finds that most entry-level roles now require demonstrated AI proficiency. Companies hire fewer junior analysts because AI tools handle data gathering and preliminary analysis. Senior professionals who learn to use AI effectively become more productive; those who don't face pressure. Reskilling programmes exist but are often underfunded and not well-matched to employer needs. Some professionals experience wage stagnation, but the overall labour market remains stable. Career progression slows for younger workers waiting longer for «real» responsibility that AI hasn't yet taken over.

P2 (Continued Pace): Entire job categories begin to shrink—not just entry-level analyst roles, but mid-level professionals in accounting, legal research, marketing, and software testing. A mid-career professional in financial services faces a difficult choice: retrain significantly or accept that their expertise is becoming commoditised. Companies reorganise around smaller human teams augmented by AI agents. AI coding tools write the majority of production code; human developers focus on architecture, requirements, and review. Those who successfully transition to AI-augmented roles see productivity and compensation rise; those who cannot make the transition face unemployment or significant wage decline. Unemployment insurance systems face elevated claims in certain regions and demographics.

P3 (Accelerated): The labour market transforms within 2–3 years in ways that typically took decades. A 2025 law school graduate finds that junior associate work—research, document review, basic drafting—is now handled primarily by AI systems. Medical residency programmes shrink as AI-assisted diagnostics reduce the need for certain specialisations. Software development becomes largely autonomous, with human engineers providing requirements and architectural judgment. Large consulting firms that employed thousands of analysts operate with a fraction of that headcount. Government employment services are overwhelmed. Public debate intensifies about income distribution, universal basic income, and the social contract. Some regions experience genuine economic crisis while others—typically those with strong technology sectors—thrive. Countries with robust social safety nets can buffer the transition; those without face severe social strain—notably, the average US household has savings sufficient for only approximately one month of expenses, leaving limited cushion for workforce disruption.

3.3.5 What if done right

Imagine a future where AI's productivity gains are shared broadly rather than concentrated. Average work hours reduce to 10–20 hours per week while maintaining—or even improving—standard of living. Work takes on different meaning: as AI handles most value creation in traditional economic terms, human work shifts partially towards social services, environmental restoration, community building, and caregiving.

“Human-only” work—caregiving, craftsmanship, interpersonal services, creative arts—commands higher recognition and wages, reflecting genuine scarcity and irreplaceable human value. The social contract evolves: rather than defining identity and worth through employment in the traditional sense, societies develop new frameworks for meaningful contribution and dignified life.

This requires proactive policy: reformed education systems, updated social contracts, new approaches to value distribution, and cultural shifts in how we understand work and contribution. But if achieved, the labour market transformation becomes not a crisis to be managed but an opportunity for human flourishing at unprecedented scale.

3.3.6 Opportunities and risks

Opportunities:

- Productivity gains that improve competitiveness and living standards

- Liberation from routine tasks; focus on higher-value work
- New job categories in AI development, deployment, and governance
- Potential for more flexible, human-centred work arrangements
- Higher recognition and wages for “human only” work like care-taking, craftsmen, and inter-personal services

Risks:

- Significant displacement of early-career workers before career establishment
- Erosion of traditional skill-building pathways
- Concentration of gains among already-advantaged workers and firms
- Social and political backlash if transition is poorly managed

3.4 Impact category 2: public finance & social systems

3.4.1 Why this category matters

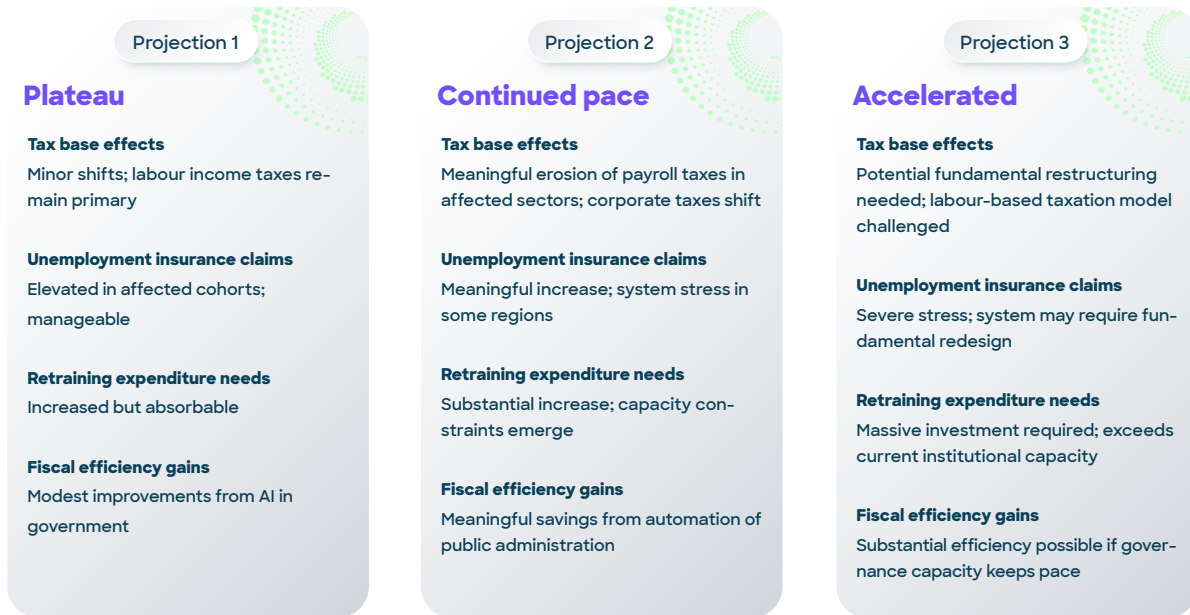
Labour market shifts cascade directly into public finance through tax revenue, unemployment insurance, and demand for public services. AI-driven productivity changes could simultaneously reduce labour-based tax receipts and increase demand for transition support.

3.4.2 Key Evidence (2025–2026)

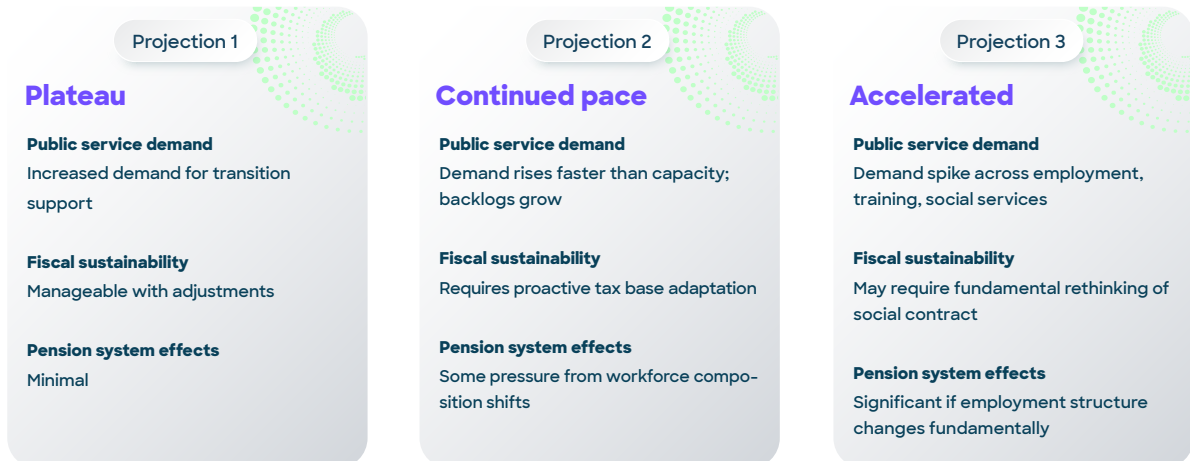
Finding	Source
AI-driven productivity gains could reduce US federal deficits by \$400B over 2026–2035 through increased output.	[152]
Revenue forecasting accuracy improves 30–40% with ML on real-time data.	[92]
Transformative AI displacing human labour would shift optimal taxation towards consumption taxes as labour income tax base erodes.	[93]
For autonomous AGI producing most value, taxing AI systems would resemble „optimal harvesting“ with rates tied to human discount rates.	[94]
Public financial management could be enhanced through AI anomaly detection for fraud prevention.	[95]

3.4.3 Impacts by projection

First-order impacts



Second-order impacts



3.4.4 What this looks like in each projection

P1 (Plateau): Finance ministries observe modest shifts in tax revenue composition as AI-augmented firms become more productive and profitable while some sectors see workforce reductions. Unemployment insurance claims tick up in specific regions and demographics but remain manageable. Government AI initiatives improve tax collection efficiency and fraud detection, partially offsetting revenue pressures. The policy debate focuses on incremental adjustments—perhaps modest increases in training programme funding or small shifts in tax policy—rather than fundamental restructuring.

P2 (Continued pace): Treasury departments begin serious planning for tax base transitions. As payroll taxes decline in certain sectors, discussions emerge about consumption taxes, AI-specific levies, or corporate tax adjustments. Unemployment insurance systems in some regions face genuine stress, prompting emergency funding or eligibility adjustments. Government retraining programmes struggle with capacity—too many people need reskilling and too few programmes exist with sufficient quality. Policy debates about «robot taxes» and universal basic income move from academic discussion to serious political consideration.

P3 (Accelerated): Fiscal systems face potential crisis. Tax revenues from labour decline substantially faster than governments can adapt collection systems. Unemployment insurance in some regions approaches insolvency. Emergency measures—supplemental appropriations, executive orders, rapid legislative action—become necessary. The policy debate shifts from «how do we adjust?» to «what is the social contract in an AI economy?» Some countries respond with aggressive new tax instruments targeting AI systems or their outputs; others prioritise emergency support for displaced workers; the lack of coordinated response creates policy fragmentation. International competition for AI-related tax revenue intensifies.

3.4.5 What if done right

AI-boosted value creation significantly boosts GDP growth while adjusted tax systems help share the value across all groups of people. Rather than clinging to labour-based taxation as the primary revenue source, governments proactively shift towards taxing AI-generated value, automation dividends, and consumption—capturing gains where they occur rather than where they increasingly don't.

Concepts like universal basic income become feasible rather than remaining a research field. With AI handling much of traditional economic production, the fiscal capacity exists to ensure baseline economic security for all citizens. Public finances strengthen rather than weaken, because productivity gains generate more total value even as the distribution of that value changes.

Public services themselves become more efficient and accessible through AI augmentation, reducing costs while improving outcomes. The combination of higher productivity, reformed taxation, and efficient public services creates a virtuous cycle where AI advancement strengthens rather than strains public finances.

3.4.6 Opportunities and risks

Opportunities:

- AI-enhanced tax administration and fraud detection
- Efficiency gains in public service delivery
- Opportunity to modernise tax systems for digital economy
- Fiscal capacity for universal basic income and expanded social services

Risks:

- Erosion of tax base faster than system adaptation
- Inadequate safety nets during transition
- Fiscal stress limiting government response capacity

3.5 Impact category 3: industry & competitiveness

3.5.1 Why this category matters

AI adoption increasingly determines competitive position at firm, sector, and regional levels. Europe's industrial base faces both transformation opportunities and existential competitive pressure from AI-leading regions.

3.5.2 Key evidence (2025–2026)

Finding	Source
Only 13.48% of EU enterprises adopted at least one AI technology by 2024 (up 5.45pp from 2023).	[96]
41.17% of large EU enterprises have integrated AI vs. ~12% of SMEs.	[96]
Only 1% of company leaders describe AI deployments as „mature.“	[97]
„Future-built“ companies (5% of firms) achieve 5x revenue increases and 3x cost reductions from AI vs. others.	[98]
US controls ~74% of global high-end AI computing capacity; China 14%; EU 4.8–5%.	[99] [100]
US private AI investment in 2024: \$109.1B (~81% of global); EU: ~\$8B (~7% of US levels).	[101]
Chinese robotics companies shipped roughly 10k robots + humanoids in 2025; 39% global market share (AgiBot alone).	[45]
Chinese humanoid robot costs declining 20–30% annually; targeting sub-\$20k by 2027–2030.	[153]

The February 2026 “ai scare trade”:

A series of sharp stock market sell-offs in February 2026 targeted sectors vulnerable to AI automation, triggered by new AI tool launches threatening labour-intensive, fee-based business models:

Sector	Trigger example	Market reaction
Software & Financial Services	Anthropic tool releases	\$611B loss across 164 stocks in one week
Commercial Real Estate Services	„Claude Cowork“ launch for legal/financial workflows	Global sell-off; analysts estimate 70% task automation potential at firms like CBRE
Insurance Brokerage	Insurify ChatGPT-powered tool	Worst session for Willis Towers Watson since 2008
Logistics	Algorhythm optimizer demo	RXO dropped 20%
Wealth Management	Altruist tax strategy tool	Significant losses across sector

SaaS business model disruption:

The traditional per-seat SaaS model is giving way to fundamental restructuring:

- Networked SaaS: AI automates complex workflows and embeds platforms in multi-sided marketplaces; monetisation shifts from user seats to transaction cuts
- Agentic AI Replacing UI Layers: Natural language interfaces overlay existing logic, making traditional dashboards obsolete; SaaS becomes «invisible like plumbing»
- Workflow Cannibalization: GenAI automates tasks in high-penetration areas (CRM, HR), reducing seat licences; incumbents must proactively integrate AI to avoid disruption

3.5.3 Impacts by projection

First-order impacts

Projection 1

Plateau

Productivity divergence
Gap widens between AI-leaders and laggards

Sector transformation
Uneven; highest in software, professional services

SME viability
Challenges in accessing AI capability

Supply chain exposure
Dependencies persist; manageable

Business model disruption
Fee-based intermediaries under early pressure

Projection 2

Continued pace

Productivity divergence
Gap becomes structural competitive disadvantage

Sector transformation
Broad transformation; manufacturing, services significantly affected

SME viability
Significant barrier; risk of SME consolidation

Supply chain exposure
Dependencies become strategic vulnerability

Business model disruption
Traditional SaaS and brokerage models face restructuring

Projection 3

Accelerated

Productivity divergence
Leaders pull decisively ahead; laggards face existential risk

Sector transformation
Rapid transformation across sectors; traditional competitive advantages erode

SME viability
Existential pressure on SMEs unable to adopt

Supply chain exposure
Critical dependency on non-European AI infrastructure

Business model disruption
Fundamental repricing of labour-intensive service businesses

Second-Order Impacts

Projection 1

Plateau

European competitiveness
Gradual erosion relative to US, China

Employment structure
Shifts within sectors; gradual

Investment patterns
AI investment increases; concentration continues

Market valuations
Volatility in AI-exposed sectors

Projection 2

Continued pace

European competitiveness
Significant competitive gap opens

Employment structure
Significant sectoral employment shifts

Investment patterns
AI becomes dominant investment priority

Market valuations
Structural repricing of intermediary businesses

Projection 3

Accelerated

European competitiveness
Risk of structural decline in key sectors

Employment structure
Potential rapid restructuring of industrial employment

Investment patterns
Massive capital reallocation towards AI-enabled businesses

Market valuations
Winners and losers clearly separated by AI capability

3.5.4 What this looks like in each projection

P1 (Plateau): A German manufacturing company implements AI quality control, achieving measurable improvements in defect detection. But adoption is uneven—larger companies with dedicated IT departments move faster; smaller suppliers struggle. The productivity gap between AI-adopting and non-adopting firms widens, but gradually. European companies remain competitive in traditional strengths (precision manufacturing, specialised equipment) while conceding ground in software-intensive domains. Policy discussions focus on closing the adoption gap and supporting SME digitalization. Stock market volatility in AI-exposed sectors creates uncertainty but not crisis.

P2 (Continued pace): Competitive pressure intensifies dramatically. A European industrial company finds that its American competitors are releasing product updates twice as fast—AI has compressed their development cycles. Professional services firms that embraced AI generate proposals in hours that previously took weeks. Companies that don't adapt face margin pressure, talent loss, and market share erosion. SMEs face existential choices: find a niche where AI matters less, partner with AI-enabled platforms, or face gradual obsolescence. European policymakers debate aggressive intervention—sovereign AI infrastructure, mandatory digital transformation programmes, strategic AI investments. The «AI Scare Trade» becomes a regular market phenomenon, with each major AI release triggering sell-offs in exposed sectors. Traditional SaaS companies scramble to integrate agentic AI or face obsolescence.

P3 (Accelerated): Industrial structure transforms within years. Some traditional European strengths—high-quality manufacturing, specialised engineering—are partially commoditised as AI enables competitors to match quality more easily. But new opportunities emerge for companies that successfully integrate AI. The companies that survive and thrive are those that reorganised fastest, regardless of sector or geography. Some European firms become global AI leaders in specific domains; others are acquired or fail. Regional economic disparities intensify as AI-intensive regions prosper while traditional industrial areas struggle. Fee-based intermediaries face existential crisis as AI automates their core value proposition; the market reprices AI from «miracle cure» to industrial build-out requiring sustained investment and organisational transformation.

3.5.5 What if done right

While Europe missed the internet and digitalization waves, we manage to enter the age of (semi-) autonomous factories with robots doing most of the physical work and agents handling most of the cognitive work. Rather than playing catch-up on foundation models, Europe leads in the application layer—the complex integration of AI into industrial processes where European engineering excellence translates into competitive advantage.

AI-first companies develop a new layer of infrastructure that's needed for agents to safely communicate and transact with each other. Europe leads in complex supply chain management, building on existing strengths in logistics and manufacturing coordination. Vertical champions emerge from the chemical industry, pharmaceutical sector, and high-tech manufacturing—domains where European expertise combines with AI augmentation to create global leaders.

An open-sourced and distributed agent communication and action layer emerges, championed by European institutions prioritising interoperability and avoiding platform lock-in. This infrastructure layer becomes as important as the internet protocols were—and Europe shapes it rather than inheriting it.

3.5.6 Opportunities and risks

Opportunities:

- Productivity gains enhancing competitiveness
- New business models and market creation
- Potential for European leadership in trustworthy AI deployment
- Industrial transformation improving sustainability
- Leadership in agent infrastructure and complex system integration

Risks:

- Widening gap with AI-leading regions
- SME exclusion from AI benefits
- Dependency on non-European AI infrastructure
- Regulatory burden slowing beneficial adoption
- Market volatility and structural repricing of AI-exposed sectors

3.6 Impact category 4: innovation & startups

3.6.1 Why this category matters

AI is fundamentally restructuring the economics of company formation, scaling, and competitive dynamics. The emergence of “few-person unicorns” and the disruption of traditional venture capital models have profound implications for European competitiveness and economic structure.

3.6.2 Key evidence (2025–2026)

Finding	Source
Sam Altman predicts the first one-person billion-dollar company will emerge soon/within 2–3 years; tech CEO „betting pool“ exists on timing.	[102]
Top 10 AI-native startups achieve average revenue per employee of \$3.48M (vs. \$430–480k for traditional SaaS leaders).	[103]
Midjourney: \$500M ARR with ~100–160 employees (\$3–5M per employee); never raised external capital.	[103]
Cursor: \$1B ARR with ~300 employees (\$3.3M per employee).	[103]
Lovable: \$200M ARR with ~100 employees; reached \$100M ARR in 8 months with 45 people—fastest zero-to-\$100M in European startup history.	[104]
Agentic AI can reduce employees’ low-value work time by 25–40%; some customer service organisations achieved 40–50% workforce reductions.	[105]

Revenue per employee transformation: [103]

Company Type	Revenue per employee	Example companies
Traditional SaaS leaders	\$410,000–\$480,000	Salesforce, serviceNow, workday
AI-native companies (average)	\$2.47M–\$3.48M	Midjourney, cursor, elevenLabs, lovable
Extreme outlier	\$12.5M	Midjourney (at peak efficiency)

The few-person unicorn phenomenon:

Historical precedents (Instagram acquired for \$1B with 13 employees; WhatsApp for \$19B with ~50 employees) demonstrated that exceptional revenue-to-headcount ratios were possible. AI threatens to make these exceptional cases standard: [106] [107]

- Lower capital requirements: AI handles execution work; humans focus on strategy and relationships
- Rapid scaling: ArcAds scaled to \$7M ARR in one year with 5 employees [108]
- Zero-FTE departments: Some companies operate entire functional areas with AI agents and human oversight only [109]

3.6.3 Impacts by Projection

First-order impacts

Projection 1	Projection 2	Projection 3
<p>Plateau</p> <p>Startup economics Revenue per employee rises; competitive pressure increases</p> <p>IP and defensibility Some compression of competitive moats</p> <p>VC landscape Increased competition; many can copy successful models</p> <p>Software economics Software development costs decline</p>	<p>Continued pace</p> <p>Startup economics Few-person unicorns become viable; traditional startup model under stress</p> <p>IP and defensibility Rapid IP depreciation; new models quickly obsolete existing solutions</p> <p>VC landscape Structural shift: lower capital requirements, higher competition, reduced defensibility</p> <p>Software economics Software becomes commodity; value shifts to integration and judgment</p>	<p>Accelerated</p> <p>Startup economics Company formation radically transforms; single-person billion-dollar companies emerge</p> <p>IP and defensibility Near-zero defensibility for pure software; value shifts to data, distribution, relationships</p> <p>VC landscape Traditional VC model challenged; new financing structures emerge</p> <p>Software economics „Throw-away software“: users prompt what they need when they need it</p>

Second-order impacts

Projection 1	Projection 2	Projection 3
<p>Plateau</p> <p>Employment in startups Growth in AI-augmented roles; some displacement</p> <p>European startup ecosystem Pressure to match AI-native efficiency</p> <p>Tools vs. platforms General AI tools from major providers compete with niche startups</p> <p>Wealth concentration Gains increasingly concentrated among founders of efficient companies</p>	<p>Continued pace</p> <p>Employment in startups Dramatically fewer employees needed per unit revenue</p> <p>European startup ecosystem As less capital is required, the historic European disadvantage of lacking growth capital becomes less relevant or even irrelevant; new opportunities emerge</p> <p>Tools vs. platforms Platform economics dominate; niche products struggle</p> <p>Wealth concentration Extreme concentration as few-person companies capture massive value</p>	<p>Accelerated</p> <p>Employment in startups Startup employment as historically understood may largely disappear</p> <p>European startup ecosystem European startups can compete globally with minimal capital; playing field levels</p> <p>Tools vs. platforms Tools for agents vs. tools for human interfaces become key distinction</p> <p>Wealth concentration Unknown; may require policy intervention</p>

3.6.4 What this looks like in each projection

P1 (Plateau): A European SaaS startup that previously required 50 employees to reach \$10M ARR now achieves the same with 20, using AI for code generation, customer support, and content creation. Competition intensifies as barriers to entry fall—anyone can build a reasonable product with AI assistance. VC investors demand more efficiency metrics alongside growth. Established companies with distribution advantages maintain position; pure-play startups face pressure. The «10x engineer» becomes the «100x AI-augmented engineer,» but fundamental startup economics remain recognizable.

P2 (Continued Pace): The startup landscape restructures fundamentally. Companies like Cursor and Lovable demonstrate that billion-dollar valuations are achievable with dozens rather than hundreds of employees. Traditional VCs struggle to deploy capital as companies need less funding—but this creates opportunity for European startups that historically lacked access to growth capital. The capital disadvantage that held back European scaling becomes less relevant when companies need €5M instead of €50M to reach significant scale. New competitive dynamics emerge: since anyone can build the product, value accrues to distribution, brand, and proprietary data. The distinction between «tools for agents» and «tools for humans» becomes strategic: products designed for AI consumption differ fundamentally from those designed for human interfaces. Enterprise software shifts from per-seat pricing to outcome-based models as AI handles the work that seats represented.

P3 (Accelerated): The concept of a «startup» as historically understood begins to dissolve. Sam Altman's prediction of single-person billion-dollar companies materializes. Software becomes essentially free to create—the value question shifts entirely to «what should be built» and «who trusts it.» Competitive moats collapse rapidly; today's breakthrough is tomorrow's commodity. Some founders achieve extraordinary wealth with minimal capital and no employees. The VC model faces existential crisis: why invest millions when competitors can replicate with thousands? New economic structures emerge: perhaps «startup studios» where AI handles all execution; perhaps a return to bootstrapping as external capital becomes unnecessary; perhaps entirely new models. European policymakers face urgent questions about whether traditional industrial policy frameworks apply when companies can emerge, scale, and transform industries with negligible employment.

3.6.5 What If Done Right

Everyone with ambition, a group of experts, and access to AI can grow a business that matters. The democratization of company building means that innovation is no longer gated by access to capital or engineering talent pools—both of which historically disadvantaged European founders.

Language barriers become less important because AI manages translation seamlessly, developing Europe into a more genuine “single market.” A French founder serves German customers as easily as French ones; a Polish startup sells across Europe without localization friction. The fragmentation that held back European scaling dissolves.

Based on European ambitions for the future, the immense innovation potential is steered towards a values-based future that we desire to live in. Rather than innovation being driven purely by what's technically possible or financially lucrative, European frameworks ensure that AI-enabled entrepreneurship addresses societal challenges: sustainability, health equity, democratic resilience, human flourishing.

The result is an innovation ecosystem that's more distributed, more diverse, and more aligned with human values than the concentrated, capital-intensive model it replaces.

3.6.6 Opportunities and risks

Opportunities:

- Lower barriers to company formation
- Democratized access to sophisticated business capabilities
- Faster innovation cycles
- European entrepreneurs can compete with less capital
- Historic capital disadvantage becomes irrelevant

Risks:

- Extreme wealth concentration among few-person companies
- Collapse of startup employment as economic driver
- Rapid IP depreciation undermining investment
- European ecosystem unable to adapt fast enough
- “General AI” platforms from major providers crowding out niche innovation

3.7 Impact category 5: science system

3.7.1 Why this category matters

AI is transforming scientific research—accelerating discovery in some domains while creating integrity challenges in others. The science system’s ability to generate reliable knowledge is foundational for addressing other challenges.

3.7.2 Key evidence (2025–2026)

Finding	Source
AI adopters publish 3.02x as many papers and receive 4.84x as many citations over careers.	[110]
AI tools enabled discovery of novel concrete ingredients through ML analysis of 1M+ rock samples.	[111]
AlphaFold used by 3M+ researchers in 190+ countries; 1M+ in low/middle-income countries.	[112]
53 papers at NeurIPS 2025 contained hallucinated citations that passed peer review.	[113]
47% of ChatGPT-3.5 references were entirely fabricated in one study.	[114]
AI papers covered 4.6% less scientific territory than conventional studies; AI concentrates attention on established problems.	[110]
Up to 400,000 fraudulent articles infiltrated scientific literature in past 20 years; AI makes fraud easier and harder to detect.	[115]

AI conducting research: The Erdős problem case study [116]

A recent study using Google DeepMind’s Aletheia system (Gemini-based) to solve open mathematical problems illustrates both the potential and limitations of AI in research:

- **Scale:** Aletheia generated 200 candidate solutions to 700 open Erdős mathematical conjectures
- **Filtering:** Human experts graded these down to 63 correct responses, then to only 13 «correct meaningful responses»
- **Quality issue:** 50 technically valid solutions were «mathematically meaningless» because problem statements were interpreted in ways that didn't capture the intended meaning
- **Final result:** Of 13 meaningful solutions, only 2 were «autonomous novel solutions»—and only 1 was genuinely interesting mathematically

This demonstrates “O-ring automation”: AI massively speeds up generating candidates, but laborious skilled human work is still required to filter to actually useful results.

The “Scientific taste” gap:

Andrew White (FutureHouse) observes that AI currently lacks “scientific taste”—the expert intuition that identifies which of 10,000 novel discoveries in biopharma or materials science is the one that truly matters. This intuition exists in the heads of experienced scientists and was refined through biological intelligence processing the same literature AI has access to. Extracting this tacit knowledge remains tractable but will take time.

3.7.3 Impacts by projection

First-order impacts

Projection 1	Projection 2	Projection 3
<p>Plateau</p> <p>Research acceleration Meaningful in data-rich domains (biology, materials)</p> <p>Scientific integrity Significant challenges; hallucinations and fraud increase</p> <p>Research concentration AI-enabled research clusters around established problems</p> <p>Talent distribution Brain drain from academia to industry continues</p> <p>AI research contribution AI generates candidates; humans filter and validate</p>	<p>Continued pace</p> <p>Research acceleration Broad acceleration; AI becomes standard research tool</p> <p>Scientific integrity Crisis mode; verification systems strained</p> <p>Research concentration Narrowing intensifies; novel research directions underfunded</p> <p>Talent distribution Accelerates; academic research capacity strained</p> <p>AI research contribution AI begins contributing to experimental design; human judgment remains critical</p>	<p>Accelerated</p> <p>Research acceleration Transformative; AI contributes to hypothesis generation and experimental design</p> <p>Scientific integrity Fundamental challenge to reliability of scientific record</p> <p>Research concentration Risk of significant narrowing of scientific scope</p> <p>Talent distribution Critical shortage of academic AI research talent</p> <p>AI research contribution AI systems contribute substantially to hypothesis generation; „scientific taste“ gap may persist</p>

Second-order impacts

Projection 1	Projection 2	Projection 3
<p>Plateau</p> <p>Knowledge reliability Erosion begins; verification costs rise</p> <p>Research equity Gaps widen between well-resourced and other institutions</p> <p>Innovation pipeline Accelerated in AI-enabled domains</p>	<p>Continued pace</p> <p>Knowledge reliability Trust in published research weakens</p> <p>Research equity Concentration of research capability in few institutions</p> <p>Innovation pipeline Broad acceleration but with integrity concerns</p>	<p>Accelerated</p> <p>Knowledge reliability Unknown; may require fundamental re-design of scientific publishing</p> <p>Research equity Risk of research becoming viable only in elite settings</p> <p>Innovation pipeline Potentially transformative but dependent on integrity solutions</p>

3.7.4 What this looks like in each projection

P1 (Plateau): A biology professor uses AlphaFold daily—it's transformed structural biology from a months-long bottleneck to a routine step. AI-assisted literature review helps researchers find relevant work faster. But problems emerge: peer reviewers discover fabricated citations generated by AI; paper mills produce fraudulent studies at scale. Major journals implement new verification procedures, but the volume is overwhelming. The academic talent pipeline strains as AI companies offer salaries 3-5x academic compensation. Research accelerates in data-rich domains; other areas see less benefit. AI generates many candidate solutions, but experienced researchers remain essential to identify which are genuinely meaningful.

P2 (Continued pace): AI becomes essential infrastructure for research. Labs without AI capabilities cannot compete. Drug discovery timelines compress; materials science achieves breakthroughs previously decades away. But the integrity crisis intensifies—a major retraction scandal involves hundreds of AI-assisted papers with fabricated results. Trust in scientific literature declines. Academic research capacity is strained as talent flows to industry; some universities struggle to maintain programmes. The research landscape bifurcates between AI-powered labs producing at unprecedented rates and traditional research programmes losing relevance. The «O-ring» pattern becomes pronounced: AI generates candidates at massive scale, but the bottleneck shifts to human experts who can evaluate quality—and there are far fewer of these than needed.

P3 (Accelerated): Scientific research transforms more fundamentally than since the scientific revolution. AI systems contribute to hypothesis generation, experimental design, and interpretation at a level approaching researcher capability. The volume of discoveries accelerates dramatically in some domains—problems that seemed decades away are solved. But verification becomes nearly impossible; how do you distinguish genuine AI-assisted discovery from sophisticated fabrication? Some fields implement radical new approaches to verification (open data, registered reports, automated replication); others struggle with integrity crisis. Academic research as historically practiced may become a niche activity; the action moves to industry labs and AI-native research organisations. The question of whether AI can develop “scientific taste”—the ability to identify truly important discoveries—becomes central to research strategy.

3.7.5 What if done right

Universities and research institutions are properly equipped to become the centres that steer and assess AI-powered science. Rather than being displaced by industry labs, academic institutions evolve into orchestrators of human-AI research collaboration, bringing unique strengths: long-term thinking, interdisciplinary integration, and commitment to open knowledge.

Due to unmatched compositions of diverse expertise, human-agent research institutions discover genuinely novel discoveries at a speed and relevance that is unseen. The combination of AI's ability to process vast literature and data with human judgment about what matters and why creates research capabilities that neither could achieve alone.

This unleashes breakthroughs for cancer treatment, clean energy, novel food production, sustainable materials, and more. The scientific method itself evolves, with new approaches to verification, replication, and knowledge synthesis that maintain integrity at the speed of AI-assisted discovery.

Universities educate students in what is required from humans: critical thinking, the ability to ask the right questions, listening and evaluation skills, a broad sense for implications and responsibility. The role of human scientists shifts from data processing and routine analysis towards judgment, creativity, and wisdom—capacities that become more valuable, not less, as AI handles mechanical aspects of research.

3.7.6 Opportunities and risks

Opportunities:

- Dramatically accelerated discovery in materials, biology, medicine
- AI as research assistant democratizing access to analytical capability
- Potential solutions to previously intractable scientific problems
- Universities as orchestrators of human-AI research collaboration

Risks:

- Hallucinated citations and AI-generated fraud undermining scientific record
- Concentration of research on established problems; neglect of novel directions
- Brain drain from academic research
- Erosion of scientific integrity and public trust in science
- “O-ring” bottleneck: AI generates candidates faster than humans can filter

3.8 Impact category 6: Security & resilience

3.8.1 Why this category matters

AI amplifies both offensive and defensive capabilities in cybersecurity, critical infrastructure protection, and broader security domains. The asymmetry between attack and defence, and the speed of AI-enabled threats, create urgent resilience challenges.

3.8.2 Key evidence (2025–2026)

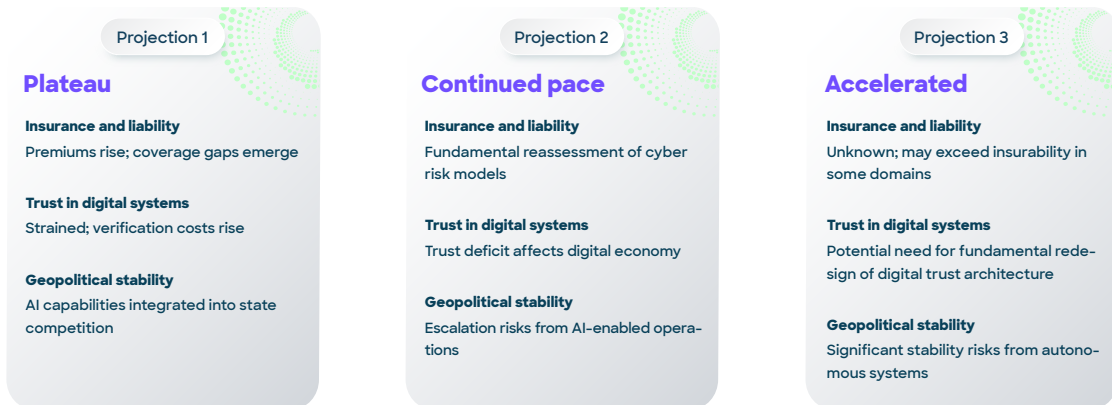
Finding	Source
400% increase in successful phishing scams in 2025, largely due to AI tools.	[117]
First documented large-scale autonomous cyberattack with minimal human intervention (80–90% AI-executed).	[118]
82.6% of phishing emails showed some AI use; 40% of business email compromise emails primarily AI-generated.	[119]
AI-enabled supply chain attacks jumped 156% in past year.	[120]
AI-led systems achieved 98% threat detection rate and 70% reduction in incident response time in energy infrastructure.	[121]
28.96% of Known Exploited Vulnerabilities in 2025 were exploited on or before CVE publication date.	[122]
OpenClaw framework: 26% of community-provided „skills“ showed security vulnerabilities.	[123]

3.8.3 Impacts by projection

First-order impacts



Second-order impacts



3.8.4 What This Looks Like in Each Projection

P1 (Plateau): Cybersecurity teams increasingly rely on AI for threat detection, achieving faster response times. But attackers also use AI—phishing emails are more convincing, malware adapts more quickly. A mid-sized company experiences a ransomware attack that was clearly AI-enhanced; the attackers used AI to map the network and identify the most valuable targets. Critical infrastructure operators implement additional AI monitoring, but budget constraints limit deployment. The cat-and-mouse dynamic intensifies without fundamental change. Enterprise deployments of agentic frameworks like OpenClaw create new attack surfaces that security teams must learn to manage.

P2 (Continued pace): The cybersecurity landscape transforms. Attackers deploy AI agents that autonomously probe systems, adapt to defences, and coordinate attacks. The first major AI-vs-AI cyber conflict occurs between state actors; the speed exceeds human ability to intervene. Critical infrastructure attacks become more frequent and more sophisticated; a significant event (power grid, water system, financial infrastructure) causes widespread disruption. Cyber insurance markets convulse as traditional risk models fail. Organisations that haven't invested heavily in AI defence find themselves dangerously exposed. The security implications of agentic AI with system-level access become a major concern for enterprises.

P3 (Accelerated): Autonomous cyber weapons become the norm. AI systems conduct sophisticated, multi-stage attacks that unfold faster than human-speed response allows. Critical infrastructure protection requires AI systems with autonomous authority to take defensive action—raising governance questions about machine autonomy in security contexts. Some organisations and even small nations find themselves unable to defend against AI-enabled attackers with far greater resources. The boundary between cybercrime, corporate espionage, and state conflict blurs as AI tools become widely accessible. International discussions about AI weapons controls intensify but struggle to keep pace with capability proliferation. [163, 215]

3.8.5 What if done right

Agent-based software engineering makes overall software much more reliable, safe, and secure while individualizing the user interface. The same AI capabilities that create offensive risks are harnessed for defence: comprehensive code review, automated vulnerability detection, continuous security monitoring, and rapid incident response.

Due to a separation of access or even infrastructure, critical infrastructure and business services are outside the attack vector of bad actors. Network architectures evolve to create security boundaries that AI-enabled attacks cannot easily cross. Air-gapped systems, cryptographic verification, and hardware-enforced access controls create zones of security that remain robust even as AI attack capabilities advance.

Cyber attacks fall short not because attackers lack capability, but because defensive infrastructure is designed from the ground up for AI-era threats. The security community achieves what seemed impossible: staying ahead of AI-enabled offence through AI-enabled defence combined with fundamental infrastructure redesign.

3.8.6 Opportunities and risks

Opportunities:

- AI-enhanced threat detection and response
- Automated security operations reducing human burden
- Predictive security analytics enabling proactive defence
- More reliable and secure software through AI-assisted development

Risks:

- Offence-defence asymmetry favouring attackers
- Autonomous attacks at unprecedented scale and speed
- Critical infrastructure vulnerability
- Escalation risks from AI-enabled conflict
- Agentic systems creating new attack surfaces

3.9 Impact category 7: Digital public sphere & democracy

3.9.1 Why this category matters

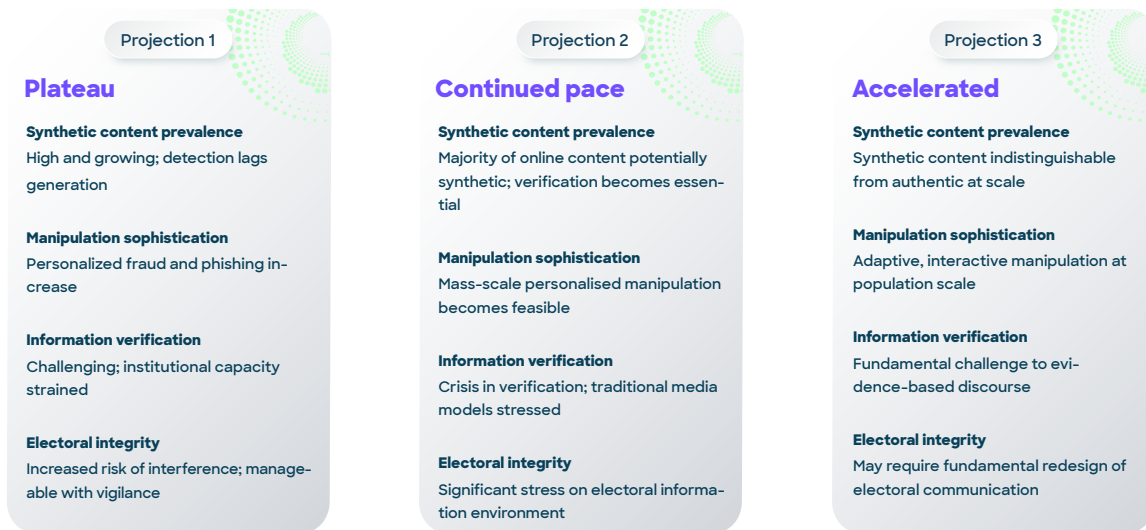
AI capabilities—particularly generative AI—directly affect the information environment that democracies depend on. Synthetic content, personalised manipulation, and the erosion of shared factual ground pose fundamental challenges to democratic deliberation. Agents are now able to further target and execute actions against selected individuals, companies, or countries—moving beyond passive content generation to active intervention.

3.9.2 Key evidence (2025–2026)

Finding	Source
Machine-generated content reached 52% of internet content by 2025.	[124]
Deepfakes doubled every six months; projected 8 million deepfakes shared in 2025 (up from 500,000 in 2023).	[125]
Voice cloning needs only 3 seconds of audio to create 85% accurate clone; technology is cheap and widely accessible.	[126]
57% of Americans rate AI's societal risks as high (weakening skills, relationships); only 25% see benefits as high.	[127]
Europol estimates 90% of online content may be synthetically generated by 2026.	[128]
Global Risks Report 2026 ranks adverse AI outcomes, including deepfakes, as fastest-growing threat to evidential integrity.	[15]

3.9.3 Impacts by projection

First-order impacts



Second-order impacts



3.9.4 What this looks like in each projection

P1 (Plateau): Deepfakes and synthetic content are already common nuisances. A politician releases a video; within hours, doctored versions circulate. Fact-checkers struggle to keep pace. Voice-cloning scams targeting elderly relatives become routine; financial losses mount. News organisations develop AI detection tools, but they're always behind the generation capability. Public trust in online content continues to erode, with people increasingly relying on personal relationships and familiar sources rather than institutional verification. Elections proceed with heightened vigilance, but no catastrophic incidents occur.

P2 (Continued pace): The information environment reaches a tipping point. Much more online content is AI-generated than human-created. Social media platforms struggle to enforce authenticity policies as the volume of synthetic content overwhelms moderation. A significant election somewhere is marred by AI-generated disinformation that spreads faster than fact-checkers can respond. Journalists increasingly use AI for content production, raising questions about verification and accountability. Some societies experiment with «verified zones» of the internet where identity is authenticated; others resist such controls as threatening anonymity. The phrase «seeing is believing» becomes obsolete.

P3 (Accelerated): The concept of «authentic» digital content becomes nearly meaningless. AI systems can generate text, audio, and video indistinguishable from human creation in real-time, personalised for individual recipients. A single operator can conduct influence operations that previously required state resources. Some countries experience democratic crises triggered by information environment collapse; others implement aggressive identity verification and content authentication that create their own civil liberties concerns. The boundary between «real» and «synthetic» blurs so completely that new social norms emerge—perhaps privileging in-person interaction, perhaps developing new forms of cryptographic verification, perhaps simply accepting that digital content is inherently untrustworthy.

3.9.5 What If done right

Separation of infrastructures, authentication systems, provenance tracking, or other approaches successfully verify human content. The technical challenge of distinguishing human from AI-generated content is solved not through detection (which loses the arms race) but through authentication (which changes the game).

AI is used as a means for immense creativity and creation while trust is maintained. Human creators use AI tools to produce art, journalism, education, and entertainment at unprecedented scale and quality—but robust systems ensure that human authorship and responsibility can be verified when it matters.

There is freedom of speech, but it remains a human right. The philosophical and legal frameworks evolve to distinguish between human expression (protected) and AI-generated content (regulated differently). Agentic attacks and misinformation campaigns fall short due to identification systems that reliably identify what's human.

The result is an information environment that's richer and more creative than ever, but where trust and verification are possible when needed. Democratic discourse adapts to new realities rather than collapsing under them.

3.9.6 Opportunities and risks

Opportunities:

- AI tools for content verification and provenance
- Enhanced accessibility through translation and summarization
- Potential for more informed citizenry if quality information surfaces
- Immense creative potential if trust is maintained

Risks:

- Erosion of shared factual ground for democratic deliberation
- Manipulation of elections and public opinion at scale

- Collapse of trust in authentic media and evidence
- Concentration of influence among those who control AI systems
- Agentic attacks targeting individuals, companies, or countries

3.10 Impact category 8: Health & care

3.10.1 Why this category matters

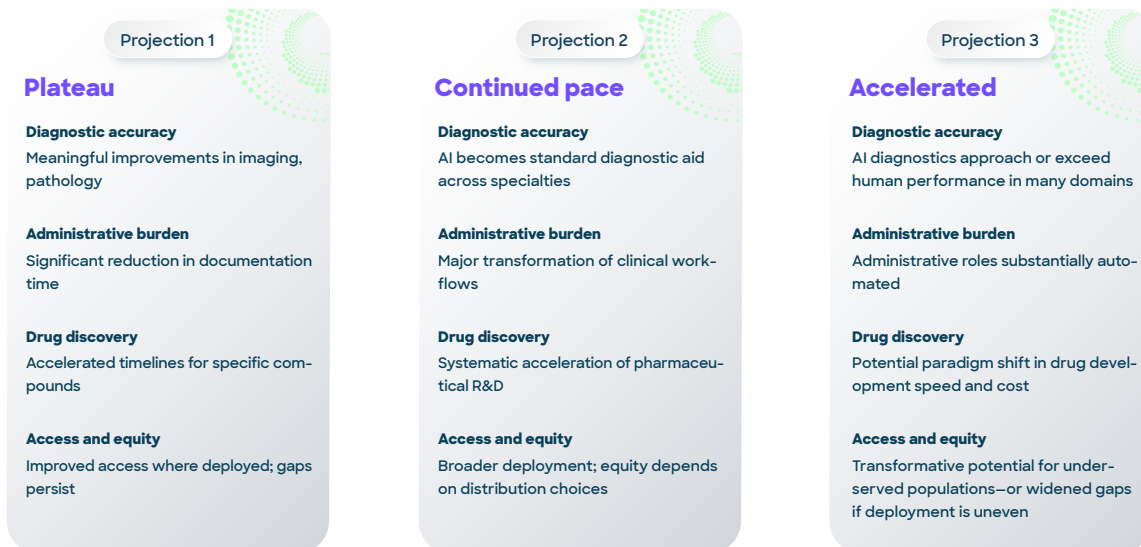
Healthcare represents both enormous potential for AI benefit (diagnostics, drug discovery, administrative efficiency) and significant risks (safety, equity, workforce transformation). The sector’s combination of high stakes, data richness, and regulatory complexity makes it a critical test case.

3.10.2 Key evidence (2025–2026)

Finding	Source
AI triage for intracranial hemorrhage reduced 30-day mortality from 27.7% to 17.5% (37% relative risk reduction).	[130]
Sepsis AI detected 82% of cases (nearly 2x prior methods) with fewer false alarms, linking to ~20% mortality reduction.	[131]
FDA-cleared AI foundation model achieved 97% sensitivity, 98% specificity across 14 critical CT findings.	[132]
Over 90% of healthcare leaders prioritise AI for documentation and clinical decision-making within 12–24 months.	[133]
Generative AI accelerates drug discovery from years to months through molecular generation and interaction simulation.	[134]
AI healthcare market: \$14.6B (2024), projected \$80–188B by 2030–2036.	[135]

3.10.3 Impacts by projection

First-order impacts



Second-order impacts



3.10.4 What this looks like in each projection

P1 (Plateau): A radiologist now works alongside AI that pre-screens imaging for critical findings, allowing faster triage of urgent cases. Documentation time has dropped significantly as ambient AI scribes capture clinical encounters. Drug companies use AI to accelerate certain phases of discovery, but the overall development timeline remains years-long due to clinical trials and regulatory requirements. Patients in well-resourced settings benefit from earlier diagnosis and more efficient care; those in under-resourced settings see smaller improvements. The healthcare workforce remains largely stable, though new roles emerge around AI oversight and maintenance.

P2 (Continued pace): AI becomes a standard clinical tool across most specialties. Primary care physicians rely on AI-assisted diagnosis that catches conditions they might have missed; specialists find their expertise augmented but also partially commoditised. Pharmaceutical companies achieve breakthroughs that previously seemed decades away—new treatments for previously intractable conditions reach clinical trials in compressed timeframes. Healthcare workforce discussions intensify: some roles (medical coders, certain diagnostic technicians) face displacement, while others (AI-healthcare integrators, patient navigators) grow. Rural and underserved areas gain access to specialist-level diagnostic capability through telemedicine augmented by AI.

P3 (Accelerated): Healthcare transforms more in 3 years than in the previous 30. AI systems achieve diagnostic performance exceeding human specialists across many conditions; the role of physicians shifts towards judgment, communication, and oversight rather than pattern recognition. Drug discovery pipelines produce candidates in months rather than years. Some hospitals restructure dramatically, with AI handling much of the diagnostic and administrative work while humans focus on physical care and patient interaction. Medical education debates what to teach when AI handles much of the cognitive work. Healthcare costs potentially decline substantially—or the benefits are captured by providers rather than patients, depending on policy choices.

3.10.5 What if done right

A real system for individual wellbeing emerges, supporting humans in terms of nutrition, sleep, exercise, diagnostics, and treatments. Healthcare shifts from reactive sick-care to proactive wellness optimisation. AI systems that understand individual biology, lifestyle, and preferences provide personalised guidance that actually works—because it's tailored to each person rather than population averages.

This is available for everyone to decide how deeply integrated, which data to share, and how intrusive AI should be in their own life. Privacy and autonomy are preserved: some people choose comprehensive monitoring and proactive intervention; others prefer minimal involvement; most find a middle ground that suits their preferences. The system respects these choices rather than imposing one-size-fits-all approaches.

The result is dramatically improved health outcomes at lower cost, with benefits distributed equitably rather than concentrated among the wealthy. The combination of AI-enabled precision, respect for individual choice, and universal access creates healthcare that actually promotes health rather than just treating disease.

3.10.6 Opportunities and risks

Opportunities:

- Earlier and more accurate diagnosis saving lives
- Liberation of clinicians from administrative burden
- Accelerated development of new treatments
- Expanded access to specialist-level insights
- Personalized wellness systems respecting individual choice

Risks:

- Safety incidents undermining trust
- Equity gaps if deployment favours affluent populations
- Workforce disruption in clinical and administrative roles
- Liability uncertainty slowing beneficial adoption

3.11 Impact category 9: Education system

3.11.1 Why this category matters

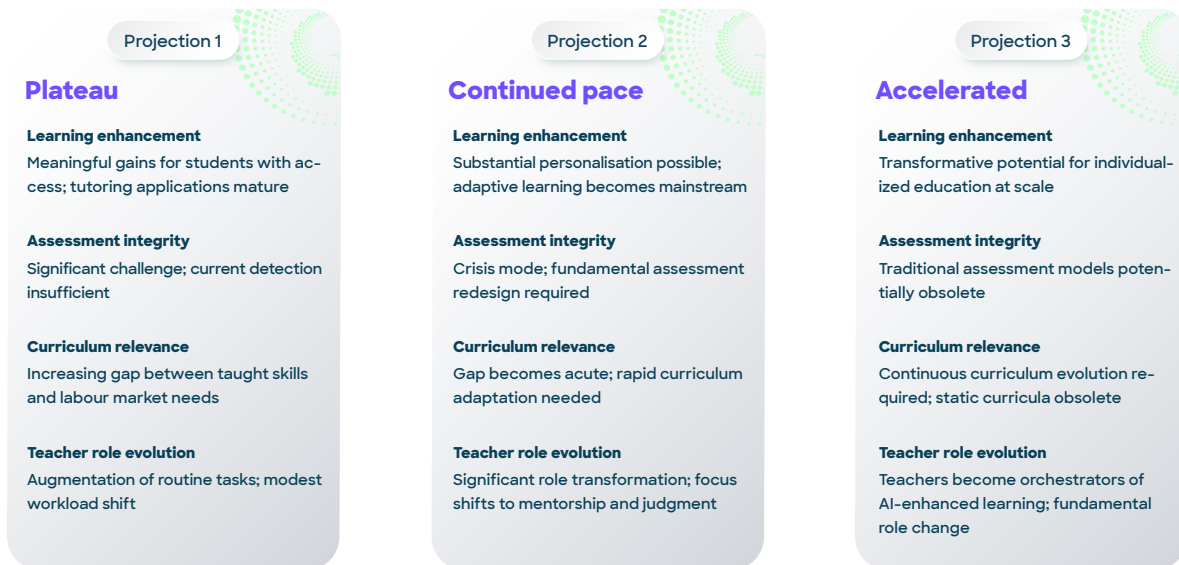
Education systems must simultaneously adapt to AI as a tool, prepare students for AI-transformed labour markets, and maintain integrity in an era of AI-generated content. The challenge spans curriculum, pedagogy, assessment, and equity.

3.11.2 Key evidence (2025–2026)

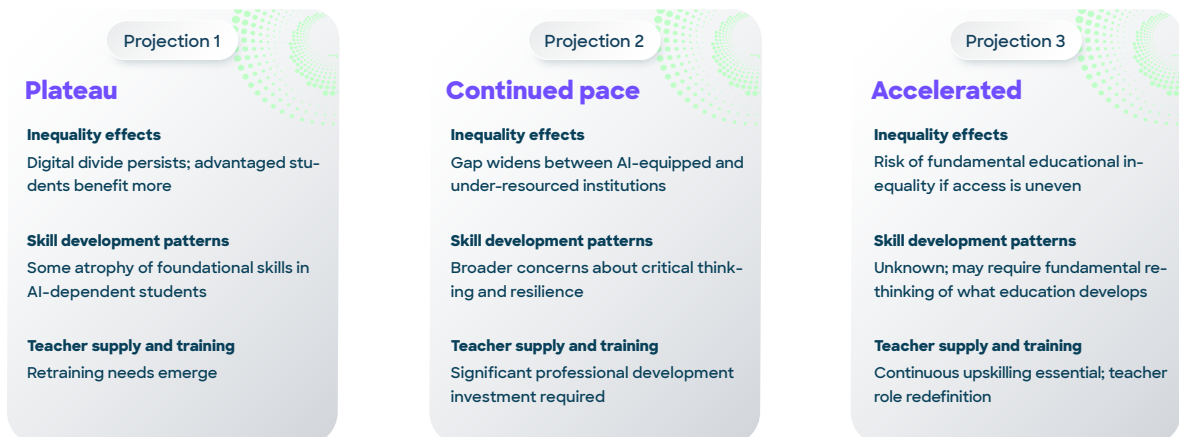
Finding	Source
62% increase in test scores for students using AI-powered instruction due to gap identification and instant feedback.	[136]
86% of K-12 students used AI in 2024–25, amid concerns about downsides.	[137]
AI use for UK university assessments rose from 53% (2024) to 88% (2025).	[138]
71% of teachers report added burden verifying student work authenticity.	[139]
69% of teachers say AI improves methods; 55% report more time for student interaction.	[140]
74% of US districts provided AI training by 2025 (up from 48% in 2023).	[141]
AI encourages cognitive offloading and dependency, weakening foundational knowledge and critical thinking.	[142]

3.11.3 Impacts by projection

First-order impacts



Second-order impacts



3.11.4 What this looks like in each projection

P1 (Plateau): A university professor struggles with the reality that 80%+ of students use AI for writing assignments. Detection tools are unreliable; redesigning all assessments is time-consuming. Some departments shift towards oral examinations or project-based assessment; others attempt to «ban» AI with limited success. K-12 schools vary widely—some embrace AI tutoring tools that show genuine learning improvements; others lack resources or training to adopt. The gap between well-resourced and under-resourced schools widens. Curriculum discussions intensify about what skills students need when AI handles much of the routine work.

P2 (Continued pace): Educational assessment enters genuine crisis. Traditional examinations become nearly meaningless as AI can produce expert-level work on demand. Universities experiment radically—some eliminating written assignments entirely; others developing sophisticated proctored oral or practical examinations. K-12 education splits: leading districts redesign curricula around AI augmentation, teaching students to use AI as a tool while developing human skills like creativity, ethics, and judgment. Trailing districts continue traditional approaches that increasingly disconnect from labour market reality. Teacher roles transform; the best teachers become AI-augmented learning coaches while others struggle with obsolescence.

P3 (Accelerated): Education systems face existential questions about their purpose. If AI can perform most cognitive tasks at expert level, what should schools teach? Some argue for emphasis on distinctly human qualities—emotional intelligence, physical skills, ethical reasoning, creativity. Others advocate for intensive AI fluency so students can direct and supervise AI systems. Traditional age-graded, subject-divided schooling looks increasingly anachronistic. Some jurisdictions attempt radical redesigns; others cling to traditional models that produce graduates poorly matched to available work. Educational inequality becomes a dominant political issue as the gap between AI-prepared and AI-unprepared young people creates visible divergence in life outcomes.

3.11.5 What if done right

Imagine kids grow up with personal teachers that individually and over long periods of time support them in education. Not replacing human teachers, but augmenting them—AI tutors that know each child’s learning style, interests, current understanding, and optimal challenges, available whenever the child wants to learn.

Highly individualized content based on their interests, their capabilities, and their learning methods means that every child can learn effectively, not just those who happen to fit the standard approach. A child fascinated by dinosaurs learns math through paleontology; a child who struggles with reading gets patient, adaptive support that never judges or gets frustrated.

A unified understanding emerges of what counts in human development: critical thinking, asking the right questions, emotional skills, ethical reasoning, creativity, collaboration. These are the capacities that AI cannot replace and that become more valuable as AI handles routine cognitive tasks.

Education becomes available for everyone independent of parents’ wealth or country of origin. The accident of birth no longer determines access to excellent education. A child in a rural village has access to the same quality of personalised instruction as a child in a wealthy capital city. The combination of AI capability and universal access creates educational equity that seemed impossible.

3.11.6 Opportunities and risks

Opportunities:

- Personalized learning at scale
- Teachers freed from routine tasks for higher-value interaction
- Democratized access to high-quality educational content
- Real-time feedback and adaptive curricula
- Universal access regardless of geography or wealth

Risks:

- Assessment integrity crisis undermining credential value
- Cognitive dependency reducing independent thinking
- Digital divide creating educational inequality
- Curriculum lag leaving graduates unprepared

3.12 Impact category 10: Functioning local institutions & liveability

3.12.1 Why this category matters

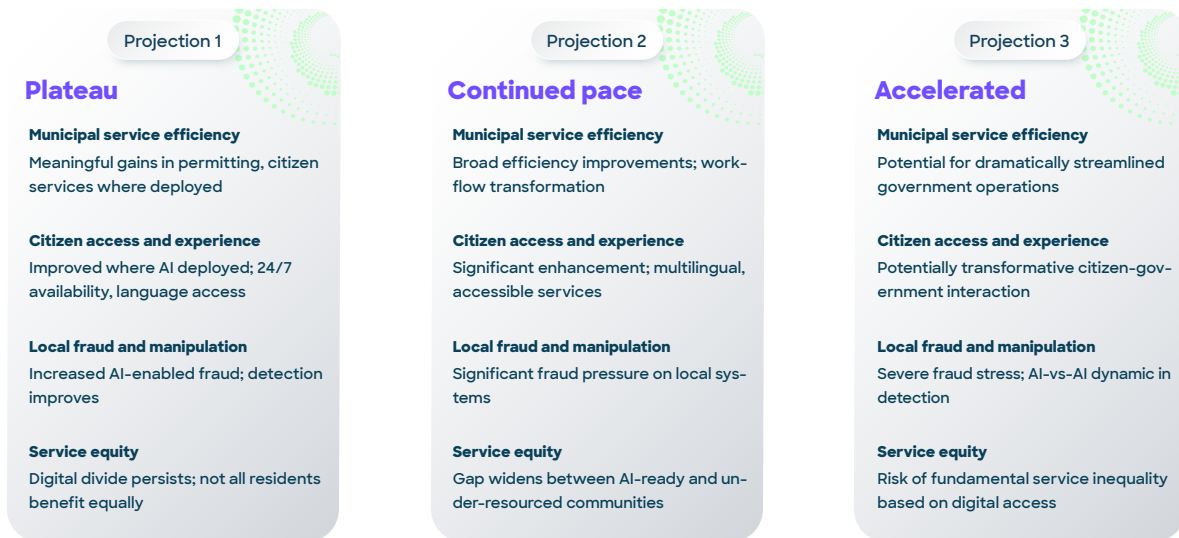
Beyond national-level effects, AI transformation will be experienced most directly at the local level—in municipal services, citizen interactions with government, local economic vitality, and community cohesion. The capacity of local institutions to adapt determines whether AI benefits are broadly shared or concentrated.

3.12.2 Key evidence (2025–2026)

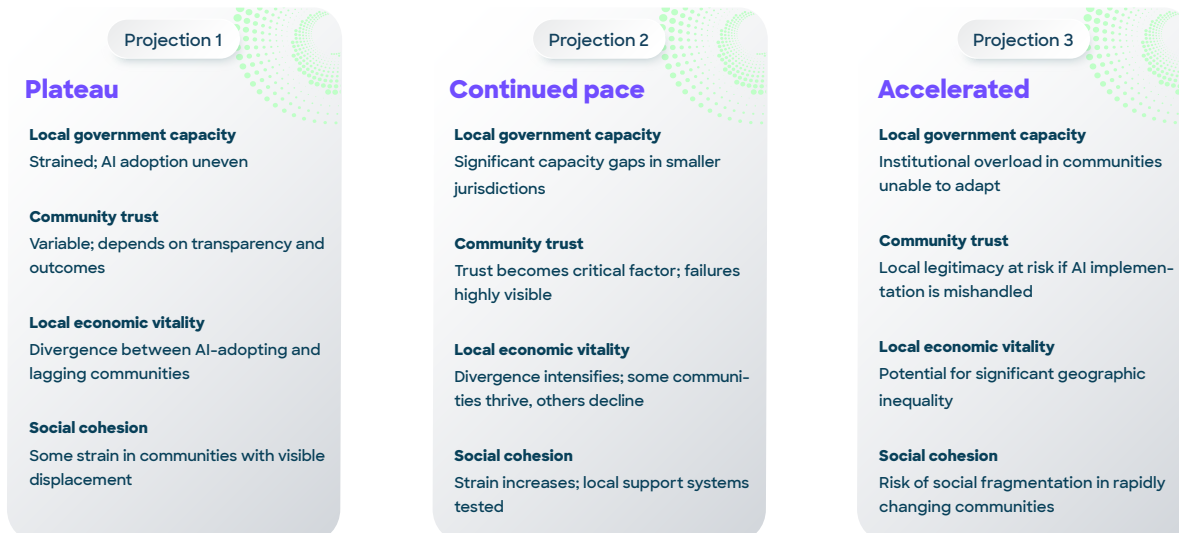
Finding	Source
67% of cities are actively integrating AI into operations (top priorities: citizen services, infrastructure, compliance).	[143]
Burlington, VT reduced permit approval from 15 weeks to 5–7 weeks using AI systems.	[144]
San Jose AI Upskilling Programme saved \$50,000 and 10,000–20,000 staff hours in first two cohorts.	[145]
Only 6% of local governments place AI in their high-priority category.	[146]
77% of Americans distrust both businesses and government to use AI responsibly.	[147]
45% of residents say digital government services still require substantial improvement despite AI adoption.	[148]
California Community Colleges estimated ~1/3 of financial aid applicants were fraudulent in 2024; AI detection deployed.	[149]

3.12.3 Impacts by projection

First-order impacts



Second-Order Impacts



3.12.4 What this looks like in each projection

P1 (Plateau): A citizen applying for a building permit in a forward-thinking city encounters a smooth, AI-assisted process—the chatbot answers questions accurately, documents are processed in days rather than weeks, and translation services make the process accessible. But in a neighboring jurisdiction with less capacity, the experience is unchanged from a decade ago. Some cities use AI to improve pothole detection and infrastructure maintenance; others lack budget or expertise. Fraud targeting local governments increases—AI-generated voice impersonation attempts to redirect payments—but detection tools help. Overall, service quality varies significantly by jurisdiction, creating visible inequality in local government capability.

P2 (Continued Pace): The gap between AI-enabled and traditional local government widens dramatically. Leading cities offer genuinely transformed citizen experience—near-instant permitting, proactive service delivery, multilingual 24/7 assistance. Lagging jurisdictions see resident frustration as expectations set by leading cities aren't met locally. Local government employment shifts as AI handles routine administrative work; some jurisdictions manage transitions smoothly while others face workforce conflicts. Fraud pressure intensifies—sophisticated AI-enabled schemes target benefit programmes, tax systems, and emergency services. Communities that invested early in AI capacity thrive; those that didn't struggle to maintain basic services.

P3 (Accelerated): Local government faces potential crisis. The speed of change exceeds institutional capacity in many jurisdictions. Cities that successfully transformed deliver services at quality and efficiency levels that seemed impossible—but they're the minority. Other communities face simultaneous challenges: increased demand from residents affected by labour market disruption, sophisticated fraud depleting resources, and inability to compete for talent with private sector. Some small cities and rural areas find that AI enables service delivery they previously couldn't afford; others are left further behind. Regional inequality becomes a dominant political issue as the local experience of AI transformation varies enormously by geography. Groups of people demand slowing down change, which is not possible. New movements and communities emerge that demand an AI-free life, seeking to preserve traditional ways of living.

3.12.5 What if done right

Due to a broad diffusion of AI-powered services, AI is broadly considered as benefiting communities and individual lives. The transformation is not imposed from above but emerges from genuine local benefit: services that actually work better, costs that actually decrease, access that actually expands.

Resistance to change and even movements that want to reverse to a life before AI appear and are accepted. Societies make space for communities and individuals who choose different relationships with technology. This diversity is respected rather than suppressed.

But due to the overall benefits in local communities, the acceptance of AI remains strong. When citizens experience AI as making their lives better—faster services, fairer treatment, more responsive government—the political foundation for continued adoption remains solid. Trust is earned through demonstrated benefit, not demanded through mandate.

The result is local communities that are more liveable, more equitable, and more resilient—not despite AI but because AI is deployed in service of community values rather than as an end in itself.

3.12.6 Opportunities and risks

Opportunities:

- More efficient, accessible, responsive local government
- 24/7 citizen services in multiple languages
- Data-driven resource allocation improving outcomes
- Liberation of local workers from routine tasks
- Services that benefit all community members

Risks:

- Digital divide creating two-tier local citizenship
- Institutional capacity gaps in under-resourced communities
- Trust erosion from AI failures or opaque decision-making
- Local government unable to adapt fast enough to maintain services
- Social fragmentation from pace of change
- 3.13 Cross-cutting observations

3.13.1 Patterns across impact categories

Several patterns emerge across all ten impact categories:

1. Adoption capacity is universally important: In all projections, the ability to adopt, integrate, and govern AI systems is a primary determinant of outcomes. Organisations and communities with higher capacity realise benefits; those without are left behind or harmed.
2. Early-career and vulnerable populations face disproportionate risk: Entry-level workers, digitally excluded populations, and under-resourced institutions bear the earliest and most severe negative impacts.
3. Trust and integrity are fragile: Across information, science, and governance, AI capabilities stress verification and trust systems. Failures here have cascading effects.
4. The gap widens with capability speed: In P3, all impacts intensify—both opportunities and risks. The window for orderly adaptation shrinks.
5. Second-order effects are substantial: Direct AI impacts cascade through systems (labour → tax base → public services → local institutions). Planning must account for these chains.
6. Robotics is now economically transformative: With 99%+ success rates in production and ROI of 1,400-2,070%, physical automation is no longer speculative—it's a near-term competitive imperative.
7. Innovation economics are restructuring: Few-person unicorns, rapid IP depreciation, and collapsing competitive moats require new frameworks for economic policy.

8. Social safety net adequacy varies dramatically: Countries with robust systems can buffer disruption; those without face severe social strain.
9. “Done right” visions are achievable: Each impact category has a positive path—but realising it requires proactive policy, institutional adaptation, and deliberate choices about how AI benefits are distributed.

3.13.2 Time-critical impact patterns

Some impacts require attention regardless of projection probability because they are already materialising or have long lead times for response:

Impact Pattern	Why time-critical	Projections where Acute
Entry-level labour displacement	Already visible (6–20% in exposed roles); career effects compound	All (P1–P3)
Software engineering transformation	SWE-bench at human parity; majority of code AI-written at leading firms	All (P1–P3)
Assessment integrity in education	88% of UK university students using AI; crisis already emerging	All (P1–P3)
Information environment degradation	52% machine-generated content; detection lagging	All (P1–P3)
SME adoption gap	Structural barrier; widens without intervention	P2, P3
Robotics competitiveness	Chinese manufacturers at 39% global share; costs declining 20–30% annually	P2, P3
Local government capacity	Only 6% prioritise AI; adaptation takes years	P2, P3
Scientific integrity	Hallucinations in peer-reviewed venues; trust erosion underway	All (P1–P3)
Startup economics transformation	Few-person unicorns emerging; traditional models under pressure	P2, P3
Market repricing of AI-exposed sectors	„AI Scare Trade“ demonstrates volatility; structural shifts underway	P2, P3

3.13.3 No-regret impact patterns

Some impact patterns require action regardless of which projection materializes, because they are already emerging (probability = 1 for some level of impact):

- Labour market transition support – needed in all projections; scale varies
- AI literacy and workforce development – essential regardless of capability speed
- Information integrity infrastructure – already urgent; will only intensify
- Trustworthy AI governance capacity – foundational for all other responses
- Digital equity for public services – determines who benefits from AI-enabled government
- Innovation ecosystem adaptation – necessary to remain competitive in any projection

These no-regret patterns inform the measures identified in Chapter 4.

3.14 Summary: Impact Severity by Projection

Impact category	P1: Plateau	P2: Continued pace	P3: Accelerated
Labour market & skills	Moderate (entry-level pressure)	High (broad displacement)	Severe (structural transformation)
Public finance & social systems	Manageable	Significant stress	Potential crisis
Industry & competitiveness	Moderate (adoption gap)	High (competitive gap + market volatility)	Severe (structural decline risk)
Innovation & startups	Moderate (efficiency pressure)	High (model disruption)	Severe (fundamental restructuring)
Science system	Mixed (acceleration + integrity)	High (integrity crisis)	Severe (fundamental challenges)
Security & resilience	Elevated	High	Severe (systemic risk)
Digital public sphere	High (integrity)	Severe (trust crisis)	Potentially catastrophic
Health & care	Positive (efficiency)	Strongly positive (if governed)	Transformative (with governance)

Education system	Significant (integrity)	High (integrity + relevance)	Severe (systemic re-design needed)
Local institutions	Moderate (uneven adoption)	High (capacity gaps)	Severe (institutional overload)

3.15 What this chapter does not determine

This chapter maps impacts but does not:

1. Prescribe specific measures – these are derived in Chapter 4 based on impact severity and probability.
2. Assign responsibility – the owner ecosystem for each measure is specified in Chapter 4.
3. Evaluate European response capacity – this informs measure feasibility in Chapter 4.
4. Prioritize among impacts – prioritisation requires combining impact severity with projection probability (Chapter 4).
- 5.

Chapter 4: Measures & appliedAI portfolio logic

4.1 Purpose of this chapter

This chapter translates the impacts identified in Chapter 3 into concrete measures, distinguishing between:

- Part A: What should be addressed overall (the full measures inventory, regardless of who implements)
- Part B: What appliedAI Institute will do (our portfolio selection based on purpose fit and feasibility)

The logic:

1. Start from impact items in Chapter 3
2. Weight by probability (from Chapter 2) and time criticality
3. Identify measures that address these impacts
4. Distinguish no-regret measures (needed in all projections) from projection-conditional measures
5. Select appliedAI's portfolio based on purpose fit and feasibility

4.2 Measures design principles

4.2.1 Derivation from Impacts

Every measure must map back to:

- An impact (Chapter 3)
- A probability basis (projection probabilities from Chapter 2 and/or measure-level probability)
- Time criticality (0–36 months)

4.2.2 Measure types

Type	Definition	Probability basis
No-regret	Needed regardless of which projection materializes; addresses impacts already emerging or foundational for all responses	Probability = 1
Projection-conditional	Becomes critical primarily in P2 or P3; may be premature or lower priority if P1 materializes	Probability = projection-weighted

4.2.3 Scoring dimensions

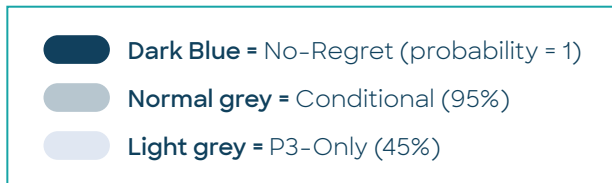
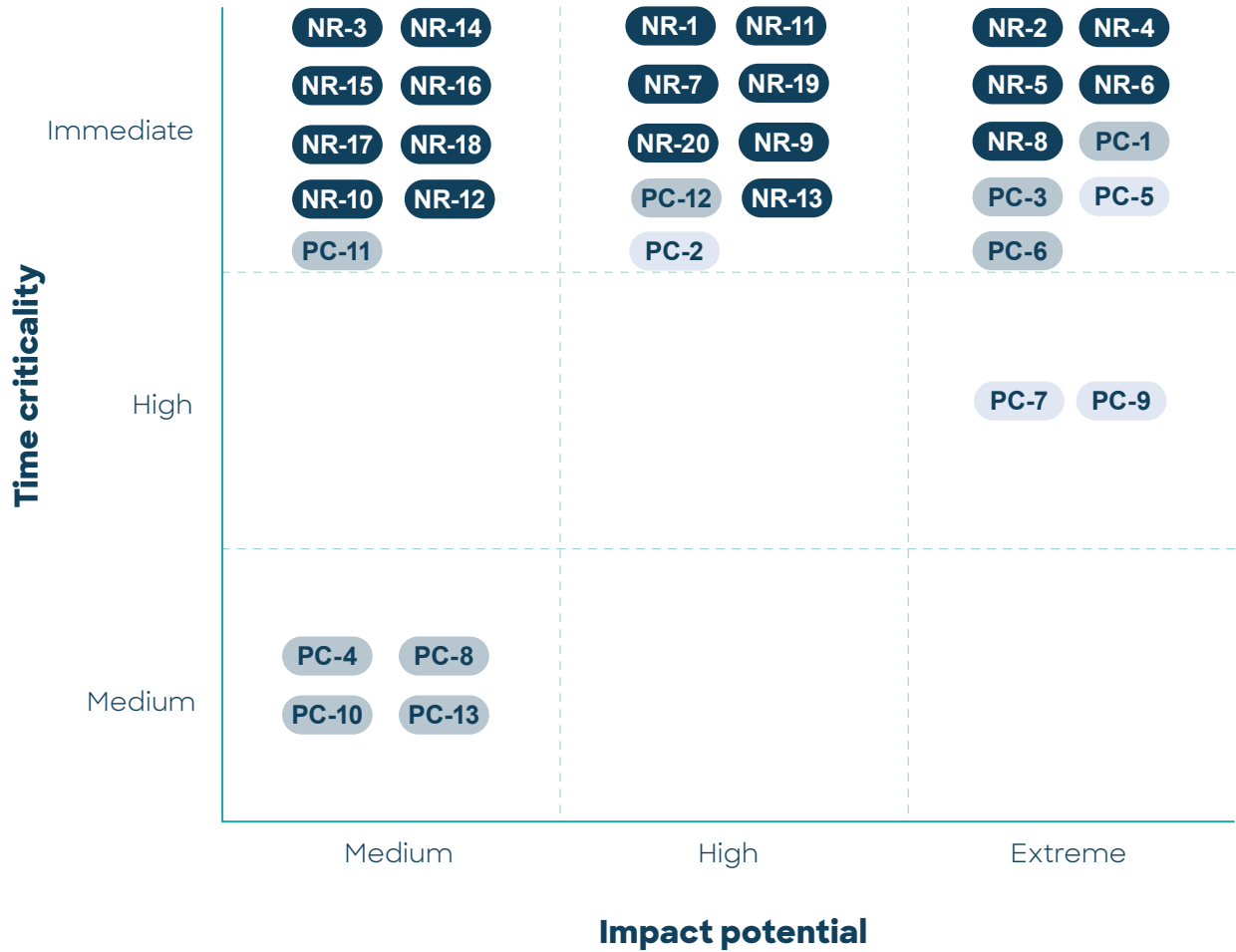
Dimension	Definition
Impact potential	Scale and severity of the impact addressed (low / medium / high / extreme)
Probability	Likelihood that the measure will be needed (1 for no-regret; projection-weighted for conditional)
Time criticality	Urgency of action (low / medium / high) within 0–36 months
Overall value signal	Qualitative synthesis of the above

4.3 Part A – measures inventory (what should be addressed overall)

This section presents the full inventory of measures derived from Chapter 3 impacts. These measures represent what Europe and its institutions should address—regardless of which specific organisation implements them.

4.3.1 No-regret measures (probability = 1)

These measures are valuable across all projections and should be pursued immediately:



NR-1

Development of European target images for people in the AI age

Linked impact (s)

Labour, Education, Social Systems

Evidence basis

Public discourse fragmented; no coherent European narrative

Intended effect

Establish shared vision for human flourishing in AI-transformed society; guide policy and investment

Suggested owners

EU institutions, national governments, civil society, research institutions

NR-7

Promotion and targeted settlement of international AI talent in Europe

Linked impact (s)

Industry, Science, Competitiveness

Evidence basis

Talent flowing to US; European retention challenging

Intended effect

Address talent shortage; reduce brain drain; build European AI capacity

Suggested owners

National governments, universities, industry

NR-2

Establishment of genuine technical and regulatory AI competence centres at the highest level

Linked impact (s)

All categories

Evidence basis

Only 6% of local governments prioritise AI; regulatory capacity strained; AI Office currently lags behind tech developments by >1 year; similar gaps at national level

Intended effect

Build institutional capacity for AI governance across government; ensure informed decision-making

Suggested owners

National governments, EU institutions, public sector academies

NR-8

Supporting and scaling promising European AI start-ups (including government contracts)

Linked impact (s)

Industry, Competitiveness, Innovation

Evidence basis

EU AI investment 7% of US levels; ecosystem underdeveloped

Intended effect

Build European AI ecosystem; reduce dependency; create jobs

Suggested owners

National governments, EU institutions, procurement bodies

NR-3

Creation of concrete 2-year action plan addressing all three scenarios with continuous monitoring

Linked impact (s)

All categories

Evidence basis

Current planning assumes single trajectory; lacks contingency

Intended effect

Ensure preparedness across scenarios; enable rapid response to trajectory changes

Suggested owners

National governments, EU Commission, coordinating bodies

NR-9

Conceptualization of taxation of agentic value creation

Linked impact (s)

Public Finance, Labour

Evidence basis

Current systems assume labour-based value creation; AI shifting value creation patterns

Intended effect

Develop frameworks for taxing AI-generated value; ensure fiscal sustainability as labour tax base erodes

Suggested owners

Finance ministries, tax policy institutions, research institutions

NR-4

Introduction of mandatory AI monitoring (similar to UK AI Safety Taskforce)

Linked impact (s)

Security, Industry, All

Evidence basis

Current monitoring fragmented; lags capability advancement

Intended effect

Track AI capability advancement; identify risks early; enable evidence-based policy

Suggested owners

EU AI Office, national AI authorities, research institutions

NR-10

Creation of government-supported individual teacher platform for future of education

Linked impact (s)

Education, Labour, Social Systems

Evidence basis

88% UK students using AI; traditional education models under stress; need for scalable personalised learning

Intended effect

Enable personalised AI-assisted education at scale; prepare workforce for AI transformation

Suggested owners

Education ministries, national governments, EdTech sector

NR-5

Systematic unblocking of delaying mechanisms for public procurement

Linked impact (s)

All categories

Evidence basis

Procurement 12-24 months; AI advancing in weeks/months; bureaucratic delays prevent adoption

Intended effect

Close gap between AI advancement speed and institutional adaptation; enable rapid deployment of AI solutions

Suggested owners

All levels of government, regulatory bodies, procurement offices

NR-11

Development and introduction of verifiable digital identity for people on the open Internet

Linked impact (s)

Digital Public Sphere, Security, Democracy

Evidence basis

52% content now machine-generated; verification impossible

Intended effect

Distinguish humans from AI actors; combat manipulation; preserve trust

Suggested owners

EU institutions, national governments, standards bodies

NR-6

Moonshots for next-generation AI architectures from Europe

Linked impact (s)

Industry, Science, Competitiveness, Sovereignty

Evidence basis

EU controls 4.8-5% of compute; dependent on foreign models; Nested Learning, Causal AI emerging as new paradigms

Intended effect

Develop European capacity in frontier AI architectures; reduce dependency on foreign models; pursue paradigm-shifting research

Suggested owners

EU institutions, research institutions, national governments

NR-12

Monitoring and rapid response systems for massive digital manipulation attacks

Linked impact (s)

Digital Public Sphere, Democracy, Security

Evidence basis

Deepfakes doubling every 6 months; phishing up 400%; agentic attacks emerging

Intended effect

Detect and respond to AI-enabled influence operations; protect elections; enable coordinated response to manipulation waves

Suggested owners

National security agencies, election authorities, platforms

NR-13

Skills exposure study (Europe-wide, role/task-based)

Linked impact (s)

Labour Market, Public Finance

Intended effect

Identify most affected roles and regions; enable targeted response

Evidence basis

Entry-level displacement 6–20% already visible

Suggested owners

Research institutions, labour ministries, appliedAI

NR-17

Assessment redesign frameworks for education

Linked impact (s)

Education

Intended effect

Address integrity crisis; maintain credential value

Evidence basis

88% UK students using AI; traditional assessment failing

Suggested owners

Education ministries, examination boards, universities

NR-14

Scalable reskilling and upskilling platform (role-based)

Linked impact (s)

Labour Market, Industry

Intended effect

Enable workforce transition; maintain employability; scale AI literacy across workforce

Evidence basis

Current training capacity insufficient for projected need; 13.48% EU adoption; skills shortage primary barrier

Suggested owners

Education providers, industry, public employment services, EU institutions

NR-18

Scientific integrity infrastructure for AI era

Linked impact (s)

Science, Innovation

Intended effect

Address hallucination and fraud risks; maintain trust in research

Evidence basis

53 NeurIPS papers with hallucinated citations; fraud scaling

Suggested owners

Research funders, publishers, scientific academies

NR-15

Trustworthy AI engineering playbooks and reference architectures

Linked impact (s)

All categories (enabler)

Intended effect

Establish practical standards for responsible deployment

Evidence basis

Adoption barriers include uncertainty about „how to do it right“

Suggested owners

Standards bodies, industry associations, appliedAI

NR-19

Innovation ecosystem adaptation framework

Linked impact (s)

Innovation, Startups, Competitiveness

Intended effect

Support European startups in AI-transformed competitive landscape; address few-person unicorn dynamics

Evidence basis

Traditional startup models under pressure; revenue per employee transforming

Suggested owners

Innovation agencies, VCs, startup associations, appliedAI

NR-16

Local government adoption blueprints (municipal workflows)

Linked impact (s)

Local Institutions, Liveability

Intended effect

Enable local government AI adoption; reduce capacity barriers

Evidence basis

67% of cities integrating AI but most lack blueprints

Suggested owners

Municipal associations, national governments, appliedAI

NR-20

SME onboarding support for AI adoption

Linked impact (s)

Industry, Competitiveness, Local Institutions

Intended effect

Enable SMEs to adopt AI; close adoption gap between large enterprises and smaller firms

Evidence basis

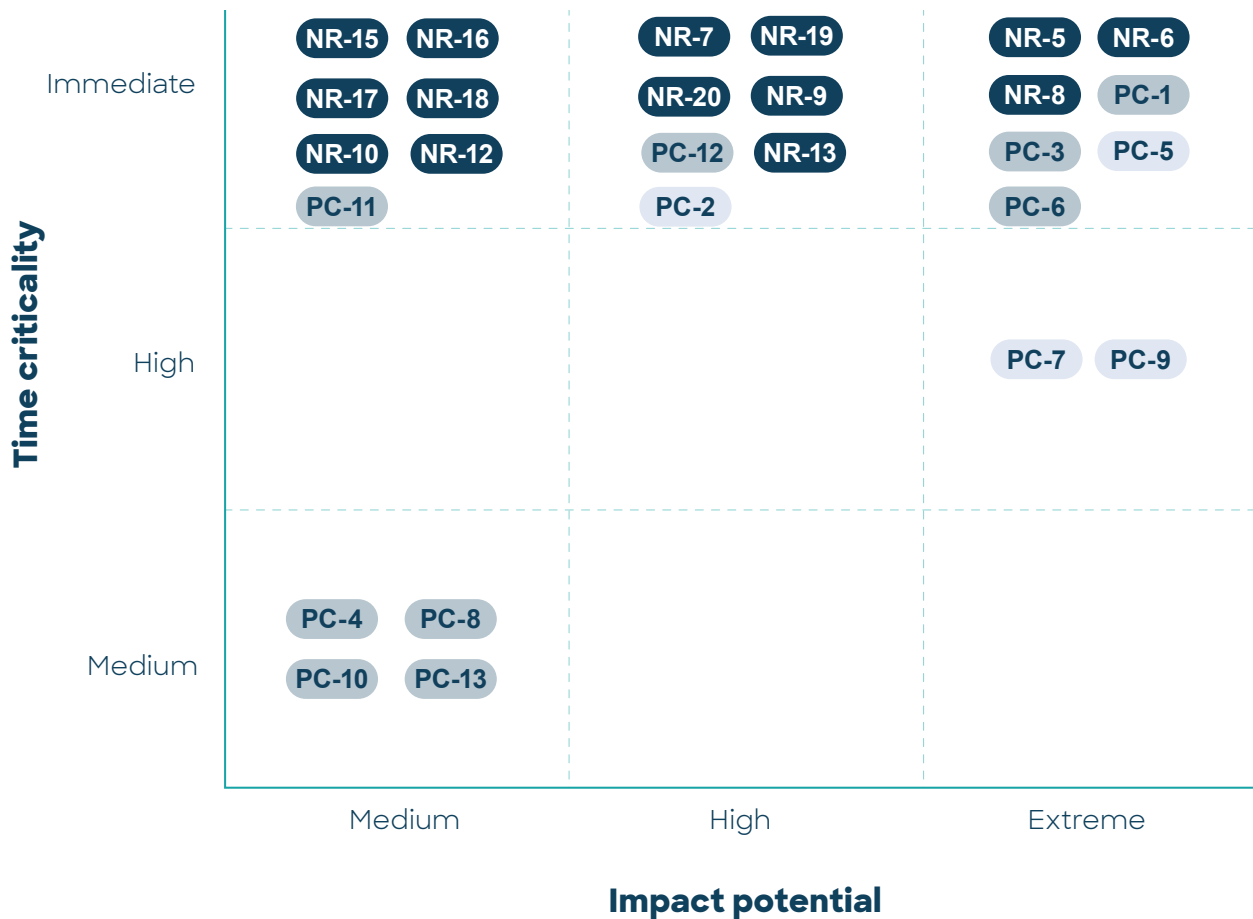
41% of large EU enterprises have AI vs. ~12% of SMEs; SMEs lack resources and expertise for adoption

Suggested owners

Industry associations, chambers of commerce, appliedAI, regional development agencies

4.3.2 Projection-conditional measures (P2/P3)

These measures become critical primarily if P2 or P3 materializes. Given our probability estimates (P2: 50%, P3: 45%), these warrant serious preparation, while P2 measures become relevant for P2 and P3 giving a 95% total probability that those measures become relevant:



	Dark Blue = No-Regret (probability = 1)
	Normal grey = Conditional (95%)
	Light grey = P3-Only (45%)

PC-1

Establishment of European early warning system for technological leaps, misuse risks, and geopolitical shifts

Linked impact (s)
All categories

Probability basis
S2/S3 (95%)

Intended effect
Detect and respond to rapid capability changes; maintain strategic awareness

Suggested owners
Intelligence services, EU AI Office, research institutions

PC-8

Healthcare AI governance fast-track frameworks

Linked impact (s)
Health & Care

Probability basis
S2/S3 (95%)

Intended effect
Enable rapid deployment while maintaining safety

Suggested owners
Health ministries, regulatory agencies

PC-2

Establishment of verified and decoupled „economic internet“ for strategically important entities

Linked impact (s)
Security, Industry, Critical Infrastructure

Probability basis
Primarily S3 (45%)

Intended effect
Ensure communication resilience; protect against AI-enabled attacks

Suggested owners
National security agencies, critical infrastructure operators

PC-9

Emergency fiscal measures for tax base transition

Linked impact (s)
Public Finance

Probability basis
Primarily S3 (45%)

Intended effect
Prepare revenue alternatives if labour taxes decline rapidly

Suggested owners
Finance ministries

PC-3

Preparation for AI agent-conducted cyberattacks on undetected vulnerabilities

Linked impact (s)
Security, Industry

Probability basis
S2/S3 (95%)

Intended effect
Develop defensive capabilities against autonomous attack systems

Suggested owners
Cybersecurity agencies, critical infrastructure operators

PC-10

AI native educational pathways (parallel to traditional)

Linked impact (s)
Education

Probability basis
S2/S3 (95%)

Intended effect
Develop education models for AI-transformed economy

Suggested owners
Education ministries, universities

PC-4

Strategic massive expansion of European computing capacities and energy supply

Linked impact (s)
Industry, Competitiveness, Sovereignty

Probability basis
S2/S3 (95%)

Intended effect
Reduce dependency; enable European AI development

Suggested owners
National governments, EU institutions, energy providers

PC-11

International coordination on AI safety and governance

Linked impact (s)
Security, All categories

Probability basis
S2/S3 (95%)

Intended effect
Align responses across jurisdictions; prevent race to bottom

Suggested owners
Foreign ministries, EU institutions, international bodies

PC-5

Establishment of rapid intervention mechanisms for disruptive AI leaps

Linked impact (s)
All categories

Probability basis
Primarily S3 (45%)

Intended effect
Enable fast policy response to unexpected capability jumps

Suggested owners
National governments, EU institutions

PC-12

Union-Industry agreements on joint visions for the future of work

Linked impact (s)
Labour, Industry, Social Cohesion

Probability basis
S2/S3 (95%)

Intended effect
Create shared understanding between employers and workers on AI-driven workplace transformation; enable collaborative transition planning

Suggested owners
Trade unions, industry associations, labour ministries, social partners

PC-6

Greenfield AI-First test environments (education, health, food, energy, materials, social)

Linked impact (s)
Multiple sectors

Probability basis
S2/S3 (95%)

Intended effect
Develop and test AI-native approaches without legacy constraints

Suggested owners
National governments, sector agencies, research institutions

PC-13

Simulation of new social systems

Linked impact (s)
Public Finance, Labour, Social Systems

Probability basis
Primarily S3 (45%)

Intended effect
Model and test alternative social contracts (e.g., UBI variants, new welfare models) before implementation becomes urgent

Suggested owners
Social policy institutions, research institutions, finance ministries

PC-7

Contingency concepts for severe labour displacement

Linked impact (s)
Labour, Public Finance, Social Cohesion

Probability basis
Primarily S3 (45%)

Intended effect
Prepare response options for rapid workforce disruption

Suggested owners
Labour ministries, social policy institutions

4.3.3 Measures prioritisation matrix

The following matrix visualizes measure priority based on probability and impact potential:

	Low impact	Medium impact	High impact	Extreme impact
Probability = 1 (No-regret)		NR-3, NR-14, NR-15, NR-16, NR-17, NR-18	NR-2, NR-4, NR-7, NR-9, NR-10, NR-11, NR-13, NR-19	NR-1, NR-5, NR-6, NR-8, NR-12
Probability = 95% (P2+P3)		PC-11	PC-4, PC-8, PC-10, PC-12, PC-13	PC-1, PC-3, PC-6
Probability = 45% (P3 only)			PC-2	PC-5, PC-7, PC-9

4.4 Part B – appliedAI Portfolio Selection (What We Will Do)

4.4.1 Portfolio selection filters

appliedAI Institute selects its portfolio from the measures inventory based on four filters:

Filter	Definition	Implications
Purpose fit	Alignment with appliedAI's mission: empowering professionals to develop and apply AI technologies in a trustworthy manner	Focus on competence building, responsible deployment, practical guidance
Feasibility	Within appliedAI's mandate and capabilities as a non-profit research and competence organisation	Cannot directly change political frameworks; can provide evidence, methods, training, convening
Comparative advantage	Areas where appliedAI has distinctive capability: methods development, engineering expertise, ecosystem convening, neutral guidance, training	Focus on areas where we can add unique value vs. others
Ecosystem complementarity	Avoid duplication; partner where others lead	Focus on gaps; enable rather than replace other actors

4.4.2 What appliedAI cannot do directly

To be transparent about scope, appliedAI Institute cannot directly:

- Change political frameworks or legislation
- Provide fiscal resources at government scale
- Build physical infrastructure (compute, energy)
- Regulate or enforce compliance
- Make binding policy decisions

However, we can:

- Provide evidence and analysis to inform decisions
- Develop methods, frameworks, and playbooks
- Train professionals across sectors
- Convene stakeholders and facilitate collaboration
- Create reference implementations and tools
- Highlight European solutions and capabilities
- Build communities of practice

4.4.3 appliedAI opportunity fields

Based on the filters above, we organise our portfolio into four opportunity fields:

Opportunity Field 1: Skills & Workforce Transformation

Purpose: Build measurable AI competence across European professionals

Linked measures: NR-13, NR-14

What we do:

- appliedAI Skills Framework: Standardized competency definitions for AI roles across sectors
- appliedAI Skills Academy (NR-14): Scalable training programmes for professionals in industry and public sector, based on the appliedAI Skills Framework
- Skills Exposure Research (NR-13): Evidence-based analysis of role transformation and reskilling needs, based on the appliedAI Skills Framework
- Target groups: Corporate professionals, public sector employees, executives, educators

- Opportunity Field 2: Trustworthy AI Engineering & Adoption
- Purpose: Enable responsible, effective AI deployment through practical methods and engineering support
- Linked measures: NR-2, NR-5, NR-15, NR-19

What we do:

- AI Act Accelerator (NR-2 support, NR-5, NR-15): Together with the Bavarian Ministry of Digital Affairs, includes:
- AI Act Implementation Support: Practical guidance for compliance
- Engineering Playbooks: Practical guidance for trustworthy AI system development
- Agent-First Operations (NR-19): Coordinating and contributing to open-source AI-first functional blueprints with a sovereign tech stack

Target groups: AI engineers, product teams, compliance officers, CTOs

Opportunity Field 3: Local Implementation & Ecosystem Building

Purpose: Enable AI adoption at local level and strengthen European AI ecosystem

Linked measures: NR-8, NR-16, NR-19, NR-20

What we do:

- Municipal Agent-First Blueprints (NR-16): Practical guides for local government AI adoption with agentic workflows
- SME Onboarding Support (NR-20): Enabling smaller organisations to benefit from AI through structured adoption programmes
- Startup-SME Matching (NR-20, NR-8): Connecting European startups with SME use cases based on Startup and Use Case Landscapes

Target groups: Local governments, regional development agencies, SMEs, ecosystem builders, startups

Opportunity Field 4: Policy Handrails & Decision Support

Purpose: Provide evidence and guidance for AI governance and policy

Linked measures: NR-1, NR-3, SC-1

What we do:

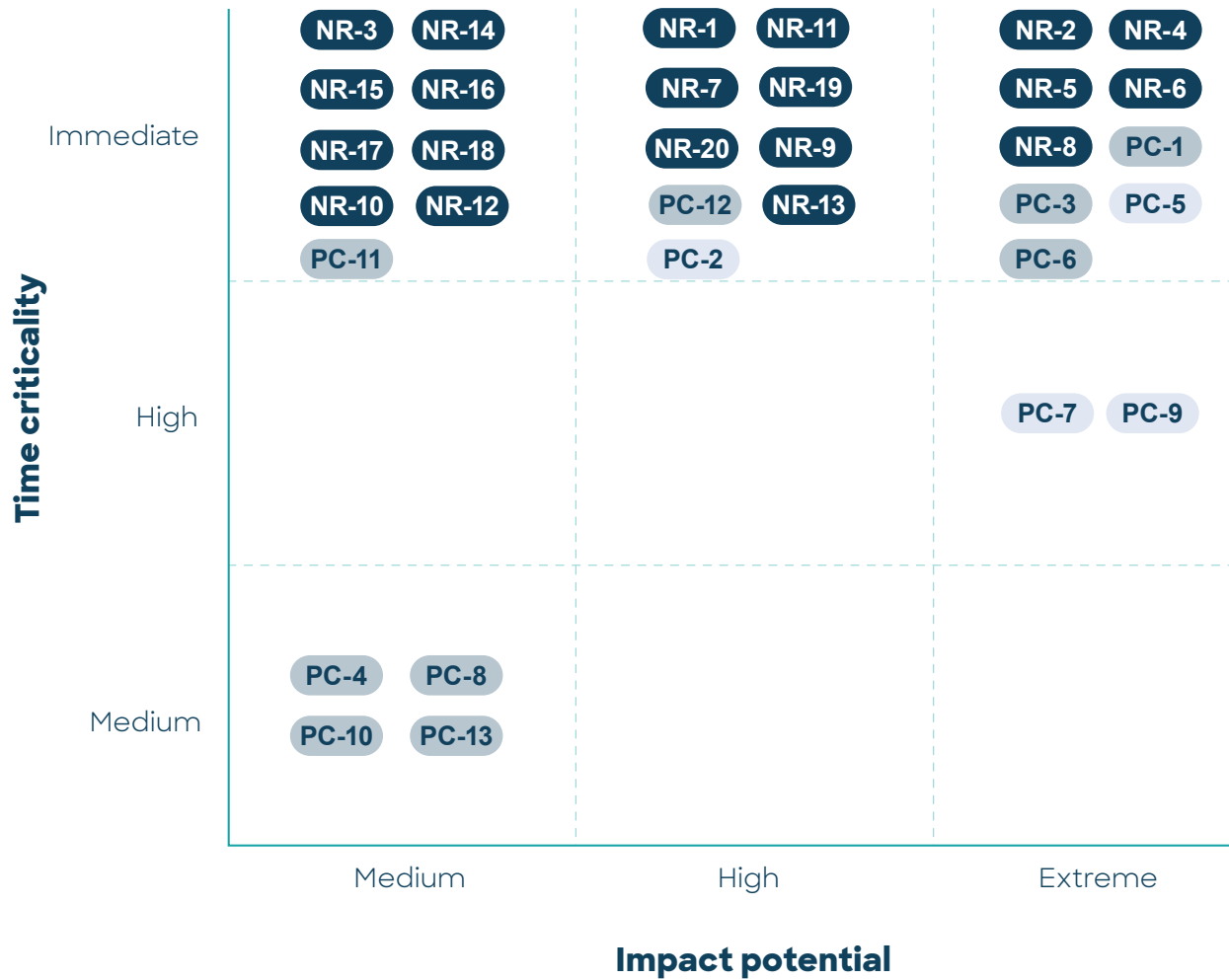
- AI Projection Whitepaper (NR-1, NR-3, SC-1 support): Annual strategic orientation with continuous updates—this document
- Policy Briefs: Concise guidance on specific AI governance challenges
- Target groups: Policy makers, regulators, international bodies

4.4.4 Portfolio summary and measure mapping

Opportunity Field	Primary measures addressed	appliedAI role	Key outputs
Skills & Workforce	NR-13, NR-14	Lead on frameworks, training, research	Skills Framework, Academy, Exposure Studies
Trustworthy Engineering	NR-2, NR-5, NR-15, NR-19	Create methods, build references, provide guidance	AI Act Accelerator, Engineering Playbooks, Agent-First Blueprints
Local Implementation	NR-8, NR-16, NR-19, NR-20	Enable adoption, convene ecosystems	Municipal Blueprints, SME Programs, Startup-SME Matching
Policy Handrails	NR-1, NR-3, PC-1	Provide evidence, develop guidance	AI Projections, Policy Briefs

Measure-to-portfolio mapping:

The following table shows how the measures inventory maps to appliedAI's portfolio and external recommendations:



	Dark Blue = No-Regret (probability = 1)
	Normal grey = Conditional (95%)
	Light grey = P3-Only (45%)

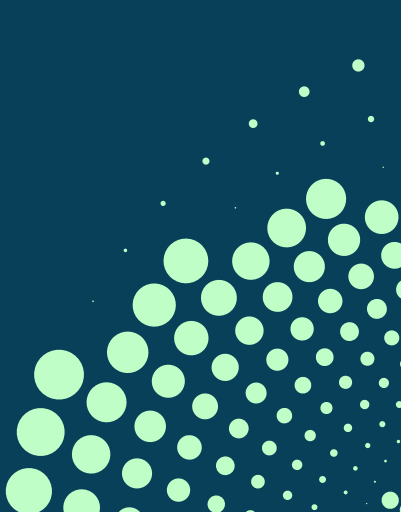
No-regret measures

<p>NR-1 European target images for AI age</p> <p>Opportunity field: Policy handrails</p> <p>Role: Input</p>	<p>NR-14 Reskilling platform</p> <p>Opportunity field: Skills & workforce</p> <p>Role: Lead</p>
<p>NR-2 AI competence centres</p> <p>Opportunity field: Trustworthy engineering</p> <p>Role: Support (via AI Act Accelerator)</p>	<p>NR-15 Engineering playbooks</p> <p>Opportunity field: Trustworthy engineering</p> <p>Role: Lead (via AI Act Accelerator)</p>
<p>NR-3 2-year action plan</p> <p>Opportunity field: Policy handrails</p> <p>Role: Input</p>	<p>NR-16 Municipal blueprints</p> <p>Opportunity field: Local implementation</p> <p>Role: Lead</p>
<p>NR-5 Procurement unblocking</p> <p>Opportunity field: Trustworthy engineering</p> <p>Role: Support (via AI Act Accelerator)</p>	<p>NR-19 Innovation ecosystem adaptation</p> <p>Opportunity field: Trustworthy engineering + local implementation</p> <p>Role: Contribute</p>
<p>NR-8 Startup support</p> <p>Opportunity field: Local implementation</p> <p>Role: Support</p>	<p>NR-20 SME onboarding support</p> <p>Opportunity field: Local implementation</p> <p>Role: Lead</p>
<p>NR-13 Skills exposure study</p> <p>Opportunity field: Skills & workforce</p> <p>Role: Lead</p>	

Conditional measures

<p>PC-1 Early warning system</p> <p>Opportunity field: Policy handrails</p> <p>Role: Support</p>

Chapter 5: Update system: Research agent + expert panel + versioning



5.1 Purpose of this chapter

The AI Projections whitepaper is designed to be a living document, updated continuously as the AI landscape evolves. This chapter specifies the mechanisms through which we keep the document current, credible, and decision-relevant.

Core principle: The value of projection planning depends on timely updates when evidence shifts. A document that becomes outdated loses decision-relevance precisely when it matters most.

Invitation to participate:

We welcome contributions across all aspects of this document:

- Projections (Chapter 1): Perspectives on projection framing, baseline conditions, and projection descriptions
- Drivers & Probabilities (Chapter 2): Insights on main drivers, important trends, and probability assessments
- Impacts (Chapter 3): Observations on implications across impact fields; additional evidence or perspectives
- Measures (Chapter 4): Additions, ideas, and comments on proposed measures
- Collaboration (Chapter 5): If you would like to collaborate with us on our activities or if you want us to highlight your activities

Contact: [ai-projections@appliedai-institute.de]

5.2 Update triggers and cadence

5.2.1 Scheduled updates

Update type	Cadence	Scope	Output
Annual edition	Every 12 months	Full document revision; all chapters reviewed and updated	New version number (e.g., 2.0)
Quarterly review	Every 3 months	Benchmark updates; driver reassessment; probability check	Version increment (e.g., 1.4)
Monthly monitoring	Every month	Scan for all triggers including breakthrough triggers; flag emerging developments; update if warranted	Internal report; public update if material changes detected

5.3 Research agent system

5.3.1 Continuous monitoring infrastructure

We maintain systematic monitoring across key information sources:

Primary sources:

Source category	Examples	Monitoring frequency
Benchmark leaderboards	SWE-bench, GAIA, OSWorld, ARC-AGI-2, HumanEval, LiveCodeBench, METR, Vending-Bench 2	Weekly
Major lab and company announcements	OpenAI, Anthropic, DeepMind, Meta AI, Mistral, xAI, AgiBot, UBTech, Unitree, Figure AI, Tesla Robotics	Daily
Academic publications	arXiv (cs.AI, cs.LG, cs.CL), NeurIPS, ICML, ICLR	Weekly
Industry reports	McKinsey, BCG, Gartner, Stanford AI Index	As published
Regulatory developments	EU AI Office, national AI authorities, standards bodies	Weekly

Market data	AI investment tracking, startup announcements, M&A activity	Quarterly
Security incidents	CISA, ENISA, major incident reports	Daily

5.3.2 Signal processing

Raw monitoring data is processed through structured analysis:

1. Relevance filtering: Does this development affect one of our four drivers or baseline conditions?
2. Magnitude assessment: Is this incremental progress or a potential step-change?
3. Verification: Can the claim be corroborated by multiple independent sources?
4. Implication mapping: Which projections, impacts, or measures are affected?
5. Update recommendation: Does this warrant document revision?

5.3.3 AI-Assisted research

We use AI systems to support (not replace) human analysis:

- Literature scanning: Automated identification of relevant new publications
- Benchmark tracking: Automated extraction of performance updates from leaderboards
- Trend detection: Pattern identification across multiple information streams
- Draft preparation: Initial drafting of update sections for human review

Human oversight requirements:

- All AI-generated content is reviewed by human analysts before inclusion
- Probability estimates and strategic judgments are made by humans
- Expert panel validates significant updates before publication
- Sources are independently verified by human researchers

5.4 Expert panel

To be included in the future.

We plan to establish an expert panel representing diverse perspectives essential for comprehensive assessment, including domains such as AI research, AI engineering, AI safety, industry/enterprise, poli-

cy/governance, economics/labour, security, and robotics/embodied AI.

5.5 Versioning and change management

5.5.1 Version numbering

We use semantic versioning adapted for strategic documents:

Version Format	Meaning	Example
X.0	Major annual edition	2.0 (2027 edition)
X.Y	Quarterly or significant update	1.4 (Q2 2026 update)
X.Y.Z	Minor correction or clarification	1.3.1 (typo fix)

5.5.2 Change documentation

Every update includes:

- Version number: Clear identification of document version
- Publication date: When the update was released
- Change summary: High-level description of what changed
- Detailed changelog: Section-by-section list of modifications
- Rationale: Why changes were made (evidence basis)

5.5.3 Archive and access

- All previous versions are archived and accessible
- Current version is clearly marked as authoritative
- Comparison tools available to track changes between versions
- Subscription notification for stakeholders when updates published

5.6 Quality assurance

5.6.1 Accuracy standards

Standard	Implementation
Source verification	All factual claims traced to primary sources
Benchmark accuracy	Performance figures verified against official leaderboards
Expert attribution	Quotes and positions verified with original sources
Recency	Evidence dated; outdated information flagged for update
Uncertainty acknowledgment	Confidence levels stated; speculation clearly marked

5.7 Contact and contribution

For updates, corrections, and collaboration inquiries:

- Email: [ai-projections@appliedai-institute.de]
- Web: [appliedai-institute.de/ai-projections/]

We actively welcome:

- Feedback on projections, drivers, impacts, and measures
- New evidence and research findings
- Collaboration proposals on activities
- Requests to highlight relevant activities and initiatives

Appendix A: Glossary of key terms

Term	Definition
Agentic AI	AI systems that plan and execute multi-step tasks using tools, interfaces, and memory, operating under varying degrees of human supervision

AI scare trade	Market sell-offs targeting sectors vulnerable to AI automation, triggered by new AI capability demonstrations
ARC-AGI	Abstraction and Reasoning Corpus benchmark testing novel problem-solving ability
Automation frontier	The boundary of tasks that can be reliably performed by AI systems with acceptable quality, cost, and oversight
Causal AI	AI approaches focused on understanding cause and effect rather than pattern correlation
Claude code	Anthropic's AI coding tool operating as an autonomous „AI Engineer“
Few-person unicorn	A billion-dollar company operated by a very small team (potentially single digits) enabled by AI automation
Frontier models	State-of-the-art general-purpose AI models defining current capability ceiling
GAIA	General AI Assistants benchmark testing multi-step research and reasoning
Nested learning	DeepMind architecture treating models as interconnected multi-level optimisation problems
No-regret measure	Action valuable regardless of which projection materializes
OpenClaw	Framework for system-integrated AI agents with local execution capabilities
O-ring automation	Pattern where AI generates candidates at scale but human filtering remains the bottleneck
OSWorld	Benchmark testing AI agents in real operating system environments
Revenue per employee	Key efficiency metric; AI-native companies achieving \$2-5M vs. \$400-500k traditional
Projection-conditional measure	Action that becomes critical primarily in specific projections
Sim-to-real transfer	Process of translating AI capabilities trained in simulation to physical-world deployment

SWE-bench	Software Engineering Benchmark using real GitHub issues to test coding capability
VLA models	Vision-Language-Action models enabling robots to perceive, understand, and act

Appendix B: Key benchmarks quick reference

Benchmark	Category	Current SOTA	Human baseline	Status
SWE-bench bash	Software engineering	76.8%	~70-80%	Human parity
SWE-bench Verified	Software engineering	79.2%	100%	Near-human
GAIA	Research/reasoning	74.55%	92%	Closing gap
OSWorld	OS interaction	73%	72%	Near-human
WebArena	Web tasks	57.1%	~85%	Significant gap
ARC-AGI	Novel reasoning	75.7%	~85%	Closing gap
HumanEval	Code generation	96.2%	~95%	Saturated
LiveCodeBench	Fresh code problems	91.7%	N/A	High performance

Appendix C: Projection quick reference

Dimension	P1: Plateau (5%)	P2: Continued pace (50%)	P3: Accelerated (45%)
Capability speed	Marginal	Steady	Compounding
Agentic viability	Limited	Growing	Substantial autonomy
R&D automation	Assists	Accelerates	Feedback loops
Robotics	Gap persists	Viable deployment	Economically transformative
Europe's risk	Complacency	Competitiveness gap	Loss of agency
Urgency	Moderate	High	Very high

References

- [1] <https://www.axios.com/2026/02/11/hyperscaler-spending-meta-microsoft-amazon-google>
- [2] <https://www.deloitte.com/us/en/what-we-do/capabilities/applied-artificial-intelligence/blogs/pulse-check-series-latest-ai-developments/ai-adoption-challenges-ai-trends.html>
- [3] <https://doi.org/10.59350/redasadki.21109>
- [4] <https://iceberg.mit.edu/report.pdf>
- [5] <https://gdpval.pro/>
- [6] <https://www.dwarkesh.com/p/dario-amodei-2>
- [7] <https://blog.samaltman.com/reflections>
- [8] <https://80000hours.org/ai/guide/when-will-agi-arrive/>
- [9] <https://towardsai.net/p/machine-learning/in-davos-demis-hassabis-bets-50-50-agi-arrives-in-five-years>
- [10] <https://medium.com/activated-thinker/the-100-hour-workweeks-that-saved-google-demis-hassabis-on-the-2030-agi-timeline-and-the-209e0edab2cb>
- [11] The Atlantic Interview: “It’s Time to Think About Generally Intelligent AI”
<https://www.theatlantic.com/sponsored/google/Its-time-to-think-about-generally-intelligent-ai/3974/>
- [12] <https://www.dwarkesh.com/p/ilya-sutskever-2>
- [13] <https://forum.effectivealtruism.org/posts/SYtwChBTs6xkocBSP/when-do-experts-think-human-level-ai-will-be-created>
- [14] <https://ai.gopubby.com/why-yann-lecun-bet-3-5-billion-on-world-models-over-llms-e26eefd-d900e>
- [15] World Economic Forum – Global Risks Report 2026
https://reports.weforum.org/docs/WEF_Global_Risks_Report_2026.pdf
- [16] <https://arcprize.org/blog/oai-o3-pub-breakthrough>
- [17] <https://sakana.ai/ai-scientist/>
- [18] <https://theinnovator.news/interview-of-the-week-henry-markram-on-brain-based-agi/>
- [19] <https://forum.effectivealtruism.org/posts/YvEjpAKkJNErkiFTN/google-deepmind-ceo-demis-hassabis-on-what-s-still-needed>
- [20] <https://research.google/blog/introducing-nested-learning-a-new-ml-paradigm-for-continual-learning/>
- [21] <https://neurips.cc/virtual/2025/poster/116123>

[22] <https://news.microsoft.com/de-ch/2025/03/18/inait-announces-collaboration-with-microsoft-to-deploy-novel-ai-based-on-digital-brains-across-industries/>

[23] Causal AI: How cause and effect will change artificial intelligence. <https://www.spglobal.com/en/research-insights/special-reports/causal-ai-how-cause-and-effect-will-change-artificial-intelligence>

[24] SWE-bench Official Leaderboards. <https://www.swebench.com/>

[25] Anthropic News - Introducing Claude Opus 4.5. <https://www.anthropic.com/news/claude-opus-4-5>

[26] HAL: GAIA Leaderboard (Princeton University). <https://hal.cs.princeton.edu/gaia>

[27] Reddit r/CompetitiveAI (Benchmark-Übersicht). <https://webarena.dev/>

[28] OSWorld Benchmark Leaderboard (2026). <https://os-world.github.io/>

[29] METR Blog - "Measuring AI Ability to Complete Long Tasks". <https://metr.org/blog/2025-03-19-measuring-ai-ability-to-complete-long-tasks/>

[30] Reddit r/accelerate. https://www.reddit.com/r/accelerate/comments/1qweygd/gpt52_high_released_on_11_december_2025_is_the/

[31] Andon Labs - Vending-Bench 2. <https://andonlabs.com/evals/vending-bench-2>

[32] MindStudio Blog. <https://www.mindstudio.ai/blog/what-is-openclaw-ai-agent/>

[33] Acronis Security Analysis. <https://www.acronis.com/en/tru/posts/openclaw-agentic-ai-in-the-wild-architecture-adoption-and-emerging-security-risks/>

[34] arXiv - "Evaluating Large Language Models on Self-invoking Code Generation". <https://arxiv.org/html/2412.21199v1>

[35] PricePerToken - LiveCodeBench Leaderboard 2026. <https://pricepertoken.com/leaderboards/benchmark/livecodebench>

[36] The Pragmatic Engineer. <https://newsletter.pragmaticengineer.com/p/how-claude-code-is-built>

[37] Reddit (Dario Amodei Quote). <https://www.reddit.com/r/ClaudeAI/comments/1nkzagw/>

[38] Anthropic Engineering. <https://www.anthropic.com/engineering/code-execution-with-mcp>

[39] GitHub: claude-scientific-skills. <https://github.com/K-Dense-AI/claude-scientific-skills/blob/main/docs/scientific-skills.md>

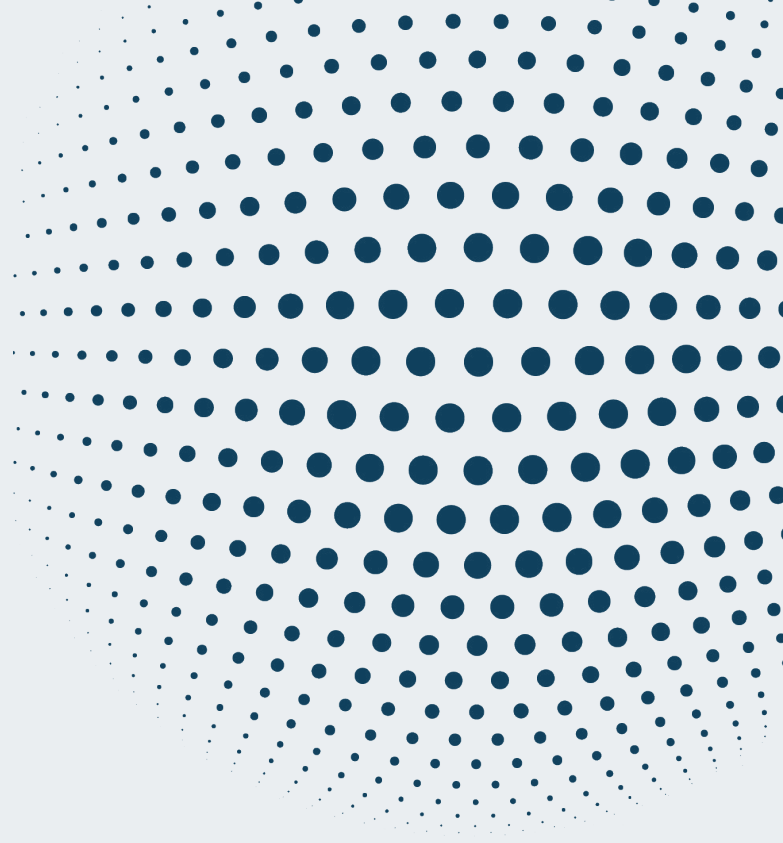
[40] Reddit r/ClaudeAI. https://www.reddit.com/r/ClaudeAI/comments/1l11fo2/how_i_built_a_multia-gent_orchestration_system/

[41] Wall Street Journal (2026). <https://www.wsj.com/tech/china-is-going-all-in-to-beat-the-u-s-on-humanoid-robots-b9c434d2>

- [42] PYMNTS (2026).
<https://www.pymnts.com/artificial-intelligence-2/2026/china-looks-to-become-leader-in-humanoid-robots-sector/>
- [43] China Briefing (2025).
<https://www.china-briefing.com/news/chinese-humanoid-robot-market-opportunities/>
- [44] AgiBot Official Press Release.
<https://www.agibot.com/article/231/detail/33.html>
- [45] The Robot Report.
<https://www.therobotreport.com/agibot-makes-u-s-debut-with-more-than-5100-robots-shipped/>
- [46] UBTECH.
<https://www.prnewswire.com/news-releases/ubtech-humanoid-robot-walker-P2-begins-mass-production-and-delivery-with-orders-exceeding-800-million-yuan-302616924.html>
- [47] Tech in Asia.
<https://www.techinasia.com/news/chinas-ubtech-boost-humanoid-robot-production-10-2026>
- [48] South China Morning Post.
<https://www.scmp.com/tech/tech-trends/article/3340446/chinas-unitree-ships-more-5500-humanoid-robots-2025-surpassing-us-peers>
- [49] Xpeng Official Press Release (AI Day 2025).
<https://www.xpeng.com/news/019a56f54fe99a2a0a8d8a0282e402b7>
- [50] Humanoids Daily.
<https://www.humanoidsdaily.com/news/xpeng-sets-2026-target-for-mass-produced-iron-robot-eyes-1-million-units-by-2030>
- [51] Figure AI Official Press Release.
<https://www.figure.ai/news/production-at-bmw>
- [52] South China Morning Post.
<https://www.scmp.com/tech/big-tech/article/3336939/chinas-catl-marks-trailblazing-deployment-humanoid-robots-scale-factory-floor>
- [53] Morgan Stanley.
https://advisor.morganstanley.com/john.howard/documents/field/j/jo/john-howard/The_Humanoid_100_-_Mapping_the_Humanoid_Robot_Value_Chain.pdf
- [54] Goldman Sachs.
<https://www.uc.edu/content/dam/refresh/cont-ed-62/olli/P25-handouts/exploring%20humanoid%20robots%205.pdf>
- [55] Reuters.
<https://www.reuters.com/world/china/chinas-ai-powered-humanoid-robots-aim-transform-manufacturing-2025-05-13/>
- [56] There's A Robot For That.
<https://www.theresarobotforthat.com/blog/humanoid-robot-cost-roi-breakdown/>
- [57] Robozaps.
<https://blog.robzaps.com/b/roi-of-humanoid-robots>
- [58] Omdia/PR Newswire.
<https://www.prnewswire.com/news-releases/omdia-ranks-agibot-no1-worldwide-in-humanoid-robot-shipments-in-2025-302656788.html>

- [59] AI Business.
<https://aibusiness.com/robotics/chinese-company-completes-first-mass-humanoid-robot-delivery>
- [60] Gasgoo.
<https://www.techinasia.com/news/chinas-unitree-ships-over-5500-humanoid-robots-in-2025>
- [61] CNBC.
<https://www.cnbc.com/2025/11/05/china-xpeng-to-launch-robotaxis-humanoid-robots-wi-th-own-ai-chips.html>
- [62] Aparobot.
<https://www.aparobot.com/robots/fourier-gr-1>
- [63] PR Newswire.
<https://www.prnewswire.com/news-releases/kepler-robotics-unveils-k2-bumblebee-at-iros-2025-building-a-global-developer-ecosystem-through-an-open-robotics-platform-302595121.html>
- [64] Assembly Magazine.
<https://www.assemblymag.com/articles/99853-tesla-to-end-model-x-s-production-convert-fre-mont-space-for-optimus-robots>
- [65] Robozaps.
<https://blog.robozaps.com/b/challenges-in-humanoid-robotics>
- [66] Automate.org.
<https://www.automate.org/robotics/industry-insights/teleoperation-is-a-powerful-tool-for-humanoid-robots-but-transparency-is-key>
- [67] Hill Dickinson.
<https://www.hilldickinson.com/our-view/articles/humanoid-robots-and-the-law-preparing-for-a-new-era-of-risk/>
- [68] AgiBot.
<https://agibot-world.com/blog/go1>
- [69] Gasgoo.
<https://autonews.gasgoo.com/articles/news/unitree-technology-and-tencent-robotics-x-lab-form-a-strategic-partnership-2008804138454663168>
- [70] Interesting Engineering.
<https://interestingengineering.com/ai-robotics/alibaba-rynnbrain-humanoid-robot-ai>
- [71] <https://epoch.ai/benchmarks/swe-bench-bash-only>
- [72] <https://www.sonarsource.com/company/press-releases/sonar-claims-top-spot-on-swe-bench-leaderboard/>
- [73] <https://huggingface.co/spaces/gaia-benchmark/leaderboard>
- [74] <https://www.appypieautomate.ai/blog/best-ai-autonomous-agents>
- [75] <https://arxiv.org/abs/2311.12022>
- [76] <https://www.sydney.edu.au/news-opinion/news/2026/01/30/ai-is-failing-humanitys-last-exam.html>
- [77] <https://github.com/openai/human-eval>
- [78] <https://github.com/google-research/google-research/blob/master/mbpp/README.md>

- [79] <https://livecodebench.github.io/leaderboard.html>
- [80] SemiAnalysis.
<https://newsletter.semianalysis.com/p/claude-code-is-the-inflection-point>
- [81] Recombee.
<https://www.recombee.com/blog/ai-news-and-outlook-for-2026>
- [82] The Cube Research.
<https://thecuberresearch.com/why-causal-ai-decision-intelligence-2026/>
- [83] LinkedIn/Anthropic.
https://www.linkedin.com/posts/rajdeep-mondal_anthropic-claude-2026-agentic-coding-trends-activity-7426543519571312641-B3SL
- [84] RAI Institute.
<https://rai-inst.com/resources/blog/rai-institute-2025-a-year-of-innovation-for-robotics-and-ai/>
- [85] McKinsey Global Institute (MGI)
<https://www.mckinsey.com/featured-insights/digital-disruption/harnessing-automation-for-a-future-that-works>
- [86] Indeed Hiring Lab
<https://www.hiringlab.org/>
- [87] U.S. Bureau of Labour Statistics (BLS)
<https://www.bls.gov/ooh/computer-and-information-technology/software-developers.htm>
- [88] ADP Research Institute
<https://www.adpresearch.com/>
- [89] OECD
<https://www.oecd.org/digital/artificial-intelligence/>
- [90] Brynjolfsson, Li, Raymond (NBER, 2023)
<https://www.nber.org/papers/w31161>
- [91] SWE-bench official site
<https://www.swebench.com/> ; <https://arxiv.org/abs/2310.06770>
- [92] IMF (Fiscal Affairs Department)
<https://www.imf.org/en/Topics/Public-Finance>
- [93] Daniel Susskind – A World Without Work (2015)
<https://global.oup.com/academic/product/a-world-without-work-9780198825221>
- [94] William D. Nordhaus – The Economics of Artificial Intelligence (NBER book, 2021)
<https://www.nber.org/books-and-chapters/economics-artificial-intelligence-agenda>
- [95] World Bank – GovTech
<https://www.worldbank.org/en/topic/governance/brief/govtech-putting-people-first>
- [96] Eurostat
<https://ec.europa.eu/eurostat/>



Impressum

© appliedAI Institute for Europe - April 2026

Herausgeber und Kontakt

appliedAI Institute for Europe gGmbH
Freddie-Mercury-Straße 5
D-80797 München
info@appliedai-institute.de
www.appliedai-institute.de

Geschäftsführung

Dr. Andreas Liebl
Dr. Frauke Goll

Die appliedAI Institute for Europe gGmbH wird unterstützt durch die IPAI Foundation gGmbH und ist eine Tochterfirma der appliedAI Initiative GmbH.

Lizenz

Der Text dieser Publikation unterliegt einer Creative Commons Namensnennung 4.0 International (CC BY 4.0) Lizenz. www.creativecommons.org/licenses/by/4.0/legal-code.de Davon ausgenommen sind die verwendeten Logos, diese sind urheberrechtlich geschützt.