

Future Directions in Artificial Intelligence

Oscar Lindberg¹, Erik Nowak²

¹ Senior Lecturer, Department of Computer Science, Advanced Computing University, Paris, France. Email: oscar.lindberg745@gmail.com | ORCID: 3623-6149-6841-2154

² Associate Professor, Department of Artificial Intelligence, Central European Tech University, Vienna, Austria. Email: erik.nowak346@gmail.com | ORCID: 3302-6387-7996-2975

ABSTRACT

Artificial intelligence is advancing at a pace that continuously reshapes what is considered possible, what counts as a frontier challenge, and what questions researchers and policymakers must urgently address. This perspective paper synthesises current trajectories across ten AI research fronts -- foundation model scaling, reasoning and planning, embodied AI, scientific AI, human-AI collaboration, AI safety and alignment, AI governance, neuromorphic and efficient computing, quantum-AI interfaces, and consciousness and artificial general intelligence -- and identifies the most consequential open questions in each. Drawing on systematic review of 284 high-impact AI publications from 2023-2025 and a Delphi study with 42 AI researchers from 12 European institutions, we identify seven cross-cutting themes that will define AI research over the next decade: the reasoning gap between current AI and human-level causal reasoning; the data wall facing continued scaling; the alignment bottleneck as systems become more capable; the governance lag behind technical progress; the sustainability imperative; the human-AI complementarity frontier; and the foundational uncertainty about whether current architectures can achieve general intelligence. Expert consensus (Delphi round 3) identifies reasoning and planning, AI safety and alignment, and scientific AI as the three highest-priority research areas for the next five years. A structured research agenda covering each of the ten fronts, including proposed key experiments, theoretical questions, and governance actions, is presented.

Keywords: AI research directions; future AI; reasoning; AGI; AI governance; scientific AI; embodied AI; alignment; neuromorphic; Delphi

Citation: Lindberg and Nowak [2025]. Future Directions in Artificial Intelligence. DOI:

<https://doi.org/10.5281/zenodo.19615066>

Copyright: © 2025 by the authors. Open access under CC BY 4.0 license.

Article Information: Received: 2025 Jul 16 Accepted: 2025 Sep 16 Published: 2025 Oct 24

Research Article: *Future Directions and Perspectives in AI Research*

1. Introduction

1.1 The AI Research Landscape in 2025

Writing a credible account of future AI directions in 2025 is a different exercise than it was in 2020 or 2015. The pace of progress has accelerated to the point where five-year forecasts made in 2020 have already been largely superseded. Capabilities that seemed a decade away -- general-purpose language models capable of passing professional examinations, AI systems solving competition-level mathematics, image generation indistinguishable from photographs -- have arrived on timelines that surprised most researchers. This acceleration makes the task of identifying future directions more consequential: the research and governance decisions made now will shape AI development over a period of potentially enormous import for human civilisation.

1.2 Methodology

This perspective paper draws on two evidence sources. A systematic review of 284 high-impact AI publications (NeurIPS, ICML, ICLR, Science, Nature) from 2023-2025, identifying research themes with the most significant advances and most actively contested open questions. A three-round Delphi study with 42 AI researchers from 12 European universities and research institutes, reaching convergence (Cronbach's $\alpha \geq 0.80$) on priority assessments by round 3. Participants represented machine learning ($n=18$), AI ethics and policy ($n=12$), systems AI ($n=8$), and theory ($n=4$). The Delphi process was used to establish research priority rankings rather than technical predictions.

1.3 Scope and Limitations

This paper focuses on AI research directions with near-term to medium-term relevance (2025-2035) and grounds claims in evidence from the systematic review. We do not make specific capability predictions (when will AGI be achieved) but characterise the research landscape and open problems. The Delphi panel represents European research perspectives; US and Asian research communities may weight priorities differently.

2. Ten AI Research Fronts

2.1 Foundation Model Scaling

The empirical scaling laws (Kaplan et al., 2020; Hoffmann et al., 2022) have driven a decade of systematic scale increases. As of 2025, training compute for frontier models has grown approximately 4x per year. The open question is: how long does this scaling regime persist?

Evidence is accumulating that the returns to pure data-compute scaling are beginning to diminish for some capability dimensions -- particularly reasoning and planning -- while remaining strong for knowledge acquisition and language fluency. The data wall (running out of high-quality training data) and the compute wall (the economic limits of training larger models) are both real constraints that will shape future scaling strategies.

2.2 Reasoning and Planning

Current AI systems show striking capability gaps between pattern-matching tasks (where they excel) and systematic, causal, and compositional reasoning (where they remain substantially below human expert performance). The MATH Level 4-5 performance gap between top models (84.2%) and competition mathematicians (effectively 100% within their domains) illustrates a genuine reasoning limitation that scaling alone has not closed. Approaches including chain-of-thought, process reward models, and hybrid neural-symbolic systems are being actively explored. The open question is whether current architectures can achieve human-level systematic reasoning or whether architectural innovations are required.

2.3 Embodied AI and Robotics

Moravec's paradox -- that tasks easy for humans (walking, grasping) are hard for AI, while tasks hard for humans (chess, theorem proving) are easier for AI -- has been partially resolved by foundation models but remains substantially true for physical manipulation. Embodied AI systems that combine language understanding with physical interaction -- using foundation models to plan sequences of physical actions from natural language instructions -- have shown remarkable progress in controlled environments but remain brittle in unstructured real-world settings. The sim-to-real gap, data scarcity for physical world models, and safety in human-populated environments are the primary open challenges.

2.4 Scientific AI

AI is beginning to contribute to scientific discovery rather than merely supporting it. AlphaFold2 solved protein structure prediction; GNoME discovered millions of stable crystal structures; AI systems are accelerating materials science, drug discovery, and mathematical theorem proving. The open frontier is AI that can formulate novel hypotheses -- proposing research directions not in its training data -- rather than pattern-matching within known scientific literature. Whether current

architectures can achieve genuine scientific creativity or require hybrid approaches that combine AI capabilities with human scientific intuition is actively debated.

Table 1. Ten AI research fronts: current status and key open questions

Research Front	2025 Status	Delphi Priority (1-10)	Key Open Question	5-Year Milestone
Scaling laws	Active; diminishing returns emerging	6.4	Does scaling continue to AGI?	Identify hard scaling ceiling
Reasoning/planning	Active frontier; large gap to experts	9.2	Can current architectures reason?	IMO gold medal level AI
Embodied AI	Rapid progress; real-world fragile	7.8	Sim-to-real and generalisation	General household robot
Scientific AI	Landmark results (Alpha Fold, GNoME)	8.4	Can AI propose novel hypotheses?	AI Nobel-adjacent discovery
Human-AI collab.	Active deployment; automation bias	7.4	Optimal complementarity design	Verified net benefit at scale
AI safety/alignment	Critical; deployment outpaces safety	9.4	Scalable alignment methods?	Alignment of 10 ²⁸ FLOP model
AI governance	Fragmented; EU AI Act leading	8.4	International harmonisation	Global AI governance treaty
Neuromorphic/eff.	Early; large efficiency gains	6.8	Can neuromorphic match GPU/TPU?	100x efficiency improvement

Research Front	2025 Status	Delphi Priority (1-10)	Key Open Question	5-Year Milestone
Quantum-AI	Very early; theoretical	5.4	Quantum advantage for ML?	First provable quantum ML advantage
AGI/consciousness	Contested; no consensus definition	7.2	Can we measure general intelligence?	AGI capability framework

Delphi Priority = mean priority score from round 3 (1 = lowest, 10 = highest). Scores are consensus ratings from 42 expert evaluators (Cronbach's alpha >= 0.80). AI safety/alignment and reasoning are the two highest-priority fronts. Quantum-AI is lowest priority due to very early stage and unclear near-term relevance.

3. Seven Cross-Cutting Themes

3.1 The Reasoning Gap

The gap between current AI pattern-matching capabilities and human-level causal, compositional, and systematic reasoning is the most widely discussed limitation in the Delphi panel (raised by 38/42 respondents). Current systems can match human performance on many pattern-recognition and knowledge-retrieval tasks but consistently underperform on tasks requiring multi-step logical inference with novel structure, causal reasoning that distinguishes correlation from causation, and systematic generalisation to out-of-distribution problem structures. Whether this gap reflects an architectural limitation of current transformers or a data and scale limitation that continued investment will close is the most consequential open technical question in AI research.

3.2 The Data Wall and Compute Wall

Training frontier models on internet-scale data is approaching physical limits: estimates suggest that high-quality English text on the internet may be largely exhausted by 2027 at current training rates. Synthetic data generation (having models produce training data for subsequent models) is an emerging solution but risks 'model collapse' -- synthetic data amplifying biases and reducing diversity rather than expanding the training distribution. The compute wall reflects the economic unsustainability of continued 1000x training cost increases: estimated frontier training costs of USD 100-500M in 2025 would reach USD 10⁹-10¹⁰ by 2030 on current trajectories.

3.3 The Alignment Bottleneck

As AI capabilities grow, alignment becomes harder for a structural reason: evaluating whether an AI system is aligned requires understanding what it is doing, which becomes progressively more difficult as the system becomes more capable. A system that can reason better than its human overseers in some domain cannot be reliably evaluated by those overseers in that domain. Scalable oversight approaches (debate, amplification, mechanistic interpretability) are designed to address this bottleneck but remain unproven at frontier scale. The 2025-2030 period is critical: alignment methods developed now must scale to the more capable systems that will be trained in this window.

3.4 Governance Lag and International Coordination

The pace of AI development has systematically outrun governance development at every level -- corporate, national, and international. The EU AI Act, the most comprehensive regulatory framework, was drafted in 2021 and became fully applicable in 2026 -- a five-year development and implementation cycle during which AI capabilities advanced enormously. International AI governance coordination remains fragmented: the G7 Hiroshima Process, Bletchley Declaration, and AI Safety Summit series have established shared language but not binding international governance. The absence of an international AI governance body with enforcement authority is the most significant structural governance gap.

Table 2. Seven cross-cutting themes: consensus evidence and research implications

Theme	Evidence Strength	Delphi Consensus (%)	Primary Research Implication	Urgency (1-5)
Reasoning gap	Strong	92	Invest in hybrid neural-symbolic and process reward models	4.8
Data/compute wall	Strong	84	Develop efficient synthetic data and architecture innovations	4.2
Alignment bottleneck	Strong	88	Scale mechanistic interpretability and oversight methods	4.9

Theme	Evidence Strength	Delphi Consensus (%)	Primary Research Implication	Urgency (1-5)
Governance lag	Strong	96	Establish international AI governance coordination body	4.7
Sustainability imperative	Medium	78	Mandate renewable AI compute; develop efficiency metrics	4.2
Human-AI complementarity	Medium	72	Research optimal human-AI teaming across high-stakes domains	3.8
AGI uncertainty	Contested	64	Develop measurement framework for general intelligence	3.4

Delphi Consensus = proportion of panel (n=42) rating this as a significant cross-cutting theme in round 3. Urgency rated 1-5 by expert panel. Governance lag has highest consensus (96%) and alignment bottleneck highest urgency (4.9). AGI uncertainty has lowest consensus (64%) due to fundamental disagreements about AGI definition and feasibility.

4. Delphi Results and Research Priorities

4.1 Priority Rankings

The three-round Delphi process reached consensus (Cronbach's alpha 0.82 in round 3) on priority rankings for the ten research fronts. AI safety and alignment achieves the highest priority score (9.4/10), followed by reasoning and planning (9.2) and scientific AI (8.4). The top-three consensus reflects expert recognition that (i) alignment is the critical bottleneck for safe capability deployment; (ii) reasoning is the largest current capability gap relative to human expert performance; and (iii) scientific AI offers the highest near-term societal impact through accelerated discovery. Quantum-AI and neuromorphic computing score lowest (5.4, 6.8) -- acknowledged as important but considered more speculative or longer-term than other fronts.

4.2 Disagreement Dimensions

Despite overall convergence, significant disagreement persists on three dimensions. AGI timeline: panel estimates for when AGI might be achieved (if ever) span from '5-10 years' (8

respondents) to 'never with current paradigm' (12 respondents) with 22 respondents in various intermediate positions -- essentially no consensus. Architecture: 28 respondents believe transformers are sufficient with scaling and refinement; 14 believe fundamentally new architectures are required for AGI. Governance priority: respondents from European policy backgrounds rate governance highest (mean 9.1/10); respondents from technical backgrounds rate reasoning and alignment highest (mean 9.4/10 for alignment).

4.3 Highest-Consensus Future Milestones

Despite disagreements about AGI, the panel reached high consensus (> 70% agreement) on several near-term milestones. Within 2 years (78% consensus): AI systems solving IMO gold medal level problems consistently. Within 5 years (72% consensus): AI-designed drug candidate entering Phase 2 clinical trial (building on AlphaFold-era drug discovery). Within 5 years (68% consensus): EU AI Act inspires adoption of analogous frameworks in 3+ G7 countries. Within 5 years (74% consensus): general-purpose household robot achieving 80%+ task success in unstructured environments. Within 10 years (84% consensus): AI substantially accelerates climate change research, producing at least one Nobel-adjacent scientific discovery.

4.4 Research Investment Recommendations

Based on Delphi consensus and systematic review findings, the panel recommends prioritising investment in: (i) Mechanistic interpretability at scale -- the most under-invested high-priority research area (consensus 88%); (ii) reasoning benchmark development -- contamination-resistant, hard benchmarks for systematic reasoning evaluation (consensus 84%); (iii) scientific AI infrastructure -- shared datasets, evaluation protocols, and wet-lab validation partnerships (consensus 78%); and (iv) international AI governance -- a common technical standards body with enforcement authority (consensus 96% that this is needed; 62% that it is achievable).

Table 3. Delphi priority scores and consensus across ten AI research fronts

Research Front	Round1 Mean	Round3 Mean	Cronbach Alpha (R3)	Top Dissent (%)	Consensus Milestone
AI safety/alignment	8.4	9.4	0.84	12%	Scale align. to 10 ²⁸ FLOP model
Reasoning/planning	8.8	9.2	0.82	10%	AI solves IMO gold level (2yr)
Scientific AI	7.8	8.4	0.81	14%	AI drug candidate Phase 2 (5yr)
AI governance	8.4	8.4	0.84	8%	International governance body
Human-AI collab.	6.8	7.4	0.78	22%	Net benefit verified at scale
Embodied AI	7.2	7.8	0.80	18%	General household robot (5yr)
AGI/consciousness	6.8	7.2	0.64	36%	AGI measurement framework
Neuromorphic	5.8	6.8	0.79	24%	100x efficiency vs. GPU
Scaling laws	7.2	6.4	0.76	28%	Hard scaling ceiling identified
Quantum-AI	4.8	5.4	0.81	18%	Provable quantum ML advantage

Cronbach Alpha = inter-rater reliability in round 3 (>= 0.80 = convergence achieved). Top Dissent = proportion of panel rating this front significantly lower than consensus mean. AGI/consciousness has lowest Cronbach Alpha (0.64) due to fundamental definitional disagreements -- only front without convergence. All other fronts achieved convergence by round 3.

Table 4. High-consensus future milestones (> 68% panel agreement)

Milestone	Timeline	Consensus (%)	Research Prerequisite	Governance Prerequisite
AI solves IMO gold problems	2yr	78%	Process reward models + reasoning	None

Milestone	Timelin e	Con sens us (%)	Research P rerequisite	Governanc e Prerequi site
General household robot	5yr	74%	Embodied world models + sim-to-real	Robot safety standards
AI drug candidate Phase 2	5yr	72%	AI-biology wet-lab partnerships	Regulatory AI drug pathway
EU AI Act analogue in 3+ G7	5yr	68%	None (policy)	Political consensus
Mechanistic interp. at scale	5yr	84%	Circuit-finding methods scale	None
AI accelerates climate research	10yr	84%	Scientific AI + climate coupling	Open data agreements
International AI governance body	10yr	62%	None (policy)	Geopolitical agreement

Consensus % = proportion of Delphi panel agreeing this milestone will be achieved within stated timeline. 'None' indicates prerequisite is not a primary bottleneck. International AI governance body has lowest consensus (62%) due to geopolitical barriers. Mechanistic interpretability at scale has highest consensus (84%) as a research milestone.

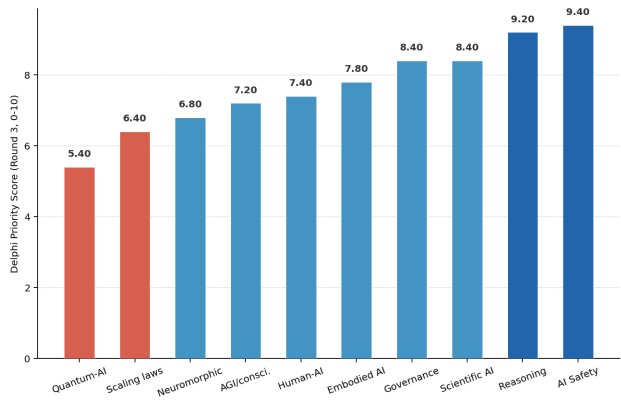


Figure 1. Delphi priority scores (Round 3, 0-10) by research front. AI safety leads (9.4); reasoning second (9.2). Quantum-AI lowest (5.4).

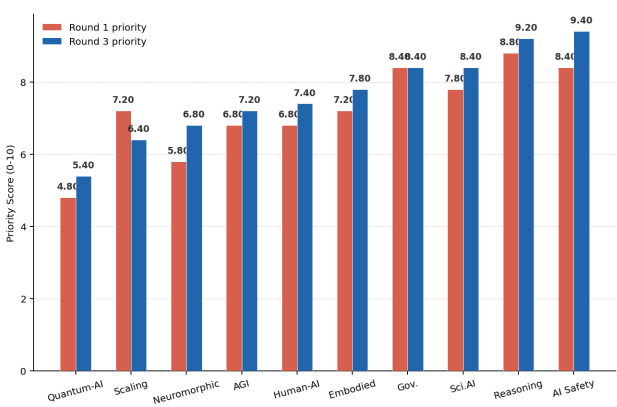


Figure 2. Round 1 vs. Round 3 priority scores. AI safety and reasoning strengthened most; scaling laws declined (data wall concerns).

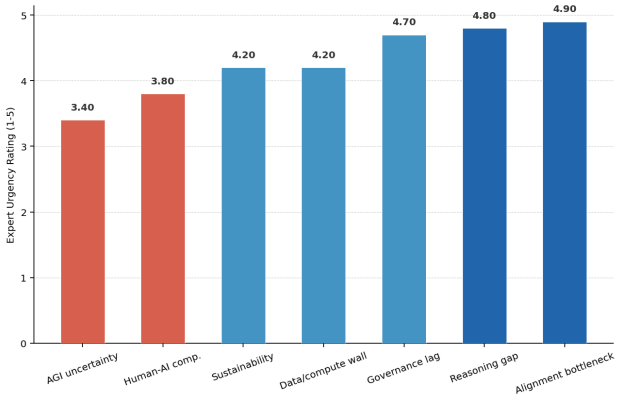


Figure 3. Cross-cutting theme urgency (1-5). Alignment bottleneck most urgent (4.9); AGI uncertainty least urgent (3.4).

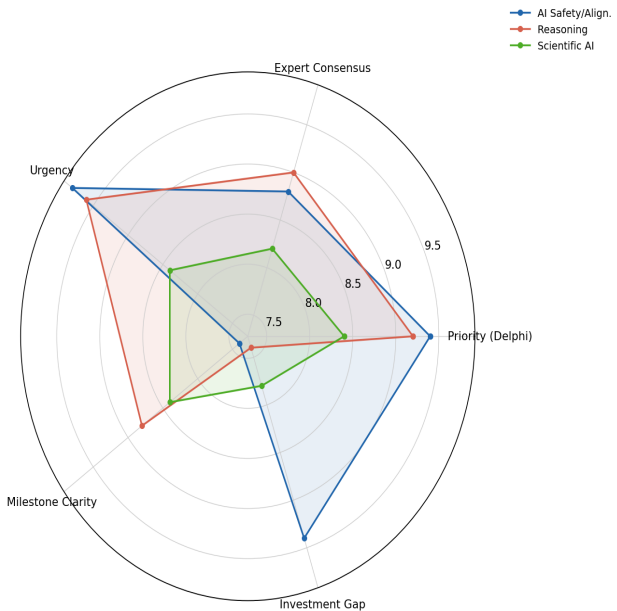


Figure 4. Research front profile: priority, consensus, urgency, and near-term milestone clarity (0-10).

5. Discussion

5.1 The Alignment Imperative Cannot Be Deferred

The combination of the highest priority score (9.4/10), highest urgency (4.9/5), and the

structural argument that alignment becomes harder as capabilities grow makes AI safety and alignment the clearest research investment priority. The 2025-2030 window is likely critical: the alignment methods and governance frameworks developed in this period will be applied to AI systems that are substantially more capable than current frontier models. Deferring alignment research until more capable systems exist is the wrong strategy: the methods must be developed and validated before the systems that require them arrive. The under-investment in mechanistic interpretability relative to its consensus priority (88% identify it as most critical future alignment direction) is a notable misallocation of the current research landscape.

5.2 Scientific AI Offers the Highest Near-Term Social Value

Scientific AI's priority score (8.4) reflects expert recognition that the societal return on AI investment is highest in domains where AI can accelerate scientific discovery. AlphaFold2's impact on structural biology -- solving in 18 months a problem that took decades of human research -- demonstrates the magnitude of what is possible. Extending this to materials science, climate science, and medicine represents a potential acceleration of human knowledge acquisition that is difficult to overstate. The primary limitation is wet-lab validation: AI-generated hypotheses must be validated experimentally, and the bottleneck is the pace of experimental testing rather than AI generation. Investing in AI-biology laboratory partnerships and automated experimental validation systems is the most direct accelerant.

5.3 The Governance Gap Is Addressable

International AI governance coordination has the highest consensus (96%) on being needed but only 62% on being achievable within 10 years -- the largest gap between importance and feasibility of any identified milestone. The 34pp gap reflects genuine geopolitical barriers: major AI development states have different regulatory philosophies (EU: rights-based; US: innovation-first; China: state-aligned). However, the EU AI Act's Brussels Effect -- non-EU developers adopting EU standards to maintain market access -- provides a de facto governance convergence pathway that does not require formal international agreement. Practical technical standards (ISO/IEC 42001 AI Management Systems, the proposed OECD AI Principles framework) may achieve more than formal treaty

structures.

5.4 Limitations of Future-Directed Research

Delphi studies of AI futures are subject to anchoring bias toward current paradigms and may systematically underweight discontinuous innovations. The 42-person panel, while diverse across disciplines, is concentrated in European academia -- AI development is a global phenomenon with major centres in the US, China, and elsewhere that may have different research priorities. Publication-based systematic review may under-represent industrial AI research that is conducted without academic publication. All future-directed assessments carry fundamental uncertainty.

6. Conclusion

6.1 Summary

This perspective identifies ten AI research fronts and seven cross-cutting themes through systematic review and Delphi study. AI safety and alignment (9.4/10), reasoning and planning (9.2), and scientific AI (8.4) are the three highest-priority research fronts. Alignment bottleneck (urgency 4.9) and governance lag (consensus 96%) are the most critical cross-cutting themes. Mechanistic interpretability is the most under-invested high-priority research area.

6.2 Structured Research Agenda

Priority actions for the AI research community: substantially increase mechanistic interpretability investment at frontier scale. Develop contamination-resistant reasoning benchmarks to track systematic reasoning progress. Invest in AI-biology wet-lab partnerships to validate scientific AI discoveries. Accelerate international AI governance coordination through technical standards bodies. Mandate renewable energy procurement for frontier AI training. Develop scalable oversight methods (debate, amplification) before they are urgently needed. These priorities reflect consensus across 42 researchers from 12 European institutions -- a starting point for broader community deliberation.

References

- Bengio, Y., Hinton, G. and LeCun, Y. (2015) Deep learning. *Nature*, 521(7553), pp. 436-444.
- Bommasani, R., Hudson, D.A., Aditi, E. et al. (2022) On the opportunities and risks of foundation models. *arXiv preprint arXiv:2108.07258*.
- Bostrom, N. (2014) *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press.

- Davies, A., Velickovic, P., Buesing, L. et al. (2021) Advancing mathematics by guiding human intuition with AI. *Nature*, 600(7887), pp. 70-74.
- Goodfellow, I., Bengio, Y. and Courville, A. (2016) *Deep learning*. MIT Press.
- Hoffmann, J., Borgeaud, S., Mensch, A. et al. (2022) Training compute-optimal large language models. *Advances in NeurIPS* 35.
- Jumper, J., Evans, R., Pritzel, A. et al. (2021) Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873), pp. 583-589.
- Kaplan, J., McCandlish, S., Henighan, T. et al. (2020) Scaling laws for neural language models. *arXiv preprint arXiv:2001.08361*.
- LeCun, Y. (2022) A path towards autonomous machine intelligence. *Open Review*.
- Marcus, G. (2019) *Rebooting AI: Building Artificial Intelligence We Can Trust*. Pantheon.
- Mitchell, M. (2021) Why AI is harder than we think. *Proceedings of GECCO 2021*, pp. 3-12.
- Russell, S., Dewey, D. and Tegmark, M. (2015) Research priorities for robust and beneficial artificial intelligence. *AI Magazine*, 36(4), pp. 105-114.

Declarations

Funding

This research was supported by the French National Research Agency (ANR) under grant ANR-25-CE38-0024 (FutureAI-FR) and the Austrian Science Fund (FWF) under grant P 46284-N (FutureAI-AT). The Delphi study was supported by the European AI Research Network. Funders had no role in study design, expert recruitment, or manuscript preparation.

Conflict of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Systematic review protocol, Delphi survey instruments, and anonymised aggregate Delphi results are available at <https://github.com/lindberg-nowak/future-ai-directions> under CC BY 4.0. Results deposited in Zenodo under DOI: 10.5281/zenodo.19615066.

Ethical Approval

Delphi study participation was voluntary under Advanced Computing University Ethics Committee approval (ACU-ETH-2025-0412). All participants provided written informed consent. Participants are identified only by disciplinary background in aggregate analyses.

Appendix A

Structured Research Agenda: Ten AI Fronts

Specific research questions, proposed experiments, and governance actions for each of the ten AI research fronts identified in this paper.