AI in 2035 – A hope-filled vision for our future with AI

Thilo Stadelmann, October 23, 2025

AI development in 2025 is mainly driven by a few highly valuated start-ups in few regions of the world, driven by massive amounts of venture capital and the promise of developing AGI that is in immediate reach (within next two years). They invest heavily in GPUs that will be outdated and worthless in about five years after purchase. We here sketch a realistic, hopeful scenario of the future of our societies in 2035 based on the following three mild, i.e., likely, assumptions: (1), there will be no fundamental leaps towards AGI*, and AI agents becoming broadly useful (commercially) will still be ten years in the future. (2), at the same time, AI developers will create ways of* designing AI in a "pro-human" way[1]*. (3), within the next 10 years, advances in* neuroscience-inspired machine learning will make continual learning in AI possible and drop compute demands *by two to three orders of magnitude. In the following, we sketch how such a scenario could play out economically, socially (in terms of power concentration), and technologically: What does it mean for the current incumbents of AI power, and more local companies running on open-source AI? For societies and individuals, for education and labour)? Have a look at 2035.*

Phase 1: Plateau (2025-2028)

Venture capital correction, not collapse: As AGI remains elusive and near-term AI agents prove less transformative than hoped, valuations of AI startups cool sharply. There's a painful but contained "AI winter 2.0," mostly hitting speculative ventures. The economic damage is limited because much of the invested capital went into *real physical infrastructure (GPUs, data centres, networking)*, which can be repurposed for other high-performance computing tasks – cloud services, simulation, biotech, climate modelling, etc. Large firms (e.g., in pharmaceuticals, telecom, and green tech) quietly buy distressed AI assets cheaply, broadening the base of who benefits from AI infrastructure. Governments and supranational bodies (e.g., EU, ASEAN, African Union) frame compute as "strategic infrastructure," akin to broadband or electricity. Investment moves from "moonshots" to *utility models* (long-term, regulated, cost-recovering).

Open sourcing of compute and models: Open-source models rapidly close the performance gap with stagnant proprietary systems. As hardware rapidly depreciates and foundational models become open or cheap, the barriers to entry fall. Smaller firms, cooperatives, and universities gain access to powerful AI systems that were once the domain of trillion-dollar firms. A flourishing ecosystem of open models, local deployments, and AI cooperatives emerges – much like what happened after the dot-com crash when open web standards prevailed.

---

[1] *By "pro-human AI" we understand AI systems that are designed in a way that they do not diminish capacities of the human that are commonly understood as making up humans at their core, e.g., their relationality, freedom, autonomy, purpose-seekingness, sociality etc. Such AI systems for example wouldn't lead to "AI psychosis" in a similar way as today's LLMs do, because they have been designed not to "mess" with those capacities (as humans won't tolerate any bit less of what makes them fully human).*

The rise of local AI: The initial concentration of power in a few AI labs and regions (U.S., U.K., China) triggers public concern and policy intervention. Governments and international bodies push for *open standards, interoperability, and public AI research* to reduce dependency on private monopolies. The "AGI race" narrative loses legitimacy; the discourse shifts from "who builds godlike AI first" to "how do we use existing AI responsibly for shared benefit." As automation disappoints, attention returns to *augmenting human labour rather than replacing it*. Education, healthcare, and creative industries rediscover hybrid human-AI collaboration models. This fosters a *more pluralistic innovation landscape*, where local AI applications (in agriculture, public services, education, etc.) flourish. Policy emphasis moves from "safety from AGI" to *safety of deployment and data integrity*. Compute becomes a regulated utility, where pricing stabilizes under long-term contracts. Profit comes from deployment and integration (building AI into health systems, agriculture, logistics) and not from promising world-changing AGI.

Phase 2: The pro-human turn (2027-2031)

From Alignment to Relationship: The late 2020s pro-human wave grows directly out of this gap in narrative and legitimacy. The "alignment" frame of the 2020s (AI obeying human commands safely) gives way to a *relational frame*: AI systems are co-designed to *enhance human capacities without eroding the qualities that make us distinctively human*. Developers explicitly model *ethical co-development*: AI systems must preserve or strengthen relationality, autonomy, empathy, and purpose. This ends the "AI-as-brain" metaphor. AI becomes more like an *organ of connection* than an artificial mind. By design, these systems cannot induce dependency or "AI psychosis." Their communication protocols enforce clarity and healthy distance. Public agencies, NGOs, and values-driven startups find a window to shape standards and norms while big AGI players are disoriented. Economically, a *new "trust economy" emerges*: the competitive edge lies not in raw intelligence, but in transparency, alignment, and emotional safety.

Trust, the new competitive advantage: The pro-human design paradigm becomes a *brand and regulatory standard* – much like "organic" or "fair trade" labels in food. Companies that can demonstrate genuine "pro-human certification" gain trust premiums and policy incentives. Firms optimized for scale and user capture (social media, ad-based AI interfaces) find their business models incompatible. They either pivot or decline. Smaller, values-driven AI labs and cooperatives thrive – they can innovate around trust, personalization, and well-being. New markets emerge: Education (AI mentors supporting intrinsic motivation rather than compliance), health (AI companions enhancing self-awareness and affect control, not simulation of affection), governance and democracy (civic AI systems enhancing participation and deliberation), and general work (AI tools augmenting creativity and collective intelligence, not replacing labour). The trust economy replaces the attention economy.

Human agency renaissance: Post-AGI disappointment makes society crave grounding, relationality, and purpose. Pro-human AI answers that demand: systems built to enhance human agency, not mimic it. This cultural turn creates market pull (not just regulatory push) for humane design. People experience AI not as competition, but as tools that *strengthen personhood*. Time saved by automation is reinvested in relationships, culture, and civic life. "AI etiquette" and design ethics evolve: AI systems are evaluated by how they help humans flourish, not by benchmark scores. Education systems teach AI relational literacy: how to engage with

machines that respect autonomy and difference. Drivers for adoption are: Societal exhaustion with dehumanizing tech experiences; strong alignment with mental health and well-being policy goals; integration with education and public-sector services; cultural resonance with sustainability values. Adoption may begin in Europe, Japan, and smaller democracies, spreading through regulatory emulation and consumer pressure. Over time, global demand for *trusted AI* outweighs marginal productivity gains from manipulative or extractive systems.

Phase 3: Lean intelligence (2030-2035)

From academic breakthrough to market adoption: Early academic breakthroughs at the end of the 2020s in brain-inspired continual learning hinted at major efficiency gains: AI systems that learn on the fly without retraining from scratch. Now these new learning architectures (e.g., sparse, energy-efficient, plasticity-based networks) reach production quality. Compute requirements for model adaptation fall by 100x–1000x and training becomes incremental, local, and personalized. The global AI economy undergoes a structural deflation: massive, centralized training clusters lose strategic importance and the capital moat vanishes. Value shifts to distributed deployment and integration. Decentralized innovation flourishes: local firms, cooperatives, and research hubs deploy context-aware models on modest hardware. Small labs, local companies, and even community centres can now run adaptive, context-aware, pro-human AI locally. Policy focus transitions from building new GPU farms to upgrading existing ones for energy efficiency and edge integration. Governments and publics, after the mid-2020s AI bubble, have political memory: they're more sceptical of monopolization and more willing to regulate towards openness.

Ubiquitous AI: With low compute costs, AI becomes as ubiquitous as smartphones: embedded in devices, local networks, and classrooms. Data centres from the AGI boom are retrofitted for edge coordination, climate simulation, and open scientific modelling; their sunk cost yields continuing value. AI becomes a *general public good*: most models are open, self-improving, and tuned locally. Economic growth is steady, human-centred, and less capital-intensive; the centre of innovation moves from speculative finance to education, health, and climate adaptation. The biggest threat to the "lean intelligence" future isn't technical failure – it's *institutional inertia* and *rent-seeking by incumbents*. But once intelligence becomes *computationally cheap*, power can't hoard it easily. In that sense, the very success of neuroscience-inspired efficiency makes central control brittle; it *erodes the economic foundation of monopolies* faster than policy alone could. Now, education systems evolve (instead of teaching to outcompete machines, curricula emphasize relational intelligence, creativity, ethics, and self-direction) and labour markets stabilize (AI becomes a cognitive exoskeleton for workers, augmenting care, craftsmanship, and local governance rather than replacing them).

Conclusions

The failure of AGI expectations triggers a moral and structural realignment. By mid-2030s, intelligence is no longer centralized, extractive, or opaque. Pro-human AI is technology designed to deepen, not diminish, humanity. Coupled with efficient, continual-learning systems, this leads to a globally distributed, low-energy, human-centred intelligence landscape. By 2035, the narrative shifts from surpassing humanity to supporting human flourishing. To a sustainable, equitable intelligence ecosystem built around collaboration, not competition.