AI-Augmented Organisations

A Strategic Framework for Implementing AI
to Augment Employees and Organisations in Europe

For CIOs, CTOs, and Technology Leaders

March 2026 · v1.0

Michael Eriksson

Founder & CEO, Liminal Discovery

Liminal Discovery SL

www.liminaldiscovery.com

Contents

1 Executive Summary **3**

2 Introduction **5**

2.1 The productivity evidence—and its limits 5
2.2 The neurodiversity opportunity 5
2.3 The European regulatory and institutional context 6
2.4 Purpose and structure of this report 7

3 The Evidence Base **8**

3.1 What determines AI implementation success? 8
3.2 The workforce transformation thesis 9
3.3 Designing effective human–AI collaboration 9
3.4 Readiness is socio-technical, not purely technical 10
3.5 The labour market evidence: complementarity dominates substitution 10
3.6 AI augmentation across organisational levels 11
3.7 Explainability as strategic capability 11
3.8 Governance and the EU AI Act 11
3.9 The unmeasured dividend: AI and neurodivergent productivity 12
3.10 Evidence confidence levels: a boardroom reference 14
3.11 CEO ownership and the investment trajectory 16
3.12 Summary tables 17

4 Case Studies **19**

4.1 Case Study 1 (Success): Octopus Energy — AI as Augmentation, Not Replacement 19
4.2 Case Study 2 (Success): DBS Bank — A Decade of Deliberate Capability-Building 20
4.3 Case Study 3 (Success): Siemens — Industrial AI as Workforce Augmentation . . . 21
4.4 Case Study 4 (Failure): Klarna — The Cost of Automation Without Augmentation 23
4.5 Case Study 5 (Failure): IBM Watson Health — The $4 Billion Cost of Skipping Co-Development 24
4.6 Case Study 6 (Failure): Zillow Offers — When Data Dependency Is Underestimated 26
4.7 Cross-case analysis 28

5 Synthesis and Gap Analysis: Where Theory Meets Practice **30**

5.1 Convergence: what both streams agree on 30
5.2 Divergence one: structured capability-building vs. emergent 31
5.3 Divergence two: the measurement vacuum 32
5.4 Divergence three: the shadow AI paradox 33

5.5 Divergence four: the AI intensification problem . . . 34
5.6 Divergence five: the unmeasured neuroinclusion dividend . . . 35
5.7 Organisational scale and the implementation gap . . . 37
5.8 The meta-finding: implementation is the strategy . . . 38
5.9 Technical architecture implications for CTOs . . . 38
5.10 The conclusions in brief . . . 40

6 Recommendations for European Technology Leaders . . . 41

6.1 Reframe the AI budget: 50/30/20 . . . 41
6.2 Select initial use cases using the "strategic friction" principle . . . 42
6.3 Design for cognitive diversity from day one . . . 42
6.4 Implement "AI Practice" norms to prevent work intensification . . . 43
6.5 Address shadow AI through enablement, not prohibition . . . 43
6.6 Foundation practices . . . 44
6.7 Summary of recommendations . . . 45

About the Author . . . 46

About Liminal Discovery . . . 46

7 Appendix: Glossary of Key Terms and Abbreviations . . . 47

7.1 AI and technology terms . . . 47
7.2 Organisational and implementation terms . . . 49
7.3 Neurodiversity and cognitive diversity terms . . . 51
7.4 Regulatory and institutional terms . . . 52
7.5 Abbreviations reference . . . 54

1. Executive Summary

This report identifies three findings that are absent from the current AI implementation literature and analyses why they matter for European technology leaders.

Finding 1: The unmeasured neuroinclusion dividend. Generative AI may be the most significant neuroinclusion intervention ever deployed at scale—and it has been deployed almost entirely by accident. Neurodivergent professionals (an estimated 15–20% of the workforce) are 55% more likely to adopt AI tools than neurotypical colleagues, and 88% report productivity gains. General AI productivity research consistently shows bottom-quintile workers improving by 30–35%. Yet no controlled study has directly measured AI's differential impact on neurodivergent versus neurotypical workers—a gap that represents both an evidence failure and a first-mover opportunity for organisations willing to measure what others assume.

Finding 2: AI intensifies work rather than reducing it. Ethnographic evidence published in HBR (2026) documents that AI augmentation without deliberate work design produces not the promised productivity-and-leisure dividend, but expanded workloads, fragmented attention, and cognitive fatigue. This finding contradicts the prevailing narrative and has direct implications for European organisations operating under labour frameworks that foreground worker wellbeing.

Finding 3: The five divergences between theory and practice. Academic research and business practice agree on fundamentals (organisational capabilities matter more than models; training has the highest ROI; readiness is socio-technical) but diverge consequentially on how capabilities should be built, how success should be measured, how shadow AI should be governed, and whether neuroinclusion and work intensification are even on the agenda. These divergences are where implementation fails, and understanding them is the analytical core of this report.

The shared ground. These findings sit atop a substantial evidence base reviewed in Section 3. European firm-level data estimates a 4% average productivity uplift from AI adoption, but gains accrue disproportionately to firms that invest in complementary assets—above all workforce training, where an additional percentage point of expenditure yields nearly six percentage points of productivity gain. MIT research finds that 95% of generative AI pilots fail to deliver P&L impact, driven by structural failures: wrong use cases, generic tooling, deferred governance, and an inability to measure outcomes. BCG's research finds that organisations which allocate over 50% of their AI budget to workforce development dramatically outperform those that prioritise technology spend. The EU regulatory environment, when treated as a design input rather than an obstacle, produces augmentation architectures that are simultaneously compliant and effective.

Implications. The bottleneck is not access to AI models; it is the ability to reshape work. Training expenditure delivers the highest measurable return of any AI-related investment. The window for competitive positioning is open but closing: waiting for "proven" approaches is itself a high-risk strategy.

This report draws primarily on enterprise-scale research and international case evidence. A planned companion report will translate these findings into SME-specific guidance through primary research with European mid-market technology leaders.

---

2. Introduction

Artificial intelligence has become an infrastructure capability—available through APIs, embedded in productivity suites, and increasingly accessible as open-source tooling—that is profoundly rewiring how work is performed in European firms of every size. Yet the dominant boardroom narrative remains unhelpfully polarised between promises of full automation and anxiety about displacement. This polarisation obscures the question that actually matters.

That question is not "Should we adopt AI?" The competitive, regulatory, and labour-market pressures bearing down on European organisations have already answered it. The real strategic question—and the one this report addresses—is: ***How do we use AI to augment people and organisations in a way that is economically meaningful, culturally sustainable, and legally robust within the European institutional context?***

2.1. The productivity evidence—and its limits

European firm-level studies estimate a 4% average productivity uplift for AI-adopting firms, with task-level gains of 10–65% depending on domain and baseline skill. These gains accrue disproportionately to firms investing in complementary assets—above all workforce training, where an additional percentage point of expenditure yields nearly six percentage points of productivity gain. But the pathway from task-level improvement to organisational performance runs through mediating factors poorly captured by these headline figures: workflow integration quality, employee trust, governance effectiveness, and adaptive capacity. MIT's finding that 95% of generative AI pilots fail to deliver measurable P&L impact confirms that the technology is not the bottleneck—the organisation is.

2.2. The neurodiversity opportunity

AI augmentation is reshaping what "talent advantage" can mean—and doing so in ways that most organisations have not yet grasped, let alone measured. An estimated 15–20% of the general population is neurodivergent—encompassing ADHD, autism spectrum conditions, dyslexia, dyspraxia, and other cognitive profiles. In knowledge-intensive and technical workforces, the proportion is often higher. Neurodivergent professionals frequently bring distinctive strengths in pattern recognition, systems thinking, creative problem-solving, and deep focus—capabilities that become more valuable in an AI-augmented environment where routine cognitive tasks are increasingly handled by machines.

Yet these strengths have historically been underleveraged, because traditional workplace designs impose disproportionate friction on neurodivergent workers. Emerging evidence suggests that AI tools can materially reduce this friction—and the early quantified data is striking. EY's study of 317 neurodivergent and disabled employees using Microsoft 365

Copilot found that 88% reported feeling more productive and 96% reported time savings. EY's larger Global Neuroinclusion at Work Study (2,111 professionals across 22 countries) found that neurodivergent workers are 55% more likely to use AI at work than neurotypical colleagues, and 81% rated AI "very or extremely effective" at improving information access. The UK Department for Business and Trade's Copilot pilot found neurodiverse employees were 25% more satisfied with AI tools than neurotypical respondents.

These neurodivergent-specific findings are reinforced by a broader pattern in the AI productivity literature: AI disproportionately benefits lower-performing workers, compressing productivity distributions. In the most rigorous field study to date, Brynjolfsson, Li, and Raymond found that bottom-quintile customer-support agents improved by 30–35% with AI assistance while top performers saw near-zero gains. In a BCG experiment, the performance gap between top and bottom consultants shrank from 22% to just 4% with AI access. The implication for neurodivergent workers whose traditional performance metrics were suppressed by environmental friction—not by lack of capability—is profound: AI tools may be removing precisely the barriers that produced the underperformance.

For European organisations facing acute skills shortages and demographic contraction, this is a strategic lever: designing AI augmentation with cognitive diversity as a default parameter expands the effective talent pool, improves retention, and builds organisational resilience.

2.3. The European regulatory and institutional context

European organisations operate under an institutional regime materially different from the environments described in most US- or Asia-centric AI implementation literature. The EU AI Act (in force August 2024, phased compliance through 2026) introduces risk-based classification with stringent requirements for high-risk systems—including AI used in employment decisions—covering data quality, transparency, human oversight, and documentation, with penalties of up to €35 million or 7% of global turnover. But for CIOs and CTOs, this is not merely a compliance exercise. The EU AI Act, together with GDPR, national labour laws, and works councils, collectively define the design space within which viable augmentation strategies must operate. The requirements for human oversight, explainability, and data governance that the EU AI Act mandates are closely aligned with the design principles that research identifies as prerequisites for successful human–AI collaboration. Organisations that treat regulation as a design input—rather than an obstacle—tend to build more transparent, trustable, and effective AI-augmented systems.

2.4. Purpose and structure of this report

This report provides a rigorous, practitioner-relevant analysis of how to implement AI in ways that systematically augment employees and organisations in European settings. It integrates three analytical lenses: a synthesis of recent academic research on AI implementation capabilities, human–AI collaboration, organisational agility, and neuroinclusion; a distillation of findings from major global consultancies and leading firms that have moved beyond pilots to scale AI augmentation; and six detailed case studies—three successes and three failures—analysed for root causes and transferable lessons.

The central thesis is that AI transformation is, at its core, workforce transformation. The organisations that derive durable competitive advantage from AI are not those with the most sophisticated models or the largest data estates. They are the ones that systematically build augmentation into roles, skills, structures, and governance—explicitly including cognitive diversity—rather than chasing automation for its own sake.

The report is structured as follows. Section 3 synthesises the academic and business evidence base, with analytical commentary on what each body of research underweights. Section 4 presents six case studies with structured analysis connecting each to the report's analytical framework. Section 5 provides the analytical core: a synthesis identifying where theory and practice converge, diverge, and why the gaps matter, concluding with the report's key findings. Section 6 translates these into concrete recommendations. An appendix provides a glossary of key terms.

---

3. The Evidence Base

This section synthesises two complementary bodies of knowledge: recent academic research on AI implementation, and operational findings from major consultancies and leading firms. The academic literature establishes what works and why at a theoretical level; the business evidence reveals how these principles play out—and often break down—in practice. Rather than presenting these as separate literature reviews, the discussion is organised thematically around the questions most relevant to European technology leaders making implementation decisions.

3.1. What determines AI implementation success?

Organisational capabilities, not algorithmic sophistication, are the primary determinant. Weber et al. (2023) identify four capabilities that address AI's distinctive properties—inscrutability and data dependency: *AI project planning* (use-case identification anchored in workflow bottlenecks), *co-development* (integrating domain expertise throughout, not just at requirements), *data management* (making data findable, clean, and governed), and *AI model lifecycle management* (disciplined progression from experimentation through deployment to retirement).

MIT's NANDA initiative validates this powerfully: **95% of enterprise generative AI pilots fail to deliver measurable P&L impact**—not because the technology fails, but because organisations lack the disciplines to make AI work at scale. The four structural failure modes: *friction avoidance* (choosing low-stakes use cases), *generic tooling* without contextual adaptation, *learning gaps* (no feedback loops), and *shadow AI economies* (ungoverned workarounds).

The counter-intuitive implication: **the most productive path is not to minimise friction but to choose the right friction**—selecting strategically important use cases and building feedback loops.

What this finding underweights: Both Weber and the MIT research treat "organisational capabilities" as though they develop uniformly across the workforce. They do not. An organisation's ability to co-develop, for example, depends on whether its co-development processes accommodate different communication styles, processing speeds, and modes of expertise contribution. The capability frameworks are necessary but incomplete—they describe *what* to build without fully addressing *for whom* it needs to work.

3.2. The workforce transformation thesis

AI transformation and workforce transformation are inseparable. BCG's AI Radar 2026, surveying over 1,800 executives, finds that roughly 70% of AI's value comes from rethinking the people component—roles, skills, workflows, governance—while only 30% comes from the technology itself. Trailblazers allocate approximately 60% of their AI budget to workforce development (versus ~15% for Followers), have upskilled 70% of their workforce (versus under 20%), and pursue end-to-end transformation rather than use-case-by-use-case thinking. The performance gap is driven almost entirely by people strategy, not technology choices.

The European evidence reinforces this: an additional percentage point of training expenditure increases AI-related productivity gains by nearly six percentage points. PwC's 2025 Global AI Jobs Barometer documents that the AI skills premium has more than doubled in a single year, from 25% to 56%.

What this thesis assumes but does not examine: "Workforce development" is treated as a monolithic category—as if training 70% of a workforce on AI is a uniform intervention. It is not. The 70% includes individuals with vastly different cognitive profiles, learning styles, prior technology experience, and psychological orientations toward AI. The workforce transformation thesis correctly identifies *where* to invest but says little about *how* that investment should be differentiated across the actual cognitive diversity of the workforce it targets.

3.3. Designing effective human–AI collaboration

Kolbjørnsrud et al. (2024), in the *California Management Review*, propose six principles for "intelligent organisations": *addition* (more intelligent actors increase collective capability), *relevance* (matching intelligence type to problem type), *substitution* (replacing lower intelligence with higher, in both directions), *diversity* (combining different forms of intelligence), *collaboration* (developing collaborative skills in both humans and AI), and *explanation* (investing in transparency and interpretive capacity).

McKinsey's 2025 research operationalises these through a three-phase workflow evolution: Phase 1, stand-alone AI agents on discrete tasks; Phase 2, groups of AI agents completing end-to-end processes under human oversight; Phase 3, fully autonomous agentic systems. Most organisations should focus on Phases 1 and 2. The practical guidance: select two or three strategically important workflows, map them in detail, and redesign them around a human–AI collaboration model rather than deploying AI widely and generically.

What this framework overlooks: The six principles assume that "collaboration" and "explanation" mean the same thing to all users. They do not. An explanation that provides clarity for one cognitive profile may produce information overload for another. The collaboration design that suits a highly verbal, linear thinker may exclude someone whose strength is pattern recognition or visual reasoning. The principles are sound in the abstract but require differentiation in application—a gap that becomes visible only when implementation encounters the actual diversity of a workforce.

3.4. Readiness is socio-technical, not purely technical

People, process, and data readiness matter more than technology readiness. Uren and Edwards (2023) propose an extended readiness model finding that barriers to AI adoption are located predominantly in non-technical dimensions. This is corroborated by a 2025 executive survey in which 91% identified cultural barriers as greater impediments than technological limitations. Complementary research on cognitive personas in AI transformation (IEEE, 2025) shows that employees differ not just in skills but in psychological orientation toward AI—trust levels, tolerance for uncertainty, and preferred interaction modes. One-size-fits-all rollout strategies systematically underperform.

What the readiness models miss: The four-dimensional framework (technology, people, process, data) treats "people readiness" as a single dimension. In practice, it is at least three: skills readiness (can people use AI tools?), psychological readiness (do they trust and want to?), and cognitive readiness (does the implementation accommodate how they actually process information?). Organisations that assess only the first two will consistently underestimate resistance and underperformance in the third—particularly among the 15–20% of the workforce whose cognitive processing diverges from the neurotypical assumptions embedded in most AI interfaces.

3.5. The labour market evidence: complementarity dominates substitution

Mäkelä et al. (2024), analysing 12 million job vacancies, find that **complementarity effects are 1.7 times larger than substitution effects**. AI-focused roles are nearly twice as likely to require complementary skills—resilience, analytical thinking, teamwork, ethics, agility—and these skills command significant wage premiums. The net effect: AI creates more demand for human skills than it destroys, provided those skills complement AI capabilities.

The underexamined implication: If the skills commanding premiums in an AI-augmented market are contextual judgment, ethical reasoning, and collaborative intelligence, then the labour market is effectively pricing in the *human* capabilities that AI cannot replicate. This has a second-order consequence that the complementarity literature does not address: neurodivergent workers often possess precisely these premium capabilities (systems thinking, pattern recognition, deep-focus analysis) but have been systematically excluded by workplace designs that penalise their processing differences rather than leveraging their strengths. The complementarity thesis implicitly argues for cognitive diversity without ever naming it.

3.6. AI augmentation across organisational levels

Prikshat et al. (2023) map AI augmentation across three levels: *micro* (individual task automation, decision support), *meso* (team coordination, knowledge sharing), and *macro* (organisational productivity, innovation, agility). The critical finding: most implementations achieve micro-level gains only. Macro-level integration—connecting AI tools to organisational strategy—is where implementations fall short and where systemic value lies.

The analytical gap: This micro-meso-macro framework correctly diagnoses the problem (organisations stop at micro) but does not adequately explain *why*. The transition from micro to meso requires that AI tools work across different team members' cognitive styles, communication preferences, and processing speeds. A team-level AI integration that assumes uniform interaction patterns will fragment at exactly the point where it needs to cohere. The jump from micro to macro is not just an architectural challenge—it is a human-systems design challenge that requires accounting for cognitive diversity at the team level.

3.7. Explainability as strategic capability

Shafiabady et al. (2024) find that the strongest predictors of organisational agility are not technological but organisational: decentralised decision-making, active learning cultures, ethical awareness, and customer orientation. AI enhances agility only when it is explainable. Opaque AI produces short-term efficiency at the cost of long-term organisational fragility. **Explainability is the critical mediating factor** between AI deployment and organisational agility—reinforcing the strategic logic of the EU AI Act's explainability requirements as design enablers rather than bureaucratic overhead.

The unasked question: If explainability mediates AI's contribution to agility, then explainability *for whom*? The literature treats explainability as a property of the system. In practice, it is a property of the interaction between the system and the user. An explanation that is transparent to a verbal, sequential processor may be opaque to a visual, pattern-based thinker. Explainability requirements under the EU AI Act are necessary but insufficient unless they account for the cognitive diversity of the humans who must actually understand and oversee AI outputs.

3.8. Governance and the EU AI Act

Deloitte's State of AI in the Enterprise (2026) finds that **governance is the gating factor for AI scaling, not technology**. Nearly 60% of AI leaders cite compliance concerns as primary barriers; 67% of enterprise AI initiatives that stalled in late-stage deployment cited regulatory risk overtaking technical feasibility.

The EU AI Act defines specific governance requirements: risk classification early (high-risk systems face full enforcement from August 2026), human oversight designed in from the outset, non-negotiable documentation and transparency, and provisions for smaller organisations including regulatory sandboxes. The strategic insight: governance integrated from the outset functions as an accelerator—organisations that classify use cases early and design oversight into workflows avoid the costly late-stage failures that derail those treating governance as an afterthought.

What the governance literature neglects: The EU AI Act's human oversight requirements implicitly assume a neurotypical overseer—someone who processes information in the default modality, at the default speed, through the default interface. For organisations where 15–20% of the workforce processes information differently, "meaningful human oversight" requires interface diversity, not just process compliance. This is not a speculative concern—it is a gap in how governance is currently operationalised, with direct implications for the organisations required to demonstrate non-discrimination under the Act.

3.9. The unmeasured dividend: AI and neurodivergent productivity

Generative AI may be the most significant neuroinclusion intervention ever deployed at scale—and almost entirely by accident. AI-powered writing assistants, summarisation tools, task decomposition agents, and personalised interfaces are being adopted across millions of workplaces for general productivity reasons, yet their architecture maps precisely onto the cognitive accommodations that neurodivergent workers have historically lacked: structured information, reduced executive-function overhead, multiple processing modalities, and real-time scaffolding for tasks that impose disproportionate friction on non-neurotypical cognitive profiles.

The evidence now spans three tiers: neurodivergent-specific survey research, general AI productivity studies showing disproportionate gains for lower-performing workers, and pre-AI neurodivergent productivity benchmarks. The convergence across these tiers is compelling. But **no controlled study has yet directly measured AI's differential productivity impact on neurodivergent versus neurotypical workers performing identical tasks**—a gap that is itself a finding, with significant implications for both research and competitive strategy.

Neurodivergent-specific AI evidence. The EY–Microsoft Copilot Accessibility Study (December 2024) surveyed 317 employees who self-identify as disabled or neurodivergent across 17 organisations and 7 sectors, all using Microsoft 365 Copilot. 88% reported feeling more productive, 87% said Copilot reduced the mental energy demanded by tasks, 96% reported time savings, and 91% viewed Copilot as a valuable assistive technology. On specific cognitive functions, 80% reported improved written communication, 59% improved memory and recall, and 48% improved concentration and focus. These are perception metrics from a convenience sample with no control group, but the consistency across functions is notable.

EY's Global Neuroinclusion at Work Study (June 2025)—the largest in this space, surveying 2,111 professionals (1,603 neurodivergent, 508 neurotypical) across 22 countries—found that neurodivergent workers are 55% more likely to use AI at work (79% of neurodivergent vs. 51% of neurotypical professionals). Among neurodivergent AI users, 81% rated AI "very or extremely effective" at improving information access, 77% at improving work quality, and 71% at enhancing communication quality. The study also found that neurodivergent professionals who feel "truly included" report up to 10% higher skill proficiency across the World Economic Forum's 10 fastest-growing skills.

The UK Department for Business and Trade's Copilot Pilot (2024–2025) deployed 1,000 licences with 300 participants evaluated via diary studies, interviews, and usage dashboards. Neurodiverse employees were 25% more satisfied with AI assistants than neurotypical respondents, with statistically higher satisfaction (at 90% confidence) and stronger willingness to recommend (at 95% confidence). General time savings of 1.3 hours per drafting task and 0.8 hours per research summarisation task were reported across all users, though the study found no evidence of department-level productivity gains overall.

AI as equaliser: the proxy evidence. The strongest quantified productivity evidence comes from general AI studies that consistently show AI disproportionately benefits lower-performing workers—a finding widely extrapolated to neurodivergent populations whose traditional performance metrics may have been suppressed by environmental friction rather than lack of capability. Brynjolfsson, Li, and Raymond (published *Quarterly Journal of Economics*, 2025) studied 5,172 customer-support agents and found bottom-quintile workers improved by 30–35% while top-quintile workers saw near-zero improvement. In a BCG experiment with 758 consultants, Dell'Acqua et al. (2023) found bottom-half performers saw a 43% quality improvement versus 17% for the top half, compressing the performance gap from 22% to just 4%. In randomised trials with approximately 4,867 developers, Cui et al. (2024) found junior developers improved 35–39% in completed pull requests versus 8–16% for senior developers.

Pre-AI neurodivergent benchmarks. Several widely-cited figures establish that neurodivergent workers can dramatically outperform when properly matched to roles. JPMorgan Chase's Autism at Work programme reported that autistic employees in certain technology roles were 90–140% more productive than neurotypical colleagues, with a 99% retention rate across 200+ employees in 8 countries. Hewlett Packard Enterprise reported neurodiverse testing teams were 30% more productive in preliminary results from its Australian Department of Human Services programme—a figure subsequently generalised by Deloitte Insights and the World Economic Forum. These benchmarks should be cited with the caveat that they reflect role-matching effects in highly structured programmes, not generalised productivity differences.

AI tools can function as "cognitive prosthetics" that reduce the friction neurodivergent workers experience in neurotypically designed workplaces. Microsoft's research (2026) documents how AI features—real-time transcription, intelligent summarisation, task decomposition, focus-mode interfaces—enable neurodivergent professionals to work more effectively. PwC's analysis argues that many neuroinclusive design principles—clear information architecture, multiple modalities, structured task management—improve outcomes for all employees through the "curb-cut effect."

For CIOs and CTOs, the practical implication is that **neuroinclusive design should be a default parameter in AI augmentation strategy, not a specialist accommodation.** When selecting or configuring AI tools, "Does this work for people who process information differently?" is a design quality indicator that correlates with better outcomes for the entire workforce. But the strategic implication runs deeper: organisations that systematically measure AI's differential impact on neurodivergent employees will generate proprietary insights unavailable to competitors. The evidence gap is not just a research curiosity—it is a first-mover opportunity. Any European organisation that conducts even a modest controlled comparison of AI-augmented task performance across cognitive profiles will produce the study that does not yet exist, positioning itself at the intersection of an emerging evidence base with direct implications for workforce strategy, talent acquisition, and inclusion policy.

3.10. Evidence confidence levels: a boardroom reference

The neurodiversity–AI evidence assembled above spans three tiers of varying rigour. Technology leaders presenting these findings to boards or investment committees should be explicit about what is established, what is strongly indicated, and what remains an inference requiring validation.

Evidence tier	What it shows	Strength	Limitations	Confidence
Tier 1: Neurodivergent-specific AI studies (EY–Microsoft, UK DBT pilot)	88% self-reported productivity gain; 55% higher adoption; 25% higher satisfaction among ND workers	First quantified data on ND–AI interaction; consistent across multiple studies and geographies	Self-reported; convenience samples; no control groups; funded by technology vendors	**Moderate** — consistent signal, but not yet validated by independent controlled studies
Tier 2: AI-as-equaliser studies (Brynjolfsson et al., Dell'Acqua et al., Cui et al.)	Bottom-quintile workers gain 30–35%; performance gap compression from 22% to 4%	Rigorous methodology (RCTs, large samples, published in top journals); replicated across domains	Did not measure neurodivergent status; the extrapolation to ND populations is an inference, not a finding	**High** for the general finding; **Moderate** for the ND extrapolation
Tier 3: Bridge inference — combining Tiers 1 and 2	If AI disproportionately helps lower performers, and ND workers underperform due to environmental friction (not capability), then AI should disproportionately help ND workers	Logically coherent; consistent with all available data; no contradicting evidence	Untested directly; relies on the assumption that ND underperformance is environmentally produced; could be confounded by other factors	**Plausible but unvalidated** — the hypothesis is strong; the direct evidence does not yet exist
Tier 4: Pre-AI benchmarks (JPMorgan, HP Enterprise)	90–140% ND productivity in matched roles; 30% higher in structured programmes	Demonstrates ND capability when friction is removed	Role-matching effects in highly structured programmes; not generalisable to all roles or contexts	**High** for matched-role contexts; **Low** as general productivity claims

The honest summary for a board presentation: The evidence that AI tools benefit neurodivergent workers is consistent, multi-sourced, and growing. The evidence that this benefit is *larger* than for neurotypical workers is strongly indicated but not yet proven by controlled study. The strategic recommendation to design for neuroinclusion is justified on three independent grounds: (1) the evidence that does exist, (2) the curb-cut effect (what helps ND workers helps everyone), and (3) the first-mover opportunity to generate the evidence that closes the gap. An organisation does not need to wait for proof to act—but it should be transparent about what is proven and what is inferred.

3.11. CEO ownership and the investment trajectory

BCG's AI Radar 2026 reveals a structural shift: CEOs are now the primary owners of AI strategy, displacing CIOs and CTOs as lead decision-makers. Companies plan to double AI spending in 2026, and 94% plan to maintain elevated investment even without immediate ROI. For technology leaders, the optimal role is as the bridge between business ambition and technical reality—ensuring AI strategy is grounded in realistic assessments of data readiness, organisational capability, and regulatory constraints.

3.12. Summary tables

Academic research findings:

Research stream	Core finding	Implication for technology leaders
Organisational capabilities (Weber et al.)	Four capabilities predict implementation success	Build deliberately; don't assume they emerge from project experience
Phased transformation (Holmström & Magnusson)	AI transformation requires strategic sequencing	Resist skipping from experimentation to scaling without groundwork
Intelligent organisation design (Kolbjørnsrud et al.)	Six principles govern effective human–AI collaboration	Design for complementarity; invest in explainability
Socio-technical readiness (Uren & Edwards)	People, process, and data readiness matter more than technology	Start with readiness assessment, not technology procurement
Labour market complementarity (Mäkelä et al.)	Complementarity effects are 1.7× larger than substitution	Invest in complementary skills; augmentation creates more value than automation
Multilevel HRM augmentation (Prikshat et al.)	Most organisations only achieve micro-level augmentation	Design for integration across organisational levels
Agility and explainability (Shafiabady et al.)	Explainable AI enhances agility; opaque AI undermines it	Treat explainability as strategic capability
Neurodiversity and AI (the unmeasured dividend)	88% of ND workers report productivity gains with AI; 55% higher adoption; AI equalises bottom-quintile performance by 30–35%; no controlled ND/NT differential study exists	Make neuroinclusion a default design parameter; measure differential impact; the evidence gap is a first-mover opportunity

Business best practices:

Practice	Evidence base	Implication
Invest 50–60% of AI budget in workforce development	BCG AI Radar 2026	The most reliable differentiator
Start from workflows, not models	McKinsey 2025	Select 2–3 strategic workflows; redesign end-to-end
Navigate friction, don't avoid it	MIT NANDA 2025	Choose strategically important use cases
Build contextual intelligence, not generic tooling	MIT NANDA 2025	Adapt AI to your specific data and workflows
Integrate governance from day one	Deloitte 2026, EU AI Act	Classify use cases early; design oversight in
Lead with coalition of willing, not mandate	McKinsey 2025	Empower early adopters; build trust through evidence
Design for cognitive diversity by default	EY 2024/2025, Microsoft 2026, PwC 2024, UK DBT 2025	Audit tools for accessibility; measure ND/NT differential impact
Position CIO/CTO as bridge	BCG 2026	Own integration of business strategy with capability and regulatory readiness

4. Case Studies

This section presents six case studies—three successes and three failures—selected for their relevance to the challenges facing European technology leaders. While several of these organisations are larger than typical enterprises, the dynamics they illustrate are scale-independent: the same structural factors that produced success or failure operate in organisations of any size. Each case study follows a consistent structure: context, what happened, root cause analysis, and transferable implementation principles.

4.1. Case Study 1 (Success): Octopus Energy — AI as Augmentation, Not Replacement

Context

Octopus Energy is a UK-based renewable energy supplier that has grown rapidly to serve over 10 million retail customers globally. The company built its brand on exceptional customer service in an industry notorious for poor service. As the company scaled, it faced the classic tension: how to maintain service quality while managing the costs of rapid growth. The company chose an augmentation-first approach that has made it a widely cited example of effective human–AI collaboration.

What happened

Following the release of ChatGPT in November 2022, Octopus Energy moved quickly. By January 2023 it had deployed an AI customer service assistant for a small percentage of customers; by May 2023, nearly half of all customer communications were AI-assisted. The AI was trained on the company's best service interactions, learning the empathetic, helpful communication style that had earned Octopus its reputation.

The results were striking. AI-assisted responses achieved an 80–85% customer satisfaction rating, compared with 65% for responses composed entirely by human agents—a counter-intuitive finding that challenged the assumption that human service is inherently superior. The AI handled billing questions, scheduled engineer visits, and resolved account issues, while seamlessly transferring complex or emotionally charged cases to human agents. Octopus now holds more five-star reviews on Trustpilot than any other company in the UK across all sectors.

Critically, Octopus did not use AI to replace its customer service team. The company used AI to augment agents—handling routine queries, drafting responses for agent review, and freeing human staff to focus on complex cases where empathy, judgment, and contextual understanding add the most value.

Root cause analysis: why it worked

- **Started from the workflow, not the model.** Octopus identified the specific tasks within customer service where AI could add value (routine queries, drafting, triage) and designed the human–AI collaboration around those tasks, rather than deploying a generic chatbot.

- **Trained AI on its own best practice.** By training the model on Octopus's own highest-rated interactions, the company ensured contextual fit—the AI sounded like Octopus, not like a generic assistant.
- **Maintained meaningful human oversight.** Complex cases were escalated to human agents automatically. The AI was positioned as a tool that makes agents better, not a replacement that makes agents unnecessary.
- **Built on a strong data foundation.** Octopus's Kraken platform integrates meter readings, billing, customer history, and communications in a single system—providing the clean, accessible data that AI systems need to perform well.
- **Measured what mattered.** Customer satisfaction was tracked rigorously from the outset, providing real-time evidence that the augmentation was working and building trust among sceptical staff.

Transferable lessons for European organisations

- AI augmentation in customer-facing roles works best when the AI is trained on the organisation's own data and best practices, not on generic datasets.
- Meaningful human oversight is not a constraint on AI value—it is a contributor to it. The combination of AI speed with human judgment produces outcomes that neither achieves alone.
- A unified data platform is a prerequisite for effective AI augmentation. Fragmented systems produce fragmented AI performance.

4.2. Case Study 2 (Success): DBS Bank — A Decade of Deliberate Capability-Building

Context

DBS Bank, headquartered in Singapore, is Southeast Asia's largest bank by assets. While not European, its AI transformation journey is arguably the most thoroughly documented example of the organisational capability-building approach that the academic literature (Section 3.1) prescribes. DBS has achieved S$1 billion in annual economic value from AI—a figure that makes it one of the few organisations globally to have demonstrably scaled AI beyond pilots to P&L-relevant impact.

What happened

DBS began its AI journey in 2012 with small experiments and accelerated from 2018 with a strategic commitment to data-driven transformation. By 2025, the bank had deployed over 1,500 AI and machine learning models across more than 370 use cases, spanning customer-facing businesses, risk management, fraud detection, and operational support. AI-powered fraud detection now vets 100% of technology change requests, producing an 81% reduction in system incidents. Generative AI tools in customer service cut call handling times by 20%.

The workforce strategy is equally significant. DBS cultivated a 700-person "Data Chapter" of professionals and identified over 12,000 employees for upskilling or reskilling since early 2025. Nearly all have commenced learning roadmaps covering AI and data competencies. The bank simultaneously reduced approximately 4,000 temporary and contract positions over three years—but reported no AI-related layoffs of permanent staff, suggesting that AI is changing job composition rather than job quantity.

CEO Tan Su Shan has been explicit: "It's not hope. It's now. It's already happening."

Root cause analysis: why it worked

- **Deliberate capability-building over a decade.** DBS did not attempt to achieve AI transformation in a single initiative. It built data infrastructure, cultivated data talent, established governance, and developed use cases progressively over more than ten years—precisely the phased approach that Holmström and Magnusson's framework (Section 3.2) prescribes.
- **All four organisational capabilities developed in parallel.** AI project planning (370+ use cases systematically identified), co-development (Data Chapter embedded across business units), data management (unified data platform), and model lifecycle management (1,500+ models in production with monitoring).
- **Experimentation culture normalised.** DBS fostered a culture that treats experimentation as operational necessity, not a luxury. This cultural foundation enabled rapid scaling when generative AI capabilities became available.
- **Workforce transformation treated as core strategy, not afterthought.** Upskilling 12,000 employees is not a training programme; it is a strategic commitment that signals to the workforce that AI is a tool for their development, not a threat to their employment.

Transferable lessons for European organisations

- AI transformation that delivers P&L impact requires years of capability-building, not months of tool deployment. organisations that begin building data, skills, and governance foundations now will be positioned to scale when the technology matures further.
- The ratio of permanent workforce investment to temporary workforce reduction at DBS illustrates the augmentation thesis in practice: AI changes what people do, not whether people are needed.
- A unified data platform—even a modest one—is a prerequisite for scaling beyond isolated use cases.

4.3. Case Study 3 (Success): Siemens — Industrial AI as Workforce Augmentation

Context

Siemens, the German industrial conglomerate, provides a European case study of AI augmentation in manufacturing—a sector where precision matters, error tolerances are narrow, and the consequences of AI failure are physical, not just financial. Siemens CEO Roland

Busch has framed the company's approach as "AI-powered factories need humans, just fewer of them"—a formulation that acknowledges automation while centring augmentation.

What happened

Siemens developed "Industrial AI models"—foundation models augmented with decades of manufacturing data from thousands of factories. The company built data-sharing alliances with German machine builders including Trumpf and DMG Mori, pooling operational information to train models that understand not just generic factory processes but the specific characteristics of individual equipment.

The results are measurable. PepsiCo deployed Siemens Digital Twin Composer to simulate and optimise production line changes before physical modifications, achieving a 20% increase in throughput at a Gatorade facility within three months while reducing capital expenditure by 10–15%. Siemens' predictive maintenance solutions, deployed through its MindSphere platform, enable factories to anticipate equipment failures before they occur, reducing unplanned downtime.

The Siemens Industrial Copilot, developed in partnership with Microsoft, addresses the skilled labour shortage directly. At thyssenkrupp, the copilot enables less experienced engineers to write automation code in significantly less time, compensating for the retirement of experienced workers. The copilot is being rolled out globally across manufacturing, energy, and infrastructure sectors.

Root cause analysis: why it worked

- **Domain-specific, not generic.** Siemens explicitly rejected the approach of applying general-purpose AI to industrial contexts. As Busch noted, a 70% accuracy rate that is acceptable for consumer chat applications would be catastrophic on a pharmaceutical production line. Industrial AI requires domain-specific models trained on domain-specific data.
- **Augmentation framing from the outset.** The Industrial Copilot is designed to make less experienced workers more capable, not to replace experienced workers. This framing has enabled workforce buy-in and union cooperation in the German institutional context.
- **Data pooling as strategic capability.** By building data-sharing alliances with other manufacturers, Siemens overcame the data scarcity problem that limits individual-firm AI performance—a model that European organisations could replicate through industry consortia.
- **European regulatory alignment.** Siemens' emphasis on precision, transparency, and human oversight aligns naturally with the EU AI Act's requirements, positioning the company as a beneficiary rather than a victim of European regulation.

Transferable lessons for European organisations

- In domains where errors have physical or financial consequences, domain-specific AI models trained on domain-specific data are essential. Generic AI tools are insufficient.

- The skilled labour shortage is a strategic context that makes augmentation the rational choice. AI that makes existing workers more capable is more valuable than AI that attempts to replace workers who cannot be recruited.
- Data-sharing alliances between non-competing firms in the same sector can overcome the data scarcity that limits individual organisational AI performance.

4.4. Case Study 4 (Failure): Klarna — The Cost of Automation Without Augmentation

Context

Klarna, the Swedish fintech company, is a cautionary tale that is particularly relevant for European organisations because it illustrates what happens when an organisation optimises for automation and cost reduction rather than augmentation and quality. The case is especially instructive because the initial results appeared overwhelmingly positive—before the consequences of the approach became visible.

What happened

In early 2024, Klarna deployed an OpenAI-powered AI assistant that handled two-thirds of customer service chats in its first month—2.3 million conversations. The company claimed the AI was doing the equivalent work of 700 full-time agents. Resolution times dropped from 11 minutes to 2 minutes. Repeat inquiries fell 25%. The system operated in 23 markets and 35 languages. Klarna projected $40 million in profit improvement for 2024.

CEO Sebastian Siemiatkowski announced that AI could perform the jobs humans do. Klarna paused hiring for over a year and reduced its workforce from 5,500 to 3,400 employees.

Then the consequences emerged. By mid-2025, Klarna began experiencing a surge in customer complaints, declining satisfaction scores, and growing frustration with AI responses that were too generic, repetitive, or unhelpful for complex issues. The AI struggled with nuanced support tasks requiring empathy, discretion, or deeper understanding of customer situations. Siemiatkowski publicly acknowledged that overdependence on automation had led to "lower quality" customer experience. Klarna began quietly rehiring human agents.

Root cause analysis: why it failed

- **Automation logic, not augmentation logic.** Klarna framed AI as a replacement for human workers, not as a tool to make human workers more effective. The 700-agent equivalence was presented as a cost saving, not as a redeployment of human capability to higher-value tasks.
- **Speed metrics mistaken for quality metrics.** The reduction from 11-minute to 2-minute resolution times was celebrated without adequate attention to whether the faster resolutions actually resolved customer problems. The 25% reduction in repeat inquiries was an early positive signal, but it masked a growing category of customers whose issues were closed by the AI but not genuinely resolved.

- **Complex and emotional cases mishandled.** Customer service in financial services frequently involves stress, confusion, and emotional vulnerability. The AI lacked the capacity for empathy, discretion, and contextual judgment that these interactions require—precisely the human capabilities that the augmentation thesis identifies as irreplaceable.
- **Workforce reduction eliminated the escalation path.** By reducing headcount so aggressively, Klarna compromised its ability to handle the cases that AI could not. When the volume of complex cases exceeded the reduced human team's capacity, service quality collapsed.
- **No governance feedback loop.** The initial positive metrics created organisational complacency. There was no systematic mechanism for detecting the slow accumulation of customer dissatisfaction until it reached crisis levels.

Transferable lessons for European organisations

- The automation-first framing is seductive but dangerous. Short-term cost savings from replacing human workers with AI can be overwhelmed by medium-term costs of declining quality, customer attrition, and reputational damage.
- Resolution speed is not resolution quality. Any metric that measures efficiency without measuring effectiveness is a potential trap.
- Human oversight and escalation capacity must be maintained as a structural capability, not treated as a legacy cost to be eliminated.
- In regulated European markets—where the EU AI Act mandates human oversight for high-risk applications and where works councils have consultation rights on workforce changes—Klarna's approach would face additional legal and institutional obstacles.

4.5. Case Study 5 (Failure): IBM Watson Health — The $4 Billion Cost of Skipping Co-Development

Context

IBM's Watson Health initiative, which consumed an estimated $4 billion in investment before being sold off in 2022, is the most extensively documented AI implementation failure in corporate history. While the scale dwarfs anything a smaller organisation would attempt, the root causes are directly relevant because they map precisely onto the organisational capability gaps identified in Section 3.1.

What happened

IBM positioned Watson as a revolutionary AI system that could transform cancer treatment by analysing patient data, medical literature, and clinical records to recommend personalised treatment plans. The system was developed in collaboration with Memorial Sloan Kettering Cancer Center and deployed at hospitals worldwide, including MD Anderson Cancer Center in Houston.

The project failed at multiple levels. At MD Anderson, the "Oncology Expert Advisor" was terminated after four years and $62 million in spending. Watson could not reliably interpret doctors' notes and patient histories. When MD Anderson migrated to a new electronic health record (EHR) system, Watson lost access to patient data entirely. Physicians grew frustrated, wrestling with the technology rather than caring for patients.

More broadly, Watson for Oncology's treatment recommendations were found to be based primarily on training data from a single institution (Memorial Sloan Kettering), making them unreliable in different clinical contexts with different patient populations, treatment protocols, and data formats. The system's recommendations sometimes contradicted local clinical practice without providing adequate explanation for the discrepancy.

Root cause analysis: why it failed

- **Catastrophic co-development failure.** IBM built Watson largely in isolation from the clinical professionals who would use it. Physician involvement was inadequate throughout the development process, resulting in a system that did not fit clinical workflows, could not interpret the messy reality of patient records, and produced recommendations that clinicians neither trusted nor found useful.
- **Data management failure.** Watson's training data was narrow (primarily from one institution), unrepresentative (US-centric treatment protocols), and unable to adapt to the data formats and quality levels of different hospital systems. When MD Anderson changed its EHR, the entire system broke.
- **AI project planning failure.** IBM chose "the highest bar possible, real-time cancer diagnosis, with an immature technology," as a Harvard Business School case study later concluded. The ambition of the use case was radically mismatched with the maturity of both the technology and the organisation's implementation capabilities.
- **Lifecycle management failure.** Watson was deployed as a static product rather than a learning system. It could not adapt to new clinical evidence, changing treatment protocols, or the accumulated feedback of physicians who used it.
- **Marketing-led rather than science-led.** IBM's public communications consistently overstated Watson's capabilities, creating expectations that the technology could not meet and eroding trust when reality became apparent.

Transferable lessons for European organisations

- Co-development with domain experts is not optional. AI systems built by technologists without deep, continuous involvement of the people who will use them will fail—regardless of the sophistication of the underlying models.
- Data representativeness matters as much as data volume. An AI system trained on data from one context will not generalise reliably to different contexts without deliberate adaptation.
- Ambition must be calibrated to capability. Starting with a use case that exceeds the organisation's implementation capabilities produces expensive failures. The "strategic friction" principle (Recommendation 7.3) is the antidote.
- Transparency about AI limitations builds trust; overpromising destroys it. In the European context, where the EU AI Act mandates transparency, this lesson is both strategically and legally relevant.

4.6. Case Study 6 (Failure): Zillow Offers — When Data Dependency Is Underestimated

Context

Zillow, the US online real estate marketplace, provides a stark illustration of what happens when an organisation builds a major business operation on AI predictions without adequately addressing the data dependency challenge identified in the academic literature (Section 3.1).

What happened

Between 2018 and 2021, Zillow purchased approximately 27,000 homes through its "Zillow Offers" iBuying programme. The programme used Zillow's proprietary Zestimate algorithm to predict home values, purchasing properties at algorithm-determined prices, renovating them, and reselling them for a projected profit.

The Zestimate was not accurate enough for the purpose. While the algorithm achieved a median error rate of 1.9% for on-market homes, the error rate for off-market homes was 6.9%—and in volatile markets during the pandemic, errors of 20% or more were common. In Phoenix, 90% of purchased homes were listed at a loss. In Q3 2021 alone, Zillow recorded $421 million in losses, leading to the shutdown of Zillow Offers and a 25% workforce reduction.

Root cause analysis: why it failed

- **Data dependency underestimated.** Zillow's algorithm could not account for hyperlocal variability—neighbourhood-level factors, property-specific conditions, supply-chain delays, and contractor shortages that affect renovation costs and timelines. The data the model consumed was insufficient for the predictions it was asked to make.
- **Concept drift unmanaged.** During the pandemic-era housing boom, the relationships between variables that the model had learned from historical data changed rapidly. This "concept drift"—where the statistical patterns that a model relies on shift over time—rendered the model's predictions systematically wrong. Zillow had no adequate mechanism for detecting or adapting to this drift.
- **Operational decisions automated beyond the model's accuracy envelope.** Purchasing a home is a high-stakes, low-reversibility decision. Using an algorithm with a known 6.9% error rate for off-market properties to make purchasing decisions worth hundreds of thousands of dollars each was a fundamental mismatch between model accuracy and decision stakes.
- **No meaningful human override mechanism.** The system was designed to operate at speed and scale, which required minimising human intervention in purchasing decisions. This removed the human judgment that could have caught the systematic errors accumulating in the portfolio.

Transferable lessons for European organisations

- Data dependency is not just a technical challenge; it is a business risk. Any AI system that informs high-stakes decisions must be evaluated not just for average accuracy but for worst-case error rates and the conditions under which errors become systematic.
- Concept drift is inevitable in any domain where the underlying conditions change over time. AI systems that are deployed without monitoring and adaptation mechanisms will degrade—often silently—as the world shifts beneath them.
- Human oversight is essential for high-stakes decisions, regardless of the efficiency cost. The EU AI Act's requirement for meaningful human oversight in high-risk applications is not bureaucratic caution; it is the codification of a lesson that Zillow learned at a cost of $500 million.
- The appropriate question before deploying AI for operational decisions is not "How accurate is the model on average?" but "What happens when the model is wrong, and can the organisation survive the worst case?"

4.7. Cross-case analysis

The six case studies, viewed together, reveal patterns that reinforce the report's central findings:

Dimension	Successes (Octopus, DBS, Siemens)	Failures (Klarna, IBM Watson, Zillow)
Framing	Augmentation: AI enhances human capability	Automation/replacement: AI substitutes for human judgment
Co-development	Deep involvement of domain experts and frontline workers	Technology-led with minimal user involvement
Data strategy	Unified platforms, domain-specific training data, data-sharing alliances	Narrow training data, poor contextual fit, unmanaged drift
Human oversight	Maintained as structural capability; escalation paths preserved	Minimised for efficiency; eliminated as cost
Measurement	Customer satisfaction, business outcomes, employee experience	Speed and cost metrics without quality controls
Governance	Integrated from design stage; aligned with regulatory expectations	Deferred or absent; compliance treated as afterthought
Workforce strategy	Upskilling, redeployment, capability-building	Headcount reduction, hiring freezes, workforce contraction
Capability maturity	Built deliberately over time; phased progression	Skipped directly to deployment; capabilities assumed rather than built

The pattern is unambiguous. The organisations that succeeded treated AI as a capability to be built through deliberate, phased investment in people, data, governance, and workflows. The organisations that failed treated AI as a product to be deployed for cost reduction, and in each case, the specific failure mode maps directly onto one or more of the four organisational capability gaps identified by Weber et al.: weak project planning (IBM), weak co-development (IBM, Zillow), weak data management (IBM, Zillow), or weak lifecycle management (Klarna, Zillow).

Strategic lessons connecting to the five divergences (Section 5):

Each case study illuminates at least one of the divergences analysed in the Synthesis. Octopus Energy demonstrates "structured improvisation" (Divergence 1)—building capabilities through the work itself rather than in the abstract. DBS Bank exemplifies what closing the measurement vacuum (Divergence 2) looks like: a decade of disciplined metrics enabled AI scaling that most organisations cannot replicate. Siemens shows neuroinclusive design principles operating at industrial scale through its Digital Industries Academy, though it does not name them as such (Divergence 5). Klarna is the canonical case of work intensification without work design (Divergence 4)—initial productivity gains followed by quality erosion, customer backlash, and workforce damage. IBM Watson Health is a $4 billion case study in friction avoidance (Divergence 1)—choosing technically ambitious but organisationally disconnected use cases. And Zillow Offers demonstrates what happens when shadow AI logic (Divergence 3) scales to enterprise level: algorithmic decision-making without the governance infrastructure to contain it.

The case studies do not describe technology failures. They describe organisational failures. The technology, in every case, performed roughly as designed. What failed was the organisational system surrounding it.

5. Synthesis and Gap Analysis: Where Theory Meets Practice

This section is the analytical core of the report. It moves beyond summarising what academics find and what consultancies recommend, to examine where these two bodies of knowledge converge, where they diverge, and—critically—why the gaps exist. Understanding these gaps is essential for CIOs and CTOs because the gaps themselves are often where implementation fails. An organisation that follows academic prescriptions without understanding business constraints will build elegant frameworks that nobody uses. An organisation that follows consultancy playbooks without understanding the research will repeat predictable mistakes with expensive consequences. The synthesis that follows aims to equip European technology leaders with the analytical depth to navigate both.

5.1. Convergence: what both streams agree on

Before examining the tensions, it is worth establishing the shared ground quickly. These convergence points represent the highest-confidence findings—areas where academic research and business practice independently reach the same conclusions. They are summarised here rather than elaborated, because a CIO or CTO reading this report in 2026 has likely encountered most of them. The value lies not in these findings themselves but in the divergences that follow.

Convergence point	Academic evidence	Business evidence	Confidence
Organisational capabilities > algorithmic sophistication	Weber et al.: four implementation capabilities predict success; Uren & Edwards: socio-technical readiness	BCG: Trailblazers distinguished by organisational disciplines; MIT: 95% pilot failure driven by structural, not technical, factors	Very high
Human capital is the highest-leverage investment	European training multiplier: 6× productivity return per percentage point of expenditure	BCG: Trailblazers allocate 60% of AI budget to workforce development; PwC: 56% wage premium for AI-skilled workers	Very high
Readiness is socio-technical	Four-dimensional model: technology, people, process, data; people and process are the binding constraints	"AI transformation is workforce transformation" (BCG, McKinsey, Deloitte)	High
Phased implementation outperforms big-bang deployment	Path Framing → Narrating → Stretching (Holmström & Magnusson)	McKinsey's three-phase workflow evolution; Deloitte's phased maturity model	High

The divergences that follow are where implementation actually succeeds or fails.

5.2. Divergence one: structured capability-building vs. emergent

learning

The first significant divergence concerns **how organisations should develop AI implementation capabilities**.

Academic research prescribes deliberate, structured capability-building. Weber et al. argue that the four AI implementation capabilities must be consciously designed and developed, because they address challenges (inscrutability, data dependency) that are qualitatively different from those in conventional IT. The implication is that organisations should invest in building these capabilities before or alongside AI deployment, through dedicated training, process design, and governance structures.

Business practice tells a different story. In the majority of organisations—particularly SMEs—capabilities emerge through project experience rather than deliberate design. Teams learn by doing. Data management improves because a specific project exposes data quality problems. Co-development practices evolve because a specific model fails in production and forces domain experts and data scientists to communicate more effectively.

Why the gap exists: The academic prescription assumes a degree of organisational slack—time, budget, leadership attention—that most smaller organisations simply do not have. They cannot afford to build capabilities in the abstract before they begin creating value. They need to generate returns from AI quickly enough to justify continued investment, which means learning and capability-building must be embedded in the work itself, not separated from it.

What this means for CIOs/CTOs: The synthesis suggests a "structured improvisation" approach. Do not attempt to build all four capabilities before launching AI initiatives, but do not rely on purely emergent learning either. Instead, select initial use cases that will naturally exercise and develop the weakest capabilities. If data management is the binding constraint, choose a use case that forces the organisation to address data quality. If co-development is weak, choose a use case that requires deep collaboration between technical and domain teams. Use projects as deliberate learning vehicles, not just value-delivery mechanisms.

5.3. Divergence two: the measurement vacuum

The second divergence concerns **how organisations measure AI implementation success**.

Academic frameworks emphasise structured measurement: AI maturity models, readiness assessments, capability audits. The recent gap analysis framework by Ciupek et al. (2025) reveals the scale of the problem: despite 78% of organisations reporting AI adoption, only 1% characterise themselves as "mature" in AI deployment, and fewer than 20% track any AI-specific performance indicators at all.

This is an extraordinary finding. It means that the vast majority of organisations making significant AI investments have no systematic way of knowing whether those investments are working. They can report adoption rates and pilot counts, but they cannot connect AI deployment to business outcomes.

Business consultancies acknowledge this gap but approach it differently. McKinsey recommends establishing clear metrics from the outset—productivity, quality, speed, employee satisfaction, customer outcomes. BCG measures AI maturity through organisational archetypes (Follower, Pragmatist, Trailblazer). But both acknowledge that most organisations lack the measurement infrastructure to evaluate their own AI implementations rigorously.

Why the gap exists: Measurement is hard because AI's impact is rarely attributable to a single system or deployment. Productivity gains may be real but distributed across many small workflow improvements. Cost savings may be offset by increased spending on infrastructure, training, or governance. And the most important benefits—improved decision quality, enhanced employee capability, greater organisational agility—are inherently difficult to quantify.

What this means for CIOs/CTOs: Establish measurement from day one, but be pragmatic about what you measure. At minimum, track three categories: (i) task-level metrics (time saved, error rates, throughput) for specific AI-augmented workflows; (ii) adoption and usage metrics (who is using AI, how often, for what tasks); and (iii) employee experience metrics (satisfaction, trust, perceived value). These three categories provide a reasonable proxy for AI's impact without requiring the sophisticated attribution infrastructure that only large enterprises can afford.

5.4. Divergence three: the shadow AI paradox

The third divergence is perhaps the most practically urgent for CIOs and CTOs: **the gap between official AI strategy and actual AI use**.

Academic literature on AI implementation largely assumes that organisations control which AI tools are used, by whom, and for what purposes. Business reality has thoroughly demolished this assumption. IDC's Global Employee Survey (April 2025) reveals that in EMEA, 39% of employees use free AI tools at work and a further 17% use AI tools they pay for personally—meaning that over half of employees are using AI that their organisation neither provides nor governs. KPMG finds that almost half of employees admit to uploading sensitive company data to unauthorised AI platforms. Menlo Security's 2025 report documents that 68% of employees use free-tier AI tools via personal accounts, with 57% inputting sensitive data.

MIT's research confirms the pattern: even when organisations fail to provide enterprise AI tools, 90% of employees report using personal generative AI at work. This "shadow AI" economy is not a marginal phenomenon; it is the dominant mode of AI adoption in most organisations.

Why the gap exists: The shadow AI phenomenon is driven by a combination of employee initiative, organisational inertia, and perverse incentives. Employees adopt AI because it makes them more productive, and in many cases because they feel their jobs depend on meeting expectations that have been implicitly raised by the existence of AI tools. Organisations fail to provide sanctioned alternatives quickly enough, or the alternatives they provide are less capable, less user-friendly, or more restrictive than consumer tools. The result is a paradox: the organisations with the most "official" caution about AI often have the most ungoverned AI use.

What this means for CIOs/CTOs: Shadow AI cannot be eliminated by prohibition; it can only be channelled. The strategic response is threefold. First, acknowledge that employees are already using AI and treat this as a signal of demand, not a compliance failure. Second, provide sanctioned AI tools that are at least as capable and user-friendly as consumer alternatives—an increasingly achievable standard. Third, establish clear, simple governance guidelines that employees can actually follow: what data can be used with AI, what tools are approved, how outputs should be reviewed, and where to report problems. The goal is to bring shadow AI into the light, not to drive it further underground.

5.5. Divergence four: the AI intensification problem

A fourth tension, barely visible in the academic implementation literature but increasingly prominent in business reality, is the phenomenon of **AI-driven work intensification**.

Harvard Business Review's March 2026 study, based on an eight-month ethnographic investigation at a 200-employee technology company, finds that AI does not reduce workload. It intensifies it. The study documents three mechanisms:

- **Task expansion:** AI makes unfamiliar tasks feel accessible, causing employees to absorb responsibilities outside their traditional roles. Product managers write code. Researchers handle engineering tasks. Workers absorb tasks that would previously have justified additional headcount.
- **Blurred boundaries:** AI's conversational interface and ease of use causes work to spill into breaks, evenings, and mornings. The informality of AI interaction makes work feel less like work, eroding the boundaries that protect recovery time.
- **Cognitive overload through multitasking:** AI accelerates task-switching, causing workers to juggle multiple AI-driven workflows simultaneously. The result is fragmented attention, reduced deep focus, and a paradoxical decline in the quality of decision-making even as the quantity of output increases.

The study concludes that without deliberate "AI Practice"—a set of intentional norms and routines that structure how AI is used, when it is appropriate to stop, and how work should and should not expand—AI will simply amplify existing incentive systems that reward more output, leading to burnout, cognitive fatigue, and weakened judgment.

Why the gap exists: The academic AI implementation literature is oriented toward the challenge of getting organisations to adopt AI; the intensification problem is a consequence of successful adoption. It appears in the gap between what the academic literature studies (implementation) and what organisations experience (the ongoing consequences of implementation). The consultancy literature is beginning to recognise this—McKinsey's emphasis on "reconfiguring work" gestures in this direction—but the intensification problem is not yet adequately addressed in any mainstream implementation framework.

What this means for CIOs/CTOs: AI augmentation strategy must include explicit work design principles that prevent intensification. This means setting boundaries on task expansion (just because AI makes a task possible does not mean it should be added to a role), protecting recovery time (AI availability does not mean employee availability), and monitoring for cognitive overload (tracking not just how much people produce with AI, but how they experience the work). For neurodivergent employees, who may be particularly vulnerable to sensory and cognitive overload, these protections are not optional—they are essential for sustained performance.

5.6. Divergence five: the unmeasured neuroinclusion dividend

A fifth gap—and arguably the one with the greatest untapped strategic value—concerns the treatment of neurodiversity in AI implementation.

The paradox is this: AI tools are being deployed at scale across European workplaces for general productivity reasons, yet the architecture of these tools—structured outputs, reduced executive-function overhead, multiple processing modalities, real-time scaffolding—maps precisely onto the cognitive accommodations that neurodivergent workers have historically lacked. Generative AI may be the most significant neuroinclusion intervention ever deployed, and it has been deployed almost entirely by accident. Nobody designed it for this purpose; nobody is measuring whether it is working for this purpose; and yet the circumstantial evidence is substantial.

Academic research on AI implementation (Sections 3.1–3.5) treats the workforce as cognitively homogeneous. Readiness models assess "people readiness" without considering that individuals process information, uncertainty, and ambiguity in fundamentally different ways. Capability frameworks address inscrutability and data dependency as organisational challenges without acknowledging that explainability requirements are not uniform across cognitive profiles—what constitutes a transparent system for a neurotypical user may be opaque or overwhelming for someone with ADHD, and vice versa.

Business practice is ahead but still fragmented. EY's studies (2024, 2025) now provide the first quantified data on neurodivergent AI experiences—88% self-reported productivity gains, 55% higher adoption rates, 25% higher satisfaction in the UK government pilot. But even with these advances, neuroinclusion remains a separate workstream addressed by HR or DEI functions rather than integrated into AI implementation strategy. The critical gap remains: **the bridge between self-reported neurodivergent satisfaction data and the rigorous "AI as equaliser" productivity studies (Brynjolfsson et al.'s 30–35% bottom-quintile gains, Dell'Acqua et al.'s 22%→4% gap compression) has not been built.** The proxy inference—that AI's equalising effects disproportionately benefit neurodivergent workers whose underperformance was environmentally produced—is highly plausible but unvalidated.

Why the gap exists: The AI implementation literature and the neurodiversity literature developed in parallel academic disciplines (information systems vs. organisational psychology), published in different journals, and read by different audiences. The interdisciplinary connection—that AI tools designed for augmentation simultaneously function as neuroinclusive accommodations, and that neuroinclusive design principles improve AI tools for everyone—has not been made systematically in either body of work. This is a classic case of disciplinary silos suppressing a finding that sits at the intersection.

What would need to be true for the bridge inference to hold—and what could falsify it. The inference rests on three assumptions, each of which is testable: (1) that a material portion of neurodivergent underperformance in traditional metrics is environmentally produced (supported by pre-AI matched-role studies at JPMorgan and HP Enterprise, which show 90–140% ND productivity when friction is removed); (2) that AI tools reduce the specific environmental frictions that suppress ND performance (supported by the EY/Microsoft and UK DBT data on reduced executive-function overhead, improved information access, and higher satisfaction); and (3) that the "equaliser effect" documented by Brynjolfsson et al. and Dell'Acqua et al. operates through the same mechanism—removing friction for those most burdened by it—rather than through a different mechanism such as compensating for lower baseline skill. If assumption (3) is wrong—if AI helps low performers primarily by substituting for missing skills rather than removing environmental barriers—then the extrapolation to neurodivergent workers would be substantially weaker. This is testable, and testing it is precisely the opportunity this report identifies.

Why this matters strategically: This is not merely an academic gap. It is a competitive intelligence opportunity. Any European organisation that conducts even a modest controlled comparison of AI-augmented task performance across cognitive profiles will produce the study that does not yet exist—generating proprietary workforce insights while contributing to an evidence base with direct implications for talent strategy, tool selection, and regulatory compliance under the EU AI Act's human oversight requirements. The practical steps are modest (auditing tools for cognitive accessibility, including neurodivergent employees in pilot cohorts with consent, measuring differential outcomes) but the positioning value—as a thought leader at the intersection of AI implementation and neuroinclusion—is significant, precisely because the intersection is currently empty.

5.7. Organisational scale and the implementation gap

The divergences analysed above have particular salience for European smaller and mid-sized organisations, because they face a distinctive set of constraints that amplify some gaps and reduce others:

Dimension	Large enterprise reality	Smaller organisation reality	Implication
Capability building	Can invest in dedicated AI teams, centres of excellence, and structured programmes	Must build capabilities through project work; limited slack for abstract capability development	"Structured improvisation": select projects that deliberately exercise weak capabilities
Measurement	Can invest in sophisticated attribution and ROI modelling	Must rely on proxy metrics and qualitative assessment	Establish minimum viable measurement from day one; don't let perfect be the enemy of good
Shadow AI	Can deploy enterprise-wide AI platforms with centralised governance	Often cannot provide tools faster than employees find them independently	Channel rather than prohibit; provide good-enough sanctioned tools quickly
Work intensification	Can implement formal work-design policies through HR and compliance functions	Work design is more informal; boundaries are set by culture, not policy	Leadership must model AI boundaries; make "AI Practice" norms explicit and visible
Neurodiversity	Can fund specialist DEI programmes and dedicated accommodations	Must integrate neuroinclusion into general AI design; no budget for separate programmes	Treat neuroinclusion as a default design parameter, not a specialist add-on
Regulation (EU AI Act)	Dedicated legal and compliance teams; regulatory expertise in-house	Must develop regulatory understanding within existing roles; limited access to specialist advice	Use AI Act requirements as design guidelines; leverage provisions for smaller organisations and sandboxes

The central insight of this synthesis is that the gaps between academic prescription and business reality are not random—they are structurally produced by resource constraints, incentive misalignments, and disciplinary silos. The strategic response is not to close every gap simultaneously (which is impossible) but to identify which gaps are most consequential for the organisation's specific context and address them deliberately, while accepting pragmatic imperfection in the others.

5.8. The meta-finding: implementation is the strategy

Both the academic and business literatures converge on a meta-finding that deserves explicit statement: **in AI augmentation, there is no meaningful distinction between strategy and implementation**.

The academic literature frames AI transformation as a strategic journey requiring deliberate capability-building. The business literature frames it as a change management challenge requiring workforce transformation. Both are correct, and both are incomplete in isolation. The synthesis is that AI augmentation is a form of organisational learning in which strategy emerges through disciplined implementation, and implementation is shaped by strategic intent.

For a CIO or CTO, this means that the traditional sequential model—first strategy, then implementation—does not apply. AI strategy is discovered through implementation, refined through feedback, and validated through measurable outcomes. The organisations that wait for a perfect strategy before acting will be overtaken by those that begin implementing with strategic intent and learn their way to excellence.

This is not an argument for recklessness. It is an argument for what the academic literature calls "path framing" and what business practice calls "starting small with strategic ambition." Begin with use cases that matter, build capabilities through the work, measure relentlessly, adapt continuously, and treat every friction point not as a failure but as information about where the organisation needs to grow.

5.9. Technical architecture implications for CTOs

The preceding analysis has focused primarily on organisational capabilities, workforce strategy, and governance—the CIO's natural territory. CTOs reading this report will rightly ask: what does this mean for our technical architecture? This section addresses the engineering and infrastructure implications directly.

Data architecture must support differential measurement. The neurodiversity evidence gap identified in Section 5.6 cannot be closed without data infrastructure that can track AI tool usage, task performance, and user experience at the individual level while maintaining privacy and consent compliance under GDPR. This means designing AI analytics pipelines that capture not just aggregate productivity metrics but disaggregated outcomes—by team, role, and (with consent) cognitive profile. Most enterprise AI platforms do not support this natively. CTOs should evaluate whether their current data architecture can answer the question: "Is this tool working differently for different user populations?" If not, this is an infrastructure investment, not just a policy decision.

Model selection should optimise for interaction diversity, not just benchmark performance. The evidence on explainability (Section 3.7) and cognitive diversity (Section 3.9) implies that model selection criteria should extend beyond accuracy, latency, and cost. The relevant additional criteria include: output customisability (can users control format, verbosity, and modality?), explanation granularity (can the model provide reasoning at different levels of detail?), and interaction flexibility (does the API support different interaction patterns—conversational, structured, visual?). In practice, this means multi-modal models with strong instruction-following capabilities will outperform narrower models for augmentation use cases, even when the narrower model scores higher on domain-specific benchmarks.

Integration architecture determines whether augmentation scales or fragments. The shadow AI problem (Section 5.4) is fundamentally an integration architecture problem. When enterprise AI tools are siloed from the workflows where employees actually work, employees route around them. The technical response is not to lock down endpoints but to invest in middleware and API layers that embed AI capabilities into existing tools—email clients, project management systems, documentation platforms, communication tools. The McKinsey three-phase workflow model (Section 3.3) implies an integration architecture that evolves from Phase 1 (discrete API calls from individual applications) through Phase 2 (orchestration layers coordinating multiple AI agents within a workflow) to Phase 3 (autonomous agents with human oversight hooks). Most organisations should be designing for Phase 2 now while their governance catches up.

EU AI Act compliance is an architecture decision, not a bolt-on. The risk classification system (Section 3.8) has direct implications for how AI systems are designed, deployed, and documented. High-risk systems require audit trails, human oversight mechanisms, data quality assurance, and bias monitoring—all of which must be built into the system architecture from the outset. Retrofitting compliance onto an existing AI deployment is substantially more expensive and less effective than designing for it. CTOs should ensure that their AI platform architecture includes: logging and provenance tracking at the inference level, configurable human-in-the-loop checkpoints, automated bias monitoring for protected characteristics, and documentation generation that maps to EU AI Act requirements.

The build-vs-buy calculus has shifted. Two years ago, the technical advantage lay with organisations that could build custom models. Today, foundation model capabilities have commoditised to the point where the differentiating technical investment is in fine-tuning, prompt engineering, workflow integration, and the data infrastructure described above—not in model training. For most European organisations, the CTO's highest-value contribution is not selecting which model to use but designing the integration and measurement architecture that makes any model effective within the organisation's specific context.

5.10. The conclusions in brief

The five divergences, taken together, yield a set of conclusions that should function as decision criteria for European technology leaders. These are falsifiable claims, not platitudes:

AI augmentation is not a technology category; it is an organisational design paradigm. Treating AI as a procurement decision leads to marginal, fragile gains. Treating it as a redesign of how work is structured, governed, and experienced leads to durable value creation. For European organisations, this is simultaneously sobering (it cannot be delegated to IT) and liberating (the primary determinant is organisational capability, not budget).

Training expenditure has the highest demonstrable return of any AI-related investment. The European 6× training multiplier, BCG's Trailblazer allocation data, and PwC's 56% wage premium all point in the same direction: every euro diverted from technology procurement to structured workforce development delivers substantially higher returns. This is not a soft finding—it is a hard economic conclusion backed by firm-level econometric data.

The European regulatory environment is a net advantage for augmentation strategies. The EU AI Act requires human oversight, explainability, data quality, and transparency. These are precisely the properties that academic research and business practice identify as prerequisites for effective human–AI collaboration. Regulation and good augmentation design are aligned, not in tension.

Neuroinclusion is a strategic multiplier, not a cost centre. The curb-cut effect means that AI features designed for neurodivergent workers—customisable interfaces, multiple modalities, structured task management—improve outcomes for all users. The investment is modest (selection criteria and configuration, not bespoke development); the returns are substantial. In a tight European labour market, neuroinclusive AI augmentation expands the effective talent pool by enabling the 15–20% who are neurodivergent to contribute at their full potential. This is not compassion; it is competitive arithmetic.

AI augmentation without work design is a recipe for intensification, not improvement. The remedy is not to slow AI adoption but to pair it with explicit "AI Practice" norms: boundaries on task expansion, protection of recovery time, monitoring for cognitive overload. For neurodivergent employees susceptible to hyperfocus-and-burnout cycles, these norms are not optional—they are a prerequisite for sustainable augmentation.

The window for competitive positioning is open now—and closing. Eurostat data shows 20% of EU enterprises used AI in 2025; BCG projects companies will double AI spending in 2026. Forrester predicts 2026 marks the end of the hype period and the beginning of ROI-focused scrutiny. The organisations that build augmentation capabilities now will be significantly better positioned than those that wait. The evidence is clear: there is no risk-free path, and waiting is itself a high-risk strategy.

6. Recommendations for European Technology Leaders

The synthesis in Section 5 identifies where theory and practice diverge. This section translates those findings into what to do. The recommendations are designed for European technology leaders—CIOs, CTOs, and heads of digital—operating under EU regulatory frameworks and navigating the resource constraints, talent pressures, and competitive dynamics characteristic of European markets. Each recommendation is actionable and grounded in the evidence reviewed throughout this report. While the principles apply across organisational scales, a planned companion report will translate them into SME-specific implementation guidance through primary research with European mid-market technology leaders.

6.1. Reframe the AI budget: 50/30/20

Allocate approximately 50% of the AI budget to workforce development, 30% to technology and infrastructure, and 20% to governance and measurement.

This ratio inverts the default allocation in most organisations, which typically spends 70–80% on technology and treats training and governance as residual line items. The evidence is unambiguous that this default allocation is wrong. BCG's data shows that Trailblazer organisations allocate the majority of their AI budget to workforce development and achieve dramatically better outcomes than those that prioritise technology spend. The European training multiplier (6× return per percentage point of expenditure) confirms that human capital investment delivers the highest measurable return.

For an organisation investing €500,000 annually in AI, this means:

- **€250,000 on workforce development:** Structured upskilling programmes, protected learning time, AI literacy for non-technical staff, advanced training for technical teams, and neuroinclusive training design.
- **€150,000 on technology and infrastructure:** AI platform licensing, data infrastructure improvements, integration with existing systems, and contextual adaptation of AI tools.
- **€100,000 on governance and measurement:** EU AI Act compliance, use-case classification, human oversight mechanisms, documentation, and a minimum viable measurement framework.

This allocation is not a fixed rule; it should be adjusted based on the organisation's specific maturity. Organisations with very weak data infrastructure may need to temporarily increase the technology share. But the principle holds: workforce development should never be the residual category.

6.2. Select initial use cases using the "strategic friction" principle

Choose first use cases that are strategically important and that will force the organisation to develop its weakest capabilities—not use cases that are easy to implement but irrelevant to core operations.

MIT's research demonstrates that pilot purgatory is caused primarily by friction avoidance: organisations select use cases specifically to minimise disruption, which produces pilots that succeed on their own terms but never scale because they were never connected to the organisation's real work.

The counter-intuitive recommendation is to embrace strategic friction. Select use cases where the workflow is **strategically important** (not a back-office curiosity), AI augmentation requires **cross-functional collaboration** (forcing co-development capabilities to develop), the **data challenge is real but bounded**, **employees are motivated** to participate, and the outcome is **measurable**.

For a typical European organisation, strong initial candidates include customer support or sales operations (high visibility, clear metrics, motivated employees), product development or engineering workflows (AI-literate workforce, high strategic value), and financial planning and analysis (data-intensive, decision-critical, connected to leadership visibility).

6.3. Design for cognitive diversity from day one

Treat neuroinclusion as a default design parameter in all AI tool selection, configuration, and workflow design decisions.

The evidence shows that neuroinclusive AI design benefits the entire workforce while unlocking the distinctive strengths of the 15–20% who are neurodivergent.

Practically, this means: **when evaluating AI tools**, add cognitive accessibility to the selection criteria—does the tool support multiple modalities, customisable interfaces, structured task decomposition? **When redesigning workflows**, include neurodivergent employees in the design process; their experience of friction points reveals issues that affect everyone. **When designing training**, offer multiple formats and learning speeds. **When monitoring AI impact**, track cognitive load and wellbeing alongside productivity—AI that increases output while increasing stress is intensification, not augmentation.

The measurement opportunity is significant: any organisation that includes cognitive profile as a variable in its AI pilot measurement will generate data that does not yet exist in the research literature. This is a modest investment with disproportionate strategic and reputational returns.

6.4. Implement "AI Practice" norms to prevent work intensification

Establish explicit organisational norms that govern how AI is used, when it is appropriate to stop, and how work should and should not expand as AI capabilities grow.

The evidence from HBR's 2026 ethnographic study is that AI augmentation without work design produces intensification, not improvement. The recommendation is to implement "AI Practice" norms—explicit agreements that structure the organisation's relationship with AI tools:

- **Task boundaries:** Define which tasks AI should be used for and which it should not. Just because AI makes a task possible does not mean it should be added to a role. Resist the pull of "task creep."
- **Time boundaries:** AI availability does not mean employee availability. Establish clear expectations about when AI-assisted work should and should not happen—particularly outside core working hours.
- **Quality over quantity:** Make explicit that the goal of AI augmentation is better work, not more work. Measure outcomes by quality and impact, not volume of output.
- **Cognitive load monitoring:** Train managers to recognise signs of AI-driven cognitive overload—fragmented attention, excessive multitasking, declining quality despite increasing output—and to intervene before burnout occurs.
- **Neurodivergent-specific protections:** Ensure that AI Practice norms explicitly address the risks of hyperfocus-and-burnout cycles that neurodivergent employees may experience when AI tools make deep-focus work feel frictionless but unsustainable.

6.5. Address shadow AI through enablement, not prohibition

Acknowledge that employees are already using AI, provide sanctioned alternatives that are at least as capable, and establish simple governance guidelines.

IDC's data shows that over half of EMEA employees use unsanctioned AI tools at work, with nearly half admitting to uploading sensitive company data to these platforms. Prohibition does not work—it drives shadow AI further underground and creates ungoverned risks.

The recommended approach: (1) **Audit current shadow AI use** through anonymous surveys—understand which tools, for what purposes, and why. (2) **Provide sanctioned alternatives** that match or exceed the capability and usability of consumer AI tools. (3) **Publish simple governance guidelines** specifying what data may and may not be used, which tools are approved, and how AI outputs should be validated. (4) **Create a positive feedback loop** by celebrating employees who identify productive uses for sanctioned tools, rather than punishing prior unsanctioned use.

6.6. Foundation practices

The following four practices are well-established in the implementation literature and should be treated as foundational infrastructure. They are presented here in compressed form because the principles are widely understood; the challenge is execution, not awareness.

Practice	Core action	Key principle
Readiness assessment	Assess readiness across four dimensions (technology, people, process, data) before selecting any AI vendor. Ask five honest questions about AI literacy, workflow clarity, data quality, infrastructure, and regulatory classification.	The binding constraints are almost never technological. Start with the weakest dimension.
EU AI Act compliance by design	Classify every use case under the EU AI Act risk framework at the design stage. Build human oversight, explainability, and documentation into the design process, not the approval process. Leverage sandbox provisions.	Retrofitting compliance costs 5–10× more than building it in. Compliance-driven design produces better systems.
Minimum viable measurement	Track three categories from day one: task-level performance (time, errors, throughput), adoption and usage (active users, breadth), and employee experience (trust, cognitive load, satisfaction). Establish baselines before deployment.	Only 1% of AI-adopting organisations are "mature" in measurement. Without baselines, you cannot distinguish success from slow failure.
Cross-functional governance	Create a small (3–7 person) cross-functional AI task force including technology, business, HR/people, and domain leads—plus employee representation. Do not create a separate AI department.	AI expertise concentrated in a silo creates dependencies and distances AI from domain knowledge.

6.7. Summary of recommendations

#	Recommendation	Distinctiveness	Timeline	Expected impact
6.1	Reframe AI budget: 50/30/20	Evidence-backed reallocation; inverts industry default	Immediate	Highest-ROI allocation
6.2	Select use cases by "strategic friction"	Counter-intuitive; directly addresses 95% pilot failure rate	Months 1–2	Avoids pilot purgatory
6.3	Design for cognitive diversity from day one	Novel in AI implementation literature; generates proprietary data	Ongoing	Improves outcomes for all; first-mover positioning
6.4	Implement "AI Practice" norms	Addresses emerging intensification risk most frameworks ignore	Months 2–3	Prevents burnout; sustainable augmentation
6.5	Address shadow AI through enablement	Reframes dominant prohibition approach	Months 1–3	Brings ungoverned AI into managed framework
6.6	Foundation practices (readiness, compliance, measurement, governance)	Well-established; challenge is execution	Month 1 onwards	Foundational infrastructure

About the Author

Michael Eriksson is the Founder and CEO of Liminal Discovery, a digital and AI transformation consultancy based in Europe. With twenty years of experience leading digital businesses, including serving as President of Europe and Asia for DMI, a global digital transformation consultancy, he has overseen the design and delivery of large-scale, end-to-end transformation programmes for enterprises across multiple industries.

For the past four years, Michael has dedicated his work to exploring how AI can augment organisations and their people to generate measurable business impact. He sees AI as a key technology in the continuation of digital transformation efforts, and his work sits at the intersection of AI strategy, workforce transformation, and neuroinclusion, the core themes of this report. He is also a guest lecturer at CREA Genève, teaching AI, digital marketing, and sustainable business models.

Email michael@liminaldiscovery.com
LinkedIn linkedin.com/in/michaeleriksson

About Liminal Discovery

Liminal Discovery is a European consultancy specialising in human-first AI implementation. The firm works with organisations navigating the transition from experimentation to strategic AI adoption, with a core conviction: AI should augment people and organisations, not replace them.

Liminal Discovery supports technology leaders in designing AI strategies that build organisational capability, respect regulatory requirements, including the EU AI Act, and deliver sustainable business outcomes. The consultancy draws on deep expertise in digital transformation, workforce enablement, and the emerging intersection of AI and neuroinclusion.

www.liminaldiscovery.com

7. Appendix: Glossary of Key Terms and Abbreviations

This glossary defines the key terms, abbreviations, and concepts used throughout this report. It is designed for CIOs, CTOs, and senior technology leaders who are familiar with general technology concepts but may not have deep specialist knowledge of AI, neurodiversity, or EU regulatory terminology. Terms are organised into four categories: AI and technology terms, organisational and implementation terms, neurodiversity and cognitive diversity terms, and regulatory and institutional terms.

7.1. AI and technology terms

Term	Definition
Agentic AI	AI systems designed to autonomously pursue complex goals and multi-step workflows with limited direct human supervision. Agentic AI can plan, execute, and adapt sequences of actions, distinguishing it from reactive AI that responds to single prompts. Corresponds to Phase 3 of McKinsey's workflow evolution model (Section 4.2).
AI model	A mathematical system trained on data to recognise patterns, make predictions, or generate outputs. Models range from simple classifiers to large language models with billions of parameters.
Algorithm	A set of rules or instructions that a computer follows to solve a problem or perform a task. In the AI context, algorithms are the mathematical procedures used to train and run models.
API (Application Programming Interface)	A standardised interface that allows different software systems to communicate with each other. APIs are the primary mechanism through which organisations access cloud-based AI services.
Concept drift	The phenomenon where the statistical relationships a model has learned from historical data change over time, causing the model's predictions to degrade. Concept drift was a critical factor in the Zillow Offers failure (Case Study 6).
Copilot	A category of AI assistant designed to work alongside a human user within a specific application or workflow, providing suggestions, drafting content, or automating subtasks while the human retains control. Examples include Microsoft Copilot and GitHub Copilot.
Digital twin	A virtual replica of a physical asset, process, or system that is continuously updated with real-world data. Used in manufacturing and infrastructure for simulation, optimisation, and predictive maintenance. Referenced in the Siemens case study (Case Study 3).

Fine-tuning The process of adapting a pre-trained AI model to a specific domain or task by training it on additional, domain-specific data. Fine-tuning enables generic models to perform well in specialised contexts.

Foundation model A large AI model trained on broad data at scale that can be adapted to a wide range of downstream tasks. Examples include GPT-4, Claude, Gemini, and Llama. Foundation models are the basis for most current generative AI applications.

Generative AI (GenAI) AI systems that can generate new content—text, code, images, audio, video—rather than simply classifying or predicting based on existing data. The current wave of AI adoption is driven primarily by generative AI capabilities.

Hallucination A phenomenon where an AI model generates outputs that are plausible-sounding but factually incorrect or fabricated. Hallucination is a direct consequence of the inscrutability challenge described in Section 3.1.

Human-in-the-loop (HITL) A design pattern in which a human operator reviews, validates, or overrides AI outputs before they take effect. Required by the EU AI Act for high-risk AI systems.

Inscrutability The challenge posed by the probabilistic and often opaque nature of AI outputs. Unlike deterministic software, AI models produce approximate outputs whose reasoning may be difficult to explain or predict. One of the two fundamental AI properties identified by Weber et al. (Section 3.1).

Large language model (LLM) A type of foundation model trained on large volumes of text data, capable of understanding and generating human language. LLMs power most current generative AI applications, including chatbots, writing assistants, and code generators.

Machine learning (ML) A subset of AI in which systems learn patterns from data rather than being explicitly programmed. Machine learning encompasses supervised learning, unsupervised learning, and reinforcement learning.

Natural language processing (NLP) The branch of AI concerned with enabling computers to understand, interpret, and generate human language.

Predictive maintenance The use of AI and sensor data to predict when equipment will fail, enabling maintenance to be performed before failure occurs. Reduces unplanned downtime and extends asset life. Referenced in the Siemens case study (Case Study 3).

Prompt engineering The practice of crafting input text (prompts) to elicit desired outputs from generative AI models. A key skill for employees working with LLM-based tools.

RAG (Retrieval-Augmented Generation) An architecture that enhances LLM outputs by first retrieving relevant information from external knowledge sources, then using that information to generate more accurate, grounded responses. RAG reduces hallucination and enables AI to work with organisation-specific data without retraining the model. glean

SHAP (Shapley Additive Explanations)	A method from explainable AI (XAI) that attributes the contribution of each input feature to a model's output, enabling humans to understand why a model made a specific prediction. Referenced in the organisational agility research (Section 3.7).
Shadow AI	The use of unsanctioned AI tools by employees without organisational knowledge, approval, or governance. Analogous to "shadow IT." Identified as a major governance challenge in Section 5.4.
XAI (Explainable AI)	A set of techniques and design principles aimed at making AI model outputs understandable to humans. XAI is both a research field and a regulatory requirement under the EU AI Act for high-risk systems.

7.2. Organisational and implementation terms

Term	Definition
AI augmentation	The use of AI to enhance human capabilities rather than replace them. Augmentation positions AI as a tool that makes employees more effective, not a substitute that makes employees unnecessary. The central concept of this report.
AI automation	The use of AI to perform tasks previously performed by humans, without ongoing human involvement. Distinct from augmentation in that the human is removed from the task rather than enhanced.
AI implementation gap	The persistent difference between AI's demonstrated potential in controlled experiments and its actual impact when deployed in organisational settings. The gap is primarily organisational, not technological.
AI literacy	The ability to understand what AI can and cannot do, to interact effectively with AI systems, and to evaluate AI outputs critically. A prerequisite for effective human–AI collaboration at all organisational levels.
AI maturity model	A framework that describes progressive stages of AI capability within an organisation, typically ranging from ad hoc experimentation to fully integrated, strategically governed AI operations.
AI Practice	A set of intentional norms and routines that structure how AI is used within an organisation—including boundaries on task expansion, time limits, and quality standards. Proposed as a remedy for AI-driven work intensification (the synthesis).
Change management	The structured approach to transitioning individuals, teams, and organisations from a current state to a desired future state. In AI contexts, change management addresses the human and organisational dimensions of technology adoption.
Co-development	The integration of domain expertise, technical knowledge, and workforce perspectives throughout the AI development process. One of four critical organisational capabilities identified by Weber et al. (Section 3.1).

Complementarity (human–AI)	The principle that humans and AI have different strengths, and that combining them produces better outcomes than either alone. The labour market evidence shows complementarity effects are 1.7× larger than substitution effects (Section 3.5). arxiv
Contextual intelligence	AI capability that is adapted to an organisation's specific data, workflows, decisions, and operating environment. Distinguished from generic AI capability that works in controlled settings but fails in real-world contexts. The 5% of successful AI pilots build contextual intelligence; the 95% that fail do not.
Cross-functional AI task force	A small team drawn from technology, business, HR, and domain functions, tasked with coordinating AI initiatives. Recommended as an alternative to dedicated AI departments for SMEs (Recommendation 7.7).
Curb-cut effect	The phenomenon where accommodations designed for specific populations benefit everyone. Named after the observation that pavement curb cuts, designed for wheelchair users, also benefit cyclists, parents with pushchairs, and delivery workers. Applied in this report to neuroinclusive AI design (Section 3.8).
Data dependency	The fundamental reliance of AI performance on the quality, volume, representativeness, and timeliness of input data. One of the two distinctive properties of AI identified by Weber et al. (Section 3.1).
Path Framing / Narrating / Stretching	The three phases of the AI Transformation Framework proposed by Holmström and Magnusson (2025). Path Framing defines "what" (strategy), Path Narrating defines "when" (sequencing), and Path Stretching defines "how" (scaling). (Section 3.2).
Pilot purgatory	The state in which an organisation runs multiple AI pilots that may succeed technically but never scale to production or deliver measurable business impact. MIT research finds 95% of generative AI pilots remain in this state.
Regility	A combination of resilience and agility in organisational design. AI can enhance regility when systems are explainable and support decentralised decision-making, but can undermine it when AI centralises control or creates opaque dependencies (Section 3.7).
Strategic friction	The deliberate selection of AI use cases that are strategically important and that force the organisation to develop its weakest capabilities—as opposed to selecting easy use cases that avoid organisational disruption (the strategic friction recommendation).
Work intensification	The phenomenon where AI increases the volume, pace, and scope of work performed by employees, leading to cognitive overload, blurred work–life boundaries, and burnout. Documented by HBR (2026) and addressed in the synthesis.

7.3. Neurodiversity and cognitive diversity terms

Term	Definition
ADHD (Attention Deficit Hyperactivity Disorder)	A neurodevelopmental condition characterised by patterns of inattention, hyperactivity, and/or impulsivity. Individuals with ADHD often exhibit strengths in creative thinking, hyperfocus on topics of interest, and rapid ideation. AI tools that support task structuring, prioritisation, and focus management can significantly benefit ADHD professionals.
Autism spectrum condition (ASC)	A neurodevelopmental condition characterised by differences in social communication, sensory processing, and patterns of thought and behaviour. Autistic individuals often bring strengths in pattern recognition, systematic thinking, attention to detail, and deep focus. Also referred to as autism spectrum disorder (ASD) in clinical contexts.
Cognitive diversity	The variation in how individuals perceive, process, and respond to information. Cognitive diversity encompasses both neurodivergent and neurotypical variation and, when leveraged effectively, improves team problem-solving, innovation, and decision quality. lexxic
Cognitive prosthetic	A tool that compensates for specific cognitive challenges while preserving the user's broader capabilities. In this report, AI tools that reduce executive-function friction for neurodivergent workers function as cognitive prosthetics (Section 3.8).
Dyscalculia	A neurodevelopmental condition affecting the ability to understand and work with numbers and mathematical concepts.
Dyslexia	A neurodevelopmental condition that primarily affects reading, writing, and spelling. Individuals with dyslexia often bring strengths in visual-spatial reasoning, creative thinking, and holistic pattern recognition. AI tools that provide text-to-speech, dictation, and visual representations of written content can reduce dyslexia-related friction.
Dyspraxia (Developmental Coordination Disorder)	A neurodevelopmental condition affecting motor coordination, spatial awareness, and sometimes organisation and planning. AI tools that support task sequencing and provide structured workflows can benefit individuals with dyspraxia.
Executive function	The set of cognitive processes—including working memory, flexible thinking, and self-control—that enable planning, organising, prioritising, and completing tasks. Many neurodivergent individuals experience differences in executive function that are not reflective of intelligence or capability but create friction in conventionally designed workplaces.
Hyperfocus	A state of intense, sustained concentration on a single task or topic, commonly experienced by individuals with ADHD and autism. Hyperfocus can be a significant strength in knowledge-intensive work, but without boundaries, it can lead to exhaustion and neglect of other responsibilities. AI Practice norms (Section 7.8) should account for hyperfocus patterns. lexxic

Neuroinclusion	The practice of designing workplaces, tools, processes, and communications to be accessible and effective for people with all cognitive profiles—not just those who are statistically typical. Neuroinclusion goes beyond accommodation of diagnosed conditions to designing for cognitive diversity as a default. lexxic
Neurodivergent	Describes an individual whose neurological functioning differs from the statistical norm. Neurodivergent conditions include ADHD, autism, dyslexia, dyspraxia, dyscalculia, and Tourette syndrome, among others. An estimated 15–20% of the general population is neurodivergent.
Neurotypical	Describes an individual whose neurological development and functioning fall within the statistical norm. The term highlights that "typical" is a statistical category, not a qualitative standard.
Neurotypical tax	An informal term used in this report to describe the additional cognitive effort that neurodivergent workers must expend to navigate workplace systems—communication norms, information formats, process structures—that were designed for neurotypical cognitive profiles.
Sensory overload	A state of overwhelm caused by excessive sensory input, commonly experienced by autistic individuals and others with sensory processing differences. AI tools that simplify interfaces, reduce notification frequency, and offer "focus mode" settings can mitigate sensory overload in digital workplaces.

7.4. Regulatory and institutional terms

Term	Definition
Conformity assessment	A formal evaluation process required under the EU AI Act for high-risk AI systems, verifying that the system meets all applicable requirements before it can be placed on the market or put into service. gdprlocal
Data governance	The framework of policies, processes, and responsibilities that ensures data is managed consistently, securely, and in compliance with applicable regulations. Data governance is both a regulatory requirement (GDPR, EU AI Act) and an organisational capability essential for AI performance.
EU AI Act (Regulation (EU) 2024/1689)	The European Union's comprehensive regulation on artificial intelligence, entered into force on 1 August 2024 with phased compliance obligations through 2027. Establishes a risk-based classification system for AI, with obligations ranging from outright prohibition (unacceptable risk) to voluntary codes of conduct (minimal risk).
GDPR (General Data Protection Regulation)	The EU's regulation on data protection and privacy (Regulation (EU) 2016/679), applicable since May 2018. GDPR governs the processing of personal data and is directly relevant to AI systems that use personal data for training or inference.

High-risk AI system	Under the EU AI Act, an AI system classified as posing significant risks to health, safety, or fundamental rights. Includes AI used in employment and worker management, education, access to essential services, law enforcement, and critical infrastructure. Subject to stringent requirements for data quality, transparency, human oversight, robustness, and documentation.
Human oversight (EU AI Act)	The requirement that high-risk AI systems be designed to allow effective oversight by natural persons during the period of use. Oversight includes the ability to fully understand the AI system's capabilities and limitations, to properly monitor its operation, and to decide not to use the system or to override its output.
Limited-risk AI system	Under the EU AI Act, an AI system that poses limited risk and is subject primarily to transparency obligations. Users must be informed that they are interacting with an AI system. Includes chatbots and AI-generated content.
Minimal-risk AI system	Under the EU AI Act, an AI system that poses minimal or no risk and is subject to no specific regulatory obligations. The majority of current AI applications fall into this category. Voluntary codes of conduct may apply.
Regulatory sandbox	A controlled environment established by a national regulatory authority where AI systems can be tested under regulatory supervision before full market deployment. The EU AI Act mandates that Member States establish at least one regulatory sandbox. SMEs are given priority access in several jurisdictions.
Risk-based classification	The EU AI Act's approach of assigning regulatory obligations in proportion to the potential harm posed by an AI system. Four categories: unacceptable risk (prohibited), high risk (strict compliance), limited risk (transparency obligations), and minimal risk (no specific obligations).
Transparency obligation	A requirement under the EU AI Act that providers and deployers of certain AI systems inform users that they are interacting with AI, disclose the AI system's capabilities and limitations, and provide adequate documentation. Applies to high-risk and limited-risk systems.
Unacceptable-risk AI system	Under the EU AI Act, an AI system deemed to pose an unacceptable threat to fundamental rights. Includes social scoring systems, manipulative AI, exploitation of vulnerable groups, and certain real-time biometric identification uses. These systems are prohibited.
Works council	An employee representation body, common in many European countries (particularly Germany, the Netherlands, Austria, and Scandinavian nations), with legal consultation and co-determination rights on matters including the introduction of new technologies. Works councils are relevant to AI implementation because they may have formal rights to be consulted on AI deployment affecting workers.

7.5. Abbreviations reference

Abbreviation	Full term
ADHD	Attention Deficit Hyperactivity Disorder
AI	Artificial Intelligence
API	Application Programming Interface
ASC	Autism Spectrum Condition
BCG	Boston Consulting Group
CIPD	Chartered Institute of Personnel and Development
CIO	Chief Information Officer
CTO	Chief Technology Officer
DEI	Diversity, Equity, and Inclusion
EHR	Electronic Health Record
EU	European Union
GDPR	General Data Protection Regulation
GenAI	Generative Artificial Intelligence
HITL	Human-in-the-Loop
HBR	Harvard Business Review
HR	Human Resources
HRM	Human Resource Management
LLM	Large Language Model
ML	Machine Learning
MIT	Massachusetts Institute of Technology
NLP	Natural Language Processing
P&L	Profit and Loss
PwC	PricewaterhouseCoopers
RAG	Retrieval-Augmented Generation
ROI	Return on Investment
SHAP	Shapley Additive Explanations
SME	Small and Medium-Sized Enterprise
TRL	Technology Readiness Level
XAI	Explainable Artificial Intelligence