

The agent does the junior's work, not the apprenticeship

In the short term, everything works. In the long term, no one is left who can take back control.

An engineer signing what they did not write

Inside a team, an engineer validates, day after day, what an agent has produced. They sign, the code ships, the file moves forward. Nobody holds it against them: everything works. Except they never wrote that part of the chain themselves, nor the part before, nor the part after. They are not supervising, they are trusting. And above them, those who could still verify are thinning out, because they too, by now, have not done it.

That is the subject, and it is simpler than the concepts usually attached to it. The problem is not that AI decides badly. The problem is that an organisation can keep deciding correctly while gradually ceasing to manufacture the people capable of taking back control when the tool steps outside its frame. Junior training is the place where this shows up first, but the stake is the continuity of the entire chain.

How an organisation stops manufacturing its supervisors

The mechanism is slow, and that is why it goes unseen. It unfolds in a few steps, each of them perfectly reasonable.

Agents first absorb the entry-level tasks: research, the first draft, debugging, data reconciliation. These are precisely the tasks on which juniors used to learn the real craft. The junior therefore touches the real chains less. Onboarding (when it still happens) no longer quite produces autonomy, it produces familiarity with a tool. In the short term, nothing raises an alarm: productivity rises, juniors look augmented, and a few seniors compensate by validating faster.

Then those seniors get overloaded, and we end up relieving them in turn through assisted validation. When one of them leaves, we do not replace them by manufacturing a successor, we buy one on the market. Step by step, the proportion of people who have actually walked through the chain they validate keeps decreasing. Supervision does not disappear all at once: it becomes nominal. We still sign, but we redo less and less.

The labour market gives the first sign of this tilt. In France, APEC measures a decline in junior executive hiring of roughly 16 % in 2025, after roughly 19 % in 2024, concentrated in IT, engineering and consulting, which happen to be the sectors that used to train the most inexperienced profiles; consulting itself contracts by about 10 % (Syntec Conseil). The same movement is visible across the Atlantic, where demand for seniors has surged while junior hiring has shrunk (Revelio Labs). One caveat is in order: APEC primarily attributes this decline to the macroeconomic cycle, not to AI. These figures therefore do not prove the mechanism, they reveal the terrain: organisations that already buy their seniority rather than producing it.

Sources: APEC, executive recruitment 2025; Syntec Conseil, consulting market 2024–2025; Revelio Labs.

Where the question matters, and where it does not

Before going further, the scope has to be drawn. In many activities, growth, marketing, content, support, an organisation can perfectly well accept that it will never be able to redo everything itself, and it will often be right: error there is small, frequent, reversible, and speed beats control. The reasoning that follows holds fully only where small errors are expensive and unrecoverable: health, defence, cybersecurity, systemic finance, heavy industry, critical infrastructure, regulated systems. This note speaks to those systems. Elsewhere, it describes a risk that many will be entitled to ignore.

What the junior was building without anyone noticing

It must be said without idealising the old days: those entry-level tasks also produced fake seniors, surface competence, seniority mistaken for judgement. Ten years in the chain is not in itself a guarantee of understanding. But, for many, they produced one precise and costly thing: the capacity to sense that a result rings false before it becomes an incident, and to take back control when needed. That capacity is deposited by doing, by failing, by watching a senior hesitate, by paying once the price of an error. It was the slow way an organisation manufactured people capable, later, of judging and correcting other people's work.

The only question that counts

Hence the question, and it is less obvious than it looks. Supervising does not necessarily mean redoing everything. Sometimes a partial reconstruction suffices, sometimes a statistical challenge, sometimes a cross-check by others. What matters, in the end, is that some way of taking back control remains: individual, collective, or mixed. The question is therefore not "are we keeping people who can redo everything alone", but "are we keeping, by one means or another, the capacity to take back control when the tool gets

it wrong or steps outside its frame". Does an organisation remain capable of sustaining that capacity, as it automates its lower layers? This question applies at every scale: the individual, the team, the organisation, society, and indeed civilisation.

The junior capitalises, the agent is rented

The question needs to be pushed a notch further, because it is also economic. A junior is "expensive" at first, then changes nature: they learn, become autonomous, transmit, and supervise in their turn. The initial outlay turns into capital, an asset that stays in the house and whose marginal cost decreases over time.

The agent, by contrast, does not capitalise in that way. It is re-rented at every request. It depends on an infrastructure that is not owned: compute, a vendor, a closed model, a roadmap decided elsewhere, an inference price that fluctuates. And its cost rises precisely when one asks of it what this note demands: context, traceability, supervision, the capacity to take back control. *Contestable cognition is not cheap cognition.*

That last sentence is a thesis in itself, and it will not be treated here: its gradients, its insurance trade-offs, its cost models call for a more economic text than this one. Three things should be retained.

1. The human is not a naturally virtuous form of capital: many people capitalise next to nothing, some skills age fast, and certain AIs do mutualise massive costs.
2. The asymmetry is a tendency, not a law. Some organisations will choose rental knowingly, more flexible, lighter, and the problem is not that they are wrong, but that they rarely make that trade-off explicitly.
3. Finally, whoever no longer produces their own cognition becomes dependent on it, and the loss of capital becomes a loss of sovereignty.

When assistance precedes autonomy

This text has so far assumed humans trained the old way, then assisted. A deeper rupture is approaching with newcomers who will never have known research without automatic synthesis, writing without completion, debugging without a copilot, or the slow construction of a long line of reasoning.

The problem is then no longer only transmission inside the organisation. It touches the foundation: what capabilities does a human system stop internalising when their cognitive cost is permanently externalised?

This must be said carefully, because the terrain is mined. The claim is not that a generation would think less well, which would be both false and easy. The claim is a structural displacement: educational and professional architectures may optimise

assisted orchestration before autonomous reconstruction. One then becomes excellent at piloting a tool, and weaker at taking back control without it. Organisations externalise their operational cognition; individuals externalise their reconstructive cognition; and each move reinforces the other. The question remains measurable: above what level of permanent assistance does a system stop producing people capable, if need be, of functioning without it? The AF 447 Rio-Paris crash is a tragic illustration.

Why no one sees it coming, and what reveals it

None of this is visible in the short term, and that is the trap. Productivity rises, costs fall, juniors look augmented, the seniors already in place compensate, the dashboards are green. The bill, if it comes, arrives five to ten years later, when not enough people, nor enough devices, are left to take back control.

The real test is therefore not performance when everything is fine, it is degraded mode. A vendor outage, a cloud disruption, a security incident, an unavailable context, a regulatory constraint, an isolated environment, a cyberattack: the day assistance is missing or out of frame, one suddenly discovers who can still function without it. That is precisely what is never measured. For individuals as for organisations, the same lure: smooth and fast, until the first off-script event.

The serious objection

One could answer that we will redo the apprenticeship differently, through simulation. Medicine has institutionalised this: since the 2012 report of the French National Authority for Health (HAS), the rule is "never the first time on a patient", and more than a hundred centres now train gestures and critical procedures on simulators before reality, with benefits established by systematic reviews. Simulation does therefore train. But look at what it trains: situations one knew how to script because one had foreseen them. There are excellent surgeons performing robotic surgery who find themselves lost the moment the robot arms freeze inside the patient's abdomen and they have to revert to the ancient scalpel to perform an extended laparotomy and extract the arms. The scripted crisis, yes; the drift no one thought to script, no. The question then becomes who spots it, and nothing is settled in advance: certain drifts, statistical, weak, multidimensional, will be better seen by adversarial architectures or multi-agent monitoring than by any human; others will not. No one has a monopoly on off-frame vigilance.

Sources: HAS, "never the first time on a patient"; systematic review, healthcare simulation and patient outcomes.

None of this is written in advance

If all these mechanisms pointed the same way, this text would describe a fate. It is not one. The gains are real, and silencing them would be dishonest: externalised cognition can bring statistical robustness, mutualise scarce skills, accelerate, raise the average level of quality, open access to knowledge previously reserved to few. An organisation can lose reconstructive autonomy and gain a great deal elsewhere; the trade-off only makes sense if both columns are written down.

And the capacity to take back control need not remain individual. Supervision can be collective, distributed, redundant, and sometimes more robust that way than when carried by a single brain. No individual redoes the entire chain, and the system still holds, sometimes better. The condition is always the same, and it is the only one that counts: that the system, by one means or another, retains the capacity to challenge and to take back control, not merely to produce. The risk described here is not a mechanical effect of AI. It is the default option when no one chooses the other.

One possible answer, among others

The other option has a shape, and it is a working hypothesis of the Twingital Institute, not a closed solution. It consists in deliberately maintaining two cognitive modalities, the human and the artificial, designed so that each corrects the other's weaknesses, and in keeping in parallel two pipelines rather than one: an AI pipeline equipped with guardrails, supervision, traceability and takeover conditions, and a long-term pipeline of human skill development, the very one that the mechanism described here is allowed to silt up. The "port of promotion" is nothing other than the name of that second pipeline.

It deserves the same suspicion as any solution that fits its problem a little too snugly. Opposing a human pipeline to an AI pipeline is convenient for thinking, but real systems hybridise, interweave, co-produce; in the end, taking back control will mean less reverting to the human and more reconfiguring a distributed cognition. The two-modality model is a starting point, not an end, and it is not the only conceivable answer.

It is also a terrain, not only a principle. Within the TweenMe team, young data scientists train under the guidance of seniors while AI serves as a production assistant in a reasoned way: the two modalities coexist by construction, the tool remaining an assistant rather than a substitute for the judgement one is trying to form. Like any implementation terrain, it shows that such an architecture can stand up, without by itself constituting a general proof.

What an executive should be asking

Most organisations steer revenue, margin, productivity. In systems where small errors are expensive, very few know how to answer a handful of concrete questions: how many of our seniors did we manufacture, against how many did we buy; what fraction of the chain has the person who validates actually walked through; if we cut the tool tomorrow, who would know how to take back control; and are we genuinely maintaining two modalities, or one disguised as two. These are not HR indicators, they are governability indicators.

A decision can be wrong; that is correctable. What is not correctable is the day the system steps outside its frame and no one, neither human nor device, has kept what it takes to take back control.