

L'agent fait le travail du junior, pas ses classes

À court terme, tout fonctionne. À long terme, plus personne ne sait reprendre la main.

Un ingénieur qui signe ce qu'il n'a pas fait

Dans une équipe, un ingénieur valide chaque jour ce qu'un agent a produit. Il signe, le code part, le dossier avance. Personne ne le lui reproche : tout fonctionne. Sauf qu'il n'a jamais écrit lui-même cette partie de la chaîne, ni celle d'avant, ni celle d'après. Il ne supervise pas, il fait confiance. Et au-dessus de lui, ceux qui pourraient encore vérifier se raréfient, parce qu'eux non plus, désormais, ne l'ont pas faite.

Voilà le sujet, et il est plus simple que les concepts qu'on lui accole. Le problème n'est pas que l'IA décide mal. Le problème est qu'une organisation peut continuer à décider correctement tout en cessant, peu à peu, de fabriquer les gens capables de reprendre la main quand l'outil sort de son cadre. La formation des juniors est l'endroit où cela se voit en premier, mais l'enjeu est la continuité de toute la chaîne.

Comment une organisation cesse de fabriquer ses superviseurs

Le mécanisme est lent, et c'est pour cela qu'on ne le voit pas. Il se déroule en quelques pas, chacun parfaitement raisonnable.

Les agents absorbent d'abord les tâches d'entrée : la recherche, la première version, le débogage, la réconciliation de données. **Ce sont précisément celles sur lesquelles les juniors apprenaient le métier réel. Le junior accède donc moins aux chaînes réelles. L'onboarding (quand il peut avoir lieu) ne produit plus tout à fait de l'autonomie, il produit de la familiarité avec un outil.** À court terme, rien n'alarme : la productivité monte, les juniors paraissent dopés, et quelques seniors compensent en validant plus vite.

Puis on surcharge ces seniors, et on finit par les soulager à leur tour par de la validation assistée. **Quand l'un part, on ne le remplace pas en le fabriquant, on l'achète sur le marché. Étape après étape, la proportion de gens ayant réellement traversé la**

chaîne qu'ils valident diminue. La supervision ne disparaît pas d'un coup : elle devient nominale. On signe toujours, mais on refait de moins en moins.

Le marché de l'emploi donne le premier signe de ce basculement. En France, l'APEC mesure un recul des recrutements de cadres débutants d'environ 16 % en 2025, après environ 19 % en 2024, concentré dans l'informatique, l'ingénierie et le conseil, soit les secteurs qui formaient le plus de profils peu expérimentés ; le conseil lui-même recule d'environ 10 % (Syntec Conseil). Le même mouvement s'observe outre-Atlantique, où la demande de seniors a bondi tandis que le recrutement junior se contractait (Revelio Labs). Une précaution s'impose : l'APEC impute d'abord ce recul à la conjoncture, pas à l'IA. Ces chiffres ne prouvent donc pas le mécanisme, ils en montrent le terrain : des organisations qui, déjà, achètent leur séniorité au lieu de la produire.

Sources : [APEC, recrutement de cadres 2025](#) ; [Syntec Conseil, marché du conseil 2024-2025](#) ; [Revelio Labs](#).

Où la question se pose, et où elle ne se pose pas

Avant d'aller plus loin, il faut délimiter. Dans bien des activités, croissance, marketing, contenu, support, une organisation peut parfaitement accepter de ne jamais pouvoir tout reprendre en main, et elle aura souvent raison : l'erreur y est faible, fréquente, réversible, et la vitesse vaut mieux que le contrôle. Le raisonnement qui suit ne vaut pleinement que là où l'erreur faible coûte cher et ne se rattrape pas : santé, défense, cybersécurité, finance systémique, industrie lourde, infrastructures critiques, systèmes régulés. C'est à ces systèmes que cette note s'adresse. Ailleurs, elle décrit un risque que beaucoup pourront ignorer sans tort.

Ce que le junior fabriquait sans qu'on le voie

Il faut le dire sans idéaliser l'ancien temps : ces tâches d'entrée produisaient aussi des faux seniors, de la compétence de surface, de l'ancienneté prise pour du jugement. Et y passer dix ans ne suffit pas à comprendre. Mais, chez beaucoup, elles fabriquaient une chose précise et coûteuse : la capacité de sentir qu'un résultat sonne faux avant qu'il ne devienne un incident, et de reprendre la main quand il le faut. On la dépose en faisant, en se trompant, en voyant un senior hésiter, en payant une fois le prix d'une erreur. C'était la manière, lente, dont l'organisation se fabriquait des gens capables, plus tard, de juger et de corriger le travail des autres.

La seule question qui compte

D'où la question, et elle est moins évidente qu'elle n'en a l'air. Superviser ne veut pas forcément dire tout refaire. Parfois une reconstruction partielle suffit, parfois une

contestation statistique, parfois une vérification croisée par d'autres. **Ce qui compte, à la fin, c'est qu'un moyen subsiste de reprendre la main : individuel, collectif, ou mixte.** La question n'est donc pas « gardons-nous des gens capables de tout refaire seuls », mais « gardons-nous, par un moyen ou un autre, la capacité de reprendre la main quand l'outil se trompe ou sort de son cadre ». Une organisation reste-t-elle capable d'entretenir cette capacité, à mesure qu'elle automatise ses couches basses ? Cette question doit se poser à toute échelle : individuelle, équipe, organisation, société et même civilisation.

Le junior capitalise, l'agent se loue

Il faut pousser la question d'un cran, car elle est aussi économique. Un junior coûte «cher» au début, puis il change de nature : il apprend, devient autonome, transmet, supervise à son tour. Sa dépense initiale se transforme en capital, un actif qui reste dans la maison et dont le coût marginal baisse avec le temps.

L'agent, lui, ne capitalise pas ainsi. Il se reloue à chaque requête. Il dépend d'une infrastructure qu'on ne possède pas : du calcul, un fournisseur, un modèle fermé, une feuille de route décidée ailleurs, un prix d'inférence qui varie. Et son coût grimpe précisément quand on lui demande ce que cette note réclame : du contexte, de la traçabilité, de la supervision, de la capacité à reprendre la main. **La cognition opposable n'est pas la cognition bon marché.**

Cette dernière phrase est une thèse à elle seule, et on ne la traitera pas ici : ses gradients, ses arbitrages assurantiels, ses modèles de coût appellent un texte plus économique que celui-ci. On se contente d'en retenir trois choses.

1. L'humain n'est pas un capital naturellement vertueux : beaucoup ne capitalisent presque rien, des compétences vieillissent vite, et certaines IA mutualisent au contraire des coûts massifs,
2. L'asymétrie est une tendance, pas une loi. Certaines organisations choisiront la location en connaissance de cause, plus flexible, plus légère, et le problème n'est pas qu'elles aient tort, mais qu'elles tranchent rarement cet arbitrage explicitement,
3. Enfin, qui ne produit plus sa cognition en dépend, et la perte de capital devient une perte de souveraineté.

Quand l'assistance précède l'autonomie

Ce texte a supposé jusqu'ici des humains formés à l'ancienne, puis assistés. Une rupture plus profonde s'annonce avec les entrants qui n'auront jamais connu la recherche sans synthèse automatique, l'écriture sans complétion, le débogage sans copilote, ni la construction lente d'un raisonnement long.

Le problème n'est alors plus seulement la transmission dans l'organisation. Il touche le socle : quelles capacités un système humain cesse-t-il d'internaliser quand leur coût cognitif est externalisé en permanence ?

Disons-le avec soin, car le terrain est miné. Il ne s'agit pas de prétendre qu'une génération penserait moins bien, ce serait faux et facile. Il s'agit d'un déplacement structurel : les architectures éducatives et professionnelles peuvent optimiser l'orchestration assistée avant la reconstruction autonome. **On devient alors excellent à piloter un outil, et plus faible à reprendre la main sans lui.** Les organisations externalisent leur cognition opérationnelle ; les personnes externalisent leur cognition reconstructive ; et chaque mouvement renforce l'autre. La question reste mesurable : à partir de quel niveau d'assistance permanente un système cesse-t-il de produire des gens capables, au besoin, de fonctionner sans elle ? L'accident de l'AF 447 Rio-Paris en est une illustration tragique.

Pourquoi personne ne le voit venir, et ce qui le révèle

Rien de tout cela ne se voit à court terme, et c'est le piège. La productivité monte, les coûts baissent, les juniors semblent augmentés, les seniors en place compensent, les tableaux de bord sont au vert. La facture, si elle arrive, arrive cinq à dix ans plus tard, quand il ne reste plus assez de gens, ni de dispositifs, pour reprendre la main.

Le vrai test n'est donc pas la performance quand tout va bien, c'est le mode dégradé. Une panne fournisseur, une coupure cloud, un incident de sécurité, un contexte indisponible, une contrainte réglementaire, un environnement isolé, une cyberattaque : le jour où l'assistance manque ou sort de son cadre, on découvre d'un coup qui peut encore fonctionner sans elle. C'est précisément ce qu'on ne mesure jamais. Au niveau des personnes comme des organisations, le même leurre : fluide et rapide, jusqu'à la première sortie de scénario.

L'objection sérieuse

On répondra qu'on referra les classes autrement, par la simulation. La médecine l'a institué : depuis le rapport de la Haute Autorité de santé de 2012, la règle est « jamais la première fois sur le patient », et plus d'une centaine de centres forment gestes et conduites critiques sur simulateur avant le réel, avec des bénéfices établis par revues systématiques. La simulation forme donc vraiment. Mais regardez ce qu'elle forme : des situations qu'on a su mettre en scénario parce qu'on les avait prévues. Il existe d'excellents chirurgiens utilisant la chirurgie robotisée qui sont perdus dès lors que les bras du robot s'immobilisent dans l'abdomen du patient et qu'il faut passer à l'antique scalpel pour faire une laparotomie étendue et extraire les bras. La crise scénarisée, oui ; la dérive que personne n'avait pensé à scénariser, non. Reste à savoir qui la repère, et là

rien n'est joué d'avance : certaines dérives, statistiques, faibles, multidimensionnelles, seront mieux vues par des architectures adversariales ou du monitoring multi-agent que par n'importe quel humain ; d'autres, non. Personne n'a le monopole de la vigilance hors cadre.

Sources : [HAS, « jamais la première fois sur un patient »](#) ; [revue systématique, simulation en santé et résultats patients](#).

Rien de tout cela n'est écrit

Si tous ces mécanismes pointaient dans le même sens, ce texte décrirait une fatalité. Ce n'en est pas une. Les gains sont réels, et les taire serait malhonnête : une cognition externalisée peut apporter de la robustesse statistique, mutualiser des compétences rares, accélérer, relever la qualité moyenne, ouvrir l'accès à des savoirs jusque-là réservés. Une organisation peut perdre de l'autonomie reconstructive et gagner beaucoup ailleurs ; l'arbitrage n'a de sens qu'en portant les deux colonnes.

Et la reprise de contrôle n'a pas besoin de rester individuelle. Une supervision peut être collective, distribuée, redondante, et parfois plus robuste ainsi que portée par un seul cerveau. Aucun individu ne refait toute la chaîne, et le système tient quand même, parfois mieux. La condition est toujours la même, et c'est la seule qui compte : que le système, par un moyen ou un autre, garde la capacité de contester et de reprendre la main, pas seulement de produire. Le risque décrit ici n'est pas un effet mécanique de l'IA. C'est l'option par défaut quand personne ne choisit l'autre.

Une réponse possible, parmi d'autres

L'autre option a une forme, et c'est une hypothèse de travail du Twingital Institute, pas une solution close. Elle consiste à entretenir délibérément deux modalités cognitives, l'humaine et l'artificielle, conçues pour que chacune corrige les faiblesses de l'autre, et à maintenir en parallèle deux pipelines plutôt qu'un seul : un pipeline d'intelligence artificielle outillé de garde-fous, supervision, traçabilité, conditions de reprise, et un pipeline de développement des compétences humaines en temps long, celui-là même que le mécanisme décrit ici laisse s'envaser. Le « port de promotion » n'est rien d'autre que le nom de ce second pipeline.

Il faut s'en méfier comme de toute solution qui tombe trop juste avec son problème. Opposer un pipeline humain à un pipeline IA est commode pour penser, mais les systèmes réels s'hybrident, s'intriquent, se co-produisent ; à terme, reprendre la main vaudra moins dire revenir à l'humain que reconfigurer une cognition distribuée. Le modèle à deux modalités est un point de départ, pas une fin, et ce n'est pas la seule réponse imaginable.

C'est aussi un terrain, pas seulement un principe. Au sein de l'équipe TweenMe, de jeunes data scientists se forment sous la conduite de seniors pendant que l'IA sert d'assistant de production, de manière raisonnée : les deux modalités y coexistent par construction, l'outil restant un assistant et non le substitut du jugement qu'on cherche à former. Comme tout terrain d'implémentation, il montre qu'une telle architecture tient debout, sans en faire à lui seul une preuve générale.

Ce qu'un dirigeant devrait se demander

La plupart des organisations pilotent le chiffre d'affaires, la marge, la productivité. Dans les systèmes où l'erreur faible coûte cher, très peu savent répondre à quelques questions concrètes : combien de nos seniors avons-nous fabriqués, contre combien achetés ; quelle part de la chaîne celui qui valide l'a-t-il réellement parcourue ; si l'on coupait l'outil demain, qui saurait reprendre la main ; et entretenons-nous vraiment deux modalités, ou une seule déguisée en deux. Ce ne sont pas des indicateurs de ressources humaines, ce sont des indicateurs de gouvernabilité.

Une décision peut être fautive ; cela se corrige. Ce qui ne se corrige pas, c'est le jour où le système sort de son cadre et où plus personne, ni humain ni dispositif, n'a gardé de quoi reprendre la main.