

Version en ligne : aipm-standard.com/aipm/cas/anthropic-enquete-interne · édition du 30/09/2026. Le statut de preuve peut évoluer : la page en ligne fait foi.

Cas approfondi

Chez l'éditeur lui-même, la délégation complète reste l'exception

En août 2025, Anthropic interroge 132 de ses ingénieurs et chercheurs et analyse 200 000 sessions internes de Claude Code. Les usagers déclarent se servir de Claude dans 59 % de leur travail, contre 28 % un an plus tôt. Et pourtant, plus de la moitié disent ne pouvoir en « déléguer entièrement » que 0 à 20 %. Même chez les plus avancés, l'agent est un collaborateur supervisé, pas un exécutant autonome.

Toute la banque de cas (aipm-standard.com/aipm/cas) · La règle de preuve (aipm-standard.com/aipm/methode-de-preuve)

Statut de preuve : Confirmé opérationnel. Ce statut n'est pas une opinion sur l'entreprise ou sur la technologie : il décrit ce que les sources consultées permettent d'affirmer, et rien de plus. La règle qui l'attribue est publiée sur la page Méthode de preuve (aipm-standard.com/aipm/methode-de-preuve). Il porte sur le constat central de la fiche ; les chiffres cités à l'intérieur conservent le statut de leur propre source, et ceux qui proviennent du seul éditeur restent « annoncés », y compris dans une fiche confirmée.

Le contexte

Qui, quand, quel système

Le 2 décembre 2025, Anthropic publie, dans sa rubrique de recherche « Societal Impacts », une étude sur son propre personnel technique. Méthode : une enquête auprès de 132 ingénieurs et chercheurs en août 2025, 53 entretiens qualitatifs approfondis, et l'analyse de 200 000 transcriptions internes de Claude Code de février et d'août 2025, par un outil d'analyse préservant la confidentialité.

Les auteurs reconnaissent d'emblée leur position : étudier l'effet de l'IA dans une entreprise qui la fabrique, avec des ingénieurs ayant un accès anticipé aux meilleurs outils, est une situation privilégiée. Ils ajoutent que leurs résultats « ne se généralisent probablement pas » aux autres organisations, et publient une section entière de limites.

La décision, et qui l'a prise

La Directive de fait

La décision documentée est méthodologique : croiser le déclaratif (enquête, entretiens) avec l'observé (journaux d'usage). C'est ce croisement qui donne sa valeur au cas, parce que les deux sources ne disent pas la même chose avec la même force.

Du côté des répondants, la décision qui se lit dans les entretiens est celle de ce qu'on délègue. Les ingénieurs confient à Claude les tâches faciles à vérifier, à faible enjeu, ou ennuyeuses ; ils gardent la conception de haut niveau et les décisions qui demandent du contexte ou du « goût ».

Ce que le système a réellement fait

Côté déclaratif, les répondants estiment utiliser Claude dans 59 % de leur travail, contre 28 % douze mois plus tôt, pour un gain de productivité moyen de +50 %, contre +20 %. Les auteurs rapprochent ce chiffre d'une hausse de 67 % des demandes de fusion intégrées par ingénieur et par jour, observée lors de l'adoption de Claude Code dans l'organisation d'ingénierie. Mais plus de la moitié des répondants ne peuvent « déléguer entièrement » que 0 à 20 % de leur travail.

Côté observé, en six mois, le nombre maximal d'appels d'outils consécutifs sans intervention humaine par transcription passe de 9,8 à 21,2 ; le nombre moyen de tours humains par transcription passe de 6,2 à 4,1. La complexité moyenne estimée des tâches monte de 3,2 à 3,8 sur une échelle de 1 à 5.

Les entretiens ajoutent ce que les chiffres ne montrent pas : la crainte d'une atrophie des compétences nécessaires pour superviser l'agent, que les auteurs nomment « paradoxe de la supervision ».

L'état de la preuve

Ce que les sources établissent, et ce qu'elles n'établissent pas

Ce que la source établit

- L'enquête a été conduite en août 2025 auprès de 132 ingénieurs et chercheurs, complétée de 53 entretiens
- Plus de la moitié des répondants déclarent ne pouvoir déléguer entièrement que 0 à 20 % de leur travail
- La part déclarée du travail faite avec Claude passe de 28 % à 59 % en un an, selon les souvenirs des répondants
- Les journaux d'usage montrent une autonomie accrue par session : de 9,8 à 21,2 appels d'outils consécutifs au maximum par transcription, de 6,2 à 4,1 tours humains
- Les auteurs publient leurs limites : échantillon de convenance, réponses non anonymes, biais de rappel

Ce que la source n'établit pas

- Une productivité mesurée : le +50 % est auto-déclaré, et les auteurs citent une étude externe où des développeurs ont surestimé leur gain
- L'origine vérifiable du +67 % de demandes de fusion : il ne provient pas de l'enquête, et la source qu'Anthropic cite est un article payant de tiers, non lu ici
- Des volumes absolus : l'échantillonnage des transcriptions ne mesure que des

proportions

- La transposabilité hors d'une entreprise qui fabrique l'outil et dont les salariés y ont un accès anticipé
- Le sens exact de « déléguer entièrement », que les répondants interprètent différemment

Sources consultées : Anthropic — How AI is transforming work at Anthropic (source primaire 2025-12-02)

Ce que ce cas dit du référentiel

Ce cas tempère le récit d'une délégation déjà généralisée, et c'est pour cela qu'il compte. Il donne une assise au **principe 4**, « L'autonomie se mérite, jamais par défaut » : les ingénieurs décrivent eux-mêmes une progression de confiance, des tâches simples vers des tâches complexes, et une frontière de délégation renégociée à chaque modèle. Les journaux confirment la trajectoire — plus d'appels d'outils, moins de tours humains — sans que la délégation complète devienne la norme.

Leurs critères de délégation recoupent le domaine **4.3 Cadrage & filtre de délégabilité** : facile à vérifier, faible enjeu, bien délimité. C'est un filtre de délégabilité formulé spontanément par les praticiens.

Le « paradoxe de la supervision » est une forme précise du risque **7.5 Dépendance excessive** : superviser l'agent exige les compétences que l'usage de l'agent érode. Il pèse directement sur le rôle du **Directeur-Vérificateur**, dont la capacité à juger doit être entretenue, pas présumée.

Enfin, la façon dont la source traite ses propres chiffres est un modèle du **principe 7** : elle sépare ce qui est déclaré de ce qui est observé et publie ses biais.

La consigne

Ce qu'un lecteur fait différemment après avoir lu cette fiche

Mesurez la délégation complète, pas l'usage. « Nos équipes utilisent l'IA » ne dit rien. Demandez quelle part du travail part sans vérification humaine, et sur quels types de tâches.

Croisez toujours déclaratif et journaux. Si vous faites une enquête interne, adossez-la à une donnée d'usage observée ; la productivité ressentie n'est pas une métrique AIPM.

Protégez la compétence du vérificateur. Prévoyez, pour chaque Directeur-Vérificateur, des tâches faites sans agent, à intervalle régulier. Un vérificateur qui ne sait plus faire ne sait plus juger.

Ce qui invaliderait cette lecture. Une réédition de l'étude, par Anthropic ou un tiers, montrant qu'une majorité de professionnels délèguent entièrement plus de la moitié de leur travail : le constat « la délégation complète reste l'exception » serait périmé. Ou une étude anonyme, à instruments validés, qui ne retrouverait pas la tendance mesurée dans les journaux.

Cette fiche suit le gabarit des cas approfondis d'AIPM : huit questions, toujours les mêmes. Voir les autres fiches (aipm-standard.com/aipm/cas-approfondis).

© 2026 DYONYSOS — Tous droits réservés · Bexo® est une marque déposée de DYONYSOS · aipm-standard.com. Reproduction interdite sans autorisation, hors courtes citations avec mention de la source (aipm-standard.com).