NOTE DE TERRAIN

19 novembre 2026
6 min de lecture

J'ai bâti une équipe d'agents IA. Puis je suis devenu le goulot d'étranglement.

Quatorze agents IA travaillent dans nos opérations, et les étapes importantes attendent une seule personne. Ce que dit la recherche, et ce que nous changeons.

RÉDIGÉ PAR

Mo Kahlain

Lire en ligne: https://inocta.io/fr/labo/notes-de-terrain/je-suis-devenu-le-goulot-de-mes-propres-agents-ia/

Le 3 novembre, j'ai demandé à l'un de nos agents IA de réviser avec moi un brouillon de ce texte. Il m'est revenu avec une révision soignée : six remarques, et chacune était une chose à vérifier de mon côté. Le travail était bon. L'agent me le renvoyait quand même.

Ma semaine tient dans ce moment. Chez inocta.io, quatorze agents IA travaillent dans nos propres opérations. Ils rédigent, développent du code, s'occupent d'une partie de la comptabilité et font le suivi des clients. Au cours des 30 jours qui ont précédé le 3 novembre, ils ont enregistré 1 042 fermetures de tâches. Une bonne part de ce qui compte attend pourtant une seule personne avant d'avancer. Cette personne, c'est moi.

01

Le piège a un nom

Dans la Harvard Business Review, Sangeet Paul Choudary décrit ce qu'il appelle l'illusion de « l'humain dans la boucle ». Bien des entreprises placent des personnes autour des tâches qu'une machine ne sait pas encore faire, comme une couche de sécurité posée sur l'automatisation. Il montre pourquoi cette couche s'affaiblit avec le temps. Devant un système fiable, les personnes qui le surveillent relâchent leur vigilance. Une mauvaise réponse de l'IA paraît aussi assurée qu'une bonne. Chaque correction sert ensuite à entraîner le système, si bien que le rôle humain rétrécit sans cesse. Pour lui, ces points de contrôle sont à la fois peu fiables et appelés à disparaître d'eux-mêmes.

Une recherche publiée par Anthropic le 27 octobre va plus loin. Cinquante analystes professionnels ont travaillé avec un agent IA à qui les chercheurs avaient donné, à dessein, une préférence. Sur quatre choix d'analyse, les analystes ont suivi la préférence de l'agent 86 % du temps, alors que le hasard en donnerait environ 50 %. L'expérience ne les a pas protégés, et leur confiance était la même dans un cas comme dans l'autre. L'étude porte sur une seule tâche, dans un seul domaine, avec un seul modèle, et la préférence avait été renforcée exprès. Elle me dit tout de même une chose inconfortable : ma signature ne prouve pas que j'ai exercé mon jugement.

02

Comment nous avons construit cette boucle nous-mêmes

Personne n'a décidé de me placer dans chaque boucle. Nous avons écrit nos règles selon le type d'action. Tout ce qui envoie, publie, déploie, dépense de l'argent ou modifie une règle de fond exige mon accord. Prise isolément, chaque règle se défendait. Mises ensemble, elles dirigent chaque étape importante vers un seul bureau.

Cette boucle est aussi moins visible qu'on le croit. Le 3 novembre, notre propre système de billets affichait deux éléments en attente de ma part. Le vrai compte était plus élevé. Cinq changements de code attendaient que je les fusionne. Quatre avaient été révisés en cours de

route. Aucun n'avait de révision réussie sur sa version finale. Plusieurs billets indiquaient « en attente de Mo », alors que le champ qui désigne le prochain intervenant nommait un autre agent.

Le plus difficile pour moi, c'est que chaque agent applique correctement ses règles. Il termine le travail et le vérifie. Puis il demande la permission avant l'étape qu'il n'a pas le droit de franchir. La friction vient de la conception, et c'est donc à moi de la changer.

03

Ce que nous avons changé, et ce que nous testons

Le premier changement a retiré la vérification de mon bureau. Depuis le 26 octobre, une tâche à fort enjeu ne peut être fermée sans révision indépendante réussie ou raison écrite et comptée. Ce réviseur est un autre modèle d'IA ou un agent distinct. Choudary demande davantage : des règles intégrées au travail et vérifiées en continu. C'est un premier pas, et le premier décompte n'est pas flatteur. Notre registre des révisions a commencé le 26 octobre. Au 3 novembre, 31 des 358 fermetures enregistrées depuis portaient une révision. Sur les 38 fermetures visées par la nouvelle règle, 28 ont utilisé la raison écrite. Maintenant, nous voyons l'écart.

Le deuxième changement est sur mon bureau en ce moment : un approbateur désigné qui n'est pas le PDG. Prenons un changement interne dont la révision est approuvée et dont les tests passent, sans aucun lien avec un client ni avec de l'argent. L'agent qui en est responsable le fusionnerait lui-même. Je verrais la liste chaque jour, et chaque fusion pourrait être annulée. La révision approuvée doit porter sur la version finale : aucun de ces cinq changements n'aurait été admissible.

Le troisième porte sur la confiance qui se mérite. Chaque agent a un score bâti sur son historique. Nous avons décidé que ce score n'ouvre encore aucune porte. Pendant 30 jours, il notera seulement ce qu'il aurait permis, à côté de ce qui s'est réellement passé. Ce registre n'a pas encore commencé. Je trancherai ensuite sur des cas réels, pas sur un chiffre.

Certaines décisions reposent sur mon jugement, comme un prix ou un conseil à un client. Pour celles-là, nous préparons un dernier test. L'agent présenterait d'abord le meilleur argument en faveur de chaque option, puis son propre choix. Quand mon avis est justement ce qui compte, il me le demanderait en premier.

04

Donner une destination au temps libéré

Rien de tout cela ne compte si le temps gagné se perd dans encore plus de vérifications. IKEA montre une autre voie. De 2021 à 2023, le robot conversationnel du service à la clientèle d'Ingka Group a réglé 3,2 millions d'interactions et a permis d'économiser près de 13 millions d'euros. Ingka a aussi formé 8 500 employés de centre d'appels. Ils ont appris le design d'intérieur à

distance, la vente au détail numérique, l'établissement de relations et la résolution de problèmes complexes. À la fin de son exercice 2022, les ventes réalisées par ses points de rencontre à distance atteignaient 1,3 milliard d'euros. Le robot a retiré du travail. La valeur est venue de la décision prise sur ce que les personnes feraient ensuite.

Cette décision a besoin d'un responsable. Notre méthode suit quatre étapes : Cartographier, Structurer, Déployer, Gérer. Nous cartographions d'abord le travail tel qu'il se fait aujourd'hui, puis nous structurons qui décide de quoi. Ensuite seulement, nous déployons une application ciblée, gérée par un responsable désigné. C'est à l'étape Structurer que se fixe l'essentiel de la charge de l'humain dans la boucle, souvent avant que quiconque s'en rende compte.

05

Cinq questions pour votre prochaine réunion de direction

Avant d'étendre une initiative d'IA, je mettrais ces questions sur la table :

1. Quel problème client avons-nous maintenant les moyens de régler ?
2. Qui est responsable de ce résultat et a l'autorité pour l'améliorer ?
3. Quelles décisions l'agent peut-il prendre seul, lesquelles exigent un approbateur désigné, et cet approbateur est-il quelqu'un d'autre que le PDG ?
4. Quels signes nous feraient arrêter, modifier ou étendre le processus, et qui met à jour les consignes quand la même correction revient ?
5. Où ira le temps libéré, et comment saurons-nous qu'il a créé de la valeur ?

Si, à la troisième question, votre réponse honnête est « moi » pour la plupart des décisions, vous avez déplacé le travail sur votre propre bureau. Je le sais, parce que c'est exactement ce que j'ai fait.

Lire en ligne: https://inocta.io/fr/labo/notes-de-terrain/je-suis-devenu-le-goulot-de-mes-propres-agents-ia/