

# Puissants et risqués, les agents IA arrivent

**TECHNOLOGIE** OpenAI a récemment dévoilé ses assistants capables de tout faire, entre planification de mariage, activités bureautiques et organisation de vacances. Tous les géants du domaine s'y intéressent, mais les produits sont loin d'être parfaits

ANOUCH SEYDTAGHIA

Pour certains, c'est une vision du paradis technologique. Pour d'autres, c'est plutôt un aperçu de l'enfer numérique. Les agents d'intelligence artificielle sont en train d'arriver sur le marché, et suscitent autant d'espoirs que de craintes. Ces nouveaux outils sont des petits robots appelés à tout faire, ou en tout cas à accomplir des tâches qu'effectuent aujourd'hui des humains: faire des recherches sur le web, compiler des résultats, effectuer des études de marché, des réservations de voyage, des publications sur des réseaux sociaux, ou encore répondre à des clients. Les possibilités sont infinies et tous les géants de l'intelligence artificielle (IA) investissent des milliards dans ces agents.

Tout récemment, OpenAI a dévoilé ses premiers agents destinés au grand public. La société dirigée par Sam Altman dévoilait jeudi dernier ce qu'elle appelle «l'agent ChatGPT» appelé à «faire le lien entre recherche et action». Selon OpenAI, «désormais, ChatGPT réfléchit et agit. À l'aide de sa boîte à outils, il choisit de façon proactive les bonnes compétences agentiques pour réaliser la tâche demandée». La société évoque un «système agentique unifié», ChatGPT étant capable d'interagir avec des sites web et d'effectuer des recherches avec son «propre ordinateur individuel». «En suivant vos instructions, ChatGPT se charge du travail à votre place afin d'accélérer votre production, tout en vous libérant du temps pour vous consacrer à d'autres aspects de votre quotidien», affirme OpenAI.

## «Une étape clé»

Des exemples? Selon la société, l'agent peut analyser une invitation à un mariage, puis aider l'utilisateur à trouver une tenue appropriée, réfléchir à des idées de cadeaux et réserver un hôtel à proximité. L'agent peut créer diaporamas et graphiques. Il est aussi capable d'utiliser une photo pour ensuite en commander des autocollants basés sur ces clichés sur des sites spécialisés, puis

## «Ces agents ne se contentent plus d'analyser ou de répondre: ils sont désormais capables d'agir»

ALAIN FRIGERIO, ANALYSTE CHEZ REYL INTESA SANPAOLO

envoyer ces derniers à des adresses postales. Ces agents, disponibles pour l'heure aux utilisateurs payants des services de ChatGPT aux Etats-Unis – et sans doute dans les prochains mois en Europe – peuvent être à tout moment arrêtés ou reconfigurés par les utilisateurs.

Il s'agit donc de donner les clés de son ordinateur, de son agenda, de son navigateur web et de ses comptes en ligne à ChatGPT, qui devient une sorte d'assistant multifonction. Est-ce une évolution si importante? «Le lancement des agents IA par OpenAI et d'autres acteurs majeurs du secteur marque une étape clé dans l'évolution vers des IA plus autonomes, proactives et opérationnelles. Ces agents ne se contentent plus d'analyser ou de répondre: ils sont désormais capables d'agir, en exé-

cutant des tâches concrètes sans intervention humaine directe», estime Alain Frigerio, analyste chez Reyl Intesa Sanpaolo.

Selon le spécialiste, «l'IA devient un acteur, capable de prendre en charge des processus, de coordonner des actions, et de s'intégrer plus profondément dans les environnements professionnels. A

terme, ces agents pourraient automatiser des tâches complexes et redéfinir des métiers entiers. Leur capacité à passer de l'analyse à l'action transforme l'IA d'un outil conversationnel en un véritable collaborateur numérique, avec un potentiel d'impact considérable sur la gestion de projets, l'analyse concurrentielle ou la coordination d'équipes hybrides.»

## Une mise en garde de Sam Altman

Et il faut se préparer à ce que ces agents d'IA envahissent notre quotidien, estime Christophe Pouchoy, gérant auprès de La Financière de l'Echiquier: «La création d'agents d'IA est la nouvelle vague de monétisation de cette technologie et elle va s'étendre dans les logiciels à tous les segments de marché. De la réservation de

voyages à l'exécution de process ou tâches manuelles, les agents d'IA proliféreront dans l'économie. Ils requièrent toujours une personnalisation plus au moins importante en fonction de la tâche ou de l'application souhaitée mais celle-ci peut être standardisée par sous-segment de marché.»

Christophe Pouchoy y voit aussi un intérêt financier direct pour OpenAI. «L'éditeur de ChatGPT pourrait utiliser les agents d'IA pour générer davantage d'activités commerciales au sein de son écosystème. Cela lui permettrait à la fois de facturer ce service aux utilisateurs – en plus des abonnements de base – et de prélever une commission auprès de sociétés dont les produits et services seraient achetés via OpenAI.»

Qui dit délégation de tâches entières à l'IA dit bien sûr risque

d'erreurs. Et les agents sont loin d'être infaillibles, a reconnu Sam Altman: «C'est une technologie de pointe et expérimentale. C'est l'occasion d'essayer (sic) l'avenir, mais je ne l'utiliserais pas encore pour des utilisations à fort enjeu ou avec beaucoup d'informations personnelles tant que nous n'aurons pas eu l'occasion de l'étudier et de l'améliorer». OpenAI ne permettra par exemple pas, au début au moins, aux agents d'effectuer des opérations financières. «De la même manière que les premiers modèles de langage avaient des taux d'hallucination élevés à leurs débuts, les agents d'IA pourraient commettre des erreurs. Cependant, les optimisations progressives de ces agents devraient les rendre de plus en plus intelligents, réduisant le nombre de fautes», affirme Christophe Pouchoy.

## Quelle tolérance aux erreurs?

Alain Frigerio estime que «les erreurs sont un frein, mais pas un obstacle insurmontable». Selon lui, «quand on confie à une IA des tâches critiques – envoyer un e-mail, réserver un vol, modifier un fichier, ou même prendre une décision basée sur des données –, la tolérance à la faute est très faible. Et aujourd'hui, même les meilleurs agents peuvent mal interpréter une consigne, rater une subtilité contextuelle, ou agir sur une information incomplète. Cette fragilité perçue va ralentir leur adoption dans certains secteurs sensibles, comme la finance, le juridique ou la santé. Mais dans d'autres contextes, où les conséquences d'un bug sont moindres ou facilement rattrapables, les agents vont s'imposer plus vite — notamment pour des tâches de support, de coordination ou d'analyse».

L'analyste chez Reyl Intesa Sanpaolo conclut que «ce qui fera la différence, ce sera la capacité à instaurer un cadre de confiance: transparence sur les actions de l'agent, possibilité de valider ou d'annuler ses décisions, traçabilité de ce qu'il a fait et pourquoi, et surtout, apprentissage continu à partir des retours humains». ■



(S. SNJAYTUTURKHI/GETTY IMAGES)