

# Le Sénat juge que l'expérimentation de l'IA générative dans le public passe à côté des vrais sujets

Emile Marzolf

Image d'illustration générée par Midjourney.

La délégation à la prospective du Sénat a rendu son [tout premier rapport sur l'intelligence artificielle](#) (IA) dans le service public. Avec un volet important sur l'IA générative et ses sous-jacents, les grands modèles de langue (LLM) qui, selon les 2 auteurs du rapport, Didier Rambaud (RDPI, majorité) et Sylvie Vermeillet (Union centriste), *“excellent dans le traitement du langage naturel et des données non structurées et hétérogènes, notamment les textes”*, là où les techniques plus classiques d'apprentissage automatique sont plus adaptées pour traiter des données *“structurées et normalisées”*. Ces dernières sont déjà utilisées depuis 2013 par les services des impôts pour détecter la fraude.

Mais malgré les expérimentations nombreuses engagées dans le secteur public, l'IA générative serait encore testée de façon trop *“timide”*, sans s'attaquer aux vrais sujets cœurs de métier de la direction générale des finances publiques (DGFIP), de la douane ou encore des administrations sociales que sont l'Urssaf, la Caisse nationale des allocations familiales et l'assurance vieillesse (Cnav). *“Pour le service public, tout ceci n'est pas seulement un enjeu d'efficacité : c'est aussi un enjeu d'équité, d'accessibilité, et donc d'humanité, estiment les sénateurs. L'IA, et particulièrement l'IA générative, permet de simplifier, de personnaliser, d'expliquer et de rapprocher le service public.”*

## Utilisation “superficielle”

Les sénateurs se sont attachés à identifier les cas d'usage de l'IA générative à plus fort potentiel, qu'il s'agisse de l'utilisation des IA génératives “sur étagère”, comme le “Copilot” de Microsoft, ou bien des modèles de langage réentraînés sur un sujet spécifique. Pour les premières, les rapporteurs voient un gisement de productivité essentiellement en matière d'informatique pure, pour aider à écrire du code, à “debugger” des applications ou à rédiger automatiquement de la documentation à partir du code...

*“À elle seule, la DGFIP utilise près de 700 applications métiers différentes, et compte 5 200 agents dans ses services informatiques. La Douane, pour sa part, utilise 200 applications et compte 420 informaticiens. L'Urssaf assure la gestion d'environ 100 applications. Ici, l'IA générative pourrait radicalement changer la donne, compte tenu de ses performances impressionnantes en matière de génération de code informatique et d'assistance aux développeurs”*, écrivent les sénateurs. D'ailleurs, l'assistance aux développeurs fait partie des cas d'usage prioritaires identifiés

par la DGFIP, notamment pour convertir des programmes écrits dans un langage devenu obsolète, et sur lequel ne sont pas formées les nouvelles recrues. Problème : l'utilisation des outils sur étagère comme Github Copilote sont strictement interdits pour des raisons de sécurité et de confidentialité des données. Une interdiction qu'il convient de lever, selon les sénateurs, en adoptant une *“approche différenciée en fonction du niveau de risque”*.

Des cas d'usage un peu plus poussés ont bien été développés, comme l'outil d'attribution et de résumé d'amendements parlementaires de la DGFIP, “Llamandement”, mais restent superficiels, selon les sénateurs. Cette expérimentation ne concerne pas le “cœur de métier” de l'administration fiscale : l'outil n'est pas intégré aux “grands” systèmes d'information de la DGFIP (gestion de l'impôt, du recouvrement, etc.) et n'exploite aucune donnée individuelle.

### **Pas encore de *chatbot* au contact de l'utilisateur**

Les sénateurs regrettent également que les expérimentations de *chatbots* “dopés” à l'IA générative se limitent à des usages marginaux et ils ciblent en particulier l'expérimentation d'aide à la réponse aux “expériences” partagées par les usagers sur la plate-forme Services publics+. *“C'est l'intérêt même d'un tel outil qui pose question : il porte sur un cas d'usage tout à fait marginal à l'échelle du service public (répondre à quelques dizaines de commentaires sur Internet), et il apporte une aide relativement limitée aux agents dans leur travail, et aucune aide aux usagers directement”*, pointent les rapporteurs. Mais il fallait bien commencer quelque part, et cette expérimentation a eu le mérite de permettre aux directions interministérielles du numérique et de la transformation publique, mais aussi aux agents des services publics partenaires de Services publics+ de mettre les mains dans le cambouis.

Et si l'expérimentation de l'IA générative “Albert”, pour aider les agents des maisons France services à répondre aux questions des usagers leur semble plus prometteuse, *“il n'est pas question qu'un usager puisse, depuis chez lui, s'adresser directement à Albert pour obtenir des conseils personnalisés et des explications détaillées, et encore moins pour faire les choses à sa place – soit précisément la rupture technologique introduite par les assistants IA comme ChatGPT”*, déplorent les sénateurs. Qui ne sous-estiment pas pour autant le manque de fiabilité de ces outils, encore sujets à de nombreuses approximations, voire à des réponses complètement inventées.

Ce qui les amène d'ailleurs à remettre en question le choix de modèles de langue ouverts, comme Mistral ou LLama2 pour faire tourner leurs IA génératives. *“Au nom de la « souveraineté », il a été décidé de privilégier un modèle open source, français de surcroît, mais ancien, peu performant (7,3 milliards de paramètres) et nécessitant une étape de fine tuning\* complexe et coûteuse pour l'entraîner sur des données spécifiques, alors que les modèles plus gros et sans fine tuning, par exemple GPT-4 (1,7 milliard de milliards de paramètres) ou même le plus récent Mistral Large, offrent des performances incomparablement supérieures et peuvent être utilisés directement”*, observent-ils. Certes, mais pour un coût autrement plus élevé, et avec des risques pour les données sensibles de l'État et des usagers.

Quoi qu'il en soit, les sénateurs appellent à passer à la vitesse supérieure et à ne pas se limiter à ces IA génératives "superficielles", qui ne font que traiter des masses de données non sensibles, et surtout non individuelles. *"Pour le service public, la véritable plus-value de l'IA est ailleurs"* et ne *"peut pas venir de l'extérieur"* : dans son intégration au cœur du système d'information pour traiter des données individuelles et métiers. Avec un potentiel autrement plus important pour préremplir les démarches dans la logique du "Dites-le-nous une fois", voire pour réaliser ces démarches à la place de l'utilisateur, et même carrément les supprimer.

*\* Technique consistant à spécialiser un modèle d'IA préentraîné à l'accomplissement d'une tâche spécifique (source : Cnil).*