

« IA générative :  
face au tourbillon,  
les clés d'une  
adaptation rapide  
pour les DSI »



# Sommaire

- 3 Avant-propos**  
Quels changements au sein des DSI pour mieux accélérer sur l'IA?
- 4 Synthèse**  
Les DSI face au tourbillon de l'IA générative
- 6 Interview de Julien Largillière**  
Veolia est partie tôt pour tenir le marathon de l'IA générative dans la durée
- 9 Interview de Lionel Chaine**  
L'IA générative, un tournant majeur pour le métier de Bpifrance
- 11 Interview partenaire**  
**Philippe Battel et Philippe Wojcik.** Pour Oracle, les stratégies d'IA générative devront plus que jamais tenir compte du ratio prix-performance
- 14 Interview partenaire**  
**Stéphane Estevez.** L'AI Ops, intimement liée au déploiement de l'observabilité, analyse Splunk
- 16 À propos**  
Oracle, Splunk, Ippon Technologies & Alliancy

# Avant-propos

## Quels changements au sein des DSI pour mieux accélérer sur l'IA ?

L'année 2023 a vu la découverte et la prise de conscience par le grand public de l'importance de l'intelligence artificielle générative (IAG) et des *large language models (LLM)*. Dans les milieux professionnels, les annonces se sont multipliées, car les organisations n'avaient pas le choix : il fallait montrer qu'elles prenaient le train en route, à leurs clients comme aux investisseurs. Les mois à venir vont donc être déterminants : il est maintenant essentiel pour les organisations de pouvoir révéler comment ces avancées sont intégrées à leurs stratégies IA existantes. Au-delà de faire des « coups de communication », elles doivent donc confirmer comment elles peuvent embarquer, efficacement, à la fois les équipes IT, data et métier dans les nouvelles transformations à venir. Choix technologiques à activer, maturité de la coopération IT-métier sur le sujet, transformation de la gouvernance pour industrialiser les projets... : ces thèmes clés ont été au cœur des premiers *Data Excellence Sprint* d'Alliancy consacrés à l'IAG.

Les directeurs des systèmes d'information réunis par Alliancy ont ainsi pu partager leurs expériences sur cette première année d'IAG. Lors de deux ateliers, ils se sont confrontés aux expériences de deux de leurs pairs, nos « *team leaders* » qui avaient travaillé sur le sujet : Julien Largillière, *group chief technology officer* de Veolia, et Lionel Chaine, directeur des systèmes d'information de Bpifrance. Puis, ils ont écouté les partenaires de la rencontre, Oracle, Ippon Technologies et Splunk, exposer leur vision du marché et des transformations en cours. Lors d'un diner-débat animé, tous ont ensuite pu poser leurs questions, « se défier » et se projeter sur les priorités pour 2024.

Quels changements de gouvernance data mettre en place pour pouvoir industrialiser ses projets IA ? Comment dépasser l'effet paillettes de ChatGPT pour montrer des gains concrets sur les activités métiers ? Quelles sont les urgences ? Et, comment les distinguer : entre IA « sauvage », IA « embarquée », IA « hybride » et, peut-être demain, le développement de leurs propres *LLM* par les entreprises ? Autant de questions qui ont été au cœur des débats. Nous avons réuni dans les pages qui suivent certains des enseignements clés de cette soirée. Et bien sûr, 2024 ne manquera pas d'occasions de faire le point sur la vitesse à laquelle les entreprises auront su se transformer avec l'IAG. À commencer lors de la deuxième édition des *Data Excellence Sprint* d'Alliancy.

Bonne lecture.

Dorian Marcellin, rédacteur en chef

# Les DSI face au tourbillon de l'IA générative

Alors que ChatGPT est considéré comme un véritable «game-changer» par les directeurs des systèmes d'information (DSI), la question de la confidentialité des données et de l'adaptation de l'entreprise à ce nouvel environnement se pose.

À l'occasion du premier *Data Excellence Sprint* d'Alliancy, événement qui a réuni une quarantaine de DSI et de décideurs technologiques le 21 novembre 2023 à Paris, le sujet de l'impact de l'IA générative (IAG) sur la direction des systèmes d'information et sur les entités métiers a été débattu lors d'ateliers.

L'IAG est-elle une technologie de rupture? Même si la plupart des participants rappellent que ce n'est qu'une technologie d'IA parmi d'autres (sur laquelle de très nombreux chercheurs travaillent depuis plusieurs décennies), ils constatent qu'elle apporte de véritables nouveautés, notamment à travers les récentes avancées proposées par ChatGPT.

## ChatGPT : un nouveau paradigme

«ChatGPT apporte une interface et une façon d'utiliser de l'algorithmique en temps réel qui sont de vraies nouveautés. Cela perturbe tout le monde, car ce n'était pas très visible jusqu'à présent. Et surtout, nous arrivons aujourd'hui à

mixer plusieurs approches : texte, image, voix, compréhension, vidéo... Donc cela remue au sein même des métiers de la direction des systèmes d'information : génération de code, transcription d'un code à un autre... Et, dans les entités métiers, ces algorithmes vont donner lieu à une multitude d'usages, à des rythmes différents», déclare une participante qui évolue dans le secteur de l'énergie.

Un autre participant, dont l'entreprise est, elle aussi, dans le secteur de l'énergie, met en valeur des apports supplémentaires de ChatGPT : «Avant tout, c'est la manière dont nous lui parlons, via les *prompts*, qui a changé. ChatGPT se souvient par ailleurs du contexte de vos demandes. L'esprit de synthèse est également une nouveauté; tout comme la créativité, c'est-à-dire la capacité à sortir des informations que vous lui avez données quelque chose qui n'existait pas avant. Quand vous additionnez tous ces bénéfices, vous obtenez une technologie qui change tout», analyse-t-il.

## La confidentialité des données au cœur des enjeux

Mais la confidentialité des informations injectées dans les *large language models (LLM)*, et celle des informations qui en sortent, est un enjeu hautement stratégique. «La question est de savoir comment garder ses propres savoir-faire dans un contexte comme celui-ci. Parce que si vous demandez à ChatGPT d'analyser votre compte de résultat, par exemple, cela revient à le donner à votre concurrent direct», fait remarquer un participant.

Dans ce contexte, quelles sont les options qui s'offrent aux entreprises? Une des participantes rappelle que nous possédons en France une réelle expertise, avec des start-up, comme LightOn, qui



permettent de déployer son propre *LLM* au sein de ses infrastructures (en mode *on-premise* ou sur un cloud privé). Avec son modèle *open source* Alfred et sa plateforme Paradigm, LightOn fournit en outre aux entreprises des solutions d'IAG garantissant la confidentialité de leurs données.

### Fine-tuning ou RAG? Telle est la question...

Pour aller plus loin encore dans leur démarche et personnaliser les résultats obtenus, les entreprises peuvent mettre en place du *fine-tuning*. Cette pratique consiste à adapter un modèle de langage préentraîné à des tâches et des données spécifiques. Les bénéfices sont nombreux : les projets tirent profit des connaissances accumulées par le modèle, tout en obtenant des résultats plus précis.

Autre option : mettre en œuvre une architecture *RAG* (*retrieval-augmented generation*). Le *RAG* permet au *LLM* d'extraire des informations de sources externes, comme des corpus organisationnels ou des bases de données documentaires internes à l'entreprise. Le *RAG* permet ainsi au modèle de mieux comprendre le contexte des requêtes et de générer des réponses plus pertinentes.

### Quelle place donner à l'IA ?

Mais au-delà des différentes options technologiques qui s'offrent aux entreprises, un acteur français de l'assurance se pose la question de savoir jusqu'où l'IA doit être mise en avant. « La vraie question qui nous taraude aujourd'hui est de savoir quelle place nous voulons vraiment donner à l'IA dans la relation à l'humain – à la fois vis-à-vis des conseillers en relation avec les sociétaires et dans la relation même avec ces derniers. Nous voyons bien qu'il serait aujourd'hui facile d'imaginer un avatar capable de prendre la place d'un conseiller sur les canaux digitaux. Mais est-ce réellement ce que nous voulons? », confie-t-il.

Un autre participant remet, lui aussi, l'IA à sa place : « L'IA doit rester un copilote, c'est-à-dire un composant en apprentissage permanent, situé à côté du pilote, à savoir l'être humain. Il ne faut pas inverser les rôles. L'IA est une aide, un appui,

nous devons constamment rester en mesure de la contrôler et de la maîtriser. »

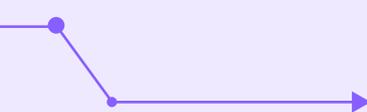
### Le DSI, un garde-fou avant tout

Pour un autre participant encore, nous devons revenir aux fondamentaux : l'IAG n'est que du calcul de probabilité de mots. « Il n'y a aucune capacité de raisonnement et la technologie est soumise à des biais et des hallucinations. Le mot *copilote* est donc extrêmement bien choisi : il s'agit d'aide à l'être humain, pas de quelque chose qui va remplacer l'expert », note-t-il.

Ces réflexions en amènent d'autres, davantage centrées sur les nouvelles compétences qui vont émerger. « Les modèles divergent dans le temps, car les données et les pratiques bougent. Il faut donc développer des compétences capables d'évaluer un modèle, c'est-à-dire : de savoir si je peux lui faire confiance, d'identifier les biais et les moments de divergence. C'est une façon de penser très différente des pratiques IT traditionnelles », précise une des intervenantes.

Le mot de conclusion revient à l'un des participants : « Le rôle de DSI est celui de garde-fou. Nous retournons aux philosophes : 'Qui gardera les gardiens ?' Et attention de ne pas nous prendre pour des gardiens absolus nous-mêmes. »





# Veolia est partie tôt pour tenir le marathon de l'IA générative dans la durée

*Julien Largillière a été nommé chief technology officer (CTO) du groupe Veolia en 2022, quelques semaines avant l'arrivée de la première version publique de ChatGPT. Il a été l'un des deux team leaders d'Alliancy pour mener les discussions auprès des participants de la première édition du Data Excellence Sprint, en s'appuyant sur son expérience unique.*

## Que se passe-t-il quand vous êtes nommé CTO du groupe Veolia en 2022, juste avant le « boom » de l'IAG ?

Quand vous prenez un poste de CTO, comme ce fut mon cas en 2022, on attend de vous une vision et une feuille de route claire. Le groupe Veolia est une grande maison aujourd'hui, avec des technologies hyper-hétérogènes, et ce, d'autant plus depuis le rachat du groupe Suez. Nous comptons 220 000 collaborateurs et il faut pouvoir apporter de la cohérence dans leurs usages technologiques. Ma nomination a correspondu, en effet, au moment où l'IA générative [IAG] était soudainement mise en évidence, mais également au moment où le groupe lançait son nouveau plan stratégique 2024-2027, après la fusion. La dynamique était donc très marquée et positive, avec la nomination récente d'une nouvelle directrice générale en la personne d'Estelle Brachlianoff, mais également la structuration d'une nouvelle direction *Digital business & technology*.

Rapidement, tout le monde a pu constater les premières alertes sur l'usage, par les salariés, des versions publiques de ChatGPT et de Bard [récemment renommé Gemini]. Je me souviens d'avoir reçu des newsletters en interne, nous expliquant à quel point l'outil était génial pour répondre à des appels d'offres ou faire des devis à la volée ! L'urgence était très claire, car personne ne semblait se rendre compte que tout ce qui était partagé sur ces modèles publics, pouvait servir à



Julien Largillière, *group chief technology officer* de Veolia

enrichir les versions suivantes... La priorité était donc d'offrir une perspective qui permette de reprendre le contrôle, mais aussi de faciliter ces nouveaux usages.

## Quelle stratégie avez-vous adoptée ?

Veolia utilise déjà l'intelligence artificielle depuis longtemps. Pour de la maintenance prédictive, de la prévision des volumes d'eau consommés, ou même sur de la reconnaissance d'objets dans le cadre de la gestion des déchets, par exemple. Ces cas d'usage sont déployés et industrialisés. Par ailleurs, l'intelligence artificielle générative en soi n'est pas une nouveauté. Mais « l'effet ChatGPT », c'est d'avoir touché d'un seul coup des dizaines de milliers de collaborateurs au quotidien. Nous avons ainsi atteint près de 30 000 connexions par jour sur les IAG publiques fin 2022. Nous avons donc dû rapidement rédiger des documents de position pour expliquer ce qu'était vraiment l'IAG, et en parler au plus grand nombre, que ce soit sur le terrain ou au sein du comité exécutif. Pour beaucoup de personnes, le fonctionnement de l'IAG reste très nébuleux. Il y a une sorte de mythe

autour du sujet, des raccourcis qui sont pris avec l'impression que tout est possible, y compris avec de la data déstructurée... Les sujets clés de la gouvernance de la donnée, de sa qualité, de sa lisibilité et de son accès ne sont pas des points qui vont de soi : on doit encore trop souvent se battre pour en parler.

## Par quoi avez-vous commencé ?

Nous avons voulu agir très vite sur l'expérience digitale de l'employé : avec de l'acculturation et de la formation, mais aussi avec un accompagnement pour communiquer sur les impacts juridiques potentiels. La direction informatique de Veolia est assez aguerrie dans l'accompagnement des enjeux majeurs du cloud public et de la data... il fallait que nous montrions que nous étions aussi au rendez-vous pour cette « vague IAG ». C'est pour cela que nous avons monté rapidement une offre interne, Veolia Secure GPT, visant à garantir le cloisonnement des informations. Il nous fallait faire une application Web agnostique [capable de fonctionner avec tous les systèmes d'exploitation] pour ne pas être complètement liés à un offreur en particulier ; mais qui s'appuierait sur nos stacks techniques habituelles afin d'aller vite, ne serait-ce que contractuellement. L'idée était d'avoir une fondation solide pour ensuite se garder la possibilité, en fonction des *personas* [types d'utilisateurs] et des cas d'usage, de pouvoir se rattacher aux *large language models* [LLM] que nous souhaiterions.

## Comment s'est déroulé le projet Veolia Secure GPT ?

Nous avons décidé de le lancer fin mars 2023 et il a été livré en juin de la même année, grâce à un fonctionnement en mode « *task force* produit ». Nous sommes rarement allés aussi vite sur un produit de ce genre ! Fin 2023, nous en étions à 13 000 utilisateurs quotidiens, avec 50 000 connexions et 40 millions de requêtes stockées dans BigQuery. Nous continuons aujourd'hui le déploiement avec l'ambition d'adresser tous les « cols blancs » du groupe, dans le monde entier.

## À quels autres enjeux êtes-vous confrontés ?

L'avantage du projet Veolia Secure GPT, c'est

qu'il a permis de nous rendre légitime. Il a fallu ensuite s'organiser pour aller plus loin, afin de servir des ambitions business bien plus précises. L'IAG va devoir jouer tout son rôle dans le plan stratégique de Veolia, notamment en matière de décarbonation et de gestion des ressources. Il faut donc se doter des capacités pour répondre aux besoins, car 2023 a créé un effet de traction très important. Une première étape a été franchie en « augmentant » nos équipes techniques avec l'IAG. Des développeurs aux experts cybersécurité, nous leur avons permis d'accéder à tout l'outillage pour améliorer la qualité du code, générer des storytellings en amont de développement, mieux analyser les *logs*, etc. En quelques mois, c'est entré dans les mœurs. En 2024, l'objectif est de passer à l'échelle en adressant davantage de *personas*.

Un autre point sur lequel nous devons être proactifs, c'est la gestion des offres du marché. On voit, en effet, que tous les acteurs du SaaS embarquent déjà de l'IAG. Mais derrière ces ambitions, il y a souvent une problématique de modèle de prix. Depuis trois ans, les études montrent que l'augmentation de la tarification des SaaS est deux fois plus importante que l'inflation. Et les éditeurs sont loin d'être à l'aise sur les implications de l'IAG en matière de modèle économique, car ils ne savent pas trop ce que cela va leur coûter, en termes de *GPU* [processeur graphique] par exemple. Certains en viennent donc à encapsuler des modèles très légers pour ne pas impacter le coût à la licence... Car les CIO font vite le calcul : quand on doit payer 30 dollars supplémentaires par utilisateur et par mois, les

Après être passé par HSBC, Accor et Edenred, Julien Largillière a rejoint Veolia en 2011, entreprise pour laquelle il a exercé des responsabilités à la fois au niveau de l'exploitation et de l'architecture IT. Il a également dirigé pendant trois ans et demi Veolia Water Information Systems, avant de devenir le chief technology officer du groupe en 2022. Julien Largillière est diplômé de SUPINFO et de l'Institut des Hautes Études de l'Entreprise (IHEE).

effets sur la productivité dans un grand groupe ont intérêt à être clairement démontrés. La proactivité de la direction des systèmes d'information doit donc aussi permettre d'apporter le sceau d'une garantie sur certains outils et usages, afin de créer une dynamique de confiance. Pour beaucoup de collaborateurs, cela va être un gage de valeur alors que les offres vont se multiplier partout.

## Comment mesure-t-on l'apport réel de l'IA générative pour une DSI ?

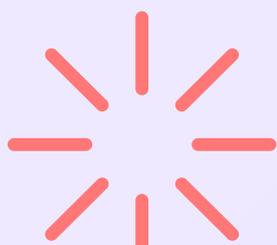
Pour les solutions à base d'IAG chargées « d'augmenter » les développeurs, nous avons surveillé plusieurs types d'indicateurs de performance : la durée des *sprints*, la vélocité, le *time-to-delivery*, le volume de tests unitaires menés, la diminution du nombre de *rollbacks* [retours en arrière] à des fins de sécurité... Aujourd'hui, nous avons une chaîne CI/CD complète [chaîne d'intégration, de distribution et de déploiement continu], sur laquelle nous pouvons observer des indicateurs globaux. Nous avons aussi mené des tests autour de vrais cas d'usage, en comparant les actions d'un développeur « augmenté par l'IAG » à celles d'un autre qui ne l'est pas. Il faut le dire pour ceux qui n'en ont pas fait l'expérience : c'est le jour et la nuit. On est sur des ratios d'environ 70 % de rapidité en plus. Ce que l'on observe, c'est que l'IAG peut aider des développeurs moyens à s'améliorer. Par ailleurs, un bon développeur deviendra lui aussi bien meilleur : plus productif, avec un code mieux documenté, etc. Ce n'est pas une chimère. C'est la vraie vie et c'est impressionnant de voir la vitesse à laquelle ce métier change.

## Que dire aux entreprises qui semblent surtout craindre l'impact de l'IAG en interne ?

On retrouve, dans l'IAG, les mêmes problématiques qu'autour du cloud : *Patriot Act* [loi antiterroriste], Schrems II [arrêt invalidant le « bouclier de protection des données » entre l'Europe et les États-Unis], RGPD [Règlement général sur la protection des données]... À minima, tout doit donc être bien cadré avec le département juridique, au niveau contractuel. L'idée chez Veolia, c'est de ne pas diaboliser par principe : cela fait treize ans que nous sommes sur le cloud public, et tout se passe bien. Par ailleurs, à partir du moment où l'on

entre dans des usages B2B autour de l'IAG, nous sommes très loin des pratiques et des problèmes du « ChatGPT public ». Mais il faut prendre des dispositions plus larges : données anonymisées, *prompts* conservés seulement trente jours sur le cloud, pseudonymisation des *logs* pour pouvoir analyser les principaux usages, etc.

Depuis que notre groupe a mis en œuvre sa stratégie *SaaS-first*, nous avons mis en place une méthodologie globale de « *design & legal assessment* », pour mieux prendre en compte des critères très variés dans l'évaluation des prestataires. Ceux-ci vont de la tarification à la cybersécurité, en passant par la maturité de la structure, ou encore la dépendance économique... Nous nous sommes donné les moyens de « noter » tous les acteurs qui émergent. C'est utile face à la dynamique du marché sur l'IAG. Quant à la pertinence des *LLM* en tant que tels, ce n'est évidemment pas le CTO qui va l'évaluer seul dans son coin. Je m'appuie sur les *data scientists*, les architectes, les développeurs... Sans être sur une équipe à temps plein, nous devons être plus d'une soixantaine à travailler sur tous les aspects du sujet. C'est une « *guilde* » transverse sur l'IAG qui s'est formée, avec des rôles répartis dans l'organisation. Nous n'avons pas l'ambition de créer une entité verticale fermée qui incarnerait la démarche IAG, comme par le passé il y a pu avoir avec la mode du *chief digital officer* dans certaines entreprises, par exemple. Aujourd'hui, on retrouve le numérique partout, cela paraît complètement naturel. Il en sera de même avec l'intelligence artificielle générative.



# L'IA générative, un tournant majeur pour le métier de Bpifrance

*Le DSI de Bpifrance nous livre sa vision quant aux impacts de l'IA au sein des directions informatiques. Il détaille également la feuille de route qu'il a initiée dans son organisation, tout en décrivant les grandes familles d'usages qu'il anticipe à plus ou moins long terme.*

## Les DSI ont-ils réellement pris la mesure de la révolution provoquée par l'IA générative ?

Après avoir échangé avec un certain nombre de mes collègues DSI, je suis assez surpris de constater que la plupart d'entre eux n'ont pas pris suffisamment le temps d'entrer dans une démarche de questionnement sur ce que l'IA générative [IAG] permet de faire et de ne pas faire. Ils ont soit des croyances très fortes, soit l'intime conviction que c'est une technologie de plus, comme la *blockchain*.

J'ai de mon côté la conviction que l'IA générative ne ressemble en rien à la *blockchain*. L'IAG est la dernière version de l'IA qui, elle-même, représente une vague dont l'ampleur est du même niveau que celle des débuts d'Internet. C'est donc un tournant majeur de notre métier.

Certains DSI n'ont, selon moi, pas encore perçu l'impact que cela allait avoir sur leur métier et sur les métiers de leurs métiers. Je suis persuadé que l'IA va, par exemple, changer en profondeur une partie de la médecine et qu'elle va avoir un impact majeur dans le domaine des matériaux, des nouvelles molécules et de l'ingénierie des usines [...]

## Quelles décisions avez-vous d'ores et déjà prises pour vos collaborateurs ?

J'ai décidé de déployer, d'ici la fin de l'année, pour tous mes équipiers, les outils Copilot de Microsoft. C'est, en effet, la première fois que je suis face à un outil qui va potentiellement nous permettre d'obtenir des gains de productivité compris entre 20 et 30 %. Par le passé, certains vendeurs de



Lionel Chaîne, directeur des systèmes d'information de Bpifrance

solutions m'ont fait des promesses qui ne se sont jamais concrétisées. Mais là, même si je ne peux pas en être sûr à 100 %, il est assez facile de constater à quel point nous sommes bousculés. À titre personnel, j'expérimente déjà des gains de productivité avérés compris entre 8 et 10 %, et je n'en suis qu'au premier rang.

Cette décision d'équiper mes collaborateurs des outils Copilot correspond à ce que j'appelle la phase du « Pourquoi ? ». En général, j'aime en effet beaucoup commencer par cette question quand je suis confronté à une nouvelle technologie ou un changement de paradigme. Pourquoi cela fonctionnerait, ou non ? Pourquoi maintenant ?

## À l'inverse, quelles limites voyez-vous à l'IA générative ?

La question que vous posez correspond non plus au pourquoi, mais au comment. Une fois que le

pourquoi a été traité, le comment se pose en effet assez rapidement. Et, force est de constater qu'il existe aujourd'hui énormément de freins. Des freins réglementaires, mais aussi des freins liés à la peur, comme celle de perdre son travail.

À mon avis, la plupart des DSI qui s'intéressent à l'IAG sont confrontés à ces sujets-là. Ils se battent notamment avec leur département conformité pour lui expliquer qu'il est possible de trouver des solutions permettant de minimiser les risques inhérents à cette technologie, notamment à la protection et à la souveraineté des données.

Il faut cependant avoir en tête que l'IAG a jusqu'à présent été entraînée avec les données du Web, soit environ 25 % des données totales disponibles. Or, ces 25 % sont de mauvaise qualité. Les 75 % restants, présents dans les entreprises, ont des niveaux de qualité qui n'ont rien à voir. Le potentiel de l'IAG sur les données d'entreprise est donc trois fois supérieur à ce qui est déjà sorti sur le marché et sans doute beaucoup plus en termes de valeur. Je veux dire par là que les entreprises sont la véritable cible des grands acteurs de ce marché naissant.

Il est donc essentiel que les DSI se concentrent sur ce sujet afin de capter le maximum de la valeur qui va se présenter à eux. Ils doivent se mettre en ordre de marche pour lancer les bons projets au bon moment. Au niveau opérationnel, ils doivent notamment se poser les questions suivantes : quelles nouvelles compétences acquérir, comment gérer les hallucinations de l'IA, faut-il un grand ou un petit LLM [*large language model*], quid des aspects liés au *copyright*, à l'*open source* et à la souveraineté ?

### Quels grands usages de l'IA générative vont émerger pour les DSI ?

Il existe, selon moi, quatre grandes classes d'usages pour les entreprises. J'ai appelé la première l'IA « sauvage ». Elle signifie que les collaborateurs vont utiliser l'IA générative pour créer ou résumer des textes et documents, pour cataloguer, pour segmenter, etc. En un mot, ils vont faire du « chat » sous une forme ou sous une autre pour leurs propres usages. La question qui se pose aux DSI est de savoir s'il faut autoriser ou non ces utilisations de l'IAG, selon quel périmètre de données et selon quelles bonnes pratiques.

La deuxième grande famille d'usages est celle de l'IA « embarquée ». Cette IA correspond à celle que nous trouvons dans certaines solutions comme Microsoft 365 (Copilot), ainsi que dans le CRM ou l'*ITSM* [*IT service management*] des entreprises. Cette IA est relativement simple à gérer et il faut, selon moi, aller assez vite. C'est avant tout un sujet de conduite du changement, d'appropriation. Les questions qu'il faut se poser sont de savoir quels collaborateurs vont pouvoir en bénéficier et si les gains attendus seront au rendez-vous.

Vient ensuite l'IA « hybride », que nous pouvons aussi appeler RAG [*retrieval-augmented generation*]. Elle consiste à utiliser un agent conversationnel avec ses propres données. C'est ici qu'il y a, je pense, le plus de valeur, car nous pouvons produire – via de grands LLM – des IA verticales grâce aux données de l'entreprise. Les données étant entièrement maîtrisées, il n'y a presque plus de problématiques d'hallucination. Les questions qu'il faut se poser sont : comment obtenir des données extrêmement propres et des sets de données de taille suffisante ?

Enfin, le quatrième sujet, qui n'est pas d'actualité pour le moment, concerne les applications que nous pourrions nous-mêmes créer avec de l'IA, en mode *fine-tuning*. Même si ce n'est pas pour tout de suite, il est intéressant de se dire qu'un jour, il y aura des facteurs d'accélération des systèmes et des infrastructures qui nous aideront à y parvenir.

Lionel Chaine est le directeur des systèmes d'information de la Banque publique d'investissement Bpifrance depuis juin 2020. Il a précédemment mené sa carrière au sein du groupe La Poste, où il a notamment été DSI du pôle Colis et directeur du SI finance du pôle Courrier. À ces postes, il a été particulièrement actif sur la transformation agile, à l'échelle IT et métier, mise en œuvre dans le groupe. Lionel Chaine a été formé à CPE Lyon et à l'University College London; il est également titulaire d'un certificat d'administrateur de sociétés, d'administration et de gestion des affaires, de Sciences Po.

# Pour Oracle, les stratégies d'IA générative devront plus que jamais tenir compte du ratio prix-performance

Oracle était partenaire du premier Data Excellence Sprint d'Alliancy organisé fin 2023 et consacré au sujet de l'intelligence artificielle générative (IAG). Philippe Battel, VP artificial intelligence EMEA et Philippe Wojcik, cloud solution engineering director, reviennent sur les échanges qu'ils ont eus à cette occasion avec les directeurs des systèmes d'information présents, et sur l'enjeu de l'IAG pour un acteur systémique dans l'écosystème IT comme Oracle.

## Quel niveau de maturité constatez-vous chez les DSI français quand il est question d'IA générative ?

**Philippe Wojcik.** Cela dépend évidemment de la taille des entreprises, mais, de manière générale, nous avons pu voir que de nombreux acteurs se mobilisent déjà fortement. Les exemples de mise en production de *large language models* [LLM] sont visibles et certaines très grandes entreprises, comme EDF, se positionnent même pour aller jusqu'à créer leurs propres LLM. Autrement dit, les DSI ne sont pas dans la théorie, ils ont mis les mains dans le cambouis depuis plusieurs mois. C'est d'autant plus vertueux que les questions auxquelles ils doivent répondre sont nombreuses. Par exemple : comment réconcilier les applications d'entreprises avec l'IAG ? Et, comment utiliser les données de l'entreprise pour nourrir ces IAG efficacement ?

**Philippe Battel.** Un autre sujet qui revient fréquemment dans les échanges, et de façon légitime, est : comment mesurer le succès des projets d'IAG ? S'agit-il de se concentrer sur les gains de productivité ou le niveau d'adoption des outils d'IAG ? Faut-il multiplier les indicateurs de performance ? De manière générale, tout le monde s'entend, au sein des DSI, pour pointer une quinzaine de grands cas d'usage clairs pour l'IAG. Par exemple : l'automatisation de tâches administratives, la création automatique de



Philippe Battel, VP artificial intelligence EMEA chez Oracle

campagnes marketing ou l'analyse de milliers de documents financiers. Cela dépasse, de loin, l'usage de ChatGPT pour les entreprises, grâce à la naissance d'outils d'IAG spécialisés sur une industrie ou des processus horizontaux (tels que les RH ou la finance) et avec l'emploi d'API qui permettent d'intégrer les IAG de centaines de fournisseurs (face à une seule avec ChatGPT). Les expérimentations se multiplient, et les comités de direction commencent à identifier des indicateurs de performance pertinents pour le succès ou l'échec des projets ; ce qui permet de passer à un déploiement à grande échelle.

## Pourquoi vous paraît-il important que les DSI se mobilisent sur ce point ?

**Philippe Battel.** Ils vont jouer un rôle central dans l'adoption des solutions d'IAG dans l'entreprise. Tout d'abord, pour éduquer et obtenir le support de leur comité exécutif, en leur faisant comprendre l'impact croisé que l'IAG va avoir tant sur l'IT que sur les *business units*. Les nombreux DSI avec lesquels nous travaillons nous racontent qu'ils ne peuvent pas s'opposer au caractère « viral » de la diffusion de l'IAG. Ils décident, pour la grande majorité, de fournir un environnement informatique permettant l'identification des cas d'usage



Philippe Wojcik, cloud solution engineering director chez Oracle

prioritaires, l'expérimentation et la diffusion de bonnes pratiques (notamment en matière de cloisonnement et de gestion des risques). Un environnement qui favorise l'accompagnement au changement et la formalisation de la gouvernance qui va accompagner cette nouvelle vague technologique majeure pour toutes les entreprises.

### En quoi le fait de mettre l'IAG au cœur des applications est selon vous le grand défi 2024 des DSI ?

**Philippe Wojcik.** Il s'agit d'amener de la valeur au métier, sans se contenter de faire du ChatGPT. Donc naturellement, c'est l'intégration et la mesure des bénéfices de l'IAG par rapport à la réalité des environnements de chaque entreprise qui va faire la différence. Est-ce que l'IAG va rester cantonnée à une simple amélioration des processus ou est-ce que l'on va parvenir à passer un cap sur des cas d'usage métier ? La façon d'intégrer l'IAG dans les applications va être un révélateur important de la relation IT-métier pour les mois à venir.

**Philippe Battel.** Il y a un risque que l'IAG n'apporte pas de gain réel pour l'entreprise, surtout s'il n'y a pas, en amont, un travail pour identifier les cas d'usage les plus pertinents. Le plus important est donc d'inclure l'IAG dans la stratégie IA globale de l'entreprise, notamment

en incluant la gouvernance autour des données – sujet extrêmement stratégique – et l'intégration avec les applications existantes (qu'elles soient industrielles, fonctionnelles ou techniques).

### Ces dernières années, les entreprises sont passées successivement de sujet en sujet : data warehouse, data lake, data mesh... Est-ce que l'IAG va vraiment permettre de passer un « cap data » ?

**Philippe Battel.** Pour qu'une solution d'IAG soit pertinente, il faut avoir des données organisées, de qualité et en volume. Tous les « chantiers data » sont aujourd'hui à nouveau prioritaires pour les CEO en raison de leur impact potentiel fort, que ce soit en matière d'innovation ou d'optimisation. Les entreprises viennent nous voir toutes les semaines avec des scénarios IAG, certains simples et d'autres très avancés. Par exemple, nous travaillons avec plusieurs sociétés qui veulent utiliser l'IAG pour gagner plusieurs mois, voire années, sur le lancement d'un nouveau médicament. Pour rappel, le coût d'un médicament est de plus d'un milliard d'euros, et les brevets pharmaceutiques sont délivrés pour vingt ans, avec une durée de création de dix à quinze ans. Des gains de temps dans le processus de R&D et sur les essais cliniques sont donc stratégiques. Pour les comités exécutifs, on ne fait pas « de la data pour de la data », mais on fait atterrir ces projets data structurants pour répondre à des cas d'usage sur l'IAG.

**Philippe Wojcik.** Beaucoup de DSI travaillent depuis longtemps sur ces fondations data. Souvent, ils nous disent ensuite que la difficulté vient du temps mis pour démontrer le retour sur investissement pour les métiers. L'un des atouts de l'IAG pour ces DSI, c'est de pouvoir beaucoup mieux montrer à leurs interlocuteurs la pertinence des investissements data. Plus les données seront propres, moins le projet IAG coûtera cher. Par conséquent, pour réussir, la pertinence de l'entraînement de l'IA sur les besoins métiers est un facteur tout aussi important que la volumétrie des données.

### Quels changements a provoqués le « boom de l'IAG » pour la stratégie d'Oracle ?

**Philippe Battel.** Cela fait déjà de nombreuses années qu'Oracle intègre l'intelligence artificielle

dans ses produits, que ce soit dans les bases de données, les solutions spécialisées IA (comme les *chatbots*), au sein de nos applications métiers, ou sur notre offre Oracle Cloud Infrastructure (OCI) en partenariat avec NVIDIA. Pour les bases de données, Oracle Autonomous Database s'autogère (correctifs, etc.) et offre plus de trente algorithmes de *machine learning* pour des cas d'usage comme la maintenance prédictive ou la détection de fraude. Quant aux applications Oracle Apps, elles offrent des dizaines de «copilotes» d'IAG permettant aux entreprises d'automatiser des centaines de processus métiers (ressources humaines, vente, finance) ou d'avoir des indicateurs prévisionnels, par exemple, sur la gestion des stocks. Tout ceci sans quitter l'application métier Oracle, sans migrer ni donner accès aux données critiques des apps, et donc en offrant une meilleure sécurité. De nombreux clients nous font confiance comme BBVA, HSBC, Aon, Twist, National Grid, iSIZE, Etisalat et des centaines d'autres.

Mais aujourd'hui, la brique la plus importante de notre stratégie est OCI Generative AI, bâtie en collaboration avec la start-up Cohere. Il s'agit d'un service géré qui permet aux utilisateurs d'intégrer des LLM dans leurs applications via une API – que ce soient les modèles Cohere, Llama de Meta ou un autre LLM. Cette approche est prévue pour s'intégrer facilement avec AI Vector Search, une fonctionnalité d'Oracle Database qui permet la RAG [*retrieval-augmented generation*]. En effet, il est nécessaire pour beaucoup d'entreprises de pouvoir combiner des LLM préentraînés avec leurs propres données, afin d'augmenter la pertinence des réponses fournies par l'IAG sur leurs domaines de spécialisation. C'est un moyen fort pour s'emparer des usages de l'IAG tout en préservant la sécurité et la confidentialité des données. Et, pour nous, c'est un vrai facteur de différenciation.

**Philippe Wojcik.** Par ailleurs, OCI Generative AI nous permet l'ajout de fonctionnalités IAG dans nos applications métiers. Nombreuses sont les personnes à ne pas avoir encore conscience que l'on embarque déjà de la sorte aujourd'hui ce «ChatGPT industriel» dans les workflows de nos applications, par exemple, dans Oracle Human Capital Management (HCM) ou bien dans Oracle Advertising and Customer Experience (CX). Notre approche est de proposer de l'IAG

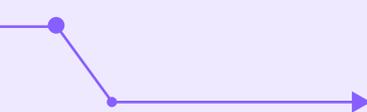
directement, en achetant des modèles, mais aussi de proposer aux entreprises d'en faire elles-mêmes, en utilisant les bonnes infrastructures. Oracle a facilité le multicloud ces dernières années, notre stratégie IAG s'inscrit dans cette continuité. Nous voulons faire en sorte que l'on puisse librement et facilement s'emparer du sujet de «l'IAG à l'échelle», en étant agnostique des plateformes, quel que soit le fournisseur de cloud.

## Quelle perception ont les DSI de ce positionnement d'Oracle sur l'IAG ?

**Philippe Battel.** Nous avons déjà beaucoup de clients, mais nous avons aussi conscience que nous ne sommes pas encore toujours attendus sur ce sujet. Ce que nous voulons montrer, c'est que le ratio prix-performance que nous proposons est aujourd'hui le meilleur du marché. Lorsque nous sommes dans une phase de *proof of concept*, le retour des clients est très positif. De plus, nous avons de plus en plus de projets d'IAG de type LLM, déployés notamment sur l'infrastructure Oracle AI (GPU). Vu la dynamique de l'IAG aujourd'hui, la croissance des usages, l'adaptation rapide du marché et le coût d'un déploiement à grande échelle, cette proposition de valeur nette, en matière de prix-performance, prend un poids de plus en plus important. Les nouveaux leaders de l'IA, quels qu'ils soient, vont avoir besoin de capacité pertinente et bon marché.

Avec xAI par exemple, Elon Musk construit ses LLM chez nous ! Et OCI est utilisée par Microsoft afin d'inférer des modèles d'IA optimisés pour les recherches conversationnelles de Bing. En mars 2023, OCI était déjà devenu le premier fournisseur de cloud à très grande échelle à proposer NVIDIA DGX Cloud. Depuis, la *marketplace* intègre aussi NVIDIA AI Enterprise pour permettre aux entreprises d'effectuer des entraînements sur modèle volumineux pour leurs applications d'IAG. En résumé, il faut percevoir que l'IAG pour Oracle s'étend de l'infrastructure aux applications métiers, en passant par l'amélioration du *fine-tuning* des entreprises.





# L'AI Ops, intimement liée au déploiement de l'observabilité, analyse Splunk

Comment accompagner les DSI dans les transformations provoquées par la montée en puissance de l'IA? Quels choix technologiques privilégier? Les réponses avec Stéphane Estevez, EMEA director of product marketing, observability & IT markets de Splunk.

## De quelle façon Splunk accompagne-t-elle les DSI dans leur intégration des dernières avancées de l'IA ?

La thématique de l'IA est présente chez Splunk depuis 2015, principalement à travers le *machine learning* [l'apprentissage automatique, en français]. Historiquement, notre rôle a été d'aider les DSI à devenir plus efficaces dans leur métier. Pour adresser les enjeux liés à la cybersécurité, par exemple, nous avons mis en place des solutions d'analyse comportementale. Sur le volet «AI Ops» [artificial intelligence for IT operations], nous avons déployé du *machine learning* pour les aider à nettoyer le très grand nombre d'événements qui se produisent et à réaliser des analyses prédictives sur les incidents.

Récemment, nous avons mis à jour nos algorithmes pour absorber les nouvelles volumétries de données liées aux microservices, aux conteneurs et aux technologies *serverless*. Malgré tout, il existe encore beaucoup de direction des systèmes d'information qui n'ont pas terminé leur démarche «AI Ops». Certes, ces dernières mettent en place des actions pour réduire le bruit sur la gestion des incidents. Mais, avec l'avènement d'environnements de plus en plus distribués et dynamiques, elles s'aperçoivent que les équipes de production ou les équipes «cyber» ne peuvent plus absorber les volumes de données actuels (logs, traces, métriques, événements, etc.).

Ce phénomène est intrinsèquement lié aux dernières tendances de l'observabilité. Obtenir toutes les données nécessaires pour savoir ce



Stéphane Estevez, EMEA director of product marketing, observability & IT markets de Splunk

qui se passe à l'intérieur du système génère un tsunami de data. D'où le renforcement de l'approche *machine learning*. Plus vous faites de l'observabilité, plus vous avez besoin d'AI Ops.

## En tant que plateforme, comment vous positionnez-vous par rapport à ces enjeux liés à l'IA ?

Splunk est effectivement née comme plateforme data. Cela signifie que nous n'avons créé des cas d'usage (sécurité, IT, *DevOps*, *IoT*, etc.) que dans un second temps. Malgré le fait que nous proposons de l'IA «prête à l'emploi», c'est-à-dire embarquée dans ces différents cas d'usage, notre approche de type «plateforme» permet également de faciliter l'accès à la donnée par les *data scientists* de nos clients. Ils peuvent ainsi se concentrer sur la création des algorithmes d'IA – sans trop se préoccuper des étapes de collecte, de normalisation et de préparation des données.

Autre tendance de fond : les utilisateurs en interne ont envie d'interfaces « *ChatGPT-like* ». Nous nous efforçons donc de simplifier la consommation de données grâce à des assistants. Nous avons récemment lancé un « assistant IA » qui permet de chatter avec Splunk. Au lieu de lancer des requêtes, n'importe quel ingénieur peut désormais discuter en langage naturel avec Splunk pour faire avancer ses projets. Cette démocratisation de l'accès aux données cible aussi les métiers. L'enjeu est de permettre à tout « *citizen data scientist* » d'élaborer ses propres tableaux de bord, par exemple.

### Comment Splunk aide-t-elle les DSI à transformer leur gouvernance pour industrialiser leurs projets ?

Il existe énormément d'outils, aujourd'hui, dans les directions des systèmes d'information, liés à la transformation de leur gouvernance. Il y a quelques années, une étude que nous avons réalisée nous a permis d'en identifier entre trente et cinquante, selon les entreprises. Dans ce contexte, obtenir une vue de bout en bout de son système d'information est très compliqué. Et, pour les DSI, présenter à leur comité de direction une vision à 360 ° de leur système d'information, c'est presque mission impossible.

Grâce à l'approche « plateforme », il est possible de fédérer cette vue à la fois du côté technique et au niveau des processus métiers ; ces derniers pouvant commencer par une application dans le cloud et finir dans SAP, en passant par une étape d'IloT [*Industrial Internet of things*].

### Auriez-vous un cas client à partager avec nous ?

Lors du dernier symposium de Gartner, nous avons présenté un cas client, celui de BMW. Cet acteur majeur du marché automobile est en pleine transformation de son modèle économique : demain, vous ne négocierez plus vos prix avec le concessionnaire, mais directement avec le constructeur. Pour gérer efficacement cette migration, il faut synchroniser à la fois l'environnement omnicanal de configuration des véhicules, le parcours lié au financement (MyBMW Financial Services) et la partie industrielle – pour vérifier si le véhicule a bien été fabriqué et quand il pourra être livré.

Dans cet exemple, Splunk est intervenue – en tant que plateforme – soit pour aller chercher elle-même les données, soit pour se connecter aux outils existants. Le but était d'élaborer un tableau de bord permettant de vérifier si le processus global de configuration, financement et livraison des véhicules fonctionnait, mais également de suivre le chiffre d'affaires associé et de surveiller qu'aucune interruption de service ni aucun ralentissement n'ait lieu. Cela impliquait des dizaines d'applications différentes, qui ne discutaient pas entre elles, mais que l'on peut maintenant « désiloter » et rendre « observable » via Splunk.

### Quelles priorités en matière de choix technologiques les entreprises doivent-elles activer ?

Quelle que soit leur approche de l'IA, le carburant nécessaire au moteur, c'est la donnée. Je recommanderais donc de s'intéresser à OpenTelemetry. Ce projet *open source* devient progressivement le standard de facto de la collecte de l'ensemble des données nécessaires à l'observabilité. Avec lui, vous n'êtes plus dépendant d'un éditeur, vous pouvez librement choisir chez qui vous envoyez vos données, que ce soit Splunk ou un autre acteur du marché.

Par ailleurs, le code étant accessible, vous pouvez aller y jeter un œil en cas de doute autour de la sécurité ou de potentiels biais. Enfin, le dernier avantage est que vous réalisez des économies, car ce type de solution consomme beaucoup moins de ressources qu'une configuration où vous avez un agent par outil.

Mon conseil est donc de démarrer avec OpenTelemetry, puis de créer une infrastructure propre pour collecter la donnée et avoir la maîtrise totale de ses données. Et je dirais aussi qu'il est nécessaire de se renforcer sur l'*AIOps*, cela permet notamment d'obtenir des gains rapides pour améliorer la résilience digitale de son entreprise.



# À propos

**ORACLE**

Oracle propose des suites d'applications intégrées ainsi qu'une infrastructure sécurisée et autonome dans Oracle Cloud. Pour en savoir plus sur Oracle (NYSE : ORCL), rendez-vous sur [www.oracle.com/fr/](http://www.oracle.com/fr/)

**splunk**>

Splunk contribue à la construction d'un monde numérique plus sûr et plus résilient. Les organisations comptent sur Splunk pour éviter que les problèmes de sécurité, d'infrastructure et d'application ne deviennent des incidents majeurs, résister aux effets néfastes des perturbations numériques et accélérer leur transformation numérique. Pour en savoir plus, rendez-vous sur [www.splunk.com](http://www.splunk.com)

 **ippon**

Créé en 2002, Ippon Technologies est un cabinet de conseil et d'expertise en technologies, international et indépendant. Ippon accompagne la transformation numérique des entreprises, en les aidant à concevoir leur stratégie et à déployer leur *roadmap* à l'échelle, afin de délivrer rapidement la valeur attendue. Guidé par une énergie collective et un goût prononcé pour le challenge, Ippon Technologies a plus de 700 salariés passionnés de tech dans le monde. Pour en savoir plus, rendez-vous sur [ippon.tech](http://ippon.tech)

**Alliancy**

Média sur la transformation numérique des entreprises, engagé en faveur du « plus forts ensemble », Alliancy sélectionne pour vous les témoignages et les retours d'expériences de femmes et d'hommes actifs dans une dynamique d'innovation. Vision stratégique, gouvernance et modes d'organisation, utilisation de la donnée, enrichissement et animation d'écosystèmes (innovation, ressources humaines, systèmes d'information, finances et achats...), cloud, sécurité et pilotage du système d'information... Nous nous intéressons à ces différents leviers de la transformation pour qu'ils s'invitent peu à peu dans votre quotidien. Stimulez votre curiosité, gagnez du temps et identifiez les clés de la réussite de votre entreprise. Pour en savoir plus, rendez-vous sur [alliancy.fr](http://alliancy.fr)



# Alliancy

Mentions légales et droit d'exploitation  
Alliancy - 32, rue des Jeûneurs - 75002 PARIS  
SARL au capital de 167 550 €  
792 635 138 R.C.S. Paris  
[alliancy.fr](http://alliancy.fr)

Directeur de publication : Sylvain Fievet  
Coordination rédactionnelle : Dorian Marcellin et Fabrice Deblock  
Journalistes : Dorian Marcellin et Fabrice Deblock  
Design graphique : Coralie Fau  
Relecture : Virginie Gaillard Mesplarau

Mars 2024

Toute reproduction des textes publiés dans ce carnet est interdite sans autorisation explicite de la rédaction. Pour tout renseignement, vous pouvez adresser vos questions à l'adresse suivante : [redaction@alliancy.fr](mailto:redaction@alliancy.fr)

