

# L'IA générative en entreprise à l'épreuve du réel

EMERTONDATA



## LIVRE BLANC

JUIN 2024



"**Les limites de notre langage signifient les limites de notre propre monde**", affirmait Wittgenstein. En matière d'IA, il est indéniable que l'apparition des grands modèles de langage (LLM) a fait reculer, en quelques mois seulement, les frontières du possible dans le monde de l'entreprise et au-delà, puisque **cette révolution** a atteint le grand public à une vitesse inimaginable il y a seulement deux ans.

**L'évolution des modèles** et leurs performances comparées est désormais un sujet d'actualité quasi-quotidien : Mistral semble surpasser Claude, qui talonne lui-même le dernier né d'Open AI... L'apparition de nouvelles fonctionnalités innovantes dans les services des GAMAM, ainsi que les investissements records, qui atteignent des centaines de millions de dollars, nourrissent une considérable cacophonie qui empêche parfois de distinguer ce qui relève du long terme, voire de la science-fiction, de ce **qui est bien réel et tangible**.

C'est pourquoi, le temps de ce livre blanc, nous proposons de mettre sur pause les discussions sans fin sur le potentiel de valeur, la vitesse de l'innovation, et les fantasmes sur la singularité et le remplacement de l'humain, pour nous intéresser aux **questions concrètes qui se posent quotidiennement dans les entreprises**, qu'il s'agisse de nos fleurons industriels ou de nos startups.

Nous avons mené une enquête auprès de nombreux acteurs du secteur, que nous avons choisi d'identifier clairement, afin de fournir une photographie la plus contextualisée possible de l'IA générative dans les entreprises. Les résultats sont édifiants : un an et demi après l'apparition de ChatGPT, toutes les entreprises interrogées ont industrialisé des cas d'usage d'IA générative, et **80 % d'entre elles indiquent que la priorité est à la formation à grande échelle des collaborateurs**. Toutefois, le foisonnement des idées et la prolifération des applications qui en découle fait peser un risque sur l'atteinte d'un véritable bénéfice économique à court terme.

Ce livre blanc, qui rassemble de nombreux témoignages, vise à répondre à plusieurs **questions fondamentales** :

- Comment faire les meilleurs choix de technologies et de cas d'usage ? A quels besoins métiers répondre **en priorité** ?
- Quels **"paris"** faire sur les IA génériques ou spécialisées et cœur de métier ?
- Comment assurer **le contrôle, la robustesse et la confiance** dans ces nouvelles technologies ?
- Comment **préparer les organisations** à l'évolution du travail, par une intégration profonde de l'IA générative dans les processus décisionnels et fonctionnels ?

Nous espérons que ce livre blanc fournira des **réponses claires et des orientations précieuses** pour permettre aux entreprises de naviguer efficacement dans ce nouveau paysage économique, tellement changeant et, il faut le reconnaître, tellement enthousiasmant.

Bonne lecture.



**Aimé Lachapelle,**  
Managing Partner,  
Emerton Data



**Perle Bagot,**  
Associate Director,  
HUB Institutue



**Yannick Léo,**  
Partner & Director of Data  
Science,  
Emerton Data

# SOMMAIRE

---



## 01

EDITO

---



## 02

SYNTHÈSE

---



## 03

ENTRETIENS ET TABLES  
RONDES

1. Construire la prochaine pépite : stratégies pour l'essor des startups en IA générative
  2. Ecosystème open source et IA générative : insights de Hugging Face
  3. Quels sont les cas d'usage d'IA générative à prioriser pour maximiser l'impact ?
  4. Vers une IA souveraine et plus durable ? Exploration des défis et enjeux
  5. Future of Work : demain, tous "augmentés" avec l'IA générative ?
  6. L'IA générative est-elle déjà un facteur de transformation profonde des banques et assurances ?
  7. Comment assurer la robustesse et la conformité des IA génératives ?
-

# SYNTHÈSE

L'ensemble des interviews que nous avons conduites, la diversité des profils, des entreprises sondées nous ont permis de parvenir aux conclusions suivantes :

- **L'IA générative représente déjà un fort vecteur de transformation pour les entreprises.** En particulier, la quasi-totalité de nos intervenants déclare avoir déjà industrialisé un cas d'usage impliquant l'IA générative, et tous envisagent de le faire à brève échéance. De plus, interrogés sur la vitesse à laquelle cette technologie sera adoptée par leurs équipes, tous ont estimé que l'IA générative sera pleinement intégrée dans les opérations quotidiennes de leurs collaborateurs sous 3 ans, la moitié d'entre eux pensant même que cette adoption sera effective sous 1 an.
- **L'IA générative trouve un large spectre de débouchés pour nos interlocuteurs.** L'ensemble des principales fonctions des entreprises semblent devoir être touchées par la transformation induite par l'IA générative. Nous constatons également cette grande diversité, s'agissant de ce qui pousse les responsables data & IA interrogés à investir sur cette technologie. Si les gains de productivité constituent leur motivation la plus fréquente, sont cités ensuite, à la fois amélioration de l'expérience client, meilleure expérience collaborateur et aide à l'innovation produit.
- **L'ensemble de nos intervenants témoigne d'une attitude positive et optimiste de leurs collaborateurs envers l'IA générative,** eux-mêmes qualifiant régulièrement leur propre jugement sur cette technologie comme "enthousiaste" et "pragmatique". Interrogés explicitement sur le sujet, aucun de nos intervenants n'a confié voir l'IA générative comme un risque pesant sur l'emploi français.
- **Nos intervenants ont également identifié d'importants besoins de formation et d'accompagnement, sujet que presque tous considèrent comme très important dans leur feuille de route des prochains mois.** Le premier volet de cet accompagnement concerne évidemment les chatbots généralistes, comme ChatGPT, qu'une proportion déjà significative (entre un quart et la moitié) du personnel des entreprises sondées semble déjà utiliser.
- Enfin, **les principaux points d'attention identifiés par nos intervenants sont la sécurité des données, l'éthique et les biais de modèles, et les coûts de mise en oeuvre de ces technologies.** Nous notons à cet égard que de nombreuses versions internes et sécurisées de ces chatbots, appelées SecureGPT, ont déjà été déployées au sein des grandes entreprises françaises.

# ENTRETIENS ET TABLES RONDES

## 1. Construire la prochaine pépite : stratégies pour l'essor des startups en IA générative

**Guillaume Avrin**, Coordinateur national pour l'intelligence artificielle

**Vincent Rapp**, Responsable Sectoriel Numérique @ *Bpifrance*

**Emmanuel Papadacci-Stephanopoli**, Chief Product Officer @ *La Fabrique by CA*

**Pascal Simon**, Investment Strategy @ *Kaukana Ventures*

## 2. Ecosystème open source et IA générative : insights de Hugging Face

**Yoach Lacombe**, ML Engineer @ *Hugging Face*

## 3. Quels sont les cas d'usage d'IA générative à prioriser pour maximiser l'impact ?

**Harriet Rees**, Chief Information Officer @ *Starling Bank*

## 4. Vers une IA souveraine et plus durable ? Exploration des défis et enjeux

**Olivier Abecassis**, CEO @ *Qwant*

**Catherine Dubeau**, Responsable département data science et gouvernance @ *Safran Aircraft Engines*

**Lamiae Azizi**, Professeur @ *College of Computing, UM6P*

## 5. Future of Work : demain, tous "augmentés" avec l'IA générative ?

**Renata Alves Santos**, Deputy Chief Data Officer @ *TotalEnergies*

**Maxime Marembaud**, Group Chief Data & Digital Officer @ *Sodexo*

**Claire Lebarz**, Data & AI Officer @ *Malt*

## 6. L'IA générative est-elle déjà un facteur de transformation profonde des banques et assurances ?

**Ahmed Besbes**, Chief Data Officer @ *Direct Assurance*

**Maxime Havez**, Chief Data Officer @ *Crédit Mutuel Arkéa*

## 7. Comment assurer la robustesse et la conformité des IA génératives ?

**Jean-Marie John-Mathews**, Ethics Researcher & Co-founder @ *Giskard*

# CONSTRUIRE LA PROCHAINE PÉPITE: STRATÉGIES POUR L'ESSOR DES STARTUPS EN IA GÉNÉRATIVE



**Guillaume Avrin**

Coordinateur national pour l'intelligence artificielle



Le gouvernement a annoncé un plan d'investissement de près de 4 milliards d'euros pour soutenir les technologies numériques souveraines. Quelle est votre stratégie pour l'essor des start-ups françaises en IA générative, et quelle part de ce budget leur sera réservée ?

Au niveau de la stratégie, nous essayons de réunir l'ensemble des ingrédients nécessaires pour développer des IA génératives compétitives en France :

- Nous devons nous assurer que l'écosystème français dispose de bonnes infrastructures de calcul. Nous avons notamment financé l'extension **du supercalculateur Jean Zay, un super calculateur public et gratuit hébergé par le CNRS**. Tout le monde peut s'en servir pour entraîner des modèles d'IA générative, la seule condition étant de publier, ce qui ne signifie pas que le modèle doit être open source. C'est un point crucial, car ces entraînements de modèles peuvent coûter des millions d'euros sur des infrastructures classiques. Nous sommes en train d'intégrer ce supercalculateur dans un "continuum" de calcul en lien avec les acteurs privés du cloud, comme Scaleway, OVHcloud et OUTSCALE, **pour s'assurer que l'ensemble des acteurs ait accès à une puissance de calcul suffisante**, aux différents stades du cycle de vie d'une IA : recherche, développement, adaptation et *fine-tuning*, déploiement, maintien à jour.
- Au niveau des données, la France coordonne une initiative européenne portée par 16 Etats-membres pour **regrouper les bases de données et l'ensemble des ressources nécessaires à l'entraînement d'IA génératives** : des données qui valorisent le patrimoine national, représentative des différentes langues et des spécificités culturelles de l'Union européenne. C'est important pour nos entreprises locales, car les entreprises extra-européennes entraînent leurs modèles sur des données qui peuvent ne pas répondre parfaitement à nos besoins.
- Nous soutenons également deux dispositifs opérés par Bpifrance :
  - Le premier vise à créer **des communs numériques de l'IA générative**, réunissant toutes les briques technologiques partagées au sein de l'écosystème IA, par les développeurs et les utilisateurs – les modèles de fondation, les modèles fine-tunés, les API et les outils permettant l'évaluation des modèles ;
  - Le second, appelé **"Accélération des usages de l'IA générative"**, finance la mise en place de partenariats entre les développeurs et les utilisateurs de l'IA générative, pour accélérer la diffusion de cette technologie.

Nos grandes écoles et universités ont une très bonne qualité de formation et de recherche, mais les talents formés ici sont évidemment tentés de rejoindre les GAMAM ou autres grandes start-ups de l'écosystème américain. Comment s'assurer de les retenir en France ?

Je vais peut-être me faire l'avocat du diable, mais c'est **un sujet complexe**. L'analyse même de la situation, sur les impacts positifs ou négatifs de ces passages par les entreprises américaines, n'est pas simple.

Rappelons d'abord que l'ensemble des cofondateurs des principales entreprises IA françaises – Mistral AI, H, etc. – sont passés par des entreprises américaines. Ainsi, d'une certaine façon, **le passage par une Big Tech contribue à la formation de nos talents français de l'intelligence artificielle**. Si on juge les start-ups à l'aune des fonds qu'elles ont réussi à lever, le meilleur parcours est clairement de profiter de la formation d'excellence en mathématiques et en ingénierie que l'on a en France, puis de travailler quelques années chez un GAMAM, et enfin revenir créer son entreprise ici.

Le principal intérêt de passer par les GAMAM est que les ressources de calcul mises à disposition des profils tech sont sans commune mesure avec ce qu'on trouve ici. Un entrepreneur français venu créer son entreprise d'IA en France après un passage chez OpenAI m'a confié qu'il disposait dans son ancienne entreprise de l'équivalent de l'ensemble du supercalculateur Jean Zay. **Les personnes qui ont accès à de telles puissances de calcul obtiennent une expérience du supercalcul à l'échelle pour l'entraînement des modèles d'IA extrêmement rare et prisée**. On dit qu'il y a 50 talents de l'IA générative dans le monde, et ce sont en fait les 50 personnes qui ont eu accès à des supercalculateurs pendant un temps suffisamment long.

Et parmi ces 50 personnes, il y a des Français qui sont revenus créer des start-ups en France. La réponse n'est donc pas évidente. On veut bien sûr conserver nos talents, mais **je veux aussi insister sur l'importance de rapatrier des talents qui présentent un réel intérêt pour notre écosystème**. Nos principaux concurrents, les Etats-Unis, la Grande-Bretagne et la Chine, ont une politique très permissive vis-à-vis des déplacements de leurs talents, et il serait étonnant qu'on adopte ici une attitude contraire.

On dit souvent que l'argent est la chose la plus importante pour faire revenir nos talents, mais ce n'est pas totalement vrai : **il faut aussi leur donner les ressources nécessaires, notamment de la puissance de calcul, pour qu'ils puissent continuer à travailler ici** sur des sujets à la pointe de la technologie.

Le dernier point, c'est la simplification : il faut que nos meilleurs talents ne soient pas absorbés par des tâches, des procédures administratives qui n'ont rien à voir avec leurs recherches, **il nous faut faciliter l'obtention de financement, simplifier le cumul d'activités** pour faciliter les allers-retours entre les sphères publique et privée.



## Vincent Rapp

Responsable Sectoriel Numérique – Coordination IA –  
Direction de l'Innovation



### Quelle est la vision de Bpifrance de l'impact sur l'économie française portée par l'IA ?

On constate dorénavant que cet impact est grandissant, mais la technologie se heurte quelque peu à un tissu industriel et économique qui **n'est pas encore assez digitalisé**. Or, la digitalisation est le premier pas vers l'utilisation de l'IA générative. Une très récente étude Bpifrance montre d'ailleurs que 70% des patrons de PME n'utilisent pas l'IA et 60% n'envisagent pas de le faire à court terme. **L'adoption est donc encore largement perfectible**. La formation des patrons est, à cet égard, un élément important de la stratégie de Bpifrance. A l'inverse, l'IA générative peut, dans un certain nombre de secteurs, être un incitateur à la digitalisation des entreprises.

Cela étant, nous constatons qu'après de nombreux projets lancés rapidement autour de l'IA générative, **les demandeurs temporisent quelque peu leurs développements**, car les ROIs ne sont pas encore garantis, et les mises en œuvre souvent assez complexes. L'offre doit se structurer, avec notamment plus de solutions sur l'étagère.

### Êtes-vous optimiste quant au positionnement français sur le sujet? Quelles sont pour vous les conditions permettant la création de pépites français dans l'écosystème de l'IA générative?

Bpifrance est par nature optimiste, c'est la condition pour prendre du risque et investir!

La condition clef pour permettre la création de ces pépites, c'est d'abord le talent. Nous avons la chance d'avoir en France une très bonne formation, notamment en mathématiques, mais cela ne suffit pas. Il manque ici **une expérience d'opérations, des process**, qui est beaucoup plus présente aux Etats-Unis, notamment grâce aux GAMAM. Ce n'est d'ailleurs pas un hasard si les entreprises qui marchent le mieux ici sont celles dont les fondateurs (ou C-level) sont passés par les Etats-Unis, ou sont passés par des incubateurs ayant cette culture.

De plus, **l'IA nécessite des investissements considérables**. Tout d'abord, des moyens financiers, et la France est de ce point de vue très bien placée en Europe, en termes de puissance capitalistique, notamment grâce à France 2030, même si elle reste encore largement en-dessous des Etats-Unis. L'IA requiert également **des données, de la puissance de calcul**. Réfléchissons donc à notre législation pour mettre à disposition assez de données aux entreprises qui souhaitent créer des LLMs, par exemple. Pour ce qui est de la puissance de calcul, le sujet est sans doute à traiter au niveau européen.



Le développement de l'IA, c'est aussi **la structuration de l'écosystème**, un lien à établir entre l'offre, d'une part, et la demande, le tissu économique, d'autre part. Sur ce point, nous sommes très en retard sur les Américains, qui prennent beaucoup plus de risque et chez qui ce lien se construit de façon beaucoup plus fluide.

Je vais revenir sur la question des investissements. Bpifrance poursuit son rôle de soutien public à l'innovation. En revanche, je trouve que la **France a beaucoup de retard sur l'investissement privé**. Même nos plus grandes pépites – comme Mistral AI, très soutenue par l'écosystème public français – lèvent une grande partie de leurs **fonds privés à l'étranger**. C'est sans doute le point sur lequel je suis un peu pessimiste: la capacité des investisseurs privés et des grands groupes à prendre le relais de l'investissement public, et à soutenir le développement de nos start-ups.

Quelles sont pour vous les qualités et critères essentiels pour envisager d'y investir ?

Encore une fois, **les qualités personnelles des fondateurs et des premiers employés** de la start-up sont un élément déterminant dans notre processus de décision.

Au-delà de ce simple critère, le positionnement de la start-up sur la chaîne de valeur de l'IA est très important. Nous ne financerons pas 15 entreprises comme Mistral AI, qui demande énormément de capitaux. Nous cherchons notamment **à combler les failles, les faiblesses de l'écosystème français existant**, par exemple sur le hardware – pas sur les GPUs, marché dominé par Nvidia, mais sur les CPUs et les middlewares.

Les grandes entreprises se mobilisent pour évaluer l'impact de l'IA générative sur leur activité et prioriser des premiers cas d'usage, mais comment s'assurer que les ETI, PME et petites entreprises ne seront pas les laissés-pour-compte de cette révolution ?

Je suis convaincu que 90% des usages finaux, de la transformation que l'IA permettra dans les entreprises se jouera dans ce tissu de petites et moyennes entreprises et des ETI. Il est donc vital de mettre en place des plans **d'accompagnement et de diffusion de l'IA**, et c'est ce que nous avons commencé à faire avec le programme **IA Booster France 2030**, qui s'adresse spécifiquement aux PME et ETI de tout secteur d'activité.

Face à la volatilité des profils et à l'appétit des grands acteurs américains, ne craignez-vous pas que ces startups peinent à constituer leurs équipes et atteindre donc leurs objectifs?

Pour le moment, le contexte nous est plutôt favorable. Beaucoup de chercheurs, qui étaient partis travailler à l'étranger, notamment aux Etats-Unis, cherchent à **revenir en France**. Les raisons en sont multiples. Tout d'abord, **les salaires qu'offre la France redeviennent compétitifs sur le marché**. Mais, au-delà de cet aspect, il y a donc beaucoup de turnover aux Etats-Unis, beaucoup ne sont pas entièrement satisfaits de leur situation là-bas: la concurrence entre développeurs est très rude, la qualité de travail n'est pas parfaite, et beaucoup de profils seniors peinent, dans les énormes équipes qui se sont constituées là-bas, à comprendre la vision globale de l'entreprise et à trouver leur place au sein de celle-ci. Certains hauts profils acceptent de baisser significativement leur salaire pour **retrouver cela ici**.

Cependant, cela reste très compliqué pour beaucoup de start-ups françaises de **recruter les meilleurs profils**, très opérationnels ou ayant une très forte expertise scientifique.



**Emmanuel Papadacci-Stephanopoli**  
Chief Product Officer

Comment la Fabrique by CA s'est-elle structurée au sein du Crédit Agricole, et quel est son modèle opératoire ? Comment passez-vous de l'identification d'une opportunité à une décision d'investissement et de lancement d'une nouvelle start-up ?

La Fabrique by CA est le start-up studio fintech du groupe Crédit Agricole. Le principe des start-up studios existe depuis une dizaine d'années. Sa principale spécificité est que **nous sommes à la fois investisseurs et créateurs d'entreprises**, là où, dans les incubateurs notamment, l'idée initiale est portée, apportée par les fondateurs.

Cela nous permet donc de mutualiser certaines fonctions entre les start-ups et d'accumuler une certaine expérience de création d'entreprises, ce qui **se reflète sur le taux de survie de nos start-ups : 80% de survie à 3 ans**, là où l'on se situe plutôt normalement autour de 10%.

Nos actionnaires sont les entités du groupe Crédit Agricole, ce qui nous donne deux particularités par rapport aux autres studios : les start-ups doivent s'inscrire dans la stratégie du groupe CA, et ont nécessairement, dès le premier jour de leur existence, des **synergies avec un grand groupe bancaire** (une start-up n'est lancée que si un partenariat interne est déjà fortement pressenti), alors que la priorité de beaucoup de start-ups est plutôt leur valorisation, avec la perspective d'un rachat.

Cependant, les start-ups ne travaillent pas seulement avec le groupe : une partie importante du chiffre d'affaires de nos start-ups se fait à l'extérieur du groupe. Cette capacité à aller chercher des contrats sur le marché est un indicateur très important de la pérennité de la start-up.

Envisagez-vous de structurer un pôle d'excellence en IA générative, mutualisant l'expertise pour chacune des start-ups de votre écosystème ?

L'IA est un sujet que je suis depuis mon master sur le sujet en 2002. A l'époque, les technos étaient déjà disponibles, mais assez peu déployées – les seuls débouchés concrets étaient en médecine et sur les jeux vidéos. En ce qui concerne les services financiers, il y avait jusque très récemment deux types d'acteurs : les fintechs qui basent tout leur produit sur l'IA, et celles qui, à l'inverse, sont très en retard sur le sujet.

Lorsque je suis arrivé à la Fabrique by CA, j'ai donc voulu rattraper ce retard, en **complétant l'équipe tech avec des profils IA, pour positionner l'équipe beaucoup plus fortement sur le sujet.**

Ensuite, l'IA générative est arrivée. De mon point de vue, cette technologie est très plug and play. Je ne veux donc pas tant créer un pôle d'excellence « académique » sur l'IA générative, mais plutôt une capacité et un réflexe chez les CTO des start-ups de la Fabrique : **restez connectés, comprenez la tech, testez et exploitez les derniers modèles avant les concurrents.**

Un exemple très concret : nous explorons l'opportunité de créer une start-up qui vise à aider les personnes en situation de dépendance à comprendre à quelles aides, provenant de différents niveaux administratifs ou sociaux, elles ont le droit. Beaucoup de personnes ont déjà essayé de résoudre ce problème par des techniques classiques, avec des arbres de décision. Il me paraît évident que c'est l'IA générative qui permettra de résoudre ce use case, en exploitant la masse d'informations textuelles émanant de l'ensemble des acteurs sociaux du pays.

J'ai un autre exemple intéressant : notre dernière start-up (GENIAL, ex-CibleR) aide des acteurs territoriaux à déployer des programmes d'argent fléchés sur le territoire, en particulier pour les collectivités publiques et pour les banques régionales. Son rôle consiste à orienter ces fonds, et nous avons rapidement réalisé que l'IA générative facilitait la mise en relation et la sensibilisation des bénéficiaires éligibles aux avantages auxquels ils ont droit. Cette technologie permet de le faire **de manière naturelle et personnalisée en utilisant le langage naturel, auprès d'un grand nombre de commerçants et de lieux culturels sur un territoire donné.** La start-up rencontre d'ailleurs un beau succès, et a notamment gagné un appel d'offres d'Atout France pour France.fr, site qui accompagne les Jeux Olympiques, afin de déployer un module d'accompagnement qui offrira des suggestions touristiques personnalisées aux touristes qui visiteront le site.

Nous ne créerons jamais une start-up dédiée à l'IA générative. Ce n'est pas notre métier, mais **nous voulons exploiter cette technologie pour résoudre des problèmes métier spécifiques.** Nous avons 12 start-ups actuellement, je suis persuadé qu'une majorité d'entre elles utiliseront très prochainement un algorithme d'IA générative « off-the-shelf », sur l'étagère.

Rétrospectivement, quels ont été les autres facteurs de succès des start-ups technologiques que vous avez créées ?

Quant aux facteurs du succès, ils ne concernent pas uniquement l'IA. Pour qu'une start-up fonctionne, il faut être bon sur tous les aspects : la tech, le réglementaire, le go-to-market, l'offre, le suivi, la finance. Si je reviens sur la data, une start-up qui n'a pas compris que la data n'est pas la clef de son modèle économique, qui dépend trop d'un tiers pour cet aspect de sa technologie, finira tôt ou tard par devenir trop dépendante de cet acteur tiers.

**L'IA devient une norme aujourd'hui, une obligation,** ce qui n'était pas encore le cas il y a deux ans.



**Pascal Simon**  
Investment Strategy



Kaukana Ventures est un start-up studio, fondé par Emerton Data. Pouvez-vous nous présenter son fonctionnement, ce qui vous a poussé à créer ce start-up studio, et une première success story issue de ses rangs ?

Pour nous, l'élément fondamental est notre logique d'écosystème. Kaukana Ventures fait partie d'un groupe qui comprend à la fois des équipes de conseil en stratégie, et une practice de conseil data & digitale. Nous sommes convaincus que la combinaison et la disponibilité immédiate des expertises scientifique, technique et business que nous avons développées, et l'écosystème que nous avons créé sur ces sujets peuvent **aider nos start-ups à accélérer, à trouver leur positionnement sur le marché, et à accéder à un vaste éventail d'options technologiques possibles**. Sur certains secteurs d'activité, comme l'industrie ou l'énergie, l'expérience que nous avons accumulée et notre proximité avec les principaux acteurs du marché sont un facteur clé de réussite pour l'adoption de nos solutions par le marché. Nous pouvons nous appuyer sur **notre équipe de consultants en stratégie, data scientists et software engineers** pour parvenir à cette rapidité d'exécution et pour prendre rapidement les bonnes décisions.

Nous avons **un fonctionnement hybride, assez agnostique**, dans le sens où nous avons déjà creusé, développé et accéléré certaines opportunités business à partir de nos propres idées, alors que, dans d'autres cas, nous collaborons avec des entrepreneurs qui viennent à nous avec leurs propres projets, nous discutons ensemble des sujets sur lesquels il serait pertinent de joindre nos forces, et nous travaillons alors main dans la main pour accélérer ces premières idées.

Notre première expérience, qui est aussi une vraie success story, est **Agrisight**. Agrisight développe un logiciel de logistique end-to-end pour l'agriculture – le secteur céréalier en particulier – qui permet d'optimiser à la fois les marges opérationnelles et l'empreinte carbone des acteurs agroalimentaires. Ce projet **a été incubé chez Kaukana Ventures** jusqu'au développement d'un premier produit et l'obtention de premières références commerciales, puis nous avons laissé la start-up se joindre à une autre agritech pour continuer sa croissance.

Kaukana Ventures est un projet hybride aussi parce qu'il arrive assez fréquemment que des consultants du groupe Emerton travaillent pour le start-up studio et en fassent profiter leur expertise. La concurrence est rude pour attirer les meilleurs data scientists ; à cet égard, cette opportunité qu'ont vos data scientists de travailler pour Kaukana Ventures représente-t-elle un atout ?

Nous croyons fermement que cette perméabilité, ces passerelles laissées entre ces deux types de parcours et de compétences sont un avantage, un élément différenciant valorisé par nos consultants. Les data scientists apprécient cette possibilité d'acquérir de l'expérience sur différents "tracks", sans avoir à choisir une spécialisation définitive. Nous pensons qu'il est important pour des profils récemment diplômés **d'être exposés à différents types de projets**, d'intervenir sur différents rôles avant de s'engager dans un "track" en particulier.

Nous leur apportons dans un même écosystème **l'opportunité de se former et d'entreprendre**. Certains de nos employés pourront ensuite prendre un **rôle plus entrepreneurial**, en prenant les rênes d'une start-up - c'est comme ça, et ça peut d'ailleurs devenir une opportunité de partenariat pour Kaukana Ventures.

Les volumes d'investissements dans l'IA sont astronomiques, avec par exemple un récent tour de seed à 200 millions d'euros. Pensez-vous qu'il existe une bulle de l'IA ?

Bien sûr, je n'ai pas de boule de cristal, mais je vais répondre par une métaphore : **je ne pense pas qu'on puisse parler de bulle, je préfère plutôt parler de vague**. La vague est probablement un peu haute, un peu gonflée, mais lorsqu'une vague redescend, elle ne se retire jamais complètement. En conclusion, je dirais qu'il s'agit surtout d'une question de timing, et qu'il faut savoir laisser cette vague passer.

# ECOSYSTÈME OPEN SOURCE ET IA GÉNÉRATIVE : INSIGHTS DE HUGGING FACE



**Yoach Lacombe**  
ML Engineer



**Hugging Face**

Pouvez-vous nous présenter Hugging Face ?

Notre but est **promouvoir et démocratiser une IA éthique**, en développant une plateforme sur laquelle les acteurs privés ou publics de l'IA peuvent héberger et diffuser leurs modèles d'IA, ainsi que les datasets associés, du code de démonstration, des publications académiques sur le sujet, etc.

Vous êtes passé par les grandes écoles françaises, et vous travaillez maintenant pour Hugging Face, une entreprise américaine. Pouvez-vous nous expliquer les raisons de votre choix, ce qui vous a attiré chez Hugging Face ?

Tout d'abord, il y a bien sûr un certain nombre de considérations matérielles : l'aspect salarial, mais aussi les moyens ITs mis à disposition, en particulier le nombre de GPUs.

Mais au-delà de ça, je dois dire que l'engagement pris par Hugging Face sur l'**open source** est un élément très important pour moi, et le fait que, en réalité, l'activité d'Hugging Face consiste à **donner au grand public les moyens de communiquer sur leurs travaux, leurs modèles**. Je pense que c'est un point sur lequel beaucoup d'entreprises françaises devraient travailler – laisser leurs employés faire la publicité de leurs travaux. C'est tout l'inverse chez Hugging Face, Mistral AI, ou dans les grandes entreprises américaines, et d'ailleurs, une partie de mon travail consiste en fait à promouvoir mon travail et **construire mon propre portfolio**. Ce point est vraiment très important aujourd'hui pour les chercheurs comme les entrepreneurs, c'est un élément déterminant de l'attractivité des entreprises AI.

Effectivement, je pense que cette réflexion intéressera les grandes entreprises françaises. Pouvez-vous nous en dire plus sur l'ensemble des outils que Hugging Face apporte à la communauté IA ?

Tout d'abord, il y a bien sûr le **hub des modèles d'IA**, qui est la fonctionnalité phare de Hugging Face. Le hub s'adresse vraiment à tous les membres de la communauté IA, qu'il s'agisse des créateurs de modèles (comme les chercheurs) ou des utilisateurs de ces modèles. L'éventail de ce qu'on peut héberger sur Hugging Face est très large : les checkpoints des modèles, leur licence commerciale permissive, des jeux de données associés, le code ayant permis d'entraîner les modèles et les bibliothèques externes qui permettent leur entraînement et leur diffusion.

Nous proposons également un **ensemble d'interfaces permettant de tester et de jouer avec les modèles que nous hébergeons**. On s'adresse encore une fois à tous les types de profils, même ceux qui ont très peu d'expérience avec l'IA. Ces interfaces de démonstration sont gratuites, ne nécessitent pas de puissance de calcul externe, et peuvent gérer de nombreux types de modèles – génération d'images, modèles de langage, les derniers modèles de Mistral AI, par exemple –, développés par des équipes du monde entier. Grâce à ces interfaces, **chacun peut maintenant rester à jour des dernières avancées en IA générative**.

Le dernier point que je veux mettre en exergue est que **nous mettons gratuitement à disposition des GPUs pour des acteurs publics**, comme des chercheurs, qui autrement n'auraient pas les fonds pour accéder à la puissance de calcul nécessaire à l'entraînement de leurs modèles. Ils peuvent donc utiliser ces GPUs pour promouvoir leurs travaux et développer leurs propres cas d'usage, qui bénéficieront en retour à l'ensemble de l'écosystème.

Beaucoup de gens pensent l'IA générative est en quelque sorte une technologie "plug-and-play", donc je voudrais vraiment vous demander : est-ce que c'est si simple et quelles compétences cela requiert-il de construire un modèle d'IA générative, modèle de fondation ou spécialisé, et de le déployer ?

C'est vraiment très complexe ! Considérons à la fois l'entraînement et le déploiement. **L'entraînement demande une grande expertise technique**, qu'on construit par beaucoup de recherches internes et d'expérimentation. Cela est très chronophage, parce qu'il y a en fait beaucoup de choses à apprendre, de petites subtilités techniques à comprendre. Je vais vous donner un exemple. Avant d'envoyer les données à un modèle d'IA, nous devons transformer chaque donnée sous une forme que le modèle pourra comprendre, soit une suite de nombres. On appelle ce processus la **tokenization**, et, bien qu'il soit relativement simple de tokenizer du texte, il est beaucoup plus compliqué de le faire avec d'autres types de données, comme des fichiers audio. Comment peut-on transformer de façon optimale un fichier audio en cette suite de nombres que le modèle peut gérer ? Les équipes de recherche de Meta ont travaillé pendant des années sur cette simple question, et c'est uniquement la première étape pour créer des modèles d'IA générative sur des données audio !

Un autre défi lié à l'entraînement est que les modèles doivent être entraînés sur des milliers de GPUs en parallèle. **Comment coordonner tous ces GPUs**, en particulier, comment le gradient de l'optimisation doit être transmis à chacun des GPUs ? Il y a plein de défis de la sorte, et l'expertise opérationnelle, le nombre d'expérimentations nécessaires sont de plus en plus grands : c'est pourquoi **seuls quelques experts** sont désormais capables d'entraîner des modèles à grande échelle, comme chez Mistral AI ou Llama.



Une fois que le modèle est entraîné, **maintenir le modèle en production est encore plus complexe**, et nécessite beaucoup de ressources humaines, mais je dirais qu'on a ici besoin d'ingénieurs au profil un peu plus traditionnel. Gérer un grand nombre de requêtes en parallèle sur un grand nombre de GPUs reste néanmoins difficile, et ce n'est qu'une partie du problème, et c'est ce qui fait toute la valeur d'entreprises comme Mistral AI ou Open AI aujourd'hui.

# QUELS SONT LES CAS D'USAGE D'IA GENERATIVE À PRIORISER POUR MAXIMISER L'IMPACT ?



**Harriet Rees**  
Chief Information Officer



Pouvez-vous détailler votre stratégie pour prioriser les projets d'IA générative ?

Notre objectif principal est d'identifier des cas d'usage qui **améliorent les expériences clients et rationalisent, accélèrent nos processus internes**. Nous évaluons constamment les possibles cas d'application, en travaillant en étroite collaboration avec diverses unités commerciales pour comprendre les défis et les objectifs qui leur sont propres. Cette approche collaborative garantit que nos projets seront en phase avec les besoins réels de nos équipes et seront en mesure de leur fournir des avantages tangibles.

Comment parvenez-vous à estimer l'efficacité de ces projets, compte tenu de la nature relativement nouvelle de la technologie d'IA générative ?

Nous utilisons une combinaison de métriques quantitatives et qualitatives. Pour les projets axés sur l'efficacité, nous observons simplement les indicateurs de performance déjà existants et les comparons aux valeurs qui faisaient référence avant le déploiement du projet. Pour les initiatives visant à améliorer la qualité de nos services, **nous nous appuyons sur les retours des utilisateurs, des évaluations d'experts et des procédures de test rigoureuses**. Nous prenons également en compte des facteurs tels que la précision du modèle, sa fiabilité et sa capacité à gérer les cas limites.

Compte tenu de l'évolution rapide de l'IA générative, un déploiement rapide est-il toujours l'objectif principal ?

Bien que nous ayons tout à fait conscience de l'importance de rester à la pointe de l'innovation, nous pensons également qu'il **est nécessaire de laisser sa place à une expérimentation réfléchie, rationnelle**. Nous encourageons donc nos équipes à explorer de nouveaux outils, de nouvelles technologies, pour favoriser une culture de l'innovation et de l'apprentissage au sein de l'entreprise. Nous pensons que cette approche nous permettra de construire des solutions plus robustes et plus impactantes à long terme.

Pouvez-vous partager des exemples de défis que vous avez rencontrés lors du déploiement de l'IA générative ?

A force de pratique, nous avons compris que certains de nos cas d'usage ne doivent pas nécessairement être résolus par de l'IA générative, en particulier, concernant notre expérience client, ceux impliquant des scénarios complexes ou très sensibles. Nous avons pris délibérément la décision d'exclure certaines applications de notre périmètre actuel et de **nous concentrer sur des domaines où cette technologie peut vraiment exceller.**

Quels sont les éléments de décision à avoir en tête pour déterminer si un modèle d'IA générative est prêt à être déployé ?

Pour nous, **il est primordial de comprendre dans le détail les capacités du modèle** : ce qu'il peut faire, et ses propres limites. Cela suppose donc un processus intensif de test et de validation, pour garantir la précision et la fiabilité de nos modèles. Un autre critère primordial pour nous est l'adoption des modèles par les utilisateurs, et les retours que ceux-ci nous adressent. Nous nous efforçons de créer des solutions intuitives et parfaitement intégrées aux flux de travail existants.

Quels sont les principaux domaines sur lesquels votre équipe se concentre pour l'année à venir ?

Notre objectif est de développer **un assistant général pour notre personnel, intelligent et interactif, capable d'interagir de façon autonome avec son utilisateur**, et de lui fournir un support de classe mondiale. En parallèle, nous explorons d'autres applications qui exploitent les LLMs et leur compréhension pour offrir la meilleure expérience possible à nos clients.

# VERS UNE IA SOUVERAINE ET PLUS DURABLE ? EXPLORATION DES DÉFIS ET ENJEUX



**Olivier Abecassis**  
CEO



Pouvez-vous nous présenter Qwant, et ce qui fait sa spécificité parmi les moteurs de recherche ?

Comme quasiment tous les moteurs de recherche grand public, notre modèle est basé sur la publicité. Cependant, **le respect de la vie privée étant dans notre ADN**, nous n'avons aucun contexte sur nos utilisateurs, car nous ne stockons aucune de leurs données, aucun historique. Deux personnes différentes effectuant la même recherche auront donc les mêmes résultats, et les mêmes publicités, quel que soit leur historique. Nous nous sommes construits en opposition au modèle américain d'ultra-tracking des données personnelles. Notre vision est que, même si nous n'arriverons jamais à être aussi bons que Google, **nous pouvons apporter des résultats de recherche suffisamment pertinents à l'utilisateur**, qui doit prendre conscience que la qualité de Google s'achète par l'utilisation de ses propres données personnelles.

De plus, **les données sont traitées sur des serveurs français d'OVHcloud**, et nous avons développé une surcouche qui garantit le respect de la vie privée des utilisateurs, même lorsque nous utilisons des technologies étrangères, comme Bing, auquel nous allons fortement réduire notre dépendance en fin d'année. Je crois que les entreprises européennes ne doivent pas s'interdire de travailler avec la tech US pour des raisons de souveraineté, mais plutôt utiliser cette tech à bon escient, pour ne pas être à la traîne, **en rajoutant les surcouches de confidentialité nécessaires en amont**.

Quelle est la stratégie de Qwant vis-à-vis de l'IA générative ? L'IA générative représente-t-elle un atout ou un défi pour vous ?

L'IA générative représente un atout pour Qwant, car elle permet aux moteurs de recherche de passer d'une **recherche thématique** (j'obtiens des pages relatives à un thème donné), à **une interaction de type question/réponse**. Nous avons donc introduit dans notre page de recherche un module qui permet de répondre à un certain nombre de requêtes sous forme de question/réponse. **L'IA générative nous a permis d'enrichir considérablement notre expérience utilisateur**.

Avez-vous un premier retour d'expérience depuis votre lancement de cette nouvelle fonctionnalité, qui intègre GPT3.5 d'Open AI ?

D'abord, comment de quoi s'agit-il? Il s'agit d'apporter, plutôt qu'une liste de pages web, **une réponse à une question posée par l'utilisateur**, bien sûr en utilisant en interne notre moteur de recherche. Lorsque nous recevons une requête, nous commençons par identifier les documents pertinents grâce à notre moteur de recherche, puis nous ajoutons un moteur de pertinence, qui nous permet d'identifier les documents les plus pertinents. Ensuite, nous utilisons un LLM qui va générer une réponse à partir de ces documents. Aujourd'hui, nous utilisons GPT-3.5, et demain, des modèles comme Mistral ou LLama. Nous devons quand même filtrer certaines requêtes (comme des demandes de suicide, etc.) et **ne pas générer trop d'hallucinations**, ce qui est par exemple un défaut de certains modèles LLM.

Le niveau de réponse est très satisfaisant et les retours sont très positifs. En revanche, le coût de la réponse est élevé, ce qui nous oblige à **optimiser le traitement des requêtes**, en travaillant sur la sélection des données, et en se bornant à apporter des réponses courtes.

Qwant a-t-il développé d'autres briques d'IA générative dans son moteur de recherche ?

Oui, nous apportons notamment **un résumé des pages citées en résultat de recherche**, qui est généré par IA générative. Il est important de ne pas perdre de vue la cohérence produit de Qwant. Nous restons prudents et souhaitons ne déployer que **des solutions vraiment abouties**, pour conserver la confiance des utilisateurs. Par ailleurs, nous essayons de réduire notre dépendance à Microsoft Bing d'ici la fin de l'année, et l'IA nous y aidera énormément.

Votre proximité avec OVHcloud vous donne-t-elle un avantage sur ce critère de souveraineté?

Évidemment, OVH nous aide énormément. C'est un service cloud très orienté sur la souveraineté – **tous ses serveurs sont en France** – et sur la durabilité, **les data centers d'OVH étant les moins consommateurs en électricité de tous les fournisseurs cloud**.

C'est un partenaire très important, qui vient en complément de la surcouche logicielle permettant d'éviter la fuite des données. Il faut être pragmatique sur la souveraineté. La souveraineté totale, en France, sera toujours un vœu pieux, car on n'en a pas toujours les moyens. Il nous faut donc avancer progressivement.

Enfin, quel regard portez-vous sur les enjeux de durabilité ? Quelle est votre vision sur ce sujet?

Actuellement, les moteurs de recherche maintiennent leur connaissance du web en crawlant, en scannant constamment l'ensemble des adresses IP disponibles. Cela est **extrêmement énergivore**, d'autant que 80% des documents du web sont du spam, et que les éditeurs de LLM eux-mêmes scannent de plus en plus le web pour se constituer des bases d'entraînement.

Nous pensons que cette méthode n'est plus adaptée, et cherchons à développer un modèle alternatif, dans lequel **les sites pousseraient eux-mêmes leurs nouvelles pages** ou leurs modifications vers les moteurs de recherche. L'économie, en argent et en électricité, est énorme, car il n'y a d'échange de données que si cela est nécessaire. Il y a une vraie prise de conscience des acteurs sur ce sujet, qui ont d'ailleurs le même intérêt que nous. **Le passage à ce modèle de push nous paraît, ainsi qu'à d'autres acteurs de l'industrie, inévitable.**



**Catherine Duveau**

Responsable de Département Data Science  
& Governance



Pouvez-vous nous présenter les principaux enjeux actuels pour Safran Aircraft Engines ?

**Safran Aircraft Engines est un motoriste mondial majeur**, avec trois missions principales : faire voler nos clients en toute sécurité, offrir à chacun un environnement de travail sûr et motivant, et mobiliser nos équipes en tant qu'acteurs de notre transformation et notre responsabilité société et environnementale.

Actuellement, **quatre activités en croissance sont à mener de front** : dans le domaine civil avec une augmentation de livraison des moteurs LEAP, dans le domaine militaire avec le triplement de la cadence de production du M88 entre 2022 et 2027 pour soutenir les livraisons France et export, dans les services avec le quadruplement de l'activité LEAP entre 2022 et 2027, et en R&T pour préparer le futur de la société en développant les programmes Rise et SCAF.

Pour accompagner cette croissance et relever ces différents défis, de nombreux recrutements ont eu lieu ces dernières années.

En quoi l'IA peut être utile, en particulier pour une entreprise comme Safran Aircraft Engine ?

Les applications de l'IA sont très variées, et touchent pour beaucoup à notre cœur de métier, à la fois sur les produits et les processus : **assistance à la fabrication et au contrôle de nos pièces, analyse de nos essais, de nos calculs, maintenance prédictive, logistique** (prévision moyen-terme de la demande commerciale), etc. Ensuite, de manière générale, l'IA, intégrée dans les outils du quotidien, représente une source non négligeable de gains de productivité pour l'ensemble des métiers et activités.

Nous avons, au sein la direction technique, des équipes ayant des compétences fortes en IA. **Ces compétences sont nécessaires, car elles nous permettent de bien prioriser les sujets sur lesquels nous souhaitons investir**, de pouvoir répondre aux exigences réglementaires à venir autour de l'IA de confiance et aussi d'être force de proposition pour l'innovation. Cette équipe participe également à la **montée en compétence interne de nos collaborateurs sur l'IA**, à travers des formations, et des sessions de coaching.

Pour développer ces compétences, nous nouons des partenariats, dans le monde académique, mais aussi avec des startups et d'autres sociétés.

Pouvez-vous nous expliquer comment vous abordez le sujet de l'IA générative chez Safran Aircraft Engines et où en êtes-vous dans votre feuille de route ?

Il y a deux types de projets au sein de l'engineering : d'abord, nous avons déployé **des solutions clefs en main** au sein de Safran, comme Copilot ou un chatbot sécurisé ; en parallèle, nous travaillons sur un certain nombre de cas d'usages internes pilotés par la valeur, qui utilisent des solutions open source. Nous avons initié une dizaine de preuves de concept, et en avons déjà industrialisé certaines.

Par ailleurs, nous établissons actuellement une feuille de route dans laquelle certains cas d'usage sont très classiques (liés au texte, au code), et d'autres assez spécifiques, **liés notamment à notre activité de conception de pièces**.

La question du « make or buy » est très importante pour nous. Pour y répondre, il nous faut développer une vision précise du marché et des opportunités techniques actuelles et futures.

Dans le secteur de l'aéronautique, la souveraineté est un enjeu majeur, en particulier dans le domaine militaire. Pourriez-vous nous rappeler pourquoi la souveraineté est cruciale et comment vous y avez répondu jusqu'à présent en termes d'infrastructures et de données ?

Sur le domaine militaire, il faut évidemment **une infrastructure IT dédiée** répondant à des exigences plus strictes.

Au-delà du militaire, se pose la question de la confidentialité des données et de la maîtrise des modèles pour répondre aux exigences de l'IA Act, qui sont déclinées par l'EASA pour l'aéronautique. Nous devons disposer d'infrastructures adaptées à chaque niveau de sensibilité des données : par exemple, nous devons garder **une traçabilité très stricte des traitements et entraînements réalisés sur chaque donnée** pour certains cas d'applications.

L'IA générative est arrivée et de nombreux cas d'usage se présentent. Avez-vous défini des règles sur la mise en place de ces cas d'usage ? Quels sont les défis techniques que vous anticipez pour assurer la souveraineté ?

Nous appliquons nécessairement les mêmes règles de gouvernance que sur le reste des données et applications IA. Il faut être encore plus vigilant avec l'IA générative.

Concernant les aspects techniques, l'IA générative requiert encore plus besoin de GPUs que les applications IA classiques. Par ailleurs, ces modèles étant construits, entraînés sur des quantités énormes de données, **se pose donc**, en parallèle de la protection des données, **la question de la protection des modèles**.



Avec tous ces défis, pouvons-nous dire que la souveraineté est synonyme de retard dans l'innovation?

Il est vrai que, ne pouvant pas utiliser les solutions cloud dans tous les cas, nous sommes dans une certaine mesure ralentis dans nos déploiements. Mais cela est également l'occasion de se poser quelques bonnes questions : comment ces modèles sont structurés ? quelles sont nos priorités ? comment les utiliser de façon plus parcimonieuse ? **Cela nous force à monter en compétence nos équipes techniques**, et à développer une vraie stratégie sur ces sujets

Dans notre domaine, tous nos compétiteurs sont soumis aux mêmes exigences – même si les Américains conservent un avantage, car ils disposent d'AWS et d'Azure.

A l'inverse, la souveraineté est-elle synonyme de durabilité? Quelle est votre vision sur le sujet?

Oui, car, ces enjeux de souveraineté nous poussent à être **beaucoup plus conscients de ce que nous consommons**. Nous devons toujours nous demander : où ces modèles ont-ils le plus de pertinence ?

Cela étant, cette question doit être traitée dans sa globalité, celle de la durabilité du numérique. C'est encore un sujet émergent, mais qui finira par devenir un sujet d'importance dans les prochaines années.



**Lamiae Azizi**  
Professeure College of Computing

Pouvez-vous nous présenter l'UM6P, votre rôle et évoquer la place de l'IA au sein de l'UM6P ?

L'UM6P, Université Mohammed VI Polytechnique, est **une institution dédiée à la recherche et à l'innovation**. Créée en 2017, elle regroupe 6000 étudiants, provenant principalement du Maroc et du continent africain, et vise à se positionner parmi les universités à renommée mondiale, à l'avant-garde du développement de l'Afrique sur certains domaines-clefs, comme **l'IA, la sécurité alimentaire, l'industrialisation, le développement durable, les sciences comportementales et sociales**.

L'université est **très engagée pour le développement durable**, atteignant le niveau Silver en 2020, et Gold en 2023 du classement STARS, un protocole d'évaluation de la durabilité des universités.

En ce qui me concerne, après avoir occupé différents postes académiques et industriels et mis en place des écosystèmes de collaboration en recherche et développement sur l'IA et technologies dérivées entre les universités, institutions publiques et le privé en Australie, en Grande-Bretagne et en France, j'ai rejoint récemment **le College of Computing, qui vise à développer l'éducation et la recherche sur l'ensemble des sciences du numérique** : IA et science des données, cybersécurité, modélisation et simulations, télécommunications, etc. Ce collège a une place centrale dans la stratégie de l'UM6P, et compte parmi ses rangs de nombreux Marocains, binationaux et étrangers formés dans des grandes universités internationales à la renommée mondiale dans leurs domaines. En parallèle, la création **du centre « AI Movement »**, ouvert en 2021 par l'UM6P à Rabat, **et reconnu par l'UNESCO comme établissement de catégorie 2**, concrétise encore davantage cette ambition.

En quoi, pour un pays comme le Maroc et des institutions comme l'UM6P, la souveraineté sur l'IA est un sujet important ?

La nature des enjeux est probablement la même ici que dans d'autres pays, mais le Maroc a fait, encore plus qu'ailleurs, de la souveraineté des données et de l'IA un sujet majeur. Les lois de protection et échange de données au Maroc sont assez strictes, imposant un traitement local des données. Ceci impose donc **de disposer d'une infrastructure et de solutions locales**, afin de permettre le développement d'un marché local de solutions IA, et une stratégie de formation adaptée avec la réglementation en place.

A cet égard, **l'UM6P dispose de son propre datacenter** et a mis en place le supercalculateur Toubkal, le plus puissant d'Afrique, afin de permettre aux différents acteurs de l'écosystème marocain et africain (entreprises, startups ou institutions publiques) **d'entraîner localement leurs propres modèles d'IA**.

Quelle est la place de la formation et de la culture des talents dans la quête de l'innovation et de la souveraineté ? Comment l'UM6P s'y prend pour faire émerger ses talents et les retenir, et répondre à ces enjeux ?

**La formation et la culture des talents, ici comme ailleurs, en France notamment, est la clef de la réussite en termes de souveraineté et d'innovation.** A ce propos, on pense d'abord à la formation initiale, qui se concentre sur ceux qui développeront ces technologies. Mais il ne faut pas perdre de vue également la formation, ou reformation, des professionnels actuels, dont les métiers seront, a minima, profondément transformés par les innovations des prochaines années – et c'est aussi sur cette reconversion que l'IA générative a un rôle à jouer. A l'UM6P et au College of Computing, **nous offrons notamment un ensemble de formations continues et exécutives qui répondent à ces besoins**, comme la formation « IA pour les leaders », lancée récemment par le College et qui permet aux cadres exécutifs de différentes entreprises et institutions publiques de bien appréhender l'IA, ses implications concrètes, notamment en ce qui concerne la communication avec leurs équipes techniques.

A travers cette formation et bien d'autres, l'UM6P vise à développer **un pôle de rayonnement pour le Maroc, mais également pour l'ensemble de l'Afrique.** Offrir aux Marocains et aux Africains une formation d'excellence ici, au Maroc, puis accompagner les chercheurs ou étudiants qui souhaitent entreprendre, créer leurs solutions ici, est **la meilleure façon de conserver ces talents.** De plus, l'UM6P développe de nombreux partenariats avec plusieurs écosystèmes à l'étranger, notamment grâce à la création **de l'UM6P France et Canada.** En même temps, l'UM6P a mis en place une stratégie, qui s'avère porter ses fruits, pour faire revenir certains Marocains, binationaux ou Africains formés ou installés à l'étranger, qui apporteront leurs expertises dans le domaine et aider à construire un système porté par les Africains pour l'Afrique.

Pour conclure, j'ajouterais que la stratégie de l'UM6P n'est pas d'empêcher ou de décourager les gens de partir travailler à l'étranger dans le domaine de l'IA, mais plutôt d'en tirer parti, d'y voir **l'opportunité de créer un réseau** dont le pays et le continent profiteront tôt ou tard.

De votre point de vue, le Maroc est-il bien positionné pour tirer son épingle du jeu sur l'IA générative ?

Sur l'IA générative, comme sur le reste des sujets technologiques de pointe, le Maroc peut s'appuyer sur **une mine de talents marocains et africains.** Le continent est, comme vous le savez, très jeune, en forte croissance, et foisonne de talents. La formation locale est très bonne, en particulier en mathématiques et informatique, qui sont les bases fondamentales de l'IA.

De plus, observons que la recherche sur l'IA ne nécessite pas de moyens colossaux, mis à part la puissance de calcul mentionnée plus haut. De plus, les algorithmes pré-entraînés permettent de réduire encore plus le coût de certains cas d'usage. **Le principal investissement est l'investissement humain,** contrairement à d'autres technologies d'avenir, comme le calcul quantique, qui requièrent un important investissement matériel initial.

**La recherche en IA est donc décentralisée** en quelque sorte, facile à mettre en place, à bootstrapper, une fois que les acteurs ont été correctement formés. Je ne vois donc pas d'obstacle à ce que **le Maroc et l'Afrique se positionnent à l'échelle mondiale en termes de solutions locales basées sur l'IA et l'IA générative.**

Concrètement, avez-vous lancé des projets de création d'IA génératives souveraines ? Quelles sont les start-ups championnes sur l'IA au Maroc ? Quel potentiel voyez-vous dans l'IA pour le Maroc ?

Bien sûr ! Nous avons par exemple **créé des chatbots** (Multilingual LM) adaptés à la langue et la culture marocaines, utilisés par des institutions publiques et capables d'interagir en arabe dialectal. Nous travaillons également sur des solutions IA génératives pour des secteurs de souveraineté tels que la cybersécurité et pour des applications d'intérêt national et continental tels que les énergies renouvelables, l'agriculture et la santé. **De nombreuses start-ups émergent par exemple au Maroc**, citons par exemple :

- **Aerodrive Engineering Services**, qui fournit des services de création de cartographie 2D et 3D par drones au profit d'acteurs de référence dans divers secteurs : énergie, agriculture, BTP, etc.
- **AgriEdge**, qui, grâce à ses efforts de R&D, aide les agriculteurs à prendre de meilleures décisions opérationnelles des solutions digitales d'optimisation des opérations pour l'agriculture (fertilisation, irrigation, récolte)

L'UM6P est fière de contribuer au développement de ce tissu de start-ups, notamment via son **fonds d'investissement, UM6P Ventures.**

Je pense que le Maroc a le potentiel pour être un levier de développement de l'IA sur le continent africain. **Le Maroc est au centre des opportunités entre Europe et Afrique**, et le caractère décentralisé de la recherche sur l'IA et l'infrastructure qui se met en place permettront au pays et au continent d'en tirer pleinement parti.

Pour conclure, que pouvons-nous vous souhaiter pour les prochaines années au sein de l'UM6P et en tant que Marocaine sur ce sujet de l'IA ?

Que la vision pour laquelle je suis venue ici se concrétise, c'est-à-dire **de permettre au Maroc et à l'Afrique de former les prochains leaders mondiaux de l'IA**, qui pourront créer des solutions innovantes et locales prenant **en compte les besoins spécifiques du continent.** Je souhaite vraiment que des leaders dans leurs domaines grandissent ici et émergent du Maroc, de l'Afrique et de l'UM6P, et je ne vois aucun obstacle à cela.

# FUTURE OF WORK : DEMAIN, TOUS "AUGMENTÉS" AVEC L'IA GÉNÉRATIVE ?



**Renata Alves-Santos**  
Deputy Chief Data Officer



Pouvez-vous nous présenter vos responsabilités et vos principales missions au sein de Total Energies ?

Je suis **Deputy Chief Data Officer de Total Energies**. J'ai dans mon périmètre la stratégie, gouvernance data et toutes les utilisations de la data au sein du groupe.

En parallèle, je travaille dans **une task force sur l'IA générative**. Porté par l'esprit pionnier qui l'anime depuis sa création il y a 100 ans, TotalEnergies s'appuie sur une tradition d'innovation, dont **l'usage des technologies d'intelligence artificielle dans un cadre industriel** est une composante depuis plusieurs années. Cette task force, qui vise à faire de l'IA générative un levier de transformation pour accompagner la transition énergétique, représente la vision portée par notre PDG, Patrick Pouyanné.

Mise en place l'année dernière, elle comprend deux axes :

- **Un travail sur les applications de l'IA générative pour nos métiers**, avec l'identification de 7 grand thèmes et plus de 150 cas d'usage, dont certains en cours de développement ;
- **Le déploiement et l'adoption de l'assistant GenAI de TotalEnergies**, Microsoft Copilot, qui sera déployé auprès de 30.000 collaborateurs à travers le monde.

Quel a été le processus de déploiement de Microsoft Copilot ?

Nous avons d'abord réalisé un **déploiement-pilote** auprès de 300 collaborateurs, qui a montré **des résultats très positifs**. Nous avons donc décidé d'étendre le déploiement aux collaborateurs utilisateurs des outils Microsoft 365 au quotidien.

Outre les gains de temps et de productivité permis par Copilot, nous constatons des **effets positifs plus indirects** : l'outil aide nos collaborateurs à monter en compétence sur certaines tâches dont ils sont peu familiers, et permet à des collègues de différents pays de travailler ensemble, ce qui est très important pour une multinationale comme la nôtre.

## Quel accompagnement avez-vous mis en place pour former vos collaborateurs à ce nouvel outil?

Lors du déploiement-pilote de Copilot, nous avons compris l'importance de l'accompagnement des utilisateurs, si nous voulions que chaque collaborateur puisse vraiment tirer profit de l'outil. Nous avons donc mis en place **une formation pratique**, lors de laquelle sont présentées toutes les fonctionnalités de MS Copilot, et l'ensemble des cas d'usage possibles déjà identifiés.

Sur ce sujet, **un exemple parlant est l'utilisation de compte-rendu de réunions Teams générés par IA**. Beaucoup de collègues de notre panel de test ont été très convaincus par cette fonctionnalité. Une autre fonctionnalité appréciée par nos collaborateurs est la création de mails dans une langue étrangère, proposée par Outlook. Pour d'autres éléments de la suite Microsoft, comme Excel, l'outil reste perfectible.

Nous avons donc décidé de mettre cette formation en place comme prérequis au déploiement à grande échelle de Copilot. Pour chaque outil de la suite, nous présentons une démonstration, les principaux cas d'usage, et, en retour, cherchons à collecter les besoins et idées des participants. **Le processus d'acculturation est donc à double sens : la formation des collaborateurs va de pair avec le recueil de leurs besoins**. L'équipe de formation a une composition très transverse : digital, data, IT, cyber, RH, pour pouvoir répondre à toutes les questions remontant lors des formations.

Enfin, nous avons également mis en place une large communauté, sur divers sujets – partage de bonnes expériences, de prompts, de nouveautés, etc. – pour faire circuler l'information et favoriser l'innovation. A cet égard, **on peut dire que l'IA générative permet de faire se croiser et dialoguer des gens qui ne se seraient peut-être jamais rencontrés** au sein de l'organigramme TotalEnergies !

## Avez-vous proposé d'autres outils que MS Copilot aux collaborateurs ?

Oui, lors de la même formation, nos collaborateurs vont découvrir **la Microsoft Power Platform**, des outils low-code ou no-code pour le développement d'applications et l'automatisation en toute autonomie.

Il est important, d'une part, de bien expliquer aux collaborateurs que **l'IA générative n'est pas la solution à l'ensemble de leurs cas d'usage**, et d'autre part, à cet effet, de mettre en valeur les outils dont nous disposons déjà, pour que chaque collaborateur puisse décider : ceci est un cas d'usage Copilot, ceci ne l'est pas. Dans ce dernier cas, il est important que les collaborateurs puissent aussi être aiguillés vers d'autres équipes (BI, Digital Factory, etc.).

Quels ont été les plus grands défis que vous avez pu rencontrer lors de votre mise en place de Copilot ?

Tout d'abord, **bien comprendre, et bien positionner ce que l'IA générative peut faire**, car les collaborateurs ont parfois des incompréhensions ou des appréhensions sur l'outil, qu'il est important de dissiper rapidement.

A l'inverse, **il ne faut pas non plus trop promettre**, en expliquant que tout ne sera pas ou n'est pas encore faisable aujourd'hui et que nous ne maîtrisons pas l'agenda de développement de Microsoft.

Enfin, il est crucial d'embarquer certaines fonctions, comme les RH, la cyber et la data pour la réussite de l'initiative.



**Maxime Marembaud**  
Group Chief Data & Digital Officer



Pouvez-vous nous présenter brièvement le groupe Sodexo et la façon dont vous avez structuré les activités d'IA générative en son sein ?

Le groupe Sodexo est présent dans **45 pays**, et développe ses activités de façon très décentralisée. Nous proposons différentes typologies de services, pour des profils de consommateurs très différents.

La restauration collective est l'aboutissement d'un certain nombre d'étapes et de processus : le personnel qui produit, la compréhension des préférences des consommateurs, leur satisfaction, l'optimisation des flux, la gestion des stocks et de l'approvisionnement, la définition d'un prix juste... Étant donnée la nature de notre activité, la priorité est donnée en interne aux équipes locales opérationnelles, qui sont in fine les garantes de la qualité de nos services, les outils pour répondre au mieux à leurs enjeux sur le terrain. Nous avons donc choisi **de construire une organisation digitale et IA appuyée par des experts techniques**, qui reporte directement au COMEX de Sodexo. L'équipe est globale, implantée dans les principales régions du groupe, et centrée sur le delivery des produits digitaux et IA, pour soutenir la production opérationnelle au plus près de nos consommateurs.

Dans ce cadre, l'IA générative intervient sur 3 grands piliers :

1. **Une gouvernance efficace et un plan d'acculturation global**, qui nous permet de bien **communiquer auprès des collaborateurs de nos régions**. C'est la garantie d'une utilisation optimale des outils, tout en protégeant les données propriétaires de l'entreprise, étant donnée la multitude des outils d'IA générative actuellement disponibles au grand public. Dans une perspective d'amélioration des usages, nous animons différentes communautés, qui s'adaptent aux spécificités des fonctions au sein du groupe.
2. L'utilisation opportuniste de **solutions « off-the-shelf »** :
  - a. Nous sommes en train de **déployer un assistant IA, Microsoft Copilot + Github Copilot**, en commençant par une phase pilote impliquant 1000 collaborateurs. L'enjeu de cette phase pilote est double : bien identifier les usages les plus pertinents, en les mappant aux différents personas de l'entreprise, pour prioriser le déploiement à grande échelle (le groupe Sodexo emploie 400.000 collaborateurs), et bien estimer les besoins de conduite du changement associés au déploiement.
  - b. Nous testons aussi d'autres **outils off-the-shelf d'IA générative**, en support aux activités commerciales, qui pourront être notamment intégrées à notre CRM.
3. Le **delivery de produits, de cas d'usage customs** impliquant l'IA générative, s'appuyant sur notre équipe développant historiquement le reste de nos cas d'usage ML.



Sur ce dernier point, pouvez-vous nous donner des exemples concrets de cas d'usage sur lesquels vous travaillez ?

Citons 3 cas d'usage sur lesquels nous avons déjà beaucoup travaillé :

1. Un cas d'usage de support aux ventes, visant à apporter **un premier niveau de réponse assisté par IA générative aux appels d'offres auxquels nous répondons**. Nous travaillons à la fois sur l'efficacité opérationnelle et sur la qualité de nos propositions commerciales, en assurant une bonne customisation et une bonne contextualisation de cette proposition ;
2. Un cas d'usage **de gestion de la lecture et de la compréhension des contrats** définissant les conditions d'opération d'un site Sodexo ;
3. Un cas d'usage de **support à la construction de notre contenu marketing et communication**, une fonction très impactée par l'IA générative, notamment par l'analyse du retour des consommateurs sur la base de leurs commentaires.

L'IA générative nous aide également à traiter certaines de nos données : par exemple, sur l'enrichissement des données de nos fournisseurs, ou pour l'amélioration de la performance de nos algorithmes de matching de tickets de caisse.

Principalement dans les opérations de cuisine et sur les sites des clients, quelles sont vos convictions à long terme sur l'impact de l'IA générative, et potentielle de l'AGI, sur une entreprise comme Sodexo ?

A moyen terme, j'ai deux grandes convictions.

**Certaines fonctions vont devoir complètement se transformer** pour bénéficier au maximum de la valeur de l'IA générative, dans des logiques d'automatisation, de personnalisation. Je pense, dans le cadre de Sodexo, aux fonctions de sales, du marketing, et des ressources humaines, en particulier en ce qui concerne les relations avec les employés. Il est important d'anticiper dorénavant l'impact qu'aura l'IA générative pour ces métiers, dans une double logique de gain de productivité et d'augmentation de la valeur produite.

Sur les sites, au cœur des opérations, **l'IA générative pourra simplifier le rôle de manager de site**, en fluidifiant l'accès à l'information émanant du groupe et de ses différentes entités, et, en nous aidant à mieux comprendre les besoins de nos consommateurs localisés sur chaque site, nous permettra d'**ajouter de la personnalisation à nos recettes**, tout en restant en compliance avec nos catalogues standard.



**Claire Lebarz**  
Chief Data & AI Officer



Tout d'abord, pouvez-vous nous présenter Malt ?

Malt a été fondée en 2013 par Vincent Huguet (co-CEO) et Hugo Lassiège, avec Alexandre Fretti pour co-CEO. **Forte d'une communauté de plus de 700 000 talents indépendants** mettant sur sa marketplace leurs compétences et expertises au service des organisations à la recherche de talents externes, Malt est désormais **la plateforme leader du freelancing en Europe**. Malt compte déjà plus de 70 000 clients, PME comme grands groupes. Depuis l'acquisition de Comach en 2022 et le lancement de Malt Strategy, Malt propose également l'expertise de son réseau de consultants indépendants et managers de transition. Malt est présent dans 9 aires géographiques (Allemagne, Belgique, Émirats arabes unis, Espagne, France, Pays-Bas, Scandinavie, Royaume-Uni et Suisse).

La fonction RH semble offrir de nombreuses opportunités avec l'intégration de l'IA générative. Quelles sont les opportunités spécifiques que vous identifiez pour Malt et, plus largement, pour la fonction RH ?

A Malt, l'IA est au cœur du produit depuis des années. Malt, c'est plus de 700.000 talents, nous avons donc besoin d'apprentissage machine pour proposer les talents les plus pertinents à travers l'Europe. Depuis un an, **nous avons amélioré la performance du matching avec une vector database et en utilisant des modèles de langage**. La sémantique est importante pour nous car nous avons beaucoup de texte sur les profils (la description, les expériences passées notamment) et également sur la description des missions. Pour prendre un exemple, les termes « data science » et « machine learning » seront considérés comme très proches, alors qu'ils n'ont pas du tout les mêmes racines étymologiques. Cela nous permet de projeter les freelances dans un « espace » des talents, de projeter les missions déposées sur Malt dans ce même espace et de regarder ensuite la distance entre les talents et la mission.

En parallèle, l'IA nous aide à améliorer l'efficacité et l'expérience de travail de nos collaborateurs. Pour cela, nous avons mis à disposition une "AI as a service" en construisant une intégration avec Dust. **Cette plateforme permet à toutes les équipes de Malt de créer ses propres agents d'IA générative de façon complètement autonome**. Citons quelques assistants : documentation de l'équipe engineering, connaissance interne du produit, soutien RH aux managers, assistant à la création de contenu pour les designers... La performance (taux de bonnes réponses) de chaque assistant est monitorée. L'adoption est très bonne, organique, alors même que nous n'avons pas encore fait de communication générale sur le sujet !

En tant qu'acteur central dans le domaine des ressources humaines, quelles tendances majeures avez-vous observées depuis l'émergence de l'IA générative ?

Ce qui est clair, c'est que toutes les entreprises se posent la question de l'IA, y compris les PME. Nous recevons beaucoup de demandes de missions d'éducation, d'identification d'opportunités IA, d'accompagnement à la stratégie, même dans des petites entreprises. Il y a eu énormément de missions de ce genre en 2023. **Pour 2024, les entreprises se projettent plus dans l'exécution, le déploiement des solutions.** Certaines entreprises ont une très bonne vision des choses et une forte stratégie autour des compétences, alors que d'autres restent surtout cantonnées au chatbot interne, n'ont pas encore les investissements data nécessaires à l'IA ou souffrent d'un manque d'expertise d'ingénierie en IA.

Comment ces tendances transforment-elles le marché du travail et les attentes des employeurs et des freelances ?

Actuellement, nous ne constatons pas chez Malt de baisse de la demande pour les talents. Les études qui sont sorties en 2023 aux Etats-Unis sur le sujet me paraissent trop précoces, et semblent attribuer l'impact du contexte macroéconomique et resserrement des financements à ChatGPT. Cela étant, **sur les profils non-tech, on observe que beaucoup d'entre eux commencent à ajouter dans leur CV des compétences autour de l'IA générative**, compétences qui représentent un avantage compétitif sur leur marché. C'est plutôt le designer expert des outils d'IA générative qui remplacera le designer qui ne maîtrise pas ces outils, plutôt que les outils eux-mêmes.

# L'IA GÉNÉRATIVE EST-ELLE DÉJÀ UN FACTEUR DE TRANSFORMATION PROFONDE DES BANQUES ET ASSURANCES ?



**Ahmed Besbes**  
Chief Data Officer



Vous êtes CDO de Direct Assurance, pouvez-vous nous présenter vos principales missions à long terme ?

J'ai trois principales missions:

- le développement **d'une gouvernance des données et la protection des données personnelles,**
- le déploiement et la maintenance **des plateformes IT servant aux cas d'usage data,**
- enfin, bien sûr, le développement de l'expertise data science et IA et la contribution à **l'effort d'innovation de l'entreprise.**

Comme vous pouvez vous en douter, l'IA générative a un impact sur chacune de ces missions.

En interne, quel regard portent les actuaires sur le changement de paradigme que peut représenter l'IA générative ?

Chez Direct Assurance, les pratiques de data science sont déjà intégrées depuis longtemps au travail des actuaires. D'une certaine manière, nos actuaires sont également data scientists, et nos data scientists déjà actuaires. Ces deux profils travaillent donc déjà **main dans la main.**

**Les actuaires portent un regard enthousiaste sur l'IA générative,** comme sur l'IA en général, ils veulent comprendre et tester les opportunités que cette technologie leur apporte. Notre mission est de leur apporter les outils, l'accompagnement et l'expertise IA dont ils ont besoin pour parvenir à cet objectif.

Direct Assurance est une filiale au sein du groupe Axa, quels sont selon vous les sujets qui doivent être portés au niveau du groupe, et ceux qu'il est plus pertinent de traiter directement en filiale ?

Il y a **des synergies assez fortes entre les filiales** et les projets qu'elles mènent, qui sont plus faciles à identifier au niveau du groupe. Nous créons une communauté de praticiens de l'IA, au sein de laquelle chaque filiale peut initier des premiers cas d'usage et obtenir un retour d'expérience qui sera utile pour tous.

Certains sujets sont communs entre plusieurs filiales, et nous préférons alors **nous coordonner pour les traiter plus efficacement**. Citons par exemple le SecureGPT, qui a été déployé par Axa à l'ensemble du groupe, ou les sujets relatifs aux call centers, qui concernent la plupart des filiales. Le groupe permet **de passer à l'échelle les sujets intrinsèquement transverses, ou des cas d'usage locaux à forte valeur**, et met à disposition des filiales ses infrastructures IT.

Quels sont selon vous les principaux défis et opportunités liés à l'IA générative pour les assurances ? Diriez-vous que l'IA générative amène une transformation profonde de ce secteur?

La promesse que porte l'IA générative est énorme, avec un grand nombre d'opportunités. Prenons quelques exemples :

- **les call centers**, qui sont un gros poste de coût pour nous, et qui manipulent de nombreuses données (texte, voix) que **l'IA générative aidera à structurer**;
- **la gestion des sinistres**, sur lesquels il y a de gros enjeux d'exploitation de documents et d'images

Je ne pense pas que l'IA générative remplacera des postes, mais sera **un outil de productivité pour nos collaborateurs**. Il y a donc pour nous un gros enjeu de formation et d'accompagnement de l'ensemble du personnel de Direct Assurance. Je ne pense pas non plus que l'IA générative soit pour le moment assez mature pour être mise en interaction directe avec les clients, ni que ce soit une direction qu'on envisage aujourd'hui pour notre expérience client. C'est moins souvent noté, mais l'IA générative représente également pour nous un défi, via **l'utilisation que des tiers malveillants peuvent faire de cette technologie**, notamment en termes de fraude et de cybersécurité. Il faut donc être conscient des évolutions de l'IA générative, pour affronter ces risques nouveaux.

Quelles sont selon vous les compétences-clefs que doivent développer les assurances pour intégrer pleinement l'IA générative ?

Nous organisons un développement des compétences à deux échelles :

- un premier volet de **formation et d'acculturation de l'ensemble des collaborateurs**,
- un second volet du développement de **notre expertise interne sur ces sujets** (comment les modèles fonctionnent, quelles sont leurs limites, leur modèles de coût, quelles sont les architectures IT possibles), pour permettre à mon département de réaliser les bons choix techniques. L'IA générative évoluant très vite, il y a donc un travail constant de veille à réaliser.

Il faut aussi **être complètement conscient des enjeux réglementaires** pour pouvoir déployer l'IA générative de façon complètement maîtrisée.

L'assurance est un secteur très réglementé, dans lequel la gestion de la donnée a une place prépondérante. Quel rôle doivent jouer les régulateurs et gouvernement dans l'encadrement des usages de l'IA pour les secteurs financiers ?

Il est important **d'établir un cadre pour le respect de la vie et**, quelque part, ce cadre est déjà établi par le **RGPD et l'AI Act**, qui le renforce. De manière générale, je suis promoteur d'une approche par le risque, qui responsabilise les acteurs, pour parvenir à conserver une liberté de tester, d'innover, tout en étant conscience des impacts possibles sur la vie privée des personnes.

Que pensez-vous de la maturité du secteur quant à l'IA générative et les transformations qu'elle permet ?

Notre secteur en est encore au début des transformations induites par l'IA générative. Beaucoup d'idées circulent en ce moment, je ressens **un enthousiasme assez fort de la part des acteurs du secteur**. L'enjeu est de saisir les premiers cas d'usage qui auront un impact réel sur le business.



**Maxime Havez**  
Chief Data Officer

Pouvez-vous nous présenter le groupe Crédit Mutuel Arkéa et les principales missions de son data office ?

Le groupe Crédit Mutuel Arkéa est un groupe **mutualiste de bancassurance**. Sa quarantaine de filiales lui permettent de couvrir l'ensemble des métiers de la banque et de l'assurance, notamment **la banque de détail** (avec les Fédérations du Crédit Mutuel de Bretagne et du Crédit Mutuel du Sud-Ouest), **la banque en ligne** (avec Fortuneo), **la banque d'entreprise**, **l'assurance IARD et l'assurance-vie**.

Le data office du Crédit Mutuel Arkéa, dont je suis en charge, travaille de façon transverse pour l'ensemble des entités du groupe au travers de 5 missions principales :

- **L'animation de l'IA factory et le développement des cas d'usage** liés à l'exploitation de l'IA au service des différents métiers du groupe : nous avons mis en place une task force sur le sujet depuis un an, qui a permis d'identifier plus de 70 cas d'usage liés à l'exploitation des IA génératives ;
- **Le développement de la plateforme data & IA** et sa mise à disposition pour nos entités ;
- **L'animation de l'innovation au niveau du groupe**, assurant notamment le pilotage du comité de direction générale dédié à l'innovation ;
- Le copilotage d'un programme de **collecte des données extra-financières** ;
- **L'animation d'une démarche visant à diffuser une approche data & IA responsable** – nous avons d'ailleurs reçu dans ce cadre la certification IA de confiance décernée par le LNE.

Au-delà des cas d'usage, vous avez déclaré observer un changement de paradigme lié à l'arrivée de l'IA générative. Pouvez-vous nous en dire plus ?

Jusqu'à présent, c'étaient les structures de type **IA Factory ou datalabs** qui étaient à l'initiative pour catalyser l'usage de l'IA dans les process des grandes entreprises, en mettant en place des sessions d'acculturation, des ateliers d'idéation, etc. De ce processus relativement « top-down », on essayait de faire émerger des cas d'usage à forte valeur ajoutée pour les métiers.

L'IA générative est arrivée, soudainement, avec **des outils très grand public, faciles à prendre en main**, comme ChatGPT. Cela a permis aux collaborateurs des entités métiers de se projeter beaucoup plus facilement dans les perspectives que cela pouvait ouvrir dans leur missions quotidiennes. La relation entre ces équipes métier, opérationnelles et notre IA Factory s'est inversée : ce sont elles qui désormais attendent que nous développiions ces cas d'usage pour répondre à leurs enjeux du quotidien.

C'est finalement la grande vertu de l'IA générative : **avoir permis de motiver, d'embarquer beaucoup plus profondément les opérateurs vers la transformation IA**, en leur fournissant un cas d'usage (le chatbot) généraliste entre les mains. D'ailleurs, il arrive que nos équipes formulent parfois des demandes de nouveaux cas d'usage basés sur l'IA générative, et qui sont finalement adressables avec une IA plus traditionnelle.

Quel est pour vous le principal défi à relever, le principal enjeu lié à l'IA générative ?

Je pense qu'il s'agit pour nous de **réussir à définir la vision** que l'organisation souhaite donner sur l'adoption de l'IA.

Nous avons dû nous positionner, choisir entre des déploiements massifs à l'échelle du groupe, ou des déploiements plus maîtrisés, en rapport avec notre qualité d'entreprise à mission.

Notre conviction est que nous souhaitons **développer une IA maîtrisée, sobre, la plus souveraine possible**, et cela a bien sûr un impact sur la façon dont l'IA sera déployée à l'échelle. Cela implique de devoir développer certains projets de R&D, comme le fine-tuning d'un LLM que nous avons réalisé pour réduire notre consommation énergétique, modèle que nous avons d'ailleurs [publié sur HuggingFace](#).

Comme je l'ai dit plus haut, le paradigme a changé, et l'adoption, l'utilisation des IA génératives n'est plus du tout réservée aux seuls experts. Il est donc crucial que **cette vision** (et les éléments de décision qui en découlent quant aux modèles à privilégier) **soit bien comprise par tous au sein du groupe**. Ce sera déjà une grande victoire que d'arriver à **définir clairement ces éléments de décision et à les diffuser largement**.

A cet égard, l'exigence de sobriété, découlant de la qualité d'entreprise à mission du Crédit Mutuel Arkéa, a des conséquences très concrètes. En effet, les solutions « sur l'étagère » des GAFAM sont parfois moins chères que des modèles open source ou des modèles fine-tunés, car les coûts liés à l'hébergement des modèles déployés on-premise ne sont évidemment pas nuls. **Pour autant, l'impact en termes de sobriété de ces solutions n'est pas égal, et ne doit pas être négligé**. Nous devons a minima sensibiliser nos équipes sur ces sujets, en particulier celles qui seront chargées de choisir les modèles que nous utiliserons.



# COMMENT ASSURER LA ROBUSTESSE ET LA CONFORMITÉ DES IA GÉNÉRATIVES ?



**Jean-Marie John-Mathews**  
CAI Ethics Researcher & Co-founder



Qu'est-ce que la robustesse d'un système d'IA, et en quoi serait-ce différent de tout autre système informatique ? Si l'on parle d'IA générative, quels en sont les enjeux plus spécifiquement ?

Dans le ML traditionnel, c'est-à-dire avant l'arrivée de l'IA générative, dire qu'un modèle est **robuste signifiait qu'un petit changement en entrée du modèle ne provoque pas un changement trop significatif dans ce que le modèle produit**. Avoir un modèle robuste est évidemment un enjeu pour l'adoption des modèles par les équipes métier. Nous avons par exemple travaillé récemment avec une entreprise retail sur un algorithme de pricing: les équipes savaient qu'en changeant une des variables, comme la qualité d'un produit, en entrée, le prix devait nécessairement augmenter.

Il arrive que la robustesse soit également **liée à l'impératif de non-discrimination**: par exemple, en credit scoring, on doit vérifier que le sexe de la personne n'ait pas d'influence sur l'octroi de crédit. Ce sujet est à la frontière entre la non-discrimination et la robustesse de l'IA.

Pour ce qui est de l'IA générative, la robustesse au sens strict est bien plus complexe à appréhender, car la nature textuelle et qualitative des entrées et des sorties s'y prête peu. Les tests qu'on réalise sont alors **plus complexes, plus «humains»**: on paraphrase des questions, notamment, en vérifiant que la réponse apportée reste toujours la même. Il existe beaucoup de tests de cette nature, devenant toujours plus complexes. On utilise alors l'IA générative pour tester la robustesse d'un modèle d'IA générative.

Est-ce qu'il n'est pas alors risqué de tester un LLM par un autre LLM ?

C'est effectivement un motif d'inquiétude assez fréquent. Cela étant, les LLM qui permettent de faire les tests sont très spécifiques, souvent conçus pour ce type d'usage. Les modèles open source peuvent être également **testés avec les meilleurs modèles propriétaires, plus performants et plus chers**, comme GPT-4. J'insiste: le modèle utilisé pour tester n'est pas le même que le modèle testé. Cette méthode, qu'on appelle **LLM as a judge**, s'impose actuellement comme le standard pour tester la robustesse des modèles.

## Comment contrôler les risques d'hallucination, dont on parle si souvent ? De quoi sont issues ces hallucinations ?

Il faut d'abord rappeler que **l'hallucination est intrinsèque aux modèles d'IA générative**. En effet, ces modèles sont entraînés par LHF (Learning from Human Feedback), on demande donc au modèle de se calibrer au feedback humain, quitte à l'encourager à produire des réponses qui vont au-delà de sa base d'apprentissage initiale. C'est ce qui fait d'ailleurs leur force, mais c'est aussi ce qui explique ces hallucinations.

De manière générale, pour réduire les risques d'hallucination, on utilise une **base d'informations tierce, utilisée comme référence** pour quantifier la propension du modèle à produire ces hallucinations. Ces bases peuvent être une API vers Google, une calculatrice pour du calcul, un autre LLM, ou des RAGs.

Il y a beaucoup de types d'hallucination différents, avec leurs tests spécifiques. Chez Giskard, nous contrôlons par exemple **la sycophancy, type d'hallucination dans lequel le LLM tente de plaire à l'utilisateur** (on peut penser à une forme de démagogie du LLM), quitte à finir par mentir. Dans ce cas, la méthode de test consiste à poser des questions au modèle en avançant des faits inexacts, et observer si le modèle va « nous croire sur parole ».

## Dans ce cas, comment avoir une approche systémique du contrôle des hallucinations ?

Pour vous donner quelques statistiques que nous observons, on estime actuellement que **5 à 10% des réponses des modèles bruts sont hallucinatoires**, ce qui est substantiel. Il est alors possible, en mettant en place un ensemble de tests et d'actions correctrices, **de faire tomber ce chiffre à 2%**. Pour ce faire, lorsqu'un test n'est pas concluant, on va modifier, encapsuler le prompt initial, pour attirer l'attention du LLM sur l'erreur qu'il pourrait commettre, tout en veillant à ce que la performance générale du modèle ne soit pas dégradée. Ce travail, très itératif, permet de réduire grandement les risques.

## La robustesse des modèles d'IA génératives est-elle suffisante pour permettre aux entreprises de les mettre au cœur des interactions avec les clients ?

Actuellement, les applications principales sont **la recommandation produit, des créations de FAQ**. Les acteurs ont bien conscience **des risques réputationnels que les chatbots exposés aux clients** peuvent causer et c'est la raison pour laquelle elles mandatent des parties tierces comme Giskard.

Nous **évaluons alors les modèles**, attribuant un score sur plusieurs critères. Lorsque l'évaluation est trop faible, nous pouvons définir un plan de résolution avec le client. Il nous est même déjà arrivé plusieurs fois de stopper des projets qui présentaient des risques trop importants.

Nous venons d'explorer les enjeux liés à la robustesse des IA génératives, qui nécessitent la mise en place de contrôles. Ces dernières années, l'Europe a décidé de réguler l'usage des IA. Mais réguler l'IA, c'est réguler quoi selon vous?

Il peut s'agir tout d'abord de normer les process, **en utilisant des standards** comme ISO-42001, par exemple. Mais on peut également normer les modèles, ce qui se développe de plus en plus, notamment et effectivement avec l'AI Act européen, très centré sur le modèle. Dans ce cas, la question est: régule-t-on uniquement tout le modèle, ou l'ensemble du cycle d'apprentissage qui l'a produit?

Ensuite, on peut aussi parler de la **« supply chain » des modèles**: qui est responsable de quoi, sur la chaîne de valeur de l'IA? Il est très difficile de parvenir à un standard consensuel pour le moment.

L'UE a-t-elle raison de réguler l'IA? Et pensez-vous que la régulation soit une bonne réponse au besoin de robustesse?

Je pense que oui. Il est vrai qu'on craint souvent la régulation, qu'on pourrait voir comme une entrave à l'innovation. Cela étant, il ne faut pas oublier que dans de nombreux secteurs, comme la banque, l'assurance, la santé, **les acteurs sont habitués à la régulation** et ne craignent pas l'AI Act. Son article 50, qui formule **des obligations en termes de transparence**, imposera de mettre en place une certaine forme de documentation, notamment sur la robustesse des modèles. Néanmoins, cette obligation ne sera que déclarative, et les tests de robustesse resteront entrepris par les éditeurs des modèles eux-mêmes.

Vous avez publié une partie de votre code en open source. Quelle place tient l'open source dans votre stratégie?

L'actif principal que développe Giskard, c'est **un vaste ensemble de tests de qualité**, très divers. L'open source est donc très important pour nous, tout simplement car nous avons besoin d'obtenir de nouvelles contributions, d'augmenter encore la couverture de nos tests.

Cependant, on pourrait alors se dire: en mettant en avant notre façon de tester les modèles, ne risque-t-on pas de laisser les éditeurs de modèles s'adapter, contourner nos tests? Mais je ne suis pas très inquiet sur ce point, et je suis convaincu que l'open source est **un élément-clef pour construire notre fiabilité à long terme**. Prenons un exemple célèbre: Linux, qui est open-source, et que toutes les DSI utilisent désormais, notamment pour sa sécurité.

Les auteurs tiennent à remercier chaleureusement les personnes interviewées pour leur contribution à ce livre blanc :

**Guillaume Avrin**, Coordinateur national pour l'intelligence artificielle – [Direction Générale des Entreprises](#)

**Vincent Rapp**, Responsable Sectoriel Numérique – Coordination IA – Direction de l'Innovation – [Bpifrance](#)

**Emmanuel Papadacci-Stephanopoli**, Chief Product Officer – [La Fabrique by CA](#)

**Pascal Simon**, Investment Strategy – [Kaukana Ventures](#)

**Yoach Lacombe**, ML Engineer – [HuggingFace](#)

**Harriet Rees**, Chief Information Officer – [Starling Bank](#)

**Olivier Abecassis**, CEO – [Qwant](#)

**Catherine Dubeau**, Responsable de Département Data Science & Governance – [Safran Aircraft Engines](#)

**Lamia Azizi**, Professeure College of Computing – [UM6P](#)

**Renata Alves Santos**, Deputy Chief Data Officer – [TotalEnergies](#)

**Maxime Marembaud**, Group Chief Data & Digital Officer – [Sodexo](#)

**Claire Lebarz**, Chief Data & AI Officer – [Malt](#)

**Ahmed Besbes**, Chief Data Officer – [Direct Assurance](#)

**Maxime Havez**, Chief Data Officer – [Crédit Mutuel Arkéa](#)

**Jean-Marie John-Mathews**, AI Ethics Researcher & Co-founder – [Giskard AI](#)

Les auteurs voudraient remercier les personnes suivantes, ayant contribué à la rédaction et la révision de ce livre blanc :

**Aimé Lachapelle**, Managing Partner – [Emerton Data](#)

**Ugo Philippart**, Associate Partner – [Emerton Data](#)

**Franck Abihssira**, Partner – [Emerton Data](#)

**Yannick Léo**, Partner and Data Science Director – [Emerton Data](#)

**Constance Frécenon**, Principal – [Emerton Data](#)

**Pierre Six**, Case Team Leader & Lead Data Scientist – [Emerton Data](#)

**Bailey Aaron**, Data Scientist & Strategy Consultant Intern – [Emerton Data](#)

Les auteurs voudraient également remercier les personnes suivantes, ayant aidé à l'organisation de la série de tables rondes "The Real Deal of Generative AI" du 4 juin 2024, à laquelle les personnes interviewées ont participé :

**Perle Bagot**, Associate Director – [HUB Institute](#)

**Baptiste Pellé**, Partnerships Manager – [HUB Institute](#)

**Alexia Gineste**, Editorial Content Manager – [HUB Institute](#)

Toutes les équipes du [HUB Institute](#)

# REMERCIEMENTS

# EMERTON DATA

## Contributeurs

Aimé Lachapelle  
Yannick Leo  
Franck Abihssira  
Ugo Philippart  
Constance Fréceçon  
Pierre Six

[aime.lachapelle@emerton-data.com](mailto:aime.lachapelle@emerton-data.com)

Emerton Data est le cabinet de conseil spécialiste des sujets digitaux, data et IA.

Nous aidons les entreprises à définir leurs stratégie data & IA, proposons le développement de modèles et de solutions spécialisés, et incubons des solutions SaaS basées sur l'IA, grâce à notre start-up studio, Kaukana Ventures.



Fondé en 2012, le HUB Institute, avec plus de 20 communautés professionnelles, est le tiers de confiance du marché sur les sujets de transformation digitale et durable des entreprises.

Notre mission est de permettre aux grands groupes de comprendre et d'anticiper les tendances d'aujourd'hui et de demain, d'échanger avec les acteurs les plus innovants et d'accélérer leur croissance et leurs ambitions business.

## Contributors

Perle Bagot  
Baptiste Pellé  
Alexia Gineste

# À PROPOS DE NOUS

EMERTONDATA



# L'IA générative en entreprise à l'épreuve du réel