

Génération

# IA

#2

*Back to School*

10 JUIN 2025

SYNTHÈSE DES  
ATELIERS

# Un nouveau tournant s'annonce : celui de la "Genisation".

**D**ix ans après le phénomène "d'Uberisation" de l'économie décrit par Maurice Lévy, un nouveau tournant s'annonce : celui de la "Genisation".

À la différence près que l'Uberisation désigne la désintermédiation et la facilité d'accès à un service par la mise en œuvre d'une plateforme digitale, alors que le phénomène actuel lié à l'arrivée des IA génératives et des LLM (modèles de langage de grande taille) met sur le devant de la scène, non pas les usages ou les avantages, mais la technologie elle-même.

Il faut dire que recenser et synthétiser les usages des IA génératives est complexe tant ils apparaissent infinis ! La magie technologique que l'on constate à chaque génération de texte ou autres contenus par un LLM nous dépasse et interroge sur ses limites, mais aussi sur les nôtres.

Résultat : après seulement 3 années et demi, le monde ne parle plus que d'IA Génératives et les grandes puissances mondiales & géants du numérique s'approprient à investir en cumulé plus de 1000 milliards de \$ à horizon 2030 !

La France figure parmi les plus grands générateurs de startups IA (grâce au soutien considérable de BPI France). Des entrepreneurs passionnés et convaincus ont bien compris le potentiel de l'intégration des LLM et des perspectives de marché considérables ! 750 Mds de \$ dans le monde à horizon 2027 – alors qu'il était quasi inexistant, il y a si peu de temps !

Mais comment attaquer le monde depuis la France ? Faut-il se revendiquer Français, voire arguer d'être souverain ? Est-ce une carte à jouer ou un fil à la patte ?

Une France, surtout, terre d'accueil des investissements futurs liés à l'IA - désormais près de 125 Mds € annoncés d'ici 5 ans (annonces du Sommet de l'IA et de Choose France). À noter que ces investissements viendront à -90 % d'Arabie Saoudite et d'Amérique du Nord et s'orienteront massivement vers la construction de supercalculateurs.

Si d'un côté ces investissements sont une menace directe pour les hébergeurs européens qui ne sont pas ou peu représentés dans le tour de table, cela peut malgré tout être l'opportunité d'émergence d'une filière de construction et d'exploitation de Datacenters performante et respectueuse de l'environnement (c.f. le *Référentiel Datacenters Hautes Exigences Environnementales de l'ADIRA*).

Paradoxalement, notre pays demeure en bas du classement des pays européens quant à l'adoption des IA génératives. Si elles sont au cœur de toutes les réflexions pour accélérer la productivité et la compétitivité, on tâtonne, on s'interroge.

Faut-il construire une stratégie IA ? À l'échelle de l'entreprise ? À l'échelle des activités ? Faut-il nommer quelqu'un en charge, un Chief IA Officer par exemple ? Est-ce aux métiers, ou à la DSI, ou la Digital Factory ? On écrase la roadmap en cours ou bien, on ajoute de l'IA de-ci de-là ? Au fait, le PoC de chatbot avec le service client, il en est où ?

Les entreprises ont encore des difficultés à trouver leur cap.

C'est pour répondre à ces questions que nous organisons ce 4ème événement dédié aux IA génératives, et le 2nd opus de Génération IA. Notre ambition : vous inspirer à travers un concentré de cas d'usage concrets, de démonstrateurs, et de partages d'expérience pour imaginer ensemble les applications de demain.

Un immense merci aux partenaires de l'évènement qui ont identifié les thèmes d'intérêt et se démènent depuis des mois pour répondre au mieux à vos attentes, de même pour les intervenants inspirants que nous avons le plaisir d'accueillir et toute l'équipe qui a conçu avec amour cet événement !

Bon retour sur les bancs de l'école !



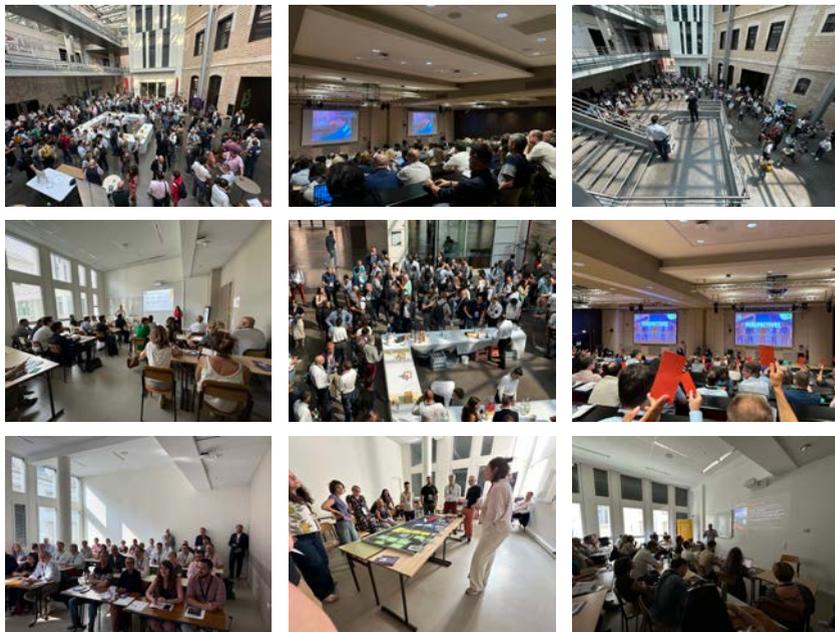
**Pierre-Antoine TROUBAT**  
Délégué Général, ADIRA

# GÉNÉRATION IA #2

Le 10 juin dernier, Génération IA a une nouvelle fois tenu ses promesses : **plus de 400 participants**, une scène inspirante, des retours d'expérience concrets et un public au rendez-vous pour explorer ensemble les usages réels des IA génératives. Merci à nos sponsors qui ont rendu tout cela possible et bravo à tous les intervenants pour leur expertise et leur énergie.

L'après-midi a été rythmée par des talks, des démos, des échanges passionnants, et s'est clôturée en beauté par une soirée conviviale où les discussions se sont prolongées. Entre curiosité, pragmatisme et innovation, cette édition a brillamment rempli sa mission. On repart avec des idées pleines la tête, de belles rencontres, et l'envie de recommencer.

*Quelques photos souvenirs à découvrir ci-dessous !*



Grâce à notre partenaire **Noota**, les ateliers ayant rythmé l'après-midi ont été enregistrés puis retranscrits sous forme de comptes-rendus. L'ADIRA s'est ensuite occupé de la reformulation et de la mise en page pour vous offrir ce livrable.

## ANTEMETA

Arnaud SENANEDJ  
arnaud.senanedj@antemeta.fr

## ASI

Franck CHEKKI  
fchekki@asi.fr

## ASKILLS

Anthony FRADERA  
anthony.fradera@askills.fr

## AVANADE

Lionel VRAC  
lionel.vrac@avanade.com

## AVISIA

Hamza MOULIM  
hmoulim@avisia.fr

## DELL TECHNOLOGIES

Barbara RENAC  
barbara.Renac@dell.com

## DEVOTEAM

Carole HERNANDEZ  
carole.hernandez@devoteam.com

## PRODATASEC

Sabrina JAFFRES  
sabrina.jaffres@prodatasec.com

## SPIE ICS

David MATRAT  
david.matrat@spie.com

## WIKIT

Yvan Levacher  
yvan.levacher@wikit.ai

## NOOTA

Alexandre DUFFAUT  
alexandre.d@noota.io



## Appels d'offres : répondez 3x plus vite, sans rien sacrifier à la qualité

L'exemple présenté illustre un usage commercial concret avec l'automatisation de la rédaction de réponses à des appels d'offres complexes. L'objectif est d'optimiser le temps des équipes tout en garantissant rigueur et conformité. La solution ne remplace ni la compétence humaine, ni la responsabilité juridique. Elle automatise ce qui est répétitif, structuré ou documentaire. L'humain reste au centre des arbitrages fins.

### FONCTIONNEMENT

- Analyse automatique du cahier des charges (dates clés, contraintes réglementaires, etc.)
- Suggestion des chapitres pertinents à intégrer dans la réponse à partir d'une base documentaire personnalisée
- Génération d'un mémoire technique complet, prêt à être relu, incluant le texte existant (non modifié) et des propositions IA (surlignées en jaune pour relecture humaine)
- Ajout de contenus manquants (ex. : chapitre RSE, certifications ISO)
- Remplissage automatique de questionnaires de conformité ou sécurité (RGPD, HDS, etc.)
- Agent conversationnel intégré pour enrichir, reformuler ou synthétiser à la demande
- Historique des interventions et traçabilité des sources utilisées

### AVANTAGES

- Une solution industrielle, évolutive et modulaire avec une **intégration simple** au système d'information
- Bibliothèque de **modèles IA prêts à l'emploi** (généralistes ou spécialisés)
- **Personnalisation par métier** (agents pour commerciaux, RSSI, support, RH, etc.)
- **Suivi du ROI** avec un gain de temps mesurable, taux d'utilisation, évolution continue via retour utilisateur
- **Modèle économique OPEX** (abonnement mensuel, engagement sur 36 mois possible)

### Les piliers d'une solution souveraine, sécurisée et maîtrisée :

# 1

#### Souveraineté

Déploiement sur appliance locale, dans l'infrastructure client, sans transfert de données vers le cloud.

# 2

#### Sécurité

Accès restreint (VPN), bases documentaires maîtrisées, conformité RGPD et ISO.

# 3

#### Contrôle humain

Les contenus générés doivent toujours être validés, notamment les éléments engageants (juridique, pénalités, prix).

## Support IT : automatisez vos niveaux 0, 1 et 2 avec l'IA et gagnez en qualité & réactivité

Cette solution clé en main combine une **puissance de calcul maîtrisée**, une **plateforme évolutive avec modules IA intégrés** et un **accompagnement complet** (configuration, exploitation, maintenance, gouvernance, adoption). Le cas d'usage démontré est un support IT automatisé, capable de traiter directement les demandes reçues par mail ou autres canaux.

### FONCTIONNEMENT

- Lecture du message entrant (texte, capture d'écran, image)
- Analyse automatique pour catégoriser la demande (problème de mot de passe, accès SharePoint, incident de cybersécurité, etc.)
- Réponse générée automatiquement en moins de 2 minutes, dans le ton et le format de l'entreprise
- Création de tickets en cas d'incident plus complexe avec extraction des données critiques et envoi aux responsables (mail, Teams, SMS)

### APPROCHE GOUVERNÉE ET PERSONNALISABLE

- **Capitaliser** sur l'existant (connaissances internes)
- **S'intégrer** aux outils métiers via API ou connecteurs (264 disponibles)
- **Adapter** ses réponses grâce aux retours utilisateurs (système de notation, ajustement des modèles)
- **Gérer** des tâches automatisées (exécution de scripts, accès aux datasheets techniques)

Antemeta propose une interface de chat IA interconnectée aux bases de données internes capable de :

- retrouver des tickets en cours ou passés
- générer des rapports
- s'adapter à un langage naturel en conversation

La démonstration a mis en lumière la capacité de l'IA à interroger des bases de données en temps réel, à intégrer des réponses dans des présentations et à s'adapter à des requêtes métier précises.

**L'IA AU SERVICE DE LA GOUVERNANCE, ILLUSTRÉ À TRAVERS LA SOLUTION : STEROÏD.**

Cette offre d'« IA fermante », conçue pour exécuter localement des applications d'intelligence artificielle sans exposition des données à l'extérieur. Reposant sur des technologies open source et déployée directement sur l'infrastructure du client, cette solution garantit une confidentialité renforcée, une conformité réglementaire et une maîtrise totale des données. Une approche souveraine pensée pour répondre aux exigences de sécurité des organisations les plus sensibles.



Développée par H2 Studio, Steroïd est une plateforme d'IA conçue pour assister les dirigeants, membres de comités, administrateurs et actionnaires dans leurs prises de décision.

L'outil permet d'interroger en langage naturel l'ensemble de la base de connaissances de l'entreprise et de générer, en temps réel :

- des analyses expertes
- des rapports de synthèse
- des comptes rendus de réunion
- des listes de tâches structurées avec échéances et responsables

L'intégration de l'IA a été conçue pour s'inscrire naturellement dans les outils utilisés au quotidien (Microsoft Word, Excel, PowerPoint), sans ajout de surcouche logicielle complexe. Les collaborateurs peuvent ainsi interagir avec des agents IA spécialisés, directement depuis leur environnement de travail, via des modules développés sur mesure ou par H2 Studio.

L'outil se distingue par sa capacité à décomposer les questions complexes, à vérifier ses sources (fact-checking) et à personnaliser les réponses selon les profils utilisateurs et les niveaux d'habilitation.

**Veille stratégique augmentée : l'IA qui éclaire vos décisions en temps réel**

Antemeta, spécialisée dans la résilience et la donnée, est une société experte en sécurisation des environnements numériques.

La plateforme Steroïd est une IA souveraine, hébergée localement ou sur un data center dédié, conforme aux normes françaises et européennes. Elle est conçue pour des usages métiers concrets avec une interconnexion aux bases internes de l'entreprise. La tarification est fixe quel que soit le nombre d'utilisateurs. La plateforme est scalable et résiliente.

**FONCTIONNALITÉS PRINCIPALES**

- Injection et vectorisation de toutes les données internes.
- Analyse croisée des documents, avec génération de réponses structurées.
- Justification des réponses par les sources chaque affirmation est cliquable avec surlignage dans le document d'origine.

**AGENTS SPÉCIALISÉS**

- Checker / cross-checker : vérifie les réponses en accédant directement à la bonne ligne dans les documents PDF.
- Fast-checker : identifie les erreurs ou zones floues dans les contenus.
- Compte-rendu automatique : enregistrement d'une réunion via l'outil Record, qui produit un CR en deux minutes (actions, échéances, intervenants).
- Créateur de slides ou rapports : exports automatiques au format Word ou PowerPoint, chartés aux couleurs de l'entreprise.

**AVANTAGES**

- Possibilité de définir des droits d'accès très fins selon les services, jusqu'à un niveau de fichier précis.
- Chaque service a ses propres instructions de réponse (niveau de synthèse, ton, format...).
- Toute question et réponse reste strictement personnelle à l'utilisateur.

## Finis les comptes rendus fastidieux : vos réunions résumées automatiquement, en toute souveraineté

La plateforme AntemetA AI est une solution IA comprenant plusieurs modules interconnectés et intégrés à la suite Office (Word, Excel...). Elle permet d'enregistrer une réunion via micro, fichier ou capture vidéo puis d'en générer une transcription textuelle complète et une synthèse automatique. Le contenu est ensuite structuré selon les templates de l'entreprise. Elle se distingue par une approche pragmatique et modulaire.



L'IA s'intègre directement dans les processus métiers (RH, commerce, réunions, appels d'offres...). Tout se fait en quelques clics et sans compétence technique.



### SOUVERAINETÉ ET SOBRIÉTÉ

Contrairement aux solutions cloud, les données ne sortent jamais de l'infrastructure du client :

- Aucune interconnexion avec Gmail, Office 365 ou autres services en ligne
- Possibilité de fonctionner en réseau fermé, avec stockage et un traitement en local
- Modèle économique stable, sans surcoût en fonction du nombre d'utilisateurs ou du nombre de requêtes IA
- Modèles spécialisés et optimisés, privilégiant l'usage de plusieurs petits modèles adaptés plutôt qu'un modèle généraliste



### INTEROPÉRABILITÉ ET TRAÇABILITÉ

Grâce à un portail unique (Smart), les utilisateurs accèdent à l'historique des réunions, peuvent revoir, éditer ou enrichir les comptes rendus.

Chaque transcription est archivée, horodatée et non falsifiable, permettant de garantir la véracité des échanges.

Un chatbot interne permet également de re-questionner les contenus générés, grâce à des agents IA personnalisables selon les métiers.

## Incidents & NIS2 : formalisez vos déclarations en quelques minutes grâce à l'IA

Cet atelier présente une solution d'intelligence artificielle développée pour faciliter la conformité aux réglementations normées comme NIS2, à travers deux outils principaux : **un agent conversationnel intelligent** et **un assistant de remplissage automatique de formulaires normés**.

**NIS2 ?** | NIS2, comme le RGPD à l'époque, impose aux entreprises une obligation de transparence et de réaction rapide en cas d'incident impactant la cybersécurité ou la production.



### OUTILS 1 : LE "CHAPS"

Cet agent conversationnel IA intégré dans une interface graphique permet :

- De répondre à des questions réglementaires complexes
- De croiser les données connues de l'entreprise avec les critères de la directive (CA, effectifs, secteur...)
- D'accompagner les utilisateurs dans la gestion d'incidents

L'IA s'appuie sur des documents réglementaires, des bases de données internes, et des formulaires d'entreprise déjà intégrés à son socle de connaissances.



### OUTILS 2 : NORME

Cet outils est un add-in Word permettant de :

- Pré-remplir automatiquement des formulaires complexes
- Extraire les bonnes informations depuis des mails ou documents internes validés
- Gagner du temps et réduire les erreurs en situation d'urgence

Ex : en cas de ransomware, l'IA complète instantanément les champs nécessaires (SIRET, usine concernée, impact opérationnel...) grâce à sa base de connaissances.



### AVANTAGES

- Automatisation des tâches administratives post-incident
- Réduction des délais de réaction réglementaire
- Simplicité d'usage via des interfaces familières (chat, Word)
- Sécurité des données respectée : tout est hébergé localement ou dans des environnements cloisonnés

## Souveraineté & IA : retour d'expérience avec Mistral AI

L'usage massif de ChatGPT, notamment par des profils métiers non techniques, a bouleversé les repères. Les directions métiers ont pris les devants, obligeant les experts techniques à repenser leur posture et leurs offres. Un POC est mené avec la CCI Auvergne-Rhône-Alpes pour créer un assistant intelligent destiné aux conseillers. L'objectif est de proposer des prestations plus adaptées aux besoins exprimés par les clients.

### D'UNE POSTURE EXPLORATOIRE À UNE STRATÉGIE R&D STRUCTURÉE

Face à l'adoption rapide des outils génératifs et à la pression des directions métiers, une démarche de R&D est lancée pour :

- Comprendre les enjeux métiers, techniques et réglementaires.
- Proposer une offre sur-mesure mêlant IA générative et souveraineté des données

Deux grandes solutions sont identifiées :

- **Des solutions propriétaires** (API payantes, peu modifiables).
- **Des solutions open source**, potentiellement hébergées en interne pour plus de maîtrise.



### MISE EN ŒUVRE CONCRÈTE

Architecture mise en place :

- Interface web via Streamlit.
- Base vectorielle Chroma contenant le catalogue des prestations.
- LLM Mistral (d'abord 7B, puis Small 24B).
- Moteur d'embedding Jina "Javanese", optimisé pour le français.

Fonctionnement :

L'utilisateur saisit une demande. Elle est enrichie sémantiquement puis la recherche est effectuée dans la base vectorielle. Le top 10 des prestations est extrait puis est filtré par le LLM pour évaluer la pertinence de chaque prestation. Les résultats sont, ensuite, renvoyés avec un argumentaire à un conseiller.

La solution a rapidement trouvé sa place sur le terrain avec plus de 80 % des recommandations générées ont été validées par les conseillers. L'interface, intuitive et simple à prendre en main, a facilité l'adoption. Le dispositif montre également un fort potentiel d'extension, notamment vers d'autres profils métiers comme les centres de contacts clients.

### Et demain ? Vers une architecture d'agents hybrides ?

Un mix technologique entre des modèles légers hébergés localement, pour la gestion de données sensibles, et des modèles plus puissants en cloud pour une recherche plus large et contextuelle. Cette approche vise à offrir un système à la fois performant, sécurisé et modulable, capable de s'adapter aux enjeux spécifiques de chaque métier.

## Gouvernance des données : pour une IA sécurisée et efficiente

Alors que d'ici 2026, 80 % des applications devraient intégrer de l'IA, sans données fiables, l'IA n'est qu'une illusion de performance. Une façade brillante, mais fragile. Pour éviter les dérives et exploiter l'IA de manière durable, la gouvernance des données doit être repensée. Protéger les données, c'est aussi protéger l'identité des utilisateurs et les droits d'accès aux ressources. La gouvernance ne se limite donc pas aux fichiers mais englobe aussi les identités, les devices, et les usages.



La plupart des entreprises voient le potentiel de l'IA, mais peinent à le maîtriser

- **Confiance & transparence** : garantir une IA explicable et non discriminatoire
- **Sécurité & conformité** : Respect des réglementations comme le RGPD, l'AI Act ou des normes sectorielles
- **Éthique** : L'IA doit être alignée avec les valeurs de l'entreprise.
- **Performance** : L'innovation est possible qu'avec des données exploitables et bien gouvernées.



- Pertes, duplications, accès non maîtrisés des données.
- Biais algorithmiques (données incomplètes, non représentatives).
- Fuites de données, provoquées par des erreurs internes ou des usages imprudents de l'IA
- Usage non éthique ou illégal : une IA peut générer des réponses inappropriées.
- Risque humain : 93 % des organisations craignent les erreurs internes liées à l'IA générative.

Pour répondre à ces enjeux, ASI propose la plateforme complète de Microsoft Purview :

- **Cartographie des données sensibles** (scans réguliers sur l'environnement Microsoft 365),
- **Applique des labels de classification** (confidentiel, public...),
- **Contrôle les accès aux données**, même lorsqu'elles sont interrogées via Copilot ou d'autres IA,
- **Bloque ou alerte** en cas de comportements à risque (tentative de résumer un document confidentiel via une IA tierce),
- **Surveille les usages des IA internes ou externes** (Copilot, ChatGPT Enterprise, etc.)..

## Y a-t-il (vraiment) une place pour la durabilité avec l'IA ?

L'IA générative remplace progressivement les moteurs de recherche dans certaines recherches, avec une approche plus structurée et des contenus plus riches. L'effet « wow » de l'IA cache une complexité plus profonde. Il est important de dépasser la fascination et se concentrer sur l'impact réel. Il faut donc être réaliste, il n'y a pas de retour en arrière possible. Nous devons mesurer les impacts pour prendre les bonnes décisions et pour progresser par itérations, en intégrant des exigences éthiques, environnementales, sociales.

### “LA DURABILITÉ EST LE FAIT DE SATISFAIRE LES BESOINS DU PRÉSENT SANS COMPROMETTRE CEUX DES GÉNÉRATIONS FUTURES.”

(Commission Brundtland, 1987)

#### • Encadrement & régulation

L'AI Act européen marque une avancée majeure dans la régulation de l'intelligence artificielle, avec une approche fondée sur les niveaux de risque. Certaines pratiques sont strictement interdites, telles que la notation sociale, la surveillance de masse ou les manipulations comportementales à grande échelle.

#### • Vers un numérique responsable

L'intégration de l'IA ne peut se faire sans une démarche d'éco-conception et une réflexion globale sur un numérique responsable. Il faut anticiper les impacts environnementaux et sociaux dès la phase de conception, à l'image de ce qui a été initié avec le RGPD. Des référentiels comme ceux de l'INR offrent des repères précieux pour orienter les choix techniques et éthiques.

### QUELQUES CHIFFRES

Temps pour atteindre 1M d'utilisateurs:

- Netflix : 3,5 ans
- Facebook : 10 mois
- Instagram : 2,5 mois
- ChatGPT : 5 jours.

Étude Bpifrance (2025) :

- 32 % des PME/ETI utilisent l'IA générative.
- Près de 50 % ont déjà lancé une stratégie IA.

Quels sont les principaux usages :

- 70 % pour la création de contenu
- 60 % pour la recherche/analyse de données
- 36 % pour la traduction

- Une seule requête GPT représente environ 4,3g CO2, soit 3 à 5 fois plus qu'une recherche Google ([news](#))
- Les datacenters pourraient devenir la 3e ou 4e source mondiale de consommation d'électricité d'ici 2030.

## IA & accessibilité : l'innovation au service de l'inclusion

L'accessibilité numérique vise à rendre les services numériques utilisables par tous. Ce n'est plus une option, mais une exigence. L'IA accélère son déploiement et facilite la mise en conformité, tout en réduisant les coûts, accélérant l'implémentation et en industrialisant les bonnes pratiques.

Capacité d'un service numérique à être utilisé par tout le monde, quels que soient leurs handicaps.

### DÉFINITION

- **Visuel** : malvoyance, cécité → lecteurs d'écran, dispositifs braille.
- **Auditif** : surdité, perte auditive → transcription automatique, sous-titrage.
- **Moteur et cognitif** : troubles moteurs, autisme, dyslexie → contenus simplifiés, navigation adaptée

### COMMENT L'IA ACCÉLÈRE L'ACCESSIBILITÉ ?

#### • Génération de contenus alternatifs

→ description automatique d'images via IA générative, amélioration des expériences pour les utilisateurs de lecteurs d'écran...

#### • Adaptation dynamique des interfaces

→ L'IA est capable de modifier automatiquement les contrastes et les tailles de police selon le handicap détecté.

#### • Automatisation des tests d'accessibilité

→ L'IA permet de tester rapidement des centaines de pages selon les normes WCAG/RGAA.

*L'humain reste indispensable pour évaluer la nuance, la pertinence et l'expérience réelle de l'utilisateur.*

### Obligations légales :

- Loi du 11 février 2005 : égalité des droits et des chances.
- Loi pour une République numérique (2016)
- European Accessibility Act (en vigueur au 28 juin 2025) : S'applique aux services publics, entreprises > 250M € CA et progressivement aux PME.

### Outils cités :

- Figma et plugins d'audit d'accessibilité.
- Copilotes IA pour tests automatisés (ex. copilote Microsoft, IA cloud).
- LexiVoice (2025) : traduction en temps réel avec voix naturelle pour utilisateurs multilingues.
- Space Synapse : traduction automatique en langue des signes via avatar.

### UN CHANGEMENT CULTUREL : FINI LES POC, PLACE À LA PRODUCTION

Fin 2023, le groupe Adéquat fait face à une pression croissante de son conseil de surveillance : il faut accélérer sur l'intelligence artificielle générative.

Jusqu'alors, seuls quelques projets ponctuels étaient menés (prédiction de chiffre d'affaires, OCR...), sans réelle industrialisation. L'arrivée en force des éditeurs d'IA a convaincu la direction générale : il ne s'agissait plus de tester, mais d'agir. Une consigne claire a été donnée : il faut faire de l'IA, rapidement et à grande échelle.



1400 utilisateurs actifs sur 1500 collaborateurs avec une moyenne de 700 utilisateurs actifs par semaine.

Les métiers eux-mêmes proposent de nouveaux cas d'usage, inversant la logique traditionnelle de l'IT descendante.

### UN ASSISTANT IA DÉPLOYÉ EN 6 SEMAINES

- **Avril 2024** : décision de créer un assistant interne basé sur l'IA générative,
- **Juin 2024** : déploiement éclair auprès de 1 500 collaborateurs, sans formation ni onboarding formel,
- 

Architecture technologique :

- Stack Microsoft, notamment Azure et Copilot.
- Utilisation de LLM variés (OpenAI,...).
- Intégration de bases vectorielles pour la recherche documentaire.
- Approche modulaire avec des agents spécialisés selon les cas d'usage.

Le groupe a alors opéré un virage stratégique majeur avec l'abandon des POC classiques au profit d'un passage direct en production et l'usage réel devient le nouvel indicateur de réussite. Nous remarquons que l'adoption côté métiers est forte, portée par des usages concrets et accessibles. En revanche, les équipes IT, plus habituées aux cycles traditionnels de développement manifestent parfois une forme de résistance culturelle. Un autre enjeu clé est de mieux intégrer l'IA dans les chaînes de production existantes, en la considérant comme une brique cognitive parmi d'autres, au même titre qu'une API ou un microservice.

## Gouvernance, sécurité et frugalité : découvrez le cockpit IA Synrun

ProoveUp et Askills adoptent une démarche résolution pragmatique et frugale pour déployer l'IA dans les organisations. Leur conviction est que l'IA ne doit pas être un gadget mais un levier concret d'amélioration continue. Tout commence par une cartographie des processus internes, au plus près du terrain. En identifiant les tâches répétitives, chronophages ou à faible valeur ajoutée, l'objectif est de prioriser les cas d'usage simples, utiles et mesurables, là où l'impact est immédiat.



### Synrun, la plateforme qui structure l'action

Pour rendre cette approche opérationnelle, les deux partenaires ont développé Synrun, une boîte à outils IA sur-mesure. Elle permet de :

- Recueillir les besoins métiers directement auprès des équipes via un assistant conversationnel intégré,
- Structurer les données (temps passé, étapes, documents utilisés, contraintes identifiées...),
- Cartographier les processus de façon visuelle et collaborative,
- Quantifier les gains potentiels à chaque étape (économies de temps, amélioration de la qualité, réduction des coûts...)

Synrun intègre également des briques IA concrètes et actionnables, capables d'interagir avec les documents (analyse automatique d'un appel d'offres PDF, par exemple) ou d'effectuer des actions ciblées (comme un inventaire de produits à partir d'une simple photo).

L'objectif est de maîtriser les données, privilégier l'hébergement local quand c'est possible et concevoir des outils adaptés au terrain, en étroite collaboration avec les utilisateurs.

- **Le processus est collaboratif** : les outils sont coconstruits avec les équipes métiers, pour éviter les usines à gaz mal adoptées.
- **L'approche est modulaire** : l'accompagnement peut se faire avec ou sans l'équipe de ProoveUp, selon la maturité de l'organisation.
- **Un volet gouvernance est intégré dès le départ** : cartographie des données utilisées, gestion des accès, encadrement des usages IA.

⚠ Beaucoup d'entreprises abordent l'IA par imitation. Elles souhaitent reproduire ce qu'elles ont vu chez Google ou Microsoft, sans réel cadrage préalable. Résultat, les projets sont lourds, coûteux et souvent voués à l'échec. Gartner estime que 80 % des projets IA échouent ou sont abandonnés.

## Peut-on "contrôler" une IA ? Poursuivra-t-elle nos objectifs ou les siens ?

L'alignement est un concept issu de la recherche anglo-saxonne qui repose sur trois niveaux :

### QU'EST CE QUE L'ALIGNEMENT ?

- **Alignement des intentions** : l'IA comprend-elle correctement les consignes données ?
- **Alignement des comportements** : ses réponses respectent-elles les valeurs humaines, le droit, l'éthique ?
- **Alignement à long terme** : l'IA contribue-t-elle positivement à la société ?



### PAPERCLIP PROBLEM

Les intervenants ont illustré les dérives possibles via une expérience de pensée célèbre : le "paperclip problem" où une IA dédiée à fabriquer des trombones finirait par transformer toutes les ressources (humaines incluses) pour atteindre son but. S'il s'agit d'une caricature, elle soulève des questions fondamentales :

**une IA peut-elle mentir ?  
Tromper ? Se préserver  
en contournant les  
consignes ?**

Les grands laboratoires comme OpenAI ou Anthropic intègrent déjà des méthodes avancées de contrôle des modèles comme le RLHF (Reinforcement Learning with Human Feedback). Cette approche se déploie en trois étapes clés :

1. **Supervised Fine-Tuning** : des experts humains génèrent des réponses modèles pour guider l'IA.
2. **Classement et alignement** : les réponses sont évaluées et classées selon leur qualité ou leur conformité à des principes éthiques, soit par des annotateurs humains, soit par des systèmes de type Constitutional AI.
3. **Reward Model** : un modèle de récompense est ensuite entraîné pour affiner les comportements du modèle principal, en valorisant les réponses alignées.

Au-delà de ces approches en laboratoire, les organisations doivent mettre en place des garde-fous concrets pour encadrer l'IA dans leurs usages internes. Des dispositifs d'encadrement se mettent progressivement en place. Cela passe par :

- Des **instructions claires** dans les prompts, pour limiter les dérives dès la formulation des requêtes.
- Des **systèmes de surveillance** comme Guardrails.ai ou MemoGuardrails, qui analysent les inputs/outputs des modèles en temps réel et détectent d'éventuelles violations (données sensibles, biais, langage inapproprié...).
- Des **outils d'audit et de robustesse** comme Giskard, Holistic Evaluation ou Walid, qui permettent de tester les modèles, vérifier leur fiabilité, et formaliser des critères d'évaluation adaptés aux règles internes de l'entreprise.

## L'anti-Ikigai ou que puis-je donner à l'IA pour m'aider au quotidien ?



Ikigai : la boussole du sens au travail, concept japonais, désigne votre raison d'être professionnelle, la zone dans laquelle vous êtes le ou la plus épanoui(e).

Et si l'automatisation par l'IA devenait une occasion de recentrer son énergie sur ce qui nous anime vraiment ? Cet atelier propose une méthode originale pour choisir les tâches à automatiser tout en préservant son "Ikigai", cet équilibre entre ce qu'on aime, ce qu'on sait faire, ce dont le monde a besoin et ce pour quoi on peut être payé.

### MAIS ALORS POURQUOI PARLER DE L'ANTI-IKIGAI POUR L'IA ?

À l'inverse, l'anti-Ikigai est constitué de toutes les tâches que vous pourriez transférer à une IA pour vous recentrer sur ce qui a du sens pour vous.

- **Ce qui prend du temps** : tâches longues, chronophages, fastidieuses
- **Ce qui est source d'erreurs** : erreurs de saisie, oublis, doublons...
- **Ce que vous n'aimez pas faire** : tâches répétitives, rébarbatives, peu motivantes



### MÉTHODOLOGIE : CARTOGRAPHIER ET DÉLÉGUER INTELLIGEMMENT

*Les participants ont été invités à remplir une fiche avec ces trois cercles et à lister des exemples de tâches correspondant à chaque zone, voire à des zones qui se recoupent.*

L'objectif :

- Identifier les tâches prioritaires à automatiser
- Visualiser où l'IA pourrait être utile, sans toucher à ce qui fait la richesse du métier

**L'IA n'est pas là pour remplacer votre ikigai, mais pour vous en rapprocher.**

## Comment foirer un projet GenAI ?

À contre-courant des discours enchantés sur la GenAI, l'atelier propose un regard lucide, nourri de retours terrain, sur les échecs les plus fréquents. Un bon projet nécessite du cadrage, des cycles d'évolution, et une vraie gouvernance. Le projet ne peut pas être uniquement porté par l'IT, les métiers doivent être partie prenante.

### TROIS SCÉNARIOS D'ÉCHEC (TROP) FRÉQUENTS :



#### L'ENTHOUSIASME SOLITAIRE

Un responsable innovation, séduit par les capacités de ChatGPT, lance un projet de chatbot client. Le projet est rapide mais réalisé sans coordination avec la DSI, sans un cadrage métier et sans réflexion sur la sécurité. Résultat, l'outil est lent, imprécis, coûteux et expose des données sensibles. Le projet est abandonné.



#### MENÉ UNIQUEMENT PAR L'IT

Cette fois, tout est structuré avec des développeurs et data scientists, une infrastructure solide et sécurisée, des milliers de documents intégrés. Sauf que le corpus est mal sélectionné, les données obsolètes, et surtout les utilisateurs n'ont jamais été consultés. L'outil n'est donc pas utilisé et il meurt à petit feu malgré les moyens engagés.



#### L'IMPORTANCE DU SUIVI

Ici, le projet est bien lancé. Les métiers sont impliqués, le cas d'usage pertinent et les résultats au rendez-vous (gain de temps, meilleure qualité, forte adoption). Mais six mois plus tard, le moteur IA a évolué, les documents de référence ne sont plus à jour. Et sans processus d'entretien ou de supervision, le système devient progressivement inutilisable.

C'est un système vivant, qui nécessite :

- La sécurité est centrale, il est impératif de penser "privacy by design" dès le départ.
- Une stratégie doit être définie avec des cas d'usage précis. Le tout porté par une collaboration étroite entre les métiers, l'IT, la sécurité et le juridique avec une architecture sobre mais robuste.

L'IA générative a un potentiel immense, mais elle ne fait pas tout toute seule. Pour concevoir des outils utiles, fiables et adoptés, il faut avant tout du collectif, de la méthode, et du bon sens.

### FACE À CES CONSTATS

## GenAI : de l'idée à la mise en production

Aujourd'hui, 80 % des projets GenAI ne passent pas le cap de la production et seuls 30 % démontrent une vraie valeur au bout d'un an. Ces chiffres ne traduisent pas un échec de la technologie, mais bien la nécessité d'un cadre solide, d'un collectif engagé, et d'une stratégie claire. L'enjeu n'est pas de "faire de l'IA", mais de l'intégrer de manière cohérente et maîtrisée dans les processus métiers.



Réussir un projet GenAI, c'est conjuguer vision stratégique, excellence technique, ancrage métier et la capacité à itérer. C'est aussi adopter une posture d'humilité : l'IA générative ne remplace pas l'humain, elle le prolonge. **La technologie est prête à nous de l'utiliser avec méthode, sens et responsabilité.**

### 1 Le rôle central du corpus documentaire :

Le moteur LLM n'est aussi bon que les données qu'on lui fournit. Plutôt que de miser sur un volume massif de documents, mieux vaut :

- Sélectionner un corpus limité, propre, structuré et à jour,
- Définir une charte documentaire claire (ton, formats, exclusions...),
- Anticiper les évolutions et mises à jour nécessaires.

### 2 Des cas concrets et une méthode éprouvée :

Avanade accompagne ses clients sur toute la chaîne, de l'acculturation à l'industrialisation :

- Cadrage avec l'identification des cas d'usage, l'évaluation de la maturité, la formation des équipes,
- POC (architecture sécurisée, indicateurs de performance, premières intégrations),
- Déploiement et supervision (intégration SI, outils de monitoring, sécurité RGPD),
- Industrialisation (adaptation continue, FinOps, gestion du changement).

L'approche intègre aussi des outils comme les LLMs spécialisés par fonction (rédaction, analyse, support...), ou le "LLM as a judge", où un modèle évalue la qualité des réponses d'un autre, pour fiabiliser les résultats.

### 3 Une équipe interdisciplinaire au service du projet :

Un projet GenAI mobilise une diversité de profils : architectes cloud, data engineers, juristes, experts sécurité, UX designers, mais aussi des utilisateurs finaux. Ce collectif est essentiel pour ancrer la solution dans les usages réels, anticiper les contraintes et garantir l'adoption.

## Agents AI : attention, peinture en cours de séchage...

Présenter l'état de l'art et les perspectives autour de l'agentique, c'est-à-dire les systèmes d'agents intelligents capables d'interagir, d'orchestrer des tâches et de collaborer entre eux, dans une logique de multi-agent. L'équipe a partagé son retour d'expérience sur des projets concrets, ses réflexions sur les architectures cibles, ainsi que les enjeux techniques, fonctionnels et organisationnels de ces systèmes.

### QU'EST-CE QUE L'AGENTIQUE ?

Contrairement aux assistants classiques qui réagissent en temps réel à des requêtes, un agent IA est :

- Un système autonome ou semi-autonome, capable de réaliser des tâches à la place de l'humain,
- Un outil programmable, déclenché par un orchestrateur ou d'autres agents,
- Un acteur d'un système distribué, en lien avec d'autres agents spécialisés (logique multi-agents),

Les agents combinent ainsi automatisation (type RPA) et technologies GenAI, ouvrant la voie à une nouvelle forme d'intelligence collective entre machines.



#### CAS PRATIQUE 1 : LE SERVICE CLIENT

1. Un client se plaint via chatbot.
2. Un agent IA détecte le sentiment de la demande.
3. Un autre agent évalue le contexte client (facture, service, niveau de satisfaction) en se connectant à plusieurs SI.
4. Une proposition personnalisée est générée (remise, réponse adaptée au ton).
5. La réponse finale est reformulée automatiquement en fonction du ton du client.

Chaque étape repose sur des agents spécialisés qui collaborent, ce qui permet un traitement plus fin, plus rapide, et adapté.



#### CAS PRATIQUE 2 : APPLICATION RH

Dans les RH :

- Un assistant central peut sous-traiter à différents agents les réponses sur la paie, les congés, la formation, etc.
- Ces agents interrogent en parallèle différentes bases et restituent une réponse agrégée à l'utilisateur.

Dans l'industrie ou la finance :

- Les agents permettent de monitorer des chaînes logistiques, d'identifier des anomalies, ou d'optimiser des décisions métier

Les agents reposent sur une logique modulaire (type microservices) hébergés dans des environnements sécurisés. Des normes inter-agents (orchestration, sécurité, protocole de communication) sont en cours de définition mais encore immatures. Avanade et Accenture développent des agents types (extraction, prédiction...) en boucle de code réutilisable.

### Avizia x Manutan : l'IA au service de la gestion massive de fiches produits

Manutan est un distributeur B2B, surnommé l'"Amazon des pros", avec plus d'un million de références produits allant des stylos aux machines industrielles. Il reçoit 20 000 nouveaux produits par mois, à intégrer rapidement et correctement sur son site e-commerce.

#### LE PROBLÈME ?

Avec une gestion manuelle, lente et peu fiable, les fiches produits reçues sous forme de fichiers Excel hétérogènes, sont mal structurées. Les fournisseurs ne respectent pas toujours de normes communes. Le temps de traitement est donc de 15 jours à 1 mois, via des prestataires externes ou équipes internes. La qualité de données est insuffisante pour un bon référencement produit.

1

#### Une plateforme sur mesure :

Développée en Python, avec une interface en Streamlit pour usage interne, elle est hébergée sur Azure, utilisant les modèles OpenAI hébergés en France (conformité souveraine) avec un upload direct des fichiers Excel par les utilisateurs métier.

2

#### Injection du fichier Excel:

Traitement par un LLM configuré pour :

- Réécrire les titres produits de façon claire, normée et SEO-friendly
- Générer des bullet points de description et fourniture de titres exemples pour guider la rédaction
- Respecter les contraintes métiers (longueur, vocabulaire, absence d'abréviations)
- Utilisation du RAG (Retrieval Augmented Generation)
- Recherche vectorielle dans la base produit pour trouver des titres similaires
- Génération de toutes les déclinaisons produit d'un coup avec un contrôle manuel final pour validation

3

#### Défis et apprentissages :

Les modèles de langage (LLM) ne devinent pas les normes produits. Ils doivent être explicitement guidés. Sans contexte suffisant, le modèle peut générer des informations erronées ou inventées.

- L'utilisation de prompts structurés avec l'intégration d'exemples concrets issus de la base interne,
- L'ajustement des consignes est progressif en lien étroit avec les équipes métier grâce à une fourniture maximum de contexte. Des instructions claires dans le prompt permettent de limiter les dérives

⚠ Les LLM n'ont pas de réelle perception des longueurs et des limites de caractères. Ils fonctionnent en tokens. Pour garantir, par exemple, un titre de 60 caractères maximum, une double étape de génération et une boucle de réécriture automatique ont été mises en place.

## GEN BI : qui est le meilleur ami du data analyst ?



GenBI est un assistant intelligent conçu pour simplifier l'accès aux données et automatiser les analyses métiers, sans besoin de compétences techniques. Pensé comme le meilleur allié du data analyst, il permet de gagner du temps, fluidifier les prises de décision et valoriser la donnée déjà présente.

L'équipe à l'origine travaille depuis 15 ans dans l'univers de la data et a vu l'émergence de l'IA générative comme une opportunité. GenBI permet de poser une question métier en langage naturel (« Quel est le chiffre d'affaires du mois dernier ? ») et d'obtenir instantanément :

- Une réponse chiffrée ou visuelle (graphique, tableau...)
- Le code SQL utilisé (utile pour les experts)
- Une interaction conversationnelle fluide, comme avec un véritable analyste

#### FONCTIONNEMENT

L'architecture repose sur plusieurs agents :

- Un **agent de validation** de la demande
- Un **générateur de requêtes SQL** basé sur le schéma de données et enrichi par des exemples métiers
- Un **moteur de restitution** qui reformule les résultats de façon claire, contextualisée et visuellement exploitable

Tout cela s'appuie sur des technologies standards : LLM (comme GPT via Azure), BigQuery, Streamlit ou Teams selon les clients, avec intégration possible à l'existant (Google Chat, site web interne...).



#### CONDITIONS DE RÉUSSITE

- Une bonne organisation des données (qualité, structuration, catalogage)
- Un cadrage clair des cas d'usage à couvrir (périmètre raisonnable, données prêtes à l'emploi)
- Une phase d'apprentissage continue (validation manuelle, enrichissement des règles, indicateurs de confiance)
- Une interface user-friendly et évolutive, pensée pour s'adapter au public cible.
- Les COMEX peuvent disposer de chiffres actualisés en temps réel



Chaque déploiement est sur mesure, selon les infrastructures, les cas d'usage et les métiers du client. L'idée n'est pas de remplacer les data analysts, mais de leur offrir un portail d'entrée intelligent, pour que chacun puisse accéder à la donnée sans dépendance technique.

## L'IA révolutionne le Datacenter : le retour d'expérience d'une grande banque française



En adoptant le refroidissement liquide pour un projet IA de type HPC, une banque hexagonale :

- A multiplié par quatre sa puissance de calcul.
- A remboursé l'investissement de la plateforme en moins de cinq ans via les économies d'énergie.

Points de vigilance et recommandations

- **Anticiper la compatibilité** des composants entre eux (refroidissement, format, densité).
- **Travailler avec des spécialistes** (plombiers industriels, ingénieurs thermiques) pour sécuriser les installations.
- **Délimiter clairement les responsabilités** contractuelles entre l'hébergeur, le fabricant et le client, notamment via un échangeur thermique comme point de référence.

L'IA générative a profondément bouleversé le marché du compute, générant de nouveaux besoins en serveurs accélérés, notamment équipés de GPU. En France, ce marché représente désormais plus de 130 millions d'euros par trimestre, sans cannibaliser le marché traditionnel.

Il s'agit bien d'un marché additionnel, tiré par les investissements publics et l'urgence des entreprises à intégrer l'IA. Cette révolution technologique nécessite une vraie maturité d'infrastructure. Une montée en puissance électrique difficilement soutenable. Les dernières générations de composants affichent des puissances énergétiques inédites :

- Jusqu'à 1 000 watts par CPU
- Plus de 1 200 watts par GPU

⚠ Or, les datacenters ont été conçus il y a 10 ou 12 ans pour des charges bien moindres. Cette évolution ultra-rapide crée un décalage critique. Les infrastructures ne sont plus adaptées à ces nouvelles charges thermiques et électriques, ce qui rend le refroidissement par air rapidement insuffisant.

Il faut s'orienter vers des solutions de refroidissement liquide (direct liquid cooling), plus efficaces et plus économiques à long terme afin :

- **De supporter des densités plus élevées** (jusqu'à 160 kW/rack)
- **De faire baisser le PUE** (Power Usage Effectiveness) de 1,15 à 1,07
- **D'atteindre un ROI rapide** (entre 7 mois et 5 ans)

Le refroidissement liquide est donc non seulement plus performant, mais aussi plus rentable, notamment dans des projets IA à forte intensité.

## Optimisez l'IA avec les RAG : génération augmentée, impact maximal

Lors de cet atelier, le concept de RAG (Retrieval-Augmented Generation) nous a été démystifié avec une approche très concrète pour contourner les limites bien connues de l'IA générative : hallucinations, biais, ou encore perte de contrôle sur les données. Car oui, entre les jurisprudences inventées par ChatGPT ou les réponses absurdes (voire discriminantes) de certains outils, les dérives existent. Il est important d'agir avec prudence.

### LE RAG, UNE SOLUTION MALIGNE ET EFFICACE

Au lieu de recréer un modèle ou de le réentraîner (ce qui coûte cher, prend du temps et demande des compétences rares), il suffit de connecter un modèle existant à sa propre base documentaire. Ainsi nous obtenons, des réponses bien plus fiables, avec une solution rapide à déployer, économique et souveraine puisque les données restent en interne. Cette solution s'adapte divers contextes, les cas d'usage ne manquent pas !

### COMMENT ?

Le processus s'articule en plusieurs étapes :

- Découpage des documents (PDF, PowerPoints, etc.) en petits blocs de texte.
- Vectorisation de ces blocs via un moteur d'embedding, pour leur donner une "empreinte numérique".
- Stockage dans une base vectorielle.
- Lorsqu'une requête est envoyée au LLM, cette requête est elle aussi vectorisée, puis comparée aux documents via une recherche sémantique (par distance cosinus) pour trouver les passages les plus pertinents.
- Le LLM répond ensuite en s'appuyant sur les documents pertinents, garantissant ainsi une réponse contextualisée.

### DÉMONSTRATION

L'intervenant a présenté deux outils open source :

- **LM Studio**, pour télécharger et interagir avec des modèles LLM localement.
- **AnythingLLM**, pour simuler facilement un RAG avec une base documentaire personnalisée.

Même sans infrastructure lourde, il est possible de construire rapidement un assistant intelligent fiable.

Dell Technologies accompagne ses clients via son framework IA Factory, une méthode structurée allant :

- de l'identification du cas d'usage,
- à la qualification des données,
- jusqu'au déploiement de l'infrastructure et de l'écosystème technologique (avec des partenaires comme Mistral, Hugging Face, Scaleway, ou Pointy2).

## Révolutionnez votre approche des données avec Dell AI Factory & Dell Data Lakehouse

Une démarche IA générative pensée comme un levier stratégique, où la donnée joue le rôle central, à travers deux cadres méthodologiques:

- Les frameworks AI Factory
- L'approche Dell Data Lakehouse

Nous recentrons notre attention sur la donnée pour la placer au cœur du jeu. C'est un peu le soleil autour duquel gravitent tous les usages métier et technos. Le message est limpide : ce ne sont pas les modèles d'IA qui font la réussite d'un projet, mais bien la qualité, la gouvernance et l'exploitation intelligente des données. La célèbre maxime "Garbage In, Garbage Out" illustre bien ce constat: sans données fiables, même le meilleur algorithme s'égare.

Structuration, gouvernance, sécurité, souveraineté ou encore monétisation des données deviennent alors les vrais piliers de performance. Quant aux technologies comme la 5G ou le quantique, elles sont vues comme de simples accélérateurs, pas le cœur du réacteur.

### QUELQUES CHIFFRES

- 73 % des projets IA échouent à cause de problèmes liés à la donnée.
- 78 % du temps est encore absorbé par la gestion de la donnée au détriment de l'innovation.
- Jusqu'à 92 % des projets IA échouent dans les secteurs sensibles (comme la défense) en raison de contraintes réglementaires.

### UN TRIPTYQUE GAGNANT

L'approche proposée repose sur: un cas d'usage métier clair (chatbot, computer vision, génération de contenu...), des sources de données adaptées et enfin une infrastructure technologique appropriée, qu'elle soit frugale ou très puissante. Le framework AI Factory s'appuie sur plus de 800 cas d'usage étudiés et une trentaine de déploiements en production dans le monde, notamment dans des environnements critiques comme la presse ou les agences gouvernementales américaines.



Un point important souligné concerne les silos de données (data warehouses, data lakes, fichiers plats...), souvent hétérogènes, peu interopérables et mal gouvernés. Le Dell Data Lakehouse se positionne ici comme une solution unifiée, permettant d'ingérer, nettoyer, orchestrer et valoriser les données (structurées ou non) via des pipelines complets, tout en intégrant des outils de monitoring, de versioning et de visualisation.

## Orienter l'innovation IA par la contrainte carbone-énergie : exploration d'un cas d'usage et de ses dépendances

Comment penser l'IA à l'heure de l'urgence climatique, en s'appuyant sur les travaux du Shift Project ?

### LE CADRE DE RÉFLEXION DU SHIFT PROJECT

- Une dépendance persistante aux énergies fossiles, dont les réserves s'épuisent,
- Une nécessité absolue de réduire les émissions de gaz à effet de serre pour freiner le réchauffement climatique.
- L'enjeu est donc clair : maintenir une société fonctionnelle tout en réduisant drastiquement notre empreinte. Le numérique et plus encore l'IA s'inscrit pleinement dans cette tension.

Ces besoins appellent de nouvelles infrastructures (réseaux, stockage)

De nouveaux usages génèrent de nouveaux besoins,



Celles-ci rendent possibles de nouveaux usages

### Limiter l'impact ?

Deux leviers sont complémentaires, d'un côté, optimiser l'existant (batteries plus petites, calcul local, stockage réduit) et de l'autre, interroger les besoins réels. Tous les capteurs ou fonctions sont-ils indispensables ?

*La contrainte peut devenir un moteur d'innovation sobre, comme l'a montré la crise des semi-conducteurs en 2021.*

### Penser au-delà du carbone ?

- Les limites planétaires (Stockholm Resilience Centre) qui fixent les seuils à ne pas franchir pour préserver l'environnement,
- La théorie du donut (Kate Raworth), qui invite à concilier plancher social et plafond écologique.

## Cyberlake ou comment améliorer sa posture sécurité grâce à l'IA et la Data

CyberLake est une plateforme pensée pour faire face à la forte augmentation des menaces cyber, en combinant centralisation des données, détection intelligente des signaux faibles et automatisation par l'IA. Entre l'essor des attaques géopolitiques, la généralisation du cloud et la multiplication d'outils cloisonnés, les entreprises doivent traiter des volumes gigantesques de logs (jusqu'à plusieurs dizaines de To par jour), avec des ressources humaines et budgétaires souvent limitées. **Résultat** : stocker, croiser et analyser les données devient un casse-tête.

### L'APPROCHE DATA-DRIVEN APPLIQUÉE À LA CYBERSÉCURITÉ

Inspirée du modèle des data lakes, CyberLake permet de centraliser toutes les données de sécurité (logs, événements, traces) dans un entrepôt unique, économique et pérenne. Cette base unifiée est enrichie de modèles IA pour :

- Identifier des signaux faibles (attaques zero-day, comportements anormaux),
- Croiser automatiquement les données issues de différentes sources (IAM, WAF, réseau...),
- Alléger la dépendance aux outils propriétaires souvent coûteux et peu interopérables.

L'objectif est de faire mieux, plus vite, et à moindre coût, sans remplacer l'humain, mais en le renforçant. La solution s'articule autour de trois grandes étapes.

1

**Stockage économique** : les logs sont collectés et archivés durablement dans des solutions cloud optimisées.

2

**Détection automatisée** : des modèles ML spécialisés analysent les flux ; les logs suspects sont comparés à des référentiels vectorisés comme MITRE ATT&CK.

3

**Assistance à l'humain** : les analystes reçoivent des alertes qualifiées, et peuvent interroger les données via un chatbot IA (RAG) convertissant les requêtes en langage naturel.

### AVANTAGES

- Optimisation des coûts (stockage, analyse),
- Meilleure couverture des menaces (zéro-day, signaux faibles),
- Moins de complexité et d'outils à gérer,
- Renforcement de la posture cyber sans alourdir les effectifs.

## "2min chrono" : un podcast généré par l'IA pour couvrir les JO de Paris 2024

Cette présentation retrace la genèse et la réalisation d'un podcast IA automatisé diffusé pendant les 19 jours des Jeux Olympiques, à raison d'un épisode de deux minutes publié chaque matin à 6h. Le projet, mené par une petite équipe de trois personnes, a permis de tester et déployer une chaîne de production audio générative de A à Z en conditions réelles.



L'idée est née de produire un podcast quotidien relayant l'actualité sportive française malgré :

- un délai de 30 jours pour concevoir et déployer le projet,
- une période estivale (juillet) peu propice à la mobilisation des équipes,
- une réticence culturelle forte dans les rédactions sportives à intégrer l'IA.

### MISE EN ŒUVRE

Le projet s'est articulé autour de plusieurs briques technologiques :

- **Scraping intelligent** du site L'Équipe, via des marqueurs CSS spécifiques
- **Système de scoring** basé sur l'origine des athlètes (focus sur les Français) et leur performance (médaillon, buzz),
- **Génération automatique de texte** (IA générative),
- **Voix de synthèse réaliste** via ElevenLabs,
- **Publication automatisée** sur les plateformes de podcast.

Objectif assumé : zéro intervention humaine, hormis le clic de publication à 6h du matin.

### RÉSULTATS

- 19 épisodes générés automatiquement,
- Une moyenne de 3 000 auditeurs par jour, avec des pics à 14 000 écoutes,
- Une portée bien supérieure aux podcasts traditionnels de l'équipe,
- Aucun incident technique ni remontée négative notable.



Ce projet démontre le potentiel créatif des IA génératives dans les médias et le sport. Il ouvre la voie à des usages similaires dans d'autres domaines : politique, juridique, business.. L'équipe a déjà testé une extension sur le basket, avec faux matchs et commentaires simulés par IA. Le seul frein aujourd'hui reste l'imagination humaine.

## Rex Audit IA – Du cadrage à la restitution, ce que nous apprennent les audits IA en conditions réelles

Dans un contexte de croissance rapide portée par une stratégie d'acquisitions :

- Outils disparates, processus non harmonisés avec des données dispersées dans un environnement réglementaire mouvant.
- Une charge de travail élevée concentrée sur des tâches répétitives, un turnover préoccupant et un impact tangible sur la qualité de vie au travail.

### CONSÉQUENCES:

- Perte d'efficacité opérationnelle,
- Risque d'erreurs,
- Prestations non facturées,
- Expérience client dégradée,
- Démotivation des équipes.

### AMBITIONS:

- Retrouver de la cohérence, de la performance et du sens.
- Soulager les collaborateurs,
- Recentrer les efforts sur les missions à valeur ajoutée,
- Offrir une meilleure expérience clients et sécuriser les revenus.



L'IA peut être un puissant levier de transformation à condition d'être intégrée avec discernement. Une trajectoire réaliste et progressive a été construite, à leurs côtés qui s'inscrit dans les usages existants. Une approche qui place les collaborateurs au cœur du changement.

### COMMENT ?

L'objectif est de comprendre les vrais irritants métiers, y compris ceux dont les équipes n'osent pas toujours parler en groupe et identifier ce qui peut être automatisé sans altérer la qualité humaine de la relation client.

- Sensibilisation des dirigeants à l'IA
- Ateliers collectifs et entretiens individuels avec les équipes de chaque service (expertise, juridique, RH),
- Analyse des outils et processus existants,
- Cartographie des cas d'usage potentiels.
- Collecte automatisée de documents dès réception d'un email ou dépôt dans la GED,
- Pré-remplissage automatique des ERP,
- Analyse prédictive de chiffres et de risques
- Réduction des oublis de facturation en automatisant la traçabilité des prestations

## Défi IA dans l'Aérospatial : comment entraîner des modèles sans données suffisantes ?

Un grand industriel aéronautique, fabricant des moteurs CFM et LEAP pour Airbus et Boeing, a fait appel pour automatiser une étape clé de la maintenance : **l'identification des numéros de série des pièces**. Jusqu'ici effectuée manuellement, cette tâche, lente et coûteuse, devenait un goulot d'étranglement face à l'explosion du trafic aérien.

### LE DÉFI ?

L'industriel ne pouvait fournir qu'un faible volume de photos réelles de numéros de série comprenant des données sensibles, non exploitables sur le cloud, trop peu nombreuses pour entraîner un modèle fiable et mélangées avec des pièces avion et hélicoptère aux conditions de vol différentes). Des marquages sont souvent illisibles par les outils OCR classiques car ils sont gravés sur des pièces usées.

### COMMENT ?

Face à l'absence de données suffisantes et à la difficulté de lire des numéros de série gravés sur des pièces usées, l'IA va donc créer des images synthétiques réalistes. Grâce à un outil d'annotation sécurisé, déployé localement en usine et au modèle open source Stable Diffusion. Il est désormais possible de simuler une infinité de cas d'usure et d'altération. Les techniciens ont validé la qualité des images générées, assurant leur crédibilité opérationnelle.

L'infrastructure IA a été installée sur les serveurs internes du client, sans connexion à l'extérieur. Des modèles Transformer, adaptés à la vision par ordinateur, ont été utilisés. Le prompt engineering permet de générer précisément les images nécessaires, comme "un chiffre 4 abîmé sur du métal oxydé".

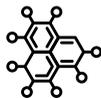


### DES RÉSULTATS CONCRETS RAPIDEMENT OBSERVÉS:

- Une précision de lecture parfois supérieure à celle de l'humain,
- Une heure de traitement économisée par opérateur chaque jour,
- Une solution scalable, prête à accompagner la hausse du volume de maintenance,
- Une méthodologie réutilisable pour d'autres cas, comme la détection de défauts mécaniques.

## Chimie R&D - L'IA au service de la chimie verte : découverte accélérée de polymères biodégradables

Aujourd'hui, la majorité des polymères sont issus de ressources pétrochimiques, avec un impact environnemental important. Le défi est de développer des alternatives durables, capables par exemple de se décomposer naturellement. Dans les centres R&D du client, les chimistes doivent formuler des matériaux aux propriétés précises (résistance thermique, élasticité, biodégradabilité...). Un processus complexe, long et très exigeant sur le plan scientifique, que l'IA vient désormais enrichir en explorant des combinaisons moléculaires inédites.



Pour que l'IA soit capable de générer des polymères pertinents, elle a été entraînée à partir d'un corpus très riche de données scientifiques (brevets, articles, bases publiques et privées). Les polymères ont été convertis en représentations textuelles (formules SMILES) compatibles avec les modèles de langage. Un travail d'extraction automatique d'information a été mis en place pour détecter, dans les documents scientifiques, les formules, propriétés physiques et chimiques.

### COMMENT ?

Une fois les données collectées, un modèle transformeur a été entraîné pour générer de nouvelles formules. Pour éviter les hallucinations (propositions irréalistes), chaque formule générée est ensuite analysée et validée via un module de chimie computationnelle, capable de prédire ses propriétés. Ce double mécanisme (génération puis validation) permet de garantir que les molécules proposées sont à la fois innovantes et scientifiquement crédibles.

Cette approche peut être transposée à d'autres secteurs : génération de modèles 3D dans la CAO, rédaction automatisée de brevets ou d'articles scientifiques, simulation de structures complexes. Elle ouvre la voie à une généralisation de l'IA générative dans les processus industriels, bien au-delà de la chimie.



### DES RÉSULTATS CONCRETS RAPIDEMENT OBSERVÉS :

- **Gain de temps considérable** : certaines tâches qui prenaient plusieurs semaines ont été réduites à 24 heures.
- **Amélioration du processus de sélection** : les chimistes reçoivent des centaines de formules qu'ils peuvent ensuite tester et affiner.
- **Rôle valorisé du chercheur** : l'IA donne de nouveaux leviers. Elle permet d'être encore plus stratégique, décideur, garant de la faisabilité.

## Cyberdéfense assistée par IA : comment les RSSI contre-attaquent ?

A travers des cas d'usage concrets, l'atelier visait à démontrer comment l'IA s'intègre aux architectures de cyberprotection et aux missions du RSSI pour améliorer la détection, l'analyse et la réponse aux menaces, tout en facilitant la conformité et en augmentant la résilience des organisations. Fortinet et SPIE ICS proposent une approche de cyberdéfense augmentée : automatisée, accessible et adaptée à la complexité croissante des infrastructures informatiques.

 Qu'il s'agisse de machine learning ou de deep learning, l'éditeur développe depuis plusieurs années des solutions reposant sur ces technologies, avec plus de 1 300 brevets déposés dont 558 en lien avec l'IA.

Aujourd'hui, Fortinet structure sa gamme de solutions autour de trois grands usages de l'IA appliquée à la cybersécurité :

- **FortiProtect** : pour la détection proactive des menaces (malwares, attaques zero-day), la sécurisation des applications, ainsi que la prévention des usages malveillants de l'IA générative.
- **FortiAssist** : un assistant intelligent dédié à la configuration, à la gestion des incidents et au threat hunting, avec la génération automatique de règles de sécurité et de rapports.
- **FortiSecure** : pour sécuriser les environnements IA eux-mêmes (LLM, conteneurs, données sensibles), en assurant confidentialité, supervision et résilience.



### L'IA OFFENSIVE

Attention car l'IA peut également être utilisée à des fins malveillantes. Des outils comme KaliGPT, une version générative de Kali Linux, permettent de concevoir plus facilement des attaques, même pour des profils peu techniques, comme des fraudes via deepfake ayant entraîné des pertes de plusieurs millions d'euros.

## Frugalité et souveraineté de l'IA : les chercheurs lyonnais progressent !

Cet atelier, co-animé par SPIE ICS et l'INSA de Lyon, met en lumière la richesse de l'écosystème R&D Lyon-Saint-Étienne, composé de plusieurs laboratoires spécialisés dans : les télécommunications (Laboratoires CITI), le traitement massif des données (Laboratoire LIRIS), l'imagerie (LIRIS, CREATIS), la modélisation physique (LabHC), la supply chain (DISP) ou encore les humanités numériques (ERIC).

SPIE et l'INSA collaborent au travers d'une chaire de recherche sur l'IA embarquée dans les infrastructures, autour de grands axes :

- **L'IA appliquée** aux flots de données avec un traitement en continu, en temps réel, et non sur des ensembles de données figés.
- **L'IA frugale**, visant à développer des modèles sur-mesure, moins énergivores, adaptés aux contraintes matérielles et embarquées localement sur les équipements informatiques
- **La souveraineté des données**, qui implique de garder les traitements et les décisions dans les infrastructures internes de l'entreprise, sans dépendance excessive au cloud.
- **La maîtrise et l'explicabilité des décisions** de l'IA, afin de renforcer la confiance dans les résultats produits.



Pour renforcer la dynamique de recherche appliquée, Spie ICS s'appuie sur ces thèses CIFRE :

- **Micro data centers intelligents** : dimensionnement en temps réel des ressources (calcul, stockage, réseau) selon la charge.
- **Détection proactive des anomalies complexes** : analyse des signaux faibles dans les infrastructures hétérogènes pour identifier les causes racines d'incidents.
- **Cybersécurité** : analyse des flux pour l'assistance intelligente à la détection de variations de comportements



SPIE appelle les acteurs présents à faire un don de leurs données pour la recherche, dans un cadre sécurisé et conventionné. Ces données permettront d'entraîner des modèles plus pertinents, dans un esprit de co-construction. Le message est clair : l'IA ne doit pas être développée de manière isolée ou centralisée, mais dans une logique collaborative, raisonnée, souveraine. C'est aussi un modèle français d'innovation.

### POURQUOI UNE IA RESPONSABLE ?

Parce que l'IA actuelle consomme massivement énergie et données, tout en renforçant notre dépendance technologique. Adoptons une IA en circuit court : locale, sobre et maîtrisée.

## GenAI : innovez dans l'art de questionner vos assistants virtuels

Chez SPIE, les assistants et les agents IA deviennent des outils du quotidien, en particulier dans les équipes techniques et métiers. Leur objectif est d'apporter des réponses précises, contextualisées, sécurisées, en s'appuyant sur des modèles génératifs maîtrisés. Mais pour en tirer toute la valeur, encore faut-il bien savoir... "prompter" !



### LES PRINCIPES DU PROMPT

- **Clarté** : éviter les formulations floues ou trop générales
- **Spécificité** : poser une question ciblée, avec une attente claire
- **Contexte** : indiquer le secteur, le rôle ou la finalité métier
- **Instructions** : demander un format (ex. : tableau, plan d'action)
- **Itération** : tester, affiner, reformuler au besoin

SPIE a développé HubIA, une plateforme interne sécurisée, coupée du web, avec plusieurs assistants spécialisés :

#### Assistant généraliste SPIE

- Basé sur un LLM allégé, pour une IA utile, rapide et économe
- Utilisé pour produire des schémas, plans d'action, résumés

#### Assistant QHSSE

- Analyse automatique d'images (ex : photo d'un local technique)
- Génère un tableau de risques, des mesures correctives, exportables en CSV

#### Assistant RH

- Connecté à SharePoint via RAG
- Répond à des questions internes (procédures, congés...) avec 240 lignes de pré-prompting pour ajuster ton, langage, et cadre légal

SPIE démontre que l'IA générative peut être déployée de façon souveraine, sobre et pertinente. La vraie valeur ne vient pas du modèle lui-même mais de l'ingénierie humaine qui l'entoure. Un bon assistant commence toujours par un bon prompt et sur des cas d'usage à forte finalité.

## Automatiser le support utilisateur avec un chatbot interne (rex du Groupe Martin Belaysoud)

Le Chief Data Officer du groupe Martin Belaysoud est venu partager les coulisses de la mise en place d'un assistant conversationnel IA interne déployé auprès de 3 000 collaborateurs, en réponse à une surcharge des canaux traditionnels (mail, téléphone, Teams...). Tout a commencé lors d'une rencontre avec Wikit lors de l'édition 2024 de l'événement Génération IA organisé par l'ADIRA. En juillet, un atelier multi-services a été organisé en interne pour identifier les tâches répétitives à automatiser. Le choix s'est porté sur Wikit, éditeur de la plateforme Wikit Semantics, déjà éprouvée par de nombreuses entreprises dans le secteur public et privé.

### MISE EN ŒUVRE STRUCTURÉE

L'équipe a lancé un appel d'offres formalisé via un cahier des charges, avec des critères de choix clairs :

- Simplicité d'administration, sans technologie complexe ou de "boîte noire",
- Interopérabilité avec l'intranet en refonte et les outils du système d'information,
- Déploiement rapide pour coller au calendrier interne,
- Modèle économique flexible,
- Sécurité et souveraineté des données, hébergement en France (data center BRH),
- Possibilité d'intégration avec l'ITSM pour escalader les demandes bloquantes.



### FONCTIONNEMENT DE LA SOLUTION

L'assistant, baptisé Martin est intégré à l'intranet via un composant web. Il fonctionne selon une architecture en trois temps :

- Saisie de la demande par l'utilisateur,
- Traitement via un moteur sémantique vectoriel alimenté par les bases internes (50 bibliothèques SharePoint indexées),
- Génération de réponse avec sources explicites, ou escalade vers l'humain si besoin (via un formulaire, bientôt un ticket connecté ITSM).

Les contenus sont prétraités, fragmentés, vectorisés et reliés à des gardes-fous pour éviter les dérapages (langage agressif, absence de sources...).

### DÉPLOIEMENT EXPRESS ET STRUCTURÉ

Le projet a été lancé le 28 janvier et mis en production le 18 mars, soit moins de deux mois pour :

- Sélectionner la solution, indexer les contenus, intégrer l'assistant et effectuer les tests et le paramétrage.

L'adoption de l'outil a été rapide, portée par une intégration fluide. Dès les premières semaines, le taux de bonnes réponses atteignait déjà 60 à 70 %, grâce à une exploitation intelligente des contenus intranet fraîchement mis à jour. Un système simple de retours utilisateurs (👍👎) permet un suivi qualitatif en continu.

## Agents IA : mythe ou réalité ?

Structurellement, un agent se compose d'un **cerveau** (raisonnement, planification), d'une **mémoire** (historique des actions, apprentissage) et d'**outils** (accès à des bases, API, commandes système). Les agents IA sont une évolution naturelle du LLM, capables de passer à l'action. Ils ouvrent la voie à des applications plus intelligentes, plus autonomes et plus utiles au quotidien, en entreprise comme pour le grand public. Pour en tirer le meilleur parti, il faut maîtriser leur conception et leur pilotage, choisir des cas d'usage pertinents et accepter une phase d'expérimentation.

### L'IA ENTRE DANS UNE NOUVELLE ÈRE

Dans un système avancé, plusieurs agents IA spécialisés peuvent coopérer de manière fluide et autonome. Par exemple, l'un recherche une solution adaptée, un autre exécute ou simule une action, tandis qu'un troisième ouvre un ticket ou escalade le cas si nécessaire. Chaque agent remplit un rôle précis et leur interaction en temps réel permet une réponse plus riche, rapide et adaptée au contexte métier.

### LES DÉFIS TECHNIQUES À RELEVÉR

Cette approche, prometteuse reste jeune et soulève plusieurs enjeux.

- **La mise à l'échelle** car plus le nombre d'outils à intégrer augmente, plus la coordination devient complexe.
- **Le temps de réponse** peut aussi s'allonger, notamment si les agents doivent interagir avec des systèmes métiers hétérogènes.
- **L'interopérabilité** reste un point sensible car tous les outils ne sont pas conçus pour dialoguer entre eux.

De nouveaux protocoles émergent, à l'image du Model Context Protocol (MCP). Ce standard en construction vise à structurer les échanges entre agents IA, un peu comme le protocole HTTP l'a fait pour le web, en facilitant l'orchestration, la gestion du contexte, et la supervision des interactions multi-agents.



### UN SYSTÈME DOTÉ DE PLUSIEURS CAPACITÉS FONDAMENTALES :

- Autonomie dans l'exécution de tâches.
- Proactivité (capacité à agir sur son environnement).
- Réactivité (capacité à s'adapter à l'environnement).
- Rationalité, mémoire et collaboration avec d'autres agents.

## Comment intégrer l'IA dans vos outils métiers ? Rex de la société ISILOG

Intégrer l'IA là où elle est invisible mais utile. Plutôt que de former les collaborateurs à utiliser des outils IA externes, Wikit défend une approche plus fluide : intégrer l'intelligence artificielle directement dans les interfaces métier, là où les utilisateurs travaillent déjà. Cela permet un usage naturel de l'IA, sans rupture d'expérience.

**Le cas d'Isilog:** Isilog propose une plateforme de ticketing et de gestion des services ESM (Enterprise Service Management) utilisée par plus de 800 clients dans la francophonie. Face à la complexité croissante des demandes clients, à l'exigence de réactivité et à la pression budgétaire, l'enjeu était clair : renforcer l'efficacité du support sans alourdir les équipes.

Wikit a accompagné Isilog pour intégrer des briques IA dans leur outil. Deux types d'apps ont été créés et déployés :

### DES APPS DE GÉNÉRATION

Le technicien gagne ainsi du temps sur ses tâches rédactionnelles : génération de réponses automatiques aux tickets, reformulation de messages, rédaction de FAQ...

### DES APPS D'AUTOMATISATION

Les fonctionnalités permettent de réduire le temps de traitement et d'augmenter la qualité des réponses: classification automatique des tickets, suggestion de résolution, détection de cas similaires ou encore résumé d'un dossier complexe.

### UN FOCUS SUR LA SOUVERAINETÉ ET LA SÉCURITÉ

La plateforme Wikit est hébergée en France (OVH), avec des options pour connecter des modèles open source ou propriétaires selon le cas d'usage. Les données exploitées (tickets, FAQ, documents techniques..) sont strictement maîtrisées. Elles ne sortent jamais de l'environnement client et sont utilisées via des pipelines RAG (Retrieval-Augmented Generation), assurant confidentialité et mise à jour en temps réel.