[sf≡ir]

WENVISION

L'IA générative pour les CIO et les CTO Release 11

TECHWAVES - #GENAI FOR CIO & CTO

- TechWaves #GenAl: 01. des LLM aux systèmes agentiques
- Notre vision de la plateforme #GenAl O2. d'entreprise
- La #GenAl en entreprise : 03. bonnes pratiques et retours d'expérience

Avant-propos : "Bienvenue dans l'ère cognitive !"



Didier GIRARD DG de SFEIR & WENVISION didier@wenvision.com

Préparer son SI pour 2030

En quelques décennies, sommes passés de l'ère mécanique (les cartes perforées) à l'ère visuelle, constituée d'interfaces graphiques homme-machine et de formulaires à remplir. Il faudra beaucoup moins de temps que cela pour que nous passions à l'ère cognitive. Une ère centrée sur le raisonnement et les interactions conversationnelles.

Le "chat" vient évidemment à l'esprit, mais lui-même ne sera quère suffisant pour assurer la montée en puissance de l'IA au sein du SI. Nous dialoguons aujourd'hui avec des assistants demain, nous exprimerons une intention à des agents. Ces agents vont interagir avec système d'information, et potentiellement interagir avec d'autres agents.

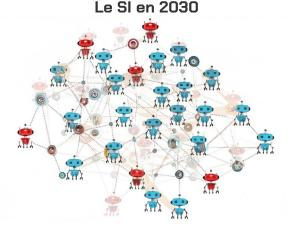
Autrement dit, pour envisager 2030, il va falloir penser moins en termes d'IHM et davantage en termes d'API.

L'IA comme système nerveux du SI

Du point de vue technologique, la feuille de route des CIO doit comporter a minima ces 3 grands volets:

- Des données accessibles, bien gouvernées : un accès différencié à des produits data de qualité est fondamental pour créer de la valeur avec les produits GenAl.
- Un SI jumeau numérique de **l'entreprise** : le système d'information doit offrir une réplique fidèle et interrogeable de l'entreprise, pour faciliter les interactions et le dynamisme.
- Une plateforme IA système nerveux de l'entreprise : la plateforme doit accueillir, exécuter, gérer, surveiller les personas IA qui vont émerger et se multiplier. La puissance de ce système nerveux aura un impact direct sur la création de valeur.







Les agents viennent compléter les assistants



Mylène MAIGNANT Consultante Data et IA mvlene@wenvision.com

Les agents IA marquent une étape dans l'évolution des technologies d'assistance.

Ils dépassent le cadre de l'interaction conversationnelle pour s'inscrire dans une approche agentique, où l'IA agit de manière autonome. De plus en présents dans métiers environnements et techniques. ils permettent d'automatiser des processus complexes de bout en bout notamment dans le développement logiciel — et ouvrent de nouvelles perspectives dans des domaines tels que le support client, l'analytique ou la gestion documentaire.

Leur montée en puissance dessine nouvelle phase transformation numérique, fondée non plus sur la seule assistance, mais sur la délégation intelligente de tâches complexes à des systèmes capables d'apprendre, de s'adapter et d'interagir avec leur environnement.

Dimension	Assistant	Agent
Nature	Modèle probabiliste statique	Système autonome interactif
Exemple	ChatGPT, Microsoft 365 Copilot, SFEIR Raise	Claude Code, OpenAl Codex
Finalité	Générer du texte	Atteindre un objectif
Mémoire	Se limite au contexte de la conversation	Mémoire persistante, adaptation de leur comportement dans la durée
Autonomie décisionnelle	Absente	Présente : l'agent planifie ses actions
Interaction	Dialogue linéaire	Boucles perception-action
Architecture	Monobloc (prompt → réponse)	Modulaire (environnement, outils, prompts, planification, mémoire)

Assistants vs agents : des usages différenciés

Les assistants assistent

Les assistants sont des systèmes interactifs conçus pour aider un utilisateur à accomplir des tâches variées à travers le langage. Ils aioutent aux LLM un cadre conversationnel structuré et peuvent exploiter des données contextuelles pour formuler des réponses.

Les agents agissent

Les agents sont des systèmes autonomes conçus pour percevoir leur environnement, prendre des décisions et agir pour atteindre des objectifs spécifiques. Fondés sur des LLM pour assurer la partie dialogue. les agents intègrent des capacités supplémentaires telles que la mémoire, la planification, l'exécution d'actions et l'interaction avec des outils externes.

Exemple de prompt : « Organise ma journée de demain. »

L'assistant propose une organisation générique, sans connaître les contraintes personnelles de l'utilisateur si elles ne lui ont pas été précisées.

L'agent consulte le calendrier de l'utilisateur, identifie les rendez-vous et envoie des rappels tout en s'adaptant aux modifications éventuelles.









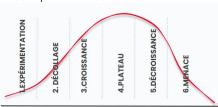
La méthodologie de nos Techwaves

Les Techwaves sont le résultat d'une méthodologie éprouvée chez SFEIR et WEnvision, pour sélectionner les technologies les plus pertinentes au sein d'un domaine. Celles qui permettront de faire émerger les ambitions des entreprises sans accroître leur dette technique.

Nous évaluons divers critères. comme la facilité de prise en main, la constitution d'une communauté, la disponibilité de modules de formation voire de livres, etc.

Un exercice possible grâce au crowdsourcing interne au groupe, qui nourrit aussi bien notre veille technologique que nos analyses.

Le résultat de ces analyses permet de positionner les technologies sur notre courbe. découpée en 6 secteurs



EXPÉRIMENTATION

Concerne les développeurs et entreprises innovantes. start-up, qui cherchent des moyens d'obtenir un avantage décisif grâce à la technologie. Ils testent beaucoup, réalisent des "proofs of concept".

2.

DÉCOLLAGE

La technologie est adoptée par les start-up et DSI qui sont prêts à miser beaucoup dessus afin d'obtenir un avantage différenciant. Pour eux, il est vital que le projet aboutisse.

3.

CROISSANCE

Les entreprises un peu plus prudentes. mais qui souhaitent tout de même pouvoir s'appuver rapidement sur des technologies innovantes, adoptent à leur tour la technologie.

4.

PLATEAU

La technologie a trouvé son marché. Les plus conservateurs, qui souhaitent partir sur des technologies éprouvées. l'ont adoptée. Le marché est capable de répondre à la demande.

DÉCROISSANCE

De moins en moins de nouveaux projets adoptent cette technologie. L'usage est en décroissance. L'utilisation de cette technologie peut être assimilée à de la dette.

6.

MENACE

L'utilisation de cette technologie sur un nouveau projet est à proscrire. Peu d'acteurs sont encore impliqués dessus.



TechWave : les assistants et agents spécialisés dans le code



Définition

QI

Les assistants et agents IA spécialisés dans le développement sont des outils fondés sur des LLM, entraînés pour le développement. Ils sont conçus pour automatiser des tâches de programmation. Tandis que les assistants se contentent de proposer des suggestions, généralement au sein de l'environnement de développement, les agents peuvent interagir avec des fichiers, des API ou des dépôts Git pour exécuter des actions complexes. Cette évolution s'inscrit dans une approche agentique, où l'IA devient un partenaire autonome capable de collaborer sur l'ensemble du cycle de développement.

Cas d'usage



Les assistants et agents spécialisés dans le code permettent d'automatiser un ensemble de tâches techniques, telles que la complétion de code, la génération de tests, le refactoring de code existant, la rédaction de documentation ou la gestion de pull requests. Les plus avancés peuvent prendre en charge l'intégralité du développement à partir de spécifications écrites ou d'écrans dessinés.

GitHub Copilot

CURSOR

W Windsurf

Croissance

Plateau

Décroissance

OpenAI Codex

JetBrains Al

Décollage

Notre œil d'expert



Si les agents IA spécialisés dans le développement présentent un fort potentiel, leur déploiement en entreprise présente encore - à l'heure actuelle - plusieurs limites.

La première concerne la maîtrise des coûts. Contrairement à GitHub Copilot, qui repose sur un forfait avec quota intégré, des outils comme Windsurf ou Cursor fonctionnent par appels API successifs. Chaque tâche peut mobiliser plusieurs itérations, ce qui complique l'estimation du coût final

La deuxième limite réside dans le manque de recul et l'absence de benchmarks fiables. Peu d'évaluations indépendantes permettent aujourd'hui de mesurer avec précision la performance ou la robustesse de ces agents en contexte réel.

La problématique majeure concerne la qualité de ce qui est produit et notre capacité à en juger et l'améliorer. Les spécifications doivent être claires (possiblement optimisées grâce à un assistant IA) et le développeur doit être en capacité de donner des directives d'architecture et de valider le code généré.

Focus sur...



Accessible au travers d'un navigateur et alimenté par le modèle codex-1 (basé sur o3), il peut développer des fonctionnalités complètes, corriger des bugs, naviguer dans des bases de code et créer des pull requests de manière autonome. Codex ne remplace pas les développeurs, il les libère des tâches répétitives.

base de code et gère

de développement. En

effet, il est capable de

s'intégrer avec GitHub,

pour gérer l'ensemble

au développement.



S'intégrant directement dans le terminal et sans configuration complexe, cette solution comprend automatiquement votre l'ensemble du workflow GitLab et tous vos outils en ligne de commande des actions nécessaires

L'extension VSCode propose trois modes distincts: Ask pour les questions techniques, Edit pour les modifications ciblées et Agent pour les tâches complexes autonomes. Le mode Agent va bien au-delà de la génération de code en analysant l'ensemble du proiet et en planifiant les étapes de développement.



Expérimentation



Menace

TechWave: les frameworks pour construire des agents



Définition

Q

Les frameworks de création d'agents sont des outils (SDK, plateformes, librairies) permettant de **concevoir, orchestrer et superviser des agents autonomes ou multi-agents**. Un agent ici est une entité IA capable d'observer un contexte, de prendre des décisions, d'appeler des outils ou d'autres agents, et de mémoriser des expériences passées pour accomplir une tâche. Ces frameworks offrent des briques comme :

Mémoire (short-term, long-term)

OpenAl Agents SDK

Décollage

Croissance

Agent Development Kit

- Reasoning/Decision making (chaînes d'action, graphes, rôles)
- Orchestration d'outils (API, bases de données, fonctions métiers)
- Communication multi-agents (collaboration entre plusieurs IA)
- Gouvernance et monitoring (contrôle des actions prises)

Cas d'usage



Les frameworks de création d'agents permettent par exemple d'automatiser la création d'ensemble d'actions et de tâches complexes comme la recherche d'informations, l'exécution de workflows métiers, l'assistance client, la coordination d'équipes virtuelles ou l'enrichissement dynamique de bases de connaissances.

mann

Agent Builder

Plateau

Swarm[®]

Décroissance

haystack

Notre œil d'expert



Les frameworks d'agents IA offrent une flexibilité impressionnante, potentiellement accessible à des collaborateurs non IT, grâce à des environnements no code. Par rapport à des workflows classiques, leur potentiel réside dans leur flexibilité, leur capacité à orchestrer des agents spécialisés et à s'intégrer dans des environnements métiers variés. Toutefois, leur mise en œuvre nécessite une attention particulière.

Typiquement, enchaîner plusieurs décisions autonomes multiplie les risques d'erreurs invisibles sans monitoring renforcé. Cette imprévisibilité complique l'adoption à grande échelle. Un deuxième enjeu concerne la maîtrise des coûts. Chaque agent peut mobiliser de nombreux appels API et outils externes, rendant difficile l'anticipation des dépenses, surtout dans des architectures multi-agents.

Enfin, les questions de sécurité et de confidentialité sont prégnantes : l'entreprise doit avoir une certaine maturité en matière de sécurité des données et de contrôle d'accès.

Focus sur...



CrewAl est un framework permettant d'orchestrer des équipes d'agents IA, chacun avec un rôle et outils dédiés. Il facilite la collaboration agents entre pour automatiser des tâches complexes, comme la génération de contenu ou l'exécution workflows. Simple, flexible et adapté environnements aux multi-agents.



Smolagents est un framework de Hugging Face pour créer rapidement des agents IA autonomes en Python. Minimaliste et flexible, il permet aux agents de générer et d'exécuter du code, avec compatibilité LLM et outils externes. Idéal pour le prototypage rapide et les projets légers.



OpenAl Agents est un SDK qui permet de créer agents capables d'utiliser des outils, de aérer des tâches de complexes et de façon raisonner autonome, directement via l'API d'OpenAI. Il s'intègre facilement à l' écosystème ChatGPT, avec mémoire, fonctions encore actions personnalisées.



smolagents

Expérimentation

Menace



Trois options pour interagir avec l'IA générative, combinables selon les besoins

Trois manières d'interagir avec l'IA générative

A - API:

Intégration directe dans les outils métiers (CRM, ERP) via des API, offrant flexibilité et personnalisation.

B - Intégration :

Incorporation dans les outils bureautiques et de développement (Office, IDE), facilitant l'adoption sans changer les habitudes.

C - Interface de chat :

Interaction en langage naturel via un chat, permettant aux utilisateurs d'obtenir rapidement des réponses sans compétences techniques.

API:

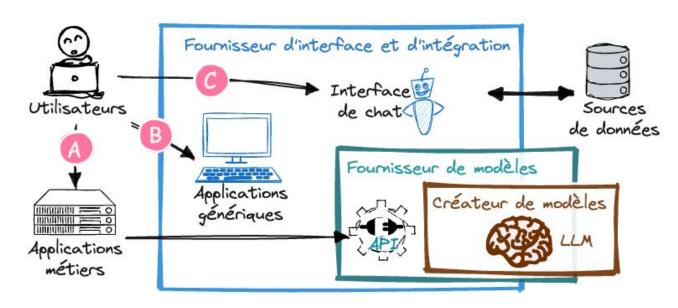
Au travers d'API pour une intégration dans les outils métiers

Intégration :

directement intégré dans nos outils habituels

Chat:

via une interface de chat et des interactions en langage naturel



Un nouveau rôle pour les développeurs



Aurélien PELLETIER Consulting director -**Platforms** aurelien@wenvision.com

Devenez des managers d'agents IA

Si vous avez déjà expérimenté Cursor, Windsurf ou Claude Code, vous le savez, une transformation radicale du métier de développeur est en marche. L'écriture de code, de plus en plus confiée à l'IA générative, va se réduire, laissant place à un rôle de quide et de garant de la qualité.

Avec la GenAl, le développeur de demain sera

- Un relecteur expert : son œil critique sera indispensable pour analyser le code généré par l'IA, identifier les erreurs, les vulnérabilités potentielles et les axes d'optimisation.
- Un chef d'orchestre : il devra maîtriser l'art de formuler des instructions précises détaillées ("prompts") pour orienter l'IA vers l'atteinte des objectifs.
- Un garant de la qualité : il définira les standards de code, les bonnes pratiques à respecter et s'assurera de leur intégration par l'IA.

Apprenez à collaborer dès maintenant avec l'IA

Cette évolution rappelle celle des créateurs de projets open source. Initialement seuls à coder, ils deviennent, avec l'arrivée de contributeurs, les gardiens du projet, validant les "pull requests" et passant plus de temps à relire, commenter et quider qu'à coder eux-mêmes.

Nous allons tous devenir des "mainteneurs" de nos projets, assurant la cohérence et la qualité du code produit par l'IA.

Qu'est-ce que l'arrivée de l'IA générative implique pour vous, les développeurs?

- Adaptabilité: la familiarisation et la maîtrise des outils d'IA sont devenues indispensables.
- Montée en compétences : la connaissance approfondie des architectures logicielles, des bonnes pratiques de développement et de la sécurité est plus cruciale que jamais.
- Communication : la capacité à formuler des instructions claires et précises pour l'IA est une compétence clé.
- Une nouvelle forme de créativité : l'IA est un outil puissant, et c'est à nous de l'utiliser pour explorer de nouvelles frontières et repousser les limites du possible.

À lire aussi sur ce sujet, notre tribune sur le "vibe coding" dans CIO-Online:

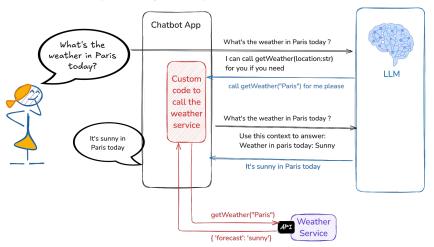
https://www.cio-online.com/actualites/lire-le -vibe-coding-bouleverse-la-notion-de-produ ctivite-des-developpements-16298.html



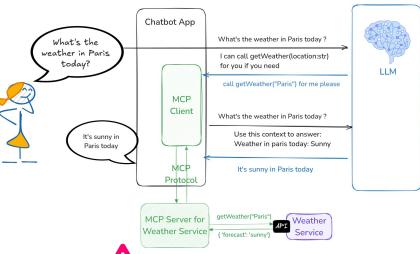
MAJ. A

Des protocoles type MCP pour des interactions inter-IA

Sans MCP: Chaque utilisateur d'un service écrit son code spécifique pour intégrer chaque service avec un LLM



Avec MCP: Le fournisseur du service écrit une fois le code pour intégrer son service à un LLM



Faire interagir un LLM avec un autre service demande beaucoup d'efforts

Faire interagir un LLM avec le SI d'une entreprise ou une application externe nécessite de passer par une API. Deux mauvaises solutions s'offrent alors :

- demander au LLM de lire la spécification et de construire la requête à la volée : manque de fiabilité, insoutenable à l'échelle;
- écrire un code spécifique pour chaque API qu'il faudra interroger : demande du temps, manque de fiabilité, difficile à maintenir, restreint l'IA.

MCP permet de mutualiser ces efforts

Le principe du "Model Context Protocol" consiste à publier des serveurs MCP correspondant chacun à une API spécifique, que des clients MCP pourront appeler, sans avoir besoin de connaître les spécifications de l'API.

Introduit en fin d'année dernière par Anthropic, MCP a engrangé de nombreux soutiens, dont ceux de Microsoft et OpenAl, tandis que plus d'un millier de serveurs MCP sont déjà disponibles.

A

Un protocole jeune, à la sécurité incertaine

À date, MCP priorise l'intégration facile au détriment de la sécurité, sans standard d'authentification, ni chiffrement de contexte, ni vérification d'intégrité des outils. Ces lacunes exposent les entreprises à des risques significatifs de fuites de données et d'intrusions.

Nul doute que ces problèmes seront adressés dans les prochains mois. Inutile d'attendre pour commencer à expérimenter!



MAJ. All

GenAI: une gouvernance spécifique s'impose



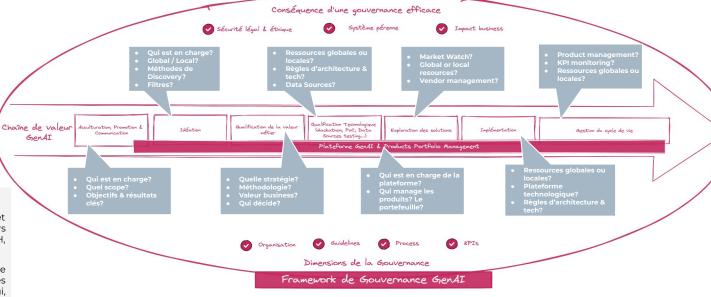
Consulting Director Strategy
olivier@wenvision.com

Une gouvernance à créer

L'IA générative est un sujet poussé par tous les acteurs de l'entreprise : IT, RH, Comex, salariés...

Comment s'assurer que toutes les initiatives d'assistants aujourd'hui, d'agents demain, servent bien les objectifs stratégiques de l'entreprise, mais aussi qu'elles respectent les règles de sécurité et de conformité ?

La gouvernance GenAl doit s'occuper de ces questions, indépendamment de la gouvernance data.



Une chaîne de valeur à maîtriser

Au-delà de la mise à disposition d'une plateforme (avec interface de chat, analyse de documents, etc.), son cycle de vie et l'intégration des capacités de la plateforme avec des outils métiers doivent être gérés comme des produits, avec des phases d'exploration impliquant largement les utilisateurs. En gardant en tête que pour une roadmap GenAl, il faut penser en termes de domaines à explorer plutôt que de promettre telle fonctionnalité ou tel ROI.

Des principes simples mais forts

La mise en place de la gouvernance GenAl est un moment fort, qui encadre la manière dont l'entreprise va aborder le sujet. Elle doit pouvoir s'incarner dans 3 à 5 principes directeurs, clairs et faciles à retenir, qui permettront de rassurer les collaborateurs et de les embarquer. Ces grands principes immuables sont accompagnés de règles d'implémentation ainsi que d'un ensemble de recommandations pour déployer plus facilement à l'échelle.



Réorienter la stratégie GenAI : priorité à l'adoption avant les cas d'usage



Aude DEFRETIÈRE Consultante Orga & Product aude@wenvision.com

Alors que de nombreuses organisations se lancent dans des démarches IA générative en cherchant à tout prix des cas d'usage "rentables", cette stratégie se révèle souvent contre-productive.

La focalisation prématurée sur le ROI conduit à des projets déconnectés des réalités métiers, conçus en circuit fermé par des profils techniques, sans implication des utilisateurs finaux. Résultat : une faible appropriation, des outils sous-exploités et une perte de confiance dans la technologie.

À rebours de cette approche, **une stratégie d'adoption progressive et massive s'impose** comme levier structurant.

Former, équiper, accompagner les collaborateurs dans l'exploration autonome de la GenAl permet une acculturation réelle et durable. Ce n'est qu'à cette condition que des usages pertinents émergent, portés par ceux qui maîtrisent leurs contraintes opérationnelles.

L'enjeu ne réside pas dans la détection des bons cas d'usage, mais dans la création des conditions d'une transformation organique, portée par les utilisateurs eux-mêmes - au sein du cadre fixé par l'organisation.

Une dynamique d'appropriation collective, distribuée et incrémentale, est seule à même de produire des impacts mesurables à long terme.

"Et si on arrêtait de prendre les plateformes d'IA générative pour des produits comme les autres pour lesquels il faudrait justifier de cas d'usage et de ROI pour accepter d'investir sur le sujet ?

Et si, pour une fois, on assumait plutôt le postulat suivant : l'IA générative est une révolution en tant que telle et pas juste un produit dans lequel on investit.

Alors pour une fois, j'ai envie de dire à tous : oubliez vos KPI de rentabilité et intéressez-vous seulement à cet indicateur : l'adoption !"



Aude accompagne plusieurs clients dans leur démarche d'adoption de l'IA. Sa conviction : chercher à mesurer le ROI de la GenAl, c'est mesurer un tsunami avec une règle de 30 cm!

L'importance d'organiser l'accompagnement au changement



Romain MAZUIR Consultant - Head of Product romain@wenvision.com

L'adoption réussie de l'IA générative dépend étroitement de la capacité de l'entreprise à organiser et piloter efficacement le changement.

Voici trois leviers incontournables pour accompagner avec succès processus:

- Créer une vision stratégique claire cohérente. largement communiquée au sein de toutes les strates hiérarchiques. Cela constitue la première étape indispensable pour assurer l'adhésion de l'organisation. Il est impératif que cette vision précise les bénéfices attendus, les objectifs concrets à atteindre, et explicite comment l'IA Gen s'intègre aux processus existants. Cette clarté rassure, donne du sens aux équipes, et garantit que tous avancent dans même direction une avec conviction.
- Stimuler et encourager les usages. Pour réussir l'intégration technologique, il est essentiel de créer rapidement des cas d'usages concrets. Favoriser la constitution d'une communauté d'early adopters constitue un levier puissant.

Ce groupe pilote permettra de développer et tester des ressources pratiques comme des formations ciblées, des applications métiers spécifiques et des contenus pédagogiques adaptés.



Progressivement, cette communauté initiale deviendra le moteur d'une adoption plus large, diffusant les bonnes pratiques et assurant le relai de communication au sein de l'organisation grâce à une équipe d'animation dédiée.

durablement Ancrer technologie. La conduite changement ne doit pas être perçue uniquement comme une initiative temporaire, mais comme une démarche structurante visant à intégrer durablement l'IA générative dans les processus quotidiens de l'organisation. Il s'agit donc de planifier et de piloter cette intégration sur le long terme, avec un suivi régulier, des ajustements permanents et une attention continue portée aux retours terrain.

En institutionnalisant l'IA dans les méthodes de travail, l'organisation assure ainsi sa pérennité et maximise la valeur tirée de ces nouvelles capacités technologiques.

Parmi les facteurs clés pour garantir cette pérennisation, on retrouve explicite l'inscription de la technologie dans le plan stratégique l'entreprise, ainsi ajustement de l'organisation interne en conséguence. Cela lui conférera le poids décisionnel nécessaire, grâce à une représentation dédiée clairement établie.

Intégrer des agents IA : du test à la transformation globale



Nicolas Suchaud Consultant - Head of AI Transformation nicolas@wenvision.com

"Les agents marquent le début d'une nouvelle ère. Les entreprises qui sauront les intégrer avec méthode, éthique et agilité auront un temps d'avance. Le futur ne sera pas automatisé ; il sera orchestré. Le défi n'est pas uniquement technologique. Il est stratégique, organisationnel, culturel.

Il s'agit pour les organisations de bâtir un nouveau mode de travail où le collaborateur de demain pourra être agentique."



1. Aligner et ancrer dans la stratégie

Le point de départ, c'est l'alignement. Il s'agit de connecter l'IA aux priorités métiers via des cas d'usage concrets, à fort impact culturel. Des ateliers dédiés avec les directions métiers permettent de cartographier les processus clés, d'impliquer les équipes et de poser les bases d'une adoption engagée. En parallèle, il est essentiel d'installer une gouvernance claire, avec des rôles définis, des règles d'usage et une supervision humaine des cas sensibles. Cette riqueur structure la confiance.

2. Repenser l'organisation autour des agents

Les agents IA doivent s'intégrer nativement aux systèmes internes et interagir entre eux. Cela suppose une architecture modulaire, interopérable, ouverte à l'orchestration. Mais le vrai défi est organisationnel : il faut redéfinir les processus et repositionner les rôles humains. L'agent devient un opérateur, et l'humain un superviseur, un garant de sens. Ce nouveau tandem doit être accompagné par un plan de montée en compétence ambitieux, pour permettre à chacun de prendre part activement à la transition.

3. Piloter, ajuster, amplifier

La transformation doit être mesurée. Suivre des KPI ciblés (efficacité, qualité, satisfaction) permet de documenter les apprentissages et d'itérer rapidement. Car le changement ne se décrète pas : il se pilote, se teste, se renforce chaque jour, avec méthode et conviction.

Vos 4 fondations pour démarrer

- La donnée avant tout : l'avantage concurrentiel vient de l'usage que vous avez de vos données.
- L'humain au centre : l'IA utile est celle qui renforce les équipes, pas qui les remplace.
- Un usage vertical ciblé: mieux vaut un cas d'usage maîtrisé qu'un outil générique sans impact.
- Commencer ciblé, élargir ensuite : partez d'un besoin précis, puis déployez progressivement.





[Sf=ir] WENVISION

TECHWAVES - #GENAI FOR CIO & CTO

sfeir.com

wenvision.com