

# Chapitre 3

## L'écosystème IA de Microsoft

### 1. Choisir, c'est renoncer

Voici un aperçu des «quelques» services IA représentant l'écosystème IA de Microsoft :



Comme vous pouvez le voir, un grand nombre de services couvrent de larges domaines de l'intelligence artificielle, du RAG, au machine learning en passant par la vision et la conception d'agents. Il peut donc être très difficile de s'y retrouver dans cette galaxie nébuleuse.

Dans ce chapitre, nous allons détailler ensemble les principaux services Microsoft utilisés par les entreprises.

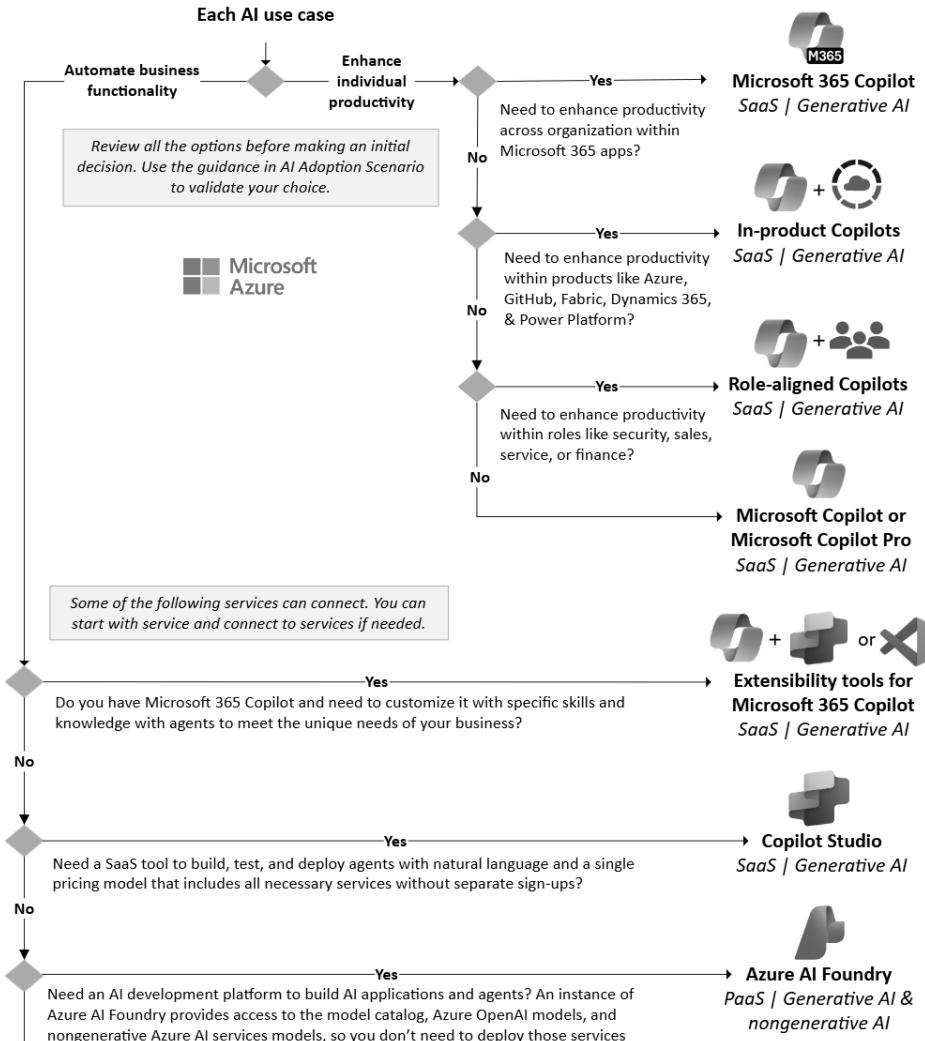
Mais comment s'y retrouver ? Peut-être en partant de votre propre objectif...

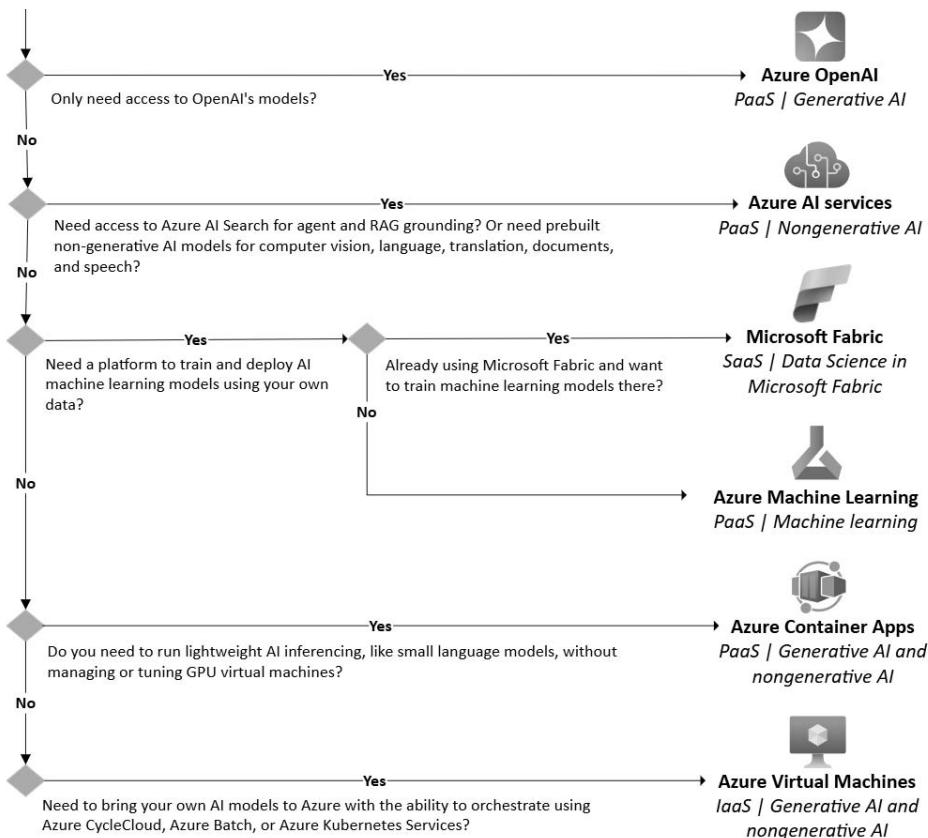
- Cherchez-vous à améliorer votre productivité, celle de vos salariés ? Par exemple en automatisant des tâches répétitives ?
- Avez-vous pris en compte les commentaires de vos clients pour découvrir les cas d'usage qui amélioreront leur satisfaction ?
- Que proposent vos concurrents pour résoudre les problèmes identiques aux vôtres ?

Pour chaque cas d'usage identifié, il convient ensuite de décrire précisément l'objectif, les métriques de réussite et les moyens pour y parvenir (formation, embauche, etc.).

Puis vient la phase du choix (choisir, c'est renoncer...).

L'arbre de décision ci-dessous proposé par Microsoft depuis le lien : <https://learn.microsoft.com/fr-fr/azure/cloud-adoption-framework/scenarios/ai/strategy> est un excellent point de départ :





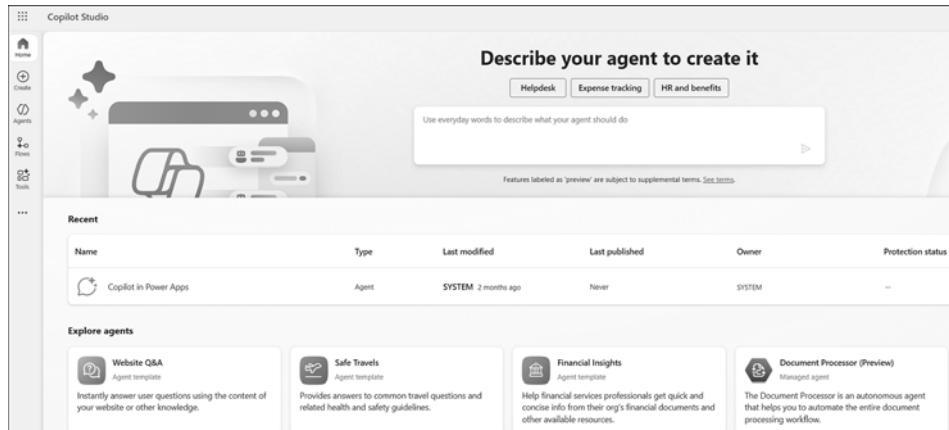
Si l'objectif est d'améliorer votre productivité ou celle de vos salariés, dirigez-vous vers Microsoft 365 Copilot dans le cas où vous êtes déjà fort consommateur des applications de la firme de Redmond, telles que Word, Excel, PowerPoint...

### Remarque

*Copilot est aussi proposé pour des produits spécifiques tels qu'Azure, GitHub, Fabric, Dynamics 365 ou encore Power Platform.*

Des agents Copilot sont aussi désignés pour répondre aux enjeux spécifiques tels que la cybersécurité, la vente, le service ou encore la finance...

Si l'objectif est d'automatiser des fonctionnalités métier précises, utilisez le portail Copilot Studio (<https://copilotstudio.microsoft.com/>) en mode SaaS (*Software as a Service*). L'interface très simple et accessible vous permettra la création et le déploiement simplifié d'agents, en langage naturel ou via une configuration rapide.



### ■ Remarque

*Nous détaillerons Copilot Studio dans le chapitre suivant, Copilot Studio, le copilote simple et intuitif.*

Choisissez Microsoft Foundry (<https://ai.azure.com>) si vous cherchez à intégrer une IA dans votre application métier. La plateforme de développement vous proposera un accès API à un large panel de modèles (OpenAI, Grok, DeepSeek, Mistral, etc.).

The screenshot shows the Microsoft Foundry platform interface. On the left, there is a sidebar with navigation links: Collections, Capabilities, Source, Inference tasks, and Fine-tuning methods. The main area displays a grid of 14 AI models, each with a small icon, name, and description. The models listed are:

- gpt-5.2-codex Responses
- DeepSeek-V3.2 Chat completion
- gpt-5.2-chat Chat completion,Responses
- gpt-5.2
- Cohere-rerank-v4.0-fast Text classification
- Kimi-K2-Thinking Chat completion
- claude-opus-4-5 Messages
- claude-sonnet-4-5 Messages
- gpt-5.1-codex Responses
- DeepSeek-V3.1 Chat completion
- gpt-5-chat Chat completion,Responses
- claude-haiku-4-5 Messages
- claude-opus-4-1 Messages
- grok-4 Chat completion
- embed-v-4-0 Embeddings,Summarization
- gpt-5.1-chat Chat completion,Responses

Si vous avez uniquement besoin d'accéder aux modèles OpenAI (GPT-4o, sora, etc.), déployez une instance Azure OpenAI depuis le portail Azure (<https://portal.azure.com>) puis utilisez un modèle depuis le portail Microsoft Foundry que nous détaillerons dans le chapitre Microsoft Foundry - la fabrique d'IA de l'entreprise.

The screenshot shows the Microsoft Azure portal interface. At the top, it says "Microsoft Azure". Below that, the URL "AI Foundry | Azure OpenAI" and the subtext "The Lord of the Rings". There is a search bar and several buttons: "Générer des modèles Deform pour les ressources sélectionnées", "Générer du code", "Rechercher", "Créer", "Manage deleted resources", "Gérer la vue", "Actualiser", and "Exporter au format".

The left sidebar has a tree view with categories: Accueil > AI Foundry > Use with AI Foundry > Azure OpenAI (which is selected). Other options include All resources, Use with AI Foundry (Azure AI Foundry(preview), AI Hubs), AI Search, More services, and Classic AI services.

The main content area shows a table of Azure OpenAI resources. The columns are: Nom, Genre, and Emplacement. The data is as follows:

Nom	Genre	Emplacement
aoaillivre	OpenAI	East US
chatgpt4ppdemo	OpenAI	Canada East
GPT-4o-GENAI	OpenAI	Sweden Central
openai-demo-pp	OpenAI	East US
speechtextppdemo	OpenAI	West Europe

Si vous avez besoin de modèles non génératifs, utilisez les services Azure AI. Ils vous donneront accès à un panel de fonctionnalités IA supplémentaires, comme la vision, le speech-to-text, le doublage des vidéos, etc., qui seront détaillées dans le chapitre Les IA non génératives de Microsoft.

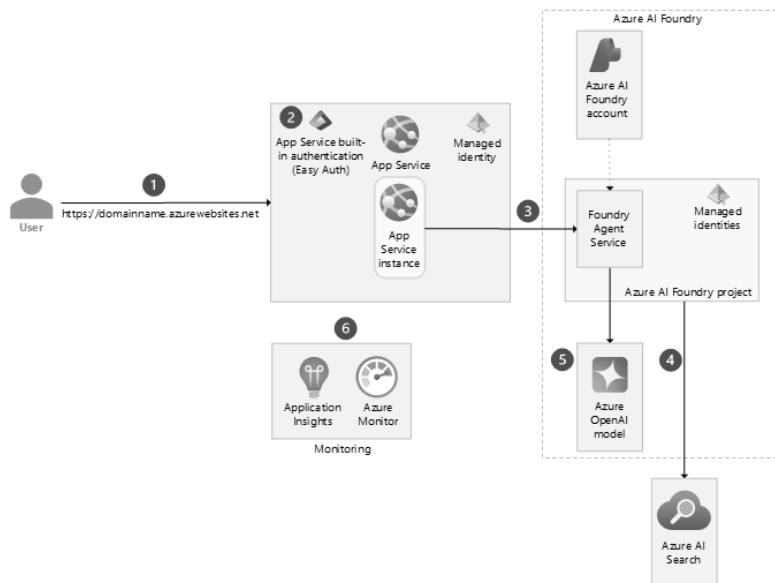
Si vous devez entraîner et déployer des modèles de Machine Learning avec vos propres données, privilégiez Fabric si avez des compétences dans la plateforme d'analytique unifiée en mode SaaS de Microsoft. Sinon, utilisez Azure Machine Learning directement.

Utilisez Azure Container Apps pour l'inférence d'IA légère (SLM) sans gérer l'infrastructure GPU.

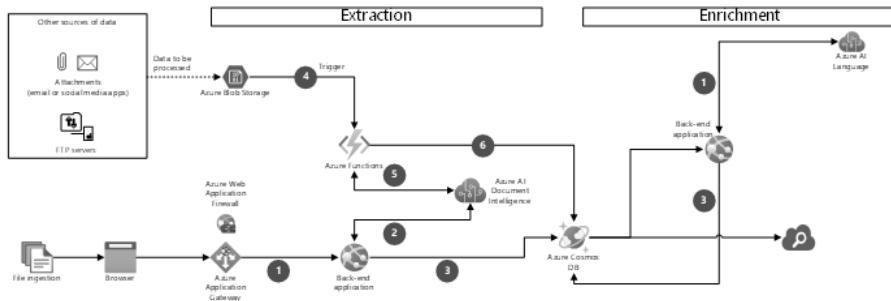
Si vous devez utiliser vos propres modèles et les orchestrer avec Azure Kubernetes, utilisez des machines virtuelles Azure.

## 2. Autopsie d'un projet IA

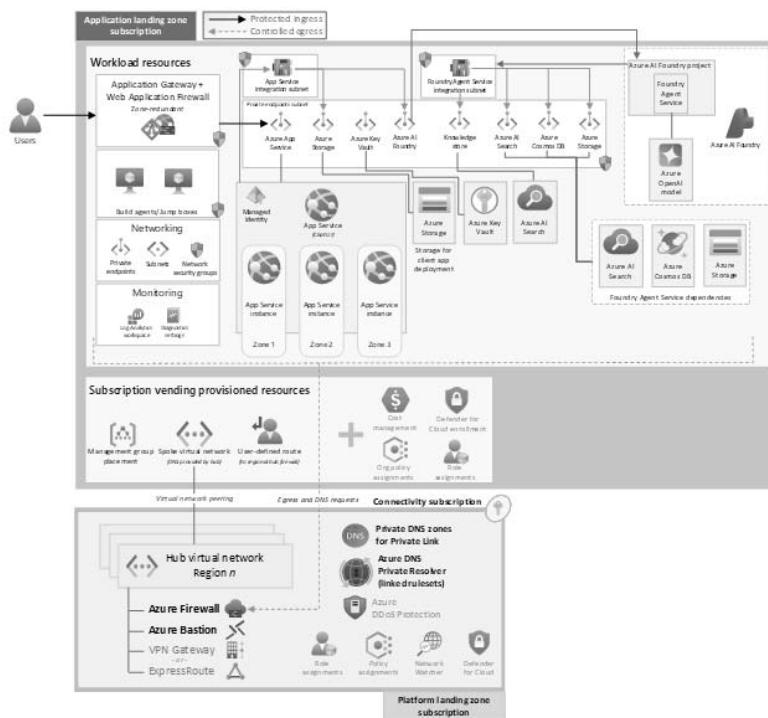
Un projet IA simple (agent qui vous permet d'interroger vos données) ressemble techniquement à cela :



Il peut être un peu plus complexe dès que l'objectif est d'effectuer un traitement d'images et de textes via la multimodalité :



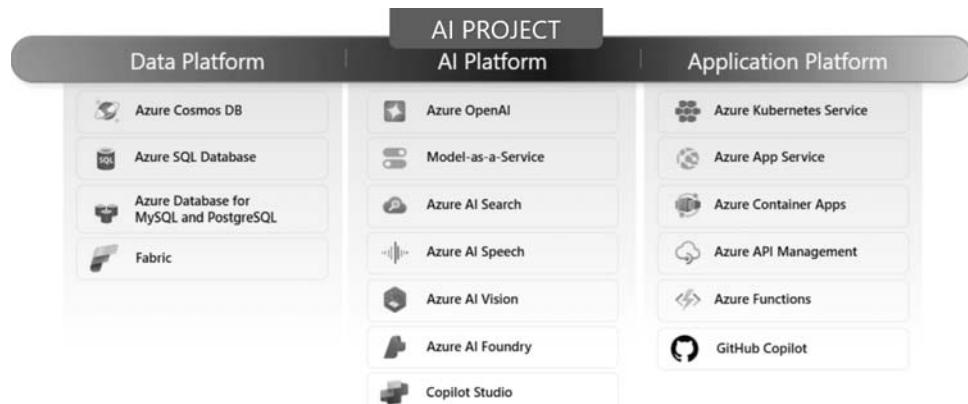
Ce projet IA est souvent intégré plus largement au système d'information de l'entreprise via une zone d'atterrissement (ou *Landing Zone*), dans une topologie de type Hub and spoke, par exemple le schéma ci-dessous :



Ainsi, un projet IA n'est pas simplement représenté par un seul modèle déployé, mais bien par une multitude de services contenus dans trois piliers, chacun ayant une fonction précise : celui représentant la donnée (Data), la couche applicative (Application) et enfin l'IA.

Dans chacun d'entre eux, des services cloud de Microsoft sont déployés.

L'image suivante n'est pas exhaustive mais liste les principales ressources utilisées dans un projet IA déployé depuis le cloud public Azure :



Dans le pilier Data, nous retrouvons souvent la gestion des données depuis une base de données relationnelles ou non, ou encore depuis Fabric.

## 2.1 Microsoft Fabric : la plateforme unifiée de données pour l'IA

Microsoft Fabric est une plateforme de données unifiée conçue pour l'ère de l'IA. Proposée en mode SaaS, elle regroupe toutes les étapes du cycle de vie des données : ingestion, préparation, stockage, analyse, machine learning et visualisation.

Son ambition est de fournir une expérience unique englobant les fonctions auparavant dispersées entre plusieurs services (Azure Data Factory, Power BI, Synapse Analytics, etc.).