

RÉFÉRENTIEL DE CLASSIFICATION

LE BESTIAIRE DE L'IA

Cartographie fonctionnelle et technique du bestiaire
de l'Intelligence Artificielle

SOMMAIRE

7 espèces d'IA à comprendre pour
décider en connaissance de cause



- | | | | |
|----|---|----|---------------------------------|
| 01 | Sortir de l'illusion de l'IA monolithique | 05 | Multimodal et Robotique (VLA) |
| 02 | Les LLM — La mécanique de la probabilité | 06 | Les World Models |
| 03 | IA Génératrice d'Images et Vidéos | 07 | Synthèse Stratégique Européenne |
| 04 | Systemes Sonores et Musicaux | | |

« Comprendre les espèces pour décider en connaissance de cause »

01

CHAPITRE 01

Sortir de l'illusion de l'IA monolithique

L'IA n'est pas une entité unique mais un bestiaire
technologique



L'IA n'est pas un bloc unique



Erreur stratégique majeure : appréhender l'IA comme un bloc technologique unique se paie en CAPEX gaspillés et en échecs opérationnels.



L'IA-washing

Le marketing présente l'IA comme une entité omnisciente, masquant une fragmentation technique profonde.



Un bestiaire technologique

Chaque famille d'IA possède un mécanisme propre et des capacités non interchangeables.



Préalable décisionnel

Comprendre que nous manipulons des outils aux natures divergentes est le préalable à toute décision éclairée.

Le panorama commence par l'espèce la plus bruyante : celle qui traite le langage. Ne pas distinguer les différentes "espèces" d'IA, c'est s'exposer à des échecs opérationnels et à une mauvaise évaluation des risques.

02

CHAPITRE 02

Les LLM

La mécanique de la probabilité textuelle

Les modèles de langage simulent la pensée par le traitement statistique



Les LLM simulent la pensée par le traitement statistique



Probabilité (Tokens)

Le LLM décompose le langage en **"tokens"** (jetons). Il prédit statistiquement le jeton suivant en fonction du contexte précédent, **sans planification globale initiale**.

Chaque mot est une probabilité calculée.



Raisonnement (Conversation)

Ce que nous percevons comme une réflexion est une **orchestration** : le modèle engage une "chaîne de pensée" — une conversation avec lui-même — pour stabiliser sa trajectoire.

La cohérence se construit au fil de l'eau.



Agentivité (Action via Code)

L'agentivité est une **extension fonctionnelle**, pas une nouvelle espèce. En produisant des jetons de code ou des "tags" MCP, le LLM déclenche des outils tiers.

Recherche web, bases de données, action sur le monde réel.

Contrairement aux idées reçues, un LLM ne "réfléchit" pas avant de parler ; il génère sa cohérence au fil de l'eau, token par token.

Le concept de « Délire » vs Hallucination

LE DÉLIRE

Pourquoi « délire » plutôt qu'« hallucination »

Puisque le raisonnement se construit **pendant la pose des tokens**, une erreur initiale force le modèle à poursuivre sa génération de manière **logiquement cohérente** avec cette erreur.

Exemple : si le modèle affirme qu'une tasse en porcelaine est en métal, il maintiendra une structure logique sur cette base erronée.

Il ne voit pas de choses inexistantes ; il maintient une logique sur une base fausse.

L'EFFET MIROIR

Les biais ne sont pas une défaillance

Les biais sont le **reflet exact du corpus d'entraînement humain**. Le modèle reproduit les statistiques de ses données sources par pur mécanisme probabiliste.

95%

des récits de médecins célèbres dans les données sources concernant des hommes — le modèle reproduira cette statistique.

03

CHAPITRE 03

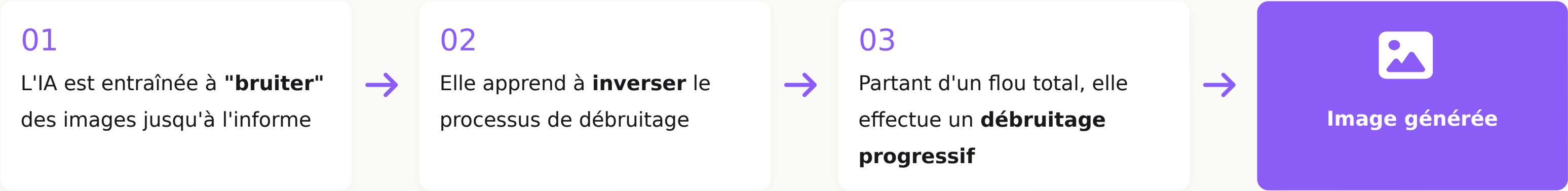
IA Génératrice d'Images et Vidéos

La puissance du débruitage progressif

Du bruit informe à l'image nette

Concept	Le Token (LLM)	Le Pixel débruité (Image Gen)
Mécanisme	Prédiction statistique de la suite d'une séquence textuelle	Transformation d'un bruit informe en image nette guidée par un prompt

Le mécanisme de diffusion



Tagage humain massif : des opérateurs doivent documenter chaque source (ex: "un homme tenant un mug") pour que l'IA sache dans quelle direction "nettoyer" le bruit. La vidéo étend ce principe en calculant les images suivantes pour assurer une cohérence temporelle.

04

CHAPITRE 04

Systèmes Sonores et Musicaux

L'espèce hybride : séquençage et texture



Le son exige séquençage et texture

ESPÈCE HYBRIDE

Le son est une structure complexe exigeant à la fois du **séquençage** (rythme) et de la **texture** (timbre).

C'est une espèce hybride : elle utilise des **tokens audio** (fragments sonores assemblés comme des mots) combinés à des **logiques de diffusion** pour la clarté fréquentielle.

Tokens audio

Fragments sonores assemblés séquentiellement comme des mots dans un LLM



Diffusion fréquentielle

Logique de débruitage pour la clarté et la qualité du signal sonore

RUPTURE MAJEURE

Le Real Time (vocal en temps réel)

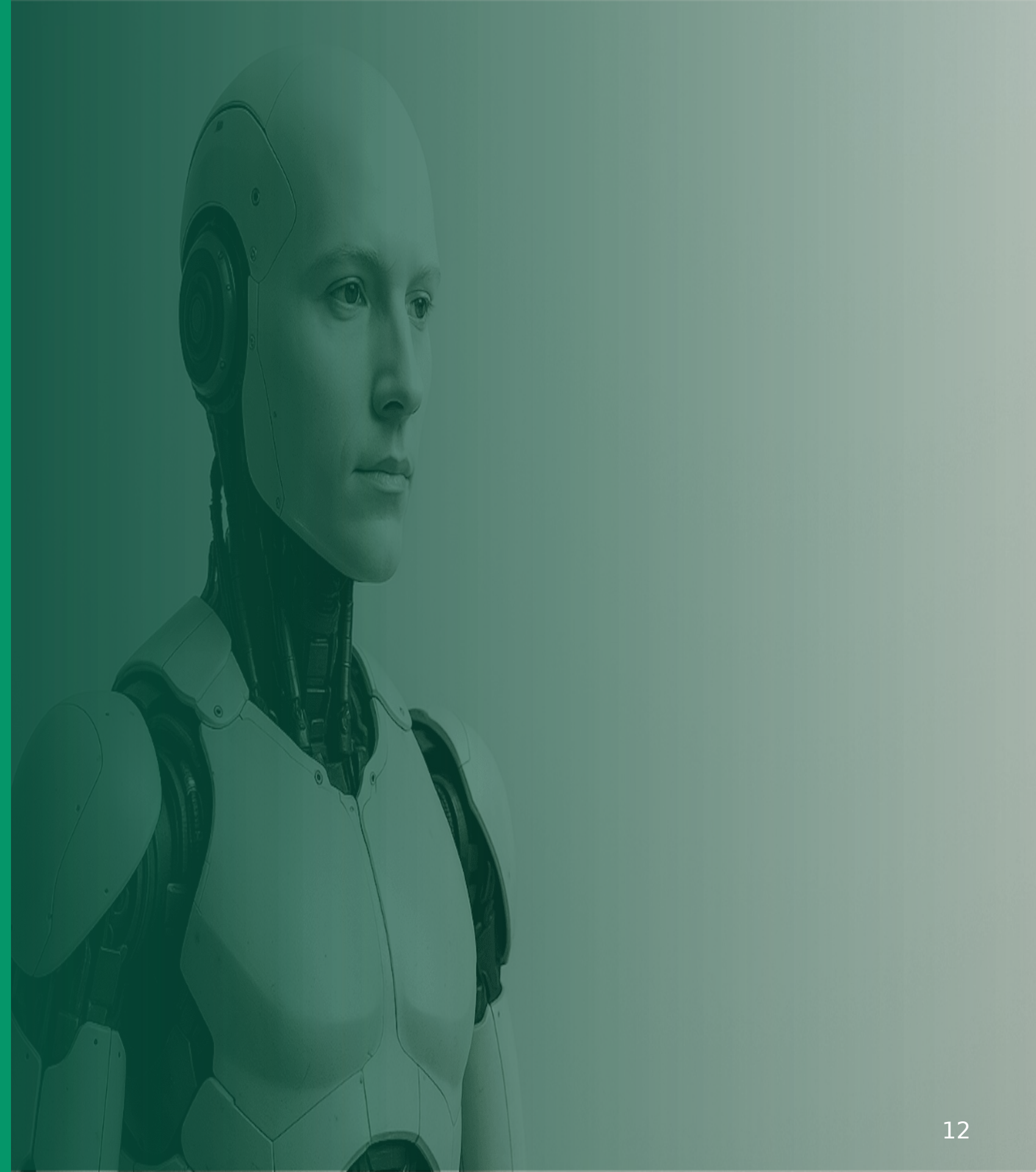
Les tokens sonores transitent **directement** entre l'utilisateur et le modèle. En éliminant la couche intermédiaire de conversion "Text-to-Speech", on réduit la latence et on permet **l'interruption naturelle** de l'IA, rendant l'échange véritablement fluide.

05

CHAPITRE 05

Multimodal et Robotique (VLA)

L'agrégation des compétences



VLA — Vision, Language, Action



Impératif : distinguer l'**interface** (le "wrapper" comme ChatGPT) du **modèle** sous-jacent.

La multimodalité = orchestration

La multimodalité est souvent une **orchestration de modèles spécialisés**. Un LLM comprend votre demande et rédige un prompt optimisé qu'il envoie à un modèle d'image.

Ex: Mistral × Black Forest Lab

MODÈLE VLA



Vision

+



Language (LLM)

+



Action (Gestes)

= **Robotique humanoïde**

POINT DE FRICTION

La documentation des gestes physiques

Contrairement au texte, il n'existe pas de base de données universelle expliquant "**comment**" **saisir un objet**. Cela impose de nouvelles méthodes de capture de données humaines pour entraîner les modèles de gestes.

06

CHAPITRE 06

Les World Models Prédire la réalité physique

Le prochain saut paradigmatique de l'IA



Des vecteurs pour calculer l'évolution physique

Les World Models ne manipulent pas des jetons textuels, mais des **vecteurs** pour calculer l'évolution physique de l'environnement. Si vous lâchez un objet, le modèle prédit l'état suivant du monde en intégrant les lois de la physique.

Deux écoles s'affrontent

ÉCOLE VISUELLE

Fei-Fei Li / World Lab

Utilise la **génération d'images** pour prouver que le modèle comprend :

- La **persistance des objets**
- Les **changements de perspective**

"La preuve est à l'écran"

VS

ÉCOLE MATHÉMATIQUE

Yann LeCun

Un modèle **"invisible"** dont la supériorité ne se verra pas à l'écran, mais par la **performance physique des robots** :

- Rattraper un objet au vol
- Précision **surhumaine**

"La preuve est dans l'action"

07

CHAPITRE 07

Synthèse Stratégique Vers une hybridation européenne

Le Serpent et l'Oiseau : la synergie des espèces



Le Serpent et l'Oiseau



LE SERPENT (LLM)

Rampe dans les données et raisonne pas à pas, token par token.



L'OISEAU (WORLD MODEL)

Survole la réalité physique et prédit l'avenir par les vecteurs.

La stratégie souveraine européenne

Plutôt que de subir le "**monolithe intégré**" imposé par les modèles américains ou chinois, l'Europe peut miser sur **l'interopérabilité de champions spécialisés**. Cette approche "**Best-of-Breed**" est la seule voie pour éviter le verrouillage technologique (vendor lock-in).



Mistral

Le raisonnement textuel et la planification par LLM



Black Forest Lab

La génération d'images et le visuel par diffusion



Yann LeCun

La compréhension du monde physique par World Models

Recommandations pour le décideur

1

Auditez les "Wrappers"

Ne payez pas pour une interface unique sans savoir si elle cache **trois modèles différents**. L'agrégation de modèles (multimodalité) multiplie les coûts et les surfaces de risques de **fuites de données**.

2

Choisissez l'espèce selon l'usage

Ne demandez pas de **raisonnement logique** à un modèle de diffusion, et ne demandez pas de **précision physique** à un LLM sans extension agentique ou World Model. Chaque espèce a son domaine de prédilection.

3

Misez sur la synergie souveraine

Privilégiez les solutions permettant l'assemblage de briques spécialisées (Mistral + partenaires) pour garantir une **flexibilité stratégique** face aux solutions intégrées opaques.

LE BESTIAIRE DE L'IA

Comprendre les espèces pour décider en connaissance de cause

7 espèces · 7 mécanismes · 1 stratégie souveraine



Référentiel de classification de l'IA