

Outils IA & Management

SÉANCE 3

Données et création de valeur

La donnée au cœur de la valeur — le véritable carburant de l'IA.

Gouvernance et avantage compétitif.

Deux questions pour commencer

01

Qui a utilisé ChatGPT — ou un outil d'IA — au moins une fois cette semaine ?

02

Savez-vous sur quelles données ChatGPT a été entraîné ?

Deux façons d'avoir des données

DONNÉE SUBIE

Sous-produit de l'activité

Elle existe, mais personne n'en fait rien.

Dort dans un serveur.

Formats inconsistants.

Silotée par département.

→ **Aucun avantage compétitif.**

DONNÉE STRATÉGIQUE

Ressource intentionnelle

Décidée, collectée, activée.

Quoi collecter.

Pourquoi.

Dans quel but précis.

→ **Avantage compétitif durable.**

→ **Une décision managériale, pas technique.**

Quand quelque chose devient une commodité, l'avantage est dans ce qui reste rare.



Algorithme

COMMODITÉ

OpenAI, Mistral, Anthropic...
API à quelques centimes par requête.
Pas besoin de le construire.



Puissance de calcul

COMMODITÉ

AWS, Azure, Google Cloud.
GPU à la demande, sans
investissement.



Donnée

DIFFÉRENCIATEUR

Unique. Propriétaire.
Impossible à acheter.
Seul vrai avantage compétitif.

Garbage in, garbage out.

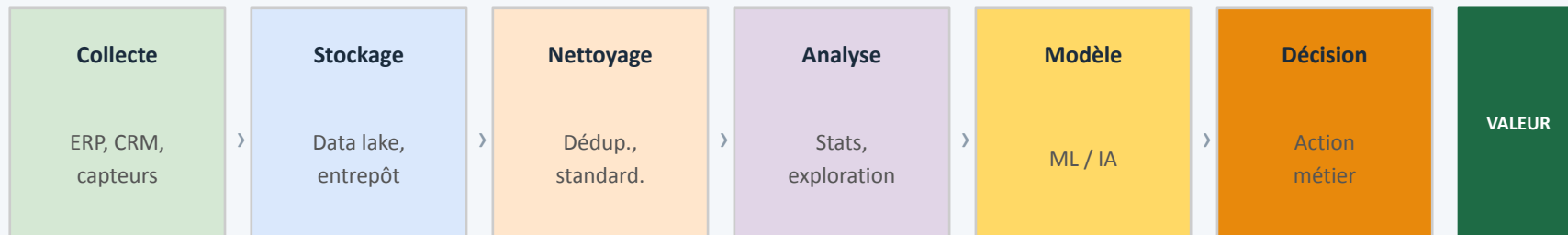
Un bon modèle entraîné sur de mauvaises données produit de mauvaises décisions — à grande échelle, automatiquement, sur des milliers de cas par jour.

Ce n'est pas un problème d'algorithme. C'est un problème de donnée.

Exemple concret

Scoring crédit entraîné sur les biais d'un conseiller des années 90 → discrimination automatisée à grande échelle.

La chaîne de valeur de la donnée

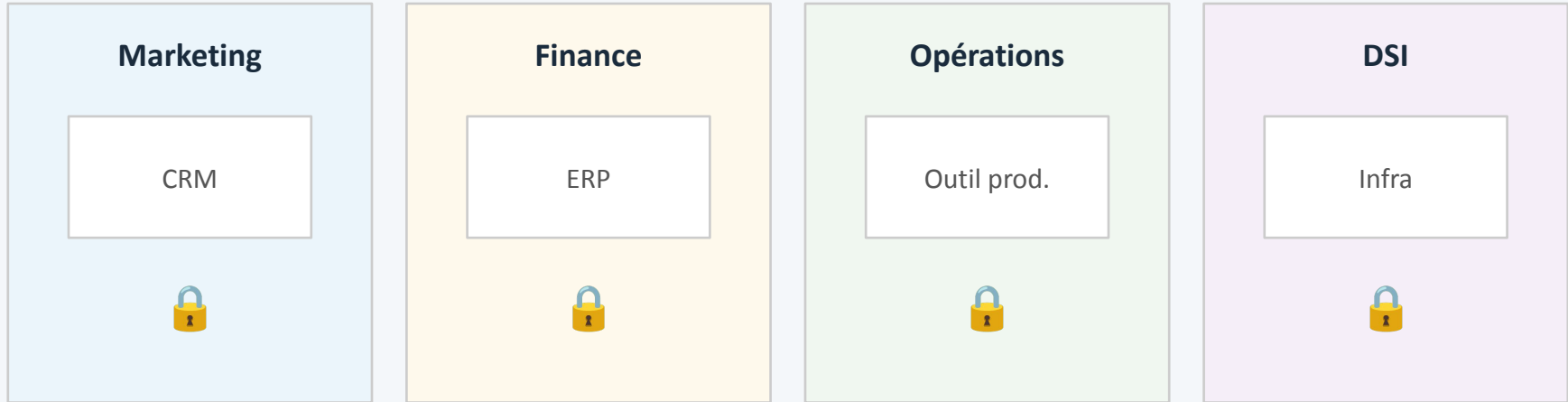


La plupart des entreprises bloquent ici

60 —
80%

du temps d'un data scientist est consacré à préparer les données — pas à modéliser.

Le problème des silos



Résultat : les projets IA qui démarrent dans ces conditions sont condamnés à rester des preuves de concept. Jamais à l'échelle.

PARTIE 2

02

L'avantage compétitif par la donnée

Data moat · Flywheel · Valeur × Risque × Dépendance · 4 modèles

Le data moat — le fossé défensif

La donnée propriétaire crée une barrière à l'entrée que vos concurrents ne peuvent pas acheter.

Amazon

300M+ clients
25 ans d'historique d'achat

Recommandation imbattable

Doctolib

Tous les RDV
médicaux de France

Capital impossible à rattraper

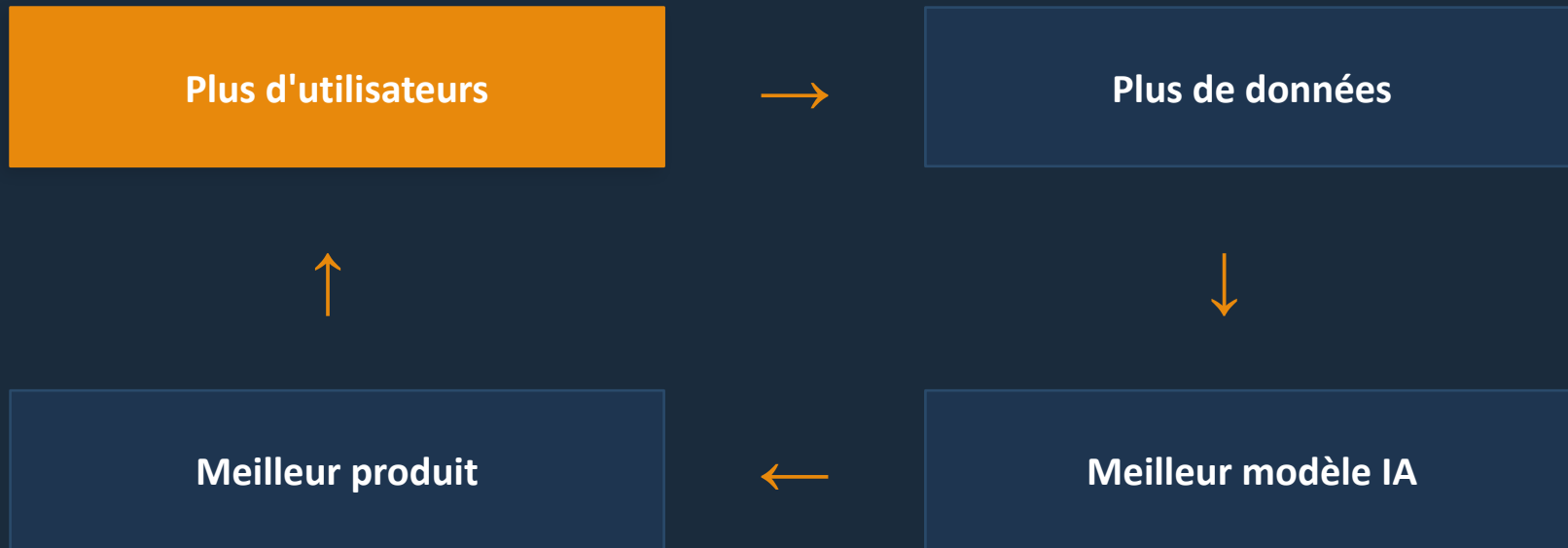
Tesla

Milliards de km
de conduite réelle

Autonomie très difficile à copier

Le pattern commun : l'avantage n'est pas dans la technologie — il est dans la donnée accumulée sur le temps long.

Le flywheel — la roue qui s'auto-alimente



Une fois lancée, cette roue est très difficile à arrêter — et impossible à reproduire rapidement par un concurrent.

Ce mécanisme est puissant. Mais regardons-le avec notre grille.

+ VALEUR

Avantage compétitif durable.
Data moat. Flywheel.
Différenciation impossible à acheter.

! RISQUE

Vol de données (cyber).
Décisions biaisées à l'échelle.
Usage non-conforme (S4).

? DÉPENDANCE

Donnée hébergée chez AWS
us-east-1.
Cloud Act américain.
Qui contrôle vraiment votre
avantage ?

4 modèles de création de valeur par la donnée

01 Monétisation directe

Nielsen, Dun & Bradstreet

Vendre la donnée agrégée à d'autres entreprises.

02 Amélioration produit

Netflix, Spotify, Amazon

Personnalisation → meilleure rétention.

03 Efficience opérationnelle

Michelin, SNCF, Airbus

Maintenance prédictive, réduction des pannes.

04 Nouveau service

John Deere, Rungis

La donnée devient le produit lui-même.

PAUSE ACTIVE

Exercice participatif

Dans votre entreprise ou stage :
quelle donnée est produite chaque jour et n'est pas exploitée ?

2 minutes de réflexion individuelle — puis 3 témoignages

PARTIE 3

03

Gouvernance & qualité des données

Définition · 4 problèmes · 2 rôles clés · Data catalog · 3 questions du manager

La gouvernance, c'est répondre à trois questions.

01 Qui a le droit de toucher à quelle donnée ?

02 Dans quel état cette donnée doit-elle être ?

03 Qui est responsable si ce n'est pas le cas ?

4 problèmes chroniques des données en entreprise

1 Silos

Chaque département a ses données. Personne ne peut croiser les sources.

2 Qualité

Doublons, champs vides, formats inconsistants, données obsolètes.

3 Traçabilité

On ne sait plus d'où vient la donnée ni qui y a touché.
(Data lineage)

4 Conformité

Les données ont une adresse légale. Droit d'utilisation à vérifier. (→ S4)

2 rôles à retenir — et le data catalog

Data Owner

Le responsable MÉTIER d'un domaine de données.

Il décide quoi collecter et dans quel but.

Ce n'est pas un profil technique.

C'est vous — directeur marketing,
directeur commercial, responsable RH.

Chief Data Officer

Pilote la stratégie données de l'entreprise.

Arbitre entre les besoins des métiers.

Rend compte à la direction.

Rôle stratégique — pas technique.

Un décideur, pas un exécutant.

Data Catalog

L'annuaire interne de vos données. Pour chaque jeu de données : ce qui existe, où c'est stocké, qui en est responsable, qui peut y accéder.

Sans catalogue, chaque projet IA recommence à zéro. Avec un catalogue : quelques minutes pour savoir ce qui est disponible.

Avant tout projet IA : les 3 questions du manager

1 Est-ce que la donnée nécessaire existe quelque part dans l'organisation ?

2 Est-ce qu'elle est propre et exploitable, ou faut-il 6 mois de travail préparatoire ?

3 A-t-on le droit légal et contractuel de l'utiliser pour ce cas d'usage précis ?

Si une seule réponse n'est pas OUI clairement — votre projet n'est pas prêt.

Aucun outil, aucun prestataire, aucun algorithme ne peut compenser ça.

Ce que vous devez retenir

I La donnée est l'actif qui ne s'use pas

Elle se valorise à l'usage. Elle se multiplie. C'est sa nature asymétrique.

II Elle crée des barrières infranchissables

Data moat + flywheel = avantage que l'argent seul ne peut pas acheter.

III Mais elle exige d'être gouvernée

Sans gouvernance → risque, non-conformité, projets IA qui échouent.

Maintenant : la réalité du terrain.

Comment une plateforme industrielle active tout ça concrètement.

Intervention Dataiku — 30 minutes

Questions & Réponses

Vos questions sur la séance ou l'intervention

À la prochaine séance :

Séance 4

Droit, éthique et responsabilité

Le cadre contraint : conformité, risques et responsabilité managériale.