

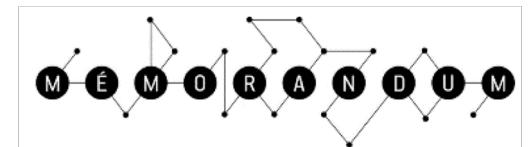
Que faire du big data dans l'entreprise ?

Romain Jouin

→ 06 52 86 87 30 – romain.jouin@memorandum.pro

Denis Oblin

→ 06 71 62 74 92 – denis.oblin@memorandum.pro



Objectif : partager et illustrer 3 propositions

1

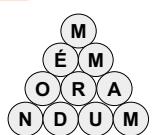
« La valeur n'attend pas le nombre des octets »

2

« Exploiter la donnée est un projet métier, pas un projet informatique »

3

« Fail fast & learn »



Objectif : partager et illustrer 3 convictions

1

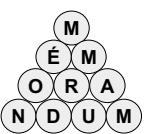
« La valeur n'attend pas le nombre des octets »

2

« Exploiter la donnée est un projet métier, pas un projet informatique »

3

« Fail fast & learn »

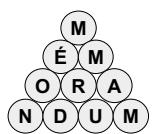


Big data : Créer la botte de foin pour trouver des aiguilles

Vous ne pouvez pas continuer à miser sur le seul bon sens pour prendre vos décisions opérationnelles.

Les réponses à vos questions les plus précises sont dans le croisement de plus en plus large de vos données

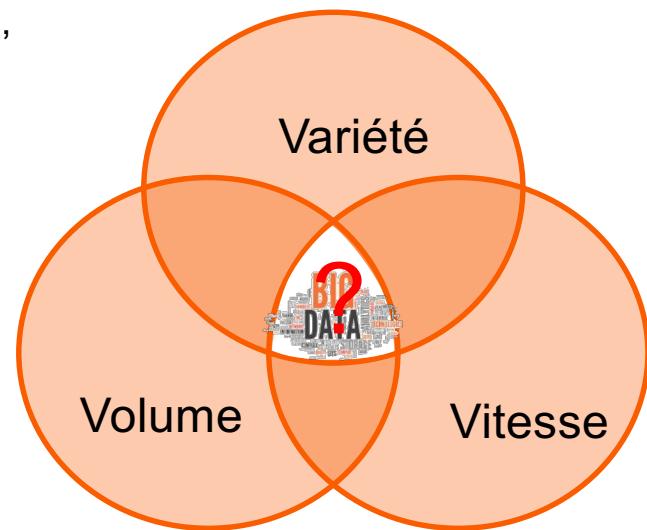
Vous renforcerez votre maîtrise sur votre activité



A partir de quel volume est on éligible aux techniques d'analyses associées au « big data » ?

Une définition fréquente : Au moins 2 des 3 « V »

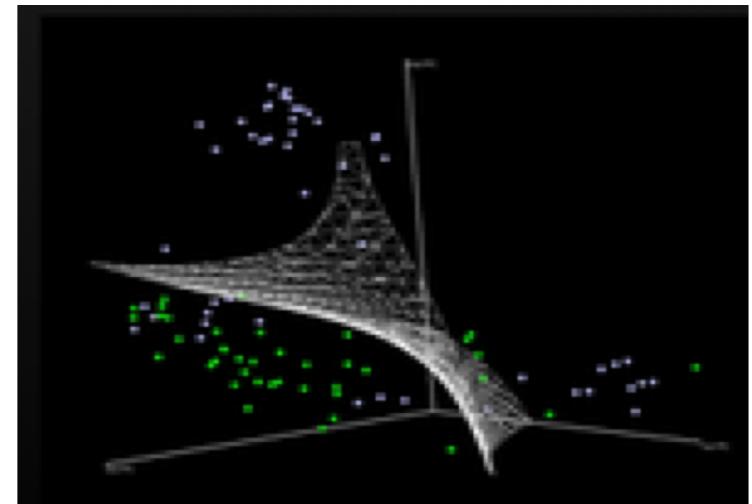
- Volume,
- Vitesse,
- Variété



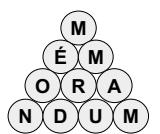
Les 3V ne sont pas un plancher d'accès mais le symbole d'un plafond constamment repoussé par la technologie

Une définition plus juste de la puissance de l'approche → 0 V

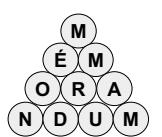
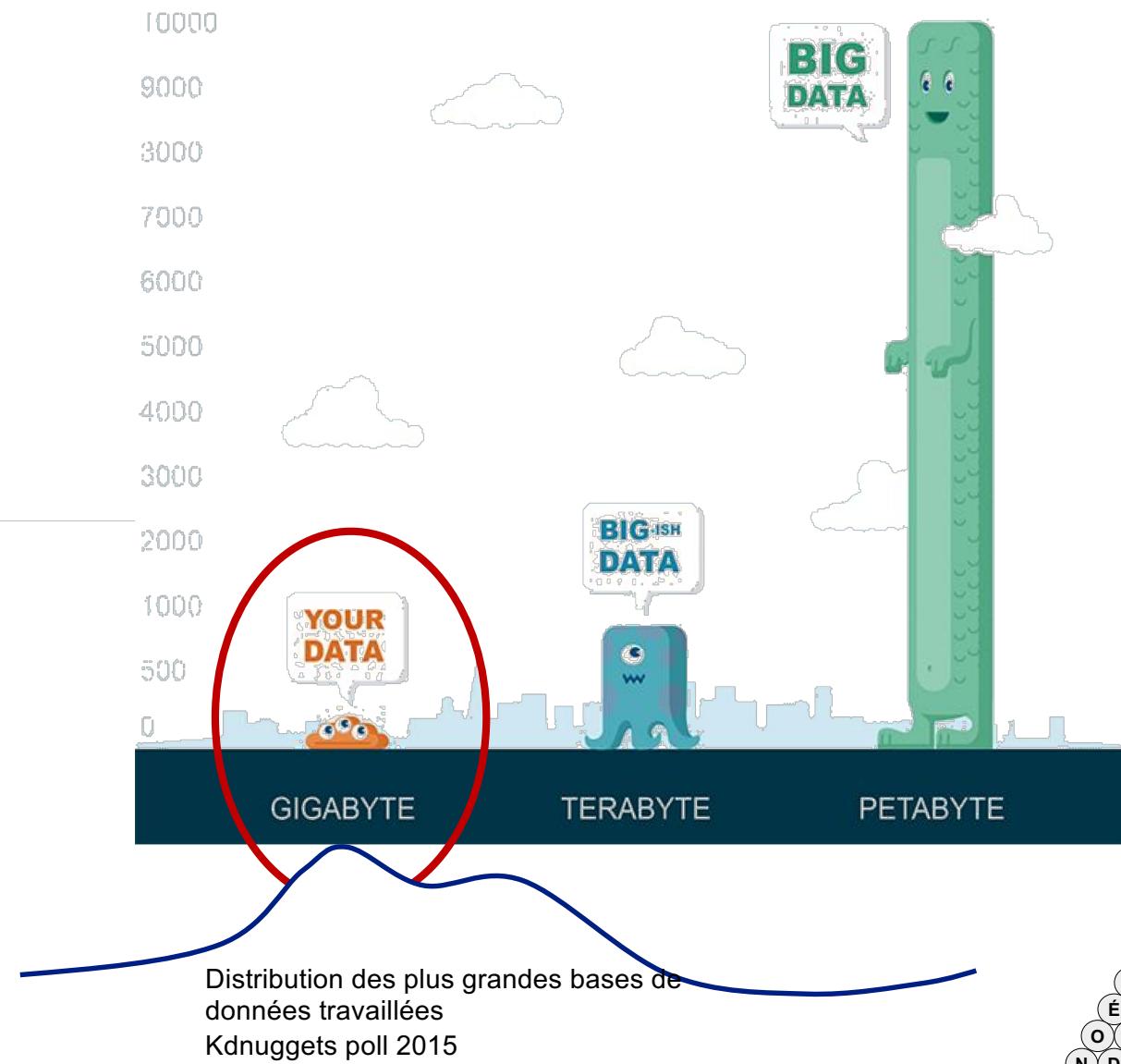
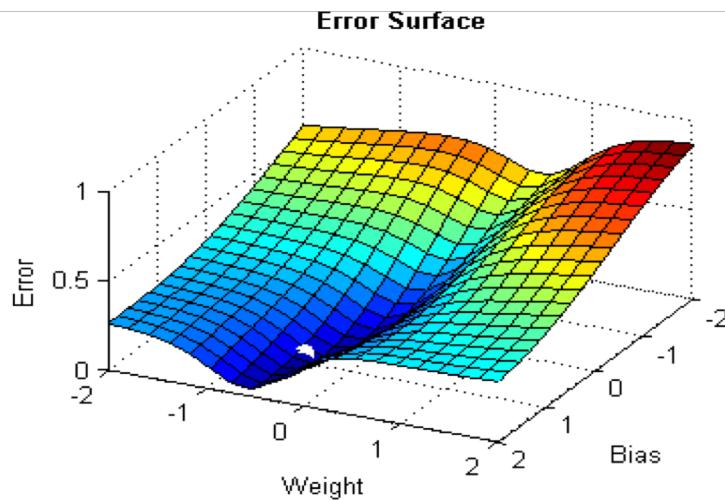
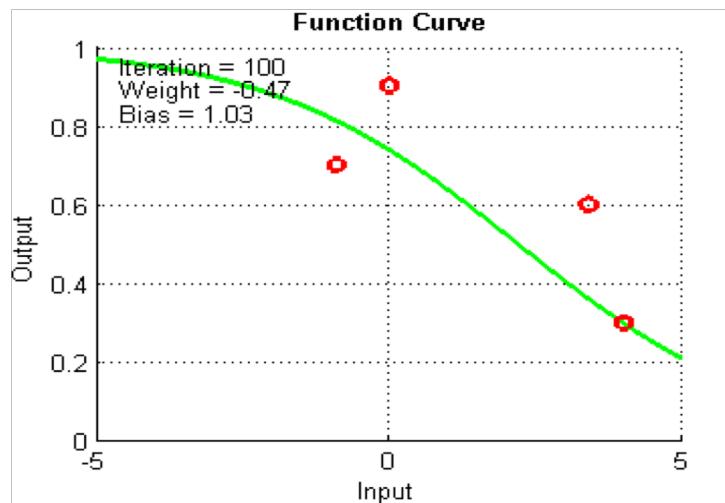
- ci-dessous : 150 observations, 3 caractéristiques → une modélisation fine !



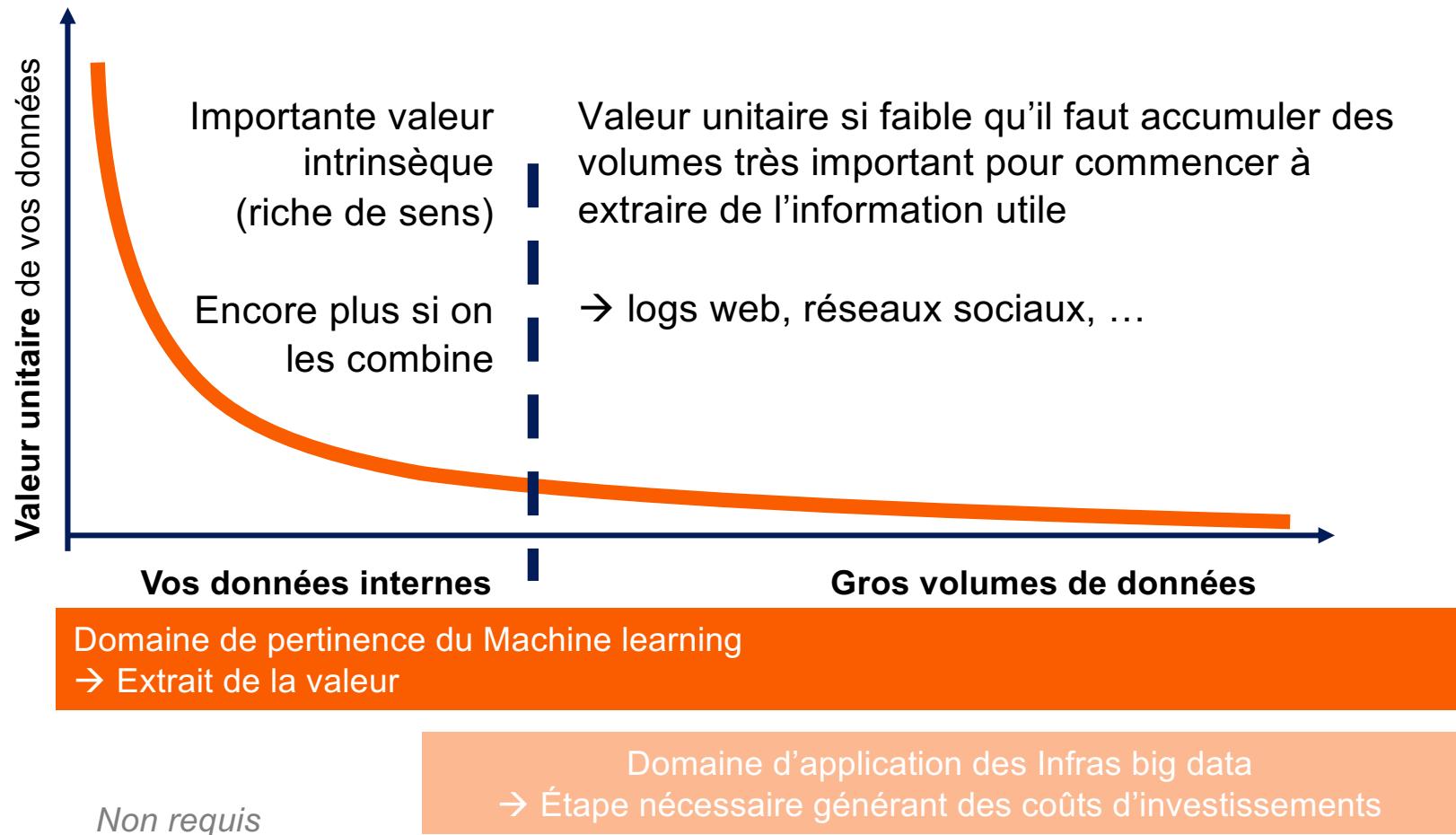
Votre entreprise a assez de données pour tester l'application des algorithmes prédictifs open source (ceux qui marchent !)



Valoriser ces données : une affaire de traitement et non de volume

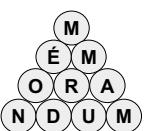


Vous êtes riches !



Vos données internes sont à la fois

- Concentrées et accessibles : chez vous, dans votre SI
- Éclatées au sein de votre organisation



Chacune de vos données embarque deux leviers de valeur

Une donnée parmi d'autres :
« Le client X a contacté le service client »

L'information pour elle-même

Déclencher l'action suivante : Répondre au client

Mettre à jour la rémunération variable du conseiller,

Alimenter les ~~reportings~~

→ **Domaine connu**

La donnée n'a été initialement produite que pour cela. Après usage cette donnée est aujourd'hui un déchet

Le contexte d'autres événements

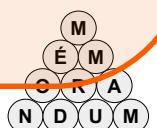
Reconstituer le contexte d'autres événements de l'entreprise, tous interdépendants à des degrés divers :

Une vente s'est faite entre un client donné et un produit particulier, dans un magasin spécifique, avec un vendeur unique, un certain jour de la semaine

Le client s'était manifesté 3 jours avant pour dire xxx

Un exemple de préoccupation

Une donnée sortie de son contexte qui éclaire le sujet



Objectif : partager et illustrer 3 convictions

1

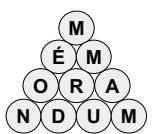
« La valeur n'attend pas le nombre des octets »

2

« Exploiter la donnée est un projet métier, pas un projet informatique »

3

« Fail fast & learn »

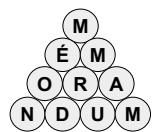


Il est plus simple de regarder ailleurs (données web) que de partager ses propres données internes

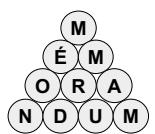
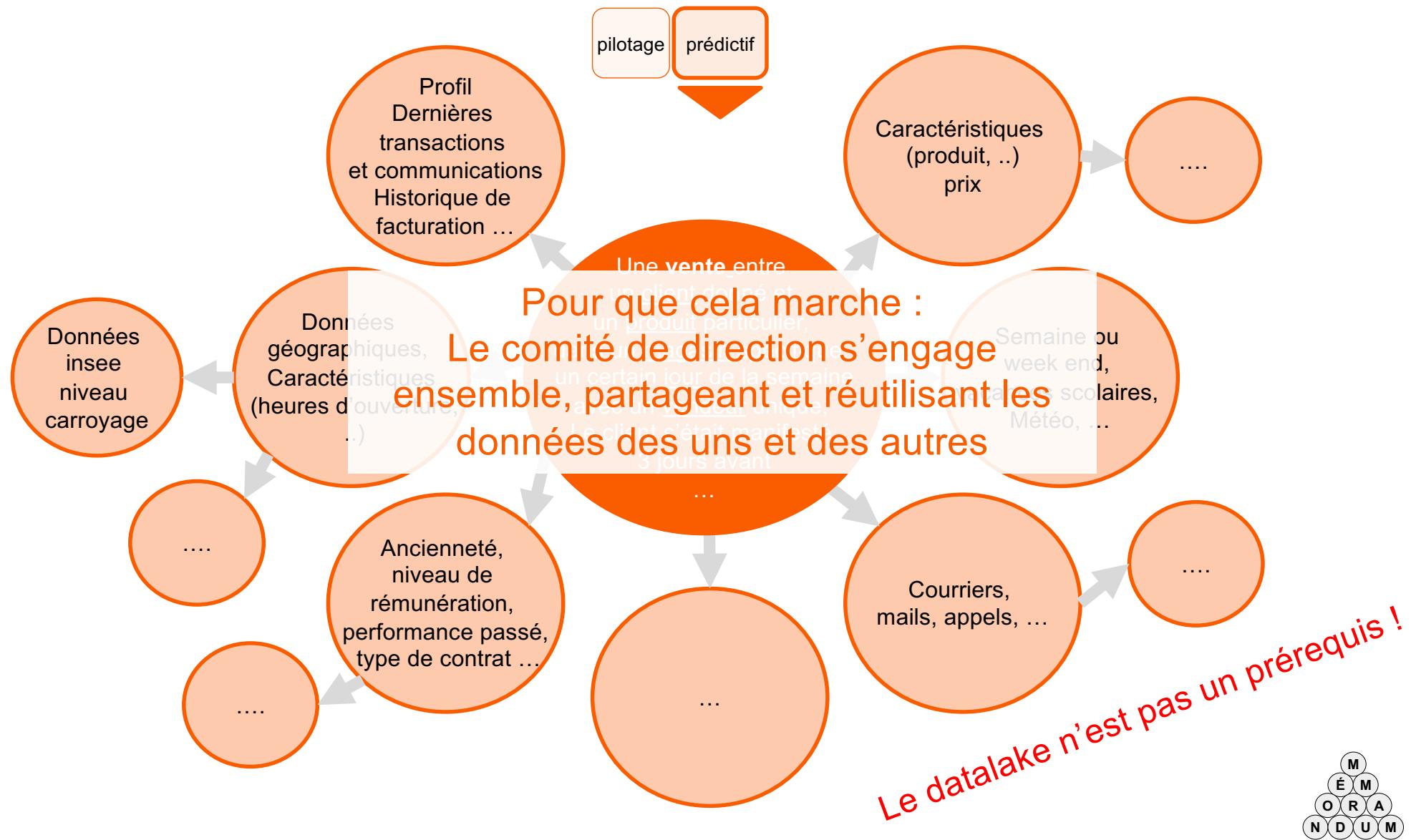
Partager ses données au sein de l'entreprise n'est pas simple

Il est souvent plus rapide de lancer un projet informatique sur les données web qu'un projet métier sur les données de chacun au sein de l'entreprise

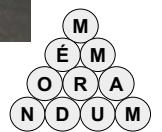
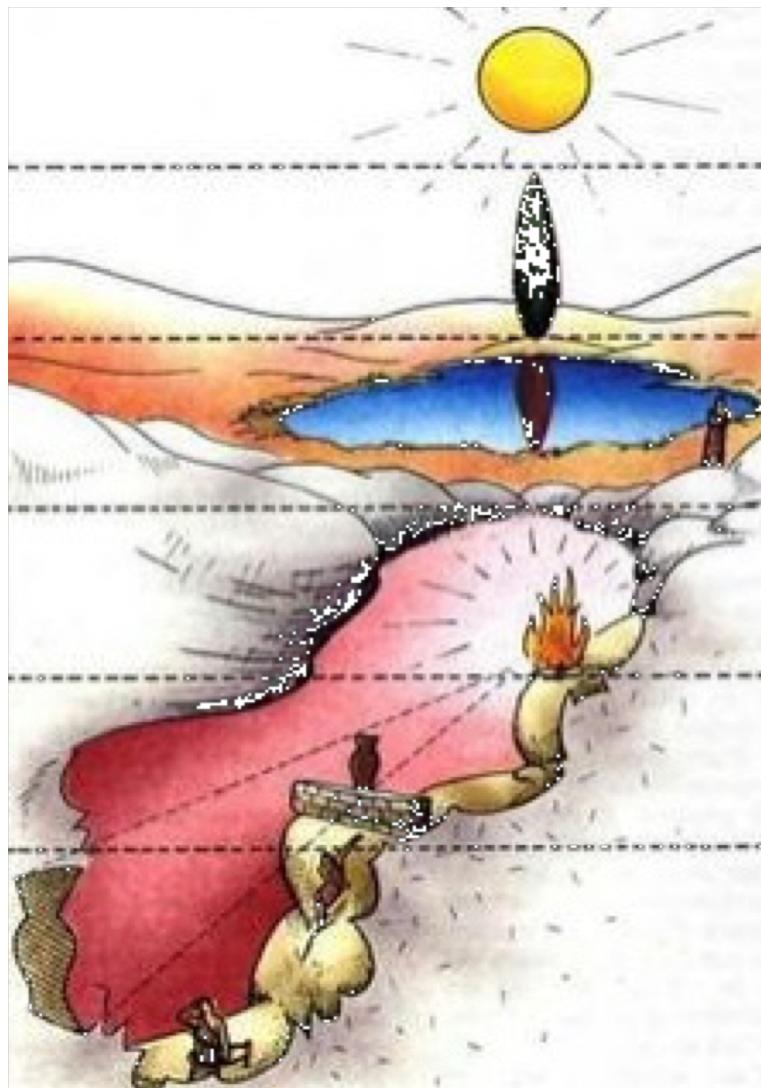
Ce miroir est pourtant un puissant levier de transformation



Les données les plus importantes sont dans les bases de données des métier



La donnée n'est que l'ombre de la réalité : la manipuler et l'interpréter mérite une supervision métier

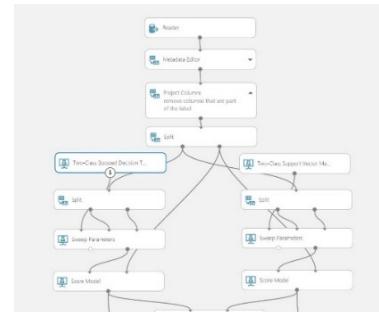


La barrière à l'entrée des techniques d'analyse se réduit progressivement

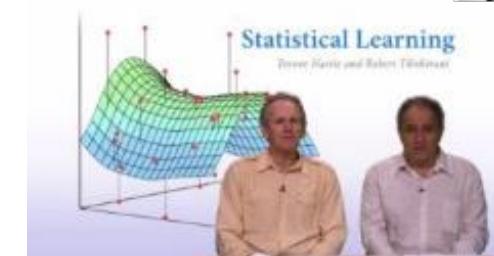
Les algorithmes de machine learning sont packagés et en libre accès dans différents environnement open source (R / Python)



Des interfaces graphiques se développent (Dataiku, Azure ML, ...) : possibilité de pratiquer sans coder

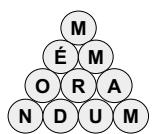


Des formations nombreuses et de bon niveau disponibles en ligne (MOOC)



Une communauté active et ouverte

...

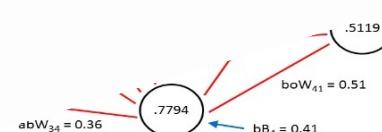


L'exploitation de la donnée en mode « big data » change profondément le processus de décision



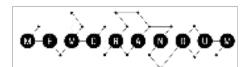
ARRÊTEZ !! LA JEUNESSE OLIVIER FABRE, SOCIALINTELLIGENCE

Mais dans nos clients nous demandons de pouvoir créer des homogènes (Personas), ayant potentiels leurs messages à leurs marchés.



Client T
Client U
Client v
Client w
Client X

Client ...



Objectif : partager et illustrer 3 convictions

1

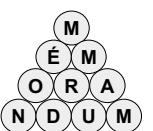
« La valeur n'attend pas le nombre des octets »

2

« Exploiter la donnée est un projet métier, pas un projet informatique »

3

« Fail fast & learn »



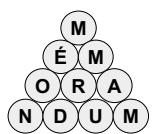
La difficulté principale est souvent plus dans la question que dans la réponse

Fail Fast

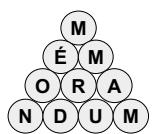
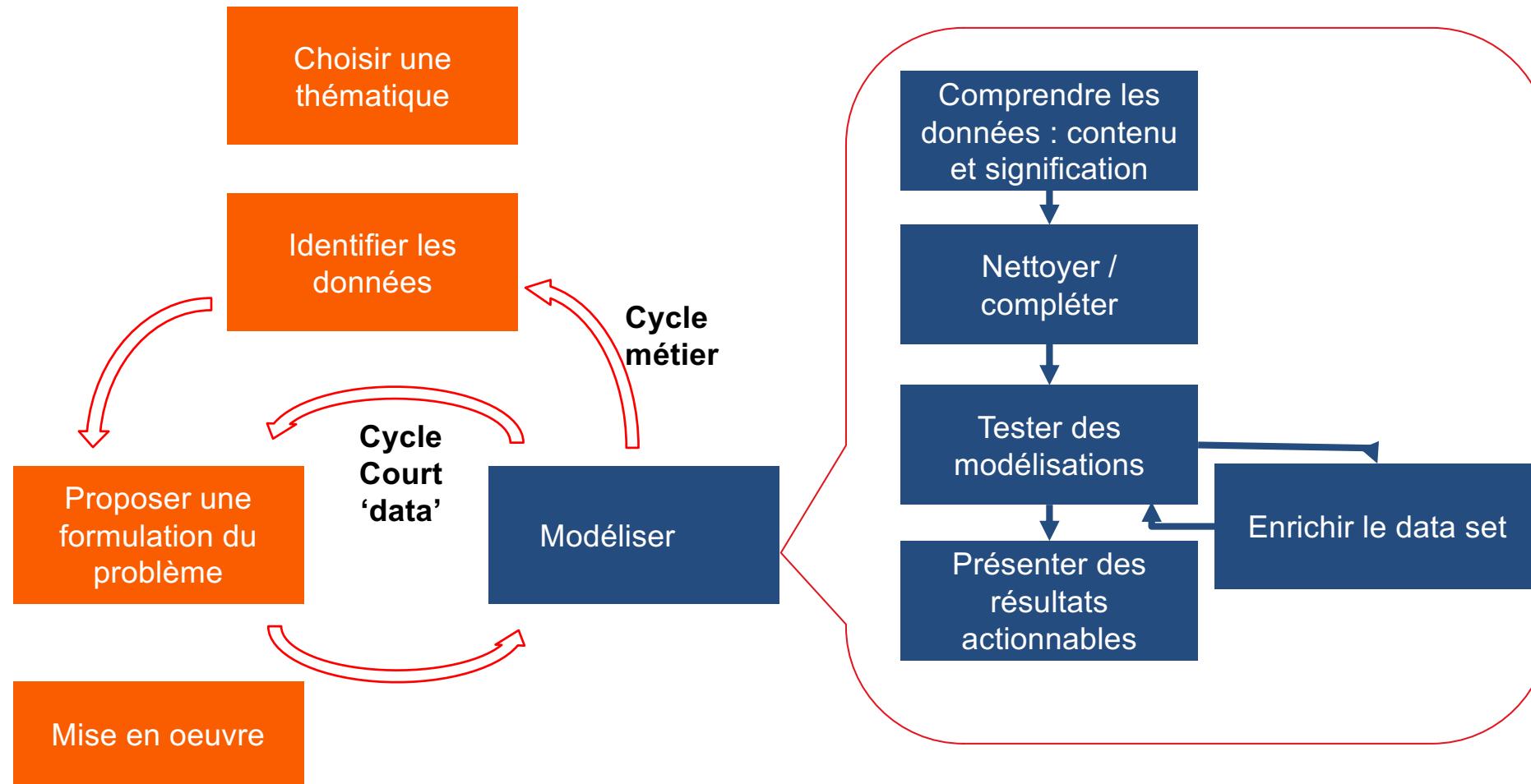
Trouver la bonne question à poser à un data scientist est aussi difficile que de choisir les questions qu'on pourrait poser à un Génie

– Espace infini de possibilités, impossibles à hiérarchiser, difficiles à formuler précisément

Heureusement avec un data scientist vous pouvez (et même devez) tester vos questions et les adapter au fur et à mesure que vous obtenez des réponses.



Des cycles courts pour des résultats rapides et une amélioration continue



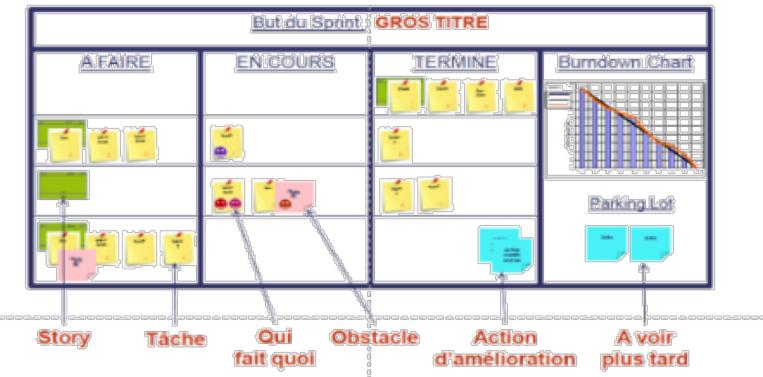
Exemple de leviers :

Structurer les méthodes de travail et les livrables

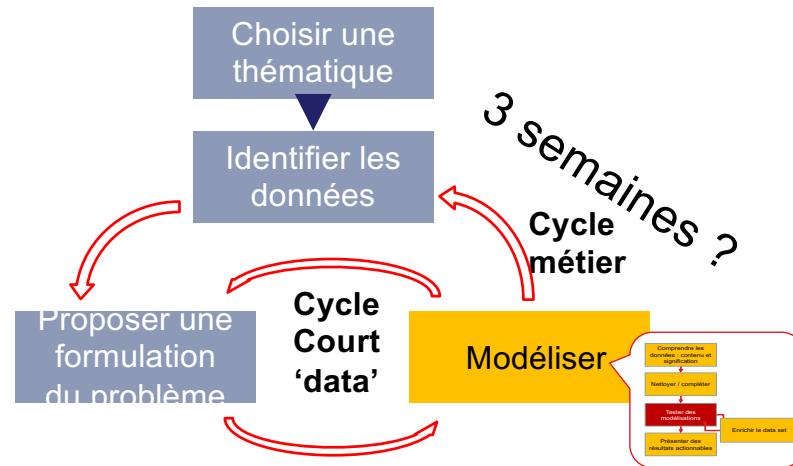
Fail Fast

Dérouler une structure de travail adaptée des méthodes agiles / scrum

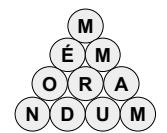
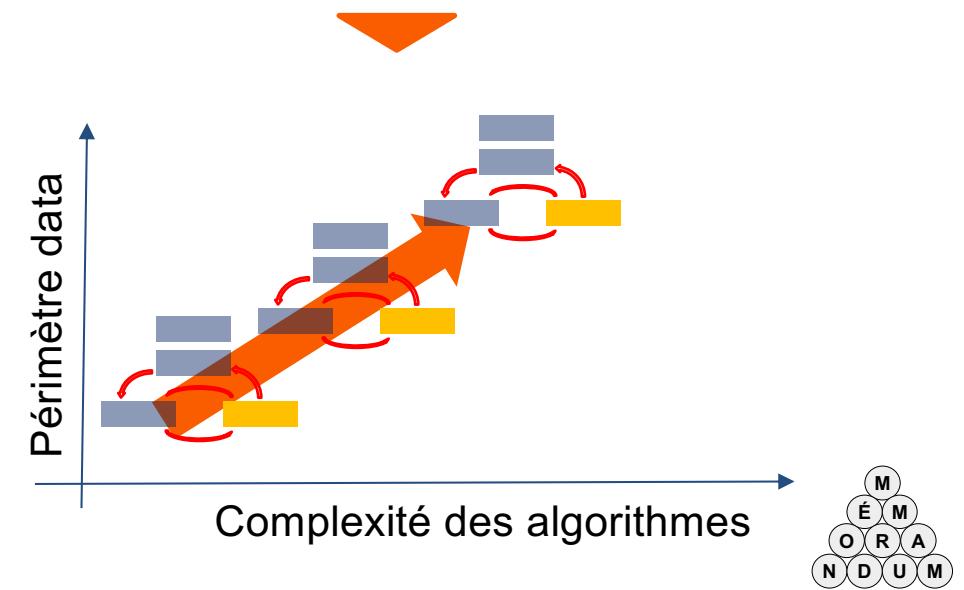
- Rend lisible l'avancée des travaux



Une démarche itérative qui fait progresser la question en même temps que la réponse



Des itérations qui partent de petits volumes et modèles simples



S'y mettre rapidement : une question vitale même pour les PME d'après BPI France

Fail Fast

Les technologies de modélisation ont occupé financièrement l'innovation historique faible pour ils se les fond concept que la n bien au marketing chain.



80% des dirigeants seront les figurants de demain.

Et vous ?

Oui. Votre business est déjà concerné.

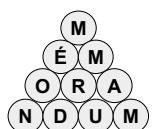
La disruption sera vecteur de compétitivité pour les entreprises prêtes à se réinventer. A condition de faire vite et de définir une vision claire, sous peine de décroître puis sombrer.

“ l'algorithme intime du temps réel : la disruption s'apprend. Heureusement, mais plus profondément encore d'entrer dans une transformation de l'expérience client sur des produits et services qui peuvent désormais bénéficier d'une évolution constante. **”**

Au fond, les modèles et algorithmes aménagent la fois des et en pour start-up. logies ement nieur, es et allant ussage nées, utili- ille et

bpifrance

Modèles, données et algorithmes
Les nouvelles frontières du numérique
Par Etienne de Rocquigny (Aremus & Associés)



Si vous ne valorisez pas vos données, d'autres le feront à votre place

Fail Fast

Le monde change

Exemple de la base adresse nationale

Au début : un actif réservé et facturé



Aujourd'hui un bien commun
(15 avril 2015)



La loi peut changer

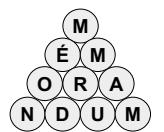
« Je souhaite créer un nouveau statut de données, les données d'intérêt général, qui sont à l'origine des données commerciales qui appartiennent à des entreprises mais qui, si elles étaient ouvertes, pourraient bénéficier à la communauté et à l'intérêt général. »

Axelle Lemaire (2015)



OpenStreetMap France
Cartographions le monde rue après rue...

Un nouvel acteur reconstitue la base en open source



Learn

Comment rater son projet ?

Sous traiter tout

Simplifier le résultat pour mieux comprendre

Séparer celui qui analyse le besoin et celui qui analyse les données

Laisser le druide seul avec sa marmite

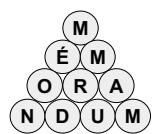
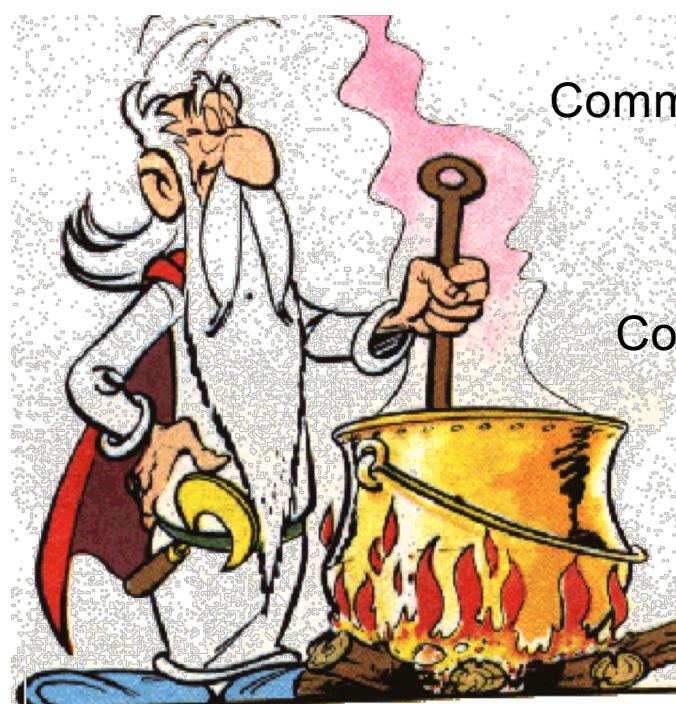
Commencer par BIG

Commencer par acheter une marmite

Commencer par recruter un druide

Commencer par les ingrédients

Poser une question et s'y accrocher coute que coûte



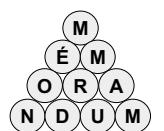
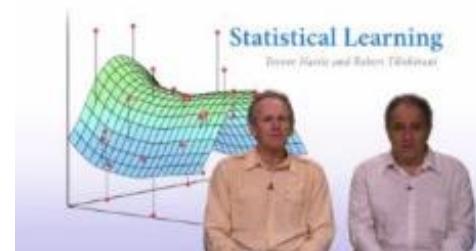
Learn ? Beaucoup de ressources à disposition

Principe lancer une formation interne s'appuyant sur un mooc, ouvert à tous, pouvant être réalisé en partie sur le temps de travail

- Tout le monde apprend un peu
- Quelques uns vont au bout
- Certains sont bons et veulent continuer

L'entreprise

- a passé son message interne sur l'importance qu'elle accorde aux données
- Donne sa chance à tous
- Fait émerger un pool de salariés motivés .. Et reconnaissants



Objectif : partager et illustrer 3 propositions

1

« La valeur n'attend pas le nombre des octets »

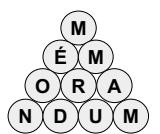
2

« Exploiter la donnée est un projet métier, pas un projet informatique »

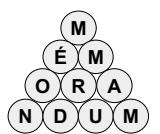
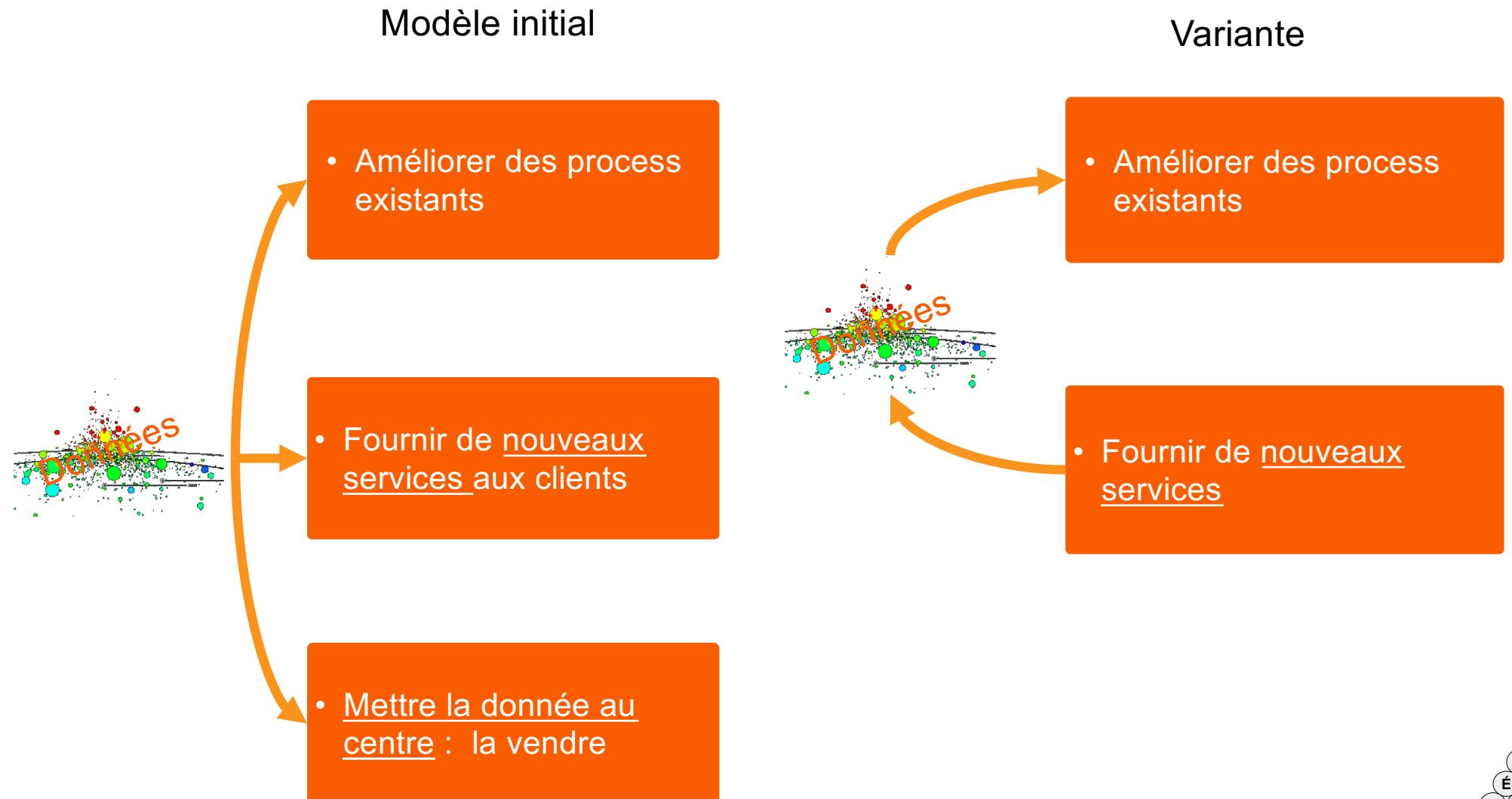
3

« Fail fast & learn »

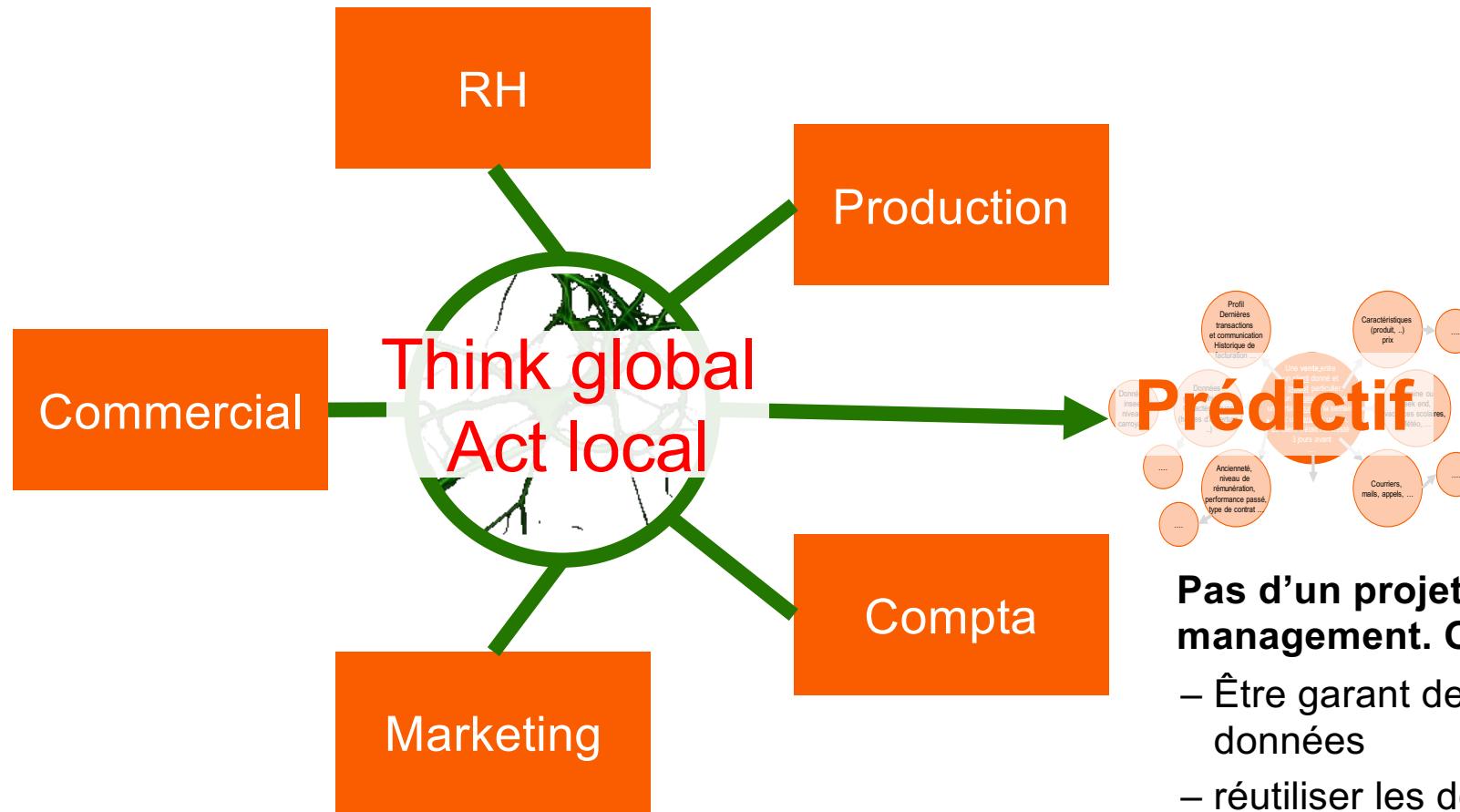
Pour quoi faire ?



Pour quoi faire ?



La donnée : un vrai projet d'entreprise Vos données aussi doivent être animées en transverse !



Pas d'un projet informatique mais de management. Objectif managers :

- Être garant de la disponibilité de ses données
- réutiliser les données des autres directions

