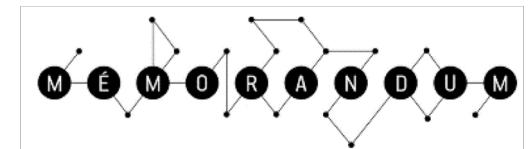
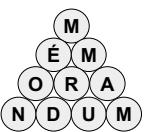
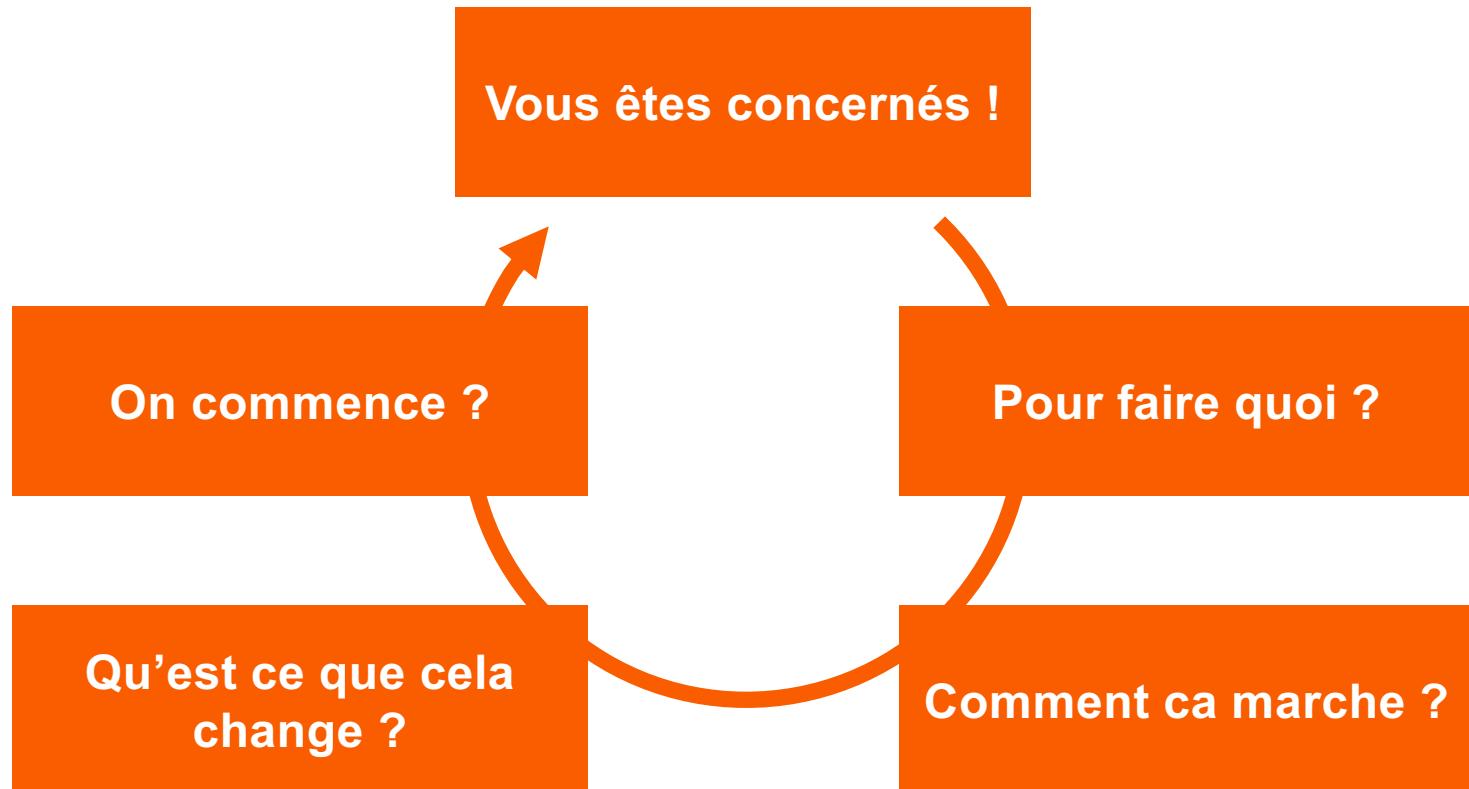


La révolution des Datas dans l'entreprise

Denis Oblin
Romain Jouin

3 mars 2016





Vous êtes
concernés !

On
commence ?

Pour quoi
faire ?

Qu'est ce que
cela change ?

Comment ca
marche ?

Extraits du livre blanc de bpifrance pour promouvoir le big data auprès des PME

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



Aremus & Associés
CAPITAL-INNOVATION

80% des dirigeants seront les figurants de demain.

Et vous ?

Oui. Votre business est déjà concerné.

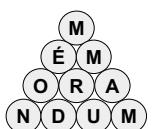
La disruption sera vecteur de compétitivité pour les entreprises prêtes à se réinventer. A condition de faire vite et de définir une vision claire, sous peine de décroître puis sombrer.

“ l'algorithme intime de De plus en plus temps réel : la disruption s'apprend. 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. ”

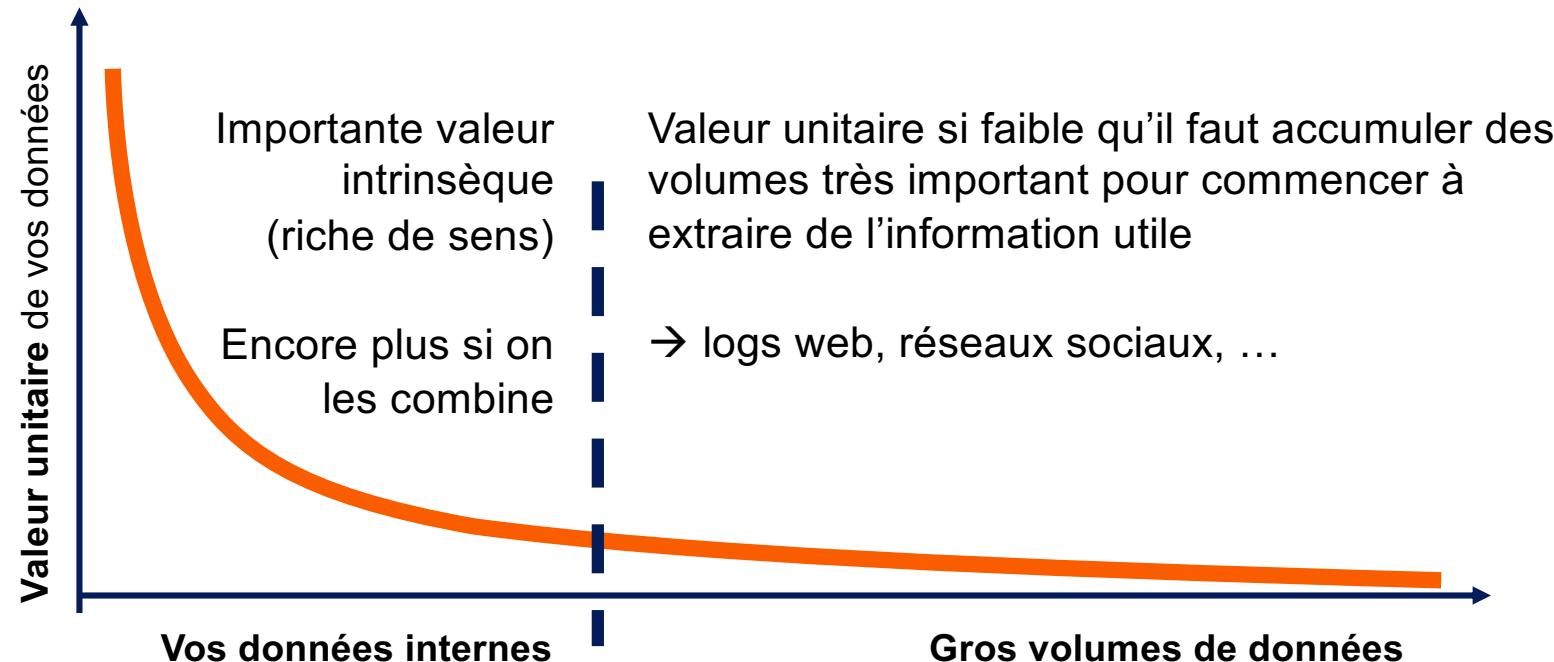
Au fond, les modèles et algorithmes aménagent la fois des et en pour start-up. logies ement nieur, es et allant usage nées, utili- ille et

bpifrance

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



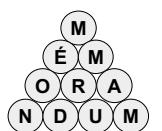
Vous êtes riches !



Les projets de valorisation des données focalisent souvent sur les sources de gros volumes (justifiant l'infrastructure) au détriment des données à forte valeur intrinsèques, déjà disponibles et accessibles

Vos données internes sont à la fois

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



On commence ?

Pour quoi faire ?

Qu'est ce que cela change ?

Comment ça marche ?

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 ~~rapports~~

→ **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

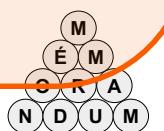
Reconstituer le contexte d'autres événements de l'entreprise, tous interdépendants à des degrés divers :

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

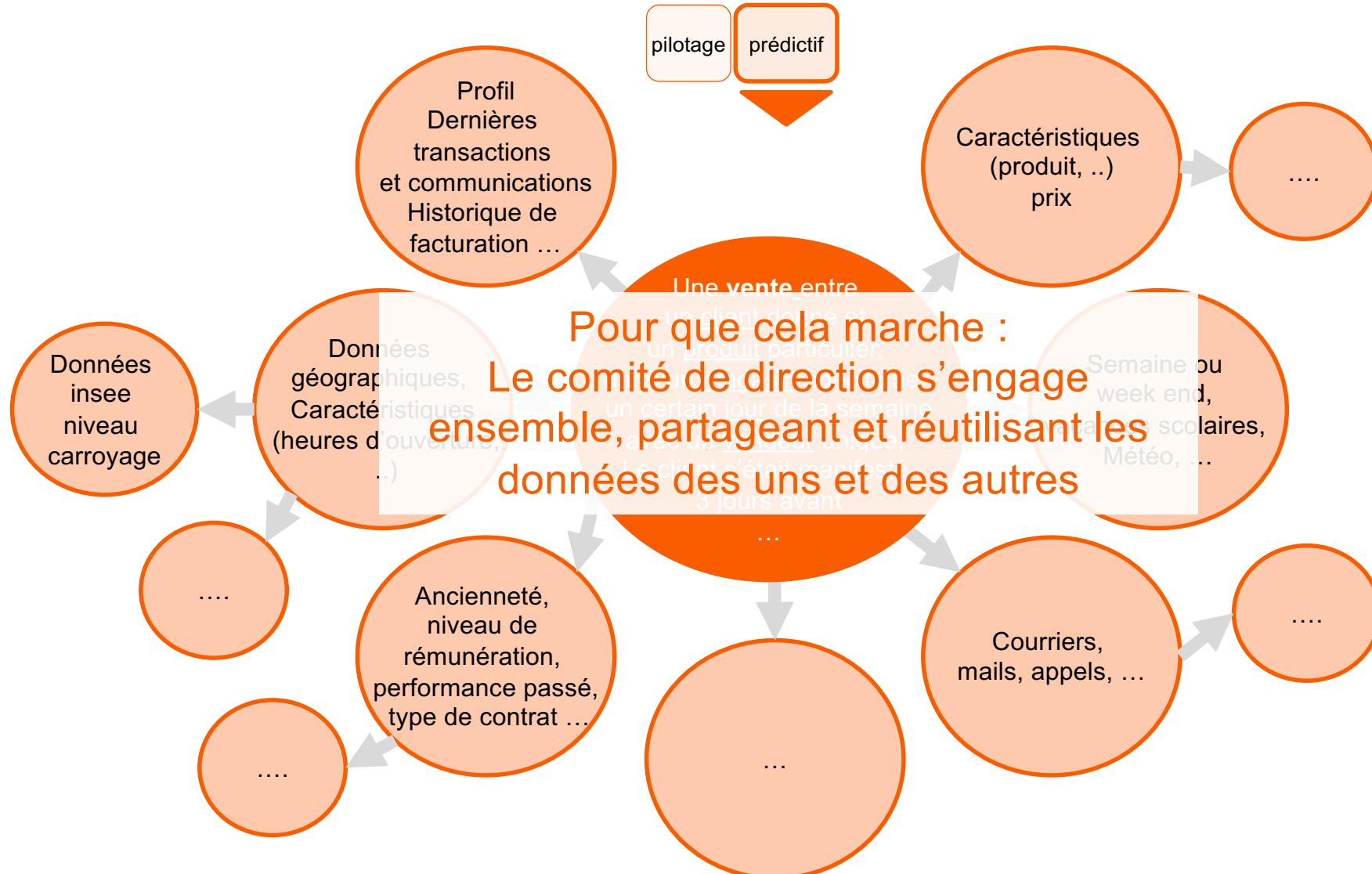
Le client s'était manifesté 3 jours avant pour dire xxx

Un exemple de préoccupation

Une donnée sortie de son contexte qui éclaire le sujet

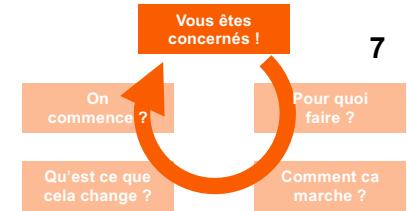


L'approche *machine learning* permet de reconstituer un contexte global



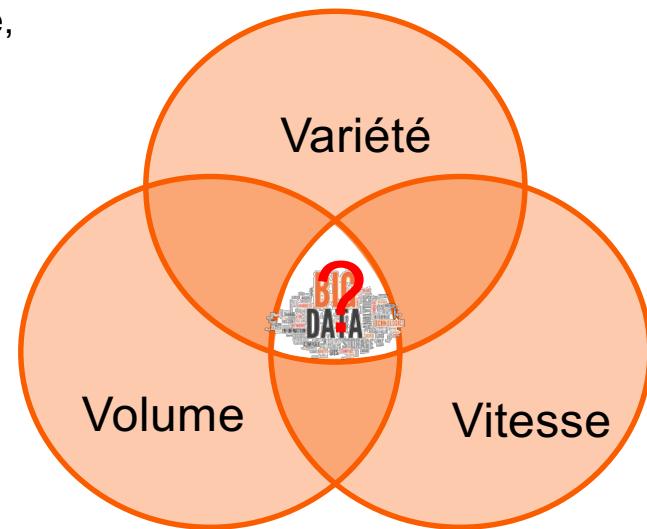
6

A partir de quel volume est on éligible aux techniques du big data ?



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

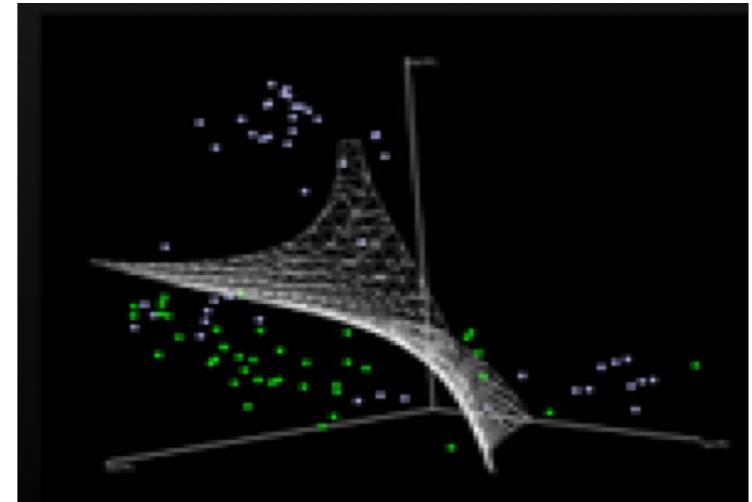
- Volume,
- Vitesse,
- Variété



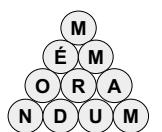
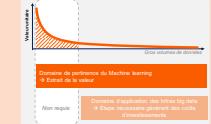
Les 3V ne sont pas un plancher d'accès mais le symbole d'un plafond constamment repoussé par la technologie

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 !)



Vous êtes
concernés !

On
commence ?

Pour quoi
faire ?

Qu'est ce que
cela change ?

Comment ca
marche ?

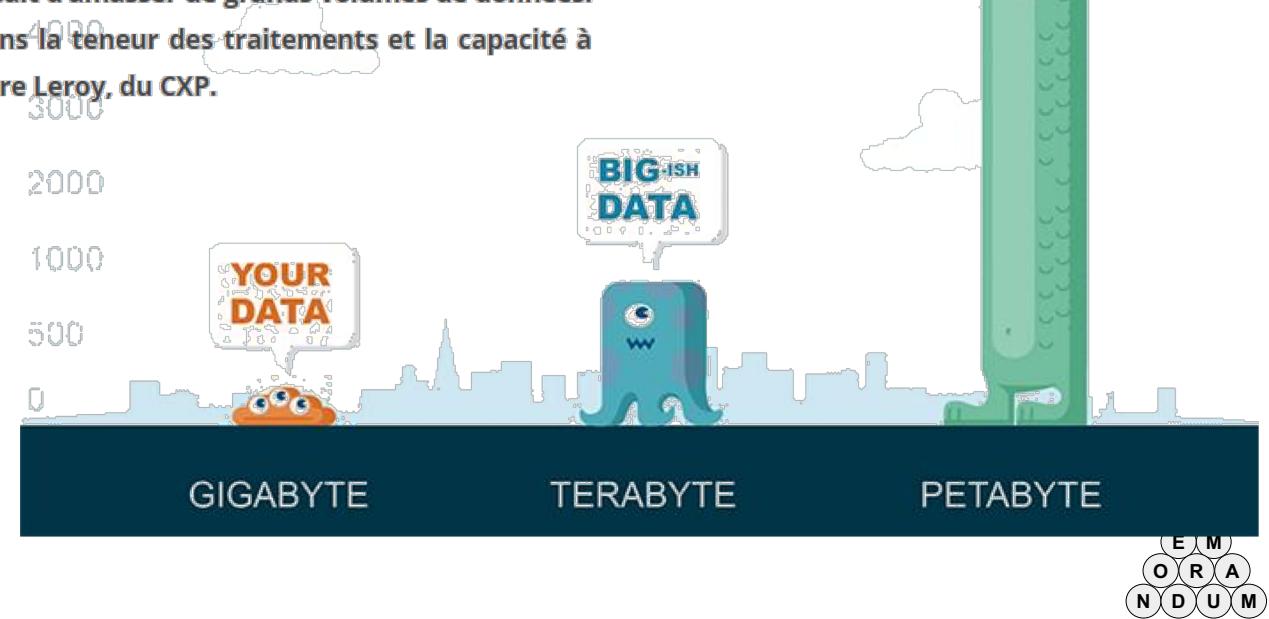
Big Data : une affaire de traitement et non de volume (tribune)

La Rédaction, 26 novembre 2015, 9:32

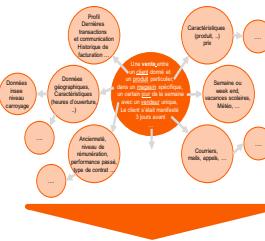
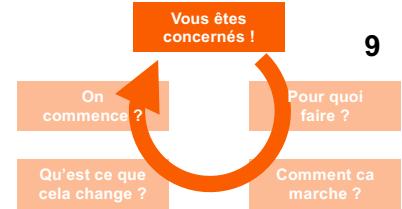


Le terme de Big Data a pu laisser penser qu'il s'agissait d'amasser de grands volumes de données.

Mais la valeur de ces applications réside plutôt dans la teneur des traitements et la capacité à évaluer la pertinence des informations, analyse Claire Leroy, du CXP.



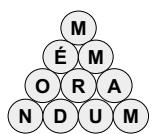
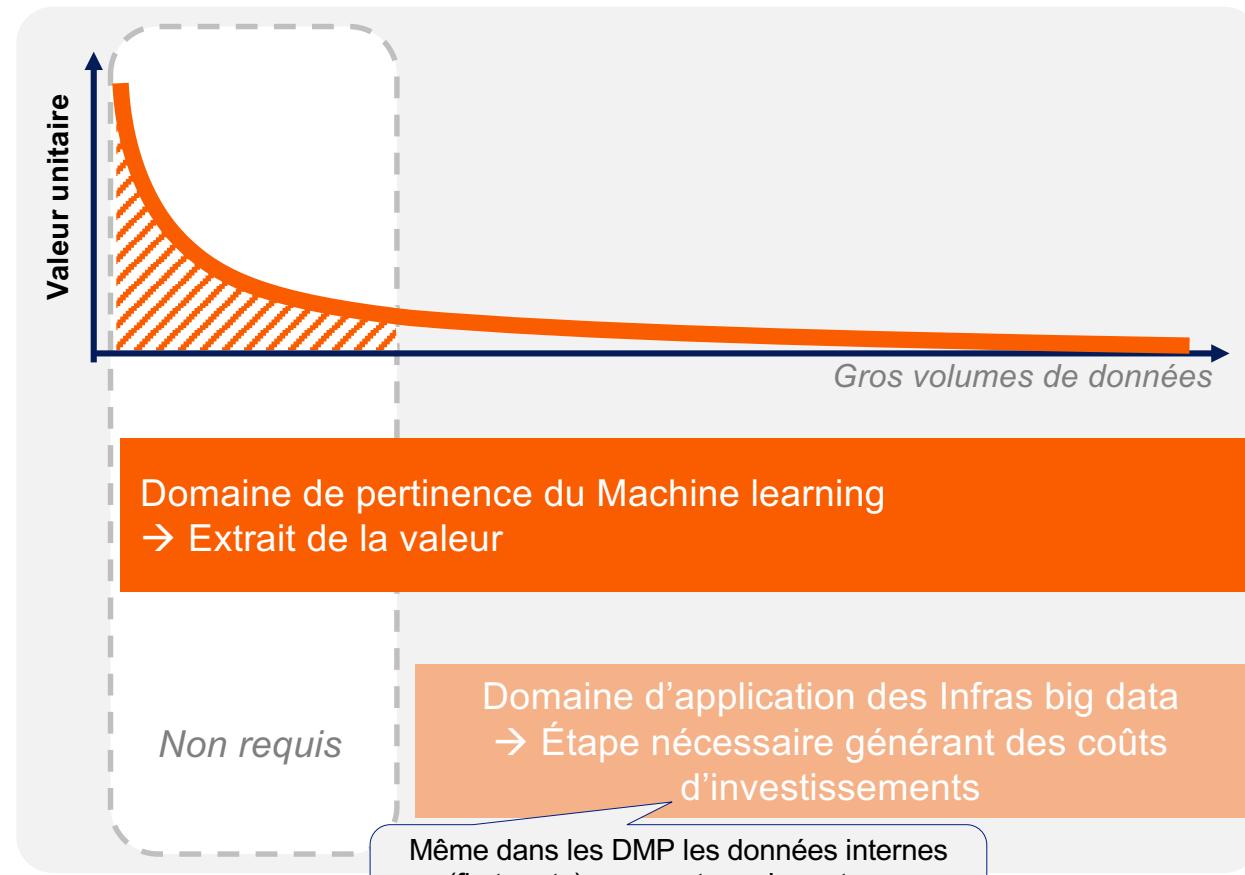
L'exploitation des données internes de l'entreprise cumule 3 avantages

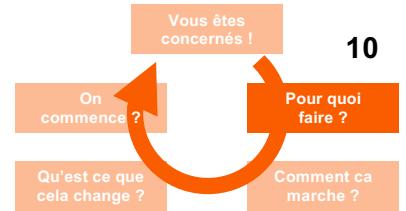


1) Travail sur des données naturellement riches et accessibles

2) Pertinentes pour les algorithmes d'analyse

3) Sans besoin d'infrastructure





Quels contextes souhaitez vous reconstituer : là est la question difficile

Churn

Appétence

Ajustement démarche de vente

Filtrage des risques de fraude

Présélection de cv

Audit comptable

...



**Quoi,
Quand,
Comment
?**

(exemple : détection de chuter à 1 mois ou à 3 mois ?)

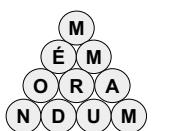
**Quelle
décision
sera modifiée
quand vous
aurez le
résultat ?**

Quels process dans votre entreprise ne rentre pas dans le scope du machine learning ?

**Prédire
ou
Comprendre
?**

Le bon contexte ne se trouve pas du premier coup

Le big data n'est pas un outil de décision stratégique : c'est une stratégie de la micro décision



Applications sectorielles pour les PME, collectées par bpifrance

Agriculture

La collecte de produits pondéreux est couteuse et complexe. Chaque % gagné est une victoire, nouveaux modèles pour comprendre et prescrire les bons comportements exploitants

prévoir les incidents de chantiers.

les données remontées par les engins de chantier sont nombreuses.

bâtiment

E commerce

Moteur d'affectation automatique de produits par catégories sur la base d'un réseau bayésien. algorithme d'optimisation des parcours de mes coursiers.

les évaluations et prévisions sont générées en temps réel

Evaluer le financement nécessaire pour financer ce r

Sans compter les applications sur les services supports (compta, finance, RH, ...)

Services

fois moins cher et une qualité de déménagement optimale.

Passer d'une rente passive d'infrastructure en une plateforme biface simplifiant en temps réel la vie de l'automobiliste et optimisant l'investissement immobilier par l'usage

Recommander des contenus à nos clients avec une finesse inégalée Face à la concurrence d'Amazon, nous ne pouvons plus nous permettre de rester simples éditeurs de livres

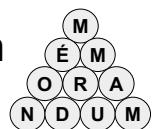
Média

Transport

Maîtrise accrue des aléas en dynamique pour ouvrir de nouvelles manières de servir le client calcule en temps réel la consommation d'essence

La prédiction du churn est un élément clé de notre métier

Telecom



Exemples
Entreprise gérant des interventions techniques
 (plusieurs centaines de techniciens sur toute la France)

Vous êtes concernés !

12

On commence ?

Pour quoi faire ?

Qu'est ce que cela change ?

Comment ça marche ?

Contexte initial

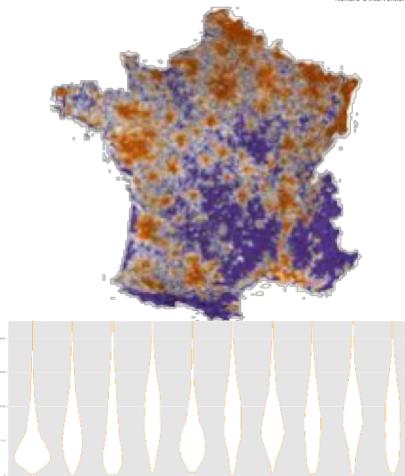
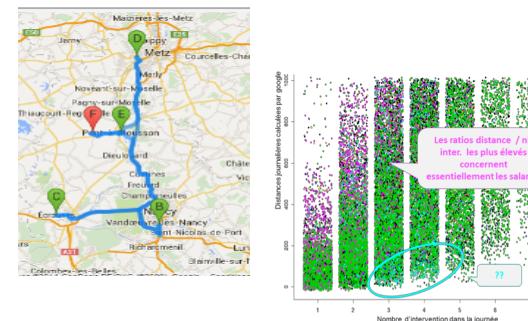
10 ans d'historique
 d'interventions de techniciens,
 jamais exploités

Contexte social tendu sur la
 gestion des planifications de
 tournées

Situation concurrentielle difficile :
 trouver de nouveaux leviers de
 rentabilité

Démarche

Observation des données et
 entretiens avec le comité de
 direction
 – 6 pistes de travail data



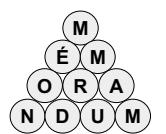
Résultat

Processus de planification
 refondu (forte évolution des
 principes sous-jacents)

Anticipation des interventions
 difficiles (modèle prédictif)

Aide à la décision sur la
 modification de l'implantation du
 réseau

Identification de nouveaux leviers
 de négociation des contrats
 (modèle prédictif de coûts complets)



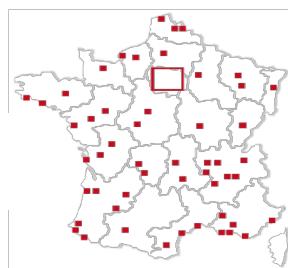
Exemples
Réseau de travail temporaire

Contexte initial

Un réseau de plusieurs centaines d'agences

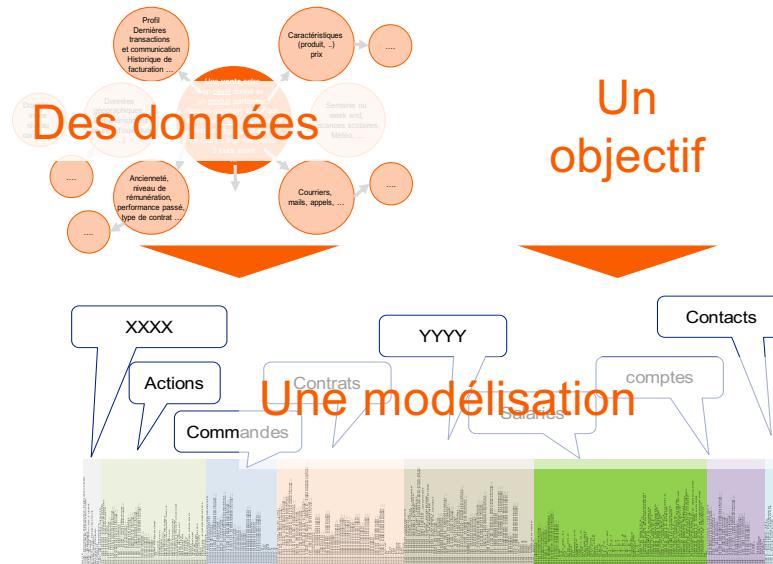
Un objectif opérationnel :
optimisation de la marge brute
par tête

→ Une recherche des leviers d'action cachés



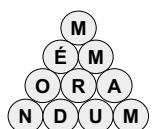
Démarche

- Constitution d'un large spectre de données opérationnelles (RH/commercial/ marketing / .)
- Modélisation de l'indicateur cible
- Analyse des principaux axes du modèle
 - Les premiers sont évidents
 - Les suivants sont en général une surprise ... le travail commence là



Résultat

Exemple : ajuster des heures d'ouverture ou un mode de gestion du flux d'appel en fonction de caractéristiques locales



Vous êtes concernés !

On commence ?

Pour quoi faire ?

Qu'est ce que cela change ?

Comment ça marche ?

Exemples Filiale assurance vie d'un groupe d'assurance

Contexte initial

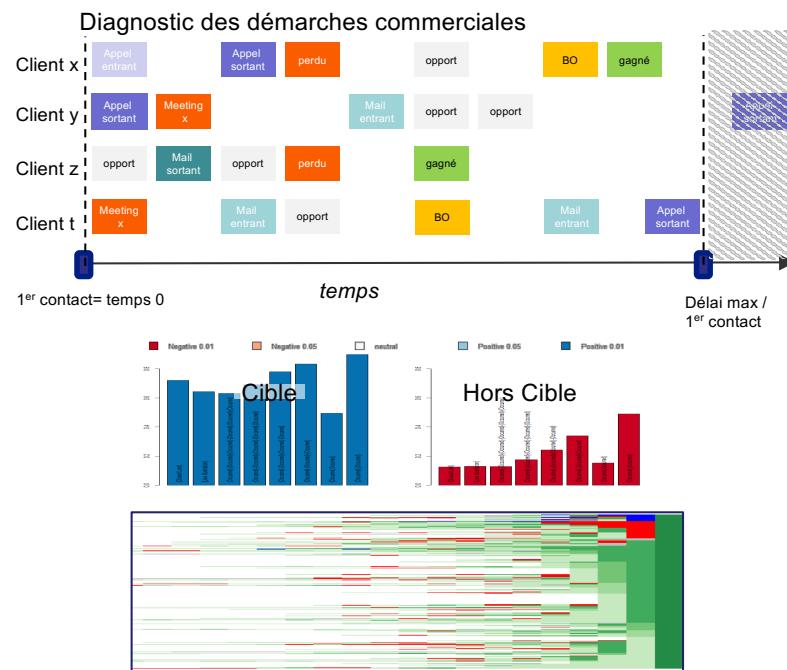
Une jeune équipe de data scientists tout juste formée

Un premier objectif opérationnel sur l'optimisation des forces de vente sur le marché UK

Démarche

Analyse des séquences commerciales sur les 18 derniers mois (salesforce)

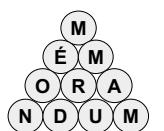
- Visualisation des données
- Modélisation supervisée (probabilités de succès des démarches en cours, faire des recommandations d'action)



Résultat

Montée en compétence rapide de l'équipe (acquisition de données, construction d'interface graphiques, présentation de résultats actionnables)

Prise de conscience par les managers commerciaux de la réalité des démarches commerciales



Vous êtes concernés !

On commence ?

Pour quoi faire ?

Qu'est ce que cela change ?

Comment ça marche ?

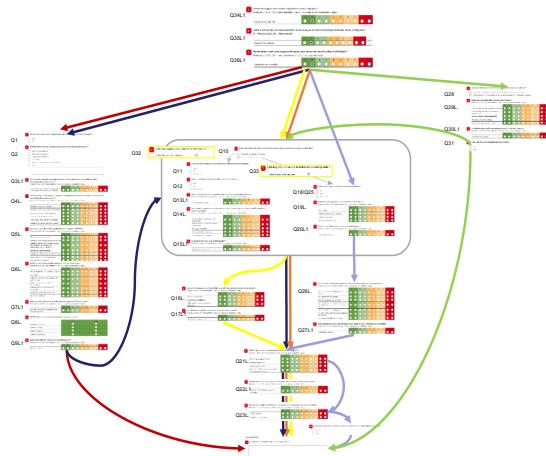
Exemples Opérateur d'enquête de satisfaction client

Contexte initial

Des questionnaires de satisfaction administrés au fil de l'eau

Des centaines de séries temporelles à interpréter !

- Par magasin
- Par question
- Par type de client



Démarche

Identifier les profils opérationnels, leurs objectifs spécifique (retour du client, rentabilité, ..) et leurs leviers d'action

Modéliser les objectifs de chacun : identifier les indicateurs de satisfaction en lien direct avec leurs objectifs

Transformer les données de satisfaction en recommandations opérationnelles pour chaque profil

Evaluer en amont les objectifs de chaque profil interne
Exemple centre de vacances



Profil

Directeur régional

Objectif

rentabilité

Exemple leviers

Organisation

Directeur de site

Retour client

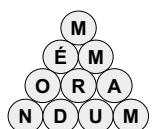
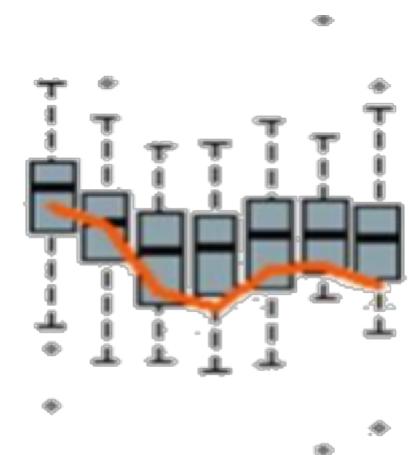
Parcours d'accueil

Chef d'étage NPS

Propreté, ...

Résultat

Un outil de management opérationnel



Vous êtes concernés !

On commence ?

Pour quoi faire ?

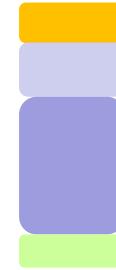
Qu'est ce que cela change ?

Comment ça marche ?

Exemples Centres de contacts → une mine !

Mieux évaluer la performance des conseillers

Événement précédent du client



Contacts du conseiller X

Événement suivant du client



Adapter les démarches de vente au client en ligne



Le client est en fin de droit et ...
Le client est débile et ne comprends rien à ce que je lui explique ...

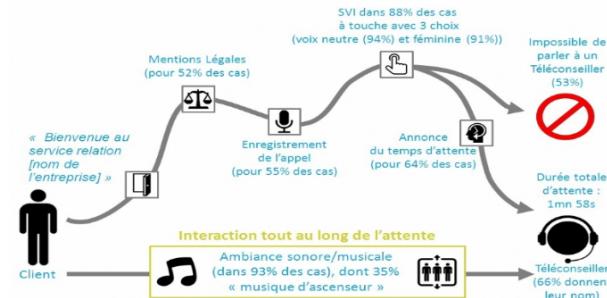
Le client se fait opérer de la prostate et sera indispo ...

Monsieur XX est en prison pour X années

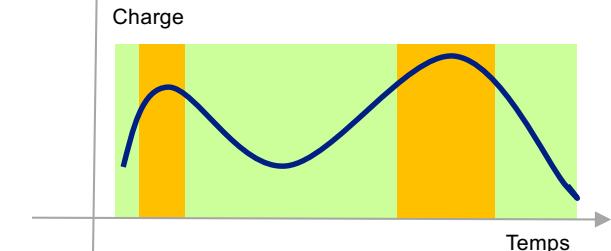
Le client n'a pas de ressources suffisantes pour faire une demande ...
Le client n'a pas de temps pour nous écouter ...
La discussion avec le client n'a pas permis d'aboutir ...
La discussion est difficile ...

Identifier et reformuler des commentaires désobligeants sur les clients

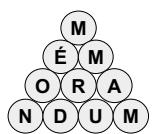
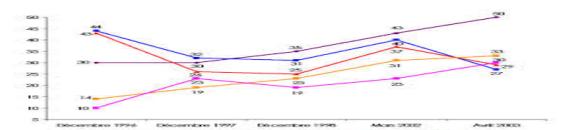
Améliorer la parcours de prise en charge des clients

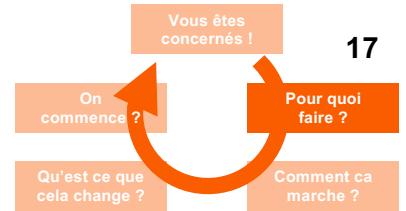


Prévision de flux intégrant le trafic web, la météo, les match de foot, ..

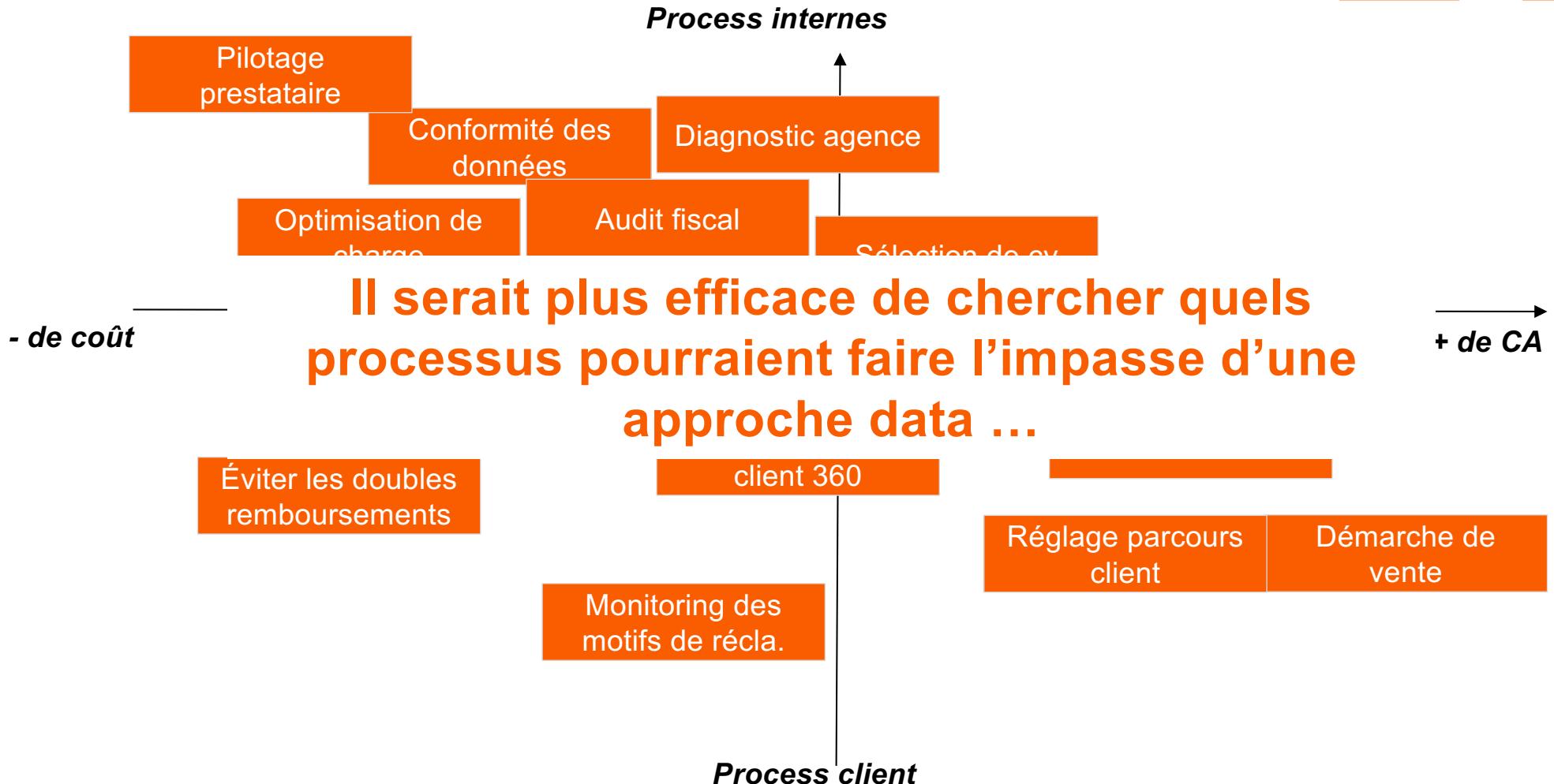


Identifier les motifs corrélés aux pics de charge : les anticiper

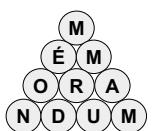




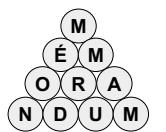
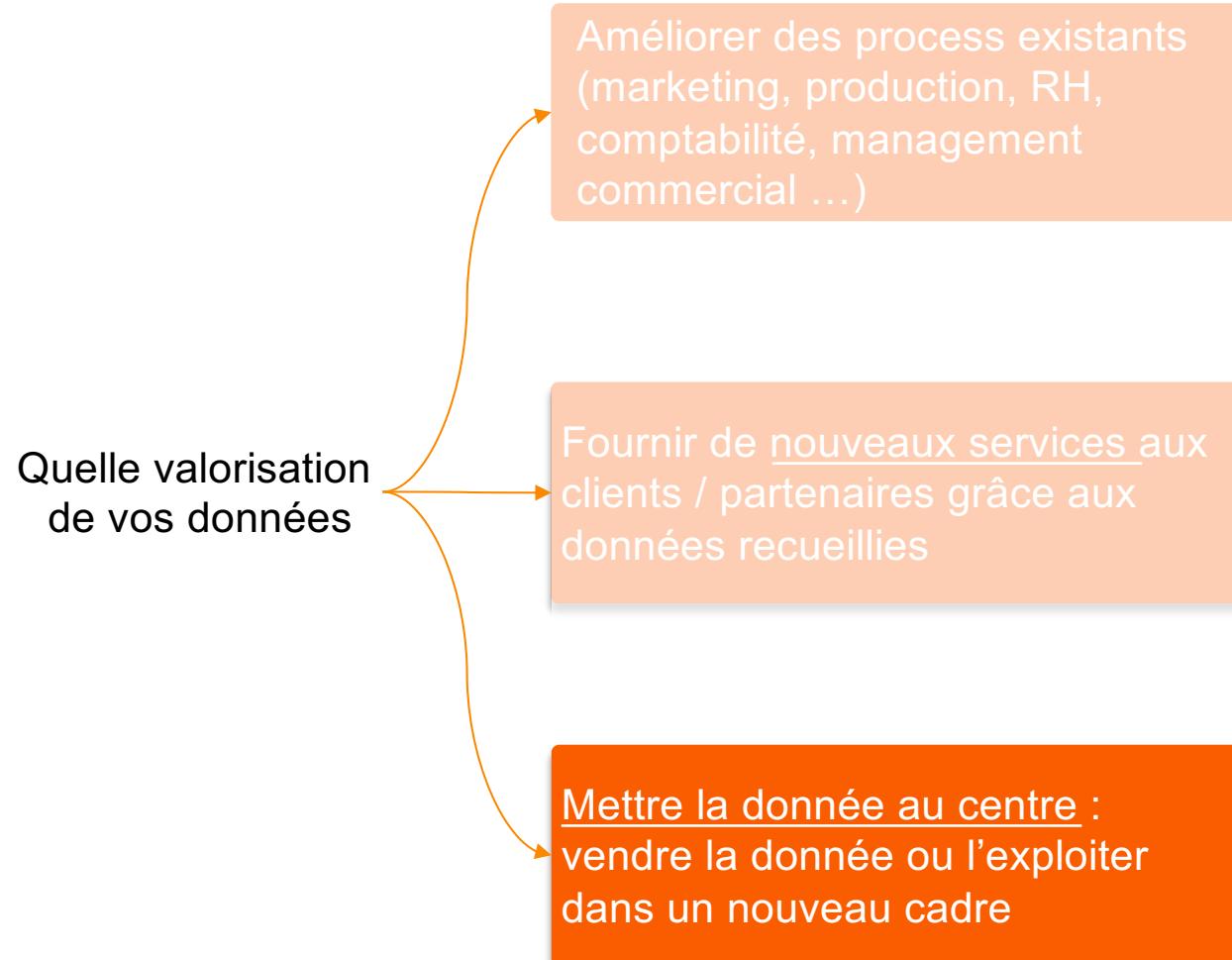
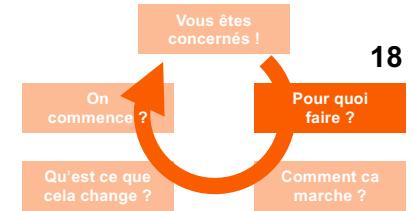
Des cas d'usage qui touchent à tous les process de l'entreprise



Pas des projets data mais une approche data des projets



La donnée peut servir encore dans des contextes très nouveau



D'une stratégie à l'autre : inverser la logique proposer des services qui produisent des données

Vous êtes concernés !

19

On commence ?

Pour quoi faire ?

Qu'est ce que cela change ?

Comment ça marche ?

Modèle initial

- Améliorer des process existants

- Fournir de nouveaux services aux clients

- Mettre la donnée au centre : la vendre

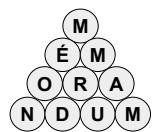


Données

Variante

- Améliorer des process existants

- Fournir de nouveaux services



Vous êtes concernés !

On commence ?

Pour quoi faire ?

Qu'est ce que cela change ?

Comment ça marche ?

Comment ça marche ?

Poser une question : aussi délicat que de trouver une réponse

Une question ...

.... Qui a besoin d'être précisée

Exemple

Repérer les churners ?

Churn ? résiliation administrative ? A d'usage d'un service, ... ?

Tous ? Ou uniquement un certain

sûr :
évaluation de la
change la réponse

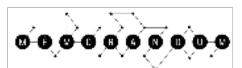
Peut être ch

- Optimiser
- Comprendre
- churners : ag
-

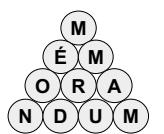
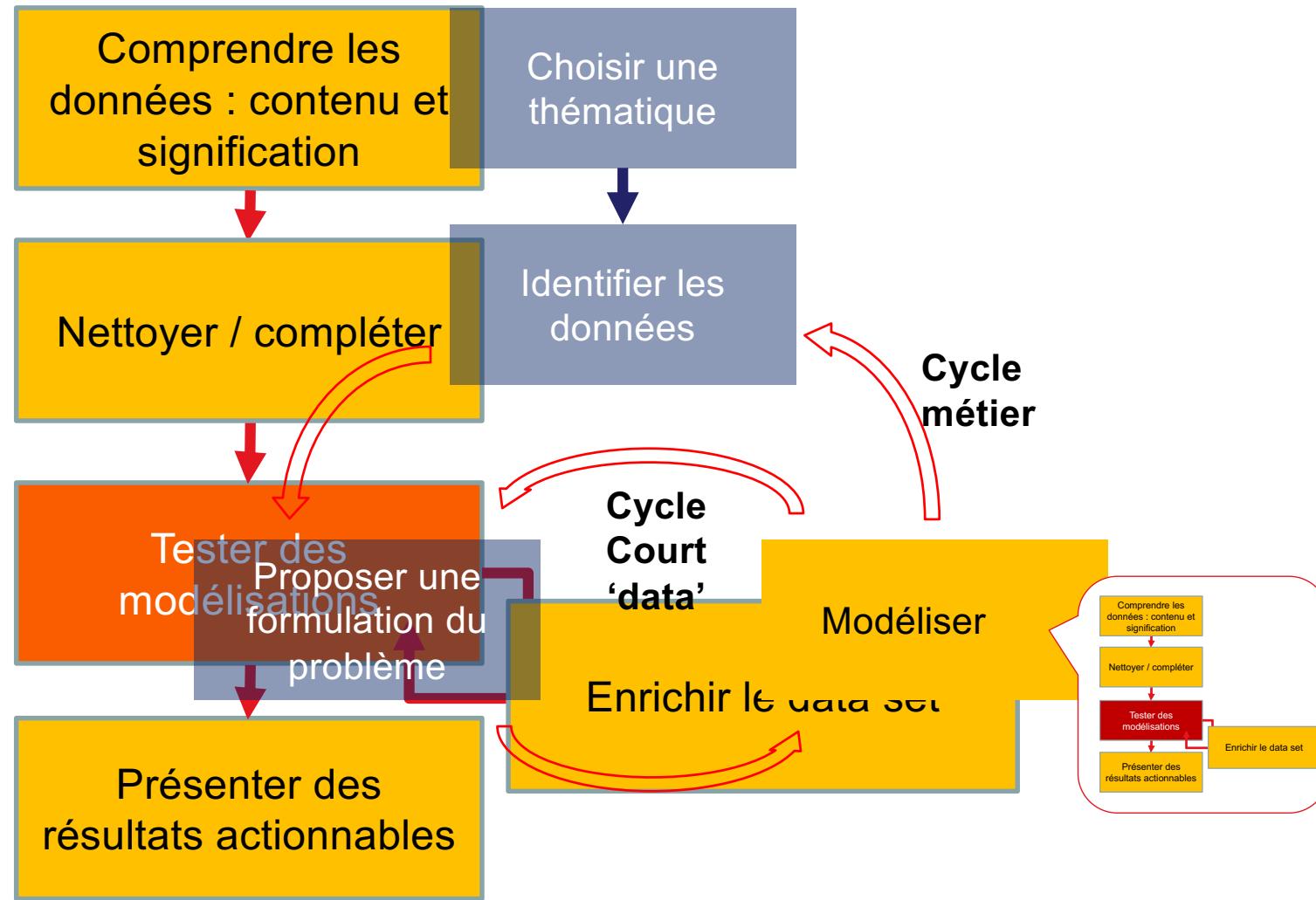
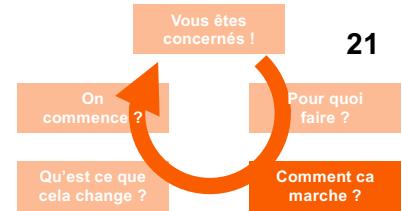
Pistes pour mieux cerner ce qu'on cherche

- Que fera t on avec la réponse ? (courrier, appel, offre, ...)
- Comment mesurera t on la performance ?
- Questions de base :

- churn = choix rationnel du client ou résultat d'un démarchage
- Combien de temps avant le client prend il sa décision



Comment ça marche ? faire progresser question et réponse par itération



Vous êtes
concernés !

On
commence ?

Pour quoi
faire ?

Qu'est ce que
ça change ?

Comment ça
marche ?

Comment ça marche ? Avec quelles compétences ?

Comprendre
le métier,
Préparer les
données ?

Produire un
algorithme

Machine Learning



what society thinks I
do



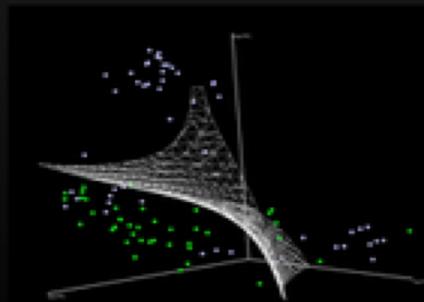
what my friends think
I do



what my parents think
I do

$$\begin{aligned} L_r &= \|\mathbf{w}\|^2 + \sum_i \alpha_i y_i (\mathbf{x}_i \cdot \mathbf{w} + b) + \sum_i \alpha_i \\ \alpha_i &\geq 0, \forall i \\ \mathbf{w} &= \sum_i \alpha_i \mathbf{x}_i, \sum_i \alpha_i y_i = 0 \\ \nabla \hat{g}(\theta_t) &= \frac{1}{n} \sum_{i=1}^n \nabla \ell(x_i, y_i; \theta_t) + \nabla r(\theta_t), \\ \theta_{t+1} &= \theta_t - \eta_t \nabla \ell(x_{t(0)}, y_{t(0)}; \theta_t) - \eta_t \cdot \nabla r(\theta_t) \\ E_{t(0)}[\ell(x_{t(0)}, y_{t(0)}; \theta_t)] &= \frac{1}{n} \sum_i \ell(x_i, y_i; \theta_t). \end{aligned}$$

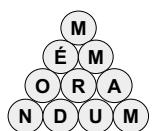
what other programmers
think I do



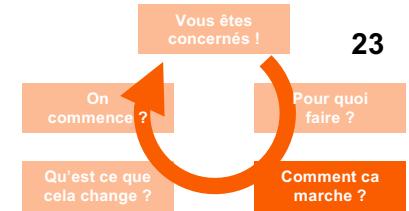
what I think I do

```
>>> from scipy import SVM
```

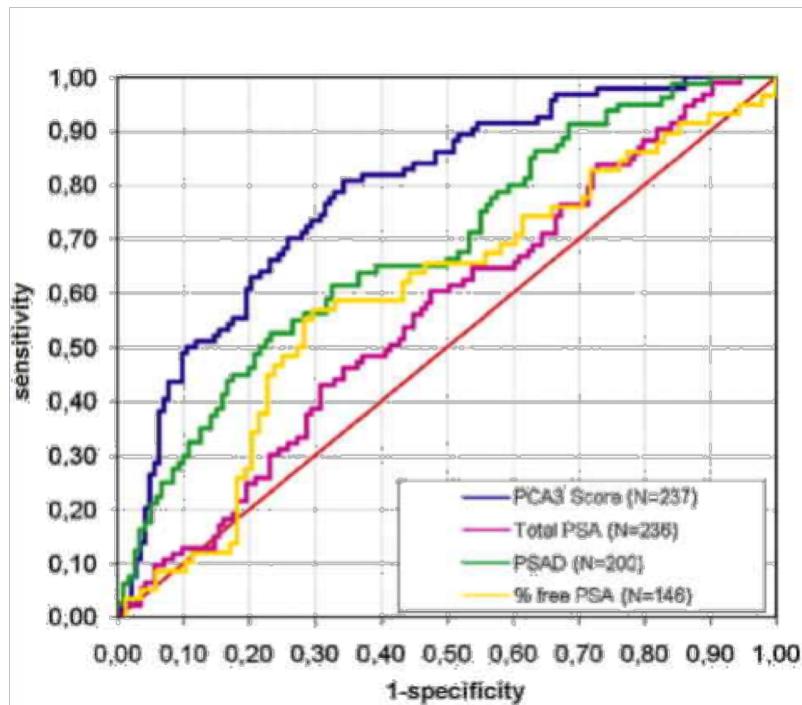
what I really do



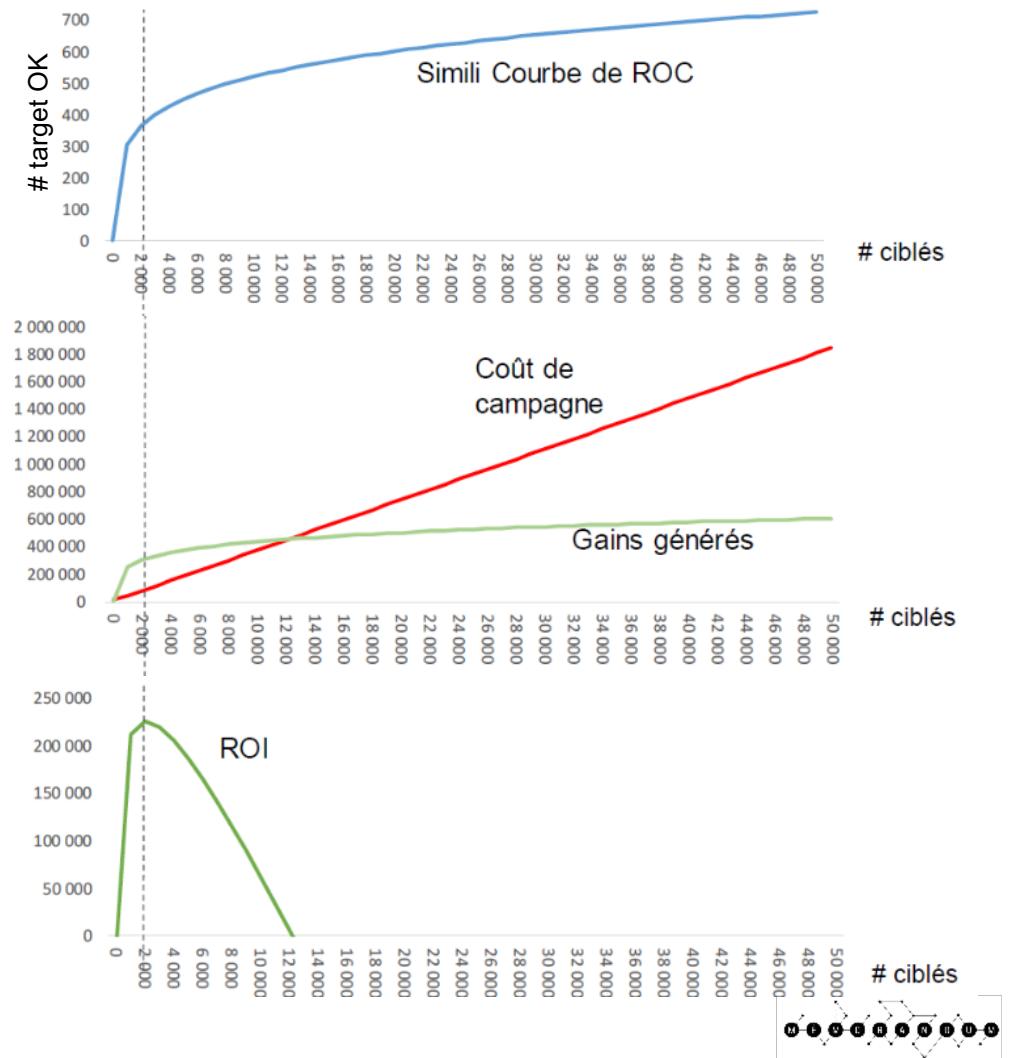
Comment ça marche ? Transformer le résultat statistique en scénario métier



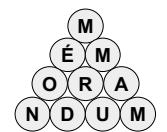
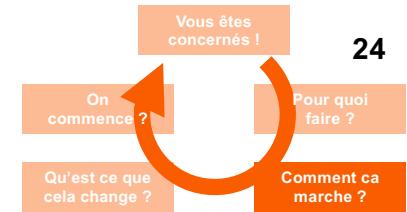
Une courbe ROC ?



Ou une courbe ROI ?



Comment ça marche ? Planter un projet



Vous êtes concernés !

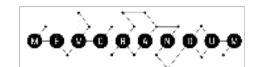
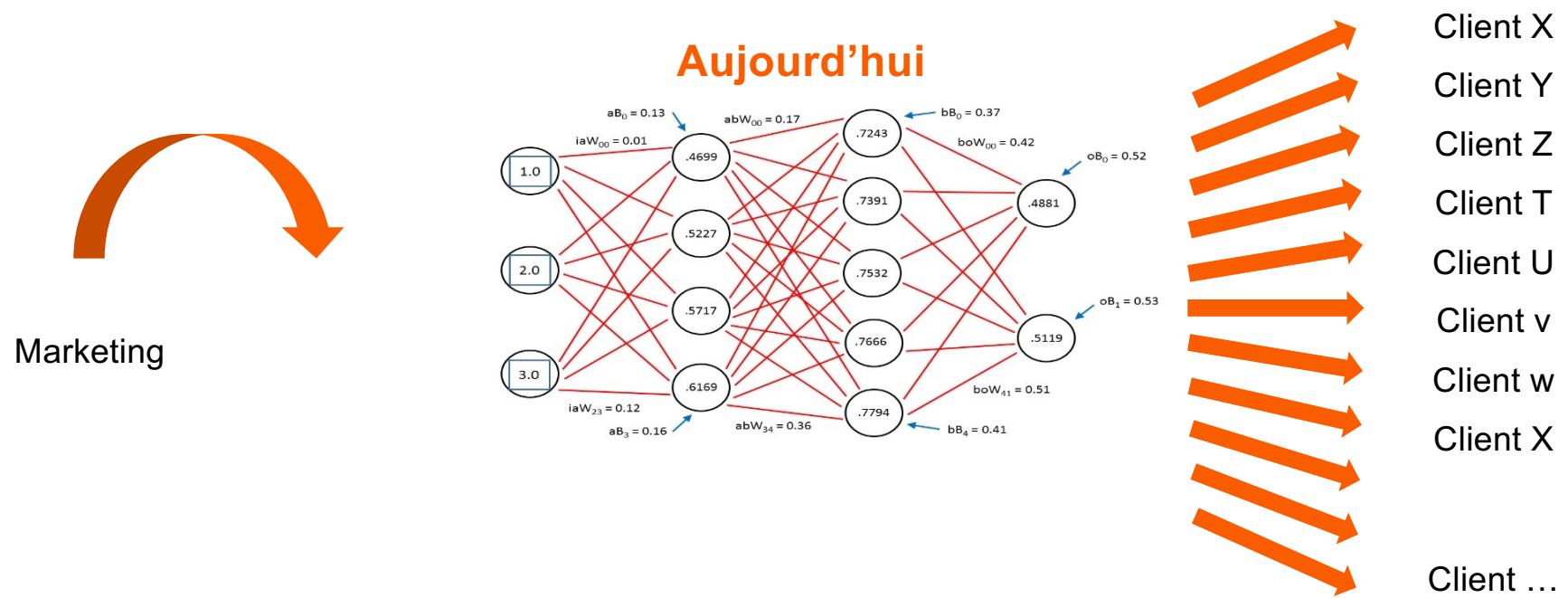
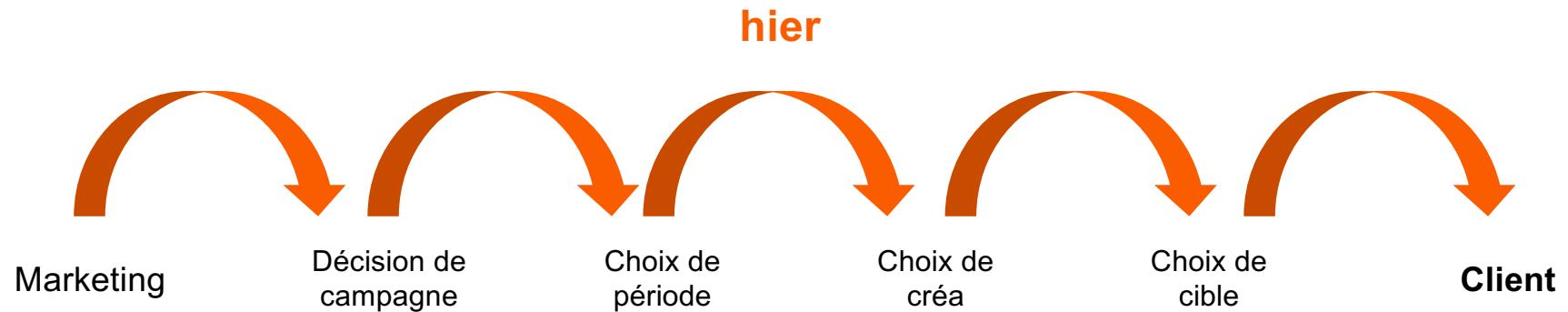
On commence ?

Pour quoi faire ?

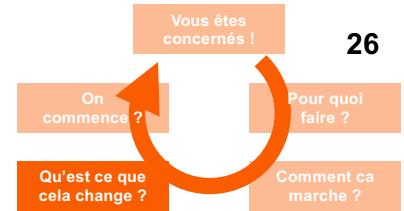
Qu'est ce que cela change ?

Comment ça marche ?

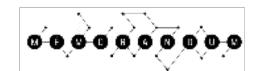
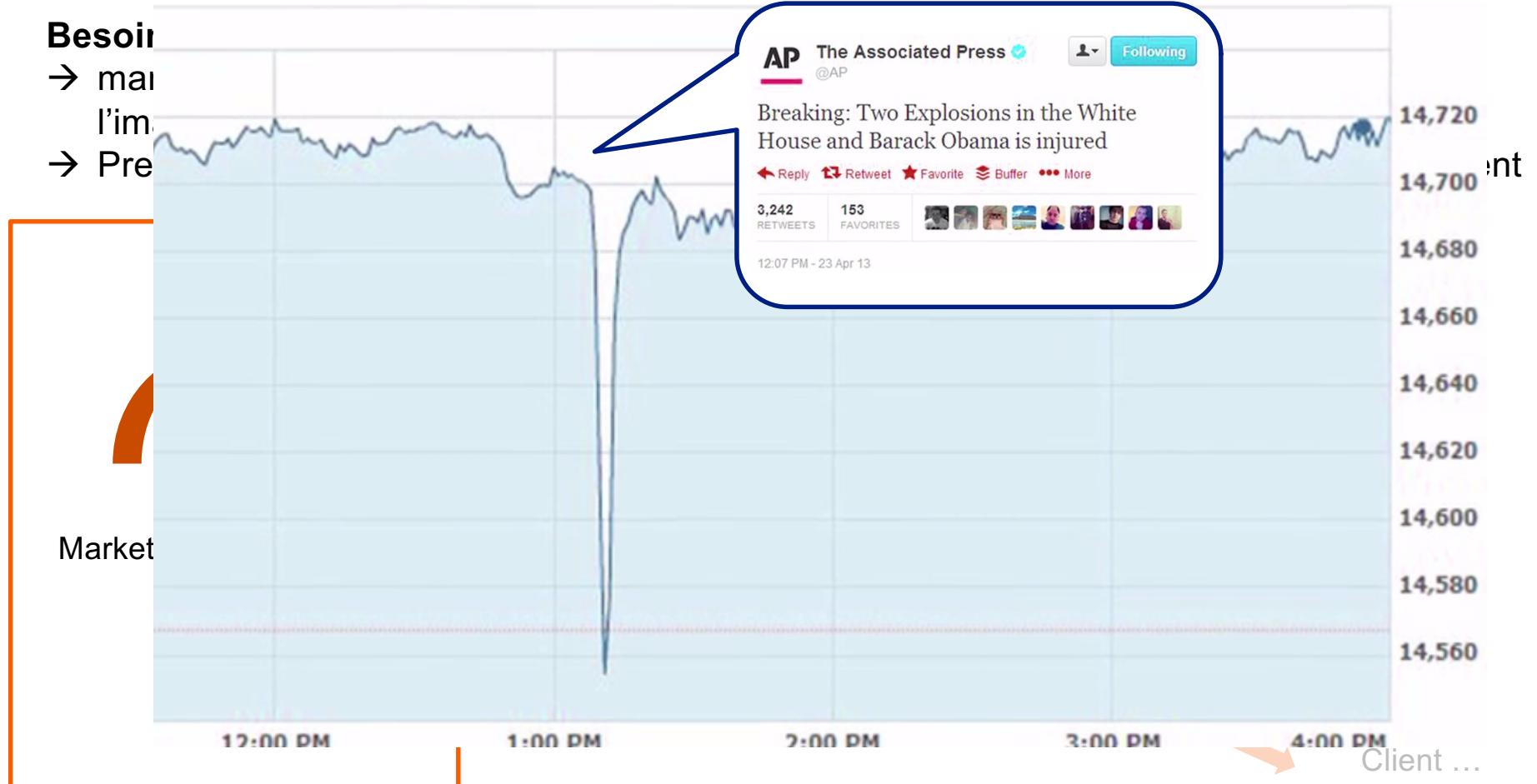
Un circuit de décision transformé



Un circuit de décision transformé → un besoin de cadrage accru en amont



Besoin
→ mai
l'im
→ Pre



Vous êtes concernés !

On commence ?

Pour quoi faire ?

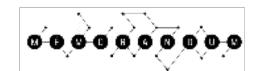
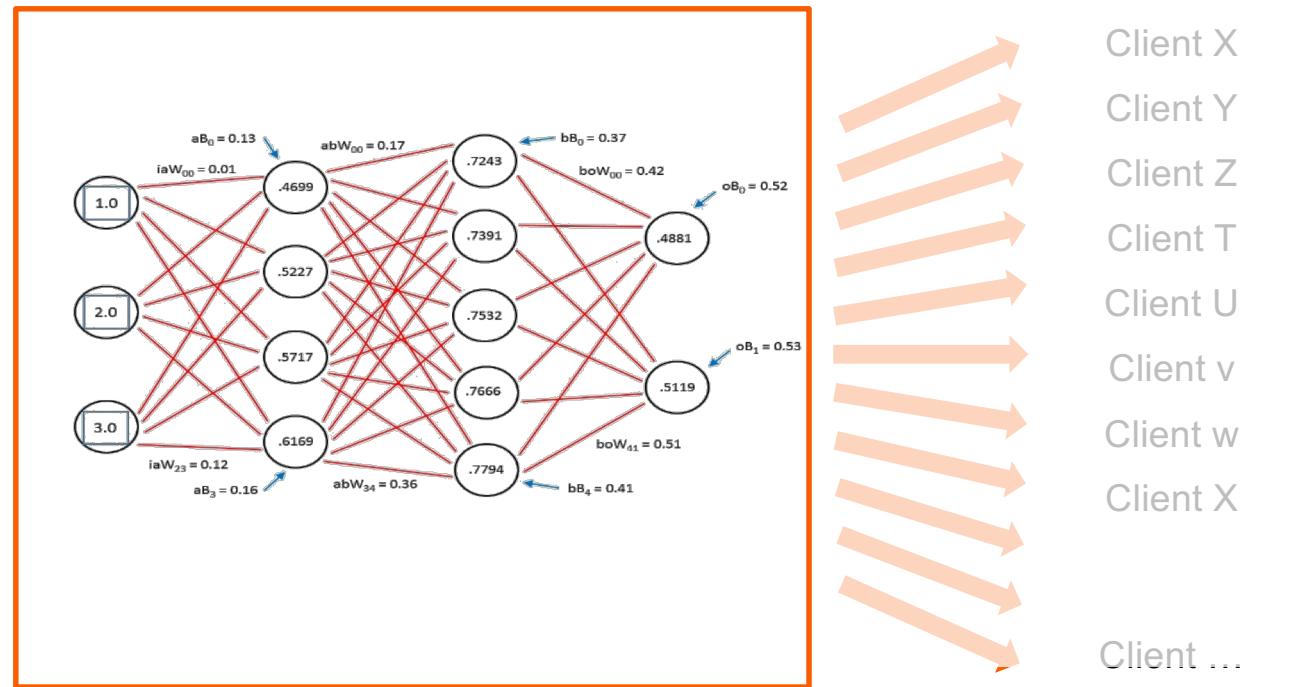
Qu'est ce que cela change ?

Comment ça marche ?

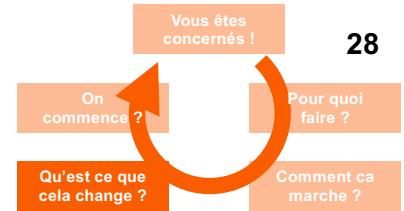
Un circuit de décision transformé → une culture data essentielle

Penser données

- Savoir identifier quand on manque de données, et qu'on a des données biaisées (exemple score crédit) et prendre les moyens de les acquérir : phase d'apprentissage, nouveaux services, crowd sourcing, open data, 2nd party, ...
-

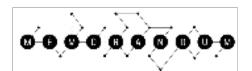
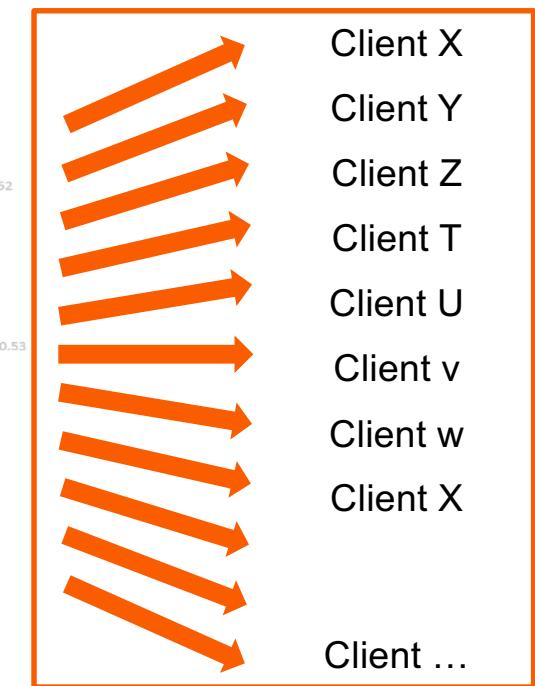
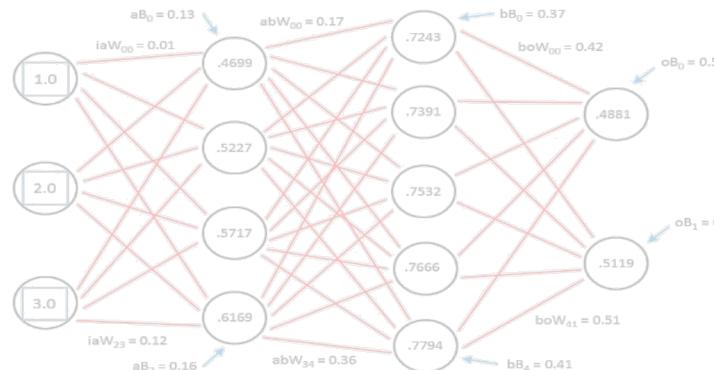


Un circuit de décision transformé → un pouvoir accru de ceux qui sont au contact du client



Pouvoir des vendeurs et des conseillers

- Les acteurs en bout de chaîne (vendeurs, conseillers téléphoniques, ..) ont un pouvoir croissant : plus informé sur leurs clients, plus



Vous êtes concernés !

On commence ?

Pour quoi faire ?

Qu'est ce que cela change ?

Comment ça marche ?

Managez vos données comme vous managez vos équipes

Vos données sont rangées comme vos organisations

| Commercial | Marketing | RH | Compta | Production |
|------------|-----------|----|--------|------------|
| | | | | |

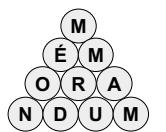
Vos données doivent être animées comme votre management : transverse



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

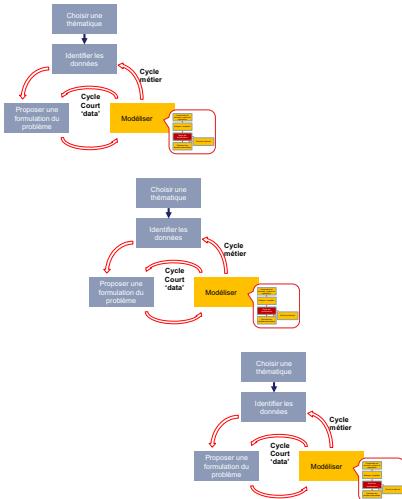
- Être garant de la disponibilité de ses données
- réutiliser les données des autres directions

Vos données ne travailleront ensemble que sous l'impulsion de votre Comité de Direction



Chaque projet data laisse des bénéfices durables

Gains cumulatifs sur les process travaillés : meilleur ciblage, correction des éléments de diagnostic, ..



Développement de la culture data des managers

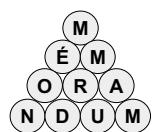
- Un travail de long terme, d'importance croissante chaque année qui passe
- No free lunch theorem : importance du contexte
- Savoir lire un résultat, son domaine de fiabilité, ...

Prise de conscience accrue de l'interdépendance de toutes les directions de l'entreprise

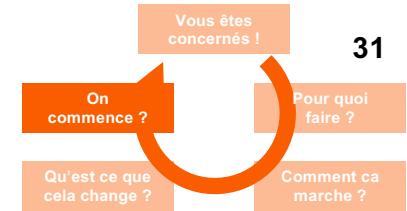
- Les données de chacun influencent les process de tous
- Un chantier RH, qui fait prendre conscience que l'entreprise est un tout organique :

Sauvegarde du patrimoine de données :

- les fuites ne sont détectées que quand on ouvre le robinet. Plus vous tardez plus vous appauvrissez
- Chaque projet permet d'identifier des lacunes dans le management des données : inputs pour les prochaines évolutions SI des systèmes



Commencer ?



**La difficulté n'est pas de voir ce que personne n'a jamais vu
mais de penser comme personne n'a jamais pensé
au sujet de choses que tous voient**

Arthur Schopenhauer

