



Introduction Big Data & IA

Alison PATOU
Patou.alison@gmail.com



Programme

- Introduction
- What is data ?
- Big Data
- Machine Learning, Data Science, IA
- Focus sur l'Intelligence Artificielle



A savoir

- Tout est sur mon GitHub :
<https://github.com/apatou>
- (L'essentiel à retenir du cours, les dataset, exercices, ...)
- Merci de m'envoyer à chaque fin de séance vos TPs : patou.alison@gmail.com
- 1 projet en groupe

1

Introduction

Objectifs de ce cours

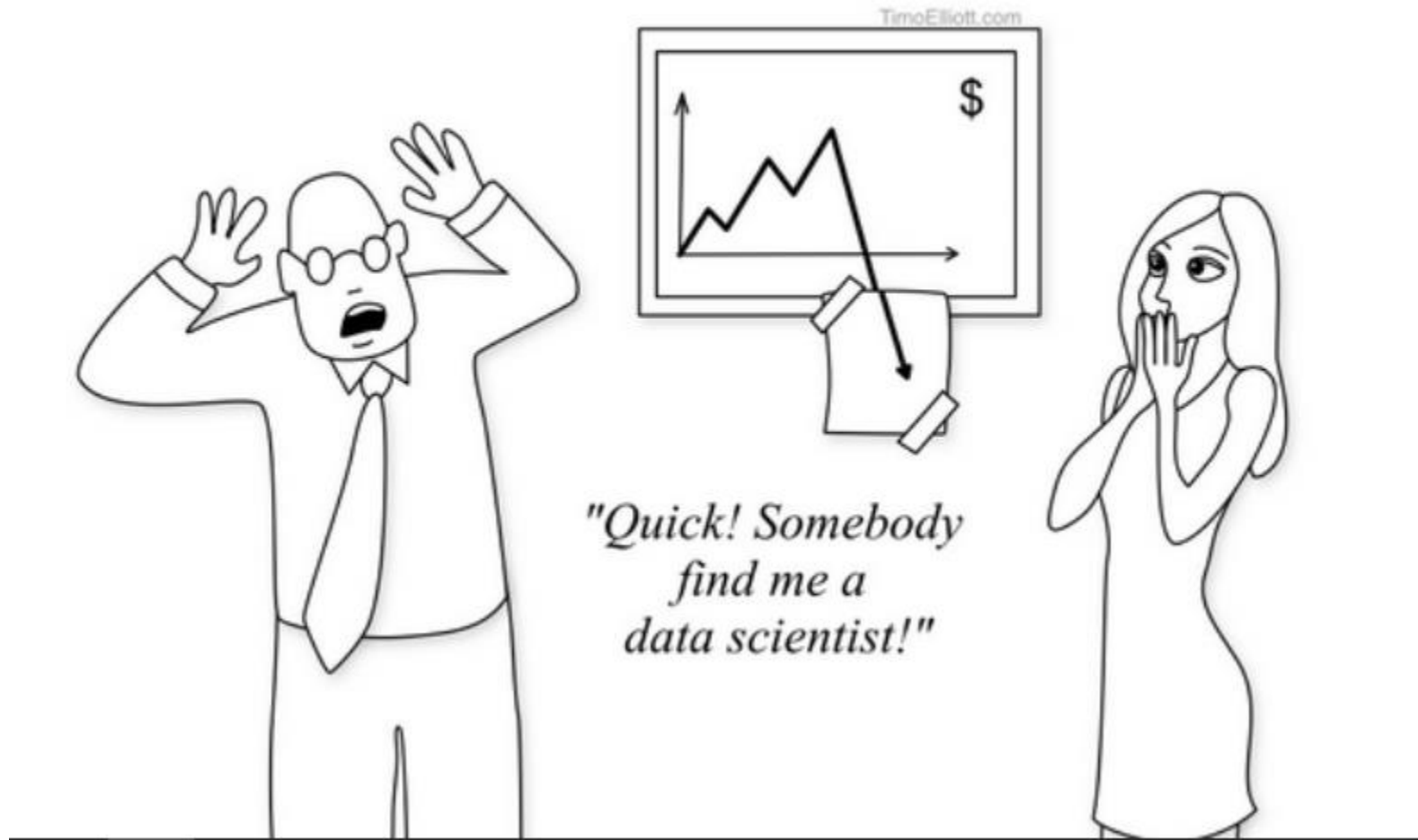
Savoir expliquer les concepts du Big Data et de l'IA

Comprendre les enjeux de l'IA

Convergence entre IA&Big Data

La notion éthique

La **Data Science** et le **Big Data** sont devenus depuis quelques années la solution à tous les problèmes.



Cela vous
évoque quoi?



Big Data



Data Science



Intelligence
Artificielle

Cela se passe aujourd'hui...



24 PO (24 000 000 000 000 000 000) / j
(environ 30 000 disques durs de PC)



10 000 000 photos / jour
3 000 000 000 'like' / jour



2015 : 500 000 000 tweets / jour



800 000 000 utilisateurs / mois

Cela se passe aujourd'hui...



24 PO (24 000 000 000 000 000 000) / j
(environ 30 000 disques durs de PC)



10 000 000 photos / jour
3 000 000 000 'like' / jour



2015 : 500 000 000 tweets / jour



800 000 000 utilisateurs / mois

Explosion du volume des données numériques

Particuliers



Entreprises



Objets connectés



Open Data

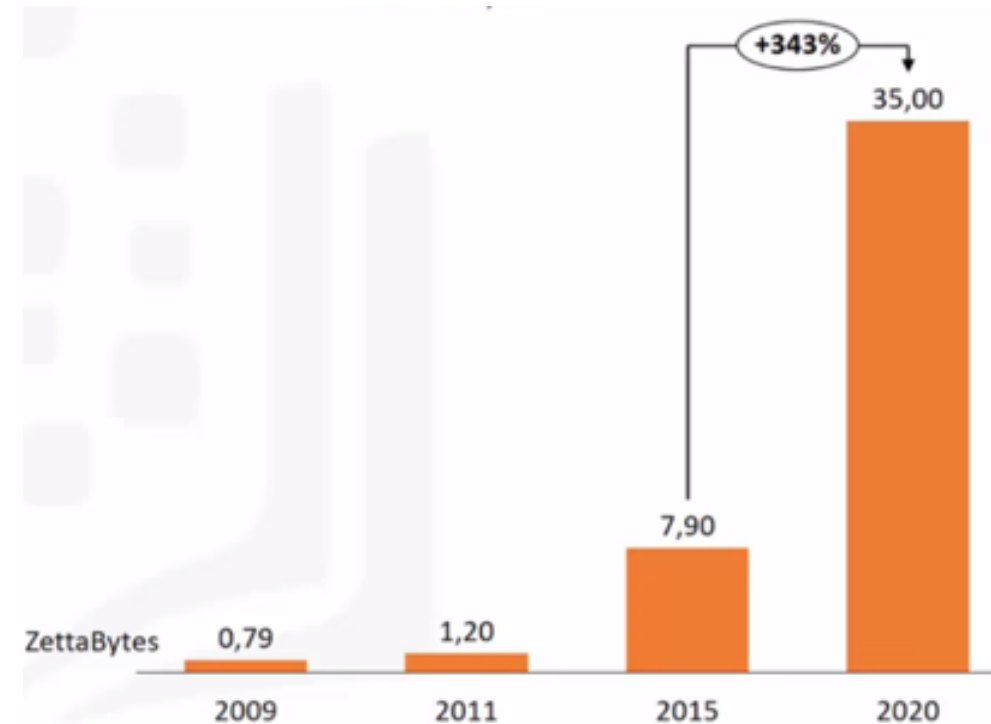


Internet



Des faits :

- Plus de données ont été créées durant les deux dernières années que dans toute l'histoire de l'humanité auparavant
- En 2020, 1,7 megabytes de nouvelles données seront créées chaque seconde par personne dans le monde
- En 2020, les données que nous créons et copions atteindront 35 zettabytes, alors que nous sommes actuellement à environ 7,9 zettabytes (+343%)



2

What Is Data

A word cloud featuring various French terms in different colors and sizes, arranged in a diagonal shape from top-left to bottom-right. The words include:

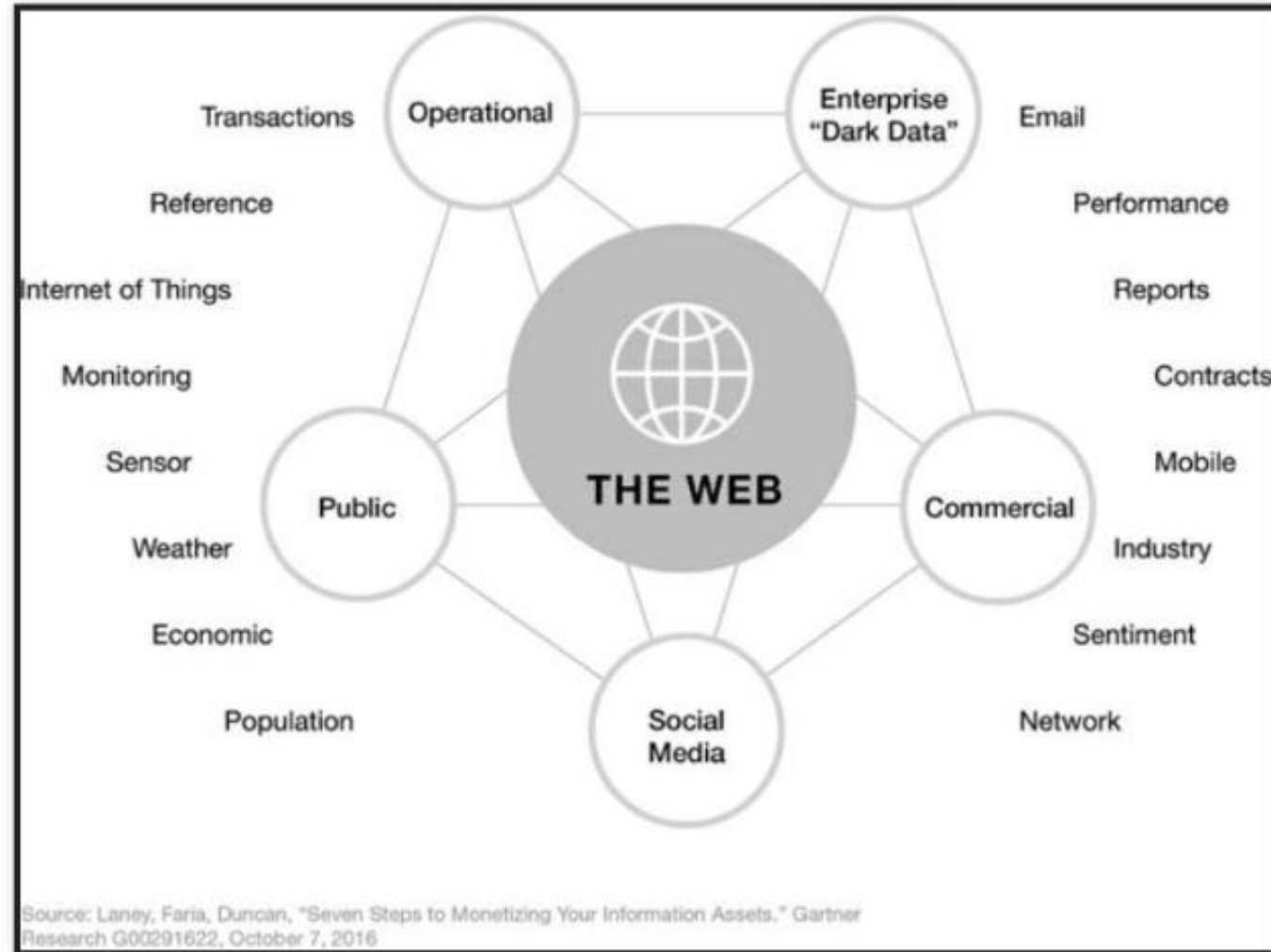
- couleur (blue)
- temperature (brown)
- longitude (large, brown)
- relation (blue)
- fichier (blue)
- duree (purple)
- age (purple)
- date (blue)
- son (teal)
- prix (large, light blue)
- note (brown)
- liste (large, blue)
- latitude (brown)
- adn (green)
- commentaire (blue, vertical)

On pense maintenant aux données au sens large

Les données ne sont plus seulement numériques, ni des chiffres dans un fichier Excel.

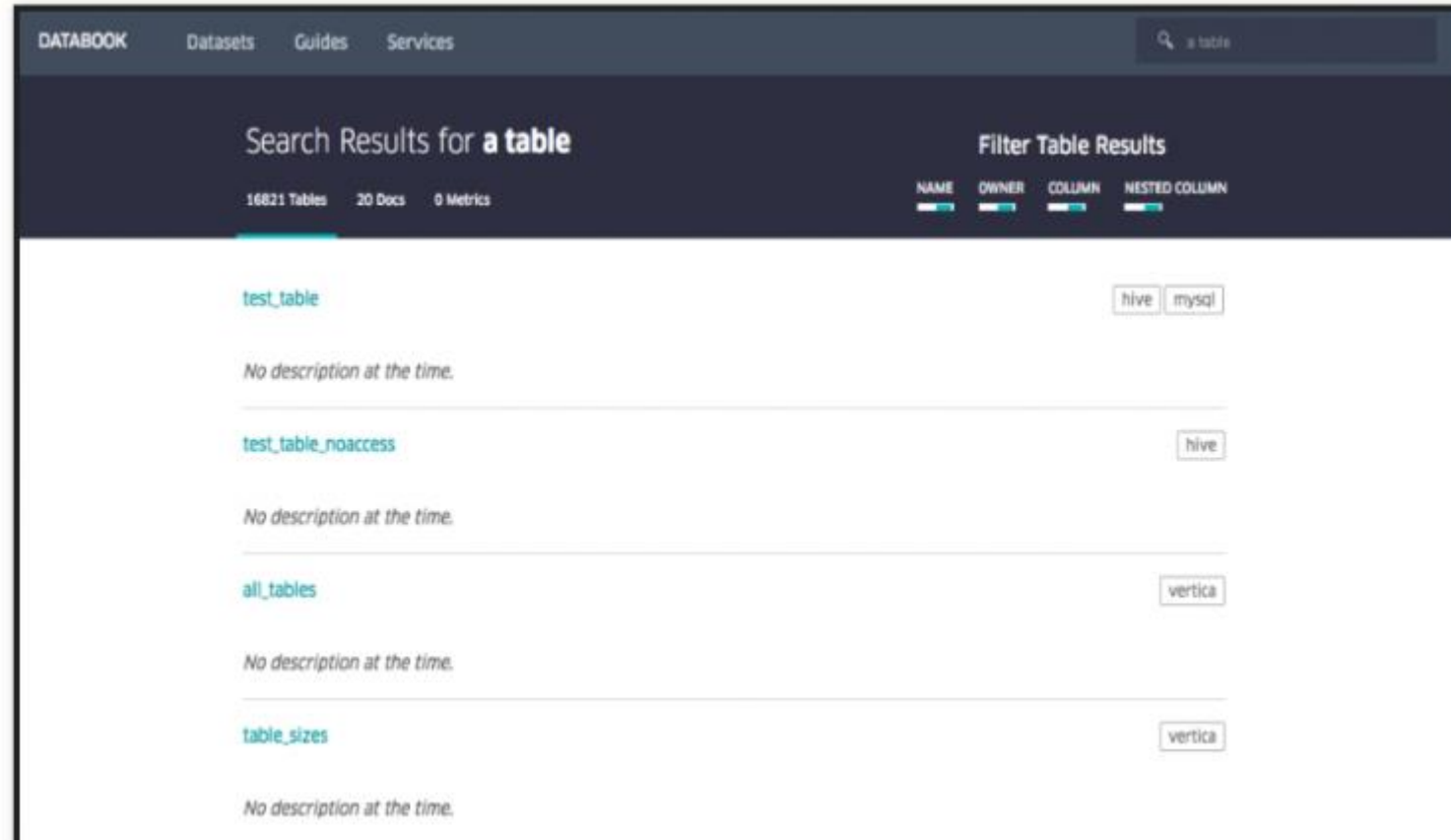
- Les images sont des données
- La langue est une données
- Les relations sont des données
- Les préférences, les états émotionnels, l'écriture sont des données
- ...

Où se trouvent les données ?



Où se trouvent les données ?

Peu d'entreprises ont mis en place une "gestion des données" pour s'assurer que les données sont connues, répertoriées, de qualité et exploitables. Une exception positive intéressante: Uber a mis en place un "databook" stockant les métadonnées relatives aux jeux de données dans l'entreprise:



De quoi s'agit-il ?



De quoi s'agit-il ?



- Support numérique
- Deux types de données : texte & image
- Le texte est formaté en texte brut
- Le codage du texte standard (UTF-8)
- Ce tweet fait partie d'une liste de tweets. La liste représente la structure de données de l'ensemble de ces données.
- Ce tweet est stocké sous forme d'image (.png)

Taille de données

Sources: McKinsey report – Digital era, data growth, estimates and, Gartner, IDC estimates

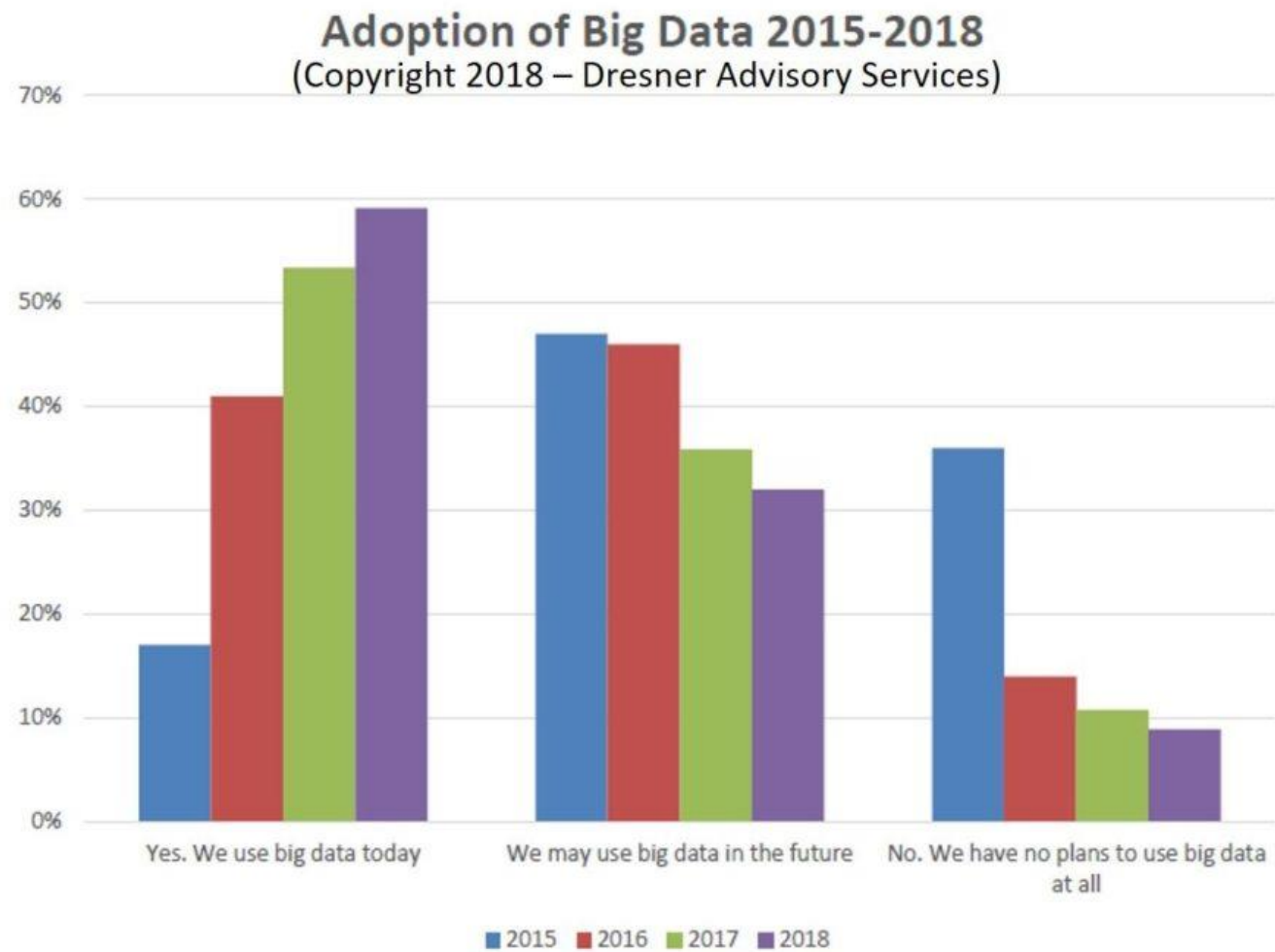




3

Big Data

Le big data devient un indispensable



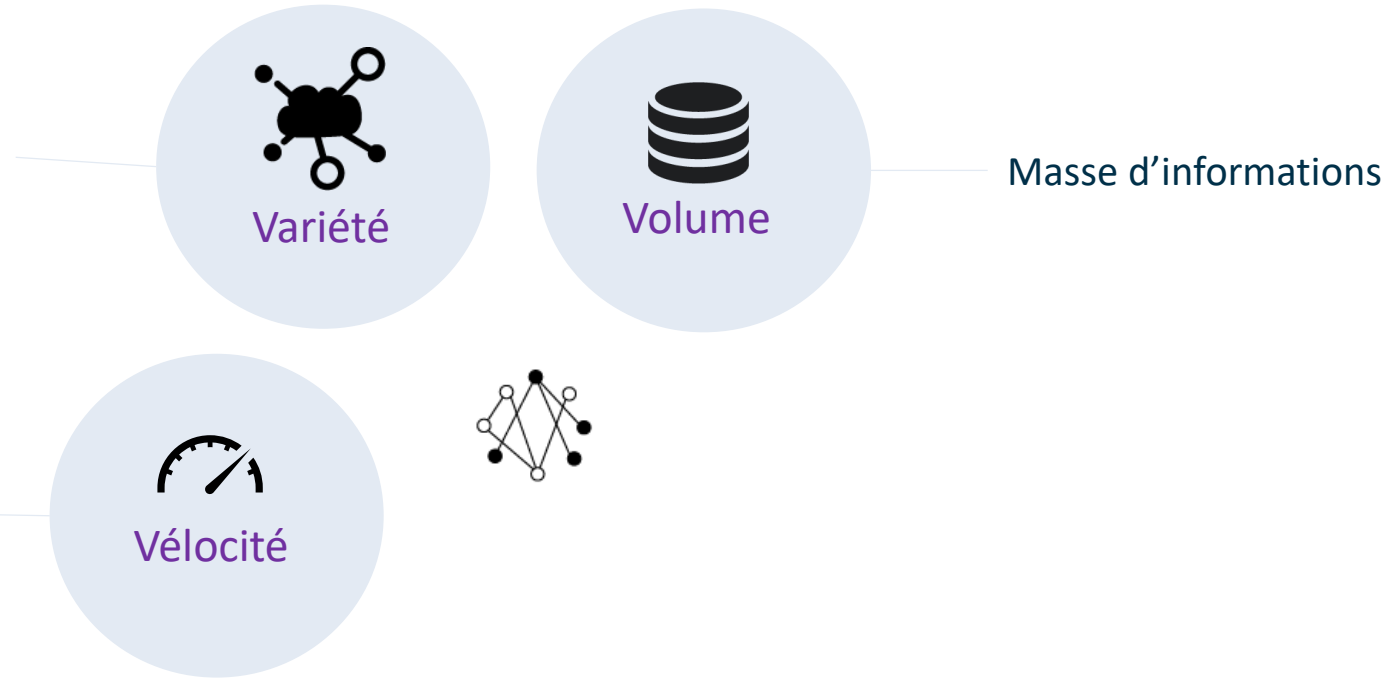
Le big data : Les 3V

Dans la littérature, on parle donc des **3 V** :

Différents types de données :

- Structurées (transactions de vente par exemple),
- Non structurées (images, vidéos, voix, textes,...)

Rapidité de l'élaboration et du déploiement de nouvelles données



V

VOLUME DES DONNÉES



Le volume des données stockées est en **pleine expansion** : les données numériques créées dans le monde seraient passées de 1,2 zettaoctet par an en 2010 à 2,8 zettaoctets en 2012 et s'élèveront à 40 zettaoctets en 2020.

À titre d'exemple, Twitter générerait en janvier 2013 7 téraoctets de données chaque jour et Facebook 10 téraoctets. En 2014, Facebook Hive générerait 4 000 To de data par jour.

V

VARIÉTÉ DES DONNÉES



Il ne s'agit **pas de données relationnelles traditionnelles**. Ce sont des données complexes provenant du web au format texte et images ou vidéo, de capteurs, d'objets connectés...

V

ÉLOCITÉ :



C'est la fréquence à laquelle les données sont à la fois générées, capturées, partagées et mises à jour. On parle ici de « **temps réel** »



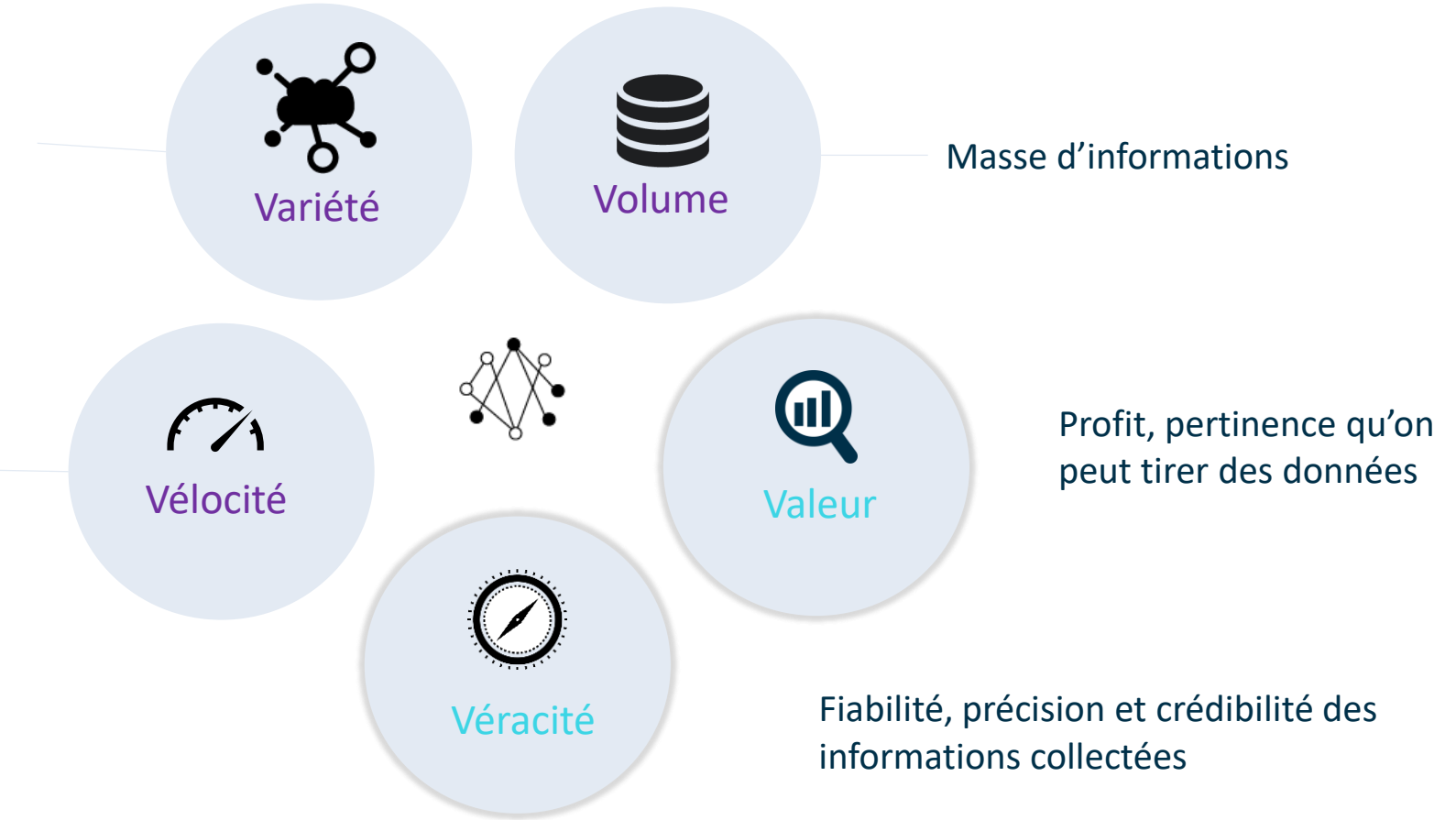
Big data : Les 3V deviennent les 5V

La notion des 3V peut-être étendue à **5V**, si on rajoute deux dimensions :

Différents types de données :

- Structurées (transactions de vente par exemple),
- Non structurées (images, vidéos, voix, textes,...)

Rapidité de l'élaboration et du déploiement de nouvelles données





VERACITE



La véracité ou la fiabilité des données est notamment menacée par les comportements déclaratifs (sur formulaires) ou encore les saisies manuelles (création de cartes fidélités par exemple), par la diversité des points de collecte, la multiplication des formats de données et par les faux profils innombrables sévissant sur Internet.

Des études ont montré qu'un leader business sur trois n'a pas confiance en ses données !



VALEUR



Dans un contexte d'infobésité, il s'agit d'être capable de se concentrer sur les données ayant une réelle valeur et étant actionnables



Définition du big data

Qu'est ce que le Big Data ?



Le big data, littéralement « grosses données », parfois appelées données massives, désigne des ensembles de données devenus si volumineux et si variés qu'ils dépassent l'intuition et les capacités humaines d'analyse et même celles des outils informatiques classiques de gestion de base de données ou de l'information.



Conséquence :

Pour répondre aux problématiques big data il est donc nécessaire de repenser l'architecture, les supports de stockage et les outils.

Définition du big data

- Il n'y a pas de « taille minimale » pour être dans un contexte Big Data.
- Il s'agit plutôt d'une notion relative : ce sont des données massives que les systèmes informatiques, outils actuels, ont du mal à stocker/gérer les données qu'elles doivent traiter.
- Le Big Data est donc une notion relative.

Définition du big data

Ce qui était considéré comme Big Data dans les années 2000 est considéré comme « small data » aujourd'hui, car nous avons une meilleure puissance de stockage des données, de calcul et de traitement.

Définition du big data

Et c'est sans fin : à mesure que les systèmes informatiques s'améliorent pour faire face aux « big data actuels », les données sont générées dans des volumes encore plus importants, ce qui nécessite de nouveaux progrès / innovations pour les gérer.

Variétés des sources

Démocratisation des objets dits « smart »



Le big data en entreprise

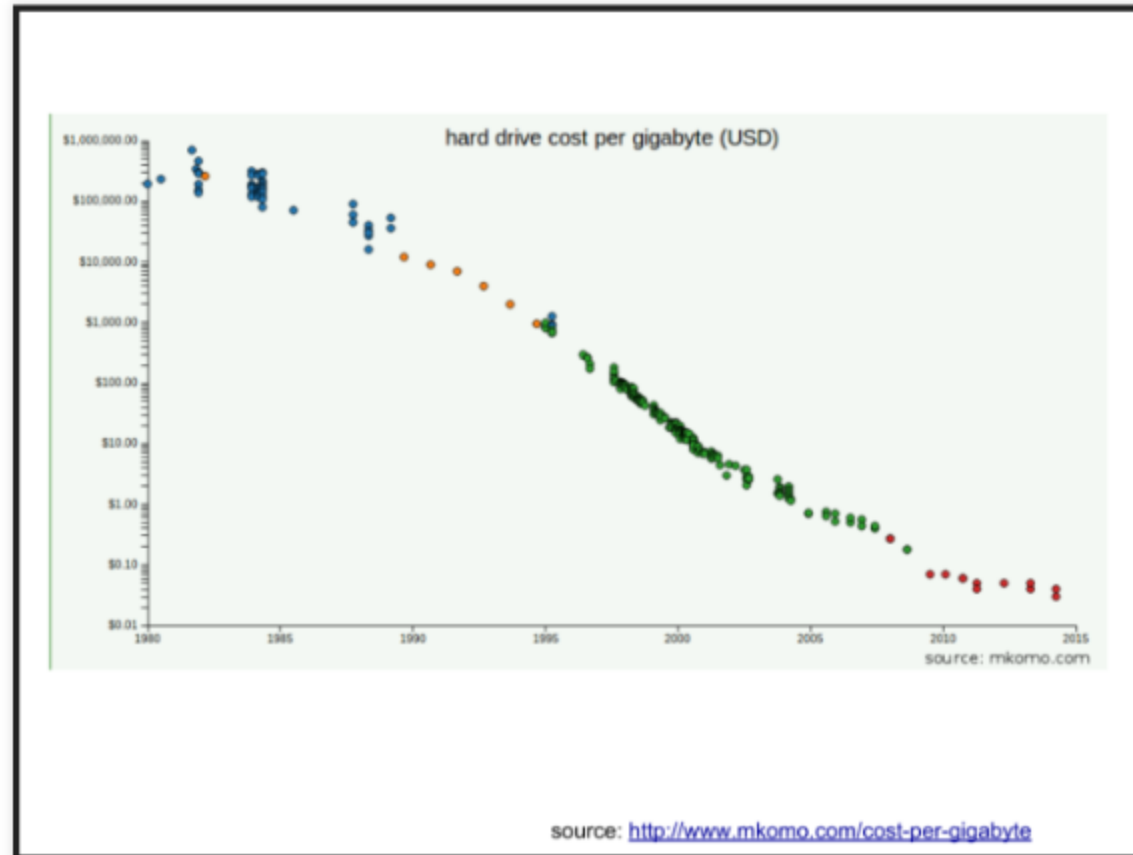
Les entreprises ont compris que leurs données avaient de la valeur et cherchent à les exploiter



En 2016 :

- ✓ 51% des entreprises sont en train de faire du Big Data (POC, industrialisation, ...)
- ✓ 41% en sont qu'au stade d'évaluation, d'apprentissage

La démocratisation du big data en entreprise



Le coût du stockage est de moins en moins cher

La démocratisation du big data en entreprise



La mise à disposition d'Open Data pour enrichir

Les Open Data

“

”

Une Open Data (donnée ouverte) est une **donnée numérique** d'origine publique ou privée et diffusée de manière à garantir son libre accès et sa réutilisation par tous, sans restriction technique, juridique ou financière.

Exemples : données sur le transport , la cartographie, les statistiques, la géographie, la sociologie, l'environnement, etc.

Les Open Data

Depuis les années 2000, il y a un élan gouvernemental de mettre à disposition les données :

Projet data.goven 2009 aux USA

Directive européenne en 2003 sur la réutilisation des données publiques

Etalab en France (2011) est chargée de créer et alimenter le portail de données publiques ouvertes data.gouv.fr.

Données de la métropole lyonnaise : <https://data.grandlyon.com/accueil>

Bénéfices pour le secteur public :

Transparence, réduction de coûts, amélioration des services

Bénéfice économique:

Accès aux données, notamment pour les PME/TPE

La démocratisation du big data en entreprise



La mise à disposition d'outils Open Source

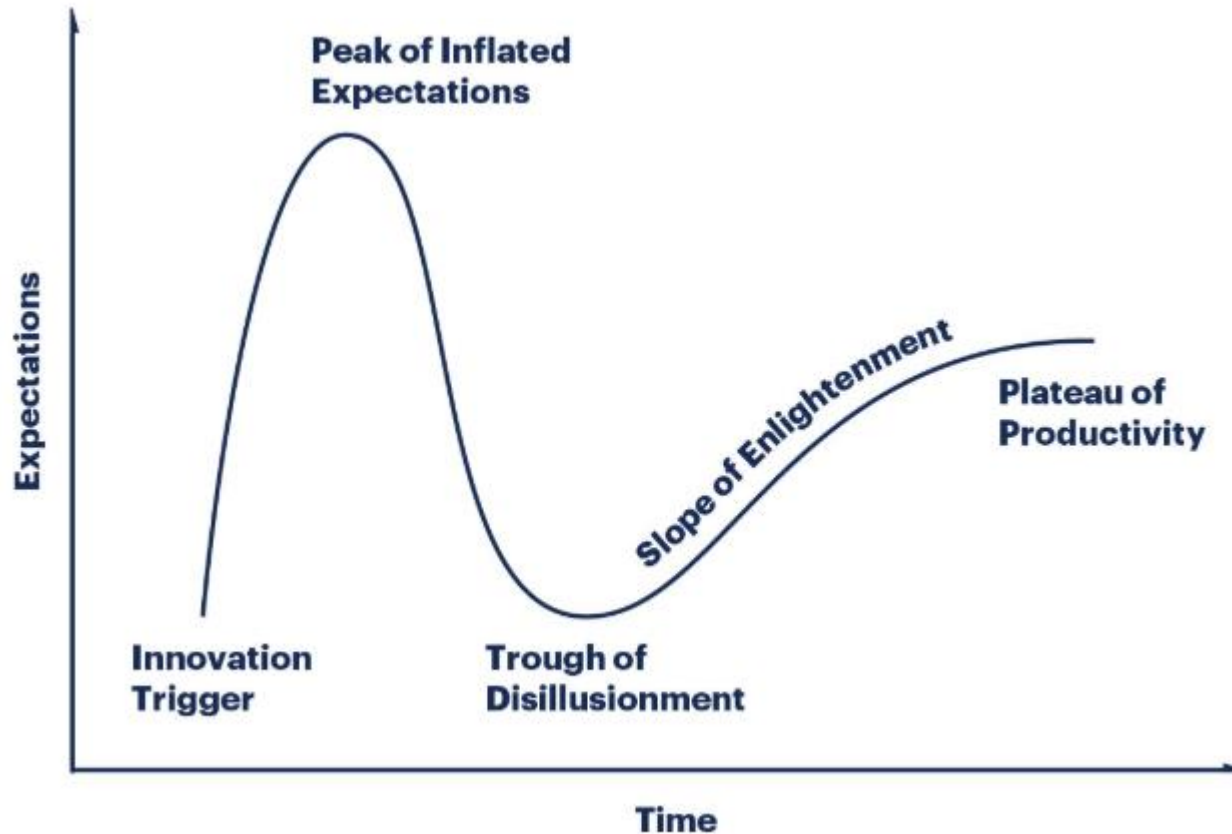
Promesses et attentes exagérées

On parle beaucoup de projets Big Data en entreprise.

Or :

- Le « Big Data » pour les PME n'est pas le même que pour les entreprises du CAC40
- On met en place des projets Big Data (infrastructure, etc.) pour des données qui ne sont pas toujours des « Big Data »
- La majorité des données des entreprises sont structurées.

Promesses et attentes exagérées



Gartner Hype Cycle

Innovation Trigger : Le potentiel d'une technologie perçe et l'intérêt des médias augmentent significativement l'intérêt sur le sujet

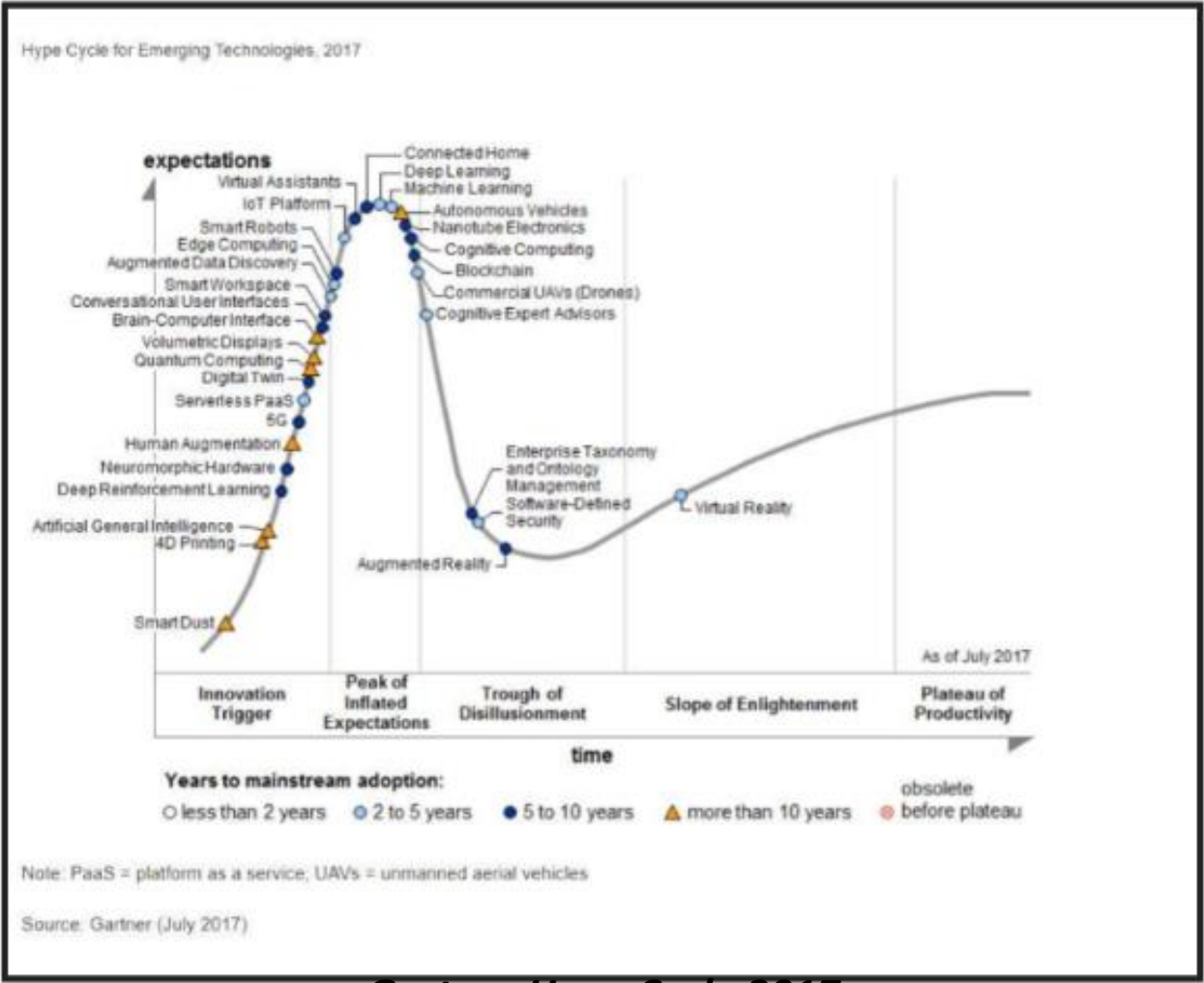
Peak of Inflated Expectations : La publicité précoce produit de nombreuses success story, souvent accompagné des scores d'échecs.

Trough of Desillusionment : L'implémentation et les REX sont peu concluants.

Slope of Enlightenment : La majorité sont au stade de POC, les entreprises conservatives rappellent d'être prudents

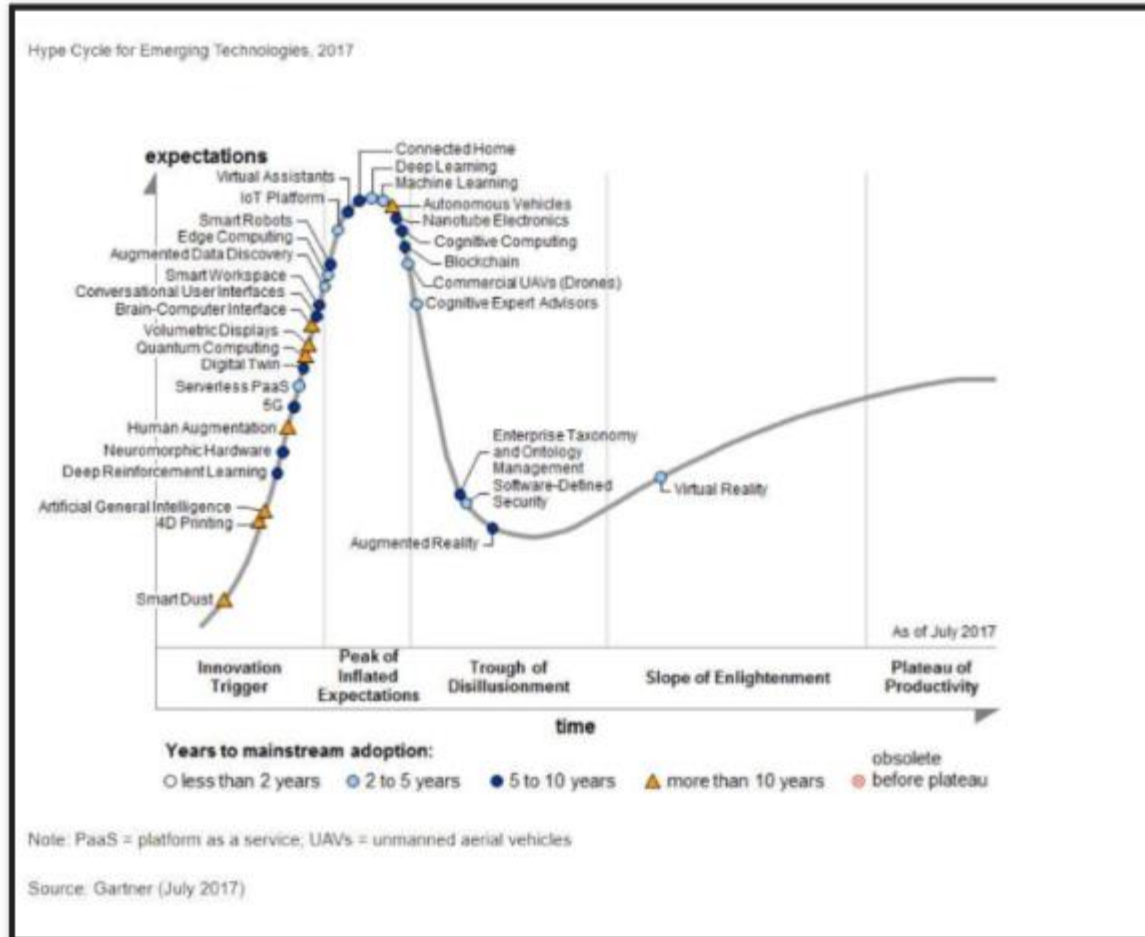
Plateau of Productivity : Les critères d'efficacité et de viabilité sont clairement définis. Et les applications commencent à payer

Promesses et attentes exagérées



Gartner Hype Cycle 2017

Promesses et attentes exagérées



En 2017, le «big data» est toujours au top des technologies hype, mais se décompose en Deep Learning et en Machine learning. Notez également la catégorie "Intelligence artificielle générale"

Gartner Hype Cycle 2017

Potentiel des analyses Big Data

Analyser et améliorer le service :

- Recommandation/Segmentation clients
- Recherche d'emploi
- Analyse de la réputation, des réseaux sociaux d'une marque

Améliorer la performance

- Anticiper les défaillances
- Economiser les ressources
- Accompagner la prise de décision

Prédire et prévenir

- Anticiper les épidémies
- Assurer la sécurité
- Anticiper les besoins
- Prévenir les maladies

Projet simple?



Les métiers fleurissent autour du Big Data :

- Les industries de gestion d'infrastructure
- Les cabinets de conseils
- Les fournisseurs d'outils / de logiciel / d'application métier



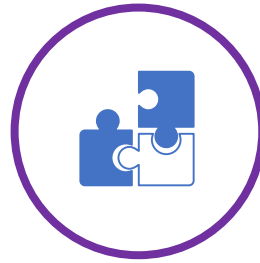
En résumé

Les projets BIG DATA visent en général à :



Analyser

une grande variété de données
(fichiers, vidéos, audio, photos, ...),



Compiler

plusieurs données en les
recoupant ou en les rapprochant.



Capter

la donnée en mouvement,

L'objectif est de collecter les données pour **donner du sens** et ainsi **anticiper les actions**.

Big Data : pour qui, pour quoi ?

Industrie



- Produit comme un service
- Qualité, innovation R&D
- Maintenance préventive

Distribution



- Offres temps réel et service personnalisés
- Optimisation de l'expérience magasin
- Pricing dynamique

Banques



- Parcours clients multi-canaux
- Fraude, anti blanchiment
- Partage des données consommateurs pour personnalisation

Assurance



- Fraudes et risques
- Recommandation client
- Tarification à l'usage, personnalisation

Santé



- Gestion des effets indésirables
- Traitements personnalisés.
- Amélioration des diagnostics

Transports, loisirs



- Planification et gestion des evts liés à la logistique
- Service client temps réel
- Economie d'énergie
- Pricing dynamique

Secteur public



- Services informationnels
- Fraudes, abus
- Sécurité publique
- Personnalisation de la relation citoyen

Telecom



- Parcours clients multi-canaux
- Partage de données de géo localisation
- Fraudes et analyse du comportement client

Produits gde conso.



- Analyse de sentiments et retour produits
- Relation personnalisée avec le consommateur
- Produit comme un service

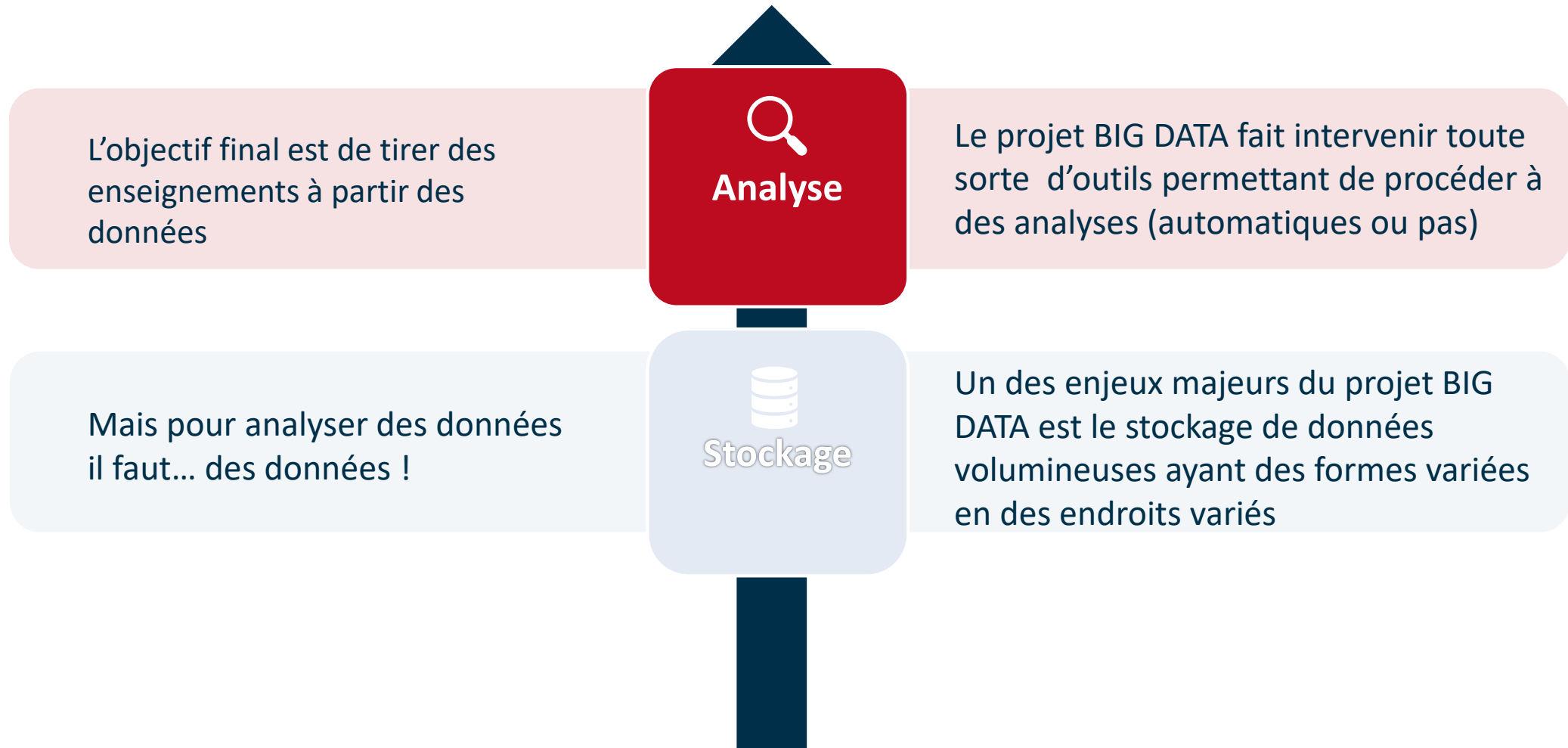
C'est là que la Data Science intervient



Les briques fonctionnelles d'un projet Big data



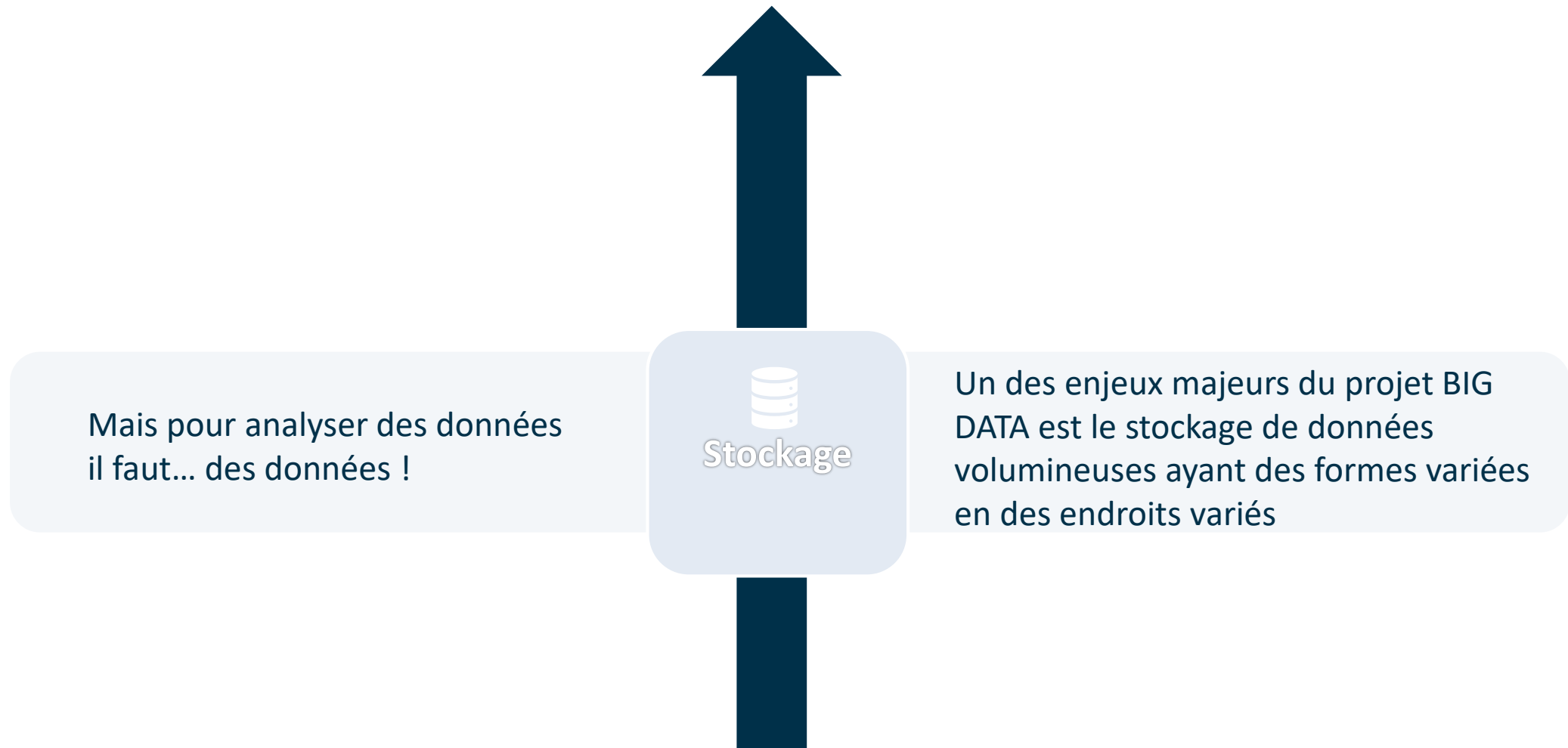
LES BRIQUES FONCTIONNELLES D'UN PROJET BIG DATA



LES BRIQUES FONCTIONNELLES D'UN PROJET BIG DATA



LE STOCKAGE



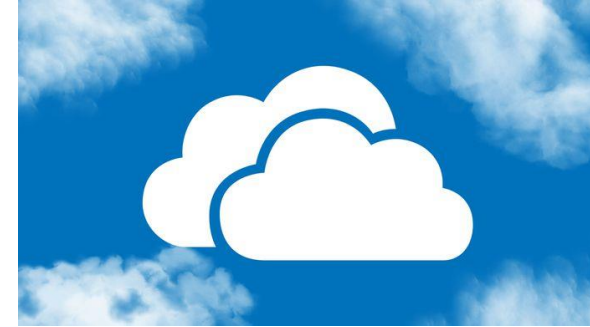
LE STOCKAGE

Le stockage de données a plusieurs formes :

- sur des serveurs physiques
- sur des Machines Virtuelles (VMs)
- sur le Cloud



LE CLOUD



Le terme CLOUD a été rendu populaire par Amazon en 2016 avec leur service Amazon Elastic Compute Cloud

Le cloud

Selon McKinsey (2013), l'émergence du cloud computing a beaucoup contribué au lancement de l'ère du Big Data:

- Le cloud computing permet aux utilisateurs d'accéder à des ressources de calcul et de stockage hautement évolutives à travers internet
- En utilisant le cloud computing, les entreprises peuvent utiliser la capacité des serveurs selon leurs besoins et l'étendre rapidement à l'échelle requise pour traiter de gros jeux de données et exécuter des modèles mathématiques complexes
- Le cloud computing réduit le coût d'analyse des big data car les ressources sont partagées entre de nombreux utilisateurs qui ne payent que pour la capacité qu'ils utilisent réellement.



LE CLOUD

Le cloud consiste à exploiter la puissance de calcul ou de stockage de [serveurs informatiques](#) distants par l'intermédiaire d'un réseau, généralement [Internet](#). Les serveurs sont loués à la demande, le plus souvent par tranche d'utilisation, selon des critères techniques (puissance, [bande passante](#), etc.), mais, également, au forfait.



Facile d'utilisation : pas besoin d'être technique



Facturé à la demande sur le temps d'utilisation



Architecture 'élastique' : évolue en fonction des besoins

LE CLOUD PUBLIC OU MUTUALISÉ

3 leaders :

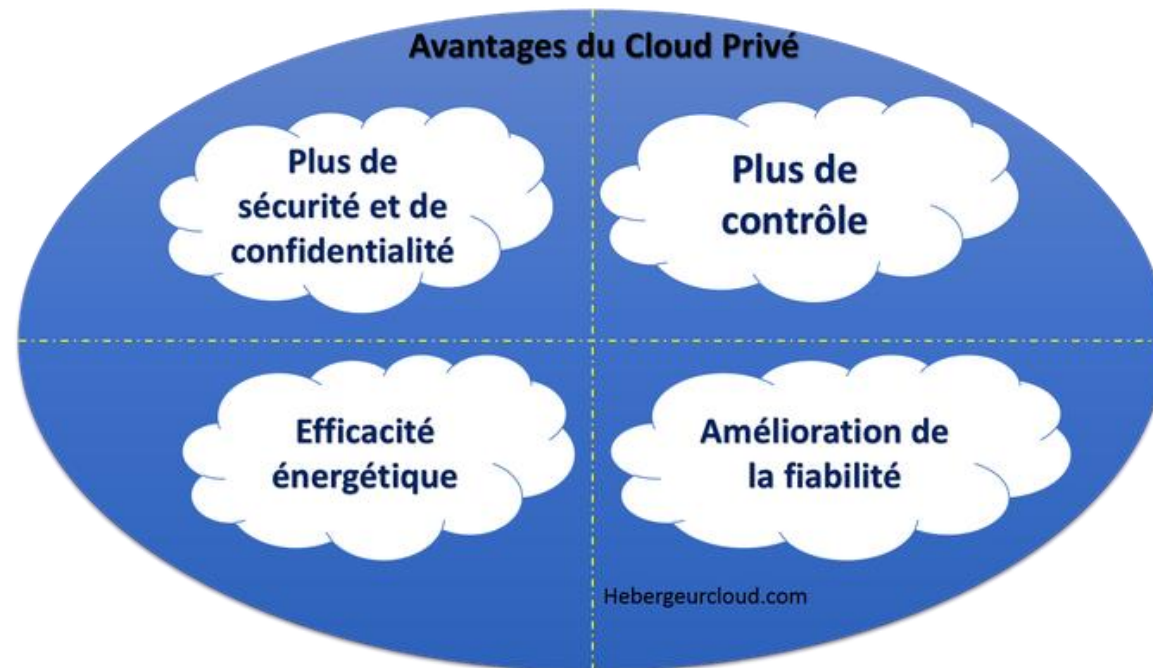


Google Cloud



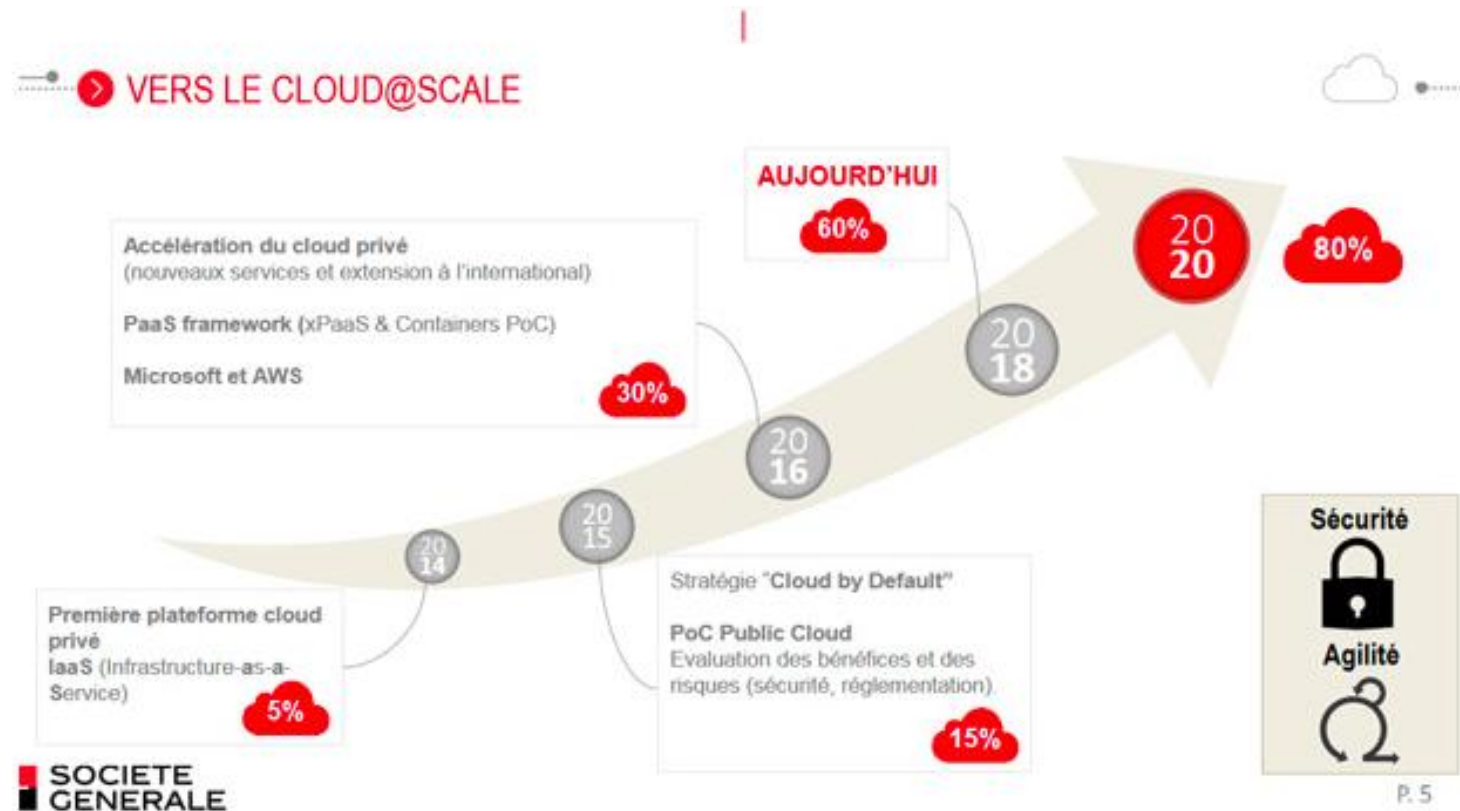
LE CLOUD PRIVÉ

Dans tout système de Cloud, seule votre entreprise et vos collaborateurs vont pouvoir accéder à vos ressources. Peu importe le terminal utilisé, que ce soit un ordinateur, une tablette ou un mobile, que ce soit dans les locaux de l'entreprise ou à l'extérieur, vous allez vous connecter à votre serveur de Cloud.



LE CLOUD HYBRIDE

Société Générale : exemple de stratégie de Cloud Hybride
2018 : 60% des infras dans le Cloud (55% Privé et 5% Public)
Cible 2020 : 80% (55% Privé et 25% Public)



4

Machine Learning, Data
Science, IA

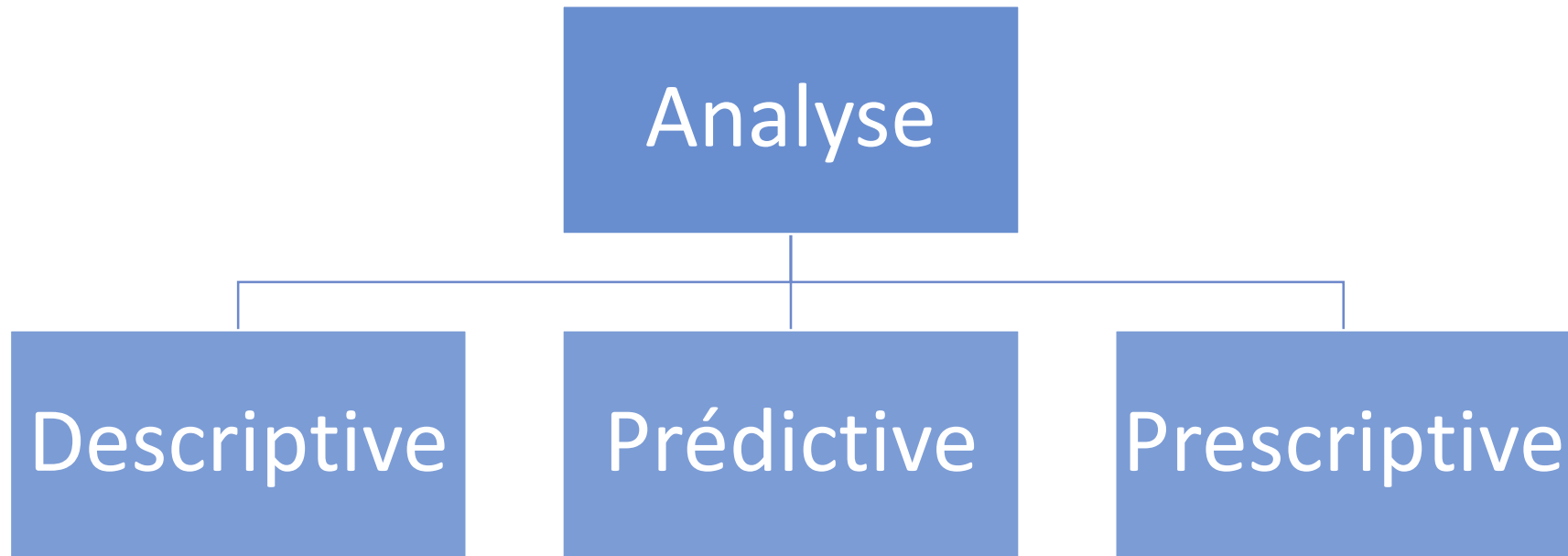
L'analyse

Stocker des données de types différents, provenant de sources différentes sur des technologies différentes c'est bien

MAIS

Ce qui est encore plus intéressant est d'en tirer de la valeur grâce à l'analyse

Découpage de l'analyse



Ce qu'il s'est passé
Ex : corrélations, moyenne, ...

Ce qu'il se passera
Ex : Détection fraude,
prédiction météo

Optimisation
***Pourquoi cela se passera et
recommandation d'actions pour
tirer un bénéfice des prédictions***

Des cas d'usages

Cas d'usage 1

Recommandation de produits en ligne

Exemple de recommandation sur Amazon

Les clients ayant acheté cet article ont également acheté



Le degré zéro de l'écriture
► Roland Barthes
★★★★☆ (7)
Poche
EUR 6,27



Qu'est-ce que la littérature
Jean-Paul Sartre
★★★★☆ (3)
Poche
EUR 7,98



Figures I
Genette
Poche
EUR 7,70



Littérature : 140 textes
théoriques et critiques
Jacques Vassevière
★★★★☆ (8)
Broché
EUR 24,13

Des cas d'usages

Cas d'usage 1

Recommandation de produits en ligne

- ✓ Historique des achats
- ✓ Autres produits achetés lors de la commande par d'autres utilisateurs
- ✓ Produits similaires basés sur des caractéristiques
- ✓ « Traduction » et analyse des clics sur le site d'achat en temps réel

Exemple de recommandation sur Amazon

Les clients ayant acheté cet article ont également acheté



Le degré zéro de l'écriture
Roland Barthes
★★★★☆ (7)
Poche
EUR 6,27



Qu'est-ce que la littérature
Jean-Paul Sartre
★★★★☆ (3)
Poche
EUR 7,98



Figures I
Genette
Poche
EUR 7,70

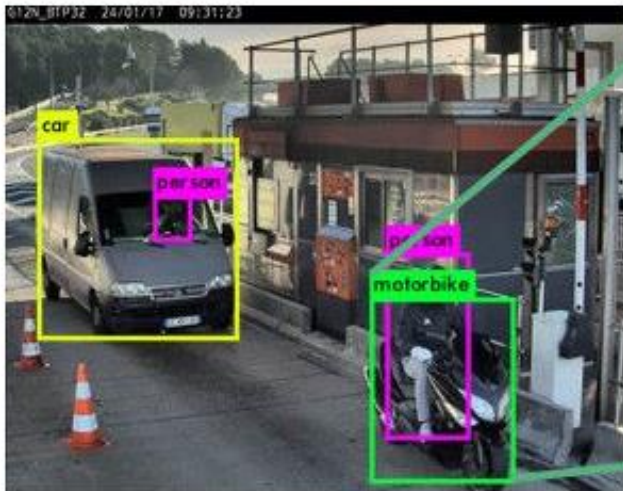


Littérature : 140 textes théoriques et critiques
Jacques Vassevière
★★★★☆ (8)
Broché
EUR 24,13

Des cas d'usages

Cas d'usage 2 Une société d'autoroute veut mieux classifier les véhicules au péage et éviter les interventions des agents lors des appels par interphone des usagers.

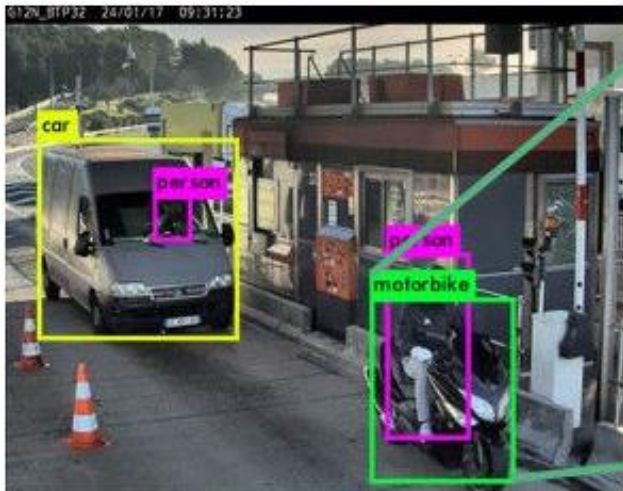
De l'analyse d'images aux Péages



Des cas d'usages

Cas d'usage 2 Une société d'autoroute veut mieux classifier les véhicules au péage et éviter les interventions des agents lors des appels par interphone des usagers.

De l'analyse d'images aux Péages



- ✓ Données météorologiques
- ✓ Calendrier (vacances)
- ✓ Analyse du flux en temps réel
- ✓ Historique des images, passages au péage, ...

Des cas d'usages

Cas d'usage 3 Une société de chauffeur privé souhaite ajuster ses prix en fonction de la demande et des temps de trajet.



Des cas d'usages

Cas d'usage 3 Une société de chauffeur privé souhaite ajuster ses prix en fonction de la demande et des temps de trajet.



- ✓ Données de trafic
- ✓ Données d'évènements
- ✓ Analyse des réseaux de transport public des villes

Des cas d'usages

Cas d'usage 4 Compteur intelligent pour adapter l'offre, détecter les défaillances et le risque d'attrition



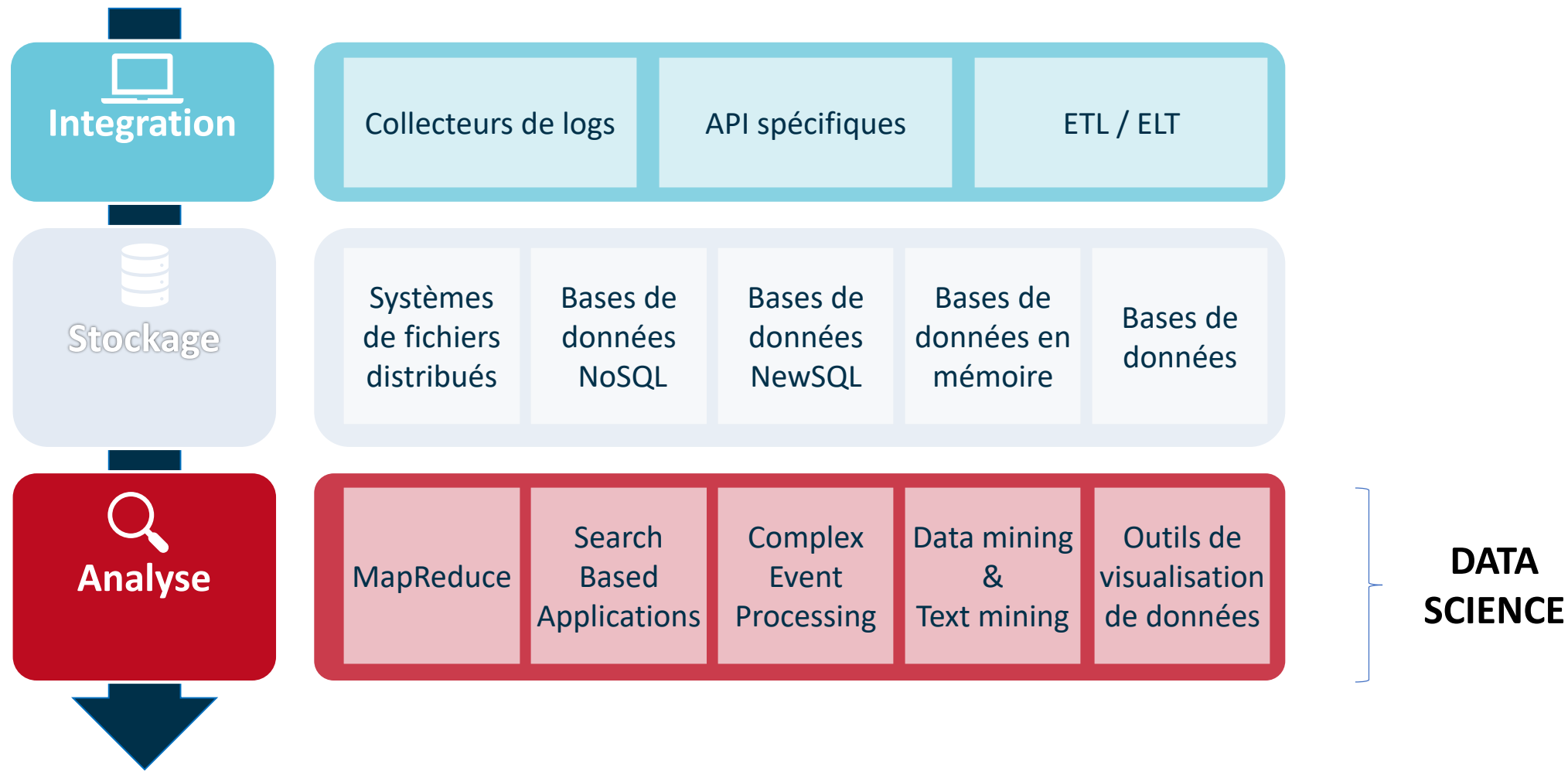
Des cas d'usages

Cas d'usage 4 Compteur intelligent pour adapter l'offre, détecter les défaillances et le risque d'attrition



- ✓ Données de consommation
- ✓ Données des clients
- ✓ 1 relevé par jour de ses 11 millions de clients

Tour d'horizon technologique - Big data



La data science

➤ Définition de la Data Science

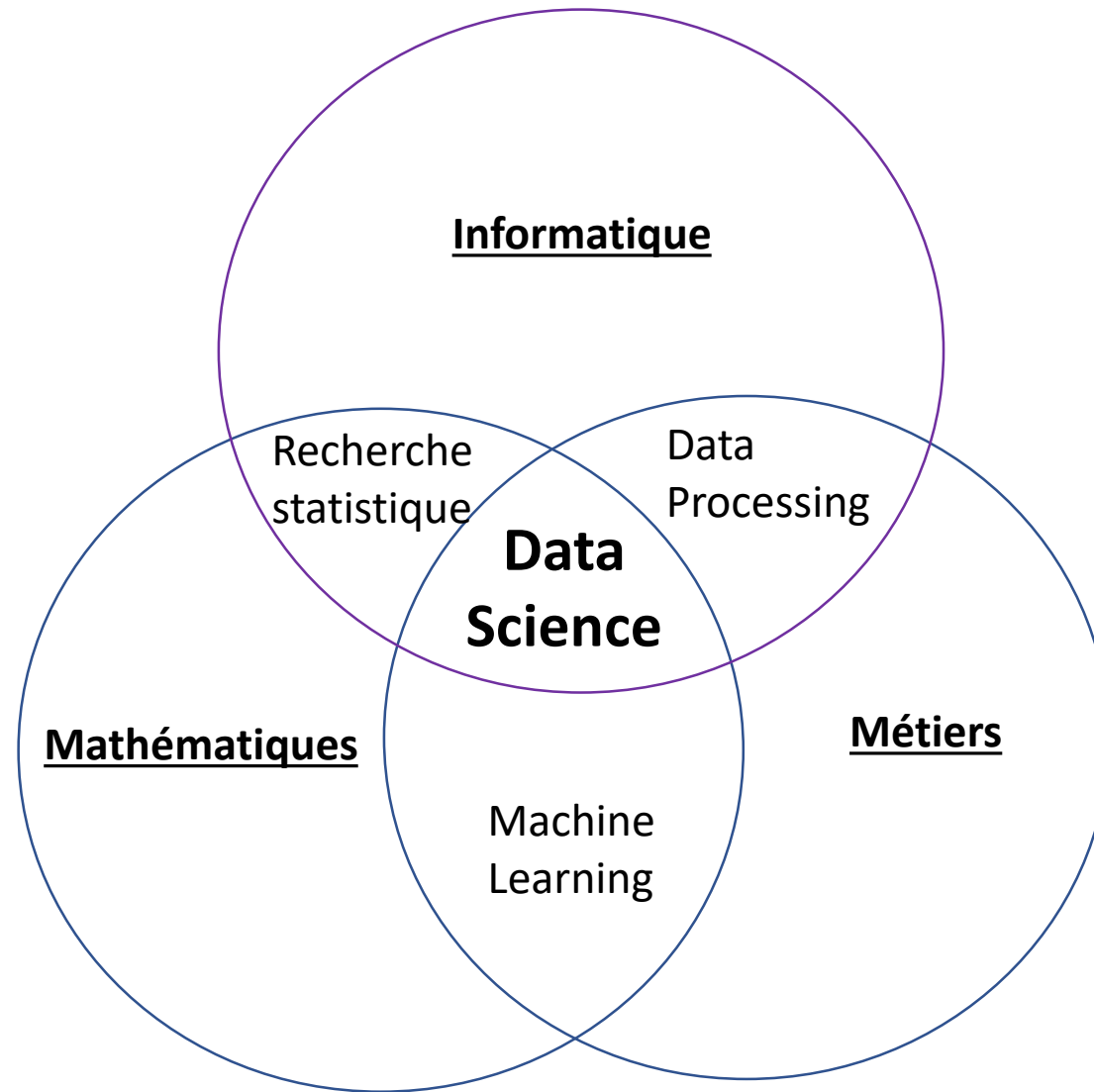
Objectif : Extraire des connaissances d'ensembles de données

La Data Science est un **mélange disciplinaire** entre les statistiques, le développement d'algorithmes et la technologie, dont le but est la **résolution de problèmes analytiques complexes**.

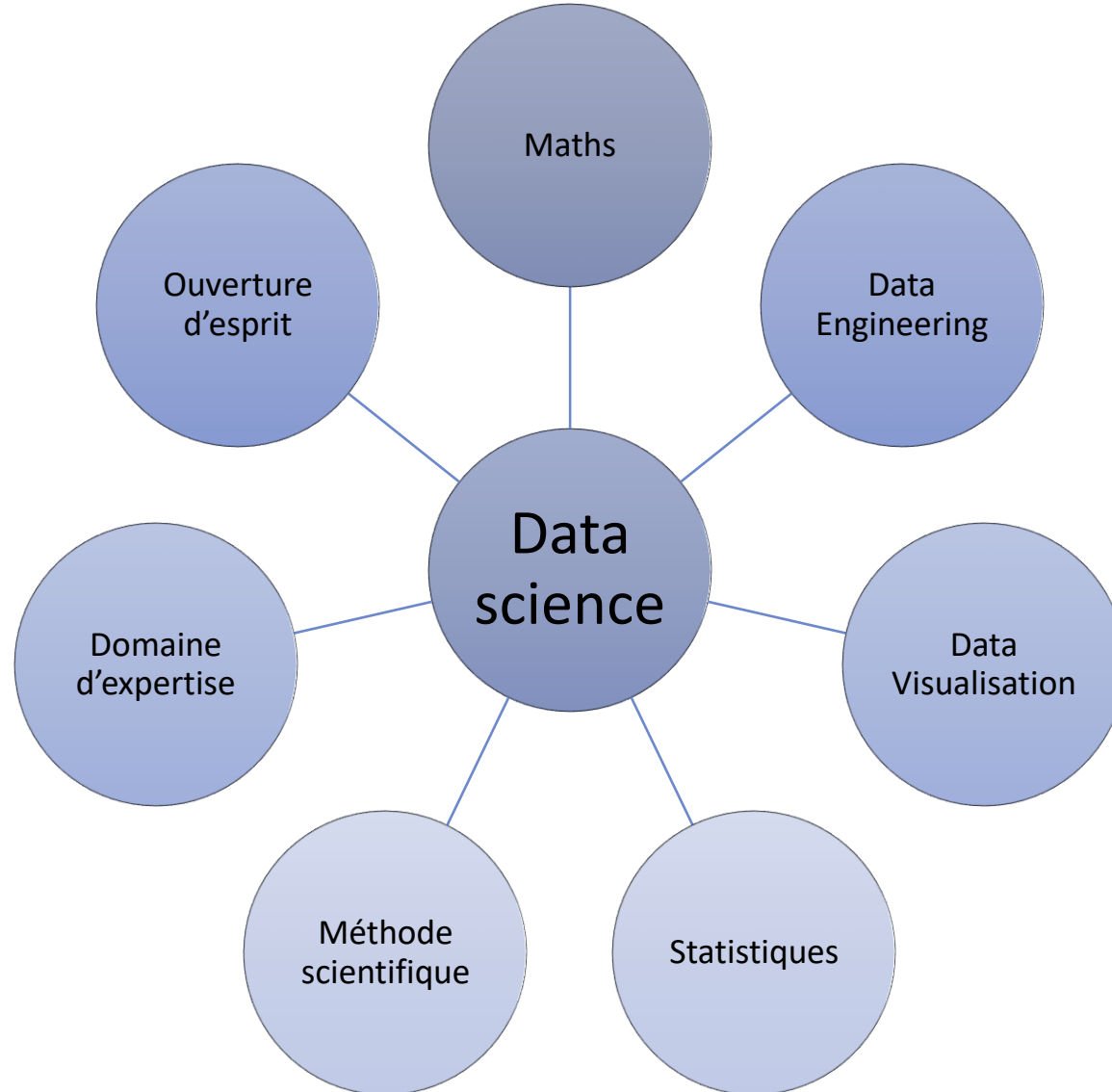


En étudiant les données, le Data Scientist peut **découvrir et comprendre des tendances et des comportements complexes**. Il aide à extraire des informations permettant aux entreprises de **prendre des décisions** plus adaptées à leurs différentes problématiques.

Qu'est-ce que la Data Science

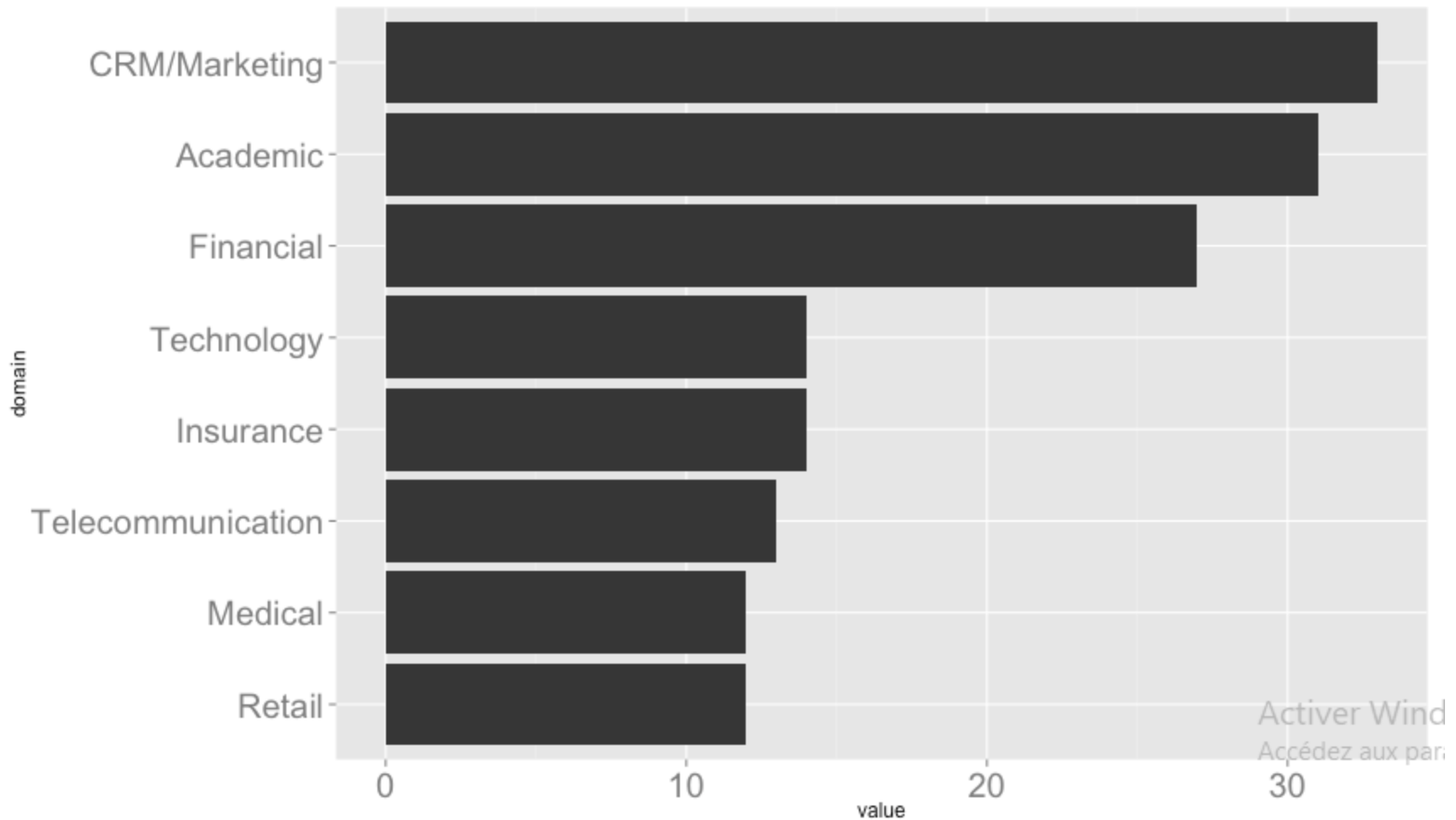


Qu'est-ce que la Data Science



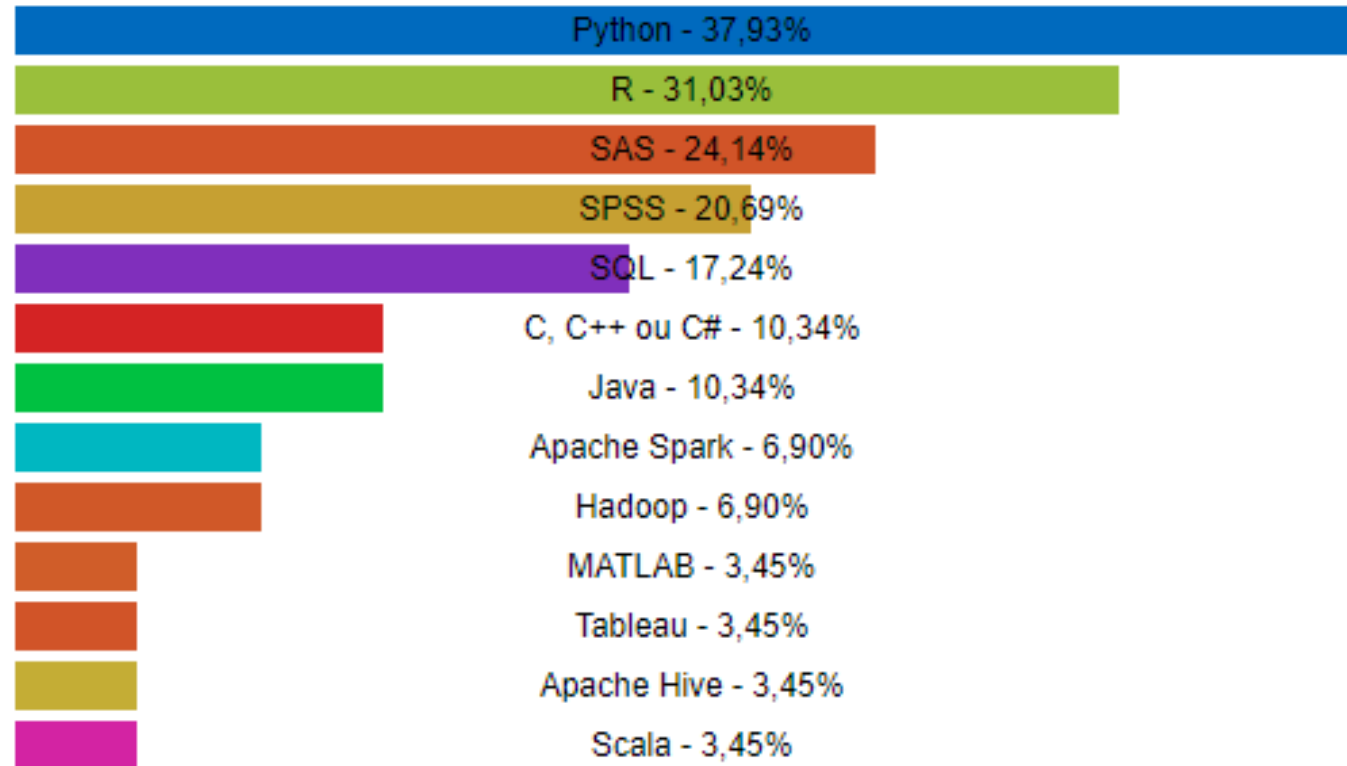
La data science

Les domaines où la Data Science est la plus utilisée



La data science

Les langages préférés des Data Scientist



La data science

Le Big Data

- ✓ Le mot tendance qu'on entend partout
- ✓ Englobe la partie architecture, data science et data visualisation

Data Science

- ✓ Résolution de problématiques analytiques complexes
- ✓ Mélange disciplinaire de statistiques et d'algorithmie

Le Machine Learning

- ✓ Un domaine de la Data Science
- ✓ Apprentissage par le biais d'exemples déjà présents dans la base de données

La data science

Et le data mining ?

- ✓ L'association : la recherche de schémas selon lesquels un événement donné est relié à un autre
- ✓ L'analyse de séquence ou de cheminement : recherche de schémas selon lesquels un événement donné conduit à un événement ultérieur
- ✓ La classification : recherche de nouveaux schémas
- ✓ L'agrégation : recherche de groupes de faits jusqu'alors non identifiés.
- ✓ La prévision : découverte au sein des données de schémas susceptibles de conduire à des pronostics raisonnables



Le data-mining est une partie d'exploration des données, que l'on peut également faire via de la data visualisation

La data science

Et le data mining ?

Applications du data mining dans les directions Marketing / Commerciale

- **Ciblage** : maximisez le rendement des actions de marketing direct, ciblez une offre vers les clients qui ont une forte probabilité d'y répondre, personnalisez votre marketing suivant le profil d'achats de vos clients
- **Fidélisation** : réduisez le départ de vos clients en identifiant les clients susceptibles de partir à la concurrence, réalisez des actions préventives sur vos clients fragiles.
- **Ventes additionnelles, ventes croisées** : comprenez et pilotez les migrations dans le cycle de vie des clients, dans les gammes de produits ou services, analysez la co-détention de produits pour accroître vos ventes
- **Efficacité de la force de vente** : aidez vos commerciaux à cibler leurs clients, affectez des scores d'appétence produits à chaque clients, anticipez le nombre d'appels sur vos plate-formes téléphoniques
- **Personnalisation de l'offre** : mettez en évidence des segments de clientèle caractérisés par un profil détaillé, des centres d'intérêt et des motivations, un potentiel commercial
- **Relation client** : étudiez l'usage des différents canaux de communication (e-mail, courrier, face à face, téléphone, etc.), utilisez le canal le plus adapté au profil de vos clients
- **Typologie d'agences** : analysez la clientèle et les résultats de chacune de vos agences, réalisez une typologie de vos agences et pilotez efficacement la définition de leurs objectifs commerciaux
- **Enquête de satisfaction** : mettez en évidence et hiérarchisez les principaux leviers de satisfaction et d'insatisfaction

La data science

La Data Science : La vraie valeur ajoutée du Big Data

➤ Qu'est ce que le Machine Learning ?

« **L'apprentissage automatique**, c'est la capacité d'un ordinateur à apprendre sans avoir été explicitement programmé. »

Arthur Samuel, pionnier du Machine Learning

C'est un outil complémentaire du Data Mining. Il consiste en un **apprentissage par le biais d'exemples** présents dans la base de données.

Cela consiste en la mise en place d'**algorithmes** ayant pour objectif d'obtenir une **analyse prédictive** à partir de données, dans un but précis.



Principe : **Bâtir des modèles qui établissent les relations entre de nombreux facteurs** permettant l'évaluation des risques ou celle des opportunités associés à un ensemble spécifique de conditions, afin **d'orienter la prise de décision** vers telle ou telle action.

Machine learning

En entreprise, le Machine Learning doit satisfaire plusieurs critères :

- Le meilleur ratio « Meilleur résultat » / « temps/complexité de calcul »

Contrairement à la recherche, la précision du résultat est évidemment important mais nous devons satisfaire un calcul pas trop long/complexe et pas trop couteux en terme de ressource

- Possibilité de parallélisation

Pour réduire les temps de construction/calcul, on doit paralléliser sur plusieurs machines.

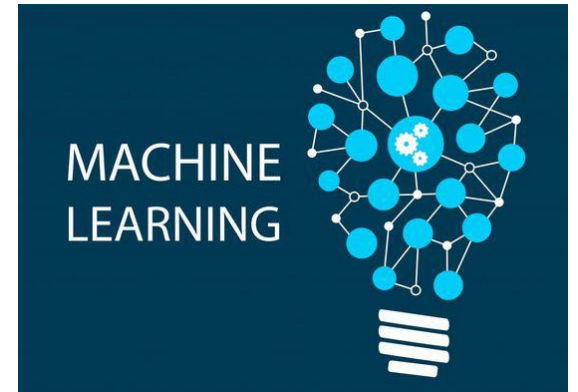
Machine learning

➤ Les différents types de Machine Learning

Apprentissage supervisé

On parle notamment de **classification** et de **régression**.

Ex : Prédiction de prix de maison en fonction de ses caractéristiques
Classifier un email en spam ou non

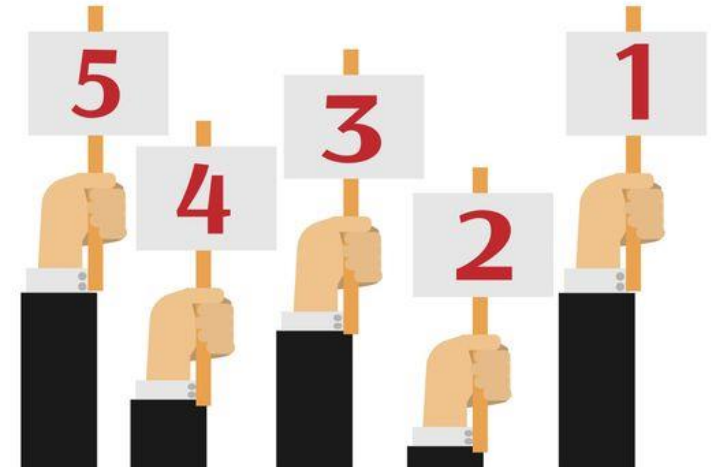


Machine learning

➤ Exemple de classification

On souhaite scorer les clients d'une banque selon leur capacité à rembourser un crédit.

En se basant principalement sur les informations bancaires et les données comportementales du demandeur, le score risque permet de déterminer la probabilité que celui-ci assure le remboursement de ses échéances jusqu'à leur terme.



Machine learning

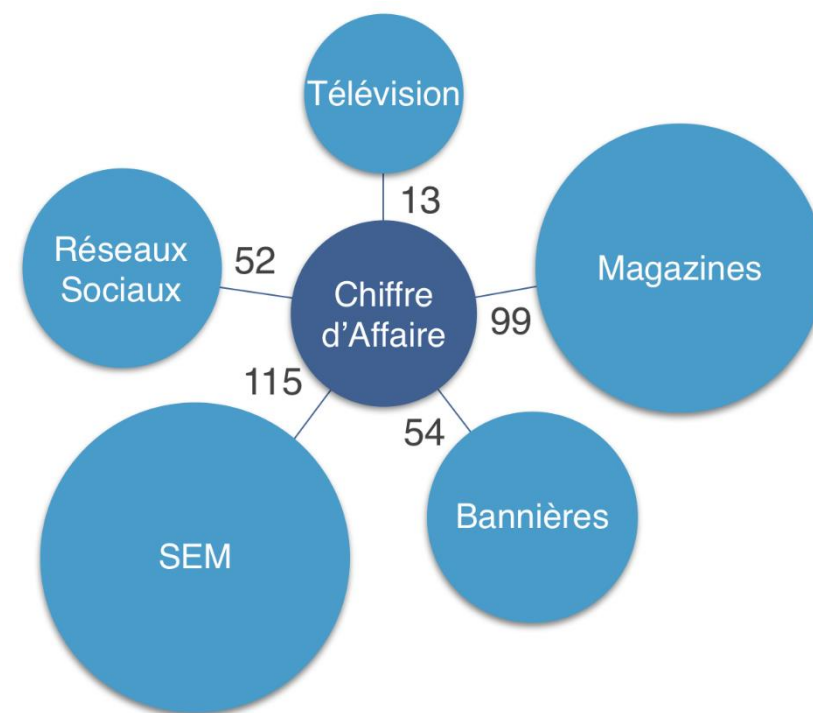
➤ Exemple de régression

Déterminer le budget idéal à allouer à chacun des médias de communication.

On récupère les budgets attribués à chaque média ainsi que les chiffres d'affaires réalisés par l'entreprise sur les 12 derniers mois.

Un modèle mathématique simple attribuera un coefficient à chaque média pour mesurer son impact sur le chiffre d'affaire.

Plus le coefficient relatif au média sera élevé, plus l'impact de ce média sur le chiffre d'affaire sera grand et plus il sera intéressant d'investir un budget supérieur le trimestre prochain.



Machine learning

➤ Les différents types de Machine Learning



Apprentissage non supervisé

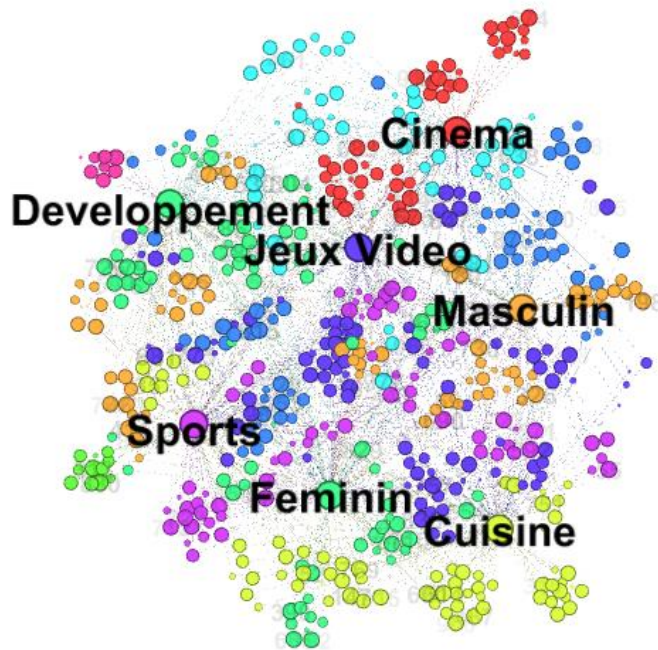
On parle notamment de **clustering**.

L'algorithme va regrouper les données par similarités, selon une notion mathématique de distance. Par exemple on fournit un ensemble de photos d'animaux, l'algorithme va regrouper les photos de chats dans un même cluster et celles des chiens dans un cluster différent.

Machine learning

➤ Exemple de clustering :

On souhaite établir des groupes de personnes selon les thématiques qu'ils évoquent sur Twitter



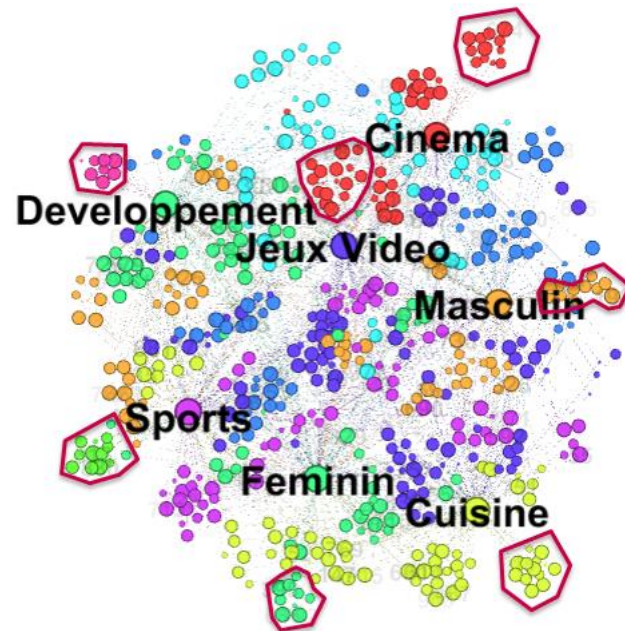
Ici, les cercles sont de deux types : les plus gros sont les thématiques, tandis que les plus petits sont des utilisateurs. Si un utilisateur a partagé un lien associé à une des thématiques, il est relié à celle-ci par un arc dans le graphe.

On va pouvoir isoler des utilisateurs spécifiques à une seule thématique, les influenceurs. Les individus qui leurs sont proches seront sûrement intéressés par des contenus de la thématique en question : voilà la base d'un système de recommandation

Machine learning

➤ Exemple de clustering :

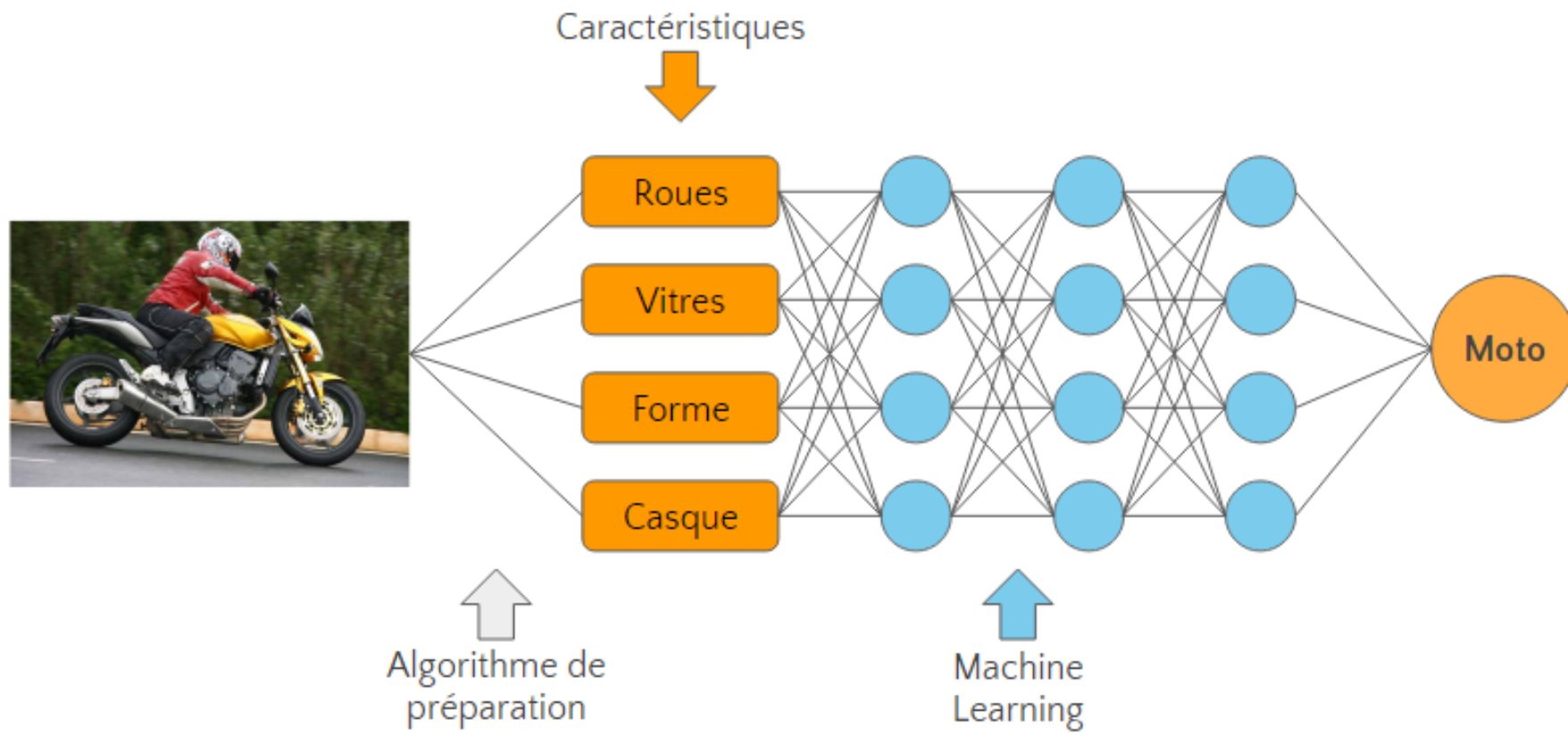
On souhaite établir des groupes de personnes selon les thématiques qu'ils évoquent sur Twitter



Idées d'utilisation du clustering :

Proposer du contenu aux influenceurs
sur leur sujet de prédiction

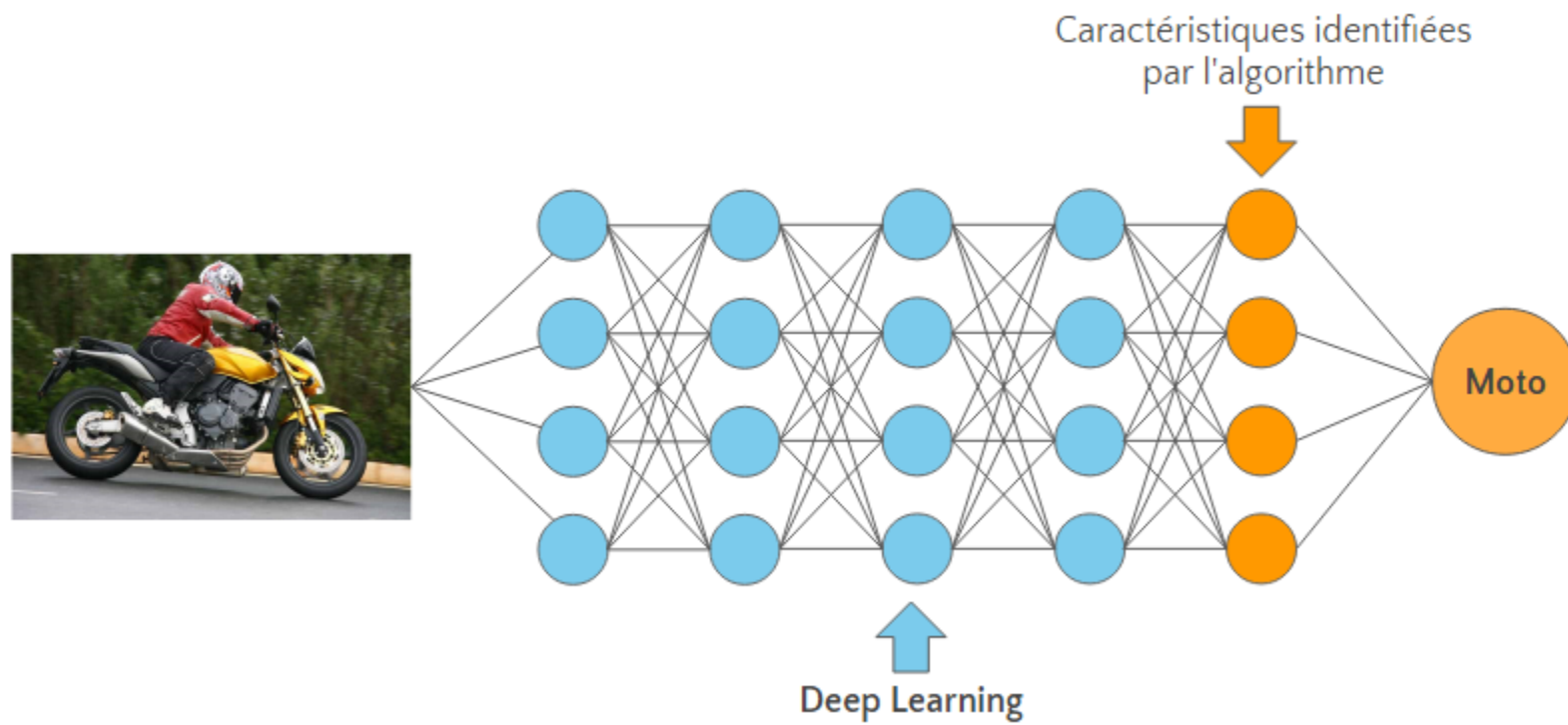
Machine learning



Deep learning

- ✓ Les caractéristiques essentielles du traitement ne seront plus identifiées par un traitement humain dans un algorithme préalable, mais directement par l'algorithme de Deep Learning.
- ✓ Construit, à partir des exemples à disposition, ses propres caractéristiques
- ✓ Le but est d'être le plus similaire à l'identification humaine

Deep learning



Deep learning

Chihuahua vs. Muffin



Deep learning

Visualisation souvent utilisée
pour représenter l'IA
(Intelligence Artificielle)

Derrière : Algorithme de Deep
Learning

Chihuahua vs. Muffin



chocolate cookie



fawn smooth Chihuahua



baked blueberry muffin



white chihuahua



fawn smooth Chihuahua



blueberry muffin



[unknown]



brown coated Chihuahua



baked muffin



beige short coated puppy

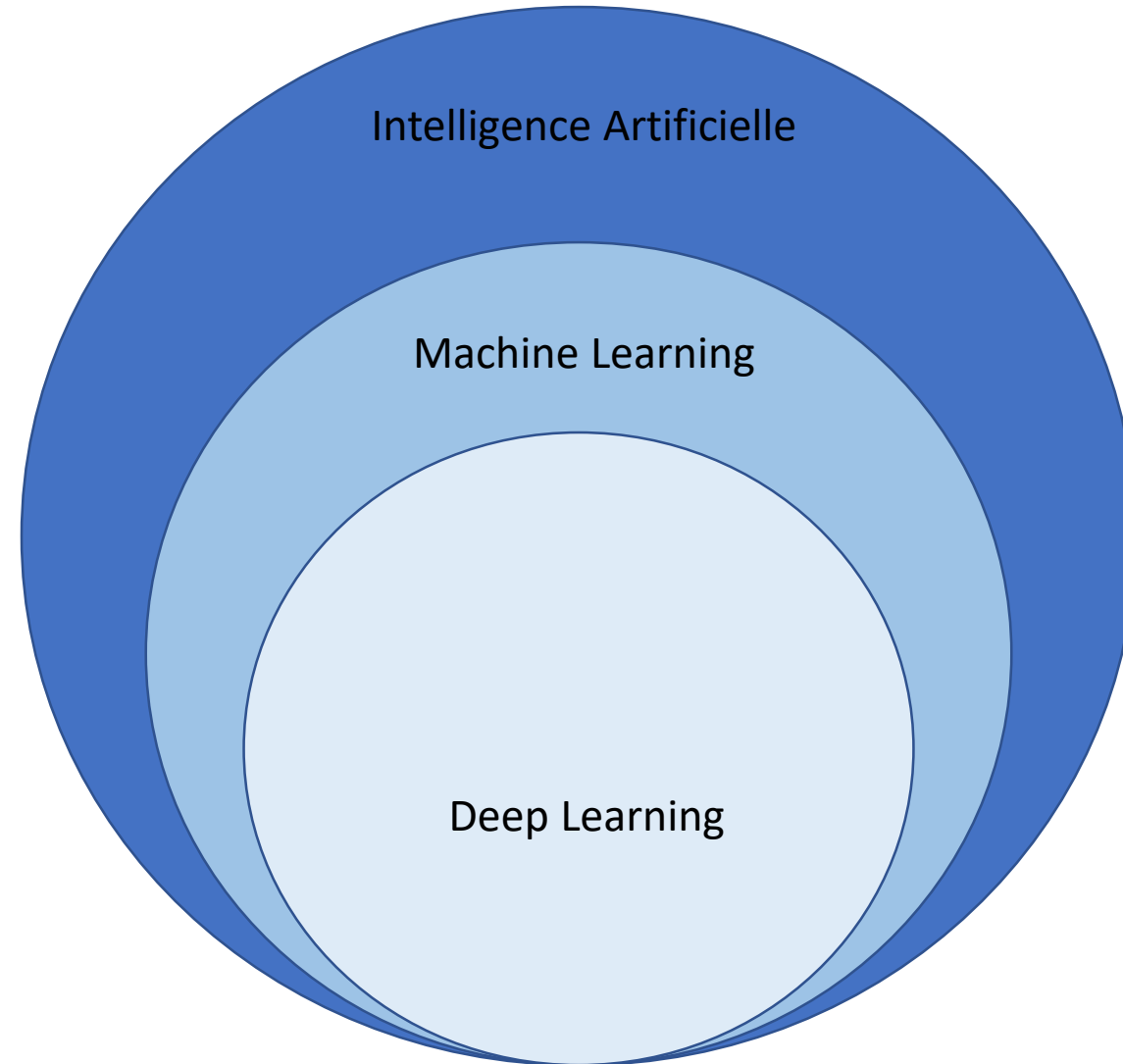


tan smooth Chihuahua puppy



blueberry cupcakes

In fine



4

Focus IA

Test de Turing

Article de 1950 : Computing Machinery and Intelligence.

Idée : ne pas définir l'IA mais la tester.

Protocole : un individu communique à l'aide d'un terminal d'ordinateur avec un interlocuteur invisible. Il doit décider si l'interlocuteur est un être humain ou un système d'IA imitant un être humain.

Construire un cerveau artificiel en imitant moins la physiologie du modèle humain, que son comportement logique, les "états de pensée" étant équivalents aux instructions de la machine.

Prédiction de Alan Turing : 30 % de réussite pour une machine discutant avec une personne non avertie pendant 5 minutes.

Les principales composantes d'un système d'IA doivent être les connaissances, le raisonnement, la compréhension du langage naturel et l'apprentissage. **TURING**



L'Intelligence Artificielle

L'explosion de la puissance de calcul des machines et des données dont nous disposons ont fait basculer l'Intelligence Artificielle, dans les années 2010, d'une idée de science-fiction à une réalité de plus en plus proche, devenue enjeu scientifique majeur.

L'Intelligence Artificielle (IA, ou AI en anglais pour *Artificial Intelligence*) consiste à mettre en œuvre un certain nombre de techniques visant à permettre aux machines d'imiter une forme d'intelligence réelle. L'IA se retrouve implémentée dans un nombre grandissant de domaines d'application.

L'IA est déjà partout

Actuellement, on a déjà de l'intelligence artificielle un peu partout autour de nous



I'm listening...

Hey Siri[™]
tap to edit



Waze



12:32



240 m

Jefferson Ave

13:04PM 21mi. 32min. >





Internet est

internet est **lent**

internet est **mort**

internet est **il un espace privé**

internet est **lent windows 10**

internet est **un service du web**

internet est **dangereux**

internet est **un moyen de communication et de liberté**

internet est **la première chose que l'homme a créé sans la comprendre**

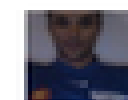
internet est **long**

internet est **attaqué**

Recherche Google

J'ai de la chance

[En savoir plus](#)



Yohann Polron

FAVORITES

News Feed

Messages

Events

LISTS

Close Friends

Family

20+

[See all lists...](#)

PAGES

Blog des Nouveaux...

1

APPS

Bejeweled Blitz

Live Holdem Pro

App Requests

3

Edgeworld

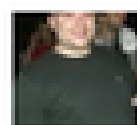
GROUPS

Create Group...

 You are currently offline.
To chat with your
friends, go online.

Update Status Add Photo Ask Question

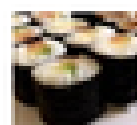
RECENT STORIES



Alexandre Ferchaud > Olivier Luquero

Bon anniversaire! Biz

Like - Comment - See Friendship - 8 minutes ago via mobile -



Chloé Vilalta

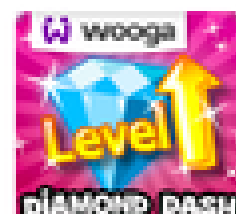
according to my driving instructor, I am a very good driver!!!

Like - Comment - 12 minutes ago -

Justine Ranger likes this.



Marie-France Bolssinot

Marie-France a atteint un nouveau niveau dans
Diamond Dash!Marie-France est maintenant au niveau 8 a atteint un
Bonus de 10%. Like - Comment - Débloque un niveau - 17 minutes ago via Diamond
Dash - 

Maxime Jaquet

Petit entretien avec mon chef pour voir la suite des événements ... a suivre !!!

Like - Comment - 23 minutes ago -



Maxime Jaquet

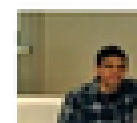
Caroline Baudry commented on
her own photo: "Clin d'oeil au
canadien :) une..."Chloé Vilalta likes Raphaël
Lootvoet's post on Coin Jeunes
Mignaloux Beauvoir's Wall.Caroline Baudry commented on
Julie Chatelain's post on Caroline
Baudry's wall: "Je suis au buffet
du lux cafe..."Mthid Htkns recommends a link:
Nouvelle République : Motards :
le rappel à l'ordre du maire de
Niort - société.

Arnaud Lièvre and Olivier Luquero



Create an Event

People To Subscribe To

[See All](#)

Blake Ross

Director of Product at Facebook
1 friend is subscribed.

Subscribe



Kevin Rose

Co-Founder/CEO at Milk Inc.
2 friends are subscribed.

Subscribe

Ou encore :

- ✓ **Les recommandations de musique sur Spotify**
- ✓ **La détection de spam dans les mails**
- ✓ **Autofocus de l'appareil photo de votre smartphone**
- ✓ **...**

L'Intelligence Artificielle ce n'est pas que des robots, le mythe de l'Intelligence Artificielle qui remplace l'homme est un fantasme

Concrètement, c'est quoi l'IA ?

Il existe deux types d'Intelligence Artificielle :

- L'IA faible ou verticale :

- n'a pas vocation à évaluer
- ne va pas évoluer
- vise à effectuer une ou plusieurs tâches

-> La machine apprend à reproduire un comportement humain, elle peut faire cette tâche mieux qu'un humain mais elle ne sait rien faire d'autres que la tâche qu'on lui aura apprise.

- L'IA forte ou horizontale :

- approche qui cherche à comprendre et construire
 - * Des systèmes de plus en plus autonomes
 - * Des algorithmes capable de résoudre des problèmes
 - * La programmation d'un apprentissage

-> Reproduit toute l'intelligence humaine, avec sa capacité de conscience et de pouvoir réappliquer la tâche que l'on a appris à tout autre chose.

Environnement technique adéquat



Des algorithmes de Data Science innovants

Analyse de
texte

VoiceMining

Reconnaissance
d'image

Deep/Machine
Learning

Théorie du
signal

IA, Machine Learning

L'Intelligence Artificielle est la capacité des machines à se comporter d'une manière que nous considérons comme «intelligente», notamment les fonctions cognitives.



Le Machine Learning utilise l'intelligence artificielle pour traiter d'importants volumes de données, permettre à la machine d'apprendre par elle-même et de s'améliorer avec l'expérience

Objectifs à long terme de l'IA

Construire des agents intelligents

- capable de résoudre des problèmes variés
- de façon autonome
- en s'adaptant au contexte rencontré

Principales difficultés

- proposer des méthodes génériques
- sans trop d'hypothèses sur les problèmes à résoudre
- mais suffisamment précises pour pouvoir proposer des théories et des algorithmes

Place au Travaux Dirigés

Enoncé disponible :

<https://github.com/apatou>