

SUPPLY CHAIN 4.0



SEANCE 4-A

Big Data, Data Mining Data Analytics

Outils technologiques

4 PRINCIPAUX MOTEURS DE CHANGEMENT



94%

Plateformes/
solutions de
Visibilité Supply
Chain



90%

Big Data/
Analytics



81%

Outils de
simulation



80%

Cloud

MIEUX EXPLOITER LES DONNÉES

"Mon entreprise a accès à la plupart des données issues de ma supply chain étendue."



15% aujourd'hui

54% d'ici 2020

"La plupart des données issues de ma supply chain étendue sont analysées et utilisées dans le cadre de processus décisionnels."



23% aujourd'hui

68% d'ici 2020

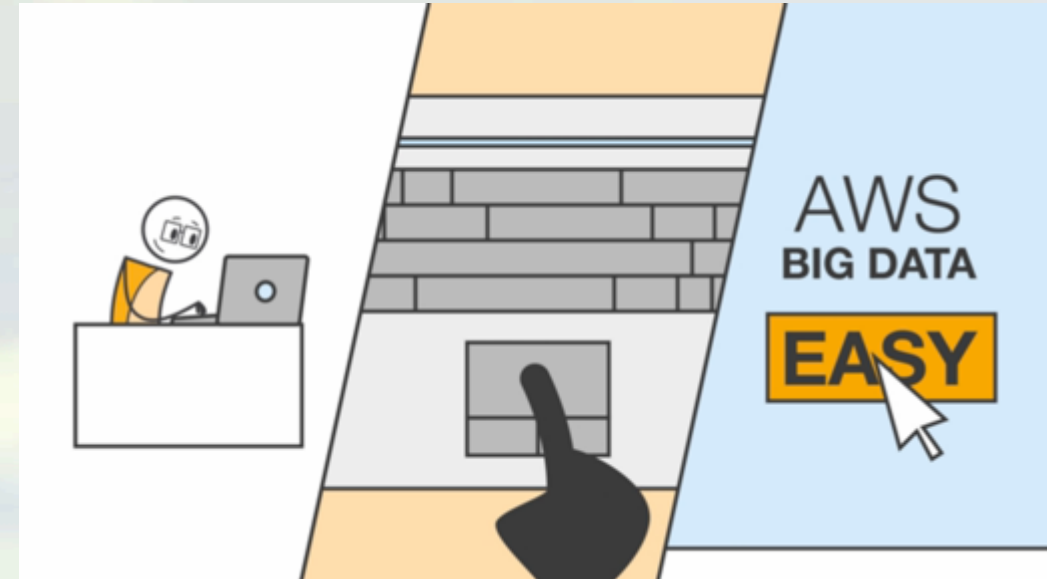
EVOLUTIONS DE LA SUPPLY CHAIN

<https://www.lebigdata.fr/>

<https://www.lebigdata.fr/analytics/data-analytics>

BIG DATA

Il s'agit de gros volumes de données (structurées ou non) qui alimentent l'activité quotidienne d'une entreprise. Mais ce n'est pas la quantité des données qui compte, c'est ce que les entreprises en font. Il est possible d'analyser les big data pour en tirer des enseignements permettant de prendre de meilleures décisions et de lancer des initiatives stratégiques.



<https://aws.amazon.com/fr/big-data/>

applications d'analyse de Big Data

LES 3 CARACTERISTIQUES DU BIG DATA

LES 3V

Volume. Les entreprises collectent des données de diverses sources : les transactions commerciales, les médias sociaux et les informations issues de capteurs ou de transactions de machine à machine. Auparavant, le stockage de ces données aurait posé problème, mais cette tâche est désormais simplifiée par les nouvelles technologies, notamment **Hadoop**.

Vitesse. Les données nous parviennent à un rythme de plus en plus soutenu et doivent être traitées rapidement. Balises RFID, capteurs et compteurs intelligents déversent des torrents d'informations qu'il faut analyser en temps quasi réel.

Variété. Les données revêtent tous types de formats : des données numériques structurées dans des bases de données traditionnelles aux documents texte non structurés, en passant par les e-mails, les vidéos, les fichiers audio, les données boursières et les transactions financières.

Hadoop est un framework libre et open source écrit en Java destiné à faciliter la création d'applications distribuées et échelonnables permettant aux applications de travailler avec des milliers de nœuds et des pétaoctets de données. Ainsi chaque nœud est constitué de machines standard regroupées en grappe.

A quoi servent les big data ?

- réduire les coûts
 - gagner du temps
 - développer de nouveaux produits et optimiser les offres et
 - prendre des décisions avisées. *En associant les big data à de puissants outils d'analyse, vous pouvez accomplir les tâches essentielles pour l'entreprise*
-
- Déterminer en temps quasi réel les causes premières des défaillances, problèmes et défauts.
 - Générer des bons de réduction sur le point de vente en fonction des habitudes d'achat du client.
 - Recalculer le risque sur des portefeuilles complets en quelques minutes.
 - Détecter un comportement frauduleux avant qu'il n'ait une incidence sur votre entreprise.

REX UPS – BIG DATA



PROJET

ORION

(On-Road Integration Optimization and Navigation)

Ce projet s'appuie essentiellement sur des données cartographiques en ligne pour réorganiser en temps réel les tournées des chauffeurs.

ECONOMIE DE PRES DE
32 millions de litres de carburant

RACCOURCISSEMENT
des tournées quotidiennes de
127 millions de kilomètres.

UPS estime que diminuer les
parcours de **1,5 km par jour et
par chauffeur** permet à la société
de réaliser **30 millions de dollars
d'économies.**

DATA ANALYTICS

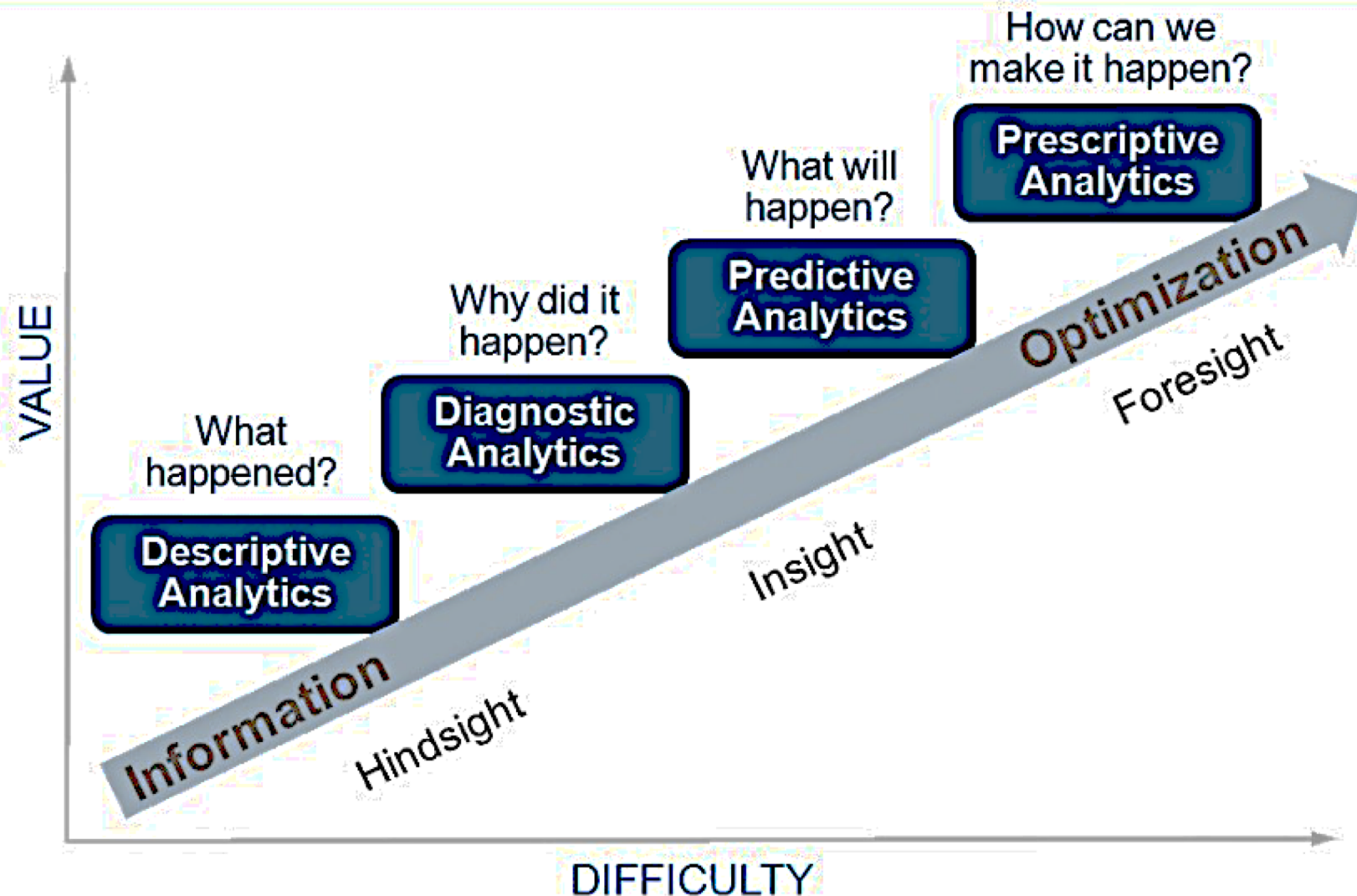
Le Data Analytics, en abrégé par **DA**, est une science qui permet d'examiner des données brutes, pour tirer des conclusions à partir des informations exploitées.

Le Data Analytics est utilisé dans de nombreuses industries et organisations : il permet de **prendre de meilleures décisions**.

C'est le **Data Analyst** qui met en œuvre.



Analytic Value Escalator



La valeur de l'analyse des données est d'autant plus forte, lorsque cette analyse est complexe.

Comment peut-on faire pour que cela se passe ?

PREVOIR

Qu'est-ce qui va se passer ?

PRESENT

Qu'est-il arrivé ? Que s'est-il passé ?

RETROSPECTIVE

Qu'est-il arrivé ?

EDA, CDA et QDA

Le Data Analysis se décompose en 2 catégories

- L'Exploratory Data Analysis (*analyse de donnée exploratoire*), acronyme **EDA**, aide à **découvrir de nouveaux éléments** dans les données.
- La Confirmatory Data Analysis (*analyse de données confirmatoire*), acronyme **CDA**, permet de **prouver si des hypothèses existantes sont vraies ou fausses**.

La Qualitative Data Analysis (*analyse de données qualitative*), acronyme **QDA**, est utilisée dans les sciences sociales pour tirer des conclusions de données non numériques telles que les mots, les photographies ou la vidéo.

Informatique & Data Analytics

Lorsqu'il est utilisé dans un contexte informatique il s'agit d'un audit informatique.

Il s'agit dans ce cas de déterminer si les systèmes d'information sécurisent au niveau souhaité les données

EXEMPLES D'UTILISATIONS

E-COMMERCE

Les banques et les entreprises de cartes de crédit, des acteurs du e-commerce tels que STRIPE, PAYPAL, PAYPLUG, HIPAY, GREEN PAY, analysent les transactions et les dépenses pour **empêcher les fraudes et les usurpations d'identité**.

Les sites de Ecommerce analysent le trafic d'un site web ou les cycles de navigation pour déterminer quels consommateurs sont susceptibles d'acheter un produit ou un service en fonction de leurs précédents achats ou des pages qu'ils ont consultées.

REAL-TIME ANALYTICS

Les Data Analytics utilisent des tableaux de bord d'information et KPI alimentés par des flux d'information en temps réel.

Le Real-Time Analytics c'est **l'analyse et le rapport dynamique** basé sur des données inscrites dans un système moins d'une minute avant le moment de l'utilisation.



La startup **UntieNots** a utilisé sa technologie d'analyse de données basée sur l'intelligence artificielle pour comprendre les habitudes de consommation des supporters et des téléspectateurs de la dernière coupe du monde de football



Apache HBase est une base de données non relationnelle (NoSQL) faite pour les analyses Big Data. Elle autorise le traitement rapide d'immenses volumes de données de différentes sources et de structures diverses.

LES LOGICIELS POUR LE BIG DATAT

<https://www.capterra.fr/directory/31077/data-analysis/software>



Créé par Apache, c'est un framework open source qui permet de traiter de grands ensembles de données par l'intermédiaire de centaines de serveurs qui travaillent en parallèle. Il permet de trier et d'analyser, de traiter des données en les rendant disponibles sur des machines locales.



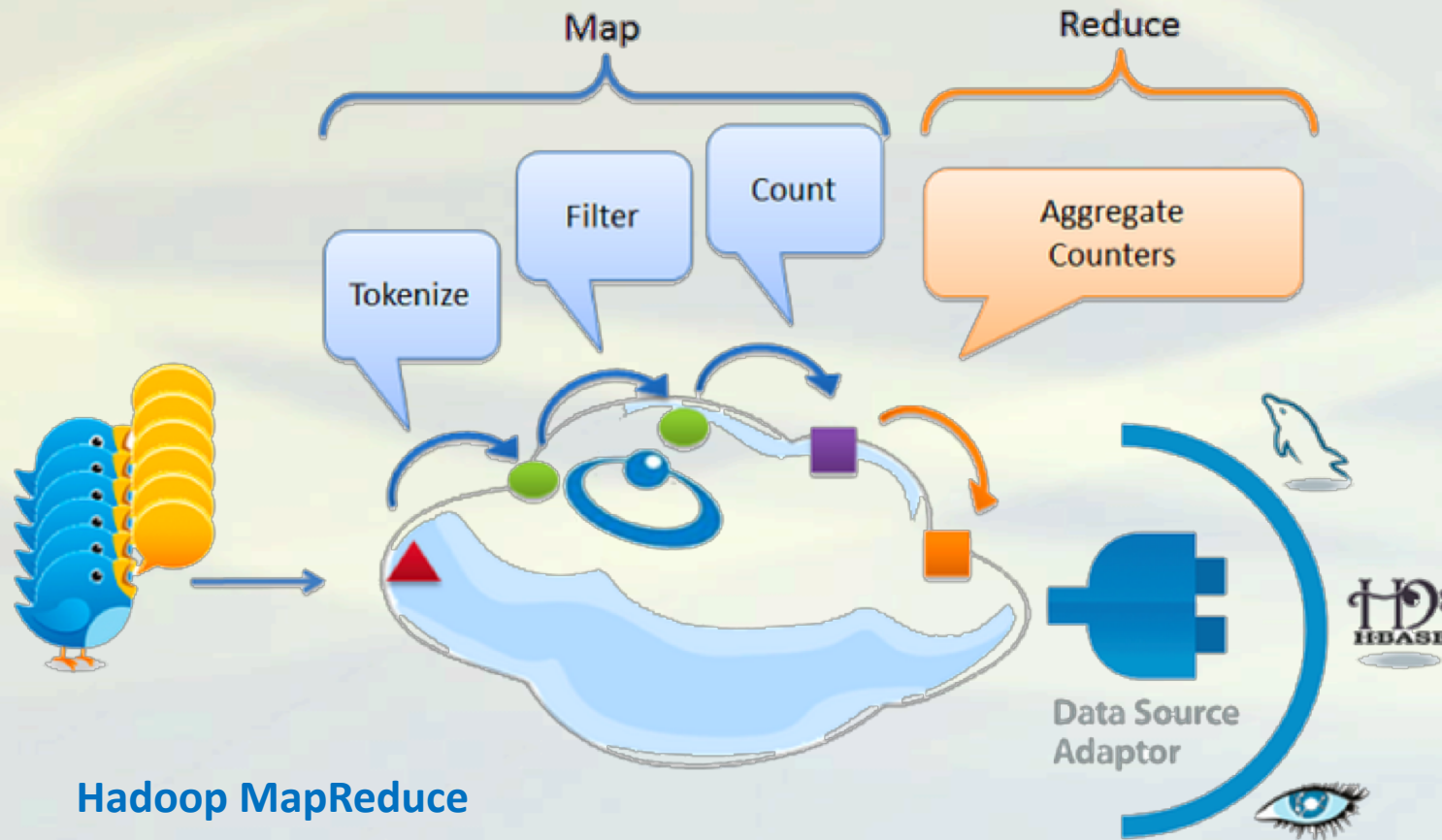
APACHE
STORMTM
Distributed • Resilient • Real-time

Créé par Apache. C'est un open source qui permet de traiter du big data en temps réel. Compatibles avec tous les langages de programmations. Storm s'adapte au Machine Learning en ligne.



Apache Cassandra est une base de données NoSQL hautement scalable. Elle peut surveiller de larges ensembles de données répartis sur divers clusters de serveurs et sur le Cloud.

LES LOGICIELS POUR LE BIG DATAT



C'est un modèle de programmation et un framework logiciel pour créer des applications de traitement de données.

Il autorise le traitement rapide et parallèle de gros ensembles de données sur des clusters de nœuds.

Ce framework a été créé par Google il permet :

- Le mapping pour séparer les données à traiter
- La réduction pour analyser les données

LES LOGICIELS POUR LE BIG DATAT



OpenRefine est également open source, spécialement conçu pour les données désordonnées.

Il permet de nettoyer rapidement des ensembles de données et de les transformer dans un format exploitable.

Des utilisateurs sans compétences techniques peuvent se servir de cette solution.

OpenRefine permet aussi de créer instantanément des liens entre les ensembles de données.



C'est un open source capable de prendre en charge des données non structurées, tels que des fichiers texte, des logs de trafic et des images.

C'est une plateforme de science des données reposant sur la programmation visuelle pour les opérations.

Les points forts sont :

La manipulation, l'analyse, la création de modèles, l'intégration rapide dans les processus de business process



Base de données NoSQL open source hautes performances, disponibilité élevée, forte scalabilité.

Adaptée aux langages de programmation JavaScript, Ruby et Python.

MongoDB est facile à installer, à configurer, à maintenir et à utiliser.

Data Analytics & Data Mining

« Les **Data Miners** trient de vastes ensembles de données en utilisant des logiciels sophistiqués pour identifier des patterns (modèles) non découverts et établir des relations cachées.

Le **Data Analytics** se concentre sur **l'inférence et sur le procédé permettant de tirer une conclusion** basée uniquement sur ce qui est déjà connu par le chercheur. »

Le data mining permet de mieux comprendre les liens entre des phénomènes en apparence distincts et d'anticiper des tendances encore peu discernables.

Exemples :

Comportement des acheteurs
Caractéristiques des produits
Historisation de la production



5S – Data Mining

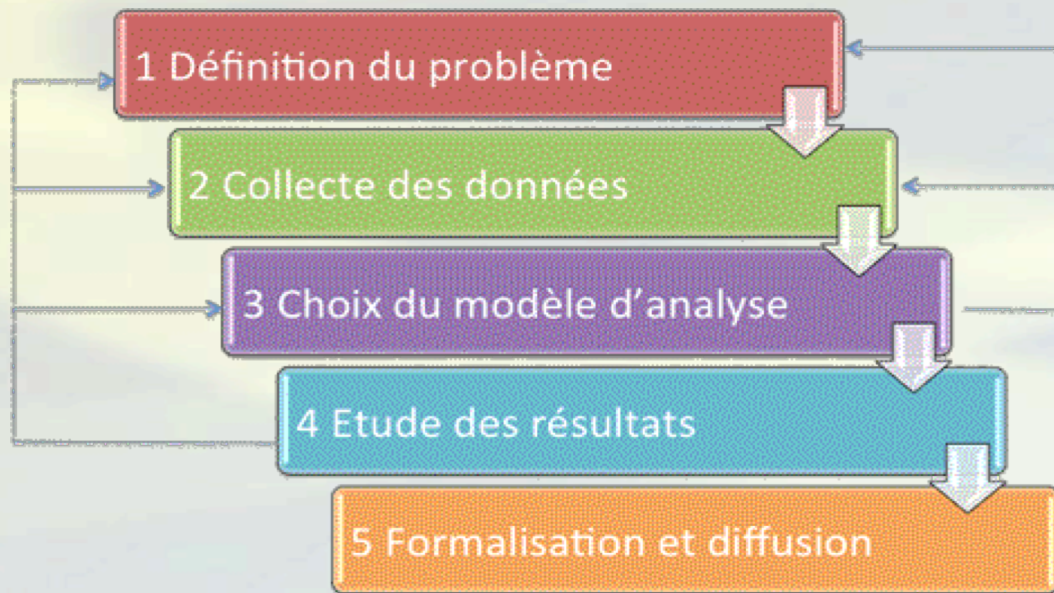
Nettoyer les données avant toute démarche Data Mining

Data Mining = Comportement des utilisateurs et des clients



**Le Data Mining ne dit
cependant pas
POURQUOI ?**

PROCESSUS DATA MINING



Comment mettre en œuvre le DATA MINING ?

L'utilisation de différentes techniques dépend de la nature des données et du type d'étude que l'on souhaite mener.

- Méthodes utilisant des principes d'arbres de décision assez proches des techniques de classification
- Méthodes fondées sur des principes et des règles d'associations ou d'analogies
- Méthodes exploitant les capacités d'apprentissage des réseaux de neurones
- Algorithmes génétiques pour les études d'évolution de populations

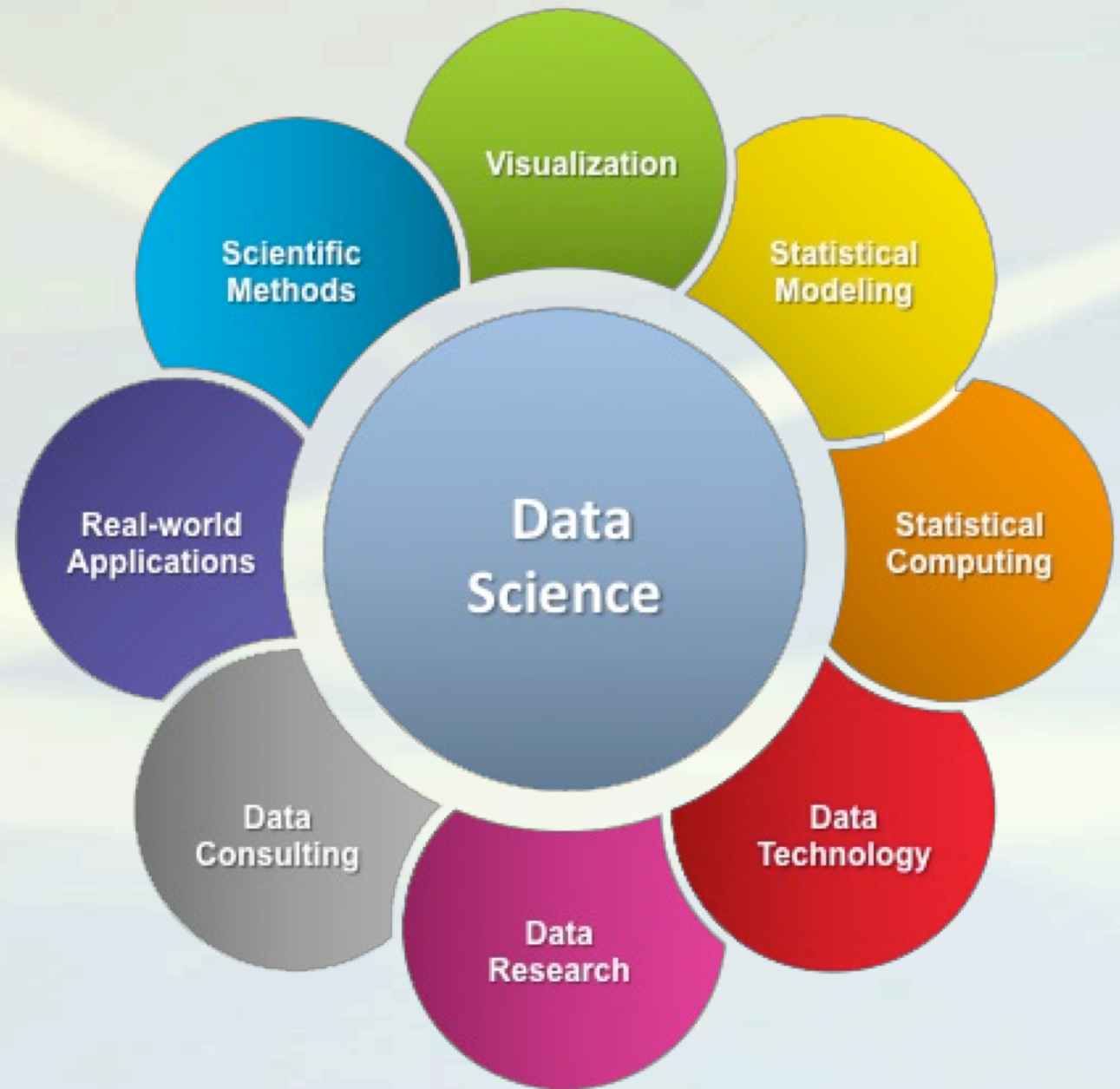
Algorithmes Naïve Bayes (modèle à caractéristiques statistiquement indépendantes), séries chronologiques, régression linéaire...

LA DATA SCIENCE

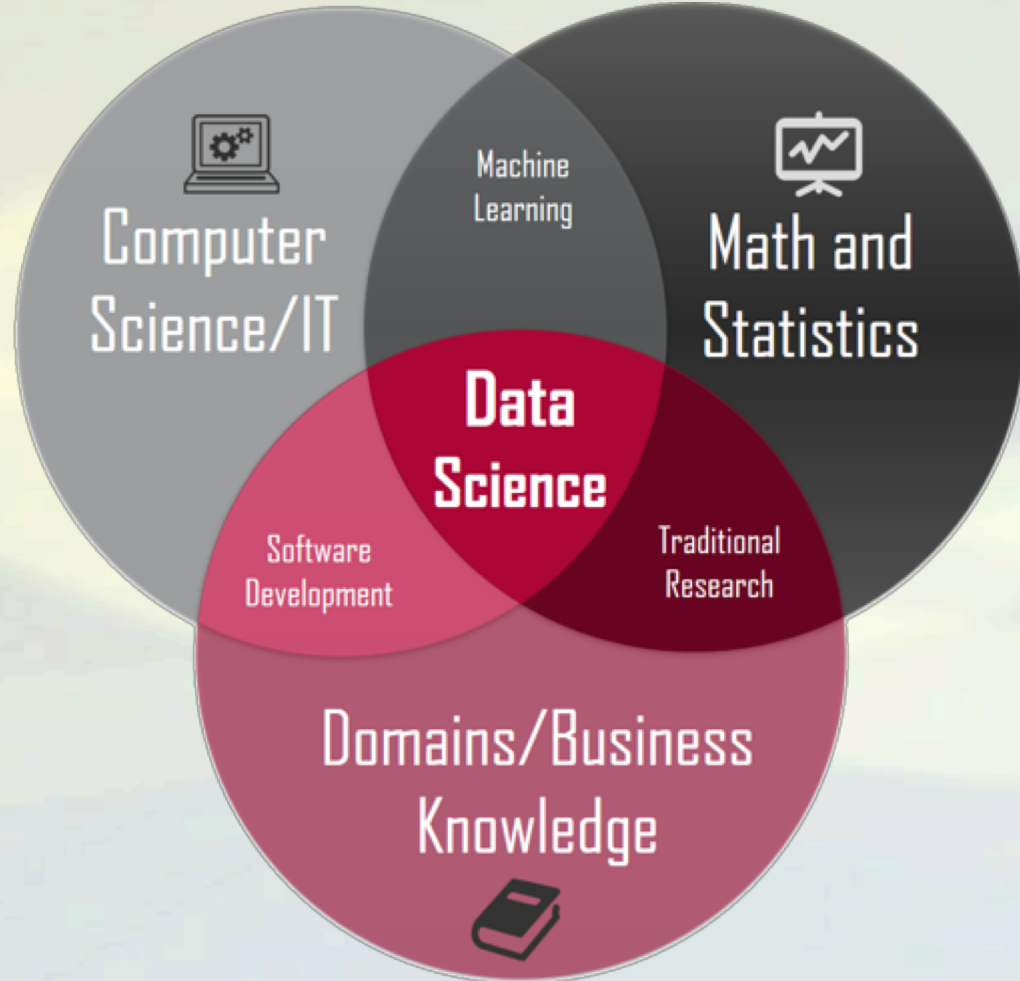
C'est l'ensemble des méthodes scientifiques, des mathématiques, statistiques, et autres outils utilisés sur les ensembles de données pour en extraire des connaissances et des idées.

La Data Science extrait et utilise les informations les plus pertinentes du Big Data.

Les Data Scientists sont les experts concernés.



Data Science **VS** Data Analytics



La data science est un secteur primordial pour le machine learning et l'intelligence artificielle.
Explorer, Connecter de gigantesques quantités de données.

KAGGLE acteur de la Data Science, acheté par Google

Kaggle est une communauté en ligne organisant des compétitions de data science et de machine Learning

<https://www.kaggle.com>

DS et DA permettent d'explorer du Big Data pour trouver des informations utiles.

La différence principale n'est autre que la façon de procéder et l'objectif à atteindre.

Big Data VS Intelligence Collective

L'intelligence Collective est comme son nom l'indique une intelligence de groupe, c'est-à-dire une intelligence partagée grâce à une collaboration étendue.

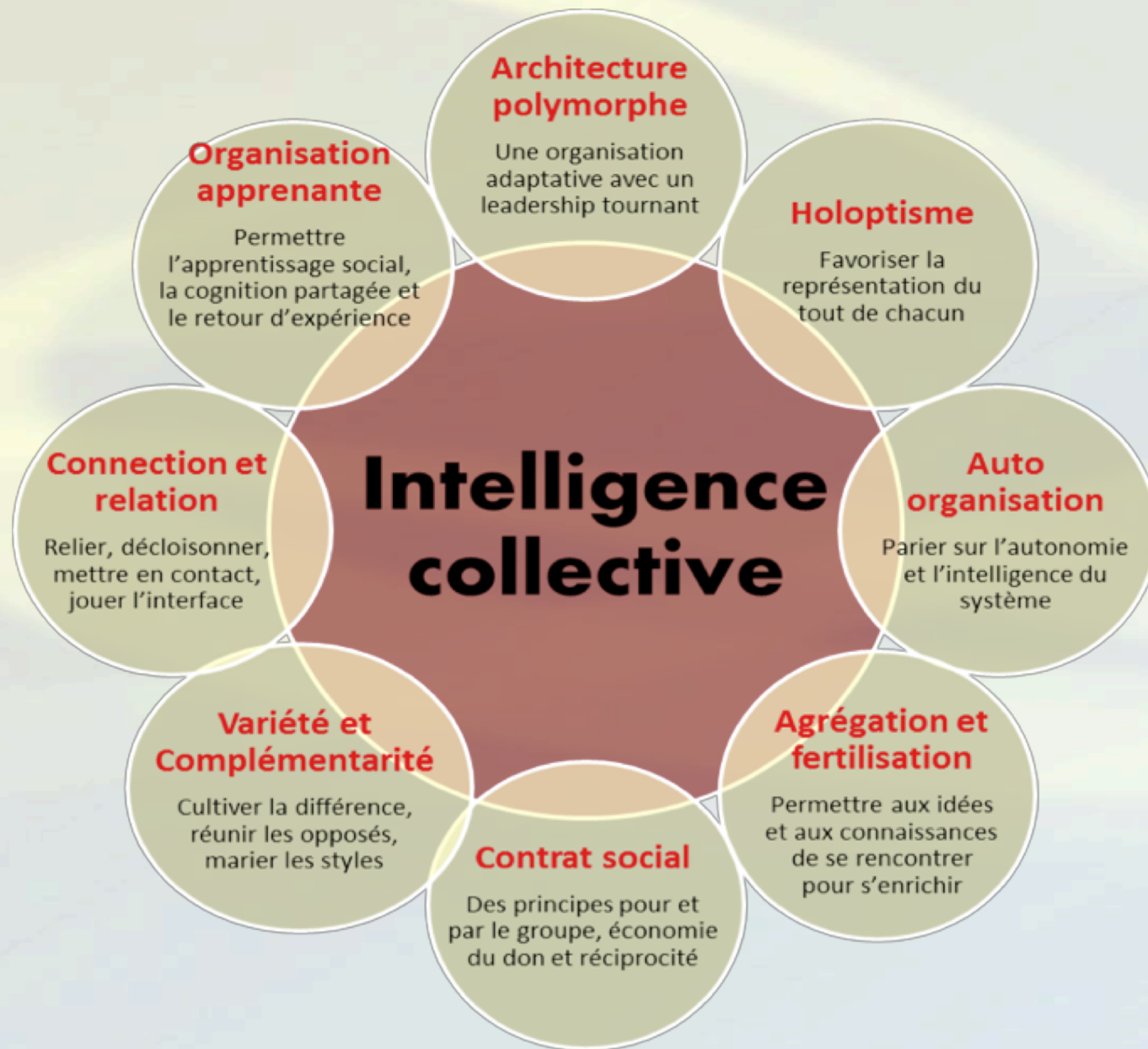
Dans ce cadre les décisions seront prises par consensus.

On peut évoquer le terme « d'intelligence Open Source »

<https://www.journaldunet.fr/management/guide-du-management/1204891-l-intelligence-collective/>



Le concept d'intelligence collective remonte à 1785. Le **Marquis de Condorcet** souligne que si chaque membre d'un groupe a de plus fortes probabilités de ne pas prendre une décision correcte, la probabilité pour que le vote majoritaire de ce groupe soit la décision correcte augmente avec le nombre de membres du groupe. Il s'agit du théorème du jury.



Exemple d'une situation pour laquelle les prédictions réalisées par l'intelligence collective s'avèrent plus utiles que les prévisions menées par le Big Data

Dans les années 2000, la demande mondiale en produits laitiers a soudainement été multipliée par trois en l'espace de quelques mois. Après une décennie de stabilité, les producteurs de produits laitiers ne pouvaient plus se baser sur les modèles prédictifs basés sur les données. Par conséquent, **les acteurs de l'industrie se sont appuyés sur l'intelligence collective des agriculteurs, plus proches des consommateurs, pour mieux comprendre et modéliser les facteurs de cette nouvelle demande.**

Intelligence Artificielle



Intelligence Collective

IEML Information Economy Meta Language

Logiciel d'intelligence collective qui implémente le langage IEML

Logiciel développé par Pierre Lévy

« Son **objectif** n'est pas de rendre les ordinateurs plus intelligents, mais d'utiliser les ordinateurs pour rendre les humains plus intelligents. »



Nombreux détracteurs de cette approche

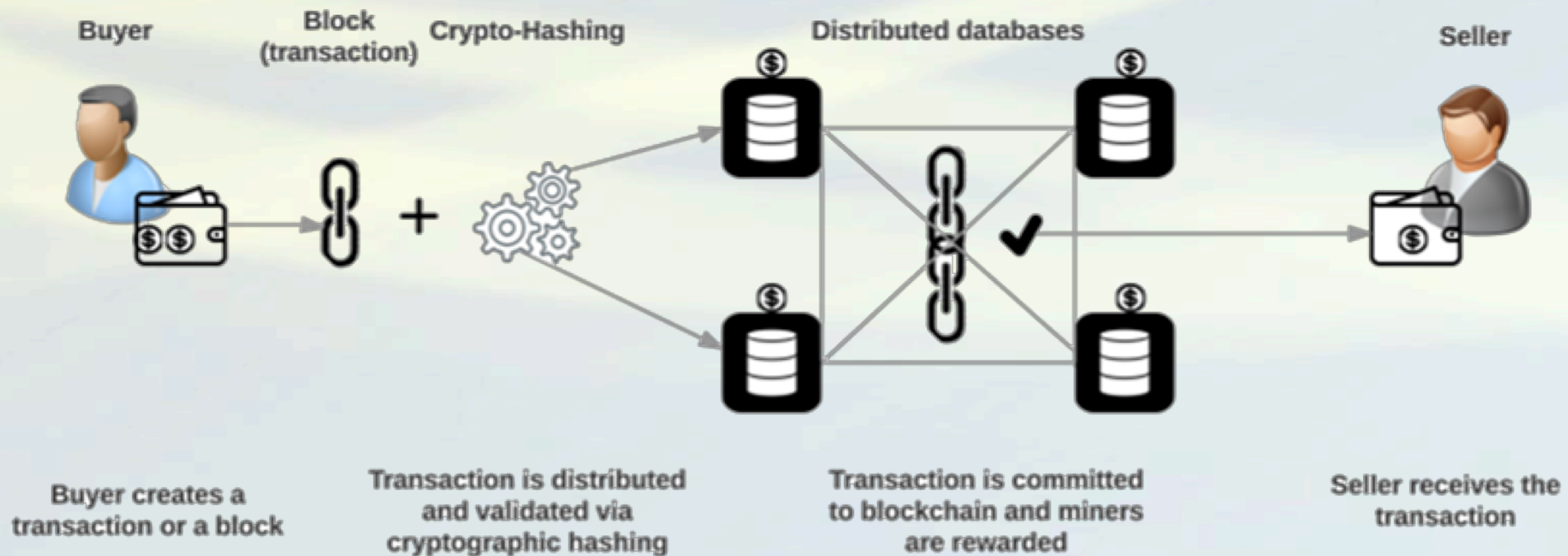
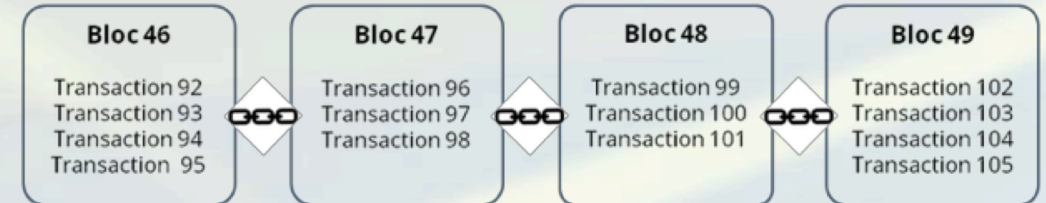
SEANCE B

BLOCKCHAIN

Qu'est-ce que la blockchain ?

La blockchain est une technologie de stockage et de transmission d'informations, transparente, sécurisée, et fonctionnant sans organe central de contrôle (*définition de Blockchain France*).

Par extension, une blockchain constitue une base de données qui contient l'historique de tous les échanges effectués entre ses utilisateurs depuis sa création. Cette base de données est sécurisée et distribuée : elle est partagée par ses différents utilisateurs, sans intermédiaire, ce qui permet à chacun de vérifier la validité de la chaîne.



Comment ça marche ?

Une blockchain publique fonctionne avec une monnaie ou un token (jeton) programmable.

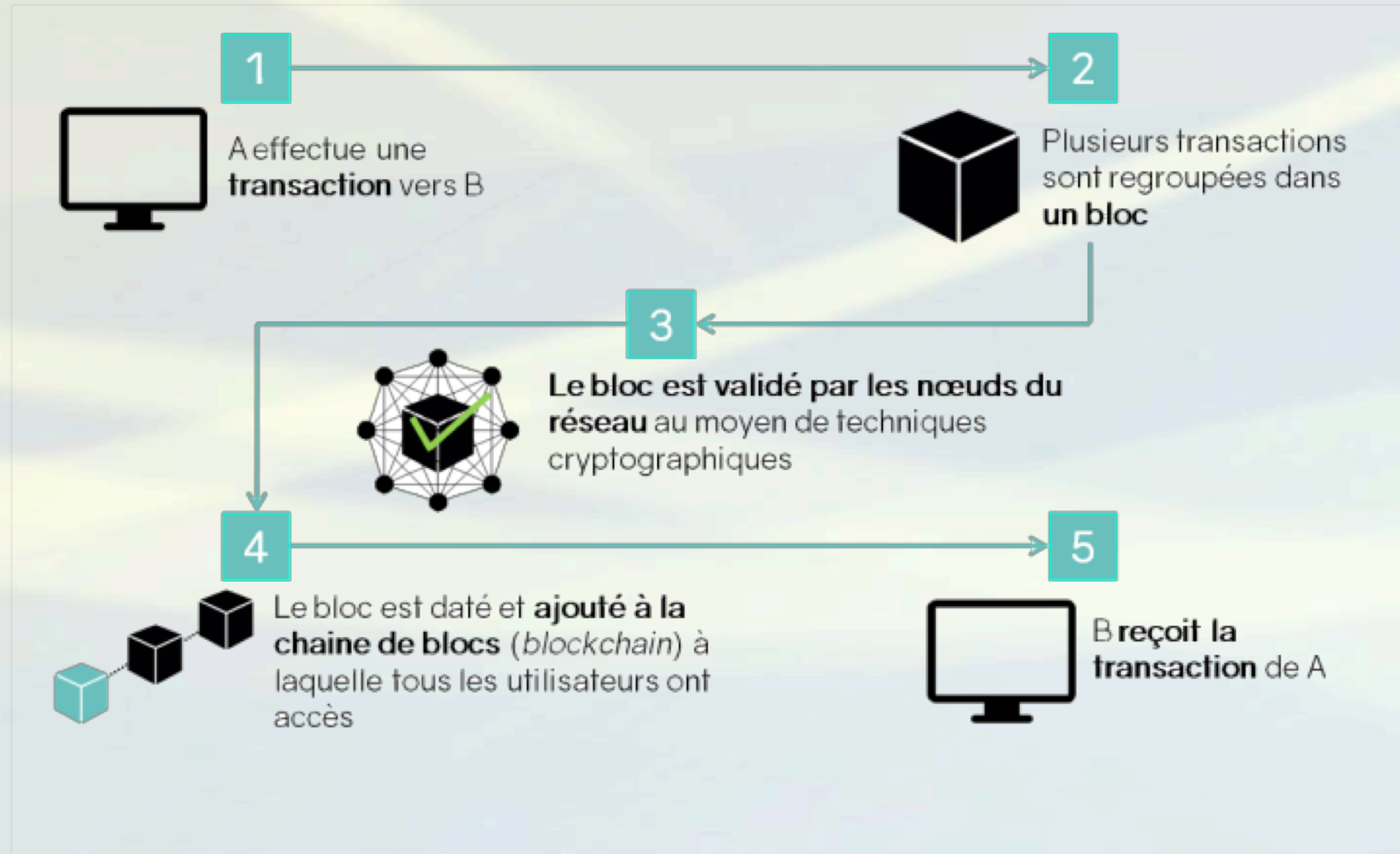
Bitcoin est un exemple de monnaie programmable.

Les transactions effectuées entre les utilisateurs du réseau sont regroupées par blocs.

Chaque bloc est validé par les nœuds du réseau appelés les “mineurs”, selon des techniques qui dépendent du type de blockchain.

Dans la blockchain du bitcoin cette technique est appelée le “Proof-of-Work”, preuve de travail, et consiste en la résolution de problèmes algorithmiques.

Une fois le bloc validé, il est horodaté et ajouté à la chaîne de blocs. La transaction est alors visible pour le récepteur ainsi que l'ensemble du réseau.



POTENTIEL DE DEVELOPPEMENT DE LA BLOCKCHAIN

<https://blockchainfrance.net>

Sécurité & transparence, la blockchain ne se limite pas au secteur de la finance

Principales utilisations de la blockchain

- Les applications pour le transfert d'actifs (utilisation monétaire, mais pas uniquement : titres, votes, actions, obligations...).
- Les applications de la blockchain en tant que registre : elle assure ainsi une meilleure traçabilité des produits et des actifs.
- Les smart contracts : il s'agit de programmes autonomes qui exécutent automatiquement les conditions et termes d'un contrat, sans nécessiter d'intervention humaine une fois démarrés.

Tous les secteurs d'activités sont, seront concernés : banques, assurance, santé et industrie pharmaceutique, supply chain ,agroalimentaire, luxe, commerce international, distribution, vins, aéronautique, automobile..., industrie musicale, énergie, immobilier, vote...

Grâce à la blockchain un nouveau web émerge: **le web décentralisé**

Une nouvelle économie apparaît : l'économie numérique, **la token économie**

Aller plus loin sur le sujet des cryptoactifs

<http://blockchainpartner.fr/crypto-nos-10-propositions-france/>

LOGISTIQUE, TRANSPORTS, SUPPLY CHAIN & BLOCKCHAIN

La blockchain permet le stockage et la transmission d'informations de manière transparente, sécurisée, sans intermédiaire et en temps réel.

- La gestion des flux financiers : vérification financière des partenaires, désintermédiation, cohérence en temps réel entre les flux physiques et financiers, etc.
- La gestion des flux d'informations : traçabilité (agro-alimentaire , GD Walmart, biens rares), certification de données (import/export), instantanéité en cas d'alerte, lutte contre la contrefaçon, suivi des actifs (ex : suivi de containers à Rotterdam), etc.
- La gestion des flux physiques : suppression des flux et des stocks liés au contrôles à réception, facilitation les processus décisionnels (ex : réaction face à un aléa transport), etc.

BIG METAMORPHOSE ...

lorsque la blockchain effacera les intermédiaires.

Dans ce cas l'économie circulaire et collaborative générera davantage de marge.

« La blockchain instaure la confiance, sécurise et trace la donnée... elle fait donc de plus en plus parler d'elle.
Alors... la blockchain marquera-t-elle la fin des intermédiaires, d'Uber et des ERP ? »

L'AVENIR DE LA BLOCKCHAIN

Naissance de marketplaces
où les individus et les
entreprises s'adonneraient
au commerce de données



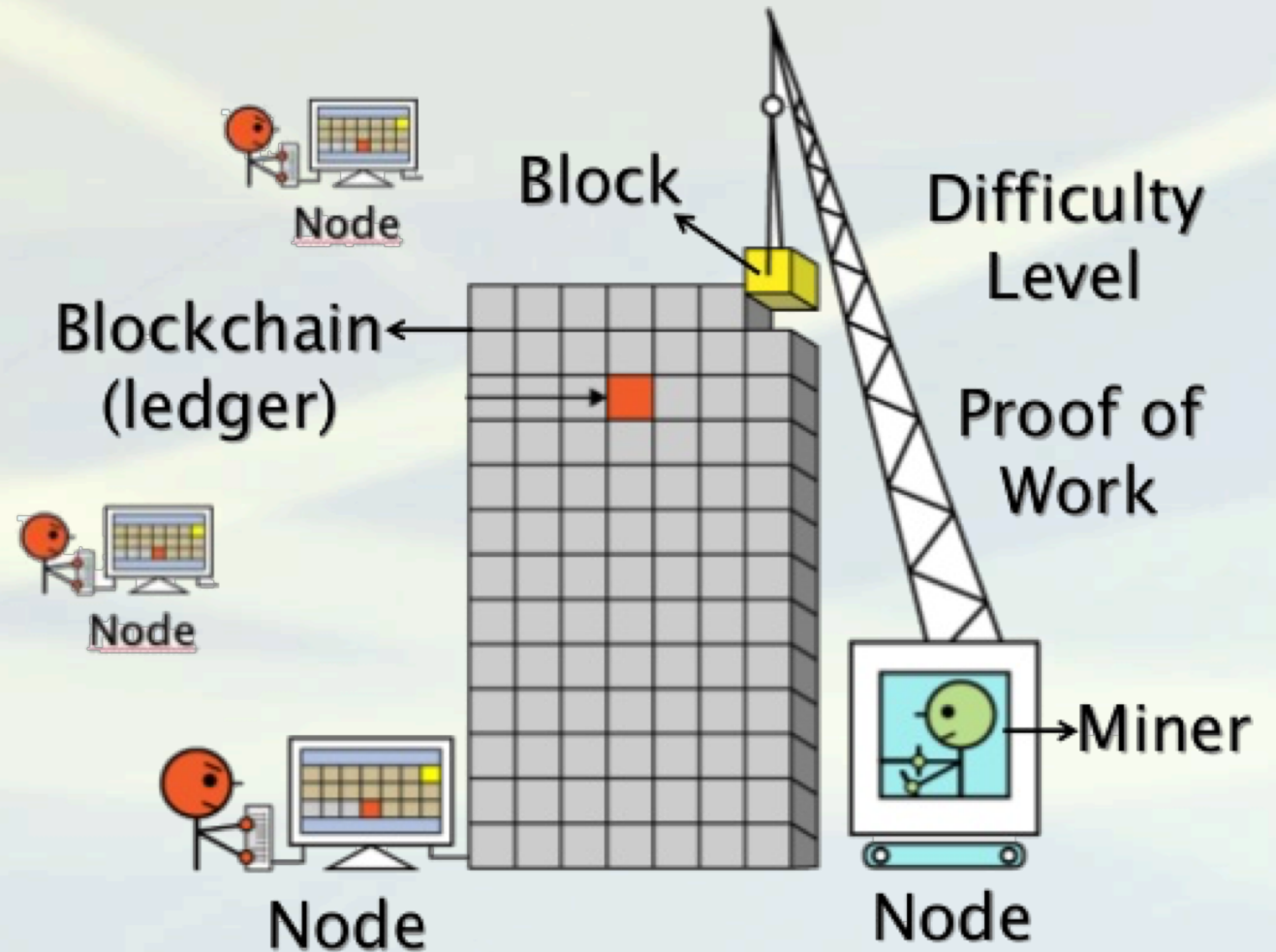
La Blockchain pourrait
représenter **20%** du marché
Big Data en 2030, avec un
revenu annuel de plus de
100 milliards de dollars

Les analyses Big Data seront alors indispensables

Emergence de nouveaux services de
Data Intelligence

COMMENT DECELER SI UNE ENTREPRISE EST PRETE POUR LA « BLOCKCHAIN » ?

- **Ordinateur partagé** accessible à tous
- **Une seule et même source d'informations** fiables au sein de laquelle il est possible de stocker les événements, les participations détenues, les activités,
- **Exécuter des flux de travail** impliquant une multitude de parties sans avoir recours à des systèmes et des bases de données distincts, et sans la nécessité de procéder à une réconciliation.



Autre révolution en cours, la *blockchain*, cette technologie qui sous-tend les monnaies cryptographiques comme le bitcoin, tisse sa toile. En tout cas, telle est l'ambition de la startup parisienne Ownest.

« Au départ, il s'agit de responsabiliser les acteurs de **votre supply chain en traçant les actifs en temps réel** », indique Clément Bergé-Lefranc, cofondateur et président de cette jeune société - elle a été créée l'année dernière -, qui a décroché Carrefour comme premier client. L'enseigne doit gérer chaque jour 10.000 rolls-conteneurs qui transitent en Île-de-France.

« Au lieu de leur coller des étiquettes électroniques, on leur applique le principe de la "patate chaude" ! », sourit Clément Bergé-Lefranc.

En d'autres termes, **lorsqu'un chauffeur arrive dans un entrepôt pour prendre cinq rolls, le responsable d'expédition lui transfère sur son smartphone cinq « unités de responsabilité blockchain ». À son tour, en restituant à Carrefour les cinq rolls physiques, le chauffeur se débarrassera de ses cinq unités de responsabilité.** Et le tour est joué.

« Chacun est en concurrence pour partager l'objectif commun de faire réussir la transaction », poursuit le jeune dirigeant qui s'appuie sur des blockchains publiques, comme Ethereum Classic ou Litecoin, tout en restant prêt à changer de système si une meilleure technologie open source arrive sur le marché.

De fait, cette architecture, qui trace non pas des produits en direct mais des responsabilités, s'annonce comme **une solution prometteuse pour remplacer les EDI** (échanges de données informatisés), inaccessibles auprès des PME et des TPE. Et Clément Bergé-Lefranc d'ajouter :

« Nous travaillons avec le ministère des Transports ainsi qu'avec les principales fédérations de transporteurs pour étudier la possibilité de généraliser ce type de solution. »