

Evolution du SI

V. 1.0 juillet 2019

Evolution du SI

- La fonction IT
- La stratégie IT
- L'urbanisation
- Les principes d'évolution des SI
 - cartographie du système d'information
 - Open data, Open innovation, Bring Your Own Device

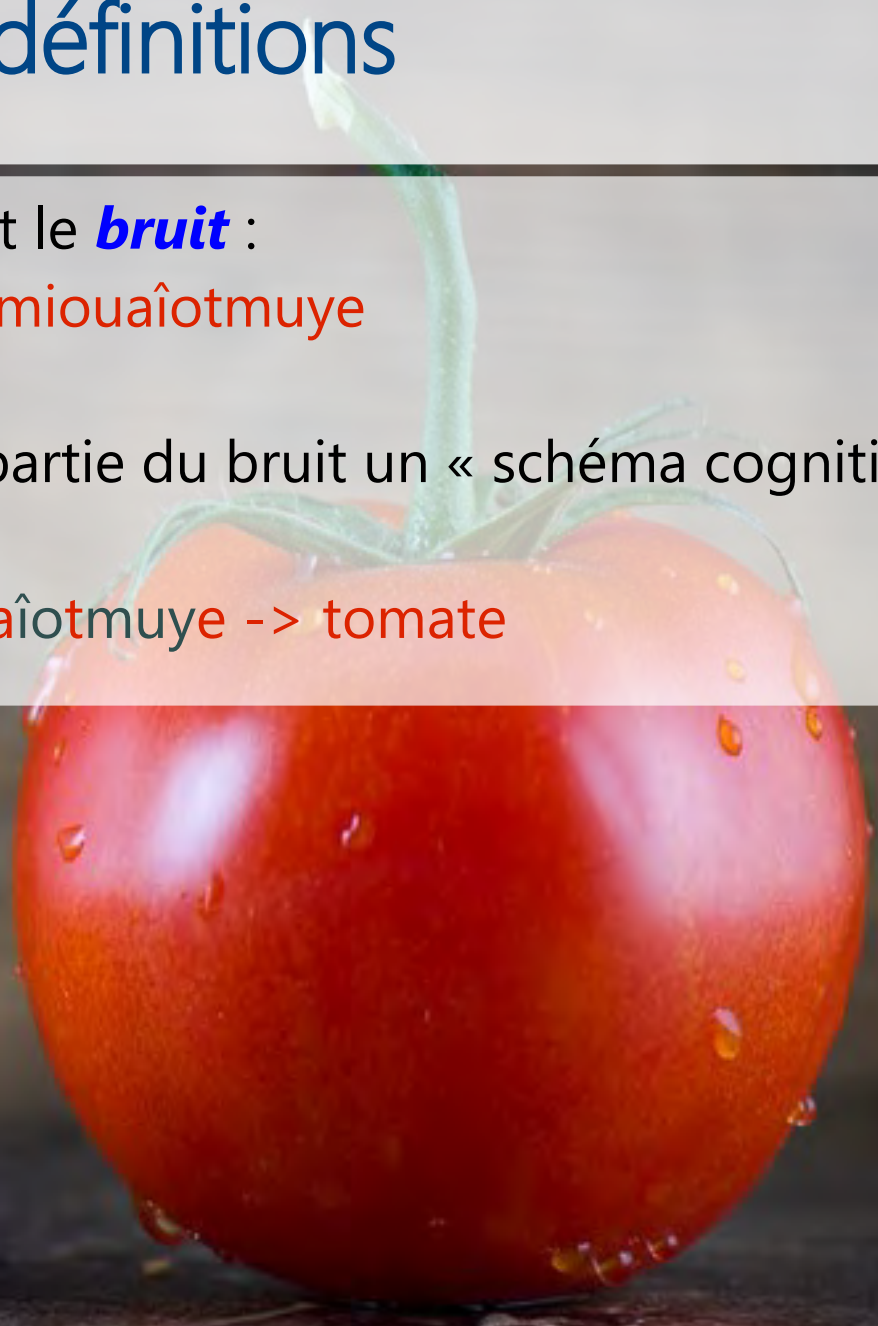
Data : définitions

Le niveau zéro de l'information est le **bruit** :

tzeoqzemiouaîotmuye

Lorsque l'on peut associer à une partie du bruit un « schéma cognitif » on obtient une **donnée** :

tzeoqzemiouaîotmuye -> tomate



Data : définitions

Les données deviennent **information** lorsqu'elles sont assemblées en un tout cohérent, qui peut être relié à d'autres informations : **une tomate est un fruit**

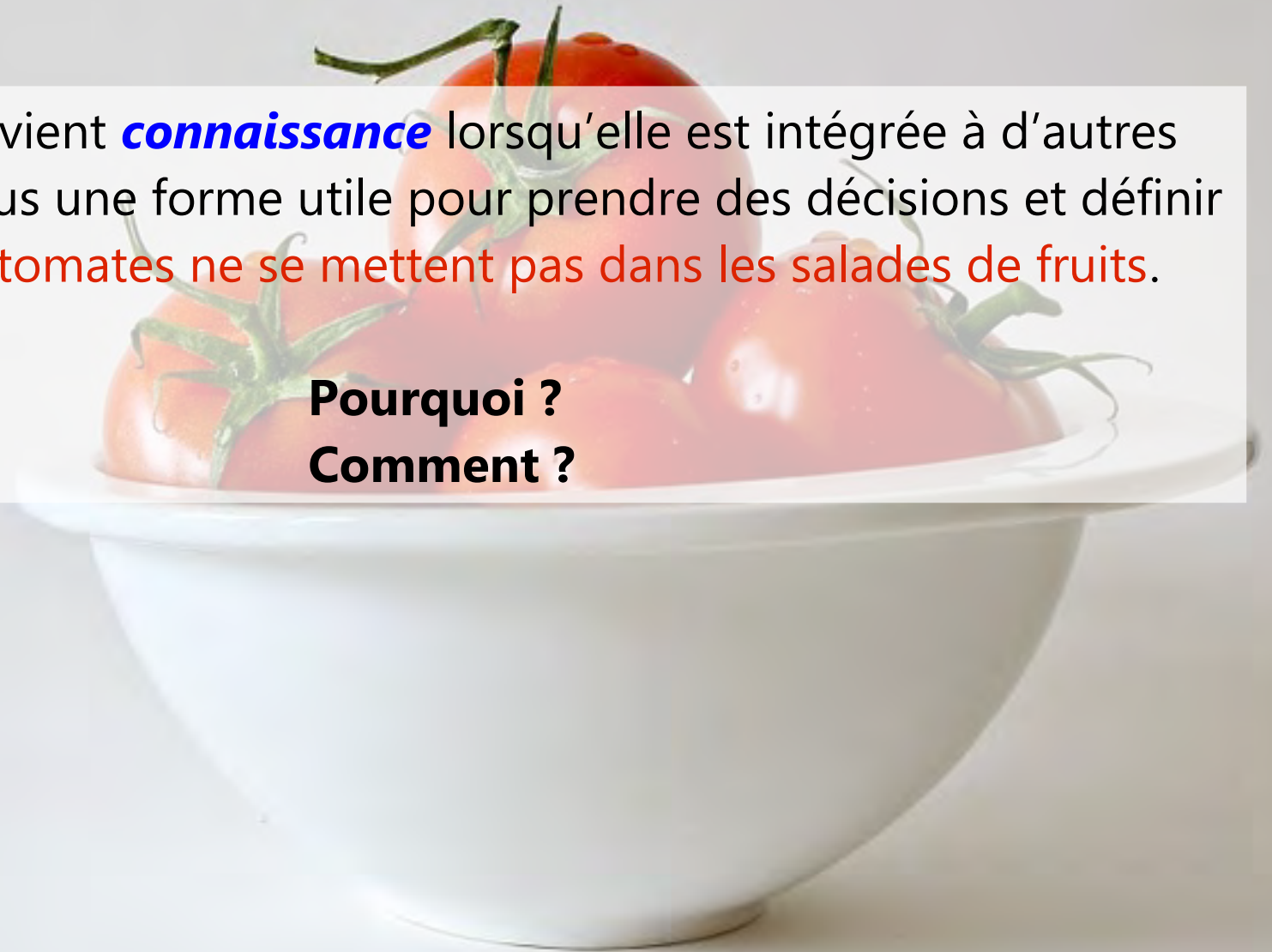
Qui ?
Quoi ?
Quand ?
Où ?

Data : définitions

L'information devient **connaissance** lorsqu'elle est intégrée à d'autres informations, sous une forme utile pour prendre des décisions et définir des actions : **les tomates ne se mettent pas dans les salades de fruits.**

Pourquoi ?

Comment ?



Data : définitions

La connaissance devient **compréhension** lorsqu'elle est reliée à d'autres connaissances, d'une manière qui permette d'anticiper, de juger et d'agir : les tomates font de mauvaises salades de fruit pour des raisons de goût (pas sucré) et d'habitude culinaire.

**Pour quelles raisons ?
Par quels mécanismes ?**

Data : définitions

La compréhension devient **sagesse** lorsqu'elle est modulée par une finalité, une éthique, des principes, une mémoire et une projection : l'usage des tomates est une illustrations des conventions d'un groupe social, conventions modulées par l'histoire, parfois par les croyances (dogmes), ces conventions pouvant être supprimées lors de grands changements.

**Quelles sont les lois universelles (science) ?
Quels sont les fondements éthiques ?**

Data : définitions

Bruit

Donnée

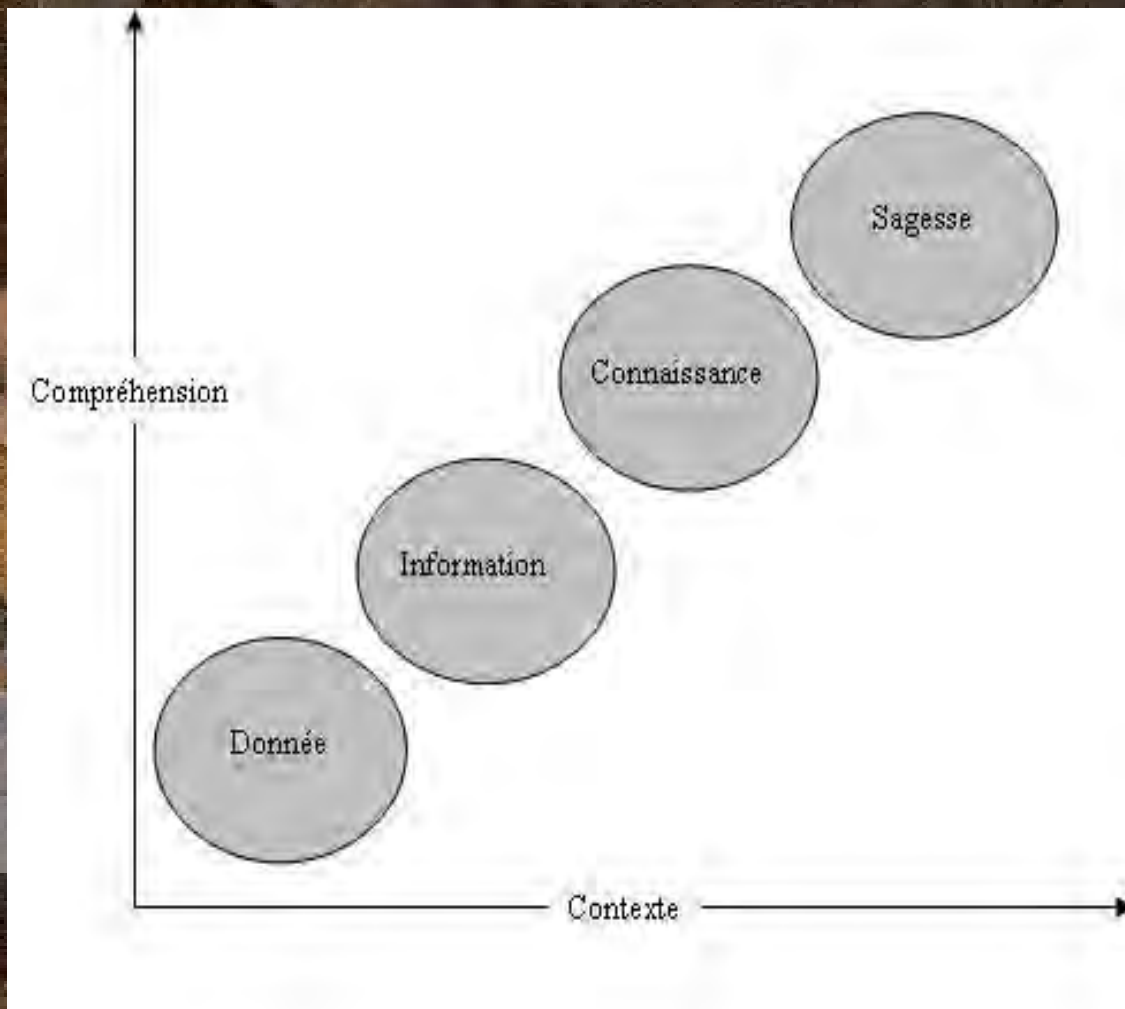
Information

Connaissance

Et sagesse sont des concepts liés les uns aux autres.



Data : définitions

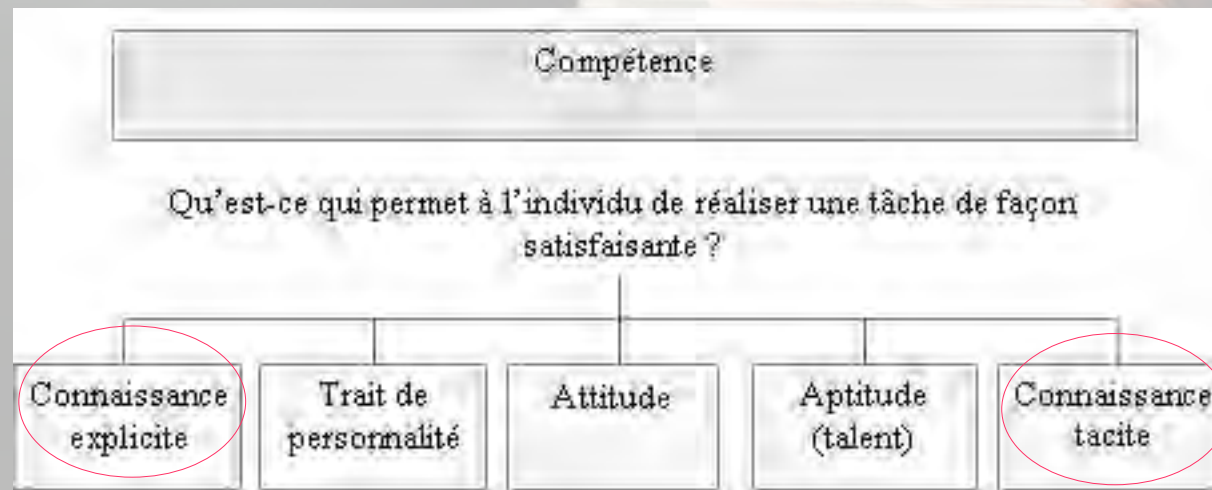


Data : définitions

Connaissance et compétence

« La **compétence** désigne la capacité d'un individu à réaliser de façon satisfaisante une tâche déterminée.

Dans le cadre professionnel, la compétence est considérée comme un ensemble d'aptitudes et de talents, de traits de personnalités et de connaissances acquises pour mener à bien les tâches assignées à un individu [C. Lévy-Leboyer, Le bilan de compétences , Éditions d'Organisation, 1993]. »



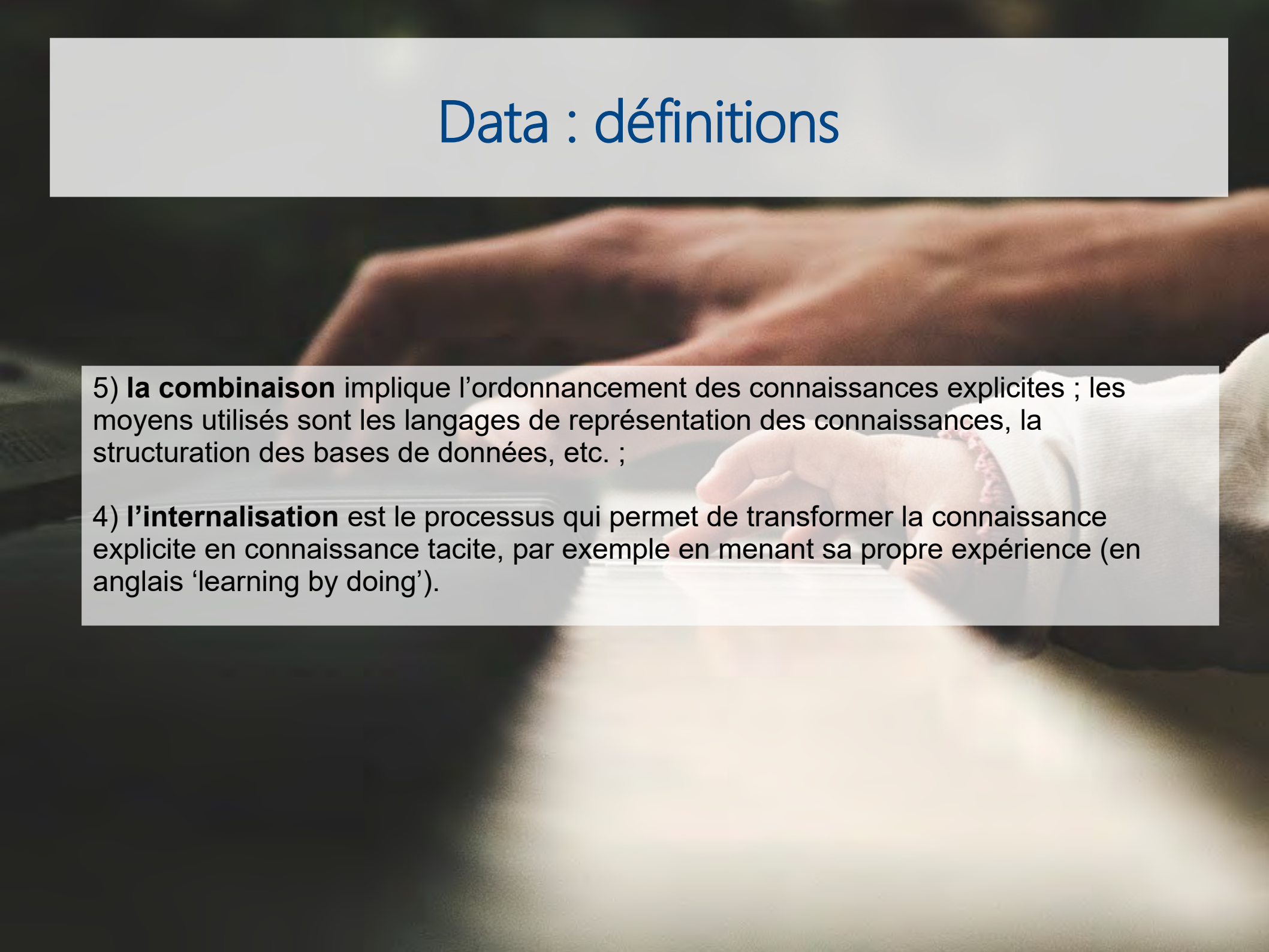
La **connaissance** est liée à l'individu, mais pas seulement.

Data : définitions

Ikujiro Nonaka et Hirotaka Takeuchi ont décrit le processus de création des connaissances comme un phénomène en spirale se déroulant en quatre phases [I. Nonaka, H. Takeuchi, The Knowledge-Creating Company, Oxford University Press, 1995]:

- 1) **la socialisation** permet à un individu d'acquérir des connaissances tacites en étant en contact direct avec d'autres personnes ;
- 2) **l'externalisation** est l'articulation des connaissances tacites en connaissances explicites : les personnes formalisent les connaissances en exploitant des métaphores, en développant des histoires, en concevant des modèles, en ayant recours à des analogies ; dès que la connaissance est explicitée, elle peut être modélisée et partagée sous une forme tangible;

Data : définitions

- 
- 5) **la combinaison** implique l'ordonnancement des connaissances explicites ; les moyens utilisés sont les langages de représentation des connaissances, la structuration des bases de données, etc. ;
- 4) **l'internalisation** est le processus qui permet de transformer la connaissance explicite en connaissance tacite, par exemple en menant sa propre expérience (en anglais 'learning by doing').

Data : définitions

L'organisation informatique des données : le stockage « simple »

L'un des objectifs des systèmes d'exploitation est de gérer les données à travers des systèmes de fichiers :

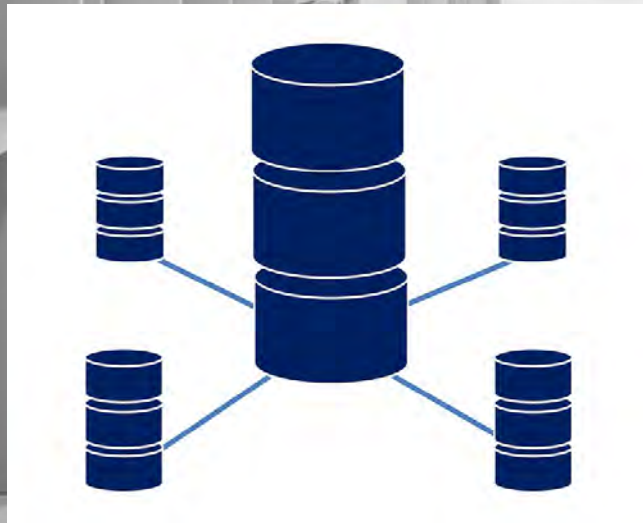
- Un ensemble cohérent de bits (0/1) codant l'information crée un fichier
- Les fichiers sont organisés en arborescence
- Il existe des moyens de « naviguer » sur cet arborescence pour retrouver les fichiers à partir de leur nom.

Ce système montre cependant rapidement ses limites. Il est alors utile d'ajouter des outils supplémentaires pour rechercher ses fichiers...

Data : définitions

Mais à nouveau ce n'est pas suffisant : rechercher un fichier ne veut pas dire être capable de faire un traitement sur l'information qu'il contient et le nom de fichier est une information très pauvre pour retrouver quelque chose dont on n'a qu'une vague idée.

Deux grands types de solutions sont possibles selon que l'on choisit d'organiser a priori les données ou pas.



Data : bases de données

L'organisation informatique des données : le stockage en base de données classiques

Avant de stocker les informations, on va faire un travail d'analyse, de typologie, de mise en évidence de relations entre données, etc.

Puis on va synthétiser ce travail dans un « dictionnaire » qui va définir la structure de stockage des données.

Ce dictionnaire sera figé mais son intégrité et sa cohérence garanties.

C'est le modèle des SGBDR (système de gestion de base de données relationnelles).

Data : SGBDR

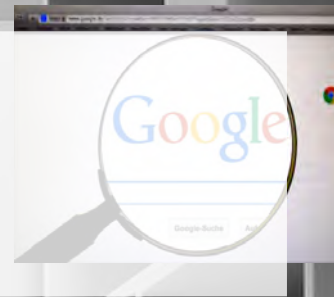
Les bases de données relationnelles

Parmi les composants logiciels les plus importants du système d'information : ils stockent les données qui sont la vie de l'entreprise.

Le premier SGBD a été créé en 1960 et la technologie s'est répandue à partir des années 1970. COBOL est un langage adapté aux bases de données a été créé en 1965.

Le modèle relationnel (SGBDR) a été créé en 1970 et les premiers SGBDR sont apparus en 1978 : IBM System R et Oracle V2.

Data : stockage internet



L'organisation informatique des données : le stockage «internet»

Internet permet de retrouver des fichiers à travers un moteur de recherche complété par un mécanisme d'enrichissement de l'information stockée.

Un robot va « visiter » les fichiers sur le réseau, il les analyse et tente de leur assigner des informations supplémentaires les plus pertinentes possibles sous forme de « tags » (**mot clés**).

Ensuite, le moteur de recherche va essayer d'analyser le contexte de la requête (tags de la recherche) et va essayer de retrouver les informations les plus pertinentes liées à ces tags de recherche.

Ici, l'humain a besoin de la machine car ses capacités d'analyse a priori ou de recherche a posteriori sont définitivement insuffisantes.

Data : stockage internet



Les bases de données utilisées dans ce genre de cas diffèrent des SGBDR pour des raisons de rapidité et de nombre d'accès concurrents.

On parle alors de bases de types No SQL (not only SQL).

Ces bases fonctionnent en « agrégats » sur un grand nombre de machines en même temps.

Parmi ces bases on compte : Cassandra, Couchbase, Elasticsearch, Hbase, MongoDB, Redis, Riak, ...

Data : du « gros » au « big »



Depuis l'arrivée d'internet est son explosion vers 2008, un nouveau problème s'est posé : comment stocker et valoriser une masse toujours plus grande de données hétérogène ?

C'est le domaine du Big Data.

Data : du « gros » au « big »

Le Big Data se caractérise par ses 5 V :

- * Le volume des données est considérable
- * La vitesse / vélocité à laquelle il faut intégrer ces données est un défi
- * La variété / hétérogénéité est la règle
- * La véracité des données est problématique
- * La valeur est ce que l'on cherche à trouver

Toutes les données du Big Data arrivent dans un « data lake ».

Sources



Données non structurées



Données semi-structurées



Données structurées



Données personnelles

Data Lake

Big Data
/
Hadoop

SQL

NoSQL

Meta-Data Management



Analyse temps réel



Analyse prédictive



Data Discovery

Industrialisation

Traitements ETL



Data Warehouse



Reporting d'entreprise

Data : 3 types de connaissance



Data : 3 types de connaissance

On recense trois grands types de connaissances dans une entreprise.

Ce sont celles-ci qui vont intéresser le ***Knowledge Manager*** :

- Connaissances **stratégiques** : clients, fournisseurs, concurrents
- Connaissances **métiers** : compétences et savoirs faire
- Connaissances **opérationnelles** : savoirs et expériences

Ces connaissances seront déclinées en fonction des entreprises.



donnez des exemples pour un cabinet de RH
et une société de pêche en mer.

Synthèse

Toute entreprise acquière, gère et enrichit de la connaissance (**savoir et savoir-faire**).

Cette connaissance est un des éléments essentiels de la **richesse** de l'entreprise, c'est également un élément de **compétitivité** et d'**innovation**.

L'exploiter correctement (**décisions plus rapides, plus riches et mieux fondées**) est un objectif majeur. C'est l'ambition du Knowledge Management depuis la fin des années 1990.

Synthèses

Pour diriger une entreprise, il faut être capable d'obtenir le maximum d'éléments permettant de faire les bons choix.

Ces éléments peuvent être extraits de l'intérieur de l'entreprise comme de l'extérieur.

Cette extraction est plus ou moins naturelle, plus ou moins maîtrisée.

L'un des buts du management de la connaissance est d'extraire le maximum de choses du « bruit » pour atteindre le niveau maximal de richesse utile à la direction de l'entreprise.

Tout ceci s'appuie sur des outils informatiques.

Open Data : définitions

Dans son Rapport au Premier ministre sur la gouvernance de la donnée, l'Administrateur général des données de l'État propose définition de la donnée numérique :

Une donnée numérique est la description élémentaire de nature numérique, représentée sous forme codée, d'une réalité (chose, événement, mesure, transaction, etc.) en vue d'être :

- collectée, enregistrée ;
- traitée, manipulée, transformée ;
- conservée, archivée ;
- échangée, diffusée, communiquée

Open Data : définitions

Selon leurs destinations, les données peuvent être

fermées (réservées à quelques personnes ou à des organisations),

partagées (leur usage est conditionné par une licence contractuelle ou des conditions générales d'utilisation)

ou **ouvertes** (ouvertes à tous utilisateurs et à tous usages légaux).

Les **données** doivent être traitées, manipulées, transformées ou encore croisées pour aboutir à une **information** qui a du sens. Elles nécessitent une interprétation pour devenir une information.

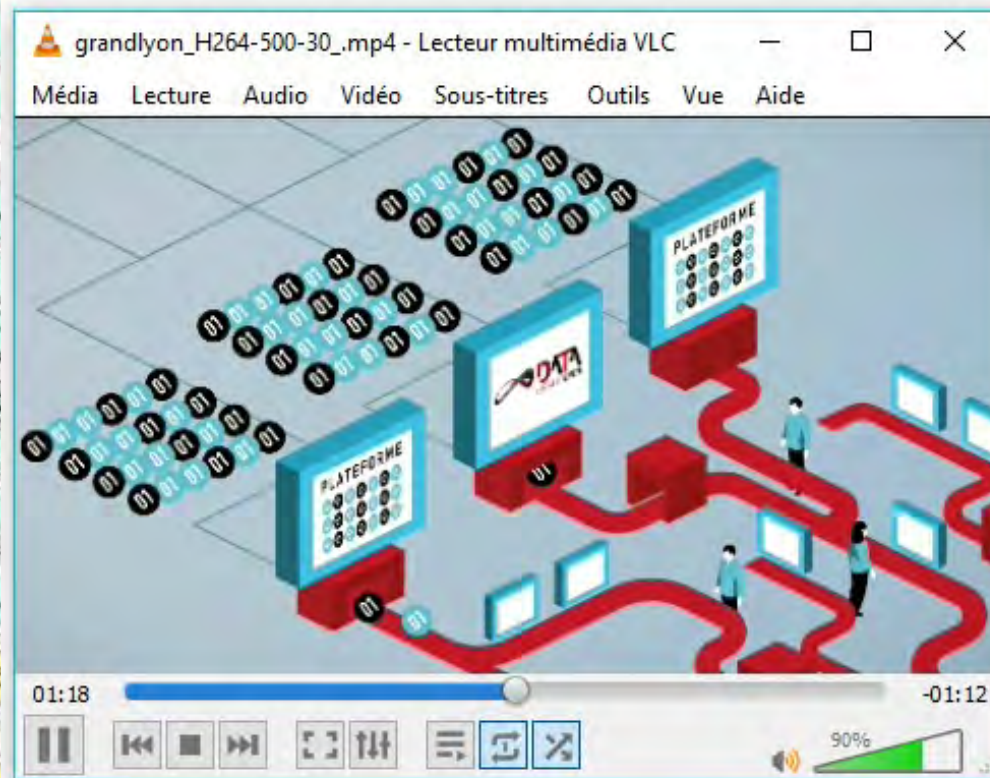
Open Data : définitions



Notre production / collecte journalière de données est énorme.

Sources <https://www.cerema.fr>

Open Data : définitions



C'est un enjeu majeur pour les acteurs publics.

sources <https://www.cerema.fr>

Dans chacune des situations suivantes, indiquez s'il s'agit d'une création de données ou d'une information reçue.



Je valide mon titre de transport.

C'est une



Je consulte la carte de vigilance de Météo-France.

C'est une



Je vois un message affiché sur un panneau de l'autoroute :

C'est une



Un géomètre est au travail dans la rue :

C'est une



Une voiture Street-view circule dans la rue :

C'est une



Ce plan du métro vous indique l'itinéraire à suivre entre deux stations :

C'est une



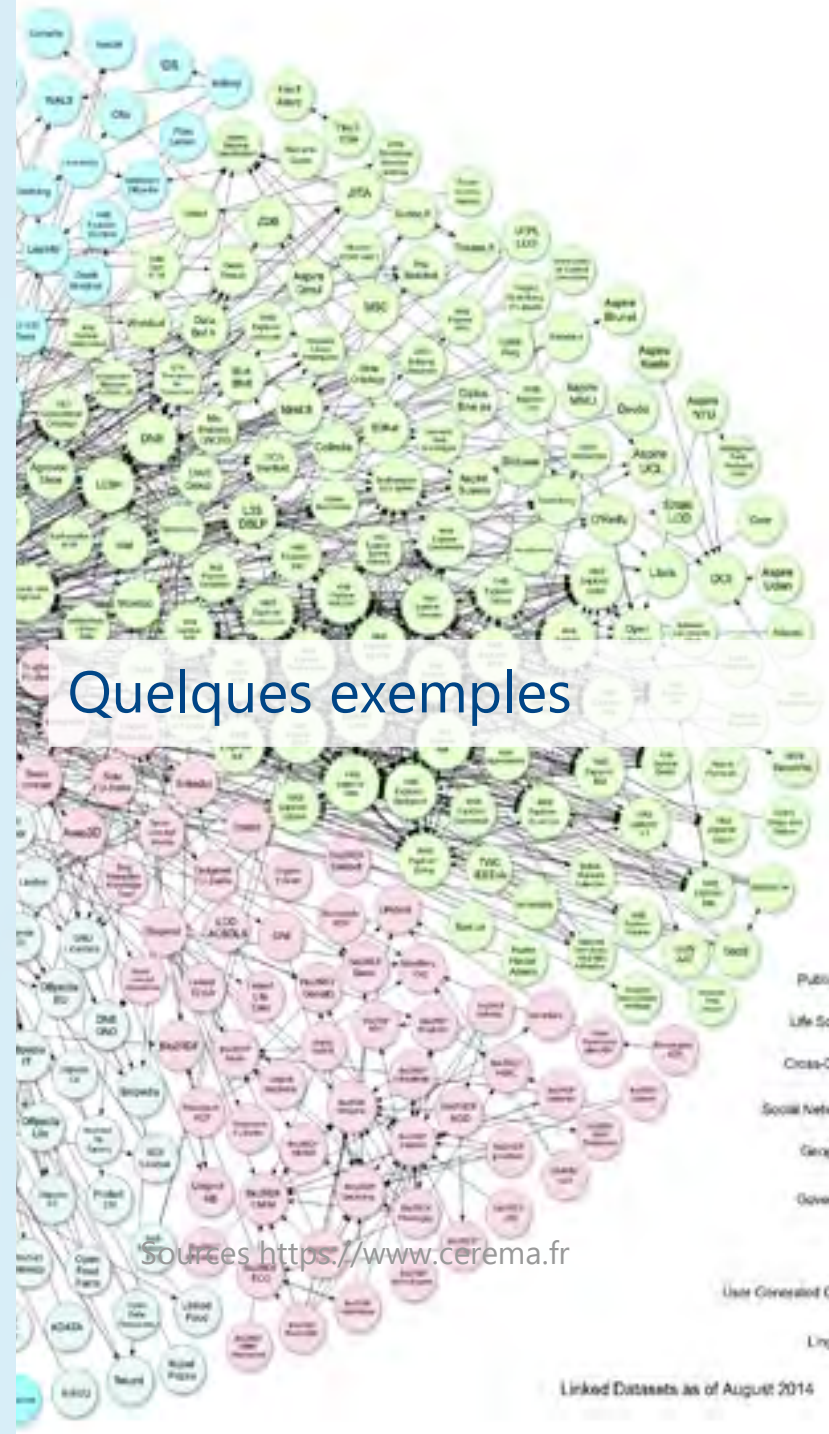
Vous remplissez le formulaire d'inscription sur un site internet :

C'est une



Aie ! Vous avez été flashé sur la route :

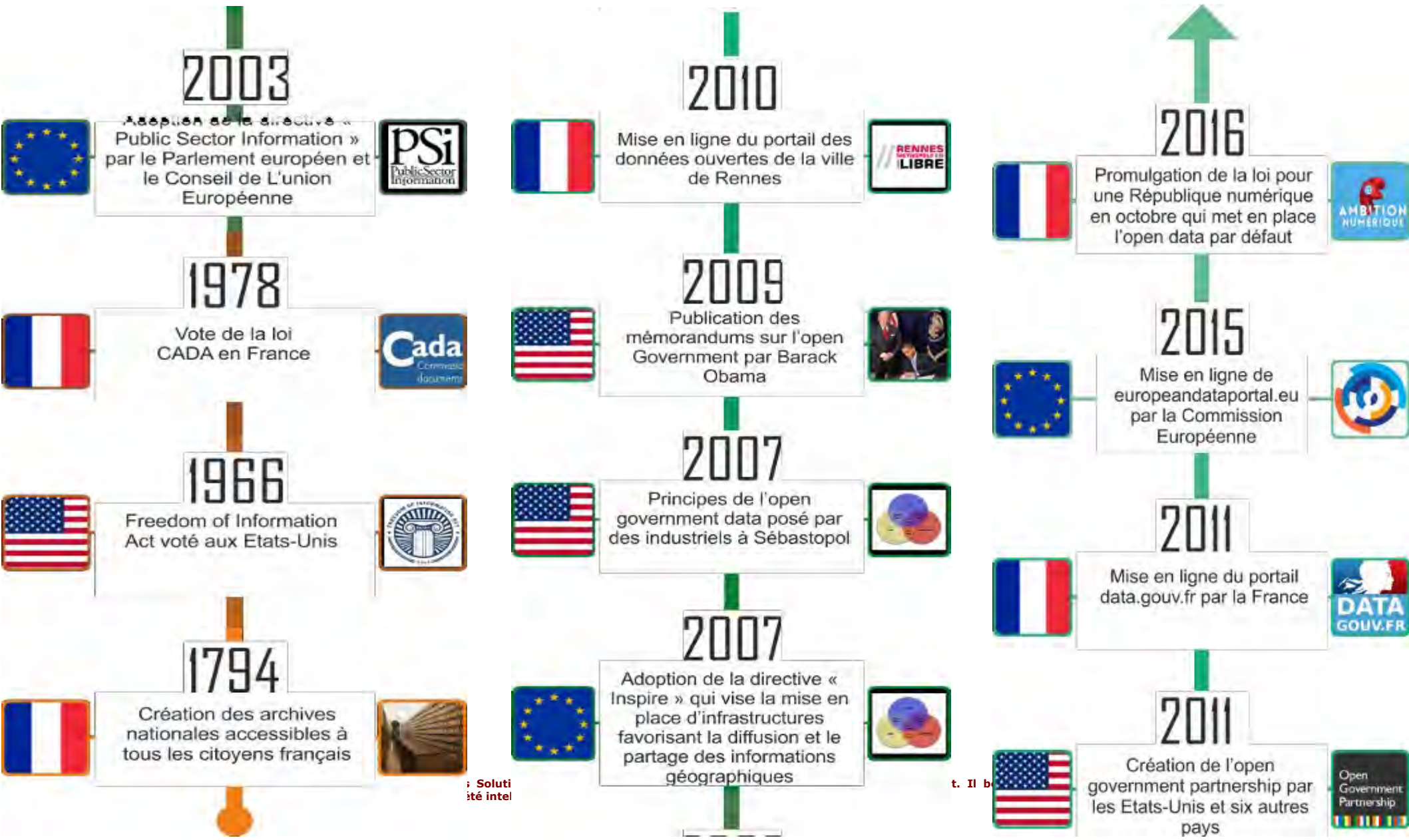
C'est une



Quelques exemples

Sources <https://www.cerema.fr>

Open Data : histoire



Open Data : Enjeux

La démarche Open Data s'appuie sur la **libre disponibilité pour tous des données, sans restriction** de droit d'auteur, brevets ou autres mécanismes de contrôle.

L'Open Data désigne une **dynamique** tendant à favoriser l'accès aux données sous une licence permettant la libre reproduction, redistribution, modification et réutilisation, y compris à des fins commerciales.

Inspirée de l'Open source et initiée à la fin des années 2000, elle tend à se globaliser en couvrant de plus en plus de domaines d'application.

Open Data : Enjeux

Les données sont considérées comme ouvertes, et donc relevant de l'Open Data, si elles répondent à 8 principes définis en décembre 2007

- 1) Les données sont **complètes** : toutes les données doivent être rendues disponibles sauf les données pouvant porter atteinte à la vie privée des citoyens ou à la sécurité.
- 2) Les données sont **primaires** : les données doivent être brutes, telles qu'elles ont été collectées à la source, non agrégées, non modifiées.
- 3) Les données sont **tenues à jour** : elles doivent être rendues disponibles aussi vite que possible afin de préserver leur valeur.
- 4) Les données sont **accessibles** : les données sont disponibles au plus large spectre d'utilisateurs pour l'usage le plus large.

Open Data : Enjeux

Les données sont considérées comme ouvertes, et donc relevant de l'Open Data si

- 5) Les données doivent permettre un **traitement automatisé** : elles doivent être structurées et documentées afin de permettre un traitement informatisé.
- 6) L'**accès aux données est non discriminatoire** : elles sont disponibles à tout le monde de façon anonyme ne nécessitant pas d'enregistrement.
- 7) Les données sont disponibles dans un **format non propriétaire** : elles doivent être rendues disponibles au moins dans un format sur lequel aucune entité ne détient le monopole (ex : non PDF, non Excel).
- 8) Les données sont **libres de droits** : elles ne doivent pas être l'objet de droits d'auteurs, marques déposées, brevets, etc...

Open Data : Enjeux

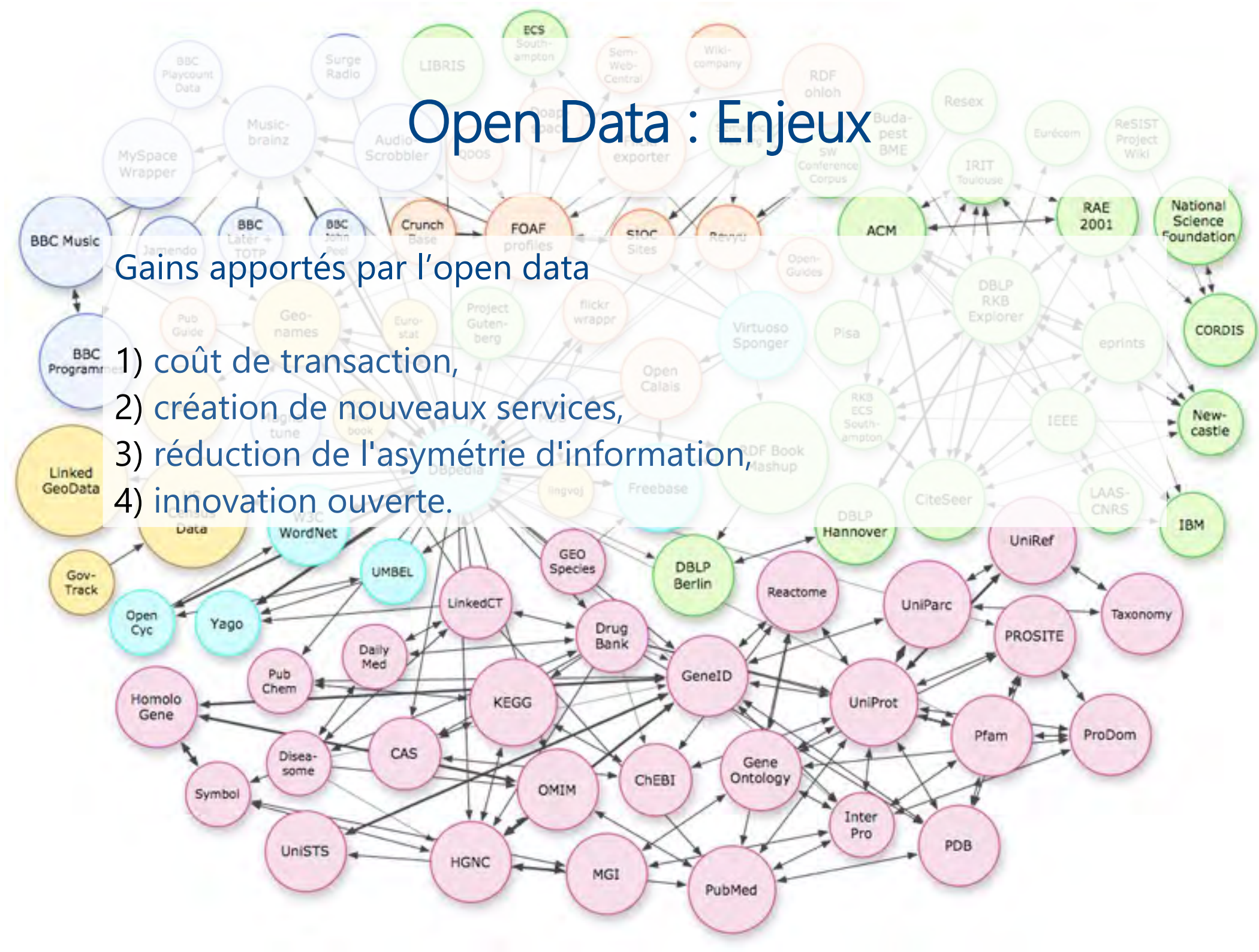
Les trois choses les plus importantes :

- 1) Les données sont **accessibles** : les données sont disponibles à tous les utilisateurs pour l'usage le plus large.
- 2) Les données doivent permettre un **traitement automatisé** : elles doivent être structurées et accompagnées d'une documentation claire afin de permettre un traitement informatisé.
- 3) Les données sont tenues **à jour** : elles doivent être mises à jour aussi souvent que possible afin de favoriser le maintien des services associés aux données.

Open Data : Enjeux

Gains apportés par l'open data

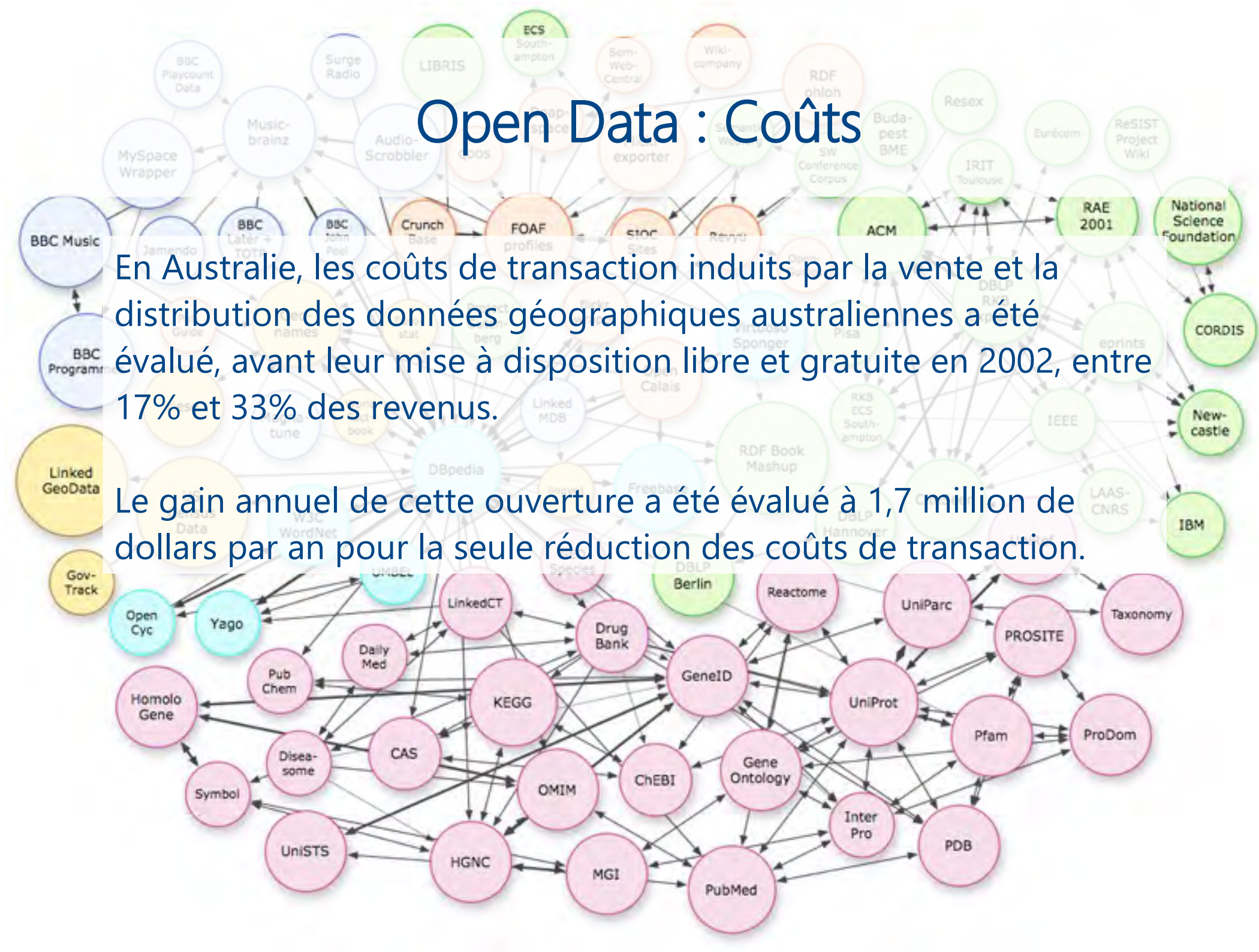
- 1) coût de transaction,
- 2) création de nouveaux services,
- 3) réduction de l'asymétrie d'information,
- 4) innovation ouverte.



Open Data : Coûts

En Australie, les coûts de transaction induits par la vente et la distribution des données géographiques australiennes a été évalué, avant leur mise à disposition libre et gratuite en 2002, entre 17% et 33% des revenus.

Le gain annuel de cette ouverture a été évalué à 1,7 million de dollars par an pour la seule réduction des coûts de transaction.



Open Data : Coûts

Au Danemark, le gouvernement a lancé un programme nommé « Basic Data ». Il s'agit de mettre en place une infrastructure informationnelle autour de trois bases de données de référence (registres-clés), librement et gratuitement disponibles.

Les gains de ce projet sont estimés à 35 millions d'euros annuels pour le secteur public (meilleure efficacité) et 70 millions d'euros pour le secteur privé.

Les gains de ce projet sont estimés à 35 millions d'euros annuels pour le secteur public (meilleure efficacité) et 70 millions d'euros pour le secteur privé.

Open Data : nouveaux services

Aux Pays-Bas, l'ouverture des données météorologiques a permis la création d'un écosystème de ré-utilisateurs professionnels très dynamique : le revenu des acteurs privés a augmenté de 400%, le nombre d'utilisations de ces données de 300%.

Au final, ces activités ont généré un retour de 35 millions d'euros pour les finances publiques néerlandaises, sous la forme d'impôts et de taxes additionnels.

Open Data : nouveaux services

Aux Pays-Bas, l'ouverture des données météorologiques a permis la création d'un écosystème de ré-utilisateurs professionnels très dynamique : le revenu des acteurs privés a augmenté de 400%, le nombre d'utilisations de ces données de 300%.

Au final, ces activités ont généré un retour de 35 millions d'euros pour les finances publiques néerlandaises, sous la forme d'impôts et de taxes additionnels.

Open Data : nouveaux services

Aux Pays-Bas, l'ouverture des données météorologiques a permis la création d'un écosystème de ré-utilisateurs professionnels très dynamique : le revenu des acteurs privés a augmenté de 400%, le nombre d'utilisations de ces données de 300%.

Au final, ces activités ont généré un retour de 35 millions d'euros pour les finances publiques néerlandaises, sous la forme d'impôts et de taxes additionnels.

Open Data : asymétrie

Il y a asymétrie d'information quand un acteur possède une information plus complète ou de meilleure qualité que les autres acteurs participant à une transaction ou une communication. Cela aboutit à des situations de déséquilibre. Les données ouvertes permettent de réduire ces asymétries à plusieurs niveaux.

Au niveau macroéconomique, la transparence est un outil de **lutte contre la corruption** reconnu notamment par la Banque mondiale.

Au niveau microéconomique, la mise en ligne de données sur les **marchés publics** permet à tous les acteurs de disposer du même niveau d'information. Les répondants peuvent connaître le dernier attributaire d'un marché public et les conditions du marché, leur permettant ainsi de mieux dimensionner leur réponse.

Open Data : la loi

Le 7 octobre 2016, la loi pour la République Numérique a étendu cette obligation de diffusion à tout type de données disponibles sous forme numérique pour la majorité des autorités publiques, dont les collectivités territoriales de plus de 3500 habitants et les administrations de plus de 50 agents.



Open Data : la loi

Le 7 octobre 2016, la loi pour la République Numérique a étendu cette obligation de diffusion à tout type de données disponibles sous forme numérique pour la majorité des autorités publiques, dont les collectivités territoriales de plus de 3500 habitants et les administrations de plus de 50 agents.





PRINCIPES À RETENIR DE LA DEFINITION **OPENDATA**



ACCESSIBLES

Les données sont disponibles à tous les utilisateurs pour l'usage le plus large.



AUTOMATISEES

Les données doivent être structurées et accompagnées d'une documentation claire afin de permettre un traitement informatisé.



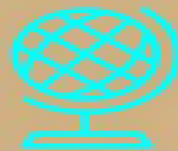
COMPLETES

Toutes les données doivent être rendues disponibles sauf les données pouvant porter atteinte à la vie privée des citoyens ou à la sécurité.



TENUES A JOUR

Les données doivent être mises à jour aussi souvent que possible afin de favoriser le maintien des services associés aux données.



LIBRES DE DROIT

Les données ne doivent pas être l'objet de droits d'auteurs, marques déposées, brevets, etc.



NON DISCRIMINATOIRE

Les données sont disponibles à tout le monde de façon anonyme ne nécessitant pas d'enregistrement.



Open Data : Exemples

L'ambroisie est une plante envahissante dont le pollen est fortement allergisant. Agir contre l'expansion de cette plante est un enjeu majeur de santé publique.

Pour renforcer la connaissance des sites d'implantation de cette plante, l'Agence Régionale de Santé et la région Rhône-Alpes ont financé une application pour smartphone permettant à chacun de **signaler la présence d'ambroisie** de manière simple et rapide. Les signalements sont reçus par le référent de la commune concernée qui coordonne alors les actions de lutte pour éliminer l'ambroisie. Le signalement peut également se faire par mail ou par téléphone. La pratique consistant à faire appel à un grand nombre de personnes, et notamment au grand public, pour participer à la création massive de données s'appelle la **production participative** (ou **crowdsourcing** en anglais).

Open Data : Exemples

Le Ministère de la Cohésion du Territoire, en partenariat avec l'IGN, a mis en place depuis 2016 le **Géoportail de l'urbanisme** (GPU). Il permet aux collectivités territoriales, aux services de l'État et aux gestionnaires de Servitudes d'Utilité Publique de publier sous forme numérique les documents d'urbanisme suivants :

- Les Plans Locaux d'Urbanisme (PLU)

- Les Plans de Sauvegarde et de Mise en Valeur (PSMV)

- Les Plans d'Occupation des Sols (POS)

- Les Cartes Communales (CC)

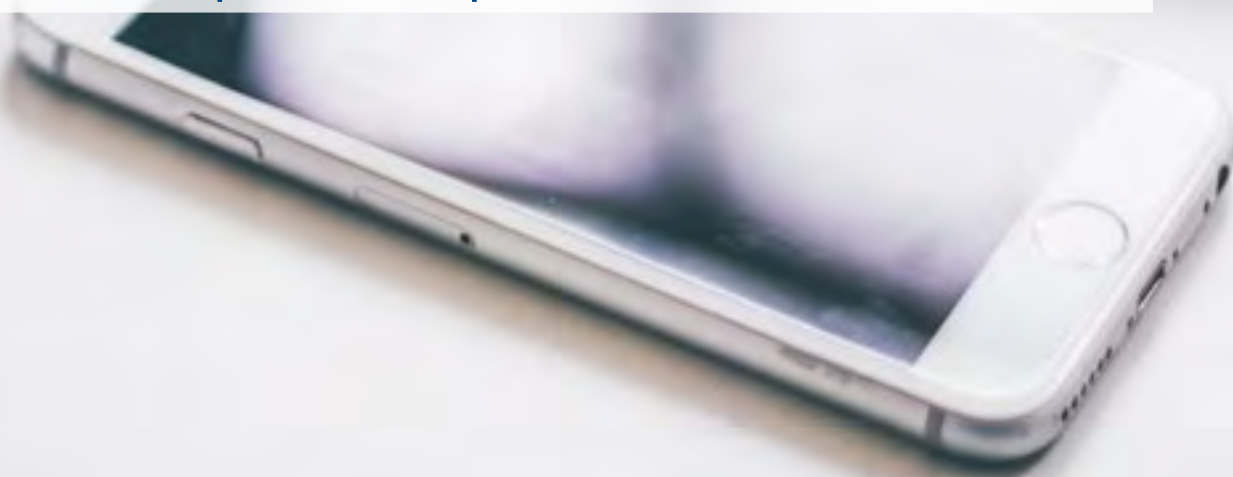
- Les Servitudes d'Utilité Publique (SUP)

- Les Schémas de Cohérence Territoriale (SCOT)

Open Data : Exemples

AtlaSanté est une démarche de gestion mutualisée du patrimoine de données géographiques en santé, pour la France métropolitaine et d'Outre-Mer. L'objectif est d'appuyer les prises de décisions pour la politique de santé (ARS, directions de l'administration centrale du ministère de la santé, partenaires, etc.).

AtlaSanté est une infrastructure de données géographiques couvrant le champ de la santé, développée sous Prodiges et complétée de l'outil géostatistique Geoclip.



Open Data : Exemples

Datacity Paris a pour objectif d'utiliser les données privées et publiques afin de produire et d'inventer des solutions innovantes aux problématiques des villes. Il a pour ambition :

- d'expérimenter sur le terrain des solutions concrètes autour des services urbains
- de participer activement à la définition de problèmes à résoudre / à l'émergence de nouvelles idées en relayant les besoins des citoyens
- de permettre un réel impact de solutions technologiques appliquées à la ville
- de sécuriser l'investissement dans de futurs dispositifs coûteux pour la collectivité
- de valoriser des jeux de données publiques
- de créer de nouvelles applications et services

Il intervient sur les thèmes suivants :

- mobilité : affluence dans les transports, déplacement des salariés, fiabilisation des données routières...
- aménagement urbain : escalator, parcours sportif, accessibilité...
- environnement : réemploi des encombrants, valorisation des déchets, optimisation des comportements, application visant à informer en temps réel les concierges, les entreprises et les citoyens, sur l'heure de passage des éboueurs, afin de limiter le temps passé par les poubelles sur les trottoirs de Paris.
- énergie : lisser la consommation d'énergie...



Open Data : Exemples

<https://opendata.reseaux-energies.fr/pages/accueil/>

<https://www.insee.fr/fr/accueil>

<https://www.data.gouv.fr/fr/>

<https://www.datara.gouv.fr/accueil>

