
ARCHITECTURE DE DONNÉES ET DE L'INFORMATION

OBJECTIFS D'APPRENTISSAGE DE LA PRÉSENTE SÉANCE

Objectif principal : mieux vous faire comprendre ce que vous devez prendre en compte sur le plan des données afin de mener, de gérer, de créer et d'initier des projets axés sur les données fructueux.

Objectifs complémentaires :

- être exposé à de nombreux concepts et termes en jeu dans le monde des données d'aujourd'hui – p. ex. modèle de données, relations multivoques, bases de données, systèmes noSQL, requêtes, gestion des données;
- faire l'expérience de la création d'un modèle conceptuel et de sa transformation en un type particulier de modèle de données, selon un cas d'utilisation.

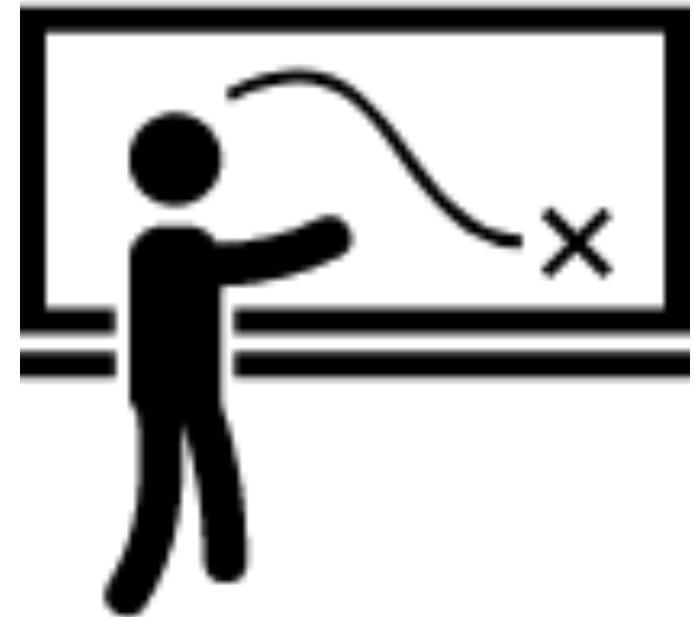
DU VIEUX OU DU NOUVEAU?

Puisque vous avez tous été choisis pour le flux de données, vous êtes peut-être déjà des « personnes responsables des données ».

Vous connaissez sans doute déjà bon nombre de ces concepts et compétences.

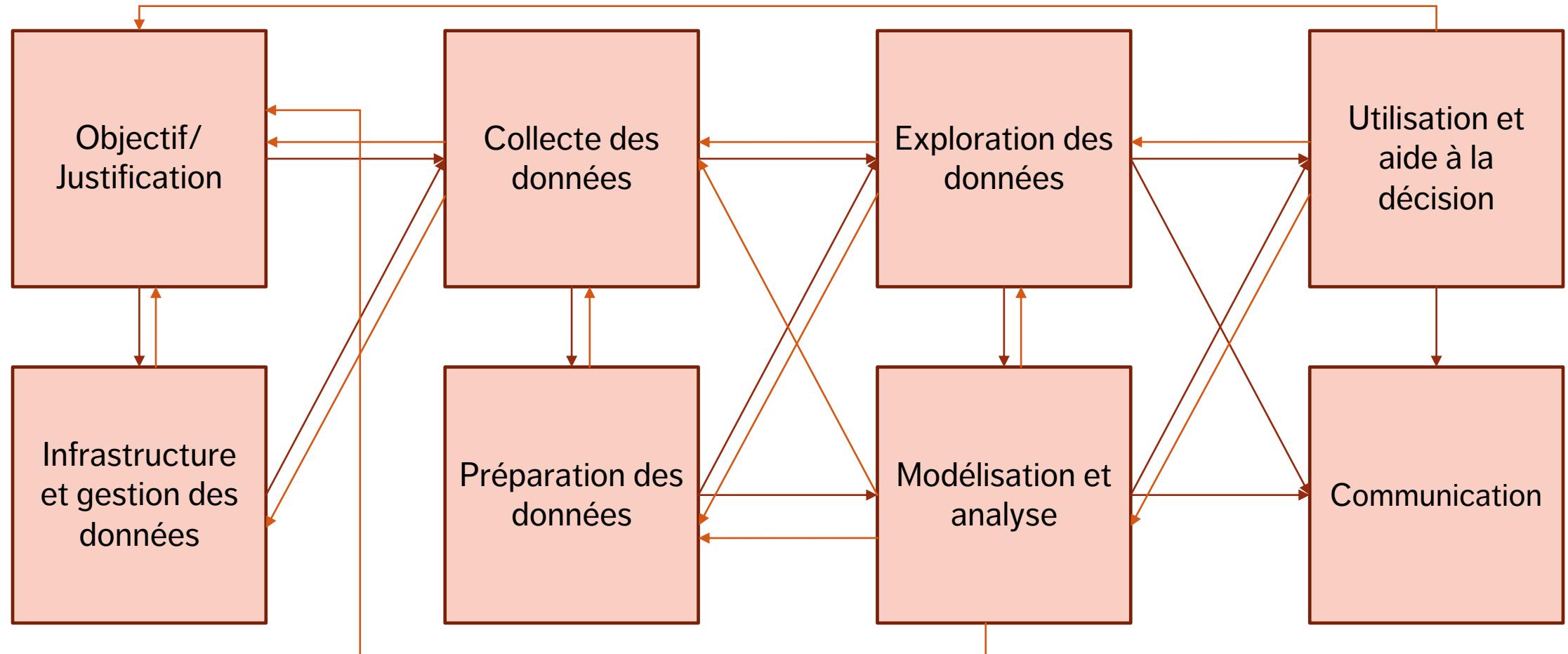
Cependant, comme ces concepts d'architecture de données et de connaissances sont essentiels à l'élaboration d'un projet axé sur les données, nous ne pouvons pas simplement supposer que vous les connaissez!

Espérons que certains de ces concepts seront encore nouveaux pour tout le monde! Ontologies, quelqu'un?



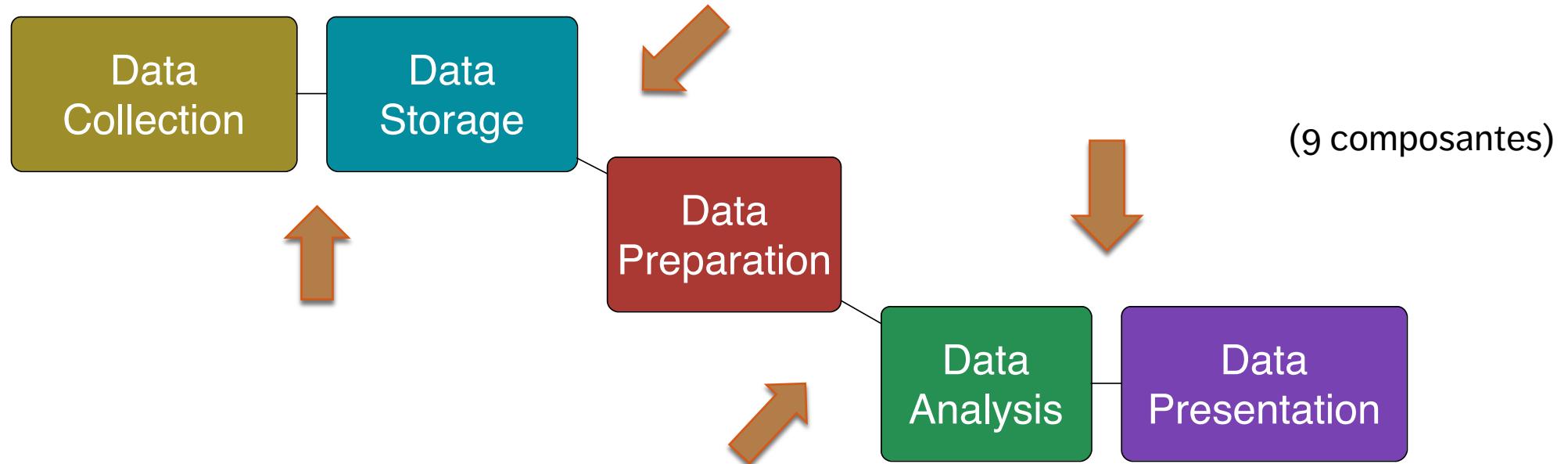
QUELQUES ÉLÉMENTS DE CONTEXTE

LE PROCESSUS D'ANALYSE (DÉSORDONNÉ)



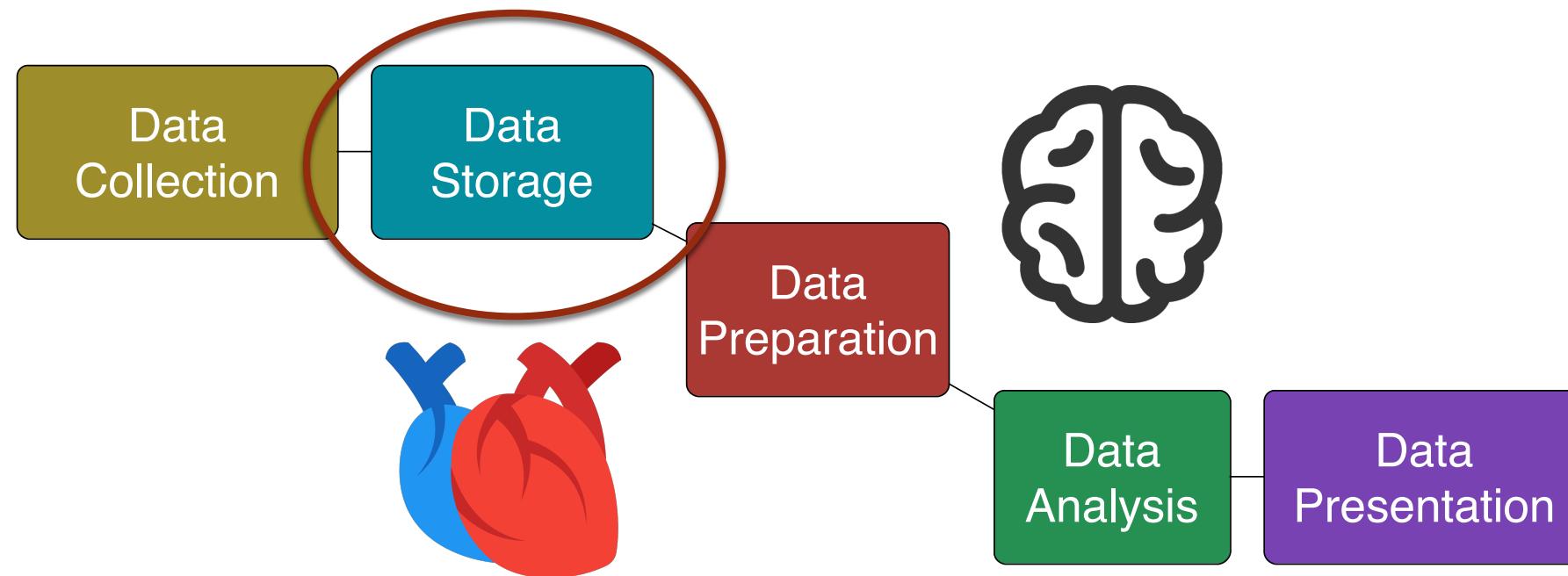
UN PIPELINE DE DONNÉES (AUTOMATISÉ) MIS EN ŒUVRE

Dans le **contexte de la prestation de services**, il se peut que vous souhaitiez avoir l'une de ces composantes...



(Comme toujours, attention au modèle « drift »)

UN PIPELINE DE DONNÉES (AUTOMATISÉ) MIS EN ŒUVRE



SOURCES DE DONNÉES

DISCUSSION EN GROUPE : QUELQUES QUESTIONS DE DONNÉES FONDAMENTALES

- **Pourquoi** recueillons-nous des données? Que pouvons-nous **faire** avec des données?
- D'où proviennent les données?
- En supposant que nous recueillons des données pour pouvoir obtenir une *collection* de données – à quoi ressemble une « collection » de données? Comment en décririez-vous une?
- Faut-il faire la distinction entre les données, les renseignements, les connaissances?

DE QUOI S'AGIT-IL?



MOTIVATIONS POUR LA COLLECTION

Trois fonctions, historiquement :

- Tenue de dossiers (gestion des personnes et de la société!)
- Science - Nouvelles connaissances générales
- Renseignements - commerciaux, militaires, de police, sociaux? Au Canada? Personnel!



MOTIVATIONS POUR LA COLLECTION

Chacune de ces trois fonctions a traditionnellement utilisé des **sources** d'information différentes.

Elles ont recueilli **différents types de données**.

Elles ont aussi eu des **cultures de données** et des terminologies différentes!

En science des données (interdisciplinaire!), nous pouvons les rencontrer toutes dans le cadre d'un même projet.



DIFFÉRENTES CULTURES DE DONNÉES, DIFFÉRENTS TERMES

Renseignements opérationnels :

- entrepôt de données + mini-entrepôt de données
- « dimension » de données (= ensemble de données)
- données hiérarchiques (tranches)
- élément de données
- tableau des dimensions + tableau des faits

Données

Sciences/statistiques :

- données expérimentales
- essais
- participants
- variables
- corrélation

Connaissances

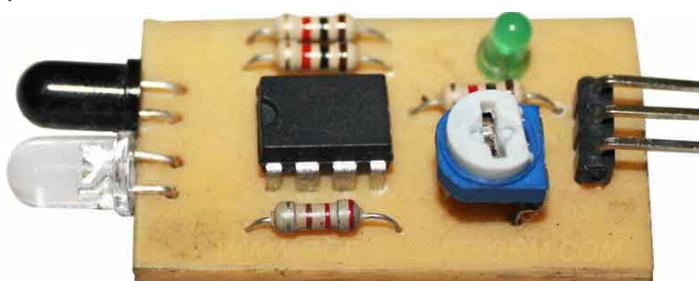
Gestion des documents :

- architecture de l'information
- plan de classement
- ressource d'information
- champ
- formulaire
- objet

Renseignements

QUELQUES SOURCES DE DONNÉES

- comptes rendus des activités
- observations (scientifiques)
- capteurs + surveillance
- ordinateurs!



ENTREZ DANS LE MONDE DE LA SCIENCE INFORMATIQUE! ET DE L'INFORMATIQUE!

L'informatique (et les sciences de l'information) a son propre point de vue théorique **fondamental** sur les données et l'information.

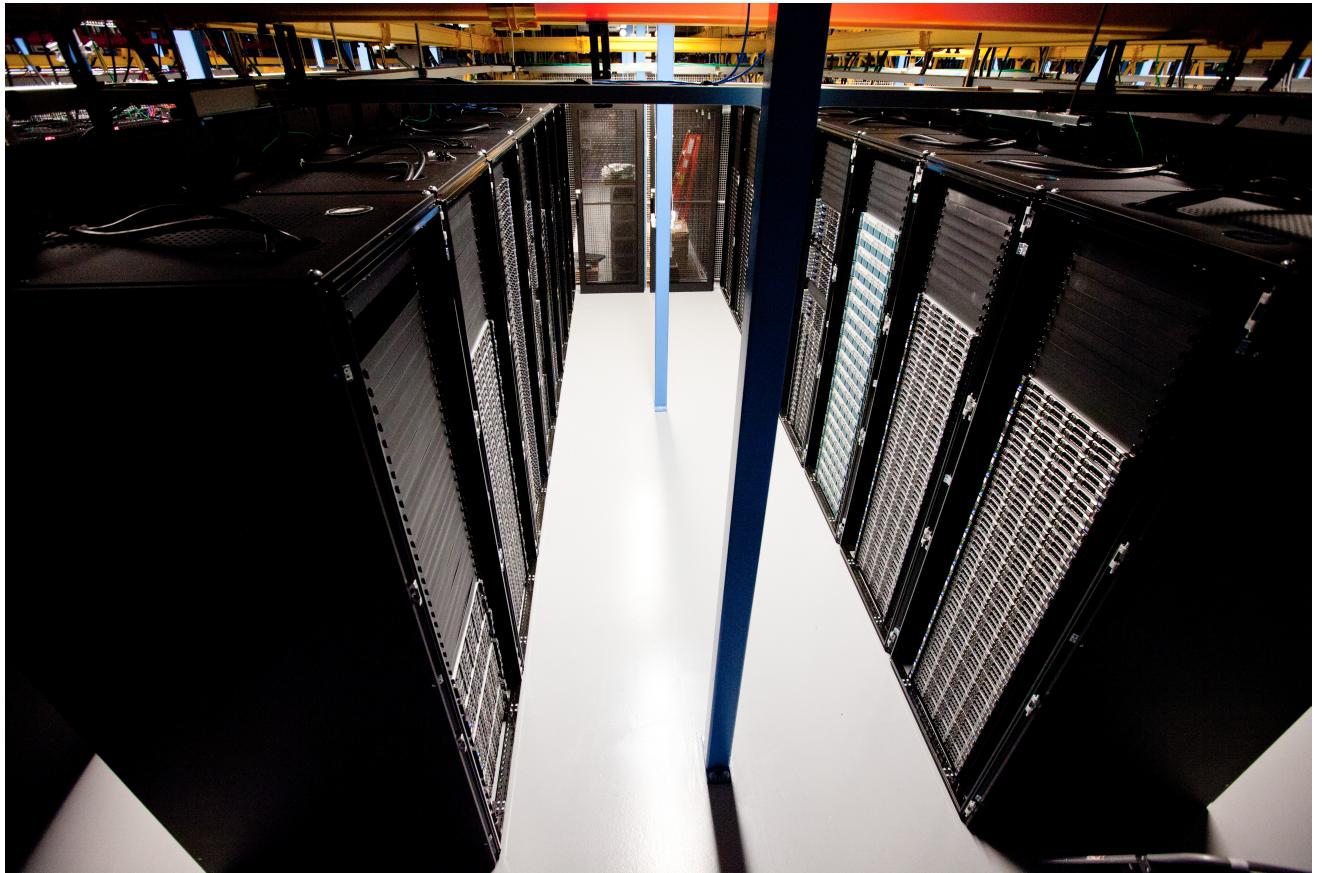
Essentiellement, les ordinateurs fonctionnent avec des données - des 1 et des 0 qui représentent des chiffres, des lettres, etc. - **les données deviennent numériques.**

De façon pragmatique, les données sont maintenant stockées sur des ordinateurs, et elles sont accessibles par l'intermédiaire de notre réseau informatique mondial.



LES DONNÉES SONT RÉELLES!

- Les données sont une représentation, mais elles sont tout de même physiques!
- Les données ont des propriétés physiques - elles occupent un espace physique! Il faut de l'énergie pour travailler avec elles.



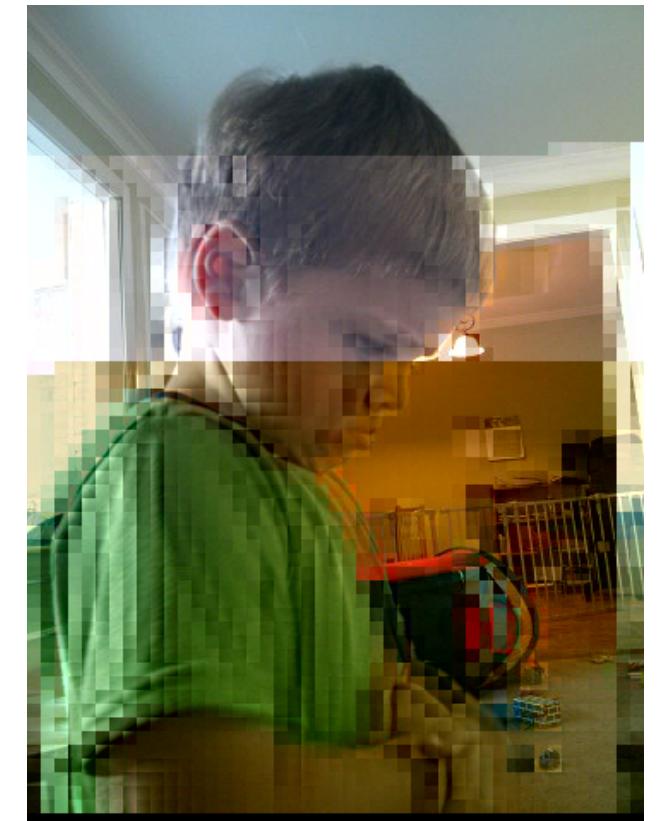
DÉGRADATION DES DONNÉES

Les données vieillissent avec le temps - elles ont une durée de vie.

Nous utilisons les expressions « données pourries », « données en dégradation » :

- **littéralement** - dans le sens où le support de stockage de données pourrait se dégrader;
- **métaphoriquement** - lorsque les données ne **représentent** plus exactement les relations et les objets pertinents ou même lorsque ces objets n'existent plus de la même manière.

Vos données doivent être « fraîches », « actuelles », et non « périmées » (en fonction du contexte et du modèle!)



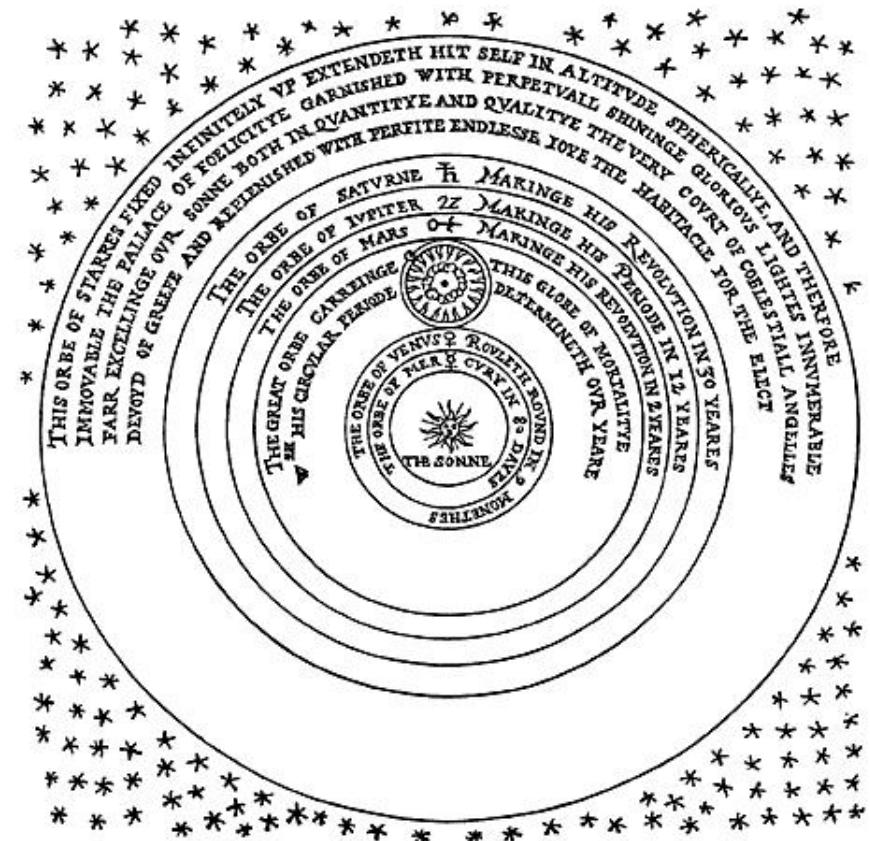
AVANT LES DONNÉES

CONCEPTS DE BASE

Comment recouper les différentes disciplines qui utilisent des données?

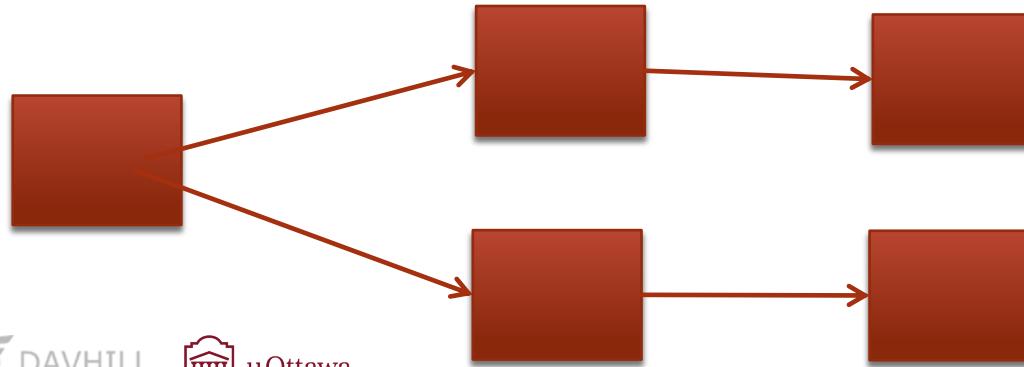
Concepts ou éléments de base (systèmes) :

- objet - attributs (concrets ou abstraits)
- objets multiples - **liens** entre ces objets/attributs
- façon dont ces éléments évoluent au fil du temps



LIENS AU SEIN DES SYSTÈMES

- quelques liens fondamentaux :
 - en partie
 - est une
 - est un type de
 - cardinalité (un à un, un à plusieurs, plusieurs à plusieurs)
- quelques autres liens liés à l'objet :
 - propriété,
 - liens sociaux
 - devient,
 - mène à



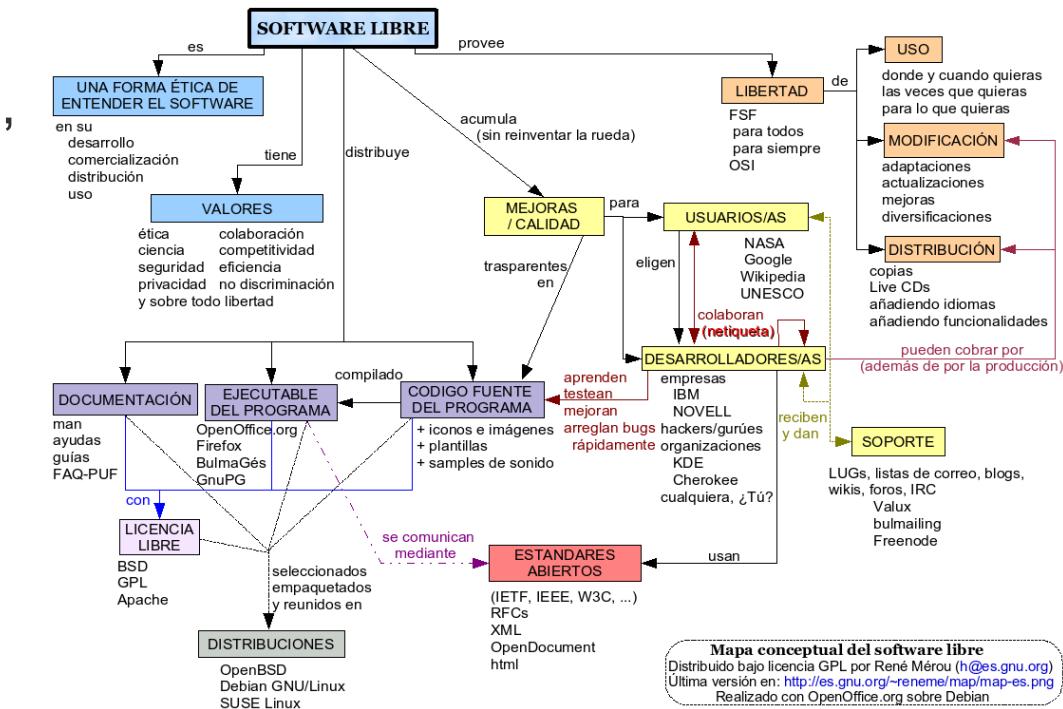
MODÈLE CONCEPTUEL D'UN SYSTÈME

Un modèle conceptuel est, grosso modo :

- un modèle qui n'est pas mis en œuvre (p. ex. sous forme de modèle à l'échelle ou de code informatique), mais qui n'existe que sur le plan conceptuel;
- un diagramme ou une description verbale d'un système (p. ex. boîtes et flèches, cartes mentales, listes, définitions).

L'accent n'est pas nécessairement mis sur la saisie de comportements spécifiques, mais sur les **états possibles**.

L'accent est mis sur les types d'objets, pas sur des instances spécifiques – le but est l'abstraction!



Mapa conceptual del software libre
Distribuido bajo licencia GPL por René Mérou (h@es.gnu.org)
Última versión en: <http://es.gnu.org/~reneme/map/map-es.png>
Realizado con OpenOffice.org sobre Debian

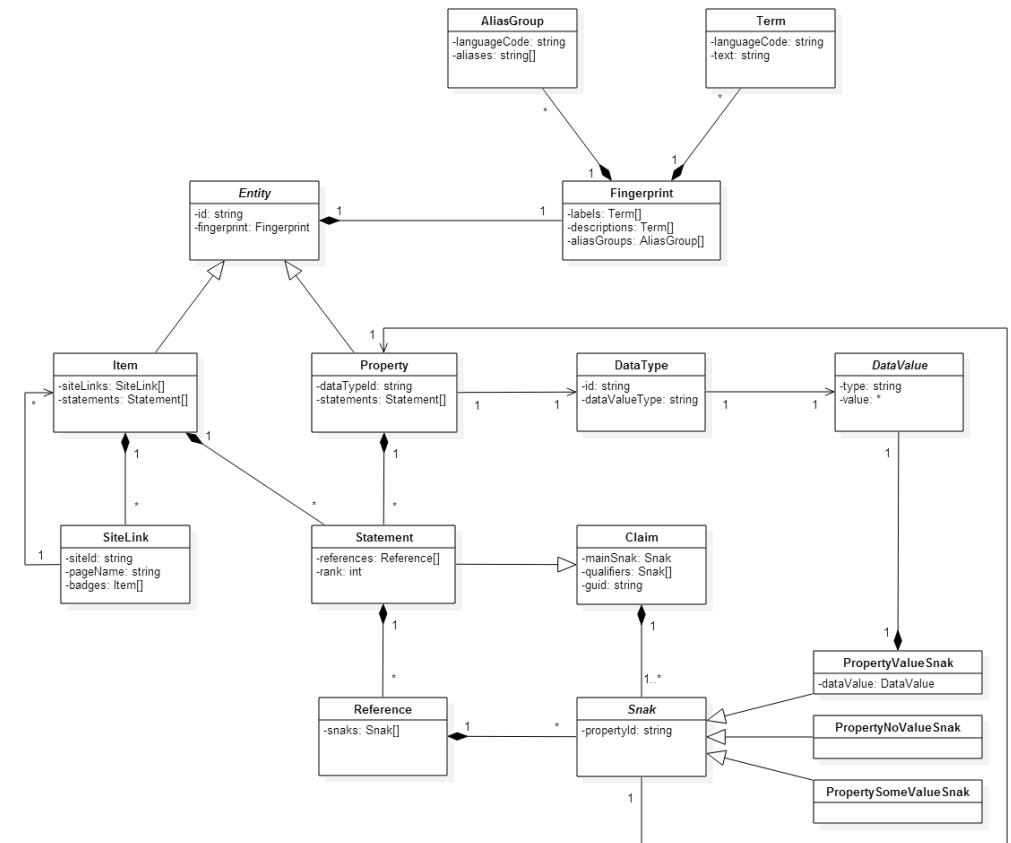
LANGAGES FORMELS DE MODÉLISATION CONCEPTUELLE

La modélisation conceptuelle n'est pas une science exacte – il s'agit plutôt de rendre vos modèles conceptuels internes **explicites** et **tangibles**.

Elle vous donne l'occasion d'examiner et d'explorer vos idées et vos hypothèses.

Cela dit, divers efforts ont été faits pour formaliser la modélisation conceptuelle :

- UML - Langage de modélisation unifié
- Modèles de rapport entre entités - mais ils sont généralement connectés à des bases de données relationnelles



MISE EN PRATIQUE

Choisissez un système qui vous convient et créez un modèle conceptuel qui englobe :

- des objets - en particulier les **objets clés** (peuvent être abstraits ou concrets);
- les propriétés de ces objets et les valeurs de ces propriétés;
- liens entre les objets (partie entière, est un, propre à l'objet, un à plusieurs);
- les liens entre les propriétés dans toutes les *instances* d'un type d'objet.

MODÉLISATION DE DONNÉES ET DE CONNAISSANCES

MODÈLES MÉTHODOLOGIQUES

Modèle conceptuel



Modèle mathématique

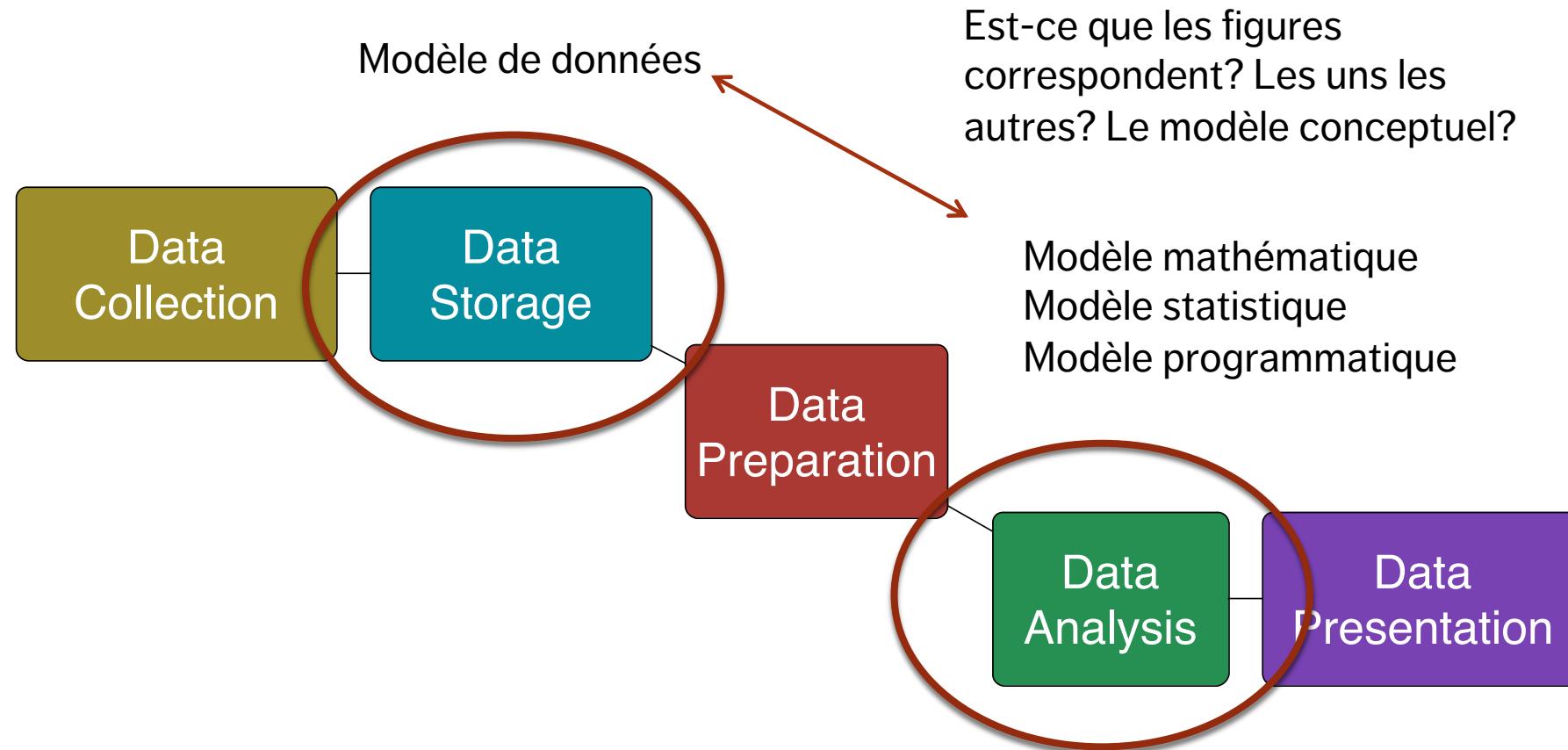
Modèle statistique

Modèle programmatique (informatique)

Ontologie (modèle de connaissances)
(Exercice de simulation)

Modèle de données

UN PIPELINE DE DONNÉES (AUTOMATISÉ) MIS EN ŒUVRE



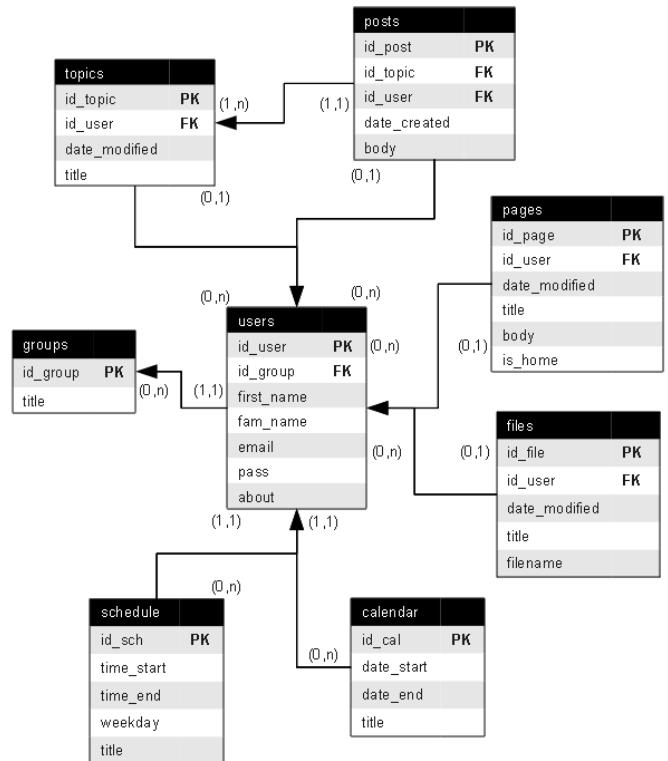
ACCROCHEZ-VOUS À CES CONCEPTS FONDAMENTAUX

Comment structurer vos **données** et vos **connaissances** pour qu'elles puissent être :

- stockées de façon utile
- ajoutées à
- extraites de façon utile et efficace de la mémoire en question (extraction – transformation – chargement)
- exploitées par des **humains** et des **ordinateurs** (programmes, robots, IA).

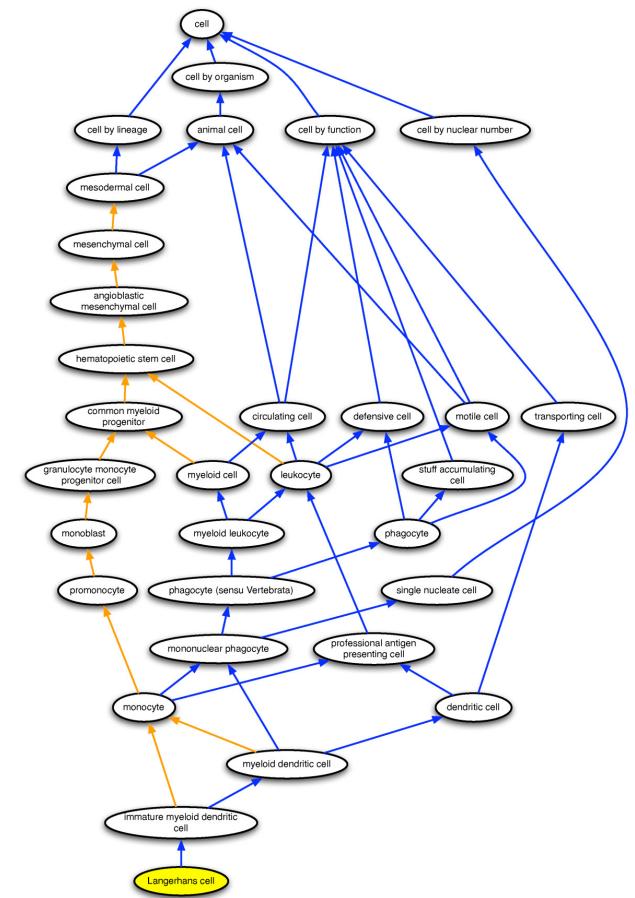
QU'EST-CE QUE LA MODÉLISATION DE DONNÉES?

- Un modèle de données est une description abstraite (logique) d'un système, construite en des termes qui peuvent ensuite être mis en œuvre comme la structure d'un type de logiciel de gestion de données.
- Vous pourriez soutenir que la modélisation de données est à mi-chemin entre un modèle conceptuel et une mise en œuvre de base de données.
- Les données elles-mêmes concernent les **instances** – le modèle concerne les **types d'objets**.
- Une autre option vaut la peine d'être considérée à ce sujet – **les ontologies**.



ONTOLOGIE – MODÈLE DE CONNAISSANCES

- Une collection structurée et lisible par machine de **faits** sur un domaine.
- Motivée par le désir de créer des données de plus en plus lisibles par machine, mais toujours complexes sur le plan conceptuel.
- Vous pourriez décrire l'ontologie, en plaisantant un peu, comme « un modèle de données gonflé aux stéroïdes ».
- Une tentative de se rapprocher du niveau de détail d'un modèle conceptuel complet.



MÉTADONNÉES POUR FOURNIR UN CONTEXTE

- Nous perdons quelque chose lorsque nous passons de notre modèle conceptuel à un modèle de type spécifique – p. ex. le modèle de données ou de connaissances.
- Une façon de conserver le contexte est de fournir des métadonnées (riches, si possible) – des données **sur** nos données!
- Les métadonnées sont essentielles lorsqu'il s'agit de mettre en œuvre des stratégies pour travailler d'un ensemble de données à l'autre.
- L'ontologie peut aussi jouer un rôle ici!

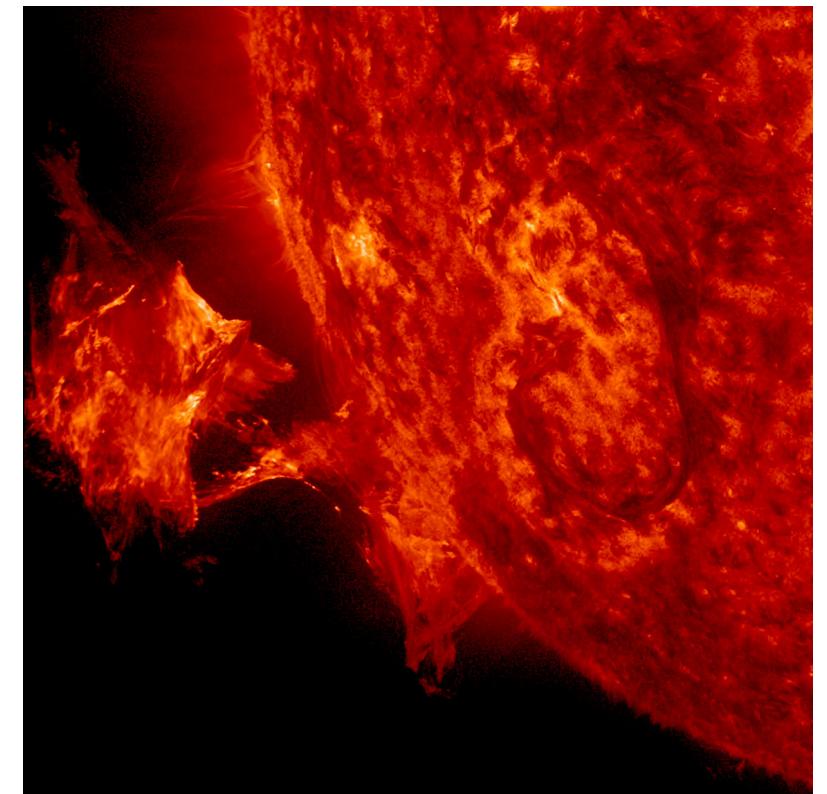
DONNÉES STRUCTURÉES PAR RAPPORT AUX DONNÉES NON STRUCTURÉES

La disponibilité croissante de données non structurées et de grands objets binaires « **blob** » est l'une des principales motivations de certains des nouveaux développements dans les types de bases de données et autres stratégies de stockage de données.

Données structurées : étiquetées, organisées, discrètes, selon une structure limitée et prédéfinie

Données non structurées : non organisées, pas de modèle de données structuré prédéfini précis – p. ex. texte dans un document

Données « Blob » : grand objet binaire – images, audio, multimédia



QU'EST-CE QUE LA MODÉLISATION DE DONNÉES?

Nous allons examiner quatre options différentes qui sont actuellement populaires en matière de modélisation ou de stratégies de structuration **de données et de connaissances** fondamentales :

- paires de valeurs clés (p. ex. JSON)
- triples (p. ex. modèle RDF)
- bases de données graphiques
- bases de données relationnelles

noSQL

MÉMOIRES DE VALEURS CLÉS ET MÉMOIRES TRIPLES

Voici des moyens relativement peu structurés de stocker les données :

- **Valeur clé** : toutes les données sont simplement stockées sous la forme d'une liste géante de clés et de valeurs, dans laquelle la clé est un nom ou une étiquette (possiblement d'un objet) et la valeur est une valeur qui y est associée.
- **Triple** : les données sont stockées en tant que sujet – prédicat – objet

EXEMPLES

type de pomme - couleur de la pomme

Granny Smith - verte

Red Delicious - rouge

mot - définition

URL - page Web

Nom de rapport - rapport (dossier de documentation)

personne - pointure de chaussures

Jen Schellinck - femmes, pointure 7

Colin Henein - hommes, pointure 10

Les triples ajoutent un verbe au mélange :

Personne - est - âge

Objet - est de couleur - couleur

RÉGLER LES DÉTAILS

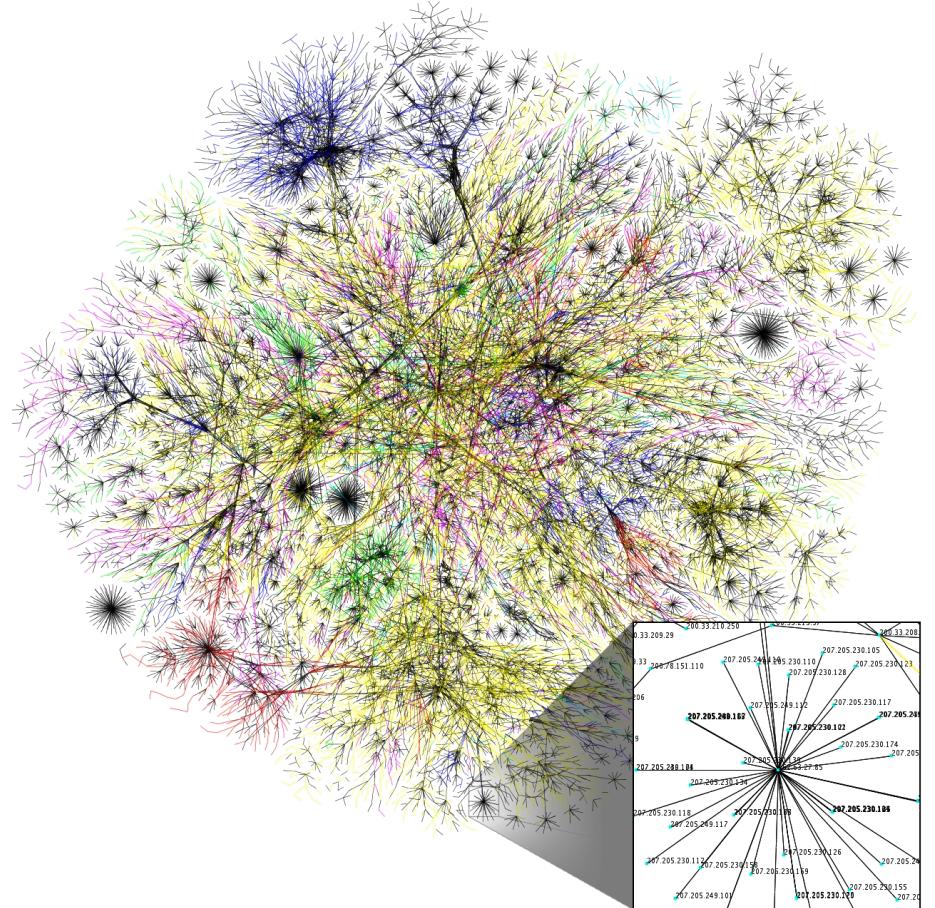
- Les deux stratégies se traduisent par une grande flexibilité en ce qui concerne la « conception » du stockage des données.
- Pour ce qui est de la **mise en œuvre** de ces modèles de données, le problème réside dans les détails, et leur extrême souplesse peut ajouter à ce défi.
- Voir par exemple :
<https://softwareengineering.stackexchange.com/questions/321534/what-json-structure-to-use-for-key-value-pairs>

BASES DE DONNÉES GRAPHIQUES

Accent mis sur les **liens** entre les différents types d'objets, plutôt que sur les liens entre un objet et les propriétés de cet objet.

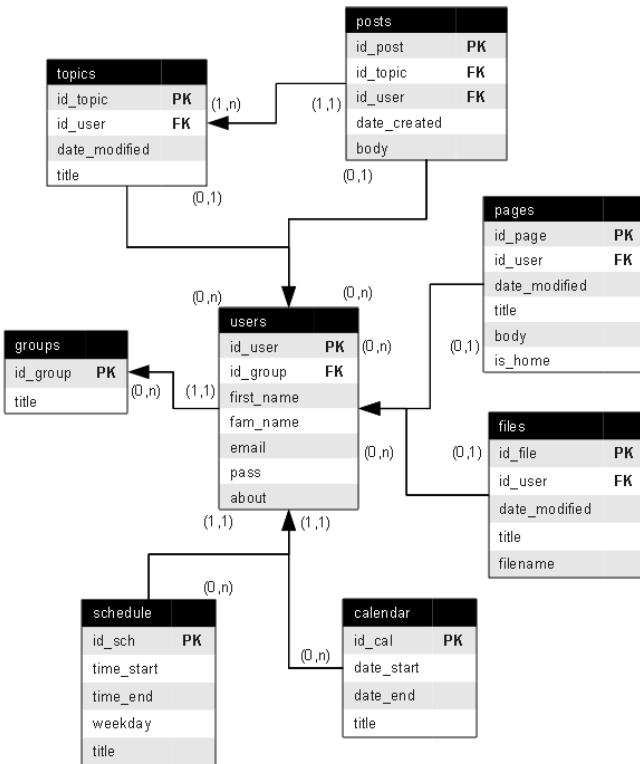
Le modèle de données :

- objets représentés par des nœuds
- liens entre ces objets représentés par des bords
- les objets peuvent avoir un lien avec d'autres objets du même type - la personne est le frère ou la sœur de l'autre personne



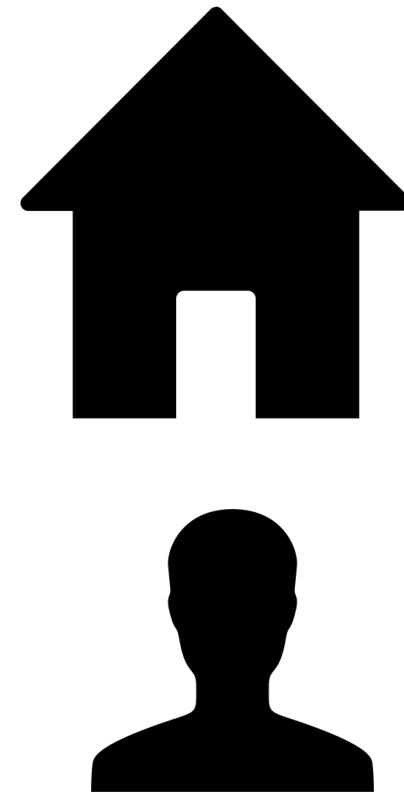
BASES DE DONNÉES RELATIONNELLES

- Données stockées dans une série de tableaux.
 - En gros, chaque tableau représente un objet et des propriétés liées à cet objet.
 - Des colonnes spéciales dans les tableaux relient les instances d'objets dans le tableau.



BASES DE DONNÉES RELATIONNELLES

- Exemple : Je vis dans une maison qui a une adresse particulière.
- Parfois, la propriété de la maison sera stockée dans le tableau qui stocke les informations me concernant.
- Dans d'autres cas, il sera plus logique de stocker les informations sur la maison dans son propre tableau.
- La forme des bases de données relationnelles est déterminée par le fait qu'il s'agit ou non de liens un à un, un à plusieurs ou plusieurs à plusieurs (cardinalité).



MISE EN PRATIQUE

Imaginez un exemple de lien un à un, un à plusieurs et plusieurs à plusieurs. Dans quelle mesure la cardinalité de ce lien est-elle claire?

Quelle est la cardinalité du lien père-enfant?

AVANTAGES, INCONVÉNIENTS ET CAS D'UTILISATION DE CHACUN

Base de données relationnelle : largement soutenue, bien comprise, fonctionne bien pour de nombreux types de systèmes et de cas d'utilisation. Base toutefois difficile à changer une fois mise en œuvre; ne gère pas bien les liens (malgré son nom).

Mémoires de valeurs clés : peuvent prendre n'importe quel type de données; nul besoin de beaucoup de renseignements sur l'avancement de sa structure. Si vous avez beaucoup de valeurs manquantes, ces mémoires ne prendront pas de place. Peuvent toutefois être désordonnées et mystérieuses; difficile d'y trouver des données.

Bases de données graphiques : rapides et intuitives si vous utilisez des données fortement axées sur les liens; pourraient être la seule option si vos données sont ainsi parce que les bases de données traditionnelles peuvent ralentir énormément. Sont toutefois probablement trop spécialisées si vos données ne sont pas ainsi, pas encore supportées à grande échelle.

REMARQUE SUR LES FICHIERS NON HIÉRARCHIQUES ET LES FEUILLES DE CALCUL

Qu'en est-il de la conservation de vos données dans un seul tableau géant (feuille de calcul)? Ou plusieurs feuilles de calcul? Ça ne peut pas être si terrible que ça!

Wayne Eckerson a inventé le terme « spreadmart » (par opposition à un mini-entrepôt de données) pour décrire une situation avec plusieurs feuilles de calcul (ponctuelles) comme stratégie de données.

Date	Con	Lab	LDs	SNP	UKIP	Greens	Con av	Lab av	LD av	SNP av	UKIP av	Green av
15 September 2017	41	41	5	4	5	3	40.7	41.4	6.8	3.3	4	2.7
15 September 2017	39	38	8	3	6	4	40.7	41.7	7	3.2	3.8	2.6
13 September 2017	41	42										
10 September 2017	42	42										
1 September 2017	38	43										
Date	Con	Lab	LDs	SNP	UKIP	Greens	Con av	Lab av	LD av	SNP av	UKIP av	Green av
15 September 2017	41	41	5	4	5	3	40.7	41.4	6.8	3.3	4	2.7
15 September 2017	39	38	8	3	6	4	40.7	41.7	7	3.2	3.8	2.6
13 September 2017	41	42	7	4	3	2	40.9	42.2	6.8	3.3	3.5	2.4
10 September 2017	42	42	7	3	4	3	40.9	42.2	7	3.2	3.5	2.4
1 September 2017	38	43	7	3	1	4	40.9	42.3	7	3.2	3.4	2.3
31 August 2017	41	42	6	4	4	2	41	42.1	7.1	3.2	3.9	2
22 August 2017	42	42	7	2	3	3	41	42.2	7	3.1	4	2
22 August 2017	41	42	8	4	4	1	40.8	42.5	7	3.3	3.9	1.8
18 August 2017	40	43	6	4	4	2	40.5	42.9	6.8	3.3	3.9	1.8
11 August 2017	42	39	7	2	6	3	40.6	42.9	6.9	3.2	3.8	1.8
1 August 2017	41	44	7	3	3	2	40.5	43	6.9	3.2	3.4	1.7
19 July 2017	41	43	6	4	3	2	40.3	43.1	6.7	3.2	3.6	1.7
18 July 2017	41	42	9	3	3	2	40.3	43.4	6.7	3.1	3.5	1.6
11 August 2017	42	43	7	3	3	2	40.3	43.6	6.4	3.1	3.4	1.5
1 August 2017	41											
19 July 2017	41											
18 July 2017	41											
16 July 2017	42											
15 July 2017	39											
14 July 2017	41											
11 July 2017	40											
6 July 2017	38											
3 July 2017	41											
30 June 2017	41	40	7	2	2	2	40.8	43.5	6.4	2.9	2.7	1.8
6 July 2017	39	45	5	3	5	2	40.7	44.2	6.3	3.0	2.8	1.7
3 July 2017	41											
30 June 2017	41	40	7	2	2	2	40.8	43.5	6.4	2.9	2.7	1.8
29 June 2017	39	45	5	3	5	2	40.7	44.2	6.3	3.0	2.8	1.7

REMARQUE SUR LES FICHIERS NON HIÉRARCHIQUES ET LES FEUILLES DE CALCUL

- **Avantages** - très efficace si vous recueillez des données une seule fois, sur un type particulier d'objet (p. ex. études scientifiques!); comme certains types d'analyse exigent que vous ayez toutes les données en un seul endroit, vous allez devoir générer un fichier non hiérarchique de toute façon. Facile à lire dans le logiciel d'analyse (p. ex. R) et il est facile d'effectuer des opérations quant à l'ensemble de données
- **Inconvénients** - il est très difficile de gérer l'intégrité des données à long terme si vous recueillez et travaillez continuellement avec les données. Ne fonctionne pas bien si vous traitez des données sur un système comprenant de nombreux types d'objets différents et leurs liens. Peut être très difficile d'exécuter des opérations d'interrogation de données

QUELQUES OUTILS ET MOTS À LA MODE

- MongoDB, ArangoDB
- Entrepôt de documents
- JSON, YAML
- API, GraphQL
- Données interreliées
- Web sémantique
- Langage d'ontologie Web (OWL)
- Protected

MISE EN PRATIQUE

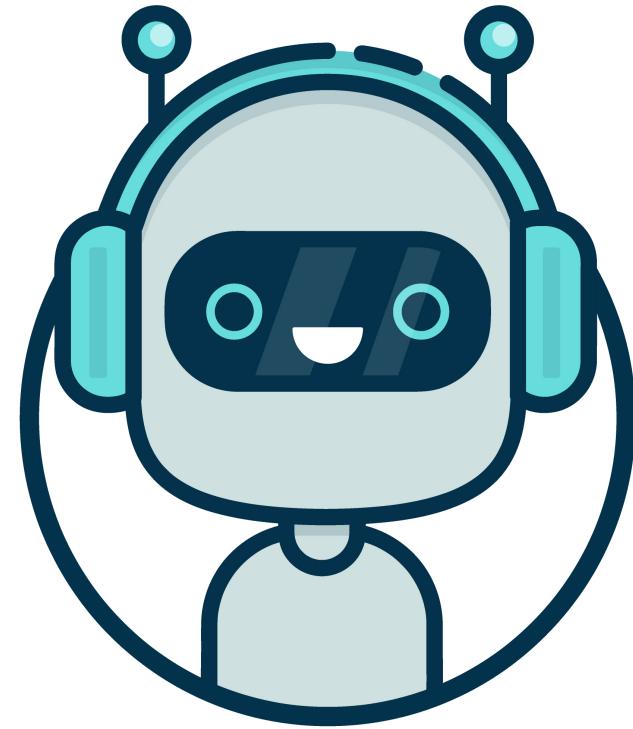
Jetez un coup d'œil au modèle conceptuel que vous avez créé pour votre système. Choisissez une façon d'utiliser les données ou les connaissances sur ce système. Selon vous, quel type de données ou modèle de connaissances est le plus approprié pour ce système et ce cas d'utilisation? Créez un modèle de données ou de connaissances approprié à partir de votre modèle conceptuel.

MISE EN ŒUVRE DE VOTRE MODÈLE

MISE EN ŒUVRE DE VOTRE MODÈLE

- Pour mettre en œuvre votre modèle de données/connaissances, vous devez avoir accès à un **logiciel de stockage et de gestion de données**.
- L'accès pourrait représenter un défi pour vous en tant que personne, parce que traditionnellement ces logiciels fonctionnent sur des serveurs.
- Les serveurs sont bons parce qu'ils permettent à plusieurs personnes d'accéder à une seule base de données en même temps, à partir de différents programmes clients. Il est toutefois difficile de « jouer » avec une base de données.
- SQLite à la rescousse!

LIEN ENTRE LE PROGRAMME ET LES DONNÉES



RÔLE DU LOGICIEL DE GESTION DES DONNÉES

Le logiciel de gestion des données offre aux humains un moyen facile d'interagir avec leurs données.

Il s'agit essentiellement d'une interface – de données – humaines.

Grâce à cette interface, vous pouvez :

- ajouter des données à votre collection de données;
- extraire des sous-ensembles de données de votre collection en fonction de certains critères;
- supprimer des données de votre collection ou en modifier.



UNE DONNÉE QUOI???

Auparavant :

- Base de données
- Entrepôt de données
- Mini-entrepôts de données
- Système de gestion de bases de données
- (SQL)

Maintenant :

- Ensemble de données
- Bassin de données
- Marais de données?
- Cimetière de données?
- (noSQL)

De plus en plus de distinction entre la mémoire de données et le logiciel de gestion de données

DANS CE COURS : SQLITE À LA RESCOUSSE

SQLite est une base de données relationnelle **autonome** qui est installée sur vos appareils Surfaces.

Jetons maintenant un coup d'œil rapide à l'installation de SQLite!

DU MODÈLE DE DONNÉES À LA MISE EN ŒUVRE

Que faisons-nous une fois que nous avons terminé un modèle de données (logique)?

- instancier le modèle dans le logiciel de votre choix (p. ex : créer des tableaux dans MySQL);
- charger les données.
- **demander** les données :
 - Les bases de données relationnelles utilisent le langage d'interrogation structuré (SQL)
 - D'autres types de bases de données utilisent des langages de requête totalement différents (AQL, moteurs sémantiques, etc.) ou s'appuient sur des programmes informatiques sur mesure (p. ex. écrits en R, Python).

GESTION DE LA COLLECTION DE DONNÉES

Une fois que vous avez créé une collection de données, vous devez également la gérer! Qu'est-ce que cela signifie? Essentiellement, cela signifie qu'il faut maintenir les données dans la base de données :

- exactes,
- précises,
- uniformes,
- complètes.

Ne laissez pas votre lac de données se transformer en marais de données...

SUPPLÉMENTS

BASES DE DONNÉES RELATIONNELLES – EN PROFONDEUR

Apprendre à bien structurer une **base de données relationnelle** est une compétence utile - il existe de nombreuses pratiques exemplaires à cet égard.

Cela dépasse le cadre de ce cours!

Cependant, comme activité supplémentaire, vous pouvez essayer de mettre en œuvre votre modèle de données dans votre base de données SQLite.

INDICATION DES SOURCES ET RÉFÉRENCES

SOURCES DES IMAGES (I)

- Personne avec graphique : https://www.flaticon.com/free-icon/person-explaining-strategy-on-a-board-with-a-sketch_38792
- Signal numérique : par Mcanet - Own work, CC BY 3.0, <https://commons.wikimedia.org/w/index.php?curid=6024833>
- Capteur infrarouge :
https://commons.wikimedia.org/wiki/File:Infrared_Transceiver_Circuit.jpg#/media/File:Infrared_Transceiver_Circuit.jpg
- Salle de serveurs : https://upload.wikimedia.org/wikipedia/commons/9/98/Wikimedia_Servers-0051_10.jpg
- Image retournée sur deux bits : par Jim Salter - Own work, CC BY-SA 4.0, <https://commons.wikimedia.org/w/index.php?curid=64046994>
- Système solaire selon Thomas Digges : par Jeangagnon - Wikipedia anglais, domaine public, <https://commons.wikimedia.org/w/index.php?curid=3155015>
- Diagramme UML du modèle de données Wikibase :
https://upload.wikimedia.org/wikipedia/commons/4/4a/Wikibase_Conceptual_Data_Model.png
- Carte conceptuelle du « logiciel libre » (espagnol) : par René Mérou h(at)es.gnu.org - <http://www.es.gnu.org/~reneme/mapa-conceptual-software-libre.png> (http://www.es.gnu.org/), CC BY-SA 3.0, <https://commons.wikimedia.org/w/index.php?curid=1377588>

SOURCES DES IMAGES (II)

- Entité – Modèle de liens : [https://commons.wikimedia.org/wiki/File:Entity-Relationship_Model_\(diagram\).png](https://commons.wikimedia.org/wiki/File:Entity-Relationship_Model_(diagram).png)
- Une représentation ontologique améliorée des cellules dendritiques comme paradigme pour tous les types de cellules : par Masci AM, Arighi CN, Diehl AD, Lieberman AE, Mungall C, Scheuermann RH, Smith B, Cowell LG - , CC BY 2.0, <https://commons.wikimedia.org/w/index.php?curid=7355297>
- Une goutte de matière solaire : par le Goddard Space Flight Center de la NASA de Greenbelt, MD, USA - Twisting Blob of Plasma, domaine public, <https://commons.wikimedia.org/w/index.php?curid=51482685>
- Carte partielle de l'internet : par The Opte Project - À l'origine de Wikipedia en anglais; la page de description est/était ici, CC BY 2.5, <https://commons.wikimedia.org/w/index.php?curid=25698718>
- Astronomes ottomans : <https://commons.wikimedia.org/w/index.php?curid=4932717>