

**Collectif francophone pour l'enseignement libre de l'informatique**

# **Théorie et modèles relationnels**

## *Introduction*

**TMR\_01**

Christina KHNAISSER ([christina.khnaisser@usherbrooke.ca](mailto:christina.khnaisser@usherbrooke.ca))

Luc LAVOIE ([luc.lavoie@usherbrooke.ca](mailto:luc.lavoie@usherbrooke.ca))

*(les auteurs sont cités en ordre alphabétique nominal)*

*CoFELI/Scriptorum/TMR\_01-Introduction (v105), version 1.0.0.b, en date du 2024-09-02*

*— document de travail, ne pas citer —*

# Plan

Introduction .....	3
1. L'information, les données et leur traitement.....	4
2. Principe d'adéquation informatique.....	7
3. Caractérisation d'un problème .....	11
Conclusion .....	24
Références.....	25

# Introduction

Le présent document a pour but de présenter une vue générale de la discipline « Modélisation, conception et exploitation de données » et, plus particulièrement, de la situer relativement à la science (l'informatique) et aux technologies de l'information.

# 1. L'information, les données et leur traitement

## 1.1. Données et information

- Des données représentent un fait à propos d'une entité.
- L'information est une interprétation des données selon un domaine d'application.

Par exemple, l'ensemble des personnes étudiantes inscrites à une activité peut être représentées par les données suivantes, la signification précise de chacune des données et les liens entre elles (l'information) devant être précisée, par ailleurs, à l'aide de prédicats — d'où l'importance qui sera accordée aux prédicats dans la suite du module.

*Tableau 1. Étudiant*

matricule	nom	ville
15113150	Paul	>Δ <sup>ς</sup> σ <sup>ς</sup> ᵇ
15112354	Éliane	Blanc-Sablon
15113870	Mohamed	Tadoussac
15110132	Sergeï	Chandler

## 1.2. Traitement

L'informatique est la science du traitement rationnel et automatique de l'information. Son objet est donc l'information et son traitement.

- Mais pourquoi rationnel ?
- Pourquoi automatique ?
- Pourquoi restreindre le traitement à l'intersection de ces deux caractéristiques ?

## 2. Principe d'adéquation informatique

En regard d'exigences exprimées en termes d'un modèle bien fondé (c'est-à-dire reposant sur un cadre théorique convenu et documenté), un artéfact informatique est adéquat s'il répond aux huit propriétés suivantes :

## *Propriétés absolues*

Les propriétés absolues découlent des exigences fonctionnelles. L'artéfact doit être

### 1. **valide**

- conforme à la théorie soutenant le modèle à l'aide duquel les exigences sont formulées (les solutions apportées sont correctes);

### 2. **efficace**

- conforme aux traitements requis par les exigences ;

### 3. **cohérent**

- ne pas comporter pas de contradiction (ne permet pas d'en induire).



## *Propriétés relatives*

Les propriétés relatives découlent des exigences *non* fonctionnelles et sont subordonnées aux propriétés absolues. L'artéfact doit être

### 4. **complet**

- offrir une couverture « suffisante » du problème ;

### 5. **efficace**

- utiliser « bien » les ressources ;

### 6. **évolutif**

- être « aisément » adaptable aux changements.

## *Méta-propriété*

Les méta-propriétés découlent de l'épistémologie et de l'éthique. L'artéfact doit être

### **7. réfutable**

- apte à permettre l'invalidation ;

### **8. acceptable**

- conforme aux principes et règles de conduite propres à une société humaine de référence.

# 3. Caractérisation d'un problème

## 3.1. Besoins

Les organisations humaines ont besoin de traiter, de conserver et d'analyser de (très) grandes quantités de données dans tous les domaines, par exemple :

- gouvernements (recensement, impôts, santé...);
- recherche scientifique (astronomie, chimie, génétique...);
- télécommunications;
- banques et assurances;
- secteurs de production (énergétique, industriel, manufacturier...);
- grande distribution;
- agences de propagande et de marketing.

## 3.2. Comment parler du problème ?

Comment caractériser les problèmes de traitement de données afin de déterminer les solutions les plus adéquates ?

- L'approche descriptive des **8V**
- La hiérarchisation des modèles

## 3.3. L'approche descriptive des 8 V

### *Les quatre critères classiques*

- **volume**
  - quantité de données devant être stockées ;
- **variété**
  - diversité et complexité des types utilisés par le modèle ;
- **vélocité**
  - caractérisation du débit de données entrantes et sortantes ;
- **véracité**
  - caractérisation, voire évaluation, de l'incertitude.

## *Les quatre critères supplémentaires*

- **valeur**

- richesse analytique des données, mais aussi valeur économique, politique, sociale ou sociétale ;

- **variabilité**

- relativité des données : temps, espace, agent ;

- **virtualité**

- distribution des sources de données ;

- **vertu**

- gouvernance des données, rendre compte des lois et règlements quant à la protection et l'utilisation éthique des données.

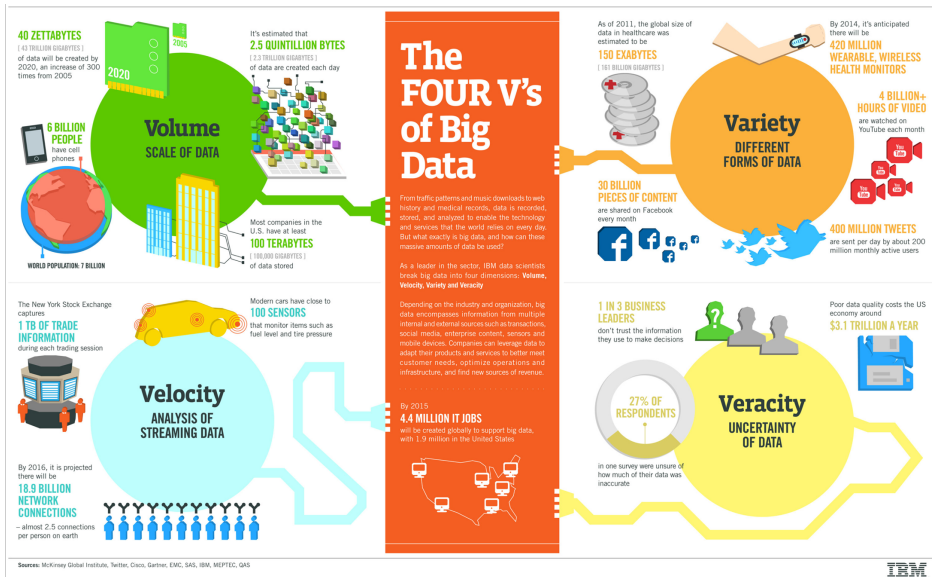


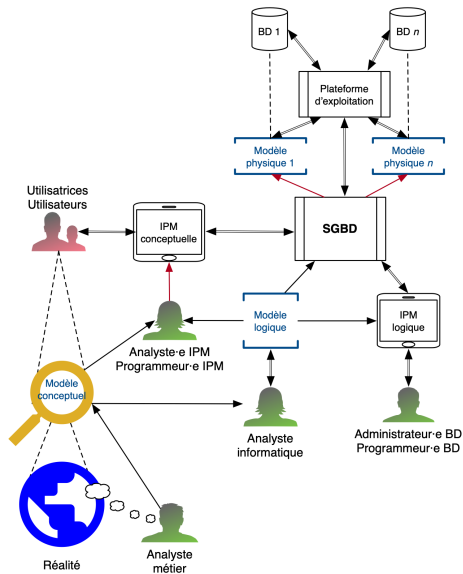
Figure 1. Illustration des 4V selon IBM (2013)

## 3.4. La hiérarchisation des modèles

Avec l'approche tri-schématique:

- Modèle physique
- Modèle logique
- Modèle conceptuel





*Figure 2. Illustration de l'approche tri-schématique (tradition)*

### 3.4.1. Modèle conceptuel

Le modèle «conceptuel» (ou modèle d'information) est la représentation de la portion intéressante de l'univers (de la réalité).

Les caractéristiques généralement souhaitables sont:

- la conformité à un méta-modèle conceptuel: entité-association (étendu), UML, Merise, etc.
- la capacité d'induire (automatiquement, quasi-automatique et quasi-complète d'un modèle logique.

## 3.4.2. Modèle logique

Le modèle logique est une représentation intermédiaire permettant de découpler la réalité de la représentation opératoire.

Les caractéristiques généralement souhaitables sont :

- un fondement mathématique solide permettant de formuler et de démontrer certaines propriétés, dont l'intégrité ;
- la conformité à un méta-modèle logique : relationnel, objet, graphe, etc.
- la capacité d'induire (automatiquement) un modèle physique « raisonnable ».

### 3.4.3. Modèle physique

Le modèle « physique » détermine le choix des représentations opératoires en regard d'un automate particulier (un ordinateur).

Les structures d'accès :

- Adressage dispersé
- Arbre de recherche (B-Tree), etc.

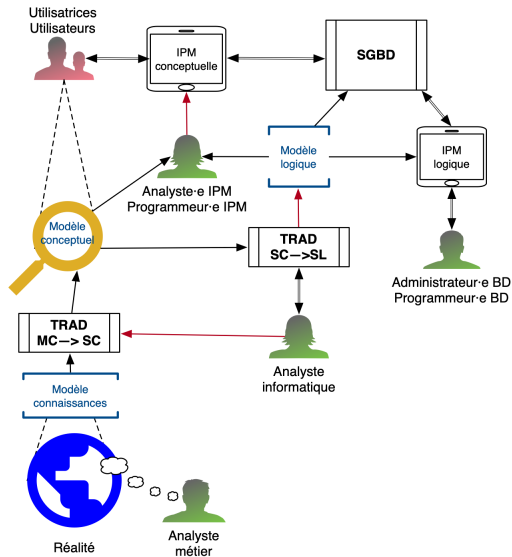
Les techniques de stockages :

- Stockage horizontal
- Stockage vertical
- Stockage mixte, etc.

### 3.4.4. Modèle de connaissances

Beaucoup de chercheurs estiment aujourd'hui qu'un quatrième niveau est nécessaire : celui des modèles de connaissances.

- Le modèle de connaissances est plus souvent informel ; il est *dans la tête* des experts et des parties prenantes, au mieux dispersé dans une collection de documents non totalement répertoriés, pas forcément à jour, ni cohérents entre eux.
- Les ontologies appliquées ont connu de très beaux succès dans le domaine scientifique.
- Il vaut toujours mieux un modèle (formel), même mauvais, que pas de modèle !



*Figure 3. Illustration de l'approche tri-schématique (tendance)*

## 3.5. Tendances

Intégrer différents types de données

- Structurées
- Semi-structurées
- Non structurées

Intégrer différentes sources de données

- Internet des objets
- Médias sociaux

# Conclusion

Une base de données est un modèle d'un domaine d'application. C'est une solution pour conserver et traiter une (très) grande quantité de données pour produire de l'information.



# Références

## [Date2014a]

Chris J. DATE, Hugh DARWEN, Nikos A. LORENTZOS;

*Time and Relational Theory: Temporal Databases in the Relational Model and SQL*;

Morgan Kaufmann, Waltham (MA, US), 2014;

ISBN 978-0-12-800631-3.

## [DoD2010a]

Deputy Chief Information Officer;

*DoDAF — DoD Architecture Framework*;

Version 2.02, U. S. Department of Defense, 2010;

<http://dodcio.defense.gov/Library/DoD-Architecture-Framework/> (consulté le 2024-05-30).

## [Elmasri2016]

Ramez ELMASRI et Shamkant B. NAVATHE;

*Fundamentals of database systems*;

7<sup>th</sup> Edition, Pearson, Hoboken (NJ, US), 2016;

ISBN 978-0-13-397077-7.

Produit le 2025-01-07 11:03:05 UTC



**Collectif francophone pour l'enseignement libre de l'informatique**