

No 11-522-XIF au catalogue

**La série des symposiums internationaux  
de Statistique Canada - Recueil**

**Symposium 2005 : Défis  
méthodologiques reliés aux  
besoins futurs d'information**



2005



Statistique  
Canada

Statistics  
Canada

Canada

# CONSTRUCTIONS COMMUNES DE MÉTADONNÉES CONCERNANT LES DONNÉES STATISTIQUES

Daniel W. Gillman<sup>1, 2</sup>

## RÉSUMÉ

Quelles que soient les particularités d'un schéma de métadonnées, il existe des constructions communes de métadonnées qui servent à décrire les données statistiques. Le présent document donne un aperçu des différentes approches adoptées pour atteindre l'objectif commun de fournir des renseignements cohérents servant à étayer l'interprétation des données statistiques et fait état des parallèles et des similitudes entre ces approches. Il présente ensuite l'approche adoptée par la norme ISO/IEC 11179.

MOTS CLÉS : DDI; ISO/IEC 11179; Neuchâtel; SDMX; Théorie de la terminologie

## 1. INTRODUCTION

Depuis plus de quinze ans, les bureaux de la statistique du monde entier se préoccupent beaucoup de la gestion des métadonnées. Entre autres utilisations, les métadonnées statistiques, c'est-à-dire les métadonnées concernant les données et les processus statistiques, servent à bonifier la recherche et la compréhension des utilisateurs, à améliorer le traitement des données d'enquête au sein de chaque bureau et à faciliter l'harmonisation des données statistiques. Il s'agit donc d'un secteur propice à la recherche et au développement. Bon nombre de bureaux utilisent un système de métadonnées pour automatiser certaines parties du processus d'enquête (Johanis, 2000; Oakley, 2004).

Avant de construire un système de métadonnées, il faut comprendre et assimiler plusieurs notions. D'abord et avant tout, il faut savoir ce qui constitue des métadonnées pour le problème à l'étude. Le concept de métadonnées n'est pas absolu. Les données ne sont pas des métadonnées en vertu de propriétés inhérentes, mais en vertu de leur utilisation. La notion de métadonnées est donc relative. Les données deviennent des métadonnées lorsqu'elles sont mises dans une relation descriptive avec d'autres données (Gillman, 2005; Farance et Gillman, 2006).

Après avoir compris les éléments nécessaires aux métadonnées, on peut construire un modèle des métadonnées à utiliser. Le modèle est un cadre général d'organisation des métadonnées dans une base de données; on en optimise souvent la structure pour améliorer les utilisations de la base de données (Date, 2003). Les constructions communes aux modèles et à leurs attributs font l'objet du présent document, les constructions de métadonnées étant des composantes des modèles.

La plupart des situations exigent un certain travail de modélisation. Il est rare qu'un modèle utilisé dans un bureau de la statistique fonctionne sans modification dans un autre bureau. Au moment de concevoir un modèle, il convient de se poser la question suivante : « Quelqu'un d'autre s'est-il penché sur ce problème et existe-t-il une solution que je pourrais appliquer à ma situation? » Les modèles existants peuvent ne pas fonctionner du tout; ils peuvent fonctionner pour certaines fins, mais pas pour d'autres; ou ils peuvent fonctionner intégralement. Dans le cas des modèles qui conviennent partiellement, on peut les faire fonctionner s'il est possible de les modifier, ce qui est habituellement le cas.

---

<sup>1</sup> Daniel W. Gillman, Bureau of Labor Statistics, Office of Survey Methods Research, Room 1950, 2 Massachusetts Ave. NE, Washington, DC 20212, USA, [Gillman\\_D@BLS.Gov](mailto:Gillman_D@BLS.Gov)

<sup>2</sup> Les opinions exprimées dans ce document sont celles de l'auteur et ne reflètent pas nécessairement les politiques du Bureau of Labor Statistics.

Où trouver des modèles qui conviennent? Il existe quatre réponses possibles : les autres bureaux de la statistique, les logiciels commerciaux, les ouvrages ou documents publiés et les normes. Les autres bureaux de la statistique constituent une excellente source de modèles de métadonnées, car ils ont mis au point plusieurs bons modèles (Johannis, 2000). Les logiciels commerciaux comportent rarement des modèles de métadonnées convenables, car les besoins des bureaux de la statistique sont trop spécialisés et la commercialisation de produits spécialisés n'est pas rentable. Les modèles de métadonnées décrits dans les ouvrages et les documents sont de niveau trop élevé, donc plus ou moins utiles à l'élaboration de systèmes. Toutefois, ils sont utiles à l'élaboration d'un cadre conceptuel commun. Enfin, les normes constituent une bonne source de modèles de métadonnées, car elles sont très détaillées et reposent sur le consensus d'un vaste groupe d'utilisateurs. Les normes sont souvent établies par une communauté de praticiens, c'est-à-dire des personnes d'entreprises semblables ou ayant des préoccupations semblables. C'est ainsi qu'on élabore des normes répondant aux besoins de groupes spécialisés, par exemple la norme Data Documentation Initiative (ICPSR, sans date).

Comme les normes et les autres bureaux de la statistique semblent constituer les meilleures sources de modèles de métadonnées pertinents, nous analyserons quatre schémas de métadonnées empruntés à chacune de ces sources. Notre analyse portera en partie sur les constructions communes. Quelles que soient les particularités d'un schéma de métadonnées, il existe des constructions communes de métadonnées qui servent à décrire les données statistiques. Le présent document donne un aperçu des différentes approches adoptées pour atteindre l'objectif commun de fournir des renseignements cohérents servant à étayer l'interprétation des données statistiques et fait état des parallèles et des similitudes entre ces approches. Il présente ensuite l'approche adoptée par la norme ISO/IEC 11179.

Le document est organisé en sections. Nous commençons par une section sur la théorie de la terminologie, qui fournit un cadre général à notre analyse. Puis, nous abordons les données statistiques en nous fondant sur la théorie de la terminologie. Nous décrivons ensuite les constructions les plus importantes des quatre schémas suivants : la norme Data Documentation Initiative (DDI); la norme ISO/IEC 11179; les modèles de variables et de classification de Neuchâtel; et l'initiative Échange de données et de métadonnées statistiques (SDMX). Enfin, nous abordons les constructions communes, comparons les différents modèles et proposons un cadre d'utilisation de ces modèles dans un bureau de la statistique.

## 2. THÉORIE DE LA TERMINOLOGIE

### 2.1 Définitions de base

Pour commencer, nous décrivons certaines constructions utiles empruntées à la théorie de la terminologie, issues de plusieurs sources (Sager, 1990; ISO, 1999; ISO, 2000) et définies ci-dessous :

- *propriété* – observation servant à décrire ou distinguer un objet (par ex., « Daniel a les yeux gris-bleu » signifie que « les yeux gris-bleu » est la propriété de Daniel; celle-ci est ramenée à une caractéristique, la couleur des yeux, d'une personne – voir caractéristique);
- *objet* – chose imaginable ou perceptible;
- *caractéristique* – abstraction d'une propriété d'un ensemble d'objets;
- *concept* – unité de connaissance créée par une combinaison unique de caractéristiques;
- *intension* – somme de caractéristiques qui constituent un concept;
- *extension* – ensemble d'objets auxquels correspond un concept;
- *définition* – expression d'un concept au moyen du langage naturel, qui spécifie une intension et une extension uniques;

- *système conceptuel* – ensemble de concepts structurés en fonction des relations entre eux;
- *désignation* – représentation d'un concept par un signe qui le désigne;
- *concept général* – concept auquel correspondent deux ou plusieurs objets (par ex., planète, tour);
- *concept individuel* – concept auquel correspond un seul objet (par ex., Saturne, la tour Eiffel).

Il existe trois types de désignation : un *terme* est une désignation verbale d'un concept général, une *appellation* est une désignation verbale d'un concept individuel et un *symbole* est toute autre désignation.

Les philosophes grecs de l'Antiquité furent les premiers à étudier la terminologie et la formation des concepts langagiers (Wedberg, 1982) et découvrirent une relation utile entre désignation, concept, objet et définition, qui est illustrée dans la figure 1 (CEN, 1995). Ce diagramme, si on exclut la partie *définition*, est souvent appelé triangle d'Ogden (Ogden et Richard, 1989).

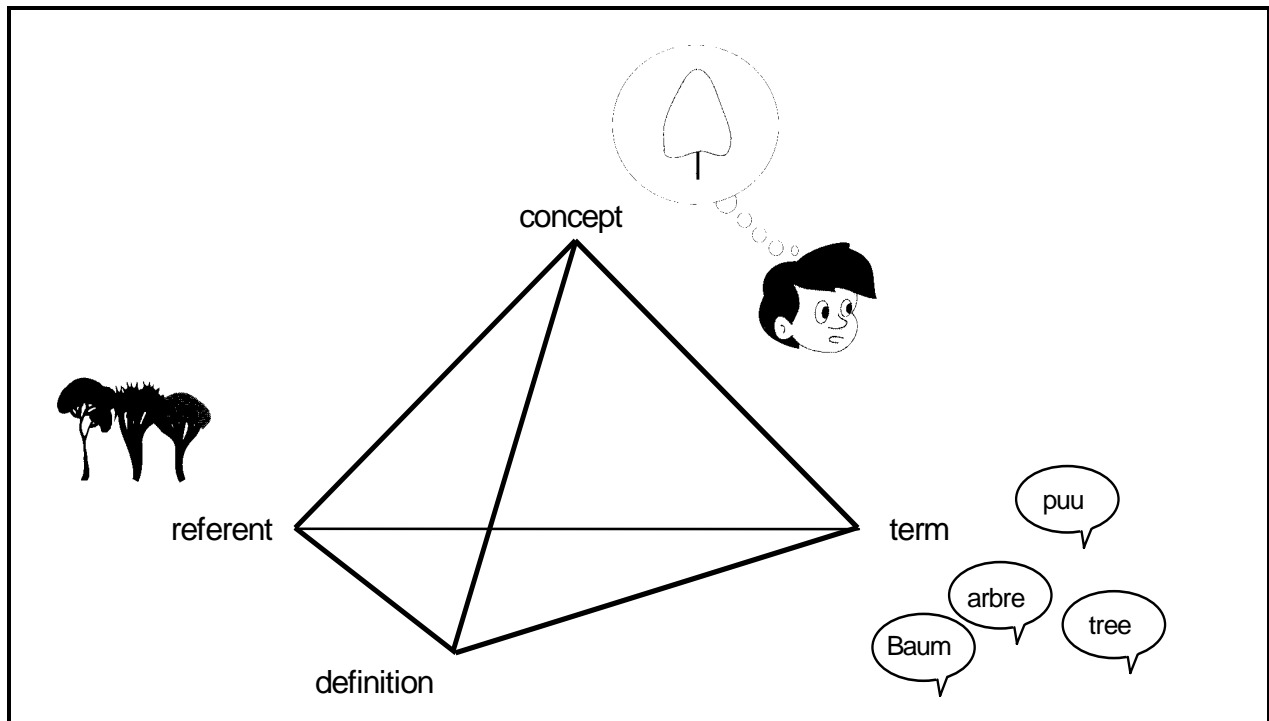


Figure 1. Relations entre référents (objets), concepts, termes (ou, plus généralement, désignations) et définitions.

Il importe de souligner que les concepts sont des notions humaines (Lakoff, 2002). Il est souvent impossible de donner une description complète d'un concept, si bien défini soit-il. L'identification des caractéristiques pertinentes dépend de la culture. Aussi certains objets compris dans l'extension d'un concept, appelés prototypes, correspondent-ils mieux que d'autres aux caractéristiques (Lakoff, 2002). Par exemple, un aigle correspond mieux qu'un pingouin aux caractéristiques d'un oiseau.

## 2.2 Données

Il importe également de noter qu'une donnée est une désignation (Farance et Gillman, 2006). Une valeur obtenue au moyen du processus de l'interview ou enregistrée dans une base de données (exemple d'un signe – voir la section 2.1) représente un concept; c'est donc une désignation, fort probablement un symbole.

Les statisticiens considèrent une donnée comme une désignation d'une classe dans une partition d'une population d'objets, où la partition<sup>3</sup> est définie pour une caractéristique de la population (Froeschl, Grossmann et Del Vecchio, 2003). Ici, dans une optique qui diffère légèrement de la théorie traditionnelle de la statistique, la population est un concept général ou individuel et les objets sont l'extension de ce concept. En voici deux exemples :

#### Exemple 1 : Concept général de population

Population : Ensemble d'adultes âgés de 16 ans et plus au Canada en 2006
Caractéristique : Sexe
Partition : {hommes, femmes}
Désignations : 0 pour les hommes 1 pour les femmes

#### Exemple 2 : Concept individuel de population

Population : Ensemble d'adultes âgés de 16 ans et plus au Canada en 2006
Caractéristique : Proportion de femmes
Partition : $\{x \mid 0 \leq x \leq 1\}$
Désignations : Nombres réels entre 0 et 1, avec une précision à trois décimales

De plus, on emploie le terme « population » dans deux sens. D'une part, une population est un concept représenté par une définition ou une description. D'autre part, elle est l'extension du concept, soit l'ensemble d'objets au sujet duquel nous recueillons ou observons des données.

Le concept de population est individuel ou général. Dans le cas des microdonnées, les populations sont des concepts généraux – voir l'exemple 1. Toutefois, les données agrégées, ou macrodonnées, exigent qu'une population corresponde à un seul objet. Dans l'exemple 2, la caractéristique « proportion de femmes » s'applique à l'ensemble contenant l'ensemble d'adultes âgés de 16 ans et plus au Canada en 2006, et non à chaque personne. L'ensemble composé de « l'ensemble d'adultes âgés de 16 ans et plus au Canada en 2006 » compte un seul élément, mais c'est cet agrégat qui possède la caractéristique « proportion de femmes », et non chaque personne. La « proportion de femmes » n'est pas une caractéristique des personnes, mais le « sexe » en est une. De même, le « sexe » n'est pas une caractéristique de l'agrégat, mais la « proportion de femmes » en est une.

Lorsqu'on évoque un objet donné, la notion de cet objet est un concept individuel. Cela signifie qu'à chaque objet correspond un concept individuel. Les données liées à un objet donné sont descriptives de cet objet; il s'agit donc de métadonnées. Les données ne sont des métadonnées que lorsqu'elles servent à décrire un objet. Toutefois, les métadonnées sont semblables à des données agrégées; leurs populations sont des concepts individuels. Cela implique qu'au point de collecte, toutes les données sont des métadonnées!

Le cadre général que nous venons de décrire sert de base à la description de toutes les données statistiques. Dans les sections qui suivent, nous allons décrire dans quelle mesure chaque schéma de métadonnées suit ce cadre général.

## 2.3 Mise en œuvre

La norme ISO/IEC 11179 met en œuvre la théorie terminologique des données de manière très directe. À chacune des constructions de la théorie correspond une construction dans la norme (ISO, 2005)

---

<sup>3</sup> Une partition est un ensemble non vide de sous-ensembles incompatibles et exhaustifs d'un autre ensemble. Le nombre de sous-ensembles n'est pas nécessairement fini.

Sans entrer dans les détails du modèle décrit dans la norme ni définir tous les termes, nous allons appairer les termes de la norme à ceux de la théorie terminologique. Pour plus de détails, voir la norme ISO/IEC 11179-1 : *Cadre général* (ISO, 2005).

Voici la liste; les termes de la norme ISO/IEC 11179 figurent à gauche :

- |                          |                     |
|--------------------------|---------------------|
| • classe d'objet         | population          |
| • propriété <sup>4</sup> | caractéristique     |
| • partition              | domaine conceptuel  |
| • catégories             | domaines de valeur  |
| • désignations (valeurs) | valeurs admissibles |

On pourrait appairer bien d'autres termes de la théorie terminologique à ceux de la norme, mais il faudrait les aborder de manière beaucoup plus approfondie, alors que le présent document se veut un aperçu.

### 3. SCHÉMAS DE MÉTADONNÉES

Dans la présente section, nous décrivons et comparons quatre schémas de métadonnées.

#### 3.1 DDI

Data Documentation Initiative (DDI) est une initiative internationale consistant à établir une norme de métadonnées en langage XML pour le contenu, la présentation, le transport et la préservation de la documentation concernant des ensembles de données en sciences sociales. Les spécialistes des sciences sociales doivent consigner et communiquer de manière directe toutes les caractéristiques importantes des données empiriques dont ils sont responsables. Tel est l'objet de la norme DDI.

Lancée par le Inter-university Consortium for Political and Social Research, la spécification de métadonnées DDI fait maintenant l'objet d'une alliance (<http://www.icpsr.umich.edu/DDI/org/index.html>) d'environ 25 établissements nord-américains et européens. Elle repose sur le principe du « livre de code » électronique, dont elle conserve les capacités, mais en accroît les possibilités en améliorant la rigueur et en étendant la portée. La norme DDI est en usage dans de nombreuses archives de données en sciences sociales et dans des bureaux de la statistique du monde entier.

La norme DDI est représentée en version XML DTD et XML-Schema (W3C, 2004). La norme DDI-DTD est répartie en cinq grands chapitres :

- Description du document – description du document XML proprement dit;
- Description de l'enquête – description de l'enquête à l'origine des données;
- Description des fichiers – disposition physique du ou des ensembles de données décrits;
- Description des données – description conceptuelle des données;
- Autres documents – descriptions des données et des documents connexes.

La norme DDI s'applique principalement à l'enquête et à l'ensemble de données du point de vue de la statistique des sciences sociales. Seul chapitre exigé, « Description de l'enquête » représente une description d'ordre supérieur des données. Cela signifie que les archives peuvent renfermer des descriptions individuelles de chaque ensemble de données qu'elles gèrent, mais aussi que certaines métadonnées concernant une série d'ensembles de données provenant de la même enquête ou du même programme doivent être répétées.

La norme DDI comporte un riche ensemble d'éléments consacrés aux besoins des statisticiens et d'autres utilisateurs de données statistiques. C'est le seul des quatre schémas qui est expressément conçu pour décrire des enquêtes

---

<sup>4</sup> Ce terme est très mal choisi. Dans la prochaine édition (la troisième), il sera remplacé par « caractéristique ».

statistiques. Le Groupe de Neuchâtel, décrit plus loin, produit également des normes directement liées aux bureaux de la statistique, mais elles sont beaucoup plus spécialisées.

Dans la section « Description des variables » du chapitre « Description des données », on trouve certains éléments permettant de saisir les concepts, mais peu en fait de gestion des concepts. Cette lacune est due en partie à la conception. Le langage XML étant hiérarchique, il est difficile de modéliser des structures de relations complexes. Les révisions apportées à la norme DDI devraient y remédier en partie.

### 3.2 ISO/IEC 11179

La norme ISO/IEC 11179 – *Registres de métadonnées* – est une spécification de métadonnées consacrée à la sémantique des données. Elle contient également un modèle et un aperçu d'une méthode d'enregistrement, d'où le mot « registres » dans le titre, mais elle porte surtout sur la sémantique.

La norme comporte six parties :

- Partie 1 – Cadre général – aperçu de la norme et de la méthodologie qui sous-tendent la sémantique des données;
- Partie 2 – Classification – présentation d'un modèle de gestion d'un schéma de classification, notamment pour lier les éléments de données (variables) les uns aux autres;
- Partie 3 – Métamodèle et attributs de base – présentation du modèle complet pour la sémantique, la classification et l'enregistrement des données;
- Partie 4 – Formulation des définitions de données – principes de formulation des définitions de données;
- Partie 5 – Principes de dénomination et d'identification – présentation d'une convention de dénomination pour chacune des principales parties de la sémantique des données;
- Partie 6 – Enregistrement – méthodes d'enregistrement des éléments de données.

La dernière version publiée de la norme ISO/IEC 11179 est la deuxième édition, celle de 2005. Toutes les parties publiées les plus récentes de la norme sont accessibles sans frais sur le site Web du Groupe spécial Technologies de l'information formé par ISO<sup>5</sup> et la CEI<sup>6</sup> ([http://isotc.iso.org/livelink/livelink/fetch/2000/2489/Itf\\_Home/ITTF.htm](http://isotc.iso.org/livelink/livelink/fetch/2000/2489/Itf_Home/ITTF.htm)).

L'unité de base de la norme ISO/IEC 11179 est l'élément de données (variable). Le modèle spécifié dans la norme montre comment décrire un élément de données. Il s'agit d'un modèle conceptuel qui suit le cadre général de la théorie terminologique des données, décrite plus haut.

Toutefois, la norme ne porte pas sur les données statistiques en soi. Elle présente une description générale des données, sans plus. Elle ne décrit même pas la notion d'ensemble de données.

### 3.3 Groupe de Neuchâtel

Le Groupe de Neuchâtel est composé de représentants des bureaux de la statistique des Pays-Bas, de la Norvège et de la Suède, de l'Office fédéral de la statistique de Suisse et du Bureau of Labor Statistics des États-Unis, ainsi que d'un petit fabricant allemand de logiciels, run-Software AG, qui construit des logiciels de mise en œuvre des spécifications.

Une version antérieure du groupe a mis au point un modèle de gestion des systèmes de classification utilisés dans les bureaux de la statistique, appelé le modèle de classification du Groupe de Neuchâtel. Il est utilisé dans de nombreux bureaux de la statistique en Europe, en Amérique du Nord, en Australie et en Nouvelle-Zélande.

Le modèle de variables du Groupe de Neuchâtel est encore en cours d'élaboration. La principale construction qu'il décrit est évidemment la variable ou élément de données. Le modèle contient un grand nombre des mêmes composantes de données que la norme ISO/IEC 11179. La première version de la spécification sera publiée d'ici la

---

<sup>5</sup> Organisation internationale de normalisation.

<sup>6</sup> Commission électrotechnique internationale.

fin de l'année. Elle est fortement inspirée de la norme ISO/IEC 11179, tout en présentant aussi des différences importantes. Tout d'abord, elle est formulée à l'aide de termes statistiques et contient des détails sur les données statistiques qu'on ne trouve pas dans la norme ISO/IEC 11179. De plus, outre les variables de base, elle décrit des bases de données, des fichiers et des formats.

Le modèle de variables n'offre pas la capacité d'enregistrement, ni la souplesse qu'offre la norme ISO/IEC 11179. Toutefois, en ce qui concerne la description d'un élément de données, les deux spécifications sont très semblables.

Puisque la spécification de Neuchâtel est encore en cours d'élaboration, le groupe refuse de distribuer le document à l'heure actuelle; c'est pourquoi nous n'y faisons aucun renvoi. De plus, ses ressources étant limitées, le groupe ne possède pas de site Web. Ce dernier est également en cours d'élaboration.

### **3.4 Échange de données et de métadonnées statistiques (SDMX)**

L'initiative d'échange de données et de métadonnées statistiques (*Statistical Data and Metadata Exchange*, ou SDMX) est parrainée conjointement par sept organismes statistiques et financiers internationaux : la Banque mondiale, la Banque des règlements internationaux (BRI), le Fonds monétaire international (FMI), l'Organisation de coopération et de développement économiques (OCDE), Eurostat, la Banque centrale européenne (BCE) et la Division de la statistique des Nations Unies. Chaque organisme a besoin de décrire, de partager et de transférer des statistiques et leurs métadonnées.

Lancé en 2001, le projet en est à sa deuxième version, disponible sur son site Web (<http://www.sdmx.org>). Il fera également l'objet d'une norme internationale, la norme ISO 17369.

Le modèle SDMX est très perfectionné et très général. Il n'a pas encore fait l'objet d'une analyse complète, mais le fait qu'il n'a pas été conçu expressément pour la statistique est préoccupant. Il n'y a pas assez d'attributs pour situer le modèle directement dans le domaine de la statistique. Par conséquent, l'aspect général du modèle risque de satisfaire les besoins de nombreux types d'entreprise. Il peut s'agir aussi d'un avantage, car les concepteurs peuvent utiliser le modèle pour transférer plusieurs types de données et de métadonnées. Toutefois, l'appellation SDMX ne convient peut-être pas au modèle.

La principale construction du modèle SDMX est la *famille clé* et les catégories connexes. Bon nombre de composantes des données y sont décrites. De plus, les auteurs affirment que tout le contenu de la norme ISO/IEC 11179 peut être apparié au modèle SDMX. On ne l'a pas vérifié de manière rigoureuse, mais il existe un lien étroit entre les normes.

L'enregistrement fait également partie du modèle SDMX. Les concepteurs ont suivi la conception de la spécification de registre ebXML définie par le consortium OASIS (Organization for the Advancement of Structured Information Standards) (OASIS, 2002). Généralisée d'après la notion d'enregistrement énoncée dans la norme ISO/IEC 11179, cette spécification de registre est semblable sur le plan conceptuel.

### **3.5 Comparaison**

La principale construction commune à ces modèles est la variable statistique. Alors que le modèle de variables du Groupe de Neuchâtel est axé sur les microdonnées et que le modèle ISO/IEC 11179 décrit n'importe quel type de données, le modèle SDMX se concentre sur les macrodonnées et les séries chronologiques, quoique les modèles de SDMX et de Neuchâtel permettent aussi de décrire d'autres types de données. La norme DDI décrit des microdonnées et des macrodonnées, mais la description n'est pas aussi détaillée.

Comme la norme ISO/IEC 11179 contient la description la plus générale des données, il s'agit du principal modèle parmi les quatre. Comme nous l'avons montré dans la section 2.3, la norme ISO/IEC 11179 met en œuvre la description terminologique des données. D'après l'exposé de la section 2.2, le modèle ISO/IEC 11179 est alors fondamental. Cela ne signifie pas pour autant que les modèles de SDMX et de Neuchâtel n'offrent pas la même fonctionnalité. Toutefois, chacun d'eux est consacré à un domaine spécialisé, celui de la statistique. Ils sont conçus pour offrir aux statisticiens d'autres fonctions que la description de données statistiques.



Comment apparie-t-on ces normes? Une bonne façon consiste à considérer la norme ISO/IEC 11179 comme le moyeu d'une roue et les autres spécifications comme les rayons de cette roue. Pour apparier deux normes, on apparie d'abord une norme radiale à celle du centre, puis à celle à laquelle on veut l'apparier. Comme il y a seulement quatre modèles en présence, il est facile de créer des appariements par paire (il y en a six.). Toutefois, il n'y a que trois appariements par rayon.

Un bureau de la statistique peut se servir de ces quatre spécifications pour créer une spécification complète de métadonnées. La norme ISO/IEC 11179 décrit l'enregistrement et la sémantique des données. Le modèle de variables du Groupe de Neuchâtel décrit la sémantique des données, les ensembles de données, les cubes et les bases de données. La norme DDI décrit les ensembles de données, les cubes et les statistiques. La norme SDMX décrit l'enregistrement et le transfert de données ou de métadonnées. La combinaison offre un cadre général aux multiples possibilités.

## RÉFÉRENCES

- CEN. (1995), *Medical Informatics - Categorical Structures of Systems of Concepts*. Ébauche, Brussels: European Committee for Standardization.
- Date, C. (2003), *An Introduction to Database Systems (8<sup>th</sup> ed)*. Addison Wesley.
- Farance, F. et Gillman, D. (2006), "The Nature of Data", Manuscrit soumis pour publication.
- Froeschl, K., Grossmann, W. et Del Vecchio, V. (2003), *The Concept of Statistical Metadata*. Deliverable #5 pour projet MetaNet. Obtenu juillet 2004 du site [http://www.epros.ed.ac.uk/metanet/deliverables/D5/IST-1999-29093\\_D5.doc](http://www.epros.ed.ac.uk/metanet/deliverables/D5/IST-1999-29093_D5.doc).
- Gillman, D. (2005), "Data Semantics", Dans D. Schwartz (ed.) *Encyclopedia of Knowledge Management*. Hershey, PA: Idea Group.
- ICPSR (Inter-University Consortium for Political and Social Research). (n.d.), *Data Documentation Initiative*, Obtenu juillet 2004 du site <http://www.icpsr.umich.edu/ddi>.
- ISO. (1999), *ISO 704: Principles and methods of terminology*. Geneva: International Organization for Standardization.
- ISO. (2000), *ISO 1087-1: Terminology – Part 1: Vocabulary*. Geneva: International Organization for Standardization.
- ISO. (2005), *ISO/IEC 11179 - Metadata registries (All Parts)*. Geneva: International Organization for Standardization and International Electrotechnical Commission.
- Johanis, P. (2000, novembre), *Statistics Canada's Integrated Metadatabase: Our Experience To Date*. Article invité #3 présenté au "UNECE Workshop on Statistical Metadata", Washington, DC.
- Lakoff, G. (2002), *Women, Fire, and Dangerous Things* (Édition ré-impression). University of Chicago Press.
- Oakley, G. (2004, février), *Using ISO/IEC 11179 to help with metadata management problems*. Article invité #9 présenté au "Joint UNECE, Eurostat, OECD Workshop on Statistical Metadata", Geneva.
- OASIS. (2002), *OASIS/ebXML Registry Information Model, v2.0*. Organization for the Advancement of Structured Information Standards. Obtenu janvier 2006 du site <http://www.oasis-open.org>.

Ogden, C. et Richard, I. (1989), *The Meaning of Meaning: A Study of the Influence of Language upon Thought and of the Science of Symbolism*. Harcourt.

Sager, J. (1990), *A Practical Course in Terminology Processing*. John Benjamins.

Wedberg, A. (1982), *A History of Philosophy – Vol. 1: Antiquity and the Middle Ages*. Oxford: Clarendon Press

W3C. (2004), *Extensible Markup Language*. XML 1.1 reference specification. Obtenu juillet 2004 du site <http://www.w3c.org>.