

Chapitre 3 : LES TECHNIQUES DE REPRESENTATION DE L'INFORMATION

Introduction:

Les données traitées par les ordinateurs peuvent varier en types, incluant texte, nombres, images, sons, vidéos, etc. Cependant, quelle que soit leur forme originale, elles sont toujours converties en une représentation numérique par l'ordinateur. Le processus de codage consiste à faire correspondre la forme externe de l'information (comme du texte ou une image) avec sa représentation interne dans la machine, qui est toujours une série de bits.

1. Notion d'information

Une information est un renseignement qui accroît la connaissance concernant une personne, un objet ou un événement déterminé. L'information peut être :

- objective, quand elle reflète un ensemble de données porteur de sens ;
- subjective, quand elle résulte de l'interprétation d'un ensemble de données.

2. Notion de donnée

Une donnée ne devient une information que quand elle trouve son sens par rapport à un référentiel (un contexte, un système de valeurs, un problème à résoudre. . .). Dans les systèmes informatiques, la donnée est par ailleurs la traduction codée d'une information.

3. Les caractéristiques de l'information

L'information a des caractéristiques de contenu, de coût et de qualité.

3.1. Le contenu sémantique de l'information

Une information est plus ou moins sélective (quantité de connaissances nouvelles apportées par l'information), synthétique et précise. De plus elle est positionnée dans le temps.

3.2. Le coût et la valeur de l'information

a) Le coût de l'information

Le coût d'une information correspond à l'ensemble des sommes engagées par l'entreprise pour sa recherche, son traitement, son stockage et même sa destruction.

b) La valeur de l'information

Une information a de la valeur si elle permet de réduire l'incertitude, d'améliorer la décision et d'éviter de commettre des erreurs. Cette valeur est appréciée sur la base de trois critères de qualité : la pertinence, la fiabilité et la disponibilité de l'information.

3.3. Qualité de l'information

Il y a trois critères essentiels pour caractériser la qualité d'une information:

- La fiabilité (est-elle exacte et à jour?)
- La disponibilité (parvient-elle au bon moment, aux bons destinataires et sous une forme directement et rapidement exploitable?)

- La pertinence (est-elle autorisée par la législation et ni redondante ni calculable à partir d'autres informations?).

Exemple :

Si la date de naissance d'un individu est déjà mémorisée dans une base de données, son âge n'est pas une information pertinente puisqu'il est calculable à partir de la date de naissance.

4. Les principaux rôles de l'information

4.1 Support des processus de gestion

Traitement des commandes, tenue d'une comptabilité, programmation d'une action, ... Nécessitent que l'on dispose d'informations.

4.2 Outil de communication

- L'échange d'informations est nécessaire pour la réalisation de nombreuses activités.
- La technologie a également beaucoup fait évoluer cet aspect, (aux systèmes de messagerie et courrier électronique, aux réseaux (Internet bien sûr, mais aussi l'intranet d'entreprise, ...))

4.3 Outil de connaissance individuelle

La capacité cognitive des organisations est aussi celle des individus qui en font partie.

4.4 Outil de liaison avec l'environnement

Une information de qualité aura un effet positif sur le climat social au sein des organisations.

5 Les Sources D'information

5.1 Sources externes (Environnement du système) :

- Flux en provenance des partenaires (Clients, Fournisseurs, Administration,)
- Être à l'écoute pour anticiper les changements et adapter son fonctionnement

5.2 Sources Internes :

- Flux générés par les acteurs du système (Approvisionnements, Production, Gestion des salariés, Comptabilité, Vente).
- Flux formalisés par des procédures.
- Flux informels (climat social, savoir-faire, ...). Ils sont difficiles à recueillir et à exploiter, mais grande importance.

6. REPRESENTATION DE L'INFORMATION

Afin de pouvoir représenter l'information dans des structures informatiques, nous allons lui donner une forme particulière appelée Données. En informatique de gestion, il existe trois grandes structures de données :

- Les tableaux
- Les fichiers.
- Les bases de données.

6.1 Les formes d'information

L'information peut être écrite, picturale, orale... Dans le cas où l'information résulte d'un traitement, on parle d'informations structurées (ou formalisées ou documentées).

6.2 La structure de l'information

L'information peut être représentée par différentes formes :

- Le caractère : une lettre, un chiffre, un symbole
- Le mot : assemblage de plusieurs caractères qui prennent une signification.
- La rubrique : assemblage de plusieurs mots qui prennent un sens. La rubrique est le plus petit ensemble d'information qui puisse être utilisé de façon autonome.
- L'article : ensemble de rubriques représentant les propriétés d'un objet; l'article permet son identification.

6.3 LES MANIPULATIONS DE L'INFORMATION

Les actions de base de l'information sont résumées dans ce qui suit :

1. Création : On crée de l'information en décidant ou bien en réutilisant ce qu'on l'on sait par ailleurs.
2. Recopie : Consiste à faire passer l'information d'un support vers un autre sans rien lui ajouter.
Exemple : copie de supports informatiques.
3. Traduction : Recopie avec modification de l'information. Elle s'applique dès qu'il y a un code d'expression, une langue. Exemple : traduire une disquette vers une imprimante, c'est à dire, on traduit le contenu de la disquette en un texte imprimé sur papier.
4. Mémorisation : c'est la conservation de l'information dans le temps. Elle suppose un rangement sur un support. Exemple : les fichiers sur disque, disquettes, CD, papier, ...
5. Transport de l'information : déplacer l'information d'un endroit à un autre en utilisant les réseaux de transmission, les bus d'E/S, les ondes, ...
6. Traitement : En informatique, le processus de traitement peut se traduire par un algorithme.
Remarque : l'information traitée n'est pas l'information créée.
Exemple : $2+3 = 5$; 5 n'est pas une information nouvelle, mais on pourra toujours la retrouver à partir des informations 2 et 3.
7. Destruction de l'information : c'est la faire disparaître. En gestion, on doit se prémunir de la destruction malchanceuse.

8. Compréhension : C'est le processus par lequel on accède à un sens de l'information. C'est un processus complexe et particulièrement difficile à réaliser par l'ordinateur. Par exemple :
 - a. La reconnaissance de la parole
 - b. La reconnaissance des formes
9. Résumé : Résumer un ensemble d'informations et réduire son volume en tentant de garder son sens. Il suppose la compréhension de l'information afin d'appliquer le procédé du résumé.

6.4 ÉTUDE DE L'INFORMATION

6.4.1 Classes d'information et réalisation de classe

Un type ou classe est un ensemble d'éléments ayant les mêmes caractéristiques. Une occurrence d'un type ou d'une classe est un élément particulier à cet ensemble. Exemple : La classe des étudiants (chaque étudiant est considéré comme une occurrence de la classe des étudiants).

6.4.2 Description de classe

Une description de classe ou de type regroupe l'ensemble de toutes les rubriques permettant de décrire une classe d'informations. Pour représenter une description de classe on utilise un descripteur qui sera composé de :

- * Nom de la classe
- * Suivie entre accolades ({}) de la liste des rubriques concernées en les séparant par des virgules.

Exemple :

AFFECTATION_ETUDIANT {Matricule, Filière, Année, Groupe}

CLIENT {CodeClient, NomClient, ADRESSE {Rue, CodePostal}}

ETUDIANT {matricule, Nom_Pren, MODULE {Code_Module, Coefficient, EMD {NumEMD, NOTE}}}

Adresse est une rubrique groupée : si l'on considère qu'un client n'a qu'une et une seule adresse. Elle deviendra une rubrique « groupée » et « répétitive » si l'on considère qu'un client peut avoir plus d'une adresse.

Module et EMD sont des rubriques répétitives de groupe.

Module est une nouvelle classe subordonnée à la classe Étudiant.

EMD est une nouvelle classe subordonnée à la classe Module.

6.4.3 **Article** : Un article est l'ensemble ordonné des occurrences des rubriques d'une description de classe.

6.4.4 **Fichier** : Un fichier est l'ensemble de tous les articles correspondant à une classe donnée. Exemple : Le fichier Étudiant.

7. CODIFICATIONS DE L'INFORMATION

La codification a un rôle très important d'organiser et structurer les informations préalablement pour que l'ordinateur puisse traiter automatiquement de grandes quantités d'information, à grande vitesse et dans d'excellentes conditions de sécurité. La codification a pour but d'améliorer la désignation des entités décrites dans les traitements d'une manière claire et unique. Un code mal conçu est une source de retards et d'erreurs.

7.1 Notion du code

Un code est une représentation abrégée d'une information permettant de désigner un objet ou un concept de manière claire et unique. Exemple : Immatriculation d'une voiture.

7.2 Notion de la codification

La codification est l'opération qui consiste à remplacer une information sous sa forme naturelle par un code clair qui serait mieux adapté aux besoins de l'utilisateur de l'information. Dans un programme informatique, il s'agirait de noms de variables et de leurs valeurs.

7.3 Caractéristique d'une bonne codification

1. **Non ambiguë** : Un objet doit correspondre à un et un seul code et inversement.
2. **Bien adapté** : Un code doit être bien adapté à l'utilisation que l'on doit en faire.
3. **Interprétable** : Un code doit être facile à décoder. Pour cela il faut utiliser des codes faciles à lire et à interpréter tel que les abréviations ou autre.
4. **Extensible** : Un code doit offrir la possibilité d'être étendu selon la classe d'objet à modifier. Exemple : Pour les produits on parlera de référence produit comme code et pour les étudiants on parlera de matricule étudiant.
5. **Avec Insertion possible** : Un code doit prévoir suffisamment de valeurs pour couvrir l'ensemble des occurrences de la classe à laquelle il est destiné. Ceci n'est pas toujours évident et parfois il est nécessaire de faire des études statistiques pour définir par exemple des plages de valeurs possibles.
6. **Concise** : Un code doit être court et facile à utiliser pour avoir un gain de temps et de place.

7.4 Les avantages de la codification

- Identifier sans ambiguïté un individu dans un ensemble
- Réaliser des gains de place et de temps
- Représenter certaines propriétés d'un objet
- Permettre certains contrôles de forme sur l'information.

7.5 Différents types de codification

Il y a plusieurs types de codification, les plus connus sont :

1. Codification séquentielle

Elle consiste à affecter des numéros consécutifs aux objets à codifier. Exemple (001,002, etc.).

Avantages

- Simplicité
- Non ambiguïté
- Possibilité d'extension.

Inconvénients

- Non significatif
- Impossibilité d'insertion
- Pas de regroupements possibles.

2. Codification par tranches

Elle consiste à diviser l'ensemble d'objets à codifier en plusieurs catégories où chaque catégorie d'objets se verra attribuer une tranche de codes. À l'intérieur des tranches, les codes sont généralement séquentiels

Exemple1 : Gestion des stocks d'une pharmacie

No 0001→0099 : Antalgique

0100→ 0599 : Anti-inflammatoire

0600→ 0799 : Fortifiant.

No 1100→1500 : Médicament interdits sans ordonnance

Remarque : les tranches peuvent être décomposées ou non.

Exemple2 : Dans une bibliothèque, les ouvrages sont classés par catégories, comme suit : technologie, littérature, sociologie, médecine et culture générale. La codification des ouvrages peut se faire comme suit :

- De 001 à 100 : Technologie
- De 101 à 200 : Littérature
- De 201 à 300 : Sociologie
- De 301 à 400 : Médecine
- De 401 à 500 : Culture générale

Avantages

- Non ambiguïté
- Simplicité
- Possibilité d'extension et d'insertion.

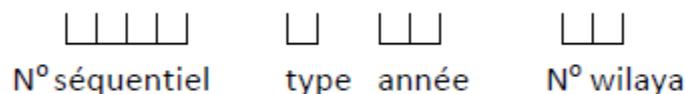
Inconvénients

- Non significatif sans table de correspondance.
- Nombre de code dans une tranche difficile à fixer.
- La répartition en catégories n'est pas toujours évidente.

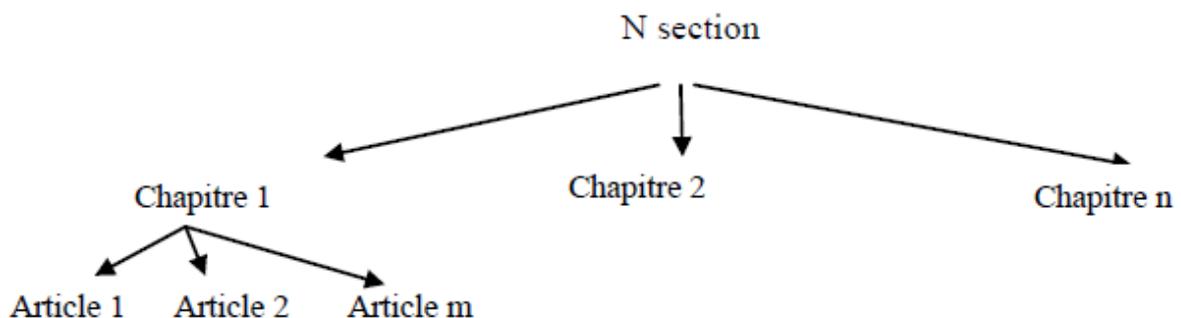
3. Codification articulée (séquentielle par blocs)

Consiste à attribuer des codes découpés en zones, chaque zone est appelée descripteur, chaque descripteur a une signification particulière relative à l'objet codifié. On distingue deux variétés de cette codification :

A) La codification juxtaposée : Le code se présente sous forme d'une série de champs les uns à côté des autres. Exemple : immatriculation d'une voiture



B) La codification hiérarchisée à plusieurs niveaux : Le code se présente sous forme d'une hiérarchie arborescente où chaque niveau de l'arbre dépend de son prédécesseur. Exemple : codification des documents juridiques tel que le journal officiel (JORA).



Avantages : non ambiguë, possibilité d'insertion et d'extension, très répondue, significative, Possibilité de regroupement et de contrôle).

Inconvénients : code long et lourd à manipulé, risque de saturation, l'instabilité : le changement d'une caractéristique de l'objet peut remettre en question toute la codification.

4. Codification mnémonique ou significative

Consiste à abrégier la désignation d'un objet à l'aide d'un ensemble réduit de caractères qui doit être évocateur de l'objet codifié. Il existe 2 sortes :

- Des codes dits « consonants », obtenus par suppression des voyelles, exemple :
FACTURE → FCTR,
- Des codes dits « abrégatifs » : obtenus par abréviation de la désignation de l'objet, exemple :
FACTURE → FACT.

Avantages : Très pratique car près du sens de l'objet.

Inconvénients : Ne se prête pas toujours au décodage automatique.

5. Codes contrôlables

Ce sont des codes dont la structure permet de déceler les erreurs de transcription. Par exemple: La numérotation d'un compte bancaire utilise une clé qui jouera le rôle de fonction de contrôle du numéro de compte. La clé de contrôle peut être un chiffre ou une lettre qui devra être déduite à partir du calcul sur la base de la clé.

Soit la fonction de contrôle « lettre modulo 23 », et le code dont le format est « 99999X ». L'occurrence 95121M donnera après calcul de $95121 \div 23$ reste 16 qui correspond à la lettre «M», d'où le code est correct.

6. Codification mixte :

Concaténation de plusieurs blocs codifiés en séquentielle ou en mnémonique. Livres : INF/001, INF/002, ...

7.6 Choix d'une codification

Pour choisir le type de codification, il faut savoir :

- De quelle manière sera utilisé le code ?
- Quel est le nombre d'information à codifier ?
- L'ensemble des informations est-il évolutif ?

8. EXERCICE d'application

On considère un établissement de formation qui offre des formations à 3 niveaux : technicien, technicien supérieur et ingénieur. Chaque niveau est organisé en 4 spécialités (gestion, finance, comptabilité et marketing). La spécialité comporte 3 sections et une section ne peut contenir plus de 90 étudiants.

- 1) Faire une codification de ces étudiants selon chaque type de codification ?
- 2) Donner les avantages et les inconvénients de chaque proposition ?

SOLUTION

Exercice 1

1. Codification séquentielle :

-Attribuer des numéros consécutifs pour tous les étudiants : $3*4*3*90=3240$

Ordre : selon l'arrivée pour inscription par exemple

1ère arrivée : code=1

2e arrivée : code = 2

.....

.....

3240^{ème} arrivée : code=3240

2. Codification par tranche

On doit choisir un critère ou une tranche pour regrouper les étudiants, prenons la spécialité comme critère. Nous avons 4 spécialités et donc 810 étudiants par spécialité au maximum Nous proposons la codification par tranche suivante

- o De 1 à 810 : tranche pour gestion
- o De 811 à 1620 : tranche pour finance
- o De 1621 à 2430 : tranche pour comptabilité
- o De 2431 à 3240 tranche pour marketing

La codification dans chaque tranche est séquentielle

Exemples

- o 1er arrivée, un étudiant de comptabilité : son code=1621
- o 2e arrivée, un étudiant de marketing : son code=2431
- o 3e arrivée, un étudiant de comptabilité : son code=1622
- o 4e arrivée, un étudiant de gestion : son code=1
- o 5e arrivée, un étudiant de comptabilité : son code=1623
- o 6e arrivée, un étudiant de gestion : son code=2
- o 7e arrivée, un étudiant de finance : son code=811 o

3. Codification articulée

Nous proposons les zones suivantes:

1 pos Alph pour le niveau (I : Ingenieur, T : technicien, S : TS)

1 pos Alph pour la spécialité (G : gestion, C : comptabilité, F : finance, M : marketing)

1 pos Num pour la section

2 pos Num pour le séquentiel

Le code proposé est donc :

Niv Spe Sec Seq

2- Les avantages et les inconvénients de chaque proposition

Tableau 2.2 Les avantages et les inconvénients de chaque proposition

1. Codification séquentielle

avantages	inconvénients
<ul style="list-style-type: none"> ○ Facile à réaliser, non ambiguë ○ Un compteur (dernier +1) ○ Permet l'extension ○ permet la connaissance de nombre d'étudiants 	<ul style="list-style-type: none"> ○ Non significative ○ Ne permet pas l'insertion ○ Pas trop utiliser ○

2. Codification séquentielle par tranche

avantages	inconvénients
<ul style="list-style-type: none"> ○ Facile à réaliser non ambiguë ○ Permet l'extension ○ Permet le groupement des étudiants selon 1 critère 	<ul style="list-style-type: none"> ○ peu significative ○ l'insertion est difficile ○ Pas trop utiliser ○ Ne permet la connaissance de nombre exacte de la population (les numéros vides laissés entre les tranches)

3. Codification articulée

avantages	inconvénients
<ul style="list-style-type: none"> ○ Facile à réaliser non ambiguë ○ Permet l'extension et l'insertion ○ Permet le groupement des étudiants selon plusieurs critères ○ Significative ○ Trop utiliser 	<ul style="list-style-type: none"> ○ Long en général ○ Le choix des zones à mettre dans le code n'est pas toujours évident ○ Instable : si un étudiant change la section le code doit changer

Exercice 2

On propose les codifications suivantes :

Numéro employé : 2 positions (n°service) + 3 positions (n°séquentiel).

Ouvrage d'une bibliothèque : 2 positions (code domaine) + 3 pos (sous domaine) + 3 pos (n°séquentiel)

1- Quel est le type de chaque code ?

2- Critiquer ces propositions, puis suggérer une solution appropriée pour chaque code ?

Solutions :

1. **Numéro employé** : 2 positions (n°service) + 3 positions (n°séquentiel).

Code articulé

Critique: Instable (si l'employé change de service le code sera erroné)

Solution : Remplacer service par année de recrutement sur 2 pos :

Le nouveau code : 2 positions (année recrutement) + 3 positions (n°séquentiel).

2. **Ouvrage d'une bibliothèque** : 2 positions (code domaine) + 3 pos (sous domaine) + 3 pos (n°séquentiel)

Code articulé par niveau

Critique: long (8 position pour un ouvrage)

La décomposition en domaine suivi de sous domaine est exagérée

Solution : le domaine sur 3 pos suffira largement :

Le nouveau code : 3 positions (domaine) + 3 positions (n°séquentiel).