Chapter II: Information representation techniques

1. Concept of information

Information is a concept of the discipline of information and communication sciences(SIC). In the etymological sense, "information" is what gives shape to the mind. It comes from the Latin verb "informare", which means "to give form to" or "to form an idea of" [16].

In computing and telecommunications, information is an element of knowledge (voice, data, image) capable of being preserved, processed or transmitted using a medium and a standardized coding method [17].

2. Forms of information

Information can take several forms, namely:

- Textual
- Digital
- Alphanumeric
- Visual
- Sound
- Audiovisual

3. Coding of information

Codification is an operation which consists of replacing natural information with abbreviated conventional representations (codes) making it possible to better carry out automatic processing on this codified information and to designate it in a precise, clear and unique way.

3.1. Objectives of coding
Good coding makes it possible to:

- Identify the codified objects of our system without ambiguity,
- Save space and time.
- Control information easily.

- Represent some properties of the objects.

3.2. Criteria for good coding

Good coding must fulfill the following criteria [18]:

- *Uniqueness*: the coding must make it possible to identify the codified objects without ambiguity (each object must have one and only one code and each code must be assigned to one and only one object).
- *Sustainability*: a code must normally be used for a very long time,
- *Flexibility*: possibility of extension and insertion
- Extension: the set of codified objects can increase without calling into question the chosen codification.

-Insertion: new objects can be inserted between existing objects without calling into question the chosen coding

- *Concision*: coding must avoid the manipulation of information in too many languages; to be effective, a code must therefore contain as few characters as possible.
- *Stability*: the code must be as stable as possible, that is to say that we do not change the coding each time a new object to be coded arrives in the system.

3.3. Types of coding

There are several types of codification but we are limiting ourselves here to those of functional codification which consists of giving a symbolic form to the structure of the information.

3.3.1. Sequential coding

It consists of assigning sequential numbers to the objects to be codified in the same set.

Example: Room numbers in a hotel,

Student numbers in a list…etc.

 Benefits:

- Simplicity

- Non-ambiguity

- Extension allowed (sequential numbers are unlimited)

Disadvantages:

- Not significant (numbers without meaning)

- Impossibility of insertion

- No groupings possible

3.3.2. Coding by installments

It consists of dividing objects into slices (categories). Objects in the same slice generally have a common criterion and they are coded sequentially.

Example: Coding of cosmetic products in a store

1-100: shampoos

101-200: shower gels

201-300: soaps

……etc.

Benefits

- Unambiguous coding.

- Simple coding to implement.

- Makes the search process a little easier.

- Allows extensions and insertions.

Disadvantages

- Not significant: Requires a correspondence table.

- Number of codes in a range difficult to predict and fix.

- The distribution of objects into categories is not always obvious.

- Insertion impossible if a slice is saturated.

3.3.3. Juxtaposed coding (Articulated)

The code is composed of several zones (fields) where each zone describes a particular aspect of the coded object.

Example: Barcode shave 13 digits:

• the first 2 or 3 correspond to the country of origin of the product;

• the next 4 or 5 are the member number of the company participating in the EAN system;

• the next 4 or 5 are the article number of the product;

• the last digit is a control key. Calculated from the first twelve digits, it allows you to check the validity of the barcode
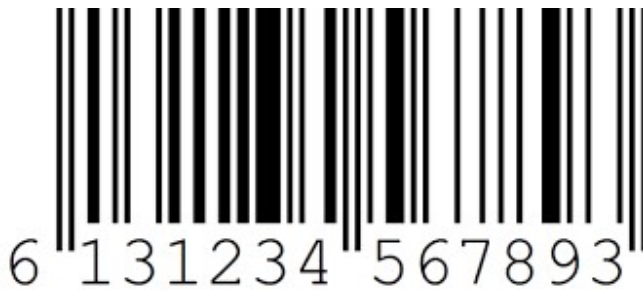
Benefits

- Unambiguous coding.

- Widely used and widely responded to around the world.

- Meaningful: we can easily interpret the areas of the code.

- Possibility of controls, statistics and grouping according to a given criterion.
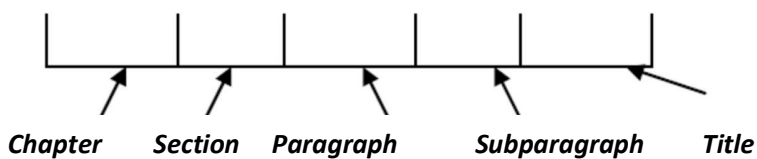
Disadvantages

- Long and heavy code

- Risk of saturation of an area

- Instability: the change of a criterion of the codified object causes the code to change.

3.3.4. Hierarchical coding (by level)

Is a coding used when there is subordination between the sections that compose it. This coding is considered a special case of juxtaposed coding.

Example: numbering sections in a book

Chapter   Section   Paragraph   Subparagraph   Title

Benefits

- Same advantages as juxtaposed coding

- Ease of searching for an element (hierarchical structure)

Disadvantages

- Same disadvantages as juxtaposed coding

3.3.5. Mnemonic coding (Significant or descriptive)

Consists of abbreviating the designation of an object using a set of characters. The coding must be evocative of the codified object.

Example: QT: to designate the quantity

 PC: postal code

Benefits

- Significant

- Easy to remember and control

Disadvantages

- Insufficient and used to codify the names of variables and not their contents.

- Risk of ambiguity

4. How to choose a coding

Coding is essential to properly prepare and control the information in our system. Therefore, a set of rules must be followed in order to complete this work successfully:

- Before committing to the proposal of a codification in an organization, it is necessary to first check whether a codification already exists. If the latter is good, it must be kept to avoid changes that can lead to instability of the information system. Otherwise:

- Identify the population to be coded

  Discuss this coding with future users in order to determine the properties desired by them.
- Study the population to be codified (individual type, number of individuals, their growth, maximum, etc.).
- Implement the coding and test it.