International Standard ISO 2788

Documentation — Guidelines for the establishment and development of monolingual thesauri

Documentation — Principes directeurs pour l'établissement et le développement de thesaurus monolingues

Second edition — 1986-11-15

ISO 2788-1986 (E)

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work.

Draft International Standards adopted by the technical committees are circulated to the member bodies for approval before their acceptance as International Standards by the ISO Council. They are approved in accordance with ISO procedures requiring at least 75 % approval by the member bodies voting.

International Standard ISO 2788 was prepared by Technical Committee ISO/TC 46, *Documentation*.

This second edition cancels and replaces the first edition (ISO 2788-1974), of which it constitutes a technical revision.

Users should note that all International Standards undergo revision from time to time and that any reference made herein to any other International Standard implies its latest edition, unless otherwise stated.

Documentation — Guidelines for the establishment and development of monolingual thesauri

0 Introduction

The effectiveness of a subject index as a means for identifying and retrieving documents depends upon a well-constructed indexing language. This applies to any system where the selection of indexing terms calls for human intellectual decisions, including those systems in which a computer is used to store and manipulate terms or to identify documents associated with terms or combinations of terms assigned by an indexer.

The compiler of a subject index faces three main tasks

    a) determining the subject matter of documents;

    b) selecting the terms which together summarize the subject;

   c) indicating relationships between the concepts represented by these terms.[1]

The first of these tasks is described separately in ISO 5963. The second and third tasks concern not only the indexer but also the user of the index. This International Standard deals with some aspects of term selection, since it contains recommended procedures for vocabulary control, but it is particularly concerned with means for establishing and displaying certain kinds of relationships between indexing terms.

Two kinds of inter-term relationships can be distinguished:

    a) syntactical or *a posteriori* relationships between the terms which together summarize the subject of a document. For example, an indexer dealing with a work on "Computers in banks in Amsterdam" may assign three terms, "Banks", "Computers" and "Amsterdam", to the document. In a post-coordinate system the relationship between these terms is not explicitly indicated, and the document would be retrieved if any or all of these terms were used as retrieval keys. In a pre-coordinated index the relationships between the terms may be conveyed in various ways, for example by symbols which express specific relationships, the positions of terms within entries, their typography and/or their accompanying punctuation. The terms in this example are not normally associated according to common frames of reference, and their interrelationships can therefore be regarded as document-dependent;

    b) those *a priori* or thesaural relationships between terms assigned to documents and other terms which, because they form part of common and shared frames of reference, are present by implication. In the example above, "Banks" would imply a broader term such as "Financial institutions"; "Computers" is mentally associated with "Data processing"; and "Amsterdam" implies the wider location "Netherlands". Any of these mentally-associated terms might serve as a user's approach to the subject index. These relationships are document-independent, since they are generally recognized and could be established through reference to standard works, such as dictionaries and encyclopaedias.

The distinction between these two kinds of inter-term relationships can be displayed as follows:
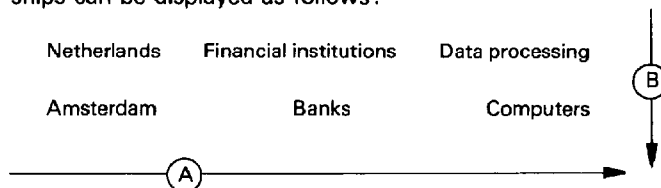
  (A) = *a posteriori* relationships between indexing terms assigned to a document

  (B) = *a priori* relationships handled by the thesaurus

This International Standard is especially concerned with those *a priori* relationships which can be displayed in a thesaurus, where they then, in effect, add a second dimension to an indexing language, as shown above.

1 Scope and field of application

1.1 The recommendations set out in this International Standard are intended to ensure consistent practice within a single indexing agency, or between different agencies (for example members of a network). They should not be regarded, however, as mandatory instructions. Optional procedures are described in many cases, for example for the display of inter-term relationships, without indicating one of these approaches as the preferred technique. The choice of procedure will vary from one indexing agency to another, depending on management decisions that fall outside the scope of this International

---

1) For practical purposes, "term" and "concept" are sometimes used interchangeably.