

INTERNATIONAL
STANDARD

ISO
1087-2

NORME
INTERNATIONALE

First edition
Première édition
2000-03-01

Terminology work — Vocabulary —

Part 2:
Computer applications

Travaux terminologiques — Vocabulaire —

Partie 2:
Applications logicielles



Reference number
Numéro de référence
ISO 1087-2:2000(E/F)

© ISO 2000

PDF disclaimer

This PDF file may contain embedded typefaces. In accordance with Adobe's licensing policy, this file may be printed or viewed but shall not be edited unless the typefaces which are embedded are licensed to and installed on the computer performing the editing. In downloading this file, parties accept therein the responsibility of not infringing Adobe's licensing policy. The ISO Central Secretariat accepts no liability in this area.

Adobe is a trademark of Adobe Systems Incorporated.

Details of the software products used to create this PDF file can be found in the General Info relative to the file; the PDF-creation parameters were optimized for printing. Every care has been taken to ensure that the file is suitable for use by ISO member bodies. In the unlikely event that a problem relating to it is found, please inform the Central Secretariat at the address given below.

PDF – Exonération de responsabilité

Le présent fichier PDF peut contenir des polices de caractères intégrées. Conformément aux conditions de licence d'Adobe, ce fichier peut être imprimé ou visualisé, mais ne doit pas être modifié à moins que l'ordinateur employé à cet effet ne bénéficie d'une licence autorisant l'utilisation de ces polices et que celles-ci y soient installées. Lors du téléchargement de ce fichier, les parties concernées acceptent de fait la responsabilité de ne pas enfreindre les conditions de licence d'Adobe. Le Secrétariat central de l'ISO décline toute responsabilité en la matière.

Adobe est une marque déposée d'Adobe Systems Incorporated.

Les détails relatifs aux produits logiciels utilisés pour la création du présent fichier PDF sont disponibles dans la rubrique General Info du fichier; les paramètres de création PDF ont été optimisés pour l'impression. Toutes les mesures ont été prises pour garantir l'exploitation de ce fichier par les comités membres de l'ISO. Dans le cas peu probable où surviendrait un problème d'utilisation, veuillez en informer le Secrétariat central à l'adresse donnée ci-dessous.

© ISO 2000

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm, without permission in writing from either ISO at the address below or ISO's member body in the country of the requester. / Droits de reproduction réservés. Sauf prescription différente, aucune partie de cette publication ne peut être reproduite ni utilisée sous quelque forme que ce soit et par aucun procédé, électronique ou mécanique, y compris la photocopie et les microfilms, sans l'accord écrit de l'ISO à l'adresse ci-après ou du comité membre de l'ISO dans le pays du demandeur.

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 734 10 79
E-mail copyright@iso.ch
Web www.iso.ch

Printed in Switzerland/Imprimé en Suisse

Contents

Page

Foreword.....	v
Introduction	vii
1 Scope	1
2 General concepts.....	1
3 Data organization.....	6
4 Filtering of terminological data	7
5 Characters	9
6 Storing data	11
7 Information retrieval	13
8 Data handling	15
9 Data validation	16
10 Operations	18
11 Related fields of application	20
Bibliography	21
Alphabetical index	23

Sommaire

Page

Avant-propos.....	vi
Introduction	viii
1 Domaine d'application	1
2 Concepts généraux	1
3 Organisation des données	6
4 Filtrage des données terminologiques	7
5 Caractères	9
6 Stockage des données	11
7 Recherche documentaire	13
8 Manipulation des données	15
9 Validation des données	16
10 Opérations	18
11 Domaines d'application associés	20
Bibliographie	21
Index alphabétique	25

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. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 3.

Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

Attention is drawn to the possibility that some of the elements of this part of ISO 1087 may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights.

International Standard ISO 1087-2 was prepared by Technical Committee ISO/TC 37, *Terminology (principles and coordination)*, Subcommittee SC 3, *Computer applications*.

This part of ISO 1087 cancels and replaces ISO 1087:1990, clause 7. ISO 1087-1, which is being prepared, will constitute a technical revision of the other clauses of ISO 1087:1990.

ISO 1087 consists of the following parts, under the general title *Terminology work — Vocabulary*:

- *Part 1: Theory and application*
- *Part 2: Computer applications*

Avant-propos

L'ISO (Organisation internationale de normalisation) est une fédération mondiale d'organismes nationaux de normalisation (comités membres de l'ISO). L'élaboration des Normes internationales est en général confiée aux comités techniques de l'ISO. Chaque comité membre intéressé par une étude a le droit de faire partie du comité technique créé à cet effet. Les organisations internationales, gouvernementales et non gouvernementales, en liaison avec l'ISO participent également aux travaux. L'ISO collabore étroitement avec la Commission électrotechnique internationale (CEI) en ce qui concerne la normalisation électrotechnique.

Les Normes internationales sont rédigées conformément aux règles données dans les Directives ISO/CEI, Partie 3.

Les projets de Normes internationales adoptés par les comités techniques sont soumis aux comités membres pour vote. Leur publication comme Normes internationales requiert l'approbation de 75 % au moins des comités membres votants.

L'attention est appelée sur le fait que certains des éléments de la présente partie de l'ISO 1087 peuvent faire l'objet de droits de propriété intellectuelle ou de droits analogues. L'ISO ne saurait être tenue pour responsable de ne pas avoir identifié de tels droits de propriété et averti de leur existence.

La Norme internationale ISO 1087-2 a été élaborée par le comité technique ISO/TC 37, *Terminologie (principes et coordination)*, sous-comité SC 3, *Terminotique*.

La présente partie de l'ISO 1087 annule et remplace l'ISO 1087:1990, article 7. L'ISO 1087-1, actuellement en cours d'élaboration, constituera une révision technique des autres articles de l'ISO 1087:1990.

L'ISO 1087 comprend les parties suivantes, présentées sous le titre général *Travaux terminologiques — Vocabulaire*:

— *Partie 1: Théorie et application*

— *Partie 2: Applications logicielles*

Introduction

The terms of this international terminology standard are given in a clustered ordering subsumed under a few general headings. Usually, terms necessary for the definition of other terms are preceding the entries of the latter. For convenience of use, however, an alphabetical index is to be found at the end of this part of ISO 1087.

The layout is designed according to ISO 10241. Thus, the elements of an entry appear in the following order:

entry number

preferred term (bold face), word class

short form (bold face)

admitted term

deprecated term: labelled "(deprecated)"

subject field (between angle brackets)

definition

examples

notes

} reference to another entry in bold face followed by entry number in brackets, when it is first mentioned

Except for entry number, preferred term and definition, elements appear only where appropriate.

Entries are nouns except where stated otherwise.

Introduction

Les termes de la présente Norme internationale terminologique sont classés par thème dans plusieurs sous-groupes généraux. En principe, les termes qui entrent dans la définition d'autres termes précèdent les entrées de ces derniers. Néanmoins, pour des raisons pratiques, un index alphabétique est fourni à la fin de la présente partie de l'ISO 1087.

La structure est établie conformément à l'ISO 10241. Les éléments de chaque entrée sont affichés dans l'ordre suivant:

- le numéro d'entrée
 - le terme privilégié (en gras), la classe de mots
 - l'abréviation (en gras)
 - le(s) terme(s) admis
 - le terme rejeté, indiqué par «(rejeté)»
 - le domaine <affiché entre ces signes>
 - une définition
 - des exemples
 - des notes
- } une référence à une autre entrée, en gras et suivie par le numéro d'entrée entre parenthèses, quand elle est mentionnée la première fois

À l'exception du numéro d'entrée, toujours affiché, le terme privilégié et les différents éléments de la définition ne sont indiqués que lorsque cela est nécessaire.

Sauf indication contraire, les entrées sont des noms.

Terminology work — Vocabulary —

Part 2: Computer applications

1 Scope

This International Standard defines terms for language and information processing for applications in terminology work and terminography.

2 General concepts

2.1 information

⟨information processing⟩ knowledge concerning such things as facts, concepts, objects, events, ideas

NOTE 1 Adapted from ISO 2382-1:1993.

NOTE 2 Information is defined differently in other application areas.

2.2 data

representation of **information** (2.1) in a formalized manner suitable for communication, interpretation, storing or processing

NOTE Adapted from ISO/IEC 2382-1:1993.

2.3 data processing DP

systematic performance of operations upon **data** (2.2)

EXAMPLE Arithmetic or logic operations upon **data** (2.2), **merging** (8.5) or **sorting** (8.4) of data, assembling or compiling of programs, or operations on **text** (2.6), such as **text editing** (11.3), sorting, merging, storing, retrieving, **displaying** (8.6), or printing.

Travaux terminologiques — Vocabulaire —

Partie 2: Applications logicielles

1 Domaine d'application

La présente Norme internationale définit les termes utilisés pour le traitement du langage et de l'information dans les applications liées à la terminologie et à la terminographie.

2 Concepts généraux

2.1 information

⟨traitement de l'information⟩ connaissance d'éléments tels que des faits, des concepts, des objets, des événements et des idées

NOTE 1 Adapté de l'ISO 2382-1:1993.

NOTE 2 Le terme information est défini différemment dans d'autres domaines d'application.

2.2 donnée

informations (2.1) représentées sous une forme conventionnelle convenant à la communication, à l'interprétation, au stockage ou au traitement

NOTE Adapté de l'ISO/CEI 2382-1:1993.

2.3 traitement des données

déroulement systématique d'opérations sur des **données** (2.2)

EXEMPLE Opérations arithmétiques ou logiques sur des **données** (2.2), **fusion** (8.5) ou **tri** (8.4) de données, assemblage ou compilation de programmes, ou opérations sur du **texte** (2.6) comme l'**édition de texte** (11.3), le tri, la fusion, le stockage, la recherche, l'**affichage** (8.6) ou l'impression.

NOTE The term should not be used as a synonym for **information processing** (2.4).

[ISO 2382-1:1993]

2.4 information processing

systematic performance of operations upon **information** (2.1) that includes **data processing** (2.3)

NOTE 1 The term must not be used as a synonym for **data processing** (2.3).

NOTE 2 Adapted from ISO 2382-1:1993.

2.5 language processing information processing (2.4) on language

EXAMPLE Proofreading: proofreading can be carried out intellectually and by using a spell-checker. Both operations are complementary.

2.6 text data (2.2) in the form of **characters** (5.1), symbols, words, phrases, paragraphs, sentences, tables, or other character arrangements, intended to convey a meaning, and whose interpretation is essentially based upon the knowledge of some natural or artificial language

NOTE Adapted from ISO 2382-1:1993.

2.7 text corpus corpus (language engineering) systematic collection of machine-readable texts (see also 6.4) or parts of text prepared, coded and stored according to predefined rules

NOTE 1 A text corpus may be limited according to aspects of subject fields, size or time, e.g. mathematical texts, certain periodicals from 1986 onwards. It is used as source material for further linguistic analysis or terminology work.

NOTE 2 See also ISO 1087-1.

NOTE Il convient que ce terme ne soit pas utilisé comme un synonyme du terme **traitement de l'information** (2.4).

[ISO 2382-1:1993]

2.4 traitement de l'information déroulement systématique d'opérations sur des **informations** (2.1), incluant également le **traitement de données** (2.3)

NOTE 1 Il convient que ce terme ne soit pas utilisé comme un synonyme du terme **traitement des données** (2.3).

NOTE 2 Adapté de l'ISO 2382-1:1993.

2.5 traitement du langage traitement de l'information (2.4) sur le langage

EXEMPLE La correction d'un texte peut être réalisée intellectuellement ou en utilisant un correcteur orthographique. Ces deux opérations sont complémentaires.

2.6 texte données (2.2) sous la forme de **caractères** (5.1), de symboles, de mots, de phrases, de paragraphes, de tableaux, ou de tout autre disposition de caractères, dont l'objectif est de véhiculer du sens, et dont l'interprétation repose essentiellement sur la connaissance d'un langage naturel ou artificiel

NOTE Adapté de l'ISO 2382-1:1993

2.7 corpus de textes corpus (ingénierie linguistique) regroupement systématique de textes exploitables par une machine (voir également: 6.4) ou de fragments de texte préparés, codés et stockés selon des règles prédéfinies

NOTE 1 Un corpus de textes peut ne regrouper que les textes correspondant à certaines des caractéristiques du domaine étudié, telles que la taille ou la période, par exemple textes mathématiques, périodiques spécifiques parus depuis 1986. Un corpus est utilisé comme le matériau source à partir duquel peuvent être réalisés des analyses linguistiques ou des travaux de terminologie.

NOTE 2 Voir aussi l'ISO 1087-1.

2.8 type

⟨language engineering⟩ linguistic unit in a **text** (2.6) representing a defined class

NOTE Such linguistic units are usually **character strings** (5.13) uninterrupted by **delimiters** (2.10).

2.9 token

⟨natural-language processing⟩ occurrence of a **type** (2.8)

EXAMPLE If the class is defined as all **word forms** (2.11) of “good”, then all the occurrences of the word forms “good”, “better” and “best” are tokens.

2.10 delimiter

separator
one or more **characters** (5.1) used to indicate the beginning or the end of a **character string** (5.13)

[ISO 2382-4:1987]

NOTE Blanks or punctuation marks often function as delimiters.

2.11 word form

morphosyntactical variant of a given word

EXAMPLE Indicate: indicates, visitor: visitors, visitor's.

NOTE In inflecting languages, word forms are often equivalent to inflected forms e.g. “go”, “goes”; but also: “go”, “went”, “gone”.

2.12 paradigm

⟨natural-language processing⟩ class of **word forms** (2.11) belonging to a given word or **multi-word term** (2.14)

EXAMPLE Sell, sells, sold, selling

2.13 base form

reference form
word form (2.11) chosen according to lexicographical conventions representing the forms of a **paradigm** (2.12)

2.8 type

⟨ingénierie linguistique⟩ unité linguistique dans un **texte** (2.6) représentant une classe définie

NOTE Ces unités linguistiques sont généralement des **chaînes de caractères** (5.13) non interrompues par des **délimiteurs** (2.10).

2.9 occurrence

⟨traitement en langage naturel⟩ nombre de fois où apparaît un **type** (2.8)

EXEMPLE Si la classe est définie comme l'ensemble des **formes lexicales** (2.11) du mot «bon», alors toutes les occasions où apparaissent les formes lexicales «bon» et «meilleur» sont des occurrences.

2.10 délimiteur

séparateur
un ou plusieurs **caractères** (5.1) utilisés pour marquer le début ou la fin d'une **chaîne de caractères** (5.13)

[ISO 2382-4:1987]

NOTE Les blancs ou les signes de ponctuation sont souvent utilisés comme délimiteurs.

2.11 forme lexicale

variante morphosyntaxique d'un mot donné

EXEMPLE Indiquer: indiquons, indiquant; instituteur: instituteurs, institutrices

NOTE Dans les langages conjugués, les formes lexicales correspondent souvent aux formes conjuguées, par exemple «être», «est», mais aussi «être», «fut», «été».

2.12 paradigme

⟨traitement en langage naturel⟩ classe de **formes lexicales** (2.11) se rapportant à un mot simple ou à un **terme composé** (2.14)

EXEMPLE Dormir, dort, dormi, dormant

2.13 forme de base

forme de référence
forme lexicale (2.11) choisie selon des conventions lexicographiques et représentant les formes d'un **paradigme** (2.12)

NOTE 1 Adapted from ISO 1087:1990.

EXAMPLE bind: bind, bound, binds, binding

NOTE 2 The term “base form” is also applied to **multi-word terms** (2.14).

2.14 multi-word term

term consisting of more than one **character string** (5.13)

2.15 compressed form

reduced form (deprecated)
character string (5.13) having undergone **compression** (7.4)

EXAMPLE The string “input/output-algorithm” becomes the compressed form: “inputoutputalgorithm”.

NOTE The term “reduced form” may be misleading in this context.

2.16 deinflection

deletion of inflectional elements in **word forms** (2.11)

2.17 deinflected word form

word part (2.18) remaining after **deinflection** (2.16)

2.18 word part

word segment
character string (5.13) taken from a **word form** (2.11) for a specific application

2.19 lemmatization

process of deriving the **base form** (2.13)

EXAMPLE “go” becomes the **base form** (2.13) of “goes”, by **deinflection** (2.16); “go” also serves as the **base form** (2.13) of “went” by virtue of irregular verb conventions that do not conform to standard inflectional practice.

NOTE The result may also be called lemmatization.

NOTE 1 Adapté de l'ISO 1087:1990.

EXEMPLE Prendre: prends, pris, prenant

NOTE 2 Le terme «forme de base» s'applique également aux **termes composés** (2.14).

2.14 terme composé

terme comprenant une ou plusieurs **chaînes de caractères** (5.13)

2.15 forme comprimée

forme réduite (rejeté)
chaîne de caractères (5.13) ayant subi une **compression** (7.4)

EXEMPLE Une fois comprimée, la chaîne «algorithme-entrée/sortie» devient: «algorithmeentréesortie».

NOTE Le terme «forme réduite» peut induire en erreur dans ce contexte.

2.16 suppression des désinences

suppression d'éléments flexionnels dans des **formes lexicales** (2.11)

2.17 forme lexicale sans désinence

partie de mot (2.18) restant à la suite de la suppression des **désinences** (2.16)

2.18 partie de mot

segment de mot
chaîne de caractères (5.13) tirée d'une **forme lexicale** (2.11) pour une application donnée

2.19 lemmatisation

processus consistant à dériver la **forme de base** (2.13)

EXEMPLE «naître» devient la **forme de base** (2.13) de «naît» par suppression des **désinences** (2.16), mais «naître» est également la forme de base de «naquit» en vertu des règles régissant les verbes irréguliers, qui diffèrent des pratiques courantes de conjugaison.

NOTE Le résultat peut également être appelé lemmatisation.

2.20 parsing

(language processing) operation segregating a given textual structure into its grammatical, categorial and/or lexical elements according to a given algorithm

EXAMPLE In the sentence “all unsaturated fatty acids are not degradable by biological methods”, “unsaturated fatty acids” and “biological methods” are identified as **multi-word terms** (2.14) and can be **extracted** (7.9).

NOTE Parsing does not necessarily provide a complete analysis of the sentence structure.

2.21 terminological data collection

collection of **data** (2.2) containing **information** (2.1) on concepts of specific subject fields

2.22 terminological entry

part of a **terminological data collection** (2.21) that contains the terminological **data** (2.2) related to one concept

2.23 homograph

each of two or more **word forms** (2.11) or words with identical spelling but representing different concepts (semantic homography) or syntactic functions (syntactic homography)

EXAMPLE lead (Pb) versus lead (guidance); bark (sound) versus bark (of a tree)

NOTE **Word forms** (2.11) having different spelling but becoming homographic after **compression** (7.4) are not called “homographs”.

2.24 disambiguation

differentiation between **homographs** (2.23) by assigning each of them to the relevant concept or, where appropriate, to the relevant syntactic function

2.20 analyse syntaxique

(traitement du langage) opération consistant à isoler dans une structure de texte donnée les éléments grammaticaux, catégoriels et/ou lexicaux en fonction d'un algorithme défini

EXEMPLE Dans la phrase «tous les acides gras insaturés ne peuvent pas être dégradés par des méthodes biologiques», «acides gras insaturés» et «méthodes biologiques» sont identifiés comme des **termes composés** (2.14) et peuvent être **extraits** (7.9).

NOTE L'analyse syntaxique ne constitue pas nécessairement une analyse complète de la structure de la phrase.

2.21 collecte de données terminologiques

regroupement de **données** (2.2) contenant des **informations** (2.1) sur des concepts propres à un domaine donné

2.22 entrée terminologique

élément d'une **collecte de données terminologiques** (2.21) contenant les **données** (2.2) terminologiques associées à un concept

2.23 homographe

chacune des **formes lexicales** (2.11) ou chacun des mots de même orthographe mais représentant des concepts différents (homographie sémantique) ou des fonctions syntaxiques différentes (homographie syntaxique)

EXEMPLES Le mousse (marin)/la mousse (végétal); louer (un appartement)/louer (le ciel).

NOTE Les **formes lexicales** (2.11) avec une orthographe différente mais devenant homographiques après **compression** (7.4) ne sont pas appelés «homographes».

2.24 désambiguïsation

opération consistant à différencier des **homographes** (2.23) en affectant chacun d'entre eux au concept correspondant ou, le cas échéant, à la fonction syntaxique appropriée

3 Data organization

3.1 sort value

sorting value
position of an element of a **character set** (5.2) according to a predefined order

EXAMPLE In French, the letter A has a lower sort value than the letter B. The question of whether or not to assign lower-case letters the same sort value as their upper-case counterpart will depend on the application. Letters with diacritics are sometimes treated like the corresponding basic letter and sometimes as a different letter.

3.2 sort key

sorting key
character string (5.13) designed to meet certain requirements for **sorting** (8.4) and **merging** (8.5) operations

EXAMPLE When bibliographical data are sorted to produce an author's catalogue, the author's surname is the primary sort key, with the given name as a secondary key. If there are two titles by the same author, the year of publication or the title is used as an additional sort key.

3.3 alphabetical ordering

arrangement of **character strings** (5.13) in such a way that the position of each **string** (5.12) in the list is uniquely determined by the appropriate **sort value** (3.1) starting from the beginning of the **string** (5.12)

NOTE 1 Special **sort values** (3.1) may have to be assigned for letters with diacritics and ligatures and for other symbols, e.g. numerals, superscripts, subscripts, etc.

NOTE 2 The rules of alphabetical ordering may differ from language to language.

3.4 reverse alphabetical ordering

arrangement of **character strings** (5.13) in such a way that the position of each **string** (5.12) in the list is uniquely determined by the appropriate **sort value** (3.1) starting from the end of the string

3.5 frequency order

order of **types** (2.8) either ascending or descending according to the frequency of occurrence of the respective **tokens** (2.9) in a specified **text** (2.6) or a **text corpus** (2.7)

NOTE The usual types of lists are **word form** (2.11) frequency lists and lemmatized frequency lists.

3 Organisation des données

3.1 valeur de tri

position de l'élément d'un **jeu de caractères** (5.2) en fonction d'un ordre prédéfini

EXEMPLE En français, la valeur de tri de la lettre A est inférieure à celle de la lettre B. La décision d'affecter aux lettres majuscules une valeur de tri identique à celle attribuée aux lettres minuscules varie en fonction du but recherché. Les lettres avec des diacritiques sont parfois traitées comme les lettres de base correspondantes, parfois considérées comme des lettres différentes.

3.2 clé de tri

critère de tri
chaîne de caractères (5.13) destinée à répondre à certaines exigences liées aux opérations de **tri** (8.4) et de **fusion** (8.5)

EXEMPLE Lorsque des données bibliographiques sont triées afin de générer un catalogue d'auteurs, le nom de famille des auteurs est la clé de tri principale, et leur prénom constitue la clé de tri secondaire. S'il existe deux titres du même auteur, l'année de publication ou le titre sont utilisés comme clés de tri complémentaires.

3.3 ordre alphabétique

arrangement de **chaînes de caractères** (5.13) de sorte que la position de chaque **chaîne** (5.12) dans la liste soit uniquement déterminée par la **valeur de tri** (3.1) appropriée, en commençant à partir du début de la **chaîne** (5.12)

NOTE 1 Des **valeurs de tri** (3.1) spéciales peuvent être affectées à des lettres comportant des diacritiques et des ligatures, ou à d'autres symboles, par exemple des numéraux, des indices et des exposants, etc.

NOTE 2 Les règles régissant l'ordre alphabétique peuvent varier d'un langage à un autre.

3.4 ordre alphabétique inversé

arrangement de **chaînes de caractères** (5.13) de sorte que la position de chaque **chaîne** (5.12) dans la liste soit uniquement déterminée par la **valeur de tri** (3.1) appropriée, en commençant à partir de la fin de la chaîne

3.5 ordre de fréquence

ordre croissant ou décroissant des **types** (2.8) selon la fréquence de leurs **occurrences** (2.9) respectives dans un **texte** (2.6) ou un **corpus de textes** (2.7) donné

NOTE Les listes les plus courantes sont les listes de fréquence de **formes lexicales** (2.11) et les listes de fréquence de termes lemmatisés.

3.6 permutation

arrangement of lexical units so that each of their constituent parts appears at the appropriate position

EXAMPLE The lexical unit “millions of instructions per second” [MIPS] would appear as: “instructions, millions of per second and “second, millions of instructions per”, to ensure that the desired constituent parts appear at the appropriate, in this case alphabetical, positions.

NOTE A given term may have only one lexically relevant constituent part.

3.7 concatenation

operation of joining two or more **character strings** (5.13) in a specified order, forming one **string** (5.12) the length of which is equal to the sum of the length of the **character strings** (5.13)

4 Filtering of terminological data

4.1 exclusion list

stop word list

list of user-selected **character strings** (5.13) to be ignored in further **data processing** (2.3) operations

NOTE 1 For terminological purposes, it may be useful to produce lists where function words (pronouns, articles, etc.) are ignored.

NOTE 2 Sometimes the character strings in the exclusion list are retained, e.g. in a **concordance** (4.4), but with no frequency given.

4.2 inclusion list

plus word list

list of user-selected **character strings** (5.13) to be retained or considered for further **data processing** (2.3) operations, while the rest is ignored

EXAMPLE All sentences containing “bank” or “credit institution” are to be extracted; all entries starting with 'M' are to be retrieved in an index of names; all words starting with “anti-” are to be selected.

NOTE When appropriate, the inclusion list can also contain **word parts** (2.18) or other **character strings** (5.13).

3.6 permutation

arrangement d'unités lexicales de sorte que chacun de leurs éléments constitutifs apparaisse à une position donnée

EXEMPLE L'unité lexicale «millions d'instructions par seconde» [MIPS] devrait apparaître comme suit: «instructions, millions par seconde» et «seconde, millions d'instructions par», afin de garantir que les éléments constitutifs souhaités apparaissent aux positions appropriées, alphabétiques dans ce cas.

NOTE Un terme donné peut n'avoir qu'un seul élément constitutif d'un point de vue lexical.

3.7 concaténation

opération consistant à joindre deux ou plusieurs **chaînes de caractères** (5.13) dans un ordre défini, pour former une **chaîne** (5.12) dont la longueur est égale à la somme des longueurs des **chaînes de caractères** (5.13)

4 Filtrage des données terminologiques

4.1 liste d'exclusion

liste de **chaînes de caractères** (5.13) sélectionnés par l'utilisateur afin d'être ignorées au cours des opérations ultérieures de **traitement des données** (2.3)

NOTE 1 Dans une optique terminologique, il peut s'avérer utile de générer des listes dans lesquelles certains mots de fonction (pronoms, articles, etc.) seront ignorés.

NOTE 2 Les chaînes de caractères de la liste d'exclusion sont parfois retenues, par exemple lors d'une **concordance** (4.4), mais dans ce cas, aucune fréquence n'est indiquée.

4.2 liste d'inclusion

liste de **chaînes de caractères** (5.13) sélectionnés par l'utilisateur afin d'être retenues ou consultées pour des opérations ultérieures de **traitement des données** (2.3), le reste étant ignoré

EXEMPLES Toutes les phrases contenant «banque» ou «établissement de crédit» doivent être extraites, toutes les entrées commençant par 'M' doivent être recherchées dans un index de noms, tous les mots commençant par «anti-» doivent être sélectionnés.

NOTE Une liste d'inclusion peut également contenir des **parties de mots** (2.18) ou d'autres **chaînes de caractères** (5.13).

**4.3
free-text search**

search (7.7) in a **text** (2.6) or a **text corpus** (2.7) enabling the user to retrieve any occurrence of any **type** (2.8)

**4.4
concordance**

⟨computer applications⟩ alphabetically ordered list of **word forms** (2.11) **extracted** (7.9) from a **text** (2.6) including the source reference for each item as well as a selectable portion of the text preceding and following the word form in question

NOTE A concordance usually has the form of a KWIC (keyword in context) concordance or of a sentence concordance.

**4.5
index**

⟨natural-language processing⟩ ordered list of **character strings** (5.13) together with keys or references to their location or locations in the a **text** (2.6) or a **text corpus** (2.7)

**4.6
term index**

index (4.5) that refers to **terminological entries** (2.22) in a vocabulary or in a **terminological database** (6.6)

NOTE The terms in a term index can be ordered alphabetically or systematically.

**4.7
alphabetical index**

index (4.5) produced by **alphabetical ordering** (3.3)

**4.8
systematic index**

index (4.5) in which all the terms are ordered to reflect the systematic relations between the concepts of a specific subject field

**4.9
permuted index**

alphabetical index (4.7) produced according to the principle of **permutation** (3.6)

**4.10
reverse index**

term index (4.6) produced by **reverse alphabetical ordering** (3.4)

**4.3
système de recherche en forme libre**

recherche (7.7) dans un **texte** (2.6) ou dans un **corpus de textes** (2.7) permettant à l'utilisateur d'extraire toute occurrence d'un **type** (2.8) quelconque

**4.4
concordance**

⟨applications logicielles⟩ liste de **formes lexicales** (2.11) classées par ordre alphabétique et **extraites** (7.9) d'un **texte** (2.6) comprenant la référence source de chaque élément ainsi qu'une portion sélectionnable du texte qui précède et qui suit la forme lexicale concernée

NOTE Une concordance se présente généralement sous la forme d'une concordance KWIC (mot-clé dans le contexte) ou d'une concordance de phrase.

**4.5
index**

⟨traitement en langage naturel⟩ liste ordonnée de **chaînes de caractères** (5.13) avec des clés ou des références à leur(s) position(s) respective(s) dans un **texte** (2.6) ou un **corpus de textes** (2.7)

**4.6
index de termes**

index (4.5) se rapportant aux **entrées terminologiques** (2.22) d'une liste de vocabulaire ou d'une **base de données terminologiques** (6.6)

NOTE Dans un index de termes, les termes peuvent être triés par ordre alphabétique ou de façon systématique.

**4.7
index alphabétique**

index (4.5) établi par **ordre alphabétique** (3.3)

**4.8
index systématique**

index (4.5) dans lequel tous les termes sont organisés de façon à refléter les relations systémiques entre les différents concepts d'un domaine spécifique

**4.9
index permuté**

index alphabétique (4.7) établi conformément au principe de la **permutation** (3.6)

**4.10
index inversé**

index de termes (4.6) établi par **ordre alphabétique inversé** (3.4)

4.11**inverted index**

index (4.5) of all **types** (2.8) derived from a **text** (2.6) together with all the references to their corresponding **tokens** (2.9)

4.12**word form index**

word index

index (4.5) of all the **word forms** (2.11) in a **text** (2.6), or a **corpus** (2.7) together with all the references to their location or locations in the text or the corpus

4.13**lemmatized word index**

word form index (4.12) in which the **word forms** (2.11) are represented by their respective **base forms** (2.13)

4.11**index à l'envers**

index (4.5) de tous les **types** (2.8) dérivés d'un **texte** (2.6) avec toutes les références correspondant à leurs **occurrences** (2.9) respectives

4.12**index de formes lexicales**

index de mots

index (4.5) de toutes les **formes lexicales** (2.11) dans un **texte** (2.6) ou dans un **corpus** (2.7) avec toutes les références correspondant à leur(s) position(s) dans ce texte ou ce corpus

4.13**index de mots lemmatisés**

index de formes lexicales (4.12) dans lequel ces **formes lexicales** (2.11) sont représentées sous leur **forme de base** (2.13) respective

5 Characters**5.1****character**

member of a set of elements that is used for the representation, organization or control of **data** (2.2)

[ISO 2382-4:1987]

NOTE Characters may be categorized as **graphical characters** (5.3) and **control characters** (5.9).

5.2**character set**

finite set of different **characters** (5.1) that is complete for a given purpose

[ISO 2382-4:1987]

EXAMPLE The international reference version of the character set of ISO 646; the 128 ASCII characters.

5.3**graphical character**

character (5.1), other than a **control character** (5.9), that has a visual representation and is produced by **writing** (8.2), printing or **displaying** (8.6)

NOTE 1 Adapted from ISO 2382-4: 1987.

NOTE 2 Graphical characters may be **alphabetical characters** (5.5), **digits** (5.7) or **special characters** (5.8).

5.4**alphanumerical character**

alphabetical character (5.5) or a **digit** (5.7)

5 Caractères**5.1****caractère**

élément d'un ensemble utilisé pour la représentation, l'organisation ou le contrôle des **données** (2.2)

[ISO 2382-4:1987]

NOTE Les caractères peuvent être classés en **caractères graphiques** (5.3) et **caractères de commande** (5.9).

5.2**jeu de caractères**

ensemble fini de différents **caractères** (5.1) considéré comme complet à des fins déterminées

[ISO 2382-4:1987]

EXEMPLES La version internationale de référence du jeu de caractères de l'ISO 646; les 128 caractères ASCII.

5.3**caractère graphique**

caractère (5.1) autre qu'un **caractère de commande** (5.9) ayant une représentation visuelle généralement **écrite** (8.2), imprimée ou **affichée** (8.6)

NOTE 1 Adapté de l'ISO 2382-4:1987.

NOTE 2 Les caractères graphiques peuvent être des **caractères alphabétiques** (5.5), des **chiffres** (5.7) ou des **caractères spéciaux** (5.8).

5.4**caractère alphanumérique**

caractère alphabétique (5.5) ou **chiffre** (5.7)

5.5
alphabetical character

graphical character (5.3) that is a letter as used in the written representation of a given language

NOTE This also includes letters with diacritical marks.

5.6
alphabet

linearly ordered set of **alphabetical characters** (5.5), the order of which has been agreed upon

NOTE 1 This definition also covers the alphabets of natural languages consisting of those **characters** (5.1) represented by letters including letters with diacritics.

NOTE 2 Adapted from ISO 2382-4:1987.

5.7
digit

numerical character
character (5.1) that represents a non-negative integer

[ISO 2382-4:1987]

5.8
special character
graphical character (5.3) that is not an **alphanumerical character** (5.4)

EXAMPLE Any punctuation mark, the ampersand (&), the percent sign (%), the plus sign (+).

NOTE Adapted from ISO 2382-4:1987.

5.9
control character
character (5.1) whose occurrence in a particular context specifies a control function

[ISO 2382-4:1987]

NOTE Control characters initiate, modify and break a **data processing** (2.3) operation.

5.10
space character
control character (5.9) that causes the print or display position to advance one position forward along the line without producing any **graphical character** (5.3)

NOTE 1 Space as **control character** (5.9) is described in ISO/IEC 646, ISO/IEC 4873 and ISO 6937-1.

NOTE 2 Adapted from ISO 2382-4:1987.

5.5
caractère alphabétique
caractère graphique (5.3) correspondant à une lettre utilisée dans la représentation écrite d'un langage donné

NOTE Cela comprend aussi les lettres avec signes diacritiques.

5.6
alphabet
jeu de **caractères alphabétiques** (5.5) présentés de façon linéaire, respectant un ordre convenu

NOTE 1 Cette définition s'applique aussi aux alphabets de langages naturels constitués de **caractères** (5.1) représentés par des lettres avec des signes diacritiques.

NOTE 2 Adapté de l'ISO 2382-4:1987.

5.7
chiffre
caractère (5.1) représentant un nombre entier positif ou nul

[ISO 2382-4:1987]

5.8
caractère spécial
caractère graphique (5.3) autre qu'un **caractère alphanumérique** (5.4)

EXEMPLES Signe de ponctuation, perluète (&), signe de pourcentage (%), signe plus (+).

NOTE Adapté de l'ISO 2382-4:1987.

5.9
caractère de commande
caractère (5.1) dont l'apparition dans un contexte déterminé provoque une fonction de commande

[ISO 2382-4:1987]

NOTE Les caractères de commande permettent de démarrer, de modifier et d'arrêter une opération de **traitement des données** (2.3).

5.10
caractère espace
caractère de commande (5.9) provoquant, sur la même ligne, l'avancement d'un cran de la position d'impression ou d'affichage sans produire de **caractère graphique** (5.3)

NOTE 1 L'espace en tant que **caractère de commande** (5.9) est décrit dans l'ISO/CEI 646, l'ISO/CEI 4873 et l'ISO 6937-1.

NOTE 2 Adapté de l'ISO 2382-4:1987.

5.11 blank

character (5.1) that represents an empty position in a **string** (5.12) of **graphical characters** (5.3)

NOTE Adapted from ISO 2382-4:1987.

5.12 string

sequence of elements of the same nature, considered as a whole

EXAMPLE **Characters** (5.1), bits or pulses.

NOTE Adapted from ISO 2382-4:1987.

5.13 character string

sequence of **characters** (5.1) uninterrupted by **blanks** (5.11)

NOTE A character string usually represents a word.

5.14 pattern

string (5.12) where some **characters** (5.1) or sequences of characters have been replaced by variables

6 Storing data

6.1 data medium

physical medium in or on which **data** (2.2) can be recorded and from which data can be retrieved

EXAMPLE Magnetic disk, CD ROM, microform.

NOTE 1 Adapted from ISO 2382-1:1993.

NOTE 2 In the field of librarianship and documentation data, medium is referred to as data carrier.

6.2 format

⟨data processing⟩ predetermined arrangement of **data** (2.2)

NOTE Formats can be subdivided into input formats, storage formats and output formats, including display formats.

5.11 blanc

caractère (5.1) qui représente une position libre dans une **chaîne** (5.12) de **caractères graphiques** (5.3)

NOTE Adapté de l'ISO 2382-4:1987.

5.12 chaîne

suite d'éléments de même nature, considérée comme un tout

EXEMPLES **Caractères** (5.1), bits ou impulsions.

NOTE Adapté de l'ISO 2382-4:1987.

5.13 chaîne de caractères

séquence de **caractères** (5.1) non interrompue par des **blancs** (5.11)

NOTE Une chaîne de caractères correspond généralement à un mot.

5.14 modèle

chaîne (5.12) dans laquelle certains **caractères** (5.1) ou certaines séquences de caractères ont été remplacés par des variables

6 Stockage des données

6.1 support de données

support physique dans lequel ou sur lequel des **données** (2.2) peuvent être enregistrées et à partir duquel des données peuvent être extraites

EXEMPLE Disque magnétique, CD ROM, microforme.

NOTE 1 Adapté de l'ISO 2382-1:1993.

NOTE 2 Le terme utilisé dans le domaine de la bibliothéconomie et de la documentation pour «support de données» est «support d'information».

6.2 format

⟨traitement des données⟩ arrangement prédéfini de **données** (2.2)

NOTE On distingue les formats d'entrée, les formats de stockage et les formats de sortie, ce dernier comprenant également les formats d'affichage.

6.3
data acquisition

process of collecting, entering and storing **data** (2.2) into a data processing system

6.4
machine-readable data

data (2.2) which are readily available for **data processing** (2.3)

6.5
database

collection of **data** (2.2) organized according to a pre-established structure

NOTE Adapted from ISO 2382-1:1993.

6.6
terminological database
database (6.5) containing terminological data

6.7
data bank
collection of **databases** (6.5) including the organizational framework for recording, processing and disseminating **data** (2.2)

6.8
term bank
terminological data bank
data bank (6.7) containing terminological **data** (2.2)

6.9
record
set of **data elements** (6.11) treated as a unit

[ISO 2382-4:1987]

6.10
file
named set of **records** (6.9) stored or processed as a unit

[ISO 2382-4:1987]

6.11
data element
unit of **data** (2.2) that, in a certain context, is considered indivisible

[ISO 2382-4:1987]

6.3
acquisition de données
processus consistant à collecter, saisir et stocker des **données** (2.2) dans un système de traitement des données

6.4
données exploitables par une machine
données (2.2) immédiatement disponibles pour le **traitement des données** (2.3)

6.5
base de données
regroupement de **données** (2.2) organisées selon une structure préétablie

NOTE Adapté de l'ISO 2382-1:1993.

6.6
base de données terminologiques
base de données (6.5) contenant des données terminologiques

6.7
banque de données
ensemble de **bases de données** (6.5) incluant la structure organisationnelle permettant l'enregistrement, le traitement et la propagation des **données** (2.2)

6.8
banque de termes
banque de données terminologiques
banque de données (6.7) contenant des **données** (2.2) terminologiques

6.9
enregistrement
ensemble d'**éléments de données** (6.11) traité comme un tout

[ISO 2382-4:1987]

6.10
fichier
ensemble identifié d'**enregistrements** (6.9) rangé en mémoire ou traité comme un tout

[ISO 2382-4:1987]

6.11
élément de données
unité de **donnée** (2.2) considérée comme étant indivisible dans un certain contexte

[ISO 2382-4:1987]

6.12**data field**

variable or fixed length portion of a **record** (6.9) reserved for a particular **data element** (6.11)

6.13**data field directory**

index (4.5) to the location of **data fields** (6.12) within a **record** (6.9), containing the **label** (6.16), length and location of each data field within the record

6.14**data category**

data element type
result of the specification of a given **data field** (6.12)

6.15**identifiant**

one or more **characters** (5.1) used to identify a name or a **data category** (6.14)

NOTE 1 An identifier may also indicate certain properties of that **data category** (6.14).

NOTE 2 Adapted from ISO 2382-4:1987.

6.16**label**

identifiant (6.15) or key within or attached to a set of **data elements** (6.11) used to identify the **data** (2.2)

NOTE Adapted from ISO 2382-4:1987.

6.17**repeatable category**

data category (6.14) which can occur more than once in the same **record** (6.9)

EXAMPLE A multilingual **terminological entry** (2.22) may contain **records** (6.9) in which certain categories are repeated for every language. Categories like synonyms, sources, context, etc. can also be repeated in monolingual terminological entries.

7 Information retrieval**7.1****information retrieval****IR**

actions, methods and procedures for obtaining **information** (2.1) from stored **data** (2.2)

6.12**champ de données**

partie d'un **enregistrement** (6.9), de longueur fixe ou variable, réservée pour un **élément de données** (6.11) particulier

6.13**répertoire des champs de données**

index (4.5) répertoriant les emplacements des **champs de données** (6.12) dans un **enregistrement** (6.9), contenant le **label** (6.16), la longueur et l'emplacement de chaque champ de données dans l'enregistrement

6.14**catégorie de données**

type d'élément de données
résultat de la spécification d'un **champ de données** (6.12) défini

6.15**identificateur**

un ou plusieurs **caractères** (5.1) permettant d'identifier un nom ou une **catégorie de données** (6.14)

NOTE 1 Un identificateur peut également indiquer certaines propriétés de la **catégorie de données** (6.14).

NOTE 2 Adapté de l'ISO 2382-4:1987.

6.16**label**

identificateur (6.15) ou clé faisant partie d'un ensemble d'**éléments de données** (6.11) ou associé à cet ensemble, utilisé pour identifier les **données** (2.2)

NOTE Adapté de l'ISO 2382-4:1987.

6.17**catégorie répétable**

catégorie de données (6.14) pouvant apparaître plus d'une fois dans le même **enregistrement** (6.9)

EXAMPLE Une **entrée terminologique** (2.22) multilingue peut contenir plusieurs **enregistrements** (6.9) dans lesquels certaines catégories sont répétées pour chaque langue. Les catégories comme les synonymes, les sources, le contexte, etc. peuvent également être répétées dans des entrées terminologiques unilingues.

7 Recherche documentaire**7.1****recherche documentaire**

recherche d'informations

ensemble des opérations, méthodes et procédures ayant pour effet d'obtenir des **informations** (2.1) à partir de **données** (2.2) rangées en mémoire

NOTE 1 Information retrieval may be performed in interactive processing or batch processing.

NOTE 2 Adapted from ISO 2382-1:1993.

7.2 query language

⟨information retrieval⟩ data manipulation language for users to retrieve stored **data** (2.2) in a **database** (6.5)

7.3 truncation

⟨information retrieval⟩ process by which a **string** (5.12) is shortened by deleting elements from the left and/or the right

7.4 compression

⟨natural-language processing⟩ operation transforming a **word form** (2.11) into a simplified form for **data processing** (2.3)

NOTE Compression usually involves the stripping of all diacritics, the suppression of **blanks** (5.11) and all **special characters** (5.8) and the elimination of the upper/lower case distinctions. Compression is used, for example, for the generation of **indexes** (4.5) to neutralize certain spelling variants.

7.5 search key

search word
⟨information retrieval⟩ **character string** (5.13) used to provide access to a **record** (6.9)

NOTE A search key may consist of a term, a word, a **word part** (2.18) or a **label** (6.16).

7.6 stem search

degradation
optional query procedure using **search keys** (7.5) automatically produced by dropping **characters** (5.1) in **character strings** (5.13) successively according to predefined rules

NOTE The order can be from left to right or right to left, depending on word formation in the different languages.

NOTE 1 La recherche documentaire peut être réalisée sous la forme d'un traitement interactif ou d'un traitement par lots.

NOTE 2 Adapté de l'ISO 2382-1:1993.

7.2 langage de requête

langage d'interrogation
⟨recherche documentaire⟩ langage de manipulation des données permettant aux utilisateurs de rechercher des **données** (2.2) stockées dans une **base de données** (6.5)

7.3 troncature

⟨recherche documentaire⟩ opération consistant à raccourcir une **chaîne** (5.12) en supprimant des éléments au début et/ou à la fin de la chaîne

7.4 compression

⟨traitement en langage naturel⟩ opération consistant à transformer une **forme lexicale** (2.11) en une forme simplifiée pour permettre le **traitement des données** (2.3)

NOTE La compression consiste généralement à supprimer tous les diacritiques, tous les **blancs** (5.11) et tous les **caractères spéciaux** (5.8) ainsi qu'à éliminer la distinction entre majuscules et minuscules. La compression est utilisée par exemple pour la génération d'**index** (4.5) afin de neutraliser certaines variantes orthographiques.

7.5 clé de recherche

critère de recherche
⟨recherche documentaire⟩ **chaîne de caractères** (5.13) utilisée pour fournir un accès à un **enregistrement** (6.9)

NOTE Une clé de recherche peut être un terme, un mot, une **partie de mot** (2.18) ou un **label** (6.16).

7.6 recherche du radical

procédure de requête facultative effectuée au moyen de **clés de recherche** (7.5) générées automatiquement par l'élimination successive de **caractères** (5.1) dans des **chaînes de caractères** (5.13) en fonction de règles prédéfinies

NOTE La recherche peut s'effectuer de la gauche vers la droite ou de la droite vers la gauche, selon les règles de formation des mots des différents langages.

7.7**search**, verb

⟨information retrieval⟩ examine automatically a set of **data** (2.2) with the aim of **extracting** (7.9) or **matching** (7.8) elements that have a given property

EXAMPLE Examining a set of names for all names beginning with J.

7.8**match**, verb

⟨information retrieval⟩ determine by comparison identity between items of one set of data and items of another set of **data** (2.2) in order to select or separate identical items

7.9**extract**, verb

⟨information retrieval⟩ select from a set of **data** (2.2) those items which meet predefined properties

7.10**ranking**

⟨information retrieval⟩ assigning weight factors to answers according to their relevance with respect to the corresponding search query

8 Data handling**8.1****read**, verb

acquire **data** (2.2) from a **data medium** (6.1), leaving the data unchanged

NOTE Adapted from ISO 2382-6:1987.

8.2**write**, verb

make a permanent or transient recording of **data** (2.2) on a **data medium** (6.1)

NOTE Adapted from ISO 2382-6:1987.

8.3**copy**, verb

read (8.1) **data** (2.2) from a source **data medium** (6.1), leaving the source data unchanged, and **write** (8.2) the same data on a destination data medium

EXAMPLE To copy a file from a magnetic tape onto a diskette.

NOTE 1 Adapted from ISO 2382-6:1987.

NOTE 2 Source and destination data medium may be physically the same.

7.7**rechercher**, verbe

⟨recherche documentaire⟩ examiner automatiquement un ensemble de **données** (2.2) pour **extraire** (7.9) ou **apparier** (7.8) des éléments qui présentent une caractéristique particulière

EXEMPLE Rechercher dans un ensemble de noms tous ceux qui commencent par la lettre J.

7.8**apparier**, verbe

⟨recherche documentaire⟩ établir une comparaison entre les éléments d'un ensemble de **données** (2.2) et les éléments d'un autre ensemble de **données** (2.2) afin de sélectionner ou de séparer des éléments identiques

7.9**extraire**, verbe

⟨recherche documentaire⟩ prendre dans un ensemble de **données** (2.2) celles qui satisfont à des conditions prédéfinies

7.10**classement**

⟨recherche documentaire⟩ affectation de facteurs de pondération à des réponses en fonction de leur adéquation avec les critères de recherche

8 Manipulation des données**8.1****lire**, verbe

extraire des **données** (2.2) d'un **support de données** (6.1) sans les modifier

NOTE Adapté de l'ISO 2382-6:1987.

8.2**écrire**, verbe

enregistrer des **données** (2.2), de façon permanente ou transitoire, sur un **support de données** (6.1)

NOTE Adapté de l'ISO 2382-6:1987.

8.3**copier**, verbe

lire (8.1) des **données** (2.2) enregistrées sur un **support de données** (6.1) source, sans altérer les données d'origine, et **écrire** (8.2) ces données sur un support de données cible

EXEMPLE Copier un fichier d'une bande magnétique sur une disquette.

NOTE 1 Adapté de l'ISO 2382-6:1987.

NOTE 2 Le support de données cible peut être le même que le support de données source.

8.4

sort, verb
order **data** (2.2) into groups according to specified criteria

NOTE Adapted from ISO 2382-6:1987.

8.5

merge, verb
combine the items of two or more sets of **data** (2.2) that are (each) in the same given order into one set in that order

[ISO 2382-6:1987]

8.6

display, verb
give a visual representation of **data** (2.2) on a screen

8.7

mask, verb
use a **pattern** (5.14) of **characters** (5.1) for controlling the retention or elimination of portions of another pattern of characters

NOTE Adapted from ISO 2382-6:1987.

EXAMPLE All vowels (a, e, i, o, u) can be replaced by the grouping character V.

8.8

convert, verb
change the representation of **data** (2.2) from one form to another, without changing the **information** (2.1) conveyed

EXAMPLE code conversion, analog-to-digital conversion, media conversion.

[ISO 2382-6:1987]

8.9

transliterate, verb
represent the **characters** (5.1) of one **alphabet** (5.6) by those of another alphabet

NOTE See ISO 2382-6:1987 for a definition applicable to **information processing** (2.4).

9 Data validation

9.1

data validation
process used to determine whether **data** (2.2) are formally accurate, consistent, correct, complete and plausible

NOTE Adapted from ISO 2382-8:1998.

8.4

trier, verbe
répartir des **données** (2.2) en groupes suivant des règles déterminées

NOTE Adapté de l'ISO 2382-6:1987.

8.5

fusionner, verbe
combinaison des éléments de plusieurs ensembles de **données** (2.2) rangés de façon comparable en un seul ensemble lui-même rangé suivant le même critère d'ordre

[ISO 2382-6:1987]

8.6

afficher, verbe
donner une représentation visuelle des **données** (2.2) sur un écran

8.7

masquer, verbe
utiliser un **modèle** (5.14) de **caractères** (5.1) pour déterminer les parties d'un autre modèle de caractères à retenir ou à rejeter

NOTE Adapté de l'ISO 2382-6:1987.

EXEMPLE Toutes les voyelles (a, e, i, o, u) peuvent être remplacées par le caractère de regroupement V.

8.8

convertir, verbe
changer le mode de représentation de **données** (2.2) sans modifier les **informations** (2.1) qu'elles contiennent

EXEMPLE Convertir un code, convertir des données analogiques en données numériques, faire une conversion de support.

[ISO 2382-6:1987]

8.9

translitérer, verbe
représenter les **caractères** (5.1) d'un **alphabet** (5.6) par ceux d'un autre **alphabet** (5.6)

NOTE Pour une définition dans le contexte du **traitement de l'information** (2.4), voir l'ISO 2382-6:1987.

9 Validation des données

9.1

validation des données
processus servant à déterminer si des **données** (2.2) sont exactes, cohérentes, correctes, complètes et vraisemblables

NOTE Adapté de l'ISO 2382-8:1998.

9.2 double-entry check

data validation (9.1) to determine whether the same data entry has been stored more than once

NOTE Sometimes it is desirable to allow double entries, e.g. if a term is used in different subject fields.

9.3 consistency check

data validation (9.1) to check whether inter-dependent entries comply with predefined conditions

EXAMPLE Verification of cross-references.

9.4 spelling check

spell check

data validation (9.1) to determine that all words comply with predefined spelling rules

NOTE See ISO/IEC 2382-23:1994 for the definition of spelling checker in the context of text editing in **information processing** (2.4).

9.5 completeness check

data validation (9.1) to see whether **data** (2.2) are present where required

NOTE 1 Adapted from ISO 2382-8:1986.

NOTE 2 Data must be present in mandatory **data fields** (6.12).

9.6 format check

data validation (9.1) to determine whether **data** (2.2) conform to a specified **format** (6.2)

NOTE Adapted from ISO 2382-8:1986.

9.7 plausibility check

reasonableness check

data validation (9.1) to determine whether **data** (2.2) conform to specified criteria

NOTE Adapted from ISO 2382-8:1986.

9.2 vérification des doubles entrées

validation des données (9.1) servant à déterminer si une même entrée de donnée a été enregistrée plus d'une fois

NOTE Il est quelquefois utile d'autoriser les doubles entrées, par exemple lorsqu'un terme est utilisé dans différents domaines.

9.3 contrôle de cohérence

validation des données (9.1) servant à vérifier si les entrées interdépendantes sont conformes aux conditions prédéfinies

EXEMPLE Vérification des références croisées.

9.4 contrôle de l'orthographe

contrôle orthographique

validation des données (9.1) servant à déterminer si tous les mots sont conformes aux règles d'orthographe prédéfinies

NOTE Pour la définition de correcteur orthographique dans le contexte de l'édition de texte et du **traitement de l'information** (2.4), voir l'ISO/CEI 2382-23:1994.

9.5 contrôle de présence

validation des données (9.1) servant à vérifier la présence des **données** (2.2) là où elles sont requises

NOTE 1 Adapté de l'ISO 2382-8:1986.

NOTE 2 Les données doivent figurer dans les **champs de données** (6.12) requis.

9.6 contrôle de format

contrôle de disposition

validation des données (9.1) servant à déterminer si les **données** (2.2) respectent un **format** (6.2) donné

NOTE Adapté de l'ISO 2382-8:1986.

9.7 contrôle de vraisemblance

validation des données (9.1) servant à déterminer si les **données** (2.2) sont conformes à des critères définis

NOTE Adapté de l'ISO 2382-8:1986.

10 Operations

10.1 batch processing

processing of **data** (2.2) or the accomplishment of jobs, accumulated in advance, in such a manner that the user cannot further influence its processing while it is in progress

10.2 interactive processing

mode of processing **data** (2.2) in a computer system in which a sequence of alternating entries and responses between a user and the computer system takes place

10.3 data protection

implementation of appropriate administrative, technical or physical means to guard against intentional or accidental disclosure, modification or destruction of **data** (2.2)

NOTE Adapted from ISO/IEC 2382-8:1998.

10.4 update, verb

to add, delete or change **data** (2.2)

NOTE Adapted from ISO 2382-4:1987.

10.5 file maintenance

activity of **updating** (10.4) and reorganizing a file

[ISO 2382-4:1987]

10.6 portability

⟨computer programming⟩ capability of a program to be executed on various types of data processing systems often involving recompiling, with little or no manual modification

NOTE Adapted from ISO 2382-1:1993.

10.7 compatibility

capability of a functional unit to meet the requirements of a specified interface without appreciable modification

[ISO/IEC 2382-1:1993]

10 Opérations

10.1 traitement par lots

traitement de **données** (2.2) ou réalisation de travaux, préalablement regroupés, effectués de sorte que l'utilisateur ne puisse plus agir sur les traitements en cours

10.2 traitement interactif

mode de traitement des **données** (2.2) sur un système informatique, pendant lequel une séquence alternative d'entrées et de réponses a lieu entre un utilisateur et le système informatique

10.3 protection des données

mise en vigueur d'un ensemble de mesures administratives, techniques et physiques pour interdire toute divulgation, modification ou destruction volontaire ou accidentelle des **données** (2.2)

NOTE Adapté de l'ISO/CEI 2382-8:1998.

10.4 mettre à jour, verbe

ajouter, supprimer ou modifier des **données** (2.2)

NOTE Adapté de l'ISO 2382-4:1987.

10.5 maintenance de fichier

ensemble des opérations de **mise à jour** (10.4) et de réorganisation d'un fichier

[ISO 2382-4:1987]

10.6 portabilité

⟨programmation logicielle⟩ capacité d'un programme à être exécuté sur différents systèmes de traitement des données, impliquant généralement une recompilation, avec peu ou pas de modifications manuelles

NOTE Adapté de l'ISO 2382-1:1993.

10.7 compatibilité

capacité d'une unité fonctionnelle à satisfaire aux exigences d'une interface donnée sans qu'il soit nécessaire d'apporter des modifications notables

[ISO/CEI 2382-1:1993]

10.8 user terminal

functional unit in a system or communication network that enables a user to communicate with a computer

NOTE Adapted from ISO/IEC 2382-1:1993.

10.9 menu

list of options **displayed** (8.6) to the user by a data processing system, from which the user can select an action to be initiated

NOTE Adapted from ISO/IEC 2382-1:1993.

10.10 data communication

transfer of **data** (2.2) among functional units according to sets of rules governing **data transmission** (10.11) and the coordination of the exchange

[ISO/IEC 2382-9:1995]

10.11 data transmission

electronic transfer of **data** (2.2) from one point to one or more other points over telecommunication facilities

NOTE Adapted from ISO/IEC 2382-9:1995.

10.12 file transfer

movement of one or more **files** (6.10) from one data processing system to another by electronic means

10.13 upload, verb

export programs or **data** (2.2) from a connected computer or from a **data medium** (6.1) to another computer

10.14 download, verb

import programs or **data** (2.2) from a connected computer to another computer or to a **data medium** (6.1)

10.8 terminal d'utilisateur

unité fonctionnelle dans un système ou un réseau de communication permettant à un utilisateur de communiquer avec un ordinateur

NOTE Adapté de l'ISO/CEI 2382-1:1993.

10.9 menu

liste d'options **affichées** (8.6) par un système de traitement des données, à partir duquel l'utilisateur peut sélectionner une action à exécuter

NOTE Adapté de l'ISO/CEI 2382-1:1993.

10.10 communication de données

transfert de **données** (2.2) entre unités fonctionnelles selon des ensembles de règles régissant la **transmission de données** (10.11) et la coordination des échanges

[ISO/CEI 2382-9:1995]

10.11 transmission de données

transfert électronique de **données** (2.2) d'un point à un autre par des moyens de télécommunication

NOTE Adapté de l'ISO/CEI 2382-9:1995.

10.12 transfert de fichier

déplacement d'un ou de plusieurs **fichiers** (6.10) d'un système de traitement des données à un autre par des moyens électroniques

10.13 charger, verbe

exporter des programmes ou des **données** (2.2) d'un ordinateur connecté ou d'un **support de données** (6.1) vers un autre ordinateur

10.14 télécharger, verbe

importer des programmes ou des **données** (2.2) d'un ordinateur connecté vers un autre ordinateur ou un **support de données** (6.1)

11 Related fields of application

11.1

computer-aided, adjective

computer assisted

pertaining to a technique or process in which part of the human work is done with the assistance of a computer

EXAMPLE Computer-aided terminology work, computer-aided translation, computer-aided lexicography.

11.2

text processing

word processing

data processing (2.3) operations on **text** (2.6)

NOTE 1 Adapted from ISO/IEC 2382-23:1993.

NOTE 2 Typical operations include entering, **editing** (11.3), **sorting** (8.4), **merging** (8.5), retrieving, storing, **displaying** (8.6) or printing.

11.3

text editing

editing

text processing (11.2) in order to manipulate **text** (2.6)

NOTE 1 Adapted from ISO/IEC 2382-23:1994.

NOTE 2 Typical manipulations include rearrangement, change, addition, deletion and reformatting.

11.4

highlight, verb

emphasize **displayed** (8.6) **text** (2.6) elements by modifying their visual attributes thus making them stand out

11.5

browse, verb

rapidly review a collection of documents or go through **records** (6.9) in a **file** (6.10) on a display screen

NOTE See ISO/IEC 2382-23:1994 for a definition in the context of **text processing** (11.2).

11 Domaines d'application associés

11.1

assisté par ordinateur, adjectif

relatif à une technique ou à un processus dans lequel une partie du travail humain est effectué à l'aide d'un ordinateur

EXEMPLE Terminologie assistée par ordinateur, traduction assistée par ordinateur, lexicographie assistée par ordinateur.

11.2

traitement de texte

opérations de **traitement des données** (2.3) sur du **texte** (2.6)

NOTE 1 Adapté de l'ISO/CEI 2382-23:1994.

NOTE 2 Le traitement de texte comprend généralement les opérations de saisie, d'**édition** (11.3), de **tri** (8.4), de **fusion** (8.5), de recherche, de stockage, d'**affichage** (8.6) et d'impression.

11.3

édition de texte

édition

traitement de texte (11.2) permettant de manipuler du **texte** (2.6)

NOTE 1 Adapté de l'ISO/CEI 2382-23:1994.

NOTE 2 Les manipulations courantes incluent la réorganisation, la modification, l'addition, la suppression et le reformatage.

11.4

mettre en évidence, verbe

faire ressortir des éléments du **texte** (2.6) **affiché** (8.6) en modifiant leurs attributs visuels afin qu'ils soient mis en relief

11.5

parcourir, verbe

consulter rapidement un ensemble de documents ou visualiser les **enregistrements** (6.9) d'un **fichier** (6.10) sur un écran d'affichage

NOTE Pour la définition du terme concerné dans le contexte du **traitement de texte** (11.2), voir l'ISO/CEI 2382-23:1994.

Bibliography

- [1] ISO/IEC 646:1991, *Information technology — ISO 7-bit coded character set for information interchange.*
- [2] ISO 1087:1990, *Terminology — Vocabulary.*
- [3] ISO 1087-1:—¹⁾, *Terminology work — Vocabulary — Part 1: Theory and application.*
- [4] ISO/IEC 2382-1:1993, *Information technology — Vocabulary — Part 1: Fundamental terms.*
- [5] ISO 2382-4:1987, *Information processing systems — Vocabulary — Part 4: Organization of data.*
- [6] ISO 2382-6:1987, *Information processing systems — Vocabulary — Part 6: Preparation and handling of data.*
- [7] ISO/IEC 2382-8:1998, *Information technology — Vocabulary — Part 8: Security.*
- [8] ISO 2382-8:1986, *Information processing systems — Vocabulary — Part 8: Control, integrity and security.*
- [9] ISO/IEC 2382-9:1995, *Information technology — Vocabulary — Part 9: Data communication.*
- [10] ISO/IEC 2382-17:1999, *Information technology — Vocabulary — Part 17: Databases.*
- [11] ISO/IEC 2382-23:1994, *Information technology — Vocabulary — Part 23: Text processing.*
- [12] ISO/IEC 4873:1991, *Information technology — ISO 8-bit code for information interchange — Structure and rules for implementation.*
- [13] ISO/IEC 6937:1994, *Information technology — Coded graphic character sets for text communication — Latin alphabet.*
- [14] ISO/IEC 6937-1:1983, *Information processing — Coded character sets for text communication — Part 1: General introduction.*

1) To be published.

Bibliographie

- [1] ISO/CEI 646:1991, *Technologies de l'information — Jeu ISO de caractères codés à 7 éléments pour l'échange d'informations.*
- [2] ISO 1087:1990, *Terminologie — Vocabulaire.*
- [3] ISO 1087-1:—¹⁾, *Travail terminologique — Vocabulaire — Partie 1: Théorie et application.*
- [4] ISO/CEI 2382-1:1993, *Technologies de l'information — Vocabulaire — Partie 1: Termes fondamentaux.*
- [5] ISO 2382-4:1987, *Systèmes de traitement de l'information — Vocabulaire — Partie 4: Organisation des données.*
- [6] ISO 2382-6:1987, *Systèmes de traitement de l'information — Vocabulaire — Partie 6: Préparation et manipulation des données.*
- [7] ISO/CEI 2382-8:1998, *Technologies de l'information — Vocabulaire — Partie 8: Sécurité.*
- [8] ISO 2382-8:1986, *Systèmes de traitement de l'information — Vocabulaire — Partie 8: Contrôle, intégrité et sécurité.*
- [9] ISO/CEI 2382-9:1995, *Technologies de l'information — Vocabulaire — Partie 9: Communication de données.*
- [10] ISO/CEI 2382-17:1999, *Technologies de l'information — Vocabulaire — Partie 17: Bases de données.*
- [11] ISO/CEI 2382-23:1994, *Technologies de l'information — Vocabulaire — Partie 23: Traitement de texte.*
- [12] ISO/CEI 4873:1991, *Technologies d'information — Code ISO à 8 éléments pour l'échange d'informations — Structure et règles de matérialisation.*
- [13] ISO/CEI 6937:1994, *Technologies de l'information — Jeu de caractères graphiques codés pour la transmission de texte — Alphabet latin.*

1) À publier.

[15] ISO 10241:1992, *International terminology standards — Preparation and layout.*

[16] ISO/TR 12618:1994, *Computational aids in terminology — Creation and use of terminological databases and text corpora.*

[14] ISO/CEI 6937-1:1983, *Traitement de l'information — Jeux de caractères codés pour la transmission de texte — Partie 1: Introduction générale.*

[15] ISO 10241:1992, *Normes terminologiques internationales — Élaboration et présentation.*

[16] ISO/TR 12618:1994, *Aides informatiques en terminologie — Création et utilisation de bases de données terminologiques et de corpus de textes.*

Alphabetical index

A

alphabet 5.6
alphabetical character 5.5
alphabetical index 4.7
alphabetical ordering 3.3
alphanumeric character 5.4

B

base form 2.13
batch processing 10.1
blank 5.11
browse, verb 11.5

C

character 5.1
character set 5.2
character string 5.13
compatibility 10.7
completeness check 9.5
compression (natural-language processing) 7.4
compressed form 2.15
computer-aided adjective 11.1
 computer assisted 11.1
concatenation 3.7
concordance (computer applications) 4.4
consistency check 9.3
control character 5.9
convert, verb 8.8
copy, verb 8.3
corpus (language engineering) 2.7

D

data 2.2
data acquisition 6.3
data bank 6.7
database 6.5
data category 6.14
data communication 10.10
data element 6.11
 data element type 6.14
data field 6.12
data field directory 6.13
data medium 6.1
data processing 2.3
data protection 10.3
data transmission 10.11
data validation 9.1
 degradation 7.6
deinflected word form 2.17
deinflection 2.16
delimiter 2.10
digit 5.7
disambiguation 2.24
display, verb 8.6
double-entry check 9.2
download, verb 10.14
DP 2.3

E

editing 11.3
exclusion list 4.1
extract, verb (information retrieval) 7.9

F

file 6.10
file maintenance 10.5
file transfer 10.12
format (data processing) 6.2
format check 9.6
free-text search 4.3
frequency order 3.5

G

graphical character 5.3

H

highlight, verb 11.4
homograph 2.23

I

identifier 6.15
inclusion list 4.2
index (natural-language processing) 4.5
information (information processing) 2.1
information processing 2.4
information retrieval 7.1
interactive processing 10.2
inverted index 4.11
IR 7.1

L

label 6.16
language processing 2.5
lemmatization 2.19
lemmatized word index 4.13

M

machine-readable data 6.4
mask, verb 8.7
match, verb (information retrieval) 7.8
menu 10.9
merge, verb 8.5
multi-word term 2.14

N

numerical character 5.7

P

paradigm (natural-language processing) 2.12
parsing (language processing) 2.20
pattern 5.14
permutation 3.6
permuted index 4.9
plausibility check 9.7
 plus word list 4.2
portability (computer programming) 10.6

Q

query language (information retrieval) 7.2

R

ranking (information retrieval) 7.10
read, verb 8.1
 reasonableness check 9.7
record 6.9
 reduced form (deprecated) 2.15
 reference form 2.13
repeatable category 6.17
reverse alphabetical ordering 3.4
reverse index 4.10

S

search key (information retrieval) 7.5
search, verb (information retrieval) 7.7
 search word (information retrieval) 7.5
 separator 2.10
sort key 3.2
sort value 3.1
sort, verb 8.4
 sorting key 3.2
 sorting value 3.1
space character 5.10
special character 5.8
 spell check 9.4
spelling check 9.4
stem search 7.6
 stop word list 4.1
string 5.12
systematic index 4.8

T

term bank 6.8
term index 4.6
 terminological data bank 6.8
terminological database 6.6
terminological data collection 2.21

terminological entry 2.22
text 2.6
text corpus (language engineering) 2.7
text editing 11.3
text processing 11.2
token (natural-language processing) 2.9
transliterate, verb 8.9

truncation (information retrieval) 7.3
type (language engineering) 2.8

U

update, verb 10.4
upload, verb 10.13
user terminal 10.8

W

word form 2.11
word form index 4.12
word index 4.12
word part 2.18
word processing 11.2
word segment 2.18
write, verb 8.2

Index alphabétique

A

acquisition de données 6.3
afficher, verbe 8.6
alphabet 5.6
analyse syntaxique (traitement du langage) 2.20
appairer, verbe (recherche documentaire) 7.8
assisté par ordinateur, adjectif 11.1

B

banque de données 6.7
banque de données terminologiques 6.8
banque de termes 6.8
base de données 6.5
base de données terminologiques 6.6
blanc 5.11

C

caractère 5.1
caractère alphabétique 5.7
caractère alphanumérique 5.4
caractère de commande 5.9
caractère espace 5.10
caractère graphique 5.3
caractère numérique 5.7
caractère spécial 5.8
catégorie de données 6.14
catégorie répétable 6.17
chaîne 5.12
chaîne de caractères 5.13
champ de données 6.12
charger, verbe 10.13
chiffre 5.7
classement (recherche documentaire) 7.10
clé de recherche (recherche documentaire) 7.5
clé de tri 3.2
collecte de données terminologiques 2.21
communication de données 10.10
compatibilité 10.7
compression (traitement en langage naturel) 7.4
concaténation 3.7
concordance (traitement en langage naturel) 4.4
concordance KWIC, voir 4.4

contrôle de cohérence 9.3
contrôle de disposition 9.6
contrôle de format 9.6
contrôle de l'orthographe 9.4
contrôle de présence 9.5
contrôle de vraisemblance 9.7
contrôle orthographique 9.4
convertir, verbe 8.8
copier, verbe 8.3
corpus 2.7
corpus de textes (ingénierie linguistique) 2.7
critère de tri 3.2

D

délimiteur 2.10
désambiguïsation 2.24
donnée 2.2
données exploitables par une machine 6.4

E

écrire, verbe 8.2
édition 11.3
édition de texte 11.3
élément de données 6.11
enregistrement 6.9
entrée terminologique 2.22
extraire, verbe (recherche documentaire) 7.9

F

fichier 6.10
format 6.2
forme comprimée 2.15
forme de base 2.13
forme de référence 2.13
forme lexicale 2.11
forme lexicale sans désinence 2.17
forme réduite (rejeté), voir 2.15
fusionner, verbe 8.5

H

homographe 2.23

I

identificateur 6.15

index (traitement en langage naturel) 4.5

index à l'envers 4.11
index alphabétique 4.7
index de formes de mot 4.12
index de mots 4.12
index de mots lemmatisés 4.13
index de termes 4.6
index inversé 4.10
index permuté 4.9
index systématique 4.8
information (traitement de l'information) 2.1

J

jeu de caractères 5.2

K

KWIC 4.4

L

label 6.16
langage d'interrogation (recherche documentaire) 7.2
langage de requête (recherche documentaire) 7.2
lemmatisation 2.19
lire, verbe 8.1
liste d'exclusion 4.1
liste d'inclusion 4.2

M

maintenance de fichier 10.5
masquer, verbe 8.7
menu 10.9
mettre à jour, verbe 10.4
mettre en évidence, verbe 11.4
modèle 5.14
mot de fonction, voir 4.1

O

occurrence (traitement en langage naturel) 2.9
ordre alphabétique 3.3
ordre alphabétique inversé 3.4
ordre de fréquence 3.5

P

paradigme (traitement en langage naturel) 2.12
parcourir, verbe 11.5
partie de mot 2.18
permutation 3.6
portabilité (traitement en langage naturel) 10.6
protection des données 10.3

R

recherche d'informations 7.1
recherche du radical 7.6
rechercher, verbe (recherche documentaire) 7.7
répertoire de champs de donnée 6.13

S

segment de mot 2.18
séparateur 2.10
support de données 6.1
suppression des désinences 2.16
système de gestion de bases de données 6.6
système de recherche en forme libre 4.3

T

télécharger, verbe 10.13
terme composé 2.14
terminal de l'utilisateur 10.8
texte 2.6
traitement de l'information 2.4

traitement de texte 11.2
traitement des données 2.3
traitement du langage 2.5
traitement interactif 10.2
traitement par lots 10.1
transfert de fichier 10.12
translitérer, verbe 8.9
transmission des données 10.11
trier, verbe 8.4
troncature (recherche documentaire) 7.3
type (ingénierie linguistique) 2.8
type d'élément de données 6.14

V

valeur de tri 3.1
validation des données 9.1
vérification des doubles entrées 9.2

