INTERNATIONAL STANDARD

ISO/IEC 13249-2

Second edition
2003-11-01

Information technology — Database languages — SQL multimedia and application packages —

Part 2: **Full-Text**

Technologies de l'information — Langages de bases de données — Multimédia SQL et paquetages d'application —

Partie 2: Texte complet

Reference number
ISO/IEC 13249-2:2003(E)

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

Contents Page

Tables

Page

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work. In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of the joint technical committee is to prepare International Standards. Draft International Standards adopted by the joint technical committee are circulated to national bodies for voting. Publication as an International Standard requires approval by at least 75 % of the national bodies casting a vote.

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

ISO/IEC 13249-2 was prepared by joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 32, *Data management and interchange*.

This second edition cancels and replaces the first edition (ISO/IEC 13249-2:2000), which has been technically revised.

ISO/IEC 13249 consists of the following parts, under the general title *Information technology — Database languages — SQL multimedia and application packages*:

— *Part 1: Framework*

— *Part 2: Full-Text*

— *Part 3: Spatial*

— *Part 5: Still image*

— *Part 6: Data mining*

Introduction

The purpose of this International Standard is to define multimedia and application specific types and their associated routines using the user-defined features in ISO/IEC 9075.

This document is based on the content of ISO/IEC International Standard Database Language (SQL).

The organization of this part of ISO/IEC 13249 is as follows:

1) Clause 1, "Scope", specifies the scope of this part of ISO/IEC 13249.
2) Clause 2, "Normative references", identifies additional standards that, through reference in this part of ISO/IEC 13249, constitute provisions of this part of ISO/IEC 13249.
3) Clause 3, "Terms and definitions, notations and conventions", defines the notations and conventions used in this part of ISO/IEC 13249.
4) Clause 4, "Concepts", presents concepts used in the definition of this part of ISO/IEC 13249.
5) Clause 5, "Full-Text Types", defines the full-text user-defined types and associated routines.
6) Clause 6, "Structured Search Pattern Types", defines user-defined types to provide for the construction of structured search patterns.
7) Clause 7, "FullText_Token Type and Routines", defines the user-defined FullText_Token type.
8) Clause 8, "SQL/MM Full-Text Thesaurus Schema", defines the SQL/MM Full-Text thesaurus schema used to define the thesaurus related routines.
9) Clause 9, "SQL/MM Full-Text Information Schema", defines the SQL/MM Full-Text Information Schema.
10) Clause 10, "SQL/MM Full-Text Definition Schema", defines the SQL/MM Full-Text Definition Schema.
11) Clause 11, "Status Codes", defines the SQLSTATE codes used in this part of ISO/IEC 13249.
12) Clause 12, "Conformance", defines the criteria for conformance to this part of ISO/IEC 13249.
13) Annex A, "Implementation-defined elements", is an informative Annex. It lists those features for which the body of this part of ISO/IEC 13249 states that the syntax or meaning or effect on the database is partly or wholly implementation-defined, and describes the defining information that an implementer shall provide in each case.
14) Annex B, "Implementation-dependent elements", is an informative Annex. It list those features which the body of this part of ISO/IEC 13249 states explicitly that the syntax or meaning or effect on the database is implementation-dependent.

In the text of this part of ISO/IEC 13249, Clauses begin a new odd-numbered page, and in Clause 5, "Full-Text Types", through Clause 12, "Conformance", Subclauses begin a new page. Any resulting blank space is not significant.

Information technology — Database languages — SQL multimedia and application packages —

Part 2: **Full-Text**

1 Scope

This part of ISO/IEC 13249:

a) introduces the Full-Text part of ISO/IEC 13249 (all parts);

b) gives the references necessary for this part of ISO/IEC 13249;

c) defines notations and conventions specific to this part of ISO/IEC 13249;

d) defines concepts specific to this part of ISO/IEC 13249;

e) defines the full-text user-defined types and their associated routines.

The full-text user-defined types defined in this part of ISO/IEC 13249 adhere to the following.

— A full-text user-defined type is generic to text handling. It addresses the need to search and retrieve information based on aspects of full-text data using patterns such as words, phrases, proximity expansion, fuzzy expansion, and thesaurus based expansions. It also addresses the need to construct such search patterns for text identification facilities and text ranking facilities.

— A full-text user-defined type does not redefine the database language SQL directly or in combination with another full-text data type.

An implementation of this part of ISO/IEC 13249 may exist in environments that also support information and content management, decision support, data mining, and data warehousing systems.

Application areas addressed by implementations of this part of ISO/IEC 13249 include, but are not restricted to, library, newspaper, multimedia, and scientific research applications.

2 Normative references

The following referenced documents are indispensable for the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO/IEC 9075 (all parts), *Information technology — Database languages — SQL*

ISO/IEC 13249-1:2002, *Information technology — Database languages — SQL multimedia and application packages — Part 1: Framework*

ANSI/NISO Z39.19:1993, American National Standard for Information Systems/National Information Standards Organization, *Guidelines for the Construction, Format, and Management of Monolingual Thesauri*