

TECHNICAL REPORT

ISO/IEC TR 9573

First edition
1988-12-01

Information processing — SGML support facilities — Techniques for using SGML

*Traitement de l'information — Facilités de support pour SGML — Techniques
d'utilisation du SGML*



Reference number
ISO/IEC/TR 9573 : 1988 (E)

Contents

0 Introduction	1
1 Scope and Field of Application	1
2 References	2
3 Definitions	2
4 Document Analysis for the Creation of a Document	
Type Definition	2
4.1 General notes	2
4.1.1 Steps of document analysis	3
4.1.2 Design issues	3
4.2 Example of the analysis of body text	5
4.2.1 Headed sections	7
4.2.2 Paragraph and note elements	7
4.2.3 List elements	8
4.2.3.1 Separately identified elements	8
4.2.3.2 Common list element	9
4.2.3.3 Discussion of the two approaches to lists	9
4.2.4 Footnote element	10
4.2.5 Figure element	10
4.2.6 Floating elements	10
4.2.7 Reference elements	11
4.2.8 Summary of the definition of the body element	11
4.2.9 Additional elements	12
4.2.10 Markup minimization	12
4.3 Example of the analysis of title matter	12
4.3.1 Title element	13
4.3.2 Instlist element	14
4.3.3 Authlist element	15
4.3.4 Date element	16
4.3.5 Abstract element	17
4.3.6 Article element	17
4.3.7 Complete definition of the title matter	18
4.4 Example of the analysis of tabular material	19
4.5 Example of the analysis of a simple mathematical formula	21
4.6 Example of the analysis of an auction catalogue	22
4.7 The DTD with relation to the size of a document	23
5 General Document	23
5.1 General notes	23
5.2 Document structure	23
5.3 Description of the tags	31
5.3.1 Major document elements	31
5.3.2 Title page elements	32
5.3.3 Front matter elements	33
5.3.4 Back matter elements	34
5.3.5 Headed sections	34
5.3.6 Topics	34
5.3.7 Tabular material	35
5.3.8 Figure	35

5.3.9	Phrases	36
5.3.10	Lists	37
5.3.11	References	39
5.3.12	Basic elements	40
5.4	Alphabetical list of the generic identifiers and their attributes	42
5.5	The document type definition	43
5.6	Commentary on the document type definition	47
5.6.1	Parameter entity declarations	47
5.6.2	Element and attribute declarations	47
5.6.3	Short references	47
5.7	National versions of the elements of the general document	48
5.7.1	French	48
5.7.2	German	51
5.7.3	Swedish	54
5.7.4	Danish	57
5.7.5	Dutch	60
5.7.6	Generic identifiers	63
5.8	Extensions to the general document	64
5.8.1	Extending the bibliography to include a bibliography list	64
5.8.2	Extending the DTD to include tables	65
6	Letter and Memorandum	66
6.1	Constituent parts of the documents	66
6.1.1	Common document elements	66
6.1.2	Elements specific to a business letter	68
6.1.3	Elements specific to an inter-office memorandum	69
6.1.4	Alphabetical list of the tags	69
6.2	Entity set	69
6.3	SGML features	70
6.4	Typical document type declarations	70
6.4.1	The document type declaration for letters	70
6.4.2	The document type declaration for memoranda	70
6.5	The common document type definition	70
6.6	A sample letter	71
6.7	A sample inter-office memorandum	72
7	Spreadsheet	72
7.1	Constituent parts of a spreadsheet	73
7.1.1	Elements of a spreadsheet	73
7.1.2	Elements of a formula	74
7.1.3	Alphabetical list of the tags	76
7.1.4	Entity sets	76
7.1.4.1	Parameter entities	77
7.1.4.2	Entities used with short references	77
7.1.4.3	Short references, notations, maps, etc.	77
7.1.5	SGML features	77
7.2	The document type definition	77
7.3	A sample spreadsheet	79
8	Mathematics	83
8.1	Concepts	83
8.2	User-interface	83
8.3	Symbols	84
8.4	Description of the elements	84
8.4.1	Formula and formula reference	84

8.4.2	Formula content	85
8.5	Short references	92
8.6	Complex examples	92
8.7	Element and entity definition	93
9	Tables	98
9.1	Table elements	98
9.2	Element definition	99
9.3	Example of a marked-up table	99
10	Integration of Computer Graphics Metafiles	100
11	Japanese	105
11.1	Special considerations	105
11.1.1	Horizontal and vertical typesetting	105
11.1.2	Sentences including words in the Latin alphabet	105
11.1.3	Ruby	107
11.2	The SGML declaration	107
11.3	The document type definition and source document	109
12	Linguistic Application — Scandinavian Runes	111
12.1	Use of tags and entities	111
12.2	Character extension techniques	113
12.3	Character extension techniques mapped onto tags using short references	114
12.4	Definition of a special character set in the SGML declaration	115
13	European Multi-lingual Document	117
14	The Use of Short References	118
14.1	Single level quotation in English	118
14.2	Limitations of the use of short references to be observed	119
15	Mixing Text in Languages Written from Left to Right and from Right to Left	120
15.1	Right to left base text mixed with text written from left to right	121
15.2	Left to right base text mixed with text written from right to left	122
15.2.1	Use of tags	122
15.2.2	Use of an 8-bit Latin/Arabic character set	123

Figures

1	Sample pages from the body of an International Standard	6
2	Sample title page	13
3	Overall structure of the basic document	24
4	Structure of the h2 element	25
5	Structure of the phrase (%m.ph), paragraph (%m.p), and paragraph sequence (%m.pseq)	

	models	26
6	Structure of the section subelements	27
7	Structure of all phrases	28
8	Structure of all paragraph/section subelements	29
9	Structure of all floating elements	31
10	Overall structure of a spreadsheet	73
11	Structure of a formula in a spreadsheet	75
12	Japanese: horizontal typesetting (text written from left to right)	106
13	Japanese: vertical typesetting (text written from top to bottom)	107
14	Japanese: source document	110

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) together form a system for worldwide standardization as a whole. 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.

The main task of a technical committee is to prepare International Standards but in exceptional circumstances, the publication of a technical report of one of the following types may be proposed:

- type 1, when the necessary support within the technical committee cannot be obtained for the publication of an International Standard, despite repeated efforts;
- type 2, when the subject is still under technical development requiring wider exposure;
- type 3, when a technical committee has collected data of a different kind from that which is normally published as an International Standard ("state of the art", for example).

Technical reports of types 1 and 2 are subject to review within three years of publication, to decide whether they can be transformed into International Standards. Technical reports of type 3 do not necessarily have to be reviewed until the data they provide are considered to be no longer valid or useful.

ISO/IEC/TR 9573, which is a technical report of type 3, was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*.

Information processing — SGML support facilities — Techniques for using SGML

0 Introduction

ISO 8879, *Information processing — Text and office systems — Standard Generalized Markup Language (SGML)*, states the rules for the description and markup of documents for their publication and interchange. A basic document type, primarily for computer-assisted publishing, is provided as an example in clause E.1 of ISO 8879, but is not explained there. It is explained in this Technical Report to assist comprehension. Also given is a variety of examples on the use of SGML. Thus this Technical Report is complementary to ISO 8879, its principal purpose being to assist in the adoption of the language.

1 Scope and Field of Application

This Technical Report complements ISO 8879 by providing additional tutorial information. It is not intended, and should not be regarded, as an extension, modification, or interpretation of ISO 8879. The SGML language contains a number of components, some of which are optional features. The tutorial information covers the main components of the language only; notably a discussion of LINK, CONCUR, and DATATAG is outside the scope of this Technical Report.

The intended audience is mainly document type designers already familiar with the basic concepts of SGML, but requiring more tutorial information on techniques for using SGML for various applications. Subclauses 5.3 and 8.4 are written in the style of a "User Guide" and can be used as a basis for end-user documentation. For an introductory tutorial on SGML the annexes of ISO 8879 can be used.

This Technical Report includes notes on the analysis of a document prior to the writing of a formal document type definition, and a series of examples.

The principal example is for a general document type, formally defined as an example in clause E.1 of ISO 8879. Others of a general nature are for letter and memorandum, spreadsheet, mathematics, and the mixing of text and graphics. Those for language applications include Scandinavian runes, Japanese, a European multilingual document, and mixing text in languages written from left to right and from right to left.

NOTE — Throughout this Technical Report terms like "is keyed in", and "with keyboarding" are used. This does not necessarily imply that the markup is to be added explicitly by a user; for text entry one would expect structured, context sensitive, editors to be used, or the markup added by application programs, e.g. in the case of interchange of spreadsheets.

2 References

ISO 646, *Information processing — ISO 7-bit coded character set for information interchange.*

ISO 2022, *Information processing — ISO 7-bit and 8-bit coded character sets — Code extension techniques.*

ISO 4873, *Information processing — ISO 8-bit code for information interchange — Structure and rules for implementation.*

ISO 6937, *Information processing — Coded character sets for text communication.*

ISO 8632-2, *Information processing systems — Computer graphics — Metafile for the storage and transfer of picture description information — Part 2: Character encoding.*

ISO 8632-4, *Information processing systems — Computer graphics — Metafile for the storage and transfer of picture description information — Part 4: Clear text encoding.*

ISO 8859-6, *Information processing — 8-bit single-byte coded graphic character sets — Part 6: Latin/Arabic alphabet.*

ISO 8879, *Information processing — Text and office systems — Standard Generalized Markup Language (SGML).*

ISO 9069, *Information processing — SGML support facilities — SGML Document Interchange Format (SDIF).¹⁾*

ISO 9070, *Information processing — SGML support facilities — Registration procedures for public text owner identifiers.¹⁾*

JIS X 0201 (formerly C 6220), *Code for Information Interchange.*

JIS X 0208 (formerly C 6226), *Code of the Japanese Graphic Character Set for Information Interchange.*

1) At present at the stage of draft.