

# INTERNATIONAL STANDARD

ISO  
9069

First edition  
1988-09-15



---

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION  
ORGANISATION INTERNATIONALE DE NORMALISATION  
МЕЖДУНАРОДНАЯ ОРГАНИЗАЦИЯ ПО СТАНДАРТИЗАЦИИ

---

## **Information Processing — SGML Support Facilities — SGML Document Interchange Format (SDIF)**

*Traitement de l'information — Bureautique — Format d'échange de document SGML (SDIF)*

STANDARDSISO.COM : Click to view the full PDF of ISO 9069:1988

Reference number  
ISO 9069: 1988 (E)

## Foreword

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

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

International Standard ISO 9069 was prepared by Technical Committee ISO/TC97, *Information processing systems*.

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

Annex A is for information only.

STANDARDSISO.COM : Click to view the full PDF of ISO 9069:1988

## Contents

<b>1</b>	<b>Scope</b>	<b>1</b>
<b>2</b>	<b>Field of application</b>	<b>1</b>
<b>3</b>	<b>References</b>	<b>1</b>
<b>4</b>	<b>Definitions</b>	<b>1</b>
<b>5</b>	<b>Description of the data stream</b>	<b>2</b>
5.1	Organization of the data stream	2
5.2	Data stream character set	2
5.3	SDIF name	3
5.4	Document descriptor	3
5.4.1	Included entities	3
5.4.2	Document character set	3
5.5	SDIF identifier	3
5.6	Entity descriptor	4
<b>6</b>	<b>Abstract data stream format</b>	<b>4</b>
<b>7</b>	<b>Encoding rules</b>	<b>5</b>
<b>Annex</b>		
<b>A</b>	<b>Relation to other standards</b>	<b>6</b>
A.1	Message-handling standards	6
A.2	Office Document Architecture (ODA)	6
A.3	OSI Presentation Services	6
A.3.1	Object identifiers	6
A.3.1.1	Abstract syntax	6
A.3.1.2	Transfer syntax	6
A.3.2	OSI Document Type Definition	7
A.3.2.1	Identifier	7
A.3.2.2	Descriptor value	7
A.3.2.3	Parameter syntax	7
A.3.2.4	Scope and field of application	7
A.3.2.5	References	7
A.3.2.6	Abbreviations	7
A.3.2.7	Document semantics	8
A.3.2.8	Abstract syntactic structure	8
A.3.2.9	Definition of transfer	8
A.3.2.10	ASE-specific operations	8

This page intentionally left blank

STANDARDSISO.COM : Click to view the full PDF of ISO 9069:1988

# Information Processing — SGML Support Facilities — SGML Document Interchange Format (SDIF)

## 1 Scope

This International Standard specifies a data structure known as the SGML Document Interchange Format (SDIF). SDIF enables a document conforming to ISO 8879, which might be stored in several entities, to be packed into a data stream for interchange in a manner that will permit the recipient to reconstitute the separate entities.

SDIF also allows related documents to be included in the data stream, such as covering letters, transmittal forms, catalog cards, formatting procedures, or the "document profile" required by a document architecture.

## 2 Field of application

The SGML Document Interchange Format shall be used solely for the interchange of SGML documents, as defined in ISO 8879, among SGML systems.

Interchange can be by means of data communications in Open Systems Interconnection or other environments, or by the exchange of storage media.

## 3 References

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

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

This International Standard uses formal syntax productions to define SDIF precisely. These are expressed in ASN.1, which is defined in the following International Standards:

ISO 8824, *Information processing systems — Open Systems Interconnection — Specification of Abstract Syntax Notation One (ASN.1)*.

ISO 8825, *Information processing systems — Open Systems Interconnection — Specification of basic encoding rules for Abstract Syntax Notation One (ASN.1)*.

## 4 Definitions

For the purposes of this International Standard, the definitions used in ISO 8879, and the following definitions, apply.

**4.1 data stream character set:** The character set used for all document markup in an SDIF data stream, and for all SDIF names and SDIF identifiers.

**4.2 document descriptor:** A component of an SDIF data stream that contains an SGML document entity and entity descriptors for all external entities that are directly or indirectly referenced from it.

**4.3 document profile:** A transmittal form, defined by a document architecture or application, that describes the characteristics of a conforming document that are relevant to its management or interchange.

**4.4 entity descriptor:** A component of an SDIF data stream that represents an external entity.

**4.5 main document:** An SGML document that is the subject of an interchange.

**4.6 omitted public text:** An entity descriptor that has no text at all; it represents public text that is part of the document, but is not being interchanged because the recipient is expected to have a copy.

**4.7 related document:** An SGML document that is used in conjunction with a main document, but is not referenced as part of its content.

NOTE — Some examples are:

- a covering letter;
- transmittal form;
- catalog card;
- formatting procedure;
- or document profile.

**4.8 SDIF:** SGML Document Interchange Format

**4.9 SDIF identifier:** A unique identifying sequence number assigned to an entity descriptor in the SDIF data stream.

NOTE — It corresponds to the position of the entity's external identifier. For example, the entity descriptor with the SDIF identifier "5" represents the entity identified by the fifth markup declaration in the data stream that has an external identifier.

**4.10 SDIF identifier reference:** An entity descriptor that contains no text of its own, but instead refers to the SDIF identifier of another entity descriptor. It represents a duplicate declaration of an entity that either has already occurred in the data stream, or that is omitted public text.

**4.11 SDIF name:** An SGML name assigned in the SDIF data stream to a document or to the data stream as a whole.

**4.12 SDIF packer:** A program that creates an SDIF data stream.

NOTE — The parameters to an SDIF packer usually include the SDIF name for the data stream, the SDIF names and system identifiers for the main and related documents, and the order in which the documents should be packed.

**4.13 SDIF unpacker:** A program that decomposes an SDIF data stream into its constituent entities.

NOTE — If necessary, the SDIF unpacker will modify the system identifier parameter of markup declarations to be consistent with storage addresses in its environment.

**4.14 SGML Document Interchange Format:** A data structure that enables a main document and its related documents, each of which might be

stored in several entities, to be combined into a single data stream for interchange in a manner that will permit the recipient to reconstitute the separate entities.

## 5 Description of the data stream

The SDIF data stream represents one or more SGML document entities, and zero or more SGML subdocument, SGML text, and data entities, as defined in ISO 8879.

### NOTES

1 These entities comprise a main document, and zero or more related documents, that are to be interchanged together. Each of the documents is subject to its own SGML declaration, whose concrete syntax definition, limits on subdocument nesting, and other restrictions apply.

2 The application-specific information parameter of the SGML declaration can be used to identify whether the document conforms to a particular document architecture.

### 5.1 Organization of the data stream

The data stream consists of a field that identifies the character set used for markup throughout the data stream, followed by document and entity descriptors that represent the main and related documents.

#### NOTES

1 The formal definition can be found in clause 6.

2 SDIF serves solely to identify and delimit the separate documents, and the entities into which the documents may be divided. No other information is represented directly in SDIF, except for the identification of the character set. In particular, SDIF has no information about the element structure of the documents, as this is found within the encapsulated SGML entities, which may also include layout and presentation information.

The document descriptors are in the order specified to the SDIF packer.

#### NOTES

1 See 5.5 for the order of the entity descriptors.

2 Documents that are not SGML documents, such as font definitions or formatting procedures, will require a simple SGML document as an identifying shell from which they can be referenced as data entities.

### 5.2 Data stream character set

This field contains one or more designating escape sequences for the data stream character set. The escape sequences shall be formulated in accordance with ISO 2022.

NOTE — It is usually the originating system character set.

### 5.3 SDIF name

The main document, related documents, and the data stream shall be assigned SDIF names when the data stream is created.

#### NOTES

1 They do not have SGML entity names, as they are not referenced from an SGML document.

2 A document's SDIF name is ordinarily derived from the document's system identifier in the originating system.

SDIF names shall conform syntactically to the rules of the SGML reference concrete syntax, and shall be coded in the data stream character set. No two SDIF names in the same data stream can be the same.

NOTE — An application or architecture can establish conventions for SDIF names just as it can for names within an SGML entity. For example, an architecture could require that the SDIF name "PROFILE" be used for the related document that contains document management attributes and similar information about the main document.

An SDIF name can be accompanied by an explanatory comment.

### 5.4 Document descriptor

#### 5.4.1 Included entities

A document descriptor contains an SGML document entity as defined in ISO 8879. The external identifiers in its markup declarations, and in declarations occurring in entities directly or indirectly referenced from it, determine the other entities for which entity descriptors will be included in the document descriptor.

#### NOTES

1 If an external entity is a subdocument, descriptors for the external entities declared within it are also included, and so on for any nested subdocuments.

2 Although a data entity contains no SGML markup, it could (in its own notation) contain references to other data entities or SGML subdocument entities. Entity descriptors for such entities are included in the data stream only if the entities are declared in some SGML entity that is included in the data stream.

No entity descriptor is included for the default entity.

NOTE — An effect of this sub-sub-clause is to require that there be an entity declaration for every entity that is to be included in the SDIF data stream, including entities referenced from data entities and entities whose references could otherwise have been defaulted.

### 5.4.2 Document character set

The document character set shall be the same as the data stream character set.

NOTE — A document character set is the only character set used for markup in a document. Any number of character sets may be used for data, and they need not be the same for all documents in the data stream.

### 5.5 SDIF identifier

NOTE — An SGML name is not sufficient for unique identification of an entity in the data stream because the same name can occur in more than one document or subdocument. Therefore, an SDIF identifier is assigned to each declared external entity in a manner that guarantees uniqueness and reflects the entity structure (in top-down, level-by-level order).

The sequence numbers for SDIF identifiers are coded in the data stream character set. They begin at 1, are incremented by 1, and are assigned in the following manner:

- a) A document descriptor is created for the first SGML document entity and its prolog is examined for external identifiers.

NOTE — The external identifiers can occur in entity, document type, link type, and notation declarations, but will be ignored in a default entity declaration.

Entity descriptors are created for them, and SDIF identifiers assigned, in the order the declarations are encountered, until the end of the prolog is reached.

The first SDIF identifier assigned in a prolog is recorded in the first identifier field. If none is assigned, the value recorded must be "0".

NOTE — The level-by-level traversal of the entity structure is an efficient way to pack the data stream. The SDIF unpacker, however, may prefer a left list traversal in order to assign system identifiers efficiently. The first identifier field permits such a traversal by identifying the children of a given document or subdocument entity.

- b) The newly created entity descriptors are examined for a subdocument entity. If one is found, its prolog is scanned as above, and entity descriptors are created for any external identifiers found.
- c) When the end of a subdocument entity prolog is reached, searching and sequential numbering resume in the prolog of the next encountered subdocument entity, if any.
- d) When the end of the data stream is reached, searching and sequential numbering continue in the prolog of the next SGML document entity, if any.

## 5.6 Entity descriptor

The type of entity descriptor is determined in the following manner:

- a) For the first occurrence of an external identifier in the data stream, if the entity text is to be interchanged, an entity descriptor is created that contains the actual text; that is, the descriptor will include an SGML subdocument entity, an SGML text entity, or a data entity.
- b) For the first occurrence of an external identifier of text that is not to be interchanged, the entity descriptor will have no text; that is, it will be an omitted public text descriptor. Such an external identifier shall be a public identifier.
- c) For occurrences of an external identifier after the first, the entity descriptor will be a cross-reference to the entity descriptor that represents the first occurrence of the external identifier; that is, it will be an SDIF identifier reference.

## 6 Abstract data stream format

```

ISO9069-SDIF DEFINITIONS ::= BEGIN

Data-stream ::= SEQUENCE {
    data-stream-character-set Character-set-designation,
    data-stream-name SDIF-name-and-description,
    related-documents-A [APPLICATION 8] IMPLICIT SEQUENCE OF
        Document-descriptor OPTIONAL,
    main-document [APPLICATION 9] IMPLICIT
        Document-descriptor,
    related-documents-B [APPLICATION 10] IMPLICIT SEQUENCE OF
        Document-descriptor OPTIONAL
}

Document-descriptor ::= [APPLICATION 11] IMPLICIT SEQUENCE {
    document-name SDIF-name-and-description,
    first-identifier SDIF-identifier,
    document-entity SGML-document-entity,
    external-entities SET OF Entity-descriptor OPTIONAL
}

Entity-descriptor ::= [APPLICATION 12] IMPLICIT SEQUENCE {
    entity-identifier SDIF-identifier,
    entity-name SGML-name,
    CHOICE {
        subdocument-structure Subdocument-structure,
        text-entity SGML-text-entity,
        data-entity External-data-entity,
        cross-reference SDIF-identifier,
        omitted-public-text NULL
    }
}

Subdocument-structure ::= [APPLICATION 13] IMPLICIT SEQUENCE {
    first-identifier SDIF-identifier,
    subdocument-entity SGML-subdocument-entity
}

```

```

Character-set-designation ::= [APPLICATION 0] IMPLICIT OCTET STRING
-- Coded character string as in ISO 2022:
-- designating escape sequences
External-data-entity ::= [APPLICATION 1] IMPLICIT OCTET STRING
-- Coded character or bit string
-- conforming to the data content
-- notation specified by the SGML
-- entity declaration.
SDIF-identifier ::= [APPLICATION 2] IMPLICIT OCTET STRING
-- Coded character string as in ISO 8879:
-- number
SDIF-name-and-description ::= [APPLICATION 3] IMPLICIT OCTET STRING
-- Coded character string as in ISO 8879:
-- name, followed optionally by a
-- description of the named object.
SGML-document-entity ::= [APPLICATION 4] IMPLICIT OCTET STRING
-- Coded character string as in ISO 8879:
-- SGML document entity
SGML-name ::= [APPLICATION 5] IMPLICIT OCTET STRING
-- Coded character string as in ISO 8879:
-- name
SGML-subdocument-entity ::= [APPLICATION 6] IMPLICIT OCTET STRING
-- Coded character string as in ISO 8879:
-- SGML subdocument entity
SGML-text-entity ::= [APPLICATION 7] IMPLICIT OCTET STRING
-- Coded character string as in ISO 8879:
-- SGML text entity

END

```

## 7 Encoding rules

An SDIF data stream shall be encoded in the manner required by the interchange medium or communication system. Where more than one encoding is permitted, the choice shall be made by agreement of the parties to the interchange.

NOTE — It is strongly recommended that a system conforming to this International Standard be able to interchange an SDIF data stream that is encoded in accordance with the ASN.1 basic encoding rules specified in ISO 8825.