



中华人民共和国国家标准

GB/T 15536—1995
ISO 9069—1988

信息处理 SGML 支持设施 SGML 文件交换格式(SDIF)

Information processing—SGML support facilities
—SGML document interchange format(SDIF)

1995-04-05 发布

1995-12-01 实施

国家技术监督局 发布

中华人民共和国国家标准

信息处理 SGML 支持设施 SGML 文件交换格式(SDIF)

GB/T 15536—1995
ISO 9069—1988

Information processing—SGML support facilities
—SGML document interchange format(SDIF)

本标准等同采用国际标准 ISO 9069—1988《信息处理 SGML 支持设施 SGML 文件交换格式(SDIF)》。

1 主题内容

本标准规定了一种称为“SGML 文件交换格式(SDIF)”的数据结构。SDIF 能使符合于 GB/T 14814 标准且可能存储在若干实体中的文件以某种方式组合到一个用于交换的数据流中,以便其接收者能够重新构造出构成该文件的各个实体。

SDIF 也允许在数据流中包含相关的文件,如附函、传输方式、目录卡片、格式化过程以及文件结构需要的“文件轮廓”。

2 适用范围

SGML 文件交换格式只用于按 GB/T 14814—93 定义的 SGML 系统之间的 SGML 文件交换。数据交换可以通过开放系统互连或其他环境的数据通信进行,也可以通过交换存储媒体进行。

3 引用标准

GB 2311 信息处理 七位和八位编码字符集代码扩充技术

GB/T 14814 信息处理 文本和办公系统 标准通用置标语言(SGML)

本标准使用形式语法产生式精确定义 SDIF。所有的语法产生式是用 ASN.1 表示的,ASN.1 在下列国际标准中定义:

ISO 8824 信息处理系统 开放系统互连 抽象语法标记法 1(ASN.1)规范

ISO 8825 信息处理系统 开放系统互连 抽象语法标记法 1(ASN.1)的基本编码规范

4 术语

GB/T 14814 中给出的术语和下列术语适用于本标准。

4.1 数据流字符集 data stream character set

在 SDIF 数据流中的所有文件置标使用的字符集以及所有 SDIF 名字和 SDIF 标识符使用的字符集。

4.2 文件描述符 document descriptor

数据流的组成部分,该数据流包含一个 SGML 文件实体以及所有直接或间接引用的外部实体的实体描述符。

国家技术监督局 1995-04-05 批准

1995-12-01 实施

4.3 文件轮廓 document profile

由文件结构或应用定义的传输格式,描述了和文件管理或交换有关的符合于文件的特性。

4.4 实体描述符 entity descriptor

SDIF 数据流的组成部分,它表示了一个外部实体。

4.5 主文件 main document

一个 SGML 文件,它是一个交换的主题。

4.6 省略公用文本 omitted public text

完全不含有正文的实体描述符,它说明此公用文本是文件的一部分,但是,由于认为接受者手中拥有备份,因此它并没有被交换。

4.7 相关的文件 related document

用来连接一个主文件的 SGML 文件,但并不做为它的部分内容被引用。

注:一些例子是:

- 附函;
- 传输格式;
- 目录卡片;
- 格式化过程;
- 文件轮廓。

4.8 SDIF

SGML 文件交换格式。

4.9 SDIF 标识符 SDIF identifier

在 SDIF 数据流中赋给实体描述符的唯一标识性顺序号。

注:它相当于实体的外部标识符的位置。例如:一个具有 SDIF 标识符“5”的外部实体表示在数据流的第五个置标声明中所标识的有外部标识符的实体。

4.10 SDIF 标识符引用 SDIF identifier reference

一种实体描述符,它不包含其自身的正文,但含有对其他实体描述符的 SDIF 标识符的引用。它表示在数据流中已出现的实体的重复声明或表示省略公用文本。

4.11 SDIF 名字 SDIF name

在 SDIF 数据流中对文件或完整数据流所指定的 SGML 名字。

4.12 SDIF 组合程序 SDIF packer

建立 SDIF 数据流的程序。

注:SDIF 组合程序的参数一般包括数据流的 SDIF 名字、主文件和相关文件的 SDIF 名字、系统标识符以及文件组合的顺序。

4.13 SDIF 分解程序 SDIF unpacker

将 SDIF 数据流分解为构成它的各个实体的程序。

注:如果必要,SDIF 分解程序将修改置标声明的系统标识符参数以与其环境中的存储地址相一致。

4.14 SGML 文件交换格式 SGML document interchange format

一种数据结构,能够使分别存储在若干实体中的主文件及其相关文件在交换时按某种方式组合成单一数据流,以便接收者重新构造出各个实体。

5 数据流描述

SDIF 数据流表示一个或多个 SGML 文件实体,零个或多个 SGML 子文件实体,零个或多个 SGML 正文实体和数据实体,它们都在 GB/T 14814 中定义。

注:① 这些实体包含一个主文件和零个或多个相关文件,它们同时被交换。每个文件受限于它自身的 SGML 声明。即它的具体语法定义,子文件嵌套限制以及其他限制。

② SGML 声明中的应用规定的信息参数用于标识文件是否符合特定的文件结构。

5.1 数据流的构成

数据流由标识数据流中置标所用字符集的字段、表示主文件和相关文件的文件描述符及实体描述符组成。

注：① 形式定义见第 6 章。

② SDIF 仅用于标识和定界各个文件以及文件可以分解成的实体。除了标识字符集之外，SDIF 中不直接表示其他信息。具体而言，SDIF 没有关于文件元素结构的信息，由于这些信息出现在 SGML 实体中，所以实体中还包括版面格式和显示信息。

文件描述符依照 SDIF 组合程序所指定的顺序排列。

注：① 有关实体描述符的顺序见 5.5 条。

② 不是 SGML 文件的文件，例如字型定义和格式化过程，需要一个简单的 SGML 文件作为标识外壳。这种文件可以通过标识外壳作为数据实体被引用。

5.2 数据流字符集

该字段包括数据流字符集的一个或多个指定的转义序列。转义序列应按 GB 2311 构造。

注：它通常是初始的系统字符集。

5.3 SDIF 名字

当建立数据流时，对主文件、相关文件和数据流应指定 SDIF 名字。

注：① 它们没有 SGML 实体名字。因为它们不会在 SGML 文件中被引用。

② 文件的 SDIF 名字一般从初始系统中文件的系统标识符得到。

SDIF 名字在语法上应遵守 SGML 具体语法的规则，并用数据流字符集编码。在同一数据流中任何两个 SDIF 名字都是不相同的。

注：应用或结构可以建立 SDIF 名字的约定，如同对 SGML 实体的名字一样。例如：一个结构要求 SDIF 名字“PROFILE”用于包含文件管理属性和类似主文件信息的相关文件。

SDIF 名字可以附有说明性注释。

5.4 文件描述符

5.4.1 包含实体

文件描述符包含 GB/T 14814 定义的文件实体。在置标声明中出现的外部标识符，以及出现在由其直接或间接引用的实体内声明中的外部标识符，决定了其他实体。对于这些实体，其描述符包含在文件描述符中。

注：① 如果外部实体是子文件，则其中声明的外部实体的描述符也应被包含，而且对于任何嵌套的子文件也应如此。

② 尽管数据实体不包含 SGML 置标，但它（以其自己的记法）包含对其他数据实体或 SGML 子文件实体的引用。仅当实体在数据流所包含的某些 SGML 实体中被声明时，其实体描述符才包含在数据流中。

默认实体不包含实体描述符。

注：本条目的作用在于要求包含在数据流中的每个实体都应有一个实体声明，包括由数据实体引用的实体以及已经缺省引用的实体。

5.4.2 文件字符集

文件字符集应和数据流字符集一致。

注：文件字符集是用于文件置标的唯一字符集。对于数据可以使用若干字符集，且数据流中所有文件使用的不必是同一个字符集。

5.5 SDIF 标识符

注：就唯一标识数据流中的实体而言，SGML 名字是不够充分的，因为在多个文件或子文件中可能出现相同的名字。因此，必须按某一方式对每个声明的外部实体指定一个 SDIF 标识符，以保证唯一性并反映实体结构（自顶向下，逐层顺序）。

SDIF 标识符的序号按数据流字符集编码。它们从 1 开始，增量为 1，并且按下列方式指定：

a. 对第一个 SGML 文件实体建立文件描述符,并针对外部标识符检查其前导说明部分。

注:外部标识符可出现在实体、文件类型、链接类型和记法声明中,但在默认的实体声明中将被忽略。

按照实体声明出现的顺序,对实体建立实体描述符并指定 SDIF 标识符,直到前导说明部分结束。

第一个标识符字段中记录前导说明部分中赋予的第一个 SDIF 标识符。如果没有赋予标识符字段,则记录的值必须是“0”。

注:实体结构的逐层遍历是组成数据流的有效方法。然而 SDIF 分解程序可能采用左列表遍历方法以有效地指定系统标识符。通过识别给定文件或子文件实体的子女,可以在第一个标识符字段上进行这种遍历。

b. 对于子文件实体,检查新建立的实体描述符。如果遇到一个,则它的前导说明部分按以上方式扫描并且对遇到的每一个外部标识符建立实体描述符。

c. 当到达子文件实体前导说明部分的尾部时,如果存在下一个子文件实体,则在其前导说明部分中继续搜寻并顺序编号。

d. 当到达数据流尾部时,如果存在下一个 SGML 文件实体,则在其前导说明部分中继续搜寻并顺序编号。

5.6 实体描述符

实体描述符的类型按下列方式确定:

a. 对数据流中首次出现的外部标识符,如果要交换实体正文,则建立包含实际正文的实体描述符,即描述符包括一个 SGML 子文件实体、一个 SGML 正文实体或一个数据实体。

b. 对首次出现而又不交换正文的外部标识符,实体描述符不包含正文。即它是省略的公用文本描述符,这种外部标识符应是一个公用标识符。

c. 对非首次出现的外部标识符,实体描述符是对表示首次出现的外部标识符的实体描述符的交叉引用,即它是一个 SDIF 标识符引用。

6 抽象数据流格式

GB/T 14814—SDIF 定义 ::=BEGIN

数据流 ::=SEQUENCE{

数据流字符集 字符集指定,

数据流名 SDIF 名字和描述,

相关文件组 A [APPLICATION 8] IMPLICIT

SEQUENCE OF

文件描述符 OPTIONAL,

主文件 [APPLICATION 9] IMPLICIT

文件描述符,

相关文件组 B [APPLICATION 10] IMPLICIT SEQUENCE OF

文件描述符 OPTIONAL

}

文件描述符 ::= [APPLICATION 11] IMPLICIT SEQUENCE{

文件名 SDIF 名字和描述,

第一标识符 SDIF 标识符,

文件实体 SGML 文件实体,

外部实体 SET OF 实体描述符 OPTIONAL

}

实体描述符 ::= [APPLICATION 12] IMPLICIT SEQUENCE{

实体标识符, SDIF 标识符,

实体名	SGML 名字, CHOICE{
子文件结构	子文件结构,
正文实体	SGML 正文实体,
数据实体	外部数据实体,
交叉引用	SDIF 标识符,
省略公用文本	NULL }
子文件结构	::=[APPLICATION 13] IMPLICIT SEQUENCE{
第一标识符	SGML 标识符,
子文件实体	SGML 子文件实体 }
字符集指定	::=[APPLICATION 0] IMPLICIT OCTET STRING ——如同 GB 2311 中的代码字符串; ——指定转义序列
外部数据实体	::=[APPLICATION 1] IMPLICIT OCTET STRING ——与 SGML 实体声明指定的数据内容 ——记法一致的代码字符或位串
SDIF 标识符	::=[APPLICATION 2] IMPLICIT OCTET STRING ——如同 GB/T 14814 中的代码字符串:数
SDIF 名字和描述	::=[APPLICATION 3] IMPLICIT OCTET STRING ——如同 GB/T 14814 中的代码字符串:名字 ——其后可跟有一描述命名对象的选项
SGML 文件实体	::=[APPLICATION 4] IMPLICIT OCTET STRING ——如同 GB/T 14814 中的代码字符串; ——SGML 文件实体
SGML 名字	::=[APPLICATION 5] IMPLICIT STRING ——如同 GB/T 14814 中的代码字符串; ——名字
SGML 子文件实体	::=[APPLICATION 6] IMPLICIT OCTET STRING ——如同 GB/T 14814 中的代码字符串; ——SGML 子文件实体
SGML 正文实体	::=[APPLICATION 7] IMPLICIT OCTET STRING ——如同 GB/T 14814 中的代码字符串; ——SGML 正文实体
END	

7 编码规则

SDIF 数据流应按照交换媒体或通信系统要求的方式编码。当允许多种编码时,这种选择应得到交换参与者的同意。

注:符合于本标准的系统能够对按照 ISO 8825 规定的 ASN.1 基本编码规则编码的数据流进行交换。

附 录 A
本标准和其他标准的关系
(参考件)

A1 信息处理标准

按照本标准构造的符合 SGML 的文件能够用 MOTIS 功能传送。MOTIS 功能在下列标准以及类似标准和 CCITT 建议中定义:

ISO/IEC 10021 信息处理系统 文本通信 面向文本交换系统的消息

A2 办公文件体系结构(ODA)

对符合 ODA 文件的 SGML 表示的文件,SGML 文件交换格式是与开放系统交换相一致的。ODA 文件在以下标准中定义:

ISO 8613 信息处理 文本和办公系统 办公文件体系结构(ODA)和交换格式

A3 OSI 表示服务**A3.1 对象标识符****A3.1.1 抽象语法**

本标准将 ASN.1 对象标识符值:

{iso standard 9069 abstract syntax}

设置为表示数据值集合的抽象语法名,其中每一个值都是在第 6 章中定义的 ASN.1 类型“GB/T 14814—SDIF 数据流”的值。对应的对象描述符值是:

“SDIF basic abstract syntax”

注:在每一种表示的数据值中,仅支持部分 SDIF 语义传输的抽象语法的说明已超出本标准的范围。

A3.1.2 传输语法

ASN.1 对象标识符:

{joint-iso-ccitt asn1 (1) basic-encoding (1)} .

和对象描述符值:

“Basic Encoding of a single ASN.1 type”

(指定给 ISO 8825 中的信息对象)能被用作代表这种抽象语法名的传输语法名。它们标识了用 ISO 8825 中说明的基本编码规则得到的传输语法。

注:对其他传输语法的描述和命名超出本标准的范围。

A3.2 OSI 文件类型定义

本条标识了适用于在 FTAM(ISO 8571)和 JTM(ISO 8832)传输中携带 SDIF 语义的 OSI 文件类型。

注:① 本条使用了术语“文件类型”,该术语只与开放系统互连相关,而不同于 GB/T 14814 中定义的“文件类型”。

② 在 OSI 文件类型登记管理机构建立以前,对于 OSI 文件类型的定义,下面小标题与在其他 OSI 标准中使用的相同。

A3.2.1 标识符

ASN.1 对象标识符值:

{iso standard 9069 document-type (2) basic (1)}

能用于标识本文件的类型。

A3.2.2 描述符值

ASN.1 对象描述符值:

“SDIF basic document type”

能用于描述本文件的类型。

A3.2.3 参数语法

参数不和 SDIF 基本文件类型一起使用。

A3.2.4 范围和应用领域

SDIF 基本文件类型定义了文卷内容,该文卷用于存储由 FTAM 存储、传输和访问以及由 JTM 传输的文卷内容。

注: SDIF 基本文件类型既不允许通过 FTAM 访问部分文件,也不允许在文件传输过程中通过 FTAM 插入检测点,它只支持为读写访问而对整个文件的非检测点传输。

A3.2.5 引用标准

ISO 8571 信息处理系统 开放系统互连 文卷传输、访问和管理

ISO 8832 信息处理系统 开放系统互连 作业传输和管理的基本类协议的说明

A3.2.6 缩略语

FTAM 文件传输、访问和管理

JTM 作业传输和操作

A3.2.7 文件语义

文件语义在本标准和 GB/T 14814 中详细说明。

A3.2.8 抽象语法结构

SDIF 信息有一种抽象语法结构,它反映了将一个或多个相关 SGML 文件构造成如在 GB/T 14814 中说明的若干实体。

注: 实体可以通过 SGML 置标进一步结构化成为可访问的单个元素和属性,这些元素和属性在 GB/T 14814 中说明。这种称为“元素结构”附加的语法结构在 SDIF 基本文件类型中是不可见的,只有实体结构在载体标准中是可见的。

A3.2.9 传输定义

传输由一数据值构成,该数据值是 ASN.1 数据类型“GB/T 14814—SDIF 数据流”的(完全)值。该数据流具有 A3 条指定的抽象语法名。建议所有实现都支持的传输语法为:

{joint-iso-ccitt asn 1 (1)basic-encoding (1)}

但也可支持其他传输语法。

A3.2.10 ASE 专用操作

FTAM 专用操作和 JTM 专用操作不适用于本文件类型。

注: 特别是没有定义文件类型与其自身的连接。

附加说明:

本标准由中华人民共和国电子工业部提出。

本标准由电子工业部标准化研究所归口。

本标准由西安交通大学、电子工业部标准化研究所负责起草。

本标准主要起草人邓良松、侯迪、陈堃铎、冯惠、邵坚。

中 华 人 民 共 和 国
国 家 标 准
信息处理 SGML 支持设施
SGML 文件交换格式(SDIF)
GB/T 15536—1995

*

中国标准出版社出版
北京复兴门外三里河北街16号

邮政编码:100045

电 话:8522112

中国标准出版社秦皇岛印刷厂印刷
新华书店北京发行所发行 各地新华书店经售

版权专有 不得翻印

*

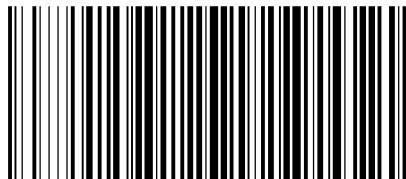
开本 880×1230 1/16 印张 3/4 字数 14 千字
1995 年 11 月第一版 1995 年 11 月第一次印刷
印数 1—1 000

*

书号: 155066·1-11958 定价10.00 元

*

标 目 275—49



GB/T 15536—1995