

电子文件管理元数据描述方式研究^{*}

张正强

摘 要 从语言学角度看,电子文件管理元数据的描述有三种方式,即语义描述、语法描述和本体描述,这三种方式在描述电子文件管理元数据方面各有异同。采用 ISO 11179 来描述电子文件管理元数据语义,如在语法上采用 XML 来描述电子文件管理元数据,只能在形式上实现电子文件管理元数据的置标机读化描述,而难以在语义层面上对其进行置标智能化描述;采用本体语言来描述电子文件管理元数据,就可以实现电子文件管理元数据语义描述与语法描述的融合集成,实现电子文件管理元数据置标智能化的描述。图 14。表 3。参考文献 5。

关键词 电子文件管理元数据 语义描述 语法描述 本体描述

分类号 G254

ABSTRACT This paper points out that from the linguistic aspect, there are three kinds of description manners for electronic record management metadata, which are semantic description, syntactical description and ontological description. They have similarities and differences. If an electronic record management metadata is described based on ISO 11179, and then described by XML, then syntactical description can only be achieved without semantic description. But if it is described in ontological language, not only syntactical description can be achieved, but also semantic description can be done, realizing the intelligent markup of electronic record management metadata. 14 figs. 3 tabs. 5 refs.

KEY WORDS Electronic record management metadata. Semantic description. Syntactical description. Ontological description.

CLASS NUMBER G254

在电子文件管理领域,随着电子文件管理元数据国家标准的即将颁布与电子文件管理元数据国家行业标准的即将出台,一个摆在我们面前迫切需要研究的问题就是如何对电子文件管理元数据进行规范、有效、全面的描述,从而最大限度地提高电子文件管理元数据的标准化程度及其功能,以充分发挥其在电子文件管理中的作用。

1 基于语言学角度的电子文件管理元数据描述方式的划分

1.1 电子文件管理元数据描述的基本涵义

电子文件管理元数据的描述,就是对电子文件管理元数据的表述与揭示。对于电子文件

管理元数据可以从不同角度加以研究,但是从语言学角度来看,电子文件管理元数据就是一种语言,且是一种人工语言,因此,从语言学角度对电子文件管理元数据进行研究是一个重要的视角。电子文件管理元数据既然是一种语言,就存在着语义与语法这两个基本方面。由此,电子文件管理元数据也就存在着从语义方面和语法方面对元数据进行表述与揭示的问题。

1.2 电子文件管理元数据语义描述与语法描述相分离的描述方式

在自然语言中,语义主要是指语词的意义,是思维理解的内容,而语法主要是指构词成句的规则,是语言构造的形式,语义与语法处于两

^{*} 本文系国家社科基金重点资助项目“基于 XML 电子文件管理元数据标准研究”(项目批准号:04ATQ002)系列研究成果之一。

个不同的层面。自然语言的语义与语法的区分可以用经典的语义三角模型来表示(图1)。

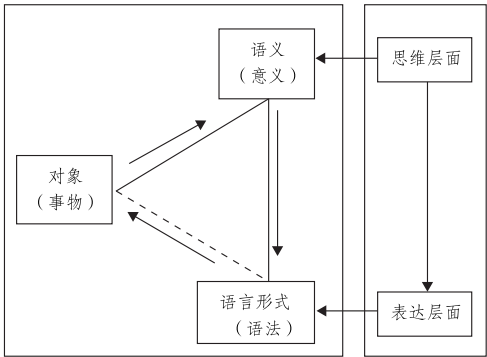


图1 自然语言语义三角模型

从图1看出,语义处于思维层面,而语法存在于表达层面,但不管怎样,都是人们用于说明、描述、指代事物所不可缺少的两个基本方面。

由于电子文件管理元数据是一种人工语言,而且是以计算机网络为应用环境,因此,在语义与语法的界定上又与自然语言语义、语法的界定有所不同。电子文件管理元数据语义是指人们对其可以直接读取的内容涵义部分;而语法则是指计算机对其可以直接读取的置标结构部分,其模型见图2。

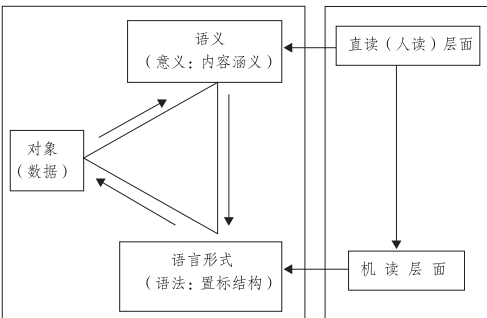


图2 电子文件管理元数据语义与语法相分离三角模型

从图2看出,电子文件管理元数据的语义是对人们直接读取而言的,而电子文件管理元数据

的语法是对计算机读取而言的。显然,没有这两方面,就很难实现电子文件管理元数据在计算机网络环境下的互操作。由此,我们根据图2模型可以划分出电子文件管理元数据描述的两种类型,即电子文件管理元数据的语义描述类型与语法描述类型。在此,我们看到了电子文件管理元数据语义描述与语法描述相分离的两种描述方式。

语言的功能在于交流。虽然从形式上看,电子文件管理元数据的语义与语法与自然语言的语义与语法划分有所不同,但其实质是一样的,都是为了实现交流。只不过自然语言的语义与语法这种划分适用于人们交流的现实世界,而电子文件管理元数据的语义与语法的划分适用于人们在计算机网络环境下交流的虚拟世界。

1.3 电子文件管理元数据语义描述与语法描述相融合的描述方式

根据对立统一法则,任何事物都不是孤立存在的。既然对电子文件管理元数据的描述存在着语义描述与语法描述分离的现象,反之,也就必然存在着将语义描述与语法描述融合集成为一体的描述。那么,在电子文件管理元数据领域,能将语义描述与语法描述融合集成为一体的描述是什么呢?这就是本体描述。本体的范围很广,但对于电子文件管理元数据的描述而言,本体主要指本体语言,如 RDF^[1] (资源描述框架)、SKOS^[2] (简约知识体系)和 OWL^[3] (Web 本体语言)等。本体描述的模型见图3。

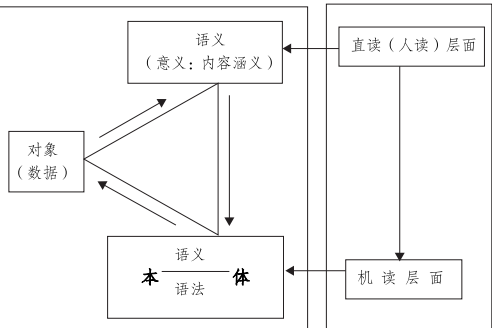


图3 电子文件管理元数据语义与语法相融合三角模型

从图3看出,采用本体语言对电子文件管理元数据进行描述,其最大特点就是在机读层面上将语义描述与语法描述融合集成在一起,使计算机既可识别电子文件管理元数据的语法,又可理解电子文件管理元数据的语义,从而形成对电子文件管理元数据置标智能化的描述,即本体描述。

在本体描述中,RDF作为本体语言在电子文件管理元数据描述方面的运用,主要是其平面三元组模型,该模型为电子文件管理元数据的描述提供了统一而又规范的抽象语法,可以达到对相同性质的元数据描述的一致性,使元数据之间关系表达更为明确。RDF的本体性主要体现在“RDF Schema(即RDF模式,以下简称RDFs)”上,因为其专门提供了在概念层次上描述语义的有关“类”和“属性”的一套词汇^[3]。

SKOS是RDFs在概念体系上的运用,是专门用于描述语义体系的“RDFs”。在本体描述中,SKOS作为本体语言在电子文件管理元数据描述方面的运用,主要是提供了一套系统而又规范的描述电子文件管理元数据语义的置标机读化词汇,以实现电子文件管理元数据语义置标的规范化与机读化。

OWL是在RDFs基础上进一步扩展了RDFs所表达的概念语义范围,从而形成了更强大的语义描述力。因此,在本体描述中,OWL作为本体语言在电子文件管理元数据方面的运用,主要是提供了更为丰富的描述语义的词汇,并且依据这些词汇,对电子文件管理元数据描述所构成的语义体系计算机可进行智能逻辑推理。

需要指出的是,采用本体语言对电子文件管理元数据进行描述,并不是采用其中一种本体语言进行描述,而是综合地、有机地采用RDFs、SKOS和OWL这三种本体语言来共同实现对电子文件管理元数据的置标智能化描述。

2 电子文件管理元数据的语义描述

电子文件管理元数据的语义描述是指对电子文件管理元数据采用直读(人读)可理解方式

所进行的语义显性化的描述,对电子文件管理元数据进行语义描述是建立“电子文件管理元数据规范(标准)”的基础。为了使电子文件管理元数据在语义理解上避免歧义、多义,从而使其具有在语义理解上的规范性与互操作性,就必须对电子文件管理元数据在语义层次上进行逐一显性化的描述,此类描述在电子文件管理元数据领域通常是根据国际标准《ISO/IEC11179-3:2003 信息技术-元数据注册-第3部分:注册元模型与基本属性》^[4](以下简称ISO 11179)来进行的。ISO 11179对元数据在语义上如何进行显性化描述做了详细规定。在ISO 11179中对电子文件管理元数据语义的描述主要分成三大类:

(1)对电子文件管理元数据元素的语义标识和语义内容方面的描述,如对电子文件管理元数据元素的标识符、优先标识、同义标识、定义、必备性、用途、注释、定义来源、提出时间、修改时间及版本等所做出的描述(表1)。

表1 基于ISO 11179对电子文件管理元数据元素在语义标识、语义内容方面的描述

类别	子类	具体描述
对元数据元素的语义标识和语义内容方面的描述	语义标识	标识符
		优先标识
		同义标识
		提出时间
		修改时间
		版本
	语义内容	定义
		必备性
		用途
		注释
		定义来源

根据表1,我们可以从语义标识和语义内容两个方面对一个具体的电子文件管理元数据元素进行描述,如对“题名”元素进行描述,其结果见图4。

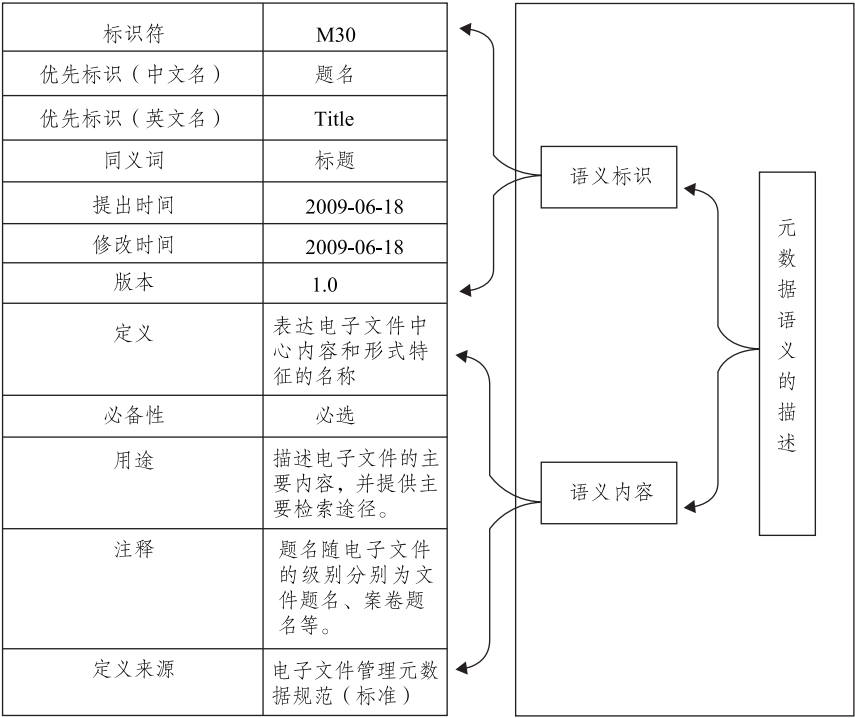


图 4 基于 ISO 11179 对电子文件管理元数据“题名”元素的语义标识和语义内容方面描述示例

(2)对电子文件管理元数据元素语义关系方面的描述,如对元数据元素之间的纵向语义关系和横向语义关系的显性化描述,其中横向语义关系还包括对相关元数据标准之间的元素的映射关系(表 2)。

表 2 基于 ISO 11179 对电子文件管理元数据元素语义关系方面的描述

类别	子类	具体描述
对元数据元素的语义关系方面的描述	纵向语义关系的描述	子元素
		子元素
		
	横向语义关系的描述	相关元素
		相关元素
		
		相关映射
		相关映射
		

根据表 2,我们从纵向语义关系和横向语义关系两个方面对一个具体的电子文件管理元数

据元素进行描述,如对“题名”这一元素进行语义关系显性化描述(图 5)。

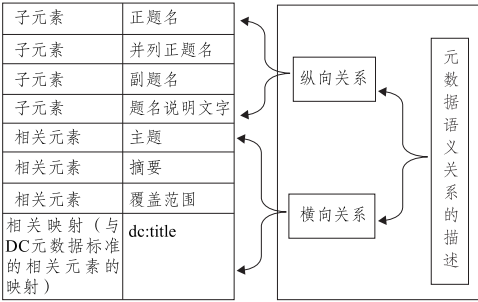


图 5 基于 ISO 11179 对电子文件管理元数据“题名”元素语义关系方面的描述示例

(3)对电子文件管理元数据元素语义值的显性化描述,如对值表示形式、值限定词、值类型、允许值、最小值、最大值、值涵义、值定义域的描述等(表 3)。

表 3 基于 ISO 11179 对元数据语义值方面的描述

类别	具体描述
对元数据语义值描述	值表示形式
	值限定词
	值类型
	允许值
	最小值
	最大值
	值涵义
	值定义域

根据表 3,我们再以“题名”这一元素为例,对其语义值进行具体的描述(图 6)。

值表示形式	文本
值限定词	《电子文件题名拟制规定 [电字(062007)]》
值类型	字符串
允许值	n/a
最小值	1
最大值	1
值涵义	n/a
值定义域	文件

元数据语义值的描述

图 6 基于 ISO 11179 对电子文件管理元数据“题名”元素的语义值描述示例

从图 4 - 图 6 看出,基于 ISO 11179 对电子文件管理元数据语义描述的三大类内容,其覆盖范围十分全面,且具体而详尽。不仅如此,从以上语义描述的具体内容还看出,如果电子文件管理元数据元素经这样描述后,由其所构成的“电子文件管理元数据规范(标准)”——“电

子文件管理元数据 Scheme”,就形成了一个元数据语义体系。因为,基于 ISO 11179 描述元数据语义时,不仅描述了元数据元素自身的语义,还描述了元数据元素之间的语义关系,由此构成的“电子文件管理元数据规范(标准)”就必然是一个既具有纵向语义关系,又具有横向语义关系的语义网,即语义体系,也就是在概念层面上形成了一个元数据概念系统。

需要指出的是,目前在电子文件管理元数据领域,在进行电子文件管理元数据语义描述时,虽然是根据 ISO 11179 来进行的,但不全面,实际上仅仅是从 ISO 11179 中选取若干项来进行描述。这有许多原因,但没有从语义角度考虑如何对电子文件管理元数据进行全面描述是其中关键的一点。如果能遵循 ISO 11179 对电子文件管理元数据语义进行如上所述的全面描述,对具体的元素而言,可以保证电子文件管理元数据语义描述的规范性和全面性,从而保证电子文件管理元数据语义描述的系统性;对元素之间的相互关系而言,不仅可以保证电子文件管理元数据语义之间的纵向等级语义关系的建立,而且可以保证电子文件管理元数据语义之间的横向相关语义关系的建立,从而保证电子文件管理元数据语义关系描述的系统性。

由此可见,从语义角度对电子文件管理元数据进行基于 ISO 11179 的语义描述,是对电子文件管理元数据语义进行全面、规范描述的关键,也是为什么要从语义角度研究电子文件管理元数据描述的原因之所在。

3 电子文件管理元数据的语法描述

电子文件管理元数据的语法描述,是指对电子文件管理元数据采用计算机可识别的方式所进行的置标结构化描述。遵循 ISO 11179 对电子文件管理元数据的语义进行描述,尽管可以建立起元数据的语义体系,或者说在逻辑层面上可以建立起元数据的概念系统,但是这种概念语义体系还只能供直读(人读),也就是只解决了电子文件管理元数据在计算机网络中跨平台交换时处于不同时间、不同地点的用户对

元数据语义理解的准确性与一致性问题——语义互操作问题,而没有解决在计算机网络中跨平台交换时机器对元数据读取的准确性与一致性问题——语法互操作性的问题。

为了实现电子文件管理元数据在计算机网络中跨平台交换时使机器具有对元数据读取的准确性与一致性,就必须将其映射为机器可读取的语言,这种映射最常见的做法就是对电子文件管理元数据进行置标,即采用置标结构在语法层面上对电子文件管理元数据进行描述。对电子文件管理元数据的置标有多种方式,其中普遍的是采用 XML(可扩展置标语言)以串行序列化的语法将“电子文件管理元数据规范”转写成“电子文件管理元数据 XML Schema,即电子文件管理元数据 XML 模式”。

如前所述,由于在电子文件管理元数据领域对电子文件管理元数据的语义描述虽然是根据

ISO 11179 来进行的,但只是从 ISO 11179 中选取了若干项来进行描述,很不全面。因此,为了全面地说明采用 XML 对电子文件管理元数据的置标结构,我们仍以前面所论述的基于 ISO 11179 语义描述的三大类内容为基础,以电子文件管理元数据“题名”元素为例,用 XML Schema 在语法上对该元数据元素进行置标描述。

(1) 用 XML Schema 对电子文件管理元数据“题名”元素的语义标识和语义内容方面进行语法描述(图 7)。

从图 7 看出,对电子文件管理元数据“题名”元素的语义标识和语义内容方面的描述,在 XML Schema 序列化语法中主要是通过注解“<xs:annotation>”这一置标标识符下,用“<em:property>”、“<em:administration>”这两组属性来描述的,如用“<em:identifier>”来表达“标识符”、用“<em:name xml:lang=“chn”>”

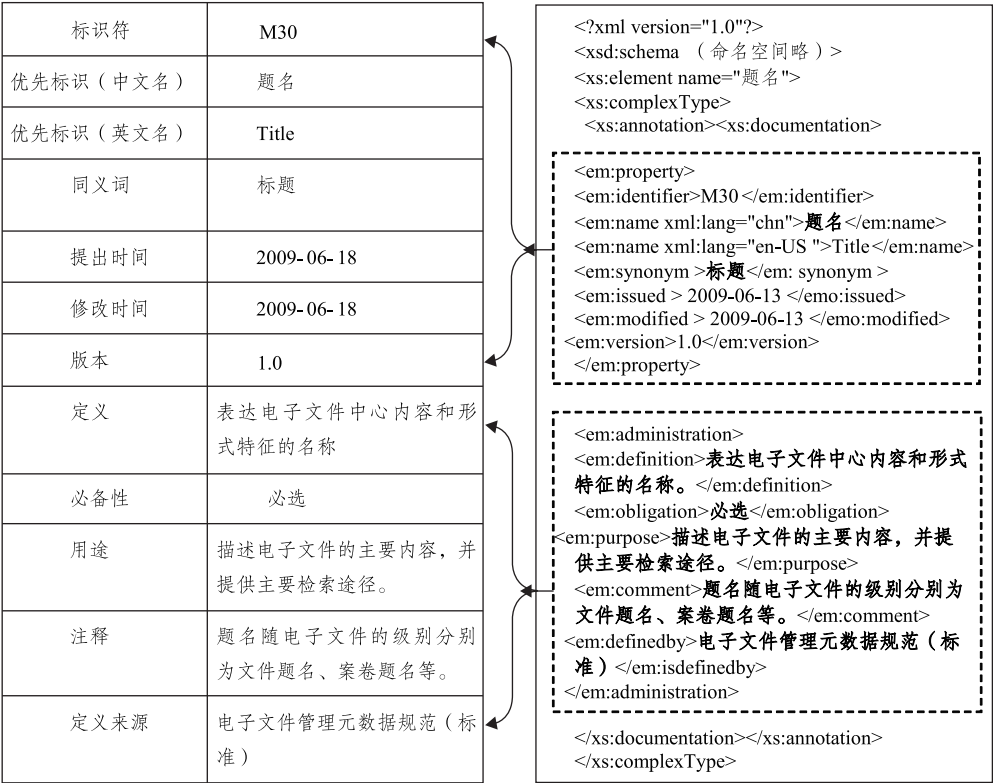


图 7 用 XML Schema 对电子文件管理元数据“题名”元素的语义标识和语义内容方面的语法描述示例

来表达“题名(中文)”、用“`<em:definition>`”来表达“定义”、用“`<em:version>`”来表达“版本”等。但通过“`<xs:annotation>`”和其下一系列置标标识符所表达的注解,其主要用途是为整个“电子文件管理元数据 XML 模式”添加注文,是供直读(人读)的信息。因此,尽管是用 XML Schema 序列化语法加以描述了,但不属于 XML Schema 定义的信息,计算机并不能理解这些信息的语义,即计算机只能在语法上识别它,而不能从语义上理解它。

此外,从图 7 还看出,在用 XML Schema 序列化语法对电子文件管理元数据“题名”元素的语义标识和语义内容的描述中,其置标标识符如 `<em:identifier>`、`<em:definition>`、`<em:version>` 等的命名主要是由“电子文件管理元数据 XML 模式”的设计者自行规定,因为这些置标标识符的命名在 XML Schema 的元规范^[2]中都没有规定。由此可见,不同设计者在对相同

的“语义标识和语义内容”的 XML Schema 置标标识符的命名上会不尽相同,这就影响了其规范化程度的提高。

(2) 用 XML Schema 对电子文件管理元数据“题名”元素之间的纵向语义关系和横向语义关系在语法上进行描述(图 8)。

从图 8 看出,在 XML Schema 序列化语法中对电子文件管理元数据“题名”元素的纵向语义关系描述上,如与“正题名、并列正题名、副题名、题名说明文字”等这些表示子元素等级关系的描述上,是通过 XML Schema 复杂类型内容模型来实现的,也就是说它是通过置标标识符“`<xs:complexType>`”下的元素组件来定义的。然而尽管在“`<xs:complexType>`”置标标识符下声明了“正题名、并列正题名、副题名、题名说明文字”这些子元素,但计算机也只能是识别它们,而不能理解它们。因为在 XML Schema 序列化语法中并没有定义“正题名、并列正题名、

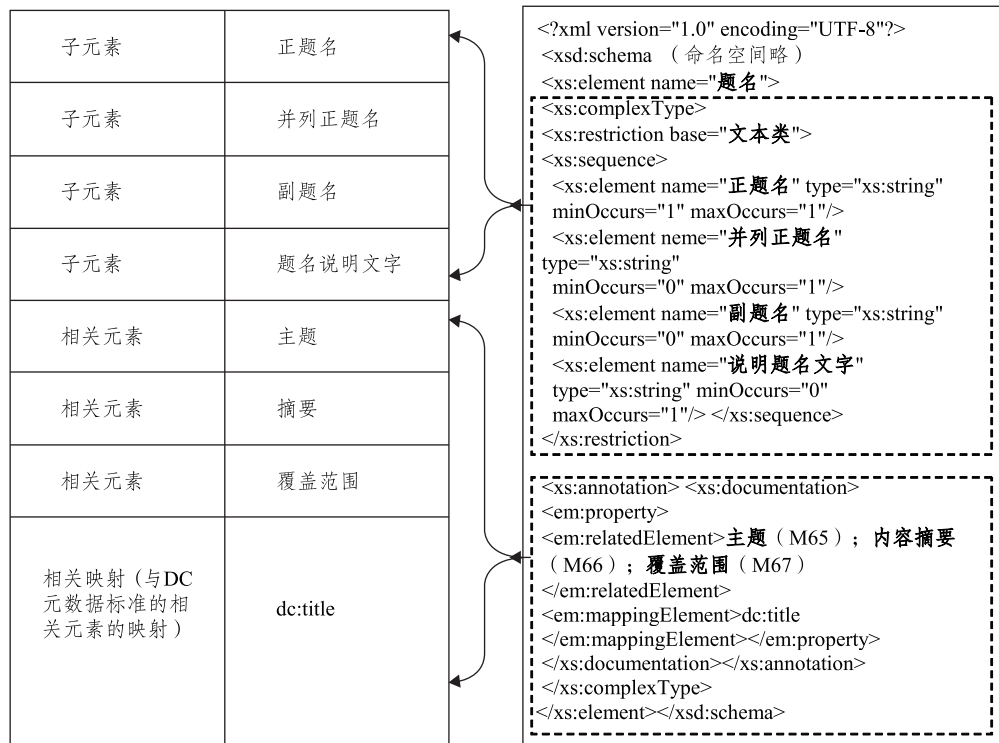


图 8 用 XML Schema 对电子文件管理元数据“题名”元素的纵向语义关系和横向语义关系的语法描述示例

副题名、题名说明文字”这些子元素在语义上是“题名”这一元素的下位元素,而仅仅是在形式上靠 XML Schema 序列化语法上的嵌套结构由元素和子元素在这一结构中的位置来表达。所以,从这个意义上讲,如果计算机通过 XML Schema 序列化语法能理解“题名”元素及其下位子元素之间的等级隶属关系,那也只是基于语法的形式上的理解,而不是基于语义的智能化理解。

从图 8 还看出,在 XML Schema 序列化语法中对电子文件管理元数据“题名”元素横向语义关系的描述上,如与“主题(M65)、内容摘要(M66)、覆盖范围(M67)”等这些表示相关元素横向关系的描述上,是通过 XML Schema 注文来描述的,即通过“<xs:annotation>”和其下一系列置标标识符所表示的注解来表达。这是因为在 XML Schema 语法中还不具有表达横向语义关系的元素,或者说 XML Schema 只能简单地表示纵向等级关系——树状关系,而难以表示横向语义上的相关关系,这样描述其结果如前所述,计算机并不能理解元素之间的横向语义关系,即只能在语法上识别它,不能从语义上理

解它。

(3)用 XML Schema 对电子文件管理元数据“题名”元素语义值进行描述,如对值表示形式、值限定词、值类型、允许值、最小值、最大值、值涵义、值定义域的定义等(图 9)。

从图 9 看出,在 XML Schema 序列化语法中对“题名”元素的语义值进行描述时,用置标标识符“<xs:restriction base = “文本型”>”来定义“值表示形式”、用“type = “xs:string””来定义“值类型”、用“minOccurs = “1”、maxOccurs = “1””来定义值域的“最小值”与“最大值”等,计算机通过这样的定义能确切地理解这些置标标识符的涵义。因此,可以看出在 XML Schema 序列化语法中,对电子文件管理元数据语义值的描述是其强项。

由此,根据图 7 – 图 9 的分析并加以综合可以看出,用 XML 来描述电子文件管理元数据元素是有其局限性的。对一个具体的元素而言,难以保证计算机对电子文件管理元数据语义理解的规范性和全面性,从而难以保证电子文件管理元数据语义描述的系统性;对元素之间的相互关系而言,不仅难以保证计算机对电子文

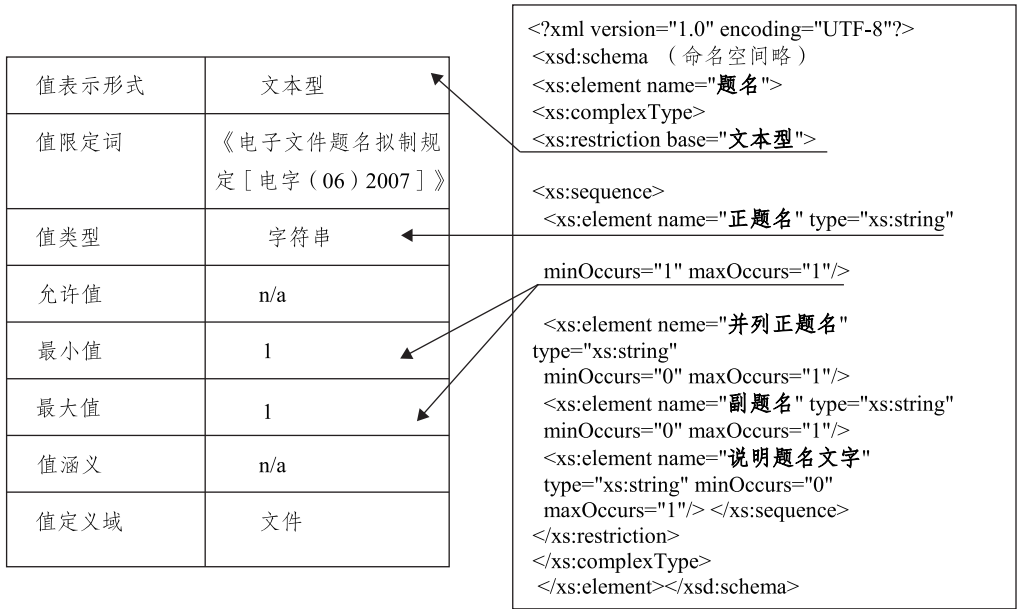


图 9 用 XML Schema 对电子文件管理元数据“题名”元素语义值的语法描述示例

件管理元数据语义之间的纵向等级语义关系的理解,也难以保证计算机对电子文件管理元数据语义之间的横向相关语义关系的理解,从而难以保证电子文件管理元数据语义关系描述的系统性。

但用 XML 来描述电子文件管理元数据,可以在形式上做到计算机可读。由此可见,采用 XML 来描述电子文件管理元数据只能实现其置标机读化的描述,而难以实现机读情况下在语义描述上的规范化与智能化,难以保证电子文件管理元数据在语法描述上的一致性,有其明显的局限性。

4 电子文件管理元数据的本体描述

电子文件管理元数据的本体描述,是指对

电子文件管理元数据采用计算机既可识别其语法又可理解其语义的方式所进行的置标智能化描述。探索采用本体方法来描述电子文件管理元数据成为继 XML 之后又一种方式。如前所述,采用本体描述,可采用的本体语言主要有 RDF^[1]、SKOS^[2] 和 OWL^[4]。

下面我们还是以电子文件管理元数据“题名”元素为例,用本体语言来描述该元数据元素的语义与语法,看其如何将语义描述与语法描述融合集成在一起的。

(1) 以本体语言对电子文件管理元数据“题名”元素的语义标识和语义内容方面进行描述(图 10)。

从图 10 看出,采用本体语言对电子文件管理元数据“题名”元素的语义标识和语义内容进行描述,其描述十分规范,主要表现在:

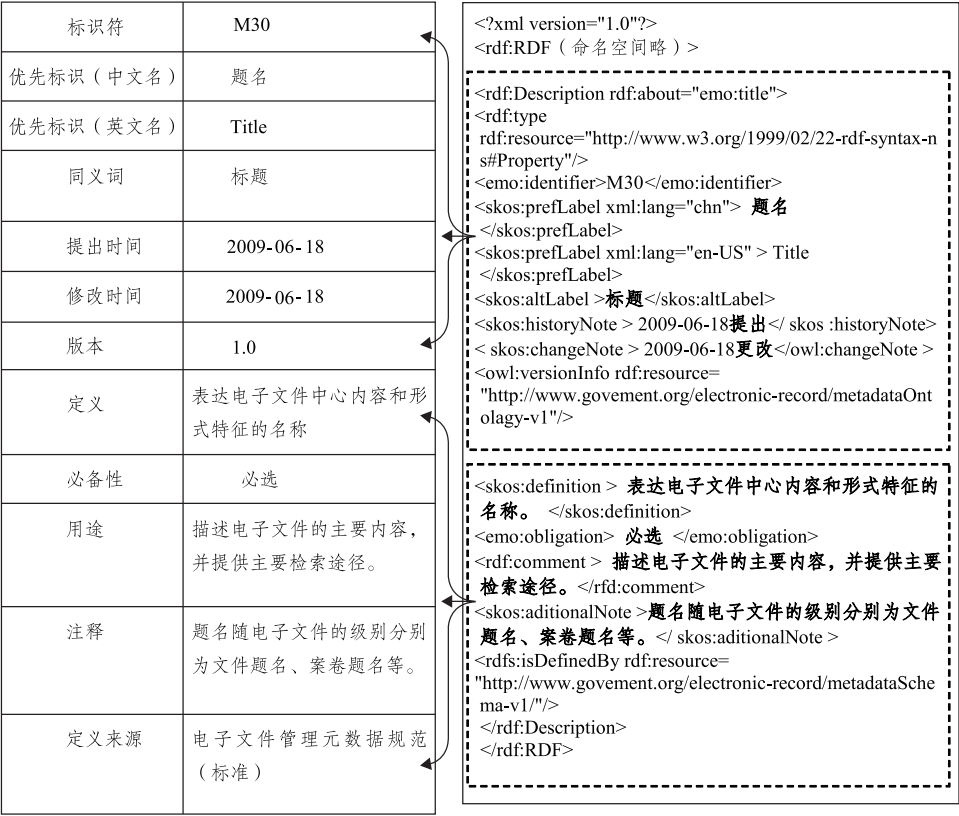


图 10 以本体语言对电子文件管理元数据“题名”元素的语义标识和语义内容方面的描述示例

第一,采用的描述语法十分规范。由于采用本体语言描述时,主要是基于 RDF 三元组抽象语法模型来描述“题名”元素的语义标识和语义内容,所以其主、谓、宾语法的构成十分有序,而且具有高度的一致性,所以语法的规范化程度很高。这是采用 XML 描述所难以企及的。

第二,采用的描述置标标识符十分规范。由于采用本体语言描述时,各类本体语言的词汇可以直接作为描述“题名”元素语义的置标标识符,如用 SKOS 的“<skos:prefLabel>”来表示优选标识“题名”,用“<skos:altLabel>”来表示同义非优选标识“标题”;用 RDF 的“<rdfs:is-DefinedBy>”来表示“定义来源”;用“<rdf:comment>”来表示“用途”、用“<owl:versionInfo>”来表示“版本”等。由于像这样的词汇在本体语言 SKOS、RDF 以及 OWL 本体中均有定义,而且其本身就是用来表示概念语义的词汇,是十分规范的置标标识符,无须“电子文件管理元数据模式——XML Schema”的设计者自行规定,可直接采用。很显然,使用这样规范的本体词汇来描述电子文件管理元数据“题名”元素的语义,其标准化程度就很高。

第三,表达的元素语义十分规范。尽管从形式上看,在采用本体语言描述题名元素的“标识符、优先标识、同义词、提出时间、修改时间、版本、定义、注释、定义来源”等这些语义标识和语义内容时与 XML 描述相类似,但由于其描述所用的语法模型、置标标识符标准化程度很高,且又是计算机可读的语法模型与置标标识符,因而又进一步保证了语义表达上的规范性,这就十分有利于在机读情况下表达语义的准确性与一致性。

(2)以本体语言对电子文件管理元数据“题名”元素之间的纵向语义关系和横向语义关系的描述(见图 11)。

从图 11 看出,采用本体语言来描述“题名”元素的关系,有助于计算机理解语义,主要表现在:

首先,语义关系的描述十分清晰。如采用 SKOS 中专门用于表达概念等级关系之间表示下位概念语义的词“<skos:narrower>”来表达“题名”元素与子元素“正题名、并列正题名、副题名、题名说明文字”之间的等级关系,就可以建立起题名“元素”之间的纵向语义关系,且描述十分清晰(图 12)。

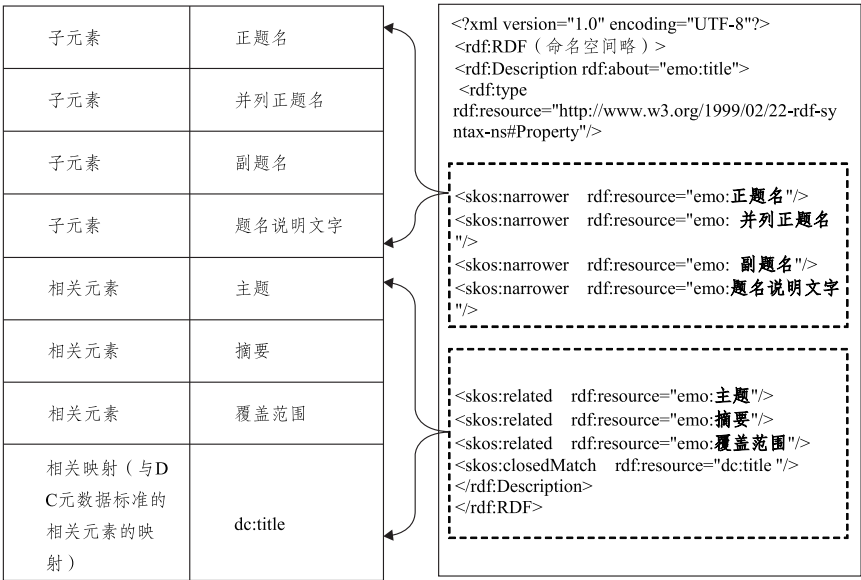


图 11 以本体语言对电子文件管理元数据“题名”元素的相关元素的描述示例

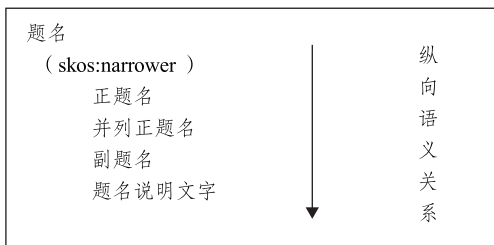


图 12 以本体语言所描述的题名
“元素”纵向语义关系

同样,采用 SKOS 中专门用于表达概念之间相关关系的词“<skos:related>”来表达“题名”元素与相关元素“主题、摘要、覆盖范围”之间的相关关系,就可以建立起题名“元素”的横向语义关系,且描述也十分清晰(图 13)。

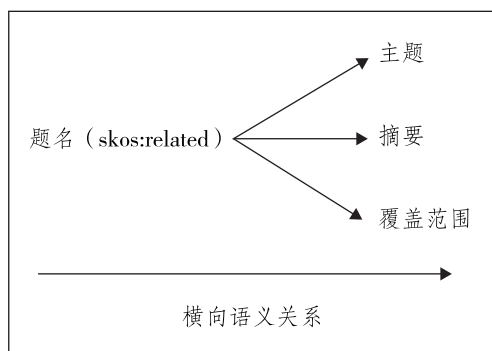


图 13 本体语言所描述的题名“元素”
横向语义关系

由此看出,采用本体语言描述电子文件管理元数据的语义关系,可以将纵向语义关系与横向语义关系描述得十分清晰。很显然,在此基础上形成的“电子文件管理元数据模式——XML Schema”就必然会形成一个纵横交错的语义网络。

其次,语义句子的描述十分清晰。如用 SKOS 的“<skos:narrower rdf:ID = “正题名”/>、<skos:narrower rdf:ID = “并列题名”/>、<skos:narrower rdf:ID = “副题名”/>、<skos:narrower rdf:ID = “题名说明文字”/>”这四个句子来定义子元素,用“<skos:related rdf:ID = “主

题”/>、<skos:related rdf:ID = “摘要”/>、<skos:related rdf:ID = “覆盖范围”/>”来定义相关元素,使计算机不仅可以识别它们,还可以理解它们,从而极大地弥补了 XML 描述机器不能理解语义关系的不足。

(3) 以本体语言对电子文件管理元数据“题名”元素语义值的描述,如对“值表示形式、值限定词、值类型、允许值、最小值、最大值、值涵义、值定义域”等的描述(图 14)。

从图 14 看出,采用本体语言在对电子文件管理元数据“题名”元素的语义值进行描述时,也可以达到与 XML 描述同样的效果。通过这样的描述,计算机就可以确切地理解这些采用本体语言的置标标识符的涵义,由此可看出本体语言在这方面的描述也是强项。

根据图 10—图 14 的分析并加以综合可以看出,用本体语言来描述电子文件管理元数据的语义有如下优势:

对一个具体的元素而言,可以保证计算机对电子文件管理元数据语义理解的规范性和全面性,从而保证电子文件管理元数据语义描述的系统性。

对元素之间的相互关系而言,不仅可以保证计算机对电子文件管理元数据语义之间的纵向等级语义关系的理解,而且可以保证计算机对电子文件管理元数据语义之间的横向相关语义关系的理解,从而保证电子文件管理元数据语义关系描述的系统性。

综上所述,采用本体语言来描述电子文件管理元数据,不仅可以在形式上做到计算机可读,而且可以在语义上做到计算机可理解。这种元数据的本体描述本质上是一种置标智能化的描述,而且是一种集语义与语法于一体的置标智能化描述,因此便于计算机实现对电子文件管理元数据在语义、语法及语义关系上的互操作,从而可以在时间和空间上最大限度地实现电子文件管理元数据的功能,达到保证电子文件管理的真实性、可靠性、完整性和可用性的长久保存的目的。

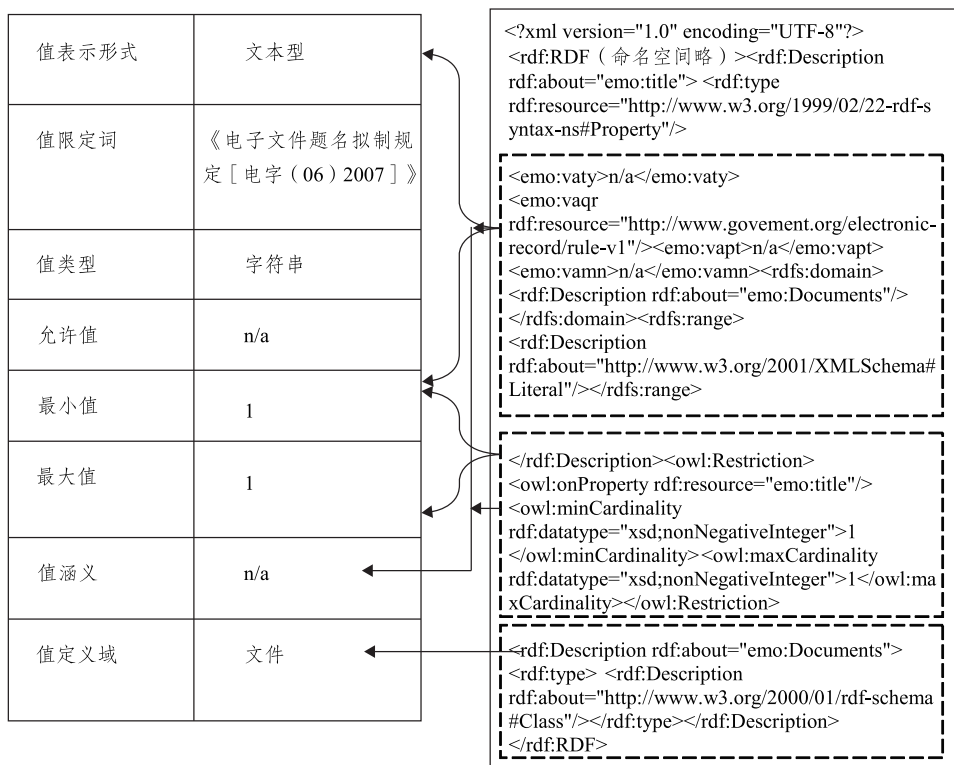


图 14 以本体语言对电子文件管理元数据“题名”元素语义值定义的描述示例

参考文献:

- [1] W3C. Resource Description Framework (RDF): Concepts and Abstract Syntax [S/OL]. W3C Recommendation 10 February 2004. [2009-11-03]. <http://www.w3.org/TR/rdf-concepts/>.
- [2] W3C. SKOS Simple Knowledge Organization System Reference [S/OL]. W3C Recommendation 18 August 2009. [2009-11-03]. <http://www.w3.org/TR/2009/CR-skos-reference-20090317/#L1693>.
- [3] W3C. OWL Web Ontology Language Overview [S/OL]. W3C Recommendation 10 February 2004. [2009-11-03]. [http://www.w3.org/TR/owl-](http://www.w3.org/TR/owl-features/)

features/.

- [4] ISO/IEC. ISO/IEC11179 - 3:2003 信息技术 - 元数据注册 - 第 3 部分:注册元模型与基本属性[S]. 2003-02-15.
- [5] W3C. XML Schema Definition Language (XSD) 1.1 Part 1: Structures [S/OL]. W3C Candidate Recommendation 30 April 2009. [2009-11-03]. <http://www.w3.org/TR/xmlschema11-1/>.

张正强 南京政治学院上海分院信息管理系教授,博士生导师。通讯地址:上海四平路 2575 号。邮编:200433。
(收稿日期:2009-09-07;修回日期:2009-11-05)