

【自动化技术】

# 军械保障阵地设施表格文档信息标准化

盖炳良<sup>a</sup>, 唐金国<sup>b</sup>, 孙媛<sup>b</sup>

(海军航空工程学院 a. 研究生管理大队; b. 科研部, 山东 烟台 264001)

**摘要:**军械保障阵地设施文档涉及大量表格信息,将表格信息结构化,能提高信息检索与再利用的效率。面向军械保障实际应用,借鉴相关标准,提出了设施表格文档信息的通用模型和信息结构模型。通过信息描述模型和表示方法确立了设施表格文档信息标准化方法。

**关键词:**阵地设施;表格文档;信息标准化;通用模型

**中图分类号:**TP311

**文献标识码:**A

**文章编号:**1006-0707(2011)09-0098-04

军械保障阵地设施是军械保障阵地中军械装备及其配套物资进行存放、保管、保养、维修、训练等所需的建筑、场所及其配套设施的统称。文档是归档文件最小的文件单元,可以是一个独立的自然件,也可以是组合文件中的某一自然件或某一自然件的附件<sup>[1]</sup>。军械保障阵地设施表格文档是描述军械保障阵地设施属性信息的表格式文档。在信息采集阶段,作为阵地设施信息采集模板。随着各类网络平台的应用,表格信息以电子文件为载体通过网络传递,而这类平台往往基于 doc 等字处理文件格式,其中包含的信息是非结构化的,难以让计算机快速处理。故本文对设施表格文档信息进行标准化研究,包括对设施表格文档信息内容、描述方式及表示方式的标准化。

## 1 相关概述

立足信息标准化,本研究借鉴和应用科学数据共享工程技术标准、国家档案行业标准《文书类电子文件元数据方案》、CCTS、HL7 CDA R2 以及 XML 等关键技术,确立保障设施信息文档通用模型,建立设施信息文档的信息标准化方法。

科学数据共享工程技术标准《元数据标准化基本原则和方法》提供了共享参考元数据所定义的元数据框架,包括标识信息、内容信息、分发信息、数据质量信息、数据表现信息、参照系信息、图示表达目录信息、元数据扩展信息、数据模式信息、限制信息和维护信息等 11 个元数据子集。在制定领域和领域专用元数据标准时,需要对这 11 个元数据子集进行选择,并对其中所包含的元数据实体和元数据元素进行扩展和裁减<sup>[2]</sup>。本研究根据实际应用需要,选择标识信息、内容信息、限制信息、维护信息等元数据。标识信息包含了唯一标识数据集的信息。维护信息包含有关元数据或数据集的更新频率及更新范围的信息。限制信息包含访问和使用元数据或数据集的限制信息。限制信息包含如下几方面的

信息:使用方面的限制信息,安全方面的限制信息,法律方面的限制信息。内容信息提供描述数据集内容特征的信息。

CCTS(core components technical specification)是联合国促进贸易与电子商务中心用于规范电子信息交换和共享的信息标准,描述电子信息的组成结构<sup>[3]</sup>。CCTS 的关键是基于核心组件(core component)及业务信息实体(business information entities)2个概念的抽象。核心组件是建立互操作的业务实体和文档的关键。核心组件按照功能分为聚集核心组件(aggregate core component, ACC)、联合核心组件(association core component, ASCC)和基本核心组件(basic core component, BCC)3类。其中,聚集核心组件是电子信息组成结构中的最顶层,是一组相关联信息的聚合,类似于实体(对象)类的表示。联合核心组件定义为聚集核心组件间关联的角色,包含自身属性及上一层聚集核心组件的对象类,其与聚集核心组件在本质上是相同的,区别仅在于它们在电子信息中处于不同的层次。基本核心组件表示聚集核心组件的属性,不可向下分解,类似于实体(对象)类的属性。核心数据类型(core data types, CDT),用来描述基本核心组件,常用类型有文本、数值、日期时间等。三者的关系模型如图 1 所示。

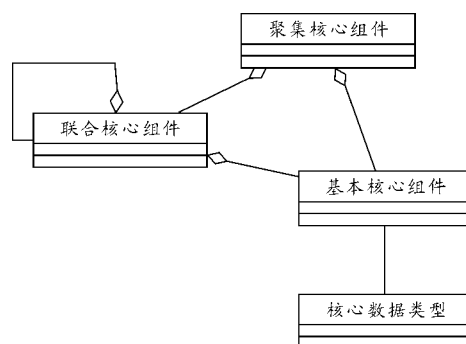


图 1 核心组件间关系模型

收稿日期:2011-07-11

作者简介:盖炳良(1981—),男,硕士研究生,主要从事装备保障信息化研究。

## 2 设施表格文档信息通用模型

HL7 (health level seven)是目前世界上最大的卫生信息标准的国际开发组织,其开发的临床文档标准 HL7 CDA R2 称为临床文档架构,是文档标记标准,用于限定信息交换的结构和语义<sup>[4]</sup>。借鉴 HL7 CDA R2 临床文档架构模型,军械保障阵地设施表格文档由文档头(Header)和文档体(Body)2 大部份构成,文档头在设施表格文档的头部分,需要记录信息的内容有 2 类:文档信息和设施基本信息。其中文档信息包含文档标识信息,文档维护信息,文档密级等相关信息。由此可见,按照元数据的类别,文档头信息表述阵地设施的标识信息以及文档维护信息及限制信息。

文档体(Body)是设施信息文档的主体部分,表达元数据的内容信息。记录的信息内容应该涵盖设施信息的各个方面,从军械保障实际应用需求出发,文档体内容应着力体现设施的功能信息,及设施的质量状态。

因此,设施表格文档信息通用模型如图 2 所示。

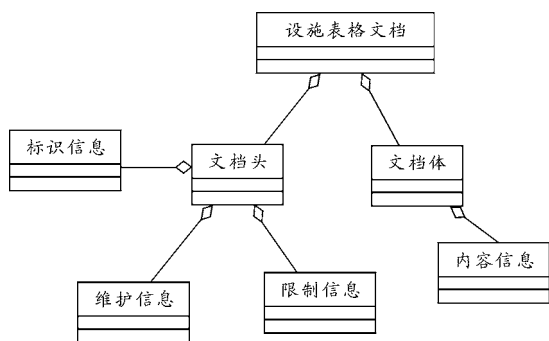


图 2 设施表格文档信息通用模型

## 3 设施表格文档信息标准化方法

表格文档信息结构模型是核心,文档信息描述和表示及其它处理都是围绕着信息结构模型进行的<sup>[5]</sup>。依据图 1 和图 2,得到设施表格文档信息结构模型如图 3 所示。

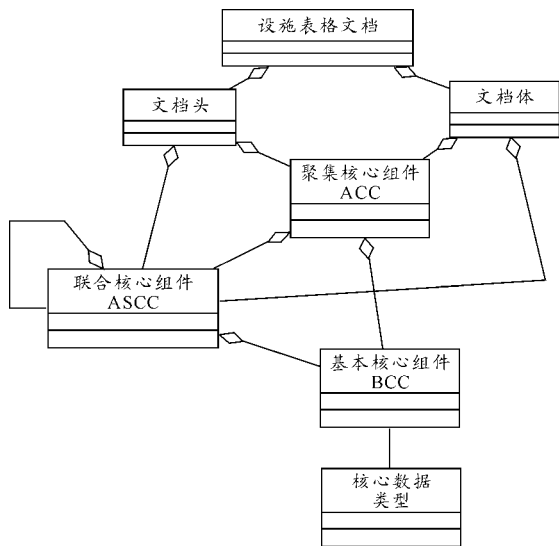


图 3 设施表格文档信息结构模型

### 3.1 描述模型

本文采用数据元作为描述表格文档信息的数据单元。数据元是用一组属性来描述其定义、标识、表示和允许值的数据单元,由对象类、特征及表示 3 部分组成。三者之间的关系是:对象类具有属性,特征具有表现样式<sup>[6]</sup>。例如“建筑时间”是设施基本信息中的一个数据单元,用数据元进行描述,即“建筑时间”是对象类,它具有“数据类型”特征,“数据类型”的表现样式为日期时间。

依据图 2 信息通用模型和图 3 信息结构模型,得到核心组件和数据元的对应关系,从而通过对应关系建立表格文档信息的描述模型。采用聚集核心组件描述标识信息,维护信息,限制信息,内容信息等最顶层的根聚合数据元,按照图 1 核心组件内部关系分别将它们逐级分解成单一数据元和聚合数据元。采用联合核心组件描述聚合数据元。采用基本核心组件描述单一数据元,采用核心数据类型描述其表现样式<sup>[7]</sup>。

下面以标识信息为例,建立表格文档信息描述模型。

首先用聚集核心组件描述标识信息,然后将其分解为文档标识信息等单一数据元,设施标识信息等聚合数据元。设施标识信息又分为设施名称,设施编码,设施所属单位三个单一数据元,设施基本信息一个聚合数据元。继续对设施基本信息进行分解,分别得到建筑时间,竣工时间,启用时间,建筑形式,建筑结构,用途分类,主要用途,建筑面积,使用面积,使用率,层数,房间总数等单一数据元及建筑尺寸聚合数据元,建筑尺寸又分解为长度,宽度,高度三个单一数据元。照此分解完毕后,即可得到标识信息描述模型。见图 4 所示。

在设施基本信息聚合数据元分解时,对于基本信息确定充分参考和借鉴了已有相关规则。2008 年 6 月住房和城乡建设部发布的《国家机关办公建筑和大型公共建筑能耗监测系统分项能耗数据采集技术导则》中,建筑基本情况数据采集指标根据建筑规模、建筑功能、建筑用能特点划分为基本项和附加项,其中基本项信息为建筑规模和建筑功能等基本情况的数据,建筑对象的基本信息包括建筑名称、建筑地址、建设年代、建筑层数、建筑功能、建筑总面积、空调面积、采暖面积、建筑空调系统形式、建筑采暖系统形式、建筑体型系数、建筑结构形式、建筑外墙材料形式、建筑外墙保温形式、建筑外窗类型、建筑玻璃类型、窗框材料类型、经济指标(电价、水价、气价、热价)、填表日期、能耗监测工程验收日期等<sup>[8]</sup>。

在《广东省国家机关办公建筑和大型公共建筑基本信息及建筑能耗统计人员工作手册》中,其建筑基本信息包含建筑物详细名称,建筑详细地址,竣工时间,建筑类型,建筑功能,建筑层数,建筑面积,供冷方式等信息<sup>[9]</sup>。

在 2008 年开始的全国中小学校舍基本信息调查中制定的《全国中小学校舍基本信息调查表》中,基本项信息为建筑物名称,建筑物编号,主要功能,设计抗震烈度取值,结构类型,基础形式,平面形式,楼板形式,产权性质,竣工年份,最高层数,建筑面积,固定资产价值,国家级专项资金类别,危

房情况等信息。其中,危房情况又分为危房等级,危房面积,鉴定机构名称,鉴定年份,使用状态等信息<sup>[10]</sup>。

分析上述3类具体应用中,首先梳理共有项,得到建筑名称,建筑层数,建筑面积,建筑(主要)功能等四项,其中名称已经作为单一数据元采用,因此,采用建筑层数,并在建筑层数基础上,扩展军械保障关注的房间总数一项。采用建筑面积,并在建筑面积上扩展使用面积和使用率2项,以掌握设施更详细的使用情况。不同于上述三类实际应用,阵地设施的功能用途为核心关注,因此本文采用设施用途分类来概括性描述,采用主要用途准确定位设施功能。

3类应用中分别采用的“建设年代”、“竣工时间”、“竣工年份”,意在掌握设施的时间要素,但表述含糊,不全面,因此本文将设施时间扩展为建筑时间,竣工时间,启用时间3项,较为全面的反映设施的时间要素。3类应用中,分别有“建筑结构形式”、“建筑类型”、“结构类型”等项,依据军事设施实际,本文采用建筑形式来描述设施是处于地面或是地下等状态,采用建筑结构掌握设施的构筑。

另外,3类应用中都缺乏对建筑设施的物理形状尺寸的描述,在军事应用中,物理形状是定位识别的重要指标,因此本文增加建筑尺寸一项。

由此通过借鉴分析,面向军械保障实际应用需要,选取建筑时间,竣工时间,启用时间,建筑形式,建筑结构,用途分类,主要用途,建筑面积,使用面积,使用率,建筑层数,房间总数,建筑尺寸等13项。建筑尺寸又分解为建筑长度,宽度和高度。

### 3.2 表示方法

UBL(universal business language)是结构化信息标准促进组织(OASIS)制订的一项国际标准,其目的在于促进电子文件的标准化,其在 CCTS 的基础上制订了将不同数据元映射到 XML Schema 的方法。本研究利用该标准方法,确定本研究确定了数据元的表示方法。

单一数据元的表示方法:

1) 将数据元定义为 XML Schema 中的元素 element, element name 值为数据元的中文名称。

2) 采用“Type”属性定义不需要对数据类型进一步限制的数据元的数据类型。

3) 采用“simpleType”和“restriction”定义数据类型需进一步限制的数据元,“simpleType”的 name 属性是元素类型的值。

4) 采用“simpleType”、“restriction”和“enumeration”定义代码型数据元,“simpleType”的 name 属性是元素类型的值。

聚合数据元的表示方法:

1) 采用“complexType”定义聚合数据元的类型,“complexType”的 name 属性是元素类型的值。

2) 聚合数据元对应的元素的子元素应为该聚合数据元所包括的单一数据元和聚合数据元。

3) 聚合数据元所对应的元素所包含的属性应为单一数据元。

4) 聚合数据元中的数据元和聚合数据元有允许出现最大和最小次数的属性。

其映射关系如图5所示。

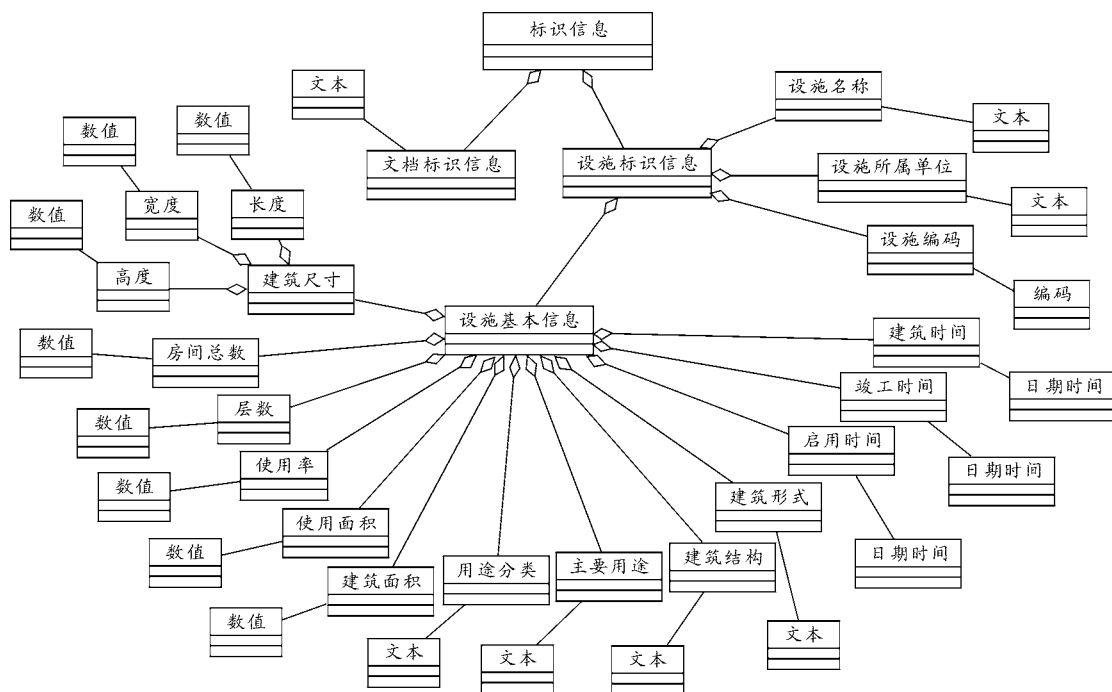


图4 标识信息描述模型

