



# 中华人民共和国国家标准

GB/TXXXXX—XXXX

## 信息技术 学习、教育和培训 数字化学习资源语义描述

Information technology – Learning, education and training -  
Semantic description of digital learning resource

(报批稿)

XXXX-XX-XX 发布

XXXX-XX-XX 实施

中华人民共和国国家质量监督检验检疫总局  
中国国家标准化管理委员会 发布



# 目 次

前言 .....	II
1 范围 .....	1
2 术语和定义 .....	1
3 缩略语 .....	1
4 数字化学习资源语义 .....	2
5 语义描述流程 .....	2
6 确定术语及其定义 .....	3
6.1 过程 .....	3
6.2 术语定义规则 .....	3
7 提取元数据 .....	3
8 规范元数据 .....	4
8.1 元数据的规范过程 .....	4
8.2 数据项 .....	4
9 描述数字化学习资源 .....	4
附录 A（资料性附录） 数据项示例 .....	5
参考文献 .....	6

## 前 言

本标准按照GB/T 1.1-2009给出的规则起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别这些专利的责任。

本标准由全国信息技术标准化技术委员会（SAC/TC28）提出并归口。

本标准起草单位：华中师范大学、中国电子技术标准化研究院、博康智能信息技术有限公司。

本标准主要起草人：吴砥、余云涛、彭娴、李莹、饶景阳、王紫琴、王杨春晓、高展、马文学。

# 信息技术 学习、教育和培训

## 数字化学习资源语义描述

### 1 范围

本标准规定了数字化学习资源语义描述的内容、流程、方法和要求。  
本标准适用于数字化学习资源、在线教育平台的设计和开发。

### 2 术语和定义

下列术语和定义适用于本文件。

#### 2.1

**数据类型** data type

已定义的特定数据结构的数据对象集合和一组允许的操作。

#### 2.2

**语义描述** semantic description

对特定的词、词组或短语所表达的概念的描述。

#### 2.3

**数字化学习资源** digital learning resources

经过数字化处理，可以在计算机或网络上使用的学习资源。

#### 2.4

**元数据** metadata

定义和描述其他数据的数据。

[GB/T 18391.1-2009, 定义3.2.16]

#### 2.5

**数据项** data item

用于描述元数据的基本数据单元。

### 3 缩略语

下列缩略语适用于本文件。

OWL: 网络本体语言 (Web Ontology Language)

RDF: 资源描述框架 (Resource Description Framework)

URL：统一资源定位符（Uniform Resource Locator）

#### 4 数字化学习资源语义

语义是语言符号所表示的内容和所传递的意义，是对客观存在抽象的、概括的反映，具有相对稳定、统一的解释，反映了人类与世界之间一种有意义的交流。

数字化学习资源（以下简称学习资源）语义是学习资源在计算机中的表示与通往现实世界的途径。对计算机而言，学习资源语义是指能够被计算机系统处理的信息；对现实世界而言，是指人借由计算机获得的理解。

示例：

在《现代远程教育资源建设技术规范》中，主要包括以下几类学习资源：媒体素材、题库、课件与网络课件以及网络课程。用 OWL 可以描述学习资源间的语义关联关系，要表达文本素材、图/图像类素材、音频类素材、视频类素材和动画类素材都是媒体素材的一种形式，可以用以下描述方式。

```
<owl: Ontology rdf: about= “统一资源定位地址，通常用 URL 表示” />
  <owl: Class rdf: ID= “文本素材” >
    <rdfs: subClassOf rdf: resource= “媒体素材” />
  </owl: Class>
  <owl: Class rdf: ID= “图形/图像类素材” >
    <rdfs: subClassOf rdf: resource= “媒体素材” />
  </owl: Class>
  <owl: Class rdf: ID= “音频类素材” >
    <rdfs: subClassOf rdf: resource= “媒体素材” />
  </owl: Class>
  <owl: Class rdf: ID= “视频类素材” >
    <rdfs: subClassOf rdf: resource= “媒体素材” />
  </owl: Class>
  <owl: Class rdf: ID= “动画类素材” >
    <rdfs: subClassOf rdf: resource= “媒体素材” />
  </owl: Class>
```

每个学习资源都可以是一个学习对象，学习对象属性的语义信息可以用学习对象元数据来表示，GB/T 21365-2008 采用 9 类特征信息来描述学习对象属性的语义信息，这九大类特征分别用数据元素：通用、生存期、元-元数据、技术、教育、权利、关系、评注和分类来表示。每个数据元素又包含子元素来表示具体的语义细节信息，如学习对象的标题、语种、关键字和制作者等基本语义信息。具体元素的描述细节参见 GB/T 21365-2008 表 1。

#### 5 语义描述流程

学习资源的语义描述一般包括 4 个步骤：选定对象、确定术语、提取元数据和用数据描述语言来描述学习资源。

选定学习资源中所要描述的对象，即选定所要描述的具体学习资源。所选定的对象可能比较具体，也可能比较笼统，即粒度有粗有细。具体对象的选择及其粒度的确定，由学习资源描述的相关方按特定需求及特定学习资源描述目的确定。本标准不对对象的选定做任何规定，只描述 4 个步骤中的后面 3 个，即：

- a) 确定所选对象的特定术语及其定义；
  - b) 针对特定术语提取并规范元数据，必要时(如,元数据很多)创建数据字典；
  - c) 利用所规范的元数据，采用适当的数据描述语言描述学习资源。
- 执行上述步骤 a) 时，其实现方法不在本标准范围内考虑。

## 6 确定术语及其定义

### 6.1 过程

确定术语及其定义是对特定对象的本质达成共识的过程。定义应准确反映对象本质，这是进一步恰当提取元数据的基础。特定对象的术语应获得大多数相关方认可。特定对象的术语及其定义一旦确定，应在其使用范围内形成规范性文件。

确定新的特定对象的术语及其定义的过程如下：

- a) 搜集所研究的专业领域的对象概念；
- b) 分析各概念的内涵和外延；
- c) 确定各概念在概念体系中的相互关系及位置；
- d) 在概念关系的基础上为概念撰写定义；
- e) 赋予每个概念确定的名称，即给出相应的术语。

### 6.2 术语定义规则

若已经有了普遍接受的术语，但概念不统一，在这种情况下应直接针对该术语规范其定义。一些相关领域已经给出了标准化的术语及其定义，凡适用者，均可直接采纳供学习资源领域使用。

**示例：**

在《现代远程教育建设技术规范》中，规范的以下术语及其定义便可直接作为描述学习资源的术语和定义：

**题库：**指按照一定的教育测量理论，在计算机系统中实现的某个学科题目的集合，是在数学模型基础上建立起来的教育测量工具。

**案例：**指有现实指导意义和教学意义的代表性的事件或现象。

**课件与网络课件：**指对一个或几个知识点实施相对完整教学的辅助教学软件，根据运行平台划分，可分为网络版的课件和单机运行的课件，网络版的课件需要在标准浏览器中运行，单机运行的课件可通过网络下载运行，以上均泛称网络课件。

**网络课程：**网络课程是通过网络表现的某门学科的教学内容及实施的教学活动的总和，它包括按一定的教学目标、教学策略组织起来的教学内容和网络教学支撑环境两个部分。

## 7 提取元数据

所确定的术语及其定义代表了所对应的对象。提取元数据，用于从不同角度描述特定对象。元数据充分且恰当的提取，有助于准确描述特定对象。

提取元数据时，除了关注对象的本质外，宜考虑特定领域的学习资源描述需求。

**示例：**

GB/T 25100-2010给出了用来支持网络课程环境的学习资源描述涉及的一些主要元数据，这些元数据有：名称、创建者、主题、描述、出版者、其他责任者、日期、类型、格式、标识符、来源、语种、关联、时空范围、权限。

## 8 规范元数据

### 8.1 元数据的规范过程

元数据的规范过程即元数据标准化过程,是在一定范围内对元数据表现形式及其传递的含义获得共识的过程。数字化学习资源相关利益方约定并达成一致,形成明文的元数据规范,对元数据的规范通过规范其数据项的方式执行。正式批准发布的元数据规范文件应纳入该文件发布者的配置管理范畴。一般是对多个元数据同时成批规范化;在元数据数量很大的情况下,往往将这些规范化的元数据编撰成数据字典。

规范元数据的过程主要包括以下步骤:

- a) 确定元数据的数据项,选定所有用于描述该元数据的必要的数据项,选择时宜考虑表现该元数据的各个维度;
- b) 为各个数据项赋值并在利益相关方范围内求得共识;
- c) 将元数据及其所有求得共识的数据项形成规范文件,即元数据规范文件;
- d) 是否执行上述步骤c),以及在什么层级形成规范文件,取决于相关方的需求。将元数据及其已赋值的数据项作为相关标准草案的组成部分,在批准发布该标准后,其中包含的元数据及其已赋值的数据项也就成为元数据规范文件。

**注:**在元数据规范语境下,也往往使用“属性”取代“数据项”,二者同义。

### 8.2 数据项

所有元数据的数据项中至少包括所描述的元数据的标识符(ID)、名称和该名称代表的元数据的定义,并指出是否为必备数据项。元数据的数据项,从约束性角度看,包括必备、可选和条件必备(在一定条件下是必备的)三类。关于必备、条件必备或可选数据项的选取,应考虑对于元数据所代表对象的重要性质表述的必要性、充分性和确切性。附录A.1是一个元数据项及其已赋值的例子。

## 9 描述数字化学习资源

自行设计或选择合适的数据描述语言对数字化学习资源进行描述。一些数据描述语言支持产生新的描述方案和描述符,以及扩充和修正现有的描述方案。运用此类数据描述语言,用户可根据需要自定义新的描述方案和描述符。以下是一个以RDF描述学习资源语义信息的例子。

示例:

```
<rdf:Description rdf:about="在线英语学习网站的地址">
  <dc:title>英语语法</dc:title>
  <dc:description>本课程介绍了高中英语常见的语法点</dc:description>
  <dc:publisher>课程的发布者或者机构名称</dc:publisher>
  <dc:date>2008-01-01</dc:date>
  <dc:type>在线视频课程</dc:type>
  <dc:format>MP4</dc:format>
  <dc:language>中文</dc:language>
</rdf:Description>
```

附 录 A  
(资料性附录)  
数据项示例

### A.1 数据项示例

表A.1中选取了GB/T AAAAA-AAAA的几个数据项，数据项规范如下：

标识符（或其缩略语ID）：唯一标识该元数据项的标识符；

名称：该元数据项的名称及对应的英文名称；

定义：该元数据项的释义；

大小：允许的值的个数；

值空间：该元数据项的取值范围，一般以词汇表或者引用另一个标准/规范的形式出现；

数据类型：用于呈现该元数据项的值的的数据类型，可能涉及的数据类型有多文种字符串、日期时间、持续时间、词汇表和字符串等；

约束：对该元数据项的约束状态的描述，表中以字母“M”表示必备，以字母“O”表示可选；

说明/示例：作为对定义的补充，进一步简要说明该元数据项和/或给出具体例子。

表 A.1 元数据的数据项示例

ID	名称	定义	大小	值空间	数据类型	约束	说明/示例
10000	通用	学习资源的一些通用信息。	1	---	---	M	普遍适用的数据信息
11000	标识符	资源的标号，能够通过此标识符在系统中唯一定位到此资源。	1	---	---	M	---
11100	类别	所属标识方案或编目方案的名称或指示符。一种命名方案。	1	---	字符串 (至少 1000 个字符)	M	“ISBN” “URI” “CSRC”
12000	标题	学习资源的名称。	1	---	多文种字符串 (至少 1000 个字符)	M	---
--表示不必要、不适用							

### 参 考 文 献

- [1] GB/T 13000-2010 信息技术 通用多八位编码字符集 (UCS)
  - [2] GB/T 18391.1-2009 信息技术 元数据注册系统 (MDR)
  - [3] GB/T 21365-2008 信息技术 学习、教育和培训 学习对象元数据
  - [4] GB/T 25100-2010 信息与文献 都柏林核心元数据
  - [5] GB/T AAAAA-XXXX 信息技术 学习 教育和培训 虚拟实验构件元数据
  - [6] 教育部 现代远程教育资源建设技术规范 (试行)
-