# 专题资源网站采集项目

**元数据著录规则（2019）**

**1．著录对象**

著录对象为采集的网络专题资源，著录时以单次存档的专题资源为一个著录单位。

专题：是指具有相同或相关主题的网络信息资源的集合。

专题资源：属于某个专题的具体资源对象，资源类型可以是网站、网站频道或单一网页。

**2．著录要求**

对采集的专题资源进行编目加工，要求参照著录规则进行编目。元数据著录规则如下：

| **术语** | **必备性** | **著录内容** |
| --- | --- | --- |
| 加工编号 | 必备 | 著录元数据的一个明确标识，具体规则见《专题编号及专题资源采集加工编号命名规则》。 |
| CDOI | 有则必备 | 著录所采集专题资源的唯一标识号。 |
| 资源名称 | 必备 | 著录该资源的名称，一般指网络资源正式公开的名称。 |
| 所属网站 | （网站频道或网页）必备 | 著录该资源所属网站的正式发布的规范名称。 |
| 所属专题 | 必备 | 著录该资源所属的专题编号，具体规则见《专题编号及专题资源采集加工编号命名规则》。 |
| 摘要 | 必备 | 著录该资源内容的总结概括性文字。摘要字数要求200字以内。语句简洁流畅，无语法错误。 |
| 关键词 | 必备 | 著录反映该资源主要内容的名词或名词短语。如有多个关键词，以英文半角分号间隔。 |
| 时间范围 | 有则必备 | 著录专题资源内容的时间特征。 |
| 空间范围 | 有则必备 | 著录专题资源内容涉及的空间特征。包括地点、地理坐标。 |
| 资源类型 | 必备 | 著录所保存资源的类型。如资源为单一网页则值为“网页”，如资源为某网站的某个专题频道，则值为“频道”，如资源为某个网站，则值为“网站”。 |
| 内容形式 | 必备 | 著录内容形式及内容限定。参考国家标准GB/T 3469—2013《信息资源的内容形式和媒体类型标识》取值。 |
| 媒体类型 | 必备 | 著录用以承载资源内容的载体类别。参考国家标准GB/T 3469—2013《信息资源的内容形式和媒体类型标识》取值。网络信息保存资源媒体类型统一著录为“电子”。 |
| 语种 | 必备 | 著录该资源的3位语种代码，可参考《新版中国机读目录格式使用手册》。如有多个语种，以英文半角分号间隔。 |
| 保存格式 | 必备 | 著录所采集的网站资源存档格式。统一著录为“WARC”。 |
| 采集地址 | 必备 | 著录该资源的原始访问地址。 |
| 采集日期 | 必备 | 著录该资源采集的日期。如果在审核过程中需重新采集，应对本项内容进行修改。 |
| 发布地址 | 必备 | 著录存档资源的发布地址。 |
| 发布日期 | 必备 | 著录存档资源发布的日期。 |
| 访问方式 | 必备 | 著录资源可以提供服务的范围，取值：互联网访问、数字图书馆专用网络访问、××图书馆局域网访问等。 |
| 中图分类 | 必备 | 著录该资源内容所属的中图分类号，多个分类号用英文半角分号间隔。 |
| 附注 | 有则必备 | 未在其他著录项中著录而又有必要进一步补充说明的内容，均可著录于本项。 |
| 数据提交单位 | 必备 | 著录承建馆的名称。 |
| 所属任务年份 | 必备 | 著录联建工作的任务年度，2019年度数据则著录2019。 |

注：上述编目字段中所涉及的时间日期格式为：yyyy-mm-dd，有明确开始和结束时间的填写时间段yyyy-mm-dd/yyyy-mm-dd，没有明确结束时间填写开始日期，格式为：start:yyyy-mm-dd;没有明确开始时间填写结束日期，格式为：end:yyyy-mm-dd

**3．专题编号及专题资源采集加工编号命名规则**

**专题编号**统一定长为11位。具体组成是：资源类型代码（1位，专题：Z）、采集机构代码（4位，数值取自机构代码，机构代码见图书馆机构代码表）、所属任务年份（4位）、专题序号（2位）

**专题资源采集加工编号**统一定长为15位。具体组成是：专题编号（11位）、资源记录流水号（4位）。

注释：

（1）所属任务年并非采集年，而是专题任务所属年份；

（2）专题序号是指某机构当年所采集的第几个专题，顺序排列，如“2019年两会”是2019年的第3个专题，则专题序号则为“03”；

专题编号样例：Z0100201903

（3）资源记录流水号是该条资源属于某个专题的第几个资源记录，如“两会新闻发布会召开”属于“2019年两会”专题的第98条资源记录，则该资源记录流水号是“0098”;

专题资源采集加工编号样例：Z01002019030098

**4．著录样例**

|  |  |
| --- | --- |
| 加工编号 | Z01002016030086 |
| CDOI | 108.ndlc.39.4200009031010001/T8F45.1000000004 |
| 资源名称 | 北京市全面推开营业税改增值税试点 |
| 所属网站 | 首都之窗 |
| 所属专题 | 实施营改增政策深化供给侧改革 |
| 摘要 | 2016年3月24日财政部网站公布了《关于全面推开营业税改征增值税试点的通知》，营改增试点的细则正式亮相。首都之窗策划专题“北京市全面推开营业税改增值税试点”,对此进行全媒体报道，内容包括：图解营改增、办税指南、四大行业税收指引、你问我答和政策文件等版块内容。 |
| 关键词 | 首都之窗;营改增;税收;试点;营业税;增值税 |
| 时间范围 | start:2016-03-24 |
| 空间范围 | 北京市 |
| 资源类型 | 频道 |
| 内容形式 | 多种内容形式 |
| 媒体类型 | 电子 |
| 语种 | chi |
| 保存格式 | WARC |
| 采集地址 | <http://zhengwu.beijing.gov.cn/zwzt/ygz/> |
| 采集日期 | 2016-05-17 |
| 发布地址 | http://10.102.0.0/16/36yinggaizeng/20160509023113/http://zhengwu.beijing.gov.cn/zwzt/ygz/ |
| 发布日期 | 2016-10-08 |
| 访问方式 | 数字图书馆专用网络访问 |
| 中图分类 | F812.42 |
| 数据提交单位 | 首都图书馆 |
| 所属任务年份 | [2016](http://zhengwu.beijing.gov.cn/zwzt/ygz/) |

附件1：信息资源的内容形式和媒体类型标识简表

附件1

信息资源的内容形式和媒体类型标识简表

* **内容形式词**

|  |  |
| --- | --- |
| **内容形式词** | **词的定义和范围** |
| 数据集 | 由数字编码数据表示的、要用计算机处理的内容。不包括数字形式记录的音乐、语言、声音、计算机复制的图像和文本。 |
| 图像 | 通过线条、形状、阴影等表示的、用视觉感知的内容。图像可以是静态的或动态的，二维的或三维的。 |
| 运动 | 通过动作，即物体或人改变位置的动作或过程所表示的内容。但不包括动态图像。 |
| 音乐 | 通过连续地、组合和时序关系传送有序音调或声音，产生乐曲而表示的内容。音乐可以是手写的（乐谱）、演奏的、以模拟或数字形式录制的（有节奏、旋律或和声的歌唱声、乐器声和机械声）。 |
| 实物 | 通过三维材料（自然实体或人造/机造人工制品）表示的内容。也称为三维结构或教具。 |
| 程序 | 通过用计算机处理或执行的数字编码指令表示的内容。 |
| 声音 | 通过动物、鸟类、自然噪声源，或人类声音、数字（或模拟）媒体模拟的声音而表示的内容。但不包括录制的音乐、话语录音。 |
| 话语 | 通过人类说话声音表示的内容。 |
| 文本 | 通过书写词语、符号和数字表示的内容。 |
| 多种内容形式 | 含三种或三种以上形式的混合内容。 |
| 其他内容形式 | 如果上列词语都不适用于受编资源，则使用“其他内容形式”。 |

* **内容限定**
* **类型说明**

|  |  |
| --- | --- |
| 测绘型 | 以任何比例表示地球或任何天体整体或部分的内容。 |
| 记谱型 | 为艺术目的（例如，音乐、舞蹈、舞台演出）通过符号系统以视觉感知表示的内容。 |
| 表演型 | 在一定时间、在资源上记录、以听觉形式或视觉形式表示的内容。 |

* **运动说明**

|  |  |
| --- | --- |
| 动态 | 通常通过快速的连续图像，感知动态的图像内容。 |
| 静态 | 感知静态的图像内容。 |
| 注：感知资源图像内容中有无运动，内容形式为“图像”时使用。 | |

* **维数说明**

|  |  |
| --- | --- |
| 二维 | 以二维形式感知的图像内容。 |
| 三维 | 以三维形式感知的图像内容。 |
| 注：感知资源图像内容的空间维数。若从内容形式无法推断这类内容限定时，应该使用上表所列词，内容形式为“图像”时使用。 | |

* **感官说明**

|  |  |
| --- | --- |
| 听觉 | 通过听觉感知的内容。 |
| 味觉 | 通过味觉感知的内容。 |
| 嗅觉 | 通过嗅觉感知的内容。 |
| 触觉 | 通过触觉感知的内容。 |
| 视觉 | 通过视觉感知的内容。 |
| 注：这个属性是指感知内容的人类感官。若从内容形式无法推断这类内容限定时，应该使用上表所列词。 | |

* **媒体类型词**

|  |  |
| --- | --- |
| **媒体类型词** | **适用的载体类型** |
| 音频 | 可用音频播放器播放的资源。 |
| 电子 | 计算机可用的资源。 |
| 缩微 | 可以使用缩微品阅读器的资源。 |
| 显微 | 使用显微镜的资源。 |
| 投影 | 使用投影仪的资源。 |
| 立体 | 可以使用立体观察器的资源。 |
| 视频 | 可以使用视频播放器的资源。 |
| 多媒体 | 用于三种或三种以上媒体类型适用的混合载体资源。 |
| 其他媒体 | 如果上列的词不适用于媒体类型和观看、使用或感知被著录资源内容需要的中介设备，则著录“其他媒体”这个词。 |