西安电子科技大学

信息组织实验课程实验报告

实验名称 都柏林核心元数据方案

经济与管理学院 2106011 班

成 绩

姓 名：来玉凯 学号：21069100140

姓 名：胡家苇 学号：21069100047

姓 名：吴南江 学号：21069100024

姓 名：韩铄玲 学号：21069100196

姓 名：赵红玉 学号：210619100225

实验日期 2024 年 6 月 1 日

|  |
| --- |
| 指导教师评语：  指导教师：  年月日 |

一、实验目的

（1）掌握都柏林核心元数据方案的基本概念和元素。

（2）应用DC元数据方案对不同类型的数字资源进行标引。

（3）理解元数据在信息组织和检索中的重要性。

二、实验环境

操作系统：macOS

在线XML验证工具：Code Beautify XML Validator: [https://codebeautify.org/xmlvalidator](https://codebeautify.org/xmlvalidator" \t "/Users/zhao/Documents\\x/_blank)

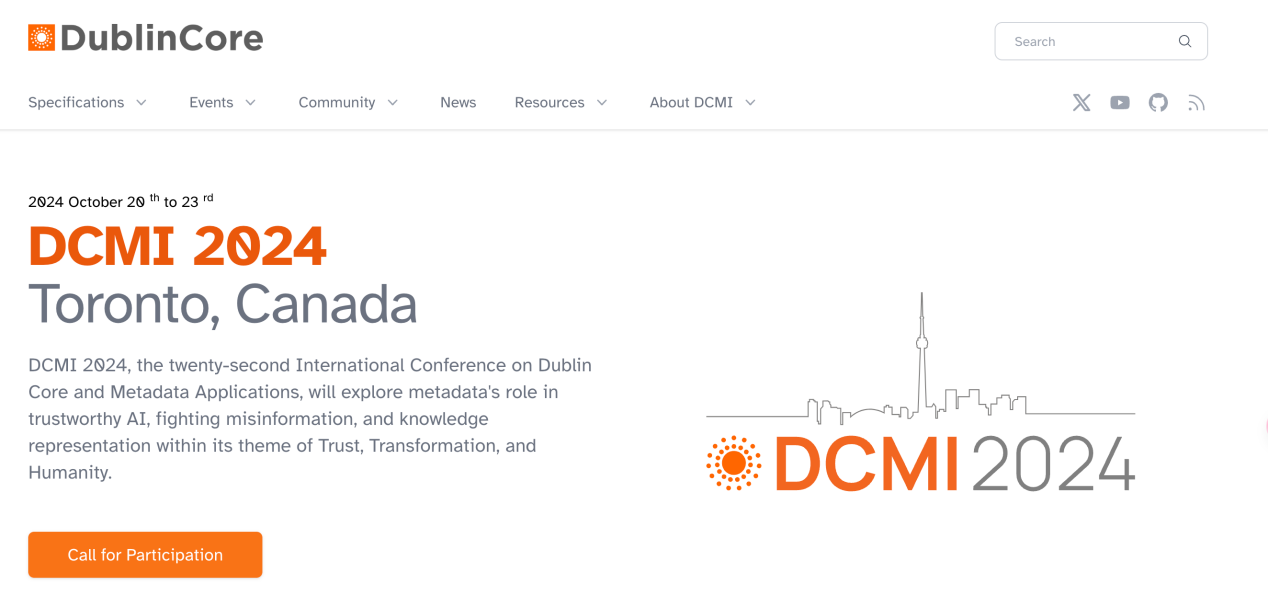
三、实验题目

都柏林核心元数据方案

四、实验操作

（1）熟悉都柏林核心元数据方案。

1.查阅dublin core官方网址（https://www.dublincore.org/），了解都柏林核心元数据方案的定义、发展历程和基本原则。



1.1定义

都柏林核心元数据方案(Dublin Core Metadata Initiative,DCMI)是一套用于描述数字资源的标准化元数据元素集合。它提供了一个核心元素集，可以用于描述各种类型的资源，例如文本、图像、音频、视频等。DC元数据方案旨在促进信息资源的发现、检索和互操作。

1.2发展历程

1995年，DCMI项目在美国俄亥俄州都柏林市启动，旨在为网络资源开发一套简单的元数据标准。

1999年，发布了第一版DCMI元数据术语，包含15个核心元素。

2003年，DCMI成立了一个正式的组织，负责维护和发展DC元数据标准。

2008年，发布了DCMI抽象模型，为DC元数据方案提供了更清晰的理论基础。

至今，DCMI持续发展和完善DC元数据方案，并推广其应用。

1.3基本原则

简单性。DC元数据方案使用简单的术语和结构，易于理解和使用。

通用性。DC元数据方案可以用于描述各种类型的数字资源。

可扩展性。DC元数据方案可以根据需要进行扩展，以满足特定领域的应用需求。

互操作性。DC元数据方案采用标准化的术语和结构，促进信息资源的共享和交换。

国际化。 DC元数据方案支持多语言和多文化环境。

1. 学习15个核心元素的含义和使用方法，并理解其应用场景。

2.1 15 个核心元素

|  |  |  |
| --- | --- | --- |
| **元素名称** | **含义** | **用法示例** |
| ****标题**** | **资源的名称** | **图书：《傲慢与偏见》** |
| ****创作者**** | **创建资源的人或组织** | **图书：《傲慢与偏见》的作者：简·奥斯汀** |
| ****主题**** | **资源的主题或主题** | **图书：《傲慢与偏见》的主题：爱情、阶级、社会** |
| ****描述**** | **对资源的简要描述** | **图书：《傲慢与偏见》的描述：一部关于爱情、阶级和社会地位的经典小说** |
| ****发布者**** | **发布资源的人或组织** | **图书：《傲慢与偏见》的出版商：企鹅出版社** |
| ****贡献者**** | **除了创建者之外，对资源创作做出贡献的人或组织** | **图书：《傲慢与偏见》的贡献者：编辑、校对员、插画家** |
| ****日期**** | **资源的创建日期或发布日期** | **图书：《傲慢与偏见》的出版日期：1813 年** |
| ****类型**** | **资源的类型，例如书籍、文章、图像或音频文件** | **图书：《傲慢与偏见》的类型：小说** |
| ****格式**** | **资源的物理或数字格式，例如纸质书、PDF 文件或 MP3 文件** | **图书：《傲慢与偏见》的格式：平装本** |
| ****标识符**** | **用于唯一标识资源的字符串，例如 ISBN 或 DOI** | **图书：《傲慢与偏见》的 ISBN：978-0141439518** |
| ****来源**** | **资源的来源，例如网站或数据库** | **图书：《傲慢与偏见》的来源：Project Gutenberg** |
| ****语言**** | **资源使用的语言** | **图书：《傲慢与偏见》的语言：英语** |
| ****关系**** | **指向相关资源的链接** | **图书：《傲慢与偏见》的相关资源：其他简·奥斯汀的小说** |
| ****覆盖范围**** | **资源涵盖的时间或空间范围** | **图书：《傲慢与偏见》的覆盖范围：19 世纪初的英国** |
| ****权利**** | **使用资源的权利信息** | **图书：《傲慢与偏见》的权利信息：受版权保护** |

2.2应用场景

（1）数字图书馆和档案馆。

用于描述馆藏资源，方便检索和发现。

（2）内容管理系统。

用于组织和管理网站内容。

（3）社交媒体。

用于丰富社交媒体帖子。

（4）教育资源库。

用于描述学习资源，方便学生查找。

（5）研究数据。

用于描述研究数据，便于共享和分析。

3.研究DC元数据方案的优势。

（1）简单易用。

DC元数据方案由15个核心元素组成，易于理解和使用，即使是非技术人员也能轻松掌握。

（2）广泛适用。

DC元数据方案可用于描述各种类型的资源，包括书籍、文章、图像、音频、视频、网站等。

（3）可扩展。

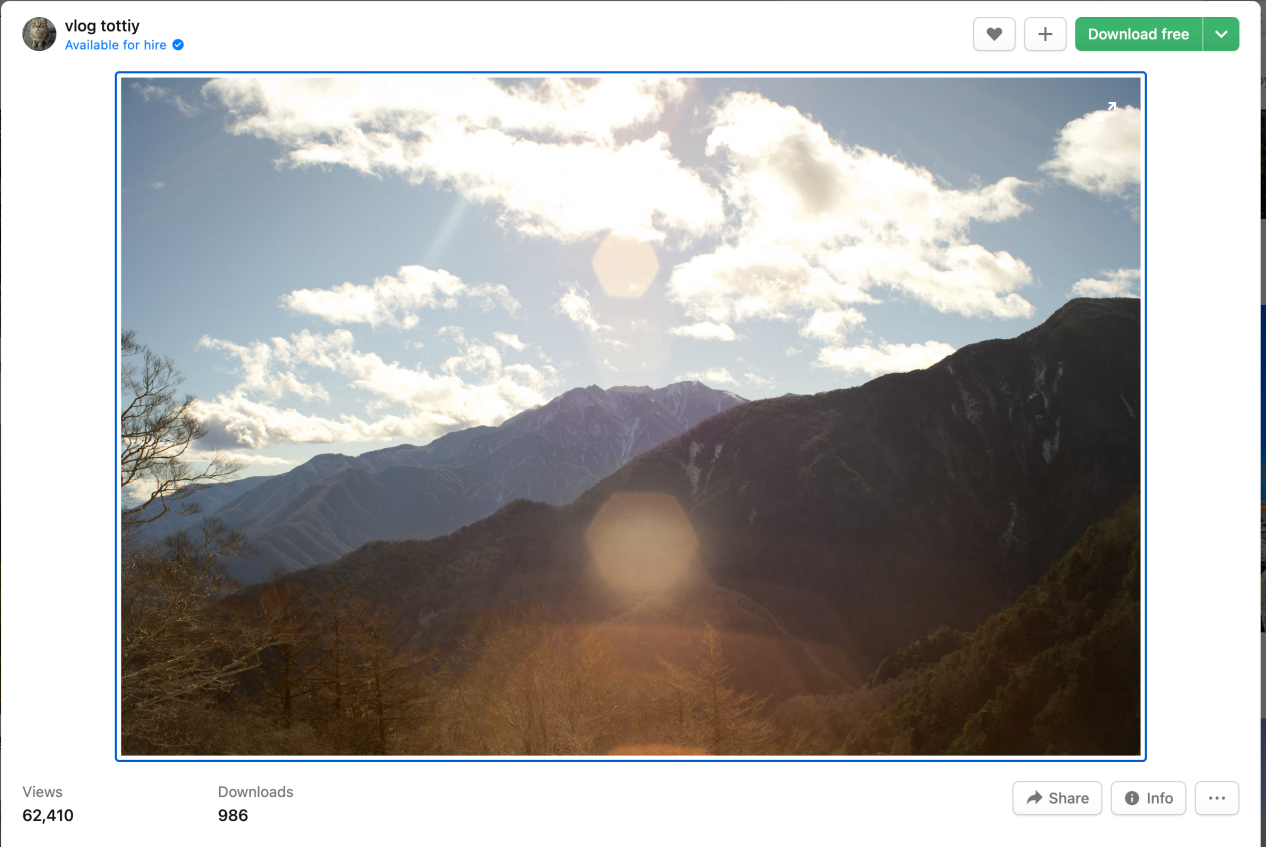
DC元数据方案可以通过扩展词汇表来满足特定需求，例如添加新的元素或修改现有元素的定义。

（4）互操作性。

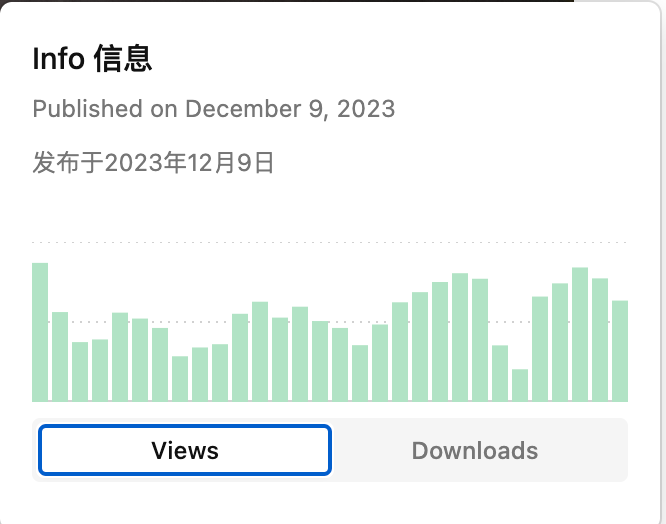
DC元数据方案与许多其他元数据方案兼容，可以轻松地在不同系统之间共享和交换资源信息。

（2）选择两类信息资源进行DC元数据标引。

1.图像资源



资源描述：一张风景照片，展现了山脉日落的美景。



DC元数据:

|  |  |
| --- | --- |
| **Title** | **日落山景** |
| **Creator** | **vlog tottiy** |
| **Subject** | **风景,山脉,日落** |
| **Description** | **一张美丽的山脉日落照片，展现了金色的阳光洒在山峰上的景色。** |
| **Date** | 2023-12-09 |
| **Type** | **Image** |
| **Format** | **JPEG** |
| **Identifier** | 山脉 |
| **Rights** | Unsplash网站 |

DC元数据(XML格式):

<?xml version="1.0" encoding="UTF-8"?>

<dc:metadata xmlns:dc="http://purl.org/dc/elements/1.1/"

xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"

xsi:schemaLocation="http://purl.org/dc/elements/1.1/ http://purl.org/dc/elements/1.1/schema.xsd">

<dc:title xml:lang="en">Mountain Sunset</dc:title>

<dc:creator xml:lang="en">vlog tottiy</dc:creator>

<dc:subject xml:lang="en">Landscape, Mountains, Sunset</dc:subject>

<dc:description xml:lang="en">A beautiful photo of a sunset over mountains, showing the golden sunlight cascading down the peaks.</dc:description>

<dc:date>2023-12-09</dc:date>

<dc:type xml:lang="en">Image</dc:type>

<dc:format>JPEG</dc:format>

<dc:identifier xml:lang="en">Mountains</dc:identifier>

<dc:rights xml:lang="en">Unsplash</dc:rights>

<dc:description xml:lang="en" dc:description.type="caption">

A group of hikers standing on a mountain ridge, admiring the breathtaking sunset view. The sky is ablaze with hues of orange, pink, and purple, while the mountains are silhouetted against the vibrant backdrop. The hikers' faces are lit up with a sense of awe and wonder, capturing the essence of this unforgettable moment.

</dc:description>

</dc:metadata>

资源类型2：文本资源

资源描述：一篇关于新冠疫苗研发进展的新闻报道。

（http://finance.people.com.cn/n1/2022/1013/c1004-32544401.html）



DC元数据:

Title: 多条技术路线并行 我国新型新冠病毒疫苗研发取得这些新进展

Creator: 田晓航、宋晨

Subject: 新冠疫苗,医学研究,公共卫生

Description: 一篇关于新冠疫苗研发最新进展的新闻报道，介绍了疫苗的研发过程、临床试验结果以及未来应用前景。

Publisher: 新闻机构名称

Date: 2022年10月13日08:53

Type: Text

Format: HTML

Identifier: http://finance.people.com.cn/n1/2022/1013/c1004-32544401.html

Language: 中文

Rights: 新华网

DC元数据(XML格式):

（3）验证元数据:

<dublincore xmlns="http://purl.org/dc/elements/1.1/">

<title>多条技术路线并行 我国新型新冠病毒疫苗研发取得这些新进展</title>

<creator>田晓航、宋晨</creator>

<subject>新冠疫苗,医学研究,公共卫生</subject>

<description>一篇关于新冠疫苗研发最新进展的新闻报道，介绍了疫苗的研发过程、临床试验结果以及未来应用前景。</description>

<publisher>新闻机构名称</publisher>

<date>2022-10-13T08:53:00Z</date>

<type>Text</type>

<format>HTML</format>

<identifier>http://finance.people.com.cn/n1/2022/1013/c1004-32544401.html</identifier>

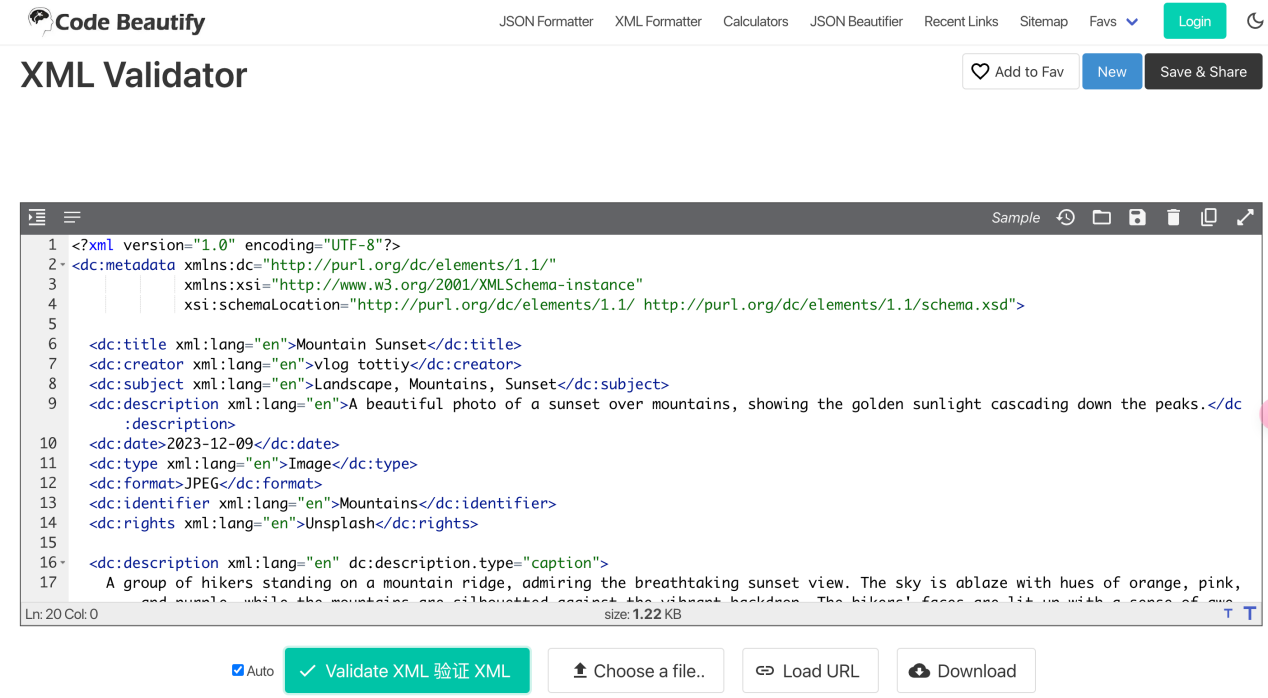
<language>zh-CN</language>

<rights>新华网</rights>

</dublincore>

使用在线XML验证工具或浏览器插件验证生成的XML文档是否符合语法规范。

图像资源：



文本资源：



五、实验体会和实验中存在的问题

实验体会:

通过本次实验，加深了对都柏林核心元数据方案的理解，并掌握了使用DC元数据方案进行资源标引的基本技能。

体会到元数据在信息组织和检索中的重要作用，能够帮助用户快速找到所需信息，并了解资源的基本属性。

意识到元数据标准化和规范化的重要性，有利于信息资源的共享和互操作。

实验中存在的问题:

对某些DC元素的理解和应用不够深入，例如“关系”和“覆盖范围”元素。

手工编写XML代码容易出错，需要仔细检查和验证。

缺乏对元数据工具的了解和使用经验，需要进一步学习和实践。

改进措施:

深入学习和理解DC元素的含义和应用场景，并进行更多实践。

学习使用元数据编辑工具，提高标引效率和准确性。

探索其他元数据方案，并比较其优缺点。