

数字对象语义关联组织的典型模型研究*

马雨萌^{1,2} 祝忠明¹

¹(中国科学院国家科学图书馆兰州分馆 兰州 730000)

²(中国科学院大学 北京 100049)

【摘要】在语义和数据驱动的开放关联环境下,数字仓储、数字图书馆等领域为支持语义计算能力,需要语义化揭示组织数字对象及其关联关系,实现知识资源的语义关联组织。研究 Fedora 通用数字对象模型、Europeana 的 EDM 模型、科研信息系统 CERIF 模型和研究对象模型 4 种数字对象语义关联组织的典型模型,并对这 4 种模型进行比较分析和选型建议。

【关键词】数字对象 语义关联组织 Fedora 通用数字对象模型 EDM 模型 CERIF 模型 研究对象模型

【分类号】G250.76

Research on Representative Semantic Models for Linking and Organizing Digital Objects

Ma Yumeng^{1,2} Zhu Zhongming¹

¹(The Lanzhou Branch of National Science Library, Chinese Academy of Sciences, Lanzhou 730000, China)

²(University of Chinese Academy of Sciences, Beijing 100049, China)

【Abstract】In the linked open environment driven by semantic and data, digital repositories, digital libraries and other domains need semantic digital objects and their linked relations to support semantic computing, making resources semantic linked and organized. This paper analyses Fedora, EDM, CERIF and Research Object which are representative semantic models for linking and organizing digital objects, and makes a comparative analysis and recommendation of model selection.

【Keywords】Digital objects Semantic linking and organizing Fedora EDM CERIF Research Object

1 引言

在语义和数据驱动的开放关联环境下,构建智能化的数字信息基础设施,提供语义计算能力,实现科学研究过程中科学数据的自动关联分析、以及相关服务和功能的智能组合,是科研和学术环境衍化的基本方向和趋势^[1]。为了构建和发展这种 e-Science 下新型学术交流模式的语义计算环境,需要实现数字对象的语义关联组织,从而支持基于语义关联的科学知识发现、检索、浏览、分析、重用、可视化等面向科学研究的服务和应用。数字对象模型的构建为语义关联组织数字对象提供了基础框架,通过对数字对象及其关系进行有效的语义化揭示、分析与组织,使大量科研知识对象成为富语义和可计算的资源,增强对资源语义关系的捕获和表达能力。

目前,数字仓储、数字图书馆、科研信息化管理以及数字化科研等领域关于数字对象语义关联组织模型的研

收稿日期:2012-12-25

收修改稿日期:2013-01-15

* 本文系中国科学院知识创新工程重要方向项目“研究所机构知识库建设”和中国科学院西部之光联合学者项目“机构知识库的语义增强方法与技术研究”的研究成果之一。

究,采用元数据描述、本体建模、关联数据等语义网技术,通过数字对象类型和关联关系的识别、语义结构的封装、知识属性和关系属性的语义描述、情境关系的捕获、关联组织框架的构建等方法,语义表达数字对象关联组织的需求,为这些应用领域提供了底层数字资源语义组织管理的模型。本文从模型的应用背景和目的、模型所构建的实体及关系框架、模型的实现方法、模型的扩展应用等角度,对目前在这些应用领域中有代表性的几种数字对象语义组织模型进行分析研究,包括 Fedora 的通用数字对象模型、Europeana 的 EDM 模型、euroCRIS 的 CERIF 模型、以及研究对象模型,并在比较分析这 4 种数字对象模型的基础上,为数字对象语义关联组织研究提供选型建议,以期为语义化知识环境及应用服务的构建提供借鉴和参考。

2 数字对象语义关联组织的典型模型

2.1 Fedora 通用数字对象模型

Fedora (Flexible Extensible Digital Object Repository Architecture) 是康奈尔大学在美国国家科学基金会 (National Science Foundation, NSF) 和美国国防部高级研究项目机构 (Defence Advanced Research Projects Agency, DARPA) 共同资助下进行的关于复合数字对象模型的研究项目^[2],旨在为数字内容、资产管理和数字资源保存等提供灵活可扩展的数字对象仓储结构。Fedora 定义了一系列表示数字对象的概念,明确了数字对象间的关系,以及连接数字对象的行为,Fedora 项目利用一个通用的数字内容管理系统来实现这些 Fedora 概念。数字对象模型是 Fedora 系统的核心,提供了一种描述、管理、传递和保存复杂数字内容的通用框架,其复合数字对象结构、模板化内容模型框架以及采用语义网技术对数字资源间复杂结构、语义、源流和管理关系的揭示,实现了数字对象模型的灵活和可扩展,使得数字仓储适应日益多样化和复杂化的数字内容管理需求。

Fedora 数字对象结构^[3]采用复合数字对象的方案,聚合了数字内容,其基本组成内容包括数字对象唯一标识符、对象属性和一系列数据流。数字对象通过数据流记录和封装数字对象的内容、元数据、关系描述等信息,其中 DC 数据流用于包含对象的元数据,AUDIT 数据流记录数字对象的所有变化。Fedora 通过数

字对象关系元数据的形式表示数字对象的关联关系,并以关系数据流的形式存储在数字对象中,RELS-EXT 数据流主要用于描述与其他数字对象的关系,RELS-INT 数据流则用来描述数字对象的内部数据流之间的关系。为了构建富语义的数字知识网络,Fedora 定义了一个基本语义关系本体来支持数字对象网络中常见关系的规范表示,包括组成关系 (hasConstituent)、成员关系 (hasMember)、衍生关系 (hasDerivation)、依赖关系 (hasDependent)、等同关系 (hasEquivalent)、描述关系 (isDescriptionOf)、标注关系 (isAnnotationOf) 等^[4],揭示了数字资源间的结构关系、源流关系和管理关系等语义关系,为数字仓储中知识网络的描述、操作和查询等问题提供了开放关联和可扩展的解决方案。此外,Fedora 可在基本关系本体的基础上,复用外部的领域本体或词汇,实现多种复杂形式关联关系的扩展^[5]。通过这种定义数字资源之间语义关系和对象之间层次关系的灵活可扩展机制,Fedora 可支持复合数字对象组成部分的重用,在互相关联资源的基础上构建数字对象网络,满足各种类型复杂数字内容的结构和语义关系的表示与组织管理。

为了提供模板化组织管理数字对象的方式,Fedora 通过引入内容模型框架 (Content Model Architecture, CMA)^[6]的概念和方法为具有相似特征的一类数字对象构建内容模型。在 CMA 中,内容模型用于定义数字内容保存和传递过程所需要的所有可能的特征,如结构、行为和语义信息,以及与其他数字对象或实体间的各类关系,为各种类型数字资源提供了一种可重用的模板,从而减少创建、摄取、存储、管理、保存、转换和获取数字内容过程中的重复工作,支持内容的分类管理,促进其分享和发现利用等。

CMA 提供了一个集成内容和行为操作的框架,改变了之前将数据、元数据和行为操作聚合在一起的数字对象结构,定义 4 类数字对象来描述数字资源的内容类型及其实现的操作行为,数字对象及其关系如图 1 所示。

数据对象 (Data Object) 表示所要组织管理的数字资源,用于存储数字内容实体;内容模型对象 (Content Model) 是数据对象的类,“hasModel”关系表示属于数据对象所遵循的内容模型;服务定义对象 (Service Definition Object) 和服务配置对象 (Service Deployment Ob-

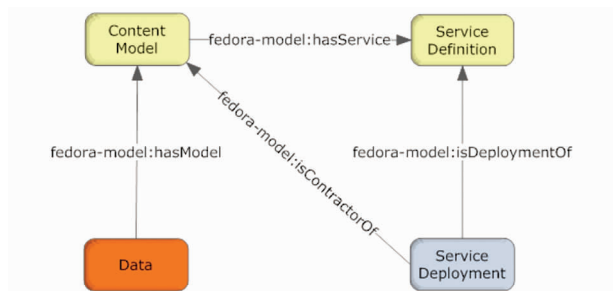


图 1 Fedora 内容模型框架中数字对象的关系^[6]

ject) 实现了内容模型对象的服务, 服务定义对象是 Fedora 中用于存储服务模型的控制对象, 定义了一个服务及其相关操作集合, 通过“hasService”关系描述内容模型对象提供的服务及相关操作, 而服务配置对象表明了内容模型对象服务的具体实现方式, 描述了 Fedora 如何执行服务定义的操作。CMA 这种数字资源表现形式和内容类型分离、表现形式和操作行为分离的机制, 不仅使得 Fedora 数字对象可遵循描述其特征的丰富内容模型, 而且提供了决定数字对象操作方法的多样行为机制, 实现各种类型数字资源及其表现形式的可定制的描述模式, 为 Fedora 数字对象的灵活扩展和重用提供基础。

作为实施知识产权长期保存和支持科研成果开放共享的主力, 数字仓储目前向基于知识语义关联的应用方向发展。面对不同类型的数字资源, Fedora 通过复合数字对象结构和内容模型框架构建了灵活可扩展的数字对象模型, 通过捕获数字对象的关联关系, 增强了数字对象语义关系的表达能力, 使其在保存管理数字内容的研究中得到越来越多的应用和扩展。

2.2 EDM 模型

为促进欧洲范围内知识文化遗产资源的共享利用, 欧盟推动建立了欧洲数字图书馆 Europeana, 提供访问和利用欧洲知识文化遗产资源的集成门户^[7]。在数字化知识文化遗产资源时, 欧洲范围内图书馆、档案馆、博物馆等机构采用了不同的元数据标准, Europeana 这种跨文化、多情境的环境要求合理组织和管理这些异构的元数据或描述信息。ESE (Europeana Semantic Elements) 数据模型的扁平结构, 造成数据提供者的元数据在映射过程中语义损失。为了有效组织管理来自不同数据提供者对文化遗产资源的描述信息, Europeana 对 ESE 数据模型进行改进升级, 采用资源聚合结

构建立了 EDM (Europeana Data Model) 数据模型, 实现对来自跨领域异构知识文化遗产元数据的有效融合、集成、重用与关联组织。

EDM 通过复用 OAI - ORE 构建了聚合模型中的资源类和情境实体类, 并捕获了类关系属性, 聚合结构使得数据提供者能够以自身的描述标准构建数据, 并且 Europeana 为这些不同来源的数据增加语义描述, 支持来自不同数据提供者的元数据的语义增强, 丰富文化遗产对象的语义描述。因此, EDM 模型包括基本结构和语义增强的扩展结构, 构建了灵活的类的层级结构以支持知识文化遗产资源的描述与组织, 包括 EDM 自身的类和从其他名称空间复用的类^[8]。EDM 基本聚合结构有三个资源核心类: 描述不同来源数据的聚合 (Aggregation), 及其所包含的文化遗产对象 (Provided Cultural Heritage Object, ProvidedCHO) 和该对象的数字形式 (Web Resource)。Europeana 需要从不同来源集成数据, 然而对于同一资源对象, 不同数据提供者采用的元数据标准和描述方式存在差异, 为了区分这些不同来源的描述, 记录数据的源流信息, EDM 引入代理 (Proxy) 表示一个特定提供者对某个 ProvidedCHO 的描述, 使用属性 ore: proxyIn 关联该提供者的聚合, 使用属性 ore: proxyFor 连接所代理的 ProvidedCHO。同时, 为了标准化和丰富文化遗产对象不同来源的描述, Europeana 对象模型使用 edm: EuropeanaAggregation 聚合类创建自身的聚合, 为文化遗产对象增加 Europeana 描述, 聚合每个数据提供者对同一对象的描述^[9]。EDM 的这种聚合结构, 能够兼容多种领域标准, 如博物馆的 LIDO、档案馆的 EAD 或数字图书馆的 METS, 实现了博物馆、档案馆和数字图书馆等领域内资源的集成关联^[7]。

为了捕获文化遗产对象的相关情境信息, 支持语义增强建模, EDM 构建了表示对象相关的情境实体 (Contextual Entities) 类来支持情境关系的建立, 包括主体 (Agent)、地点 (Place)、概念 (Concept)、事件 (Event)、自然实体 (Physical Thing)、时间范围 (Time Span)。这些情境实体本身可以拥有自身的元数据, 并使用控制词表进行规范描述, 提供细致深入的信息, 实现对象与相关的人、地点、主题等数据信息的关联, 例如, Agent 通过 VIAF 权威记录, 提供了更多有关 ProvidedCHO 作者的信息^[10]。

EDM 模型复用相关本体或词汇,通过层级结构的 EDM 属性关系(如图2)描述如下类型的资源关联:

(1)结构组成关系:核心聚合结构一方面复用 OAI-ORE 的属性 ore:aggregates,构建其子属性 aggregatedCHO 描述 Aggregation 和 ProvidedCHO 的聚合关系;另一方面将 Aggregation 和文化遗产对象的呈现形式(hasView)、所在数据提供者网站(landingPage)关联。

(2)主题类型关系:主题(dc:subject)、类型(hasType)、共现资源(hasMet)属性表示 EDM 对象主题类型描述的特征。

(3)情境关系:属性 occurredAt、happenedAt 和 wasPresentAt 提供关于主体、地点、时间范围等情境实体的描述信息。

(4)文化遗产对象间的关系:衍生(isDerivativeOf)、后继(isSuccessorOf)、包含(incorporates)等关系属性揭示了对对象间的相似关系(isSimilarTo),邻接关系(isNextInSequence)用来表示期刊等出版物序列的顺序,还包括自然实体对信息资源的实现(realizes)和资源间表现关系(isRepresentationOf)。

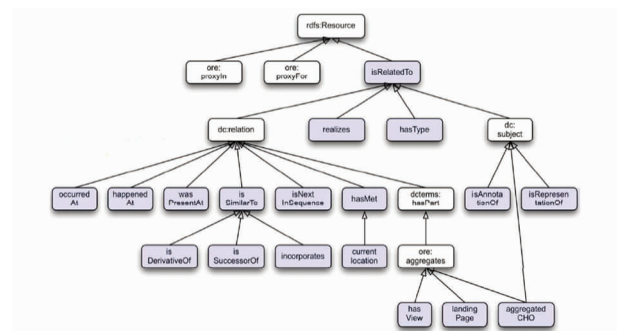


图2 EDM 属性的层级结构^[10]

EDM 遵循语义网和关联数据原则,语义化描述数字资源的聚合结构和关联关系,因此 EDM 对象易于根据关联数据原则赋予 HTTP URIs,并通过与外部 GeoNames 地点数据、DBpedia 人物数据等关联数据源的语义关联^[11],连接外部的其他项目和机构,为欧洲文化遗产数据增加更多有意义的关联关系。作为 Europeana 的一个更加灵活和明确的新数据模型,EDM 满足了跨领域知识文化遗产资源的整合和集成利用的需求,在促进 Europeana 构建语义化、网络化数据环境的同时,也为开放关联环境下数字图书馆等相关领域关联聚合资源的研究提供了借鉴。

2.3 CERIF 模型

科研信息与科学研究环境中的角色相关,能够支持科研活动的组织、管理和计划,以及科学研究过程中日益增强的合作交流,因此需要对科研信息进行组织管理和共享。CERIF(Common European Research Information Format)^[12]是 euroCRIS 推荐给其成员国的支持科研信息管理的概念模型标准,作为表示科研环境中有关实体及实体间关系的底层数据模型,应用于科研信息系统(Current Research Information System, CRIS)的建设,便于科研信息的交换和共享。CERIF 数据模型中丰富的实体、灵活的关系管理、XML 交换格式和 CERIF 语义,使其成为构建可扩展的科研信息系统的强大工具。

CERIF 从科研管理流程出发,分析和明确了研究计划、项目、成果等各个环节^[13]所涉及的主要科研实体及其科研关系,以支持科研信息的集成共享管理,如图3所示:

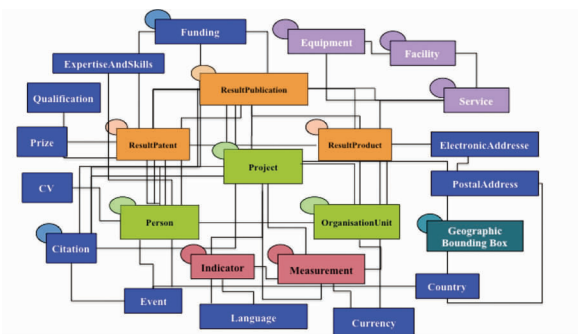


图3 CERIF 模型中的实体及实体间的关系^[12]

CERIF 的实体框架围绕人(Person)、组织单元(OrganisationUnit)和项目(Project)这三个基本实体(Base Entities)展开,他们为描述研究者及其所处科研环境开展科研活动时产生的各种交互提供基础,整个科研过程其他科研实体通过与基本实体的关联支持科研活动。基本实体产出包括出版物(ResultPublication)、专利(ResultPatent)和产品(ResultProduct)等成果实体(Result Entities)。通过成果和设备等基础设施实体(Infrastructure)的关联,CERIF 支持捕获科研产出的产生源流关系,同时为了使得科研信息系统能够提供科研评价服务,CERIF 构建评价实体(Measurement)和指标实体(Indicator)来扩展定量评价能力,这些二级实体(2nd Level Entities)通过与基本实体、成果实体的关联,描述了科学研究的情境。

CERIF 通过构建关联实体来表示实体之间的关联关系,由于科研实体在所处的科研场景发挥着特定的角色功能,并产生相应的科研关系,为实现科研领域一般关系和类型词汇的标准化,CERIF 使用标准词汇和术语构建语义层,定义和描述了科研实体在研究环境中所涉及的关系和角色,形成类的语义值和所属的类别体系,并以类别标识符和类别体系标识符标识,例如组织单元间的部分整体关系 (hasPart) 属于组织结构 (Organisation Structure) 类别体系,作者 (Author) 则属于人在出版物中所处角色 (Publication - Person Roles) 的类别体系。CERIF 关联实体以记录实体标识符的方式将两个实体相互关联,通过类别标识符和类别体系标识符关联至语义层,参考 CERIF 语义层中实体关系或角色类型的语义值^[14]。CERIF 模型构建的关于科研实体类型和角色的 CERIF 语义,为捕获科研信息管理中科研实体的功能和角色关系提供了标准化的方法。

2.4 研究对象模型

随着数据密集型计算模式逐步成为科学研究的主流范式,研究者希望能够发布、查找、共享科学研究过程中一切有价值的学术资料,包括研究方法、工具和过程。因此,支持广泛复杂类型数字知识的关联组织和结构化重用成为科学记录和学术交流出版系统适应这种研究范式演化的基本要求。显然,传统线性结构的学术出版模式使得科研过程的知识出版和传播以不可分割的单位保存,难以支持科研产出及研究方法关联重用。研究对象 (Research Object, RO) 模型从科研过程的角度,提取并集成研究数据、情境信息和研究方法等科学知识对象,通过数字知识的封装,提供了可重用科研知识的组织、发现和分享机制。

RO 是支持科学工作流分享的社交平台 myExperiment 中包 (Pack) 概念的扩展延伸^[15],封装了有关科研过程的重要信息,包括研究数据、科研方法、科研成果产出、涉及的人和机构等实体、行为操作和细节描述等注释信息以及源流、衍生等关联关系,是一种资源语义聚合的知识单元^[16]。为了实现研究学术价值链上各个环节知识产出的共享、重用和增值,RO 模型并不仅限于为科学工作流等某类科研资源提供聚合结构,还面向整个研究过程,组织管理有价值的各类科研知识产出,以多种方式支持科研知识的重用。RO 能够作为可重用 (Reusable) 的对象整体在新实验或其他 RO 中

重用,而且 RO 的聚合结构描述了组成部分之间的关系,这些组成部分可作为可重用的 (Repurposeable) 对象在研究过程中进行重用,衍生出新的 RO。RO 通过对研究过程各方面实体信息及其关系的完整记录,使得所描述的研究成为可重复的 (Repeatable)、可再生的 (Reproducible)、可重现的 (Replayable) 和可揭示的 (Revealable) 研究。同时,为了支持科研资源的发现和检索,RO 也是需要是可参考 (Referenceable) 或可引用的^[17]。这些可重用方式是可共享对象所应表现的主要行为特征,为 RO 模型的定义提供了需求基础。

为满足以上多种可重用的功能需求,RO 模型需要实现一些支持这些功能的特征。一个 RO 模型应用示例如图 4 所示:

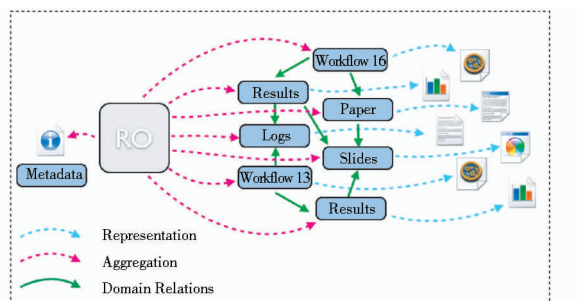


图 4 RO 模型应用示例^[16]

RO 作为一个整体可被注释或分享,在资源的基本聚合结构中,被聚合的数字对象表示了科研产出实体,聚合整体及其组成部分都可被唯一识别;RO 模型使用元数据描述 RO 中资源之间的关系,提供成果源流、衍生等关联关系的信息。RO 中捕获内部结构和资源间关系的元数据,为数字科研资源的重用提供了面向研究过程的全面视角,使得研究对象从一个简单的聚合结构成为一个可重用的对象,促进了相应附加值的产生,实现了资源重用和共享、支持科学研究的目标。

考虑到科研活动以特定序列生命周期的方式进行,RO 模型面向科研过程进行扩展,描述了科研生命周期不同阶段研究对象的状态以及执行的操作,使用版本记录对象的变化,实现其在生命周期中的动态管理。依据科研生命周期中涉及的研究活动,研究对象识别出表示资源聚合的常用类型,例如支持语义化出版的出版对象 (Publication Object)、表示科研过程中工作的动态对象 (Live Object)、记录研究方法的方法对象 (Method Object) 等^[16]。作为科研领域发布和组织

数据的方式,RO 框架提供的聚合结构可融入关联数据的语义环境中,依据关联数据规则,RO 本身作为一种资源发布为关联数据,另外研究对象可作为命名图聚合关联数据资源^[16]。RO 模型为数字化科研领域提供了一个关联组织研究过程中科研产出的框架,支持科研知识的重用和共享。

3 模型比较与选型建议

Fedora 通用数字对象模型、EDM 模型、CERIF 模型以及 RO 模型在资源语义关联组织方面各具特色,从模型的应用场景、构建标准、模型性能方面,对这 4 种数字对象模型进行比较分析,如表 1 所示,为数字对象语义关联组织模型的选择提供参考依据。

依据这 4 种数字对象模型在以上方面的对比分析,在选择数字对象模型支持数字资源语义关联组织研究时,需要考虑模型的应用领域和满足的场景需求,选择表达数字对象的模型结构,并分析所需的情境信息和所要揭示的关联关系,选型具体过程如下:

(1)若所构建模型是服务于数字化科研领域,则优先考虑 CERIF 和 RO,CERIF 为描述科研管理流程中的科研情境实体及其科研关系提供了参考,RO 则面向科学研究过程,支持科研产出及其关系的组织管理;开放关联环境下面向科学研究过程的数字化科研领域,应该对科研产出、科研情境实体以及这些实体间关联关系进行组织管理,CERIF 和 RO 的组合可以揭示科研过程的关键实体及其关系,为构建数字科研领域可重用的数字对象语义关联组织模型提供框架。若模型是应用于组织管理多种类型复杂知识对象的数字仓储等领域,可以采用 Fedora 等通用的数字对象模型,针对基于 Fedora 框架提供科研服务的仓储平台,可结合 CERIF 和 RO 扩展其数字对象模型。若模型需要服务于数字图书馆等领域有效组织管理大量异构资源时,可考虑 EDM 跨领域聚合和语义增强数据。

(2)明确模型应用需求后,考虑所构建模型是否要实现可扩展和可重用,是否特别关注数字对象的数据结构模型,以及模型所需兼容的元数据标准,从而决定是选择资源和关系描述封装的复合数字对象模型,还是资源和关系独立存在的实体关系模型,如果选择前者,则考虑采用 Fedora 和 RO 实现通用数字对象模型。

表 1 数字对象语义关联组织模型的比较分析

比较项	Fedora	EDM	CERIF	RO
适用领域	数字仓储	数字图书馆	科研管理	数字化科研
对象类型	复杂知识对象	文化遗产资源	科研信息	科研资源
实现目标	灵活、可扩展模板	兼容多领域元数据标准	科研情境描述	可重用富语义框架
数据模型	复合数字对象、资源和关系元数据封装在数字对象里	资源和关系属性	资源和关联实体独立	复合数字对象、资源和关系元数据封装在数字对象里
构建标准	DC 描述元数据,可复用 FRBR 等标准扩展关系本体	OAI - ORE 聚合资源、DC 描述元数据、SKOS 组织和表示概念词汇	DC 描述元数据	OAI - ORE 聚合资源、DC 描述元数据
关系捕获角度	数字对象之间关系	文化遗产资源聚合、组织、管理过程	科研管理流程	科学研究过程
可重用性	复合数字对象结构、模板化内容模型框架,模型重用较易	资源和关系类型针对数字图书馆组织管理文化遗产资源,模型重用较难	资源和关系类型面向科研管理,资源和关系独立存在的实体关系结构,模型重用较难	复合数字对象结构,模型重用较易;源流、衍生等关系的揭示,语义描述的资源重用较易
可扩展性	模板化框架,可定制,模型可根据所描述的资源类型扩展	模型可通过语义揭示出资源更多复杂的关联关系实现扩展	在改进 CERIF 语义和词汇,增加科研实体满足科研管理新需求等方面实现扩展	数字科研知识集成封装的框架,通过研究对象类型和更多科研关系的描述实现扩展
关系类型标准化	基本关系本体	层级结构属性	CERIF 语义层	缺乏
关系揭示复杂程度	数字对象网络中常见关系,较为简单	包括情境关系、衍生等对象间关系,较为复杂	科研角色和科研关系,较为复杂	科研产出及其源流、衍生等关联关系,较为复杂
情境描述的支持	缺乏	支持	支持	缺乏
关联数据应用的支持	缺乏	支持	正在研究	支持

(3)针对需要详细描述数字对象情境关系的应用领域需求,科研领域模型可参考 CERIF 模型,其他领域可从 EDM 情境类的角度描述服务于自身需求的场景信息。如果缺乏对模型应用领域关系类型的认识,可考虑关系类型标准化的数字对象模型,能够在选择模型关系框架的基础上,快速实现满足建模需求的关系扩展。

4 结 语

随着语义 Web 技术和方法体系逐步成熟,尤其在关联数据技术推动语义 Web 发展的背景下,科学研究过程的开放关联需要构建数字对象语义关联组织模型,实现新型学术交流模式的语义计算环境。虽然这 4 种典型数字对象语义关联组织模型目前在数字仓储、数字图书馆、科研信息系统、数字化科研领域实现了对资源的语义关联组织管理,但是面向开放关联环境下的科研领域,这些模型更多地是一种组织管理数字对象的普适框架,较少描述出符合科研产出特征的特殊属性及相应关系,缺乏基于科研过程具体场景深层次语义关联关系的揭示,其描述的关系主要还是聚合、结构等简单关系。因此,开放关联环境下科研数字对象语义关联组织模型的研究应面向研究过程各个阶段的具体场景,捕获并语义揭示整个科研过程中有价值的各类科研知识产出及其关联关系,支持科研过程和科研产出的关联交互,实现学术价值链上各个环节知识产出的共享、重用和增值。

参考文献:

- [1] Hey T. eScience, Semantic Computing and Open Science Towards a Smart Cyberinfrastructure of Software Services [EB/OL]. [2012 - 09 - 02]. http://www.powershow.com/view/f6b87-ZTQ2N/eScience_Semantic_Computing_and_Open_Science_Towards_a_Smart_Cyberinfrastructure_of_Software_Servic_flash_ppt_presentation.
- [2] The History of Fedora Repository [EB/OL]. [2012 - 09 - 02]. <http://www.fedora-commons.org/about/history>.
- [3] Fedora. Fedora Digital Object Model [EB/OL]. [2012 - 09 - 02]. <https://wiki.duraspace.org/display/FEDORA36/Fedora+Digital+Object+Model>.
- [4] Fedora. Digital Object Relationships [EB/OL]. [2012 - 09 - 02]. <https://wiki.duraspace.org/display/FEDORA36/Digital+Object+Relationships>.
- [5] Lagoze C, Payette S, Shin E, et al. Fedora: An Architecture for Complex Objects and Their Relationships [J]. *International Journal on Digital Libraries*, 2006, 6(2): 124 - 138.
- [6] Fedora. Content Model Architecture [EB/OL]. [2012 - 09 - 02]. <https://wiki.duraspace.org/display/FEDORA36/Content+Model+Architecture>.
- [7] Europeana. The Europeana Data Model for Cultural Heritage [EB/OL]. [2012 - 10 - 18]. <http://pro.europeana.eu/documents/900548/f495317b-4557-4a60-9326-723f4618b44c>.
- [8] Isaac A. The Europeana Data Model [EB/OL]. [2012 - 10 - 18]. <http://pro.europeana.eu/documents/900548/40262c74-d4a3-44d4-95ea-2f4146b50c92>.
- [9] Europeana. Europeana Data Model Primer [EB/OL]. [2012 - 10 - 18]. <http://pro.europeana.eu/documents/900548/770bdb58-c60e-4beb-a687-874639312ba5>.
- [10] Charles V. Europeana Data Model(EDM) and EDM for Libraries [EB/OL]. [2012 - 10 - 18]. http://www.jbi.hio.no/bibin/korg_dagene/korg2012/europeana_data_model_charles_korg2012.pdf.
- [11] Europeana. Technical Details [EB/OL]. [2012 - 10 - 18]. <http://pro.europeana.eu/tech-details>.
- [12] euroCRIS. CERIF 1.3 Full Data Model (FDM) Introduction and Specification [EB/OL]. [2012 - 10 - 25]. http://www.eurocris.org/Uploads/Web%20pages/CERIF-1.3/Specifications/CERIF1.3_FDM.pdf.
- [13] Jeffery K G. CRIS: Stakeholders, Benefits, History, Process, Architecture [EB/OL]. [2012 - 10 - 25]. http://www.eurocris.org/Uploads/Web%20pages/research_processCRIS/5Research%20processandCRIS.ppt.
- [14] euroCRIS. CERIF - 1.3 Semantics; Research Vocabulary [EB/OL]. [2012 - 10 - 25]. http://www.eurocris.org/Uploads/Web%20pages/CERIF-1.3/Specifications/CERIF1.3_Semantics.pdf.
- [15] De Roure D, Goble C. Lessons from myExperiment: Research Objects for Data Intensive Research [C]. In: *Proceedings of Bio Microsoft e - Science Workshop*. 2009.
- [16] Bechhofer S, Roure D D, Gamble M, et al. Research Objects: Towards Exchange and Reuse of Digital Knowledge [C]. In: *Proceedings of the Future of the Web for Collaborative Science (FWCS' 10)*, Raleigh, USA. 2010.
- [17] Bechhofer S, Ainsworth J, Bhagat J, et al. Why Linked Data is Not Enough for Scientists [C]. In: *Proceedings of the IEEE 6th International Conference on e - Science*, Brisbane, Australia. Washington, DC, USA: IEEE Computer Society, 2010.
- (作者 E - mail: mayumeng@mail.las.ac.cn)