

基于科学工作流的科研知识 产出管理和利用研究

赵迎光^{1,2} 马建霞²

(1. 中国科学院研究生院 北京 100049;
2. 中国科学院国家科学图书馆兰州分馆 兰州 730000)

摘要 随着数据密集型科研的发展,目前只关注科研产出末端的知识产出管理方式已经不能满足 e-Science 环境下对学术交流的需求。提出了从科学工作流入手构建科研知识产出管理模型的构想,通过分析国内外在科学工作流和科研知识产出方面所做的研究之后,发现其在科研过程描述方面不能满足我们的要求,因此构建了新的基于 OPM 的扩展的科学工作流对象模型。

关键词 科学工作流 科研知识产出 管理利用

中图分类号 G356

文献标识码 A

文章编号 1002-1965(2012)03-0151-06

Research on Management and Utilization of Output of Knowledge in Scientific Research Based on Scientific Workflow

ZHAO Yingguang^{1,2} MA Jianxia²

(1. Graduate University, Chinese Academy of Sciences, Beijing 100049;

2. The Lanzhou Branch of the National Science Library, Chinese Academy of Sciences, Lanzhou 730000)

Abstract With the development of data-intensive scientific research, the current management of knowledge output just focusing on the end of scientific research can not satisfy the academic communication in e-Science. There have been some programs which have done some researches in managing the output of knowledge in scientific research from the view of scientific workflow, however, their ability of describing the scientific research process is not satisfactory, and we construct an Extended Scientific Workflow Package(ESWP) based on OPM at last.

Key words scientific workflow output of knowledge in scientific research management and utilization

1 研究背景

1.1 数据密集的科研的到来 随着 e-Science 和网络技术的发展,以数据为中心的科学研究已经到来,科学数据的类型和规模急剧增多,导致了科研过程中的数据集成和发现困难重重;与此同时,目前的主要科研知识产出管理方式都集中于科研过程终端,科研前端和中间阶段大量知识产出并没有被纳入管理范围,加上共享机制不健全,造成了科研知识资产的重复生产和浪费。因此迫切需要在科研过程中采取新的信息组织形式来收集和管理数据,促进知识的发现、获取、分享和重用,为当前的科学研究范式转变^[1]提供支持和

保障。

1.2 科学工作流的发展 科学工作流^[2]是在 e-Science 环境下、在网络等先进信息基础设施平台上,利用广域分布在基础设施上的各种科学资源(主要是 Web 服务/网络服务),将具有大规模数据处理特征的科学研究过程集成起来,使科学研究过程自动化,进而加速、优化和转变科学研究过程;科学工作流系统为科研人员提供可视化的交互操作界面,负责设计、修改、运行、监控和重运行科学工作流,并对运行结果进行呈现和分析。

国外对科学工作流的研究和应用已经比较成熟, Yu^[3] 等人在 2005 年对主要的科学工作流系统进行统

收稿日期:2011-09-19

修回日期:2011-10-31

作者简介:赵迎光(1986-)男,硕士研究生,研究方向:信息检索与信息处理;马建霞(1972-)女,研究馆员,硕士生导师,研究方向:数字图书馆。

计分析时已经达到了 12 种之多;国内对科学工作流系统的研究应用相对较晚,2008 年中科院开始对科学工作流的使用进行培训^[4],通过对与科学工作流相关的文献分析,2008 年之后国内对科学工作流的研究明显增多,除了对科学工作流本身的研究^[5]之外,利用科学工作流来进行具体应用的实践也在不断发展,在天文学^[6]、海洋学^[7]等领域都有了具体的应用。但是和国外相比我国在这方面的应用还处于起步阶段,没有在各个领域中大规模广泛使用。尽管如此,在网格计算和跨学科的科研协作快速发展的今天,随着大量科学数据的不断产生,科学工作流技术在科研过程中的重要性将会不断增强,使用科学工作流系统来辅助科研和进行自动化数据处理将会是未来科研的一个发展方向^[5]。

科学工作流技术除了能够支持自动化的数据处理和科研过程之外,科学工作流描述语言可以清晰地描述科研流程,并能够在系统中以可视化形式呈现;同时科学工作流系统中的数据源流捕获技术可以捕获各个阶段的处理信息和输入输出信息,科研人员可以通过数据源流实现流程的监控和研究过程的再现。

1.3 问题的提出 在 e-Science 实践中,数字化的科研产出包括文献、数据、模型、软件、方法、工具、仪器设备使用手册、实验室记录、发表的论文、会议报告、音频、视频、成果的评价、成果的后期补充材料等。由于知识产出的类型以及在科研过程中所产生的阶段不同,通过对科研过程的分析,我们将其分为 3 个阶段:第一阶段是科研准备阶段,包括原始数据和资料的收集和整理;第二阶段是科学研究过程,包括实验的设计和执、数据的分析和转换、项目中的讨论和交流等;第三阶段是研究结束阶段,包括成果的形成和发布,图 1(左)是当前的科研知识产出的管理现状,可见在科研过程中知识产出流失严重。

因此,如何在科研知识产出管理中,结合科研流程,以语义丰富的方式实现相关的科学资源,即文献、

数据等的关联组织,并能够尊重科研人员的知识产权,提供灵活的访问授权机制,促进科学知识产出的有序化组织管理、促进科研知识产出的重用和保存,就成为目前亟待解决的问题。

而科学工作流为我们提供了解决这个问题的入口,随着科学工作流的发展,科学工作流技术日益成为科学研究过程的一个重要组件,利用科学工作流技术可以获取用于科学数据和引申产品的处理步骤,可以获取数据源流信息,让科学家以可重复的、可验证的、分布式的方式描述和开展实验过程,并能够跟踪错误、异常和误操作的来源。

对于我们所要研究的完整的科研项目流程来说,也需要实现各个阶段知识产出的捕获和管理以及利用,那么我们如何借鉴和使用科学工作流技术来实现科研知识产出的管理,将是本文的主要研究内容。图 1(右)是我们期望建立的基于科学工作流的科研知识管理的思路,图中所有的知识产出都被捕获,并且以有序的方式组织起来进入到出版交流领域。

2 相关研究项目

根据图 1(右)的思路,我们通过文献调研等方式对国内外相关的科研知识产出管理研究项目做了大量的分析调查,通过调查发现国外已经有部分研究开始关注科学工作流并将其应用到科研知识产出管理中。

JISC 和 microsoft 联合资助的科研协作项目 My-Experiment^[8]。但是 MyExperiment 将科学工作流作为一种知识资产,并且仅仅关注科学工作流本身,并不能够实现对所有科研知识产出的管理。

Elsevier 针对目前学术出版方式单一,科研过程中知识产出流失严重的情况,提出了 ExecutablePapers^[9]的设想,并在其网站公开征集解决方案。Bauer 等人^[10]通过目前的科学工作流工具实现了一个简单的 ExecutablePapers 原型。Bauer 将科学出版过程分为科学研究过程和出版过程两个阶段,在科学研究阶段,使

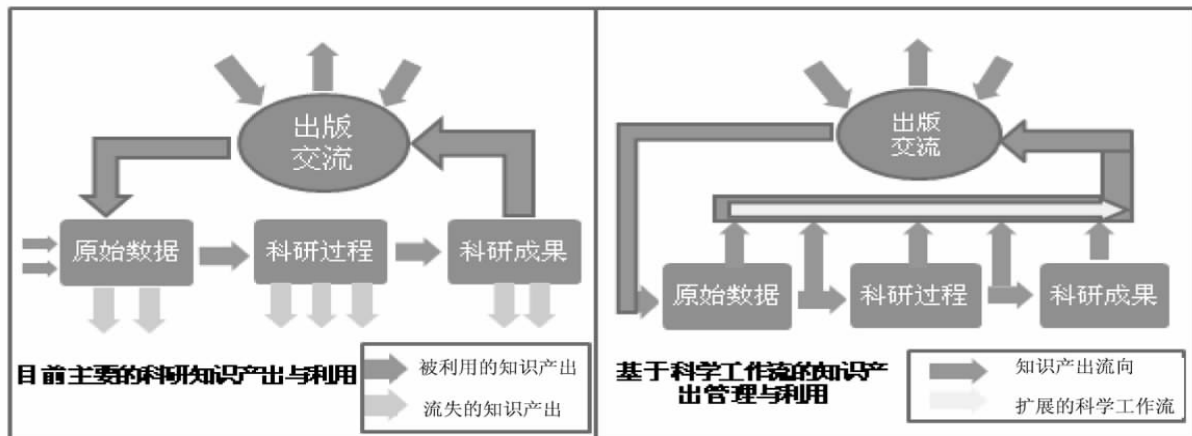


图 1 (左) 目前主要的科研知识产出与利用

(右) 基于科学工作流的知识产出管理与利用

用 VisTrails 工作流系统在数据处理过程中来获取数据源流信息;在论文出版过程中,用 XML 描述数据并作为附件和论文一起出版。该方法实现了数据和论文一起出版,并且根据数据源流来确定附加数据的语义信息,但是对于整个科研过程来说,该方法只能关注科研过程中可以使用科学工作流自动执行的阶段,而科学工作流执行过程之外科研知识产出并没有得到有效管理。

SCOPE^[11](A Scientific Compound Object Publishing and Editing System)是澳大利亚昆士兰大学开发的科学复合对象的发布与编辑系统。它使得科学家能够快速编辑内嵌了数据源流信息的科学复合数字对象,同时在该系统中也可以将除了数据源流之外的数据和文档一起打包,并把它们和 CC 协议一起发布到仓储系统中。但是 SCOPE 所描述的数据源流之外的科研知识产出只包括科研准备阶段的部分数据,而且该系统只能从输入和输出两个维度来描述知识产出,忽略了科研流程的描述以及科研过程中的知识产出。

除了具体的解决方案以外,关于科研知识产出管理的模型也不断出现。Jane Hunter 等人在总结了 Hill, Crosier, Smith, and Goodchild 的 CSCM^[12](Scientific Publish Metamodel)以及 CCLRC^[13]提出的 Scientific meta model 存在的缺陷之后,在模型中加入了工作流和其他文档资料等信息,并嵌入了语义推理功能,形成了 Scientific Publication Packages(SPPs)模型。SPPs 模型囊括了几乎所有的科研知识产出,包括数据、模型、方法、计算资源以及科学工作流等,相对于前面几种解决方案来说,SPPs 模型资源覆盖面更广,复用性强,但是 SPPs 的资源组织方式是从分类的角度对各类资源的描述,缺乏科研的流程描述信息,随着资源数量的持续增长,管理和利用的难度也会不断增加。

与我们在第一章中所提出的目标相比,这些项目尽管组织形式和思路上满足了我们的部分需求,但是和我们的最终目标还有一定差距,还需要我们对科研过程和科学工作流进行分析研究,第三章我们将从科学工作流角度入手,并结合上述项目,分析科研流程,探寻可行的知识产出组织模型。

3 基于科学工作流的科研知识产出组织管理模型的构建

科学工作流主要关注自动化的数据获取和处理,尤其是科研过程中可以通过各个 Web 服务者提供的 Web Service 来进行处理的阶段和过程,而对那些涉及了大量非结构性的科研活动来说,

是无法通过 Web 服务自动进行处理的,因此,仅仅通过科学工作流并不能完全描述科研过程中完整的科研过程和知识产出,但是科学工作流为我们提供了可以参考和借鉴的数据流模型以及其基于服务的流程描述思想。

我们期望在科学工作流模型的基础之上构建科研知识产出管理模型,不仅能够利用科学工作流对科研过程细粒度的执行过程进行捕获和描述,同时也能在粗粒度层次上对其他非结构化过程进行描述,本章将通过科学工作流的描述模型、执行模型、以及层次结构模型的分析,借鉴科学工作流的模型和设计模型,从而构建科研知识产出的组织和管理模型。

3.1 科研过程科研活动描述模型的构建

3.1.1 科研活动任务的分类和定义。科学工作流一般包括^[14]Method、Service、Workflow、Data Product。Method 指一个可以实现计算或者数据处理任务的功能或者一个步骤,Service 包含一组相关的 Method,用来执行一连串的任务。Data Product 指一个单个的数据对象或者数据对象集合。科学工作流通过有向图来表示,节点表示服务,边表示服务或者方法间的数据流。

对于我们所研究的完整的科研过程来说,其涉及的元素和科学工作流相似,我们需要解决的问题是科研流程中的非结构化的科研活动如何描述,下面将按照对科研活动进行分类和对任务节点进行描述两个步骤来解决这个问题。

上文中将科研过程分为3个阶段:科研准备阶段、科研过程阶段和科研成果阶段,在每一个阶段中都会涉及多种类型的活动。通过借鉴 SPPs 模型以及韩颖的科研过程模型^[15]并在其基础上根据上文3个阶段进行了改进,将科研活动分为调查、研究、交流、处理4种,如图2所示。

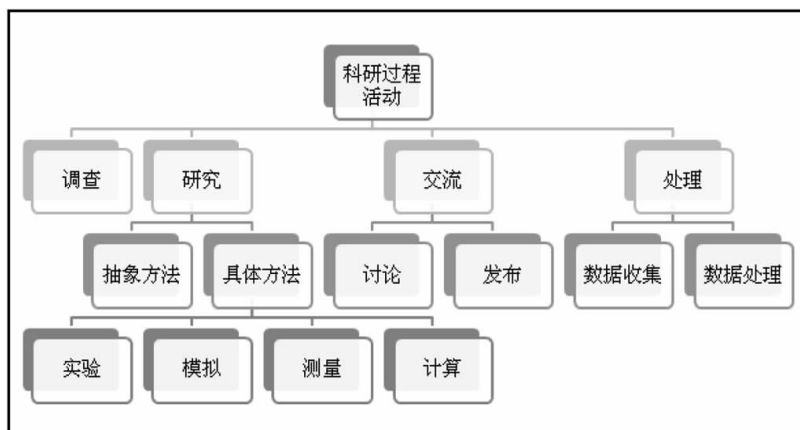


图2 科研过程中的活动

对于图3中的活动来说,由于其所处的环境、所执行的科研人员以及学科类别的差异,我们很难针对这

些活动制定统一的描述方案。但是,对于具体的项目及应用来说,例如我们将这些活动的范围缩小到某一领域,某一学科分支,甚至某一个科研团体,一个研究人员、某一项目,这些活动的标准化描述就有其存在的意义和价值。

3.1.2 科研活动过程中的知识产出描述。在科研过程描述阶段,我们可以清晰地确定并描述每一个活动所输入和输出对象的参数、类型,这些描述和科学工作流程中对端口的描述相似,这些对象不是单独孤立的单元,而是和每一个活动相连接。如果我们将前面的每个活动用线来表示,而科研知识产出用点来表示,科研活动中知识产出的转换路径便一目了然。

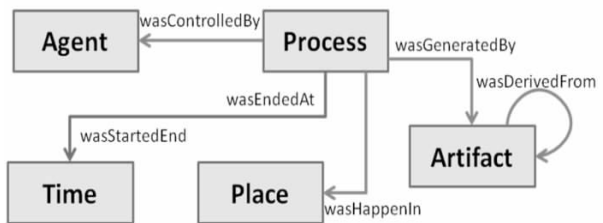


图 3 科研过程元模型示意图

相对于科学工作流程描述来说,非结构化任务描述属于粗粒度,那么我们可以将科学工作流程整体作为一个非结构化任务,加入到非结构化任务节点中,即形成了完整的科研项目描述全过程,同时也可以单独对科学工作流程节点进行查看,并查看工作流程中详细的处理单元。

3.2 科研过程执行模型的构建 在对科研过程中的任务和任务之间的连接定义好之后,科研过程就可以执行,每个科研流程定义都会有一个或多个执行过程,关于这些执行过程中的详细信息都需要执行模型进行描述。执行模型建立在描述模型基础之上,描述模型提供宏观架构,而执行模型从微观进行描述,两者相辅相成,互为补充。

本文选取了当前有代表性的 3 个科学工作流程执行模型 OPM^[16]、Provenir^[17]、CRMdig^[18] 进行分析。OPM 基于核心类 agent、process 和 artifact 构建; Provenir 在 OPM 的基础上进一步细化,在 Artifact 上进行了更为详细的描述,加入了物理特性和数据分类模型;而 CRMdig 在 OPM 和 Provenir 的基础上增加了科学测量过程中的物理环境,同时也可以映射到 OPM。

在上述 3 种模型中,OPM 是最为通用的模型,主流的科学工作流程系统 Kepler、Taverna 都支持系统数据源流模型和 OPM 的映射,因此本文在 OPM 的基础上构建科研流程的执行模型。

3.2.1 科研过程执行过程元模型的构建。OPM 模型很好地描述了 Agent、Process 和 Artifact 三者之间的关系,本文在借鉴 OPM 模型的基础上建立一个

Process 的元模型,描述了 Agent、Artifact 与 Process 之间的关系。其中的 Process 为描述模型中所构建的任务定义,在 Agent 和 Process 的共同作用下,输入的 artifact 经过处理输出了新的 artifact。

元模型是科研过程中进行过程描述的最小单元,多个元模型的组合可以实现一个具体的 service,而 service 之间也可以进行相互的调用和 service 的嵌套。

3.2.2 执行模型中的科研知识产出。在描述模型中定义知识产出的类型,那么在执行模型中我们需要管理的就是知识产出(artifact),对于每一个知识产出,都有至少以下 2 个属性:

- a. Copyright: 知识产出的版权所有者信息,一般是 agent;
- b. wasGeneratedBy: 知识产出的之间生产者,一般是 process。

除了这两个必须属性之外,还可以有时间、空间以及和具体类型相关的执行过程信息,这些信息将所有的科研知识产出以合理有序的方式连接起来,形成了知识产出关系链(网),在这个关系链(网)中,每一个知识产出将不再孤立,当这个关系网足够大时,每一个知识节点都有被发现和重用的可能,对于我们的数据密集型科研过程来说,数据发现和获取或许将不再困难。

3.3 科研过程执行模型的重用和进化 我们将各个执行模型之间的相互引用关系以及自身的变更关系称之为“进化模型”,进化模型的建立不仅能够清晰地描述科研流程的变更,更重要的是在其共享和传播过程中能够很好地促进科研生态圈的良性循环,促进科学研究的共享和重用,避免科研的重复进行。

进化模型主要分为两种类型,一是科研过程本身的进化,主要是科研设计者在科研过程中对科研过程的修改完善的记录,例如在某个科研过程执行一次之后,对某些出现的问题进行修改完善之后重新执行,那么前后两个执行模型间就是进化关系,他们都是科研过程描述模型的实例;第二种是在科研过程发布之后,其他科研项目或科研过程设计者对其进行的引用、修改和重用。它们之间的相互引用和共享,不仅扩大了科研知识产出的共享和传播渠道和范围,而且也更有利于在知识产出交流的基础上进行科研团队之间的更深入的交流和讨论,有利于构建良性可持续发展的科研生态环境。

3.4 基于科学工作流程的科研知识产出层次结构模型的构建 科研过程描述和科研模型并非独立存在的,二者具有抽象和具体,简单与复杂,类和实例多重关系,因此,如何在科研知识产出组织管理模型设计中实现科研过程的描述模型和执行模型完美融合也是一

个非常重要的方面,我们首先需要分析的是目前的科学工作流系统的总体架构模型。

3.4.1 科学工作流系统总体架构模型分析。尽管目前的科学工作流都采取不同的架构模型来实现科学工作流模型,但分层的方式依然是目前科学工作流系统的主流^[14],REDUX^[19]通过详细的多层模型来存储 provenance,包括抽象的工作流表示层、完整的实例、执行的工作流层。VisTrails^[20]的 provenance 信息也被分为 3 层存储:工作流演变层(evolution layer),和 Kepler 中的 WorkflowEvolution 类似,用来捕获工作流创建过程中的各个工作流版本之间的信息;工作流层(workflow layer)用来存储每个工作流的详细信息;执行层(execution layer)存储工作流执行过程中的信息。VIEW^[21]也包含三层模型:provenance 模型层,通过领域本体来表示科学工作流运行的数据源流信息;关系模型层用来提供 provenance 元数据的存储和查询服务;模型映射层为其他两个层提供 schema 映射、数据映射和查询映射。

3.4.2 科研过程的层次模型的构建。层次模型是描述模型和执行模型融合的最佳选择,通过对比分析,Karma 层次模型以简单、清晰的结构实现了科学工作流的有效存储、利用,同时也实现了科学工作流执行过程中各个 service 的输入和输出的捕获。Karma 模型中所有的 process 都可以做为一个服务,并通过输入和输出参数对其进行调用。这种通过对抽象(类)的服务进行注册,而在具体执行过程中只需要对服务进行调用的方式非常适合于一个科研组织内部或者一个学科领域构建科研过程。

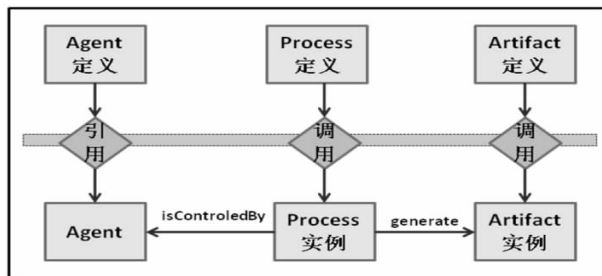


图 4 基于科学工作流的科研知识产出层次模型

图 4 是本文所构建的基于科学工作流的科研知识

产出组织管理层次结构示意图,在图的上方是登记层,个人或者组织可以在这里登记在科研过程中所使用的活动(process),artifact 等信息,并对其进行标准的定义。在层次模型的下部是科研过程的执行。在该层次模型中,用户在许可的范围内可以调用自己或者他人的服务进行科研过程的描述,并在执行过程中收集并管理每一阶段的科研知识产出,在整个科研项目的执行过程中,每个科研知识产出都是产权明晰、可追溯、可获取、可利用的。

3.5 科研知识产出组织模型的实现 由于我们所构建的模型都是建立在科学工作流的基础之上,而且我们将其作为一个完整的部分进行管理和利用,因此我们将其称之为扩展的科学工作流对象包(Extend Scientific Workflow Package,ESWP),图 5 是我们根据上述 4 个模型所构建的完整模型,包括 3 个部分,ESWP 定义(definition)、ESWP 执行(execution)、ESWP 进化(evolution)。

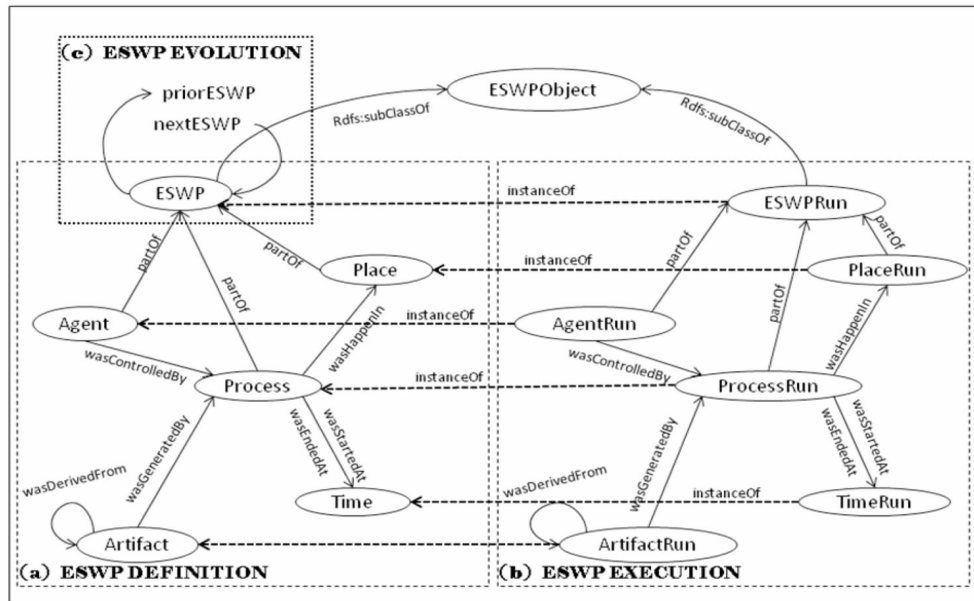


图 5 ESWP 总体模型

由于篇幅有限,图 5 以简单的方式阐述了基于科学工作流的科研知识产出组织模型,对于 Process 以及 Artifact 来说,中间还涉及 Process 之间的互相调用,前后连接,以及 Artifact 做为输入对象对 Process 的输入等。

将 ESWP 对象分为 3 大部分,很好地实现了前文提出的科研过程的层次模型,同时也便于 ESWP 文件的管理和利用。

通过上述模型以及对 3 个类的扩展,已经在很大程度上实现了本文开始时提出的构想。首先,通过科研活动定义描述了完整科研过程以及与其相关的环境,在科研活动期间,通过进行具体的科研活动实现了科研活动的运行,结果生成的每一个对象都是 ESWP

的一个实例。同时,进化模型又完整描述了 ESWP 的传播和利用的路径,不论对于单个科研项目还是某个领域的科研生态圈,这个模型都很好地满足了科研知识产出的组织管理和利用。

在当前数据密集型科研环境下,以数据为对象的研究越来越普遍,每一个数据集都可以从不同的学科、不同的视角来研究,因此对于数据集的质量、来源以及可信度都有很高的要求,而我们为科研知识产出提供了完整的全方位的信息、来源、处理路径,甚至可以通过查看使用该数据集的实验来寻找和自己的研究相关的内容。希望该模型能够在科研过程和知识产出管理中起到积极的作用,促进科研生态环境的良性健康发展。

当然,上文仅仅是一个模型,而且对于知识产出的利用方面涉及的还不够,对各个类的扩展也不是很完善,我们下一步打算基于模型建立一个知识产出组织管理和利用的原型系统,从而检验和测试模型的可行性和可用性,使得模型能够不断得到修改和完善。

参 考 文 献

- [1] Hey AJG, Tansley S, Tolle KM. The Fourth Paradigm: Data-intensive Scientific Discovery [R]. Microsoft Research Redmond, WA; 2009
- [2] 李进华. e-Science 环境下的科学工作流实现及其应用研究 (I)——概念、架构与设计 [J]. 情报科学, 2009 (9): 1394-1399
- [3] Yu J, Buyya R. A Taxonomy of Scientific Workflow Systems for Grid Computing [J]. ACM Sigmod Record. 2005, 34 (3): 44-49
- [4] 中国科学院数据应用环境. 科学工作流系统 Kepler 及其应用讲义 [EB/OL]. [2008-10-28/2011-06-21]. <http://www.csdb.cn/prohtml/0.compservice.training.material/pages/0447.html>
- [5] 张卫民, 刘灿灿, 骆志刚. 科学工作流技术研究综述 [J]. 国防科技大学学报, 2011 (3): 56-65
- [6] 邓玉坤, 王锋, 邓辉, 等. 天文轻量级科学工作流系统的实现 [J]. 天文研究与技术, 2010 (4): 338-343
- [7] 王圆. 面向 Web-MAGIS 的用户管理系统及其信息安全策略研究 [D]. 青岛: 中国海洋大学; 2009
- [8] myExperiment. myExperiment [EB/OL]. [2011/2011-05-05]. <http://www.myexperiment.org/>
- [9] ELSEVIER. Executable Paper Grand Challenge. [EB/OL]. [2011/2011-05-05]. <http://www.executablepapers.com/about-challenge.html>
- [10] USENIX. USENIX TaPP '11 Technical Program [EB/OL]. [2011/2011-06-21]. <http://www.usenix.org/events/tapp11/tech/tech.html>
- [11] Cheung K, Hunter J, Lashtabeg A, et al. SCOPE: A Scientific Compound Object Publishing and Editing System [J]. International Journal of Digital Curation. 2008, 3 (2): 1-12
- [12] Hill L L, Crosier S J, Smith TR, et al. A Content Standard for Computational Models [J]. D-Lib Magazine. 2001, 7 (6): 1082-9873
- [13] Sufi S, Matthews B, van Dam KK. An Interdisciplinary Model for the Representation of Scientific Studies and Associated Data Holdings [J]. In; 2003; 2003: 2-4
- [14] Cao B, Plale B, Subramanian G, et al. Provenance Information Model of Karma Version 3 [J]. In; 2009; IEEE; 2009: 348-351
- [15] 韩颖. 基于本体的科研数字对象情景元数据框架 [D]. 北京: 中科院文献情报中心: 中国科学院研究生院, 2010-03
- [16] Moreau L, Freire J, Futrelle J, et al. The Open Provenance Model [J]. University of Southampton. 2007
- [17] Provenir Ontology [EB/OL]. [2009/2011-09-10]. http://wiki.knoesis.org/index.php/Provenir_Ontology
- [18] CRMdig: A Generic Digital Provenance Model for Scientific Observation [EB/OL]. [2011/2011-06-21]. http://www.usenix.org/events/tapp11/tech/final_files/Doerr.pdf
- [19] Barga RS, Digiampietri LA. Automatic Capture and Efficient Storage of e-Science Experiment Provenance [J]. Concurrency and Computation: Practice and Experience. 2008, 20 (5): 419-429
- [20] Scheidegger C, Koop D, Santos E, et al. Tackling the Provenance Challenge One Layer at a Time [J]. Concurrency and Computation: Practice and Experience. 2008, 20 (5): 473-483
- [21] Lin C, Lu S, Lai Z, et al. Service-oriented Architecture for VIEW: A Visual Scientific Workflow Management System. In; 2008; IEEE; 2008: 335-342
- [22] myExperiment. myExperiment Wiki [EB/OL]. [2011/2011-07-03]. http://wiki.myexperiment.org/index.php/Main_Page
- [23] Cavalcanti MC, Mattoso M, Campos ML, et al. Sharing Scientific Models in Environmental Applications [J]. ACM, 2002: 453-457 (责编: 刘武英)
- [9] 岳洪江, 梁立明. 90 年代我国三大地带科技指标差距变动分析 [J]. 科学学研究, 2000, 18 (4): 56-62
- [10] 甘国勇, 陆艳. 我国区域科技创新差距及对策研究——基于东中西部区域比较的视角 [J]. 科技和产业, 2011, 11 (8): 95-99
- [11] 郑加强. 建设可持续发展的区域科技创新体系 [J]. 科技与经济, 2002, 15 (1): 1-5
- [12] 师萍, 李垣. 科技资源体系内涵与制度因素 [J]. 中国软科学, 2000 (11): 55-57
- [13] Huang C L, Yoon K. Multiattribute Decision Making: Methods and Applications [M]. NY: Springer-Verlag, 1981: 12-34.
- [14] Hoerl A E, Kennard R W. Ridge Regression: Biased Estimation for Nonorthogonal Problems [J]. Technometrics, 1970 (12): 55-67
- [15] Damodar N. Gujarati. 经济计量学精要 [M]. 北京: 机械工业出版社, 2004: 204-205 (责编: 王平军)

(上接第 201 页)