

科研信息化基础设施的大数据挑战

Demchenko Yuri Zhao Zhiming Grosso Paola Wibisono Adianto de Laat Cees
(阿姆斯特丹大学系统与网络工程组, 荷兰阿姆斯特丹 1098XH)

摘要: 首先探讨大数据科学 (Big Data Science) 给现代和未来科研数据基础设施 (SDI) 带来的挑战, 并参考不同科研团体的工作对数据管理、访问控制和安全要求进行定义, 然后介绍科研数据生命周期管理 (SDLM) 模型, 分析所有的主要阶段, 反映现代 e-Science 的数据管理特性, 提出 SDI 通用架构模型, 为利用现代技术和最佳实践构建以互用性数据或项目为中心的 SDI 奠定了基础, 最后对如何利用基于云的现代基础设施服务配置模型顺利实现文中提出的 SDLM 和 SDI 模型进行阐述, 解决 SDI 资源的联合访问控制问题, 为科研团体提供了一种灵活的访问控制和身份管理模型。

关键词: 大数据科学; 科研数据基础设施; 科研数据生命周期管理; 云基础设施服务; 科研信息化

中图分类号: G311

文献标识码: A

DOI: 10.3772/j.issn.1674-1544.2013.01.006

Big Data Challenges for e-Science Infrastructure

Demchenko Yuri, Zhao Zhiming, Grosso Paola, Wibisono Adianto, de Laat Cees

(System and Network Engineering Group, University of Amsterdam, Amsterdam 1098XH, Netherlands)

Abstract: This paper discusses the challenges that are imposed by the Big Data Science on the modern and future Scientific Data Infrastructure (SDI). The paper refers to different scientific communities to define requirements on data management, access control and security. The paper introduces the Scientific Data Lifecycle Management (SDLM) model that includes all the major stages and reflects specifics in data management in modern e-Science. The paper proposes the SDI generic architecture model that provides a basis for building interoperable data or project centric SDI using modern technologies and best practices. The paper explains how the proposed models SDLM and SDI can be naturally implemented using modern cloud based infrastructure services provisioning model. The paper also addresses issues with the federated access control to the SDI resources that provides a flexible access control and identity management model for scientific and research communities.

Keywords: big data science, Scientific Data Infrastructure (SDI), Scientific Data Lifecycle Management (SDLM), cloud infrastructure services, e-Science

1 引言

现代科研信息化基础设施的广泛应用解决了以往不可能出现的如基因组、气候和全球变暖等大规模数据的问题。科研信息化产生的海量数据, 需

要新型科研数据基础设施 (SDI) 进行存储、分发、处理、保存和管理^[1-2]。在科研信息化过程中, 科研数据具有复杂的内部关系。这些科研数据需要利用相应的物理或逻辑的基础设施进行存储、访问和管理。新兴的 SDI 应允许不同的研究团队围绕相同

第一作者简介: Demchenko Yuri (1958-), 男, 阿姆斯特丹大学系统与网络工程组研究员, 研究方向: 云和云际的架构, 大数据的科学基础设施, 通用安全体系结构和分布式访问控制的基础设施。

基金项目: 国家软科学研究计划资助项目“以老年人为中心的电子健康需求及其使能技术评价体系研究”(2011GXQ4K029); 国家自然科学基金重点项目“信息系统的采纳、扩散与商业价值机理研究”(70831003); 欧盟 GEYSERS—基础设施服务的通用体系架构项目 (FTP-ICT-248657); 欧盟 GEANT—高带宽的泛欧网络项目 (238875)。

收稿日期: 2012年10月15日。

的数据集开展工作, 构建自己的(虚拟)研究和协作环境, 安全地存储中间结果, 共享研究成果。新产生的数据溯源、数据安全及数据访问控制机制与工具应允许研究人员把他们的科研成果同原始数据(集)和中间数据联系在一起, 以便今后可以通过改进研究技术和工具重新使用这些数据或改变这些数据。

本文分析了大数据技术给现代科研信息化基础设施带来的挑战, 提出了应对挑战的一般方法和架构解决方案, 特别是利用云基础设施技术, 构建了新的科研数据生命周期管理(SDLM)模型和通用SDI架构模型, 为SDI异构组件的互用和集成提供了基础。本文结构是: 第二部分介绍主要研究团体和对未来SDI的总体要求; 第三部分探讨大数据科研中数据管理面临的挑战, 包括对SDLM的讨论; 第四部分介绍为应对大数据的挑战和要求而提出的e-SDI架构模型; 第五部分探讨利用云技术实现SDI; 第六部分论述构建现代和未来SDI的联合访问存取管理基础设施的具体要求, 并提出相关的建议。

2 科研信息化基础设施面临的挑战及其具体要求

目前, 大数据的不断涌现, 科研信息化基础设施一方面要处理不同研究设备产生和/或传感器收集呈指数增长的数据; 另一方面要利用适当的管理模式将信息化基础设施打造成持久的研究平台, 确保研究的连续性, 实现跨学科合作, 并能够提供持续的服务。面对这样的挑战, 需要重新考虑现有的基础设施组件、解决方案和处理过程^[2-3]。随着信息通信技术和大数据技术的迅猛发展, 现代科研信息化管理模式发生了重大变化, 主要表现在以下几点: (1) 科研信息化全部流程自动化, 包括数据收集、存储、分类和索引以及一般数据的管理和溯源等; (2) 利用多维、多层面的测量、监控方法, 将处理过程和产品转化为数字形式, 并将现有人工制品及其他内容数字化; (3) 二次研究可重新开发利用原始数据和已发布的研究数据, 并改变数据的用途; (4) 合作研究团队通过网络可利用和获取全球数据, 包括广泛的公开获取的科学数据; (5) 基础设施组件和管理工具能够为特定的研究项目和任务快速组合、调整并按需配置基础设施和服务; (6) 支持先进的安全和访问控制技术, 确保复杂的研究基础设

施和科研设备安全运行, 为合作团队和研究人员营造可靠的安全环境。

在整个数据生命周期中, 未来的SDI不仅要求大规模和长期/无限期地进行数据存储/保存、数据聚合和数据溯源, 而且要求必须确保数据的安全性(完整性、机密性、可用性和可计量性)和保护数据的所有权。处理大数据需要强大的计算能力。随着处理大数据需求的增加, 应该制定并实施数据/数据集处理策略, 能够在可靠的系统上和/或按照其他的要求处理数据。利用SDI处理研究人员的数据, 可以确保数据的安全, 不被非授权者获取。SDI具有远程分布的特性, 用户可以在不同的国家获取数据, 这就带来了数据安全性的问题。这些问题应该由SDI的重要组件来控制访问并通过计费基础设施(ACAI)加以解决^[4-5]。

近来, 笔者查阅了欧洲研究领域(ERA)^[3]中关于欧洲科研信息化基础设施已有的研究资料, 如锡耶纳项目(SIENA Project)^[6]、欧洲跨政府联合研究机构论坛联邦身份管理工作组(EIRO forum Federated Identity Management Workshop)^[5]、欧洲网络基础设施(EGI)战略报告^[7]、英国未来互联网战略组织报告(UK Future Internet Strategy Group Report)^[8]等。这些研究资料分析了各科研团队的实践, 面临的挑战, 并对科研信息化基础设施提出了新的要求。

(1) 高能物理学团体: 高能物理学团体拥有大批研究人员、独一无二的昂贵的仪器设备以及大量生成且必须连续处理的数据。该团队拥有已经投入运营的全球大型强子对撞机网格(WLCG)^[9]基础设施, 在整个科研数据生命周期中, 管理和访问数据, 保护数据的完整性。开发WLCG是推进欧洲信息化基础设施建设中的重要一步。该基础设施能够为欧洲乃至全球的多个科研团体提供服务。欧洲网络基础设施合作机构^[7]为高能物理学团体和其他团体管理欧洲乃至全球的基础设施。

(2) 材料科学、分析和低能物理学团体: 材料科学、分析和低能物理学(质子、中子、激光设备)团体的特点是既有短期项目研究又有实际操作, 用户群体不断变化。因此通过对基础设施的动态管理, 实现广泛地获取数据和分布式处理数据。

(3) 环境和地球科学团体: 环境和地球科学团体的研究瞄准了区域/国家乃至全球性的问题。他们收集来自陆地、海洋、大气和空间的大量数据,

需要不断扩大存储容量、提高计算能力。这就要求SDI能够可靠精细地对大数据集进行访问控制、对区域问题能够进行强制执行并对数据采取过滤策略(数据可能涉及与国家安全相关的信息),同时需要跟踪数据的使用并保持数据的完整性。

(4)生物和医学科学团体:生物和医学科学(也称为生命科学)团体通常关注健康、药物开发、新物种鉴定和新仪器开发等。在研究过程中产生了大量的数据,并且对分布式处理、数据共享和协作的计算能力、存储容量以及网络性能提出了更高的要求。生物医学数据(包括医疗保健、临床病例数据)是敏感型且隐私性极高的数据,必须按照欧洲个人数据处理政策进行处理^[10]。

(5)社会科学和人文学科团体:社会科学和人文学科团体及其研究项目具有多边性的特点,通常是研究人员之间的全球化合作。这就需要协作群组/团体的参与,并通过协作的基础设施实现共享数据、发现/研究成果、合作评估结果,应该允许广泛而公开地获取对所有收集到的人工制品数字化后产生的海量数据。

为此,新兴的大数据科学对SDI的一般性基础设施提出了新的要求:(1)能够支持长期运行的实验和快速生成的大数据量;(2)可以进行多层次数据分布与复制;(3)能够支持虚拟科研团体;(4)具有可靠安全的数据存储与处理环境;(5)能够保证数据的完整性、机密性和可计量性;(6)执行并完善捆绑数据的隐私保护政策。

3 大数据技术的数据管理

计算机辅助研究方法的出现改变了开展科学研究和使用科研数据的方式方法。目前,科研数据主要有以下几个类型^[4]:(1)从观察和实验中收集的原始数据;(2)经过数据过滤和处理的结构化数据和数据集;(3)已发布的科学假设、研究成果或学术观点;(4)在出版物上广泛研究整理、集成和公开发布的数据。这些数据一旦发布,科研人员就会对感兴趣的数据进行确认和再现,并产生新的研究成果。因此,SDI提供者应该深入考虑以下内容:(1)科研数据溯源:从捕捉原始数据到出版数据,其中转化过程中的信息已成为科研数据管理的重要方面^[11];(2)提供已发布数据的语义:对于已发布的数据,还要保证这些数据在科研团体的研究实践中可以重复使用。理解已发布数据的语义是实现数

据可再用性的重要环节,过去这都是靠人工方式完成的。然而,当对大数据科研中产生的空前规模的数据进行预测且有效地再利用发布时,有必要附带清晰的数据语义;(3)链接分布式数据:在通常情况下,大数据既分布在数据收集端,也分布在处理/访问端。因此,需要对数据进行收集、分配和/或复制,将分布式数据链接起来。欧盟委员会倡导开放获取公开基金支持项目中的科学数据,对于出版物以及数据的链接^[12-13],建议采用PID(永久性数据ID)和ORCID(开放研究者与贡献者ID)^[14]等方法。

在分析不同科研团体实践的基础上,笔者提出了科研数据生命周期(SDLM)模型(图1)。该模型体现了科研信息化的数据管理和处理新方法,与参考文献[15]提出的数据生命周期研究结果相符。

通常地,科研数据生命周期主要包含研究项目或实验规划、数据收集、数据处理、发布研究成果、讨论和反馈、归档(或放弃)等阶段。

在各个阶段中,SDLM都要求存储和保护数据,并允许重新使用数据或改变数据的用途,并对已处理的数据和已发布的结果进行二次研究。然而,在SDI中能够实现所有数据的识别、交叉引用和链接。在整个数据生命周期中,能够支持数据的完整性、访问控制以及可计量性。数据管理是讨论SDLM重要的组成部分,其实现的方式必须是安全的和值得信赖的。

在科研数据的生命周期内,SDLM须要确保数据的安全性并能够进行访问控制,主要包括数据(实验数据)的采集、原始数据的过滤、专业人员的处理、研究数据的存储和二次数据挖掘、数据及其研究信息的存档等。

4 SDI架构模型

本文提出的科研信息化SDI(e-SDI)框架如图2所示。主要包含以下各层。

网络基础设施层(B₁层):基于通用互联网基础设施和专用网络基础设施;

数据中心和计算资源/设施(B₂层);

基础设施虚拟层(B₃层):以支持专门的科研平台进行部署和运行的云/网格基础设施服务和中间件为代表;

科研平台与基础设施(B₄层):专门为不同研究领域提供数据共享的科研平台与基础设施;

联合决策层(B₅层):包括联合基础设施组

件，能够支持决策和协作用户；

科研应用和用户门户网站/客户端(B₆层)。

此外，还可以定义3个交叉层面：运行支持与

管理系统、安全层面及元数据与生命周期管理。

通过专有层支持SDI动态特性及其分布式多层

面团队：B₃是基础设施的有效层，典型地采用了现

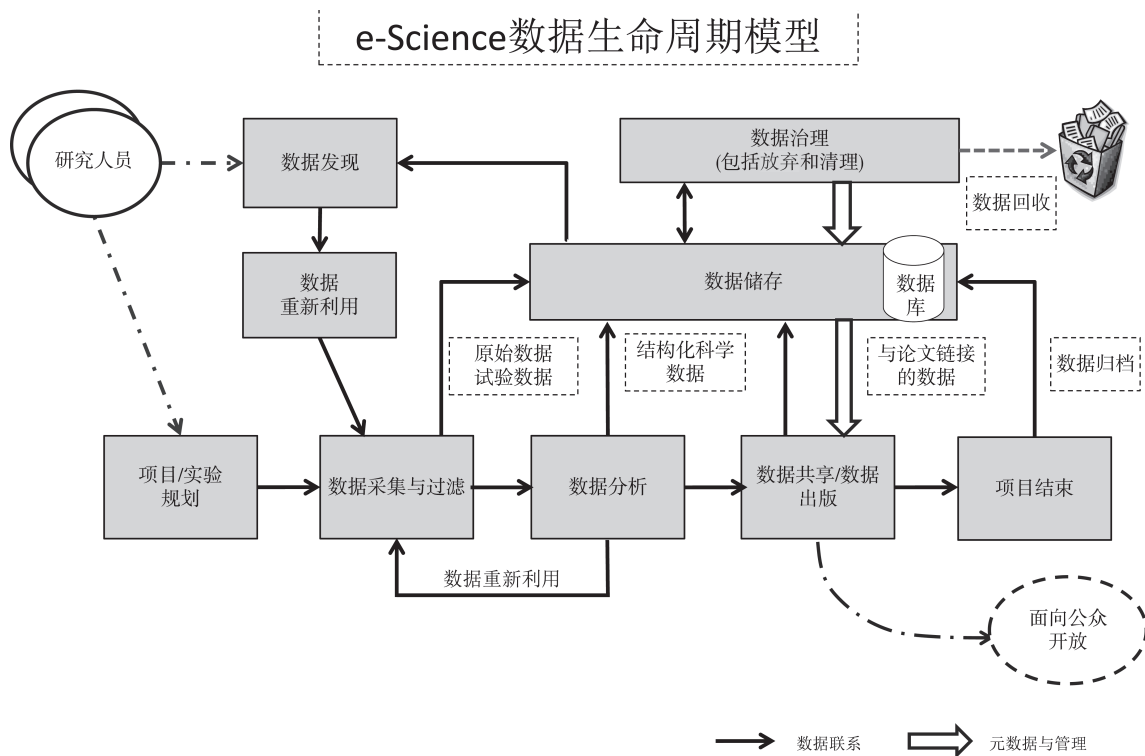


图1 科研信息化的科研数据生命周期管理

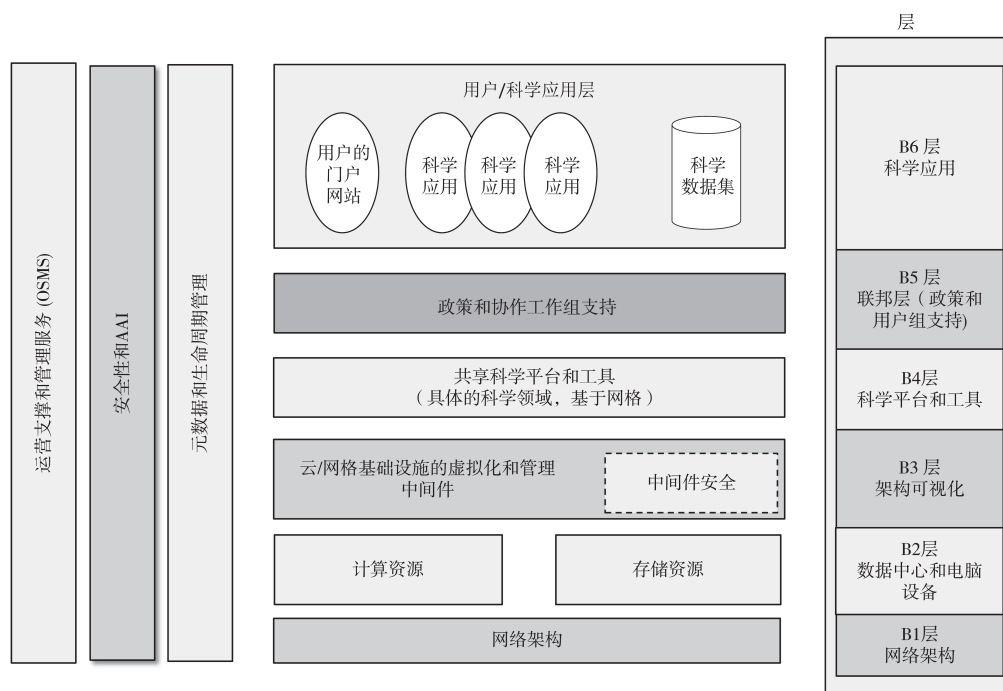


图2 本文提出的SDI架构模型

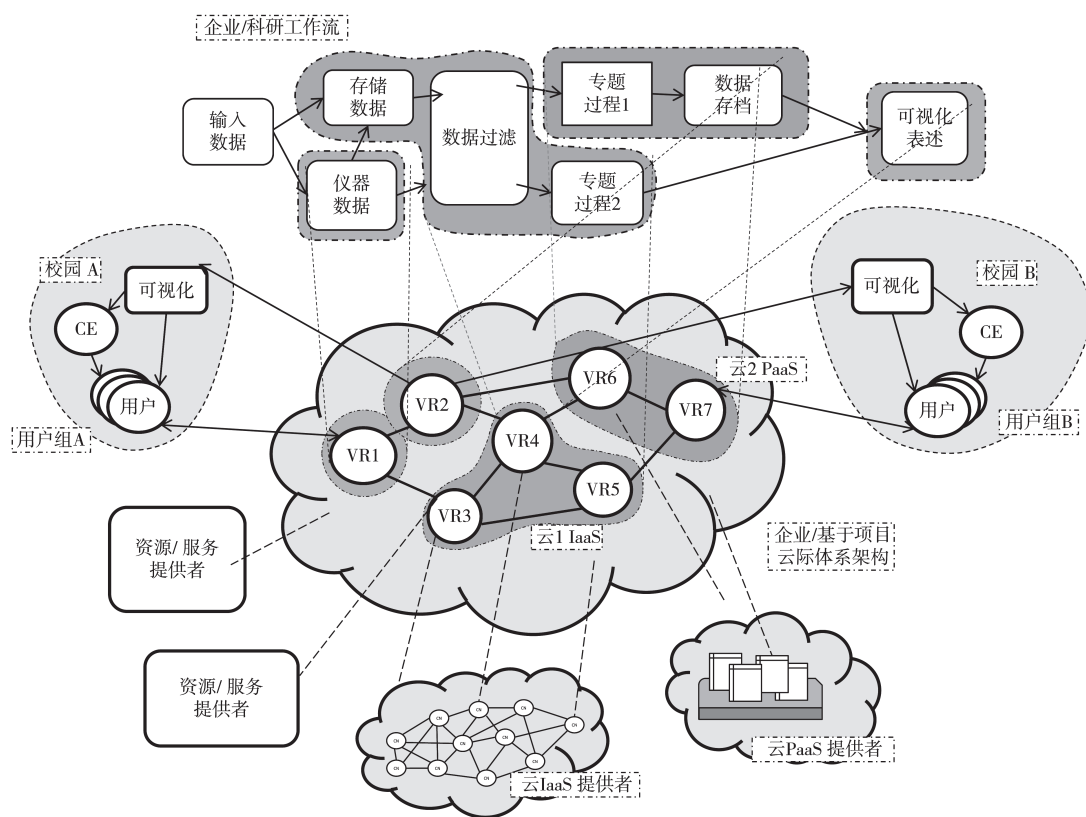


图3 从科研工作流到基于云的基础设施

代云技术；B₅是联合决策层，包含相关的联合基础设施的管理和存取技术^[4, 16-17]。引入联合决策层反映出构建和管理复杂SDI（也包括企业基础设施）的现行方法，并允许独立管理的基础设施共享资源以及支持组织间的合作。

5 云技术下的SDI基础设施服务

图3所示是一个典型的根据需要构建的科研信息化或企业协作的基础设施，包括企业所有权、基于云计算存储的资源和设备、监控系统、虚拟系统以及以客户端为代表的典型地存在于现实或虚拟环境中的用户。

企业协作或科研信息化的基础设施支持企业或科研工作流以及有关过程监控和数据处理的操作流程。云技术可以简单地构建这样的基础设施并按需要提供配置。图3以实例说明了怎样把一个企业或科研工作流映射到基于云的服务，并作为一个即时互联云基础设施进行部署和运行。主要包括云基础设施部分IaaS（VR3-VR5）和PaaS（VR6, VR7）、独立的虚拟化资源和服务（VR1, VR2）、两个相互作用的区域A和B以及相互连接的网络基础设施。

在许多情况下，需要采用专用的网络连接以保证其性能的实现。

有效地运行这样的基础设施既需要加强整体的基础设施管理，又需要提供个性化的服务以及基础设施领域间的相互交流与沟通。这已明显超出了现有云服务提供商所提供服务的范围，但却是构建未来科研信息化的SDI的宗旨所在。这是笔者另一项关于互联云框架构建的研究课题^[18-19]。在此不再赘述。

6 SDI的访问控制和计费基础设施

6.1 SDI和访问控制的一般要求

为了控制访问，确保数据的安全，未来科研数据信息化基础设施（SDI）应有相应的访问控制和计费基础设施（ACAI），以正常运行基础设施，保护资产和信息，并在分布式多组织的环境下，实现用户身份识别/鉴定和相关策略的执行。而开放获取^[12]可能需要部分改变现有的科研信息中心和图书馆的业务工作，因此，未来的ACAI应支持这样的转变，并细化访问控制，界定和控制灵活的策略实施。

鉴于未来SDI支持全部的数据生命周期，且大规模和长期/无限期地进行数据存储/保存、数据聚

合和数据溯源的要求, 未来的ACAI也应支持数据生命周期的全部阶段, 包括为数据附加策略, 以确保在线和离线处理过程中持续地执行相关的数据策略。

未来的ACAI应该具有以下特点: (1) 允许研究人员(并让他们相信)在具有数据和信息安全的大型数据中心的共享设备上执行相关数据处理。(2) 激励/确保研究人员共享和开放他们的研究环境, 为其他的研究人员提供专门/定制的预先配置的基础设施, 并允许其他研究人员使用现有的或自己的数据集。(3) 随着技术的发展, 数据保存技术本身应该确保数据的可读性和可获取性。因此, 在提供(长期)数据存档时, 应采取数据保护措施来保护数据的所有权, 确保其他关联数据(包括其他数据集和新产生的科研数据)的安全。

6.2 联合访问控制和身份管理

大数据的科研团队应该探究现有身份验证与授权基础设施(AAI)(组织、团体和国家)的联合以及常见的联合基础设施服务和策略, 包括保护相关的属性和维护信任管理权。这种联合可以在成员团体和研究组织间分担联合基础设施的管理责任。

对于用户来讲, 通过信息化基础设施进行访问, 获取相关的资源, 是最吸引人的。这种方式将消除通过信息化基础设施获取资源的障碍: 用户能够使用他们的机构账号向SDI提出注册申请(一般通过网站登录), 获取相关资源和流程。网站负责鉴别用户身份, 并将这个账户映射到被认可的合作/联合的SDI网站的本地身份证明/证书上。用户使用这个证书可以访问联合基础设施的其他网站或者获取相关的资源。

在本文提出的e-SDI架构模型中, 联合访问控制简化了对虚拟用户组的管理, 支持相应的联合基础设施。联合策略层B₅提供了许多支持其运行的功能、协议和界面: 服务注册与发现、可靠的服务代理商、身份标识提供商(IdP)、可靠的管理器/路由器、属性/名称空间解析器、互联云网关和/或属性/名称空间解码器。联合ACAI能够利用现有的平台进行联合网络的访问和联合身份的管理。这样的访问和管理将广泛地应用于多个领域和拥有多供应商的基础设施集成中(如eduGAIN^[17]、Eduroam^[20]、Shibboleth^[21]、CILogon^[22]等)。

7 未来研发

未来的研发包括进一步定义SDLM、e-SDI和

ACAI组件, 重点开发e-SDI基础设施组件。特别需要注意的是, 根据需要可以利用基于云的IaaS和PaaS平台提供SDI服务。在提供这样的SDI服务时需要进一步明确整个SDI服务周期, 特别是为支持即时科研工作流而提供的SDI服务。本文的研究也将支持相应的云和互联云框架的研究与开发, 同时支持大数据科研信息化流程和基础设施的运行。

鸣谢

本文的研究得到了“欧洲科研数据/信息资源认证、授权和计费(AAA)平台研究”特别项目的推动和部分支持。该项目由欧盟委员会委托TERENA、LIBER、阿姆斯特丹大学、德布勒森大学联合实施。本文作者对各参与成员间的现有研究基础设施的不同方面和进一步开发AAA技术进行了有价值的广泛讨论。本文提出的基于云的SDI框架是欧盟资助的FP7项目“广义动态基础设施服务架构”(GEYSERS, FP7-ICT-248657)和GEANT(Grant Agreement No. 238875)的成果。

参考文献

- [1] 全球研究数据基础设施: 全球研究数据基础设施十年展望[EB/OL].[2012-03-06].<http://www.grdi2020.eu/Repository/FileScaricati/6bdc07fb-b21d-4b90-81d4-d909fdb96b87.pdf>.
- [2] 科研数据高级专家组. 乘风破浪: 欧洲怎样从科研数据大潮中获益[EB/OL].[2010-10-06]. Available at <http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/hlg-sdi-report.pdf>.
- [3] ESFRI 2010年最新发展路线图[EB/OL].[2012-09-26].http://ec.europa.eu/research/infrastructures/pdf/esfri/esfri_roadmap/roadmap_2010/procedure_roadmap_update.pdf.
- [4] 欧洲科研数据/信息资源AAA平台研究[EB/OL].[2012-09-26].<https://confluence.terena.org/download/attachments/30474266/AAA-Study-Report-0907.pdf>.
- [5] 研究合作的联合身份管理[EB/OL].[2012-09-26].<https://cdsweb.cern.ch/record/1442597>.
- [6] SIENA欧洲e-Science网格及云标准发展路线图[EB/OL].[2012-09-26].<http://www.sienainitiative.eu/Repository/Filescaricati/8ee3587a-f255-4e5c-aed4-9c2dc7b626f6.pdf>.
- [7] 寻找新地平线: EGI为2020年努力[EB/OL].[2012-09-26].http://www.egi.eu/blog/2012/03/09/seeking_new_horizons_egis_role_for_2020.html. (下转第40页)

应用环境建设与服务”项目中,通过系统自动生成了项目内500多个数据库(数据集)的引用信息。用户进行数据检索时,在每一条检索结果中同时提供了该条数据的引用信息,以方便用户进行引用标注。

数据应用规范的推广和应用,离不开数据发布和引用系统的支持。首先必须保证数据能够被永久地标识,另外还需要有一个可靠的、可以长久定位的地址。其次,在完善科学数据发布和引用技术的基础上,需要研究和建立科学数据引用机制。通过出台鼓励政策,倡导引用科学数据的习惯,增强科学数据知识产权保护的意识。可以通过推进数据中心与学术期刊的合作,在学术文章发表的同时提交相关研究数据,建立文献与数据之间的交叉引用。

参考文献

- [1] 傅晓峰.关于促进科学数据共享管理的一些思考[J].中国基础科学,2006(6):17-19.
- [2] Chavan VS, Ingwersen P. Towards a Data Publishing Framework for Primary Biodiversity Data: Challenges and Potentials for the Biodiversity Informatics Community[J/OL]. [2012-08-20]. <http://www.biomedcentral.com/1471-2105-10/s14/s2>.
- [3] 彭洁,涂勇.科学数据引用的探讨[J].数字图书馆论坛,2008(10):14-18.
- [4] The German National Library of Science and Technology (TIB). Publication and Citation of Scientific Primary Data [EB/OL]. [2012-08-20] www.std-doi.de.
- [5] Alfred Wegener Institute for Polar and Marine Research (AWI). PANGAEA [DB/OL]. [2012-08-20] <http://www.pangaea.de/>.
- [6] NSCU LIBRARIES. How to Cite a Data Set[EB/OL]. [2012-08-20]. <http://www.lib.ncsu.edu/data/citingdatasets.html>.
- [7] King Gary. An Introduction to the Dataverse Network as an Infrastructure for Data Sharing[J]. Sociological Methods and Research, 2007(36):173-199.
- [8] Micah Altman, Gary King. A Proposed Standard for the Scholarly Citation of Quantitative Data. D-Lib Magazine [J/OL]. [2012-08-20] <http://www.dlib.org/dlib/march07/altman/03altman.html>.
- [9] 中国科学院寒旱所遥感室.世界数据中心冰川(雪冰)冻土学科中心[DB/OL]. [—]. <http://wdcdgg.westgis.ac.cn/chinese/index.htm>.
- [10] 中国科学院寒旱所环境工程研究所.中国西部环境与生态科学数据中心[DB/OL]. [2012-08-20]. <http://westdc.westgis.ac.cn/data/fecdec71-77d1-43c2-b472-8b1c729874cb>.
- [11] 中国科学院计算机网络信息中心.国际科学数据服务平台[DB/OL]. [2012-08-20]. <http://datamirror.csdb.cn/dem/files/gc.jsp>.
- (上接第35页)
- [8] 未来互联网报告[EB/OL]. [2011-05-06]. https://connect.innovateuk.org/c/document_library/get_file?folderId=861750&name=DLFE-33761.pdf.
- [9] 全球大型强子对撞机网络(WLCG)[EB/OL]. [2012-09-26]. <http://wlcg.web.cern.ch/>.
- [10] 欧洲数据保护指令[EB/OL]. [2012-09-26]. http://ec.europa.eu/justice/data-protection/index_en.htm.
- [11] Koopa, David, et al. 基于数据溯源的基础设施支持可执行文件的生命周期[EB/OL]. [2012-09-26]. <http://vgc.poly.edu/~juliana/pub/vistrails-executable-paper.pdf>.
- [12] 联合国教科文组织欧洲委员会.开放访问:机遇与挑战[EB/OL]. [2012-09-26]. http://ec.europa.eu/research/science-society/document_library/pdf_06/open-access-handbook_en.pdf.
- [13] OpenAIR - 开放访问欧洲研究基础设施[EB/OL]. [2012-09-26]. <http://www.openaire.eu/>.
- [14] 开放性研究者与贡献者ID[EB/OL]. [2012-09-26]. <http://about.orcid.org/>.
- [15] 数据生命周期模型与概念[EB/OL]. [2012-09-26]. <http://wgiss.ceos.org/dsig/whitepapers/Data%20Life-cycle%20Models%20and%20Concepts%20v8.docx>.
- [16] EGI 联合云任务组[EB/OL]. [2012-09-26]. <http://www.egi.eu/infrastructure/cloud/cloudtaskforce.html>.
- [17] eduGAIN - 网络服务和应用的联合访问[EB/OL]. [2012-09-26]. <http://www.edugain.org>.
- [18] Demchenko Y, Ngo C, Makkes M, et al. 定义互联云架构的互用性和集成性[C]. 2012年第三届国际云计算、网格和虚拟化云计算大会,法国尼斯,2012年7月22-27日.
- [19] 云参考框架[EB/OL]. [2012-06-27]. <http://www.ietf.org/id/draft-khasnabish-cloud-reference-frame-work-03.txt>.
- [20] eduroam[EB/OL]. [2012-06-27]. <http://www.eduroam.org>.
- [21] Shibboleth - 开源联合身份管理系统[EB/OL]. [2012-06-27]. <http://shibboleth.net/>.
- [22] CILogon Service[EB/OL]. [2012-06-27]. <http://www.cilogon.org/>.