

基于服务的开放数字仓储架构

——DNET解析

马建霞¹ Paolo Manghi² Wolfram Horstmann³ Friedrich Summann³

¹ (中国科学院国家科学图书馆兰州分馆 兰州 730000)

² (Istituto di Scienza e Tecnologie dell'Informazione, Consiglio Nazionale Delle Ricerche, Italy)

³ (Bielefeld University Library, Germany)

【摘要】介绍 DNET的基本功能和特点,明确 DNET解决的核心问题:对异构仓储数据的收割聚合、对分布异构的服务调度。通过对 DNET信息空间和对象模型、DNET基于服务的架构,尤其对 DNET驱动层的剖析,分析 DNET如何解决这两个核心问题。并对 DNET的适用性、数字仓储联盟建设中的标准化、数字仓储联盟可提供的服务等进行探讨。

【关键词】DRIVER DNET 数字仓储架构 基于服务的架构

【分类号】G250

Analysis of the Service - Oriented Digital Repository Architecture

——DNET

Ma Jianxia¹ Paolo Manghi² Wolfram Horstmann³ Friedrich Summann³

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

² (Istituto di Scienza e Tecnologie dell'Informazione, Consiglio Nazionale Delle Ricerche, Italy)

³ (Bielefeld University Library, Germany)

【Abstract】Based on the introduction of DNET's characteristics and function, the article analyzes in - depth the information space, digital object model of DRIVER, and the service - oriented architecture of DNET. And it finds out how to deal with the harvest information from heterogeneous repository and how to orchestrate different services for DNET. It also discusses the suitability of DNET, the role of standardization in digital repository federation and the services of digital repository infrastructure.

【Keywords】DRIVER DNET Digital repository infrastructure Service - oriented architecture

DRIVER^[1] (Digital Repository Infrastructure Vision for European Research)是欧盟 FP6、FP7连续资助的数字仓储架构的项目。其愿景是使欧洲乃至全世界的研究机构的研究出版物通过数字仓储开放获取。DNET (DRIVER Network Evolution Toolkit)为 DRIVER开放获取的数字仓储架构提供了技术平台。

欧盟认为,e - Infrastructure电子基础设施是指新的研究环境,所有研究人员无论在家或者在研究所都能共享地访问独特的分布式的科学资源(包括数据、设备、计算和通信),它将成为未来支持研究的关键因素^[2]。作为 e - Infrastructure重要组成部分的数字仓储架构,强调对人、信息内容、技术的可获得性、联通性和可访问性^[3]。从技术的角度看,它必须满足以下三个条件。

收稿日期:2009 - 12 - 11

收修改稿日期:2010 - 01 - 15

(1)能够对分布异构的知识资源实现无缝访问;

(2)提供支持管理共享的知识资源,访问和重用它们的最小功能集;

(3)支持开放性和可扩展性,该架构的新功能和资源的扩展必须很容易定制。

DRIVER 的数字仓储架构就是在这种思想指导下,以基于服务架构的 DNET 来实现,以开源软件方式发布。对于 DNET 基本架构、功能的研究评价将有利于我国数字仓储联盟在新的 e - Science 环境下的设计和选型,并借鉴其发展经验,思考我国数字仓储联盟的关键问题。

1 DNET 简介

DNET 是一个基于服务架构 (Service - Oriented Architecture, SOA) 的数字仓储注册和集中服务平台,适用于大规模数字仓储联盟架构的构建和服务^[4]。基本用户是数据提供者 (开放仓储管理者)、服务提供者 (仓储联盟架构管理者) 和最终用户。其基本功能是:

(1)允许服务提供者收割遵循 OAI - PMH 标准的数据提供者仓储中的数据,并将数据清洗聚合到一个共享的信息空间;

(2)构建和定制数字图书馆应用,提供集中的检索、浏览、推荐、定制服务等功能。

DNET1.0 的大部分服务都是基于 Perl 的。DNET1.2 完全基于 Java,将服务分解成收割服务 (Harvest Service) 和转换服务 (Transformation Service) 两部分。在转换服务中,可定义元数据映射规则,将 METS、MODS、QDC 等格式映射成 OAI - DC 的格式。这一功能将使 DNET 对于数据仓储的元数据格式的支持更加广泛。DNET2.0 目前仅是实验系统。它将支持增强出版物 (Enhanced Publication, EP) 数据模型,包括内容服务、EP 用户界面、通过 OAI - ORE 揭示 EP 中的复杂关系、相似性检索、引证分析、全文索引、全文中的书目元数据抽取等新功能。

2 DNET 的特点

2.1 分布式

DNET 的开发和运行都体现了分布式的特点。

从其开发团队的组织看,在项目管理单位 Universität Göttingen 组织下,负责架构的整体设计和服

务调度管理的意大利国家研究中心 (Consiglio Nazionale Delle Ricerche),负责数据收割和聚合的德国 Bielefeld 大学 (Universität Bielefeld)、负责架构用户界面设计的雅典国立与卡珀得斯兰大学 (National and Kapodistrian University of Athens),负责服务的认证和授权的波兰华沙大学 (Uniwersytet Warszawski) 协同工作。

从其运行方式看,DRIVER 的运行依托于 9 个主节点、25 个功能服务、36 个服务实体,而这些服务分布于不同地方的服务器。DRIVER 涉及欧盟 31 个国家的 250 个仓储^[5],并且正在扩大其国际影响,与日本的数字仓储联盟和中国科学院的数字仓储网络合作。

2.2 异构信息系统的互操作

DNET 收割的每个仓储有不同的内容政策^[6];其仓储平台多样,包括 DSpace, Eprints, OPUS, Fedora, Own 等;其内容多样,包括学术文章全文及其元数据,支持数据、图像,用增强出版物模型支持复杂对象,揭示复杂数字对象的结构和关系;内容标准多样,包括 MODS、METS、DDL、MARC、QDC、Unqualified DC 等元数据;即便对于同一个字段,不同的仓储也有不同的处理方式,例如对于语言字段,英语的标识就有 EN、ENG、en - US、en、en - CA、English、English (United States)。为了解决这些异构性的互操作问题,DRIVER 采用对遵循 OAI - PMH 的数字仓储的登记、格式验证、基于数据清洗和转换的聚合等方式降低互操作的复杂性,其中 DRIVER 指南起到了至关重要的作用。

DRIVER 指南基于欧盟多个研究项目的经验,其编辑者来自法国的 HAL,荷兰的 DARE,德国的 DNI 和 BASE,英国的 SHERPA 等,DRIVER 指南主要用于指导数据提供者定义其本地数据管理政策,用 OAI - PMH 方式开放数据;帮助数据提供者、数字仓储管理者采取措施改进数据质量,帮助数字仓储服务开发者增加新的功能^[7]。DRIVER 指南从资源集合、全文、元数据标识符、OAI - PMH 的实施、最佳实践、字典和主题分类等方面为数字仓储建设提供了详细指南。它要求元数据采用 Unqualified Dublin Core (ISO 15836:2003),也可以使用 MODS,并提供了 OAI - DC 的各个字段的详细使用说明,以及采用 MPEG - 21 DDL 封装复合数字对象的指南,并对 OAI - PMH 实施细节进行了规范。遵循 DRIVER 指南的数据将被收割并聚合到信息空间,并提供集中的检索、浏览、推荐等服务;经过

基于网络用户界面的元数据验证工具,可以为仓储管理者提供元数据质量的分析报告,帮助仓储管理者改进元数据质量。

从 DNET运行环境看,呈现出异构和多样性特点。由于 DNET依赖于对多机构开发的服务的统一调度来实现功能,尽管各项服务均运行于 Ubuntu 操作系统上,但其不同服务的运行所依赖的应用服务器环境不同。比如,驱动层服务采用了 Jetty6 应用服务器,收割聚合服务、存储、索引服务和功能层的服务采用了 Tomcat5.5 服务器。在数据库方面也不同,比如存储索引服务采用了 PostgreSQL 8.2,在驱动层又采用了 eXistXML 数据库和 MySQL 数据库。

2.3 基于服务的架构

DRIVER 不同于其他的数字仓储登记聚合系统 (Repository Aggregating System, RAS),比如 OA Ister^[8], ROAR^[9], OpenDOAR^[10], Scientific Commons^[11], BASE^[12]的重要特点在于 DNET 基于服务的架构。

它们的共同点为通过仓储注册,从 OAI-PMH 仓储中收割聚合信息,提供 OA 仓储列表和仓储内容统计功能,并提供统一的检索、浏览界面等。

RAS 一般由一个机构开发,其功能和服务运行于某个机构的服务器上,满足某个领域或者定制应用的需求,其提供的功能和服务的重用性较差,软硬件扩展会有一些的难度^[13]。

DNET 基于服务的架构,由多个机构协同开发,提供的服务运行在网上不同的站点,每个服务组件提供特定的功能,每个服务组件可以被不同的服务提供者提供,并被不同的应用共享,通过服务的注册通知和订阅机制,在系统驱动层的服务调度管理下,根据系统要求实时提供所需要的功能,同时服务可以在任何时间加入或者离开系统。这种架构可以提高软件的开放性、可重用性和扩展性^[14]。目前 DNET 的核心服务已经被重用于欧洲数字电影网关^[15]、西班牙国家数字仓储联盟^[16]、比利时国家数字仓储联盟^[17]等应用中,并将在欧盟资助的 D4Science, OpenAIRE 等项目中得到进一步扩展。

综上所述,DNET 需要解决的核心问题包括:

(1) 如何从异构、分布的数据提供者处收割数据,并提供集中检索呈现的平台。

(2) 如何对异构分布的服务通过统一调度来实现

特定功能。

本文将通过对其信息空间和对象模型、基于服务的架构,尤其是对其驱动层服务的剖析,来说明 DNET 如何解决这两个核心问题。

3 DNET 的信息空间和对象模型

3.1 DNET 的信息空间

DNET 的架构就是把异构的内容聚合起来形成一个统一的信息空间^[18]。DRIVER 信息空间 (DRIVER Information Space, DIS) 物理上是由一系列存储单元 (MDStore 资源) 组成的。一个 MDStore 资源保存着被聚合服务从一个特定的仓储服务中定期收割来的符合特定元数据格式的元数据记录。在这个信息空间 (DIS) 里,从 OAI-PMH 仓储中搜集到的元数据被转换成 DRIVER 的元数据格式 (DMF), DMF 有多个,并在不同站点运行,DRIVER 管理员服务负责从指定的仓储的 Set 集合中收割记录并定义元数据映射规则,将原始的元数据转换成 DMF,装入 DIS。

3.2 DNET 的对象模型

DRIVER 的对象模型如图 1 所示:

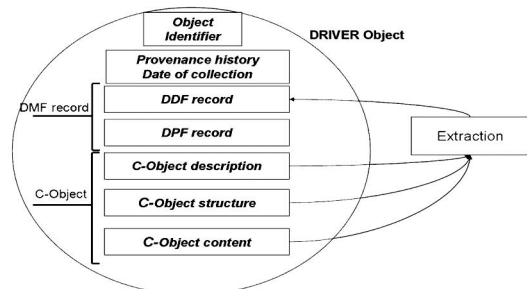


图 1 DRIVER 的对象模型^[19]

根据这个模型,任何一个数字仓储中的对象,也就是所谓的 C-Object,可以被搜集并导入到 DRIVER 信息空间,被封装成对应的 DRIVER 对象。DRIVER 对象的元数据格式 DMF 包括两部分:

(1) DRIVER 描述性元数据格式 (DRIVER Descriptive Metadata Format, DDF),它概括了描述、结构和 C-Object 的内容。

(2) DRIVER 来源元数据格式 (DRIVER Provenance Metadata Format, DPF),它记录了 C-Object 是从哪个仓储中搜集来的。

DRIVER 的元数据记录还包括一些头信息来说明

C - Object 的来源历史、DRIVER 对象的唯一标识符和 C - Object 被从原始仓储中收割的日期等。

DRIVER 信息空间的使用者是最终用户和应用。DRIVER 信息空间提供了对异构的 C - Object 的统一描述,终端应用就可以在 DMF 记录中进行高级查询操作。DRIVER 信息空间也开放基于 OAI - PMH 的接口,这样其他应用也可以连接到 DRIVER OAI - Publishing 服务来收割某段时间里 DRIVER 收割来的 C - Object 元数据。

4 DNET 基于服务的架构

DNET 系统提供了开放服务架构所需的服务,在此基础上服务可以通过彼此之间的交互来提供功能,DRIVER 利用这一架构提供从异构的外部仓储中收集内容、存储和索引、允许用户和应用集成地访问内容的功能。

4.1 DNET 服务架构概览^[19]

DRIVER 服务框架如图 2 所示:

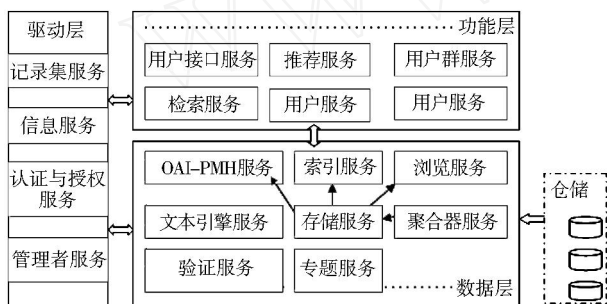


图 2 DRIVER 的服务架构^[19]

根据实现的功能不同,DNET 的服务可以分成三层:

(1) 数据层服务

聚合服务:处理数据收割和聚合。收割过程从外部仓储中收集数据,聚合过程实现收割来数据的转换和清洗。

数据存储服务:收割来数据以 DRIVER 对象形式被存储在 MDStore 数据结构资源中。

索引服务:DMF 记录、存储到 MDStore 中的记录,必须被索引数据结构资源索引,由索引服务创建。

专题服务:自动生成某个主题的数据集合。专题是 DRIVER 对象的虚拟集合,通过对 DMF 字段的预先查询获得。

浏览服务:提供按照语种、仓储、文献类型、出版日期、音序等的浏览。

OAI - PMH 服务:DRIVER 支持对外部出版的 OAI 发

布服务,因此它支持发布 UDC 格式、OAI Sets 单独的条目收割,同时也可以发布 DMF 格式,被收割的原始格式、具有 DPF 的格式^[20]。

数据验证服务:检查数据提供者提供的元数据是否符合 DRIVER 指南,保证数据质量。

文本引擎服务:提供语言探测和文本抽取功能,保证收割数据的语种和其他格式的标准规范。

(2) 功能层服务

用户接口服务:最终用户和其他应用与 DRIVER 信息空间交互的交互服务。

推荐服务:用户可以注册并描述自己的兴趣,从而获得特定主题新内容的推荐,为注册用户提供其所感兴趣的最新信息的通知。

用户群服务:允许定义某个专题集合为一个用户群,这样注册用户可以订阅该专题集合。

检索服务:为用户提供在 DRIVER 信息空间的检索功能。

(3) 驱动层服务

信息服务:跟踪注册到系统中的所有服务,并对这些服务提供与之相关的其他服务和资源状态的预订及通知功能。

管理员服务:为了最大化系统服务质量,可以调度和监测加入系统的服务。

授权和认证服务:提供了保证被授权用户和服务使用的安全机制。

记录集服务:负责管理记录集资源。记录集可以被许多服务创建和共享,是服务间互相交换数据的工具。

4.2 DNET 的数据聚合服务

DNET 的数据器服务负责数据收割和数据聚合过程。

(1) 数据收割过程

收割实体包括数据提供者仓储、主 MDStore、其他 MDStore、映射规则等。与一个仓储相关的收割活动包括:

一个注册到 DRIVER 的数据提供者仓储指定了其想要发布的元数据格式列表,也就是说用 listMetadataFormats 方法列出的记录子集,这个列表必须包括 DC。

系统创建一个与该仓储的收割相关的实体,要达到这个目标,系统为该仓储发布的每一种元数据创建一个 MDStore 来存放 DC 元数据,作为收割实体的主 MDStore。

收割实体被分配给一个聚合服务,它的配置由聚合器管理者指定映射规则、收割频率和收割开始时间来完成。收

割总是从主 MDStore 的 DC 元数据开始,然后开始其他格式元数据的收割。

(2) 数据聚合过程

聚合过程在把收割来的记录加到 MDStore 前清洗和改善收割来的记录。清洗过程调整收割来的记录并保存其原始的元数据格式,定义映射规则,完成基于词典的 dc: type、dc: date、dc: language 数据规范化,实现语言侦测、抽取文档链接等任务^[21]。这一过程的结果是感兴趣的应用可以通过连到 DRIVER,收割原始仓储服务的内容,过滤和提高其质量。

改善过程将 DRIVER 字段与收割来的记录关联起来,以便使它们成为一个对象的一部分。对于 DC 记录,定义对象的 DDF 记录,任何被清洗的记录都通过负责的聚合服务被附加上 DRIVER 元数据的唯一标识符 (dr: recordIdentifier)、OAI 条目的唯一标识符 (dr: itemIdentifier)、元数据被收割的日期 (dr: dateOfCollection)、DRIVER 对象的唯一标识符 (dr: objectIdentifier)。聚合服务对主 MDStore 进行特别的处理,MDStore 专门用于存储 DC 格式,用于在物理层展现 DRIVER 对象,被清洗的 DC 记录又被加上有关 C - Object 内容、结构和描述的 DDF 字段。

经过收割、聚合清洗调整后的数据,被存储服务存储到 MDStore 中,并以此为基础通过驱动层调度相应服务实现 DNET 的其他功能。

由此可知,DNET 的信息空间和对象模型为 DNET 从异构数据仓储中收割数据打下基础,而 DNET 的聚合服务实现了从异构数据仓储中的数据收割,同时 DRIVER 指南为 DNET 从异构系统中收割数据降低了互操作的成本。

4.3 DNET 服务调度的驱动层剖析

由于 DNET 的各项服务分别位于不同机构的不同服务器中,为了指挥调度相关的服务完成特定的需求,作为驱动层的服务起到了核心控制中心的作用。驱动层的服务根据系统要求的功能发现可获得的服务、指挥调度相关的服务合作协同,这对于异构分布式服务的实时运行和服务质量保障提供了重要的支持。以下将详细剖析驱动层的实现原理^[31]。

(1) 资源模型

允许松散耦合的服务之间合作的重要特征是发现服务及相关资源的能力,它保证了任何需要与另一个

服务的交互,以便实现期望的功能。为了达到在动态分布的环境中发现和调度相关的可获得的服务和资源,DRIVER 采用了如下思想:把相关的服务和数据结构定义为可用的资源,包括服务资源和数据结构资源,通过注册、订阅、通知机制,基于调度协议来发现和指挥可用的资源,以达到系统要求的功能。服务资源包括 4.1 节中提到的所有服务,代表运行网络服务的实例,它只能通过系统管理员来增加或者删除,而数据结构资源由服务根据相应的系统要求来创建、移除、更新数据结构,它们是被动的资源,受相关服务资源管理;系统通过调度 (Orchestration) 协议——Blackboard 协议来交互。

服务资源描述是 DRIVER 的资源在注册时提供的基本信息,可以在运行中实时更新。其目的是:代表资源的存在和可获得性;支持资源发现、存储;帮助系统正确处理资源以最大化服务质量的信息。

数据结构资源描述提供了相关资源的描述,与服务资源描述不同,它是由服务资源自己处理的。数据结构描述向负责数据结构的服务资源开放并被它处理。

(2) 信息服务

在 DRIVER 架构中,资源可以实时地被增加或者移除。信息服务 (Information Service, IS) 的角色提供在实时更新的系统架构中资源的组成途径和各个资源的运行状态情况。事实上,资源无论何时被加到系统中,IS 都会收到该资源的描述 (Profile),并注册登记。一个描述文档包括关于资源类型以及该资源实体特定属性的信息,它代表资源的存在,支持它的发现。同样,如果资源被从系统中删除或被修改,进入信息服务的描述文档也必须被删除或做相应的修改。这样信息服务就在资源注册方面扮演了非常重要的作用,资源可以查找其他与特定类型和行为属性相关的资源,在 IS 的描述文档中查找并确定哪个是所需要的服务。此外,系统也通过预订、通知工具支持资源得到某些与其他系统资源相关的事件 (包括资源描述的删除、改变和创建) 的通知。

(3) 管理服务

管理服务是 DNET 架构的指挥中心,这是为了在系统资源创建、更新和删除的过程中保证服务质量而设计的。为了达到这个目的,管理服务必须能发现可

得的服务资源,并发给它们要完成操作的指令。在 DRIVER 中,管理服务和服务资源的通信是通过调度协议来实现的,这个协议采用了信息服务的订阅和通知机制。因此,管理服务可以订阅并及时得到任何信息服务中可能的事件的变化。这种通知的频率可以定制,并根据可获得服务的情况,采取不同的响应行动。例如,在 DRIVER 中引入一个新的仓储的活动包括以下步骤:

管理系统在信息服务中查找是否有可以获得的 MDStore 服务来创建 MDStore;

根据订阅的信息服务中空闲的相关服务的状态,确定与一个空闲的 MDStore 服务进行交互,并创建 MDStore;

从信息服务中查找一个可以进行仓储聚合的聚合服务;

命令聚合服务管理仓储收割,并存入所创建的 MDStore 中。

(4) 调度协议

调度 (Ochestration) 协议是一个简单的两步协议,规定了在管理服务和服务实体之间管理信息的交换,它基于以下设想:

用“黑板”(Blackboard)的方式显示管理服务信息;

服务只要一注册,就将它的描述预订到自己的描述黑板的更新部分;

管理服务预订了所有服务的描述黑板的更新信息。

该协议定义了服务的几种状态:已分配任务 (Assigned), 任务进行中 (Ongoing), 任务完成 (Done), 任务失败 (Failure)。

Ochestration 协议的步骤如图 3 所示:

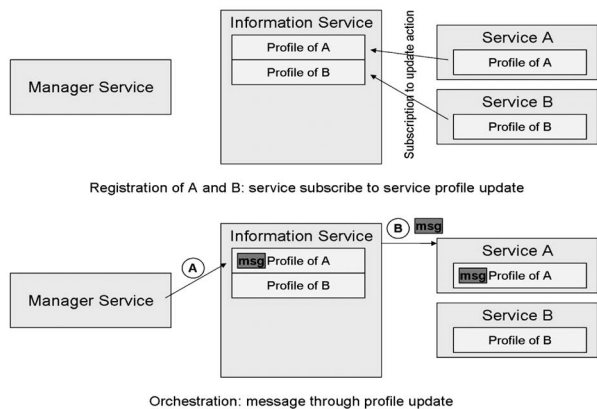


图 3 调度协议的工作过程^[19]

当管理服务决定与一个服务通信时,它发出信息

以更新 IS 中相应的描述黑板 (Step A), 相应的服务得到通知告知 IS 有改变, Step B 基于该信息中所包括的服务描述,解释该服务的功能、状态。

通过这种机制,管理服务也可以操作数据结构资源。例如,系统需要为一个上载了新的元数据的 MDStore 更新其索引数据,系统首先查找负责索引的索引服务,与之通信并在给定的索引上进行刷新索引的功能。当服务完成了被请求的操作后,它必须用一个确认信息更新它们的描述黑板,这个响应允许管理服务来检查原有的操作成功或失败,并在必要的情况下激活一个补救操作。

(5) 授权和认证服务

DRIVER 的认证和授权服务提供了利用认证和授权机制加强系统安全性的功能。该服务的架构基于 eXtensible Access Control Markup Language (XACML), XACML 处理分布的异构企业级系统的认证和授权。

5 结 语

5.1 DNET 的适用性

DNET 的设计目标是构建分布式的、异构的、动态的知识架构,其服务实体可能分布在互联网上不同的地方,通过驱动层调度这些具有不同功能的资源和服务,实现特定的功能。因而 DNET 比较适合于大规模的、需要在异地部署不同服务、强调通过动态调度符合需要的可获得的资源和服务,以实现特定功能的大规模行业或者国家级数字仓储联盟。

5.2 标准化和开放性在数字仓储联盟中的重要性

尽管 DNET1.2 将提供对不同的数字仓储开放的不同元数据格式向 OAI-DC 的映射,但标准化在对个体数字仓储数据的收割聚合中的重要性依然不能忽视。在 DRIVER 实施运行中,DRIVER 指南起到了非常关键的作用,它提供了对于元数据格式、开放元数据格式的实施细则,并对一些语义和结构信息进行了详细的规范。在数字仓储及仓储联盟建设实施过程中,无论采用什么样的软件平台,都必须遵循国际通用的标准,对标准的扩展也应该采取慎重而合乎标准要求的做法,否则必定将会造成与外部数据交换的障碍,提高数据收割聚合的复杂度和仓储联盟建设的成本。

5.3 仓储联盟可能提供的服务

仓储联盟可提供的服务是其得以存在的基本理

由。只有当它能够提供不可替代的服务的时候,其存在才具有合理性。从 DNET 当前的服务来看,主要提供集中检索、浏览、按主题浏览、按特定需求推荐订阅等基本功能,但 DNET 的后期版本将提供引用分析、增强出版物、可视化分析等服务,这将会提高其吸引力。

(致谢:本文受国家留学基金和国家科学图书馆骨干人员提升计划支持,并得到 DRIVER 项目承担者德国哥廷根大学图书馆、Bielefeld 大学、意大利国家研究中心 (CNR) 的大力支持。在此感谢 Dr Norber Lausso, Dr Dale Peter 的帮助。)

参考文献:

- [1] Lossau N, Peters D. DRIVER: Building a Sustainable Infrastructure of European Scientific Repositories [J/OL]. *Liber Quarterly*, 2008, 18 (3 - 4): 437 - 448. [2009 - 10 - 23]. <http://liber.library.uu.nl/publish/articles/000267/article.pdf>
- [2] CORDIS E - Infrastructure [EB/OL]. [2009 - 10 - 25]. http://cordis.europa.eu/fp7/ict/e-infrastructure/home_en.html
- [3] Candela L, Castelli D, Manghi P, et al Enabling Services in Knowledge Infrastructures: The DRIVER Experience [EB/OL]. [2009 - 11 - 02]. http://debs2007.cah.unipd.it/Members/parolo_manghi/ircd107_isti2.pdf/view
- [4] Horstmann W. The DRIVER Initiative for Networking Repositories [EB/OL]. [2009 - 11 - 13]. http://www.driver-support.eu/documents/DRIVER_OA6_horstmann.ppt
- [5] DRIVER Search Portal [DB/OL]. [2009 - 11 - 10]. <http://search.driver-research-infrastructures.eu/>
- [6] Horstmann W. Diversity of Digital Repositories in DRIVER [EB/OL]. [2009 - 11 - 02]. http://www.debs.info/files/pdf/DL-Foundations2008/7_HorstmannDLFoundations08.pdf
- [7] DRIVER Guidelines 2.0. Guidelines for Content Providers - Exposing Textual Resources with OA I- FMH [EB/OL]. [2009 - 11 - 02]. http://www.driver-support.eu/documents/DRIVER_Guidelines_v2_Final_2008-11-13.pdf
- [8] OA Ister [DB/OL]. [2009 - 12 - 02]. <http://www.oclc.org/oaister/>
- [9] ROA I [DB/OL]. [2009 - 12 - 01]. <http://roar.eprints.org/>
- [10] OpenDOAR [DB/OL]. [2009 - 12 - 01]. <http://www.opendoar.org/index.html>
- [11] Scientific Commons [DB/OL]. [2009 - 11 - 01]. <http://en.scientificcommons.org/register-repository>
- [12] BASE [DB/OL]. [2009 - 11 - 10]. <http://www.base-search.net/>
- [13] Manghi P. The DRIVER Repository Infrastructure [EB/OL]. [2009 - 11 - 12]. <http://www.tilburguniversity.nl/services/lis/ticer/08carte/publicat/manghi-slides.pdf>
- [14] Manghi P. The DRIVER Repository Infrastructure Digital Libraries [EB/OL]. [2009 - 11 - 02]. <http://www.tilburguniversity.nl/services/lis/ticer/08carte/.../manghi-slides.pdf>
- [15] European Film Gateway [DB/OL]. [2009 - 12 - 02]. <http://www.europeanfilmgateway.eu/>
- [16] Recolector de Ciencia Abierta [DB/OL]. [2009 - 12 - 02]. <http://search.recolecta.driver-research-infrastructures.eu>
- [17] DIAL - D   t Institutionnel de l'Acad  mie Louvain [DB/OL]. [2009 - 12 - 02]. <http://search.belgium.driver-research-infrastructures.eu/>
- [18] Artini M, Biagini F, Manghi P, et al OA I - Publishers in Repository Infrastructures [EB/OL]. [2009 - 11 - 22]. <http://www.nmis.isti.cnr.it/manghi/papers/mypapers/ircd12008.pdf>
- [19] Architecture Overview [EB/OL]. [2010 - 01 - 10]. <http://technical.wiki.driver-research-infrastructures.eu/index.php/Image:Architecture.jpg>
- [20] Artini M, Candela L, Castelli D, et al Sustainable Digital Library Systems over the DRIVER Repository [EB/OL]. [2009 - 12 - 10]. <http://www.springerlink.com/index/5770324118361628.pdf>
- [21] Summann F. BASE and DRIVER [EB/OL]. [2009 - 11 - 10]. http://www.rsp.ac.uk/events/ServiceDay2009/bath_rsp2009.pdf

(作者 E - mail: majx@lzb.ac.cn)