

全球研究型图书馆发展战略

——2008年GRL2020会议各参会机构未来发展宣言书

鲁超 尚玮娇 吴思竹 苏娜 鲁宁
王欣 孙敏杰 张红丽 编译

编者按:GRL2020(Global Research Library 2020)会议是面向未来数字图书馆发展,旨在建立和形成未来图书馆发展愿景的国际研讨会,为各图书馆提供一个平台,通过讨论交流,明确图书馆未来发展方向和所要采取的必要的计划。GRL2020计划每年在世界的一个大洲举办一次研讨会,2007年的第一届研讨会在北美洲的美国华盛顿举办,2008年在欧洲的意大利举办,2009年的GRL2020在亚洲的中国台湾举办。

每届会议参会的各图书馆及相关机构的代表向大会提交该图书馆的未来发展宣言书(Position Paper),陈述本馆未来发展思路和设想。下文为2008年GRL2020会议23个机构的宣言书的汇总报告,删除了报告前三部分的会议介绍等内容,重点对各国的发展规划部分进行编译和整理。

2009年GRL2020于今年2月底刚刚结束,我们将继续关注会议内容,并在《图书情报工作动态》随后几期发布,敬请关注。

1. 华盛顿大学2020年未来发展宣言书

Harry Bruce,华盛顿大学信息学院院长

鲁超、尚玮娇 编译

我以研究员、研究主持人和学术带头人的身份来写这份未来发展宣言书。当然,全球研究图书馆在2020年有许多方面的设想,但我想侧重于通过三个具体的机会来支持和促进学术研究和创造新工作。最近几年,我从事了由美国国家科学基金会资助的两个研究项目,这两个项目涉及到了我关于全球研究图书馆在2020年的愿景。第一个项目是集成信息检索的探索:如何进行寻找和利用那些可以满足各种信息需求和团队成员共同的问题的信息资源;第二个项目是一个研究计划,关注是个人信息管理:个人如何去组织和管理对其生活有价值的专业或者私人信息,第二个项目重点研究探讨个人如何组织和安排自己的个人信息集合。与此同时,我

还与不同的研究小组一起工作,在那里我的工作就是提供便利,支持和促进团队成员之间有关知识和创新的工作,并建立研究计划。我一直是一个研究过程的观察者,特别是对规模较大、多学科项目交叉的复杂问题。我对于怎样支持这些研究小组、他们面临的挑战、工作的流程有充分地了解。我还观察到正在形成的新的科学调查过程的趋势,这个趋势的基础是跨学科协作、计算模拟、传感器技术和操纵庞大的数据集。我认为全球性研究型图书馆在2020年的愿景是可以和这些正在形成的科学调查的未来趋势保持一致的,并作为一个关键的、积极的、有相关性的和可以增值的部分存在于研究过程中。考虑到这一点,我建议研究图书馆和大学图书馆

可以采取的一些行动。当然,这不是一套完整的行动。在此我想简单介绍一些想法,希望在 2008 年的 GRL2020 会议上进行讨论。

第一个建议是,研究图书馆必须与新一代涉及到数字时代的新用户相联系。研究图书馆历来侧重于通过深入了解用户需求和他们查找和使用信息的首选方法来满足用户的需求。在 2020 年,所有的本科生、研究生和大部分助理、副教授(即:使用研究型图书馆的大多数研究人员)将是一代在“数字”中长大的成员。正在进行的研究告诉我们,这新一代的图书馆使用者学习的知识和处理的信息是不同的。他们沟通的方式、解决问题的方式、书写的方式不同,需要使用适当的信息技术、以现在还不能预见到的方法来满足他们的需求。这合乎情理,这一代的研究图书馆用户将创造新的不同的知识,并进一步以科学咨询和学术交流的方式把知识传授给他人。他们将需要一系列非常不同的图书馆服务来支持他们的需求和工作流程。2020 年的全球性研究型图书馆必须了解和预测这新一代用户的需要——他们是如何学习和使用信息将从根本上影响未来的研究图书馆和大学图书馆的业务工作(参考咨询服务、信息提供服务、馆藏资源建设等)。

第二个建议是可能与目前的共识相冲突,关于作为场所的图书馆在数字信息世界中的地位。有的人认为,只要用户能够方便地在任何地方、任何时间获取数字信息、或者是在移动的过程中找到和使用信息,研究型图书馆就不再是一个时间或空间的概念。诚然,因为远程信息服务的发展,数字信息可以在难以置信的广阔范围内提供给用户服务。这并不意味着研究型图书馆作为一个物理空间是无关紧要的。事实上,我想强调的是,对研究图书馆进行包括合作空间在内的重新设计的重要性。到 2020 年,研究型图书馆将成为研究团队聚集的空间。解决问题、学术交流、学术研究的创造性工作过程正越来越依

赖合作的、智能化、面对面的交流合作,而这些活动都在由有着专门设计的空间中进行引导,这一空间通过信息技术来提供信息的共享访问(shared access)、利用技术捕捉新知识进行集成。的确,通过数字信息获取技术可以将研究人员聚集起来,但是很多创新活动需要将这些信息进一步转化成新的知识,而这种转化通过面对面的交流实现的。我对 GRL2020 的期待是:灵活的物理空间,可以容易地转化为研究性“合作”的场所。这些空间将由技术来支撑,装载(或技术能够容易地集成到这一空间中去)那些促进和支持共享访问和参与信息、知识创造、新知识捕获、出版和传播活动的技术。

我的最后一项建议与前两个建议相关:2020 年全球研究图书馆在形成和支持 eScience 中必须发挥重要作用。一种新的科学调查框架显示,某一个领域的科学家(现在大部分有特点的自然科学领域)已获取由环境传感器保存和捕获的大量数据集。这些原始数据在一个或多个相关学科领域中被组织、管理、计算、研究,以新知识的形式出现,增加了新的价值。然后这个新的知识成为进一步积累新知识的原材料。Tony Hey 在西雅图的第一届 GRL2020 研讨会上提到以数据中心的科学(协作、网络化、多学科的科学)。这种新科学将被新一代研究图书馆用户(数字居民)所利用。这将要求我们重新定义空间意义上的研究型图书馆,将会在图书馆传统业务的基础上(例如知识组织、馆藏建设、信息提供)增加不同的需求。同时,这还将要求研究型图书馆确立一套新的业务和机制与大量未处理的数据进行联系和管理。

我在报告中提到的 3 点想法当然无法总结或完整捕获全球性研究型图书馆的未来。相反,这 3 点的目的是希望能够有助于第二届 GRL2020 研讨会的讨论。我在这个报告中强调的观点是空间、协作和未来研究型图书馆参与的重要性。这不应该被解释为研究型

图书馆的全球性观点。我强调,在研究型图书馆中本地支持对于提高研究人员的积极性的重要性;然而,这种本地支持必须将研究过程与团队成员、以及利益相关者甚至是全球的用户都联系起来。全球性研究所需要的本

地支持随着数字资源比例的增长而有所增加。这可能不太需要面对面的交流,但现在,老一代研究人员承认喜欢面对面的接触来寻找所需要的信息。为此,我期待着第二届GRL2020研讨会的召开。

2. 荷兰莱顿大学 2020 年未来发展宣言书

Kurt De Belder, 荷兰莱顿大学图书馆馆长

鲁超、尚玮娇 编译

教师和学生为了研究和教学目的需要大量的、高质量的数字化资料,包括历史资料,最好是完整的人类记录。这些数字资源应该是可信赖的、可检索到的、并以有意义的方式组织起来。此外,教师和学生希望对这些资料进行注释、汇集、拆分、关联、插入内部链接,作为其学术研究或学习活动的一部分。而那些水平更高的研究专家们希望使用更先进的技术,如数据和文本挖掘技术、能够验证假说或提高他们研究水平的工具。他们对研究活动中接触到的相关机构(图书馆和出版商)感到困惑,这些机构分别拥有一些资源,而在格式上又不兼容,这又限制了技术上对这些资源的整合、进行专题索引等等。但另一个问题越来越明显地显现出来,他们无法用智能的方法利用以他们不懂的语言书写的论文。他们面对的现实是,图书馆和出版商仍然重视以文本为导向的资源,然而研究人员更希望得到更多的图像、数据等。此外,教师和学生也希望能够创建数字出版物,使他们不用成为技术人员或者不用必须了解如何使用复杂的XML语言就可以和外部大环境进行沟通交流。此外,他们希望与志同道合的研究人员通过如协作等方式进行交流,前提条件是这些方式增加了他们的工作附加值而不只是消耗时间。

很难试图描述在2020年图书馆将会是

什么样子的,但是更重要的是这些变化发生的大环境是什么样。过去15年里,一些颠覆性的技术已经对我们的环境产生了相当巨大的改变,并且有时候这些变化不期而至或者是我们未曾预料到的(例如:http协议、网络浏览器、短信编辑、移动技术、Google等)。与此同时,一些已经被频繁地宣传的技术(例如:电子论文)却没有产生预想中的效果。其他的正在发展的技术趋势(Web 2.0、社交网络、语义网)还没有正式形成,它们在学术信息交流方面的影响,以及如何在研究人员和学生交流、创造新知识方面的作用还不明朗。因此,换句话说,某些意想不到的技术的发展可能创造一个完全不同的背景和竞争环境,这些新的背景和环境可以为现有的和新成员、方法、概念等的创造和颠覆提供机会。让我们想一想印刷术的产生,它是经过多长时间才确定下来的,又例如文章标题、索引、版权、固定信息的表述、作者和出版者的明确定义等,涉及到宗教、启蒙、理性的话语、民主和民主式的学习等。这不仅很好地说明了技术目前的发展规模及其影响,也可以表明目前的动荡或者不确定性不仅是发展进程的必要部分,而且将可能或必然在一个完全不同的大环境下确立下来。

当然,在第一部分提到的观点我们期望在2020年可以实现。然而,如何会出现这种

情况并产生影响取决于:(1)机构,例如出版商、图书馆和大学等;(2)概念,例如版权、同行评议、信息的完整性等。(3)获取信息的商业模式、工作的安排、金融财政等方面的改变

是很难预测的。大学机构天生的保守性对其本身的发展是没有帮助的,或许可以说是毫不相关的。

3. 史密森学会 2020 年未来发展宣言书

Thomas Garnett, 史密森学会图书馆数字图书馆副馆长

鲁 超、尚玮娇 编译

在目前,即使是最先进的数字仓储系统,仍有许多需求有待解决。首先,网络服务中用于多系统之间有效消除模糊情况的工具还没有出现。例如,消除地点歧义的工具、或更具体地说通过时间和广义的电子名录寻找地理位置;消除多语种人名歧义的工具;消除生物类群歧义的工具(尽管百科全书中生物信息部分已经顺利地解决这个问题)。其他要求包括采用 Unicode 字符的光学字符识别软件来处理历史文献;特定领域文本的自动语义标记软件和一个专业型的数据库。

目前,出生在数字时代的人们和老年人之间的代沟越来越明显,体现在他们如何组织文本数据、如何管理数据、使用什么软件、甚至这些数据是如何概念化等方面。在某些学科如生物的多样性方面,这是一项耗费巨大的宏大研究。

目前解决这些问题的活动正在进行之中。这些措施包括利用 OCR 计划将有关生物多样性方面的资料数字化,提供分类的结构化检索工具使读者通过智能软件公开获取所需的资料。此外,多学科和国际合作伙伴关系可以有效地管理长期的大型研究项目的资源,也正在同时改变科学交流模式,例如科学期刊的“回溯”正在进行中。

展望 2020 年,全球性研究型图书馆的愿望应包括以下几点:

- 跨学科和国际合作伙伴关系将有效地长期管理大规模研究资源;
- 非赢利性或学术机构与商业机构开展合作;
- 去中心化的研究模式;
- 克服科学数据库集跨国、修复、长期保存中的各种障碍。包括政治障碍、知识产权障碍、学科边界障碍、机构障碍、经济障碍等。

为了实现这一目标,应采取一系列措施。首先,在适度控制下共享长期保存的研究数据将可能带来巨大的经济利益。然而,控制的程度需要确定下来。大范围传播所产生的费用造成了一定的经济问题、专业人才的职业生涯规划等也是需要解决的问题。

这种共享或聚合可以集中或分散,可以是一个组织或一系列实践活动和文本协议,甚至可以被叫作“图书馆”。这种共享要求建立一个虚拟的或合作的组织机构,需要具备专业技术的工作人员、新的职业生涯规划、需要来自机构、科学界和国家的支持,包括资金支持、愿意贡献内容的激励机制、用户服务的指导、研究团体的参与、灵活的组织模式、以及低廉的行政开销。

4. 欧洲航天局 2020 年未来发展宣言书

Luigi Fusco, 欧洲航天局高级顾问

鲁超、尚玮娇 编译

本宣言书提出了研究型数字图书馆在地球科学和地球勘测领域应该承担的主要作用和主要发展方向。地球勘测是指在适当的时间和准确的空间来大范围地观测环境现象。其他地面数据收集系统(如浮标)可以补充空间数据来满足地球科学界的研究与应用所需的资料,为科学成果和出版物中提供详细的知识,同时又可以应用到监测地球环境的实际中来。

支持类似地球科学领域的门户网站或客户端应用程序的 ICT 服务必须处理连续的不同类型的数据源,这些数据需要收集、组织并提交给它们的使用者。管理这些数据需要考虑的不仅仅是大量的信息,还必须考虑可以使用和重复使用这些内容的不同情况、以及可能适用于它们的不同分类模式。迄今为止,地球科学界的努力集中在开发和利用非常先进的标准元数据规范和服务来处理数据,例如由开放地理空间(OGC)和 INSPIRE2 所提倡的这些标准。地理空间和地球科学有关的信息提供商保存其数据档案、并从大量由复杂用户群通过网络使用的数据库中获取知识。

以地球科学领域为服务对象的研究型图书馆表现出缺乏获取和检索上述数字资源并对其进行管理和共享的信息途径。尤其是在虚拟化和个性化的环境中,这方面问题尤其突出。虽然,试图在这一特定的领域内建立专门为项目服务的数字图书馆,但是目前在大多数的服务和获取方式中最缺乏的是建立以用户为中心的地球科学信息获取方法。这是因为缺乏限制性搜索功能和从标准模板与数据库内容中定期自动产生的固定的样本信

息会产生限制作用。虽然地球科学信息的表述是动态的,根据不同的工作环境变化,根据用户的需要基于数据系统收集不同来源的数据,根据用户的使用偏好进行不同的聚合(工作原理类似于 RSS),但是这种限制是非常严格的。

在欧洲航天研究所开发和整合数字图书馆和网格技术的例子中最具代表性的项目是 DILIGENT 项目。技术整合方向是使最终用户在分布式计算和大量存储空间中得到受益;并实现获得高层次服务和检索信息。从地球科学研究领域的角度来看,研究型图书馆的发展方向应该是:

- a) 依靠无限制的硬件资源(磁盘空间、处理能力、网络容量等);
- b) 实现全文和地理空间的检索和浏览;
- c) 提供相关设施确保:基于用户需求的虚拟空间的合作,通过精选用户感兴趣的内容确定个人偏好和观点,使用语义标注信息来丰富用户的观点、检索说明和信息的融合;
- d) 通过提供用户插件、算法以及应用程序来实现兼容;
- e) 易于使用、尽可能拥有。

尤其是在 ICT 研究领域,电子化已再经常被研究人员和开发人员关注。而电子化在地球科学领域已经产生巨大的影响,每一天,世界范围内不同文化背景的科学家迫切需要使用、快速获取信息以及正确处理信息的软件设施。

DILIGENT 测试表明有用的概念(例如虚拟组织、用户工作场所、馆藏、复合型服务和日常文件等)带有创建数字图书馆的需求和设计报告的信息需求,并在虚拟研究场所中

浏览、检索和管理用户的信息。用户界面的交互性是以现代技术为基础的,这些技术基本上是动态 HTML 和具有互动性功能的内容实体,例如:在用户查询后所产生的移动互动

知识单位。这些可以被检测到,并被注释成用户定义的本体,整合和保存到用户相关信息当中,同时也可以被用户重新调出使用。

5. 洪堡大学与欧洲图书馆 2020 年未来发展宣言书

Stefan Gradmann, 洪堡大学图书情报学教授

Jill Cousins, 欧洲图书馆馆长

鲁超、尚玮娇 编译

我们目前的工作对数字图书馆的影响体现在两个主要方面。第一个是 Jill Cousins 和他的欧洲图书馆团队。第二个项目是 Europeanana, 该项目的目的创建跨领域的原型, 解决一些突出问题, 提出未来可能的研究领域。这个项目由 EDLnet 办公室负责运行, 并由一些外部专家辅助指导, 例如 Stefan Gradmann 就是其中一个团队的负责人。

1、欧洲图书馆设在荷兰国家图书馆之中, 是国家图书馆馆长会议服务的一部分。它是一个汇集性的门户, 将欧洲所有国家图书馆中数字化和非数字资源汇集在一起。欧洲图书馆提供一个统一的检索入口, 可以检索 33 个国家图书馆的馆藏和目录信息。数字资源的数量达到 600 万件。到 2008 年末, CENL 的 47 个成员馆中已经有 45 个馆加入欧洲图书馆。

2、Europeanana : 欧洲数字图书馆网络, 其目的是建立一个跨领域(博物馆、图书馆、档案馆、视听资料收藏中心)的门户。目前, 在它的网络中有 90 个这样的组织, 组成了 5 个工作小组, 旨在解决人力、机制、创建联合门户的技术、语义等问题。为 5 个以解决人权和政治、创建联合门户中的技术和语义问题为目标的工作组提供帮助。该项目始于 2007 年 7 月, 迄今已产生了一系列的成果。在 www.europeanana.eu 上的图表, 显示了用户的

基础要求。本次演示正在邀请用户进行测试, 截止到 2008 年 11 月, 以便确定哪些可以确定为原型。该原型是概念的一个证明。

未来数字图书馆的发展需要实现一些要求。对欧洲图书馆和 Europeanana 的主要要求是为使用户获取在文化遗产部门所拥有的资料, 以满足用户的需求。这需要通过先进的 W3C 的架构和语义网以及多语言互操作性来实现。这意味着找到资源比发现资源的存储地更为重要。检索应当是无缝的和检索结果应当是集成并且有序的。这三方面在跨领域研究中仍然任重而道远。

第二个要求是不论其原始格式怎样, 数据需要做到真正的互操作。这就要求: 功能和体系结构模型在 W3C 的基础标准上构建, 而不是某个图书馆自动化的解决方案; 跨越文化遗产鸿沟的数字信息的模型; 新的数据描述和可视化解决方案; 系统自动生成和改进元数据以便使基本的字段(如作者、标题、主题、描述、时间和地点)标准化并显示出来。对于所有研究来说可获得基本集中化工具是非常重要的。这包括对所有有关本体进行描述的多语种搜索工具和配合各种软件使用的工具, 并需要在线知识字典等。此外, 所有规范文档的登记将有助于资源的发展和补缺。经过许可的登记工作和授权的语义角色, 以及数字指纹登记技术和数字水印也将作为基

本集中化工具的例子成为必要的技术。

其他对数字图书馆的要求包括:不受谷歌算法的影响而按照搜索结果进行排名;数据加权和结果排名;用户使用会话 cookie 和日志分析影响排名;可扩展多语言搜索应用程序;结合了向量空间和时间的可扩展的浏览器;持续标识已有的数字化遗产和新产生的数字内容。

意识到目前的现状是非常重要的。最近完成的包括涉及到多语言网络能力的 EDL-project 在内的欧洲图书馆的活动促使欧洲图书馆网络的本地化和其多语言的能力。这是旨在改善终端用户以一种标准化的方式通过多语言界面和先进的搜索机制获得所需信息。

目前欧洲图书馆的活动包括:TELplus。其目标是做到以下几点:

- 增加和改善欧洲图书馆的内容:通过 OCR 技术增加超过 2000 万个重要的多语言页面,这些页面根据目前的技术水平仅仅可以作为图像来应用;收集国家图书馆目前所有数据库中的数据(只能通过 Z39.50 协议获取),因此收获到了更大可用性的数据,增加了保加利亚和罗马尼亚的国家图书馆作为欧洲图书馆的正式成员。

- 改进获取和利用:提高全文索引和调研自动词汇映射;通过欧洲图书馆的用户群向用户学习;创造一个模块化服务的基础设施,使用户能够把服务与欧洲图书馆门户网站整合在一起,并增加了一些新的服务;个性化的用户服务为新工作提供了方向;提供 26 种语言馆藏翻译。

- Europeana / EDLnet (Europeana. eu) 负责以下内容:

- a) 表明哪些是 www. Europeana. eu 用户所要求的

- b) 可交付的 2.2 版初始语义网技术和技术互操作性要求。

这个项目进一步探索通过数字图书馆来通报研究的现状和建议将是非常重要的。

由欧洲联盟委员会提出的愿景是更加协调地进行研究、基础设施和内容建设。这种状况正在改善,但似乎很少人对所有的有关项目以及它们如何对全局有利有全面的了解。零碎的活动很少有具体的可执行的成果,并迫使项目去符合 Europeana 的一些短期的目标利益,而且可能有长期的负面影响。由新机构建立的欧洲范围内的协调网络可以使必要的权力下放,但要保持最低限度的重复,并确保项目彼此合适。这是由 EDL 基金会启动的,但这一基金会需要可持续的资金来以继续进行该项目。这样一个机构可能与其他工作和研究领域中的数字图书馆有着密切的联系。

Europeana. eu 是走向共同的、协调地利用欧洲文化的一个具体步骤。在实现这一目标方面,Driver 项目的工作也是非常重要。此外,各方面的合作计划正在建立。正在谈判中的项目有:Arrow、Athena、EDLocal、EFG 和 EPA 等,这些项目都有助于标准的制定。接下来的步骤是根据目前的 FP7/ICT 和 eContentPlus 要求进行进一步的研究和开发。EDL 基金会将为此提供坚强的财政支持。

6. 莫纳什大学 2020 年未来发展宣言书

Catherine Harboe - Ree, 莫纳什大学图书馆长

鲁超、尚玮娇 编译

当一个信息模块与另一个相冲突时,当面临如何调整方向并在新的场景中航行时,图书馆必须挺身而出迎接这种挑战。关注图书馆的传统价值和作用,研究型图书馆必须大胆和敢于冒险。它们必须成为开拓的先锋团队,也必须突破它们传统的核心价值观,为那些信息权利受到危害的人们而努力。这种冲突挑战着研究型图书馆的维护其核心价值和提供传统服务的保守性,并且能够以新的方法和价值形式来定义这些服务。

研究型图书馆是最接近这些所谓的碰撞点的。它们是世界上最重要的学术交流活动的管理者;这一点,过去是真的,当前也是如此,很可能延续到未来。然而,随着越来越多的信息在网络中出现,图书馆的作用正发生着巨大的改变。很可能至少在可预见的未来,信息将继续被埋没于某处,并且彼此之间联系很少很难掌握。信息资源的庞杂性使研究人员更难找到他们所需要的信息,而不是更容易。研究图书馆必须找到办法来帮助他们的社区来理解和掌握这一新形式。这需要通过将系统、发现、访问路径和用户技能培养等融合起来。

研究型图书馆必须监测他们身边的变化并应对相应的用户行为的变化。而在过去,图书馆是被照顾和管理的对象,今后它们必须成为指导者和教师。在过去,图书馆提供获取信息途径;在未来他们必须帮助发现和利用信息。

在出版、管理或维护其机构研究成果等方面起到越来越重要作用的研究型图书馆正在改变着这种状况。近年来,许多图书馆已建立机构库来储存、传播和维护之前和之后

发表的文献、其他的研究性出版物、学位论文、专利和许多其他类型的数字资源。产生这一现象的动机之一是可以获得更多的科研成果,另一个动机是为其机构提供信息管理服务。

一些图书馆为某些机构提供公开的或封闭式的正式的信息发布渠道。这是一个为了学术沟通而进行研究和实验的时代。图书馆已经作为选手步入这个空间,必须对正在发生的事情评头品足,并与其他机构合作以取得最好的结果。他们必须努力避免重复传统的学术出版形式,必须和其他机构一起公开地分享信息。

研究型图书馆近年来成为数据管理领域中的领袖,数据管理领域已成为一个重大而紧迫的挑战。图书馆的传统资源和信息管理方面的专长与不断增长的数字信息所带来的知识管理相融合并承诺最高的公众获取权。数据管理需要机构间新的技术和不同的资源配置。在数据管理的利益相关群体中,图书馆具有独特视角和独特的技能。

由于现在图书馆工作人员迫切需要新的或发达的技术,图书馆、信息管理学校以及专业协会必须意识到这一新的现象并调整其课程和培训项目,研究型图书馆必须像学术图书馆那样,找到程度越来越高的专业化的应对办法。

在世界上一一直存在着信息的不均衡现象,这种现象可能存在于资源丰富和资源贫乏的机构之间,讲不同语言的人民之间,受到良好教育和未受过教育的人之间,发达国家和发展中国家之间。图书馆意识到这一点,并试图打破障碍从而获取信息。这种情景有

益也可能有害。图书馆必须努力使用新技术来改善信息的获取,确定和游说反对那些与增加公众利益相悖的法律、政策和做法。

我对于实现 2020 年研究型图书馆发展

愿景的建议是增加信息共享性、确定发展方向(例如西雅图会议宣言)、设立工作组就具体问题进行交流、年度首脑会议等。

7. 比勒费尔德大学 2020 年未来发展宣言书

Wolfram Horstmann, 比勒费尔德大学图书馆首席信息官

鲁 超、尚玮娇 编译

研究图书馆面临着远离研究人员和学生的挑战。这种远离是基于越来越多不同的服务要求和两极分化的用户群体。首先是由目前大多使用电子期刊及文章的自然科学家和计算机科学家组成的群体今后可能会使用更多的分布式和可变的互联网资源。第二个群体是由艺术和人文科学的研究人员组成的,他们通过传统的专题性资源以及其他的印本资源来满足自身的需求。第三个两极化的是年轻的学生或研究人员越来越多的使用图书馆外部的信息管理服务,如 GoogleScholar、Moodle 等,而不是使用学科数据库和“粉笔加讲课”的方式。还可以找到更多不同的用户群体的例子。为清楚起见,本文将它们定义为“常规型”和“新型”的用户。

这种情况所导致的主要问题是,无论传统的和还是新型的用户,他们的要求都必须被满足。即使图书馆正采用数字图书馆的形式服务于新型的用户,他们将不可避免地“失去”用户给其他服务提供商。在不变的预算条件下,这种情况导致图书馆不得不作出艰难的选择为两个用户类型中的一种提供服务。

在此背景下,协作和共享服务的信息管理成为一个非常重要的生存策略。正如学生们共同分享他们的信息平台,如 Facebook;又如和研究人员合作分享他们的项目数据,如

人类基因组计划或开放获取项目中的出版物,图书馆也可以共同分享他们的信息管理服务。这并不是新的想法,类似的事情在图书馆之间系统地交换印刷型论文时已经做了几十年了。在数字开放获取的时代,这作为检索服务、版权解禁、引文分析等方式广为人知。每个图书馆侧重于特定的和分享全球范围的服务原则是促使单一图书馆的有效投资和满足新型用户的需求。通过重新使用其他图书馆的服务和维持特定服务,全球范围的服务网络能够出现并满足当地需求。然而,如果商业服务供应商成为该系统的组成部分,那么这种挑战必然导致全球信息基础设施的建设。

当然,当今充分协作共享服务的例子还不是很多。共享信息管理服务的技术基础设施仍然有广泛的需求。这涉及到:

- 分布式开发、提供和部署
- 国际图书馆联盟的运行模式或商业模式
- 图书馆的技术能力已大大增加
- 图书馆员形象必须改变以适应信息管理服务的需要和多用户的需要

总之,如果研究型图书馆进行全球性合作,并拥有必要的技术和组织基础设施,那么图书馆远离研究人员和学生所带来的挑战就能应对。

8. 粮食及农业组织知识交换和建设司 2020 年未来发展宣言书

Anton Mangstl & Johannes Keizer, 粮食及农业组织知识交换和建设司主任

鲁超、尚玮娇 编译

在遥远的过去,图书馆大多局限于一种类型的知识对象:由文字信息组成的文本。最近几年,馆藏已经得到了广泛扩展,例如包括空间信息和多媒体内容。此外,图书馆目录用来处理特定纸制或印刷材料的。馆藏已经变得更加丰富了,同时内容也更加不确定。这意味着一种新型的研究馆员将必须具备一定学科的知识。将来图书馆将管理全球数据库的门户,例如统计数据、生物信息或其他科学数据,尽管它们可在网络中获得。

在未来,知识管理对任何的研究机构来说都将是一个关键问题。图书馆处在促进和推动知识管理以及机构间知识交流的最高位置。传统的出版部门大多只准备发行印刷型资源,这些资源被运到商店、图书馆、进货商或第三方出版商那里。现在,网络是所有科研机构的出版媒介,在将来也是如此。研究机构和大学图书馆应成为组织和传播知识的主体。当图书馆在研究机构中担负机构资料保存的职责时,这种主体的作用已经开始显现,但更多的变化仍然会发生。因为图书馆没有联系到具体某一个研究组或研究部门,在传播机构信息和促进机构间知识共享的过程中,图书馆正处于一个满足一般需求的关键位置。馆员具有信息科学的专业知识。这使得图书馆通过科学的组织可以实现有意义的传播行为。如果在研究性组织中的专家不接受这一挑战,那么未来那种通过机器处理实现的数据网络将永远不会产生。

另一方面,大量的科学资源呈指数增长。十五年前,科学家不仅能够而且要求对其所从事的整个学科领域有一个全面的了解。没有一个完整研究课题的文献综述的出版物是

不可能被接受的。现在就不再是那样了。科学生产的增长和大量新国家融入全球科学社会使它很难始终保持目前的发展状态。从这个意义上讲,使图书馆作为全球的知识共享和知识发现专家将是一个很好的新功能,这类专家知道方法论、技术和世界范围内的技术发展状态。未来的研究型图书馆应该能够为研究者建立其个人的信息渠道。这样做就会使大学和研究机构的图书馆从主要从事收集管理工作的机构转化为帮助科学家和学生获得网络资源的知识交流中心。

未来是否需要研究型图书馆?在今后10年内,纸制的资源仅用于存档。任何检索、查询或处理信息的工作将在全球网络中的数字内容的基础上进行。这可能是可以作为一个全球性的公益事业或有偿服务存在,但每个人都将通过自己的工作场所访问到这些资源。这种类型的访问方式需求现在已经存在了。组织中的人员很少去图书馆。毫无疑问,在获取信息过程中图书馆或图书馆员所担当的角色的作用将会减少。研究人员、技术人员和学生通过使用他们的个人电脑能够处理许多日常的知识检索问题。作为信息提供商更好地组织他们在数字图书馆中资源并使其搜索引擎的功能更强大才是更为重要的事情。

为了实现这一理想,一系列技术上的改变是必要的。未来理想的图书馆员将具备有两个条件:一个主体资格和知识组织或信息科学的资格。他/她将在为实现促进组织中的知识交流而与更多的人一起工作。他/她将有更加自觉的技术意识,并将能够使用Web技术。他/她将不仅对网络信息资源存

在的位置有具体的了解,而且还知道使用不同的工具和技术来访问它们。图书馆仍然是一个物理的地方?答案既是肯定的也是否定

的,这取决于该机构本身。理想中图书馆将物理实体场所转换成知识交流中心,是展览场所、网吧、会议室和机构所在地的统一。

9. 微软研究院 2020 年未来发展宣言书

Natasa Milic - Frayling, 微软研究院系统部主任

鲁超、尚玮娇 编译

为了保持目前的状态并给用户提供有价值的服务,图书馆应急采取了一系列的努力。“变化”是基本的前提。然而,在本文中,我的立场是,图书馆可以通过当前的两方面工作为研究提供很有价值服务:(1) 信息;(2) 跨学科之间的桥梁作用。

研究和科学发现是以大量的实验数据、理论模型和模拟为基础的。研究人员建造复杂和多功能系统来管理生产、摄取、存储过程中的数据以及出版信息和消费数据。这些系统存在并发展了很长的一段时间,并需要认真考虑存储、归档和保存的问题。

为了实现长期的影响,研究组织和社区需要的适当的工具和技能。然而,到目前为止,他们都没有明确界定角色和技能。事实上,对科学数据的新功能来讲似乎缺乏社会的广泛认同和赞赏。因此,没有既定的专业路径使希望从事科学数据信息的个人拥有各种技术。

与此同时,信息的可信性一直被认为是图书馆员工给人类知识带来的最有价值的部分。继续这一传统,图书馆可以帮助建立系统的方法、制定标准,并发展研究信息功能。在和研究型组织或社区进行协作中,图书馆可以使人员概况更加明确地界定,增强学科研究的能力和所需的信息职能。

反过来,这将导致图书馆和研究之间信息和能力的自然流通,并提供广泛的就业机会。长远的利益是通过关系密切的、经过相

关培训的、具有这些知识的人们来增强和确保研究知识。

研究型图书馆在学术工作的框架内有一个独特的地位。由于其工作的性质,图书馆员是知识传播的手段。他们具备良好的知识和经验,可以帮助其他人了解对她们而言崭新的领域。在图书馆员的协助下,学者可以跨过学科之间的界限。

然而,随着研究领域迅速的发展,在人的一生中往往没有足够的机会将各个学科之间建立联系,并从不同的知识领域中产生新的和独特的价值观念。一些专家可能与一个或多个研究领域联系起来,但往往是遗漏关系密切的情况和偶然情况。图书馆可以改变这一点。

在前面探讨的基础上,我设想研究型图书馆是由具有雄厚的专业背景知识的人构成的,并可以为学科之间建立联系。关注建立桥梁的使命,他们可以作出协调一致的努力。他们通过组织学术活动和辩论驱使和促进跨学科思想的交流。研究图书馆可以成为蒸馏、综合和传播研究性知识的催化剂。

我建议拓宽研究型图书馆关注的范围,从服务到参与的积极性和领导力、促进研究信息和跨学科知识的发展。图书馆应该扮演的角色:

- 建立专业形象并培养研究管理的能力
- 制定积极的生态系统,打通研究机构和图书馆的业务渠道。

- 为跨学科研究项目提供新的角色和能力支持
- 实现和推动不同研究社区之间的思想交流
- 承担工业界和政府关于合作的价值的

咨询工作

应采取的行动：为培养研究人员设立项目,建立标准、最佳做法和认证项目,组织会议,辩论,研讨,建立跨学科项目,出版、发起活动,为跨学科的研究人员充当中介作用。

10. 圣地亚哥超级计算中心 2020 年未来发展宣言书

Reagan Moore, 圣地亚哥超级计算中心数据和知识系统部主任

鲁超、尚玮娇 编译

科学研究项目现在依赖于大量的采集的数据分析,这种采集来自模拟输出、观测数据和实验数据。数字图书馆所需要的研究性数字馆藏同样需要关注信息来源、所表达的信息、真实性和完整性。为了科学界的方便使用,这些数据需要整合到数字参考馆藏中。为了建立对数据的信任,必须证明馆藏资源的完整性、一致性和权威性。数字图书馆所面临的一项重大挑战是辅助、管理和维护大规模的数字化馆藏的能力,这些馆藏包括数以亿计的文件。此规模上,不仅仅在辅助程序的自动化上,还在数据管理的管理任务和验证程序上需要对数据进行管理。数字馆藏的管理是一个动态的过程,需要不断监测数字馆藏。第二个挑战是大量馆藏的管理需要分派和复制数字化馆藏,以最大限度地降低数据丢失的风险。集成了以下三种软件技术的技术已经产生:把分布式数据组织成共享资源的数据网格;维护和显示数据的数字图书馆服务;为了管理技术演进而建立保存环境。未来数字图书馆需要提供这三个系统能力。第三个挑战是需要从他们的创造性环境中释放出数字资源。我们需要以数字形式独立标注原始信息结构和信息目前内容的能力。从开放网络论坛的数据格式描述语言到以 Java 为基础的多重解析技术,有多种努力已经尝试过了。

在今后三年内,技术将可自动执行管理政策,最大限度地减少管理大量数据的劳动力数量。但这种方式需要整合引擎与存储系统,管理政策以及规则,并直接强制执行。国家信息被认为是应用微观服务的成果。馆藏情况映射到查询的状态。一个例子是面向规则的集成数据系统(integrated Rule Oriented Data System, i RODS)。这种环境有能力执行保留或处置政策。今后三年的一个合理目标是能够确定政策规定(规则和相关的微观服务),实施管理政策所要求的具体社会环境,然后自动评估所需资源的有关属性。通过使用数据网格的技术,就可以采集到个机构的研究人员共同维护的国际共享的馆藏资源。

建立数字资源的参考馆藏要求科学领域的专家的参与。大多数学科积极地参与试图制订本科学领域表征信息的标准。这包括定义标准语义术语来描述物理意义的数据、标准数据格式、数据结构和标准处理服务。每个学科领域还规定了控制准入、分配、保留和审查程序的管理政策。这一过程可以通过改进表征科学数据的信息而加快。表征信息的编码需要标准的语法;数据解析需要可控的操作手册(微观服务)。从微观服务集成服务过程的能力,将扩大科学资源的使用范围。这反过来又可以体现在直接应用在存储系统中的服务器端工作流程中。

11. 科学基金会办公室 2020 年未来发展宣言书

Lucy Nowell, 科学基金会办公室项目经理

鲁超、尚玮娇 编译

到 2020 年,“数字图书馆”这一提法将是不合时宜的,因为所有的图书馆将都会有数字馆藏。对于研究型图书馆来说,组成数字资源的基本要素将是为科学研究、工程研究以及教育提供基础的数字资源。图书馆被认为是文明基石的文献的储藏地已经有 2000 多年了。馆员长期以来被认为是最有文化的人,对如何查找感兴趣的资料提供指导,以及哪一个文献可被视为值得信赖的,哪个应该对其质疑。

在数字信息时代,公共图书馆普遍为群体提供互联网接入,否则这些群体将被排除在参与全球通信之外。除了提供系统存取,一些图书馆员在数字图书馆方面已成为领导者,越来越多的图书馆学校变成信息或信息科学学校。数字资源带来的好处并不能削减对可信赖的信息资源库的需求,也不能削减对知识指南的需求,因为“信息高速公路”可以使用户南辕北辙。

过去十年情况的报告主要为数字资源的长期保存创造条件。需要强调的是,去年我

们第一次到达数字资源的年产量超过了全球存储能力(参见(<http://www.emc.com/collateral/analyst-reports/diverse-exploding-idc-exec-summary.pdf>))。我们再也无法全部保存我们所创造的一切,确定保留什么再也不能漫不经心了。找到可靠的信息仍然具有挑战性。搜索系统例如谷歌提供大量的信息,通常不考虑数据的可信赖性。此外,搜索是有优先列表,而不是客观的评估内容和价值。

所查找的文献是在实验中产生或从传感器阵列上捕获的数字数据,规模庞大的数据令人困惑。具有地理空间和属性上的数据是不可替代的,例如从地震或环境传感器得到的数据。临床试验数据和其他昂贵的人体对象研究也是几乎不可能复制。但这些数据集有多少可以保存数十年?该做什么来保存那些我们应该保存的数据?在帮助我们知道哪些应该被保护和如何保护它们并同时提供开放式访问这些方面,图书馆和档案馆都可以发挥重要的作用。

12. 世界冰川数据中心 2020 年未来发展宣言书

Mark A. Parsons, 全国冰雪数据中心/世界冰川数据中心项目经理

鲁超、尚玮娇 编译

我从三个相关的观点角度写这篇宣言书。首先,我以世界冰川数据中心的长期经理人的身份来写。从 1957 年开始作为保存国际地球物理文件和照片的图书馆,数据中心已经成为国家存储遥感和其他数字资源的大型数据中心。但我们仍然保持了文档和照

片的积极增长。第二,我是从当前的国际极地年(2007 - 2009 年)的角度来写的。极地年是一个大型的国际化的和跨学科的科学项目,已经运行了两年,4.3 亿美元的资金投入,现有的研究涉及 63 个国家与地区,这些研究包含了物理科学、生命科学和社会科学学科

等领域。因此,极地年为测试地球科学数据的管理提供了机会。最后,我从极地数据和信息服务(IP·IS)的角度来写。极地数据和信息服务是一个全球性的伙伴关系的数据中心、图书馆、档案馆以及网络工作的融合,以确保适当指导极地研究和提供相关数据。但我将把重点放在图书馆的作用上。

IP·IS的主要要求是为科学家提供数据确保基础研究。它力求:①确定哪些数据应该作为IPY的组成部分被收集;②通过相互议定书来提供数据;③全世界的图书馆和数据中心的开放档案信息系统(OAIS)、参考模型(CCSDS 2002)系统一起保存数据。

在识别数据时,我们应该考虑的由国家科学委员会定义的三个基本类别的数字资源界定标准:参考数据、资源或社会数据、研究数据,以及这些不同类别的数据是如何创建不同的政策。参考数据通常是被管理得很好并随时可用的。社会数据高度依赖于社会,但主要问题是使这些数据在原始环境之外仍然可用。研究数据是最多样化的,而且往往最难找到,特别是在多学科的极地研究中。确定数据收集的唯一方式,更不用说确保其获取和保存。这是一个图书馆可以提供优质的服务的领域。通过现有机构间的紧密联系,图书馆可以帮助研究人员描述和上交其数据到公开数据库中。更重要的是,图书馆往往非常适合教育当前以及和下一代科学家对数据进行有效的管理。

一旦数据被确定,他们必须为更广泛的社会服务。图书馆、以档案馆为基础的社区,以及国家和国际数据中心需要进行合作以便随时分享数据和分布式环境中元数据。为了便于数据共享,极地年项目已经制订了基本的元数据形式来提供多个标准。在此基础上,我们已开始通过OAIPMH和XML制订共享元数据的联合目录。下一个步骤是将其扩展到图书馆社会中。我们需要解决某些技术挑战,例如消除冲突的记录;但更重要的挑战

是协调跨不同学科却使用不同的标准、程序、词汇、隐喻和假设的情况。此外,跨学科的数据存取系统正在应付对新的用户群体。我们需要认真考虑对特定的馆藏而言,OAIS参考模型对“指定用户群”有什么要求,因为这反过来确定了许多数据保管和获取的要求。此外,我们必须认识到用户群可以随时间以意料之外的方式变化。人们已经在努力开发更先进的语义和本体,但它是一个年轻的领域,需要更多的图书馆参与才能成熟。

我们还应该考虑到这些用户群体是怎么想的。例如,David Fulkner在北极研讨会上提出的观点表明科学家的两个世界观。一种观点认为,世界是空间领域内的功能的集合体(如地理信息系统的用户);而另一种观点认为,世界是一组参数随时间变化(如气候模型)。虽然David Fulkner强调这是一个过于简化的二分法,这说明了两个基本的数据集成方式(即通过时间或空间一体化)可能适用于不同的情况。在极地年项目中出现的数据共享服务(不以FTP和HTTP为基础)也反映了这种二分法。再次,图书馆和数据中心需要在了解和获取多样性的经验方面进行合作。

最后,最重要的问题是数字资源的长期保存问题。保存可能是最好的理解问题途径,但也是最棘手的。我们知道需要做什么,但我们一直不知道怎么做。OAIS模型提供了一个良好的基线,但详细的实施才刚刚开始。这可能是最重要的和富有成果的图书馆参与的领域。图书馆不仅在数据保存方面有数百年的经验,而且它们也有必要的机构和社会的尊重和支持来提供长期可持续性。一项核心挑战是建立一个可持续的商业模式,使得整个科学界和社会在技术、社会和环境变化快速的时代可以为持续保持独特和重要数据做出贡献。我对于数据资源长期保存的愿景——建立一个核心简单的、可预见的、可靠的、具有可扩

展性、可访问性和持久性的基础设施。但就像与现有的公用设施,如供水、供电、通讯一样,其简单的表面掩盖了深刻的复杂性、结构、规划和专业精神。创造这一水平的基础设施需要协作、维护和专业的发展和认证。我们必须弥合科学学科之间的文化壁垒,例如数据管理人员和研究人员之间,图书馆和数据中心之间。

更大的跨学科的合作和协调的地球系统信息学的时机已经成熟,这是非常明显的,如国际极地年、电子地球物理年(埃及)、地球观

测系统(GEOSS),国际科学理事会的目标是在科学数据管理方面发挥主导作用。特别是,GRL2020可以建立更大的合作和信息共享平台。在eGY的主持下,我和其他人将在美国地球物理联合会秋季会议上提出一暂定题为“地球与空间科学方面的研究型图书馆和数据中心的变革”,开始正式讨论图书馆与数据中心的合作。其他学科可以考虑类似的举措。其他的合作机会包括增加图书馆之间的数据共享、数据中心通过储存机制组织资源,建立正规的数据管理课程和教育计划。

13. 欧洲委员会 2020 年未来发展宣言书

Carlos Morais Pires, 欧盟第七框架计划科学数据基础设施部门主管

吴思竹 编译

10 的 100 万次方数量极大!这是一种用于表示科学信息发生了什么变化的简单方法。通过使用新的方法、实验和测量,数字化技术正在改变着科学和研究的方式。高速的通讯网络将科学家和他们的学校和研究实验室联系在一起,使他们能够通过合作共同面对新的全球性的研究挑战。

研究的成果对于现代社会的经济竞争十分重要。在诸多和我们生存的星球、生态系统、和我们的健康和幸福相关的社会挑战面前,我们需要大量知识和全球的努力。因此欧洲希望加强它的全球合作伙伴的角色并使研究、教育和创新间的联系更加紧密。

知识时代的欧洲路线:

知识三角经常被用于描述研究、教育和创新间的联系。在上升到全球维度后,我们更愿意把这些联系看作是知识路线。一方面,这种路线使我们从科学走到了教育和创新;而另一方面,它促进了为实现共同利益的知识文明的交互合作。正如丝绸之路,促进了最大文明间的贸易交互一样。

科学信息、报纸和实验数据间的连续统一:

在很多科学领域之中,新技术在“数据工厂”的基础上创造出海量信息。数量空前的实验数据的可用性对相关的存储、管理、检索和保存等提出了新问题。它直接影响了研究成果获取的效率和将它们转换为产品和服务的效率。

欧洲已认识到这项挑战。实际上,科学数据领域是在欧洲第七研究发展框架计划中被优先考虑的事情。特别是能力计划框架,它在电子化基础框架领域中处理数字仓储、计算网格和高速网络,为科学和教育团体服务。

实验数据一直被用于限制所谓的 wet - labs 和从成果显示分离出的某些文章。今天已经不同于往常了!这些被用在‘wet - labs’、‘基于报纸’的环境中的原始期刊文章彼此不再被分离。现在,这种趋势正通过数字技术的运用朝着科学信息空间的连续统一的方向发展。

带有一种数字仓储的网络基础设施的实

验数据能同时被不同的科学家以不同方式在不同的实验网站中处理。这为重用‘原始数据’及把这些数据应用到不同科学领域,方法和增加新视角的观点创造了机会。多学科合作是科学信息连续性的另一种表现。

为欧洲科学家、教育家和学生构建最好的基础设施使他们能够利用,将带来科学方法运动,并且将得到的成果以开放的形式讨论,并通过专家同行交换建立的质量控制机制进行补充。

这就是科学信息数字化仓储的电子化基础设施在策略上对于欧洲是非常重要的原因。存储、管理和长期保存实验数据的透明的和成本效益方法能使科学家集中于他们的研究任务。

负责信息社会和媒体的 Viviane Reding 提出的这些要素贯穿于 2010 的政策框架之中,目标在于构建一个科学研究的共同信息空间;加强信息和通信技术研发及其在欧洲的部署;促进一个包容性的信息社会。

但是在信息的数字化未来里有很多同样重要的连续统一性:一种是和存储相关的连续统一,因此存储的信息在过去、现在和将来都可以访问;一种是在不同科学学科间的知识基础间,填补某些研究和教育领域中可能的缺陷的连续统一;还有一种是机构、组织和人之间的连续统一,能够确保对组织变化的全球研究团体机会的形成。

欧洲项目:

现在一些项目正在进一步推动建立科学信息的电子化基础设施。下面是一个简短的(不是很全面的)能力计划启动项目的概览。

IMPAT(生物信息学)将推进 Interpro 和它的软件,它是蛋白质家族、领域和功能型网站的一种综合文献资源,将来自 10 个主要的签名数据库的数据统一为一个单独的资源。

EURO - VO - AIDA(天文学)统一了欧洲天文学的数字化数据集。它将欧洲数据中心集成到全球虚拟天文台,创建了对最终用

户的遵从口令注册的资源和科学工具的协调发展,并在天文学团体中传播了成果。

METAFOR(气候学)目的是提供联合国气候小组第五评估报告技术发展水平的科学数据标准。面临的挑战是如何共享数据促进欧洲和国际的合作,需求元数据不仅囊括了创建数据的气候模型的整个复杂过程,并且满足了从气候科学家到气候影响评估和政策制定者等数据的广大潜在最终用户的需求。

GENESI - DR(地球观测)目的是提供开放地球科学数字仓储,欧洲和全世界的科学用户都可以访问。这将确保对来自空间、空中运动和原位传感器的数据共享。这个项目建立于现有欧洲地球观测基础设施之上并涉及地面部分协调机构中的主要合作伙伴。

NMDB(空间物理学)目的是建立一个宇宙射线数据的欧洲数字化仓储,基于现有数据存储并开发了一个来自很多欧洲中子监测站的带有观测结果的实时的数据库。

DRIVER II(电子化基础设施和信息基础设施)将提出一种泛欧洲基础设施,联邦科学存储。使用开放标准并支持复杂的信息对象。

PARSE. Insight(访问、保存和修护)目的是帮助确定需要保存的基础设施和使用我们社会依赖度越来越高的数字编码信息,然而这种依赖性是非常脆弱的。这些数字化的资源包含了我们为未来后代留下的智慧遗产,这些正是我们现在需要全面的跨领域去探测的。

全世界的研究正在经历一场加强电子化基础设施建设的革命。先进的通信基础设施如世界领先的泛欧洲研究网络 GEANT,网格基础设施促进了科学合作、资源共享及科学数据的存储的新方式,它提供了不断增长的生产率并提高了对地理位置独立研究的合作能力。这些电子化基础设施是全球虚拟研究团体的主要驱动力,因此保证了欧洲的科学卓越性和最终的经济和社会福利。

14. 国际农业研究协商小组 2020 未来发展宣言书

Enrica Maria Porcari, 国际农业研究咨询小组首席信息官

吴思竹 编译

国际农业咨询小组 (CGIAR) 系统由 15 个国际农业研究中心组成,它包含了全系统项目和计划,雇佣了 8500 位科学家和工作人员,每年的总预算约为 5.2 亿美元 (2008 年)。

CGIAR 作为国际公共商品制造者,它的使命是通过研究、建立伙伴关系、能力建设和政策支持来促进食品安全并消除发展中国家的贫困。

CGIAR 中心在这些活动中生产不同的商品。商品包装后成为成果,这些商品是具有代表性的来自研究的新知识和新技术,它们是被提高的技能和能力,有更强有力的体制安排和引导变革的政策依据。

CGIAR 最近重新将自己定位为“发展链研究”——作为国际公共商品 (IPG) 的创建者和提供者,这些公共商品可被其它人用来服务本地、国家和区域的公共和私人利益。这些国际公共商品将在与很多其他行动者和伙伴的合作过程中被生产和投入使用。

这一转变认识到了在更广泛的发展进程中其他行动者的重要作用。这种想法就是 CGIAR 中心应该较少地参加生产本地或国家公共商品。CGIAR 自己直接向农村社区和农民‘传递’成果的行动不像以前那样重要。

对 CGIAR 的信息机构而言,这意味着 CGIAR 的研究成果应该让中介者 (国家研究、推广、非政府组织等) 易于利用,他们会在自己的环境中适应,促进和运用 CGIAR 知识。

第一个国际农业研究中心成立的 40 年来,知识、信息和数据管理就已成为了 CGIAR 的重要活动。

现在,每个中心都有一个图书馆、自己的

网站,具有出版的能力和通讯项目。数以百计的研究小组发表报告和文章,生产和收集通过集成服务获得的科学和空间数据。关于植物的基因数据通过基因数据库全球系统信任保存和标引。图书馆目录通过跨库搜索引擎查询,科学家们可以通过调整期刊和数据库订购 (CGIAR 虚拟库) 来跟踪世界科学。CGIAR 组织的许多的会议和研讨会,使网络便利并支持创新的方法来进行知识共享和机构学习。总的来说它包括了一系列有益的、动态的和有创新性的知识系统。

然而,它所面临的挑战是:

——并不是 CGIAR 研究的所有产品都能以数字化格式获得。目前,尚没有适于未来存储的可靠的研究成果的存储形式。很多旧的资料仍然只能以纸质形式获得。

——虽然有大量正式生产和出版的数字形式和网络形式的报告、书籍和文章,但是其中很多对于我们发展中国家的合作伙伴来说很难访问。它们可能被获得但是却是不易被访问的。

——图书馆的目录不能标引中心的所有成果。一些成果被列在“出版物目录”中。一些在线发布但是却没有被列在任何中心的数据库中。可以通过很多方式检索到一个中心的成果;但是却没有谁能提供成果完整性的保证。

——很多出版物和图书馆目录并不是开放的和可收割的,而且其内容不易为其他服务所获得。

——测度 CGIAR 中心的研究质量的一个指标是外部同行评议的出版物的数量。这个指标的必然趋势是排除了其他的产品 (如

灰色文献),成了鼓励科学家在封闭访问服务中发表文章。

——数据和信息的生产者、CGIAR 中心和合作伙伴,使用标准化方法使得这些数据和信息能被收割、集成和易于访问,而不需要都标记它们的数据和信息。

——图书馆遭受支持综合症;它们的基础作用和功能通常都被低估了。

CGIAR 是农业研究发展信息一致性(CIARD)计划的一个合作伙伴。为了提高可用性,吸收和使用农业知识,计划的目标在于使公共领域的农业研究信息和知识能够为公众所访问。这种想法是在发展和农业应用有效性相关的组织将结合和调整它们的各种努力,向着一些关键的优先考虑的事情和在联合宣言中阐述的关键原则的方向努力工作。

一项投资计划正在制定当中,“确保来自 CGIAR 研究的思想 and 知识能够公共使用、提问和应用,尽可能更广泛的、快速有效的、更实际可行地被利用和访问。”

从 2008 年开始到 2011 年,该计划将支持 CGIAR 研究成果的最大程度的可用性和可访问性的活动。可能采取的行动包括:

——发展可用性、可访问性和应用 CGIAR 研究成果的共同愿景和国际公共商品的传递。

——建立开放和可收割的跨 CGIAR 系统的中心研究成果仓储。这些很有可能合并当前的出版物和图书目录。它显示了一种向着公共访问信息和内容的方向前进的趋势。

——教育和鼓励管理者和科学家开放他们的成果和令其可访问(开放获取),调动了政治性大宗购进和资源。

——数字化旧资料和书籍,使它们能够在网上被利用。

——修改出版运作方法以保证包含在中心机构库中的数据自动存储和信息的出版 workflow。

——实验和扩大在动态参与网络中被嵌入的研究成果和扩大其可见度的社交媒体/web2.0 应用。

——将图书馆的角色从目录和‘订购代理’转换为交流空间和融合(mash up),支持和联系科学家的信息和交流行为,并帮助合作者吸收 CGIAR 的信息和知识。

我们正在不懈地朝着一个开放的、合作的全球研究图书馆而努力。在 2011 年我们可能不会完成所有的最终的计划好的投资计划。然而,朝着这种共同的、共享愿景、并且通过组织的合作工作,我们可以预见到 2020 年可以看到一所令人兴奋、能进行有效地全球研究的“图书馆”。

15. 印度统计学院 2020 年未来发展宣言书

A. R. D. Prasad, 印度统计学院文献研究与培训中心副教授

苏娜 编译

本文试图提供一份 2020 年图书馆发展愿景框架,而不是预测,因为即使是技术预见组织也会经常出现预测失误的情况。这告诫我们无论是愿景还是预测都受当前情境的限制(“存在决定意识”)。随着环境的变化,我

们的希望也会发生变化。随着过去 50 年的巨变,我们的希望也发生了很大变化。

印度图书馆关注的典型问题也是世界各地图书馆所面对的问题。随着对实验室资助的减少,对图书馆的资助也开始减少。所有

学院的领导都声称图书馆是学院内最重要的机构,但是一旦遇到预算危机,图书馆预算将会是第一个被削减的。其他相关问题还包括电力、低带宽(尽管手机成本已经大幅降低,互联网访问成本依然很昂贵,即便是对机构来说也是如此)。

进行预测是很难的,但是我们仍然可以从一些纷繁的现象中识别当前的一些问题,期望有一个更好的未来。毫无疑问,全球不平等是存在的,委婉地来说就是不发达国家、第三世界国家、发展中国家。在科学研究和实验室、大学和学院的图书馆中也存在着不平等。尽管离实现平等还有很远的距离,但是至少可以期望缩小不平等之间的差距。

现代民主政治的历史表明,一个需要强调的重要问题是普遍教育。然而,要实现这一理念,至少在很大程度上需要实现跨政治边界,即国家的信息普遍获取。互联网就是实现这种愿望的工具之一,它促进了不同形式(如研究论文、指南、课件等)、不同多媒体格式的信息的出版与发行。互联网为发布信息提供了机会,同时也为获取信息带来了更多的挑战。过去,由于纸本空间的限制,研究者不得限制他们的出版物的大小。同行评议的价值毋庸置疑,但是它仍存在缺陷,有时会因为一些新观点与现有的观点相悖而被过滤掉。互联网克服了这些限制,同时也带来了更多的挑战。

对图书馆带来的挑战有:

1)如何测度网络所提供的信息的数量以及图书馆为用户提供的信息的数量。(即传统意义上的信息资源)

2)如何在海量信息中识别相关信息。(即传统意义上的参考咨询服务)

3)如何描述网络上的信息对象。(即传统意义上的编目。所面临的挑战是如何自动生成元数据,因为网络上大量的信息是没有元数据的)

4)如何发现信息对象之间的关系。(即

传统意义上的分类)

5)著作权法是获取信息的巨大障碍。除非著作权法更多地是面向作者和用户,而不是以出版商利益为中心,否则信息不能够从法律问题中得以脱身。

令人振奋的是,一些传统的图书馆信息处理工具开始针对网络信息,无论是将目录转变为元数据,或是将图书馆分类转变为网络本体。图书馆和情报学专家对高效网络(语义网)的开发做出了很大的贡献。

目前,我们使用 OAI - PMH 和 Z39.50 协议收割元数据,但是问题仍然是网络上有多少信息适用于这些协议。我们面临的挑战是如何利用本体/自由分类法实现元数据的自动生成和分类。毫无疑问,我们需要更有效的信息挖掘技术,需要语义网中高效的开源软件、开放链接、信息过滤的智能代理以及好的搜索引擎。

信息开放获取、开源软件、开放课程和开放标准(我更愿意称为“开放颂歌”)的进程将会开创一个更加公平的新纪元。

不平等与许多其他问题纠缠在一起,如技术问题、基础设施问题、法律问题、经济问题和政治问题(尽管并非无法完全挽回)。尽管其中的部分问题已经不是图书馆力所能及,但它们却是图书馆美好未来的前提条件——动力(电力)问题、简陋的实验室设备、低带宽等等。尽管开放课程为获取信息提供了便利,但实验室设备却是另外的一个问题。或许,基于网络的模拟实验室环境(simlabs)在一定程度上能解决这一问题。

经常被问到的一个问题是关于图书馆和情报学专家的未来的问题。只要人类存在,就会有不断增加的信息,就会有信息需求,我们就需要图书馆和情报学专家处理海量的信息。但是他们的工作形式和方法可能会改变。(经典的“形式与内容”问题)

我们期待信息的解放,它最终将会促进研究、研究图书馆和人类的普遍解放。

16. 欧洲学术出版与学术资源联盟 2020 年未来发展宣言书

David C. Prosser, 欧洲学术出版与学术资源联盟主任

苏 娜 编译

学术文献利用的谈判不应成为研究人员和图书馆员的主要关注内容。研究人员需要高效地跟踪同行的研究工作,如果我们对他们访问的内容有所限制的话,那么对他们和整个社会来说都是一种伤害。开放获取消除了在线研究与相关读者之间的经费障碍,为研究人员提供了保障,使他们可以利用到除本机构可提供文献之外的所有文献。

对图书馆而言,传统的研究期刊的采购工作将变得越来越不重要。随着越来越多的文献可以开放获取,图书馆竞争的核心不再是馆藏,而是为研究人员所提供的服务。两种需要同时开展的服务会变得越来越:传统的“进口”服务,将机构外部的内容介绍给本机构的研究人员;“出口”服务,将本机构的研究成果输出到尽可能广泛的学术领域。展示机构知识财富的机构知识库的广泛应用,正是说明了这种变化。知识库通常以图书馆为基础,并由其负责管理,使研究论文、学位论文、工作报告、数据集等可以为更多的人获取。

一旦开放获取得以广泛实施成为学术交流的标准模式,那么就可以展开一系列相关研究。越来越复杂的软件将用于文献挖掘,扫描数以万计的论文以发现论文原作者没有意识到的关联。基于复杂软件的文献挖掘工作在跨学科和多学科领域会非常重要,因为在这些领域中,一个领域的研究人员往往对本领域之外的研究进展知之甚少。研究论文、数据集、维基、博客等将被链接到一起。研究人员可以通过这些链接跟踪学术思想的发展,不会因死链或者资金障碍问题而发生中断。研究论文本身也会由目前的静态对象

发展成为一种有动态的、实时的实体,作者可以利用更多的数据、新的参考文献等不断更新论文。显然这需要解决与版本识别相关的技术和实践问题。

许多能够支撑这种开放研究环境的基础设施已经在现实中出现了。在世界范围内,正在建设越来越多的基于机构和学科的知识库(已经有 1000 多个学术知识库在 openDOAR 注册,<http://www.opendoar.org>),但是这些知识库中资源的增长速度却不快。造成这种情况的原因有多种,但最主要的就是将研究论文存储到机构库中目前还尚未成为研究人员日常工作的一部分,尽管大部分研究人员认为研究论文的存储工作是一件好事,而且也不需要太长的时间(每篇论文所需要的时间还不到 10 分钟)。目前,为应对知识库的低存储率问题,有许多研究资助团体开始使用强制存储政策。

Wellome Trust 是最早实施开放获取政策的资助机构之一。2006 年,Wellome Trust 就将开放获取作为其资助条件之一,要求受资助者要将在同行评价期刊上发表的研究论文原稿的副本存储到开放获取知识库中。紧随其后,英国研究理事会(UK Research Council)也提出了类似的资助条件。在美国国会的推动下,近期美国国家卫生研究院(NIH)也制定了开放获取强制存储政策。NIH 是世界上最大的非军事研究项目的资助机构,年预算超过了 280 亿美元。NIH 的强制存储政策每年使近 8 万篇生物医学研究论文实现开放获取。

在国家、国际范围内,支持开放获取政策的研究资助机构越来越多。因为这些研究资

助机构意识到“如果信息不能被社会广泛、方便地获取,那么他们传播知识的任务也仅完成了一半”。此外,个别机构也开始利用开放获取政策促进本机构研究的广泛传播,近期的典范便是美国的哈佛大学。

要想实现真正的开放研究环境,还需要采取进一步的措施:

- 所有相关人员都可以免费获取公共资助的研究成果。所有研究资助团体要以 Wellcome Trust、NIH 为榜样,将开放获取作为资助的条件,要求研究人员将他们最终经过同行评议的论文副本存储到开放获取的、可互操作的知识库中。

- 采取进一步措施保证公共资助研究的数据可以为公众获取。

- 重新审视现有与研究信息相关的知识产权制度。现在知识产权、版权立法中的限

制没有必要完全以研究人员利益最大化为中心。或许我们需要一种新的学术交流模型,一种既能保护权利所有人的权利,又能保持研究人员与用户访问研究成果之间的权益平衡的模型。

- 继续建立互操作标准和知识库基础设施。

- 制定相关政策,保证数字对象的长期保存。

将强制政策、卓越的开放获取平台以及资源(大部分资源是为图书馆所有,并负责管理的)结合到一起,将有助于新学术交流环境的建设。在这种新的学术交流环境中,所有人都可以访问到公共资助研究的成果,我们可以将世界所有的智慧整合到一起解决目前所面对的棘手的研究问题。

17. 联合信息系统委员会(英国)2020年未来发展宣言书

Malcolm Read,高级进修教育基金委员会行政秘书

鲁宁 编译

对于图书馆和接受过高等教育的图书馆员来说,一个重要的新机会正从机构、共享和学科知识库的发展中出现。

起初这里只是作为一个存储研究成果的地方,主要是一些学术期刊论文,作为开放获取发展趋势的一部分。但是,随着知识库可以支持和服务更大范围的信息资源,一个更有雄心的远景目标正在形成。特别是,研究数据的保存和长期保存是一个日益突出的问题,知识库也可能是教育资源的一个自然平台。

许多研究学科早已发现管理和保存研究数据的必要性,即社会科学纵向数据集,例如人口调查数据和需要花高价收集并且不可能被复制的科学数据,如气候数据。然而许多

研究数据在项目尾声时被抛弃。许多情况下,也许是大多数情况下,这些数据对以后的研究者有很少或甚至没有价值,但有一种可贵的资源在很多情况下被丢失。一项新的研究常常开始于文献调查但很少进行“数据调查”。数据以及其分析很少利于其他领域的研究者(尽管逐渐认识到发现需要链接数据)。

几乎没有一个大学拥有充足的合适的信息管理技能的职员,或者合理的结构来充分建立和开发知识库。他们需要审视和获取有合适技能的职员,自然首先看他们现有的人员。图书馆服务是一个显而易见的资源,但是图书馆和图书馆员是否能够应付这种挑战呢?当然很多人都会的。有一个令人印象深

刻的数据,各个资历水平的图书馆员是数字图书馆革命(如果它还可以称作一次革命的话)前沿的领军人物。但是这批激情绽放的专家群体已经到达足够合理的管理机构仓储的高度了吗?

组织必须考虑创建和管理知识库的许多问题:

- 政策:谁拥有数据?你如何选择哪些数据值得保存?保存多久?哪种资源应该进行开放获取?需要什么样的质量标准以及如何管理它们?

- 财政:它在现在和将来需要花费多少钱?资本和流动资金如何安排?知识库有什么好处以及对谁有好处?

- 管理:一个知识库应该是由机构内部管理、外包,还是与其他机构共享服务?应该使用什么元数据标准?如何处理著作权许可和数字权利管理?需要什么样的服务和服务质量?管理知识库是否需要必要的专业技能?

- 激励:组织应该收集什么研究数据?如何鼓励研究者和老师将资源存放到知识库中?这种存储如何吸引认可和如何奖赏?

- 互操作:个人知识库如何集成进来?我们如何从分散的机构和学科知识库构建一个国家的虚拟知识库?这些国家知识库如何进行国际连接,为学术内容提供一个无缝的

基础设施(欧洲委员会的结构路线图和 DRIVER 项目对此明显很有帮助)?研究知识库应该与其他教育和文化知识库以及相关领域如健康进行何种程度的连接?

这并不意味着大学或研究所将转向图书馆服务和图书馆员解决这些问题:尽管很多都是。并且不可能假设,图书馆专业人员将像以前那样在大学里享有同等高的地位。已经有许多研究机构资助和管理的研究数据库从图书馆中独立出来。大学管理者可能将知识库视为本质上就是计算机支持和 IT 架构的一部分,也可能对学科馆员而非一般信息管理技能有强烈需求。事实上所有的技能都是必须的,但不是所有的大学都享有集成的服务支持,图书馆专业面临相应的新技能需求的挑战,令人信服的是高级管理是适合这个工作的人选。当然在英国,我们发现研究程序(尤其是在网格或 e - science 社区),学习技术人员、管理部门和图书馆都来处理数据管理和长期保存问题,但彼此隔离。大学图书馆和图书馆员在知识库的管理中应发挥关键作用。但是不能假设这个就会发生。我怀疑,GRL2020 的参加者将是否同意知识库对研究型大学的图书馆员是一个真正的机会。但是是否所有的专业同事和机构管理层同意这个观点?你如何确保在这个日渐成长的领域发挥重要的作用?

18. 米尼奥大学(葡萄牙)2020 年未来发展宣言书

Eloy Rodrigues,米尼奥大学文献服务部主任

鲁宁 编译

在研究型大学的背景下,用户正越来越期待有一个信息丰富的环境,在这个环境中,所有需要的信息在合适的地点和时间被自由获取。信息资源也同样被期待能够在一个单独、简单的用户接口——“à la Google”中寻

找和搜索到。在葡萄牙的经验中,图书馆的设备仍然受到重视和大量使用(主要是被学生),但是在图书馆物理的业务量(借阅、参考咨询、馆际互借和文献传递)在过去的八年已经很明显的下降了。与之相反,在线的数字

服务(包括联机公共检索目录)和信息资源的使用,以及远程用户的服务请求,正在明显的上升。

另一方面,对于提供开放获取自己的出版物(在少数情况下,对研究数据)的重要性的认识在研究者中正在上升。尽管在最近几年我们做出很大努力去满足用户,例如,建设大型的数字信息资源馆藏(通过数字化馆藏,“大宗采购”和从出版商的大量采集);建立集成、跨库搜索门户网站的各种资源,对偏远地区的用户提供虚拟参考服务,建立和维护机构知识库,等等;但是图书馆还远远没有能够为研究人员提供他们所期望的服务和信息环境。

除了“传统”的中介作用和促进获取和利用外部信息来源的角色,图书馆还应该不断的关注对于它们服务的社区成员所提供的信息和数字的附加价值。在研究过程中建设、管理和维护数据库中各种形式和各个阶段产生的信息(发表研究文献、预印本和灰色文献、研究数据),提升他们的透明度和可获得性,并保证他们的保存和管理,这些都将是未来研究型图书馆最重要的任务。

在接下来的十年里,当地的研究图书馆可以而且必须成为全球研究信息提供者的网络节点。这个全球网络将由不同类型的库组成,这些库能够为他们所附属的网络社区储存、保存、检索和提供研究成果。

分布式机构知识库的全球网络为建立增值服务上的基础设施提供了可能。其中一部

分服务可以在图书馆环境下被创建和管理,而另一些则会在不同的环境中产生。但是,无论是哪种情况,这些新式的服务都是完全自动化的,并使用最好的计算工具和技术,例如,扫描、挖掘、分类、关联、统计、分级、打包和复用研究成果(出版物和数据)。这样,通过揭示研究人员以及他们的研究机构间“隐形的”和暗含关系,那些增值服务不仅会加强科学交流的效率,揭示现有信息和数据的潜在价值,也有利于个体和机构的网络协作。

为了实现这一全球网络化愿景,支持增值服务的出现,主要有2个条件:1)研究成果必须开放获取;2)必须开发使用新一代的协议、标准、工具和平台。因此,研究型图书馆必须为实现这两个目标现在就作出努力。首先,通过保证他们机构库的可获得性,帮助他们的研究人员使用机构库,倡导研究作品的开放获取。其次,支持来自不同来源(研究资助者、大学和其他研究机构)正越来越多、越来越频繁的自存储要求。同时,图书馆也应该与最近兴起的开放数据运动合作。

研究图书馆必须积极参加到当前的制订新协议、标准、工具(如开放文库计划、开放数据协议等)的活动和项目中,加强分布式机构库之间的互联和互操作性。在这方面,欧洲DRIVER是特别重要和相关的。根据DRIVER的设计,欧洲未来机构库联合和它与在世界的其他地方(美国、中国、巴西等等)正在或将要开展的相似项目的互联,将组成未来全球科研信息和数据网络的雏形。

19. 罗马第三大学国家核物理研究所 2020 年发展规划

Federico Ruggieri, 罗马第三大学国家核物理研究所主任

王欣 编译

高能物理界于九十年代后期开始参与网格计算的研究,希望网格能够解决大型强子

对撞机的计算问题,而这一方法开始评价并不太高。在那时,客户端服务器和元计算成

为了前沿,第一个实现了计算机农场的工具出现了。然而,其最大的问题是有大量的需要分析和处理的数据(每年数十千兆)。

所谓的社会化挑战就是允许来自不同国家和地区的数千万科学家们方便简单地存取这些数据。很显然,即使考虑到计算能力进化的摩尔法则,欧洲粒子物理研究所传统的专门用于计算机资源的预算也是大大不足的。现在没有显而易见的解决方法,这样一个全世界的事业也需要新方法来解决。类似与面向对象的编程和面向对象的数据库等几种“新”方法也被提出,一些研究和开发项目也致力于解决这个问题。高能物理计算从另一方面来说也是一个典型的高吞吐量的计算,他允许一个非常简单或“自然”的并行,基于复制的应用程序和基于事件的数据结构。

基本方法是分散大型强子对撞机的计算负载到各个实验室,而欧洲粒子物理研究所是数据源和这些数据的主贮存池。这个模型是由 MONARC 项目提出的,模型定义了几个层或者阶,欧洲例子物理研究室作为 0 阶,而其他地区中心作为 1 阶,更低级的作为第三阶,而网格天生就能满足这些需求。

不管怎样,大量的数据处理和分析对于大型强子对撞的网格计算和目录的应用是一种挑战,而复本和数据库标签是现成的解决办法。由于在研究人员之间存在的竞争,在整个实验期间(至少数十年),数据的存取根据数千个研究人员的身份严格控制。

在过去的几年,研究领域内外的国际合作的情况已经与迅速的与大量带宽网络的部署相结合。许多的高级服务和应用基于这些网络,开启了远程合作的新方法。网络和资源整合的环境,例如数字图书馆、计算、存储、相关的系统和工具成为了众所周知的电子信息基础设施(e - infrastructure)。在发达国家中,知识现在是经济进步和繁荣的主要元素,因而电子信息基础设施(e - infrastructure)在知识经济战略部署中是主要的促进者。另一

个方面,这也会促使发展中国家和发达国家之间的数字鸿沟不断扩大。在发达国家,知识成为商品,而且国家和公司分配在教育 and 研究开发上的预算也占预算的相当大的部分。

乍一看,发展中国家有着比建设电子信息基础设施更必要的花费项目。然而,认识到电子信息基础设施在打破潜在的循环中所扮演的角色很重要。科学在生产活动中是长期创新的基础,反过来数字基础设施允许科学家能够参与最前沿的科学活动,与全世界的同行们分享技术和经验,它是最方便的工具和方法。另一方面,在地中海地区有几个很好的例子,如埃及的亚历山大图书馆。此外,几个创新项目打算部署新的重要的研究基础设施(如约旦的 SESAME),这将会制造需要共享的大量数据和知识。中东和非洲国家的电子数据和资源库的增长将会提供大量的有关不同文化、语言和历史的独一无二的知识。

将来的电子信息基础设施应该包括几个重要的基础,在这之上应该建立一个统一的架构:战略管理,用以维护具有长期性的重要知识。应该包括一个区分可被保留和可被丢弃的数据和知识的策略。在这个早就具有高度的移动性和全世界互联的社会,数字图书馆和知识库必须有无处不在的使用接口。我们可以设想出许多为普通大众提供的新服务,而网络的连通性不再是限制因素。虚拟社区将会在全球化中扮演重要角色,研究者、兴趣小组、甚至合作中的公司需要一个个性化的接口来访问大量的知识和数据。与此相关,其也将会成为获取能够满足不同用户需求的高水平服务的接口。新的软硬件技术将会很有帮助,但是解决方案的标准化还需要很多努力。达成这个目标并不需要建立标准,而是采用一系列好的折衷措施。

要达成这一版本的目标必须采取几个步骤。e - Infrastructures 与欧洲和世界上其他

地方提供的类似服务的稳定的互操作问题是尤其重要的。探索领域并为基于高层次服务的数字图书馆架构提出可能解决方案的专门

项目。而非洲、中东和亚洲的网络社区所包含的内容将会带来不同的文化和创新想法。

20. 威斯康星大学麦迪逊分校 2020 年发展规划

Eldon Ulrich, 威斯康星大学麦迪逊分校生物磁共振数据库主任

王欣 编译

生物磁共振数据库 (BMRB : www.bmrbl.wisc.edu) 是一种主要用于存放蛋白、多肽、核酸、小分子生物学的原子水平核磁共振光谱(核磁共振)数据的设施。其中用来描述分子研究的元数据、研究所使用的实验条件和设备、与研究相关的重要出版物都被留档存储。这些数据为所研究的系统提供了化学、动力学、热力学的三维视点。大量的未经加工的时域数据、收集到的核磁共振光谱仪以及各种教育材料和各种网络资源的链接也可以从中获得。一些研究人员使用档案中的独立条目来与他们正在研究的生物系统进行对比研究,或者把他们的工作引导到与其密切相关的系统中去。在这些项目的早期,我们需要理解拥有类似特性的研究系统的技术,确定什么人已经调查研究了这样的系统,锁定这样的数据和从过去得出结论。在项目的后期,我们需要能定量的比较本地数据与存放在 BMRB 和其他公共存储的数据的工具。这些工具可能包括文件格式转换和数据可视化系统。将近发布之时(或者项目开始之时)研究者需要把他们获取的知识整合到一个宽泛的环境中。具体的目标是建立一个相关出版物的目录作为出版的参考文献。这需要对相关分子甚至看上去不太相关却拥有相似的功能或相似的动力学、热力学特性的分子做文献调研来获取信息。基因本体工程提供了一个可控的词表,这使得可以有效的检索具有相似功能和分子的基因产物。能够

跨库检索定位分子相似序列或三维构同源的工具有也是必须的。产生网络或功能同源序列的分子的应用已经存在,但是通常不对公众开放。基于网络的使作者和分子研究出版物相关联的工具以及文本的关键词搜索技术现在已经可以获得。然而分子名称的不一致和其他的含混不清会导致很多不相关的点击。更有条理的利用数据库来存储出版物信息以及所提及的检索方法将会为生物检索提供已纳入出版文献的更加精炼结果。

那些经常利用可从 BMRB 获得未经加工的数据和分配频谱参数并经常实验新方法的研究者们推动了生物核磁共振领域的发展。这包括自动核磁数据分析协议,由指定化学位移数据计算三维结构的新方法。所有这些努力包括从公共数据库中精心挑选数据构造高质量数据集。另外,能够发现并修改公共文档中的异常错误来构建既有数量又有质量的本地数据库的方法也被开发出来。这就提出了一个问题:一个公共档案是否应当尽量提供数据的对外标识,提供过滤或修改后的版本。生物核磁共振的协作计算项目和 BMRB NMR - STAR 的数据模型与国际化联原子命名规则一致,为生物分子核磁共振数据提供统一的集合和档案标准。这个标准的应用便于科学家把出版数据整合到他们的工作流中,大大提高了科学家开发新的软件应用的能力。

21. 欧洲核物理研究组织 2020 年未来发展宣言书

Jens Vigen, 欧洲核物理研究组织馆长

孙敏杰 编译

预印本作为高能物理领域(High - Energy Physics, HEP)研究者进行学术交流的主要渠道已有五十年的历史。有效的检索和获取该类信息是他们研究工作的核心组成之一。2007 年针对高能物理领域学者使用哪种信息源去获取所需信息进行了一次调查。2000 多份的答卷显示,活跃的高能物理学术团体中约有十分之一被提及。这表明学术团体主办的信息源在该领域占主体地位,商业服务仅能服务于一小部分用户。另外,高能物理领域学者似乎针对不同的信息需求使用不同的工具,这显然是重要的。调查结果阐明了在未来五年高能物理领域学者的信息需求。

多年来,高能物理领域的科学情报一直处于科技前沿。arXiv 预印本文库是该领域的生命线,早在 1991 年,互联网诞生之前,就提供联机服务。随着互联网的发展,高能物理领域建成了第一个包含预印本、文章与书籍的联机数据库。在我们的业余生活中机会是有的,比如在搜索框中输入歌词的片段并且几秒钟后就能听到播放的一首歌曲,但诸如此类的事情在我们的职业生活中是不可能的,只是输入几个选定的词组,研究人员无法找到相关文章。理想中服务的例子是:在搜索引擎中输入一个数字,就能识别出它出自哪篇文章。

目前有几个高能物理领域的科学情报系统。我们的目标是在该领域构造一个超级科学情报系统,这个系统能够满足多方的各种需求,正如当前的多个系统所能提供的服务一样。这种超级系统是实现数字科研(e - science)的电子化研究基础设施(e - infrastructure)的典型例子,它满足了当前的需求。

超级系统提供如下服务:

——能够保存高能物理领域所有开放获取文献的存储系统;

——能够提供新的全文检索和数据挖掘应用的数据库;

——能够对各个国家和机构的科学产出进行评价的系统,该评价系统可以作为费用均摊开放获取出版的基础。

——全面的、免费提供的、高能物理领域出版物的引文索引。

该超级系统能够为整个高能物理领域的科研团体提供 e - science 服务,已经超越了当前信息系统所能提供的服务:

——通过对以往读者的阅读记录、引证分析及作者网络的综合研究,能够为对某篇文章感兴趣的读者自动筛选其符合其兴趣的所有文章;

——通过全文挖掘和引证分析,能够对某篇文章的主题进行自动侦测;

——根据引证分析和合作者关系网络,能够对同行评议的文章进行自动筛选;

——引文指标系统,该系统旨在标记有影响、有声望并且热门的文献,该系统可以作为期刊影响因子的补充,影响因子是目前用于评价机构和个人科研生产力的指标,它根据已经出版的期刊引证情况进行评价;

——可以获取用于计算的大规模数据。

虽然其他学科有一段时间以来一直在探索挖掘其出版物中科研数据的方法,高能物理领域才刚刚开始建立用于检查如何保存和传播数据的模型,但两者都是为了教育目的,并且允许任何研究人员发挥其潜能。这种滞后的主要原因是高度复杂的数据集和技术限

制。

我们的目标是开展一项研究,解决从早期大型强子对撞机实验以来的数据以一种可用格式进行保存的问题。该研究将集中于几类关键用户,比如:数据保存的信息技术专家、数据传播的图书馆员、数据加密的物理学家、以及研究数据可用性的教育与通讯团体。

2020年,我们设想高能物理领域的所有科学成果,包括同行评议数据和出版物,都能以开放获取的方式提供利用。数据和出版物将会被保存在由社会经营的基础设施中,该基础设施由国际图书馆协会和与该领域关键实验有关的计算机科学家来运营。

22. Outsell 公司 2020 年未来发展宣言书

Leigh Watson Healy, Outsell 公司首席分析师

孙敏杰 编译

今天,图书馆正处于十字路口。它们服务于各种各样的用户,这些用户具有不同的需求、偏好、能力和习惯。今天的图书馆领导者正面临着做出正确选择的挑战,在最大限度发挥图书馆职能影响力的同时,要把有限的资源提供给谁。我们看到,信息产业处于富有创造性和活力的时期,其产业扩张是非线性的。对于那些要看到可能性的人来说,在创造、兴趣和扩大视野方面有着充足的能量。

对于优先图书馆投资、确定未来的发展方向和图书馆哲学来说,今天的信息用户比以往任何时候都能作为这些问题的根本驱动力。满足用户需求和期望是任何一个图书馆的基本目标,并且了解用户是实现这一目标的基本要求。用户的信息——他们的需求、习惯、行为以及喜好,随着时间的推移正在发生着变化,往往是以可预见的方式能够预期并为变更做出安排。此外,对现在的学生信息习惯的测试能够提供真正的证据表明图书馆对将来专业用户的预期。

虽然图书馆用户群在各个环节都有着无数的变化,作为他们对图书馆的一些最重要的影响以及他们将如何建设和适应未来,Outsell 公司已经确定了以下趋势:

- 用户独立性日益显著。自己查找信息是一个快速发展的偏好。当用户获得独立,他们希望图书馆扮演的角色正在发生着变化。

- 关于信息结构和信息组织方面,够用就可以了。用户更喜欢用 google 的方式查询他们所需信息,而不是在分类网站上找寻,或在局域网“滚动、指向并点击”。图书馆可能过于组织其提供的内容,或者在组织方式上没能适应用户需求。

- 点对点的信息交流越来越多,反映了用户忽视传统出版模式,支持更为直接和互动性更好的模式的意愿。对比那些被编辑过并且同行评议过的内容来说,未经过滤、未经编辑的内容被认为是限制性较小的、更前沿的、更为纯粹的。与同行直接沟通的意愿通常要重于对传统信息处理过程中的控制或质量的期望。并且在点对点的网络中,参与者互相了解,比起传统信息源来说,这样的信息可能被认为更为权威和前沿。

- 内容的购买者和检索者希望能在更小的相关数据集中获得信息。相对于非消费内容将来可能发生的情况来说,在线音乐业务提供了一个极好的类比。那就是,用户要买歌曲而不是买专辑,并且他们正在把这种偏

好带到学术生活和工作当中。用户希望他们所需的内容没有被捆绑,并且他们不希望花费额外的时间和资源去筛选非相关内容。

• 对方便性和环境的期望将增大。用户希望在需要的时候才拥有内容权力,比如在适合其工作的时间和地点。如果说内容是国王,那么环境就是女王。用户希望信息以一种对特殊任务敏感的方式呈现出来。

• 当面对着互联网、局域网、实体图书馆以及收集信息的其他选择时,用户的信息选择做得越来越好。用户每天都接收新的图片、有趣的技术和用于精确传递所需内容的小工具,这就创造了一种可能,商业、教育和研究方面的内容可以同样的方式提供给用户。

Outsell 公司已经深入研究了以上每个趋势。我们认为图书馆和信息专业人员能够积极的影响或者利用促进这些趋势,并且能够提高用户效率和用户满意度。

重要的是要注意图书馆如何回应。提高用户独立性意味着图书馆员正在建立他们的辅助研究角色和“看门人”的角色。图书馆员通过将信息直接整合到最终用户的工作和应用当中,提高用户效率和方便性。图书馆员

正在花费越来越多的时间和资源去促进用户进行点对点的信息交换,并用更加结构化的信息整合非格式化的信息。用户同时也成为信息内容的创建者,他们已经接受信息调配和互动、自由的信息共享机制。

最后,信息用户正在寻找最省力途径,并且忠诚于最终所取得的方法。图书馆员必须尽可能的理解他们的目标市场,就像设计信息产品和服务一样。

创造和再造图书馆角色,涉及新的合作伙伴、跟踪用户工作流、利用技术整合信息内容、打破传递模式的一些规则。在界定图书馆 2.0 的时候,已经呈现了清晰的特点,即图书馆实质上是要提供给用户想要的内容、用户希望采取的方式及时间:

- 以用户为中心
- 开放
- 在任何人都需要的地方
- 互动与协作
- 载体丰富
- 灵活的

对图书馆来说,新的方向看起来显著不同于传统图书馆,富有创造力和创新的激动人心的时刻就在眼前。

23. 美国国家农业图书馆 2020 年未来发展宣言书

Peter Young, 美国国家农业图书馆馆长

张红丽 编译

国家农业图书馆(NAL)发起了各相关机构联合起来协同发展国家农业数字图书馆(NDLA)的倡议,NDLA 是未来全球农业研究图书馆的雏形。NDLA 将应对来自端对端农业数据管理(从信息采集、共享、整合和管理到自然和生命科学中数据的处理、来源追踪和维护)的挑战,解决了农业相关知识在创造和使用过程中所产生的问题,提高了科学发

展应对全球健康、环境和食品安全、供应等一系列问题的能力。

通过共享全球服务网络、资源和专业化知识,NDLA 采用全球开放获取的方式提供多机构的、跨部门的、各种形式的与农业科学相关的信息内容、数据和服务。NDLA 鼓励研究者从事科学研究创新,利用全球数字化网络的潜力,改变他们在教育、研究和知识创造中

的参与能力、发现能力和学习能力。这些新一代数字系统、工具和服务的发展为研究人员、生产者和从业人员对大批科学数据的组织、管理和保存提供了便利。

NDLA 旨在建立一个开放式、可扩展、全国性的科学和工程数据收集系统。NAL 的计划中确认基于 NAL 框架在公立大学系统中建立并共享一个开放式、互操作分布联合数据储存库网络是很有必要的。该 NDLA 系统将长期提供丰富的农业信息数据集、分析工具和数据可视化工具等,来支持面向 21 世纪高等教育的学习、探索和研究活动。

未来科学研究和学习需要一个信息基础设施提供适当的服务、工具和功能,把跨学科环境中的数字化信息和数据转化为科学研究和发现过程。这点尤其适合于生物(动物和植物科学)、化学、自然、环境、工程、科学和社会学等与农业科学相关的学科(如农业经济学),因为这些学科在很多研究领域(例如:基因组学、基因测序、功能基因组学、水棚读数、生物燃料和可再生能源、水产养殖、土地和水使用的历史数据,以及地理/地理信息系统信息等)在数据访问、存储和操作等方面是相通的。

NDLA 提出的出发点是,没有一个单一的图书馆、研究机构或集团(无论是机构级别的或国家级别的)可以实现诸多交叉农业学科领域数字数据和知识内容的全方位获取、管理和维护。农业科学中的知识发现和生产力研究等工作日益需要 NDLA 即将提供的资源和服务。NAL 计划的提出源于过去几十年中在信息技术领域的领导地位。通过发展 AG-

RICOLA 数据库,建立一个农业网络信息中心联盟(AgNIC),并通过成功引入 USDA 的数字桌面图书馆(Digital Desktop Library ,DigiTop)和 AgSpace 数字图书馆,在农业信息服务中高效利用数字信息技术方面,NAL 已取得了重大进展。

创立于 1995 - 1996 年的 AgNIC 合作联盟包括多个具有合作关系的机构,方便了农业、食品等方面优质信息资源的获取,NDLA 概念的提出符合 NLA 在 AgNIC 中的领导地位。通过运用目前最先进的网络信息技术,AgNIC 提供相当快的速度实现对大量经过评估的农业信息和教育资源的访问。目前,AgNIC 已经拥有 60 多个机构成员,他们合作实现涉及农业相关主题的网站链接。联盟成立以来已经成功地应用了几代网络技术,建立了(农业的)元数据和叙词标准,并于近期通过门户网站为国家农业系统提供服务。由于信息技术及应用、科学研究发展、农业进展及美国人口压力、经济压力等各方面所产生的综合影响,NDLA 应运而生。随着十多年信息环境的发展和动态变化,图书馆一直在尝试新内容、新技术和新服务以适应用户不断变化的需求。而 NDLA 正体现了应对这些变化的适应力,试图利用大量的新机会迎接农业领域里的重大挑战。

编译自:GRL2020Event Guide and Position Papers. http://www.grl2020.net/uploads/position_papers/Event%20Guide%20&%20Position%20Papers%20Final.pdf. [2009 - 2 - 28]

(初景利 校)