

存档资源键研究*

曲云鹏^{1,2,3}

(1. 中国科学院文献情报中心, 北京 100409; 2. 中国科学院大学, 北京 100049;
3. 国家图书馆, 北京 100080)

摘要: 存档资源键 (ARK) 是一种面向数字资源长期保存的资源永久标识符系统, 由图书馆等公益保存机构开发, 申请与使用简单, 并且可以免费使用, 适合有长期保存需求的保存机构。ARK可以通过同一个标识符的不同参数指向资源的三种信息, 即资源本身、资源的元数据和提供者有关声明。ARK通过它的限定符系统为资源提供方的多种服务提供了最大的支持, 也为用户访问资源带来了方便。ARK已经在欧美的学术机构图书馆中获得了较多的应用, 值得进行研究。

关键词: 永久标识符; ARK; 长期保存

中图分类号: G250

DOI: 10.3772/j.issn.1673—2286.2014.12.005

1 前言

为了建立高质量的信息标识符命名体系, 实现对信息对象永久性的有效访问, 美国国家医学图书馆委托加州数字图书馆 (California Digital Library, CDL) 设计了存档资源键 (Archival Resource Key, ARK), 并于2003年完成。ARK是一个开放的、注重实效的、低费用的资源永久性标识解决方案, 它建立在URL、Handle System、DOI、OpenURL等成果之上, 对各类资源提供了广泛支持。ARK所支持的资源对象既包括数字文档、数据库、软件、网站以及物理对象如图书、杂志等, 也包括无形对象, 如化学物质、疾病、词汇、演出等^[1]。

ARK是一种框架, 也指遵循这种框架的唯一标识符, 可以实现对信息对象的高质量 and 持续性的标识。永久持续性是ARK服务的关键的部分, ARK框架使用灵活,

以冗余的方式最大程度上保证了资源的长期可访问性。

2 ARK的特点

2.1 ARK的优点

ARK不同于其他的唯一标识符系统, 它是图书馆行业为了保证数字资源的长期保存可用性而开发的框架, 具有以下优点:

(1) 适合图书馆提供面向长期保存的数字资源服务

ARK系统在图书馆环境中开发, 可以满足图书馆中数字存档工作人员的要求, 并针对图书馆等保存机构的资源长期保存需求, 设计了冗余的资源访问方式, 明确了组织承诺 (organisation commitment) 对于永久标识符框架的重要性。这些承诺可以帮助用户来判断资源的保存情况变化, 以保证用户有最大的可能获取到资源。

* 本研究得到国家社科基金后期资助项目“数字资源长期保存的技术研究与实践” (编号: 09FTQ005) 资助。

(2) 免费

ARK是图书馆开发的永久标识符系统,可以免费使用,降低了保存机构提供所存档资源访问服务时的成本。

(3) 简单

任何个人和机构都可以设立自己的ARK唯一标识符解析服务,同时ARK唯一标识符系统对于硬件的要求较低,而且对于数字资源的访问只需要网络服务器的“redirect”和“get”方式就可以使用。

(4) 良好的兼容性

ARK框架是一个基于标准的框架,独立于协议和技术,具有很好的兼容性。ARK可以用于封闭的环境,例如暗存档(dark archive),也可以用于开放的公众可以访问的环境。

(5) 功能多样

ARK系统中,通过ARK限定符的不同后缀,可以实现对于数据对象、元数据和机构声明等内容的访问,具有强大的功能。

(6) 灵活性

在ARK框架中,如果一个机构同时作为命名映射授权主机(NMAH, naming mapping authority hostport)和命名分配授权机构(NAA, naming assigning authority)进行工作,该机构可以对自己的标识符框架拥有完整的控制权。而且,ARK部分是独立于HTTP协议部分的,可以在不损失任何核心表示信息的情况下从一个服务器移植到另一个服务器上。

(7) 开放性

ARK标识符系统目前还在不断的发展当中,可以自由使用,欢迎各类型机构的加入。

(8) 透明性

ARK的后缀特点可以提供保存机构的更多信息,为用户进行判断提供基础。

2.2 同其他主要标识符系统的区别

(1) 同DOI的区别

ARK同知名的唯一标识符规范DOI相比,有着以下的区别:

①ARK来自于公共文化机构,比如图书馆、档案馆等等。DOI来自于出版商和电子商务团体,经常被用于识别在线的文档和出版物。②DOI标识符的生成受制于对于每个标识符的收费体系,而ARK是免费的,主要由公共文化机构使用,其持续性指的是相当长的一段

时间。③在ARK框架下,每个命名分配授权都有很大的自主权,加州数字图书馆只负责维护这种规范和命名分配授权注册,每个ARK的命名分配授权机构都可以免费定义自己的策略和服务。DOI是商业化的,是一种由出版商和在线数据提供商构成的非集中式的模型,每个DOI机构在操作层面上都有着不同的技术选择和服务提供方式。

(2) 同URL的区别

ARK和URL的区别主要体现在对于资源的永久访问上。URL是通过HTTP或者HTTPS协议对在线资源提供访问的一串字符串。通过URL,用户可以快速地对资源进行访问,但是在线资源经常会失效,因此URL也无法保证数字资源访问的长期可用性。而在ARK框架中,每个资源都会分配一个唯一的永久标识符。如果一个网站或者主机发生变化,ARK中只有命名映射授权主机部分会发生变化,唯一标识符还保留。一个使用ARK的机构,在定义了策略之后,就需要承担维护资源和标识符之间的联系的责任。如果一个资源已经被移除,机构需要给出这个资源的简单描述,和被移除的原因。ARK强制机构将资源的长期访问作为首要问题进行考虑,这是ARK同URL最根本的区别。

3 ARK的详细内容

3.1 ARK的结构与功能

3.1.1 ARK的基本结构

ARK标识符是一个包含着“ark:”标记的字符串序列,通用形式是http://NMAH/ark:/NAAN/Name。图1是一个ARK标识资源的例子:

第一部分是命名映射授权主机(NMAH, Naming Mapping Authority Hostport),在本例中是“ark.cdlib.org”。第二部分是ARK协议标识,即“ARK:”。第三部分是命名分配授权号(NAAN, naming assigning authority number),这是一个用来标识命名分配授权机构的通用唯一数字,在本例中是“13030”。第四部分是由命名分配授权(NAA, naming assigning authority)机构分配的名字,本例中是“ft4w10060w”。在ARK中,NAAN之前的部分并不用来标识或者比较对象,它表示的是一个主机名,用户可以通过这个主机访问该资源,也可以通过别的主机访问该资源,主机名只是

用来保证资源可用。NMAH部分是临时的，可以任意使用，也可以替换。当目前的命名映射授权主机失效时，新的服务提供者会替代，重新生成NMAH。这样，授权主机出现问题的时候，ARK的核心标识会很容易恢复，将由ark:/开头的ARK部分独立出来即可^[2-3]。

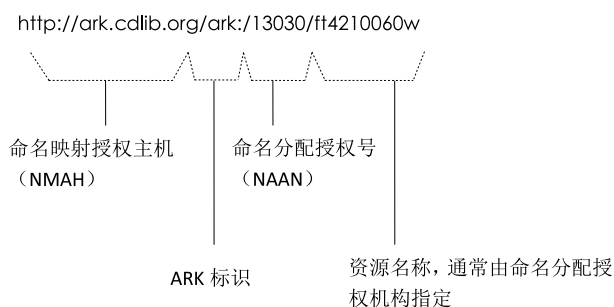


图1 ARK实例

NAA是首次给该资源命名的机构。NAAN的注册方式同URN的命名空间很相似，但是NAAN的命名只是5位或者9位的数字。前100000 (1-99999) 注册的机构可以分配到5位数字的号码，之后的机构号码就是9位的。这种规则可以减少NAAN同临近的数字之间的混淆^[4]。

3.1.2 ARK的限定符

ARK的限定符是一个强大的工具，可以通过限定符识别组件或者资源的不同版本，是ARK区别于其他永久标识符的一个明显特点。ARK的限定符是在由NAA命名之后的部分，这是一个可选的限定符部分。它将基础的ARK进行了扩展，可以指向对象内部，比如提供对于对象层级组件或者子部分的访问，或者提供对资源的各种变量，例如版本、语言、格式等信息的访问。ARK限定符的使用首先体现在通过ARK对资源的元数据和资源提供者的资源状态承诺进行访问的方式，即在ARK标识符后加“?”和“??”^[5]。

限定符可以由NAA或者任何NMA来进行设定。ARK的命名映射授权机构对于自己的ARK访问方式可以进行比较全面的定义，对于如何访问资源的一部分也有着灵活的方法。

在形式上，限定符是由组件路径 (ComponentPath)、变量路径 (VariantPath) 以及变量路径后跟随组件路径形成的。变量路径由字符“.”来隔开，变量路径由字符“/”来隔开。见如下一个例子：

http://example.org/ark:/12025/654xz321/s3/f8.05v.tiff

在本例中，字符串“/s3/f8”是组件路径，“05v.tiff”是变量路径。

使用ARK限定符可以通过在ARK标识符后增加“/”来揭示对象之间的层次关系。如以下ARK标识符：

ark:/12025/654/321

这是一个标识了一个对象的ARK标识符，事实上，它同以下三个ARK是一致的。

ark:/12025/654/xz/321

ark:/12025/654/xz/

ark:/12025/654

在ARK规范中，在名字后出现的“/”，表示它左侧的部分被包含在右侧的部分。在这个例子中，三个资源访问的都是命名为654的资源，而654资源包含在xz资源中，xz资源包含在321资源中。对于“/”本身来讲，它的左右都必须都是有效的、非结构标识的字符，在解释ARK的时候，两个连续的“/”就会被化简成为一个。

标识资源变量的ARK。在限定符的使用中，可以利用“.”来分割对象的各种变量。如果在ARK中包含了“.”，那么它左侧的部分是一个基本命名，右侧的部分一直到结尾，或者下一个“.”之间，都是一个前缀。一个资源命名中可能包含多个前缀。例如：

ark:/12025/654.24

ark:/12025/sz4/654.24

ark:/12025/654.20v.78g.f55

在前缀命名中有两个主要的规则。第一，如果两个ARK共享了同样的基本命名，但是有着不同的前缀，相应的对象被认为互为变量。第二，一个带有前缀的ARK就意味着一定有一个没有前缀的变量可以用ARK来表示。

对于变量的揭示完全由命名分配者决定。变量完全可以揭示对象的命名、多个变量的命名和相互之间的关系等等。

3.1.3 ARK的元数据和维护记录获取

ARK是同协议无关的，在HTTP和DNS环境下都可以很好地工作。很多机构都使用最常用的HTTP和DNS网络方式来进行ARK标识符的解析。

当用户通过ARK标识符访问一个资源时，如果ARK中的命名映射授权主机是有效的，用户可以被直接

引领到该有效的命名映射授权主机中。如果这个命名映射授权主机已经失效了,或者临时不可用,用户可以通过识别命名分配授权或使用加州数字图书馆维护的命名映射授权主机注册表来寻找当前正在服务的命名映射授权主机,从而定位最新的命名映射授权主机。

ARK框架同样还定义了一个简单的使用HTTP协议来传递ARK的协议,即THUMP(微型HTTP URL映射协议, Tiny HTTP URL Mapping Protocol)^[6]。这个协议允许用户在浏览器的地址栏直接输入ARK请求。只要用户知道相关命名映射授权主机的Internet主机的名称和端口号,用户就可以通过THUMP向这个主机发送请求。THUMP请求是一个以“?”结尾的字符串,其中包含了ARK请求。在THUMP的响应结果中,ARK可以解析定向到对象或者对象元数据。结果中还有一个承诺协议,包括了对对象的永久性、变化(比如对象变化的环境,比如格式迁移)以及变化历史等,这些结果会以电子资源引用符(Electronic Resource Citation, ERC)的规范返回。ERC是一种对象描述规范,使用了都柏林核心元数据。它主要使用了其中的四个核心元素来描述资源,即 who、what、when、where 在每一项中都记载了该对象相关的一些信息^[7]。

以图2为例,在URL中输入http://digital.library.unt.edu/ark%3A/67531/metadc28359/? ,这是请求NAAN为67531的机构中,名为metadc28359的资源的元数据。其响应结果如下:

```
erc:
who: Phillips, Mark Edward
what: Using Archival Resource Keys (ARKs) for Persistent Identification
when: 2008
where: http://digital.library.unt.edu3A/67531/metadc28359/
```

图2 THUMP元数据响应结果

如果输入的是http://digital.library.unt.edu/ark%3A/67531/metadc28359/??,这是请求该资源的变化情况。其响应结果如图3所示。

```
erc:
who: Phillips, Mark Edward
what: Using Archival Resource Keys (ARKs) for Persistent Identification
when: 2008
where: http://digital.library.unt.edu3A/67531/metadc28359/
erc-support:
who: University of North Texas Libraries
what: Permanent: Stable Content:
when: 20081203
where: http://digital.library.unt.edu/ark:/67531/
```

图3 THUMP文件历史响应结果

3.2 ARK的申请、使用与维护

3.2.1 ARK的申请与使用

以ARK作为唯一标识符的机构必须成为ARK框架中的命名分配授权机构,或者被授权成为一个命名分配授权机构的子授权。每个命名分配授权都通过一个或者更多的命名映射授权主机相关联。命名映射授权主机是提供服务的实体,负责提供对那些标识过的数字对象的访问、转发等等。具体的过程如下^[8]:

(1) 定义ARK命名的结构。首先,使用机构需要申请成为一个分配授权机构,获取一个分配授权号,并明确自己的ARK命名策略,即明确使用ARK来标识资源的类型,例如书目记录、物理文档、抽象的概念等。如果需要使用ARK来对应多种类型的资源,那么应该使用预定义的前缀来识别每种特殊的子集。

(2) 使用软件来生成ARK标识符。机构可以使用自己开发的工具,也可以使用开源软件NOID(Nice Opaque Identifier)^[9]来进行这项工作。开源软件NOID由加州数字图书馆设计开发,用来生成唯一标识符,在配置正确的情况下可以生成ARK标识符^[10]。另外,也可以使用加州数字图书馆提供的一种新的服务——EZID^[11]来生成和维护ARK标识符。该服务通过在线接口访问,也可以通过API访问,来自动化生成大量的标识符^[12]。

(3) 使用软件来解析ARK标识符,即需要使用一些软件将带有ARK名称的Web地址同相应的资源联系起来。该步骤中,首先要定义利用哪个主机来解析ARK标识符,即解析服务器。然后需要定义一系列的参数来让用户可以获取资源的整体或一部分,比如一页,或者针对这个资源的一些服务,例如文档的某一个版本、书目记录格式等等。例如,法国图书馆利用gallica.bnf.fr来识别数字资源标识符, catalogue.bnf.fr来识别书目记录;在ARK中使用了/fn来请求数字文档中的某一页, n就是页码。简单说,这个步骤需要你选择一个NMAH,来解析ARK,以及一个参数来获取特殊类型的服务^[13]。

3.2.2 ARK的维护

ARK框架由当前的加州数字图书馆进行维护,该机构的命名分配授权注册表中列出了命名分配授权号和对应的NMAH主机。这个注册表同样由加州数字图书馆维护,在美国国家医学图书馆有一个镜像。

注册命名分配授权表单给出了一个ARK用户群体的标识。加入该框架的成本很低，没有费用的要求。任何机构都可以通过联系加州数字图书馆获取一个命名分配授权号^[14]。

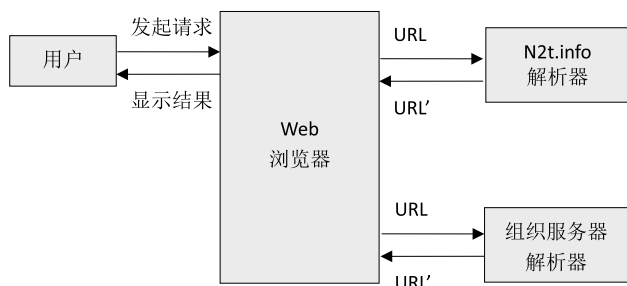


图4 N2T工作流程

在NMAH主机部分中，主机名一般可以持续几十年。当命名映射授权主机失效的时候，就需要重新寻找可以用的NMAH主机来访问资源。ARK规范中提供了三种方式来寻找NMAH。

①当需要寻找一个新的NMAH时，客户端会在ARK中抽取NAAN，然后在一个全球数据库中对NAAN进行查询，寻找所有的为该NAAN提供ARK服务的NMAH。目前这个服务器定位为N2t.info。这是一个低成本高可靠性，由类似联盟形式进行维护的NMAH，可以提供可用的基于HTTP的URL。如果希望这个寻找的过程是自动的，则需要使用一个Name-to-Thing (N2T) 解释器^[15]来进行自动化操作。

②在全球可访问的文件中查找NMAH，即在一个命名授权表中寻找一个可用的NMAH。这个表保存在加州数字图书馆，地址是<http://www.cdlib.org/inside/diglib/ark/natab>。任何机构可以从Internet上复制这个表。表中以评论行的方式(#开头的行)来解释资源的格式、文件的修改时间、地址以及NAA注册的指导。

③通过DNS寻找NMAH。这是一种比较陈旧的方法来解决寻找NMAH问题，需要基于URN解析器来进行。该方法在RFC2915^[16]中进行了描述。这种方法需要查询DNS系统。首先一个请求需要发送到DNS，请求一个可以匹配给定NAAN的解析器的列表。DNS会将这个请求分发到特殊的DNS服务器中，来提供答案，如果本地DNS缓存了这些信息，那就可以略过这个步骤。相应信息会以命名授权指针(name authority pointer)记录的形式返回，一般的结果就是一个或者多个候选的NMAH列表。

4 ARK的应用情况

目前，已经有大约150多个组织使用ARK作为唯一标识符系统^[17]。其中，较为知名的应用机构是CDL。CDL采用ARK作为它自己拥有以及它所控制数字对象的标识，采用METS描述SIP、DIP包。CDL在ARK服务方面作出如下声明：①保证CDL分配的ARK具有永久唯一性，已经发布的ARK不再重新分配；②CDL分配的ARK不包括广泛的可识别语义信息；③CDL分配的ARK包括一个校验字符，以便及时发现传输中造成的错误。CDL采用开放源码软件“NOID”(nice opaque identifiers)来进行ARK分配管理，NOID除了生成ARK标识符外，它还可以实现元数据信息与标识符的绑定，并指定建立一个URL访问接口和名称解析器。目前，CDL已经分配了80,000个ARK。在加州使用ARK的用户包括加州大学的10个校区图书馆、加州州立大学图书馆以及加州公共图书馆等^[18]。

法国图书馆使用ARK用于所有的书目信息、所有的数字化后的图书，以及这些书的图片、脚本和图等。同时ARK还被用户指向法国图书馆长期保存系统SPAR中的AIP信息包的标识，包括了其他的数字资源类型，例如ARC文件、原生的音视频和软件资源、存档记录等^[19]。法国图书馆还计划用于在线展览。由于在RDF中资源和属性必须使用URI来进行命名，法国国家图书馆使用了ARK的URI框架来识别它的数字资源和元数据记录。它的开源性和非专有性的本质和由公共机构来进行维护的特点让ARK成为了一个在数字保存环境中使用的理想的框架。ARK用于识别真实存在的对象很实用，因为它可以指向文档的一部分或者指定特殊的视图。所以ARK是法国国家图书馆命名真实的AIP或者其中的一部分并在RDF中进行描述的方法^[20]。

北德州大学图书馆使用了ARK作为所有数字对象的识别框架。除了正常访问对象、元数据和承诺的方法外，UNT图书馆还增加了一个功能，即在ARK后增加/thumbnaill/来查看预览。另外UNT图书馆使用一个工具来处理带有固定前缀的序列式标识符的生成，名字是number-server^[21]。

宾州图书馆使用ARK inspired noids作为唯一标识符，在这个scholarsphere仓储服务中，其中有亮点的是基于Rubynoid库的微服务。这些标识符用于fedora系统中，由ScholarSphere来解析^[22]。

5 总结

资源存档键ARK在制定规范时首先将资源的长期保存放在第一位考虑,对于以长期保存工作为第一任务的保存机构有着重要的意义;ARK兼容多种协议,可以利用简单的Web和HTTP方式对资源进行访问,体现了它的便利性;ARK的限定符提供了对同一资源的不同部分进行访问的功能,使保存资源的管理和访问更加容易;ARK还提供了XML方式的元数据等信息提供方式,这样元数据对于人更加易读,促进数据在不同的信息系统中的共享。ARK的这四个特点符合唯一标识符规范的发展方向,即保证资源的永久性、规范的易用性和资源的可读性。

我国对于唯一标识符系统的研究不多,主要集中在针对DOI的研究,也缺少唯一标识符的具体应用。ARK标识符充分考虑了作为非营利性公共服务的图书馆使用DOI的不便,以及资源常长期保存的需求,在免费的框架下提供了功能不弱于DOI的唯一标识系统。我国众多的图书馆和档案馆有强烈的数字资源长期保存服务需求,但是目前我国仍然缺少一个国内通用的唯一标识符规范。ARK体系为图书馆的资源标识与揭示提供了一种新的思路,值得进一步的研究与探索。

参考文献

- [1] STARR J. ARK (Archival Resource Key) Identifiers [EB/OL]. (2013-03-22) [2013-12-17]. <http://www.cdlib.org/inside/diglib/ark/>.
- [2] KUNZE J. Overview of Archival Resource Key (ARK) Tools [EB/OL]. (2005-07-01) [2013-12-17] http://dot.ucop.edu/home/jak/ark_tools.ppt.
- [3] ARK Anatomy [EB/OL]. [2013-12-17]. <https://wiki.ucop.edu/display/Curation/ARK+Anatomy>.
- [4] NAAN: Name Assigning Authority Number [EB/OL]. [2013-12-17]. <https://wiki.ucop.edu/display/Curation/NAANs>.
- [5] HAKALA J. Persistent identifiers - an overview [EB/OL] (2010-10-13) [2013-12-17]. <http://metadaten-twr.org/2010/10/13/persistent-identifiers-an-overview/>.
- [6] GAMIEL K, KUNZE J, NASSAR N. THUMP -- The HTTP URL Mapping Protocol [EB/OL]. (2007-08-24) [2013-12-17]. <https://confluence.ucop.edu/download/attachments/1802247/thumpspec.pdf>
- [7] ERC: Electronic Resource Citation and Dublin Core Kernel Metadata [EB/OL]. (2013-09-06) [2013-12-17]. <https://confluence.ucop.edu/>

<display/Curation/ERC>.

- [8] KUNZE J. The ARK Identifier Scheme [EB/OL]. (2008-05-22) [2013-12-17]. <https://wiki.ucop.edu/download/attachments/16744455/arkspec.pdf?version=1&modificationDate=1261036800000>.
- [9] KUNZE J. Noid: Nice Opaque Identifier (Minter and Name Resolver) [EB/OL]. (2012-09-04) [2013-12-17]. <https://wiki.ucop.edu/display/Curation/NOID>.
- [10] Noid-nice Opaque Identifier Generator Commands [EB/OL]. [2013-12-17] <https://confluence.ucop.edu/download/attachments/16744482/noid.pdf>.
- [11] STARR J. EZID [EB/OL]. (2013-04-30) [2013-12-17]. <http://www.cdlib.org/services/uc3/ezid/>.
- [12] Generating ARKs [EB/OL]. (2013-04-30) [2013-12-17]. <https://wiki.ucop.edu/display/Curation/Generating+ARKs>.
- [13] ARK (Archival Resource Key) [EB/OL]. [2013-12-17]. http://www.bnf.fr/en/professionals/issn_isbn_other_identifiers/a.ark_en.html
- [14] CDL Name Assignment and Support Policy Statements [EB/OL]. [2013-12-17]. <https://wiki.ucop.edu/display/Curation/CDL+Policies>.
- [15] KUNZE J. Low-Risk Persistent Identification: the "Entity" (N2T) Resolver [EB/OL]. (2006-10-10) [2013-12-17]. http://www.ipresconference.org/ipres06/presentations/John_Kunze-CDL.pdf.
- [16] MEALLING M. The Naming Authority Pointer (NAPTR) DNS Resource Record [EB/OL]. (2000-09) [2014-08-23]. <http://www.ietf.org/rfc/rfc2915.txt>.
- [17] KUNZE J. Registered Name Assigning Authority Numbers [EB/OL]. (2013-11-15) [2013-12-17]. http://www.cdlib.org/uc3/naan_table.html.
- [18] KUNZE J. University Library Experience - CDL Case Study [EB/OL]. (2005-06-30) [2013-12-17]. http://dot.ucop.edu/home/jak/cdl_case_study.ppt.
- [19] BERMES E. Persistent Identifiers for Digital Resources: The experience of the National Library of France [EB/OL]. (2006-12-30) [2013-12-17]. <http://www.ifla.org/VI/4/news/ipnn40.pdf>.
- [20] FAUDUET L, PEYRARD S. A data-first preservation strategy: data management in SPAR [C]// Proceedings of International Conference on Preservation of Digital Objects, 2010.
- [21] PHILLIPS M E. Using Archival Resource Keys (ARKs) for Persistent Identification [EB/OL]. (2008-06-05) [2014-10-02]. <http://digital.library.unt.edu/ark:/67531/metadc28359/>.
- [22] CLAIR K. Metadata for a Micro-services-based Digital Curation System [EB/OL]. [2013-12-17]. <http://dcevents.dublincore.org/index.php/IntConf/dc-2011/paper/viewFile/17/8>.

作者简介

曲云鹏，男，1980年生，国家图书馆数字资源部馆员，研究方向：网络存档、长期保存、自然语言处理，E-mail: quyp@nlc.gov.cn。

A Research on Archival Resources Key

QU YunPeng^{1,2,3}

(1. Library, Chinese Academy of Sciences, Beijing 100049, China; 2. University of Chinese Academy of Science, Beijing 100049, China; 3. National Library of China, Beijing 100081, China)

Abstract: ARK (Archival Resources Key) is a kind of persistent identifier for long-term preservation, developed by preserving institutions like libraries. It is suitable for the institutions with long-term preservation needs and can be used for free. What is different from other persistent identifiers is that ARK can point to three kinds of information of a resource: the resource itself, the metadata and the statements of the resource owners. The qualifier system gives the resource provider the most efficient support for the multiple kinds of services, easy for the users to access resources. ARK has been widely used in European and American research institute libraries and deserves more researches on it.

Keywords: Persistent identifier; ARK; Digital preservation

(收稿日期: 2014-12-01)

第五届全国知识组织与知识链接学术交流会在京召开

随着信息技术的快速进步，数字化信息资源向着两极发展：一方面，各种类型的数字资源之间聚合程度不断增强；另一方面，数字资源越来越注重内容单元的分析与呈现。在新的环境下，知识组织如何实施和发展，知识链接又该如何组织和服务？这需要不断地思考和探索。为此，中国科学技术信息研究所、国家科技图书文献中心（NSTL）和中国科学技术情报学会连续主办“全国知识组织与知识链接学术交流会”，第五届会议于2014年12月5日在北京中国科学技术信息研究所召开，共有来自全国100余所机构的300余位专家和学者参加了会议。会议共安排13个主题报告，相关信息及报告PPT已共享，欢迎访问会议网站（<http://168.160.16.186/conference/>）查看。

报告题目	发言人	发言人单位
基于关联数据的科技知识发现——以科技报告为例	夏立新	华中师范大学
复杂网络及其在情报学中的应用	王继民	北京大学信息管理系
创新性情报研究的新思考	刘细文	中国科学院文献情报中心
引文分析与引文推荐相关研究进展	万小军	北京大学计算机科学技术研究所
新时期我国科技文献服务面临的机遇和挑战	梁冰	中国科学技术信息研究所
大数据背景下的金融信息组织与服务	何华	汤森路透金融&风险产品设计部
面向医学领域的嵌入式知识服务研究	牟冬梅	吉林大学信息管理学院
中文知识组织系统形式化语义描述标准体系研究	曾新红	深圳大学图书馆
DDC与CLC先组式类目映射分析	贾君枝	山西大学管理学院
学术文献的结构抽取和功能识别	陆伟	武汉大学信息管理学院
国家叙词库框架结构设计	吴雯娜	中国科学技术信息研究所
医学知识组织系统建设与应用实践	李军莲	中国医科院医学信息研究所
农业专业领域知识资源组织与知识服务探索	赵瑞雪	中国农科院农业信息研究所