

语义 Web 模式下综合科技资源的关联策略研究*

王思丽 马建霞 祝忠明 张秀秀 马建玲

(中国科学院国家科学图书馆兰州分馆 兰州 730000)

【摘要】在开发综合科技资源集成登记系统的实践基础上,研究综合科技资源的关联策略,针对其在语义化方面的局限性,提出一种利用 D2RQ 组件将综合科技资源的关系数据源公开为 RDF 接口,进而利用 SPARQL 端点查询技术将 RDF 接口公开为 Web 服务的优化方案。试验证明,该方案合理可行,可为综合数字资源体系建设的未来发展提供思路。

【关键词】语义 Web 综合科技资源 关联策略 RDF D2RQ SPARQL

【分类号】G250.76

Study on the Linked Method of the Integrated Resources in Semantic Web Patterns

Wang Sili Ma Jianxia Zhu Zhongming Zhang Xiuxiu Ma Jianling

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

【Abstract】Based on the practices for developing integrated resources and services metadata registry, the paper studies on the linked method the integrated resources, and proposes an optimized method using D2RQ to expose relational database of integrated resources as RDF interface, then uses SPARQL points to query the RDF interface and exposes it as Web service. In fact, the method can provide a good way for the future development of constructing integrated digital resources system with its rationality and feasibility.

【Keywords】Semantic Web Integrated resources Linked method RDF D2RQ SPARQL

1 引言

以数据为中心,将语义置于数据中,是语义 Web 模式的核心视角所在^[1]。它得到了国内外数字图书馆界的广泛关注和认可,并陆续建设了一些示范性试验系统。当前,具有代表性的资源登记服务系统主要有澳大利亚的 ANDS(Australian National Data Service)^[2]、英国的 IESR(Information Environment Service Registry)^[3]、法国的 Michael(Multilingual Inventory of Cultural Heritage in Europe)项目^[4]、美国的 OCKHAM-DLSR(Ockham Digital Library Services Registry)^[5]以及我国的 OCSR(Open Collection and Service Registry)^[6]等。这些系统的共同点是都支持各类元数据的登记、发布、检索和浏览,并提供了相关轻量级协议的访问接口,在一定程度上方便了资源与服务的发布和管理。但从实际的应用和推广程度来看,这些系统大多还停留在对资源进行二维描述其基本属性关系的阶段,并

收稿日期:2011-04-15

收修改稿日期:2011-05-16

* 本文系中国科学院知识创新工程、中国科学院国家科学图书馆资助项目“综合科技资源集成登记系统”(项目编号:y000021002)和中国科学院西部之光 2009 基金项目“甘肃省综合科技资源登记示范系统建设”的研究成果之一。

没有进一步深度挖掘各类资源间的关联关系,无有效持久的工作流管理机制^[7]。因而,在语义 Web 对资源知识化和集成化组织的更高要求之下,存在着许多亟待解决的难题。综合科技资源集成登记服务系统(Integrated Resources and Services Registry, IRSR),是中国科学院国家科学图书馆“十二五”战略规划布局中综合数字资源体系建设的先导性项目,其预期目标是构建集中式资源集合元数据登记系统,为综合科技资源的语义化标注和管理提供可靠平台和技术支撑,为实现各类资源基于语义化的自动发现和关联扩展奠定基础。目前,该系统已能够支持对多种类型元数据的登记注册和管理,同时与第三方系统进行互操作的重要接口,如 Web Service、OAI-PMH、SRU 等,也正在持续开发和完善中。

本文在开发 IRSR 的实践基础上,从语义 Web 架构模式视角出发,结合当前数字资源知识化组织建设和服务的应用需求,通过分析 IRSR 的数据模型和相关的实体关系,重点研究综合科技资源的关联策略,并针对其局限性给出了实际的优化方案,从而为综合数字资源体系建设的未来发展提供思路。

2 IRSR 的关联策略分析

2.1 IRSR 的核心数据模型

IRSR 的数据模型继承了 DLSR(Digital Library Service Registry)^[8]的标准化服务描述和共享使用方案,重用并扩展定制了 DC、IESR、OCSR 等元数据标准,形成了多维度元数据应用规范和编码体系,如表 1 所示:

表 1 IRSR 基本编码体系

前缀	编码体系	URI
dc	The Dublin Core Metadata Element Set, v1.1	http://purl.org/dc/elements/1.1/
dcterms	Dublin Core Terms	http://purl.org/dc/terms/
dcmitype	Dublin Core Type Vocabulary	http://purl.org/dc/dcmitype/
cld	Collection Description Terms	http://purl.org/cld/terms/
cldtype	Collection Type Vocabulary Terms	http://purl.org/cld/cldtype/
rslpcd	RSLP Collection Description	http://purl.org/rslp/terms#
iesr	Information Environment Service Registry	http://iesr.ac.uk/terms/#
cdls	中国数字图书馆标准规范	http://cdls.nstl.gov.cn/basic/terms/
ocsr	Open Collection and Service Registry	http://ocsr.las.ac.cn/index/

该模型可抽象为由资源(Collection)、机构(Org)、服务(Service)、关系(Relations)4个核心实体元素集共同构成的元数据描述框架,如图 1 所示:

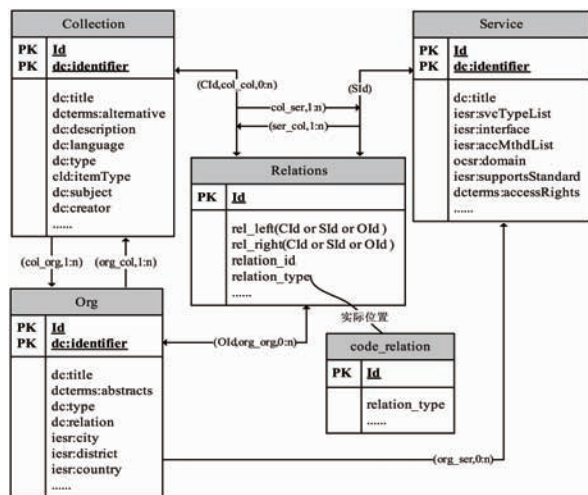


图 1 IRSR 核心数据模型

(1) 资源是指综合科技集合类资源,除了包含一些资源的基本描述信息如资源题名、交替题名、语种等之外,还包含一些版权声明、访问权限等使用信息。具体来说,共有 8 种主体资源集合类型:软件、资源集、文本集合、数据集、事件、图像/多媒体、交互资源、服务,与其相关联的资源内容类型有 65 种,每一种资源实体都被归入一种或多种相应的学科主题类型,采用中国图书馆分类法,目前共分为 16 个一级学科和 147 个二级学科,同资源内容类型一样可根据实际需求进行二次扩展。

(2) 资源、机构、服务三类实体既可单独存在,又可作为彼此的检索点,可视为一个可重用的功能组件。其中资源实体是三类实体的核心,每一种资源都给出了其对应的基本服务参数,如服务类型、服务接口地址、访问控制方式及支持协议等。IRSR 中涉及的服务类型主要有 9 种:检索服务、咨询服务、文献传递/馆际互借服务、接口服务、个性化服务、委托服务、元数据提供/收割服务、订阅服务、保存服务。同一类服务可以支持多种访问方式和协议。可根据资源实体浏览相关服务,也可根据服务的基本参数查询和使用相关服务。

(3) 机构实体同样不仅有自己的基本属性,如名称、机构 URL、所在地区、国家,可以被单独进行检索和浏览外,最重要的一个特性是作为资源、服务的管理者角色而存在。IRSR 中划分的机构类型有 10 种:研究院/所/

中心/、大学/学院/系、实验室/野外台站、学/协会、政府机构、国际组织/非盈利组织、图书馆/信息中心、博物馆/档案馆、个人、企业/公司。某种程度上讲,机构和服务的存在是从管理的角度为资源提供多样性的捕获和发现途径。

(4)关系实体,是一种规则,是关联资源、机构和服

务实体的纽带。概括来说,一个资源可以拥有 0 个或多个子集、超集及相关集合,具有一个或多个创建机构、所属机构和资助机构,具有至少一个服务。同理,机构可以拥有、创建和资助一个或多个资源,管理一个或多个服务。一个服务至少服务于一个资源。IRSR 预定义了 13 种关系结构和 6 组关系类型标识。这 13 种关系包含一组自反的关系类型(相关集合)和 6 对互逆的反向关系组(由 pair_id 标出)。具体描述如表 2 所示:

表 2 IRSR 基本关联关系

Id	relation_cn	relation_en	relation_type	pair_id	description
1	子集	dcterms:hasPart	col_col	2	所引用的资源包含在所描述的资源中
2	超集	dcterms:isPartOf	col_col	1	所描述的资源是被引用资源在物理上或逻辑上的组成部分
3	相关集合	cld:associatedCollection	col_col	3	同一个集合下,与当前资源集合平级的资源集合
4	创建机构	dc:creator	col_org	8	集合的创建和更新机构
5	所属机构	rsipcd:owner	col_org	9	资源集合的法定拥有者
6	资助机构	zydj:sponsor	col_org	10	资源集合的经济支持者
7	有服务	cld:isAccessedVia	col_ser	12	提供对资源集合进行使用的服务
8	创建的资源	dc:create	org_col	4	创建资源的 URI
9	拥有的资源	iesr:owns	org_col	5	机构合法拥有的资源的 URI
10	资助的资源	fund	org_col	6	资助资源的 URI
11	管理的服务	iesr:administers	org_ser	13	机构管理的服务的 URI
12	服务于	iesr:serves	ser_col	7	服务对应的集合资源
13	管理者	rsipcd:administrator	ser_org	11	管理服务的机构

2.2 IRSR 关联策略的优点

通过分析相关数据模型和实体关系,可以发现与先前的登记系统相比,IRSR 有了很大的改进,主要可以总结为以下两个方面:

(1)以实际的用户群需求为导向,在对综合科技资源进行描述的时候,除了遵循和定制文献资源描述领域的基本行业标准,揭示了资源自身的一些基本属性信息之外,还对资源相关的实体信息进行了多维度的关联和扩展,如提供了相关资源集合、所属的机构、资助的机构、服务的对象信息等。而相关联的实体信息在物理上或逻辑上与资源实体信息都是相对独立的,资源实体仅提供了一个或一组关联指针(Handle 标识)作为定位信息,指向已知的关联实体的详细信息点。这样不仅能够方便元数据管理者或发布者更加全面地描述、重用和更新资源,还能提高用户发现资源和选择最合适的服务的效率。

(2)面向资源可重用性,独立封装了关系实体。关系实体是 IRSR 预定义的实体关联规则,封装了资源、机构、服务实体之间的关系信息,却并不显式依赖于这三类实体。关系实体只存储了这三类实体间相互指向的

指针标识符以及对应的关系类型。这一做法从数据模型算法设计的角度看,避免了数据冗余,提高了数据更新的速度。对元数据管理者而言,同一资源如果服务于多种服务类型,或同一资源被多个机构所拥有,可以直接通过添加关系指针标识符进行引用,而不用重复描述,降低了资源重用的复杂性,提高了系统的整体性能。

2.3 IRSR 关联策略的局限性

尽管上述的资源关联策略已在 IRSR 系统中得到了很好地体现和应用,相关的评估和测试也表明了其策略的合理性,但依然存在着一一定的局限性。语义 Web 模式的另一个技术核心是机器可读、机器可理解^[9]。如何将 IRSR 的数据转换为带语义的关联数据(即 RDF 数据),使之能够通过语义 Web 数据之间的逻辑和规则进行推理,支持一定程度的替换和定制,从而在机器可读层面提高综合科技资源被访问和调用的自动化处理能力,同时增强 IRSR 与其他系统平台的互操作性,是一个问题。

3 综合科技资源关联策略的优化

针对上述问题,结合 IRSR 进行 Web 测试,本文给出了综合科技资源关联策略的优化方案。

3.1 引入 D2RQ 组件

IRSR 使用的是 MySQL 关系型数据库,可以在多用户和高强度数据负荷的条件下提供可伸缩的性能,但并不满足语义 Web 关联开放的基本原则。然而语义 Web 的实质并不是颠覆这些已有的数据方法,相反,它主要是提供 RDF 这样一种灵活的、可扩展的数据模型表示工具^[10],使资源的知识组织和语义关联更为容易。

D2RQ 就是其中最常用的组件之一,其主体框架^[11]如图 2 所示:

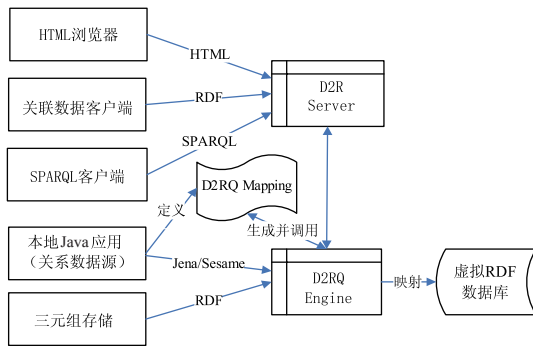


图 2 D2RQ 组件的主体框架^[11]

D2RQ 主要包括 D2RQ Server、D2RQ Engine 以及 D2RQ Mapping^[11]。

(1) D2RQ Server 是一个 HTTP Server,它的主要功能是提供对 RDF 数据的查询访问接口,以供上层的 RDF 浏览器、SPARQL 查询客户端或传统的 HTML 浏览器调用。

(2) D2RQ Engine 的主要功能是使用一个可定制的 D2RQ Mapping 文件将关系型数据库中的数据转换成 RDF 格式。它并不是将关系型数据库发布成真实的 RDF 数据,而是使用 D2RQ Mapping 文件定义的规则将其映射成虚拟的 RDF 格式。

(3) D2RQ Mapping 文件的作用是在访问关系型数据时将 RDF 数据的查询语言 SPARQL 转换为 RDB 数据的查询语言 SQL,并将 SQL 查询结果转换为 RDF 三元组或 SPARQL 查询结果。

目前 D2RQ 支持 Oracle、DB2、MySQL、PostgreSQL 等几乎所有的主流关系型数据库。

3.2 利用 D2RQ 将 IRSR 关系数据源公开为 RDF 接口

使用 D2RQ 将关系数据库公开为 RDF 接口主要有两个步骤:

(1) 定义并生成映射文件。

该映射文件用来定义数据库中的表和列与本体中的类和属性进行映射的规则,它并不生成真正的本体,主要是隐式定义并使用了一个类和属性集合的 vocab 前缀,等效于本体的命名空间。主要通过调用 D2RQ 组件 Java 类的 generate_mapping 方法实现,具体如下:

```
generate - mapping
- u root
- p ****
- d com.mysql.jdbc.driver
- o irsr.n3
jdbc:mysql://localhost:3306/irsr
```

其中, -u 表示数据库用户名; -p 是数据库登录密码; -d 是数据库驱动类名称,不同的数据库驱动格式不同,本项目是以 MySQL 为例; -o 是 D2RQ 输出的映射文件名称,后缀指定为 n3;最后一个参数设置了所映射的 MySQL 数据库的路径和名称。

生成的映射文件如下:

```
@ prefix map: < file:/F:/wsl/protocode/d2r - server - 0.7/irsr.n3# > .
@ prefix db: < > .
@ prefix vocab: < http://localhost:2020/vocab/resource/ > .
@ prefix rdf: < http://www.w3.org/1999/02/22 - rdf - syntax - ns # > .
@ prefix rdfs: < http://www.w3.org/2000/01/rdf - schema# > .
@ prefix xsd: < http://www.w3.org/2001/XMLSchema# > .
@ prefix d2rq: < http://www.wiwiss.fu - berlin.de/suhl/bizer/D2RQ/0.1# > .
@ prefix jdbc: < http://d2rq.org/terms/jdbc/ > .
map:database a d2rq:Database;
d2rq:jdbcDriver " com.mysql.jdbc.Driver";
d2rq:jdbcDSN " jdbc:mysql://localhost:3306/irsr ";
d2rq:username " root ";
d2rq:password " **** ";
jdbc:autoReconnect " true ";
jdbc:zeroDateTimeBehavior " convertToNull ";
.
#Table collection
map:collection a d2rq:ClassMap;
d2rq:dataStorage map:database;
d2rq:uriPattern " collection/@@ collection.Id@@ ";
d2rq:class vocab:collection;
d2rq:classDefinitionLabel " collection";
.
```

```
map:collection__label a d2rq:PropertyBridge;
d2rq:belongsToClassMap map:collection;
d2rq:property rdfs:label;
d2rq:pattern "collection @@ collection. Id@@ " ;
.
```

```
map:collection_Id a d2rq:PropertyBridge;
d2rq:belongsToClassMap map:collection;
d2rq:property vocab:collection_Id;
d2rq:propertyDefinitionLabel "collection Id" ;
d2rq:column "collection. Id" ;
d2rq:datatype xsd:int;
```

```
map:collection_dc_identifier a d2rq:PropertyBridge;
d2rq:belongsToClassMap map:collection;
d2rq:property vocab:collection_dc_identifier;
d2rq:propertyDefinitionLabel "collection dc_identifier" ;
d2rq:column "collection. dc_identifier" ;
.....
```

其中,d2rq:uriPattern 提供一个 URI 标识,用来指导生成所实例化实体的真实 URI,一般由关系数据库的实体表名和对应的主键组成;d2rq:class 表示该映射类所对应的实体类,其取值一般来自所应用的 OWL 本体或 RDFS Schema,也可以根据数据模型自定义新的 class,本例中的 class 是取自 IRSR 系统中的 collection 实体;d2rq:PropertyBridge 代表 OWL 本体或 ORDFS Schema 中类的属性,它具体映射到关系数据库中数据表的某一行。

(2) 调用映射文件对关系数据库进行转换与访问。

调用 D2RQ 组件 Java 类的 d2r_server 方法:d2r_server irsr. n3,启动 D2R Server,最终可将 IRSR 关系数据库中的相关实体 collection、org、service 转换为虚拟 RDF 视图的格式进行访问,服务端口为 2020,如图 3 所示:



图 3 IRSR 关系数据库实体的 RDF Schema

点击 collection,可以看到已经将其资源数据实例化为一组关联的 URI 标识,如图 4 所示:



图 4 collection 实体的 URI 标识

任意浏览其中的一个 collection URI,例如 collection 1001,详细信息如图 5 所示:



图 5 collection #1001 的 RDF 关联视图

最终构建的 RDF 接口可满足如下基本条件:

(1) 接口能够接收一个查询或者一个对返回数据的说明;

(2) 能够生成使用一个或多个一致的隐性本体描述的合法 RDF;

(3) 能够生成可以准确反映底层数据源结构的实体关系。

3.3 利用 SPARQL 端点技术查询 RDF 接口并公开为 Web 服务

在这里要用到 Jena 语义 Web 框架,因为它提供了调用 D2RQ 的强大的 API 以及基于 RDF 和 SPARQL 语言的 Java 编程环境^[12]。具体实现时,需要在 D2R Server 端建立一个基于 Query 方法的查询类,读取 irsr. n3 映射文件和 D2RQ 的虚拟 RDF 模型,构造符合 SPARQL 语法的查询语句,最后以三元组方式提供可视化的查询结果,并可定制为 XML、XML + XSLT、JSON 多种 Web 服务格式。以 collection 实体为例,查询其部分实例信息,基本 SPARQL 语法如下:

```
PREFIX rdfs: <http://www. w3. org/2000/01/rdf - schema# >
```

```

PREFIX db: <http://localhost:2020/resource/>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX map: <file:/F:/wsl/protegecode/d2r-server-0.7/irsr.n3#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX vocab: http://localhost:2020/vocab/resource/
SELECT ? dc_identifier ? dc_title ? dc_creator WHERE {
  ? collection vocab:collection_dc_identifier? dc_identifier.
  ? collection vocab:collection_dc_title? dc_title.
  ? collection vocab:collection_dc_creator? dc_creator.
} LIMIT 20
    
```

查询结果如图 6 所示:



图 6 SPARQL 查询 RDF 接口的结果

将 Results 定制为 XML 格式,示例如下:

```

<? xml version = "1.0" ? >
< sparql xmlns = "http://www.w3.org/2005/sparql-results#" >
  < head >
    < variable name = "dc_identifier" / >
    < variable name = "dc_title" / >
    < variable name = "dc_creator" / >
  < /head >
  < results >
    < result >
      < binding name = "dc_identifier" >
        < literal > C2010000006 < /literal >
      < /binding >
      < binding name = "dc_title" >
        < literal > 中国 - 欧盟生物多样性项目 < /literal >
      < /binding >
      < binding name = "dc_creator" >
        < literal > http://www.ecbp.cn/; < /literal >
      < /binding >
    < /result >
    .....
    
```

```

</results >
</sparql >
    
```

4 结 语

实际的测试和应用表明,使用上述策略的优点是:

(1)将新的数据源关联到 IRSR 系统后,只需要进行极少的配置改动,而且映射文件可根据关系数据库的实体模型自动生成。

(2)提供了 SPARQL 端点服务,方便第三方用户尤其是机器用户的自动调用。

(3)能够支持一定程度的语义查询和推理。

该策略也存在着一定的缺陷:

(1)生成的 RDF 接口几乎是关系数据库结构的一个精确映像,可能包含一些管理元数据方面的信息,而这些数据一般并不为用户所需要。

(2)SPARQL 端点技术要求用户对底层的关系数据库模式有较为细致的了解,才能构造可用的查询。因而使用对象可能会局限于科研人员、教师、学生等。

(3)关系数据库数据模式的变化会导致映射文件的重新生成和查询的改变。

在下一步的工作中,将考虑引入领域知识模型,定制更为标准的本体转换规则,以适应语义 Web 更为复杂的语义查询和推理要求,从而推动综合科技资源语义化知识服务机制的快速实现。

参考文献:

- [1] Hebler J, Fisher M, Blace R, et al. Semantic Web Programming [M]. 1st Edition. Wiley, 2009.
- [2] ANDS Guides and Other Resources [EB/OL]. [2011 - 04 - 20]. <http://ands.org.au/guides/index.html>.
- [3] Apps A. A Middleware Registry for the Discovery of Collections and Services [C]. In: *Proceedings of the 1st International Conference on E - Social Science*. 2005 : 1 - 12.
- [4] Archives:Michael and Michael Plus Projects [EB/OL]. [2011 - 04 - 28]. <http://www.michael-culture.eu/index.php?en/135/archives-michael-and-michael-plus-projects>.
- [5] The OCKHAM Initiative - Browse Digital Library Collections [EB/OL]. [2011 - 04 - 28]. <http://registry.ockham.org/collections>.
- [6] 开放式服务与资源登记系统 [EB/OL]. [2011 - 04 - 28]. <http://ocsr.las.ac.cn>.
- [7] 徐健,张智雄.数字图书馆服务登记系统的功能及应用 [J]. *图书馆理论与实践*, 2009 (5) : 64 - 67, 85.

- [8] Brogan M L. Contexts and Contributions: Building the Distributed Library [EB/OL]. [2011 - 04 - 28]. <http://old.diglib.org/pubs/dlfl06/dlfl06.pdf>.
- [9] 董慧. 本体与数字图书馆 [M]. 武汉: 武汉大学出版社, 2008.
- [10] 于晓繁, 王效岳, 白如江. 本体集成方法和工具综述 [J]. 现代图书情报技术, 2011(1): 14 - 21.
- [11] 张静, 马春娥. 如何利用 D2R 发布 Linked Data [EB/OL]. (2010 - 03 - 25). [2011 - 04 - 20]. http://www.ibm.com/developerworks/cn/web/1003_zhangjing_d2r/.
- [12] Péré J, Arenas M, Gutierrez C. nSPARQL: A Navigational Language for RDF [J]. *Web Semantics: Science, Services and Agents on the World Wide Web*, 2010, 8(4): 255 - 270.
(作者 E-mail: wangsl@llas.ac.cn)

亚马逊提升用户体验: 平台战争仍在继续

2011年2月的中上旬, 亚马逊公布了对 Kindle 的升级情况, 这次升级包括 4 个方面: 共享用户的注解内容, Kindle 用户现在可以按照以前他们分享高亮文章片段一样的方式来分享他们对文章的注解内容; 页码标记功能, 可以与图书某一特定印刷版本建立页码标记关联; 鼓励用户读完文章之后对文章进行星级标记, 并将读书状态和标记的星级发布到用户的 Twitter 或者 Facebook 上; 对杂志和报纸进行了全新的图形化布局, 能更好地展示每期的内容。这次的升级适用于第三代的 Kindle 设备, 这组功能将在随后一周内添加到其他版本设备的 Kindle 应用之中。

从这次公告(投资者并不将这次公告看成是正式新闻稿)的组成来看, 人们可能认为公众注解分享功能是最大的改进。共享是通过 Amazon Kindles 网站(<http://kindle.amazon.com>)上含有阅读、评论、识记三个选项的标签栏实现的。该网站已经展示出的新版 Kindles 有助于发现陌生或者模糊主题, 这主要通过分析那些得到 Kindle 用户最多高亮的主题和文章而得以实现。注解的分享正是这些功能的逻辑扩展。

亚马逊鼓励用户将他们的 Kindle 设备绑定到他们的 Facebook 或 Twitter 账户。然而, 实现这一功能的链接仍需要 Kindle 设备, 而不能通过社交网站来实现。

在最初的公告中最吸引眼球的是页码标记功能。在电子阅读器被作为热门礼品相互赠送的节假日期间, 主导产品缺少一样我们熟悉的功能, 如页码标记功能, 确实让很多人觉得奇怪。其竞争对手, 例如: Google 图书和 Barnes & Noble(已发现分页的实现方法)在这方面显得更有优势, 从而对 Amazon 构成了威胁。页码标记问题也说明了为什么圣诞节购物者选择有竞争力的产品, 比如新的 Nook Color 或者 iPad。这次 Kindle 最新的软件升级只是简单地为了赶上最近的竞争基准吗?

最大的阅读平台提供者——亚马逊、Google 和苹果——目前在功能上确实是处于同等水平。所不同的是各自大力宣扬是其支持和涉及专业领域。苹果的 iPad 主要定位于杂志中的图形显示, iPad 及其相关应用能利用颜色、布局和设计来吸引那些会快速把注意力放在视觉效果上的用户, 苹果同时也是多媒体平台的榜样。Google 因其长期从事图书项目、在线阅读器和其从图书馆和出版商处收集的大量的数字化产品, 在内容选择上提供了更广泛的范围。亚马逊进入市场最开始是靠着提供电子书, 更专于重点新书并拥有最主流的电子阅读设备。

亚马逊的页码标记功能与印刷版本是绑定的。Kindle 产品页面上有一块区域标明了用于 Kindle 进行页码标记的每个印刷版本的 ISBN 号。这一功能提高了教师在课堂上采用某一特定版本教材的可能性, 因为学生选择课堂参考书和引用目的书籍时, 他们喜欢选择的格式并不相同。Kindle 的检索结果中显示的页码还满足了那些只需要引用书中某段话的人的需求。但是亚马逊仍然关心课堂外用户的需求。用户为他们的 Kindle 设备所选择的字体大小不影响亚马逊的原有分页功能, 这是很多人最担心的一点。

另外一点很重要的意义是, 页码标记功能在数字版本中的增值功能是由分销平台而非版权所有者提供的。亚马逊在公告中指出, 页码标记功能应用的图书最主要考虑的是销售排名。它“为成千上万的 Kindle 图书增加了页码, 包括 Kindle 商店中前 100 名畅销书(与印刷版本进行了匹配)以及数千本最受欢迎书籍。”

曾与行业出版商和科学、技术、医学出版商都合作过的产业顾问 Joseph Esposito 这样解释: “网络环境下的基础设施很大程度上是因为考虑到一般消费者的需求而建立的。学术和专业出版商已经建立了一些数字化平台为机构内提供内容。但是他们无疑仍会期望看到为满足用户期望而创造出的来满足消费者市场的产品到底是什么。”

(编译自: <http://newsbreaks.infotoday.com/NewsBreaks/Amazon-Raises-the-Stakes-in-Your-Reading-Experience-The-Platform-War-Continues-74056.asp>)

(本刊讯)