

国家图书馆文献资源整合工作方案（2014-2015）

——“数字馆藏元数据改造”子项目介绍

国家图书馆数字资源部 冯红娟

- 1 《国家图书馆文献资源整合工作方案（2014-2015）》
- 2 数字馆藏元数据改造的必要性
- 3 我馆数字馆藏元数据基本情况
- 4 数字馆藏元数据改造的工作内容
- 5 中文数字馆藏元数据著录规范



整合方案工作目标

在2014-2015年间重点完成对古籍、民国文献以及1949年之后中外文文献资源各类元数据的全面整合，加强元数据统一检索力度，大力提升我馆文献信息整合与揭示的准确性、深度和内在关联性，并实现资源多平台统一展示与发布服务。

POINT

《国家图书馆文献资源
整合工作方案
(2014-2015)》



整合方案工作内容

- 建立超大型元数据仓储，提升文献资源利用率
- 加强元数据的规范建设，提升元数据揭示与发现功能
- 搭建资源统一发布与展示平台，提升用户操作体验满意度
- 利用数据分析技术,提升资源整合的可持续发展能力
- 加强专题资源库建设，提升馆藏资源的社会认知度与影响力

POINT

《国家图书馆文献资源
整合工作方案
(2014-2015)》



整合方案预期效果

1

在揭示层面，全面提高我馆资源的揭示与利用效率

2

经过科学梳理和有效整合文献资源，为社会公众提供更深层面、更细粒度的资源揭示，提供更快捷高效的查找、定位与获取目标信息的一站式服务，从而使海量、多样的馆藏资源得到合理开发、揭示和利用。

3

4

POINT

《国家图书馆文献资源
整合工作方案
(2014-2015)》



整合方案预期效果

1

在服务层面，为用户提供有序化、多层次的知识网络服务

2

在海量数字资源的大环境下，通过异构数字资源的融合、聚类和重组使资源从数据层的揭示与展现转向信息层、知识层的深度服务，将对国家图书馆实现公共数字文化共享、引领业界完善现代公共文化服务体系建设提供坚实有力的支撑。

3

4

POINT

《国家图书馆文献资源
整合工作方案
(2014-2015)》



整合方案预期效果

1

在展现层面，为用户提供全新的资源展示方式

2

以用户需求为导向,全面提升多媒体终端展示平台的用户体验,使图书馆所有资源的发现和传递过程得到优化,从而进一步加强国家图书馆作为业界引领者的公共文化服务能力和影响力。

3

4

POINT

《国家图书馆文献资源
整合工作方案
(2014-2015)》



整合方案预期效果

1

在管理层面，通过大数据处理和分析，实现决策管理的科学化和事业的可持续发展

2

通过对用户访问数据的收集、挖掘以及统计分析，深入了解用户需求，以及文献资源的使用情况，为文献的收集、服务管理提供依据，提升决策科学化水平。

3

4

POINT

《国家图书馆文献资源
整合工作方案
(2014-2015)》



自建资源

目前，我馆数字馆藏元数据主要调用自Aleph系统的书目数据，利用了书目数据中描述及标引资源的内容，但对实体文献成为数字文献后产生的新特性未加以描述，如分辨率、文件格式等。



征集资源

采用DC格式，但部分字段的语义不够明确、各批次征集资源的著录情况存在着差异，部分元数据还存在不规范的问题，为元数据的利用、共享及互操作造成困难。



其他

我馆数字馆藏除了CNMARC、DC格式外，还有部分元数据采用了ACCESS、EXCEL等多种格式，需进行改造，以提升元数据的揭示功能。

POINT

2

数字馆藏元数据改造的必要性

我馆**自建**数字馆藏类型多样，元数据以CNMARC为主。

■ 电子图书

■ 电子期刊

■ 电子论文

■ 古籍

■ 音视频

■ 少儿资源



我馆**征集**数字资源划分为12种资源类型，元数据以DC格式为主。

地方馆藏特色资源

家谱

少数民族

年画



地方志

老照片

少年儿童资源

科普科教



民国文献

地方视频资料

地方非物质文化遗产

历史文化



POINT

3 我馆数字馆藏元数据基本情况

我馆外购数字馆藏元数据涵盖MARC、SQL、ACCESS、TXT等多种类型。

📍 MARC

apabi 阿帕比 | 数字资源平台



ProQuest

📍 SQL Server

万方数据
WANFANG DATA
知识服务平台

bookan 博视网
.com.cn

📍 ACCESS

维普资讯
VIP

北大法意

中国人民大学
书报资料中心
INFORMATION CENTER FOR SOCIAL SCIENCES, RUC

📍 TXT

cnki 中国知网
cnki.net

POINT

3 我馆数字馆藏元数据基本情况



元数据改造工作计划

➤ 2015年元数据清洗的范围：

《中文图书》、《博士论文》、《方正电子书》、《征集资源》

后续会进一步清洗其他数字馆藏的元数据





元数据改造基本原则

改造基本原则：将描述实体资源的元数据清洗为描述数字馆藏的元数据，进一步规范元数据的内容。具体主要包含：

- 增加或完善时间、地域、分类号信息；
- 增加内容形式、媒体类型信息，替换一般资料标识信息；
- 增加描述数字馆藏特征的信息；
- 增加数字馆藏的发布地址和唯一标识符；
- 删除部分描述实体馆藏物理特征的信息。





中文图书改造规则



POINT

数字馆藏元
数据清洗工
作内容

改造前

LDR -----nam0-22----- 450-
001 003342444
005 20081121132712.0
010 |a 978-7-80158-900-2 |d CNY29.80
100 |a 20070111d2007 em y0chiy50 ea
1010 |a chi
102 |a CN |b 110000
105 |a a z 000yy
106 |a r
2001 |a 读三国悟人生智慧 |b 专著 |f 宋乃秋编著 |9 du san guo wu ren sheng zhi hui
210 |a 北京 |c 蓝天出版社 |d 2007
215 |a 263页 |c 图 |d 24cm
330 |a 本书共分四篇，内容包括：处世篇、决策篇、韬略篇和管理篇。对《三国演义》中无处不在的人生智慧作了诠释。
605 |a 《三国演义》 |x 人生哲学 |j 通俗读物 |2 ct
690 |a I207.413 |v 4
690 |a B821-49 |v 4
701 0 |a 宋乃秋 |4 编著 |9 song nai qiu
090 |a I207.41 |b snq
096 |a B821 |b snq
CAT |a QYP |b 01 |c 20070111 || NLC01 |h 1621
CAT |a QYP |b 01 |c 20070111 || NLC01 |h 1621
OWN |a ZB301
SYS 003342444



改造后

LDR -----nam0-22----- 450-
001 fa2015001013
005 20150302132712.0
010 |a 978-7-80158-900-2 |d CNY29.80
100 |a 20070111d2007 em y0chiy50 ea
1010 |a chi
102 |a CN |b 110000
105 |a a z 000yy
135 |a drbn^---auadu
2001 |a 读三国悟人生智慧 |f 宋乃秋编著 |9 du san guo wu ren sheng zhi hui
210 |a 北京 |c 蓝天出版社 |d 2007
215 |a 263页 |c 图 |d 24cm
281 |a 文本 |b 专著
282 |a 电子
307 |a 文件格式：PDF
307 |a 扫描分辨率：300DPI
324 |a 复制自：馆藏实体文献
330 |a 本书共分四篇，内容包括：处世篇、决策篇、韬略篇和管理篇。对《三国演义》中无处不在的人生智慧作了诠释。
452 |0 (NLC01)003342444
605 |a 《三国演义》 |x 人生哲学 |j 通俗读物 |2 ct
690 |a I207.413 |v 5
690 |a B821-49 |v 5
701 0 |a 宋乃秋 |4 编著 |9 song nai qiu
801 0 |a CN |b NLC
830 |a 发布系统-中文图书
8564 |u <http://mylib.nlc.gov.cn/web/guest/search/zhongwentushu/medaDataObjectDisplay?metaData.id=112446&metaData.lld=116936&ldLib=40283415347ed8bd013483174ef60002|9108.ndlc.2.1100009031010001/T1F24.003342444>



| 博士论文改造规则



改造前

LDR -----nam0-22----- 450-
001 003086430
005 20070115142241.0
010 |b 精装 |d 缴送
100 |a 20060225d2005 k y0chiy50 ea
1010 |a chi
105 |a y m 000yy
2001 |a 机器人遥操作技术的研究 |b 学位论文 |f 孙雷著 |g 刘景泰指导
209 |a 控制理论与控制工程 |c 南开大学 |d 2005
215 |a 137页 |d 30cm
6100 |a 机器人 |a 遥操作 |a 控制系统 |a 遥控系统
690 |a TP242 |v 4
690 |a TP872 |v 4
701 0 |a 孙雷 |4 著
702 0 |a 刘景泰 |4 指导
801 0 |a CN |b NLC
CAT |a LWLDC1 |b 01 |c 20060225 || NLC01 |h 0938
CAT |a LWLDC1 |b 01 |c 20061228 || NLC01 |h 1506
CAT |a LWLDC1 |b 01 |c 20061228 || NLC01 |h 1507
CAT |a LWHWY2 |b 02 |c 20070115 || NLC01 |h 1422
OWN |a LW202
SYS 003086430

改造后

LDR -----nam0-22----- 450-
001 fb2015005231
005 20150302102410.0
100 |a 20060225d2005 k y0chiy50 ea
1010 |a chi
102 |a CN |b 120000
105 |a y m 000yy
135 |a drbn^---auadu
2001 |a 机器人遥操作技术的研究 |f 孙雷著 |g 刘景泰指导
215 |a 137页 |d 30cm
281 |a 文本
282 |a 电子
307 |a 文件格式: PDF
307 |a 扫描分辨率: 300DPI
324 |a 复制自: 馆藏实体文献
328 0 |b 博士 |c 控制理论与控制工程 |e 南开大学 |d 2005
452 |0 (NLC01)003086430
6100 |a 机器人 |a 遥操作 |a 控制系统 |a 遥控系统
690 |a TP242 |v 4
690 |a TP872 |v 4
701 0 |a 孙雷 |4 著
702 0 |a 刘景泰 |4 指导
801 0 |a CN |b NLC
830 |a 发布系统-学位论文
8564 |u <http://mylib.nlc.gov.cn/web/guest/search/boshilunwen/medaDataObjectDisplay?metaData.id=410761&metaData.lld=415242&ldLib=40283415347ed8bd013483467a760008> |9 108.ndlc.2.1100009031010001/T3F24.003086430



方正电子书改造规则



改造前

LDR 00760nlm0 22002653 450
001 apabi000000376
005 20130627111310.0
010 |a 7-80081-754-7 |d CNY34.75
100 |a 20130627d1998 em y0chiy0121 ea
1010 |a chi
102 |a CN
135 |a drbn |||||
2001 |a 世界传世藏书 |b 电子资源 |f 夏于全
210 |c 蓝天出版社 |d 1998
215 |a 4册
230 |a 电子图书
304 |a 正题名取自题名屏
336 |a 文本型
337 |a 需要下载并安装APABI Reader软件阅读电子图书
6060 |a 推荐书目 |y 世界
690 |a Z835 |v 4
701 0 |a 夏于全
801 0 |a CN |b 方正APABI |c 20130627
|u http://202.106.125.14:8000/usp/nlc/pub.mvc?pid=book.detail&metai
8564 d=ISBN7-80081-754-7.3 |z 点击此处查看电子书
997 |a ISBN7-80081-754-7.3

改造后

LDR 00760nlm0 22002653 450
001 fc2015003314
005 20150302111310.0
035 |a apabi000000376
010 |a 7-80081-754-7 |d CNY34.75
100 |a 20130627d1998 em y0chiy0121 ea
1010 |a chi
102 |a CN |b 110000
135 |a drbn^---auuuu
2001 |a 世界传世藏书 |f 夏于全
210 |c 蓝天出版社 |d 1998
215 |a 4册
281 |a 文本
282 |a 电子
337 |a 需要下载并安装APABI Reader软件阅读电子图书
337 |a 通过Internet检索
6060 |a 推荐书目 |y 世界
690 |a Z835 |-v 4
701 0 |a 夏于全
801 0 |a CN |b NLC
830 |a 方正电子书
8564 |u http://202.106.125.14:8000/usp/nlc/pub.mvc?pid=book.detail&metaid=ISBN7-80081-754-7.3 |9 108.ndlc.2.1100009031010001/T3F24.003086320



征集资源清洗规则

```
<?xml version="1.0" encoding="utf-8"?>
```

```
<DC>
```

```
<item>
```

```
<名称>长江上五光而又十色的沪上文化区浅谈</名称>
```

```
<创建者 责任方式="主讲">刘锋</创建者>
```

```
<主题>沪上文化区</主题>
```

```
<描述 摘要="主讲人图文并茂地从五个方面来谈长江上五光而又十色的沪上文化区：一、面海开埠，强掳而出的万国文化；二、五光十色，海洋情调的浓烈风俗；三、江海灵气，揉荡人文的中西精髓；四、灵秀山川，独擅形胜的虹霓奇光；五、沪上文化区种种神秘现象。"></描述>
```

```
<日期 创建日期="2010-8-8"></日期>
```

```
<类型 专题类型="专题视频资料" 内容形式="图像（动态）"></类型>
```

```
<格式 文件格式="FLV" 媒体类型="电子" 载体形态="108分钟"></格式>
```

```
<标识符 CDOI="108.ndlc.2.1100009031010001/T7F9.1187291" 记录标识号="010071012010001"></唯一标识号>
```

```
<语种>chi</语种>
```

```
<权限 版权所有者="**省图书馆"></权限>
```

```
</item>
```

```
</DC>
```





改造后可达到的展示效果（时间轴）

世界数字图书馆 探索 时间线
世界历史 / 时间线

HISTORYSG

An online resource guide



3372 B.C.
玛雅历法第一天

美国

政府、政治与战争
宗教和哲学
文学与艺术
科学、技术与经济发展
旅行、探索和迁移

猜你喜欢



CLASSICAL
EMPORIUM
1299 - 1599

MARITIME RIVALRY
1600 - 1818

DPLA DIGITAL PUBLIC LIBRARY OF AMERICA

Home Exhibitions Map **Timeline** Bookshelf Apps

View: [Icons]

Search the Library

Timeline

Your search returned 8,416,553 results.

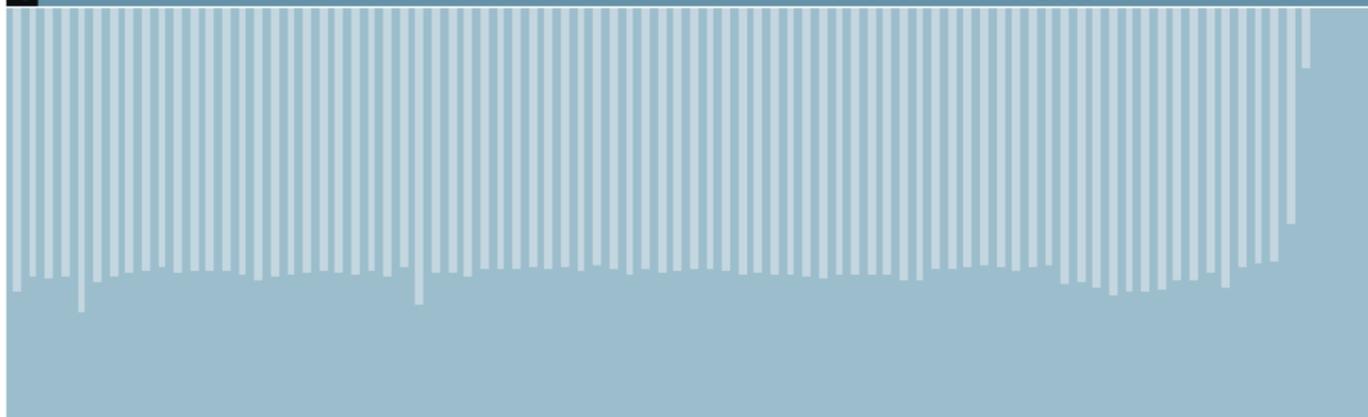
Only results with time data are shown below.

Show >

1000 1100 1200 1300 1400 1500 1600 1700 1800 1900 2000

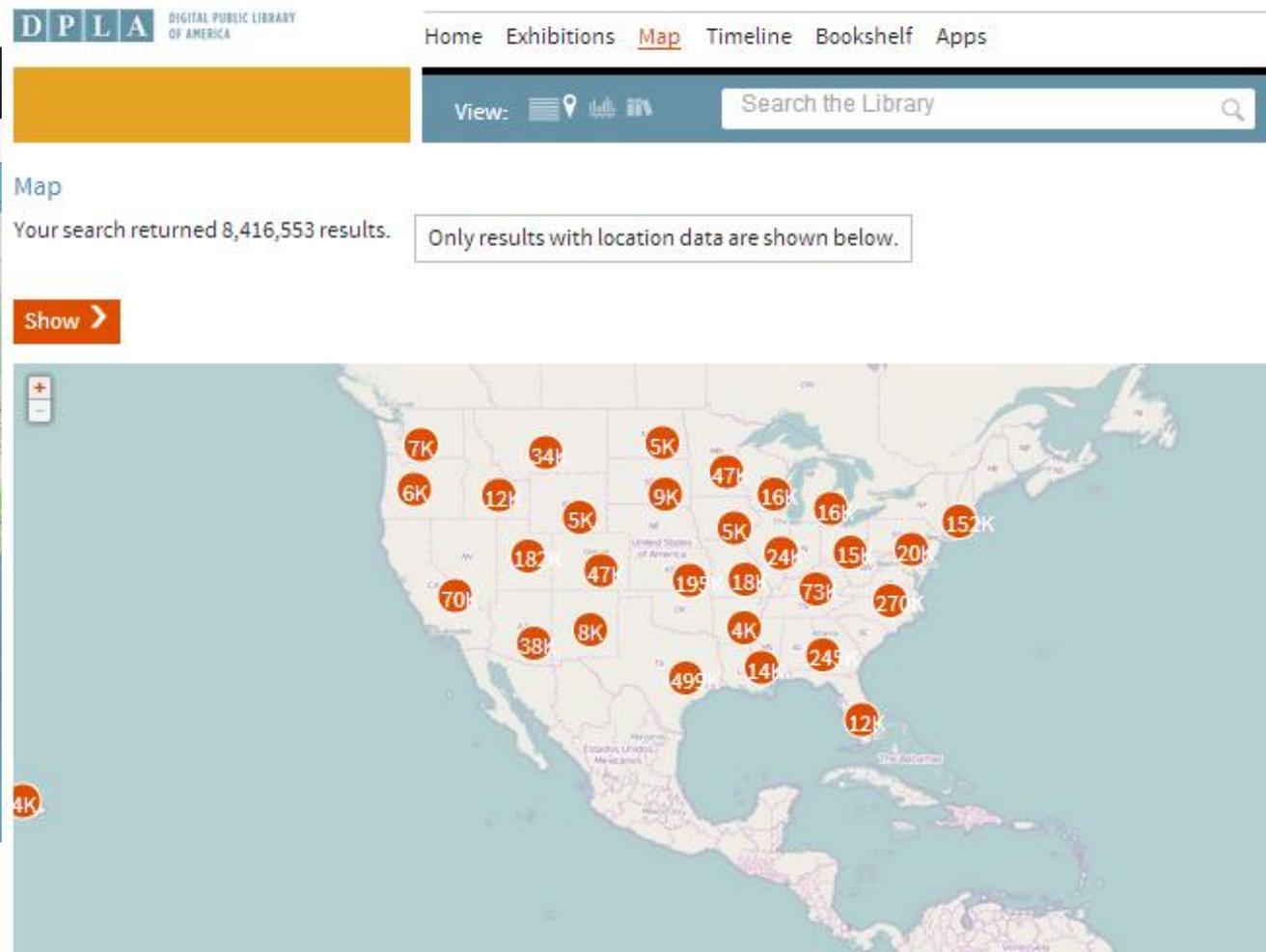
View: Decades

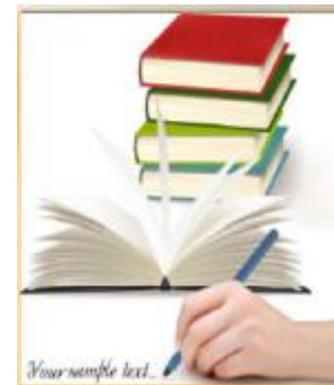
1940 1950 1960 1970 1980 1990 2000 2010





改造后可达到的展示效果（地域轴）





- 《国家图书馆中文数字馆藏元数据著录规范》的制定参考了《中国文献编目规则》、《新版中国机读目录格式使用手册》、《DCMI元数据术语集》、《信息与文献 都柏林核心元数据集》(GB/T25100-2010)、《国家图书馆元数据应用总则》、《信息资源的内容形式和媒体类型标识》(GB/T3469-2013) 等标准规范的相关内容。

- 《中文数字馆藏元数据著录规范》按元数据格式和资源类型分别起草了著录规范。
- 元数据格式：CNMARC、DC；资源类型：中文图书、学位论文、古籍及特藏、连续性资源、图像类资源、音视频资源、缩微资源。



谢谢！
Thanks