



欧洲图书馆和相关项目中的 UNIMARC

Nuno Freire
葡萄牙国家图书馆
nuno.freire@bnportugal.pt

中文翻译: 李恺 (首都图书馆, 北京)
Chinese Translator: Li Kai
Capital Library, Beijing, China

Meeting: 148. UNIMARC

WORLD LIBRARY AND INFORMATION CONGRESS: 76TH IFLA GENERAL CONFERENCE AND ASSEMBLY

10-15 August 2010, Gothenburg, Sweden

<http://www.ifla.org/en/ifla76>

摘要: 本文综览了 UNIMARC 在欧洲图书馆服务及相关项目中的使用, 特别强调了当前在这些语境中对于 UNIMARC 数据的处理和利用情况, 及对未来的展望。本文包含如下议题的信息: 在特定纲要中 UNIMARC 的转换和收割、分布式的搜索和检索、FRBR 化、版权管理以及对 Europeana 的数据提供。

介绍

欧洲图书馆项目 (European Library)ⁱ 创建于 2005 年, 它是欧洲国家图书馆会议 (Conference of the European National Libraries, CENL)ⁱⁱ 的一项服务, 意在为分布式馆藏提供通用的检索点, 这些馆藏来自于 48 个国家图书馆, 包含 32 种语言。欧洲图书馆项目当前提供整个欧洲范围内的 1 亿 5000 万条书目数据, 其中包含数字资源和模拟资源, 而且它是一项免费服务。

什么是欧洲图书馆？

Die Europäische Bibliothek L
305 The European Library De l
uropeos La Biblioteca Europe

- 面向所有欧洲国家图书馆的在线服务，提供各个图书馆目录和数字资源的检索，通过一个中心式的、多语言的网络界面（门户）实现。



当前的欧洲图书馆项目是之前若干个成功项目的结果：

- **Gabriel 项目**——欧洲各个国家图书馆的门户和桥梁（Gateway and Bridge to Europe's National Libraries）：它从 1996 年进行到 2004 年，建立了欧洲国家图书馆的联合门户网站；
- **TEL 项目**——欧洲图书馆：欧洲知识的门户（The European Library: Gateway to Europe's Knowledge）：它从 2001 年进行到 2004 年，为当前的服务奠定了基础。这个项目有 9 个国家参与：芬兰、德国、意大利（佛罗伦萨）、意大利（罗马）、荷兰、葡萄牙、斯洛文尼亚、瑞士以及英国。

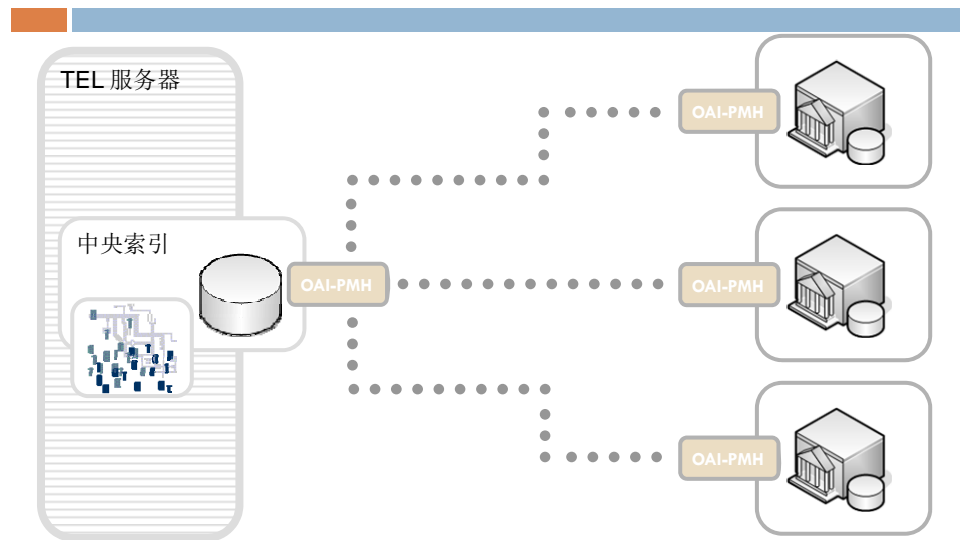
欧洲图书馆项目的成功带来了它进一步的发展和改进。国家图书馆成员的数量从最开始的 9 增加到当前的 48 个。这个项目的发展还得到了欧洲共同资助的项目的支持。比如欧洲图书馆：为网络资源提供中介的模块扩展（The European Library: Modular Extension for Mediating Online Resources, TEL-ME-MOR, 2005-2007），它的主要目标是支持、鼓励并且帮助 10 个欧盟新加盟国参与欧洲图书馆项目，这些国家也是欧洲图书馆项目的成员；以及欧洲数字图书馆项目（EDL, 2006-2007），它意在整合另外的国家图书馆，并且增强欧洲图书馆门户网站的多语种能力。这两个项目都有助于 19 个国家图书馆的馆藏整合进欧洲图书馆项目的服务中。

TELplus 项目是欧洲图书馆项目的另一个背景，这个项目是由一个 26 家成员组成的联盟承担的，它们包括国家图书馆和研究中心。TELplus 为欧洲图书馆项目提供增值的服务和产品。它的首要目标是改进服务架构的表现、增强多语种的检索功能、并且提供 OCR 处理文档的全文检索功能。

欧洲图书馆项目中元数据的使用

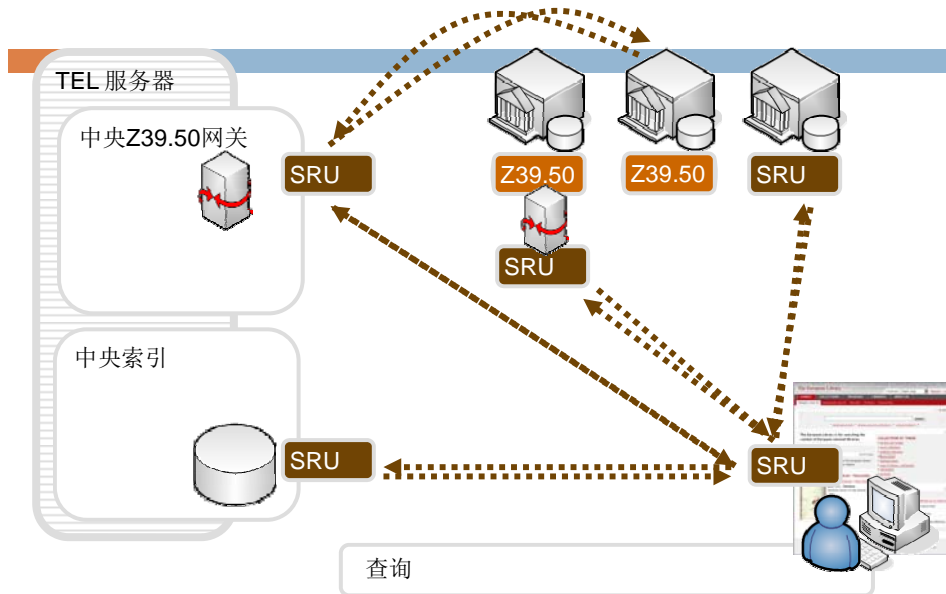
欧洲图书馆项目能够通过三种交换协议在其门户网站上提供各国国家图书馆的馆藏, 这三种协议是: 开放文献先导元数据收割协议 (Open Archives Initiative Protocol for Metadata Harvesting, OAI-PMH)、Z39.50 或者通过 URL 的搜索/检索 (Search/Retrieve via URL, SRU)。

通过OAI-PMH进行收割

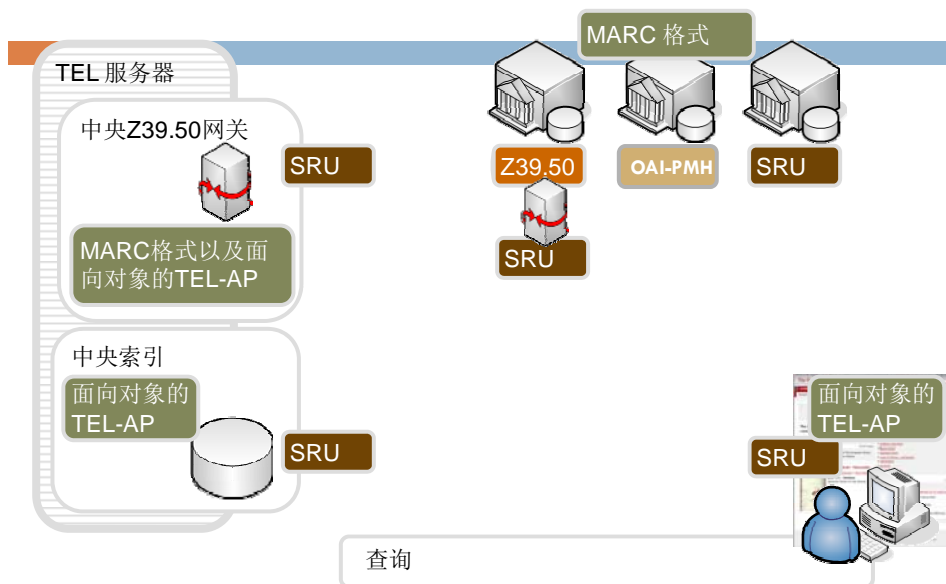


交换协议的选择极大的影响了门户网站能够提供给终端用户的功能。尽管这三个协议都提供了门户网站和图书馆系统之间的交换标准, 但它们的底层交换方式迥然有别。OAI-PMH 的设计允许门户网站把所有的元数据记录从图书馆收割到一个中央的仓储中, 而 Z39.50 和 SRU 则是为了远程的搜索和检索设计的, 所以元数据记录仍然保留在数据提供者那里。

通过SRU的搜索和检索



元数据格式



欧洲图书馆的门户网站依靠一种元数据应用纲要——欧洲图书馆对象应用纲要（The European Library Application Profile for Objects, TEL-AP），这种纲要基于 DC 元数据，以此实现欧洲各个国家图书馆所使用的不同书目元数据格式之间的互操作。门户网站和中央索引都被设计运行于这种元数据格式之上。

图书馆的元数据在总体上属于 MARC 家族，尽管它们也使用其他格式的元数据，尤其是对于数字馆藏和特殊馆藏而言。这些格式向 TEL-AP 的转换可以在各国图书馆进行，也可以在欧洲图书馆项目中进行。在第二种情况中，转换可以在中央 Z39.50 网关中进行，或者由中央索引吸收来实现。

欧洲图书馆项目中的 UNIMARC

2007 年进行的一项研究调查了各个国家图书馆对于书目数据格式的使用。ⁱⁱⁱ它研究了 29 个国家图书馆，发现 UNIMARC 及其派生的格式在其中 10 家图书馆中使用：阿尔巴尼亚、保加利亚、塞浦路斯、法国、希腊、立陶宛、葡萄牙、马其顿共和国、塞尔维亚以及斯洛文尼亚。

UNIMARC 在欧洲图书馆 中的使用

Die Europäische Bibliothek L
sos The European Library De l
uropeos La Biblioteca Europe

□ 在各个国家图书馆中

■ Christoph Schmidt-Suppran 在 2007 年进行的调查：

- 调查包括了 29 家国家图书馆：
- “.....关注数据格式和著录规则，国际标准毫无疑问占据优势。在 29 家图书馆中，10 家图书馆使用 UNIMARC 或者它的派生格式。”
- 阿尔巴尼亚、保加利亚、塞浦路斯、法国、希腊、立陶宛、葡萄牙、马其顿共和国、塞尔维亚和斯洛文尼亚。

在欧洲图书馆中，UNIMARC 在初期的项目就得到了支持，起始的 8 个国家中有 3 个国家使用 UNIMARC。UNIMARC 向 TEL-AP 的转换得到了 Z39.50 网关和 OAI-PMH 元数据收割系统的支持。这个转换被记录在《欧洲图书馆手册》(European Library Handbook) 中，我们可以在欧洲图书馆门户网站的成员区找到这个文件。

我们还有一种特殊的方法，通过 XSLT 在 UNIMARC 和 TEL-AP 之间进行转换，这种方法得到了 OAI-PMH 软件的支持，图书馆可以获得这个软件 (<http://repor.ist.utl.pt>)。

UNIMARC在欧洲图书馆 中的使用

Die Europäische Bibliothek L
os The European Library De l
uropeos La Biblioteca Europe

- UNIMARC转换为面向对象的TEL-AP
 - 中央Z39.50网关中的支持
 - UNIMARC的转换文档在《欧洲图书馆手册》（成员区）中获得
 - 转换UNIMARC的XSLT样式表可以在OAI-PMH软件中得到，图书馆可以获得这个软件
 - <http://repor.ist.utl.pt>

欧洲图书馆项目中其他最近的发展或者正在进行中的发展也包含了对 UNIMARC 的支持。我将在下文中描述这些进展。

TELplus 项目中的 UNIMARC

TELplus 项目于 2009 年 12 月结束，它有两个工作包包含对于 UNIMARC 的支持。工作包 2——“通过 OAI-PMH 一致性增加欧洲图书馆的可用性”——意在让参加欧洲图书馆项目的图书馆更多的采用 OAI-PMH。它提供了在图书馆目录和馆藏中推行 OAI-PMH 的工具和指南。它还关注在欧洲图书馆项目中 OAI-PMH 的核心用途，因为各种工具得到开发，而且这些工具被整合在中心索引中。正是在这个工作包中，通过 XSLT 从 UNIMARC 向 TEL-AP 进行转换的方法被开发出来，并且让这种方法能够在各国国家图书馆和欧洲图书馆项目中得到使用。

UNIMARC在TELplus项目工作包2 中的使用



- 工作包2——通过OAI-PMH一致性增加欧洲图书馆的可用性
 - 提供了在图书馆目录和馆藏中推行OAI-PMH的工具和指南
 - 提供扩展欧洲图书馆中央索引的工具
- 推行从UNIMARC到面向对象的TEL-AP的转换
 - 基于XSLT样式表（用于MARCXML和MarcXchange）
 - 在图书馆可以获得的OAI-PMH软件中：
 - 允许图书馆把它们的UNIMARC记录转换为面向对象的TEL-AP格式
 - 在欧洲图书馆的中央OAI-PMH 软件中：
 - 允许UNIMARC元数据被收割并被中心转换为面向对象的TEL-AP格式

TELplus 的工具包 3——“增强可检索性”——意在通过三种方式增强可检索性：增强全文检索；整合相关的资源；探索自动的词汇表映射。其中的一个任务“为欧洲图书馆建立 FRBR 仓储”使用了到 UNIMARC。这项任务得到了几个国家图书馆的支持，它们向这个实验贡献书目数据、规范数据，并且提供意见。

UNIMARC在TELplus项目工作包3 中的使用



- 工作包3，任务3.3——FRBR的聚合、搜索和浏览
 - 本任务的目的是制定解决方法，以便根据FRBR的方式，支持搜索欧洲图书馆的替代性的服务
- 这项工作关注特定的馆藏，尤其是为了更容易的对结果进行评估
 - 诺贝尔文学奖获奖者的作品
 - 概念的证明可见：
 - <http://digmap2.ist.utl.pt:8080/lgteFrbr/evaluation.html>

这项任务的首要目标是，通过在门户网站的搜索功能中推行 FRBR 模型，探索欧洲图书馆项

目内容的搜索和可检索性上可能的改进。

这项工作关注特定的馆藏而非整个目录，因为受控的馆藏能够让我们对结果进行更好的评估。最后选定的馆藏是诺贝尔文学奖获奖者的作品。

UNIMARC在TELplus项目工作包? 中的使用



- 欧洲图书馆项目成员的书目记录和规范记录都被集中到一起
- 在12家贡献馆藏的图书馆里，有5家上传的是UNIMARC的元数据记录
- 如下国家图书馆开展了UNIMARC数据的FRBR化工作：
 - 法国国家图书馆
 - 立陶宛马蒂纳斯·马齐达斯国家图书馆
 - 葡萄牙国家图书馆
 - 塞尔维亚国家图书馆
 - 比利时皇家图书馆

欧洲图书馆项目成员的书目记录和规范记录都被集中到一起，然后进行FRBR化处理。在12家贡献馆藏的图书馆里，有5家上传的是UNIMARC的元数据记录。

其中，如下图书馆的馆藏进行了UNIMARC数据的FRBR化处理：法国国家图书馆、立陶宛马蒂纳斯·马齐达斯国家图书馆（Martynas Mazvydas National Library of Lithuania）、葡萄牙国家图书馆、塞尔维亚国家图书馆以及比利时皇家图书馆。

这个任务证明的概念可以在下述网址中看到：

<http://digmap2.ist.utl.pt:8080/igteFrbr/evaluation.html>

ARROW 项目中的 UNIMARC

ARROW 项目^{iv}——版权信息和孤儿作品的可检索注册（Accessible Registries of Rights Information and Orphan Works）——关注版权信息的互操作，以及针对数字化目的的版权核准过程。它包括识别作品、阐明知识产权状况以及识别知识产权所有者三个部分。

UNIMARC在ARROW项目中的使用



- ARROW项目——面向Europeana的版权信息和孤儿作品的可检索注册
- 欧洲国家图书馆、出版者以及集体管理组织的联盟，意在：
 - 在领域内外传递可操作的版权信息基础结构
 - 有助于识别版权所有人（作者/视觉艺术家/出版者）
 - 有助于识别知识产权状态
 - 为以下内容提供测试平台
 - 数字图书馆的商业模型
 - 孤儿作品和绝版作品的注册
 - 版权核准机制
 - 由欧盟委员会高级专家组（EC High Level Expert Group, HLEG）制定的模型许可证
 - 互操作性及其标准

ARROW 依赖于现存书目数据和版权信息的可获得性。我们有已经建立好的、并且口碑不错的出版材料的信息基础结构，它包括：国家书目、在版图书和版权组织数据库。当下，这些来源是不能进行互操作的，因为它们的数据馆藏政策和数据蓝图都存在差异。书目数据库很少包括关于版权所有和使用政策的元数据。这些信息通常记录于由出版者、收藏社区和作者制作的大量其他格式中。

欧洲各国国家图书馆目录中的书目数据是 ARROW 的一个核心的数据源。因此，ARROW 构建于既存的、通过欧洲图书馆项目实现的互操作性的基础上，把它作为书目数据的核心来源。

UNIMARC在ARROW项目中的使用



- 欧洲图书馆是各国国家图书馆书目元数据的聚合器
 - 将会是识别和一个作品有关的所有载体表现的主要信息源
- 项目前期的结论认为，TEL-AP无法满足项目的要求。
 - 因此，欧洲图书馆项目正在准备一个新的中央索引，它基于MARC格式建立
 - 已经建立了对于MARC21的支持
 - 对于UNIMARC的支持正在开发中

项目前期的结论认为，TEL-AP 无法满足项目的要求。因此，欧洲图书馆项目正在准备一个新的中央索引，它基于 MARC 格式。第一家以 UNIMARC 格式在 ARROW 中建立目录的图书馆是法国国家图书馆。

向 Europeana 提供数据中的 UNIMARC

欧洲图书馆项目也是 Europeana（译者注：Europeana 是一个欧洲数字图书馆馆藏的检索平台）的聚合器。尽管到目前为止，欧洲图书馆向 Europeana 提供的绝大多数元数据都基于 TEL-AP，但在未来，数据将基于 MARC 格式。新的 Europeana 数据模型（Europeana Data Model）计划于 2011 年春季发布，它将允许 MARC 格式的元数据到达 Europeana 的时候，其语义质量得到更好的保存。UNIMARC 到 Europeana 数据模型的转换将会是欧洲图书馆项目的首要工作。

未来的工作

Die Europäische Bibliothek L
ios **The European Library** De l
uropeos La Biblioteca Europe

- 欧洲图书馆是Europeana的聚合器，在此语境
- 新的Europeana数据模型计划将在多瑙河发布（2011年春天）
- 欧洲图书馆将会支持UNIMARC向EDM的转换



参考文献：

ⁱ 欧洲图书馆门户网站的地址：<http://www.theeuropeanlibrary.org>

ⁱⁱ CENL 的成员是欧盟成员国的国家图书馆。CENL 更详细的信息可以参见：<http://www.cenl.org/>

ⁱⁱⁱ Schmidt-Supprian, Christop. 2007. 多语言联邦检索环境下的受控词汇表：欧洲图书馆的案例。

^{iv} <http://www.arrow-net.eu>