



基于IFLA-LRM的智慧图书馆数据 资源体系构建和应用探索

汇报人：苗青

课题组成员：薛崧 薛玮 袁丽霞 南春艳

复旦大学图书馆

2026年5月15日

目 录



01

研究背景



02

模型设计



03

实施路径



04

应用探索

01 研究背景

战略-趋势-挑战-路径

研究背景

01

国家：数字化转型战略

- 2015：《关于印发促进大数据发展行动纲要的通知》
- 2021：“智慧图书馆”被纳入《中华人民共和国国民经济和社会发展第十四个五年规划和2035年远景目标纲要》

02

学校：AI4S/AI4SS

- AI驱动科研创新生态建设
- AI-BEST课程体系建设
- AI推动跨学科深度融合

03

图书馆：智能基础设施

- 复旦大学图书馆在“十四五”期间，提出构建以“数字基座”为核心的图书馆服务平台，重点在于整合分散的资源与系统能力，实现数据层面的统一管理和服务支撑。

四个关键转变

资源范围

- 从传统出版物向多样化学术产出拓展，包括科研数据、预印本及其他新型资源



组织方法

- 从传统编目模式向自动化元数据处理转变
- 逐步走向以知识组织为核心



管理流程

- 从碎片化流程向一体化资源管理转变
- 强调流程整合与协同



服务模式

- 从检索与获取向智能化、知识化服务转型
- 支持深层次知识获取与分析



馆藏组织与知识服务面临的主要挑战

1

可获取性障碍

- 馆藏规模快速增长，但实体存储空间有限
- 部分资源仍采用较浅层元数据描述

- 实体资源获取受限
- 数字资源发现困难

2

数据孤岛与服务碎片化

- 资源数据分散在多个系统中
- 检索与服务体系相互割裂

- 数据整合
- 创新型知识服务

3

领域知识支撑不足

- 特色馆藏深度依赖学科知识支撑
- 图书馆在部分领域专业能力有限

- 难以支撑高水平科研与教学
- 资源价值挖掘不足

馆藏组织与知识服务发展 路径

Data → Information → Knowledge → AI-enabled Services

数据整合

- 对接国家图书馆名称规范数据
- 基于IFLA LRM的统一元数据基础

数据增强

- 基于LRM的文献实体关系网络
- 特色资源分面标引
- 资源可发现性提升

领域语料

- 考古文献库 (2024 - 至今)
- 考古报告语料库 (2024 - 至今)
- 梵文文献语料库 (2025 - 至今)

AI赋能知识服务

- 统一资源发现
- 基于图谱的导航与分面聚类
-



02 模型设计

以实体关系为中心

设计原则



设计原则



以实体-关系为中心原则

将数据组织的基本单位从“文献”转向“实体”，使每一个实体都具有独立的标识和属性。



可扩展性原则

强调模型在结构设计上的灵活性与可演化性，使其能够通过扩展实体类型、属性及关系来适应新的业务需求。



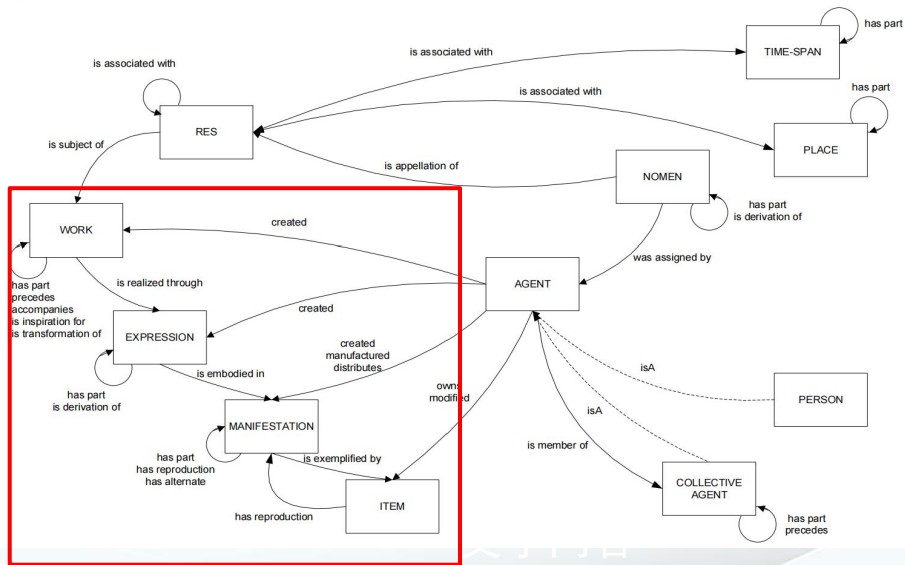
互操作原则

要求数据资源体系在设计之初即遵循国际主流标准与规范，以实现跨系统、跨平台乃至跨机构的数据共享与协同。

LRM概述

- **图书馆参考模型 (LRM)** : 将之前的概念模型——FRBR、FRAD 和 FRSAD——整合为一个统一且连贯的框架。
- **核心实体**: 作品 (Work)、内容表达 (Expression)、载体表现 (Manifestation)、单件 (Item) 四个核心实体。
- **规范相关实体**: 个人 (Person)、团体 (Corporate body)、地点 (Place)、时间 (Time-span) 等规范控制相关实体。

Figure 5.6 Overview of Relationships



--IFLA Library Reference Model: A Conceptual Model for Bibliographic Information

LRM模型选型依据

概念基础

实体分析技术

标准的一致性和延续性

ICP、ISBD、RDA

01

丰富的语义体系

LRMer、RDA

05

02

IFLA LRM提供了若干机制，允许在实际应用时进行可能的扩展。为实体“事物”定义的类别属性，允许在实施过程中为任何实体创建其有用的子类型。按照模型提供的模式可对任意或所有实体添加特定的属性，例如，覆盖特定的资源类型或为行为者提供更多的细节 —IFLA 图书馆参考模型

03 实施路径

数据分析-实体建模-数据转换

数据流程



数据源

贴源数据

明细数据

业务对象数据

应用数据

学校数据中心
图书馆业务系统
校外合作机构数据
互联网开放数据
社交媒体数据
.....

人事/学籍数据
ALEPH500系统数据
师生空间利用数据
特色资源元数据
本校学术典藏数据
教师申请专利数据
文献基础知识数据
.....

师生明细数据
馆藏书目明细数据
业务操作明细数据
师生行为明细数据
特色资源明细数据
学术成果明细数据
.....

个人实体数据
载体表现实体数据
单件实体数据
用户行为实体数据
业务操作行为实体数据
据
.....

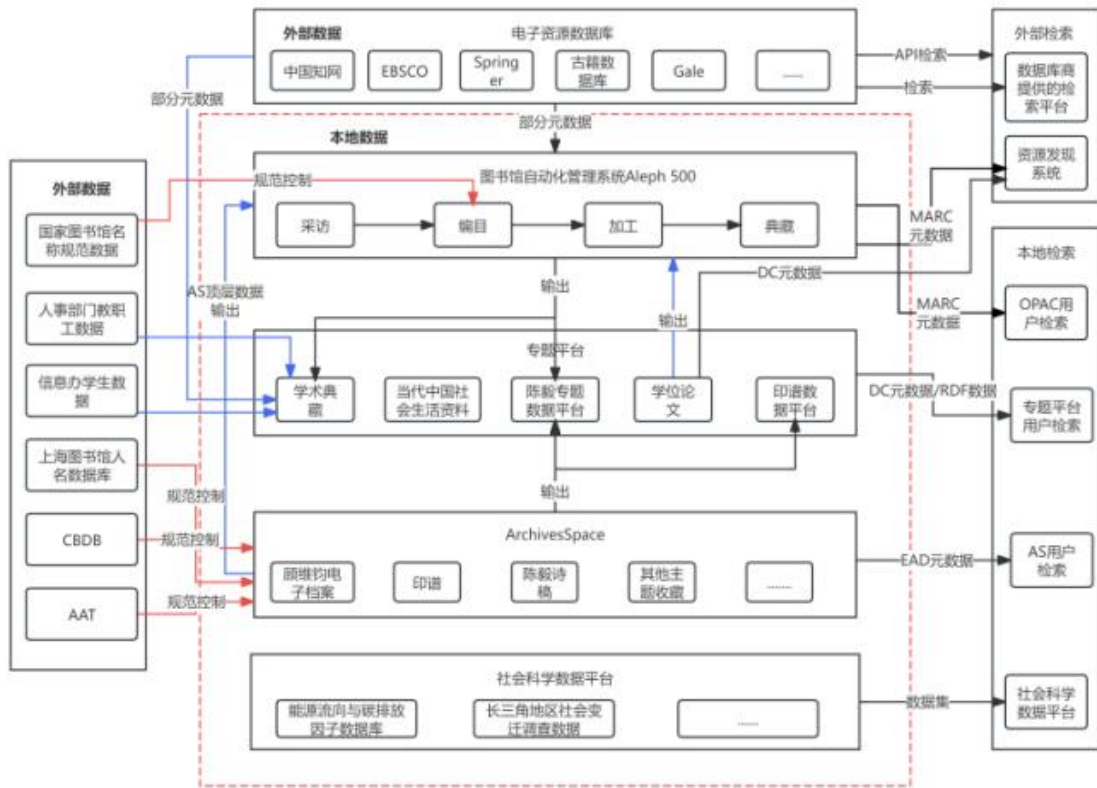
复旦学者及成果数据、
专题文献资源数据、
课程指定教参书数据
用户标签、资源标签、分级
分类标签、统计和评价指标
数据
专题文献知识图谱、馆藏评
价数据
.....

馆藏数据分析

01

文献资源数据

数据来源平台或系统；
 元数据格式；
 平台间的数据关联关系；
 文献资源对检索系统的支持方



据

期数据；
 控数据；
 据；
 行日志

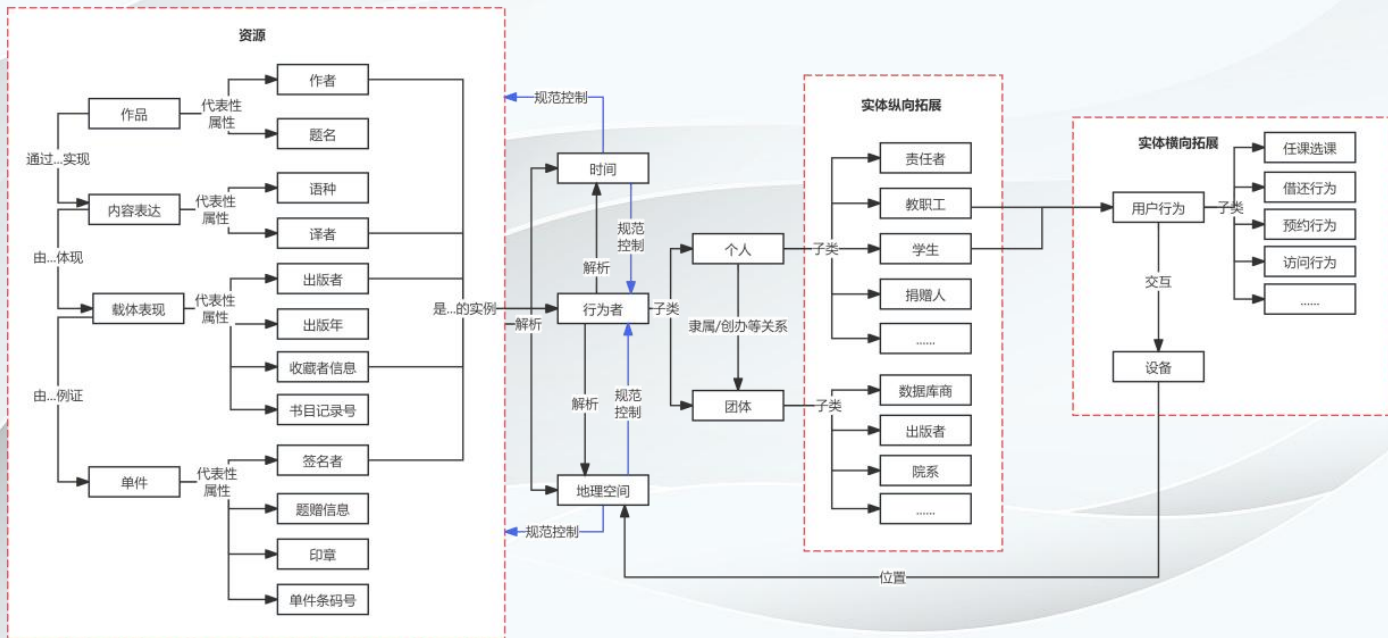
实体建模

➤ 实体的横向拓展：

- 指在LRM既有实体之外，引入新的实体类型以覆盖传统模型未能充分表达的对象。

➤ 实体的纵向拓展

- 指在LRM已有实体基础上进行细化与分层



数据处理和转换流程

关系构建

人和作品的责任关系、
主题关系；
作品和作品的改编、
翻译关系



实体抽取、映射 与消歧

对明细数据进行实体抽取并
与LRM扩展模型进行映射，
结合规范文档进行实体消歧。



明细模型

实现对不同格式标准数
据的解析和属性对齐。
如将CNMARC的701字
段，MARC21格式的100
字段解析为个人实体的
相关属性。



数据处理和转换流程



复旦-刘法迪
复旦-刘法迪
FDC-SH BUID Joint Research Center for Smart Library

首页 / 数据标准 / 属性管理

菜单搜索

首页 数据源管理 属性管理

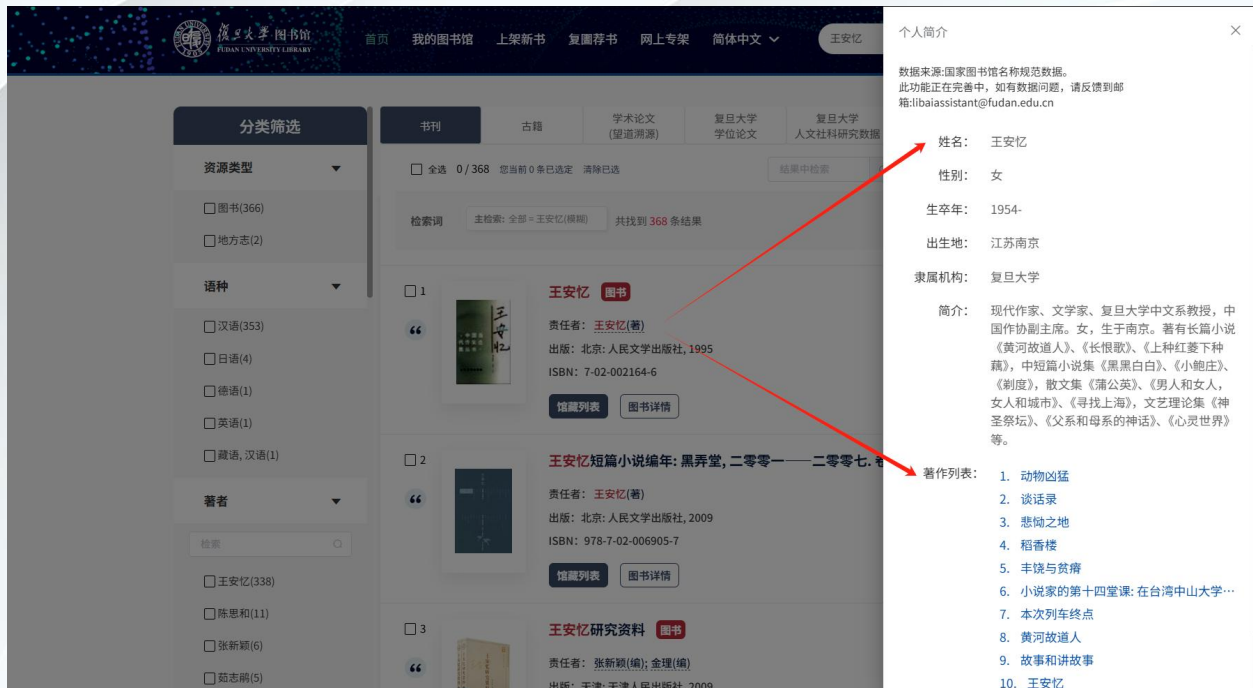
作品(3)
内容表达(1)
载体表现(10)
单件(4)
个人(5)
团体(1)
时期(1)
地点(1)
通讯(1)
馆舍空间和设施环境(1)
行为(1)
业务
科学数据(1)

编号	数据项名	中文简称	数据类型	操作
RE030101...	resource_sz...	(数字平台)载体表现层...	C	查看 编辑 删除
RE030101...	resource_zt...	载体表现层标识符	C	查看 编辑 删除
RE030101...	resource_zt...	载体表现层资源分类	C	查看 编辑 映射 删除
RE030101...	resource_zt...	载体表现层题名	C	查看 编辑 映射 删除
RE030101...	resource_zt...	载体表现层分卷标识	C	查看 编辑 映射 删除
RE030101...	resource_zt...	载体表现层分卷题名	C	查看 编辑 映射 删除
RE030101...	resource_zrz	责任者	C	查看 编辑 映射 删除
RE030101...	resource_zrfs	责任方式	C	查看 编辑 映射 删除
RE030101...	resource_b...	版本说明	C	查看 编辑 映射 删除

76% 1.9K/s 1.8K/s



基于实体关系的统一资源发现与知识服务



The screenshot shows the Fudan University Library search interface. The search results for '王安忆' (Wang Anyi) are displayed, including a list of books and a detailed '个人简介' (Personal Introduction) pop-up window. The pop-up window contains the following information:

数据来源: 国家图书馆名称规范数据。
此功能正在完善中, 如有数据问题, 请反馈到邮箱: libassistant@fudan.edu.cn

姓名: 王安忆
性别: 女
生卒年: 1954-
出生地: 江苏南京
隶属机构: 复旦大学

简介: 现代作家、文学家、复旦大学中文系教授, 中国作协副主席。女, 生于南京。著有长篇小说《黄河故道人》、《长恨歌》、《上种红菱下种藕》, 中短篇小说集《黑黑白白》、《小鲍庄》、《制度》, 散文集《蒲公英》、《男人和女人, 女人和城市》、《寻找上海》, 文艺理论集《神圣祭坛》、《父系和母系的神话》、《心灵世界》等。

著作列表:

1. 动物凶猛
2. 谈话录
3. 悲悯之地
4. 稻香楼
5. 丰饶与贫瘠
6. 小说家的第十四堂课: 在台湾中山大学...
7. 本次列车终点
8. 黄河故道人
9. 故事和讲故事
10. 王安忆

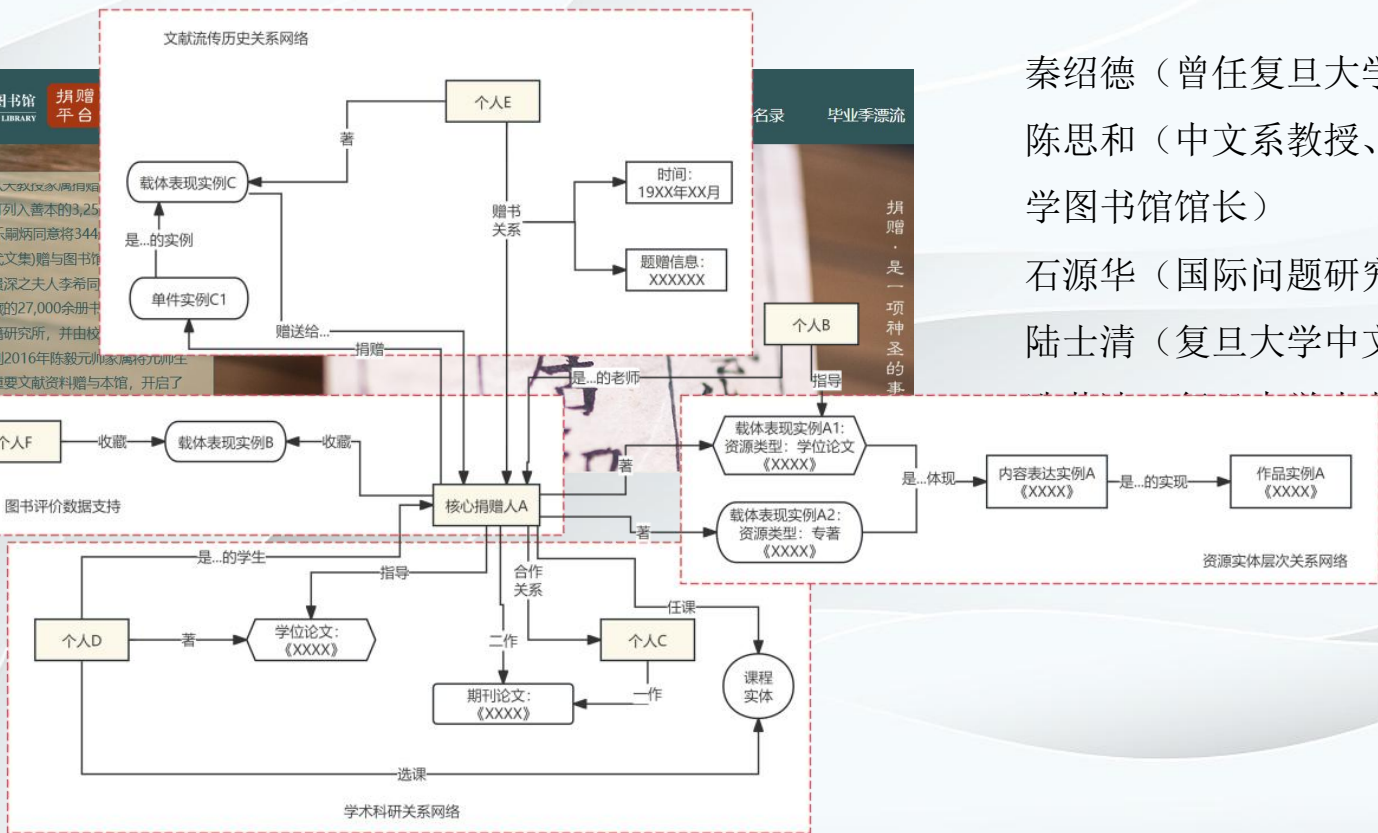
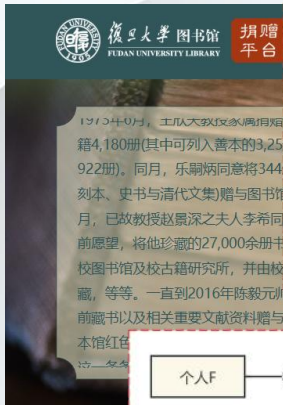
允许用户从某一作品出发, 关联至其作者个人信息、相关作品以及其他相关资源, 如高校教师的学术成果、捐赠人的藏书信息等, 实现跨实体的知识探索与发现



04 应用探索

以捐赠人为核心的知识网
络构建

以捐赠人为核心的知识网络构建



- 秦绍德（曾任复旦大学党委书记）
- 陈思和（中文系教授、曾任复旦大学图书馆馆长）
- 石源华（国际问题研究院教授）
- 陆士清（复旦大学中文系教授）
- 真学教授）

以捐赠人为核心的知识网络构建



数据集市获取相关数据

LRM

单册: 950
作品: 244
内容表达: 246
载体表现: 343
个人: 93
团体: 108

WEMI实体抽取

LRM关系统计

关系类型	数量	说明
hasSubject	4,381	主题关联
hasSeries	546	系列作品
creates	366	创作关系
publishes	352	出版关系
embodies	343	体现关系
realizes	246	实现关系
edits	37	主编关系

关系识别和发现

Person (人物) 96

关系类型统计

关系类型	数量	Neo4j关系名
合作关系	146条	HAS_COLLABORATOR
师生关系	28条	HAS_STUDENT
同行关系	9条	HAS_COLLEAGUE
同事关系	2条	HAS_COLLEAGUE
翻译关系	2条	TRANSLATES
母女关系	1条	HAS_PARENT
朋友关系	1条	HAS_FRIEND
总计	189条	

个人实体间关系说明

分析结果

研究作品分析:

- 《王安忆小说创作研究》: 研究了228个王安忆作品
- 《王安忆作品中的上海书写研究》: 研究了228个王安忆作品

被研究作品:

- 所有王安忆的作品都被研究, 包括《22岁, 莲花盛开》、《王安忆短篇小说编年》、《上种红菱下种藕》等

研究关系网络:

- 以研究作品为中心, 辐射到大量王安忆的作品
- 单向研究关系, 从研究作品指向被研究作品
- 网络高度集中, 围绕王安忆的创作展开

作品间关系说明

概已完成了作品间研究与评论关系的分析, 专注于阐明哪些作品研究王安忆作品以及其他研究关系。

分析结果

统计项	数值
研究/评论类作品	3个
研究的关系	456条
关系类型	研究关系

输出文件

研究与评论关系表已保存至: research_comments_ne4jations.csv

作品间评论研究关系识别

以王安忆为例的AI 实体识别和关系发现

研究总结与展望

01

国际标准概念模型在统一数据资源体系建设中的有效性，及对AI应用的强大支撑作用；

03

领域经典文献具有WEMI层级关系特征，实现对经典文献的组织，构建阅读指导地图。此外，WEMI层级关系及藏书共现网络也是图书评价的重要依据；

02

以捐赠人为核心的知识图谱是对捐赠行为的一种正向回馈。可以以点带面，拓展捐赠关系网络，为主动吸引优质特藏资源指明方向；

04

深入文献内容层面，探索文献资源分级体系构建及语料化的建设。

THANK!

谢谢!

苗青

qmiao@fudan.edu.cn

