

FRSAD及其对主题规范控制的影响

喻菲
国家图书馆中文采编部

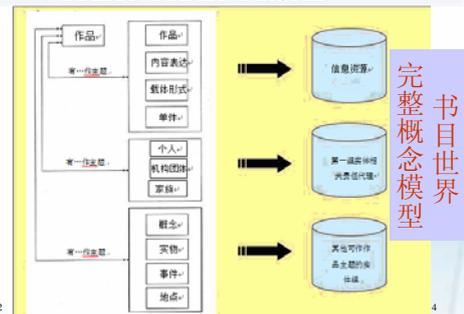
纲要

- 1. 背景资料
- 2. 书目世界完整概念模型
- 3. FRSAD概念模型
- 4. Aboutness、Ofness和Isness
- 5. FRSAD定义的实体、属性和关系
- 6. FRSAD概念模型可能存在的问题
- 7. 对主题规范控制的影响

背景资料

- 1997年，IFLA开发出书目世界概念模型FRBR。
- 1998年，FRBR最终报告发布。
- 1999年，FRANAR工作组成立，开发规范记录中实体的概念模型。
- 2005年，FRSAR工作组成立，开发主题规范记录中实体的概念模型。
- 2009年6月，FRAD最终报告发布，FRSAD全球评估草案发布。
- 2010年，FRSAD推出最终投票稿。

FRBR+FRAR+FRSAD



FRBR

- IFLA开发的概念模型
- 参考模型
- E-R模型
 - 定义书目记录的实体、属性和关系；
 - 3组实体：
 - 第1组 (FRBR的核心) = 知识或艺术创作产品：作品、内容表达、载体表现和单件
 - 第2组 = 与第1组实体有责任关系的实体：个人、团体、家族
 - 第3组 = 主题实体：概念、物体、事件、地点

2010-6-2

5

FRSAD概念模型

- 工作范围：
 - 在FRBR框架下建立同作品的Aboutness有关的第3组实体的概念模型
 - 为记录在主题规范记录中的数据以及使用这些记录的用户的需求提供一个有明确定义的、结构化的参考框架
 - 协助评估在图书馆界以及外界实现主题规范数据的国际分享与使用的潜力

FRSAD概念模型

- 例：一幅旧金山傍晚的油画
- Aboutness（关于什么主题？）：一段时间
- Ofness（画中表现了什么？）：傍晚、旧金山
- Isness（是什么？）：油画

FRSAD概念模型



- **Thema**:用于作品主题的任何实体
- **Nomen**:任何符号或符号序列（字母数字字符、符号、声音等），通过它可以知道、指示或称谓一个**thema**

传统书目数据的概念模型



2010-6-2 8

FRSAD定义的实体、属性和关系

- 2个实体
 - **Thema**
 - **Nomen**
- 属性
 - **Thema、Nomen各自带有一组属性**
- 2组关系

2010-6-2 9

Thema的属性

- **Thema属性**
 - **类型**
 - 在一些应用中，可以使用FRBR最初的实体类型：概念/实物/地点/时间。一般来说，任何FRBR和FRAD中定义的实体类型都可成为**thema**的类型，相应属性成为**thema**的属性。
 - 范围注释
 - 说明和定义**thema**，或指明它在特定主题规范系统的范围。

2010-6-2 10

Thema类型属性举例

——UMLS（统一医学医院系统）语义类型

- **实体**
 - 物理对象
 - 生物体
 - 解剖结构
 - 组织对象
 - 物质
 - 概念实体
 - 想法或概念
 - 发现
 - 生物属性
 - 智力产品
 - 语言
 - 职业或学科
 - 组织
 - 组属性
 - 组
- **事件**
 - 活动
 - 现象或过程

2010-6-2 11

Thema类型属性举例

——艺术和建筑分面叙词表（AAT）

- [AAT叙词表顶级分面]
- 相关概念
- 物理属性
- 风格和时期
- 制作者与组织
- 活动
- 材料
- 物体

2010-6-2 12

- 从UMLS和AAT可以看到定义*thema*类型的不同方法。
- *thema*类型差别很大，因此就需要有不同的属性。在第一个实例，UMLS中的“物质”（一个物理对象）和“组织”（一个概念实体）的属性就很不相同。AAT中的“风格和时期”、“制作者与组织”、“材料”的属性也一样不同。

2010-6-2

13

Nomen的属性

- **类型** 两个重要的值：标识符、受控名称
- 体系
- 参考来源
- 表现方式
- 语言
- 文字
- 转录/音译
- 编码
- 形式
- 有效时间
- 用户
- 状态

2010-6-2

14

Nomen类型属性

- 两个重要的值：
 - 标识符
 - 实体的名称，在一个域内一致并且唯一。
 - 受控名称
 - 规范控制/词表维护过程创建的名称，通常作为一个检索点。

2010-6-2

15

2组关系

- 不同实体类型之间的关系（基本的关系）
 - 作品-*THEMA*关系（重申FRBR定义的基本关系）
 - *THEMA-NOMEN*关系（FRSAD提出的新关系）
- 同一类型实体之间的关系
 - *THEMA-THEMA*关系
 - *NOMEN-NOMEN*关系

2010-6-2

16

同一类型实体之间的关系

- **Thema之间的关系**
 - 等级关系
 - 属种关系 (Generic Relationship)
 - 整体-部分关系 (Whole-Part Relationship)
 - 实例关系 (Instance Relationship)
 - 多对一等级关系 (Polyhierarchical Relationship)
 - 其它等级关系 (Other Hierarchical Relationship)
 - 相关关系
 - 其它语义关系
- **Nomen之间的关系**
 - 等同关系
 - 部分关系

2010-6-2

17

问题1: Isness是否被包含在概念模型里?

- FRSAD
 - 建立同作品的>Aboutness有关的第3组实体的概念模型
 - *Ofness*也包含在这个概念模型里
 - *Isness*不包含在这个概念模型里
- 反对意见
 - FRSAD模型应该包含*Isness*
 - 很多受控词表（特别是音乐和艺术领域）已包含了作品的体裁/形式 (genre/form) 等主题词，是重要的用户检索途径。

问题2: FRBR第3组实体是否要划分?

- FRSAD

- 对第3组实体的任何进一步划分都可能指定一种特定方法构建用于作品检索的主题语言。
- 一个好的模型应该兼容任何特定领域的结构,也应该足够灵活可以适应不同的应用。
- FRSAD模型,作为一个概念模型,不应该在主题语言在特定应用的形式上有任何约束。
- 只有一个完全独立于任何应用的、在更高的层次上处理属性和关系的、更抽象的理论模型,才能实现这个目标。
- **FRSAD模型不作任何的主题分类推荐。**

2010-6-2

19

- 反对意见

- 概念/对象/时间/地点只是划分的一种方法,同样存在阮纲纳赞或其他划分方法。但是一个模型最根本的是从这些划分方法中选择一个,遵循它实现目标。
- 至于主题的划分可能会与“应用独立”产生矛盾的担忧是没有必要的,目前的FRSAD概念模型在涉及任何应用之前还有很大的成长空间。

2010-6-2

20

问题3: 是否需要为Thema建立URI?

- FRSAD

- 为nomen建立UNI, 没有为thema建立URI。

- 问题

从全球数据共享的角度来说,如果不同系统中的nomen指向同一个thema,需要一个独立于优先显示形式的标识符(URI),将这些nomen聚集到一起。

2010-6-2

21

问题4: 重新思考thema-thema、nomen-nomen的语义关系?

- FRSAD

- 在thema层建立概念之间的**等级关系、相关关系**,在nomen层建立**等同、整体部分**关系。

- 问题?

- Thema之间可否建立**整体部分**关系,然后nomen通过推理得到该关系?
- 主题规范数据中已有的nomen之间的等级关系、相关关系如何表达?

2010-6-2

22

问题5: FRBR、FRAD和FRSAD报告之间联系?

- FRSAD

- **最新投票稿: 增加附录B. RELATIONSHIP OF FRSAD WITH FRBR AND FRAD**

FRSAD和FRBR

- FRSAD和FRBR

- 在FRBR框架中, *thema*是一个超级类,包括现有的第1组和第2组实体和其他所有可作为作品主题实体。

FRSAD和FRAD

- 相同
 - Thema-NOMEN关系符合FRAD，将事物同其所知、所指或所称（其名称、标签）区分开。
- 不同
 - 新增“发现”（Explore）用户任务。
 - *Thema*是所有可作为作品主题实体的超级类。
 - 预先没有明确定义任何实体。
 - *Nomen*被定义为单独的实体，而不是属性。

对规范数据全球共享的影响

- 主题规范的一项重要任务
 - 建立主题词之间以及它们同其他标引词之间的关系，保证在一个系统内部主题标引的规范和统一，同时方便用户从不同的入口词查找资源。
- 全球共享和利用的挑战
 - 在规范数据的全球共享和利用努力中，由于不同的系统可能使用了不同的主题规范，它们之间的共享通常是通过词表之间或词表与分类法之间的语义映射实现。

2010-6-2

26

- FRSAD“作品-thema-nomen”概念模型的提出，在各个系统的nomen之上添加了一个thema层。
- 将主题同主题所知的、所指的和所称谓的东西分离开。
- 不管各个系统中使用何种形式、语种的nomen，如果其所指是相同的，它们就指向同一个thema，从而实现主题规范数据的全球共享和利用。

2010-6-2

27

对受控词表的影响

- 如前所述，FRSAD声明仅对Aboutness、Ofness建模，Isness应由FRAD模型解决。
- 现实
 - 受控词表：很多受控词表（特别是音乐和艺术领域）已包含了作品的体裁/形式（genre/form）等受控词。
 - 如果FRSAD概念模型坚持不包括Ofness，此类词表将面临无法FRSAD同兼容的问题。

对主题标引模式的影响

- 目前主题标引模式：包含\$*a*主题因素、\$*x*通用因素、\$*y*空间因素、\$*z*时间因素、\$*j*文献类型因素等。
- 主题标引各子字段都有可能超出Aboutness的范围，这些内容将无法在主题标引模式中体现。
 - 例：《红楼梦》
 - 主题标引：\$*a*长篇小说\$*y*中国\$*z*清代
 - 例：《三国演义》
 - 主题标引：\$*a*长篇小说\$*y*中国\$*z*明代

谢谢！

喻菲

国家图书馆中文采编部

Email: yfci@nlc.gov.cn

2010-6-2

30