



《主题规范数据的功能需求 (FRSAD)》介绍 及其与 SKOS 等模型之间的映射

曾蕾, 美国 俄亥俄 肯特州立大学

Marcia Lei Zeng, Kent State University,
Kent, Ohio, USA

玛娅. 祖玛, 斯洛文尼亚 卢布尔雅那大学

Maja Žumer, University of Ljubljana,
Ljubljana, Slovenia

*在国际图联“主题规范记录的功能需求 (FRSAR)”
工作组的工作基础上撰写

中文版由曾蕾编译

Chinese version prepared by Marcia Zeng

Meeting: 200.

Classification and Indexing

WORLD LIBRARY AND INFORMATION CONGRESS: 75TH IFLA GENERAL CONFERENCE AND COUNCIL
23-27 August 2009, Milan, Italy
<http://www.ifla.org/annual-conference/ifla75/index.htm>

文摘:

“主题规范记录的功能需求 (FRSAR)”工作组成立于 2005 年, 是国际图联 (IFLA) 《书目记录的功能需求 (FRBR)》家庭的第三个工作组, 其成立是为了针对主题规范记录方面的问题, 并且调查广大用户对主题规范记录的直接和间接利用。本文介绍由 FRSAR 工作组建立的模型——《主题规范数据的功能需求 (FRSAD)》, 并将之与最近几年里定义的一些概念模型一起进行讨论, 其中包括《英国标准 BS8723-5: 信息检索的结构化词汇-指南—第 5 部分: 交换格式和互操作性协议》、W3C 的《SKOS 简单知识组织系统》、《OWL 网络本体语言》。这些模型使我们能够在更高层次上——并且独立于任何具体实施、体系、场合——来探讨主题规范数据和概念体系, 同时使我们能够专注于主题规范数据的语义、结构、互操作性问题。

国际图联(IFLA)的“主题规范记录的功能需求(FRSAR)”工作组是《书目记录的功能需求(FRBR)》家庭的第三个成员，成立于2005年4月，其责任是在FRBR框架中就FRBR第3组实体——即作品的“关于(aboutness)”方面——制定一个概念模型。本文介绍由FRSAR工作组建立的模型——《主题规范数据的功能需求(FRSAD)》，并置之於近年来随着语义网的关联的数据运动而发展起来的一些概念模型之中进行讨论。

1. 背景材料

FRBR第3组实体被认为是作品(works)(即知识或艺术的创作)的主题。除了第1组和第2组实体可以作为作品的主题外，第3组实体“代表可以作为作品主题的**其他的实体**”(IFLA, 1999: 16 and IFLA, 2008: 17, 黑体为本文作者所加以示强调)。FRBR第3组实体包括**概念**(抽象的概念或思想)、**物体**(物质的事物)、**事件**(行动或发生的事情)、**地方**(场所)(IFLA, 1999)。FRSAR工作组的成立是为了针对主题规范记录方面的问题，并且调查广大用户对主题规范记录的直接和间接利用。FRSAR工作组的职责由以下职权范围所规定：

- 在FRBR框架中就FRBR第3组实体即作品的“关于(aboutness)”方面制定一个概念模型。
- 就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、有结构的参考框架。
- 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。

FRSAR工作组成立了两个分组：“用户任务”分组重点放在确定用户群和用户的任务上；“主题实体”分组的重点则主要集中在第3组实体上，包括研究已被FRBR定义的第3组实体和其他可能替代品。两个分组在2006年和2007年进行了两次用户调查和一个小型实验，召开了大约20次会议，除此之外还有众多的联网讨论。FRSAR的顾问组曾于2008年底和2009年初对《主题规范数据的功能需求(FRSAD)》报告初稿进行了审查。修订后的报告在2009年米兰国际图联大会之前将有一个世界性的征求意见过程。

2. FRSAD模型

FRSAR工作组成立之际，在FRBR研究圈子里似乎已经达成共识，认为应当重新审核第3组实体(Delsey, 2005年)。工作组起初只是考虑怎样在FRBR第3组实体基础上去进一步完善之，很长一

段时间都耗在关于应该怎么充实实体（例如，增加 ‘时间’ 实体）的讨论上。工作组调查了其他各种模式，其中包括： <indec>模式 (Rust and Bide, 2000), 阮冈纳赞的分面理论 (Ranganathan, 1962), 由两名意大利的研究人员 Buizza and Guerrini (2002) 提出的很具体的实体名单。这些模型为 FRBR 概念模型的修改提供了非常有价值的参考资料。工作组根据上述每一个模型分析并讨论了可能的答案，包括从保守的改动（仅对 FRBR 第 3 组作出微小的修订）到彻底修改（根据上述某一个模式建立新模型）。然而，工作组发现，不管是这些模型本身还是在其基础上建立的模型，均不足以反映当今的主题规范数据的需要，特别是将不同的领域和不同的提供主题途径的工具考虑进来时更是如此。从 2007 年起，工作组把重点转移到一个全新的概念模型的发展上，在 FRBR 框架中就 FRBR 第 3 组实体 -- 即 *作品* 的“关于”方面 -- 制定了一个全新的概念模型。

在此框架内，与 FRBR 概念模型所界定的所有三组实体有关、作为查询入口的切入点都被认为有可能成为一个 *作品* 的主题。换句话说，第 1、2、3 组的所有实体与 *作品* 都可以有一个“是其主题”的关系。FRSAR 主题实体分组最终提出了一个更抽象的概念模型，并在 2007 年国际图联大会时发布。正如下图中所展现的，该模型可以从两个关键的角度来理解：

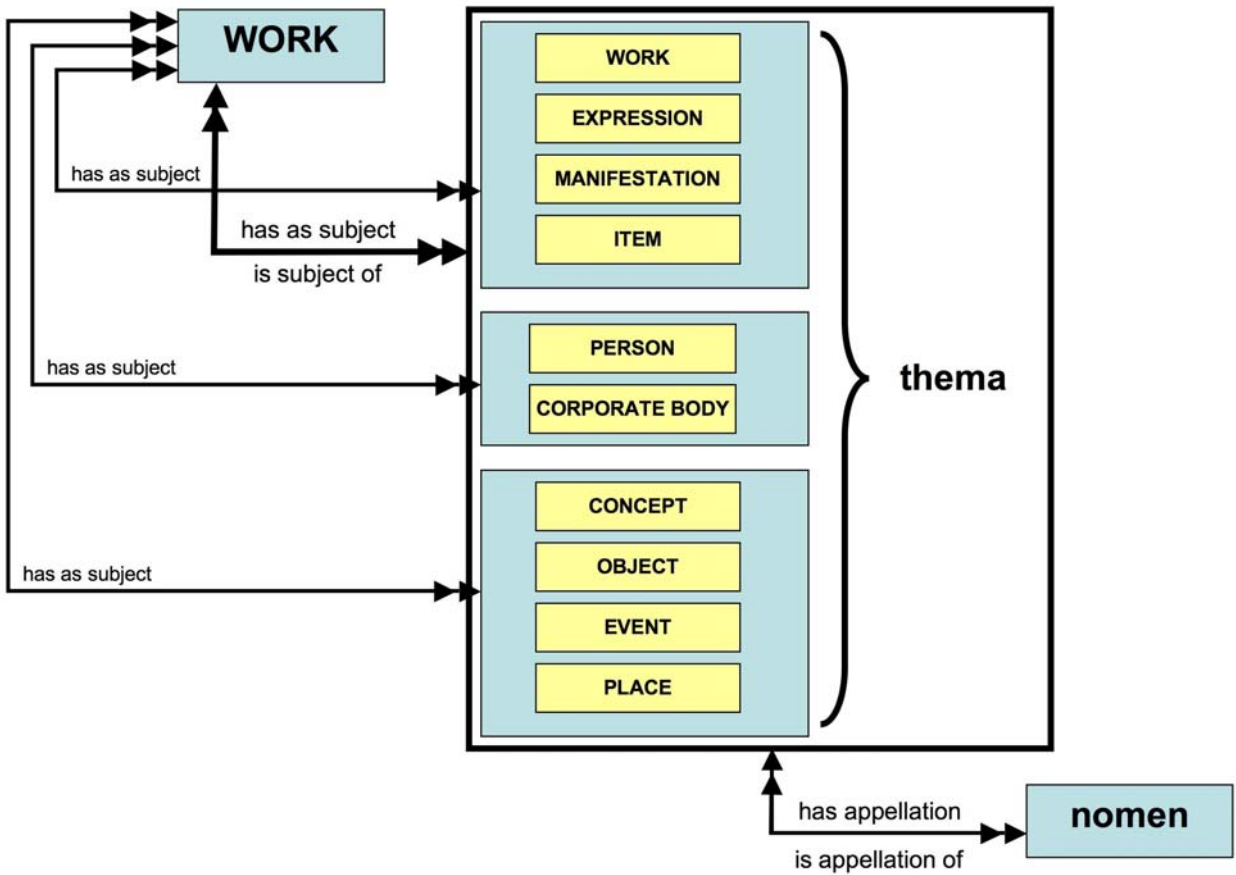
1. 该模型重申 FRBR 所定义的基本关系之一：*作品* 有主题 *希玛* (THEMA) / *希玛* 是 *作品* 的主题 (*WORK has as subject THEMA / THEMA is subject of WORK.*)。

1.1 *希玛* (THEMA) 一词用于代表可以作为 *作品* 主题的任何东西。可以说是作为 *作品* 主题的任何 FRBR 实体。

1.2 *希玛* 包括 FRBR 的任何实体 -- 现有的第 1 组和 2 组实体，以及除此之外所有的其他的 *作品* 主题。*希玛* 虽然本身是个实体，但可以被看作是一个含全部 FRBR 实体的超级实体或超级大类，这样我们得以在更普遍和抽象的水平上为实体间的关系和实体的属性建立模型。

2. 该模型还提出了一种新的关系：*希玛* 有称谓 *诺门* (NOMEN) / *诺门* 是 *希玛* 的称谓。 *THEMA has appellation NOMEN / NOMEN is appellation of THEMA.*

2.1 *诺门* 的定义是：任何众所周知、指、称 *希玛* 的标记符号 (字母数字字符，符号，声音等) 或标记符号的序列，例如“爱”，“∞”，或“595.733”。



图一. FRSAD 与 FRBR 的关系

对上述图示简化之后，FRSAD 模型可以用下图说明：



图二. FRSAD概念模型

下面进一步解释实体之间的关系：

“有主题/ 是 xx 的主题” 是多对多的关系。任何作品可以有一个以上的希玛；任何希玛可以成为一个以上的作品的主题。

一般来说（即在自然语言中或不同的词汇进行映射时），“有称谓 / 是 xx 的称谓”的关系也是一个多对多的关系。一个 *希玛* 有一个或多个 *诺门*，一个 *诺门* 可能指一个以上的 *希玛*。

这里有必要指出，在某一受控词汇或者某个领域范围内，一个 *诺门* 只应该是唯一一个 *希玛* 的称谓。

我们可以以史蒂芬·霍金的《时间简史：从大爆炸到黑洞》一书为例，这个 *作品* 有几个 *希玛*：宇宙学、空间和时间、物理学的统一、黑洞、宇宙大爆炸、时间的历史、宇宙等。还有许多其他的 *作品* 也可以有这些 *希玛*。上面列出的 *希玛* 所用的 *诺门* 是汉语词，但也可能有其他语言的 *诺门* 或根据其他受控词汇表（如主题词表，分类法等）形成的 *诺门*。《主题规范数据的功能需求 (FRSAD)》报告包括对若干问题的讨论，主要有：（一）*希玛* 的复杂性和粒度；（二）*希玛* 和 *诺门* 的属性；（三）*希玛* 与 *希玛* 之间、*诺门* 与 *诺门* 之间、*希玛* 与 *诺门* 之间的关系。

希玛 和 *诺门* 模型的重要意义在于将 *希玛*（或者说“主题”、“概念”、“（概念）的类”、“题目”等）与其称谓、其被提及或指代的形式区分开来。在规范数据的全球共享和利用努力中，有一些将重点放在 *诺门* 上（例如，元数据词汇的翻译，对称型的多语种词表，一个词汇的多入口索引），不过大多数努力都集中在概念这一层，例如两个词表之间或词表与分类法之间的语义映射。这种努力通常遇到更大的挑战，因为他们所关注的是 *希玛* 以及各 *希玛* 之间的关系。

3. FRSAD与SKOS等模型之间的映射

FRSAR 工作组的职责之三是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在性。真正意义上的分享会面对来自许多方面的挑战：异质结构、各种语言和文字、不同的建制规则和惯例、五花八门的编码模式等等。将 FRSAD 和其他模型做一个初步比较将使我们在更高层次上且不受任何具体实施、体系或者具体场合的影响来探讨主题规范数据和概念体系，使我们能够专注于主题规范数据的语义、结构、互操作性问题。

BS8723 和 ISO 25964：《英国标准 BS8723-5：信息检索的结构化词汇-指南--第 5 部分：交换格式和互操作性协议》(DD 8723-5:2008) 包括了一个模型结构，（该模型、XML 表单以及实例均可在 BS8723 官方网站上获取）。这一模型已被略加修订，继而纳入《ISO / CD 25964-1 信息和文

献-叙词表及其与其他词汇和互操作性-第 1 部分：用于信息检索的叙词表》中（2008 年：第 92 页），该部分标准已于 2009 年初经过投票过程。这个模型包括了为整个叙词表、叙词表概念系列、词条目的文件记录的建立模型所需要的元素。在这个模型中，叙词表的每个概念由一个首选词（每种语言有一个首选词）以及任何数量的非首选词来表达。类号、范围注释、上位/下位/相关词间关系都是针对整个概念（而非其首选词）而言。每个概念可以被分配给一个唯一的标识符（ISO/CD 25964-1, 2008）。总体上看，此模型和 FRSAD 模型都表达下列关系：（1）*希玛*和*诺门*的关系（一条记录记载一个概念及其*诺门*），（2）*希玛*和*希玛*的关系（等级（上位/下位/顶层概念）和相关概念），（3）*诺门*和*诺门*的关系（首选和非首选词、异体、不同语言的词等）。

SKOS：W3C 的《SKOS 简单知识组织系统》（2009）定义在知识组织系统(KOS)中表达常见特征的类型(classes)和属性(properties)，这些知识组织系统可以是叙词表、分类法、受控词单以及其他类型的 KOS 结构。“采用 SKOS，概念可用 URI 来辨识、一个概念可用一个或多个自然语言字符串作为标签，或者被配以标识符号（词汇代码），带有各种类型的说明，与其他概念相连接以组成非正式的等级和相关网络，汇总形成概念表，组合为有标记的和/或者有序的集合，并与其他概念表映射”（SKOS Reference, 2009: 概要）。作为 RDF（资源描述框架）的一种应用，SKOS 允许概念在网上组合和发布、与网络上的数据相联接、并且融入其他概念体系。每个 SKOS 的概念都被看作一宗 RDF 资源，每个概念可以带有 RDF 属性描述。SKOS 模型是建立在一个以概念为中心的词汇的观点上的，在这里原始的对象不是标签，而是标签所代表的概念。这与 FRSAD 模型所定义的*希玛*、*诺门*及其属性是相符的，SKOS 同时还采用特定的属性来表达所有的语义关系，这也与 FRSAD 相匹配。

OWL：《OWL 网络本体语言》是一种用于语义网的、带有严格定义的本体语言。本体是术语（terms）（即类和属性）的格式化的词汇，往往涉及特定领域且由某种社团的用户共享。OWL 本体通过对本体中术语间关系的描述来定义术语（OWL 2 Web Ontology Language Document Overview, 2009）。OWL 本体提供类(classes)、属性(properties)、个体(individuals)，数据值(data values)并以语义网文件形式存储之。OWL 原始版(OWL 1)主要是集中在构建有关类和个体的表达信息上。OWL 新版(OWL 2)是最新的 W3C 工作草案，提供了新的结构以表达更多的属性约束、属性的新特征、属性的不兼容性、属性链、关键属性等（OWL 2 Web Ontology Language New Features and Rationale, 2009）。OWL 2 提供了允许建立类表述（class expressions）间关系的公理(axioms)（即关于某领域中的事实的陈述），包括：是 xx 子类(SubClassOf)、等同类(EquivalentClasses)、不相交类(DisjointClasses)，不相交集(DisjointUnion)。更重要的是，OWL 2 中，类和属性表述

被用于构建类表达式(class expressions), (有时也被称为描述, 在描述逻辑文献中又称为复杂概念)。它允许列举个体以及所有的标准布尔运算符: 与(AND)、或(OR)、非(NOT)。诸如对象主体之交(ObjectIntersectionOf)、并(ObjectUnionOf)、补(ObjectComplementOf) 的类表达式提供了关于类表达式的标准形式的集合论的运算, 其 ObjectOneOf 类表达式则包括所确切指定的个体(OWL 2 Web Ontology Language Structural Specification and Functional-Style Syntax, 2009)。对于 FRSAD 涉猎的种种问题, 例如 *希玛* 的复杂性和粒度、*希玛* 之间的错综复杂的语义关系等, OWL 都有极为相称之处。

DCMI Abstract Model: 当 DCMI 抽象模型(DCMI AM) 在 2007 年成为 DCMI 推荐标准后, 其一对一原则(即每个 DC 元数据描述只描述唯一一个资源) 已被广泛承认或被其他元数据标准采用, 例如可视资源协会(VRA) 于 2007 年公布的《可视资源协会核心元数据》4.0 版(VRA Core 4.0, 2007)。按照 DCMI 模型, 一条记录(record)可以包含若干描述集(description sets), 其中可能包含由若干含属性-值(property-value)对的陈述(statements)组成的若干描述(descriptions)(DCMI Abstract Model, 2007)。这样的结果是在陈述层对信息进行处理、交换、指代、相连接。当一条记录包括对某资源的描述时, 这些描述也都可以被连接到与其属性有关的值的规范数据上, (例如, 主题规范数据, 专有名称规范数据, 或地理规范数据)。这样的信息模型是独立于任何特定的编码句法的, 因此有利于发展更好的映射和跨句法翻译。FRSAD 模型允许任何 *希玛* 独立于任何 *诺门*, 包括一个 *诺门* 可能采用的句法, 从这个角度来看, FRSAD 与 DCMI 抽象模型是相应的, 因此这一概念模式将有利于主题规范数据的共享和再利用, 不仅是主题词汇本身, 而且是元数据资源中的数据的共享和再利用。

综上所述, FRSAD 模型的建立是以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为宗旨的。有了 FRSAD 模型以及近年来随着语义网的关联的数据运动而发展起来的一些概念模型, 我们能够在更高层次上 - 完全独立于任何具体实施、体系或者场合 - 来探讨主题规范数据和概念体系, 同时能够专注于主题规范数据的语义、结构、互操作性问题。把主题规范数据放在语义网的发展环境中, 特别是从关联的数据(Linked Data)的角度来看, 根据 FRSAD 模型建模的、用 SKOS 和 OWL 编码的主题规范数据将能够成为关联的数据的一部分, 为语义网的发展做出贡献。

*鸣谢：本文基于 FRSAR 工作组的工作成果和顾问组的建议，本工作组由 IFLA 第四分部即书目控制部——特别是该部的分类和索引组——建立的。工作组的成员是：Leda Bultrini, Lois Mai Chan, Jonathan Furner, Edward O' Neill, Gerhard Riesthuis, Athena Salaba, Diane Vizine-Goetz, Ekaterina Zaytseva, Marcia Lei Zeng, Maja Zumer。顾问组的成员是：Victoria Francu, Hemalata Iyer, Dorothy McGarry, David Miller, Päivi Pekkariinen, Barbara Tillett。与 Glenn Patton 主持的 FRANAR 工作组的沟通和交流也颇有帮助。IFLA、OCLC、肯特州立大学提供了资金、设施、以及巨大的支持。

参考文献：

BS8723 Official Development Website. (2008). <http://schemas.bs8723.org/Home.aspx> (accessed May 15, 2009).

Buizza, P. and Guerrini, M. A. (2002). Conceptual model for the new "Soggettario": Subject indexing in the light of FRBR. *Cataloging & Classification Quarterly*, 34(4): 31-45.

Delsey, T. (2005). Modeling subject access: Extending the FRBR and FRANAR conceptual models. *Cataloging & Classification Quarterly*, 39(3/4): 49-61.

DCMI Abstract Model. (2007). Powell, A., Nilsson, M., Naeve, A. Johnston, P. and Baker, T. eds. <http://dublincore.org/documents/abstract-model/> (accessed May 15, 2009).

DD 8723-5:2008 Structured vocabularies for information retrieval. Guide. Exchange formats and protocols for interoperability. (2008). The British Standards Institution.

IFLA. (1998). *Functional Requirements for Bibliographic Records: Final Report*. IFLA Study Group on the Functional Requirements for Bibliographic Records. München: KG Saur. <http://www.ifla.org/publications/functional-requirements-for-bibliographic-records> (accessed May 15, 2009).

IFLA. (2008). *Functional Requirements for Bibliographic Records: Final Report*. IFLA Study Group on the Functional Requirements for Bibliographic Records. Approved by the Standing Committee of the IFLA Section on Cataloguing September 1997, As amended and corrected through February 2008. <http://www.ifla.org/publications/functional-requirements-for-bibliographic-records> (accessed May 15, 2009).

ISO/CD 25964-1 Information and documentation — Thesauri and interoperability with other vocabularies — Part 1: Thesauri for information retrieval. (2008). ISO TC 46/SC 9/WG 8. Ballot version.

OWL 2 Web Ontology Language Structural Specification and Functional-Style Syntax. (2009). Motik, B, Patel-Schneider, P.F. and Parsia, B. eds. W3C Working Draft 21 April 2009. <http://www.w3.org/TR/owl2-syntax/> (accessed May 15, 2009).

OWL 2 Web Ontology Language New Features and Rationale. (2009). Golbreich, C., Wallace, E.K. eds. W3C Working Draft 21 April 2009. <http://www.w3.org/TR/owl2-new-features/> (accessed May 15, 2009).

OWL 2 Web Ontology Language Document Overview. (2009). W3C OWL Working Group. W3C Working Draft 21 April 2009. <http://www.w3.org/TR/owl2-overview/> (accessed May 15, 2009).

OWL Web Ontology Language Reference. (2004). Dean, M. and Schreiber, G. eds. W3C Recommendation 10 February 2004. <http://www.w3.org/TR/owl-ref/> (accessed May 15, 2009).

Ranganathan, S.R., (1962). Facet analysis: Fundamental categories. In: Ranganathan, S.R. *Elements of Library Classification*. 3rd. ed., Bombay, New York: Asia Publishing House. pp. 82-89.

Rust, G. and Bide, M. (2000). The <indecs> metadata framework : Principles, model and data dictionary. Indecs Framework Ltd.

http://www.doi.org/topics/indecs/indecs_framework_2000.pdf (accessed May 15, 2009).

SKOS Simple Knowledge Organization System Reference. (2009). W3C Candidate Recommendation 17 March 2009; <http://www.w3.org/TR/2009/CR-skos-reference-20090317/> (accessed May 15, 2009).

VRA Core 4.0, (2007). Visual Resources Association Data Standards Committee.

<http://www.vraweb.org/projects/vracore4/index.html> (accessed May 15, 2009).

《书目记录的功能需求》最终报告（中文版），（IFLA Study Group on the Functional Requirements for Bibliographic Records），（1998）。王绍平等翻译；
<http://www.ifla.org/files/cataloguing/frbr/frbr-zh.pdf> (accessed July 19, 2009).