



FRBR 化：前景光明的国家书目

作者

Jan Pisanski (卢布尔雅那大学, 斯洛文尼亚)

Maja Žumer (卢布尔雅那大学, 斯洛文尼亚)

Trond Aalberg (挪威科技大学, 特隆赫姆, 挪威)

中文翻译: 蔡成普 (中国国家图书馆)

Chinese translator:

CAI Chengpu (National Library of China)

Meeting:

77. Bibliography

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>

摘要:

书目记录的功能需求(FRBR)概念模型开始得到认可。本文介绍目前 FRBR 模型表达的标准和实践的可能性和局限性。据称,国家书目从 FRBR 中受益最大,因为它们拥有很多高质量的书目数据,这些书目数据的作品许多是多版本的。本文部分以斯洛文尼亚国家书目和其他两家书目数据库的 FRBR 化结果为基础。

国家书目是一个国家知识产品的重要信息来源。然而,这些书目中的记录遵循了一些规则和实践,这些规则和实践最初是从老式卡片目录发展来的,而这限制了书目在数字环境下新应用的可用性和可再用性。为了提高国家书目的价值,针对书目资源信息,我们需要把它做成机器可解释的、可再利用的书目资源知识。其中一种可行的解决方法就是,采用书目记录功能需求(FRBR)概念模式。

FRBR

FRBR 概念模式是对书目世界的概括 (FRBR 1998)。FRBR 概念模式的基本元素，就是对当前书目记录进行逻辑分析的结果。他们由实体、实体属性和实体关系组成。第一个实体组是知识或艺术创作的产品，第二个实体组是内容、制作、传播或保管的责任者，第三个实体组是知识或艺术创作的主题。

实体组 1

知识或艺术创作的产品

- 作品—— 一个独特的知识或艺术创作
- 内容表达—— 作品的知识或艺术创作的表现方式：字母数字、音乐、舞蹈符号、声音、图像、实物、动作等，或任何这些形式的组合。
- 载体表现—— 作品内容表达的物理体现
- 单件—— 一种载体表现的单一样本

组实体 2

实体组 1 的内容、生产、传播或保管的责任者

- 个人—— 单独责任者
- 团体—— 一个组织和/或一组个人和/或机构组成的组织

实体组 3

知识或艺术创作的主题

当把 FRBR 与国家书目放在一起讨论时，我们需要注意的是，它并不是像国家书目目录那样不记录单件层次的信息。

作为一种概念模型，应当帮助使用者和专家更好地理解书目世界。虽然 FRBR 的术语可能并不常见，它代表一种必然的趋势，打破了不同概念使用同一种称谓的传统（最常见的例子是版本）。FRBR 为用户提供了更加用户友好的框架，

这种框架可以更好地展示书目数据。它能够避免书目数据库中的数据重复，还能建立顺序，并且能够借助关系，把线性的图书馆数据库变为真正的网络化环境，而这一点还能为读者提供更多相关信息。

对 FRBR 的一个最常见的误解就是，它是一个数据模型或一套编目规则，规定具体什么应该怎么做。它只不过是一个理论模型，需要在编目规则中加以阐释。并且，每一种实施将取决于系统的设计，这就意味着，除非有明确的共识规定 FRBR 的文化含义，但这是不可能的，所以该模型对于不同的领域或地区来说可以有不同的应用。正如 Maxwell (2008) 指出，我们没有任何“官方”宣传材料或实际经验，详细说明 FRBR 应当如何阐释。

然而，新英美编目条例的工作，或更确切地说，为数字世界设计的资源描述新标准，叫做资源描述与检索(RDA, <http://www.rdaonline.org/>)，把 FRBR 纳入考虑的范围。这是避免 FRBR 相关恶性循环的强有力的解脱 (Pisanski, Žumer, 2007)，并暗示着 FRBR 在图书馆界得到更多的赞赏。

把 FRBR 引入编目规则，一个重要因素就是 Carlyle (2006) 所称的“编目意识”。FRBR 应该帮助编目员更好的理解为什么他们要所做的，而这反过来还能引导更好地服务读者。虽然这个问题对共同目录更为紧迫，记录国家书目时，我们也必须牢记这一点。

RDA(资源描述与检索)遭到编目员的强烈反对，而计算机科学家和图书馆学理论家的看法通常与他们相左，这种反对的原因就是 RDA 把 FRBR 作为基础。事实上，美国国会图书馆的未来书目控制工作组曾建议停止 RDA 工作，除非 FRBR 的优点特别明显 (WoGroFuBiCo, 2008)。然而，RDA 工作未曾停止，预计 RDA 将在 2009 年通过 (Joint, 2008)。这表明，FRBR 的理论探索和实践探索，书目世界(下文即将提到)连同相关用户研究，比以往任何时间联系都紧密，尤其是因为仍没有任何正式用户研究的报告。第一批的理论用户研究表明，FRBR 普通非馆员心智模式的研究正好契合了 FRBR 的研究 (Pisanski, Žumer, 2008?)，而第一份对基于 FRBR 目录的内部用户的研究报告也显示出积极的结果 (Kilner (2005), Jepsen (2007))

还需要强调的是，在编目领域中，FRBR 和 RDA 并不是各自孤立发展。这个难题中更重要的一点是最近公布的国际编目原则声明，建立在旧的巴黎原则以及

FRBR 基础之上(2009 年报告)。

然而，即使 FRBR 通过进入新编目规则中，被引入国家书目，仍有问题存在：目前的格式不能充分表达 FRBR。尽管有很多 MARC 格式的变种，它们有一个重要的共性特征：没有简便的方法把 FRBR 纳入现有的 MARC 结构。还有个问题，就是大量的旧数据。看起来有两种可行的办法，可以把现有的数据并入 FRBR。其中一种可能是手工重编整个国家书目，这可能占用大量的时间和资源，所以，在现在的环境下并不可行；第二种方法就是从现有的书目数据中提取 FRBR 概念。第二种方法被称为 FRBR 化。

FRBR 化

作为唯一可行的解决方案，FRBR 应该被当作处理旧数据的先驱，而这些旧数据是由于 FRBR 发展而产生的。一些 FRBR 化的尝试(例如 Hegna and Murtooma, 2002; Hickey, O'Neill and Toves (2002), Hickey and O'Neill (2005))证明了它的可行性¹。然而，这些尝试也表明，想获得很高的质量是比较困难的。典型的问题包括不能充分地或者错误地确认实体和关系，以及不能充分地确认对应的实体。除此之外，许多 FRBR 化的举措只在部分 FRBR 模型中应用，或记录中部分信息的过程中。然而，完全的 FRBR 化可能带来更多问题，这些问题可能进一步降低结果的质量。需要指出的是，所有这些问题都源于现有的书目数据，以及它们的记录方式，如果这些书目数据按照一个更加结构化、更加一致的方式记录下来，FRBR 化数据时就会减少很多不能识别的问题。

虽然所有 FRBR 化算法(计算程序)的基本思想大致相同，把 MARC 记录的(子)字段与 FRBR 的实体匹配，实际产品在规格和细节方面可能发生变化，这部分归因于当地的编目考虑(例如使用格式、编目规则和标准的差异)，以及作者认为什么是可行的，什么是用户正在寻找的。

另一方面，由于受到当前原则的制约，(FRBR 化)实施的可能性会受到限制，比如 1961 年的巴黎原则，编目规则(如英美编目条例- AACR2)，数据格式(如 MARC21)，以及他们在实践中的应用，还有编目员曲解规则或犯下的其他错误。

¹ 事实上，已经有大量的 FRBR 化数据已经存在，尽管它们没有遵循 FRBR，想了解详情可以参看 Pisanski and Žumer (2007)。

使用“传统的”OPACs 时，这些问题通常被掩盖了，而使用“FRBR 化的”OPAC 时，它们就会突然显现。换句话说，如果书目记录的质量不一致，或者书目数据丢失，就很难自动提取 FRBR 模式或其他类型的结构化数据。因为许多 FRBR 的蓝本都是联合目录、一致性的数据，因而，这些原型的使用进一步减少。根据 Buchanan (2006)，依靠资源的类型和数据的质量，FRBR 实体中，多表达形式作品的自动检测的错误率高达 50%。

FRBR 与国家书目

Žumer 认为 (2007, 2008)，像联合目录和国家书目的情况一样，拥有数量庞大的作品，这些作品都有多个内容表达和载体表现，因而，引入一种概念化模式是必要的，并且，FRBR 的真正优势显而易见。同样地，根据新发布的数字时代国家书目指南 (2009)，对于 FRBR 化和与 FRBR 一致的原始编目而言，国家书目看起来是最好的起点。

有三个相互关联的原因。首先，通过大规模的书目数据，以及日积月累的且拥有大量数据的国家书目，FRBR 的一系列影响可以更好地展现。然而，这条原因自身并不能将国家书目推向比联合目录更好的位置，而其他两条原因则可以做到。

其次，因为国家书目往往比个体图书馆的目录更完整，FRBR 的影响应该更大。国家级重要的作品，FRBR 的使用应该仅次于完整的 FRBR 网络工作²，而这即使是在最大的联合目录中也不可能产生。

基于前两条原因，我们可以完全有把握的认为，逐渐积累的书目从 FRBR 的应用中受益最大。然而，人们希望是小单位(半月，月度，甚至是年度书目)，不要因实体组 1 的使用受到太大的影响。

然而，还有一个更重要的因素对国家书目有利，因为国家书目的编目水平一般都要比其他地方的水平高，从 FRBR 化获得有意义的数据的可能性就更大。国家书目的书目数据往往比联合目录的要好，由于采用更加严格的规则，它的一致性更强。相对于一些联合目录，国家书目还有一点优势就是，只使用一套编目规

² 虽然它的数量使记录完整的国家知识产品越来越不可行

则和标准，但在联合目录中，可能使用各种不同的编目规则和标准。

因此，毋庸置疑，由 Hegna 和 Murtomaa（2002）操作的第一个 FRBR 化实验在芬兰和挪威的国家书目进行。然而，他们发现，即使在那种相对有利的条件下，仍然有一些关于实体正确识别的问题。他们发现，尽管 FRBR 模型的元素一般都存在于书目记录中，编目原则不能很好的适应这个模型。他们还发现，重要的信息往往是按人容易理解的方式记录的，而不是按计算机易理解的方式。

虽然高水平的编目、相对完整的覆盖面和相对较高的一致性数据可能表明，国家书目像它们预期达到的那样完善，我们应当注意，一些秩序可能存在于这种相对完善的结构化数据中，相关信息可能还是缺失。不幸的是，许多这样的相关信息只能通过明确的关系，在不同的书目实体间才能获得，这些相关信息目前没有被记录，或者至少没有按一种容易被计算机处理的一致性的格式。

另外有一点需要说明的是，我们不能忘记 FRBR 的目标之一就是，为国家编目机构编制的记录提出一个基本的功能级别。国家书目服务国际研讨会(ICNBS)，1998年在哥本哈根举行，推荐使用 FRBR 最终报告提出的（国家书目）记录的基本级记录的组件。然而，这些建议的落实仍然取决于各国的国家书目机构，因此，我们不能期待这个问题有多少进展。

我们的发现

在最后一节我们将重点讨论 FRBR 化的结果和对 3 种不同的书目数据库 MARC 记录分析结果，这三个书目数据库是：挪威的 BIBSYS 数据库、瑞典的伯克 Burk 数据库和斯洛文尼亚国家书目（Aalberg 和 Žumer，2008）。这项研究是 NTNU（挪威特隆赫姆）和卢布尔雅那大学（斯洛文尼亚）联合工程的一部分。在这三个数据库中，只有斯洛文尼亚国家书目是国家书目，斯洛文尼亚国家书目是否适合国家书目数据库的要求，这在 FRBR 最终报告（Dimec, Žumer and Riesthuis (2005)）中已经被广泛讨论过。

NTNU 研发出一种 FRBR 化书目记录的工具，它可用于将 MARC 记录转化为 FRBR 记录的集合，而这些实体和关系只能通过输入才能得到。这些结果将通过手工检

查样品来评估。

这个工具以规则为基础，因此适用于特定集合，基于 MARC 格式的使用和其他特性，主要与编目工作有关。前文提及的三个数据库有三种各自独立的规则。基于这样的结果，我们可以说国家书目能被“FRBR 化”，因为涉及到旧数据的质量，常见问题是无法避免。

我们的研究结果支持理论看法：国家书目的编目通常比个体（普通）图书馆的目录好，因此，国家书目的 FRBR 化结果可能会优于一般图书馆的目录。例如斯洛文尼亚国家书目一贯记录所有的作者和贡献者连同适当的著作方式代码。平均来看，斯洛文尼亚国家书目记录中的责任者的数量是其他两个数据库中的两倍，这是国家书目记录采用了更严格方法的直接结果，相对于其他目录。著作方式代码，斯洛文尼亚书目一直保留，对 FRBR 内容表达的确定很有帮助。如果实体辨认过程中没有个人责任者或者团体责任者与内容表达相关联，内容表达只能在某个范畴被识别。这种情况的典型例子就是挪威和瑞典的数据库。这符合典型的 FRBR 作法，它根据语言、外形和格式，只提供成套内容表达。在斯洛文尼亚的数据库中，由于广泛使用著作方式代码，确定内容表达就能使用代理。这使得个体内容表达的确认为 FRBR 中界定的一样。由于 MARC 的结构与实体组 2 实体相关的数据往往不足以确定具体的关系，例如，常常无法确定一个人与作品或者与内容表达的关系。斯洛文尼亚书目不同于其他两个数据库，是因为它使用了 UNIMARC，并采用有包括著作方式代码在内的严格的规则。

斯洛文尼亚书目的一些原始题名出现在统一题名字段，其他的出现在附注字段。然而，由于部分使用了结构化的附注，后者相对容易确定。作为统一题名前缀介绍性词语“Prevod dela:”（作品的翻译）。然而，形态丰富的斯洛文尼亚语言（使用前缀和后缀）使一些更难以确认的作品（如一个印刷品中有一个以上的作品）。在其他两个数据库中，原始题名中可以重复的特定字段被频繁使用，因为所有的数据中，具有特定作品级题名的记录数量是比较少的。大约 80% 的记录中，题名是发现载体表现中所包含作品的唯一来源。

能从一条记录中辨认出一套实体只是 FRBR 化步骤的第一步。这项进程中比较难的就是确定演绎实体间的关系。如果所有重要的关系都被著录，这项工作就会比较容易。但是，令人沮丧的是，即使某些最明显的关系（一个作品和一个责

任者或者一个内容表达和一个责任者之间)有时也不容易确定。因为缺乏相关的著作方式代码或者类似问题。如前所述,挪威和瑞典的数据库都很少使用著作方式代码,然而,在斯洛文尼亚的国家书目中广泛使用著作方式代码。另一个问题就是,当一条记录中出现很多责任者和作品,或者内容表达时,就需要建立恰当的关系。MARC 格式,以及当前的编目工作,都没有真正地处理这样的结构性的信息,而这在 FRBR 中非常重要。但是,这不仅是格式的关键问题,在编目规则中,包括在 RDA 中,都应给予足够的重视。

此外, MARC 记录有一个适合出版物的结构,它包含单一作品的单一内容表达。含有多种内容表达的印刷品的编目工作尤为困难,在目录之间甚至是在同一条目录中都可能有一些各种各样的解决方法。

当然,基于某一个例子,推断国家书目是否适用 FRBR 化的作法是不可靠的。尤其是考虑到有些国家书目使用了不同的编目工具,可能会产生一些实际变化。

结论

采用一种概念模式,可以改变国家书目和其他书目数据库的本质,可能使它们更加用户友好,还会带来很多复杂的问题,大多源于已有书目数据记录的方式。然而,把旧数据变成与新数据匹配的唯一可行的方法就是 FRBR 化,

国家书目很是最适合 FRBR 化的书目数据库,并能从 FRBR 化获得最大利益。正因如此,它们可以成为 FRBR 化受益者的典型,可以用来说服反对者,证明书目世界概念模型的实际价值。

参考书目

Aalberg, T. and Žumer, M. (2008). Looking for Entities in Bibliographic Records. *Digital Libraries: Universal and Ubiquitous Access to Information, 11th International Conference on Asian Digital Libraries, ICADL 2008*. 327-330.

Buchanan, G. (2006). FRBR: Enriching and Integrating Digital Libraries. *Proceedings of the 6th ACM/IEEE-CS Joint Conference on Digital Libraries*. 260-269

Carlyle, A. (2006). Understanding FRBR as a Conceptual Model: FRBR and the Bibliographic Universe. *Library Resources & Technical Services*. 50 (4), 264-273.

Dimec, Z., Žumer, M. and Riesthuis, G. (2005). Slovenian Cataloguing Practice and *Functional Requirements for Bibliographic Records: A Comparative Analysis. Cataloging & Classification Quarterly*, 39 (3/4), 207-227.

(FRBR) *Functional Requirements for Bibliographic Records: final report* (1998). München: KG Saur.

Hegna, K. and Murtooma, E. (2002). *Data mining MARC to find: FRBR?* 68th IFLA Council and General Conference, August 18-24, 2002. Retrieved 21.11.2008 from: www.ifla.org/IV/ifla68/papers/053-133e.pdf

Hickey T., O'Neill, E. and Toves, J. (2002). Experiments with the IFLA Functional Requirements for Bibliographic Records (FRBR), *D-Lib Magazine*, 8 (9). Retrieved 5.1.2009 from: <http://www.dlib.org/dlib/september02/hickey/09hickey.html>

Hickey, T. and O'Neill, E. (2005). FRBRizing OCLC's WorldCat. *Cataloging & Classification Quarterly*, 39 (3/4), 239-251.

Jepsen, E. T. (2007). *Bibliographic relations*. Retrieved 21.6.2007 from: http://ru.is/kennarar/thorag/cataloguing2007/Erik_Thorlund_Jepsen.ppt

Joint Steering Committee for Development of RDA (2008). *RDA: Resource Description and Access*. Retrieved 19.1.2009 from: <http://www.collectionscanada.gc.ca/jsc/rda.html>.

Kilner, K. (2005). The AustLit Gateway and Scholarly Bibliography: A Specialist Implementation of the FRBR. *Cataloging & Classification Quarterly*, 39 (3/4), 87-102.

Maxwell, R. (2008). *FRBR: A Guide for the Perplexed*. Chicago: American Library Association.

National bibliographies in the digital age: Guidance and new directions (2009). In print.

Pisanski, J. and Žumer, M. (2007). Functional requirements for bibliographic records: an investigation of two prototypes. *Program*, 41 (4), 400-417.

Pisanski, J. and Žumer, M. (2008). How do non-librarians see the bibliographic universe? *Culture and identity in knowledge organisation: proceedings of the 10th International ISKO Conference, 5-8 August 2008, Montreal, Canada*. Würzburg: Ergon, 131-136.

(Statement) *Statement of International Cataloguing Principles* (2009). Retrieved 9.3.2009 from: http://www.ifla.org/VII/s13/icp/ICP-2009_en.pdf

(WoGroFuBiCo) *On the Record*. Retrieved 19.1.2009 from: <http://www.loc.gov/bibliographic-future/news/lcwg-ontherecord-jan08-final.pdf>

Žumer, M. (2007). FRBR: The End of the Road or a New Beginning. *Bulletin of the American Society for Information Science and Technology*, 33 (6), 27-29.

Žumer, M. (2008). FRBR... *Workshop on FRBR in The European Library, Lisbon, Portugal*.

Retrieved 3.3.2009 from: <http://frbr.bnportugal.pt/documentos/FRBR%20Zumer.ppt>