



Introducción de FRASAD y correspondencia con SKOS y otros modelos

Basado en el trabajo del FRASAR Working Group*

Marcia Lei Zeng
Kent State University
Kent, Ohio, USA
mzeng@kent.edu

Maja Zumer
University of Ljubljana
Ljubljana, Slovenia
maja.zumer@ff.uni-lj.si

Traducción:
Pascual Jiménez Huerta
Biblioteca Nacional de España

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>

Resumen:

El Functional Requirements for Subject Authority Records (FRSAR) Working Group [Grupo de Trabajo sobre Requisitos Funcionales para Registros de Autoridad de Materia] se formó en 2005, como el tercer grupo de trabajo de IFLA en el entorno FRBR, para abordar los aspectos relativos a los datos de autoridad de materia e investigar su utilización directa e indirecta de tales datos por un amplio rango de usuarios. Esta ponencia presenta el modelo FRASAD (Functional Requirements for Subject Authority Data), desarrollado por el FRASAR Working Group, y debatido en el contexto de otros modelos conceptuales relacionados definidos en las especificaciones durante los últimos años, incluyendo a las British Standard BS8723-5: Structured vocabularies for information retrieval – Guide Part 5: Exchange formats and protocols for interoperability [Vocabularios estructurados para la recuperación de información – Guía parte 5: Formatos y protocolos de intercambio para la interoperabilidad], SKOS Simple Knowledge Organization System Reference de W3C, y OWL Web Ontology Language Reference. Estos modelos permiten el estudio de las funciones de los datos de autoridad de materia y de los esquemas conceptuales a mayor nivel, que los hacen independientes de cualquier implementación, sistema o contexto específico, mientras que nos permite centrarnos en la semántica, la estructura y la interoperabilidad de los datos de autoridad de materia.

El Functional Requirements for Subject Authority Records (FRSAR) Working Group (WG) [Grupo de Trabajo de los Requisitos Funcionales para Registros de Autoridad de Materia] es el tercer grupo de trabajo de IFLA en el entorno de FRBR. Formado en Abril de 2005 tiene como tarea el desarrollo de un modelo conceptual de las entidades del grupo 3 de FRBR, dentro del marco de FRBR como ellos dicen en el “contenido” de las *obras*. Este informe nos presenta el Functional Requirements for Subject Authority Data (FRSAD), [Requisitos Funcionales para Datos de Autoridad de Materia] el modelo desarrollado por el FRSAR WG, y debatido en el contexto de otros modelos conceptuales relacionados desarrollados a lo largo del movimiento Linked Data de la Web Semántica durante los últimos años.

1.- Antecedentes

Las entidades del Grupo 3 de FRBR de la IFLA se utilizan como materias de las *obras* (es decir, los resultados del esfuerzo artístico o intelectual). “Representan un conjunto adicional de entidades que sirven como materias de las obras” (IFLA, 1999: 16 y IFLA, 2008: 17, comillas añadidas por los autores), sumadas a las entidades del Grupo 1 y 2, que también pueden ser materias de las *obras*. Las entidades del Grupo 3 de FRBR incluyen: *concepto* (una idea o noción abstracta), *objeto* (una cosa material), *acontecimiento* (una acción o suceso) y *lugar* (una localización) (IFLA, 1999). El Grupo de Trabajo FRSAR se formó para abordar los aspectos de los datos de autoridad de materia e investigar la utilización directa e indirecta de dichos datos por un amplio rango de usuarios. El papel de este Grupo se definió en los siguientes términos de referencia:

- Construir un modelo conceptual para las entidades del Grupo 3 dentro del marco de FRBR que relacione el contenido de las *obras*.
- Proporcionar un marco de referencia estructurado y claramente definido para relacionar los datos que son grabados en los registros de autoridad de materia con las necesidades de los usuarios de estos registros, y
- Ayudar en la evaluación del potencial de la utilización e intercambio internacional de datos de autoridad de materia entre el sector bibliotecario y otros sectores.

Se formaron 2 subgrupos de FRSAR. El subgrupo de “Tareas de Usuarios” se centró en los estudios sobre los propios usuarios y la definición de sus tareas. El subgrupo “Entidades de Materia” se centró principalmente en las entidades del Grupo 3, incluyendo tanto las entidades actuales del Grupo 3 de FRBR como las alternativas. Los subgrupos llevaron a cabo 2 estudios sobre tareas de usuarios y un pequeño experimento en 2006 y 2007. Se realizaron alrededor de 20 reuniones además de numerosos debates en línea. Se revisó un borrador del informe por los miembros del Grupo Asesor a finales de 2008 y principios de 2009. El informe revisado está listo para una revisión mundial antes de la Conferencia de IFLA en Milán de 2009.

2.- El Modelo FRSAD

Desde el momento en que se formó el Grupo de Trabajo de FRSAR, parece que había un acuerdo general en la comunidad de investigación de FRBR que las entidades del Grupo 3 se debían revisar (Delsey, 2005). Al principio, el Grupo de Trabajo FRSAR

tenía en mente simplemente mejorar el modelo existente basado en las entidades del Grupo 3 e ir incorporando los amplios debates que se estaban produciendo (por ejemplo “tiempo”). El Grupo investigó las propuestas de otros modelos existentes entre los que estaban: el modelo <indec> (Rust and Bide, 2000), las facetas de Ranganathan (Ranganathan, 1962), y la lista pragmática de entidades desarrollada por dos investigadores italianos, Buizza y Guerrini (2002). Estos modelos presentaban referencias sólidas para la revisión del modelo conceptual de FRBR. El Grupo de Trabajo analizó y debatió posibles soluciones basadas en cada uno de estos modelos, desde la solución más conservadora (realizando menores correcciones en el Grupo 3 de FRBR) hasta la más radical (proponiendo un modelo completamente nuevo). Sin embargo el Grupo de Trabajo encontró que ninguno de los modelos ni aquellos que estaban basados en los mismos podían ser suficientemente universales para reflejar las necesidades actuales de los datos de autoridad de materia, teniendo en cuenta especialmente diferentes ámbitos y las herramientas de acceso a las materias. En 2007, El Grupo de Trabajo FR SAR cambió su orientación al desarrollo de un nuevo modelo conceptual de las entidades del Grupo de 3 dentro del marco FRBR que relacionara el “contenido” de las *obras*.

En este marco, todos los puntos de acceso controlados relacionados con los tres grupos de entidad (como los define el modelo conceptual FRBR) tienen el potencial de ser la materia de una *obra*. En otras palabras, todas las entidades del Grupo 1, 2 y 3 pueden tener una relación de “es-materia-de” con la *obra*. El subgrupo “Materia de Entidad” de FR SAR propuso un modelo conceptual más abstracto y lo presentó en la Conferencia de IFLA de 2007. Como se ve en la siguiente Figura, el modelo se debería entender desde dos puntos de vista:

1. Este modelo confirma una de las relaciones básicas definidas en FRBR: **La OBRA tiene como materia el THEMA / El THEMA es materia de la OBRA.**
 - 1.1. **THEMA** es el término utilizado para referirse a algo que puede ser materia de una *obra*. Se define como “cualquier entidad de FRBR que se utiliza como materia de una *obra*”.
 - 1.2. **THEMA** incluye cualquiera de las entidades de FRBR – existentes en las entidades del Grupo 1 y 2 y, además, todas las materias de las *obras*. Aunque una entidad en sí misma puede ser vista como una super-entidad o super-clase de entidad de FRBR, nos permite realizar relaciones y atributos a un nivel más general y abstracto
2. Este modelo propone también una nueva relación: **El THEMA tiene como denominación el NOMEN. El NOMEN es denominación del THEMA.**
 - 2.1. **NOMEN** se define como cualquier signo o secuencia de signos (caracteres alfanuméricos, símbolos, sonidos, etc.) por los cuales un tema es conocido, referido o tratado. Por ejemplo, “amor”, “∞”, o “595.733”.

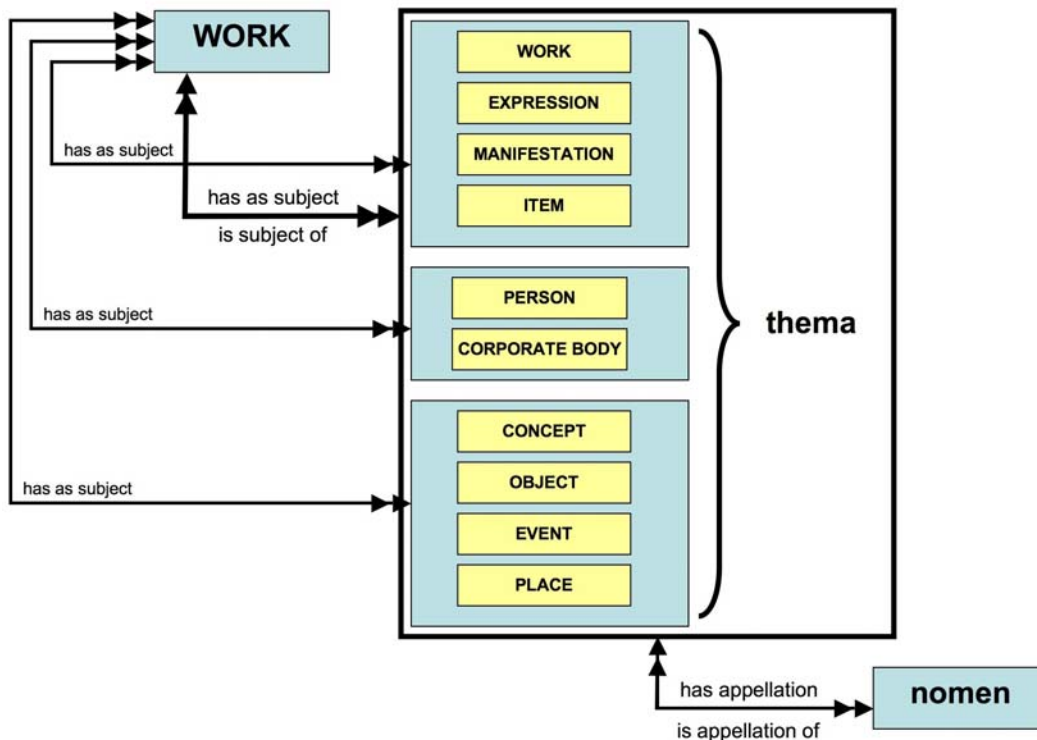


Figura 1. Relaciones de FRSAD con FRBR.

El modelo FRSAD se puede presentar de manera simplificada en la siguiente ilustración:



Figura 2. Modelo conceptual FRSAD.

Además de las relaciones entre entidades debemos explicar que:

- Las relaciones “tiene como materia/es materia de” son relaciones de varios a varios. Cualquier *obra* puede tener más de un *thema*; y cualquier *thema* puede ser materia de más de una *obra*.
- En general, (es decir, en el lenguaje natural o cuando contrastamos diferentes vocabularios) la relación de “tiene como denominación/es denominación de” es una relación de varios a varios. Un *thema* tiene uno o más *nomens* y puede haber un *nomen* referido a más de un *thema*.

- Es importante advertir que, en un vocabulario controlado dado e incluido en un dominio, un *nomem* debería tener como denominación un único *thema*.

Podemos tomar como ejemplo “Una breve historia del tiempo: del big bang a los agujeros negros” de Stephen W. Hawking. La *obra* tiene varios *themas*: cosmología, espacio y tiempo, unificación de la física, agujeros negros, big bang, historia del tiempo, universo. Hay muchas otras *obras* con esos *themas*. Para cualquiera de los *themas* de la lista, los *nomems* que aquí se presentan son términos en español¹, pero hay otros posibles *nomens* en otros idiomas o formados de acuerdo a otros vocabularios controlados como listas de encabezamientos de materia, sistemas de clasificación, etc.

En el borrador del informe FRSAD se han debatido los aspectos de (a) la complejidad y particularidad de los *themas*, (b) los atributos de *thema* y *nomen*, y (c) las relaciones entre *themas*, *nomens* y *thema-nomen*.

La importancia del modelo THEMA-NOMEM para los datos de autoridad de materia radica en separar los *themas* (o “materias”, “conceptos”, “clases (de conceptos)”, etc.) de lo que nos es conocido, referido o tratado. Entre los esfuerzos para alcanzar el intercambio y utilización a nivel global de los datos de autoridad de materia, éstos se han centrado en *nomens* (por ejemplo, un vocabulario traducido de metadatos, un tesoro simétrico multilingüe, o un índice de acceso múltiple a un vocabulario). Sin embargo, la mayoría de los esfuerzos se han centrado a un nivel conceptual, por ejemplo, la correspondencia entre dos tesauros o entre un esquema de clasificación y un tesoro. Normalmente tales esfuerzos conllevan grandes cambios ya que les afectan tanto a los *themas* como a las relaciones entre los *themas*.

3.- Correspondencia de FRSAD con SKOS y otros modelos.

El punto final de referencia del Grupo de Trabajo de FRSAR es ayudar en una evaluación del potencial de intercambio y utilización de los datos de autoridad de materia entre el sector bibliotecario y otros sectores. Los desafíos en el intercambio real afectan a muchos aspectos: estructuras heterogéneas, varios idiomas y escrituras, diferente construcción de reglas y prácticas, y esquemas codificados que se han desarrollado de forma dinámica. Una comparación preliminar de FRSAD con otros modelos nos permitirá examinar a mayor nivel que es independiente de cualquier implementación, sistema o contexto específico, y nos permitirá centrarnos en la semántica, estructura e interoperabilidad.

BS8723 e ISO 25964: Se definió un modelo para vocabularios estructurados (más específicamente, un tesoro) por la British Standard BS 8723-5: *Structured vocabularies for information retrieval – Guide Part 5: Exchange formats and protocols for interoperability* [Vocabularios estructurados para la recuperación de información – Guía. Parte 5: Formatos de intercambio y protocolos para la interoperabilidad] (DD 8723-5:2008). (El modelo, XML Schema y los ejemplos

¹ “En inglés” en el original (N. del T.)

están disponibles en la BS8723 Official Development Website). Este modelo ha sido ligeramente revisado e incluido en el ISO/CD 25964-1 *Información y Documentación – Tesoros e interoperabilidad con otros vocabularios – Parte 1: Tesoros para la recuperación de información* (2008:92), los cuales salieron por votación a comienzos de 2009. Incluye cuáles son las necesidades de un modelo: (1) un tesoro general, (2) una selección de conceptos de un tesoro, y (3) registros que documenten la entrada del tesoro. En el modelo, cada concepto del tesoro se representa por un término aceptado por el lenguaje, y por un número indeterminado de conceptos no aceptados. Las notas, notas de uso y relaciones entre términos genéricos, específicos y relacionados, se aplican al concepto como un todo, más que sólo con su término aceptado. Se asigna un único identificador a cada concepto (ISO/CD 25964-1, 2008). De forma global, este modelo y el modelo FRSAD representan estas relaciones: (1) *thema* y *nomen* (un registro que documenta un concepto y su(s) *nomen(s)*). (2) *thema-y-thema* (relaciones jerárquica (concepto genérico, específico y concepto general) y asociativa (conceptos relacionados), y (3) *nomen-y-nomen* (aceptado y no aceptado, formas léxicas variantes, y en varios idiomas).

SKOS: *SKOS, Simple Knowledge Organization System Reference* (2009) [Sistema de Referencia de Organización del Conocimiento Simple] define de forma bastante clara clases y propiedades para representar los aspectos comunes encontrados en un sistema de organización del conocimiento, como un tesoro, una taxonomía, una lista de términos controlados u otras estructuras de organización del conocimiento. “Utilizando SKOS los conceptos pueden ser identificados usando URIs, etiquetados con cadenas léxicas en uno o más lenguajes naturales, anotados (con códigos léxicos), documentados con varios tipos de notas, enlazados a otros conceptos y organizados en jerarquías informales y redes de asociación, agregados a esquemas conceptuales, agrupados en colecciones ordenadas y/o etiquetadas, y contrastados con otros conceptos en otros esquemas” (SKOS Reference, 2009: Sinopsis). Como en una aplicación de RDF (Resource Description Framework), SKOS permite redactar y publicar los conceptos en la World Wide Web, enlazarlos a otros datos en la Web e integrarlos en otros esquemas conceptuales. Cada concepto de SKOS está definido como un *recurso* RDF y puede tener *propiedades* RDF adjuntas. El modelo SKOS está basado en una concepción del vocabulario como punto central, donde los objetos elementales no son las etiquetas, sino más bien, los conceptos representados por las etiquetas. Estos se pueden ajustar a lo que se ha definido en el modelo FRSAD, con los términos de *thema*, *nomen* y sus atributos. SKOS tiene también propiedades específicas para representar todas las relaciones semánticas, que se ajustan a las definidas por FRSAD.

OWL: OWL es un lenguaje de ontologías para la Web Semántica con un significado formalmente definido. Las ontologías son vocabularios normalizados de términos (clases y propiedades), que cubren un área específica y son compartidos por una comunidad de usuarios. Especifican las definiciones de los términos describiendo sus relaciones con otros términos de la ontología (OWL 2 Web Ontology Language Document Overview, 2009). Las ontologías de OWL proporcionan clases, propiedades, personas y valores de datos, y son guardados como documentos de Web Semántica. OWL 1 estaba orientado principalmente a los constructos para expresar información de clases y personas. OWL 2, el más reciente borrador de trabajo de W3C, ofrece nuevos constructos para la expresión de

restricciones adicionales sobre propiedades, nuevas características de las mismas, incompatibilidades, cadenas y propiedades clave (OWL 2 Web Ontology Language New Features and Rationale, 2009). OWL 2 proporciona axiomas (declaraciones que expresan lo que es verdad en el dominio) que permite establecer relaciones entre expresiones de clases, incluyendo: “SubClassOf”, “EquivalentClasses”, “DisjointClasses” y “DisjointUnion”. En OWL todavía más importante es la utilización de las clases y expresiones de las propiedades para construir expresiones de clases, (a veces también llamadas descripciones, y, en la descripción de la literatura lógica, conceptos complejos). Esto se mantiene para la enumeración de personas y todos los operadores booleanos estándar AND, OR y NOT. Las expresiones de clase “ObjectIntersectionOf”, “ObjectUnionOf” y “ObjectComplementOf” proporcionan un conjunto teórico normativo de operaciones sobre expresiones de clase. La expresión de clase “ObjectOneOf” contiene exactamente las personas especificadas (OWL 2 Web Ontology Language Structural Specification and Functional-Style Syntax, 2009). OWL ofrece grandes oportunidades para los aspectos de la complejidad y particularidad de *themas* y la comprensión de relaciones semánticas entre *themas* que FRSAD intenta cubrir.

Modelo Abstracto DCMI: Cuando el DCMI Abstract Model (DCMI AM) [Modelo Abstracto DCMI] llegó a Recomendación DCMI en 2007, se había reconocido el principio “uno-a-uno” (por ejemplo, cada descripción de metadatos DC describe uno y sólo un recurso) y fue seguido por otras normas de metadatos, como por ejemplo, el VRA Core 4.0 producido por la Visual Resources Association en 2007 (VRA Core 4.0. 2007). Según el modelo DCMI, un *registro* puede contener un *conjunto de descripciones*, que a su vez pueden contener *descripciones* compuestas de *declaraciones*, que utilizan pares de *propiedades-valores* (DCMI Abstract Model 2007). Esto se traduce en información que puede ser procesada, intercambiada, referida y enlazada a nivel de declaración. Cuando un registro contiene descripciones de un recurso, las descripciones individuales también pueden ser enlazadas a los datos de autoridad que manejan los valores asociados con esas propiedades (por ejemplo, los datos de autoridad de materia, las propiedades de los datos de autoridad de nombre, o los datos de autoridad geográficos). Este modelo de información es independiente de cualquier sintaxis codificada y facilita el desarrollo de mejores mapeos y traducciones a través de la sintaxis. El modelo FRSAD se corresponde al DCMI Abstract Model permitiendo que cualquier *thema* sea independiente de cualquier *nomen*, incluyendo una sintaxis que el *nomen* debe usar. Además este modelo conceptual facilitará el intercambio y reutilización de datos de autoridad de materia no sólo entre los vocabularios de materia entre sí, sino también en los recursos de metadatos.

En conclusión, el modelo FRSAD se desarrolla con el objeto de asistir en una evaluación del potencial para el intercambio y utilización de datos de autoridad de materia tanto dentro como fuera del sector bibliotecario. El modelo FRSAD y otros modelos desarrollados junto con el progreso de la Web Semántica durante los últimos años permiten el examen de las funciones de los datos de autoridad de materia y esquemas conceptuales a un nivel superior que lo hace independiente de cualquier implementación, sistema o contexto específico, y nos permite centrarnos en la semántica, estructura e interoperabilidad de los datos de autoridad de materia. Colocando a los datos de autoridad de materia en el contexto de los desarrollos de la Web Semántica, especialmente en la perspectiva de Linked Data, los datos de

autoridad de materia que se están desarrollando basados en FRSAD y codificados en SKOS y OWL llegarán a formar parte de Linked Data y contribuirán al desarrollo de la Web Semántica.

Referencias:

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 <indec> metadata framework : Principles, model and data dictionary. Indecs Framework Ltd. http://www.doi.org/topics/indec/indec_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).

Agradecimientos*: Este informe está basado en el trabajo de del FRSAR Working Group y en las sugerencias del Advisory Group, establecido por la IFLA Division IV. Control Bibliográfico y especialmente la Sección de Clasificación e Indización. Los miembros del Working Group son: Leda Bultrini, Lois Mai Chan, Jonathan Furner, Edward O'Neill, Gerhard Riesthuis, Athena Salaba, Diane Vizine-Goetz, Ekaterina Zaytseva, Marcia Lei Zeng y Maja Zumer. Los miembros del Advisory Group son: Victoria Francu, Hemalata Iyer, Dorothy McGarry, David Miller, Päivi Pekkarinen y Barbara Tillet. La comunicación e intercambio con el Grupo de Trabajo FRANAR presidido por Glenn Patton ha sido también de gran ayuda. También han proporcionado soporte, financiación e instalaciones IFLA, OCLC, y la Kent State University.