

NAME AUTHORITY CONTROL FROM CONCEPTUAL MODEL TO THE SEMANTIC WEB AND LINKED DATA

MIRNA WILLER

Department of Information Sciences, University of Zadar, Croatia

ABSTRACT

The article focuses on the field of name authority control, specifically on personal names with the aim of describing the process of representing FRAD (*Functional Requirements for Authority Data*) conceptual model in RDF (*Resource Description Framework*), the W3C standard in order to research into the issues of its functionality in the Semantic Web and linked data environment. The practical example about the issues of expressing authority data in RDF triples is shown, while some problems encountered in this process are recognized.

KEYWORDS

name authority control, FRAD: *Functional Requirements for Authority Data*, RDF: *Resource Description Framework*, Semantic Web, linked data

Introduction

Libraries maintain and provide for their users verified and authorized metadata, the value of which is authenticated by the internationally agreed standards and national cataloguing rules. Such approach to serving users' needs should be continued by designing the next generation library services – services adapted to the Semantic Web and linked data environment. In order to meet this task it is necessary to represent bibliographic standards in the tools and standards of that environment – in the *de facto* standard of the Semantic Web, World Wide Web Consortium's (W3C) RDF (*Resource Description Framework*).¹

1 W3C. RDF: Resource Description Framework Primer, W3C Recommendation 10 February 2004 [cited: 2012-03-29]. Available at: <http://www.w3.org/TR/rdf-primer/>

The research in the field of representing IFLA's² bibliographic standards in RDF started in 2007 by its Cataloguing Section's FRBR Review Group.³ The conceptual model for bibliographic data – FRBR⁴ was the first to be considered; it was followed by the two remaining FR family of conceptual models – FRAD for authority data⁵ and FRSD for subject authority data.⁶ In 2008 the ISBD Review Group's ISBD/XML Study Group,⁷ of the same IFLA's section, started the project on representing ISBD, the standard for bibliographic description in RDF. By the end of 2011 all these standards were published,⁸ while as of February 2012, the ISBD⁹ and FRBR namespaces are providing de-referencing services to their individual class and property level. Some of the published ISBD and FR family model's namespaces are already in use in different library services which release bibliographic data as Open Linked Data.¹⁰

This article will focus on the field of name authority control, specifically on personal names with the aim of describing the process

- 2 IFLA: International Federation of Library Associations and Institutions [cited: 2012-03-29]. Available at: <http://www.ifla.org>
- 3 IFLA. Cataloguing Section. FRBR Review Group [cited: 2012-03-29]. Available at: <http://www.ifla.org/en/frbr-rg>
- 4 Functional requirements for bibliographic records : final report / IFLA Study Group on the Functional Requirements for Bibliographic Records. München : K. G. Saur, 1998. Current text including amendments and corrections to date, available at: <http://www.ifla.org/en/publications/functional-requirements-for-bibliographic-records> [cited: 2012-03-29].
- 5 Functional requirements for authority data : a conceptual model / edited by Glenn E. Patton ; IFLA Working Group on Functional Requirements and Numbering of Authority Records (FRANAR). Final report, December 2008 / approved by the Standing Committees of the IFLA Cataloguing Section and IFLA Classification and Indexing Section, March 2009. München : K. G. Saur, 2009. See also: Functional requirements for authority data (FRAD) : modifications and errata for the 2009 text, November 2011 at: <http://www.ifla.org/files/cataloguing/frad/FRADerrata2011.pdf> [cited: 2012-03-29].
- 6 Functional requirements for subject authority data (FRSD) : a conceptual model / IFLA Working Group on Functional Requirements for Subject Authority Records (FRSAR) ; editors Marcia Lei Zeng, Maja Žumer, Athena Salaba. Berlin ; München : De Gruyter Saur, 2011. Version, June 2010 available at: <http://www.ifla.org/files/classification-and-indexing/functional-requirements-for-subject-authority-data/frsd-final-report.pdf> [cited: 2012-03-29].
- 7 IFLA. Cataloguing Section. ISBD Review Group. ISBD/XML Study Group [cited: 2012-03-29]. Available at: <http://www.ifla.org/en/node/1795>. The list of presentations related to ISBD in RDF, as well as detailed reports of the Study Group's activities see at this web page.
- 8 Dunsire, Gordon; Mirna Willer. Standard library metadata models and structures for the Semantic Web. // Library hi tech news 28, 3(2011), 1-12. Earlier version available at: <http://www.ifla.org/files/hq/papers/ifla76/149-dunsire-en.pdf> [cited: 2012-03-29].
- 9 IFLA. Cataloguing Section. ISBD Review Group and ISBD/XML Study Group. ISBD namespaces published, 8 February 2012 [cited: 2012-03-29]. Available at: <http://www.ifla.org/en/news/isbd-namespaces-published>
- 10 See: University Library Mannheim, and national libraries of Germany, UK, France and Spain.

of representing FRAD conceptual model in RDF in order to research into the issues of its functionality in the Semantic Web and linked data environment. At the moment, this is the only IFLA standard for name authorities¹¹ which is represented in RDF. Namely, the work on representing *Guidelines for Authority Records and References*,¹² IFLA's document on which FRAD conceptual model is based, has not been considered, while the work on representing UNIMARC format for machine readable cataloguing of authority data¹³ has not started yet, although some general directions for representing both UNIMARC authorities and bibliographic data formats in RDF have been researched and proposed.¹⁴

Background: paradigm shift in name authority control

IFLA's work on standardizing content of bibliographic records and procedures for their functioning within the local catalogue, and for their exchange for the purpose of building local/national as well as multinational cooperative catalogues started in 1961 with the *International Conference on Cataloguing Principles* that was held in Paris.¹⁵ The concept of Universal Bibliographic Control (UBC) was adopted at the conference. The concept implied that a bibliographic entity of inter-

- 11 The term name authorities is used here to cover personal, corporate body and family names, as well as titles which are considered as access points, and taken into consideration by the FRAD conceptual model. It should be noted that IFLA's UNIMARC format for authority data includes also other types of access points (entities) such as Trademark, Printer/Publisher Device, Place and Date of Publication, Performance, Provenance, etc., Collective title, etc.
- 12 Guidelines for authority records and references / recommended by the Working Group on an International Authority System ; approved by the Standing Committee of the IFLA Section on Cataloguing and the IFLA Section on Information Technology. 2nd ed. München : K. G. Saur, 2011. Note that the relation of this document to FRAD is parallel to the one between ISBD and FRBR.
- 13 UNIMARC manual : authorities format / edited by Mirna Willer. 3rd edition. München : K. G. Saur, 2009.
- 14 Dunsire, Gordon. UNIMARC, RDA and the Semantic Web. // International cataloguing and bibliographic control (ICBC) 39, 2(April/June 2010), 37-40. Based on a paper presented to the World Library and Information Congress: 75th IFLA General Conference and Assembly, 23-27 August 2009, Milan, Italy. Available at: www.ifla.org/files/hq/papers/ifa75/135-dunsire-en.pdf [cited: 2012-03-29]; Dunsire, Gordon; Mirna Willer. UNIMARC and Linked Data. // IFLA Journal 37, 4(December 2011), 314-326. Also available at: http://www.ifla.org/files/hq/publications/ifla-journal/ifla-journal-37-4_2011.pdf [cited: 2012-03-29].
- 15 International Conference on Cataloguing Principles, Paris, 9th-18th October, 1961: Report. London : International Federation of Library Associations, 1963. Reprinted in: International Conference on Cataloguing Principles, Paris, 9th-18th October, 1961: Report. London : Clive Bingley, 1969.

est to the user of a bibliographic service should be represented by *one form of a name*, while that form of name should be established by the bibliographic agency of the entity's origin. That form of name was labelled *uniform heading*, while all other forms should be considered as variants, and should point to the uniform one. The aim of the UBC is, therefore, to enable any user anywhere in the world in any of bibliographic sources (e.g., catalogue, bibliography, list) to find all works (in whatever manifestation) by the chosen entity *collocated under that one form of name*.¹⁶

Why is an entity represented by different forms of names? The annotated edition of the *Statement of Principles* prepared by Eva Verona and associates (1971) gives a list of causes due to which variations in a personal author's name or a form of name may arise. The examples of these are: "variant spellings of a name, different romanizations of a name originally not written in the roman script, different phonetic transcriptions, different linguistic forms, use of complete and incomplete forms, change of status, arbitrary or legal change of name or form of name, use of pseudonyms, nicknames, clandestine names assumed for political activities or other assumed names, generic appellations, etc., and use of the title of another work".¹⁷ Thus, the objective of the bibliographic information organization is to recognize the causes of different forms of names, categorize them, and provide the method or mechanism to link them functionally in a particular catalogue to meet the *collocation function* of the catalogue.

Although the international agreement and published guidelines, lists and rules were declaratively accepted by the national bibliographic agencies, the real world requirements proved that strict adherence to the principles of the UBC was not sustainable. The reason for such a development can be found not only in the diversity of resources in different languages, scripts, and therefore cultural and publishing variants that become part of library collections, and which have to be recorded according to the needs and requirements of the local user (e.g., known form, or language form of a name), but also according to technical possibilities of the local system to record, index, and enable search and

16 Anderson, Dorothy. *Universal Bibliographic Control : a long term policy, a plan for action*. Pullach/ München : Verlag Dokumentation, 1974.

17 *Statement of Principles adopted at the International Conference on Cataloguing Principles Paris, October, 1961*. Annotated ed. with commentary and examples by Eva Verona assisted by Franz Georg Kaltwasser, P. R. Lewis, Roger Pierrot. London : IFLA Committee on Cataloguing, 1971. Pp. 9-11.

retrieval of specific scripts in which the original name is written. The change of the technical form of the catalogue from the card catalogue to WebPAC exposes such local solutions to network environment and global use and re-use, and their specificities become apparent more than when the catalogue served its local, relatively well defined audience. When a record becomes a part of a cooperative, often international catalogue whether physical or virtual, the differences in the treatment of forms of names as authorized or variant access points become apparent, and the need to put those forms of names for the same entity into some kind of relation becomes imperative.¹⁸

The solution to such a situation was found by two international projects which have by now turned into fully developed services: the CERL Thesaurus¹⁹ which focuses on European printed heritage in the hand-press period (up to *c.* 1830), and VIAF: Virtual International Authority File designed primarily by the national libraries and the OCLC as the technical and logistic support to this physical, cooperative authority database.²⁰ The functionality of universal name authority control as envisaged in 1961 and onward as bilateral exchange, has evolved into exposing authority data²¹ to interested parties, each of which is enabled to consult a variety of sources, and re-use a part or whole of a record, or a mash-up of data which better suits their needs and their local user requirements.

The evolution of the concept of UBC from the exchange to exposing library data can be seen as co-current at the conceptual level with the development of what is known as the Semantic Web and linked data.²² The Semantic Web is defined simply as a “web of data”,

- 18 See more on these issues in: Willer, Mirna. Name authority control paradigm shift in the network environment. // Frameworks for ICT policy : government, social and legal issues / edited by Esharenana E. Adomi. Hershey, PA : IGI Global, 2010. Pp. 182-205. DOI: 10.4018/978-1-61692-012-8.ch012
- 19 Consortium of European research Libraries. CERL Thesaurus [main information at:] http://www.cerl.org/web/en/resources/cerl_thesaurus/main; [access at:] <http://thesaurus.cerl.org/cgi-bin/search.pl> [cited: 2012-03-29].
- 20 VIAF: Virtual International Authority File [main information at:] <http://www.oclc.org/research/activities/viaf/>, [service available at:] <http://viaf.org/>. As of 4 April 2012, VIAF has transitioned from an OCLC Research prototype to an OCLC service; see: <http://www.oclc.org/viaf/default.htm> [cited: 2012-04-04].
- 21 The term which was used as opposed to *exchange* was *sharing*, but as that term still implies two parties in a process (direction one-to-many) regardless of its reiteration, the term *exposing* seems to better express the current situation.
- 22 Bizer, Christian; Tom Heath; Tim Berners-Lee. Linked data – the story so far. // International Journal on Semantic Web and Information Systems (IJSWIS) 5, 3(2009). Pre-print available at: <http://tomheath.com/papers/bizer-heath-berners-lee-ijswis-linked-data.pdf> [cited: 2012-03-29].

that is, it is “about two things. It is about common formats for integration and combination of data drawn from diverse sources, where on the original Web mainly concentrated on the interchange of documents. It is also about language for recording how the data relates to real world objects. That allows a person, or a machine, to start off in one database, and then move through an unending set of databases which are connected not by wires but by being about the same thing.”²³ Therefore, in order to enable the functioning of this “web of linked data” it is necessary to build ontologies, that is, “collections of statements written in a language such as RDF that define the relations between concepts and specify logical rules for reasoning about them. Computers will “understand” the meaning of semantic data on a Web page by following links to specified ontologies.”²⁴ IFLA’s work on building such collections of statements – that is, representations of its standards and conceptual models in RDF has been done. This is the first step in the process of building functional linked data services; the next one would be, looking from the name authority data requirements, to build links between and among authority data which point to the same bibliographic entity of interest to the user but is defined by specific/different rules applied by different bibliographic agencies. Therefore it is important to form an agreement on the type of entities of interest within the bibliographic universe the libraries have a mandate to organize, but even more so, to agree on relationships, that is, their value and granularity that are valid among these entities. The FRAD conceptual model meets these requirements as the epitome of the theory and practice of the hundred and seventy years of what is considered to be the contemporary librarianship. The remaining part of this article will present and research into issues of representing FRAD in RDF, the Semantic Web standard, specifically from the aspect of linking between and among authority data.

23 W3C Semantic Web Activity [cited: 2012-03-06]. Available at: <http://www.w3.org/2001/sw/>

24 Berners-Lee, Tim; James Hendler; Ora Lassila. *The Semantic Web*. // *Scientific American* 17 May, 2001. P. 32. See also: Berners-Lee, Tim (with Mark Fischetti). *Weaving the Web : the past, present and future of the World Wide Web by its Inventor*. London : Orion Business Books, 1999.

FRAD in RDF

Introduction, or raison d'être

If we, for the moment, put aside the need of the library community to *represent* its standards not any more in the textual or machine readable exchange format (ISO 2709) forms, but in the form or a scheme required by the advancing technology, and focus on the broader scene to which bibliographic data will be *exposed*, libraries have to make sure that these data will be recognized as authoritative and authentic by the possible users. The users of the open linked data cloud²⁵ come from the most varied communities and with the most varied information needs,²⁶ and the success of libraries will be measured not only in serving their local users, cooperating in national/international catalogues, but also in “being there” for patrons unrecognized, and unpredicted by librarians.

For example, if one searches for “hieronymus sanctus” in Google, he or she will get 204,000 results (in 0.17 seconds) (Figure 1). The results presented on the first page already show a variety of sources that compete with those from libraries. The first one is the web page of “documentacatolicaomnia”, the second one Wikipedia in Latin (la.wikipedia.org), followed by “oldedocuments” (16th century English notes in the **Sanctus Hieronymus** (1513) at the VPL), and “vialibri” (The World’s Largest Marketplace for Old, Rare & Out-of-print Books), while the last result on the first page points to the (obscure) page called “worldcat.org/identities”. If followed, the link will point to the beta version of the OCLC project of the same name – WorldCat Identities, the project which researches into integrating retrieval of the authority file with bibliographic one – the WorldCat bibliographic cooperative database/catalogue. Which one of these services will the user approach and use

25 Linked Data. Connect distributed data across the Web [cited: 2012-04-18]. Available at: <http://linkeddata.org/>

26 Consider in this respect Chris Anderson’s seminal article “The Long Tail”, and his following work. Anderson describes a business model for the digital age, in which he talks about how the Web changed the economy and therefore the culture which shifts away from “the tyranny of lowest-common-denominator fare, subjected to brain-dead summer blockbusters and manufactured pop. Why? Economics. Many of our assumptions about popular taste are actually artifacts of poor supply-and-demand matching – a market response to inefficient distribution.” The answer to this is the Long Tail of niches of resources, that can be found “out there”, and that can meet the most varied users’ needs. Anderson, Chris. The Long Tail. // *Wired* 12.20(October 2004) [cited: 2012-03-06]. Available at: <http://www.wired.com/wired/archive/12.10/tail.html>. See for further reading Mossman, Katherine. Serving the niche : viewing libraries through Chris Anderson’s & Long Tail’s lens. // *Library journal* 07/15/2006 [cited: 2012-03-06]. Available at: <http://www.libraryjournal.com/article/CA6349032.html>; Storey, Tom. The Long Tail and libraries. // *OCLC Newsletter* April/May/June 2005, 6-11.

to meet his or her specific needs is part of the spirited debate going on for almost fifteen years now since the arrival of the World Wide Web.²⁷ It is good that people engage in spirited and enduring debates because they show that something is at stake, but the problem here is that the future of libraries is at stake, and if they want to continue performing their task and mission in guarding cultural heritage and disseminating knowledge they have to evolve their services.

Google search: “out there”!

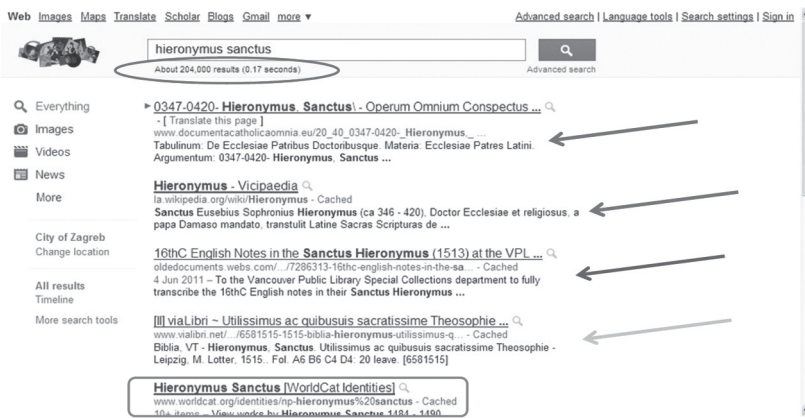


FIGURE 1.

Goggle search for “hieronymus sanctus” performed at 14th September 2011

RDF, namespaces and vocabulary management platform for publishing standards and models in RDF

RDF statement is defined as a *triple* statement. The structure of a triple is: subject – predicate – object. The subject identifies what the statement is about, such as the *thing* in general, e.g., a resource – a book, or an entity – author of a book. The predicate identifies the specific aspect of the subject being described, that is, the relation between subject and object, such as a resource *has edition statement*, or an entity *has*

27 See for example: Calhoun, Karen. The changing nature of the catalog and its integration with other discovery tools : final report, March 17, 2006 : prepared for the Library of Congress [cited: 2012-03-06]. Available at: <http://www.loc.gov/catdir/calhoun-report-final.pdf>; The Library of Congress Working Group on the Future of Bibliographic Control. On the record : report, January 9, 2008 [cited: 2012-03-06]. Available at: <http://www.loc.gov/bibliographic-future/news/lcwg-ontherecord-jan08-final.pdf>

appellation (name). The *object* identifies or presents the value of that aspect, such as “3rd ed.”, or “Hieronymus Sanctus”. The object can be a value which is identified by a controlled vocabulary, such as a personal name authorized by a specific authority file, e.g., “Hieronymus, Sophronius Eusebius” authorized by The German National Library, or a literal string, such as “3rd ed.” or a personal name without external authentication. In the case when the value of the object is identified, the object becomes the subject of another triple thus forming, in this example, two linked data triples.

RDF requires that subject and predicate of a triple be expressed as URIs: Uniform Resource Identifiers,²⁸ while object can be a literal or a string of characters for which no semantic value is expressed, or another URI. In the first case of our example, “3rd ed.” is a literal, while in the latter the object is represented by a URI the semantic value of which is expressed by the identifier of The German National Library’s record in which authorized form of a person’s name is established, i.e. “Hieronymus, Sophronius Eusebius”.

The subject of a RDF triple statement, or the *thing* the statement is about, may be defined by its type as a *class*. In the FR family model the *class* corresponds to FR entities, or the key objects of interest to users, such as work, person, family, etc. The predicate or verb phrase of a triple is expressed as a *property*, and it corresponds to entity attributes and relationships of the FR family model. Each property may declare its *domain*, that is, the class (subject) to which it belongs, and *range*, that is the class of the object or the value of the property. For example, FR attribute *title of person* is represented in an RDF triple as a property with the class *Person* declared as its domain, while the range is not defined because this property’s value can be a literal, that is, a personal name, or an identifier (URI) from an authority system. The FR relationship represented as a property, however, usually has to define both the domain and the range because FR family models define them as bi-directional between two entities. It is necessary therefore to express domain for the subject, and range for the object of such a property in a RDF triple sentence. For example, FRAD relationships *has appellation* and *is appellation of* are defined between entities *bibliographic entity* and *name*; the domain defined for RDF property *hasAppellation*²⁹ is class (subject)

28 W3C. Web naming and addressing : URIs, URLs, ... [cited: 2012-03-06]. Available at: <http://www.w3.org/Addressing/#rfc3986>

29 Notation used for the name of a class and property as represented in RDF in Open Metadata Registry.

BibliographicEntity (*frad:C1010*³⁰), while its range (object) the class *Name* (*frad:C1006*) (Figure 2). The reverse is true for property *isAppellationOf*. The service which IFLA uses as the provider of the platform for publishing standards in RDF is Open Metadata Registry (OMR) managed by Metadata Management Associates.³¹

The screenshot shows the Open Metadata Registry (OMR) interface. The header includes the OMR logo and the text "Supporting Metadata Interoperability". There are search boxes for "Search Vocabularies" and "Search Element Sets". The main content area displays "Element Sets: FRAD model" and "Elements: has appellation". Below this, there are tabs for "Detail", "Statements", and "History". The "Detail" tab is selected, showing a table of metadata:

Metadata +	
Detail	
Label:	has appellation
Name:	hasAppellation
URI:	http://iflastandards.info/ifs/frad/P2013 (RDF)
Description:	Relates a bibliographic entity to a name.
Comment:	
Type:	property
Domain:	http://iflastandards.info/ifs/frad/C1010
Range:	http://iflastandards.info/ifs/frad/C1006
Status:	Published
Language:	English
Note:	

On the right side, there is a "Browse..." section with links for "Resource Owners", "Vocabularies", "Element Sets", and "SPARQL". A "Feedback" button is also visible.

FIGURE 2.1

FRAD attribute *has appellation* published as RDF property *hasAppellation* in Open Metadata Registry

The screenshot shows the Open Metadata Registry (OMR) interface. The header includes the OMR logo and the text "Supporting Metadata Interoperability". There are search boxes for "Search Vocabularies" and "Search Element Sets". The main content area displays "Element Sets: FRAD model" and "Elements: Bibliographic Entity". Below this, there are tabs for "Detail", "Statements", and "History". The "Detail" tab is selected, showing a table of metadata:

Metadata +	
Detail	
Label:	Bibliographic Entity
Name:	BibliographicEntity
URI:	http://iflastandards.info/ifs/frad/C1010 (RDF)
Description:	An entity which reflects intellectual constructs or concepts that are integral to the rules used to create library catalogues.
Comment:	
Type:	class
Status:	Published
Language:	English
Note:	What is perceived as a specific instance of a particular entity type may vary from one set of rules to another.

At the bottom right of the table, there are buttons for "List" and "Get RDF". On the right side, there is a "Browse..." section with links for "Resource Owners", "Vocabularies", "Element Sets", and "SPARQL". A "Feedback" button is also visible.

FIGURE 2.2

FRAD entity *bibliographic entity* published as RDF class *BibliographicEntity* in Open Metadata Registry

FRBR and FRAD in RDF

FRBR, the entity-relationship conceptual model designed to represent bibliographic universe “identifies and clearly defines the entities of interest to users of bibliographic records, the attributes of each entity, and the types of relationships that operate between entities. The intent was to produce a conceptual model that would serve as the basis for relating specific attributes and relationships (reflected in the record as discrete data elements) to the various tasks that users perform when consulting bibliographic records.”³² The entities FRBR defined are *work*, *expression*, *manifestation* and *item* as Group 1 entities, *person* and *corporate body*, with a later addition of *family* as Group 2 entities, and *concept*, *object*, *event* and *place* as Group 3 entities. In RDF they are represented as classes. As the FRBR, however, focused on bibliographic data as distinct from authority data, the Study Group on the Functional Requirements for Bibliographic Records recognized “the need to extend the model at some future date to cover authority data”.³³ FRAD conceptual model was designed following the same methodology, but it covered only name authority data, excluding attributes and relationships associated with types of entities representing subject authority data, i.e., Group 3 entities. That task was dealt with by the FRAD conceptual model.

IFLA’s FRBR Review Group was formed in 2002 with the remit to support the development of the FRBR conceptual model, and promote its use as a reference model. In 2009, when all three models were publicly available, the Group took on the remit to review and maintain all three conceptual models as the “FRBR family of models”. The decision to start the project on representing the FRBR model in RDF was made already in 2007, stimulated by the report from the Data Model meeting held at the British Library in London previously that year.³⁴ The task of the project was “to define appropriate namespaces for FRBR (entity-relationship) in RDF and other appropriate

30 Shortened notation for URI <http://iflastandards.info/ns/fr/frad/C1010>. IFLA Namespaces Task Group decided to use opaque URIs for FR family models and ISBD, rather than <http://iflastandards.info/ns/fr/frad/BibliographicEntity> because of a multilingual environment in which IFLA standards are used.

31 Open Metadata Registry [cited: 2012-03-06]. Available at: <http://metadataregistry.org/>
32 FRBR. Op. cit., p. 3.

33 Ibid., p. 4; see also p. 5.

34 Data Model Meeting, British Library, London 30 April - 1 May 2007 [cited: 2012-03-06]. Available at: <http://www.bl.uk/bibliographic/meeting.html>

syntaxes”,³⁵ the approval of which marked the beginning of the work on declaring namespaces for FR family models in RDF.³⁶

Understanding of the relationship among the three conceptual models is important because in the process of representing FR family models’ entities, attributes and relationships in RDF it was necessary to agree to its methodology. As already mentioned, each FR model entity becomes a RDF class, while attributes and relationships become a RDF property. The problem to solve was the registration of namespaces for the models: should each model have its own namespace, that is, should classes and properties be represented in each model in their own right, or should they be represented in a prior model, and re-used where indicated in the FR model documentation? For example, if the FRBR entity *person* is represented in OMR as a class *frbrer:C1005*, and if its definitions in FRBR and FRAD are identical, there is no need to represent it again in the FRAD model namespace. The latter methodology was accepted by the FRBR Review Group, and the Group approved when the definitions can be considered identical. For example, the definition of the entity *person* in FRBR is “an individual”, while in FRAD it is “an individual or a person or identity established or adopted by an individual or group. [FRBR, modified].”³⁷ Although the FRBR Review Group agreed that the different definitions are sufficiently great to require FRAD to create its own class for Person,³⁸ the class Person has not been represented by its own FRAD namespace in OMR.

This issue obviously needs further consideration: the difference in definitions in this case points to different concepts which will have to be resolved on the ontological level in order to be adequately resolved in linked data systems if libraries’ data should be used as linked data outside the bibliographic universe. Namely, several questions could be therefore asked: Who or what is this instance of the class *Person* represented by the URI *frbrer:C1005*? Is it a real person with all of his or her identities (in the real world), or a person with only bibliographic identity, i.e., entity of interest to the bibliographic universe? How is such a

35 Dunsire, Gordon. Declaring FRBR entities and relationships in RDF, 2008/07/25 [cited: 2012-03-06]. Available at: <http://www.ifla.org/files/cataloguing/frbrg/namespaces-report.pdf>

36 Dunsire, Gordon. Interoperability and semantics in RDF representations of FRBR, FRAD and FRAD. // Concepts in context : proceedings of the Cologne Conference on Interoperability and Semantics in Knowledge Organization July 19th-20th, 2010 / edited by Felix Boteram, Winfried Gödert and Jessica Hubrich. Würzburg : ERGON, 2011. Pp. 133-147.

37 FRBR. Op.cit., p. 25 [3.2.5].

38 Dunsire, G. Interoperability and semantics . Op. cit., p.142.

person identified (by URI) as an information resource on the web? Can a URI of an authority record in which the name of a particular person is recorded stand instead of that non-information resource? Such questions will necessarily also be part of the discussions about the harmonization of the FR family models which is part of the strategic plan of the IFLA Cataloguing Section, to be realized by the FRBR Review Group as planned for 2012-2013.³⁹

However, within the bibliographic universe the problem is identified and defined, and can be considered to be solved at the practical level not only within a local catalogue but also at the level of the cooperative authority system VIAF. The difference in the definition of the entity *person* is a good example to test this statement. Namely, it is often the case that one bibliographic agency would recognize pseudonym(s) used by an individual as separate persona(s) and establish authorized forms for each persona with appropriate links between these forms, while other agency would recognize only one persona for an individual and the pseudonym(s) used for that person, and would establish one authorized form, while all the others would be treated as variant forms. The result of this is that in a cooperative or networked authority file the two treatments of the same “person” would clash, and a mechanism would need to be developed to resolve this issue in a transparent way for the user. A search in VIAF for one of the names retrieves two or more “headings” for the real name and the pseudonym(s): the example given in Figure 3 shows the result of the search for Mirković, Mijo – Mirković, Mijo, and author’s pseudonym Balota, Mate. Although treated differently by the agencies in case, which can be seen by detailed analysis of the records, the algorithm developed by VIAF brought together or linked person’s real name and the pseudonym, or, in FRAD terms, realized *person-to-person* relationship. Such a solution is transparent to users, and therefore could be considered satisfactory (understandable) within the bibliographic universe, that is, in the library community. FRAD, as a conceptual model for authority data recognizes such a situation⁴⁰ and that is the reason why its definition of the entity *person* is a modified FRBR one.

39 IFLA. Cataloguing Section. Strategic plan, 2011/2013: “Action plan for 2012: FRBR: Explore the preparation of a consolidated document for IFLA’s FRBR family of conceptual models in an entity-relationship formulation” [cited: 2012-03-19]. Available at: <http://www.ifla.org/en/node/1959>

40 See, for this particular case: FRAD. Op. cit., p. 25.

The screenshot shows the VIAF (Virtual International Authority File) search interface. At the top, it says 'VIAF Virtual International Authority File'. Below that is a search box with three fields: 'Select Field:' (set to 'All Headings'), 'Select Index:' (set to 'All VIAF'), and 'Search Terms:' (containing 'Mirković, Mijo 1898 1963'). A 'Search' button is to the right. Below the search box, it says '2 headings found for Mirković, Mijo 1898 1963'. A table follows with three columns: 'Heading', 'Type', and 'Sample Title'.

Heading	Type	Sample Title
1 Mirković, Mijo 1898-1963 Balota, Mate 1898-1963 Mirković, Mijo	Personal	Ekonomska historija Jugoslavije. Ekonomska historija Jugoslavije. Matija Vlačić Ilirik.
2 Balota, Mate, 1898-1963	Personal	Puna je Pula

FIGURE 3.

Result of the retrieval for Mirković, Mijo in the VIAF: Virtual International Authority File

The FRAD relationship between persons, that is *pseudonymous relationship*, is defined as the relationship “between a ‘real’ person (i.e. an individual) and persona(e) or identity adopted by that individual through the use of one or more pseudonyms”.⁴¹ This relationship is represented in RDF as two properties: *isPseudonymousPersonaOf*, and *hasPseudonymousPersona*, which is in line with the bi-directional type of *person-to-person* relationships defined in FRAD, and, indeed, as needed for expressing them as RDF properties. Consequently, both properties have defined class *Person* (*frbrer:C1005*) for their domain and range (Figure 4).

In the framework of representing the FRBR entity *person* as a RDF class, the question is whether it can be used as a class or domain for the FRAD attributes to the entity *person* expressed as properties in RDF? This is also enabled in OMR: FRAD attribute of a person *gender*, is represented as the property *hasGender* (*frad:P3029*), with the FRBR class *Person* (*frbrer:C1005*) as a domain. This is an example of the re-use of namespaces among the FR family models representations in RDF.

41 Ibid., p. 61 [5.3.1].

The screenshot shows the Open Metadata Registry website. At the top, it says "open metadata registry" with the tagline "Supporting Metadata Interoperability". There are search boxes for "Search Vocabularies" and "Search Element Sets". The main content area displays "Element Sets: FRAD model" and "Elements: hasPseudonymousPersona". Below this, there are tabs for "Detail", "Statements", and "History". The "Detail" tab is active, showing a table of metadata:

Metadata +	
Detail	
Label:	has pseudonymous persona
Name:	hasPseudonymousPersona
URI:	http://laststandards.info/inf/frad/4035 (RDF)
Description:	Relates a "real" person (i.e., an individual) and a persona(ae) or identity adopted by that individual through the use of one or more pseudonyms.
Comment:	
Type:	property
Domain:	http://laststandards.info/inf/frbr/1005
Range:	http://laststandards.info/inf/frbr/1005
Status:	Published
Language:	English
Note:	

On the right side, there is a "Browse..." section with links for "Resource Owners", "Vocabularies", "Element Sets", and "SPARQL". A "Feedback" button is also visible.

FIGURE 4.

RDA class *Person* (*frbrer:C1005*) defined as domain and the value of the range for the property *hasPseudonymousPersona*, i.e., FRAD's *pseudonymous relationship* between person and persona

FRAD in RDF

The conceptual FRAD model defines three fundamental entities, *bibliographic entities* that comprise all FRBR entities, *name* and/or *identifier*, and *controlled access point*. Additionally, entities *rules* and *agency* are defined. All these entities are represented as classes in RDF. The model defines four broad categories of relationships, the first category covering the relationships between the three fundamental entities *has appellation* and *is appellation of*, between *bibliographic entities* and entity *name*, and *is basis for* and *is based on* between entities *name* and *controlled access point*. The relationship *person-to-person*, one of which has already been mentioned, comes into to the second category of relationships.⁴² The relationships are represented as properties in RDF.

Figure 5 displays the statements about authority data, that is, RDF triples as a graph. The example shows the treatment of authorized and variant forms of a person's name by a bibliographic agency using FRAD's entities and relationships as represented in RDF. The nodes are

⁴² For detailed description of the FRAD model with reference to old and rare books see: Willer, Mirna. Conceptual model for authority data : FRAD, and its application to old books. // Summer School in the Study of Old Books : proceedings / edited by Mirna Willer and Marijana Tomić. Zadar : Sveučilište, 2010. Pp. 207-223.

used for the subject and object of a triple, the connecting lines are used for predicates, while a rectangular node is used for an object that is a literal (information for the user). The arrow at the end of a connecting line is used to indicate the direction from subject to object.

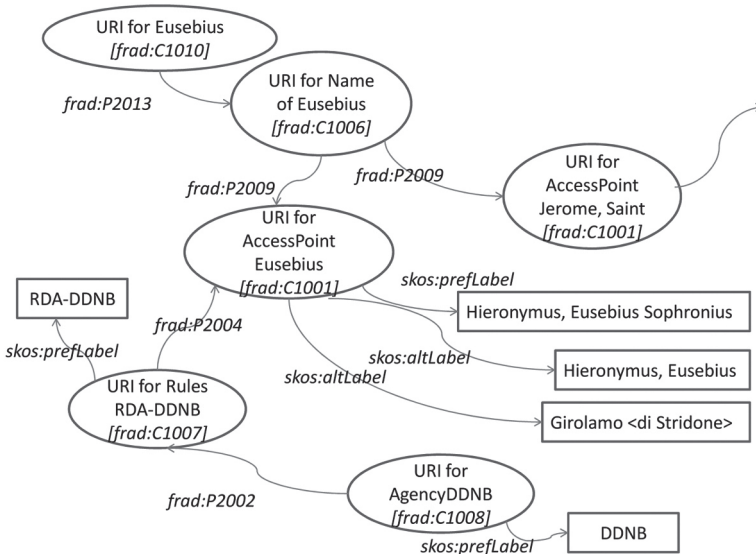


FIGURE 5. RDF graph of linked data – an example of the treatment of authorized and variant forms of names

The relationship between subject “URI for Eusebius” (BibliographicEntity [frad:C1010])⁴³ and object “URI for Name of Eusebius” (Name [frad:C1006]) is expressed by the property hasAppellation (frad:P2013), which, as already explained, has defined the class BibliographicEntity as a domain, and Name as a range. However, as the subject BibliographicEntity according to FRAD represents entities *person*, *corporate body* and *family*, it would be necessary to specify which entity is represented by the class Name; the FRAD subclass NameOfAPerson (frad:C1012) is defined in the OMR which could be used for this purpose, however the actual expression of this relationship and functionality should be further researched.

43 The reference to the class to which the subject belongs is given in the diagram in square brackets.

As a set of resources (books) that make the collection of a particular library has a particular set of forms of names by which the author is represented in those resources, and which need not be identical to a set of books in the collection of another library, these sets of names are identified in the graph by the “URI for AccessPoint Eusebius” (AccessPoint [*frad:C1001*]) and the “URI for AccessPoint Jerome, Saint” (AccessPoint [*frad:C1001*]).⁴⁴ The object of the previous triple – “URI for Name of Eusebius”, has become the subject of another triple with property *isBasisNameFor* (*frad:P2009*) between it and the objects “URI for AccessPoint Eusebius” and “URI for AccessPoint Jerome, Saint”. In the first URI we demonstrate that the treatment of access points – that is the authorized form of a name together with all variant forms, are governed by the rules (RDA-DDNB) which are applied by the agency (The German National Library: Die Deutsche Nationalbibliothek – DDNB). These two FRAD identities defined as classes in RDF, are represented as nodes in the graph with appropriate URIs: “URI for Rules RDA-DDNB” (Rules [*frad:C1007*]), and “URI for Agency DDNB” (Agency [*frad:C1008*]). The relationship among these three instances of classes is expressed in the following way: “URI for Agency DDNB” is a subject of a triple in which “URI for Rules RDA-DDNB” is the object, while the link between them is represented by the property *applies* (*frad:P2002*). This property has defined the class Agency as a domain, and Rules as a range. The object “URI for Rules RDA-DDNB” becomes a subject to a new triple to which the object is “URI for AccessPoint Eusebius”. The link is defined by the property *govern* (*frad:P2004*), with Rules as a domain and AccessPoint as a range. The value of all the three classes is expressed using notation from SKOS (Simple Knowledge Organization System)⁴⁵: *skos:prefLabel* for: Hieronymus, Eusebius Sophronius, RDA-DDNB and DDNB, and *skos:altLabel* in the case of variant forms of names: Hieronymus, Eusebius and Girolamo <di Stridone>. The identification of the authorized form as opposed to variant forms valid within a particular catalogue is thus established.

44 Each of these URIs can correspond to the URI of the authority record in the authority file of a particular local/national library.

45 W3C. SKOS Simple Knowledge Organization System [cited: 2012-03-19]. Available at: <http://www.w3.org/2004/02/skos/>

Conclusion

The new technical environment requires new standards, however, this article shows how the standards and models designed within the context of the preceding technology can evolve. The research and development of representing IFLA FR family of conceptual models in RDF, the standard of the Semantic Web has proved that possibility. However, it is necessary to understand the requirements of the new technology, and be able to re-conceptualize the existing concepts; in the case of FR family of conceptual models the asset is that they are the epitome of the cataloguing theory and practice of the last century and a half, and as such offer a firm foundation on which to build the next generation standards.

The additional value of the exercise in representing the FRAD model in RDF is that it informs back on the possible problems in the models that need further research. Namely, although mentioned in brief, FRAD entity attributes were not dealt with in this article. The attributes defined in FRAD to the entity *person* are represented in OMR as a class within the FRBR model *Person* – *frbrer:C1005*, and defined as such in OMR. However, if the application of FRAD in RDF as represented in Figure 5 is considered correct (it has followed all the rules and definitions), it shows that there is no way to attach attributes to the entity *person* as defined by FRAD, that is, properties to the class *Person* because there is no relationship or link between the FRAD class *BibliographicEntity* and FRBR class *Person*. Obviously, this problem will have to be researched from the aspect of re-use of represented classes from the prior model, and also within the project on the consolidation of FR family of conceptual models.

The article focuses on specific problems of representing instances of FRAD entity *person* as RDF triples within the bibliographic universe, however it also points to the issues that have to be discussed and researched of “positioning” bibliographic identities in relation to schemes and vocabularies of authority data providers outside the library sector.

References

- Anderson, Chris. The Long Tail. // *Wired* 12.20(October 2004) [cited: 2012-03-06]. Available at: <http://www.wired.com/wired/archive/12.10/tail.html>
- Anderson, Dorothy. *Universal Bibliographic Control : a long term policy, a plan for action*. Pullach/ München : Verlag Dokumentation, 1974.
- Berners-Lee, Tim; James Hendler; Ora Lassila. The Semantic Web. // *Scientific American* 17 May, 2001.
- Berners-Lee, Tim (with Mark Fischetti). *Weaving the Web : the past, present and future of the World Wide Web by its Inventor*. London : Orion Business Books, 1999.
- Bizer, Christian; Tom Heath, Tim Berners-Lee. Linked data – the story so far. // *International Journal on Semantic Web and Information Systems (IJSWIS)* 5, 3(2009). Pre-print available at: <http://tomheath.com/papers/bizer-heath-bernbers-lee-ijswis-linked-data.pdf> [cited: 2012-03-29].
- Calhoun, Karen. The changing nature of the catalog and its integration with other discovery tools : final report, March 17, 2006 : prepared for the Library of Congress [cited: 2012-03-06]. Available at: <http://www.loc.gov/catdir/calhoun-report-final.pdf>
- Consortium of European research Libraries. CERL Thesaurus [main information at:] http://www.cerl.org/web/en/resources/cerl_thesaurus/main; [access at:] <http://thesaurus.cerl.org/cgi-bin/search.pl> [cited: 2012-03-29].
- Data Model Meeting, British Library, London 30 April - 1 May 2007 [cited: 2012-03-06]. Available at: <http://www.bl.uk/bibliographic/meeting.html>
- Dunsire, Gordon. Declaring FRBR entities and relationships in RDF, 2008/07/25 [cited: 2012-03-06]. Available at: <http://www.ifla.org/files/cataloguing/frbrrg/namespace-report.pdf>
- Dunsire, Gordon. Interoperability and semantics in RDF representations of FRBR, FRAD and FRAD. // *Concepts in context : proceedings of the Cologne Conference on Interoperability and Semantics in Knowledge Organization July 19th-20th, 2010 / edited by Felix Boteram, Winfried Gödert and Jessica Hubrich*. Würzburg : ERGON, 2011. Pp. 133-147.
- Dunsire, Gordon; Mirna Willer. Standard library metadata models and structures for the Semantic Web. // *Library hi tech news* 28, 3(2011), 1-12. Earlier version available at: <http://www.ifla.org/files/hq/papers/ifla76/149-dunsire-en.pdf> [cited: 2012-03-29].
- Dunsire, Gordon; Mirna Willer. UNIMARC and Linked Data. // *IFLA Journal* 37, 4(December 2011), 314-326. Also available at: http://www.ifla.org/files/hq/publications/ifla-journal/ifla-journal-37-4_2011.pdf [cited: 2012-03-29].
- Dunsire, Gordon. UNIMARC, RDA and the Semantic Web. // *International cataloguing and bibliographic control (ICBC)* 39, 2(April/June 2010), 37-40. Based on a paper presented to the World Library and Information Congress: 75th IFLA General Conference and Assembly, 23-27 August 2009, Milan, Italy. Available at: www.ifla.org/files/hq/papers/ifla75/135-dunsire-en.pdf [cited: 2012-03-29]
- Functional requirements for authority data : a conceptual model / edited by Glenn E. Patton ; IFLA Working Group on Functional Requirements and Numbering of Authority Records (FRANAR). Final report, December 2008 / approved by the Standing Committees of the IFLA Cataloguing Section and IFLA Classification

- and Indexing Section, March 2009. München : K. G. Saur, 2009. Functional requirements for authority data (FRAD) : modifications and errata for the 2009 text, November 2011 at: <http://www.ifla.org/files/cataloguing/frad/FRADerrata2011.pdf> [cited: 2012-03-29].
- Functional requirements for bibliographic records : final report / IFLA Study Group on the Functional Requirements for Bibliographic Records. München : K. G. Saur, 1998. Current text including amendments and corrections to date, available at: <http://www.ifla.org/en/publications/functional-requirements-for-bibliographic-records> [cited: 2012-03-29].
- Functional requirements for subject authority data (FRSAD) : a conceptual model / IFLA Working Group on Functional Requirements for Subject Authority Records (FRSAR) ; editors Marcia Lei Zeng, Maja Žumer, Athena Salaba. Berlin ; München : De Gruyter Saur, 2011. Version, June 2010 available at: <http://www.ifla.org/files/classification-and-indexing/functional-requirements-for-subject-authority-data/frsad-final-report.pdf> [cited: 2012-03-29].
- Guidelines for authority records and references / recommended by the Working Group on an International Authority System ; approved by the Standing Committee of the IFLA Section on Cataloguing and the IFLA Section on Information Technology. 2nd ed. München : K. G. Saur, 2011.
- IFLA: International Federation of Library Associations and Institutions [cited: 2012-03-29]. Available at: <http://www.ifla.org>
- IFLA. Cataloguing Section. Strategic plan, 2011/2013 [cited: 2012-03-19]. Available at: <http://www.ifla.org/en/node/1959>
- IFLA. Cataloguing Section. FRBR Review Group [cited: 2012-03-29]. Available at: <http://www.ifla.org/en/frbr-rg>
- IFLA. Cataloguing Section. ISBD Review Group. ISBD/XML Study Group [cited: 2012-03-29]. Available at: <http://www.ifla.org/en/node/1795>
- IFLA. Cataloguing Section. ISBD Review Group and ISBD/XML Study Group. ISBD namespaces published, 8 February 2012 [cited: 2012-03-29]. Available at: <http://www.ifla.org/en/news/isbd-namespaces-published>
- International Conference on Cataloguing Principles, Paris, 9th-18th October, 1961: Report. London : International Federation of Library Associations, 1963. Reprinted in: International Conference on Cataloguing Principles, Paris, 9th-18th October, 1961: Report. London : Clive Bingley, 1969.
- The Library of Congress Working Group on the Future of Bibliographic Control. On the record : report, January 9, 2008 [cited: 2012-03-06]. Available at: <http://www.loc.gov/bibliographic-future/news/lcwg-ontherecord-jan08-final.pdf>
- Linked Data. Connect distributed data across the Web [cited: 2012-04-18]. Available at: <http://linkeddata.org/>
- Mossman, Katherine. Serving the niche : viewing libraries through Chris Anderson's & Long Tail's lens. // Library journal 07/15/2006 [cited: 2012-03-06]. Available at: <http://www.libraryjournal.com/article/CA6349032.html>
- Open Metadata Registry [cited: 2012-03-06]. Available at: <http://metadataregistry.org/>
- Statement of Principles adopted at the International Conference on Cataloguing Principles Paris, October, 1961. Annotated ed. with commentary and examples by Eva Verona assisted by Franz Georg Kaltwasser, P. R. Lewis, Roger Pierrot. London : IFLA Committee on Cataloguing, 1971.

- Storey, Tom. The Long Tail and libraries. // OCLC Newsletter April/May/June 2005, 6-11.
- UNIMARC manual : authorities format / edited by Mirna Willer. 3rd edition. München : K. G. Saur, 2009.
- VIAF: Virtual International Authority File [main information at:] <http://www.oclc.org/research/activities/viaf/>, [service available at:] <http://viaf.org/>. As of 4 April 2012, VIAF has transitioned from an OCLC Research prototype to an OCLC service; see: <http://www.oclc.org/viaf/default.htm> [cited: 2012-04-04].
- Willer, Mirna. Conceptual model for authority data : FRAD, and its application to old books. // Summer School in the Study of Old Books : proceedings / edited by Mirna Willer and Marijana Tomić. Zadar : Sveučilište, 2010. Pp. 207-223.
- Willer, Mirna. Name authority control paradigm shift in the network environment. // Frameworks for ICT policy : government, social and legal issues / edited by Esharenana E. Adomi. Hershey, PA : IGI Global, 2010. Pp. 182-205. DOI: 10.4018/978-1-61692-012-8.ch012
- W3C. RDF: Resource Description Framework Primer, W3C Recommendation 10 February 2004 [cited: 2012-03-29]. Available at: <http://www.w3.org/TR/rdf-primer/>
- W3C. SKOS Simple Knowledge Organization System [cited: 2012-03-19]. Available at: <http://www.w3.org/2004/02/skos/>
- W3C Semantic Web Activity [cited: 2012-03-06]. Available at: <http://www.w3.org/2001/sw/>
- W3C. Web naming and addressing : URIs, URLs, ... [cited: 2012-03-06]. Available at: <http://www.w3.org/Addressing/#rfc3986>

Biographical sketch

Dr. Mirna Willer is Professor at the Department of Information Sciences, University of Zadar, Croatia. Her teaching and research interests include the theory and practice of information organisation, metadata and identifiers, digital web archives, interoperability, IFLA standards and models in the semantic web. She worked as system librarian, standards officer and senior researcher at the National and University Library in Zagreb, Croatia from 1980 to 2007. Among a number of international body memberships, she was a standing member of the IFLA Permanent UNIMARC Committee from its establishment in 1991 until 2005 (chair of the Committee from 1997 to 2005), since then she has been its consultant and honorary member. She was also a member of the IFLA Working Group on FRANAR, the Working Group responsible for the development of the conceptual model FRAD, ISBD Review Group's ISBD Future Directions Working Group, chair of the ISBD/XML Study Group, and since 2011 chair of the ISBD Review Group. She wrote a book *UNIMARC in Theory and Practice*, a chapter on authority control, about 100 articles (professional and research papers,

reviews, etc.), translated texts in the field of UBC such as *UNIMARC Manual: Bibliographic Format* and E. Svenonius's *The Intellectual Foundation of Information Organization*, and edited the 3rd edition of *UNIMARC Manual: Authorities Format*.

**NADZOR NAD AUTORIZIRANIM
PODACIMA IMENA OSOBA
OD KONCEPTUALNOG MODELA DO SEMANTIČKOG
WEBI I POVEZANIH PODATAKA**

Sažetak

Rad se bavi područjem nadzora nad autoriziranim podacima imena osoba s ciljem opisa procesa predstavljanja konceptualnog modela FRAD (*Functional Requirements for Authority Data*) u RDF-u (*Resource Description Framework*), standardu W3C-a s ciljem da se istraže problemi funkcionalnosti modela u okruženju semantičkog weba i povezanih podataka. Na praktičnom primjeru prikazuje se način iskazivanja autoriziranih podataka u tripletima RDF-a, i upozorava na određene probleme koji su uočeni tijekom tog procesa.

Ključne riječi: normativni nadzor imena, FRAD: *Functional Requirements for Authority Data*, RDF: *Resource Description Framework*, semantički web, povezani podaci