Knowledge Organization and Cultural Diversity

Organizers: José Augusto Chaves Guimarães e Vera Dodebei

ISKO-BRASIL

Organizers

José Augusto Chaves Guimarães
Vera Dodebei

Knowledge Organization and Cultural Diversity

Marília, São Paulo, Brasil

Sociedade Brasileira de Organização do Conhecimento
(ISKO-Brasil)

2017

K73k Knowledge Organization and Cultural Diversity [recurso eletrônico] / José Augusto Chaves Guimarães, Vera Dodebei, organizadores. -- Pernambuco: ISKO-Brasil ; UFPE, 2017.
760 f. ; 30 cm.

ISBN: 978-85-415-0924-4

Livro digital

1. Organização do conhecimento. I. Título.

CDD 025.4

Scientific Committee

Vera Lúcia D. Louzada de Mattos Dodebei (UNIRIO)

Gercina Ângela Borém de Oliveira Lima (UFMG)

Marisa Bräscher Basílio Medeiros (UFSC)

Maria Aparecida Moura (UFMG)

Fabio Assis Pinho (UFPE)

Editors

Isadora Victorino Evangelista

Gilberto Gomes Cândido

Rafael Aparecido Moron Semidão

Rafael Cacciolari Dalessandro

Suellen Oliveira Milani

Design

Maíra Fernandes Alencar

Translation

Natália Nakano

© Reproduction of this book in whole or in part is permitted provided that the claims mentioned. The sale is prohibited.

SUMMARY

THE EPISTEMOLOGIC DIMENSION OF KNOWLEDGE ORGANIZATION

Approaches and paradigms in knowledge organization……..**30**
Leila Cristina Weiss e Marisa Bräscher (UFSC)

Historical-epistemological dimension in Knowledge Organization: Gesner's taxonomy contributions, 16th century……**37**
Andre Vieira de Freitas Araujo e Giulia Crippa (USP)

Terminological and interdisciplinary research configurations in knowledge organization (2010-2014)……**47**
Leilah Santiago Bufrem, Aline Elis Arboit e Juliana Lazzarotto Freitas (UNESP)

The conceptual dimension of Knowledge Organization in NASKO conferences: Bardian's content analysis.....**55**
José Augusto Chaves Guimarães (UNESP); Rodrigo de Sales (UFF), André Ynada dos Santos (UNESP) e Daniela Fernanda de Oliveira Matos (UNESP)

The relation between knowledge organization and information science in the Brazilian scientific community: an investigation within ISKO-Brazil……..**73**
Rodrigo de Sales (UFF)

Poole, index and fractures: indexing and serial publications in the United States of the nineteenth century……**85**
Gustavo Silva Saldanha (IBICT/Unirio) e Naira Christofoletti Silveira (Unirio)

Processes of archival knowledge representation: historical and conceptual elements for classification and description……..**94**
Natália Bolfarini Tognoli (UNESP) e Thiago Henrique Bragato Barros (UFPA)

Ontologies and Simple Knowledge Organization Systems (SKOS): similarities and differences……..**101**
Rogério Aparecido Sá Ramalho (UFSCAR)

Academic genealogy as an approach to domain analysis……....**108**
Renata Cristina Gutierres Castanha e Maria Cláudia Cabrini Grácio (UNESP)

The epistemological dimension of Document Analysis of fictional content in Knowledge Organization……..**116**
Roberta Caroline Vesú Alves e João Batista Ernesto de Moraes (UNESP)

Indexing languages and domain analysis.....**122**
Vera Lucia Ribeiro Guim e Mariângela Spotti Lopes Fujita (UNESP)

Reader-Interest classification as a method of stock control: the McClellan Legacy……..**131**
Daniel Martínez-Ávila (UNESP) e *M. P. Satija (Guru Nanak Dev University)*

Terminology, exhaustivity and specificity: a conceptual relation……**140**

Isadora Victorino Evangelista, Walter Moreira e João Batista Ernesto de Moraes (UNESP)

Aspects of subject metadata in cataloging photographs……151
Ana Carolina Simionato (UFSCar)

Documentary languages and audiovisual resources treatment: theoretical and methodological interfaces……..160
Francisnaira Cristina Ravazzi (Diretoria de Ensino) e Walter Moreira (UNESP)

Knowledge Organization and Information Design: a convergence study…….170
Natalia Nakano, Mariana Cantisani Padua e Maria José Vicentini Jorente (UNESP)

Domain Analysis of the first ISKO-Brazil chapter editions……180
Exequiel Fontans e Mario Barité (Universidad de la Republica)

Teaching in knowledge organization and representation: languages and standards……196
Cibele Araújo Marques dos Santos, Eduardo de Abreu de Jesus e João Ricardo de Luca (USP)

Genealogy of the concept of Information Science in Brazil: a discursive analysis from founding journals in the area……..206
Larissa de Melo Lima, João Batista Ernesto de Moraes e Daniel Martínez-Ávila (UNESP)

THE APPLY DIMENSION OF KNOWLEDGE ORGANIZATION

Semantic Web Technologies applied to Knowledge Organization: SKOS for the construction and use of decentralized controlled vocabularies……..220
José Eduardo Santarem Segundo (USP) e Caio Saraiva Coneglian (UNESP)

Preliminary analysis on the conversion of classification plans into controlled vocabularies……229
Luciana Davanzo e Walter Moreira (UNESP)

Thesaurus Reengineering: the Thesagro case……237
Benildes Maculan (UFMG), Gercina Lima (UFMG), Ivo Pierozzi Jr. (Embrapa) e Leandro Oliveira (Embrapa)

Conceptual Model of Organizational Competitive Intelligence.....246
Thiciane Mary Carvalho Teixeira (UECE) e Marta Lígia Pomim Valentim (Unesp)

Vocabulary control in electronic scientific journals.....256
José Carlos Francisco dos Santos e Brígida Maria Nogueira Cervantes (UEL)

Documentary Languages in archival description.....266
Maria de Fátima Santos de Lima e Francisco Aragão Pedroza da Cunha (UFBA)

Research in knowledge organization systems in DCMI conferences…….277

Felipe Augusto Arakaki, Plácida Leopoldina Ventura Amorim da Costa Santos e Rachel Cristina Vesú Alves (UNESP)

Conceptual models and descriptive representation of information……288
Elisabete Gonçalves de Souza e Wellington Freire Cunha Costa (UFF)

Nanopublishing Modeling: an experimental approach.....296
Lorena Tavares de Paula e Maria Aparecida Moura (UFMG)

Application of ontologies in the information retrieval process applied in academic settings……310
Caio Saraiva Coneglian (UNESP), Elvis Fusco (Univem) e José Eduardo Santarém Segundo (USP)

Librarian performance in the subject analysis of theses within the theoretical dimensions of subject cataloging and indexing……322
Roberta Cristina Dal'Evedove Tartarotti, Paula Regina Dal'Evedove e Mariângela Spotti Lopes Fujita (UNESP)

The contribution of archival identification to knowledge organization in personal archives……..331
Gabrieli Aparecida da Fonseca e Sonia Maria Troitiño Rodriguez (UNESP)

Document analysis and the generative trajectory of meaning in the representation of the archival document……337
Gilberto Gomes Cândido (UNESP), João Batista Ernesto de Moraes (UNESP) e Deise Sabbag (USP)

Contribution of the Generative Trajectory of Meaning for documentary reading of narrative texts of fiction……347
Deise Sabbag (USP) e João Batista Ernesto de Moraes (UNESP)

Vocabulary control and its contribution to research: the case of agrarian sciences in Uruguay……365
Lucía Simón e Mario Barité (Universidad de la República)

Rural cultural goods: an experience report of the taxonomy construction in the context of the historical farms of São Paulo.....373
Mayara Cristina Bernardino (UFSCar), Luciana de Souza Gracioso (UFSCar), Maria da Graça Melo Simões (Universidade de Coimbra) e Luzia Sigoli Fernandes Costa (UFSCar)

Harmonization of CIDOC CRM ontology in the context of archives, libraries and museums……385
Laís Barbudo Carrasco (UNESP), Silvana Aparecida Borsetti Gregório Vidotti (UNESP) e Phil Manfred Thaller (Universität zu Köln)

Cataloging of interactive resources: analysis of recommendations and practices in international catalogs……395

Víctor Amor, Daniel Martínez-Ávila (UNESP) e Rosa San Segundo (Universidad Carlos III)

Metadata structure for an image base in pathology……404
Elisabete Gonçalves Souza, Jóice Cleide Cardoso Ennes de Souza e Elan Cardozo Paes de Almeida (UFF)

Indexing through images: accessibility via OPACs……416
Rita de Cássia do Vale Caribé e Marcílio de Brito (UnB)

Interactive conceptual maps as a didactic instrument in Information and Knowledge Organization……435
Marilda Lopes Ginez de Lara, Gabriela Previdello e Nair Yumiko Kobashi (USP)

Domain analysis in Knowledge Organization: exploring thematic and citation relations……446
Bruno Henrique Alves, Ely Francina Tannuri de Oliveira e Maria Cláudia Cabrini Grácio (UNESP)

A study on domain in knowledge organization through author citation and co-citation analysis……455
Lidyane Silva Lima, Pollyana Ágata Gomes da Rocha Custódia, Ely Francina Tannuri de Oliveira e Leilah Santiago Bufrém (UNESP)

Classification approaches of the papers submitted to ENANCIB WG2: professional classification and non-professional classification……463
Walter Moreira e Isabela Santana de Moraes (UNESP)

Information behavior within ISKO - International Society For Knowledge Organization: a study for the 2004-2014 period……473
Marli Vitor da Silva e Helen de Castro Silva Casarin (UNESP)

Subject representation in Archival Science: an analysis of national and international scientific articles……..481
Graziela Martins de Medeiros, Leolibia Luana Linden, Luciane Paula Vital e Marisa Bräscher (UFSC)

Analysis of methodology in bibliometric studies: a proposal for context indicators……490
Maria Guiomar da Cunha Frota e Ana Cláudia Ribeiro (UFMG)

Knowledge Organization and Representation: contribution to metric studies ……501
Cibele Araújo Marques dos Santos (USP)

Relativizing indexes h and g: a study applied to the domain of metric studies ……..……508
Deise Deolindo Silva e Maria Cláudia Cabrini Grácio (UNESP)

Domain Analysis in knowledge organization and representation in Archival Science……515
Cynthia Maria Kiyonaga Suenaga (UNESP) e Brígida Maria Nogueira Cervantes (UEL)

The FamilySearch Indexing as a crowdsourcing initiative in the context of knowledge organization……...526
Paula Carina de Araújo e José Augusto Chaves Guimarães (UNESP)

A study on actions to make government datasets available in linked open data…….533
Fernando de Assis Rodrigues e Ricardo César Gonçalves Santana (UNESP)

Knowledge Organization and industrial heritage in São Paulo: the project Eletromemória.....544
Vânia Mara Alves Lima (USP), Marcia Cristina e Carvalho Pazin Vitoriano (UNESP) e Cristina Hilsdorf Barbanti (USP)

THE SOCIAL, CULTURAL AND POLITICAL DIMENSION OF KNOWLEDGE ORGANIZATION

Reflections on the development of a methodology for subject analysis……..585
Paula Regina Dal'Evedove (UFPE), Roberta Cristina Dal'Evedove Tartarotti (UNESP) e Mariângela Spotti Lopes Fujita (UNESP)

Information organization, representation, retrieval and access: (re)configuration of MARC21 Format and BIBFRAME for cultural diversity in digital information environments?.....592
Zaira Regina Zafalon e Marcela Cristina Nespoli (UFSCar)

Music, literature and audiovisual: the contributions of knowledge organization (KO) in the intersectional relations between the works by Dorival Caymmi and Jorge Amado……..598
Fabio Assis Pinho (UFPE), Francisco Arrais Nascimento (UFPE) e Andréa Carla Melo Marinho (UFRGS)

Mediation in Knowledge Organization Domain.....605
Mona Cleide Quirino da Silva Farias, Carlos Cândido de Almeida e Daniel Martínez-Ávilla (UNESP)

The value of information and language in consumer society……617
Luciana de Souza Gracioso (UFSCar)

Academic faculty formation in Archival Description……624
Laura Maria do Rego, José Augusto Chaves Guimarães e Natália Bolfarini Tognoli (UNESP)

PROSPECTS IN KNOWLEDGE ORGANIZATION RESEARCH IN BRAZIL

Knowledge Organization and the core of Information Science: tales of big data, computing clouds and social networks……634
Renato Rocha Souza (FGV), Mauricio Barcellos Almeida (UFMG) e Renata Abrantes Baracho Porto (UFMG)

Knowledge Organization: research and developments……644
Gercina Ângela Borém de Oliveira Lima (UFMG)

ISKO-Brazil and the research groups in knowledge organization……..663
Evelyn Goyannes Dill Orrico (UNIRIO)

Research perspectives on knowledge organization: reflections from PPGCI IBICT-UFRJ academic papers……..671
Rosali Fernandez de Souza (IBICT)

Search perspectives on knowledge organization in Brazil……683
Leilah Santiago Bufrem (UNESP/UFPE)

RESEARCH IN DOCUMENTAL ANALYSIS ON BRAZIL: THE JEAN-CLAUDE GARDIN INFFLUENCE

The search for efficiency in information and knowledge representation – later developments in Gardin's thought……700
Johanna W. Smit (USP)

Jean Claude Gardin and Document Analysis: trajectory of a semiology of representation……710
Maria de Fátima Gonçalves Moreira Tálamo e Giovana Deliberali Maimone (USP)

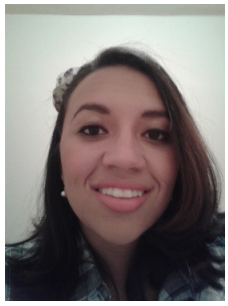
Documentary Language from J.-C. Gardin's viewpoint……..721
Marilda L. G de Lara (USP)

From Document Analysis to terminology: theoretical and methodological trajectory……730
Vania M. A. Lima (USP)

Ordering Documents: underpinnings and relations with bibliographic classification……737
Cristina D. Ortega (UFMG)

Information organization, representation, retrieval and access: (re)configuration of MARC21 Format and BIBFRAME for cultural diversity in digital information environments?

Zaira Regina Zafalon
São Carlos Federal University
zaira@ufscar.br



Marcela Cristina Néspoli
São Carlos Federal University
marcela_nespoli@hotmail.com

1 Introduction

Studies in Information Organization and Representation have been redefined over the years in search of a guarantee of information retrieval and access especially due to the diversity present in cultural heritage institutions. Such diversity can be characterized by search behavior, retrieval mode and access to information, (re)configuration of informational production and adopted technologies. In this way, concepts, models, standards and formats of information organization and representation have also been discussed. This research is part of this universe and seeks to discuss the purposes of MARC21 Format and BIBFRAME as standards of information representation in digital information environments. The discussions involve, as a background, the expansion of the universe of information search and retrieval by communities as well as metadata use and reuse by technological parks; both relate to the availability of information by cultural heritage units on the web through the description of resources. In order to meet the proposed objective, the research presents a qualitative approach, of an applied nature, with a descriptive and exploratory profile. At the same time, it uses bibliographical and documentary research. Because this is a recent theme, this paper intends to serve as a social and theoretical contribution for future research on this topic.

2 Background on representation of resources

Information organization and representation (IOR) actions are based on processes of metadata analysis, synthesis, condensation, representation, retrieval, use and reuse of information resources in order to coexist with the user and obtain documents that seek to meet their needs given the possibilities of establishing communication. We highlight the studies focused on the concepts, models, standards and formats of information organization and representation, whether they are aimed at reading, analyzing, condensing, representing and constructing documentary languages, whether they are of functional requirements, conceptual modeling, interoperability and structures of information sharing and access by humans and machines.

The result of documentary representation efforts, present in IOR studies, focuses on the bibliographic record that establishes, therefore, semantic relations, between the elements present in the represented document and their value, subsemantics, marked by the relation between the several units that compose a bibliographic record and supersemantic, among the various records that refer to the various manifestations of a work, in addition to the syntactic relations, a requirement for the registration of value in semantic structures. (ZAFALON, 2014).

Representation standards, both the content and structure ones, ensure record consistency by defining syntax and semantic rules for the description of informational support and the content of informational resources, and attest to the success in resource retrieval once it minimizes possible differences in information representation and sharing.

3 (Re)configuration of purposes by cultural diversity?

MARC21 format, as a metadata structure standard, taken as a requirement for the process of sharing bibliographic records, provides mechanisms for reading and interpreting bibliographic data by machines. Allied to the records exchange format (ISO 2709), content standards (CDD, CDU, AACR, LCSH, among others) and communication protocols and bibliographic retrieval (Z39.50), ensures the transfer of qualified data across machines. Avram (1970, 1975) points out that the development of the format was based on the philosophy of information sharing and reuse, in which

a set of data can be accessed and used by several users at the same time. Initially configured for use by the Library of Congress (LC), MARC assumes features that allow each library to have autonomy both in providing records and in sharing with other units. MARC's flexibility, according to Zafalon (2008), occurs through its marks, given the adopted conventions which make the management and retrieval of bibliographic data more efficient, resulting from the specificity they promote. MARC was, therefore, developed with the aim of standardizing the representation of informational records and, thus, allow, through cooperative actions, the efficient exchange of bibliographic data in order to avoid duplication of efforts. Disadvantages of the MARC format, discussed by Tennant (2002, 2013), one of the supporters of the MARC format substitution initiative, indicate the complexity of its adoption, reliance on markup for semantic effects, record-keeping by qualified professionals, which generates desertion and resistance from the community itself, and the difficulty of reading the records given the complex syntax, natural of the numerical scheme of the labels. The author also highlights, the inability of the format to perform document relationships that are entirely digital, such as the URLs.

In this context, the Bibliographic Framework (BIBFRAME) model, from LC initiative, has been proposed as a substitute for MARC since 2012, whose purpose is to serve as a model for which several content models can be mapped. It is being designed to, intentionally, become a formal entry point for the librarian community into a network that involves both metadata and linked documents. This tends to be made possible by the fact that BIBFRAME is based on the proposal of creating semantic relations (LinkedData) and for presenting a modeling based on Resource Description Framework (RDF), model for data interchange on the Web. In the face of information reuse paradigm at its highest level, BIBFRAME recognizes entity, attributes and relationships, through its main classes (work, instance, author and annotation). (LIBRARY OF CONGRESS, 2012).

By aligning MARC's objective with the need to make bibliographic resources available to users, to the possibility of access to bibliographic resources, and, finally, to the objective of guaranteeing access to bibliographic resources, it is understood that MARC meets its initial purpose, since it guarantees retrieval and access to the bibliographic resources described in it. However, in a broader context involving open data and Web data studies, for example, MARC appears obsolete given the difficulty in creating relationships between resources, which compromises the navigation goal

beyond the catalog. BIBFRAME, however, in terms of bibliographic data structuring, in order to allow data exchange, as well as to enable the retrieval of increasingly specific information, indicates, for the moment, a gap in the new model, which has generated debates and disagreements.

4 Conclusions

MARC format has been harshly criticized for the number of tags, fields and subfields not always used which carry unnecessary costs to the software, as well as the difficulty for catalogers to use it. BIBFRAME, because of its low specificity in record description, jeopardizes the search result and the search refining, which requires fine granularity. BIBFRAME provides the description in instances and relationships with other major classes (work, author and annotation), but not the specific description. This is due to the data modeling used, RDF, which does not involve the creation of as many predicates as necessary to achieve the specificity of MARC. Another point refers to the temporality of URLs and the standardization in the creation of URIs, in order to guarantee univocity in retrieval, as foreseen in MARC. Obtaining control of authorities and access is imperative to achieve the purpose of the model.

Even with so many issues and uncertainties, LC and Zepheria (a company specialized in library applications) are developing a demo version of the model for bibliographic resource description to structure data and allow it to be entered. It is not clear, however, how data will be stored and exchanged with other formats. The difficulty is that RDF has the function of describing generic relationships across data, allowing computers to use and exchange these data and still preserve its original meaning, but not structure the information.

In order to draw a comparison between MARC21 and BIBFRAME, it is observed that the first one is dedicated to allowing the communication of bibliographic data through exchange in order to meet any informational need, anywhere worldwide, by any library or user; and that the second one seeks to integrate various knowledge communities, through linked data. As for the structure, both are differentiated since MARC uses fields and subfields, while BIBFRAME adopts RDF triples. MARC's advantages can be observed when saving time, inserting multiple records, consistency of records, guarantee of data integrity, possibility of cooperative cataloging and guarantee of more efficient retrieval. BIBFRAME, on the other hand, presents

integration of several knowledge communities, possibility to provide better navigability, self-sufficiency in records description, library entry in the Web data. The counterpoint in using BIBFRAME focuses on the lack of familiarity by users with RDF and the fact that it is not structured for information exchange, making use of little specificity in the descriptions.

When resuming the general objective of the research, it is understood that BIBFRAME arises with the intention of integrating libraries to all knowledge communities, through the Web, since, due to its structure, MARC21 format cannot carry out such characteristic. In this way, it can be verified that the substitution proposal is based on the need to integrate libraries with the Web environment in order to allow greater navigability to the user, integrating concepts and models such as FRBR, ontologies, RDF, XML and LinkedData.

In this sense, both MARC and BIBFRAME are intended for integrating communities that are or were isolated. MARC, with a somewhat timid purpose compared to BIBFRAME (disregarding the cultural and technological configuration of the time), connected the libraries so they could exchange bibliographic records, which resulted in benefits such as data consistency, information reuse and decrease in rework. BIBFRAME, through navigability concept, tends to integrate the library into a wide range of knowledge communities on a macro level. However, the similarity of purposes is restricted only to this point, since there are still many gaps and uncertainties in the development of BIBFRAME that do not contemplate the objectives served by MARC and, therefore, do not designate the same role played by the format.

Thus, it is not envisaged, at least for now, that BIBFRAME may be a substitute for MARC. However, similarly to what occurred to MARC, many of its applications and advantages as a model of data exchange could not be felt until its stability and conception as a consolidated format. The same can happen with BIBFRAME, which, as a model, does not provide a solution to all doubts that still linger.

In short, although BIBFRAME initiative has arisen for updating MARC21, it is understood that there is no impediment of coexistence between the two, since each serves specific and complementary purposes.

References
The references was made folowing the ABNT rules.

AVRAM, H. D. MARC: its history and implications. Washington: Library of Congress, 1975.

AVRAM, H. D. The evolving MARC system: the conception of a data utility. Illinois: University of Illinois, 1970.

LIBRARY OF CONGRESS. Bibliographic Framework as a Web of Data: Linked Data Model and Supporting Services. Washington, DC. 2012. Disponível em: <www.loc.gov/bibframe/pdf/marcld-report-11-21-2012.pdf>. Acesso em: 28 abr. 2015.

TENNANT, R. MARC Must Die. Library Journal, 15 Oct. 2002. Disponível em: < http://lj.libraryjournal.com/2002/10/ljarchives/marc-must-die/>. Acesso em: 01 out. 2014.

TENNANT, R. The Post MARC Era, Part 2: Where the problems lie, part 1. Library Journal, 8 May 2013. Disponível em: < http://www.thedigitalshift.com/2013/05/roy-tennant-digital-libraries/the-post-marc-era-part-2/>. Acesso em: 25 set. 2014.

ZAFALON, Z. R. Formato MARC21 Bibliográfico: estudos e aplicações para livros, folhetos, folhas impressas e manuscritos. São Carlos: EdUFSCar, 2008.

ZAFALON, Z. R. Scan for MARC: conversão de registros bibliográficos analógicos para o Formato MARC21 Bibliográfico. São Paulo: Editora Unesp, 2014.