Agris on-line Papers in Economics and Informatics

Volume IX Number 3, 2017

# Metadata Formats for Data Sharing in Science Support Systems

Michal Stočes, Pavel Šimek, Jan Pavlík

Faculty of Economics and Management, Czech University of Life Sciences Prague, Czech Republic

## Abstract

The article deals with analysis of model data formats suitable for metadata description of digital objects (artifacts) occurring in scientific social network applications. The emphasis of analyzes is on the issue of a metadata description of artefact links to other artifacts and artifact links to individuals. The examined metadata formats include LOM (Learning Objects Metadata), MODS (Metadata Object Description Schema) and DC (Dublin Core). The article also deals with dictionaries of controlled descriptors used to refine and unify the metadata description for agricultural research.

The article presents part of the results of author's dissertation thesis.

## Keywords

Metadata, application profile, social network service for scientists, data sharing, LOM, MODS, DC, AGROVOC.

Stočes, M., Šimek, P. and Pavlík, J. (2017) "Metadata Formats for Data Sharing in Science Support Systems", *AGRIS on-line Papers in Economics and Informatics*, Vol. 9, No. 3, pp. 61 - 69. ISSN 1804-1930. DOI 10.7160/aol.2017.090306.

## Introduction

The development of information and communication technologies has greatly contributed to speeding up publication processes. Before the massive expansion of modern technology, the author of the article had to send his handwriting to the publisher, which he handed over to the writer who prepared the text for the press. After printing, materials had to be delivered to readers. This process could last several days. Now, you can do this with few simple mouse clicks from the convenience of your office, and immediately post the article on the Internet. Such readily published information can be read through the computer network by anyone almost anywhere in the world, immediately after its release. Internet users create their own content. However, this acceleration and simplification of processes also has its negatives. The Internet is overwhelmed by various types of information. The information network creates a jungle which can be hard to navigate. Mislabeled records can be easily lost in vast space of the network. The traceability of specific information depends on how well digital objects are cataloged and shared (Jarolimek and Martinec, 2016 ).

Readers must be alert when selecting relevant information. As the amount of information grows, the quality of the entire network is decreasing. The Internet is full of articles that spread half-truthfully which are not built on consistent and relevant data. For the future development of information networks, it is important to support the quality of information. We do not need quantity but quality! This issue of excess information has permeated many other areas of human activity. This paper focuses on publishing and sharing scientific information.

The phenomenon of recent times, which has significantly changed communication in human society, is social networks. This new communication platform also affected science and research. Each scientific work begins with a thorough study of the current state of the subject. To support this activity, a number of specialized mostly web-based applications have emerged recently. Examples of these applications are Social Network for Scientists, ResearchGate or the VOA3R (Virtual Open Access Agriculture & Aquaculture Repository Project) social network.

Most scientists are forced to work with multiple applications. In each application, the user is prompted to create a profile and upload metadata for each scientific publication. Each application creates its own identifier for the user and his publications. It is difficult for the enriched data created in this system to be exported or transferred from one application to another.

### Methods of writing metadata

The word "metadata" was first used by P. R. O. Badgley in the book "Extension of programming language concepts".

Metadata is most commonly and simply defined as "data about data" or "information about information", but there are many more complex definitions. The definition according to Brand (2003): "Metadata are structured data that describe, explain, localize, and facilitate the easier acquisition, use and management of an information resource." The definition states: "Metadata are structured data - descriptive information about digital objects whose primary purpose is to facilitate search in resources. They include elemental information about primary data, structured according to specific rules and standards, thereby streamlining the management of a large number of objects in data structures." One of many formal definitions states that metadata is data associated with objects that remove the need for pre-existing knowledge of the characteristics of these objects for potential users (Bartošek, 1999).

Metadata itself can be further described by other metadata. Metadata is written according to established rules and therefore machine-readable.

Metadata can be separated into following groups according to Bretherton, (1994):

- *Descriptive metadata* serves to uniquely identify a document. They include, for example, title, author name, keywords etc.
- *Structural metadata* show which parts the document consists of, such as page numbering, chapters, etc.
- *Administrative metadata* includes technical information about the document (format), access rights, etc.

Metadata serves primarily to allow search and help with search-related issues. Among the selected metadata functions are (Bartošek, 1999):

- *Documentation functions* (description of important characteristics of information source),
- *Identification functions* (unambiguous time- and space-independent identification),
- *Search functions* (discovering sources existence and its localization),
- *Selection functions* (selection of sources based on their characteristics).

### Social networking application for scientists

These are applications that support science, they work with metadata of digital artifacts, and expand metadata (enrichment) with other metadata. These applications can be divided into the following groups:

- Social networks for scientists
- Systems managing of references
- Search engines for scientific works

### Digital artifact

It is a digital form of human creation. Digital artifacts can be represented primarily in the form of texts, visualizations or sounds, or combinations of these. The term digital object can also be used.

### Controlled descriptors dictionaries

Controlled dictionaries of descriptors are also referred to as Thesauri. Thesaurus is a reference guide, a kind of dictionary that offers the user a list of synonyms, sometimes also antonyms, and often also defines dependencies between terms.

The descriptors can have defined relations of superiority and inferiority, synonyms and other related terms. In the professional literature, it is described as a controlled and changeable dictionary of descriptor and selection language arranged to explicitly capture relations between lexical units (Easylibrary, 2010).

### Identifier DOI

DOI (Digital Object Identifier) is a centralized commercial system of identifiers for digital works. The DOI is described by ISO 26324 (ISO 26324, 2012) standard. DOI ensures unambiguous identification of the digital document on the Internet and provides a permanent reference to the document. The DOI identifier is the most common and widespread system for identifying scientific publications at present time.

### ORCID identifier

ORCID (Open Research and Contributor ID) is a non-proprietary alphanumeric code that uniquely identifies academic or scientific authors and contributors. It provides people with a lasting identity identification in a similar way the digital object identifiers (DOIs) provide identification for content. ORCID is trying to encompass and merge both ResearchID and Scopus AuthorID. ORCID is managed by a non-profit organization (Nature, 2009).

Another author's identification systems include, for example: The Digital Author Identifier (DAI).

## Materials and methods

### Standard Dublin Core

For the Dublin Core standard, the abbreviation DC is used in the literature. DC is a set of fifteen metadata elements and its main purpose is to facilitate the search for electronic resources. DC was developed by professionals from various fields (computer science, librarianship). The set of DC elements is standardized in accordance with ISO 15836:2009 (International Organization for Standardization), the latest update for 2014 and ANCI / NISO Z39.85 (American National Standards Institute / National Information Standards Organization) from 2007 (ANCI/NISO, 2007). DC is currently maintained by the Dublin Core Metadata Initiative (DCMI) (DCMI, 2016). All of the fifteen metadata elements are optional.

The primary purpose of DC was to describe digital documents published on the Internet directly by the author. For its universal design it has been used by institutions dealing with the formal processing of resources (museums, libraries, universities, etc.). DC can be used to describe both digital and non-digital objects (Dublin Core: Czech, 2006).

### Qualified and unqualified Dublin Core

DC is divided into two types, the so-called Simple Dublin Core or the Unqualified Dublin Core and the Qualified Dublin Core. The simple Dublin Core element values are not limited in any way unlike to the Qualified Dublin Core, where the limitations for element values are specified using qualified terms and qualifiers. Input formats are based on generally accepted standards (Hodge et al., 2005).

### Dublin Core record formats and its elements

Dublin Core metadata entry can be created using two extended formats.

The first option is to write a record in a separate XML format (Extensible Markup Language). For each described digital object, there is one metadata file. This option is used, for example, to describe archive data. In practice, a metadata entry is stored in the database of the appropriate archive, and an XML file is generated for sharing purposes (Taheri and Hariri, 2012).

The second option is to write metadata directly into the described file. The file is usually a web site in HTML (Soundarara et al., 2010). Metadata are then written to the "head" section using the HTML tag "<meta>".

The set of fifteen metadata elements of Dublin Core can be divided into three groups: *source content, intellectual property and source identification* (Celebová, 2013).

### DCMI type dictionaries

The Qualified Dublin Core record utilizes terms of Dublin Core Metadata Initiative (DCMI). DCMI qualifiers include Collection, Dataset, Event, Image, InteractiveResource, MovingImage, PhysicalObject, Service, Software, Sound, StillImage, and Text (DCMI, 2010).

### MARC formats

MARC (MAchine-Readable Cataloging) is a standard consisting of MARC formats (see Table 1) for machine-readable cataloging (a code sample shown in Figure 1). Formats were created in the 1960s at the US Congress Library. The MARC record structure is an implementation of ISO 2709, also known as ANSI / NISO Z39.2. Data content records are defined by other standards, such as AACR2, LCSH, or MeSH. MARC comprehensively solves the problem of machine-readable cataloging, but thanks to obsolete technologies is not currently widely used. The response to technology obsolescence the more current MARC 21 format for more effective cataloging information exchange (Table 1) (Taylor and Joudrey, 2009).

| MARC format | Description |
|---|---|
| **Authorization** | Provides information on individual names, subjects, and titles. |
| **Bibliographic** | Describes the thought and physical properties of bibliographic resources (books, phonograms, sound recordings, etc.). |
| **Classification** | MARC records containing data classification. |
| **Community information** | MARC records describing the source of the provided services. |
| **Ownership** | Provides information about the printout (catalog number, shelf placement, number of pieces, etc.) |

Source: Taylor and Joudrey (2009)

Table 1: MARC – formats.

```
=LDR  05594cam  2200565 i 4500
=001  ocn798437851
=003  OCoLC
=005  20131018101403.0
=006  c\\\\\\\\d\\\\\\\\
=007  cr\\\\\\
=008  120705s2013\\\\mauaf\\\\b\\\\001\0\eng\\
=020  \\$z9780262018555 (hardcover : alk. paper)
=020  \\$z0262018551 (hardcover : alk. paper)
=040  \\$erda
=050  14$aBF503$b.A28 2013
=082  04$a153$223
=245  00$aAction science :$bfoundations of an emerging discipline /$cedited by Wolfgang Prinz, Miriam Beisert, and Arv
=264  \1$aCambridge, Mass. :$bMIT Press,$c[2013]
=264  \4$a©2013
=300  \\$a1 online resource (xi, 450 pages, 5 unnumbered pages of plates) :$billustrations (some color)
=336  \\$atext$btxt$2rdacontent
=337  \\$aunmediated$bn$2rdamedia
=338  \\$aonline resource$bcr$2rdacarrier
=504  \\$aIncludes bibliographical references and index.
=505  0\$aAcknowledgments -- Contributors -- Action science emerging : introduction and leitmotifs / Arvid Herwig, Mir
=520  \\$aThe emerging field of action science is characterized by a diversity of theoretical and methodological appro
=588  \\$aDescription based on print version record.
=650  \0$aMotivation (Psychology)
=650  \0$aCognitive psychology.
=700  1\$aPrinz, Wolfgang,$d1942-
=700  1\$aBeisert, Miriam,$d1980-
=700  1\$aHerwig, Arvid,$d1979-
=776  08$iPrint version: $tAction science.$dCambridge, Mass. : MIT Press, [2013]$w(DLC)  2012024586$w(OCoLC)798437851
=856  40$uhttp://cognet.mit.edu/book-detail/9780262312974$zMITCogNet
```

Source: Abrahamse (2013)

Figure 1: Example of record using the MARC format.

### MODS

Metadata Object Description Scheme (MODS) was developed by experts under the auspices of the US Congress Library and Marc Standard Office as a subset of MARC. The first version of the MODS metadata schema was released in 2002. The schema allows digital libraries to describe any document using a XML file. The MODS schema consists of 20 elements that are taken from other metadata schemas, making the scheme convertible into MARC 21 and Dublin Core (MODS, 2009; Svastova, 2009).

### Data model LOM

Learning Object Metadata (LOM) is a standardized model designed to describe learning objects. The model is defined by the open standard IEEE 1484.12.1 - 2002 created by IEEE (Institute of Electrical and Electronics Engineers Standards Association, New York). The relevant attributes of "learning objects" to be described include: the type of object, author, owner, distribution conditions, format and pedagogical attributes such as learning styles or interactions (Veron et al., 2016).

Learning Object Metadata is a data model, usually encoded in XML, used to describe a learning object and similar digital resources used to support learning. The purpose of object learning metadata is to promote the reusability of learning objects, to facilitate their discovery and to facilitate their interoperability, usually in the context of online Learning Management Systems (LMS) (Stoces et al., 2015).

### Other metadata formats

Other metadata formats include Darwin Core, EBU Core, PBCore, CDWA-Lite/CCO, EAD, DACS, ISAD (G), VRA Core, SPECTRUM In the area of cataloging scientific works and digital libraries, the most widely used is the DC metadata format.

### Identification of artifacts by social networks for scientists

Metadata records from institutional repositories are loaded (harvested) and collected in central repositories. Repository provides search services over metadata entries from institutional repositories. The current trend is to create a social networking site for scientific workers above the repository, which extends the primary search function of the repository to other communication functions. Within social networking applications for scientists, enriching data is generated by users, but it is not shared yet. Social networks and their data are an appropriate complement to LMS - integrating social networks into LMS to enhance the quality of sharing knowledge and communication between users and authors. It also enables users to easily access new knowledge in the field. In doing so, social networks create many metadata that enrich the original records (enriching data). Enriching data has been classified into two groups based on analyses, namely linking metadata and other metadata. The structure and function of other metadata is created by each social network separately. Examples of such data may be comments, ratings, etc. Linking metadata includes the following relations (among others):

- Digital artifact <-> Person (author, co-authors)
- Digital artifact <-> Digital artifact (citation, reference)

### Application profile

When creating a knowledge database that contains object metadata, you need to define its structure. By defining metadata elements, value rules and managed dictionaries, the so-called Application Profile (AP) is created. A large group of Application Profiles uses DC and its elements to describe objects and extends them based on application requirements. For example, the VOA3R Application Profile or EVSKP-MS (Metadata File for Electronic Higher Education Qualifications in the Czech Republic) can be used (Bratkova and Mach, 2008).

"AP is a metadata scheme that consists of metadata elements selected from one or more standard metadata schemas and is designed to allow the application to meet its functional requirements" (Heery and Patel, 2008) .

The European Committee for Standardization (CEN) defines the AP as a set of metadata elements selected from one or more metadata schemes and combined in a compound scheme. Application profiles provide means to express the principles of modularity and extensibility. The purpose of the application profile is to customize or combine existing schemes into a package that is tailored to the specific application's functional requirements while maintaining interoperability with the original schematics.

According to IMS GLC, the reasons for creating new application profiles are as follows:

- Meeting technical and other project requirements that are domain, country or region specific.
- Solving ambiguity and generality in a specification or standard.
- Support semantic interoperability, e.g. using commonly-used dictionaries.
- To facilitate compliance testing and successful collaboration.

## Results and discussion

The following section analyses selected social networking applications for scientists. Systems are analyzed (see Table 2) based on the following aspects:

- Identifier used to identify the author,
- Identifier used to identify the digital artifact,
- Ability to import metadata about digital artifact
- Ability to export metadata about digital artifact

Social networks for scientists do not allow users to export the data created within them. To identify the author they often use their own proprietary identifiers. Most of the systems analyzed allow for export of records, but no additional data such as ORCID identification of the author is enriched.

ORCID and DOI are used as the identifiers by most of the analyzed applications. The multiplicity of author's identity is still a problem, mostly in systems that publish digital artifacts (Mitrovic and Protic, 2014; Brown et al., 2016).

Appropriate metadata models for describing object metadata in social networking environments for scientists according to previous sections include LOM, DC and MODS. In the next section, the issue of describing links to the author and other digital artifacts of selected models will be discussed. The XML data format was used to write metadata in the following examples.

### Link between work and its author

#### Dublin Core

Authors are identified in the *dc:creator* element, which does not contain any extensions (Figure 2).

| Name | author ID | object ID | Import | Export |
|---|---|---|---|---|
| **Mendeley** | Scopus author ID, ORCID | DOI and others | yes, various | yes, various |
| **ResearchID/EndNote** | ORCID, researchID | DOI and others | yes, various | yes, various |
| **Google Scholar** | own | own | no | yes, various |
| **ResearchGate** | own | DOI, own | yes, various | no |
| **Academia edu** | own | own | no | no |
| **ORCID** | ORCID, Scopus author ID, researchID | According to database import | yes, various | yes, BibTex |

Source: own processing

Table 2: Analysis of applications for scientific support.

```
<dc:creator>Michal Stočes</dc:creator>
```

Source: own processing

Figure 2: dc:creator.

### Data model LOM

The problem of linking to a person is solved very complexly in the LOM model using the breakage element:lifeCycle and its descendants. There is no possibility of more detailed identification of the author in the standard (Figure 3).

```
<lom:lifeCycle>
    <lom:contribute >
        <lom:role>
            <lom:source>LOMv1.0</lom:source>
            <lom:value>Author</lom:value>
        </lom:role>
        <lom:entite>
            Michal Stočes
        </lom:entite>
    </lom:contribute >
</lom:lifeCycle>
```

Source: own processing

Figure 3: lom:lifeCycle.

## MODS schema

The MODS schema contains a description that allows a reference to specific author, through, the mods:name element. (Figure 4)

```
<mods:name type="personal">
    <mods:namePart type="family">Michal</mods:namePart>
    <mods:namePart type="given">Stočes</mods:namePart>
    <mods:role>
        <mods:roleTerm type="code" authority="marcrelator">
            aut
        </mods:roleTerm>
        <mods:roleTerm type="text" authority="marcrelator">
            Author
        </mods:roleTerm>
    </mods:role>
</mods:name>
```

Source: own processing

Figure 4: mods:name.

## Link between works (reference, citation)

In all three models, the difference between a reference and a citation is made using the references/isReferedBy quantifier.

### Dublin Core

As a record of metadata identifying the reference, Qualified Dublin Core is used. In DC standard there is no clear procedure for writing a link to a specific place in the document (Figure 5).

```
<dcterms:references xsi:type="dcterms:URI">
    http://doi:10.17221/313/2015
</dcterms:references >
```

Source: own processing

Figure 5: dcterms:references.

### Data model LOM

The IEEE LOM standard has the same drawbacks as DC - there is no unambiguous procedure for writing a reference to a specific location in the document (Figure 6).

```
<lom:relation>
    <lom:kind>
        <lom:source>LOMv1.0</lom:source>
        <lom:value>references</lom:value>
    </lom:kind>
    <lom:resource>
        <lom:identifier>
            <lom:catalog>URI</lom:catalog>
            <lom:entry>
                http://dx.doi.org/10.7160/aol.2016.080108
            </lom:entry>
        </lom:identifier>
    </lom:resource>
</lom:relation>
```

Source: own processing

Figure 6: lom:relation.

### MODS schema

In its definition the MODS schema contains a link to a specific place in the document. Further, the mods:identifier element has a type property that does not specify what values it can take, making MODS very universal in terms of unambiguous identification of the work (Figure 7).

```
<mods:relatedItem type="references">
    <mods:identifier type="doi">
        10.7160/aol.2016.080108
    </mods:identifier>
    <mods:part>
        < mods:extent unit="pages">
            <mods:start>85</mods:start>
            <mods:end>86</mods:end>
        </mods:extent>
    </mods:part>
</mods:relatedItem>
```

Source: own processing

Figure 7: mods:identifier.

## Mapping between metadata models

The following tables (3 and 4) show examples of mapping between metadata models.

| Dublin Core (DC) element | Learning object metadata (LOM)elementy |
|---|---|
| dc:identifier | /lom/general/identifier/entry |
| dc:title | /lom/general/title |
| dc:language | /lom/general/language |
| dc:description | /lom/general/description |
| dc:subject | /lom/general/keyword *nebo* /lom/classification *s* /lom/classification/purpose *equals* "discipline" *or* "idea". |
| dc:coverage | /lom/general/coverage |
| dc:type | /lom/educational/learningResourceType |
| dc:date | /lom/lifeCycle/contribute/date *when* /lom/lifeCycle/contribute/role *equals* "publisher". |
| dc:creator | /lom/lifeCycle/contribute/entity *when* /lom/lifeCycle/contribute/role *equals* "author". |
| dc:otherContributor | /lom/lifeCycle/contribute/entity *with contribution type in* /lom/lifeCycle/contribute/role |
| dc:publisher | /lom/lifeCycle/contribute/entity when /lom/lifeCycle/contribute/role equals "publisher". |
| dc:format | /lom/technical/format |
| dc:rights | /lom/rights/description |
| dc:relation | /lom/relation/resource/description |
| dc:source | /lom/relation/resource *when* /lom/relation/kind equals "isBasedOn". |

Source: own processing

Table 3: Mapping between DC a LOM.

| MODS elementy | DC elementy |
|---|---|
| <titleInfo><title> | Title |
| <name><namePart> | Creator, Contributor |
| <subject><br><topic><br><classification><br><name><br><occupation> | Subject |
| <abstract><br><note><br><tableOfContents> | Description |
| <originInfo><publisher> | Publisher |
| <originInfo><dateIssued><br><originInfo><dateCreated><br><originInfo><dateCaptured><br><originInfo><dateOther> | Date |
| <typeOfResource><genre> | Type |
| <physicalDescription><internetMediaType><br><extent><form> | Format |
| <identifier><location> <url> | Identifier |
| <language><languageTerm> | Language |
| <relatedItem> | Relation |
| <subject><br><geographic><br><temporal><br><hierarchicalGeographic><br><cartographics> | Coverage |
| <accessCondition> | Rights |

Source: own processing

Table 4: mapping between DC a MODS.

## Thesaurus AGROVOC

AGROVOC is an extensive thesaurus developed by researchers of the Food and Agriculture Organization (FAO), which is a United Nations (UN) specialized agency within the AIMS (Agricultural Information Management Standards). AGROVOC contains terms from food, nutrition, agriculture, fisheries, forestry and the environment. Thesaurus contains over 32,000 terms in 23 languages (April 2017) including the Czech language. This entire thesaurus is expressed as a Simple Knowledge Organization System (SKOS) and published as Linked Data, a data model for representing structured dictionaries. Conceptual scheme of the thesaurus AGROVOC uses three levels of display:

- terms have abstraction meanings and are also often described using the Uniform Resource Identifier (URI), e.g. for beef is used: *http://aims.fao.org/aos/agrovoc/c_861*
- terms specified by language, for instance.: حوم البقر (*Arabic*), *牛肉* (*Chinese*), hovězí maso (*Czech*), Viande bovine (*French*).
- terms have specific options (ranges) such as spelling variants or singular and plural numbers, e.g.: hen, chicken, poultry, cow, bull, cattle, etc.

This system provides for terminological relations between concepts and specific meaning. AGROVOC is thus well-suited to describe, for example, scientific research articles, expert articles, information or news from the agrarian sector, audiovisual data, etc. (Simek et al., 2013a; Masner et al. 2016).

The Czech version was prepared by the Institute of Agricultural and Food Information in 1995 and 1996 as part of the project „Czech version of the AGROVOC thesaurus“. It was a prerequisite for the creation of a national agricultural information system commissioned by the Ministry of Agriculture of the Czech Republic. Since 1997, the Czech Agriculture and Food Bibliography has been used in the processing of records in bibliography articles. (Simek et al., 2013b; Beneventano, et al., 2016).

## Conclusion

Selected DC, LOM and MODS metadata models were analyzed and the following conclusions and recommendations were found: DC is the appropriate format for writing basic metadata. This is due to its versatility and modifiability. The basic set of 14 elements is precisely defined but can be further extended by qualifiers to meet the demands and needs of different social networks for scientists. The LOM standard, complemented by the MODS element identifier, is suitable for describing links to people. Standart LOM defines a lifeCycle element that contains a comprehensive metadata entry to describe people. Entering an identifier element from MODS model extends its definition. The MODS model is suitable for describing the links to the digital artifact. The identifier element from MODS can be extended by adding attributes for various types of digital artifact description.. Schemas can be mapped to each other to allow transformation between them. The unequal identity of the author is still a problem, mostly in systems that publish digital artifacts. Identifiers that are expanding and being increasingly exploited are ORCID - Author Identification and DOI - Digital Object Identification. ORCID tries to join the two proprietary identifiers researchID and Scopus Author ID. For the proposed methodology it is recommended to provide all available identifiers to improve the resulting record. A suitable addition to metadata entries is the use of keywords from controlled dictionaries of descriptors, for the area of agriculture there is a large thesaurus AGROVOC developed by the FAO.

## Acknowledgements

The knowledge and data presented in the present paper were obtained as a result of the Grant No. 20171019 of the Internal Grant Agency titled "Options semantic and efficient storage of research results for subsequent presentation and sharing in heterogeneous environments of large networks".

*Corresponding author:*
*Ing. Michal Stočes, Ph.D.*
*Department of Information Technologies, Faculty of Economics and Management*
*Czech University of Life Sciences Prague, Kamýcká 129, 165 00 Prague, Czech Republic*
*Phone: +00420 224 382 277, E-mail: stoces@pef.czu.cz*

## References

[1] Abrahamse, B. (2013) “*Getting Started with Marc Edit*“, New England Library Association Fall 2013, Oct 2013. [Online]. Available: https://netsl.files.wordpress.com/2013/04/getting-started-with-marcedit-tuesday-nela-2013.pdf [Accessed: 19 Jun. 2017].

[2] ANSI/NISO Z39.85-2007 (2017) “Dublin Core Metadata Element Set”. ISSN 1041-5635.

[3] Bartošek, M. (1999) “Vyhledávání v Internetu a DUBLIN CORE“ (In Czech), *Zpravodaj ÚVT MU*, Vol 9. IX, No. 4, pp. 1-4. ISSN 1212-0901.

[4] Beneventano, D., Bergamaschi, S. and Martoglia, R. (2016) "Exploiting semantics for searching agricultural bibliographic data", *Journal of Information Science*, Vol. 24, No. 6, pp. 748-762. ISSN 0165-5515. DOI 10.1177/0165551515606579

[5] Brand, A., Daly, F. and Meyers, B. (2003) ”*Metadata demystified*”, Bethesda NISO Press. ISBN 1-880124-59-9.

[6] Bratkova, E. and Mach, J. (2008) “*Metadatový soubor pro elektronické vysokoškolské kvalifikační práce v ČR*” (in Czech), Asociace knihoven vysokých škol ČR. [Online]. Available: http://www.evskp.cz/standardy/evskp/ [Accessed: 10 Jun. 2017].

[7] Bretherton, P. F. and Singley, P. T. (1994) "Metadata: a user's view." In P*roceedings of the 7th international conference on Scientific and Statistical Database Management (SSDBM'1994)*, James C. French and Hans Hinterberger (Eds.), IEEE Computer Society, Washington, DC, USA, 1994, pp. 166-174.

[8] Brown, J., Demeranville, T. and Meadows, A. (2016) "Open Access in Context: Connecting Authors, Publications and Workflows Using ORCID Identifiers", *Publications 4*, Vol. 4, No. 4, Article Number: 30. ISSN 2304-6775. DOI 10.3390/publications4040030.

[9] "*Dublin Core : Czech*" (2006) [Online]. Available: http://www.ics.muni.cz/dublin_core/index.html. [Accessed: 10 Jun. 2017].

[10] Heery, R. and Patel, M. (2000) "Application Profiles: Mixing and Matching Metadata Schemas,"*ARIADNE*, Vol. 25, John Wiley & Sons.

[11] Hodge, G., Templeton, C. and Allen, R. (2005) "A metadata element set for project documentation", *Science and Technology Libraries*, Vol. 25, No.4, pp. 5-23. DOI 10.1300/J122v25n04_02.

[12] ISO 26324:2012 (2012) "Information and documentation - Digital object identifier system". [Online]. Available: https://www.iso.org/standard/43506.html [Accessed: 16 Jun. 2017].

[13] Jarolimek J. and Martinec, R. (2016) "Analysis of Open Data Availability in Czech Republic Agrarian Sector", *AGRIS on-line Papers in Economics and Informatics*, Vol. 8, No. 3, pp. 57 - 67. ISSN 1804-1930. DOI 10.7160/aol.2016.080306.

[14] Masner, J., Vanek, J. and Jarolimek, J. (2016) "Prototype of a Content Creation and Updating Application Module for Agrarian Sector and Regional Development", *Scientific conference Agrarian Perspectives - Global and European Challenges for Food Production*, *Agribusiness and the Rural Economy*, Sep 14-16, 2016, Czech University of Life Sciences Prague. ISSN 2464-4781.

[15] Mitrovic, I. and Protic, J., (2014) "Problems with affiliations, names and personal identity in the proces of evaluating higher education institutions", *6th International Conference on Education and New Learning Technologies"*, Jun 2014, Barcelona. ISSN 978-84-617-0557-3.

[16] Nature, (2009) "Credit where credit is due", *Nature*, Vol. 462, No. 7275, pp. 825. [Online]. Available: http://www.nature.com/doifinder/10.1038/462825a [Accessed: 13 Jun. 2017]. ISSN 0028-0836. DOI 10.1038/462825a.

[17] Simek, P., Stoces, M., Vanek, J., Jarolimek, J., Masner, J. and Hrbek, I. (2013b) "Using of Automatic Metadata Providing", *AGRIS on-line Papers in Economics and Informatics*, Vol. 5, No. 4, pp. 189-197. ISSN 1804-1930.

[18] Simek, P., Vanek, J., Jarolimek, J., Stoces, M. and Vogeltanzovw, T. (2013a) "Using metadata formats and AGROVOC thesaurus for data description in the agrarian sector", *Plant, Soil and Environment*, Vol 59, No. 8, pp. 378-384. ISSN 1214-1178.

[19] Stoces, M., Masner, J. and Jarolimek, J. (2015) "Mitigation of Social Exclusion in Regions and Rural Areas – E-learning with Focus on Content Creation and Evaluation", *AGRIS on-line Papers in Economics and Informatics*, Vol 7, No. 4, pp. 143-150. ISSN 1804-1930.

[20] Svastova, P. (2009) "*MODS a další metadatová schémata v oblasti digitalizace dokumentů*" (in Czech), Knihovny současnosti 2009, Vol 17, pp. 127 - 132. ISBN 978-80-86249-54-4.

[21] Taheri, S. M. and Hariri, N., A. (2012) "A comparative study on the indexing and ranking of the content objects including the MARCXML and Dublin Core's metadata elements by general search engines", *The Electronic Library*, Vol. 30, No: 4, pp. 480-491, DOI 10.1108/02640471211252193.

[22] Taylor, A. G., Joudrey, D. N. (2009) "*The organization of information*", 3rd ed. Westport, Conn.: Libraries Unlimited, Vol. 16, pp. 512. ISBN 159158700x.

[23] Veron, S., Celeste, V., Alejandra, A. M., Milagros de los , G. M. (2016) "An interoperability model based on ontologies for learning object repositories", *2016 International Symposium on Computers in Education (SIIE)*, Salamanca, pp. 1-6. DOI 10.1109/SIIE.2016.7751843.