A Game by Tjasa Jug

METADVENTURE

From Storyteller to Bestseller

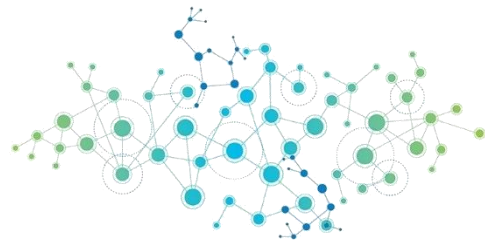## Player's guide
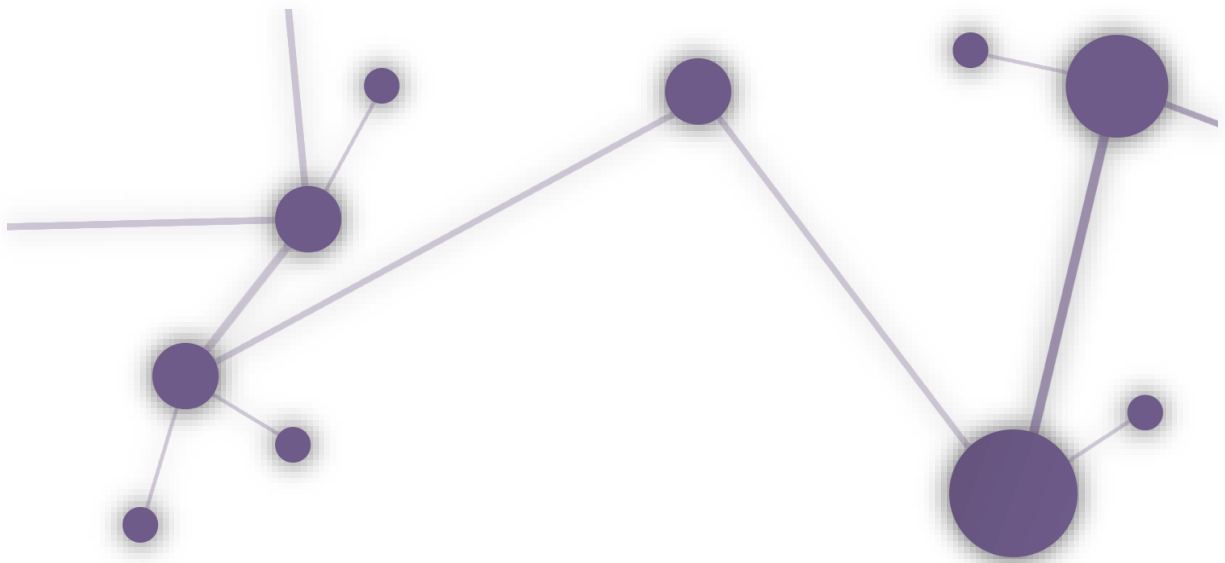
METADVENTURE

Contents:

HOW TO PLAY

In this game you are a publisher who wants to publish a book with a great content. The World Wide Web offers an excellent opportunity to market and present your title. But be aware! Your book could get lost in the jungle of different books. To avoid this, you should use rich and high quality metadata for people to find, identify and select your book.

Is your book going to stay unnoticed or will it become a bestseller?

In this game, like in real life, you can win by luck or by using your superior metadata skills. Try to be the first to reach the finish and make your book a bestseller!

Instructions:

1. Roll a six-sided die
2. Move
3. Follow the instructions on the playing board
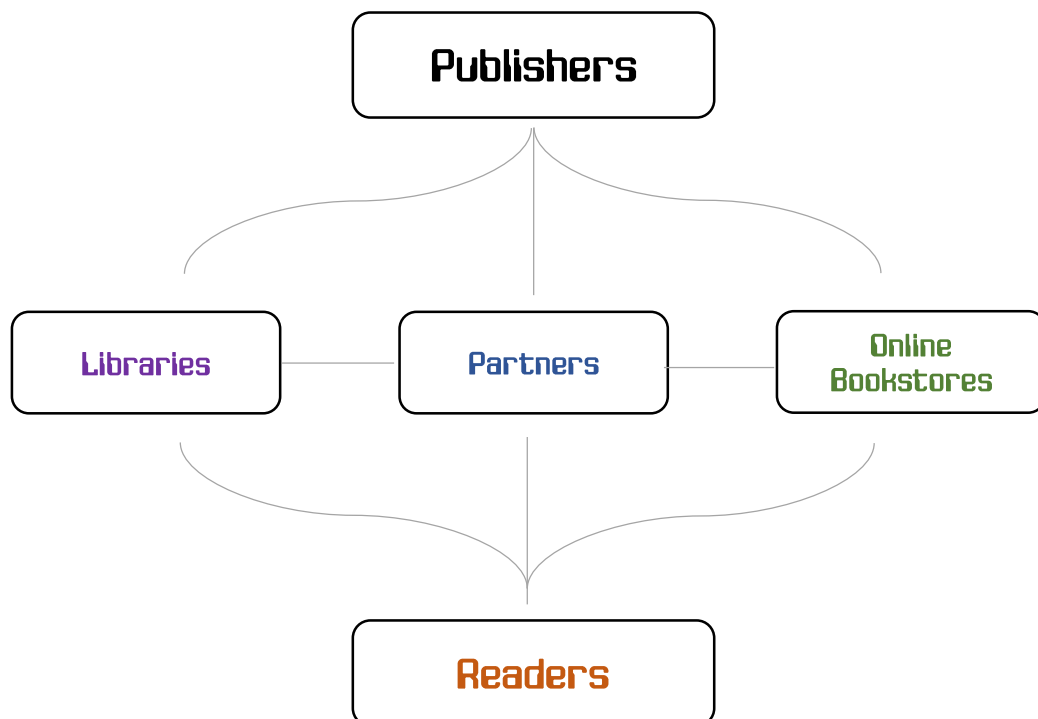4. If interested, read the explanation of spaces in this booklet

BACKGROUND

Metadata or data about data is structured information that describes, explains, locates, or otherwise makes it easier to retrieve, use, or manage an information resource. Various communities have adopted their own explanations and uses of this term. In the book market context, the term metadata means structured data presenting characteristics of books in order to help identify, discover, assess, and manage them. We are talking about bibliographic data, namely attributes of a book, such as author, title, publisher, description, physical appearance, etc.

Good book metadata is very important for different users on the book market. It enables publishers and librarians to manage some work processes before and after publication, it serves booksellers for marketing and sales, while readers use it to find, identify and obtain books.

Metadata is present in every stage of the book lifecycle. Data about a book is firstly created by publishers and is later distributed to the supply chain resellers and trading partners. End-users encounter book metadata in online bookstores and library catalogues, where they do not only see the data provided, but also create new metadata in the form of user behavior logs, reviews, and tags. However, actual practice depends on the individual publisher and book market, where some organizations or processes may be missing or are replaced by other procedures.

METADATA LIFECYCLE

```
                    ┌─────────────────┐
                    │   Publishers    │
                    └─────────────────┘
                             │
        ┌────────────────────┼────────────────────┐
        │                    │                    │
┌───────────────┐    ┌───────────────┐    ┌───────────────┐
│   Libraries   │────│   Partners    │────│    Online     │
│               │    │               │    │  Bookstores   │
└───────────────┘    └───────────────┘    └───────────────┘
        │                    │                    │
        └────────────────────┼────────────────────┘
                             │
                    ┌─────────────────┐
                    │    Readers      │
                    └─────────────────┘
```

LIBRARIES

Get Your Metadata out
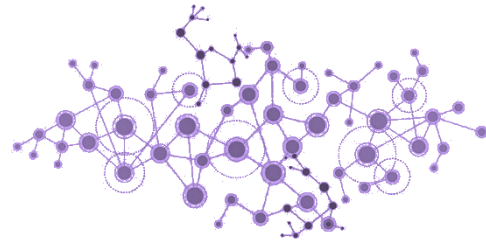
Libraries play an important role in book metadata lifecycle. They usually represent the first organization that receives book metadata, as they are responsible for the CIP creation and ISBN assignment. They can help promote books and offer enrichment of publisher's metadata description. In some countries other organizations may perform such tasks, usually for a certain price.

CIP CREATION
(Space 2)

Acronym CIP stands for Cataloging in Publication. It represents book data prepared by the national library (or an institution with a similar mission) prior to publication and is later printed on the verso of the book's title page (Library of Congress, 2017). It contains the basic metadata for book identification, such as title, author, edition, publisher, identifier, series, subjects, physical description, etc. (British Library, 2017). When the book is published, the publisher includes the CIP data on the copyright page to facilitate book processing for libraries and trading partners (Library of Congress, 2017). Its purpose is to offer an early information about publication and therefore accelerate book information flow and accessibility on the market, it enables a book to be present in the databases and aggregators, it standardizes procedures which also reduces time for preparing bibliographic descriptions, and helps booksellers to shelve books in the appropriate thematic sectors (Kanič, 2014). Therefore, every player on the book market can benefit from the CIP data. Libraries save in time and costs needed for description creation, publishers and booksellers can identify appropriate topics of their books and offer them to the relevant customers, who can also inspect CIP to learn about basic book information, which can lead to book selection (D. K. Agencies, 2017). In Slovenia, every publication should have a CIP, which is prepared by our National Library. Since 2016, Slovenian libraries are using a platform for automated e-book cataloging and CIP creation called CatSi. Its main goal is to optimize the cataloging process, facilitate book metadata management, and enrich the Slovenian union catalogue.

The most frequent problem with CIP occurs when a publisher changes essential data, such as wording of the title, after the publication of the book. That means that CIP data printed in the item and pre-publication record in the catalogues are incorrect and are therefore useless to subsequent cataloging agencies. This can lead to problems identifying and locating specific publication or its version.

Libraries play an important role in the book's life cycle. While we already carried out various studies with other players on the book market, libraries are still on our to-do list. In our future research, we would like to interview catalogers of the National Library of Slovenia and University Library of Maribor to obtain information about barriers regarding CIP creation. We are mostly interested in accuracy and completeness of metadata sent by publishers and in how catalogers cope with errors and incompleteness of the received data.

ISBN
(Space 5)

The International Standard Book Number (ISBN) is a 10 or 13 digit number that uniquely identifies books and book-like products published internationally. Each ISBN consist of 5 elements, representing country, publisher, version or format, and check digits of the specific book and therefore also serves as an EAN code (International ISBN Agency, 2017). Similar to CIP, ISBN should be assigned to the book before its publication and is later printed on the back of the book. There are over 160 ISBN agencies worldwide, which in some countries charge a fee for the ISBN creation or can assign it for free (ISBN.org, 2017). The purpose of this number is to uniquely identify and distinguish publications and their versions, allowing more efficient marketing of products by booksellers, libraries, universities, wholesalers, and distributors. Such identification is also important to the online buyers, who often need an ISBN to be sure they have purchased the correct title and an appropriate edition (Carpenter, 2010). WorldCat, the OCLC's Online Union Catalog, is also providing the xISBN service which enables the end user to retrieve a list of ISBNs associated with a submitted ISBN and therefore displays information about other versions of an individual work (WorldCat, 2017).

Some platforms allow authors and publishers to bypass the ISBN registration process, resulting in the failure of these titles to be listed in the mayor industry databases of available books and in catalogues (Register & Mcllroy, 2015) and therefore cannot be included in the book market network (NUK, 2017). What is more, different formats of a publication that are published and made separately available should be given a separate ISBN (ISBN User's Manual, 2012). However, some publishers are assigning the same ISBN to different versions of the book, which makes it impossible to identify and to distinguish between different formats and editions in which they appear (Carpenter, 2010).

In Slovenia, ISBN assignment goes hand in hand with CIP creation, since this is an activity carried out by libraries. As we mentioned before, we would like to interview catalogers responsible for creating CIP and identifiers and ask them whether they encounter any problems regarding ISBN assignment. In order to obtain information from another perspective, we would also like to interview publishers and learn what their practices are regarding ISBNs. For example, when they publish a new edition or a version of their book in a different format, do they always apply for a new ISBN or do they use the one assigned to the previous version?

AUTOMATIC CONVERSIONS
(Space 8)

Depending on their needs, different disciplines and organizations on the book market have settled on various metadata standards and schemas for data organization, which is problematic from the view of metadata exchange. Unlike publishers, who often describe and send book information in ONline Information eXchang (ONIX) (ONIX: Overview, 2017) or other internal standards, libraries are most commonly using MAchine-Readable Cataloging (MARC) format (MARC, 2017). Although these standards are structurally and semantically different, as they support different needs and communities of practice, they have similarities because they both ultimately serve to connect published work with their audience (Godby, 2010). It would therefore be useful for libraries to use metadata created by publishers in order to reduce costs and time associated with cataloguing and thus remove redundant work between publishers and libraries (Debus-López et al., 2012). Some organizations have already designed various ONIX-to-MARC mapping tools (Godby, 2012; Debus-López et al. 2012) that are able to extract MARC bibliographic description directly from the publisher-supplied ONIX (Debus-López et al. 2012). Crosswalks between these standards can also be used to facilitate the creation of CIP and to provide publishers with an XML feed of MARC description (Luther, 2009). The current use of ONIX-to-MARC converters have resulted in reduction of the time spent on creating library records while providing enriched descriptions for the greater library community (Debus-López et al. 2012). Since 2016 Slovenian publishers can also submit their e-book metadata through an online form in the system called CatSi, which automatically generates a catalogue record and CIP. The vision of CatSi is to become the basis of the future system for publishers to perform all legally necessary procedures regarding e-books, such as entering book data needed for cataloging, submitting the legal deposit copy of a book, introducing information about new titles in the Slovenian book aggregator's website, etc. (CatSi, 2016).

In our planned research with catalogers, we would also like to ask them about their experience with their new platform CatSi. We are interested if the automated process actually helps them to reduce time needed for cataloging and how much time and manual work is needed to correct generated records. Since this system is designed for publishers, their evaluation is also needed. In our past research we already conducted an interview with publishers, however, we did not get all the answers that we would like to obtain. Therefore, we would like to expand our study and conduct another round of interviews with publishers of e-books and ask them about their experience and opinions of the CatSi platform.

COLLABORATION
(Space 10)

Publishers, distributors and authors are encouraged to closely cooperate with the library sector as this leads to more complete and consistent metadata records and to a better visibility of the title on the book market (Cato & Haapamäki, 2009). Although publishers and libraries have different objectives when describing books, each can benefit from the other (Sanchez, 2011). Records originating from publishers and booksellers contain basic book information, such as authoritative information about titles, authors, publisher names, as well as some additional information that encourages sales, such as summaries, cover art, author biographies, links to reviews, sample chapters, tables of contents, and other useful material that a MARC record typically lacks (Tennant, 2006). However, publishers usually do not offer other elements, important for bibliographic description and access, and often create records that are incorrect or are not consistent in quality (Sanchez, 2011). Meanwhile libraries create valuable information and context for users to find books, for example different classifications, detailed subject headings, varying forms of the title, and authority-controlled names of authors (Godby, 2010). Thus, if data from libraries and publishers could be integrated, essential information could be transferred easily between the two metadata formats and the result might be an enhanced record that both communities would consider more valuable (Godby, 2010). OCLC has recognized this benefit and launched a project called Metadata Services for Publishers where the system accepts book data in ONIX and automatically adds it to WorldCat for immediate use by libraries. In return, some library metadata is added, making the titles more visible to the end users on the web and in online bookstores. In that manner, collaboration between publishers and libraries helps to reduce time, cost, and duplication in creating book descriptions, it increases granularity and consistency of metadata for better discovery, provides enhanced data for increase in selection of a book and opens additional channels for exposure of title metadata (OCLC Metadata Services for Publishers, 2010).

Although collaboration between publishers and libraries is recommended for better book metadata description, in Slovenia this is not the usual practice. Libraries create a bibliographic record for publishers who can later copy this data from the online catalogue. However, with CatSi and its potential to become the central system for publishers to distribute metadata, the situation can change. In our interviews with publishers and libraries, we would like to obtain their opinion and expectations about the future of this system and ask them what more could be done for better communication and collaboration between them.
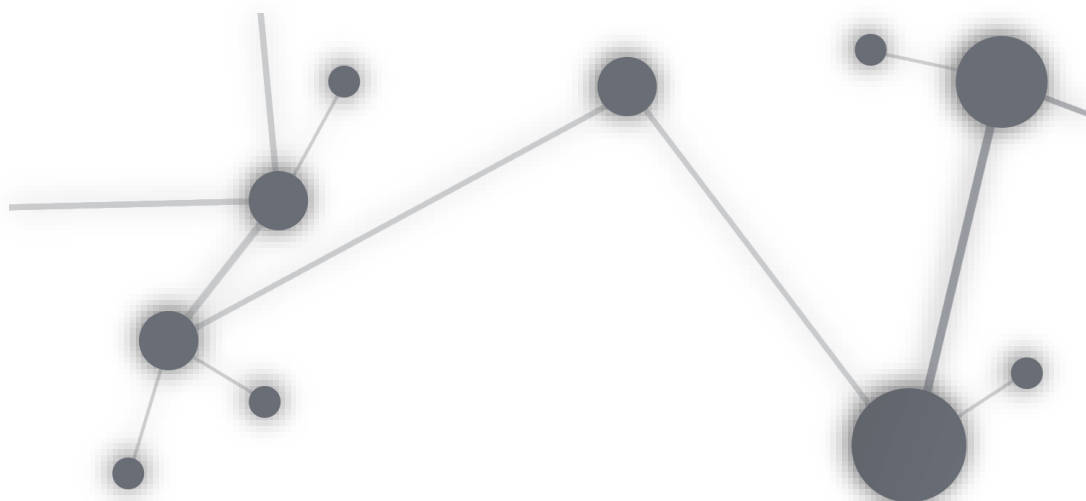
PARTNERS

Share Your Metadata
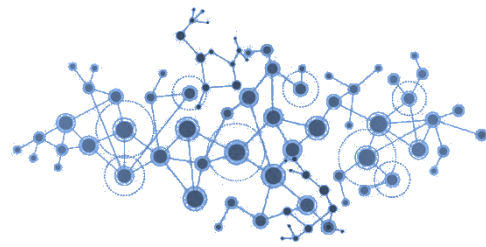
Publisher's trading partners are organizations, such as distributors, metadata aggregators, wholesalers, etc., that serve as intermediaries between publisher, libraries, and retailers. Besides books they also distribute the publisher's metadata which is later exposed in online bookstores and other book-related platforms. Different business partners might require different book information. To meet their individual needs, publishers should create and provide rich and quality metadata.

USING UNIVERSAL STANDARDS
(Space 14)

After leaving the publishing house, book data is submitted electronically to supply chain trading partners. Each of them has their own specification for how book data is submitted and usually requires certain information to be present in the record (Register & Mcllory, 2015). Besides that, there is also a great concern about the increased volume and quality of metadata shared among different stakeholders (Luther, 2009). As a solution to these problems, international standard ONIX for books was developed in 2000. As a communication format, ONIX makes it possible to deliver rich product information to all partners, from wholesalers, distributors and retailers to libraries and data aggregators. In many cases, a single data feed can be made suitable for all recipients which greatly reduces support costs, as publishers no longer need to provide data in many unique formats. ONIX is also providing a content and structural template for the quality records bringing together all the metadata needed for the description and promotion of new titles (ONIX: Overview, 2017).

ONIX for books is used by many organizations, including leading US publishers, libraries, Amazon, and Google (ONIX: Users and Services directory, 2017). However, midsized and small publishers are still using various tools for metadata maintenance, such as Excel files, proprietary formats, online forms, and print catalogs that require lots of manual intervention and are therefore susceptible to errors (Luther, 2009; Breznik Močnik, Trplan & Milohnić, 2013; Register & Mcllroy, 2015). The main problem is that the cost of metadata standardization is very high for small publishers who are also not quite aware of the long-term positive impact of using universal standard (Beky, 2004; Wayne, 2005). On the other hand, small and independent publishers have an advantage in competing with larger publishers by getting metadata right as it is easier to maintain smaller number of titles (Register & Mcllroy, 2015). By using ONIX, they could reach the worldwide marketplace equally as larger publishers, which could make them competitive and could have a great effect on their book sales (Beky, 2004; Lichtenberg, 2014).

In our research we wanted to obtain information about current publisher metadata situation on the Slovenian book market, including metadata standards. Therefore, we conducted an online interview with people working with metadata in three mayor Slovenian publishing houses. We were mostly interested in barriers and difficulties they face while creating and sharing book metadata. It was evident from their responses that Slovenian publishers do not use universal book metadata standards and that they rather rely on their internal schemas, Excel sheets, online forms etc. Although they are aware of the problems regarding non-standardized systems, they are not sure who should initiate and fund metadata standardization in our country.

COMPLETENESS OF DATA
(Space 17)

With the rise of the web, quality and accurate metadata has become a marketing tool for publishers, a shopping guide for consumers and an absolute necessity for distributors and retailers (Dawson, 2012). In the current discovery environment, it is difficult to measure what is not found and to quantify the impact and cost of poor, incomplete and missing metadata on business, which ultimately affects consumers (Luther, 2009). We should keep in mind that metadata is often the only marketing tool for digital products since users do not discover the book itself, but rather the information about it. If more information is available, the customers are more likely to find and select what they need (Dawson, 2012). Ideally, each publisher would maintain one comprehensive database that contains the entire book data needed for the various recipients. Even if this is not the case, metadata distributed into the marketplace must meet the basic needs of online e-commerce. A good book record should include metadata for content description and enhanced metadata that encourage purchasing decision, and metadata for product description needed for the identification of a specific title or version of a book. It should also contain metadata for commerce that enable business transactions. Even if a publisher meets those requirements, during the book's life cycle the metadata is still frequently changed and edited to be useful for different target groups. (Register & Mcllroy, 2015). The problem occurs when the book data becomes inaccurate due to changes made without consent or knowledge of the publisher. This can happen if metadata recipients frequently receive incomplete or incorrect metadata from a wide cross section of publishers and are trying to improve their records without full knowledge about all titles. (Development, Use, and Modification…, 2012). By providing complete and quality metadata, publishers ensure that correct information is distributed to their trading partners and consequently to the end users.

To get an overview of metadata completeness provided by publishers, we should get an insight in book records they send to their partners. Since we do not have access to their internal and confidential data, we had to examine the requirements and online forms provided by organizations to which publishers must sent their book information. We discovered that libraries and other organizations mostly require only basic book information crucial for book identification. We also conducted an expert study of online bookstores (which is described in more detail in the next chapter of this booklet) and found that publishers and online booksellers usually display only basic book information and that they generally do not provide additional book information that would encourage selection of their books.

METADATA QUALITY
(Space 20)

Complete, timely, and well-organized metadata can be very effective for finding and obtaining a resource while inaccurate product information may cause various problems, even if the publisher is providing a rich set of metadata. For example, a buyer may not find what he or she is searching or may purchase a wrong product if the title is not written correctly. In addition, incorrect data, such as wrong cover picture, target audience or related works without a clear connection, may be crucial for the book to be overlooked or may create a bad shopping experience. What is more, outdated metadata can cost booksellers sales. For example, each year thousands of orders are cancelled because metadata records indicate that the book is out of stock or there is no distributor listed (Luther, 2009). Book metadata should be updated with important changes or additions to the title information and should receive further updates to reflect important events or activities relating to the book after its publication (Register & Mcllroy, 2015). It should also be consistent, since this is crucial for a good shopping experience of end users (McKay et al., 2012). As we mentioned before, errors, and incompleteness of metadata record can be the result of different schemas and standards used by organizations in the book industry. Even if organizations are using ONIX, this format does not guarantee that metadata is correct, timely and accurate (Luther, 2009). BISG is trying to solve this problem by providing documents such as Best Practices for Product Metadata (2015), that present best metadata examples, form in which book information should be presented and a recommended set of metadata. The purpose of establishing best practices is to encourage quality, promote standardization, and create efficiencies in business practices (Register & Mcllroy, 2015). To encourage publishers to create quality metadata book industry organizations in USA (BISG, 2017), Canada (BookNet, 2017), and UK (BIC, 2017) developed certification programs with the goal of increasing metadata quality across the book industry.

While analyzing the results from the interviews with Slovenian publishers we noticed some information from which we could also assess the quality of their book metadata. They reported that they do not use any universal standards and that they are frequently creating a description for the same book multiple times. They also stated that they do not have enough metadata knowledge. That does not necessary mean that their metadata is not good, only that they do not use effective tools for metadata creation which can lead to errors. However, our study on metadata presence in online bookstores revealed that there is a great inconsistency in metadata set and form not only among different bookstores but also within individual shops. Booksellers should avoid displaying the same type of book information in different forms since this could negatively affect the performance of searching tools and can cause bad shopping experience.

BOOKSTORES

Display Your Metadata
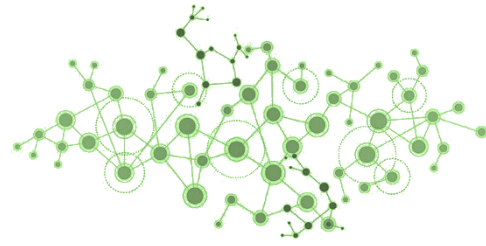
Online bookstores enable publishers to expose all of their titles, including those that are not promoted through other channels. Good metadata makes it easier to search, find, identify and obtain books, and may have a great impact on book sales. Innovative ways of metadata use, such as filtering and recommendations of similar or related works, optimizes user experience and keeps readers engaged.

SEARCH AND BROWSE
(Space 23)

Searches in a bookstore's search engine are directed specifically at the seller's database of information about books. Publishers contribute this metadata through data feeds to sellers, therefore search results will depend on the quality of provided metadata. This data is central for search engine optimization (SEO) for all books that are sold online. Taking full advantage of metadata options for product description helps to ensure results that are truly relevant to readers and keeps readers engaged with the information they find (Register & Mcllroy, 2015). Although people are used to simple search they know from Google, advanced search returns more accurate results containing the required attributes. If the elements of book metadata are not available, are not organized or are indexed incorrectly, then there will not be good retrieval and that book will not see the optimal usage (Wiley, 2015). However, since consumers most frequently discover books by browsing (Walter, 2016), a good textual description should be enriched with connections and information about related products should be provided. Many online bookstores also enable faceted browsing which offers a user relevant subcategories, so they can get an overview of the results and narrow their result list. Facets are very useful since they help users to avoid dead ends, they add to success in finding relevant results, and provide an overview and variety of the offered collection (Fagan, 2010).

Outside online bookstores, also major search engines can retrieve results pointing to books in online shops. It is not possible to control search engine algorithms, but publishers can make sure that distributed metadata is as rich as possible. While there is no guarantee that complete and quality metadata will result in book's discovery, the absence of this data will definitely guarantee its obscurity (Register & Mcllroy, 2015).

In our study where we were examining the presence of metadata elements in 33 Slovenian and 33 American online bookstores, we obtained interesting results about searching and browsing tools provided by booksellers. It seems that not all booksellers in our sample provide efficient tools for searching books. From 66 online bookstores, only 12 allow readers to use advanced search. What is more, five of the examined websites do not even provide basic search box and forces users to browse through the whole collection. We also examined metadata presence of browse elements and facets that allow users to navigate through the pages. We found that genre, target audience, author, date, books by the same author, and series are the most frequent browse categories and are equally present in both samples. Meanwhile, enriched elements and elements that represent related works or various versions of a book are more common in the sample of American bookstores.

DATA IMPORTANT TO BUYERS
(Space 27)

The role of metadata in online bookstores cannot be emphasized enough. It is the costumer's experience of shopping for books. Readers' perusal for information on the book jacket or inside the physical book (description, author information, reviewer quotes, awards, etc.) must be replicated online. Metadata should be stored and organized in a way that facilitates both user driven search and algorithms that present titles to users without a specific search, namely book metadata that includes enough information to identify and suggest similar titles (Register & McIlroy, 2015). Metadata in online bookstores should support user tasks such as discovering, identifying, and selecting books and to enable users to explore he collection (Functional…, 1998). Therefore, various metadata elements may be important to buyers in different stages of their shopping process. While elements such as title, description, and subject, play an important role in resource discovery through the keyword searching (Yang, 2016), browsing by genre, topic or books by the same author are the most common ways for people to discover new books (Ross, 2000). Basic book information, such as title, author, date, and identifiers are later crucial for the product identification. Ross (2000) also reports on the importance of information about some physical aspects of a book, such as book's dimensions. Finally, enhanced or additional book information plays an important role for book selection and purchase. The cover, long description, book reviews, author's biography, and table of contents can grab user's attention and convince them to make a purchase (Breedt & Walter, 2012). However, even if the publisher and booksellers provide a rich book metadata set, the quality of this metadata is also important. Errors and inconsistency in displayed description may cause a bad shopping experience and consequently loss of customers (McKay et al., 2012).

Metadata displayed in online bookstore is crucial for people to find, identify and select books. In our expert study on 66 American and Slovenian online bookstores we were examining the presence of different metadata elements in online stores, however, we could not be sure if these elements are actually important for buyers. Therefore, we conducted an online survey with 113 people, asking them to evaluate the importance of various book information in different stages of their buying process. The results show that basic information that is usually displayed in the detailed description page of online bookstores does not cover all attributes buyers perceive as important for book selection. There is a lack of additional book information, such as ratings or customer reviews and information about related products (for example next book in a sequel), which was especially evident in Slovenian bookstores.

INFORMATION ABOUT RELATED PRODUCTS
(Space 29)

Related book materials may appear in a form of a new edition, multiple versions of a product with the same content, other products in a collection or series etc. (Best Practices for Product Metadata, 2015). From the perspective of the Functional Requirements for Bibliographic Records (FRBR) model, related product and works also include adaptations of an original content or works where the central topic is the main product, such as forewords or essays about a title (Functional…, 1998). This information can be critical for a specific audience, such as students, who need a particular version of a book for a school reading, or collectors who would like to obtain special edition of their favorite book. Therefore, information about associated versions and works helps to ensure that the correct book is being ordered and offers customers alternatives of what they are looking for (Best Practices for Product Metadata, 2015). What is more, customers in online bookstores and library catalogues frequently want to obtain all books by the same author, on the same topic or books that are part of the series as these are one of the most important elements for fiction selection. Therefore, this information is frequently present in form of lists representing similar books or books by the same author, which helps users to explore the collection (Market Your Book, 2017). Meanwhile, if the information about related products is exposed at the right time, this can increase sales long after the book is published. Such examples are books that are turned into successful movie adaptations or plays, such as Lord of the Rings or Bridget Jones's Diary.

As we mentioned before, the results of our expert study on 66 American and Slovenian online bookstores show that there is a lack of information about related products. The most frequent metadata about related product in both samples were books by the same author and similar titles from other authors. It is interesting that information about different formats of the same title and information about other versions of a book are present only in American bookstores while Slovenian shops do not provide this kind of information. Since information about related products could be important for buyers searching for the most appropriate format or version of a book, publishers in Slovenia should make more effort in creating connections between products with the same content that appear in various forms.

# ADDITIONAL METADATA
(Space 32)

Essential metadata provides the basic information needed for efficient buying and selling processes and enables users to search and discover resources in the online environment. However, people most often decide to purchase a book based on other information, such as cover, table of contents, book reviews, content description, sample chapters, etc. This information allows people that browse for books online to share the same experience in discovering books as in physical bookstores (Market Your Book, 2017). What is more, some of this data enriches the user's experience in ways that are not possible in a printed book (Register & Mcllroy, 2015). Publishers and booksellers can offer links to various additional material, such as images or third-party sites containing applicable information, including online reviews, videos, best-seller lists of author fan sites (Best Practices for Product Metadata, 2015). Breedt and Walter (2012) report that providing enhanced book information, such as a long description, book reviews, author's biography, and table of contents, can raise the sale of a title for 55 %. Updating a book description over the life of the title with important metadata, such as rewards, new reviews and related work, contributes to continued interest and sales long after the book was published (Register & Mcllroy, 2015).

Additional book metadata is usually present in a book's detailed description page and is important for the book's selection. In the expert study on 66 American and Slovenian online bookstores we found that American bookstores provide more additional metadata than Slovenian stores. Information, such as previews and excerpts, is present in almost half of the analyzed American as well as Slovenian bookstores. Meanwhile, additional author information, professional reviews, topic, enhanced information and awards are frequent descriptive and navigational elements in American bookstores but are not usually present in Slovenian stores. In our opinion booksellers and publishers should include more additional information to their product description. The results of the online survey that we conducted with 113 people show that users perceive additional book information, such as reviews and ratings, as important for their book searching and selection. However, less than one third of the analyzed bookstores provide this type of information.
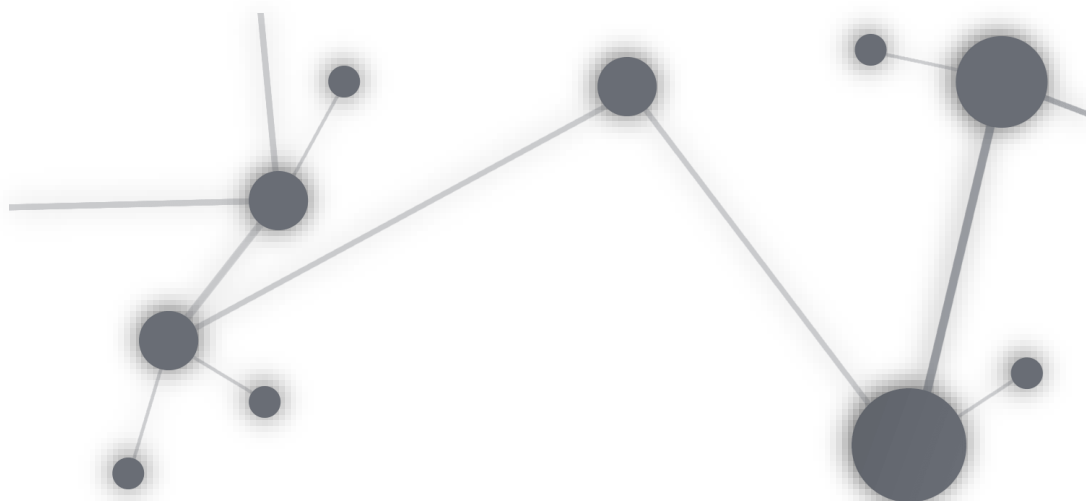
USERS

Enable Metadata Contribution
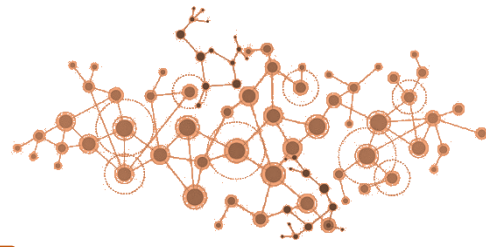
Potential buyers and readers are the key link in the book's lifecycle. Since books are made for them to read and purchase, also book metadata should be created in a way that attracts their attention. Users encounter book metadata in online bookstores, library catalogues and other book-related platforms where they do not only see book data but also create new metadata in form of user behavior logs and tags.

TAGS

(Space 35)

Tags are nonhierarchical keywords, which can be added to online objects. They represent user created metadata describing the object and therefore makes it discoverable with searching and browsing tools. Tagging is very similar to indexing made by professionals, only that tags are created by general public and are therefore more user friendly, since they reflect their interests and vocabulary (Spiteri, 2007). The main goal for people to add tags is not only category creation, but also a tendency for linking similar digital objects, such as documents, pictures etc. (Vander Wall, 2007). While the future object retrieval can be a strong motivation for users to participate in this activity, there is also a presence of social aspects and individual's personal reasons, such as desire to contribute, to attract attention or to express an opinion and self-presentation (Marlow et al., 2006). Therefore, tags represent a great and cost-free addition to online bookstores and library catalogues as they describe different aspects of a book than professionals and are keeping users engaged.

Despite their positive aspects, tags are not frequently used in professional platforms, such as library catalogues and online bookstores, as they have some shortcomings. Peters (2009) and many other experts highlight numerous disadvantages of tags and other folksonomies, which are mainly the result of the variability of human language and the lack of a controlled vocabulary. Because there are no rules governing the way people tag, folksonomies suffer from a certain degree of messiness and inconsistency. Different grammatical variations, a lack of synonym, homonym and homograph control can result in "noise" that does not bring good search results (Macgregor & McCulloch, 2006). Guy and Tonkin (2006) also note that tags are imprecise, inexact, ambiguous and overly personalized and therefore do not support searching as effectively as more controlled vocabularies do. The most problematic are misspelled tags, foreign language terms, singular versus plural form, single-use tags, personal tags etc. (Guy & Tonkin, 2006).

Our study on tagging was conducted as an extension of research on Amazon customer reviews where we found that the review organization in online bookstores is not optimal. Therefore, we wanted to test if social tagging could represent an effective tool for new review organization. In our study with 50 participants, we used an interview and a task solving method to determine whether social tagging could be an appropriate aid for this purpose. The results show that free tagging offers insight into users' vocabulary but is not optimal for online review presentation. Nevertheless, it represents a good basis for creation of categories that describe books on different levels of abstraction and could be used as a filtering tool, which would display only the reviews containing the aspects of a book an individual buyer is interested in.

REVIEWS AND RATINGS
(Space 37)

Word of Mouth (WOM) has been an important information source for buying decisions even before the online environment. This kind of messages has more influence on customer attitudes than market-generated information, as friends and other buyers represent a neutral filter and describe their personal experience with the products (Arndt, 1967). Internet has made access to information easier and, consequently, it allows us to meet a wider range of consumers and their personal opinions. It has been shown that when it comes to the decision making process, online customer reviews are one of the most trusted advertising sources (Nielsen, 2015) that also affect customer's book purchasing behaviour (Chevalier & Mayzlin 2006) and also user's borrowing intention (Huang and Yang 2010). Comments, which are marked as more helpful by other customers, are perceived as more trustworthy and therefore have significant impact on buying decisions (Chen, Dhanasobhon & Smith 2008). Furthermore, reviews that evaluate a product with high numeric rating are more common (Chevalier & Mayzlin 2006; Fowler & Avila 2009) and have a positive effect on sales of that product, while negative information usually influence purchasing decision even more (Henning-Thurau & Walsh 2003; Chevalier & Mayzlin 2006). Researchers (Kakali, 2014; Pecoskie, Spiteri & Tarulli, 2014) found that user generated content, such as online reviews and folksonomies, represent an important addition to the professional book description in library catalogues as they often provide subjective information, such as book's dynamics, emotional reading experience, similar titles, etc. They therefore suggest cooperation between libraries and online bookstores and see customer reviews as an additional promotional channel (Huang & Yang 2010; Kakali 2014).

In our study on Amazon reviews, we were interested in the aspects of a book users perceive as important and to which abstraction levels of the book they correspond. We were also interested in the relation between specific attributes in the reviews and the numeric rating of the book. Therefore, we conducted a content analysis on two random samples consisting of 600 Amazon customer reviews. The results show that most book reviews have a very high numeric star rating and usually refer to the book's content. Meanwhile, book reviews that comment on other aspects of the book, such as physical appearance or received copy, have extremely negative ratings, which indicates that buyers report on the received copy only when they have a bad shopping experience. This means that global rating of the book in online bookstore can be misleading as it can refer to characteristic of the specific version of the book, such as physical appearance of one edition, and not the quality of the content.

REVIEW ORGANIZATION
(Space 40)

Online customer reviews present one of the most important factors in book purchasing or borrowing decisions. Too many reviews can discourage customers' desire to purchase a product (Lin, Huang & Yang 2007), thus many online book platforms have integrated various sorting and filtering tools to reduce the number of reviews. For example, Amazon is already combining all formats and editions of a work under a single title so buyers can read the most helpful comments for one work. Grouping reviews for all formats is often convenient, especially when readers are interested in the content. However, book reviewers do not always focus on the work-level aspects of a book, such as content, genre, author etc., but also on the characteristics of a particular edition. What is more, they can also comment on the received copy or their buying experience (Jug & Žumer, 2016). In some cases, users can encounter reviews with inconsistent textual and numeric ratings of a product, which can result in a bad shopping experience. This phenomenon occurs due to various reasons, such as individual interpretation of ratings or mixture of positive and negative opinions about different product attributes (Mudambi, Schuff & Zhang, 2014). Book buyers might also be interested in various positive and negative aspects of a book, especially when they can choose among different versions of a book. Some researchers (Liu, Karahanna & Watson, 2011) have already found that it would be reasonable to group customer reviews into categories according to product attributes described in the reviews. In this way, users would be able to identify all the associated attributes of the product and read only those reviews, which include information about attributes they are interested in.

In our study on 600 Amazon reviews, we found that reviewers mostly comment on the book's content. However, reviews that include more book attributes, such as translation quality, physical appearance or related work, are usually marked as more helpful. What is more, in this type of comments, reviewers usually have diverse opinion about various book attributes and have problems when rating a product. Considering the fact that online reviews are one of the most important factors for book buying and borrowing decision, it would be reasonable to rethink their display and organisation. In our opinion, book reviews could be organized under categories, which would enable users to filter and read only those containing the attributes they are interested in. The results of our research with 50 participants, where we carried out interviews and used a task solving method, show that social evolvement of users could be a good approach for creation of these categories organizing reviews by attributes interesting to users.

References

Arndt, J. (1967). Role of product-related conversations in the diffusion of a new product. *Journal of marketing Research*, 291-295.

Beky, E. L. (2004). ONIX: Is there a return on investment for all publishers?. *Publishing research quarterly*, 20(2), 3-8.

Best Practices for Product Metadata: Guide for North American Data Senders and Receivers. (2015). BISG. Retrieved from https://www.bisg.org/publications/best-practices-product-metadata.

BIC. (2017). Retrieved from http://www.bic.org.uk/90/Product-Data-Excellence-Awards/.

BISG. (2017). Retrieved from http://bisg.org/page/metadata.

BookNet. (2017). Retrieved from https://www.booknetcanada.ca/standards/.

Breedt, A. & Walter, D. (2012) *White Paper: The Link Between Metadata and Sales*. Retrieved from http://www.nielsenbookdata.co.uk/uploads/3971_Nielsen_Metadata_white_paper_A4(1).pdf.

Breznik Močnik, M., Trplan, T., & Milohnić, A. (2013). *Model razvoja e-založništva v javnem interesu v Sloveniji ter primerjava s stanjem in podatki iz primerljivih evropskih držav*. Ljubljana: Mirovni Inštitut.

British Library. (2017). Retrieved from http://www.bl.uk/bibliographic/cip.html

Carpenter, T. (2010). Standards Column--eBook Identification: Disrupting the Tried and True ISBN System. *Against the Grain*, 22(2), 43.

Cato, A. and Haapamäki, T. (2009). Cooperation with publishers: Integration and sharing of metadata. In *National Bibliographies in the Digital Age: Guidance and New Directions* (pp. 103-110). München: Saur.

CatSi. (2016). Retrieved form http://catsi.info/.

Chevalier, J. A. & Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of marketing research*, 43(3), 345-354.

Dawson, L. (2012). What We Talk About When We Talk About Metadata. In McGuire, H. and O'Leary, B., *Book: A Futurist's Manifesto: A Collection of Essays from the Bleeding Edge of Publishing* (47-55). O'Reilly Media.

D.K. Agencies. (2017). Retrieved from http://www.dkagencies.com/Cataloguing-In-Publication-Data.aspx

Debus-López, K., Williamson, D., Saccucci, C., & Williams, C. (2012). Bringing Publisher Metadata Directly to the Library. *Library Resources & Technical Services*, 56(4), 266-279.

Developement, Use, and Modification of Book Product Metadata. (2012). New York: Book Industry Study Group.

Fagan, J. C. (2010). Usability studies of faceted browsing: A literature review. *Information Technology and Libraries*, 29(2), 58.

Fowler, G. A., & Avila, J. D. (2009). On the internet, everyone'sa critic but they're not very critical. *Wall Street Journal*, 1. Retrieved from http://www.wsj.com/articles/SB125470172872063071

Functional Requirements for Bibliographic Records: Final report. (1998). Munich: K.G. Saur Verlag.

Godby, C. J. (2010). *Mapping ONIX to MARC*. Retrieved form http://www.oclc.org/content/dam/research/publications/library/2010/2010-14.pdf?urlm=162946

Godby, C. J. (2012). *A Crosswalk from ONIX Version 3.0 for Books to MARC 21. OCLC Research*. Retrieved from http://www.oclc.org/content/dam/research/publications/library/2012/2012-04.pdf.

Guy, M., & Tonkin, E. (2006). *Folksonomies: Tidying up tags. D-lib Magazine*, 12(1), 1082-9873. Retrieved from http://www.dlib.org/dlib/january06/guy/01guy.html.

Hennig-Thurau, T., & Walsh, G. (2003). Electronic word-of-mouth: Motives for and consequences of reading customer articulations on the Internet. *International journal of electronic commerce*, 8(2), 51-74.

Huang, Y., & Yang, W. I. (2010). A study of internet book reviews and borrowing intention. *Library Review*, 59(7), 512-521.

International ISBN Agency. (2017). Retrieved from https://www.isbn-international.org/content/what-isbn

ISBN User's Manual. (2017). Retrieved from https://www.isbn-international.org/sites/default/files/ISBN%20Manual%202012%20-corr.pdf

ISBN.org. (2017). Retrieved from http://isbn.org/.

Jug, T., & Žumer, M. (2017). Do we need better online book review organisation?. *Libellarium: journal for the research of writing, books, and cultural heritage institutions*, 9(2).

Yang, L. (2016). Metadata effectiveness in internet discovery: An analysis of digital collection metadata Elements and Internet Search Engine Keywords. *College & Research Libraries*, 77(1), 7-19.

Kakali, C. (2014). A Utilization Model of Users' Metadata in Libraries. *The Journal of Academic Librarianship*, 40(6), 565-573.

Kanič, A. (2014). Naj kovačeva kobila ne bo bosa ... CIP in ISBN/ISMN. *Knjižničarske novice*. 24 (7/8), 4-5.

Library of Congress. (2017). Retrieved from https://www.loc.gov/publish/cip/

Lichtenberg, J. (2014). Can Onix 3.0 Create a Global Digital Publishing Industry?. *Publishers Weekly*. Retrieved from http://www.publishersweekly.com/pw/by-topic/digital/retailing/article/63923-can-onix-3-0-create-a-global-digital-publishing-industry.html.

Lin, T. M., Huang, Y. K. & Yang, W. I. (2007). An experimental design approach to investigating the relationship between Internet book reviews and purchase intention. *Library & Information Science Research*, 29(3), 397-415.

Liu, Q. B., Karahanna, E., & Watson, R. T. (2011). Unveiling user-generated content: Designing websites to best present customer reviews. *Business Horizons*, 54(3), 231-240.

Luther, J. (2009). *Streamlining Book Metadata Workflow*. Baltimore: NISO ; Dublin, Ohio: OCLC. Retrieved from http://www.niso.org/publications/white_papers/StreamlineBookMetadataWorkflowWhitePaper.pdf

Macgregor, G., & McCulloch, E. (2006). Collaborative tagging as a knowledge organisation and resource discovery tool. *Library review*, 55(5), 291-300.

MARC. (2017). Retrieved from https://www.loc.gov/marc/.

Market Your Book. (2017). Retrieved from https://www.myidentifiers.com/isbn/market_your_book.

Marlow, C., Naaman, M., Boyd, D. & Davis, M. (2006). HT06, tagging paper, taxonomy, Flickr, academic article, to read. In *Proceedings of the seventeenth conference on Hypertext and hypermedia* (31-40). ACM.

McKay, D., Buchanan, G., Vanderschantz, N., Timpany, C., Cunningham, S. J., & Hinze, A. (2012). Judging a book by its cover: interface elements that affect reader selection of ebooks. In *Proceedings of the 24th Australian Computer-Human Interaction Conference* (pp. 381-390). ACM.

NUK. (2017). Retrieved from http://www.nuk.uni-lj.si/informacije/faq.

Mudambi, S. M., & Schuff, D. (2010). What makes a helpful review? A study of customer reviews on Amazon. com. *MIS quarterly*. 34(1): 185-200.

Nielsen. (2015). *Global Trust in Advertising: Winning Strategies for an Evolving Media Landscape*. Retrieved from http://www.nielsen.com/content/dam/nielsenglobal/apac/docs/reports/2015/nielsen-global-trust-in-advertising-report-september-2015.pdf.

OCLC Metadata Services for Publishers. (2010). Retrieved from
http://oclc.org/content/dam/oclc/partnerships/material/metadata/213918usb_services_for_publishers.pdf

ONIX: Overview. (2017). Retrieved from http://www.editeur.org/83/Overview/

Register, R. & McIlroy, T. (2015). *The Metadata Handbook*. Columbus, Ohio: DataCurate.

ONIX: Users and Services directory. (2017). Retrieved from http://www.editeur.org/154/Users-and-Services-directory/.

Pecoskie, J., Spiteri, L. F., & Tarulli, L. (2014). OPACs, Users, and Readers' Advisory: Exploring the Implications of User-Generated Content for Readers' Advisory in Canadian Public Libraries. *Cataloging & Classification Quarterly*, 52(4), 431-453.

Peters, I. (2009). *Folksonomies: indexing and retrieval in Web 2.0*. Berlin: Walter de Gruyter.

Ross, C. S. (2000). Making choices: What readers say about choosing books to read for pleasure. *The Acquisitions Librarian*, 13(25), 5-21.

Sanchez, E. R. (Ed.). (2011). *Conversations with Catalogers in the 21st Century*. Oxford: ABC-CLIO.

Spiteri, L. F. (2007). The structure and form of folksonomy tags: The road to the public library catalog. *Information technology and libraries*, 26(3).

Tennant, R. (2006). Digital Libraries: The New Cataloger. *Library Journal*, 131(7), 32-33.

Vander Wall, T. (2007). *Folksonomy Coinage and Definition*. Retrieved from http://www.vanderwal.net/folksonomy.html.

Walter, D. (2016). *Nielsen Book UK Study: The Importance of Metadata for Discoverability and Sales*. Retrieved from http://www.nielsenbookdata.co.uk/uploads/10451_Nielsen_Book_UK_Study_The_Importance_of_Metadata_for_Discoverability_and_Sales_Digital_D6%20(1).pdf

Wayne, L. (2005). *Institutionalize metadata before it institutionalizes you*. Retrieved from http://wvgis.wvu.edu/data/otherdocs/metadataResources/InstitutionalizeMeta_URISA2001[1].pdf

Wiley, L. (2016). Report of the ALA ALCTS Publisher Vendor Library Relations Interest Group Meeting. American Library Association Annual Conference, San Francisco, June 2015. *Technical Services Quarterly*, 33(1), 80-87.

WorldCat. (2017). Retrieved from https://www.worldcat.org/affiliate/webservices/xisbn/app.jsp.

METADVENTURE

From Storyteller to Bestseller

A Game by Tjasa Jug

This game was made as a submission for the European Student Chapter Scientivity Contest 2017.

Special thanks to: