

Rita S. Van Duinen. *New Discoveries in the André Savine Collection: Examining the Author-Generated Metadata Contained in the Bibliographic and Biographical Record of André Savine*. A Master's Paper for the M.S. in L.S. degree. November, 2004. 28 pages. Advisor: Jane Greenberg

This paper reports on an analysis of author-generated metadata contained in a portion of the bibliographic / biographical record compiled by the late André Savine, creator of the André Savine Collection. For this study, the contents of 368 index cards created by Savine were analyzed for core bibliographic elements (author-generated descriptive metadata). Elements were identified and conceptually mapped to the English and Russian versions of the Dublin Core Metadata Element Set (DCMES). Results of this study indicate that Savine created extensive descriptive metadata and that ten of these data elements map to the DCMES. Other key metadata elements that do not map to DC were also identified. It is recommended that these additional elements also be incorporated into the schema used to facilitate resource discovery of Savine's index cards. The descriptive metadata extracted from Savine's bibliographic and biographical provide a model metadata scheme that can be used with similar projects in the future.

Headings:

Dublin Core

Information organization

Metadata

Political refugees – Soviet Union

Russians – Foreign Countries

NEW DISCOVERIES IN THE ANDRÉ SAVINE COLLECTION:
EXAMINING THE AUTHOR-GENERATED METADATA CONTAINED IN THE
BIBLIOGRAPHIC AND BIOGRAPHICAL RECORD OF ANDRÉ SAVINE

by
Rita S. Van Duinen

A Master's paper submitted to the faculty
of the School of Information and Library Science
of the University of North Carolina at Chapel Hill
in partial fulfillment of the requirements
for the degree of Master of Science in
Library Science.

Chapel Hill, North Carolina

November 2004

Approved by

Dr. Jane Greenberg

Table of Contents

1	Introduction.....	2
2	The Artifactual Importance of Savine’s Bibliographic / Biographical Record	6
3	Dublin Core and Savine’s Author-Generated Metadata	9
4	Research Objectives.....	11
5	Methodology	12
6	Results of Data Analysis.....	15
7	Significance of the Study	21
8	Conclusions / Future Research.....	22
9	Bibliography	25

1 Introduction

Beginning with the fall of Czar Nicholas II in 1917 and ending with the collapse of the Soviet Union in 1991, thousands of renowned writers, philosophers, scholars and political figures whose loyalty and ideas were unacceptable to the Soviets settled outside of Russia where they continued to live, work, and publish. These émigrés dutifully preserved and documented their Russian culture as it began to deteriorate inside their tightly controlled Motherland.

Preserving and documenting the lives of these Russian exiles was truly a labor of love for the late André Savine (1946-1999), creator of the André Savine Collection. Born in Paris, Savine was the son of a White Army soldier who fled Russia following the 1917 Bolshevik Revolution and the Russian civil war. As one of the 100,000 demoralized troops that left Russia in 1920, Savine's father endured the hardships encountered by the defeated White Army in a military camp in Gallipoli, Turkey. For André, the plight of so many Russian exiles was a very personal and integral part of his life.

André Savine's knowledge of the Russian exile experience along with his expertise as a bookdealer and bibliographer are evident throughout this collection. Together with his wife, André Savine owned and operated the Paris bookstore, *Le Bibliophile Russe*, and became a highly reputable Russian bookdealer. An avid collector, Savine spent more than 30 years acquiring hundreds of rare editions, archival materials and documents relating to worldwide Russian culture in exile. He became a member of the prestigious French *Syndicat National de la Librairie Ancienne et Moderne* (Professional Association of Antiquarian Bookdealers, Autographs and Engravings) and

had earned the title of “Libraire Expert”. His particular collecting interests were in rare editions, archival materials and documents of worldwide Russian culture.

In addition to the his bookstore stock, André Savine created his own personal collection consisting of more than 10,000 volumes of books, serials and newspapers, rare manuscripts, photographs and archives. His personal collection and the holdings of his bookstore, Le Bibliophile Russe, comprise the André Savine Collection. The University of North Carolina at Chapel Hill acquired this unique collection of historical record in 2002 with the help of a generous donation from Van and Kay Weatherspoon of Charlotte, NC.

An extremely significant part of this collection is the bibliographic and biographical record compiled by Mr. Savine. As a respected antiquarian bookdealer and as a professional bibliographer, André Savine conducted thorough research on the titles in his bookstore and in his personal collection. He took great care to extensively annotate each title resulting in a collection of thousands of index cards and notebook pages that include detailed bibliographic and biographical information. Much of the information he recorded was on particular authors, editors, illustrators, and on more obscure people who may have played key roles in shaping the history of the Russian Diaspora. It is this portion of the André Savine Collection that serves as the focus for the research discussed in this paper.

André Savine’s annotated index cards and notebook pages are rich and important resources for contemporary scholars. The complex bibliographic and biographical record he compiled not only details the holdings of his bookstore stock and personal collection, but also serves as a historical record of a specific time and culture. The information

contained within the thousands of index cards and notebook pages reflects the climate of Russian émigré culture worldwide and gives insight into his expertise as bookseller and as bibliographer and subject specialist.

The field of information and library science has well documented the historical importance of the creation of bibliographic record. Many significant contributions to the art and science of bibliography were direct results of the book trade. One notable example is the work of 16th century English bookseller, Andrew Maunsell. In his *Catalogue of English Printed Books*, Maunsell set forth detailed rules for the entry of personal names under surnames as opposed to entry under Christian names as was the common practice. Commenting on Maunsell's contributions to modern day cataloging practice, Ruth French Strout writes in *The Development of the Catalog and Cataloging Codes*, "These were radical and sudden advances in the development of cataloging, and one cannot help noticing that they came from the bookseller rather than the librarian..." (p.264).

The study of the preparation of book lists and catalogs addresses both the intellectual and artifactual value of a book's composition. This is evident in the way in which André Savine went to great lengths to create such an intricate record of items in his bookstore stock and personal collection. A close look at the level of detail contained in the index cards and notebook pages reveals intentions beyond the compilation of an inventory list used to create the sale catalogs for Le Bibliophile Russe. The complexity and richness of his bibliographic and biographical record gives insight into the relationships of his research to other people, and to the books themselves.

During his appointment as the first Librarian of the Smithsonian Institution (1848-1854), Charles Coffin Jewett once wrote, “The preparation of a catalogue may seem a light task, to the inexperienced, and to those who are unacquainted with the requirements of the learned world, respecting such works. In truth, however, there is no species of literary labor so arduous and perplexing”(qtd. in N. Baker, p.70). Evidence of Jewett’s observations is present in the massive collection of detailed information Savine assembled. His expertise and dedication to his work and to documenting Russian émigré history and culture is unmistakable throughout his bibliographic and biographical record.

The value of the information Savine compiled in the index cards and notebook pages lies within making their contents accessible to scholars through digitization efforts. With the onset of electronic publishing, current trends in the field of Information and Library Science focus on making primary and secondary resources available via the Internet, resulting in global accessibility. In his article, *The Impact of Digitization on Special Collections in Libraries*, Peter Hirtle discusses the benefits of digitizing special collections materials. He cites the increase in the use of digitized materials, the new types of research that will occur, as well as the new types of users digitized materials will attract, as the most immediate and significant results. (Hirtle, p.2-3)

The information contained in Savine’s bibliographic and biographical record is rich in historical information, and the index cards themselves constitute a unique artifactual collection. Research on its contents and structure will determine how to provide access on a global scale, after these cards have been digitized.

2 The Artifactual Importance of Savine's Bibliographic / Biographical Record

Savine's bibliographic and biographical research forms two distinct sets of records: 1) approximately 5,500 notebook pages; 2) approximately 16,000 index cards containing handwritten and typed text. Both the index cards and notebook pages are alphabetically arranged by either author entry or by title. The majority of the text is in Russian with the exception of an occasional French word or phrase.

The index cards and notebook pages that Savine created served multiple functions: as inventory lists, as working files for internal business operations, and as rich reference resources. It is interesting to note that when an item was indicated as "sold-out" from the bookstore stock, Savine handwrote the French equivalent, *épuisé*, on the card face, but did not dispose of the corresponding index card. Because he chose to retain the information even though he no longer had possession of the item attests to the many functions these resources served and of the historical documentation for this unique period of Russian history and culture.

Study of Savine's bibliographic and biographical record reveals the approximate number of titles by a given author that existed in either his bookstore or personal collection. Moreover, the level of detail Savine employed in its creation tells even more about its contents. Many titles have multiple index cards or notebook pages stapled or clipped together and often contain additional information on the verso. In some instances it appears that Savine may have supplemented the information already on a card by either taping or pasting additional details atop the text. His notebook pages consist not only of

bibliographic information, but also include photocopies of a particular work's cover or title page.

Key features of the index cards and notebooks pages include the author-generated metadata created by Savine. For the purposes of this paper, the term author-generated metadata refers to bibliographic elements Savine used to describe or identify the information resources contained in his bibliographic and biographical record.

In its simplest form, metadata is defined as 'data about data', or better yet, as 'structured data about data'. In a broader sense, metadata is also defined as, "...the sum total of what one can say about any information object at any level of aggregation" (Gilliland-Swetland, p.1).

Savine's author-generated metadata outlines the core (essential to resource discovery) bibliographic elements of each title. For example, in addition to author information, Savine recorded the birth and death dates for that person as well as any pseudonym he or she may have published under. There are also entries for editors, contributors, translators, and artists. Complete title entries are given as is publishing information including place, year, and issuing body. Page enumeration is also recorded, and in many cases, the physical format, edition, and series statements are also supplied. Savine also recorded detailed content information on individual titles and assigned his own subject areas, sometimes in multiplicity.

Many of the index cards and notebook pages contain detailed bibliographic citations where the title is referenced and/or contain detailed biographical information on the title's author, editor, or illustrator (See Figure 1).

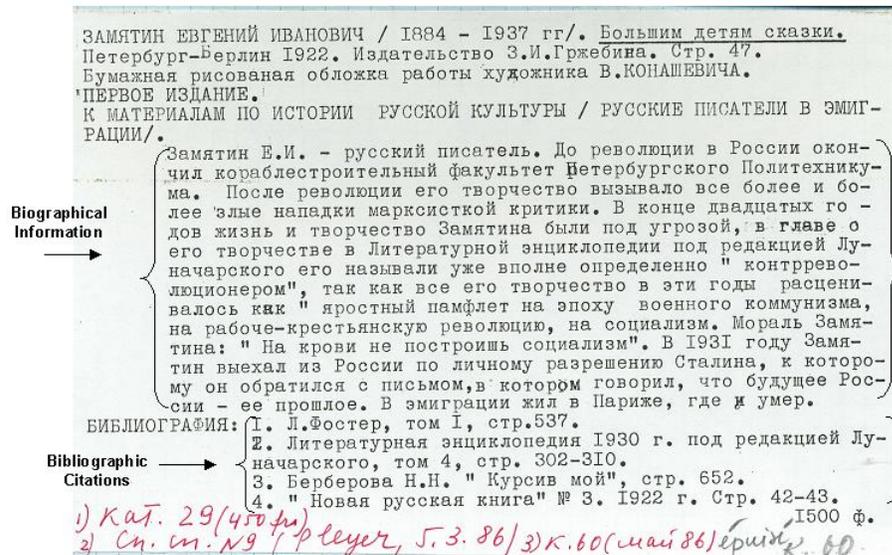


Figure 1. Example of biographical information and bibliographic citations.

In addition to bibliographic and biographical information, the index cards contain details regarding the business operations of Le Bibliophile Russe. Information on sales catalogs and prices as well as buyer information provides insight into the trends of the antiquarian book trade and into Le Bibliophile Russe as a reputable bookstore. In addition, many of the titles have specific information regarding the publication run or “tirazh” and on a particular numbered copy of a publication run.

Other features recorded by Savine include elements similar to those captured in rare book cataloging: autograph information, descriptions of a book’s dust jacket or physical condition, or statements on a particular title’s perceived rarity in the book market. These additional metadata elements warrant their own analysis for future development and use in cataloging and digitization efforts.

In his famous article entitled *Discards*, Nicholson Baker laments the decision by libraries to discard their card catalogs and replace them with electronic versions. He compares a card catalog to a personal address book, or Rolodex, in which years of history and detail are recorded. Baker asks readers to think of such files as unique pieces of literary history which contain, "...detailed descriptions of certain objects that the world has deemed worthy of organized preservation" (p.70).

Savine's compilation of index cards and notebook pages is an extensive historical and artifactual record. One can see which authors or titles received the most attention by observing their tattered, stapled corners or by noting the lack of empty space for adding any additional information or notes. The importance of Savine's index cards and notebook pages is much deeper than the bibliographic and biographical data they contain. They are essentially his handcrafted manuscripts in which he details his life as a bookdealer, and as a skilled bibliographer and historian. It is these characteristics and Savine's level of precision that makes his bibliographic and biographical record so unique and significant as a rich information resource. Furthermore, these rich information resources serve as an excellent candidate for the study of author-generated metadata and its usefulness when coupled with other metadata schemas, such as the Dublin Core.

3 Dublin Core and Savine's Author-Generated Metadata

To extend the usefulness and accessibility of Savine's bibliographic and biographical record, core bibliographic elements extracted from the index cards can be conceptually mapped to both the English and Russian versions of the Dublin Core

Metadata Element Set (DCMES) and used to facilitate resource discovery and global accessibility once digitization takes place.

The Dublin Core metadata standard consists of a set of 15 metadata elements to facilitate resource discovery for a wide range of online resources (see <http://www.dublincore.org/documents/dces/>). It is considered by many as the common semantic foundation for creating metadata for the Internet and has been adopted to describe information resources by a variety of communities due to its simplicity. All 15 Dublin Core elements are optional and repeatable, and include a limited set of element qualifiers that refine the semantic meaning of some of the basic categories of information. The DCMES and its refinements describe resources in a complete package much like the way André Savine did in his index cards.

Enhancing Savine's author-generated metadata with both the English and Russian versions of Dublin Core will add purpose, value, and meaning to the artifacts themselves. Moreover, the use of multi-lingual Dublin Core promotes the usefulness and interoperability of metadata schemas in a global community and aids in the international exchange of information and ideas related to metadata. Providing Dublin Core in the vernacular also serves the specific research needs of those scholars searching for information on the Russian Diaspora in the vernacular.

A similar use of Dublin Core is documented in a study that examined the author-generated metadata contained within ancient Marathi manuscripts from the 13-17th century. In their report, the authors describe which descriptive elements were identifiable and mapped to the DCMES in both English and in Marathi (Nagarkar, Parekh, and Manjul).

To further resource discovery and global accessibility, Dublin Core can also be used with metadata harvesting protocols such as the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH). The OAI-PMH provides a means for application-independent interoperability based on metadata harvesting and is intended to enable access to online materials through sharing metadata between services (see <http://www.openarchives.org/OAI/openarchivesprotocol.html>).

4 Research Objectives

The objective of this paper is to determine the usefulness of a multi-lingual approach to using the DCMES and its qualifiers for supporting access to the wealth of bibliographic and biographical data found in bookdealer collections. The bibliographic and biographical record compiled by André Savine was chosen for this project due to its richness as historical documentation and literary artifact as well as to its physical accessibility. The research discussed in this paper was guided by asking the following three questions:

1. Can substantive patterns of descriptive metadata elements contained in the text of the index cards be established?
2. Can the results of the content analysis be conceptually mapped to the English and Russian versions of the DCMES?
3. What remaining author-generated metadata elements that do not map to the DCMES need to be incorporated to facilitate full resource discovery once digitized?

Making André Savine's bibliographic and biographical record accessible via the Internet will enhance scholarly activity by providing researchers with the opportunity to fill historical gaps or address unanswered research questions. Identifying key author-

generated metadata elements created by Savine will prove useful in cataloging and will enhance resource discovery in the future digitization and text encoding of these artifacts.

In order for researchers to take advantage of these rich resources, some analysis of Savine's bibliographic and biographical record is needed. Evaluating these documents for textual evidence and conceptual patterns provides factual background information on the breadth and scope of the André Savine Collection, and on the circumstances under which it was amassed.

5 Methodology

The underlying research discussed in this paper was conducted using a multi-method approach of content analysis and crosswalk analysis based a sample of 16,000 index cards compiled by André Savine. Content analysis was used as the primary methodology for examining the breadth and quality of Savine's descriptive metadata. In addition, quantitative methods were used to analyze the frequency of Savine's author-generated descriptive metadata elements.

Of the 16,000 index cards, only a portion of them was sampled from a group of approximately 9,000 cards labeled "Le Bibliophile Russe Bibliographic Description of Stock". This group represents Savine's bookstore inventory and of all the index cards, they are perhaps the most detailed due to their use in the production of sale catalogs. The index cards are filed alphabetically by the author's surname or pseudonym or by title if no author is given.

The sample was selected by using a systematic sampling technique. Sample size was determined by consulting Table 3.4 in Ronald R. Powell's *Basic Research Methods*

for Librarians (p.75). As a result, every 25th index card of the 9,000 was carefully analyzed for potential descriptive metadata elements until the desired sample size of 368 was reached.

As the data elements contained in Savine's index cards were identified, categories of bibliographic data were developed based on the reoccurring patterns that emerged. This approach allowed active participation in the research environment while teasing out of the data the contextual evidence around which the data analysis was designed. As a result, theory and data collection informed each other once analysis was underway.

In the initial analysis of the 368 index cards, 24 data elements were identified as potential categories of Savine's author-generated descriptive metadata. Of the 24 elements, 9 were eliminated from the quantitative portion of the analysis after consultation with professional cataloging staff from the Academic Affairs Library at UNC-CH. Elimination was based on careful review of the nature of the content of a particular element and on the frequency of occurrence.

Of the 368 index cards selected for content analysis, a sub-sample was selected and further analyzed to determine the frequency of the 15 categories of key bibliographic elements identified in the larger sample. Using Table 3.4 (Powell, p.75) again as a guide for determining sample size, the sub-sample consists of approximately 186 index cards from the larger sample of 368. For this portion of the analyses, every 2nd index card within the larger sample was selected until 186 were reviewed. Each of the 186 index cards in the sub-sample was analyzed for the presence of any of the 15 categories identified within the larger sample. A MS Excel spreadsheet was created to track the number of elements present in each title.

Due to practical research limitations, only core bibliographic elements were analyzed and considered for mapping to the DCMES. With the exception of the ‘Pseudonym’ and ‘Contributor’ categories, those elements that were present in less than 10% of the total sample of 186 were removed, as were those elements that were contextually out of scope. For example, the category “Autograph” was created to record the presence of information on autographs found on a particular title’s cover or title page. ‘Autograph’ was eliminated since this type of information is more frequently used in describing rare books and is beyond the scope of this current research. Furthermore, ‘Autograph’ was present in only 5.9% of the total sample population.

Although they occurred in less than 10% of the sample, the ‘Pseudonym’ and ‘Contributor’ categories were retained and are considered key bibliographic access points. Since many of the people described in Savine’s bibliographic and biographical are obscure or may have published only under a pseudonym, including these two categories of descriptive data becomes necessary for resource discovery.

In addition, elements not normally considered core bibliographic data, such as the Date of Birth /Date of Death, Biographical Info., and Bibliographical Citations categories were also retained and are considered important sources of key information on specific persons, places and resources. Figure 2 details the various types of author-generated metadata elements typically found in Savine’s bibliographic and biographical record.

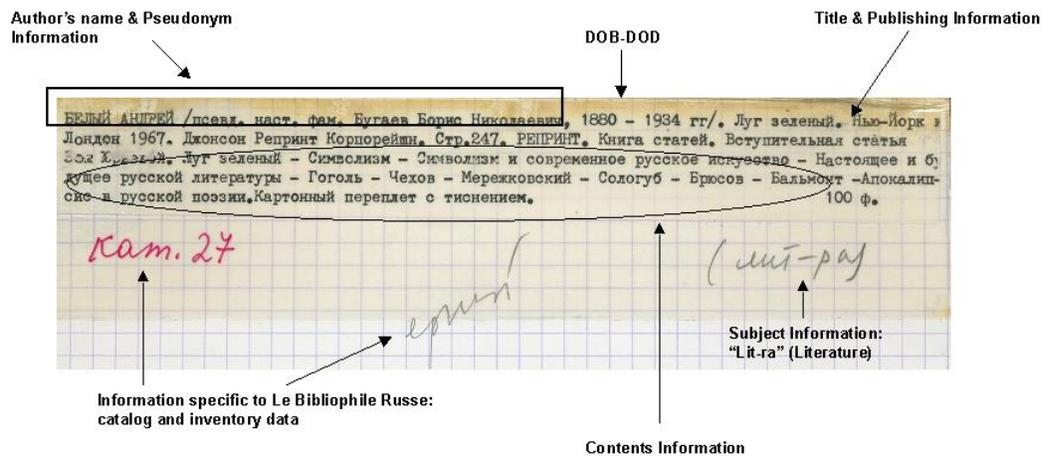


Figure 2. Example of the various author generated metadata elements created by André Savine

6 Results of Data Analysis

Results of the data analysis show that André Savine spent considerable time and effort in creating extensive author-generated metadata as recorded in the index card portion of his bibliographic and biographical record. Core bibliographic elements such as author, title, publication information, and subject area are present in almost all of the sample population.

Table 1 identifies the final 15 categories that were developed as a result of examining a portion of André Savine's index cards for author-generated descriptive metadata.

Table 1. Categories of descriptive metadata elements

Category	Description of Category
Author	Last Name, First Name
DOB	Date of Birth - Date of Death
Pseud.	Pseudonym
Title	Complete title of work
Place	Place of publication
Pub.	Name of publishing body
Yr.	Year of publication
Pgs.	No. of pages
Edition	Edition statement
Contents	TOC or summarized content information
Contrib.	Contributor(s)
Ed.(s)	Editor(s)
Subject	Subject category assigned by Savine
Biog.	Biographical information on author, editor, artist
Biblio.	Bibliographic citations

Working with the sub-sample of 186 index cards, each instantiation of the categories described above was quantified. Table 2 summarizes the number of index cards within the sub-sample that contained any of the 15 defined categories of bibliographic elements.

Table 2. Cards containing the 15 categories of bibliographic elements

Categories	No. of cards	% of total (186)
Author	152	81.7
DOB-DOD	57	30.6
Pseudonym	9	4.8
Title	186	100
Place of Publication	181	97.3
Publisher	135	72.6

Table 2 (cont.)

Categories	No. of cards	% of total (186)
Year Published	180	96.8
No.of Pages	175	94.1
Edition	24	12.9
Content Information	32	17.2
Contributor(s)	14	7.5
Editor(s)	21	11.3
Subject(s)	155	83.3
Biographical Info.	35	18.8
Bibliographic Citations	32	17.2

In order to ensure the richest delivery of author-generated descriptive metadata found in André Savine's index cards, the categories developed as a result of analysis were conceptually mapped to both the English and Russian versions of the DCMES. Element qualifiers were used where possible. The English version of the DCMES contains transliterated Russian text according to transliteration guidelines used by the Library of Congress, while the Russian version is expressed in the vernacular. Providing multi-lingual metadata facilitates global access and serves specific scholarly research needs in the vernacular.

Table 3 presents those categories of author-generated metadata identified in Savine's index cards that map to the DCMES and its refinements. Ten of the 15 identified categories of author generated descriptive metadata can be mapped to the DCMES. The elements DC.Type and DC.Language were added to further enhance resource discovery, although they were not part of the initial categories of data that were extracted and analyzed. In all, 12 DC elements can be created to facilitate resource description and discovery.

Table 3. English and Russian DC elements & refinements with Savine's Author-generated metadata

Author-generated Metadata Elements	DC Elements & Refinements – English	DC Elements & Refinements– Russian
Title	Title	Название
Author	Creator	Создатель
Pseudonym	Creator	Создатель
Subject areas assigned by Savine	Subject	Предмет
Summarized content information recorded on the index card	Description.Abstract	Описание.Реферат
Sub-units of content as recorded on the index card	Description.tableOfContents	Описание.Оглавление
Person or organization responsible for publishing	Publisher	Издатель
Person(s) responsible for contributing to the content	Contributor	Соисполнитель
Year of publication	Date.Created	Дата.Выпущен
Bibliographic citation	Identifier.bibliographic.Citation	Идентификатор* <i>*lacks the qualifier 'bibliographic.Citation'.</i>
Type of resource refers to the index cards as the genre 'text' as allowed in the DCMI Type Vocabulary	Type	Тип ресурса
Language(s) of text on index cards expressed in DC according to ISO639-2	Language	Язык

Element qualifiers were applicable in only four instances. Unfortunately, the DC.identifier element qualifier 'bibliographic.Citation' exists in the English version, but not in the Russian. Nonetheless, metadata can still be expressed for the qualified DC.Identifier.Citation element in transliterated Russian form as opposed to not using this element for resource discovery at any level.

Limitations like those of the Russian version of the DCMES are discussed in *Metadata Semantics Shared Across Languages: Dublin Cores in Languages other than English*, in which Thomas Baker stresses that "...if the Dublin Core is to meet the need of specialists across many languages, it will also have to allow users in those languages to define their own qualifiers"(par.3). Updating the current Russian version of the DCMES and its qualifiers needs to be done to promote cross- / multi-lingual interoperability of metadata.

In addition to the ten elements that mapped to the DCMES, six important categories of author-generated metadata remain for which no DCMES equivalent was identified. Table 4, column 1, lists those six categories for which a Dublin Core equivalent was not found and details their importance as well as possible solutions for future metadata mapping. Most likely, these categories would need to be incorporated into a customized namespace or application profile for use with describing resources in the André Savine Collection.

Table 4. Descriptive metadata elements for which no DC equivalent was found

Descriptive Metadata for which no DC equivalent was found	Importance / Solutions
DOB-DOD	Refinements for DC.Date are needed so that birth/death information can be captured.
Place of Publication	Place names important in scholarly research. Need way to express them using some type of metadata schema.
No.of Pages	Specific bibliographic information that is recorded in MARC, but not applicable to DC.
Edition	Used in the DC-Lib, but no Russian equivalent exists.
Editor	Refinements for DC.Creator are needed to express role, as well as structured values to express further information. They are not included in the DC-LIB application profile. Approval by DCMI of a mechanism to express these is pending.
Biographical Info.	Vital information on obscure people. Needs to be expressed in metadata, as in done in a 'Notes' field in MARC.

One approach to incorporating these additional types of author-generated metadata elements that do not map to the DCMES would be to enhance the DC-Library Application Profile (DC-LIB) already in existence (see <http://dublincore.org/documents/2004/09/10/library-application-profile/>). Other solutions include incorporating the Text Encoding Initiative (TEI) Header to record bibliographic information on both the print and electronic versions of the information

resource being described or use of the Metadata Object Description Schema (MODS) for a richer, more thorough description of digitized resources.

7 Significance of the Study

With the exception of the André Savine Collection, materials dating from the period immediately following the 1917 revolution are fragmented and scattered around the world. Because of the lack of available materials, serious study of the Russian Diaspora has been difficult due to limited access to these types of materials, and even more so, to limited bibliographic access to this genre of digitized special collections. The research conducted in this study identifies the extensiveness and quantity of author-generated metadata contained in André Savine's bibliographic and biographical record. Through analysis, the value of the information Savine created is understood, as is the usefulness of a multi-lingual approach to using the DCMES and its qualifiers.

In a discussion on the future of primary records, G. Thomas Tanselle wrote in *Literature and Artifacts*, "An object cannot be used unless it exists; and though it may exist without being used, its chances of survival are greater if it is perceived to serve a function" (p.96). The function of digitizing Savine's bibliographic and biographical record is to give unfettered access to information considered obscure or perhaps non-existent. Careful analysis of a portion of the index cards indicates that Savine created valuable author-generated descriptive metadata elements, which, once conceptually mapped to the DCMES, will help facilitate cataloging and resource discovery.

The bibliographic elements examined in the content analysis portion of this project provide access points into the collection, while also containing valuable

information on the people and places that shaped the Russian Diaspora. Extracting meaningful content and capturing the data elements from the index cards reflects the usefulness of Savine's extensive research and on the nature of the index cards as historical and literary artifacts. Utilizing the wealth of bibliographic and biographical data found in bookdealer collections such as Savine's will result in new types of research and in attracting new types of users from across the globe. Digitizing these rich resources will give access to the unique historical period and culture represented in the André Savine Collection and will reflect on the original intent of their creator.

8 Conclusions / Future Research

This study investigated the extent to which André Savine created substantive categories of author-generated descriptive metadata in the index card portion of his bibliographic and biographical record. The results indicate that Savine created extensive metadata when describing the information resources held in his bookstore's stock. The value of this data emphasizes the usefulness of these resources as research materials and confirms the need for accessibility through digitization.

A secondary purpose of this study was to evaluate the extent to which Savine's author-generated metadata could be conceptually mapped to the English and Russian versions of the Dublin Core Metadata Element Set. The results of this portion of the study indicate that ten of the 15 categories of descriptive metadata identified in Savine's index cards map to the English and Russian versions of the DCMES.

Six important categories of descriptive metadata elements remain for which no DCMES equivalent was identified. This indicates the limitations associated with using

the DCMES for resource discovery and also points to the difficulties in using a multi-lingual approach to providing metadata. This is most evident in the case of the ‘DC.identifier’ element, for which there exists the ‘bibliographic.Citation’ refinement in English, but not in Russian.

Using the DCMES to express the core bibliographic elements identified in this study promotes resource discovery, interoperability, and data sharing. Dublin Core can be combined with other metadata schemas to form more complex and richer descriptions and can be used with metadata harvesting protocols such as the Open Archives Initiative (OAI-PMH). Metadata harvesting is further enhanced by providing metadata in the vernacular and in transliterated Russian according to the Library of Congress transliteration guidelines.

Examination of André Savine’s intensive bibliographic and biographical record reveals more than just those elements considered as core bibliographic information. Other types of key metadata elements exist throughout the majority of the index cards and will need to be incorporated. The scope of information and the level of detail Savine provides warrants the development of a customized metadata application profile or namespace if all types of metadata he generated are to be used in facilitating resource discovery: bibliographic, biographical, preservation, and book trade information. Creating such an expansive schema will allow for the most complete access to these rich information resources once they are digitized and available on the Internet for scholarly use. Furthermore, developing a customized profile will compliment future digitization efforts with other sub-collections of materials contained in the André Savine Collection, and may potentially be used to facilitate collaborative or cross-institutional digital projects.

Until digitization of these rich resources begins in earnest, access to the wealth of information they contain is extremely limited. Researchers and scholars of the Russian Diaspora will benefit greatly from the extensive historical and literary record that Savine compiled. In addition, the study of Savine's bibliographic and biographical record as a literary artifact and as a commentary on the book trade and émigré publishing community will aid scholars in the field of Information and Library Science who take on similar studies or aim to make similar collections available through the Internet.

Making all parts of the André Savine Collection accessible to users will provide access to a significant source for study of the Russian Diaspora. Acquiring such a unique collection of materials gives notoriety the University scholarly community and upholds the objectives of the Library's mission statement. Utilizing the research that André Savine spent a lifetime compiling will only enhance accessibility and usage and adds value to the collection as a whole.

Through the work of André Savine, the Russian émigré experience is celebrated and freely expressed. The irony of being able to celebrate their experience lies in the circumstances under which so many devoted people became exiles – under repression and fear. Materials in the André Savine Collection give a free and open voice to the lives of those who lived through this tumultuous time in Russian history.

9 Bibliography

- Baker, Nicholson (1994). Discards. The New Yorker April 4, 64-86.
- Baker, Thomas (1997). Metadata Semantics Shared Across Languages: Dublin Cores in Languages other than English. Fourth Dublin Core Workshop. Canberra, Australia. <http://www.dublincore.org/groups/languages/mr-19970303.shtml>.
- Barysheva, Olga (2000). Kvalifikatory Dublin Core (Dublinskogo Iarda). <http://www.rba.ru:8101/rusmarc/soft/dcq.html>.
- Barysheva, Olga (1999). Nabor Elementov Metadannykh Dublin Core (Dublinskogo Iarda) Versiia 1.1: Spravochnoe opisanie. <http://www.rba.ru:8101/rusmarc/soft/dc.html>.
- Dublin Core Metadata Initiative (2003). Dublin Core Metadata Element Set, Version 1.1: Reference Description. <http://www.dublincore.org/documents/dces/>.
- Gilliland-Swetland, Anne J.(2000). Setting the Stage. Introduction to Metadata: Pathways to Digital Information. Los Angeles: Getty Information Institute http://www.getty.edu/research/conducting_research/standards/intrometadata/pdf/swetland.pdf.
- Greenberg, J., Pattuelli, M.C., Parsia B. & Davenport R.W. (2001). Author-generated Dublin Core Metadata for Web Resources: A Baseline Study in an Organization. Journal of Digital Information 2(2). <http://jodi.ecs.soton.ac.uk/Articles/v02/i02/Greenberg/>.
- Hirtle, Peter B. (2002). The Impact of Digitization on Special Collections in Libraries. Libraries & Culture 37(1). http://muse.jhu.edu/journals/libraries_and_culture/v037/37.1hirtle.pdf.
- Jewett, Charles C. (1853). Smithsonian Report on the Construction of Catalogues of Libraries and Their Publication by Means of Separate, Stereotyped Titles. Washington, D.C.: Smithsonian Institution.
- Milstead S. & Feldman S. (1999). Metadata: Cataloging by Any Other Name... ONLINE 23(1). <http://www.infotoday.com/online/OL1999/milstead1.html>.
- Nagarkar, S., Parekh H., & Manjul, V.L. (2003). Study of 13-17th century Marathi Manuscripts: Author Generated Metadata and its Mapping with Dublin Core. <http://dc2003.ischool.washington.edu/Archive-03/03nagarkar.pdf>.

- National Information Standards Organization (2004). *Understanding Metadata*. Bethesda, MD: NISO Press.
<http://www.niso.org/standards/resources/UnderstandingMetadata.pdf>.
- Open Archives Initiative (2004). *The Open Archives Initiative Protocol for Metadata Harvesting*. <http://www.openarchives.org/OAI/openarchivesprotocol.html>.
- Powell, Ronald R. (1991). *Basic Research Methods for Librarians*. 2nd edition. Norwood, NJ : Ablex Pub. Corp.
- Schneider, Georg (1961). *Theory and History of Bibliography*. Translated by Ralph Robert Shaw. New York: The Scarecrow Press.
- Strout, Ruth French (1956). *The Development of the Catalog and Cataloging Codes*. *The Library Quarterly* 26:4, 254-275.
- Tanselle, G. Thomas (1998). *Literature and Artifacts*. Charlottesville : Bibliographical Society of the University of Virginia.
- Taylor, Archer (1957). *Book Catalogues: Their Varieties and Uses*. Chicago: The Newberry Library.