

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

The Journal of Academic Librarianship

journal homepage: www.elsevier.com/locate/jacalibMultilingual Scholarship: Non-English Sources and Reference Management Software[☆]

Adam H. Lisbon

University of Colorado Boulder, University Libraries, 184 UCB, 1720 Pleasant Street, Boulder, CO 80309-0184, United States

Introduction

Research happens across borders and languages, but software for managing and citing information is not designed to consider the linguistic diversity of the world's resources. Reference management software (RMS) such as Endnote, Mendeley, and Zotero, among many others help keep track of information sources for research projects. The strength of these programs is based on automatically pulling metadata from online sources to neatly show the title, author, date of publication, etc. of books, articles, and more. RMS should then automatically output metadata into proper citation and bibliographic formats, saving researcher's time. However, when sources are not in English, and especially when they are not written in Roman characters (i.e. the Latin alphabet), RMS have no way to faithfully store the original vernacular language side by side with transliterations/translations.

For example, the author 村上春樹 must be written as “Murakami Haruki” for the sake of English speakers who do not read Japanese. Likewise, his book 海辺のカフカ would need a transliteration of “Umibe no Kafuka,” and a translation of “Kafka on the Shore.” Publishers in the English speaking world often expect some combination of vernacular, transliteration, or translation when non-English sources are cited. Current RMS do not allow for representation of a work across multiple languages or scripts.

This article investigates how multilingual researchers (MLR) use or avoid RMS. Through an in-depth survey, this study intends to understand MLRs perception of RMS and what changes such software will have to make in their functionality in order to accurately represent non-English information sources.

Literature review

Research on RMS breaks down into several major categories: technical comparisons; LIS professionals and their knowledge of RMS, educating users; user attitudes on RMS and how they use it; and most recently, whether RMS even fits into the workflow of someone's research process.

Multilingual issues are rarely mentioned in RMS research. In 2011 and 2013, Francese investigated awareness and usage at the

Universities of Tallinn, Estonia and Torino, Italy. Wu and Chen (2012) asked a small focus group of students at National Taipei University about their experiences with RMS. Sarrafzadeh and Hazeri (2014) investigated use and awareness at Persian Gulf University, Iran. Ram and John Paul Anbu (2014) conducted a similar study in India. Throughout all of these studies, concerns over the representation of materials not in the native language(s) were never noticed, considered, or asked. Only Melles and Unsworth (2015) quoted one participant in their study that

“...I work, particularly for my primary documents, essentially entirely in French, and EndNote automatically capitalises and French titles don't capitalise. And I ... tried EndNote and found it so incredibly frustrating because I would have to go back and manually change everything. I gave up on it.”

Noguchi (2009) compared RefWorks, EndNote, and Zotero's ability to import Japanese scripts from four major databases: First Search, the Library of Congress Catalog, and four Japan specific databases: WebCat, the National Diet Library (of Japan) Catalog, CiNii, and Magazine Plus. At the time of the study, only RefWorks could automatically pull Japanese language data from FirstSearch. The CiNii database was also able to interact with these RMS with varying degrees of automation. For all other instances, manual entry was necessary. However, this wasn't a true multilingual study as it only explored citation management in one language, Japanese.

While Library and Information Science literature has not addressed multilingual research and RMS, the TeX/LaTeX/BibTeX community has been exploring solutions to multilingual typesetting and bibliographies. Harders (2002) and Hufflen (2009) each explored these implementations of BibTeX. However, these documents are technical, with a steep learning curve. Developing a knowledge base of TeX, LaTeX, and BibTeX is not practical for most RMS users. Mead and Berryman (2010), Francese (2011), Francese (2013), Hicks and Sinkinson (2015), and Melles and Unsworth (2015) all point to learning curves and/or time as a major reason why users don't develop a strong foundation using RMS. Childress (2011), Ram and John Paul Anbu (2014) further point out that even LIS professionals and graduate students don't always have a strong grasp of RMS software.

The first true discussion of multilingual RMS did not arise until

[☆] This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.
E-mail address: adam.lisbon@colorado.edu.

<https://doi.org/10.1016/j.acalib.2017.12.001>

Received 20 March 2017; Received in revised form 22 November 2017; Accepted 13 December 2017
0099-1333/ © 2017 Published by Elsevier Inc.

Bennett (2013) published *Citations Out of the Box*, introducing Juris-M. At the time it was called Multilingual Zotero (MLZ). The book did not explore RMS use, or the needs of MLRs, but was a technical explanation of how Juris-M works. It explained the development of the software. Built entirely off of Zotero's open access code, it was developed to allow parallel metadata fields for vernacular, transliterated, and translated titles, authors, publishers, etc. This allowed users to accurately represent their information sources "as they are" while simultaneously recording Romanized transliterations and/or translations that publishers may require when citing those materials.¹

In studies on RMS use, Juris-M has not been mentioned. In both of Francese's studies, Melles and Unsworth, Sarrafzadeh and Hazeri, Wu and Chen, all showed EndNote having the highest rate of awareness and use. Francese, 2013's study and Melles & Unsworth found that not using any RMS was the second most common experience, 24% and 29% in their respective samples. Melles & Unsworth pointed out that further research on RMS should not be premised on adoption rates, but on the user's needs and how they conduct research.

Adoption of RMS is driven by the recommendation of colleagues (Francese, 2013) or professors' recommendations to their students (Wu & Chen, 2012). EndNote, having been around since the 1980s, has a user base large enough to make word-of-mouth an effective means of propagating its use. Resistance to experimenting with newer programs like Zotero and Mendeley was pointed out by one of Melles and Unsworth's (2015) participants:

"I just look at [other software] and think, I've spent all these hundreds of hours and money on EndNote and I'll have to start again and learn Zotero. I've looked at Mendeley, and they all promise the world and then when you go to work them, it's not as easy..."

Francese (2013) notes that 87% of participants never asked for nor received support on how to use RMS. Even when training is provided, adoption is low. Melles and Unsworth (2015) noted one participant who was unable to grasp the usefulness of RMS even though they attended a training session. Francese (2013) found the two most used features were editing citations and inserting them into papers, a feature often referred to as "write and cite." Melles and Unsworth (2015) further noted RMS users were unaware of even this core feature. Hicks and Sinkinson (2015) found Mendeley users were unaware of its social networking capabilities, one of the particular features Mendeley markets to set itself apart from other RMS.

Childress (2011) addressed the idea of teaching RMS as a set of best practices. Melles and Unsworth (2015) further concluded that RMS must be taught in the context of existing practices, with special recognition that undergraduate students are still developing their own research skills and styles. Hicks and Sinkinson (2015) took this notion further, concluding that RMS are only one part of a larger nexus of digital literacies and digital scholarship. These concepts in turn are creating fundamental shifts in perceptions of what scholarship and the research process are.

Methods

For the purposes of this study, a multilingual researcher (MLR) was defined as anyone who uses information sources that are not in English, but publishes their findings in English. Participants did not need to exclusively use non-English sources, and qualified as conducting "multilingual" research even if they had used only a few information sources

¹ Juris-M also has special metadata fields for legal citations. Legal citation management is primarily how it is marketed. Bennett is a Law professor at the University of Nagoya, Japan. Hence his particular background on the issue of multilingual legal citation management. The project to develop Juris-M began in 2009, when it was originally known as Multilingual Zotero (MLZ). The name was changed in 2013 to emphasize the software's ability to cite legal documents, and to distance itself from the original Zotero software. Juris-M is an independent project.

in foreign languages. Students who produce class papers or dissertations also qualified under this definition. On average, participants conducted research in three languages in addition to English. Because publication requirements in the English speaking world were of primary interest, researching with information sources in English and publishing in another language or any other combination of languages was not considered.

Data on how MLRs use or do not use reference management software (RMS) was gathered through a survey of primarily multiple choice and Likert scale questions, with some open text questions to understand participants' unique situations and perceptions. Open text fields were analyzed for recurring themes and normalized. There were a total of 172 participants. The survey branches at key points to identify several sub-populations including: researchers primarily using Roman script sources ($N = 82$), those using primarily non-Roman script sources ($N = 90$), RMS users ($N = 61$), and non RMS users ($N = 111$).

In this study, "Roman scripts" were defined as those that use the Latin alphabet, and primarily include Romance and Slavic languages. Non-Roman scripts are those that use some other alphabet, such as Russian, Greek, Yiddish, or Hindi, or those that use characters like Chinese and Japanese. There are several languages that cannot neatly fit into these categories. Vietnamese, for example, used a non-Roman script before colonization, but French colonizers replaced the native writing system with one based on the Latin alphabet. Azerbaijani is another example, it was written in Cyrillic during the Soviet era, and today it is written with Latin letters.

The survey was open to anyone who wished to take it, with snowball sampling employed to gather participants across disciplines and languages. Subject specialist librarians at the CU-Boulder libraries shared the survey across the humanities, social sciences, and natural sciences. Participants were encouraged before and after the survey to share it with other colleagues, their departments, email lists, etc. 172 complete responses were recorded that met the definition of an MLR who publishes in English.

Demographics

Participants in this sample represented many academic disciplines. Most were in the humanities and social sciences. Despite trying to recruit researchers in the natural sciences, only two participated. Disciplines with significant representation included art & art history, classics, Jewish/Yiddish studies, linguistics, and history. On average, participants conducted research in three languages other than English. The top 15 most interacted with languages are listed in Fig. 1.

Fig. 2 shows how participants identified their professions/roles.

Professors, librarians, and PhD graduate students represented the majority of participants. Consequently, the sample reflected an academic scholarly perspective. Participants were also asked their age (see Fig. 3), with all but one volunteering this information. The over 55 demographic constituted the largest sub population ($n = xyz$), followed by the 26–35 bracket. 18–25 year olds had the lowest response rate. Finally, this sample overwhelmingly represented the experience of people living in the US ($n = 141$).

Multilingual researchers and their citation needs

In order to understand trends in RMS use, establishing expectations for citing non-English sources was essential. To understand this, participants were asked "When you submit/hand in/publish your research (or write class papers/theses), are you required to translate the titles of your sources?"

Fig. 4 shows Non-Roman MLRs are expected to translate the titles of their sources with greater frequency than their Roman MLR counterparts. Over 50% of Roman MLRs have never done so. Most Non-Roman MLRs translate their sources at some point. All participants answered this question regardless of using RMS or not. This trend is broken down

Download English Version:

<https://daneshyari.com/en/article/6842121>

Download Persian Version:

<https://daneshyari.com/article/6842121>

[Daneshyari.com](https://daneshyari.com)