A Study of Citations in Users' Online Personal Collections

Nishikant Kapoor[1], John T. Butler[2], Sean M. McNee[1], Gary C. Fouty[2], James A. Stemper[2], and Joseph A. Konstan[1]

[1] GroupLens Research, Department of Computer Science and Engineering, University of Minnesota, Minneapolis, MN 55455, USA
{nkapoor,mcnee,konstan}@cs.umn.edu
[2] University Libraries, 117 Pleasant St SE, University of Minnesota, Minneapolis, MN 55455, USA
{j-butl,g-fout,stemp003}@umn.edu

Abstract. Users' personal citation collections reflect users' interests and thus offer great potential for personalized digital services. We studied 18,120 citations in the personal collections of 96 users of RefWorks citation management system to understand these in terms of their resolvability i.e. how well these citations can be resolved to a unique identifier and to their online sources. While fewer than 4% of citations to articles in Journals and Conferences included a DOI, we were able to increase this resolvability to 50% by using a citation resolver. A much greater percentage of book citations included an ISBN (53%), but using an online resolver found ISBNs for an additional 20% of the book citations. Considering all citation types, we were able to resolve approximately 47% of all citations to either an online source or a unique identifier.

1 Introduction

Library users increasingly have access to integrated, online tools for managing personal citation collections. These tools help users to maintain a personal citation collection, to annotate the citations, and to produce formatted bibliographies and reference lists. While the first generation of these tools were largely disconnected from library search systems - in essence they were computerized versions of index cards with citations on them - newer versions can import citations directly from such library systems, and allow users to navigate directly from their citation to the document itself in an online collection or to the library's record for that document.

We are interested in understanding how these personal citation collections can be used to personalize the services provided to library users. Personalized library services have mostly evolved around published literature, and are therefore oriented towards serving the interests of the authors with publishing history. For example, TechLens from GroupLens [1,10], TalkMine and @ApWeb from APR [2], MyLibrary[1], use keywords and the citation index in the documents to generate recommendations for the user. If you have published any articles, and the

[1] http://dewey.library.nd.edu/mylibrary/

L. Kovács, N. Fuhr, and C. Meghini (Eds.): ECDL 2007, LNCS 4675, pp. 404–415, 2007.

system knows about at least some of them, the keywords and references in those documents would serve as a representation of your profile, which can then be used to generate more relevant or matching documents.

However, for users who are not authors of any published work, and do not have a history of publications from which to build their profile, personalized recommendations are often generated using one of the techniques like, keyword searching, explicit listing of preferred documents for the user, past searches, and descriptive profiling. We believe that personal citation collections of such users are representative of their interests, and possess a great potential for offering them personalized services. In particular, we have experience building recommender systems [3] for research articles [1,4], and wondered whether personal citation collections could be used to construct such a recommender system.

In this work we examined 18,120 citations from the personal collections of 96 users of an online citation management system, RefWorks[2]. Our goal was to empirically measure how frequently we could resolve the citations in these collections to a unique identifier (which is useful for collaborative filtering recommender systems) or to an online source for the content or content metadata (which is useful for content filtering recommender systems). Specifically, we examined the following two research questions:

- **What percentage of citations in users' personal collections can be resolved to a unique identifier?** What percentage of items in the collections is of a type that even has a unique identifier? For those that do, how many include the unique identifier in the citation? How many can be found using existing citation resolvers?
- **From what percentage of citations in users' personal collections can we navigate to an online source for the content or content metadata?** Are there citations that have an online source but not a unique identifier?

The rest of this paper is organized as follows. Section 2 provides background and related work on personalized services in digital libraries. Section 3 reports on the nature of the citation collections, breaking down citations by type and users by discipline and status. Section 4 reports on citation resolvability, looking first at resolving citations to a unique identifier, and then at resolving citations to an online source. Finally, Section 5 discusses the implications of this work for future library services as well as the privacy issues raised by such services.

2 Related Work

Digital libraries have been growing rapidly since early 1990s. They have been in extensive use in various sectors ranging from academics (University of Minnesota Libraries) to public (TEL-ME-MOR - The European Library: Modular Extensions for Mediating Online Resources) to government (NLM - U.S. National Library of Medicine).

[2] http://www.refworks.com/

There are number of applications that attempt to provide enhanced, and personalized library services to users, for example, MyLibrary creates a personalized web page listing information resources available from the Libraries based on user info. TalkMine and @ApWeb from the Active Recommendation Project (ARP) [2] use prior knowledge about the user to generate useful recommendations. TalkMine uses keywords, and @ApWeb uses association between the documents authored (or co-authored) by the user to learn about the user. TechLens [1,10] uses citation co-occurrence in published literature to generate personalized recommendations. The Quickstep and Foxtrot [5] systems recommended on-line research papers to academic researchers. PYTHIA-II [6] provides recommendations to scientists trying to identify the appropriate software for their research needs. The Ilumina project [7] provides recommendations based on document metadata, available subject expert analysis of documents, resource use as discovered in logs, and user profiles for those users who are registered with the system. The Melvyl Recommender project [8] analyzed server logs captured when users chose to view detailed information about certain documents, and used those as the user profile when generating recommendations.

Most of the research in providing such personalized services in digital libraries has focused on mining the content of the papers authored by the user, or the implicitly rated citations in the reference section of those papers (i.e., public collections of rated citations). Our work focuses on citations in users' personal collections. Users' personal citation collections are an implicit means to learn about their interests, and are representative profiles. These collections, therefore, offer a great potential for systems such as a citation recommender system to offer personalized tools and services in digital libraries [9]. However, in order to build such personalized library services, we first need to be able to link the citations in these collections to their unique IDs, i.e., evaluate their resolvability. Resolvability of citations enables us to match users' collections (a) with public repositories - to help users get customized selections, (b) with metadata - to build their profile, and (c) between the users to form correlations between them. Unique identification of identical citations in different users' collections is the key to building similarities between the users. It opens the door to matching citations between collections, and to obtaining additional metadata from which to generate recommendations.

3 Citation Data

We conducted this study in April and May 2006 using the RefWorks web-based personal citation management tool installed at the University of Minnesota. Subjects who met all of the following eligibility criteria were sent an email invitation:

- Had at least 10 citations in their RefWorks citation collection.
- Had actively used their RefWorks accounts at least once within the preceding six months (i.e., added, deleted, or modified at least one citation).
- Had logged on to their RefWorks accounts at least twice within the preceding six months.

Of the 1,253 users invited, only 96 accepted the invitation (7.67%), completed informed consent, and shared their personal citation collections with us. Limitations of our dataset are further discussed in the Discussion section.

Figure 1 shows the logarithmic distribution of citations for the 96 users. Each point on the line graph represents a user. The distribution is highly skewed, with an average number of citations of 316, a median of 99, and a mode of 37. 50 users had 100 or fewer citations in their collections, and only 19 had 300 or more citations.

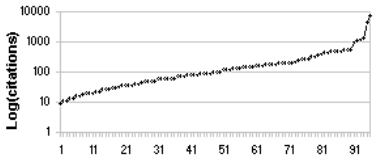

Fig. 1. Distribution of citation collection size

The collected data totaled 30,336 citations. There were, however, two high-end outliers with 7,777 and 4,439 citations in their collections respectively. Since the two outliers contained atypical data (not necessarily bad data), we conducted the data analysis with, and without the two outliers. However, due to space considerations, we are presenting only the results that do not contain the outlier data. Thus, the rest of our analysis is based on 94 users and 18,120 citations.

Figure 2 shows the distribution of citation types across our dataset of 18, 120 citations. A vast majority of citations (82.74%) are to articles in Journals and in Conference Proceedings. The high number of these citations is not surprising, given that articles in journals and conference proceedings are among the most easily found in online bibliographic search tools, and are the primary literature in many fields. Citations to books are a distant second with just over 7% of total citations.

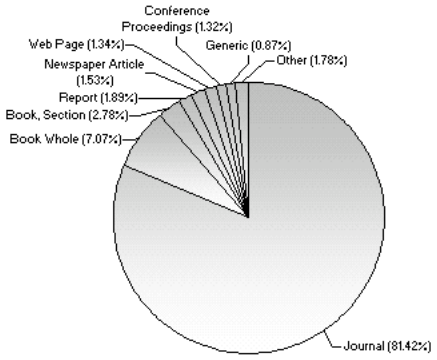

Fig. 2. Distribution of citation types

We should note that "citation type" is a property of the RefWorks citation. We were concerned that the citation type might be inaccurate if non-journal articles (e.g., book chapters or monographs) were mistakenly entered as journal articles, either because of manual entry errors or because of import filters that might improperly classify some references. To check for this possibility, we hand-validated a random sample of 100 citations classified as journal articles. We found that 98% of these citations were clearly journals, and the other two were to a university symposium series (which is published as if it were a serial) and Dissertation Abstracts International (which is a serial, if not a journal per se). Hence, we can conclude that we don't have a high percentage of mistaken classifications of non-journal entries as journal-type citations.

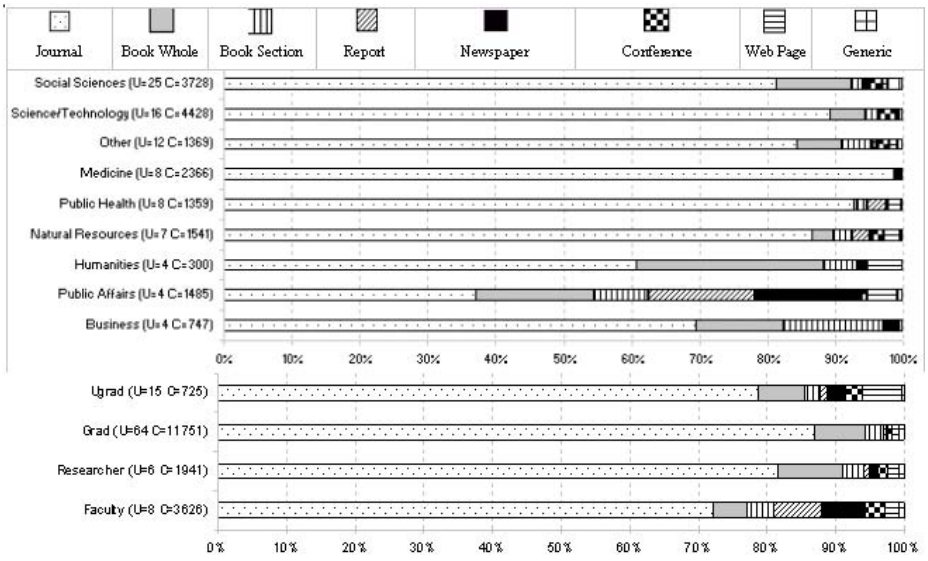

Fig. 3. Distribution of citations by (a) discipline (b) status

Figure 3 shows the citation distribution of the user disciplines and statuses, along with the number of users (U), and the number of citations (C) for the users in the discipline or status. Disciplines with fewer than four participants are not listed in the figure (and in all the subsequent tables and figures), but are aggregated into the totals.

In the dataset, the largest groups represented were graduate students (64 of 94 participants), and the disciplines of social sciences (25) and science and technology (16). Citations to Journal articles form the majority of the citations, ranging from the lowest at about 37% for Public Affairs, to the highest at about 98% for Medicine. Public Affairs has the most diverse collection - citations to Books, Newspaper articles and Reports, all hovering at about 17% of its total citations. Humanities is the largest consumer of citations to whole books at 28%.

For user statuses, the average number of citations to Journal articles is about 83% of their citations, ranging from a low of 72% for Faculty, to a high of about 87% for Graduate students. After Journal citations, whole books at about 9% form the next major share of their collections for Researchers. Within a collection, Faculty had the largest share of citations to Newspaper articles and to Reports at 7% and 6% respectively.

We analyzed the dataset by status, and found that the differences are statistically significant ($p <= .01$), except between Graduate and Researcher groups, and between Faculty and Researcher groups. Faculty had the highest number of citations per participant at 453, followed by Researchers, Graduates and Undergraduates. We carried out the similar analysis across the disciplines but there are not enough participants per discipline to generate results with confidence.

4 Citation Resolvability

When we look at citation resolvability, we are actually exploring two different questions. First we are asking whether the citation can be mapped to a unique identifier. This identifier serves as a means for determining whether different citations refer to the same underlying entity. Second, we are asking whether we can use the citation to find the entity being referenced online. Because of the different nature of books, articles, etc., we consider this type of resolvability to succeed if we can find either metadata for the entity cited, or the contents of the entity itself.

Neither of these types of resolvability automatically implies the other. A book's ISBN may serve as a unique identifier, yet we may be unable to find any online contents or metadata for that book. A technical report citation may come with a URL that leads us to the content online, yet lack any unique identifier to match it against other citations to the same report. The field as a whole, however, is moving quickly towards a greater percentage of unique IDs that also serve as pointers to online content. The DOI (Digital Object Identifier) for a journal or conference article serves both roles. We specifically explore three ways in which a citation may be resolved.

DOI: Citations to articles in Journals and in Conference Proceedings can potentially have a valid unique identifier, called Digital Object Identifier (DOI). DOIs, by their nature, serve as both unique identifiers and pointers to online content. There are three ways in which we can obtain a DOI for a citation: (a) the DOI may already be in the citation, stored in either the DOI field or the URL field; (b) we can use a DOI query at CrossRef[3], a service of the Publishers International Linking Association, which requires a journal title and author or page number, and uses other fields as provided, to look up a DOI; (c) we can construct an OpenURL - a URL to represent a reference in compliance with ANSI/NISO Standard Z39.88 and sending that OpenURL to a resolver to obtain a DOI (we used the resolver at CrossRef, which requires certain fields including author last name, title, year, volume, and issue).

ISBN: Book citations (which include whole books, monographs, and edited books) can potentially have a unique ID (a ten or thirteen digit long) called the International Standard Book Number (ISBN). There are two ways in which we can obtain an ISBN for a citations: (a) the ISBN may already be present in the ISBN field of the citation; or (b) we may be able to look up the ISBN using a resolver such as the one at WorldCat[4], a service of OCLC[5]. For this study, we did not include citations of type "Book Section" in our analysis because the ISBN uniquely identifies only a book, not a specific chapter or section.

URL: Any cited entity that is online may be found through a URL. We therefore separately examine all citations that include a value for the URL field, and validate that the URL points to an actual web page.

[3] http://www.crossref.org/
[4] http://www.worldcat.org/
[5] http://www.oclc.org/

For all the citation types in the dataset, we separated out the ones that could potentially include, or be resolved to a unique identifier, and analyzed each one of them separately to assess its resolvability. Since citations of other types such as Reports, Newspaper Articles, etc. do not necessarily have a unique ID (even though they might have an online source), they were not considered for resolvability analyses.

4.1 DOI Resolvability

Figure 4 shows the DOI resolvability breakdown by user discipline, and by user status (shaded portion). The last row of the table shows the cumulative statistics for the entire collection. The column headers for the table need some explanation before we get into detailed discussion of its contents.

Citations: Total number of citations for the discipline (or status) that could potentially have a DOI. These are citations to articles in Journals and Conference Proceedings.

Resolved: Total number of citations that resolved at CrossRef using either DOI, or OpenURL resolver. This column also shows the percentage resolvability of total citations in the discipline or the status.

RW only: Number of citations that had a valid DOI in the RefWorks record, but the CrossRef DOI resolver failed to find these. These DOIs were already available in the citations when we harvested the citation collections. These DOIs could either have been entered manually by the user at the time of building the collection, or automatically inserted by RefWorks citation management tool during import of the citation.

	Citations	Resolved	RW only	Both	CR only
Science/Technology	4,027	2,637 (65%)	10	100	2,527
Natural Resources	1,346	754 (56%)	4	76	674
Business	518	283 (55%)			283
Public Health	1,229	562 (46%)	2	58	502
Medicine	2,326	1,033 (44%)	2	109	922
Other	1,158	487 (42%)	12	81	394
Social Sciences	3,015	1,216 (40%)	1	59	1,156
Humanities	189	73 (39%)	2	14	57
Public Affairs	552	158 (29%)			158
Faculty	2,694	1,718 (64%)	1	39	1,678
Researcher	1,584	798 (50%)	14	92	692
Graduate	10,170	4,716 (46%)	18	388	4,310
Undergraduate	562	234 (42%)		11	223
Totals	**15,010**	**7,466 (50%)**	**33**	**530**	**6,903**

Fig. 4. DOI resolvability breakdown

CR only: Number of additional DOIs that we were able to fetch from CrossRef using the DOI and OpenURL resolvers. These DOIs were either not present in the original RefWorks record, or were invalid.

Both: Number of citations where the fetched DOI from CrossRef matched with the DOI already available in the RefWorks record. These DOIs already existed in the citations, and were valid.

The listing in the Figure 4 is sorted by the percentage of citations resolved for the discipline/ status. As shown, the citations from Science/ Technology have the highest resolvability at about 65%. The highest number of new DOIs (2,527) was fetched for citations from Science/ Technology; which is a clear indication that Science/ Technology is embracing the DOI standardization much faster than other disciplines. Business and Public Affairs, collectively had 441 resolved

citations, but no resolvable citations in their RefWorks records. All of the 441 DOIs were resolved by CrossRef.

The impact of DOI resolvability is clearly visible from Figure 4. Out of 15,010 potentially resolvable citations, only 566 citations (3.77%) had a valid DOI in their RefWorks account. Using CrossRef resolved another 6,903 DOIs, increasing the total resolvability to 7,466 citations (approximately 50%) - a significant improvement over the initial resolvability.

4.2 ISBN Resolvability

Figure 5 summarizes the resolvability of book and monograph citations to ISBNs. The columns in the table represent data similar to the DOI resolvability presentation, except for the following differences:

Citations: Total number of citations for the discipline (or status) that could potentially have an ISBN. These are citations to whole books, edited books and monographs.

Resolved: Total number of citations that resolved at WorldCat Libraries using the ISBN resolver.

RW only: Number of citations that had a valid ISBN in the RefWorks record when the collection data was harvested.

WC only: Number of additional ISBNs that we were able to fetch from WorldCat Libraries using the ISBN resolver.

	Citations	Resolved	RW only	Both	WC only
Other	98	83 (85%)	23	42	18
Business	97	81 (84%)	13	38	30
Public Affairs	258	215 (83%)	24	121	70
Social Sciences	414	301 (73%)	70	173	58
Science/Technology	236	164 (69%)	23	84	57
Humanities	78	53 (68%)	10	37	6
Medicine	16	10 (63%)		4	6
Natural Resources	51	22 (43%)	5	8	9
Public Health	36	7 (19%)		1	6
Graduate	916	713 (78%)	144	419	150
Faculty	180	126 (70%)	11	31	84
Undergrad	53	36 (68%)	9	11	16
Researcher	183	102 (56%)	19	68	15
Totals	**1,332**	**977 (73%)**	**183**	**529**	**265**

Fig. 5. ISBN resolvability breakdown

Both: Number of citations where the fetched ISBN from WorldCat matched with the ISBN already available in the RefWorks record.

The data in Figure 5 is sorted by the percentage of resolved citations for the discipline/ status. Each row can be interpreted in terms of its original resolvability and enhanced resolvability. For example, the total number of resolved citations for Science/ Technology is 164, of which 107 ISBNs (23 RW only + 84 Both) were already available in the original RefWorks record; and an additional 57 citations were resolved by the WorldCat ISBN resolver.

Overall, we were able to resolve 977 of 1,332 citations (73%). 183 of these citations had an ISBN in the record but were not found otherwise by the WorldCat Libraries ISBN resolver; 265 had no ISBN in the record but were found by the resolver; and 529 were found both in the record and by the resolver.

4.3 URL Resolvability

Figure 6 shows the URL validity breakdown by user discipline and user status. The data is sorted by the number of citations with a valid URL. The two columns are:

Citations: Total number of citations for the discipline (or status) that had a value in the URL field. These citations included all citations with a URL value, regardless of type.

Valid: Total number of citations that had a valid URL and were successfully traced to the web page they pointed to.

	Citations	Valid
Humanities	13	13 (100%)
Public Health	8	8 (100%)
Medicine	2	2 (100%)
Social Sciences	178	175 (98%)
Natural Resources	71	69 (97%)
Business	174	160 (92%)
Public Affairs	95	77 (81%)
Other	10	7 (70%)
Science/Technology	62	31 (50%)
Graduate	441	417 (95%)
Faculty	105	89 (85%)
Researcher	63	34 (54%)
Undergraduate	4	2 (50%)
Totals	**613**	**542 (88%)**

Fig. 6. URL validity breakdown

All citations that contained a value in the URL field were validated by following them to determine whether they point to a live web page. We examined 20 random URLs and found that 18 of them pointed to the correct content, the article. We did not attempt to find URLs for citations that did not have them, except in the case of DOIs discussed above. 613 citations had a value in the URL field, of which 7 were DOI-formatted. However, none of those 7 citations resolved using the DOI and/or OpenUrl citation resolvers at CrossRef. Of the total 613 citations, 542 were valid (about 88%). The high percentage is encouraging, given concerns that citations to web pages are likely to "go bad" over time.

At the same time, we note that faculty URLs are less likely to be valid than graduate student ones, which may be an indication that URLs go bad over time (e.g., pages go away or change address). Of particular note is the fact that science and technology URLs were the least resolvable, perhaps suggesting that they refer to more rapidly-changing and ephemeral content.

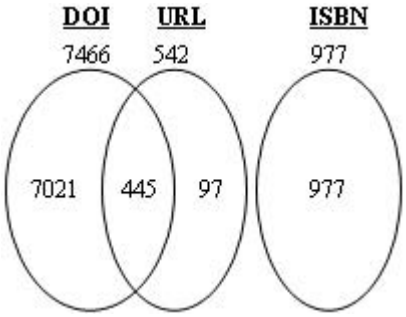

Fig. 7. Resolvability overlap

4.4 Resolvability into Multiple Identifiers

We addressed the issue of identifying citations by exploring the collections in three ways: (a) DOI resolvability, (b) ISBN resolvability, and (c) URL validity. Since DOI resolvability only looked at citations that could have a DOI, i.e., citations to articles in Journals and Conference Proceedings, and ISBN resolvability examined only the citations that could have an ISBN, i.e., citations to whole books, edited books, and monographs, there was no possibility for citation resolvability overlap between DOI-resolved citations and ISBN-resolved citations.

However, all citations could contain a URL. Accordingly, to understand the total resolvability of the dataset, we need to identify those cases where a single citation was resolved through both a URL and another mechanism. Figure 7 summarizes the resolvability overlap among the three IDs.

4.5 Cumulative Resolvability

DOI resolvers enhanced the resolvability of citations to articles in Journals and Conference Proceedings from less than 4% to about 50% (Figure 4). Figure 5 shows that using ISBN resolver enhanced ISBN resolvability from about 53% to about 73%. Figure 8 summarizes the overall resolvability of all the 18,120 citations of 94 citation collections. Using the DOI and ISBN resolvers, we were able to enhance the overall resolvability from under 13% to about 47%, gaining an impressive additional 34%. Since we did not retrieve any URLs for the citations, their resolvability remains unchanged.

5 Discussion

What percentage of citations in users' personal collections can be resolved to a unique identifier? Of all the citations that could be resolved to either a unique identifier or to an online source, we only were able to resolve certain citation types. Journal and conference papers have a unique online identifier (the DOI) (50%). Books have an ISBN, which can be used to find metadata and is a unique ID (73%). For other types of citation, the only possibility was a URL (88%), but their incidence is low (613). Citations to articles in Science and Technology have the highest resolvability at about 65%, and all 100% of citations to books resolved in Nursing. The citation collections of Faculty members are likely to be more resolvable than Researchers, Graduate and Undergraduate students.

	Number of resolved citations		
	Before	After	Gain
DOI	1,083 (7%)	7,466 (50%)	6,383 (43%)
ISBN	712 (53%)	977 (73%)	265 (20%)
URL	542 (88%)	542 (88%)	n/a
Totals	**2,337 (13%)**	**8,540 (47%)**	**6,203 (34%)**

Fig. 8. Resolvability summary

From what percentage of citations in users' personal collections can we navigate to an online source for the content or content metadata? The original dataset of users' personal citation collections was only 13% resolvable, i.e., only 2,337 citations out of the total 18,120 had either have a valid

DOI, ISBN, or URL. We found that many citations in users' personal collections were not well formed. Some were bad citations which we examined and could not resolve by hand and some were duplicate citations. We speculate that some users may bulk import the results of queries, not examine and select each citation individually. Such a practice would call into question whether the citation collections are indeed good sources of user preference information.

In this study, we analyzed resolvability of users' citation collections. There are, however, other necessary criteria that we are still studying such as, whether the resolved citations are representative of the user's interests, and how much of an overlap is there between these collections.

Privacy Concerns. Most invited subjects failed to respond at all, but 14 replied and actively declined to participate in this study, citing various concerns including privacy. This is a serious issue. While many people seem happy to share citation collections (and do so online in many venues from citeUlike to ACM DL binders), others have legitimate concerns about revealing too much information about their current interests or pursuits to possible competitors. Even if the citation collections themselves are not public, we recognize that there are scenarios in which a recommender system might compromise a user's private citations by recommending them to another user. Understanding how to protect this data is critical if recommender-enhanced library services are to gain wide acceptance. We suspect that personal collections may be used to generate profiles, but that only public citations (e.g., references in published papers) can be used to recommend citations to users.

Limitations. There are several limitations to this work. Our dataset came from volunteer participants (who may not be representative of the entire user base) and had too few representatives of some disciplines to adequately draw conclusions about those fields. We used only a single source for resolving each type of citations (CrossRef and WorldCat); while we generated significant improvements in resolvability, we believe that adding additional services for lookup (for example, Citation Matcher from PubMed) could further enhance resolvability. Other tools (e.g., Google, Google Scholar) are capable of finding unofficial copies of articles that may not be easily found through a resolver. Finally, due to the size of our dataset, we were unable to hand-validate the resolved citations. We did random spot-checks, but it is possible that some of the DOIs, ISBNs, and URLs we validated did not match the user's intended citation.

6 Conclusion

We studied 18,120 citations in the personal collections of 96 users of RefWorks citation management system. These users included graduates, undergraduates, researchers, and faculty - graduates being the majority (64). We believed that such personal collections represent users' interests and have a tremendous potential in guiding and supporting personalized services in digital libraries, such as a citation recommender system. The primary objective of this study was to evaluate the resolvability of users' personal citation collections. While fewer than 4% of citations

to articles in Journals and Conferences included a DOI, we were able to increase this resolvability to 50% by using a citation resolver. A greater percentage of book citations included an ISBN (53%), but using an online resolver found ISBNs for an additional 20% of the books. All together, we were able to resolve approximately 47% of all citations to either an online source or a unique identifier.

Acknowledgements

We are grateful to RefWorks for their assistance in carrying out this experiment. This research was funded in part by a grant from the University of Minnesota Libraries, and in part by grant IIS-0534939 from the National Science Foundation.

References

1. McNee, S., Albert, I., Cosley, D., Gopalkrishnan, P., Lam, S.K., Rashid, A.M., Konstan, J.A., Riedl, J.: On the Recommending of Citations for Research Papers. In: Proceedings of ACM 2002 Conference on Computer Supported Cooperative Work (CSCW2002), New Orleans, LA, pp. 116–125 (2002)
2. Rocha, L.M.: TalkMine: a soft computing approach to adaptive knowledge recommendation. In: Loia, V., Sessa, S. (eds.) Soft Computing Agents: New Trends For Designing Autonomous Systems, Springer Studies In Fuzziness And Soft Computing, pp. 89–116. Physica-Verlag GmbH, Heidelberg (2002)
3. Resnick, P., Varian, H.R.: Recommender systems. Commun. ACM 40(3), 56–58 (1997)
4. Torres, R., McNee, S.M., Abel, M., Konstan, J.A., Riedl, J.: Enhancing digital libraries with TechLens+. In: Proceedings of the 4th ACM/IEEE-CS Joint Conference on Digital Libraries, JCDL '04, Tuscon, AZ, USA, June 07 - 11, 2004, pp. 228–236. ACM Press, New York (2004)
5. Middleton, S.E., Shadbolt, N.R., De Roure, D.C.: Ontological user profiling in recommender systems. ACM Trans. Inf. Syst. 22(1), 54–88 (2004)
6. Houstis, E.N., Catlin, A.C., Rice, J.R., Verykios, V.S., Ramakrishnan, N., Houstis, C.E.: PYTHIA-II: a knowledge/database system for managing performance data and recommending scientific software. ACM Trans. Math. Softw. 26(2), 227–253 (2000)
7. Geisler, G., McArthur, D., Giersch, S.: Developing recommendation services for a digital library with uncertain and changing data. In: Proceedings of the 1st ACM/IEEE-CS joint conference on Digital libraries, Roanoke, Virginia, United States, pp. 199–200 (January 2001)
8. Geyer-Schulz, A., Neumann, A., Thede, A.: Others Also Use: A Robust Recommender System for Scientific Libraries. In: Koch, T., Sølvberg, I.T. (eds.) ECDL 2003. LNCS, vol. 2769, pp. 113–125. Springer, Heidelberg (2003)
9. Renda, M.E., Straccia, U.: A personalized collaborative digital library environment: a model and an application. Inf. Process. Manage 41(1), 5–21 (2005)
10. Konstan, J.A., Kapoor, N., McNee, S.M., Butler, J.T.: TechLens: Exploring the Use of Recommenders to Support Users of Digital Libraries. CNI Fall Task Force Meeting Project Briefing. Coalition for Networked Information. Phoenix, AZ (2005)