License Analysis on e-Journal Perpetual Access of Library
Mei Zhang, Kristin Eschenfelder
School of Library and Information Studies
University of Wisconsin-Madison
4207D, Helen C. White Hall
600 N. Park Street
Madison, WI, 53705
mzhang48@wisc.edu, eschenfelder@wisc.edu

ABSTRACT
This poster presents preliminary findings of a content analysis of 72 e-journal licenses in terms of perpetual access clauses. This study explored whether perpetual access clauses varied among commercial publishers, society publishers and university presses, whether clauses changed over time, and whether differences existed between consortia and site licenses. Findings showed that the licenses from commercial publisher were more likely to provide perpetual access than licenses from other two types of publishers. We observed an increasing trend in providing perpetual access over time. Consortia licenses were more likely to depend on third-party to obtain perpetual access than site licenses. This study can be helpful for libraries to understand the institutionalization of perpetual access clause in e-journal licenses.

Keywords: perpetual access, e-journal license, content analysis

INTRODUCTION
With the development of information technologies and the Internet, scholars in scientific and research fields are coming to rely on digital resources such as e-journals and e-books, and library expenditures on digital resources have increased sharply. At the same time, libraries face new challenges in the digital environment, especially regarding cost. According to Susman and Ropes’ research (2003), partly because of the merger of publishers, journal subscription fees in the digital environment inflated, which would limit users’ access to important research information. Pricing difficulties mean that libraries have to cancel digital resource subscriptions, raising the issue of “perpetual access” or to what extent authorized users might have access to previously subscribed (but now cancelled) digital materials. Because digital resources exist on publisher servers rather than library shelves, it is unclear whether or not the library can maintain access to the previously subscribed materials after their cancellation of these resources.

Digital Library Federation (DLF) Electronic Resource Management Initiative (ERMI), according to its report, developed the term, “perpetual access” (using “PA” instead), also known as “permanent access” or “post-cancellation access”, to describe the situation where the library can “permanently access licensed materials paid for during the period of a license agreement” (Jewell et al., 2004, p. 158). While cancellation is a large motivator for PA, several other factors contribute to library’s need for PA including sale of titles between publishers, mergers of publishers and insolvency of publishers. For example, North East Research Libraries (NERL) consortium states that PA should be ensured under the following contingencies: “mergers and acquisitions, insololvency, or transfers of ownership to another publisher” (NERL, 2011).

Similar to DLF’s definition, there are some other LIS professionals and publishers defined PA as libraries’ ability to get access to the previously subscribed resources after they cancel the subscriptions (Waller & Bird, 2005; Park, 2007; Rogers, 2009; Perpetual Access, n.d.; ONIX-PL, n.d.). While library and publisher stakeholders both employ this basic definition, the details differ. In “offering PA” different publishers may offer radically different services. One area of variance is the location of PA contents. PA can be provided on various media and in various locations (Waller and Bird, 2005). Another area of variance is price. There are different opinions on whether PA ought to be free or not (Beh & Smith, 2012). Moreover, what qualifies as a trigger event to provide access to PA material varies (Waller & Bird, 2005; Park, 2007; Rogers, 2009; NERL, 2011). These differences in the details about PA services may lead to different expectations about what PA services should include. Further, the differences may complicate communication about PA within the e-journal publishing community.

Among the current studies on PA issues, some researchers conducted surveys to libraries or publishers to investigate the status of PA in various libraries (Muir, 2003; Waller and Bird, 2005; Farb, 2006; Rogers, 2009; Carr, 2011). Meanwhile, some other studies have analyzed the e-journal licenses for a better understanding of PA status (Millett, 2001; Farb, 2006; Stemer and Barrieau, 2006; Rogers, 2009). A close examination of these studies reveals a significant difference in the studies’ findings concerning the percentage of publishers that grant PA to libraries. One explanation could be the different times when the research was conducted, and the different libraries and publishers included. Another reason might be the different definitions used to determine whether or not publishers provide PA. Different researchers might measure the status of PA differently. Due to the methodological issues described
earlier. This study seeks to clarify changes in PA over time and the current status of PA in licenses based on a large set of licenses spanning 2001-2009.

RESEARCH QUESTIONS
This study would be limited to digital academic resources to analyze the PA status as mentioned in their licenses. Considering the differences existing between e-books and e-journals on PA issues, the paper will only discuss academic e-journals. Electronic proceedings and monographs will not be covered by the research.

This study will answer the following research questions:
(1) How is PA defined in licenses?
(2) Do the attributes of PA offered vary in terms of:
   • Different types of publishers: commercial publishers, society publishers and university press?
   • Different time periods?
   • Different types of licenses: site license for individual library and consortium license?

METHODS
The data set analyzed in the paper is based on the licenses collected by Bergstrom, Courant and McAfee for a study of journal pricing in the United States (Bergstrom, 2009). These licenses came from 11 different publishers from 2000-2009. We received 216 licenses on a CD-ROM from Bergstrom. Due to the limited time and resources, we selected 72 licenses (33.3%) from the data set to analyze.

We used disproportionate stratified sampling (publisher types as the stratification variable) to pull out 72 licenses from the dataset. This study included 36 licenses from 8 commercial publishers, 18 licenses from 2 university presses, and 18 licenses from 1 society publisher. A simple random sampling was applied to draw samples from each group. Also, this study divided the time category of those licenses into 2 groups (2000-2004, and 2005-2009). However, the sample number of each time period is obviously skewed to the second group (2005-2009 with 40 licenses) due to the unbalanced time periods in the received data set.

In the first stage of codebook development, we developed an initial draft of codebook based on the literature review. Then we revised the codebook based on 9 rounds of coding conducted from June to August, 2011. During each round of coding, we independently coded 3 licenses pulled from Bergstrom’s dataset which were not included in the samples, and the inter-coder reliability (ICR) was calculated. In the last round of coding, we obtained 98% of ICR, higher than the targeted 90%.

In the final codebook, 39 variables were divided in three categories: definition of PA, location of PA, and services related to PA. Basically, the variables were listed as a statement, and we marked “1” for the variable if the statement was true, and marked “0” for the opposite of the statement. Otherwise, “99” was marked for the license which did not mention the statement at all. 39 licenses from 6 different publishers were assigned to each of the two coders. Both coders analyzed 6 licenses in common to calculate the ICR for the final coding.

Licenses coding ended in September, and ICR for the 6 licenses was calculated as 99%. All data collected in coding were then imported to SPSS for further analysis. In order to reveal the differences in PA clauses in terms of the type of publishers and licensee, and time variance, we employed both percentage calculation and chi-square analysis. However, due to the small sample size, we also used Fisher’s exact test as a supplementary analysis when the assumption of chi-square test (fe≥5 in all cells) was violated.

FINDINGS
How to define PA in license
We explored the ways how PA was defined in licenses. We found that most licenses included the following factors in their PA clauses: 1) PA will be provided when the license is automatically terminated due to the expiry of subscription period. Majority of the licenses did not include the situations that happened during or after the subscription period as their trigger event for providing PA. 2) Only the content that is “published for the first time during the subscription period” will be provided through PA. Therefore, back files are not covered by PA. 3) Location for PA is clarified in licenses. By specifying the location where libraries can get access to the licensed materials, licenses become easier and more practical to follow up on, which will ensure the library’s ability to get PA to some extent.

Differences in PA by publisher type
We found out there were several differences exiting in PA among different types of publishers. The most significant difference across publisher types is the number of licenses which forbade PA when the situation was beyond control. Most licenses from commercial publishers (62.9%) would not offer PA, while licenses from other types of publishers were more likely to keep silent on this issue.

Another difference is the percentage of licenses offering PA. As illustrated in Figure 1, commercial publishers were more willing to provide PA in their licenses (86.1%) compared with university press (72.2%) and society publisher (33.3%). The finding partly echoed Stemper and Barribeau’s research output (2006) that commercial publishers were more likely to offer PA than society publishers.
We also noticed that commercial publishers’ licenses would provide more detailed information in the PA clause. For instance, commercial publishers’ licenses were more likely to specify the location of PA, and only licenses from commercial publishers mentioned any particular third-party service. Moreover, commercial publishers’ licenses were more likely to provide PA in other conditions besides the automatic termination of subscription period. Finally, there were more licenses from commercial publishers to define the services related to PA in detail, such as the charges and features.

**Differences in PA over time**
As shown in Figure 2, we found changes in licenses in terms of basic attitudes towards PA due to the decreased number of licenses which did not mention PA at all and the increased number of licenses which provided PA. Moreover, there is a significant difference in terms of precluding the conditions which were beyond control from the trigger event of PA. In 2000-2004, there were only 26.3% of licenses addressing that PA cannot be provided under such situation, while the number rose to 70.3% in 2005-2009.

**Differences in PA between site license and consortium license**
The most significant difference between site license and consortium license existed in term of choosing third-party services as a way to provide PA. More specifically speaking, consortium licenses were more likely to depend on third-parties than site license (40% to 20%). Another difference between the two types of licenses in terms of charging for library’s copy (16.1% to 36%). Besides, we found the licenses from consortia are only slightly more likely to provide PA than individual libraries’ site licenses (73.3% to 66.7%).

**IMPLICATION**
This study can be helpful for libraries to understand the institutionalization of PA clause in licenses. Tolbert and Zucker, as major contributors to Institutional Theory, defined “institutionalization” as “a core process in the creation and perpetuation of enduring social groups” (Tolbert and Zucker, 1996, p. 180). Institutionalization, when applied to PA studies, can be used by libraries and publishers to examine the development of PA clause in licenses. More specifically, studying the licenses over times provides empirical evidence to reveal the different phrases of institutionalization: habitualization (new PA clause appeared in several licenses), objectification (increasing adoption of the new PA clause in licenses), and sedimentation (such PA clause is accepted by most licenses and survival for a period of time) (Tolbert and Zucker, 1996). Furthermore, comparative studies among different types of publishers, as well as different types of licenses can be used to explore the interactions between institutions (PA clause in license) and agents (different types of publishers and libraries) during the institutionalization process according to Barley and Tolbert’s sequential model of institutionalization (1997).

**LIMITATIONS AND FUTURE STUDY**
The major limitation of the study is the number of sampling. Since there are only 72 licenses analyzed in the study, the reliability of findings could be reduced to a large extent. The problem could be solved when introducing more samples into the study.

The second limitation of the study is about its validity. The data set analyzed in the research is drawn from Bergstrom’s research; however, Bergstrom’s data set does not represent a random sample, because all the licenses are from large universities in each of the responding states, and only 11 publishers responded. In addition, there is only one publisher (ACS) in society publisher group. Therefore, the findings of this study may not be able to represent the PA status of all academic e-journal licenses in the U.S., and it is very limited in what it can say about all publishers. Increasing the variety of the sample licenses in future would be helpful to improve the validity of the study.

Last but not least, lack of ways to reveal reasons for the current status of PA is also a limitation of findings in the study. License analysis itself cannot provide ways to allow researchers to investigate the dynamics among the stakeholders hidden behind the words of license. Therefore, incorporating other research methods, for instance, conducting interview or survey with stakeholders, to
explore the process of formation and development of a license could be another direction for exploring the PA issues in the future.

REFERENCES


