Exploration of Adoption of Preservation Metadata in Cultural Heritage Institutions: Case of PREMIS

Daniel Gelaw Alemneh
University of North Texas
Information Technology Services
Denton, Texas-76201, USA
daniel.alemneh@unt.edu

Samantha Kelly Hastings
Director and Professor, School of Library &
Information Science, University of South Carolina,
Columbia, SC 29208, USA
hastings@sc.edu

ABSTRACT

The challenges of long-term access issues are multifaceted, often requires a mixture of approaches. Considering the critical role of metadata in any successful digital preservation strategy, the Preservation Metadata Implementation Strategies (PREMIS) has been extremely influential on providing a "core" set of metadata elements that support the digital preservation process. However, there is no evidence, in the form of previous research, as to what factors explain and predict the level of adoption of PREMIS. This paper attempts to identify factors that affect the adoption of PREMIS in cultural heritage institutions. The study employed a web-based survey to collect data from 123 participants in 20 country as well as a semi-structured, follow-up telephone interview with a smaller sample of the survey respondents. Roger's diffusion of innovation theory was used as a theoretical framework. The main constructs considered for the study were relative advantage, compatibility, complexity, trialability, observability, and institution readiness. The analysis showed that all six factors influence the adoption of PREMIS in varying degrees. Results of a regression analysis also showed a statistically significant relationship. The R square value for the model was .528, which means that 52.8% of the variance in PREMIS adoption was explained by a combination of the six factors. This research just barely begins to show the many layers of the complex problem of digital preservation. This study has important implications for future research on preservation metadata and provides recommendations for researchers and stakeholders engaged in digital preservation and metadata standards development efforts.

Keywords

Digital preservation, PREMIS, Preservation metadata, Diffusion of innovations, Adoption of innovations.

INTRODUCTION

Today the entire information landscape has changed and continues to change at a breathtaking pace. Digital technologies are shaping creation, management, access, and preservation of information in ways that are so profound that traditional methods may no longer be effective. The main technical problems of digital preservation relate to inadequate media longevity, rapid hardware obsolescence, and dependencies on particular software products. In addition to technological issues, responsible and viable preservation planning for digital materials need to address various issues, such as policy, economic, and organizational issues.

Different communities are developing and implementing various digital preservation methods at different rates. Considering the complex set of digital preservation challenges, many researchers (Besser, 2002; Day, 2006; Hedstrom, 2003; and Lavoie, 2008) agree that there are no effective preservation methods or tools that work for all communities or types of resources. There is a fundamental need to know more about digital preservation in general. However, most agree that metadata plays a significant role in any preservation activities. There is an overwhelming consensus among experts that PREMIS provides required standards and best practices for the use of metadata in support of long-term preservation (Day, 2006; and Lavoie, 2008). Although PREMIS is becoming more popular among cultural heritage institutions, there is no evidence that explains and predicts the level of adoption

The purpose of this exploratory research is to identify factors that affect adoption of preservation metadata, specifically PREMIS, in cultural heritage institutions (libraries, archive, museums, and other repositories) using the theoretical framework provided by the diffusion of innovations theory. Bradford and Florin (2003) and Buonanno et al. (2005) say that understanding adoption of innovation in any given situation requires identification and analysis of factors that may facilitate the adoption and those that may operate as barriers to adoption. The diffusion of innovations theory provides a model for conceptualizing the acceptance of PREMIS in a cultural heritage community.

Rogers' (2003) five perceived characteristics or attributes are shown to influence the rate of adoption of innovation namely: relative advantage, compatibility, complexity, trialability, and observability. Various researchers have examined the diffusion of information technologies and related innovations using the framework from the diffusion of innovations theory. Digital preservation metadata is part of digital technology innovation, and it would be expected that factors which have been found to be related to other digital technology innovations would also explain the adoption of PREMIS in the cultural heritage institutions.
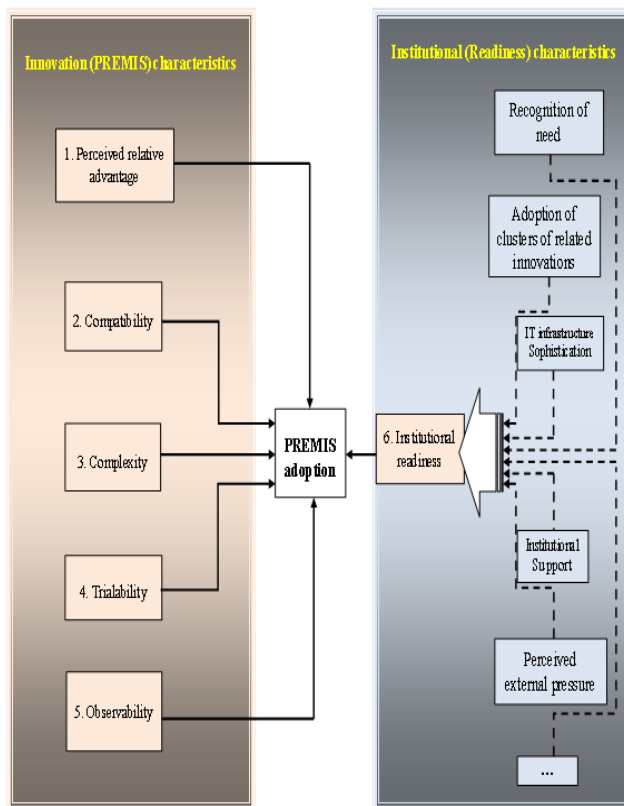

Figure 1. Factors affecting PREMIS adoptions.

METHODOLOGY OF THE STUDY

This exploratory research study examines the factors that affect adoption of PREMIS in cultural heritage institutions. To gain a broad understanding of factors that affect adoption decision in the context of cultural heritage institutions, the following two research questions are addressed:

1. What are the factors (i.e., attributes in the diffusion of innovation theory) that affect the adoption of PREMIS across the diverse cultural heritage institutions?
2. Among the diverse cultural heritage institutions that adopted or plan to adopt PREMIS, are there commonalities in factors that may affect the decision-making process?

These research questions provide the framework for identifying and examining the critical factors influencing stakeholders in the adoption decisions regarding PREMIS in cultural heritage institutions.

The literature review revealed that in addition to Rogers' (2003) five perceived characteristics, several other attributes have been added including several key organizational factors that may affect innovation adoption decision. A number of researchers (Chen, 2003; Fichman & Kemerer 1997; Gallivan, 2001; Themistocleous & Irani, 2002, among others) noted that readiness of organization is strongly associated to other parameters such as organization culture, skill sets, and IT infrastructure (architecture, sophistication, skill sets, etc.). Institutional readiness is thus conceptualized as an adoption characteristic, for the purpose of this study. Figure-1 shows the constructs utilized to understand factors affecting PREMIS adoption in cultural heritage institutions:

1. Relative advantage refers to the perceived advantages or benefits of the innovation.
2. Compatibility refers to the degree to which an innovation is perceived as being consistent with the existing values, past experiences, and needs of potential adopters.
3. Complexity is used as perceived ease of use.
4. Trialability refers to the degree to which an innovation may be experimented with on limited basis.
5. Observability, Observability refers to the degree to which the results of an innovation are visible to others.
6. Institution readiness (which includes a number of determinants, such as *availability of resources/support, IT infrastructure sophistication, adoption of related innovations etc.*).

DEMOGRAPHIC CHARACTERSTICS OF PARTICIPANTS

This research used survey questionnaires and telephone interviews to collect both quantitative and qualitative data from 123 participants in 20 countries.
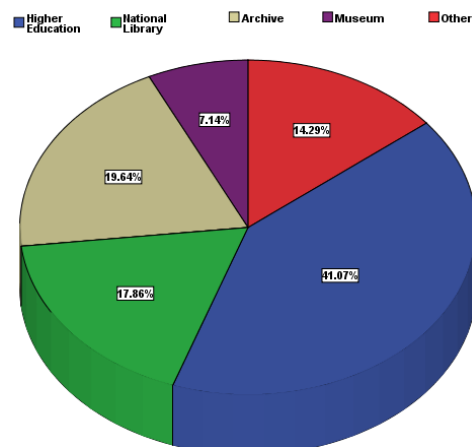

Figure 2. Survey respondents by institution type.

The researchers conducted semi-structured, follow-up telephone interviews with the survey respondents who were willing to participate in the interview to clarify and confirm the questionnaire data. By consulting and involving actual stakeholders (who are active in developing improved solutions for digital preservation challenges), factors that affect adoption of preservation metadata identified and discussed. Such triangulations of methods provided a holistic framework to identify factors and their relationship in order to understand the factors that affect adoption of preservation metadata in cultural heritage institutions.

The survey questionnaire collected several demographic characteristics of the research participants such as their institution affiliations, locations, positions, and levels of education. As can be seen from Figure 2, respondents were predominantly from higher education institutions (about 40%), followed by archives (about 18%), museums (16%), and national libraries (9%). Some of the participant institutions categorized as *others* include: government and non-government research institutes, digital documents producers (e.g., publishers, broadcasting agencies, or image service companies), non-profit art institutions, and other libraries (e.g., public, state, and charity libraries).
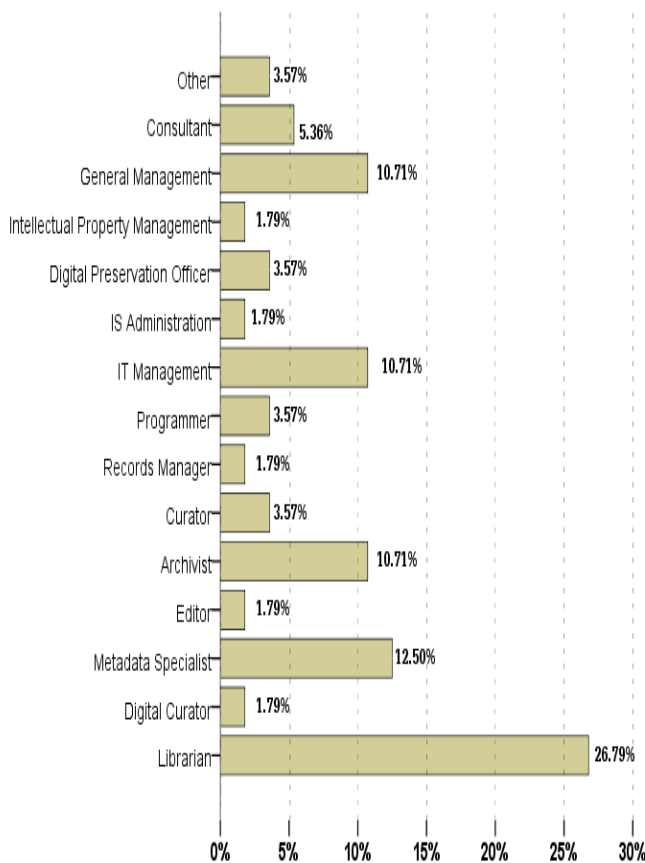

Figure 3. Distribution of the participants by specialty.

Respondents' fields of specialty were distributed as shown in Figure 3; a significant number (about one-fourth or 26%) of the respondents were librarians, while more than 20 % represented IT and general management positions. However, it is interesting to note that many respondents to the survey indicated that they regarded themselves as metadata specialists, archivists, digital curators, intellectual property managers, and digital preservation officers. As noted by Lee, Tibbo, and Schaefer (2007), in the past a data creator may have had little or nothing to do with subsequent curation, but today the digital environment demands that understanding bridge the differing roles.

PREMIS ADOPTION STATUS

Cultural heritage institutions accept the notion that maintaining usable and sustainable digital collections requires a complex set of actions. In this regard, most respondents agreed that preservation metadata is crucial to implementing reliable, usable, and sustainable digital libraries. Most (more than 91%) of the respondents in the sample believed that preservation metadata help resource managers in analyzing data and facilitating preservation decisions and actions.

However, the institutional context for preservation metadata requirements may differ across cultural heritage institutions. In this regard, despite the role of preservation metadata in digital resource life cycle management, a number of survey participants mentioned that they viewed PREMIS adoption in light of their institutions' specific characteristics. One respondent said that "PREMIS is more library-centric and our team members are cautious in recommending full PREMIS adoption." As can be seen from their statements, many emphasized that the institutional context actually matters when it comes to adopting PREMIS. The following statement from one of the interview respondents sums up the views of many participants: "While we can be informed by PREMIS and what worked elsewhere in terms adopting preservation metadata, we must take account of our own local specific conditions before implementing change."

PREMIS is relatively new innovation and perhaps growth is somewhat slow as the innovation establishes itself. Literature on diffusion of innovations suggests that time could be a vital factor in adoption of the innovation. Rogers (2003) defined five main steps in an innovation-decision process: learning of an innovation's existence and some of its functions (knowledge); forming a favorable or unfavorable attitude toward it (persuasion); engaging in activities that lead to an adopt/reject choice (decision); putting the innovation into use (implementation); and seeking information that reinforces or refutes the innovation-decision (confirmation).
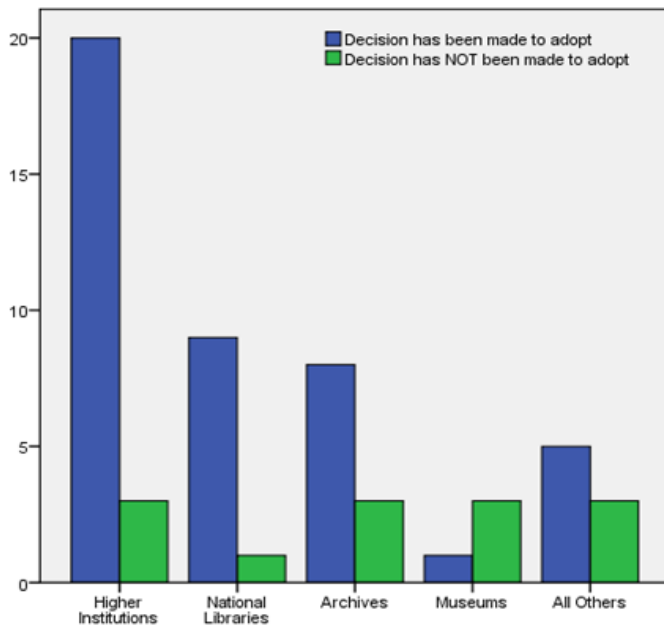
Figure 4. Distribution of the participants by specialty.

The data analysis revealed that a vast majority of the institutions had not yet reached the development stage in terms of their level of PREMIS adoption. Although academic institutions and national libraries were among the early adopters, the overall adoption was not that high. Out of the 123 participants who responded to the survey, only 4 institutions (fewer than 3% of the respondents) had fully adopted PREMIS. As can be seen from Figure-4, even though there are some commonalities, there are notable differences among cultural heritage institutions.

Barriers or Inhibiting Factors

Most frequently identified barriers or factors that discourage PREMIS adoption include:

- Lack of training/expertise (48.1%).

- Lack of integration or incompatibility with existing system (37.0%).

- We lack the knowledge necessary to be confident in our ability to implement the PREMIS (29.6%).

- Lack of interest from the decision-makers within our institute (24.1%).

- Institutions that have adopted the PREMIS cannot provide evidence of its effectiveness (20.4%).

- Usability requirements are too high (20.4%).

- Our institution prefers to take a wait-and-see approach when it comes to adopting new system (14.8%).

- Our institution has limited capacity to absorb negative consequences that might occur as a result of implementing the PREMIS (14.8%).

Facilitators or Stimulant Factors

Most frequently identified factors that facilitate PREMIS adoption:

- Adopting the PREMIS is seen as a practical necessity by our institution (58.9%).

- My institution has the resources necessary to support the adoption of the PREMIS (39.3%).

- My institution has enough technical knowledge to adopt the PREMIS (35.7%).

- Interest from the decision-makers within our institution (35.7%).

- Most cultural heritage institutions are adopting the PREMIS or seriously considering it. (25.0%)

- Benefits will outweigh costs when it comes to adopting the PREMIS at our institution (32.1%).

- The PREMIS is compatibility with existing system (21.4%).

- From a technical standpoint, it is/will be easy to implement the PREMIS (12.5%).

Other Factors

Although PREMIS played significant role in analyzing preservation requirements, it cannot accommodate all metadata requirements. As Dappert and Enders (2008) noted, no single existing metadata schema accommodates the representation of descriptive, preservation, and structural metadata. Most cultural heritage institutions have a great deal of experience with other metadata management tools (some are actually running up-to-date and fairly sophisticated tools and systems). In view of that, respondents were asked to identify tools that they use for preservation metadata generation and/or extraction of technical metadata. Most respondents use one or more format identifications tools. The open source JHOVE characterization tool was identified as one of the widely used components of many (55.4%) cultural heritage institutions' digital preservation workflows. Other format identifications tools identified by respondents include: Digital Record Object Identification (DROID), the National Library of New Zealand (NLNZ) Metadata Extraction tool, the Global Digital Registry (GDFR), the AHDS Collections Management System, DigiTool technical metadata extraction tool, and OCLC Metadata Extraction Tools. All these format identifications tools are designed to provide basic information on the file format. However, they may not provide sufficient detail to preserve the digital resource.
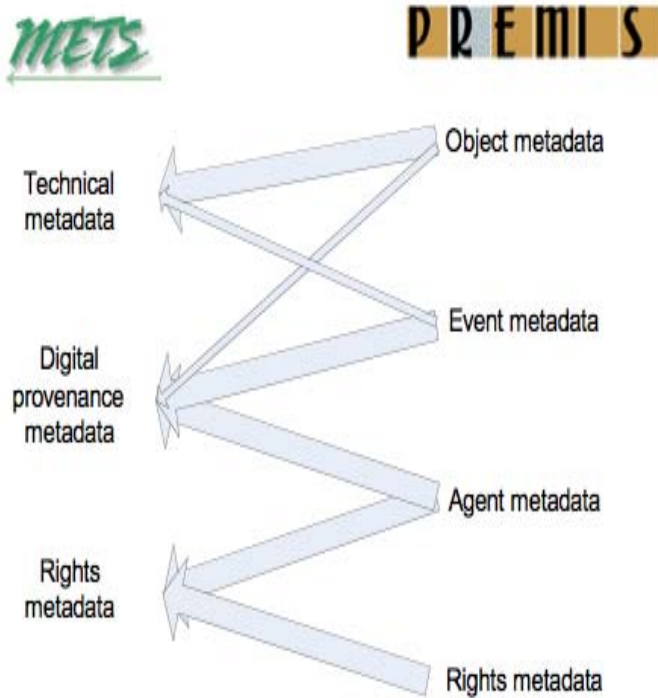
Figure 5. Mapping PREMIS Entities to METS Metadata Sections (Guenther, 2008).

Quite a number of cultural heritage institutions use a combination of metadata schemes to address their diverse metadata needs. Many respondents (about 43%) indicated that their institutions use the Metadata Encoding and Transmission Standard (METS) to implement metadata in digital library applications.

Many implementers of PREMIS have considered METS as a container to include PREMIS metadata along with other descriptive information about and links to the digital objects. In Figure 5, Guenther's mapping of PREMIS entities shows the relation to METS sections. Thick arrows show applicable subsections in METS for the named PREMIS entities; the thin arrows show links from one PREMIS entity to another METS subsection.

SUMMARY OF THE STUDY

PREMIS is a relatively new innovation and research on this topic is valuable, not only for current adopters, but also for potential adopters who are considering the opportunities and benefits PREMIS may allow them in tackling the digital projects issues.

The study yielded both qualitative and quantitative data, and analysis showed that all six factors influence the adoption of PREMIS in varying degrees. Results of a regression analysis of adoption level on the six factors

showed a statistically significant relationship. The R square value for the model was .528, which means that 52.8% of the variance in PREMIS adoption was explained by a combination of the six factors.

In many diffusion of innovations studies, perceived relative advantage, complexity, and compatibility seem positively related to adoption. Although trialability and observability are among the less-commonly used innovation attributes in some studies (Agarwal & Prasad, 1997) the relationship between perceived trialability and PREMIS adoption has been found positive. In support of this finding, Tornatzky and Klein (1982) noted that the perception of attributes of the innovation can predict the adoption, with some degree of consistency across various settings. The greater the perceived characteristics of an innovation, the more rapid its rate of adoption will be.

Figure 6 provides summary of attributes that influence PREMIS adoption in cultural heritage institutions. Based on the standardized coefficients values, among the six variables, institutional readiness, trialability, and relative advantage were the three best predictors of PREMIS adoption. As individual predictors, the trialability construct was the strongest predictor of PREMIS adoption, explaining 18.7% of the variance in predicting the PREMIS adoption. Institutional readiness was a close second, explaining 17.4% of the variance, and the relative advantage attribute was a distant third, explaining 8.5% of the variance in PREMIS adoption. The line strength in Figure 6 shows the level of influence.
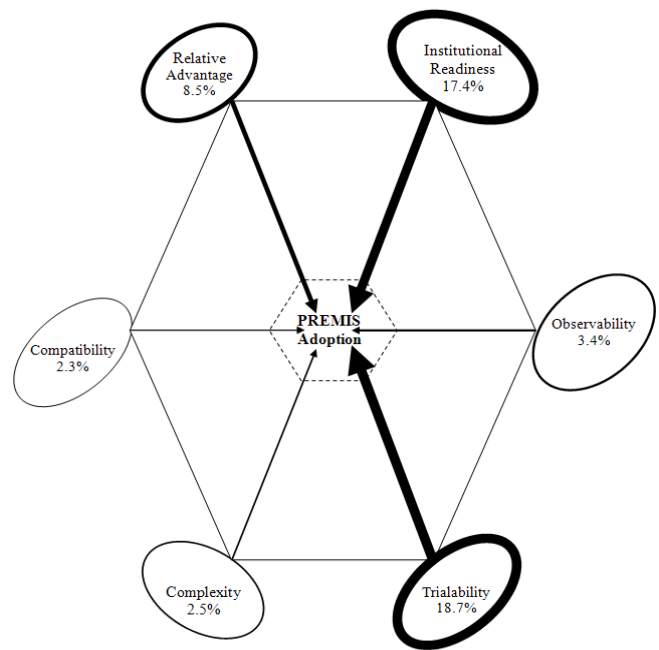

Figure 6. Summary of Attributes that Influence PREMIS Adoption

This research has demonstrated that the identification of factors influencing the adoption decisions of cultural heritage institutions is also a topic that is worth pursuing. Cultural heritage institutions have stages of adoption that map to the factors or variables seen in other innovation adoption studies. It is interesting to note the institutional differences both in number of participations, PREMIS adoption stages, and/or decision categories. Higher education institutions, national libraries, and archives tend to show relatively high level of participation and engagement in terms of PREMIS adoption compared to museums and other type of cultural heritage institutions.

Figure 7 depicts the PREMIS adoption stage category assignments as reported by the respondents, which tend toward a normal distribution. There is a strong resemblance with the innovation of adoption curve of Rogers that classifies adopters into five categories (innovators, early adopters, early majority, late majority, and laggards).
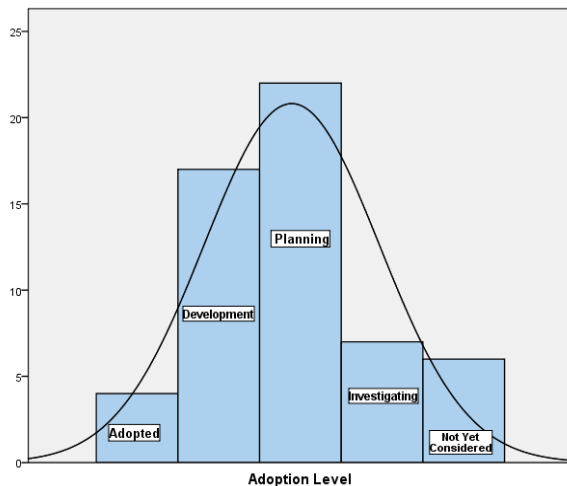

Figure 7. PREMIS Adoption Stages in Cultural Heritage Institutions

Innovators and early adopters (institutions fully adopted and in development stages) had the vision to adopt an emerging standard because of preservation needs and possible prospects. These are institutions that have prior experience with other related metadata standards or that have vested interests in digital preservation and promoting PREMIS as a standard or as viable solution to digital preservation challenges.

The early majority (planning stage) preferred to stay away from cutting edge technologies to avoid the risk associated with emerging standards. But if benefits are demonstrated by the early adopters, cultural heritage institutions that are in planning stages will be quick to adopt PREMIS. This hypergrowth phase (as depicted by the increased slope in Figure 7) has the highest adoption rate. The more institutions that adopt a standard, the faster it will be adopted by the general population due to the network externality effect.

The late majority and laggarts represent those potential adopters who dislike discontinuous Innovations and do not engage with new innovations. Cultural heritage institutions in the "Investigative" and "Not yet Considered" stages adopt PREMIS standards mainly because their major funders or collaborators require them to use the same standards. Most museums and small libraries that do not have adequate resources often belong to this category. Usually, these groups adopt when the standard has entered the stage of maturity.

CONCLUSION

The findings of this study indicate that PREMIS has not reached its full potential as a digital preservation metadata standard across diverse cultural heritage institutions. However, studying adoption of innovations requires a longitudinal study to understand an inherently complex set of issues that affect adoption.

Considering everything (the early stage of PREMIS, relatively slow adoption/participation rates, etc.), generalizations of the findings from an institutional perspective are difficult to make. For example, the data collection for this study was conducted before the release of the new version of the PREMIS. Many respondents stated that because PREMIS is still changing and because they see little success story, they wouldn't be adopting PREMIS in the near future. Although the data shows a reluctance to adopt a preservation metadata that is in revision, it is the nature of developing standards to continually revise. In other words, there is no good timing.

This research on preservation metadata adoption just barely begins to show the many layers of this complex problem. Considering the diverse needs of cultural heritage institutions and the complexity of digital preservation issues, much remains to be illuminated. Further developments in digital technologies are likely to produce new digital preservation challenges as well as opportunities.

This study has important implications for future research on preservation metadata and for researchers and stakeholders engaged in digital preservation and metadata standards development efforts. The theoretical framework selected for this research (DOI) has been identified as a suitable framework for further research in this topic area. By combining DOI with a number of other existing theories, further research in this domain will increase understanding and also clarify and expand several approaches to investigating adoption of preservation metadata in general. As some of the factors may have more weight in different institutions, future researchers may need to contact more institutions and break data down further to review the impact of the identified factors at different cultural heritage institutions.

There are also many new questions from the findings and conclusion regarding the various issues of institutional decision-making for the adoption of preservation metadata. In this regard, additional studies into the relationships among the institutional readiness characteristics would provide deeper understanding of institutional factors that explain PREMIS adoption in particular. By doing so, PREMIS can be better poised to deal with the current and future digital preservation needs of the wider cultural heritage community.

REFERENCES

Agarwal, R., & Prasad, J. (1997). The role of innovation characteristics and perceived voluntariness in the acceptance of information technologies. *Decision Sciences*, *28*(3), 557-582.

Besser, H. (2002). The Next Stage: Moving from Isolated Digital Collections to Interoperable Digital Libraries. *FirstMonday, 7*(6). Retrieved December 21, 2008, from http://www.uic.edu/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/958/879

Bradford, M., & Florin, J. (2003). Examining the role of innovation diffusion factors on the implementation success of enterprise resource planning systems. *International Journal of Accounting Information Systems, 4*(3), 205-225.

Buonanno, G. et al. (2005). Factors affecting ERP system adoption: A comparative analysis between SMEs and large companies. *Journal of Enterprise Information Management*, *18*(4), 384 – 426.

Chen, M. (2003). Factors Affecting the Adoption and Diffusion of XML and Web Services Standards for E-business Systems. *International Journal of Human Computer Studies, 58*(3), 259-279.

Day, M. (2006). The Long-term Preservation of Web Content. In Julien Masanes (Ed.), *Web Archiving* (pp.. 177-199). Berlin Heidelberg: Springer-Verlag.

Fichman, R. G., & Kemerer, C. F. (1997). The assimilation of software process innovations: an organizational learning perspective. *Management Science,* 43(10), 1345-1363.

Gallivan, M. J. (2001). Organizational Adoption and Assimilation of Complex Technological Innovations: Development and Application of a New Framework. *The DATABASE for Advances in Information Systems, 32*(3): 51-85.

Hedstrom, M. (Ed.). (2003) *It's about Time: Research Challenges in Digital Archiving and Long-term Preservation*: Final Report on the NSF Workshop on Research Challenges in Digital Archiving and Long-Term Preservation, April 12-13, 2002, Washington, DC. Retrieved December 21, 2008, from http://www.digitalpreservation.gov/library/pdf/NSF.pdf

Lavoie, B. (2008). PREMIS With a Fresh Coat of Paint: Highlights from the Revision of the PREMIS Data Dictionary for Preservation Metadata. *D-Lib Magazine*, *14*(5/6). Retrieved December 21, 2008, from http://www.dlib.org/dlib/may08/lavoie/05lavoie.html.

Lavoie, B., & Gartner, R. (2005) *Technology Watch Report: Preservation Metadata..* DPC Technology Watch Series Report 05-01. Retrieved December 21, 2008, from http://www.dpconline.org/docs/reports/dpctw05-01.pdf

Rogers, E. M. (2003). *Diffusion of Innovations* (5th ed.). New York: The Free Press.

Themistocleous, M., & Irani, Z. (2002). A model for adopting enterprise application integration technology. *The Adoption and Diffusion of IT in an Environment of Critical Change (IFIP WG8.6)*, 61-75.

Tornatzky, L. G., & Klein, K. J. (1982). Innovation characteristics and innovation adoption-implementation: A meta-analysis of findings. *IEEE Transactions on Engineering Management, 29*(1), 28-45.