An Investigation of Cyberinfrastructure Adoption in University Libraries

Jiangping Chen[1], Duha Alsmadi[1] , Xiaoyu Zhang[1]

Zhiwu Xie[2]

[1] University of North Texas, Denton Texas 76203, USA
[2] Virginia Tech, Blacksburg, VA 24061, USA
Jiangping.Chen@unt.edu

Abstract. This study aims to understand factors that affect university libraries' adoption of cyberinfrastructure for big data sharing and reuse. A cyberinfrastructure adoption model which contains 10 factors has been developed based on the technology-organization-environment (TOE) framework and the literature regarding tradeoffs of applying cyberinfrastructure. This paper describes the proposed cyberinfrastructure adoption model and explains the survey instruments. The next steps of the study are also presented.

Keywords: Library Cyberinfrastructure, technology adoption Technology-organization-environment (TOE) framework

1 Introduction

As a key component of the nation's knowledge infrastructure, libraries must continuously reinvent themselves with the emergence and the establishment of new discovery paradigms. Recent data science advancement has motivated many high-profile library big data services, notably the ambitious plan to archive all tweets at the Library of Congress [1], the heterogeneous and geographically replicated archival storage known as the Digital Preservation Network (DPN) [2], and the metadata hubs developed at the Digital Public Library of America (DPLA) [3]. Many more such big data sharing and use projects are being developed or being planned. Since building large, single-tenant data centers at each library would be prohibitively expensive, the use of widely available, shared cyberinfrastructure (CI) resources is a promising option for libraries.

CI refers to "a collection of hardware- and software-based services for simulation/modeling, knowledge and data management, observation and interaction with the physical world, visualization and interaction with humans, and distributed collaboration." [4] The shared CI may be categorized into 4 types, each with its unique strengths, weaknesses, and challenges: 1) Institutional high-performance computing (HPC), high-throughput computing (HTC) and storage facilities; 2) National HPC, HTC, and storage facilities, most notably XSEDE resources; 3) National research

clouds such as Chameleon Cloud, CloudLab, Open Science Data Cloud, etc.; and 4) Commercial clouds, such as Amazon Web Services (AWS), Rackspace, etc.

Although shared CI has been used in all stages of the research data lifecycle, this research focuses on the data sharing and reuse aspects. With different CI choices, libraries may have difficulties to find the appropriate ones. Also, it is unclear what the factors that university libraries have considered when they adopted CI are and how they should prepare for adopting CI.

This study aims to understand factors that affect university libraries when they make decisions on adopting CI for their big data sharing and reuse services. As more and more university libraries are taking research and scientific data storage and management responsibilities for their universities, this study would provide valuable insights and guidelines for university libraries to select, prepare, and apply for appropriate CI services.

2 Related Studies

Very few studies have investigated factors that affect libraries' adoption of CI. Henry [5] examined the core infrastructure elements of large, non-commercial digital libraries. The study focused on understanding different technical approaches to managing large digital libraries. Scalability was considered critical to support the long-term growth of these systems. Lyon [6] explored how libraries can re-shape to better reflect the requirements and challenges of today's data-centric research landscape. She analyzed support services of 10 libraries for research data management and pointed out that actions should be taken to meet the capacity and capability shortfall of librarians and information scientists equipped with big data management expertise.

Kim and Crowston [7] reviewed previous studies regarding cyberinfrastructure adoption and use by scientists and other special user groups. They categorized previous studies into two groups: Adoption and post-adoption (continued use) research. Initial technology adoption was found to be determined by cognitive processes, including compatibility, observability, and trainability. Two most important factors at the post-adoption stage were affective reaction and habit. (p. 9).

Oliveria and Martins [8] reviewed theories for adoption models at the firm level used in information systems literature and discussed two prominent models: the Diffusion on innovation (DOI) theory and the technology-organization-environment (TOE) framework. They concluded that the TOE framework was more complete with a solid theoretical basis, consistent empirical support, and the potential of application to information technology adoption. TOE was also applied in many different contexts, including examining small business electronic data interchange (EDI) adoption [9], evaluating and identifying factors that affecting the decision to migrate the cloud [10], determinants of Radio Frequency Identification (RFID) adoption in hospitals [11] and manufacturing industry [12] in Taiwan, and the adoption of Enterprise resource planning (ERP) in Taiwan's communication industry [13] and firms in China [14].

Additionally, Li [15] used logistic regression to examine the relationships between the technological, organizational and environmental characteristics and the enterpris-

es' adoption of e-procurement in Chinese manufacturing enterprises. (p.32) Li concluded that relative advantage, top management support, external pressure and external support are determinate factors of e-procurement adoption. (p.37). Lian, Yen, and Wang [11] indicated that the 5 top critical factors affecting the decision to adopt cloud computing in Taiwan's hospitals are "data security, perceived technical competence, cost, top management support, and complexity" (p. 28).

3 The Proposed CI Adoption Model and the Research Question

The literature indicated that multiple factors that may influence libraries' big data cyberinfrastructure (CI) adoption. Applying TOE framework as the theoretic lens, we classified these factors into 3 categories: Technological, Organizational, and Environmental. Fig.1 illustrates our CI adoption model.

Ten factors that may affect CI adoption in university libraries will be examined in this study. Accordingly, the research question is: ***What are the factors that affect university libraries' cyberinfrastructure adoption for big data sharing and reuse?*** The hypotheses to be tested in the context of university libraries (UL) include:

H1: Technology factors will significantly affect UL's CI adoption
H2: Organization factors will significantly affect UL's CI adoption, and
H3: Environmental factors will significantly affect UL's CI adoption

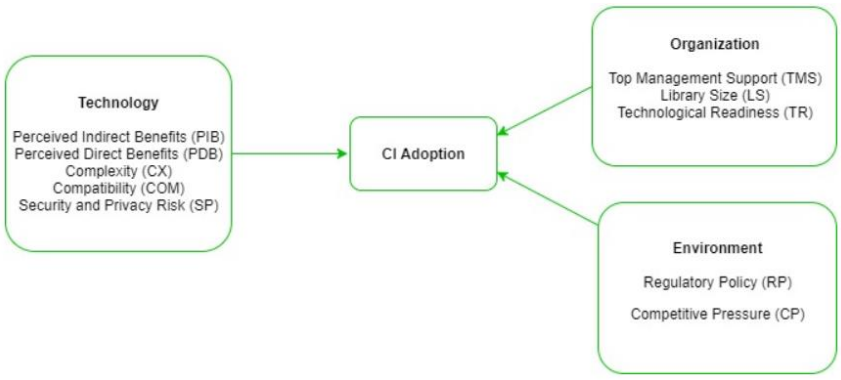

Fig.1. The CI Adoption Model

4 The Survey Constructs

Table 1 lists 10 constructs to be measured in the study, including a number of indicators, a sample indicator, and the supporting literature for each construct.

Table 1. The major survey constructs with indicators

Construct	(Number of Indicators)	A Sample Indicator	Supporting literature
Perceived direct benefits	4	I think that using CI in the library improves data accuracy	[8], [9]
Perceived indirect benefits	4	I think that using CI in the library improves organizational image	[8], [9]
Complexity	5	The skills required to adopt the CI are not complex to our employees	[11], [14]
Compatibility	4	CI is compatible with current library practices	[11], [14]
Security and privacy risk	4	I do not think it is safe to use CI because of the privacy and security concerns	[10], [11]
Top management support	6	Top management supports the adoption of CI	[11], [14]
Technological Readiness	4	We have the technical knowledge and skills to implement CI	[11]
Library Size	4	The number of employees in the library	[10], [12], [16]
Regulatory policy	4	To what extent you expect that legal implication are negatively related to the decision to adopt CI	[10], [13]
Competitive pressure	4	I think that the library faces high level of rivalry among other academic libraries	[11], [14]

5 Next Steps

We have developed the survey instrument containing 45 questions based on the proposed CI adoption model and the constructs. The survey has been approved by university IRB in October 2017. We will start to collect data from IT staff at top 200 U.S. university libraries. The collected data will be first tested on reliability and validity, and then analyzed by applying appropriate statistical approaches such as linear regression or structural equation modeling to test our hypotheses.

References

1. Zimmer, M.: The Twitter archive at the Library of Congress: challenges for information practice and information policy. First Monday 20(7), (2015).
2. Korner, B.: DPN: an overview from a technical perspective. In: 2013 Preservation and Archiving Special Interest Group (PASIG) Conference. Washington, DC (2013).
3. Altman, A., Gueguen, G., Breedlove, M.: Heiðrún: DPLA's metadata harvesting, mapping and enhancement system. In: 2015 Code4Lib Conference. Portland, OR (2015).
4. Atkins, D.: Cyberinfrastructure for research. Issues in Science and Technology, 22(2), 9 (2006).

5. Henry, G.: Core infrastructure considerations for large digital libraries. Council on Library and Information Resources 4(19), (2012).
6. Lyon, L.: The informatics transform: re-engineering libraries for the data decade. International Journal of Digital Curation 7(1), 126–138 (2012).
7. Kim, Y., Crowston, K.: Technology adoption and use theory review for studying scientists' continued use of cyber-infrastructure. Proceedings of the American Society for Information Science and Technology 48(1), 1-10 (2011).
8. Oliveira, T., Martins, M. F.: In formation technology adoption models at firm level: Review of literature. In Proceedings of the European Conference on Information Management & Evaluation. 9(1), 312-323, Lisbon, Portugal (2010).
9. Kuan, K.K.Y., Chau, P.Y.K.: A perception-based model for EDI adoption in small businesses using a technology-organization-environment framework. Information and Management 38(8), 507–521 (2001).
10. Alkhalil, A., Sahandi, R., John, D.: An exploration of the determinants for decision to migrate existing resources to cloud computing using an integrated TOE-DOI model. Journal of Cloud Computing, 6(1), 2-22 (2017).
11. Lian, J. W., Yen, D. C., Wang, Y. T.: An exploratory study to understand the critical factors affecting the decision to adopt cloud computing in Taiwan hospital. International Journal of Information Management. 34(1), 28–36, (2014).
12. Wang, Y. M., Wang, Y. S., and Yang, Y. F. Understanding the determinants of RFID adoption in the manufacturing industry. Technological Forecasting and Social Change. 77(5), 803–815, (2010).
13. Pan, M.-J., Jang, W.-Y.: Determinants of the adoption of enterprise resource planning within the technology-organization-environment framework: Taiwan's communications industry. The Journal of Computer Information Systems 48(3), 94–102 (2008).
14. Xu, W., Ou, P., & Fan, W.: Antecedents of ERP assimilation and its impact on ERP value: A TOE-based model and empirical test. Information Systems Frontiers. 19(1), 13–30, (2017).
15. Li, Y.H.: An empirical investigation on the determinants of E-procurement adoption in Chinese manufacturing enterprises. In: 2008 International Conference on Management Science and Engineering 15th Annual Conference Proceedings, pp. 32–37. IEEE, Long Beach, USA (2008).
16. McKerlich, R., Ives, C., & McGreal, R.: Measuring use and creation of open educational resources in higher education. International Review of Research in Open and Distance Learning. 14(4), 90–103, (2013).