Delineating Legacy System IT Artifacts – Applying a Systems Thinking-based Framework

Julia Stachofsky
stac0848@vandals.uidaho.edu
College of Business and Economics
University of Idaho
Moscow, Idaho 83843, USA

Abstract

As technology rapidly changes, information systems that at one time were cutting-edge have become obsolete. Many businesses are still running critical operations on these older legacy information systems. This paper applies a systems thinking-based framework to legacy information systems to identify the characteristics of legacy systems information technology artifacts. The analysis was based on data from a survey answered by information technology professionals who work with legacy systems in their organizations. The survey data indicated that regardless of artifact type, legacy information systems exhibited characteristics of integration within the organization, complexity, and communicate somewhat synchronously. The results also indicate that the framework used is valid and useful for information systems scholars to better engage with information technology artifacts in the research literature.

Keywords: Legacy Information Systems, IT Artifact, Systems Thinking, Information Systems Theory

1. INTRODUCTION

With constant advancement and change in IT, information systems that were once cutting-edge can later become obsolete. These legacy information systems are often still critical for business operations and cannot be easily modernized (Bisbal, Lawless, Wu, and Grimson, 1999). Legacy systems in an organization can be difficult and costly to maintain and are not easily replaced. Much of the existing research on legacy information systems has focused on coping strategies for legacy information systems. Little work has been done to understand the legacy systems themselves as IT artifacts (ITA). Understanding legacy systems as ITAs can help with finding theoretical similarities between different research papers on legacy systems. This could be useful for fundamentally understanding the legacy systems within an organization to inform decisions on how they should be dealt with.

In this paper the framework developed by Matook and Brown (2016) is used as a method for describing legacy information systems as an ITA. This framework uses systems thinking methodology as a basis for delineating ITAs in information systems research. The data analyzed for this paper is results from a survey answered by IT professionals about legacy information systems within their organizations. IT staff are not the only people to work with legacy systems in an organization, but they often have the most technical understanding and expertise for the information systems in their organization.

The next section provides a brief overview of the literature as it relates to legacy information systems, the ITA, and Matook and Brown's framework. Following the review, the paper discusses the methodology behind the conducted survey and how it utilizes Matook and Brown's framework. Next the results of the survey are presented before being analyzed in the discussion section. The paper concludes with a discussion of the limitations and potential research avenues related to legacy information systems in the context of the IT artifact.

2. LITERATURE REVIEW

Legacy Systems

One of the early definitions comes from Bennett (1995) "Large software systems that we don't know how to cope with but that are vital to our organization." Bennett also describes the systems as written using "outdated techniques" in languages such as assembly or Cobol. Bisbal et al. (1999) further expand on this by noting legacy information systems usually run on obsolete hardware, are difficult to extend, lack clean integration interfaces, and are costly to maintain. Both Bennett and Bisbal et al. capture this concept of legacy systems being outdated in the sense that they run on obsolete hardware and were designed with now obsolete methodologies. Ganesan and Chithralekha (2016) provide a definition that encompasses this in their survey on legacy system migration, "Legacy Systems are mission critical systems with monolithic code architecture having restrictions to archaic hardware, software and are short of resource in terms of skill sets, documentation and is therefore hard to maintain and are inflexible…"

While there exists some variation in how legacy systems are defined, Gholami, Daneshgar, Beydoun, and Fethi (2017) note an important common theme, "… the worthiness of the legacy systems and this has been the reason to keep them working in organisations. Legacy systems support business processes, maintain organizational knowledge, and provide significant competitive advantage…". The legacy systems persist in part because they provide value to the organization.

Much of the literature is focused on coping strategies. Bisbal et al. (1999) identify four coping strategies, wrapping, maintenance, migration, and redevelopment. Each of these strategies have their own lines of research but are outside of the scope of this discussion.

Information Technology Artifact (ITA)

Since the early years of the information systems discipline, there has been a stream of research and discussion focused on developing the field's identity (Benbasat & Zmud, 2003). The interdisciplinary nature of the field makes this a complex task. One particular line of research in this area is theorizing the technical artifact.

Orlikowski and Iacono (2001) kickstarted much of the current conversation about ITAs with their concern over little engagement with ITAs in the information systems literature. They found that ITAs are "… either absent, black-boxed, abstracted from social life, or reduced to surrogate measures."

There is ongoing debate for how to actually define an ITA. Orlikowski and Iacono (2001) define it as "…bundles of material and cultural properties packaged in some socially recognizable form such as hardware and/or software." While there are differences among definitions, conceptualizing IT as combined hardware and software as a basis for the ITA recurs throughout the literature (Matook & Brown, 2016).

Matook and Brown (2016) use Goldkuhl's (2013) definition of the ITA as the basis for their framework. Goldkuhl's (2013) definition conceptualizes the ITA as a physical artifact and an integrated whole of software and hardware. This is best exemplified by "without hardware, the software is just symbolic expressions. But together they are machines with the power to execute intentionally designed information-processing tasks."

Matook and Brown's Systems Thinking-based Framework for Delineating and Theorizing the ITA

Matook and Brown (2016) designed their framework to provide a way for scholars to meet Orlikowiski And Iacono's call to engage with the ITA in information systems research. The theoretical basis for this framework is systems thinking (which itself is derived from general systems theory). The concept of systems is deeply tied to the IS field. In one of the earliest foundational works of the field *Theoretical Analysis of Information Systems* the third sentence is "To emphasize the system aspect is to stress that it is the combined effect of the components that is important.". In IS research scholars study many different systems, so it makes sense to pull from systems thinking when building a framework for describing the core artifact of IS research.

Systems thinking can be understood through five key concepts

1. System Parts, Wholeness, and System Structure
2. System Boundary and Environment
3. Hierarchical Order, Wholeness, and Complexity
4. System state and change of state
5. Transformation and feedback

Matook and Brown create seven characteristics for describing ITAs based on those key concepts

1. Integration (1 and 2)
2. Connectivity (1 and 2)

3. Complexity (3)
4. State (4)
5. Adaptation (4 and 5)
6. Self-Adaptation (4 and 5)
7. Synchronicity (5)

Through using these characteristics as the framework of the ITA, Matook and Brown believe information systems scholars can work towards building a cumulative tradition of theoretical similarities and differences of ITAs. This framework provides a common language for scholars to compare results across studies and technologies.

3. RESEARCH MODEL

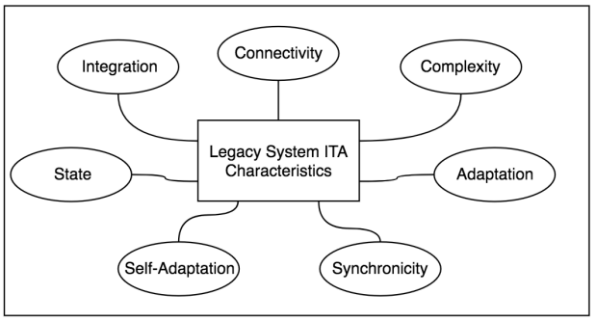

Figure 1: ITA Characteristics

Figure 1 shows a model of the legacy system ITA's characteristics. It should be noted the lines are not causal links. The lines are strictly descriptive characteristics of the ITA. Each survey response can be considered a unique ITA.

Each characteristic is measured on a scale described as follows:
Integration: Highly integrated to highly fragmented.
Connectivity: Highly connected to highly isolated.
Complexity: Highly complex to less complex.
State: Highly stateful to highly stateless
Adaptation: Highly dynamic to highly static
Self-adaptation: Highly adaptive to highly non-adaptive
Synchronicity: Highly synchronous to highly asynchronous.

4. RESEARCH METHOD

The survey for this research was conducted using Amazon's Mechanical Turk platform. This service can be used by researchers to administer surveys to workers on Mechanical Turk in trade for compensation. Additionally, Amazon offers specialized pools of workers for an additional fee.

Table 1 lists the survey questions with what characteristic it is measuring. All questions were scaled from 1 to 5 with 3 being a neutral option.

Characteristic	Survey Question
Integration	To what extent is the legacy system integrated with other systems within the organization?
Connectivity	To what extent is the legacy system connected with external systems outside of the organization?
Complexity	To what extent does the legacy system contain interdependent modules?
State	To what extent does the legacy system retain state information between sessions (such as returning customer information in an online shopping system)?
Adaptation	To what extent does the legacy system allow for system customization to meet organizational needs?
Self-Adaptation	To what extent does the legacy system change itself based on feedback from other systems or users?
Synchronicity	In what way does the legacy system communicate? (Synchronous communication is when both the sender and the receiver must be available at the same time)

Table 1: Characteristic Questions

The last question not in Table 1 asked the respondent to identify the type of system they answered the survey for (Enterprise Resource Planning, Accounting, etc.).

The full survey with answer choices can be found in Appendix A.

97 surveys responses were collected, with 75 deemed to be valid responses. 10 appeared to be automated, non-sensical responses, four surveys did not clearly state the legacy system type, and eight surveys listed multiple legacy systems instead of answering the survey about one legacy system. The survey was only available to Mechanical Turk workers whose job function is Information Technology. This criterion was selected as the focus of this research is to understand the ITA aspect of legacy systems. Individuals that work in IT have the most experience working with these systems from a technical perspective. Additional demographic information about the IT workers was not collected.

5. RESULTS

ITA Categorization

Based on an analysis of the responses to question eight of the survey, 16 distinct ITA groupings were identified.

ITA Group	Count
Accounting	24
ERP	11
Online Shopping	8
CRM	6
Medical	6
Inventory	4
Education	3
Scheduling	3
Mainframe	2
Storage	2
Documentation	1
Human Resources	1
Manufacturing	1
Networking	1
Planning	1
Security	1

Table 2: ITA Groups

Table 2 shows how many legacy systems were in each group. The most common legacy system type by far was accounting systems, with ERP systems and online shopping systems being the second and third most common respectively. The overall characteristics, as well as the top two legacy system types will be the focus of the rest of the results and analysis.

ITA Characteristics

Characteristic	Overall	Acct	ERP
Integration	3.73	3.63	3.82
Connectivity	3.16	2.92	3.55
Complexity	3.65	3.54	4.00
State	3.35	3.29	3.18
Adaptation	3.09	3.08	3.09
Self-adaptation	2.87	3.00	3.00
Synchronicity	3.53	3.63	3.72

Table 3: ITA Characteristics Mean Measurements

Table 3 shows the mean value for each ITA characteristic for all 75 ITAs listed in Table 2 as well as the mean values for accounting, and ERP system ITAs. Additional statistics can be found in Appendix B.

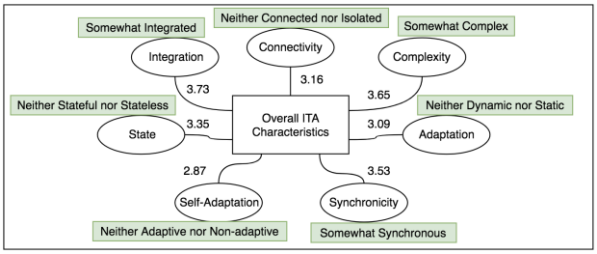

Figure 2: Overall ITA Characteristics

Figure 2 shows that on average, all of the legacy system ITAs studied were somewhat integrated, somewhat complex, and somewhat synchronous. Integration was the highest measured characteristic with a mean of 3.73, followed by a mean of 3.65 for complexity, and a mean of 3.53 for synchronicity.

The other characteristics, state, connectivity, adaptation, and self-adaptation each had a mean value closest to 3, the neutral value for a characteristic. State had a mean value of 3.35, followed by a mean of 3.16 for connectivity, a mean of 3.09 for adaptation, and a mean of 2.87 for self-adaptation. While these characteristics may have been non-neutral for individual ITAs, on average they were measured to be neutral.

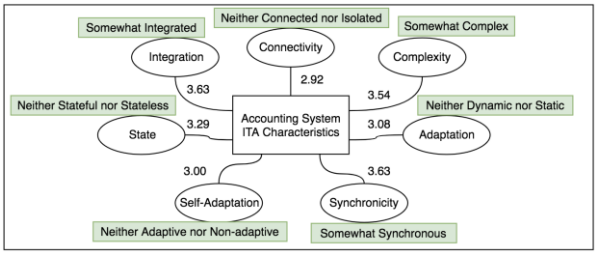

Figure 3: Accounting System ITA Characteristics

Figure 3 shows that on average, the accounting system ITAs studied were somewhat integrated, somewhat complex, and somewhat synchronous. Integration and synchronicity were the highest measured characteristics with means of 3.63, followed by a mean of 3.54 for complexity. These are the same three characteristics that were identified for the overall measurements discussed in Figure 2.

The other characteristics, state, adaptation, self-adaptation, and connectivity each had a mean value closest to 3, the neutral value for a characteristic. State had a mean value of 3.29, followed by a mean of 3.08 for adaptation, a mean of 3.00 for self-adaptation, and a mean of 2.92 for connectivity.

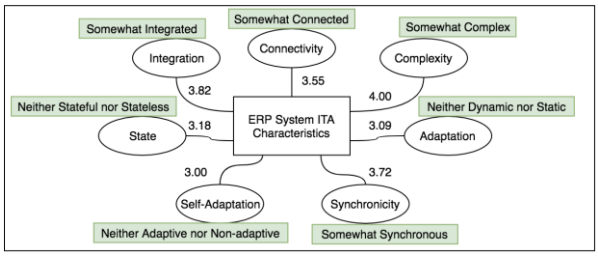

Figure 4: ERP System ITA Characteristics

Figure 4 shows that on average, the ERP system ITAs studied were somewhat integrated, somewhat connected, somewhat complex, and somewhat synchronous. Complexity was the highest measured characteristic with a mean of 4.00, followed by a mean of 3.82 for integration, a mean of 3.72 for synchronicity, and a mean of 3.55 for complexity. Integration, complexity, and synchronicity are characteristics identified in the overall and accounting system ITAs. Connectivity is an additional characteristic only identified in the ERP system ITA of the three groupings.

The other characteristics, state, adaptation, and self-adaptation each had a mean value closest to 3, the neutral value for a characteristic. State had a mean value of 3.18, followed by a mean of 3.09 for adaptation, and a mean of 3.00 for self-adaptation.

6. DISCUSSION

Overall ITA Characteristics

Four of the seven characteristics were measured to be neutral. It is likely that this can be attributed to the wide variety of ITAs that a legacy system can be. While a legacy ERP system might share some aspects in common with a legacy documentation system, the core functionality and uses of those systems leads to the ITA consisting of different characteristics.

The highest measured characteristic was integration. This high level of integration may partially explain why these legacy systems have persisted in an organization. If the system is deeply integrated with multiple parts of the organization, it makes it more difficult to upgrade or replace that system.

The second highest measured characteristic was complexity. The high level of complexity of the legacy system ITAs may also be a factor in why the legacy systems persist in the organization. If a system has many interdependent modules, changes to the system become more difficult without breaking other functionalities of the system.

The third highest, and last non-neutral measured characteristic was synchronicity. On average, the ITAs communicated somewhat synchronously. This gives the system less independence to communicate on its own terms as both the sender and receiving system must be available simultaneously to receive data (Matook and Brown, 2016).

Accounting System ITA Characteristics

The accounting systems studied exhibited the same characteristics as the overall trends seen for all of the ITAs. Given the importance of financial data to an organization, it makes sense that the accounting systems would be somewhat integrated with other systems within the organization. Accounting system complexity may be due to additional installed modules for different accounting purposes.

ERP System ITA Characteristics

The ERP systems studied exhibited the same characteristics as the overall trends, but also included the characteristic of being somewhat connected. This could be due to external systems tying into the data within an ERP system. Complexity and integration were the two highest measured averages, which is unsurprising given that ERP systems often include many different modules and are tied into numerous departments and processes across the organization.

Limitations

Given the limited amount of survey responses (especially when discussing the sub-groupings) one must be careful in inferring too much about the characteristics that make up legacy system ITAs. Additionally, it can be difficult to fully separate out what is unique to the ITA in a legacy

system context and what characteristics apply to an ITA in a general context. The survey results also only give a very surface level view of the ITA characteristics without specific contexts of how the systems are used within their respective organizations. The lack of demographic information about the survey respondents further masks possible affects these demographics may have had on the data.

The solution to these limitations is for more research to engage the ITA. Not just through isolated looks at the ITA itself, but by explicitly discussing it as a part of mainstream IS research (Matook and Brown, 2016). Matook and Brown's framework is not the only way to describe ITAs, but it offers a solid framework for IS scholars to start incorporating discussion of the ITA into their research.

7. CONCLUSION

Overall this paper contributes to the IS literature in two ways. Through the analysis of survey responses from IT professionals, it was found that legacy system ITAs exhibit characteristics of integration within an organization, complexity, and synchronous communication across the various types of ITAs. In addition to this finding, this paper also serves as an example of the validity of Matook and Brown's (2016) framework for delineating the ITA in IS research. As more IS scholars adopt this framework further progress can be made on studying the ITA, giving more theoretical legitimacy to the IS field.

8. REFERENCES

Benbasat, I., & Zmud, R. (2003, June). "The Identity Crisis within the Is Discipline: Defining and Communicating the Discipline's Core Properties." *MIS Quarterly*, vol. 27, no. 2, June 2003, p. 183., doi:10.2307/30036527.

Bennett, K. (1995). "Legacy Systems: Coping with Success." *IEEE Software*, vol. 12, no. 1, pp. 19–23., doi:10.1109/52.363157.

Bisbal, J., Lawless, D., Bing, W., & Grimson, J. (1999). "Legacy Information Systems: Issues and Directions." *IEEE Software*, vol. 16, no. 5, pp. 103–111., doi:10.1109/52.795108.

Ganesan, S., & Chithralekha T. (2016). "A Survey on Survey of Migration of Legacy Systems." *Proceedings of the International Conference on Informatics and Analytics - ICIA-16*, doi:10.1145/2980258.2980409.

Gholami, M., Daneshgar F., Beydoun G., & Rabhi F. (2017). "Challenges in Migrating Legacy Software Systems to the Cloud — an Empirical Study." *Information Systems*, vol. 67, pp. 100–113., doi:10.1016/j.is.2017.03.008.

Goldkuhl, G. (2013). The IT Artefact: An Ensemble on the Social and the Technical? – A Rejoinder. *Systems, Signs, & Actions,* vol. 7, no. 1, pp. 90-99.

Langefors, B. (1973). Theoretical analysis of information systems (4th ed.). Lund: Studentlitteratur.

Matook, S., & Brown S. (2016). "Characteristics of IT Artifacts: A Systems Thinking-Based Framework for Delineating and Theorizing IT Artifacts." *Information Systems Journal*, vol. 27, no. 3, pp. 309–346., doi:10.1111/isj.12108.

Orlikowski, W., & S. Iacono (2001, June 2). "Research Commentary: Desperately Seeking the 'IT'' in IT Research—A Call to Theorizing the IT Artifact." *Information Systems Research*, vol. 12, no. 2, pp. 121–134.

Recker, J. (2013). Scientific Research in Information Systems: A Beginner's Guide. Springer.

Appendix A: Full Survey

Choose **ONE** legacy information system used in your organization and answer the following questions about that information system.
Legacy information systems are computer systems in your organization that are critical for organizational operations, but are also obsolete and resist modernization.

Question	Response (5)	Response (4)	Response (3)	Response (2)	Response (1)
1. To what extent is the legacy system integrated with other systems within the organization?	Highly Integrated	Somewhat Integrated	Neither Integrated nor Fragmented	Somewhat Fragmented	Highly Fragmented
2. To what extent is the legacy system connected with external systems outside of the organization?	Highly Connected	Somewhat Connected	Neither Connected nor Isolated	Somewhat Isolated	Highly Isolated
3. To what extent does the legacy system contain interdependent modules?	Many Interdependent Modules	Somewhat Many Interdependent Modules	Neither Many nor Few Interdependent Modules	Somewhat Few Interdependent Modules	Few Interdependent Modules
4. To what extent does the legacy system retain state information between sessions (such as returning customer information in an online shopping system)?	Highly Stateful	Somewhat Stateful	Neither Stateful nor Stateless	Somewhat Stateless	Highly Stateless
5. To what extent does the legacy system allow for system customization to meet organizational needs?	Highly Dynamic	Somewhat Dynamic	Neither Dynamic nor Static	Somewhat Static	Highly Static
6. To what extent does the legacy system change itself based on feedback from other systems or users?	Highly Adaptive	Somewhat Adaptive	Neither Adaptive nor Non-adaptive	Somewhat Non-adaptive	Highly Non-adaptive
7. In what way does the legacy system communicate? (Synchronous communication is when both the sender and the receiver must be available at the same time)	Highly Synchronous	Somewhat Synchronous	Neither Synchronous nor Asynchronous	Somewhat Asynchronous	Highly Asynchronous

8. What type of system is the legacy system? (Examples: Enterprise Resource Planning [ERP] System, Accounting System, Medical System, Online Shopping System, etc.)

Appendix B: Survey Statistics

Overall ITA Statistics					
	Mean	**Median**	**Mode**	**Variance**	**Standard Deviation**
Integration	3.733333333	4	4	1.063063063	1.031049496
Connectivity	3.16	4	4	1.595675676	1.263200568
Complexity	3.653333333	4	4	1.364684685	1.168197194
State	3.346666667	4	4	1.445765766	1.202400002
Adaptation	3.093333333	3	4	1.599279279	1.264626142
Self-adaptation	2.866666667	3	4	1.765765766	1.328821194
Synchronicity	3.533333333	4	4	1.279279279	1.131052288

Accounting System ITA Statistics					
	Mean	**Median**	**Mode**	**Variance**	**Standard Deviation**
Integration	3.625	4	4	1.201086957	1.095941128
Connectivity	2.916666667	3	4	1.644927536	1.282547284
Complexity	3.541666667	4	4	1.563405797	1.250362266
State	3.291666667	4	4	1.346014493	1.160178647
Adaptation	3.083333333	2.5	2	1.731884058	1.31601066
Self-adaptation	3	3	4	2.086956522	1.444630237
Synchronicity	3.625	4	4	1.375	1.17260394

ERP System ITA Statistics					
	Mean	**Median**	**Mode**	**Variance**	**Standard Deviation**
Integration	3.81818182	4	4	0.763636364	0.873862898
Connectivity	3.54545455	4	4	0.672727273	0.820199532
Complexity	4	4	4	0.6	0.774596669
State	3.18181818	3	4	0.763636364	0.873862898
Adaptation	3.09090909	4	4	1.290909091	1.136181804
Self-adaptation	3	3	3	1	1
Synchronicity	3.72727273	4	4	1.018181818	1.009049958