Preprint – Article accepted to appear in *The Library Quarterly*

Digital Library Archeology: A Conceptual Framework for Understanding Library Use through Artifact-Based Evaluation

Scott Nicholson, Assistant Professor
Syracuse University School of Information Studies
scott@scottnicholson.com

Abstract

Archeologists have used material artifacts found in a physical space to gain an understanding about the people who occupied that space. Likewise, as users wander through a digital library, they leave behind data-based artifacts of their activity in the virtual space. Digital library archeologists can gather these artifacts and employ inductive techniques, such as bibliomining, to create generalizations. These generalizations are the basis for hypotheses, which are tested to gain understanding about library services and users. In this article, the development of traditional archeological methods is presented and used to create a conceptual framework for the artifact-based evaluation in digital libraries.

Introduction

For hundreds of years, scientists have worked to understand our history through the "recovery, systematic description, and study"[1,12] of artifacts left behind by a particular culture. This field of study, archeology, has several tasks, one of which is to reconstruct the "lifeways of the peoples responsible for the archeological remains" [2,6]. Archeologists may know little about the people themselves who occupied a physical space; instead, they examine artifacts left behind for patterns to understand the communities who lived there.

There are similarities between a physical place and the virtual space supported by the Internet. Communities of people form, grow, evolve, and dissipate in this virtual space. As users travel through this virtual space, they leave behind data-based artifacts. These shards of virtual pottery such as searches and browsing behavior, and burial mounds of dead discussion groups hold valuable information if the pieces can be collected, cleaned, organized, and examined. This information can give "Internet Archeologists" an idea of the communities and cultures that have existed in virtual spaces.

A framework for Internet Archeology was broadly introduced by Nicholson[3]. In order to focus this concept, the discussion of Internet Archeology is limited here to a single type of virtual place – a digital library. The Digital Library Federation defines digital libraries thusly: "Digital libraries are organizations that provide the resources, including the specialized staff, to select, structure, offer intellectual access to, interpret, distribute, preserve the integrity of,

and ensure the persistence over time of collections of digital works so that they are readily and economically available for use by a defined community or set of communities" [4]. Much like the individuals living in the past, many users of digital library services may be understood only through the artifacts left behind from “living” in the space.

What are these artifacts of digital library use? In this case, each “artifact” is one record on a server about a single access of a resource by a user. When a user visits a digital library, they will travel through several services. They may start with a library Web page, move to a proxy / login server, an OPAC, a database, a digital resource, and then go out to a vendor’s site. At each server, the user leaves behind a small data-based artifact of their step through that virtual space. By collecting individual artifacts from different servers and connecting them through date/time information or other connecting fields, the digital library archeologist can reassemble these artifacts, reconstructing the user’s path through the digital library. This can have two effects: first, it allows the library decision-makers to better understand how people are using the system; second, by connecting these data artifacts into a large data warehouse, many more patterns between library systems can be discovered through pattern-discovery statistical and analytical tools such as bibliomining [5].

Figure 1 shows artifacts as seen by a traditional archeologist and artifacts as seen by a digital library archeologist. These digital artifacts are from different

Web server logs, which include (a) who came to the digital library, (b) where that individual came from, and (c) what that user did in the library. These are just examples – other possible sources include Web proxy logs, OPAC searching histories, metadata about works in the collection, and patron demographics.

Traditional Artifacts	Digital Library Artifacts [6]
	(a) 07/09/99, 13:59:24, , 999.999.99.99, scooby.northernlight.com, crawler@northernlight.com, Gulliver/1.2 (b) 08/02/99, 12:02:35, http://ink.yahoo.com/bin/query?p= ”sample+log+file”& b=21&hc=0&hs=0,999.999.999.99, jaz.med.yale.edu (c) gateway.iso.com - - [10/MAY/1999:00:10:30 –000] “GET /class.html HTTP/1.1” 20 10000

Figure 1: Artifacts of Traditional and Digital Libraries

Just as pottery shards are gathered, washed, and reassembled, digital library artifacts will have to be gathered, cleaned, and reassembled to gain an understanding of the users who passed through that space while still protecting the privacy of the user. These artifacts may have to do with the behavior of use, the demographics of the user, or features of the item being used; when reassembled through a data warehouse, the collection of these artifacts will provide a better understanding of the digital library user and transaction than the pieces taken separately. One pioneer in digital library archeology is Joseph Zucca at the University of Pennsylvania, who has created a data farm where the artifacts of use are gathered, cleaned, matched, and privatized [7].

The completed plate from a combination of shards is akin to a reassembled session of use; not all of the pieces have to be found to provide a reasonable idea

of the original item. In fact, removing the piece that contains personally identifiable information about a user can still leave enough information to create generalizations about use while still leaving connections between other data-based artifacts (see visual analogy in Figure 2).

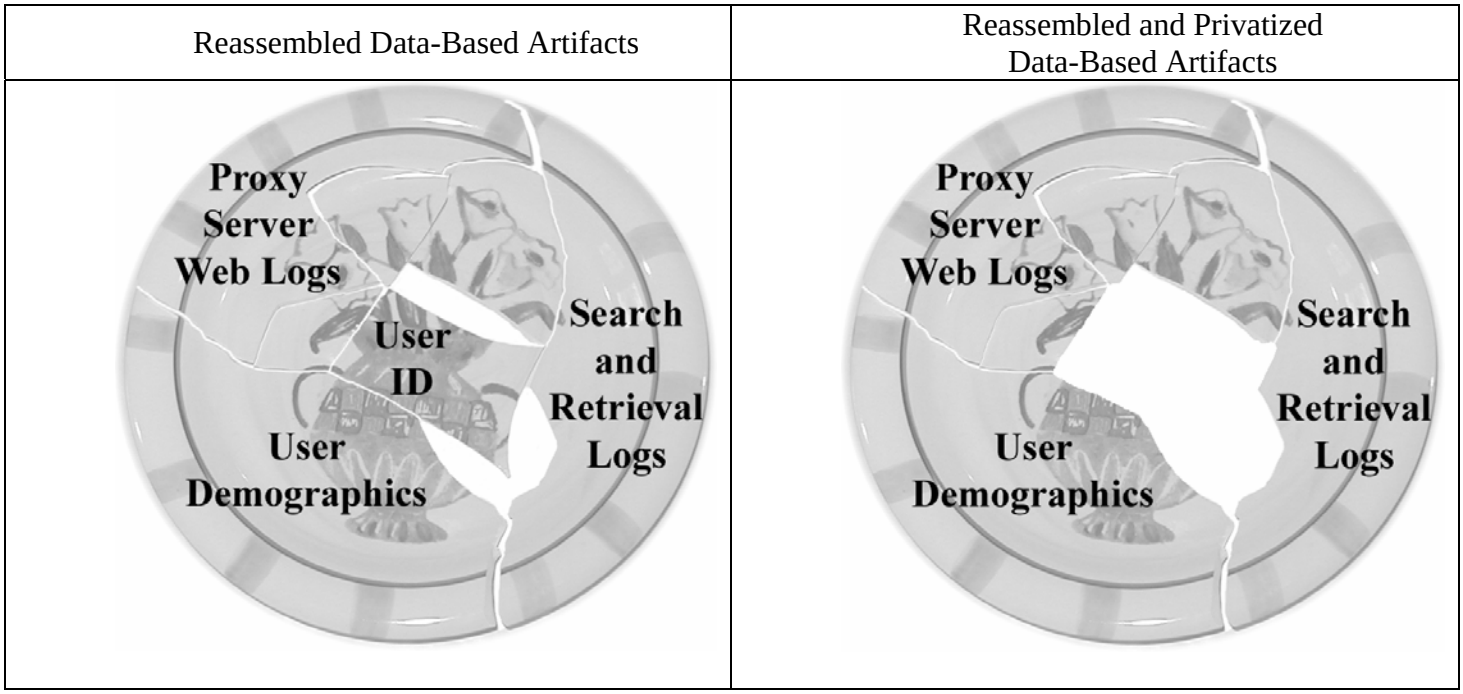

Figure 2: Visual analogy for combining data-based artifacts to understand library services and use

One definition of science is that it is a method of organizing knowledge of a subject to aid understanding [8]. The goal of this paper, therefore, is to present a framework based on archeological theories in order to aid in the understanding of

library users through the artifact-based measurement and evaluation of digital library services.

One method used in the measurement and evaluation of library services is the search for patterns; this can be done manually or with the aid of statistical and data mining tools. The concept behind *bibliomining*, or the combination of bibliometrics and data mining to understand library services, is similar, in theory, to the pattern-discovery concepts in archeology [5]. Nicholas et al., at the CIBER research center, has focused on understanding the user through the artifacts left behind in a system [9].

Review of Related Models

The concept of an information artifact for analysis of an information retrieval system was defined by Green and Benyon as "any artifact whose purpose is to allow information to be stored, retrieved, and possibly transformed" [10,803]. They proposed ERMIA, or the Entity-Relationship Modeling for Information Artifacts as a way of using data modeling to look at the key elements of a system and a user's manipulation of the system. The concept of artifact in the present archeological exploration is much more narrow, focusing on those pieces of data left in a system from a user's interaction.

Griffiths, Hartley, and Willson [11] presented a model for understanding user-system interaction through a combination of transaction log analysis and verbal protocol analysis. This technique starts with traditional log analysis techniques where the selections made by users in a system are recorded. This is supplemented with a recording of mouse movements on the screen. A talk-aloud

technique is then used at the same time to collect statements made by the user about their exploration. The result is a more complete picture of the user's interaction with an information system.

This more complete picture of user-system interaction is related to the holistic matrix of library measurement presented by Nicholson[12]. This model for library measurement and evaluation is useful in understanding where a particular type of library measurement falls in perspective with the larger body of measurement and evaluation literature. The base of this model is the measurement matrix (see Table 1), which divides measurement of library services into four classes.

Measurement	**Topic**	
Perspective	*System*	*Use*
Internal (Library System)	Procedures Standards	Recorded interactions with interface & materials**-Bibliomining**-
External (User)	Aboutness Usability	Knowledge states Value of works

Table 1: Measurement Matrix from Nicholson's holistic measurement framework

This holistic framework is related to the model of Evidence-Based Librarianship (EBL). The concept behind EBL is that librarians should seek ways to integrate the best available evidence into their decision-making processes. Library evaluation should start with a research question, seek out evidence within the literature for assistance, and use research methods that reduce bias in order to collect evidence useful not only in resolving the original research problem, but that may be useful in other decision-making situations[13].

EBL was inspired by Evidence-Based Medicine, which relied upon a large amount of medical research in order to encourage physicians and decision-makers to take advantage of the available research. Library Science has yet to build a corpus of research similar to that of medicine. Therefore, the evidence for EBL has to come from other methods of collecting evidence about library services.

The focus of the present work is to understand the interaction between a user and electronic resources through a digital library service that occurred in the past. One science used to gather evidence about the past is archeology. Therefore, in order to perform Evidence-Based Librarianship, the present work explores archeological thinking as a way of considering our exploration of the data-based artifacts left behind in a digital library system, and considers the data warehousing and data mining process known as bibliomining as a primary tool in that exploration.

Meedows considered a link between archeology and information science from the perspective of the world of scholarship[14]. His work explored the concept of a scholarly work as an artifact and the exploration of these artifacts through bibliometrics will provide an idea of the scholarly community. The current work differs from this piece in that the artifacts here are primarily of use, and the subjects of the research are the users of the digital library. Therefore, Meedows' proposal for archeology of the community of scholarship is complementary, although different, to the work presented here.

Archeological Theory

Archeology is much more than the gathering of artifacts from the past. These artifacts actually exist in the present time, and the past is only created when researchers talk about it [15]. The task of archeology is to make an educated guess about what the past was like based upon the artifacts that exist in the present. Johnson emphasizes that archeology is not the collection of materials; rather, archeology is the interpretation of the meaning behind those materials.

Differences in archeological theoretical frameworks stem from the different methods of developing these interpretations [15]. In the present article, three different frameworks that trace the development of modern archeology will be presented and applied to digital libraries: traditional archeology, new archeology, and postprocessual archeology.

Traditional Archeology

Before 1960, archeology was based in antiquarianism: it started with the collection and thorough description of large amounts of artifacts from a site. After collecting and describing, archeologists explored the implications for the cultures that occupied the area based upon these artifacts. They created chronological maps of the cultural changes in one area or diffusion maps that emphasized how one cultural change moved from place to place. The focus was mainly on the description and mapping of artifacts and a discussion of the culture [15].

This emphasis on antiquarianism was natural, as those who funded archeological expeditions desired the collection of these artifacts. As a side benefit to the collection, archeologists conjectured about culture and diffusion. However, archeologists began to feel that there was little advancement in the field. While they were gathering and describing artifacts and culture, they were not advancing archeological theory that supported how people lived [2].

Traditional Archeology and Digital Library Evaluation

This scenario is similar to the current state of digital library evaluation. Most reports of digital library usage are aggregations of user groups and materials accessed. The usage data are collected, grouped, and described much in the same way that archeological artifacts were collected, tagged, and grouped. Some discussion of the users involved might be explored, and then the results are presented or published, just as artifacts are presented for others to view and enjoy. Many contemporary digital library evaluations published today fall into this "gather and describe" cycle [16,17] although some researchers, such as [9,18] have developed more complex research explorations.

New Archeology

Traditional archeologists collected and cataloged artifact after artifact, creating more of the same types of descriptive culture maps. The addition of more data, however, was not advancing the knowledge base as little high-level analysis was occurring. Throughout the 1960s and 1970s, a shift occurred in archeological thought. The Binfords [2] are attributed with one of the first significant writings

on the topic of new archeology; in reality, the movement was a methodological adaptation that spread over time through the body of archeological thought inspired by a common dissatisfaction with the traditional methods [8].

Traditional archeology was focused on the material artifacts and the concept of a “culture” and not on the individuals behind those artifacts and those cultures. The call came for archeologists to be both more scientific and more anthropological in their studies. New archeologists no longer used “culture shift” as the only reason why anything changed; instead, archeologists were pushed to consider the individuals involved and the culture as a system made up of individual (and mutable) components [2].

There were several new types of archeological thinking that came out of this new archeology. The first was generalization through systems theory – by looking at culture as an interrelated system of sub-components, it became possible to generalize findings from one culture system to other culture systems that shared those components. Along with this new way of thinking about the internal components of a culture, archeologists began to seriously consider the external forces other than culture and other cultures [15].

Another shift was in the nature of archeological science. Instead of describing *what* was found at a site, archeologists began to ask *why* these items were found at the site. The goal became to understand the underlying process that

caused these artifacts to end up where they did rather than focus solely on the artifact. These types of questions led archeologists to begin employing the scientific method; that is, presenting hypotheses about a phenomenon and testing those hypotheses through gathered data. Employing this method also required archeologists to be explicit about their definitions and biases [15].

Hypothesis-based archeology led to a new type of artifact collection. Instead of the goal of traveling to sites to collect large numbers of artifacts, archeologists chose sites and excavation strategies with specific research questions in mind. Sampling methods were employed in site selection and efforts were made to be representative in selections. Artifacts were collected to aid understanding of an issue and not to fill the storehouses of a museum [15].

David Clark, in his classic work *Analytical Archaeology*, focused on an archeological model with three portions – data recovery, systematic description, and model/hypothesis development. In order to achieve these goals, Clark presents three objectives in the archeological process. First, researchers must define entities and processes involved with the artifacts gathered and understand how these entities and processes function in the larger cultural system. Second, archeologists hunt for patterns, or "repeated similarities or regularities in form, function, association, or developmental sequence amongst the particular entities from every area, period, and environment"[1,20]. Finally, the scientists seek to

create models and hypotheses that synthesize discoveries and are predictive in nature.

According to Kaplan [19], the discovery of patterns is one way of creating new scientific laws. Some might argue that researchers are more interested in what makes people different rather than what makes them the same. However, it is only possible to understand differences between people by contrasting them through their similarities. Kaplan states that “how we conceive of an individual is the product of generalizations . . . To understand a person or a particular configuration of an event is to know something of what kind of person or happening it is; if we have no generalizations to draw on, no kinds are available to use for knowledge of individuals.” [19,118] Therefore, discovering these patterns and using them as the inspiration for hypotheses (and eventually laws) are essential to understanding differences.

Binford [2] presented three types of theory developed from archeological studies. The first, Low Range Theory, is applicable only to a single setting and explains one aspect in that setting. A Middle Range Theory can be applied beyond the setting in question, and works to explain an aspect across settings. These theories are “generalizations that attempt to account for the regularities that occur between two or more sets of variables in multiple instances” [20,21]. Finally, Upper Range Theories apply to all settings; Binford felt that archeology could not produce Upper Range Theory.

New Archeology and Digital Library Evaluation

This evolution in archeological theory can aid researchers in thinking about ways of evaluating digital library services and understanding the use of those services. New archeology encourages researchers to move beyond the collection and description of artifacts; therefore, library evaluators need to move beyond presenting basic summaries of their data-based artifacts as the final product of an examination. The next step is to look for patterns, or "systematically correlated attributes that give recognizable group identity" [1,21] in order to seek out subgroups of users and behaviors. Digital library archeologists create a framework of understanding from these generalizations, and researchers can then develop testable models and hypotheses based on that framework to better inform the library decision-making process.

It is important to think about the library as a system of components in order to aid generalization. By focusing an examination on one component, then researchers may be able to create research questions that will lead to building theoretical knowledge for library use. Systems theory suggests that care must be taken when generalizing results; if the results from measuring one aspect of one library subsystem are affected by another subsystem in the library, then results may not be easily generalizable [21]. One way to resolve this conflict is to test the same subsystem in different library settings; if the results are consistent in different library subsystems, then researchers can be more confident in the generalization of results.

In addition, researchers should seek ways of comparing system use across different settings. One track of existing research in this vein is Evidence-Based Librarianship, which uses meta-analysis to combine studies to produce more generalizable results [13]. Through these additional steps, researchers can advance the theoretical understanding of library use and therefore advance the science of librarianship through the development of Middle Range theory.

Finally, library evaluators can benefit from hypothesis-based research as they can answer questions beyond what is available from the artifacts left in the system. The artifacts provide insight only to how the user manipulated the system; they do not allow understanding of other aspects of use, e.g., what the user was thinking during the process, why the user made those choices, what else the user wanted to do, if the user was satisfied, or if the information need was met. Since a researcher cannot determine many user issues from only the artifacts left behind, then a scientific approach involving an educated guess (a hypothesis) and a research question followed by a study where users are involved may shed light on the issue.

This need to move from a practical evaluation toward a hypothesis-based exploration to improve the science of librarianship has been voiced explicitly and implicitly by other library scientists. As McClure states, “library and information science fosters little research that is intended to produce ‘knowledge for the sake

of knowledge'"[22,17] and focuses on the gap between the generalizable research of library scientists and the applied action research desired by librarians. He argues for the need for ways to increase the impact of research on libraries. Tenopir [17] examines many significant large-scale library evaluations and find that most of them draw conclusions only about individuals or specific groups of users. She also develops numerous generalizations based on the collection of studies; this is a perfect example of the next step in generalizing results from traditional library evaluations and could serve as the inspiration for many testable hypotheses.

Combining Traditional and New Frameworks

Up to this point, two different archeological frameworks have been presented along with some discussion of how each could be applied to bibliomining. Just as archeological research based on both frameworks has added to the understanding of the past, evaluators looking to understand more about library users can employ a combination of these frameworks in analyzing the data-based artifacts of use. There is such a combined framework for scientific archeology that is appropriate (which is also used in many other branches of science) - the hypothetico-deductive-inductive scientific cycle.

Hypothetico-Deductive-Inductive (HDI) Scientific Cycle

This cycle, first outlined by Kemeny[23] , was applied by archeology by South [24] and provided the bridge between traditional and new archeology. Traditional archeology focused on describing and finding patterns within the data,

and new archeology started with a problem and sought data to support or refute hypotheses. Scientific archeology employs this HDI cycle to connect these two frameworks.

In an archeological site, just as in the logs of a digital library, there is a large amount of data available. The first step is to collect samples of data from around a site and explore those data for patterns. Inspired by these patterns, the researcher creates basic generalizations about the data. These generalizations may be based upon cultural variables or other demographic/geographic factors. These generalizations are examined for gaps and commonalities (along with data and findings from other studies) in order to develop hypotheses about the people who make up these groupings.

Research questions are created to explore the hypotheses, and then additional data are gathered to test those hypotheses. These data may come from the same source or may require different sources. This method may support or refute the hypotheses, which has the effect of building the knowledge base for the field.

New archeology focuses on the second half of this cycle and functions well when based on the artifacts from traditional archeological methods. Both methodological frameworks are useful in building knowledge from the successful exploration site – the original data power preliminary findings, and researchers go

beyond the data to create and test hypotheses. Upon moving to a new site, the process begins again, and scientists can combine knowledge from previous explorations when deducing new research questions and hypotheses to explore [24].

South presented this hypothetico-deductive-inductive cycle in a graphical chart, called the Dolphin chart, which used a dolphin hopping out of the “The Particularistic Sea of Observed Facts” into “The Nomothetic Atmosphere” and back into the sea. During its trip in the air, the dolphin goes through the stages of “Induction (Pattern Recognition), Theory (Lawlike Generalty), Deduction (Logical Analysis), Prediction (Hypothesis), and

Verification (Testing)” [24,19]

This methodology is applicable to the evaluation of digital library services. Currently, many researchers are in the first few phases, akin to traditional archeology. Most evaluators of digital library services gather data and describe the data through graphical and basic statistical analysis. Many of these evaluations stop at this point, although some go further and discover more descriptive patterns and generalizations beyond raw frequencies and averages. The challenge, therefore, is to move on to the types of tasks inspired by new archeology. These patterns and basic generalizations can be analyzed to create problem statements and hypotheses about digital library use. Research studies can be performed to explore these hypotheses and further our scientific knowledge about digital libraries.

Postprocessual Archeology

Along with postmodern movements in other fields, some archeologists claim that the rigid scientific method presented cannot capture the nuances of the individuals involved with the culture system. As Wylie said, “archeologists would not know (could not determine) whether, or in what respects, past contexts diverge from hypothetical reconstructions of them given the nature of their evidence” [25,21]. Another concern is that if the artifact data can only be interpreted in certain contexts, as the new archeology stated, then a set of data could produce different results under different contextual assumptions [25].

The term postprocessual does not refer to a single method; rather, many different frameworks of archeology take on the postprocessual label. Up until this point, there was a reasonable amount of consistency in the way archeological science was practiced, but postprocessual methods were many and varied. Postprocessual archeology is a collective term that represents many different mental models for looking at the past, such as feminist archeology, interpretive archeology, and postmodern archeology. The common factor is that they all look at new ways to add something to the traditional and new archaeologies by looking to understand the individuals that make up a “culture.” Critics of the term postprocessual state that it “is a shade arrogant in presuming what it might properly claim to complement” [26,43]; Prucell suggests that postprocessural archeology is complementary to processual forms [27]. Therefore, concepts from

postprocessual archeology will be presented here as an addition to, rather than a replacement for, other forms of archeological frameworks.

Even though postprocessual archeology encompasses many different viewpoints and frameworks, Hodder [28] presented several commonalities found in most postprocessual styles, four of which are applicable here. First, postprocessual archeology turned its focus to the theories involved in the process rather than focusing on the methods used, and, as a result, the theoretical base of archeological knowledge has grown. Second, postprocessual archeologists look to link archeology to its related fields, anthropology and history. Third, the processes within cultural systems become a target of analysis in postprocessual archeology, aligning archeology with systems theory. Finally, archeologists using these methods find themselves moving further away from the goals of funders and the interests of the public; instead, they are furthering the science of archeology. We will revisit these commonalities when considering the implications of this type of artifact-based evaluation.

Postprocessual Archeology and Digital Library Evaluation

While there are several forms of postprocessual archeological thought to choose from, postmodern archeology is one with direct application to digital library evaluation. Postmodern archeology eschews the process that produces generalizations about a culture and instead focuses on the individuals within that culture [8]. Any culture is made up of individuals, each of whom makes his/her own life decisions. In order to understand the culture, a researcher must

understand the individuals who make up that culture. Due to the focus on users by information science researchers such as Dervin and Nilan [29], there is considerable research on the importance of users in artifact-based evaluation in the library evaluation process. One good example of where this process has been effective is the development of relevance research, which has matured from a system-centered to a user-centered research stream.

This is useful in the digital library evaluation process after bibliomining. While bibliomining can aid in the understanding of what patterns of usage exist, it cannot aid evaluators in understanding what documents people find relevant (as compared to what document they used). In addition, the results from bibliomining are constrained by a system used; if a feature is not offered by a system, no information will be gathered about that feature. This is why bibliomining, like archeology, cannot produce the truth, as the truth about the information seeking situation lies within the individual doing the seeking. Bibliomining can only recover the facts about what interactions an individual had with the system and cannot recover anything about the user's mental state. Therefore, applying these methods allows the researcher to bring in qualitative elements to the quantitative bibliomining process.

Key Differences

One advantage that digital library evaluators have over many archeologists is that users of digital libraries are from the current era. In fact, depending upon

the situation, evaluators might be able to talk to the users of the system to supplement the patterns discovered with information not available through the library systems. These types of measurements are represented by the User row in the measurement matrix presented earlier (Figure 2). However, in order to guide researchers in what questions to ask users, methods presented here can aid in the development of frameworks from which to base user elicitations. Therefore, while gathering information directly from users is not bibliomining, it can be invaluable in gaining a more holistic understanding of digital library users. Only through talking with users can evaluators move from the gathering of facts to the truth behind those facts. In addition, researchers can explore the impact of the digital library on the users by examining their satisfaction, continued information seeking, and eventual use of retrieved information.

Some may wonder why the archeological framework is needed if the users are available for studies. This concern can be answered from both a practical and a research perspective. From a practical perspective, user studies take considerable resources to execute; once the data collection techniques are created, it is much easier to gather artifacts for exploration. From a research perspective, starting with generalizations from collected artifacts allows the researchers to create a framework from which to develop testable hypotheses. Without this framework, researchers may be challenged in determining what to gather in resource-heavy user studies; however, by developing hypotheses from a

conceptual framework of generalizations from gathered artifacts, researchers are more likely to gather useful information from studies of users.

There is also a difference between the two applications in regard to artifact production. In most digital library evaluation applications, there may be additional artifacts produced by the culture in question daily. This is the value of a data warehouse; it allows the researcher to capture appropriate operational data and work on that captured set without disrupting the operational systems (just as archeologists sometimes have to excavate in ways that avoid disrupting members of the current culture). At the start of the pattern-discovery process, library researchers will work with a snapshot of the data at that time and temporarily ignore the incoming data. Depending upon the situation, they might return and look at the data or integrate them into the warehouse later in the process.

Completing the Cycle

Adding postprocessual archeology to the existing hypothetico-deductive-inductive cycle allows researchers to move beyond the data in the system to gain a better understanding of their users. In order to combine them, processes are separated from products in South's cycle (presented earlier) in order to better fit a digital library. By employing the full cycle, researchers can move from describing to improving the knowledge base about library users. The full cycle including products is presented in figure 3, but the significant processes of this cycle are:

Collection: Gathering artifacts about library services, users, and resources,

Induction: Pattern recognition through bibliomining and visualization,

Deduction: Logical analysis of patterns to produce generalizations,

Prediction: Creation of hypotheses from generalities, and

Testing: Research developed to test hypotheses.

Gathering the Artifacts

The first step in this cycle is to gather artifacts. Traditional data mining requires a data warehouse, which is a collection of data cleaned for analysis. The data warehouse is then split into a development dataset and a testing dataset, and patterns are discovered in the development data. However, most libraries do not have a data warehouse, and creating a data warehouse from scratch is a daunting task. If researchers plan to build a complete data warehouse before beginning analysis, many analysis projects will never get started. The amount of time required to create a full data warehouse is a significant investment, and many library managers are reluctant to do this without some proof of concept. Therefore, the advantage to starting with easily available artifacts is lower cost; however, generalizations developed from a few hundred artifacts are more prone to bias than a larger sample of thousands.

In employing new archeology, researchers do not need to gather all artifacts before beginning, just as archeologists no longer start by excavating every artifact at a site. Just as the archeologist starts by looking at artifacts that have come to the surface over the years, the digital library archeologist can start with artifacts of use that are readily available. The amount of time spent and

sample gathered must be large enough to detect patterns of interest, but at this stage a complete data warehouse may not be needed.

Induction (Pattern Recognition)

Once artifacts have been collected, the researcher can employ pattern recognition techniques through inductive techniques such as bibliomining. Induction is “the process by which a scientist forms a theory to explain the observed facts” [23,93]. With small datasets, manual examination may suffice. Basic aggregations and graphs are typical tools used in evaluation and may allow a researcher to uncover patterns. On-line Analytical Processing (OLAP) tools can allow a researcher to interactively explore these basic aggregations and graphs; OLAP tools also speed up the process of creating reports comparing different pairs of variables.

However, these typically-employed methods will miss patterns that are hidden by aggregate methods. When a community made up of subgroups, each of whom has a different behavior, is measured through aggregates, the distinctions between the groups disappear and “average out.” This is not useful in understanding what has been really going on, and can even be detrimental in decision-making. In addition, Michael Buckland recently stated that one of the current challenges for library research is to understand the differences between user communities [30]. One method is through visualization, where a researcher explores the data to discover patterns. Another method of discovering patterns

about groups of users in the data is bibliomining, or the combination of bibliometrics and data mining [31].

Bibliomining for Library Measurement

Bibliomining is one way to measure the use of the library service from the etic perspective of the library system (as compared to the emic perspective of the user) [32] through the collection, cleaning, and mining of library artifacts. It is important to note that there are other aspects of measurement, each of which adds to the holistic understanding of the system. This research does not propose that bibliomining is the only method of measurement needed; rather, it provides one view of the way that the system is used. For example, bibliomining will not address the relevance of documents recalled by users; it will uncover only that they were used. In addition, it provides little evidence as to the knowledge-seeking state of the users; in addition, it does not allow understanding of what is missing in a library collection [12].

An important note to make during this step is that this pattern discovery can take several paths. A common way of using data mining and visualization is to do undirected exploration, wandering around for “novel and actionable” patterns [33]. However, this can produce many patterns that are neither novel nor actionable, and will take a significant amount of time to pick out the useful patterns.

A better way to start is by working with domain experts who can present topics for exploration. A librarian may have anecdotal evidence for a belief; this can inspire a fruitful pattern discovery process. Another inspiration can come from topics that are not fully understood by the domain experts. If the experts have noted a pattern of use but have no hypothesis for the reason behind that pattern, researchers might be able to use that topic as an inspiration for pattern discovery.

One of the reasons for this approach is that this research process can take considerable time and resources. By starting with problems identified by the library managers, researchers can begin to show valuable results early in the process. After building the personal and data infrastructures to work through this research process and resolving some issues inspired by the funders, then the researchers can use undirected exploration. The researchers will understand the domain more clearly and will be better able to pick out the useful patterns during the process.

Deduction (Creating Generalizations)

After using various manual and computer-aided techniques such as data examination, visualization, statistical analysis, and data mining to find a number of patterns, the researcher must then examine the artifacts, aggregates, and patterns for generalizations. Data mining collects many patterns, some of which are trivial or nonsensical, and it requires human intelligence to pick the wheat

from the chaff. While working through the patterns, the researcher then employs deductive reasoning to produce generalizations about behavior. Library professionals can prove invaluable during this stage as they can aid in the development and confirmation of these generalizations. These generalizations will be the seed from which knowledge about library use can grow.

Prediction (Developing Hypotheses)

After creating generalizations, the next step toward the understanding of library use is to develop hypotheses. The generalizations come only from *the artifacts collected*; therefore, the generalizations created from the artifacts are limited to what the user actually did with the system. These artifacts will shed little light on what the user was not able to do with the system, the relevance of the items examined in the system, or what the user did with the items retrieved in the digital library. All of these areas of knowledge would be useful in decision-making and justification involving the library services.

Therefore, the hypotheses developed by the researcher can go beyond the artifacts. The researcher will be more knowledgeable of the system because of the artifacts, and will be able to extend that knowledge by making educated guesses (i.e. hypotheses) about the use of libraries. For example, the artifacts may show that there is a pattern where users examine a similar series of objects and then leave the digital library. Researchers may examine the pattern and create a hypothesis about the relevance of the last item examined in a chain. This

hypothesis cannot be tested without talking to a user, but may be useful in the design of the digital library. Therefore, researchers should not be constrained by the data examined up to this point when creating these hypotheses; rather, they should work with librarians to craft researchable hypotheses about library use inspired by the patterns seen thus far.

If the behavioral data cannot be directly manipulated with quantitative techniques, there is another route to developing these hypotheses. The grounded theory method is a qualitative research method that starts with observations of behavior. The researcher codes the behavioral observations into concepts. These concepts are then grouped into categories. Finally, hypotheses are created that link the behavior to the categories [34]. These hypotheses can then be used in the same way that hypotheses from bibliomining are used.

Testing

The final stage in building knowledge about library use is the development of research questions to shed light on the hypotheses. Depending on the question, this may require different data sources. The researcher may need to use a subset of the data warehouse that was not used earlier in the process or may have to go back to the original source of the artifacts for previously uncaptured information. This also might require going beyond the data captured in the library system to data captured in other systems, such as citation information captured in a citation

database that can show the impact of a work. Another option is for the researcher to work with recently-produced artifacts to see if patterns have changed.

As Saracevic and Kantor explore, working directly with the users may be essential in measuring the value of a library service [35]. If the research questions involve learning more about the users than can be captured in the system, then there are several options. If the library requires logging in, then a research project could be designed to work with actual library users to integrate tracking and pattern discovery with surveys, interviews, talkalouds, or other measures designed to extract information from a user during or after the use of the library.

One problem with these methods, however, is that behaviors of users may change when measured; matching what the user did in the library with post-use interviews about how the information was used may allow a more thorough understanding of these behaviors. Therefore, in order to accurately test the hypotheses, the researchers may have to start the measurement process anew with a group of subjects and gather information from these subjects throughout their information seeking process using both qualitative and quantitative methods. Working with the same users throughout their library use can avoid many of the biases introduced when only examining part of the process. These studies then are based in the framework developed through the early part of the archeological process, but then use a more postprocessual approach to understanding users.

If privacy or other issues prevent access to real users, however, then "simulated users" must be employed; while not ideal, this can give the researcher inspiration regarding the research questions for future studies involving real users. These simulated users may be researchers using actual or simulated queries or demographically representative users working with either their own queries or queries taken from the system. While it may be appropriate to simulate user responses in an interim stage of hypothesis testing, actual system users must be employed at some point in the testing of hypotheses.

Each combination of these variables will provide a different viewpoint to the use of the system; researchers must consider their options carefully and make a choice based upon the hypothesis and research question. Binford summarized this concept from an archeological perspective as follows: "the reliability of conclusions reached by an archeologist varies directly with the degree to which the subject is removed from discussions of artifacts themselves" [2,21].

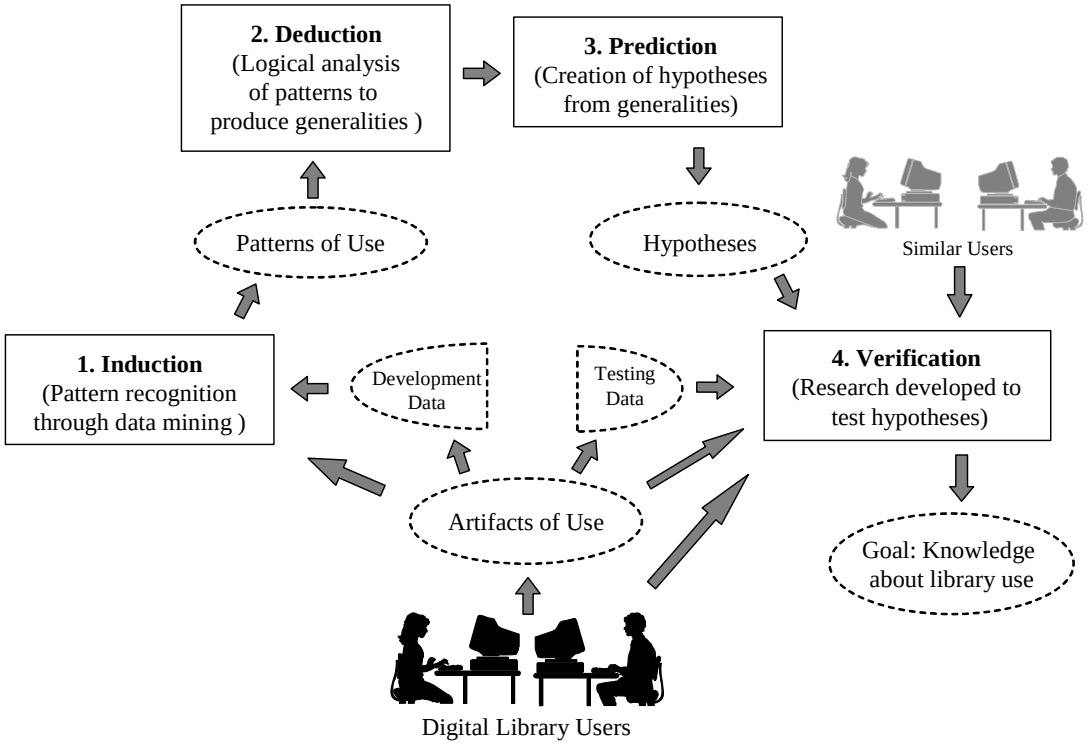

Figure 3: Artifact-Based Evaluation Cycle of Digital Library Users

The figure above brings together the processes (in boxes) along with the products used/created by the processes (in ovals). The verification process may not use all of the items shown, however, each type of resource tapped in verification allows for a more thorough understanding of use.

There are three categories of knowledge, each with a different level of reliability, which can be gained through the method presented here:

-How the users manipulated the system, which comes from the artifacts of use left behind;

-How users used the information, which comes from the user or citation artifacts after the library use;

-What users could have experienced during the library use, which comes from simulated users or real users during the process.

An important note is that one can never know what the user was thinking and experiencing from only the artifacts in the system; even talking to the user after the process will not provide an accurate portrayal of the thought process that the user went through. Part of this comes from the challenge of measuring relevancy, as a user may change their relevancy decisions about items through the searching process. An item found early and considered relevant may be deemed unneeded later in the searching process. The only way to capture this information is working with a user while he/she is searching, but that interference may change the user's behavior. Just as an archeologist does not knows the truth about the people who lived in a culture, we can not know the truth about our users from only their artifacts . We must work directly with users in order to test our hypotheses about their behavior. This is supported by Wylie's argument for postprocessual archeology: it is essential to examine different and independent sources of data and artifacts in order to test hypotheses [25].

Looking back at the commonalities of postprocessual archeology discussed earlier, we can consider a few issues regarding possible effects of implementation of this model. Just as postprocessual archeologists face challenges

in pleasing traditional audiences, “digital library archeologists” will have to ensure that the library decision-makers perceive value from this research. To strengthen results, library researchers can look to related fields for theoretical and methodological inspiration, and focus on systems theory to consider how library processes are connected. Finally, through the introduction of a theoretical focus, library researchers employing this artifact-based evaluation process will be able to expand the theoretical base of library science.

Applications of Digital Library Archeology

This framework can be used to add depth to any evaluation project. Because digital library archeology forces the investigator to think about patterns and ask follow-up questions, it is useful anytime librarians are considering library data. The key step in this process is the development of patterns or generalizations. Hypotheses can then be developed either from the patterns or from the exceptions to those patterns, which will then lead to more exploration of the data and follow-up questions to ask users.

Nicholson and Stanton [36] systematically examined the types of data that could be used in bibliomining to better understand and improve library services. As bibliomining is at the core of the framework developed here, the data sources discussed in the article would be useful as resources for digital library archeology. The article explores the available data from several perspectives and, for each perspective, discusses possible ways to explore the data.

The first perspective concerns data from the creators of the library system. Bibliographic information contains a number of fields useful in exploration, such

as subject headings, classifications, publisher, physical container, and other fields in the MARC record. The Dublin Core metadata standard contains a set of fields that can be used to describe Web-based information. Libraries have traditionally used aggregate measures to explore this information, but using bibliometric and data mining methods can tease out other patterns to inspire hypotheses and further exploration[36].

Related to bibliographic information is acquisition information. Understanding how something was brought into the library and looking for patterns between selection plans, acquisition staff, or publisher packages and use of materials can lead to important discoveries. These can fuel generalizations, which then can be tested and implemented to help selectors make better decisions and managers create more effective policies[36].

The second perspective considers data from the usage of the library system. These data include information about the users of the system, such as demographic data. It is important to protect the privacy of the user when collecting this information by collecting demographics instead of user identifiers[5]. This type of information may be impossible to collect about users of Web-based information; for decision-making, it may be important to consider a way to collect and link demographic information to use.

Another area of data from the usage is circulation, searching, and navigation information. The definition of circulation will change based upon the type of data available, but for digital media, it usually means the downloading of the full text of the piece. Using methods from transaction log analysis can extract

searching data that, when linked to other forms of data discussed here, provide rich data for digital library archeology.

The final perspective for data concerns data sources external to the system that supplies information resources. Conceptually, many people separate digital reference and interlibrary loan from the primary library services; however, it is important when exploring library use to ensure that these library services are also included as part of the library data resources. Other external resources relate to how users enter and exit digital library resources; the incoming referral URL, for example, can be used to separate users into groups to look for different usage patterns[36].

Ideally, the library will create a data warehouse that will connect all of this information to make it easy for researchers to understand the connections between what are traditionally separate data sources. Even with this, it is important to realize that bibliomining does not provide the complete picture; rather, it points the researcher toward interesting patterns to explore through other examinations of the data and discussions with users.

Applying the Framework to E-Journal Research

To demonstrate how this framework may be applied, I will reframe the components from an existing research publication. Phil Davis has done a number of interesting studies looking at the use of digital libraries, the appropriate role of consortia, and the accuracy of vendor-supplied reports about the use of these resources. His study about the behavior of Cornell chemistry journal users from

2003[37] provides an excellent backdrop for applying the archeological framework. He did not directly apply the framework in his article, but all of the components exist in this thorough work to provide a demonstration of how the archeological framework could be applied.

The key components presented by Davis include:

- Aggregate measures of journal use by department (via IP),
- Aggregate measures of article and journal use by individual (via IP),
- Relationship between number of downloaded articles/journals and unique IP addresses, and
- Consideration of user intent behind article downloads.

Putting the Components into the Framework

The first step in the process is gathering available artifacts; in this situation, these include aggregate usage statistics from the suppliers of electronic journals and other studies done about scientists' information behavior at Cornell. Initial patterns regarding chemistry journals arise from the examination of these two sources of data. This is typical at this stage; when locating patterns, one must look for subgroups of patrons or locate patterns that are hidden by general aggregate reports. Another route to locating these areas of interest is the afore-mentioned bibliomining.

The next step in the process is to make a hypothesis. This is a statement of belief that guides future steps of the research. An appropriate hypothesis at this stage is that there are facets of users of a digital library that can be predicted from

the use statistics of a set of journals. In order to examine this hypothesis, then the archeologist needs to collect artifacts that have to do with specific journals and then match those artifacts to other data sources to learn about demographics of users.

In this case, Davis determined which departments were the frequent users of these resources, that most users used only one or two journals, and that there were a few super-users that downloaded over 100 articles in a few months. He also found a strong relationship between the number of articles downloaded for a particular journal and the number of individual IPs that downloaded those articles. This pattern, like most good patterns found from data mining, makes sense; in addition, it allows the librarians to predict how many users are accessing a journal by knowing the number of article downloads of the journal.

The cycle of generalization – hypothesis – testing can then be repeated to examine new patterns that are unearthed. At some point in the study, it can be important to introduce user surveys or focus groups, which corresponds to postmodern archeology. After exploring general hypotheses, Davis learned about usage behavior directly from individuals, and learned that the format of download may be predictive with the intent of use; in this case, users downloading a PDF file were more likely doing so in order to print out the article, while users downloading the HTML version were more likely to browse on the screen. He took this information, went back into the data, and found that the super-users downloaded PDF files more often than “normal” users. This could then be the basis for another hypothesis, which would start the cycle again.

By employing the scientific process through an archeological framework, one can explore and understand more from the same type of data that other reports simply aggregate. The cycle of examine – generalize – hypothesize – test, first using quantitative data then exploring specific questions with users of the system, can create interesting theories. These theories can then be tested in other domains and settings, and move our science of librarianship forward in new directions.

Future Research – Moving Beyond Digital Libraries

This framework has applications far-reaching beyond digital libraries, as the methods can be applied to *any* user base on the Internet. This work has used digital libraries as a setting, but, with very few adjustments, it could be reapplied as a model for understanding consumer behavior, participation in online forums, or any place where users leave behind artifacts as they live in a virtual space. In all of these uses, researchers are encouraged to move beyond simple description and toward building and growing our understanding of user behavior through "Internet Archeology."

Internet Archeology is the understanding of Internet users through artifact-based evaluation. The same techniques can be used – gather easily available data, create generalities, develop hypotheses, perform a more systematic excavation, matching, and cleaning of data, and test the hypotheses – in order to gain understanding of Internet citizens in a systematic way.

Conclusion – Scientific Thinking for Librarianship

This resulting archeological framework of explanation is fundamentally an application of the traditional scientific process. Librarians and others providing Web-based information resources do not regularly use hypothesis-based research projects when evaluating their services. The result is that the science of librarianship has not grown and developed as other sciences have grown. The benefit of using the archeological approach is that it may aid those not accustomed to the scientific method to conceptualize how this method applies to their own information service. Rather than just gathering and reporting statistics, this method encourages librarians to think about the data, ask questions about why certain patterns occur, make hypotheses about these patterns, explore the data to support or refute the hypotheses, and then talk to users to better understand the phenomenon.

Archeologists recognize that their craft is as much art as science, as giving meaning to a collection of artifacts requires a number of assumptions and guesses and does not provide the truth about the mental state of the users. We can draw an essential point from this for exploration of digital libraries. Collections of data-based artifacts tell us an important part of the story, but the discovery of these patterns through bibliomining is just the beginning. By developing generalizations, creating hypotheses about library use, and testing those hypotheses through research involving data and users, researchers can move beyond descriptions and advance our understanding of library use.

Bibliography

1. Clarke, David L., and Chapman, Bob. Analytical archaeology. New York: Columbia University Press, 1978.
2. Binford, Lewis Roberts. "Archeological Perspectives." In New perspectives in archeology, edited by Sally R. Binford and Lewis Roberts Binford. Chicago,: Aldine Pub. Co., 1968.
3. Nicholson, S. "A framework for Internet archeology: Discovering use patterns in digital library and Web-based information resources." First Monday 10 (February 2005).
4. Greenstein, D. "DLF Draft Strategy and Business Plan." http://www.diglib.org/about/strategic.htm September 18
5. Nicholson, Scott. "The Bibliomining Process: Data Warehousing and Data Mining for Library Decision Making." Information Technology and Libraries 22 (December 2003): 146-151.
6. Bauer, Kathleen F. "Who goes there? Measuring library Web site usage." Online (Weston, Conn.) 24 (Jan./Feb. 2000): 25-26.
7. Zucca, Joe. "Traces in the Clickstream: Early Work on a Management Information Repository at the University of Pennsylvania." Information Technology and Libraries 22 (December 2003): 175-179.
8. Praetzellis, Adrian. Death by theory : a tale of mystery and archaeological theory. Walnut Creek, CA: AltaMira Press, 2000.

9. Nicholas, David; Huntington, Paul; and Watkinson, Anthony. "Digital journals, Big Deals and online searching behaviour: a pilot study." Aslib Proceedings 55 (2003): 84-109.

10. Green, T. R. G., Benyon D R. "The skull beneath the skin: entity-relationship models of information artifacts." International Journal of Human-Computer Studies 44 (June 1996): 801-828.

11. Griffiths, Jillian R., Hartley, R.J., Willson, Jonathan P. "An improved method of studying user-system interaction by combining transaction log analysis and protocol analysis." Information Research 7 (2002).

12. Nicholson, Scott. "A conceptual framework for the holistic measurement and cumulative evaluation of library services." Journal of Documentation 60 (2004): 164-182.

13. Eldredge, Jonathan. "Evidence-based librarianship: an overview." Bulletin of the Medical Library Association 88 (Oct. 2000): 289-302.

14. Meadows, A. J. "The scientific paper as an archaeological artefact." Journal of Information Science: Principles & Practices 11 (1986): 27-30.

15. Johnson, Matthew. Archaeological theory : an introduction. Oxford, UK ; Malden, Mass.: Blackwell Publishers, 1999.

16. Jones, Casey, and Sumner, Tamara. "Evaluation of the National Science Digital Library." http://www.uclic.ucl.ac.uk/annb/DLUsability/JonesSumner5.pdf 2004

17. Tenopir, Carol. Use and users of electronic library resources: An overview and analysis of recent research studies. Washington, D.C.: Council on Library and Information Resources, 2003.

18. Bollen, Johan; Sompel, Herbert Van de; Smith, Joan; and Luce, Rick. "Toward alternative metrics of journal impact: A comparison of download and citation data." http://www.arxiv.org/PS_cache/cs/pdf/0503/0503007.pdf 2005

19. Kaplan, Abraham. The conduct of inquiry: Methodology for behavioral science. New Brunswick: Transaction Publishers, 2002.

20. Trigger, Bruce G. A history of archaeological thought. Cambridge ; New York: Cambridge University Press, 1989.

21. Bertalanffy, Ludwig von. "General system theory: A critical review." General Systems 7 (January 1962): 1-20.

22. McClure, Charles R. "Increasing the usefulness of research for library managers: propositions, issues, and strategies." Library Trends 38 (Fall 1989): 280-294.

23. Kemeny, John G. A philosopher looks at science. Princeton, N.J.,: Van Nostrand, 1959.

24. South, Stanley A. Method and theory in historical archeology. New York: Academic Press, 1977.

25. Wylie, Alison. "A proliferation of new archaeologies: Beyond objectivism and realism." In Archaeological Theory: Who Sets the Agenda?, edited by Norman Yoffee and Andrew Sherratt, 1993.

26. Preucel, Robert W. Processual and postprocessual archaeologies : multiple ways of knowing the past. Carbondale, Ill.: Center for Archaeological Investigations, Southern Illinois University at Carbondale, 1991.

27. Renfrew, Colin, and Bahn, Paul G. Archaeology : theories, methods, and practice. New York: Thames and Hudson, 2000.

28. Hodder, Ian. "Postprocessual archeology and the current debate." In Processual and postprocessual archaeologies : multiple ways of knowing the past, edited by Robert W. Preucel. Carbondale, Ill.: Center for Archaeological Investigations, Southern Illinois University at Carbondale, 1991.

29. Dervin, Brenda, and Nilan, Michael. "Information needs and uses." In Annual review of information science and technology, v21, 1986 Knowledge Industry Publs. United States, 1986.

30. Buckland, Michael K. "Five Grand Challenges for Library Research." Library Trends 51 (Spring 2003): 675-686.

31. Nicholson, Scott. "Bibliomining for Automated Collection Development in a Digital Library Setting: Using Data Mining to Discover Web-Based Scholarly Research Works." Journal of the American Society for Information Science and Technology 54 (October 2003): 1081-1090.

32. Pike, Kenneth Lee. Linguistic concepts : an introduction to tagmemics. Lincoln: University of Nebraska Press, 1982.

33. Berry, Michael J. A., and Linoff, Gordon. Mastering data mining : the art and science of customer relationship management. New York: Wiley Computer Pub., 2000.

34. Corbin, Juliet, and Strauss, Anseim. "Grounded Theory Research: Procedures, Canons, and Evaluative Criteria." Qualitative sociology 13 (Spring 1990): 3.

35. Saracevic, Tefko, and Kantor, Paul B. "Studying the Value of Library and Information Services. Part I. Establishing a Theoretical Framework." Journal of the American Society for Information Science 48 (June 1997): 527 (516 pages).

36. Nicholson, Scott , and Stanton, Jeff. "Gaining strategic advantage through bibliomining: Data mining for management decisions in corporate, special, digital, and traditional libraries." In Organizational data mining: Leveraging enterprise data resources for optimal performance, edited by Hamid Nemati and Chris Barko. Hershey, PA: Idea Group Publishing, 2003.

37. Davis, Phil, and Solla, Leah. "An IP-Level Analysis of Usage Statistics for Electronic

Journals in Chemistry: Making Inferences About User

Behavior." Journal of the American Society for Information Science and Technology 54 (2003): 1062-1068.

Figures and Tables