Testing Our Truths: Delineating the Parameters of the Authentic Archival Electronic Record

Anne J. Gilliland-Swetland

Abstract

Based on the work of the first phase of the International Research on Permanent Authentic Electronic Record Project (InterPARES 1), this paper argues that emerging considerations specific to electronic records necessitate that the archival community delineate mechanisms for establishing, maintaining, and certifying the authenticity of preserved and reference copies of electronic records. The paper reviews how the InterPARES 1 Project identified such mechanisms through the conduct of a series of case studies of electronic information and record-keeping systems. The paper places a discussion of the case study data collection, analysis, and outcomes in the context of the quest and rationale for effective methods and a research design that would allow InterPARES researchers to discern the parameters of the electronic record and what supports a presumption of its authenticity.

Introduction

Author Timothy Garton Ash opines of historical facts that:

> Like the materials used in a collage, these pieces of evidence have different textures: here a fragment of hard metal, there a scrap of faded newspaper, there again a wisp of cotton wool. Reporters, investigators and historians will

The author gratefully acknowledges funding from the National Historical Publications and Records Commission; the Social Sciences and Humanities Research Council of Canada; and the National Research Council of Italy; as well as the support of the Banca d'Italia; Beijing Municipal Archives; the Central State Archives of Italy; the National Archives of Canada; the Netherlands Institute for Archival Education and Research; the Smithsonian Institution; the State Archives of Rome; the State Archives of China; the United States National Archives and Records Administration; the University at Albany, State University of New York; the University of British Columbia; the University of California, Los Angeles; and the University of Glasgow for the InterPARES 1 Project.

> compose widely varying collages from the same box of scraps, and further change the picture with the oil paints or watercolors of their own imagination. But there are special truth tests to which their pictures, unlike the poet's or novelist's, must always submit.[1]

Traditionally, once they have made and acted upon their appraisal decisions, archivists, for the most part, have left judgment about the trustworthiness and utilization of archival evidence to end-users, who then make those judgments in accordance with their own domain practices and needs. In so doing, archivists have relied upon their extant professional practices and procedural controls for the collective preservation, description, and reference of their holdings to support a presumption that the archival materials they provide to end-users are authentic, or else are obvious copies made for convenience or preservation purposes.[2] On those occasions when they have faced specific questions about the authenticity of a given document, archivists with the necessary skills have applied diplomatic tests to individual documents.[3] But are there occasions when archivists should also submit their evidence and how they present it to "truth tests," and if so, what should those truth tests be? While this paper is not intended to be an epistemological exposition on the nature of truth in documentary evidence, it is concerned with achieving a better understanding of the means by which archivists can ensure that the copy of the electronic record held by the archives or delivered to the user is authentic—that it is what it claims to be, that it is genuine, that it has not been counterfeited or tampered with, and that it is free of corruption.

Just as the methodologies and practices of history and other disciplines have needed to evolve over the past two centuries to address changing ideas about acceptable and sufficient documentary evidence in light of socio-political developments and shifting modes of mechanical production and dissemination, so too with the methodologies and practices of archival science. Arguably the strongest impetus for change in archival science today are the challenges associated with the long-term management and use of electronic records. While evidence-driven recordkeeping is the construction that is currently at the center of electronic records discourse, the procedural standards set by the legal

[1] Garton Ash, Timothy. *The File: A Personal History* (London: Vintage Books, 1998), 110.

[2] Records can be described as authentic in the sense that these are the same original records that were transferred to the archives and that their integrity has not been compromised in any undocumented way since they entered into the archival bond. This, of course, is not a guarantee of the accuracy or completeness of the records as received from the creator.

[3] The principles of diplomatics were first delineated in 1681 by Jean Mabillon, and have subsequently been developed and extended to address a growing body of document types. Diplomatics has been employed as a means for assessing authenticity for legal, philological, and historiographical, as well as archival, purposes.

rules of best evidence apply primarily to the *creation of reliable records* and are insufficient for *assuring the preservation and dissemination of authentic archival records*. One key motivation for the development of more systematic mechanisms for establishing and certifying authenticity as we move into an environment where archivists are increasingly preserving and providing access to electronic records, is the fact that we are unlikely to be working with original archival electronic materials. This is because all archival electronic records may be more accurately thought of as copies of original records—preservation copies made according to the particular methods and strategies that are appropriate or expedient for the records or the archives, and use copies that may be generated in many, often user-specified forms[4]. Such change also begs the question of whether we should or could extend the dominant construction to include evidence-driven record-delivery encompassing certification of the authenticity of the copy of the electronic record or specified digital components of those electronic records being provided to end-users?

Based on the work of the first phase of the International Research on Permanent Authentic Electronic Records Project (InterPARES 1), this paper argues that emerging considerations specific to the nature, preservation, and delivery of electronic records necessitate that the archival community re-visit and augment its mechanisms or "truth tests" for establishing the authenticity of preserved and reference copies of electronic records. The paper reviews how InterPARES 1 built upon extant archival and diplomatic practices and experimented with additional approaches drawn from the social sciences in order to identify, test, and formalize such mechanisms through the conduct of case studies of electronic information and record-keeping systems. The discussion of case study data collection, analysis and outcomes, therefore, is situated within a larger discussion about appropriate research design. The paper focuses, in particular, on the quest for effective methods that would allow InterPARES and other archival researchers to discern the parameters of the electronic record (for example, the complex of elements, inter-relationships, and contexts of which it is composed) and those inherent or external aspects that support a presumption of its authenticity.

[4] The research outcomes of the InterPARES Preservation Task Force indicate that, because of the inevitable reliance on migration at some level in all current and experimental preservation methods and on reproduction for display and dissemination purposes, it is inappropriate to talk in terms of originals of electronic records. As formalized in the Reference Model for an Open Archival Information System (OAIS) and implicit in the work of the San Diego Supercomputer Center's Persistent Archives technology where data collections are managed as derived data products, it will be possible to generate many types of copies and derivatives based on manipulation of specified digital components of electronic records. See the InterPARES Preservation Task Force Report, available at <http://www.interpares.org> (December 4, 2002); Reagan Moore et al. "Collection-based Persistent Digital Archives—Parts 1 & 2," *D-Lib Magazine* (March/April 2000), available at <http://www.dlib.org/> (December 4, 2002) and several papers on the OAIS Reference Model, available at <http://ssdoo.gsfc.nasa.gov/nost/isoas/ref_model.html> (December 4, 2002).

Delineating the Conceptual Requirements Necessary for Identifying and Preserving Authentic Records

The recently completed InterPARES 1 Project was a three-year international, multidisciplinary collaborative research effort that examined the preservation of the authenticity of records that are no longer needed by their creators to fulfill their own missions or purposes but that need to be preserved as archival records.[5] The work of InterPARES 1 was divided into four domains: Authenticity, Appraisal, Preservation, and Strategies, with research in each domain being conducted by a task force.[6] As part of the project's effort to develop internationally acceptable procedural rules, activity models, and system requirements, the Authenticity Task Force was charged with determining the conceptual requirements necessary for identifying and preserving authentic records.[7] At the outset of the Task Force's research, five questions were identified for study:

- What are the elements that all electronic records share?
- What are the elements that allow us to differentiate between different types of electronic records?
- Of those elements, which will permit us to verify their authenticity over time?
- Are the elements for verifying authenticity over time the same as those that permit us to verify their authenticity in time, that is, at the point at which they are originally created and transmitted?
- Can the elements be removed from where they are currently found to a place where they can more easily be preserved and still maintain the same validity?

As might be expected, however, these initial questions were considerably revised and refined, and new, unanticipated questions emerged during the course of the research. Moreover, we had an overarching immediate issue to address relating to how we arrived at an understanding of the parameters of an electronic record.

[5] The requirements and models generated by InterPARES 1 research are being further tested, refined, and augmented in InterPARES 2, a five-year continuation project that aims to develop a theoretical understanding of records generated by interactive, dynamic, and experiential systems, of their processes of creation, and of their present and potential use in the artistic, scientific, and government settings.

[6] For further detail on InterPARES 1, see the InterPARES Project website at <http://www.interpares.org> (December 4, 2002) and the US-InterPARES website at <http://is.gseis.ucla.edu/us-interpares> (December 4, 2002).

[7] Members of the InterPARES Authenticity Task Force were Heather MacNeil, University of British Columbia (Chair); Luciana Duranti, University of British Columbia; Chen Wei, Beijing Municipal Archives; Anne Gilliland-Swetland, University of California, Los Angeles; Maria Guercio, University of Urbino; Yvette Hackett, National Archives of Canada; Babak Hamidzadeh, University of British Columbia; Livia Iacovino, Monash University; Brent Lee, University of British Columbia; Sue McKemmish, Monash University; John Roeder, University of British Columbia; Seamus Ross, University of Glasgow; Wai-kwok Wan, Hong Kong Public Record Office; and Zhao Zhon Xiu, State Archives of China.

The Theoretical-Deductive Approach

In the first stage of its research, the Task Force established a theoretical-deductive framework for the analysis of various types of electronic records and the identification of those elements that need to be preserved to ensure the records' authenticity over time. Deductive research typically progresses through application of theory to the development of tightly formulated hypotheses, to data collection and analysis, and finally to confirmation or failure to confirm the initial theory. This framework, referred to as the *Template for Analysis,* decomposes an electronic record into its constituent elements.[8] The decomposition defines each element, explains its purpose, and indicates whether, and to what extent, that element is instrumental in assessing the record's authenticity. The theoretical perspective that shaped the development of the *Template* was contemporary archival diplomatics. In adopting the diplomatic approach, we were building upon efforts in recent years to adapt traditional diplomatics to contemporary recordkeeping practices, notably the work of Luciana Duranti in North America, as well as that of Paola Carucci in Italy and Dutch archivists in the Netherlands.[9] The research also built upon the outcomes of a research project carried out between 1994 and 1997 at the University of British Columbia entitled *The Preservation of the Integrity of Electronic Records* ("the UBC Project")[10] which resulted in a set of standards and rules for developing and implementing a trustworthy electronic recordkeeping system. The diplomatic approach was also augmented by theoretical aspects drawn from archival science, most notably ideas about record aggregations and an expanded notion of context.[11]

[8] The term "elements" is used differently in diplomatics to the way in which it is used in information systems design. In developing the initial research questions and the *Template for Analysis,* the Task Force used the diplomatic term "elements" to refer to both general and specific characteristics of a record that may be found in its documentary form, in annotations, or in one or more of its various contexts. As the research progressed, however, the Task Force found it necessary to narrow the scope of the concept. In the *Requirements for Authenticity,* therefore, the term "record elements" refers specifically to the intrinsic and extrinsic elements of a record's documentary form as these are identified in the *Template for Analysis.* Such redefinition is illustrative of how diplomatics continues to evolve in response to the changing nature of the record. For a more complete discussion of the development of the *Template for Analysis,* see Heather MacNeil, "Providing Grounds for Trust: Developing Conceptual Requirements for the Long-Term Preservation of Authentic Electronic Records," *Archivaria* 50 (Fall 2000): 52–78.

[9] See David Bearman and Peter Sigmond, "Explorations of Form of Material Authority Files by Dutch Archivists," *American Archivist* 50 (Spring 1987): 249–53; Peter J. Sigmond, "Form, Function and Archival Value," *Archivaria* 33 (Winter 1991-92): 141–47; Paola Carucci, *Il Documento Contemporaneo* (Rome: La Nuova Italia Scientifica, 1987); and Luciana Duranti, *Diplomatics: New Uses for an Old Science* (Lanham, Md. and London: The Scarecrow Press, Inc., 1998).

[10] For an overview of the findings of the UBC Project see Luciana Duranti and Heather MacNeil, "The Protection of the Integrity of Electronic Records: An Overview of the UBC-MAS Research Project," *Archivaria* 42 (Fall 1996): 46–67.

[11] The outcomes of the UBC Project were subsequently substantially incorporated into the Design Criteria Standard for Electronic Records Management Software Applications (DOD 5015.2-STD) promulgated by the U.S. Department of Defense.

The elements of an electronic record identified in the UBC Project provided the starting point for the identification of the InterPARES *Template* elements, which fell into four main categories: *documentary form, annotations, context,* and *medium.*[12] The first iterations of the *Template* modeled an ideal record that, based upon prior archival knowledge of record types, delineated all the possible known elements that a record may contain. The working hypothesis of the Authenticity Task Force was that, while they may manifest themselves in different ways, these same or similar elements would be present, either explicitly or implicitly in electronic as well as more traditional forms of records. However, where diplomatic typologies and analysis have in the past been developed retrospectively based upon what was known about existing records, one goal of InterPARES was to develop a predictive model that would assist archivists in identifying future record types and the necessary requirements for maintaining their authenticity over time. Based on researcher input from a range of disciplinary perspectives, as well as data collected through case studies, these original elements were revised and extended, therefore, and new elements were added to the *Template* as the research progressed.[13]

The Empirical-Inductive Approach

From our work on the development of the *Template for Analysis,* we had developed several assumptions about the nature of electronic records that we wished to examine further [see Table 1]. In order to develop a richer picture of the complex nature of electronic records, we triangulated the theoretical, deductive diplomatics-based approach with an inductive, empirical approach that was based on an examination of actual electronic records and electronic record-keeping systems. Deductive research, often referred to as a "bottom-up"

Table 1 Assumptions About the Nature of Electronic Records

- An electronic record is a complex of elements and their relationships.
- It possesses a number of identifiable characteristics, including a fixed documentary form, stable of unchangeable content, an archival bond with other records either inside or outside the system, and an identifiable context (including explicit linkages to other records through a classification system or other unique identifier scheme).
- It participates in or supports an action, either procedurally or as part of the decision-making process (meaning its creation may be mandatory or discretionary).
- At least three persons (author, writer, and addressee) are involved in its creation (although these three conceptual persons may in fact be only one physical or juridical person).
- These characteristics manifest themselves in explicit and implicit ways.

[12] For a more detailed discussion of the *Template for Analysis,* see Heather MacNeil, "Providing Grounds for Trust: Developing Conceptual Requirements for the Long-Term Preservation of Authentic Electronic Records," *Archivaria* 50 (Fall 2000): 52–78.

[13] The *Template for Analysis* is available on the InterPARES 1 website, <http://www.interpares.org/reports.htm> (December 4, 2002).

approach, typically proceeds from observation or data collection, to the discernment of patterns, to the development of tentative hypotheses, and finally to the formulation of theory. This examination of the nature of electronic records was conducted by means of purposively selected, interpretive case studies of electronic systems that contained, or were deemed likely to contain, electronic records. The objective of the case studies was to enhance our understanding of electronic records within the various juridical-administrative, provenancial, procedural, documentary, and technological contexts in which they are embedded as well as the relationships of those contexts to each other.[14]

While the addition of this "bottom-up" approach extended InterPARES research activities considerably beyond those originally envisaged, it provided a rich dataset that informed the theoretical development by indicating the increasing role of procedural and technological contexts in ensuring and maintaining the authenticity of records. At the same time, the application of the *Template of Analysis* to existing records and record-keeping systems was able to indicate which necessary elements of form were not present in systems as they were currently designed and operating, thus demonstrating potential weaknesses or deficiencies in the records or record-keeping systems examined.

Use and Selection of Case Studies

The Task Force researchers adopted a grounded theory approach in which case studies of electronic systems were examined in order to identify and describe phenomena associated with the records and their contexts. Grounded theory is a method for discovering concepts and hypotheses and developing theory directly from data under observation.[15] Cases are selected for study "according to their potential for helping to expand on or refine the concepts or theory that have already been developed. Data collection and analysis proceed together."[16] Task Force researchers, therefore, purposively identified the cases that seemed likely best to elucidate phenomena that the research was seeking to understand. For example, we were interested in what happens to active or inactive electronic records when they are migrated to new software/hardware environments.

Between Spring 1999 and Spring 2001, four successive rounds of case studies were conducted by institutional and student researchers in government,

[14] For a discussion of the embeddedness of electronic records within these contexts, see Anne J. Gilliland-Swetland and Philip B. Eppard, "Preserving the Authenticity of Contingent Digital Objects: The InterPARES Project," *D-Lib Magazine* 6 (July/August 2000). Available at <http://www.dlib.org/dlib/july00/eppard/07eppard.html> (December 4, 2002).

[15] Barney G. Glaser and Anselm L. Strauss, *The Discovery of Grounded Theory: Strategies for Qualitative Research* (Chicago: Aldine Atherton, 1967), 6–7, 46.

[16] Steven J. Taylor and Robert Bogden, *Introduction to Qualitative Research Methods: The Search for Meanings,* 2nd ed. (New York: John Wiley, 1984), 126.

university, and corporate agencies in the United States, Canada, Italy, the United Kingdom, the Netherlands, and China. The case studies included large-scale databases (such as patent and student registration systems), geographic information systems, and interactive Web-based applications. The data gathered through the case studies was then used to test and extend the *Template for Analysis*. The translation of the case study data into a form that could be analyzed diplomatically by the *Template* was achieved by coding the data for interrelated themes and concepts using a *Template Element Data Gathering Instrument (TEDGI)*.

As we anticipated, a key issue we encountered was the difficulty in identifying actual electronic records and their parameters. This issue stems from the nature of digital information systems, which are frequently multipurpose, highly networked database systems that can contain a diversity of information elements that can be compiled and presented in a variety of ways (for example, through hard-coded report formats, stylesheets, and virtual views), and which can invoke a range of functionalities, according to the needs of different users. A single system may contain only raw data or information, one or more than one type of record, or a combination of record types and data or information.

The diplomatic analysis of first- and second-round case studies indicated that few of the systems appeared to contain records that came close to the ideal promulgated in the *Template*. Some systems proved to be information systems not containing records at all, while some contained records that were able to achieve their purpose but were not intrinsically very good records. In line with the grounded theory approach, based upon what we had learned from the diplomatic and business information systems analyses of the first two rounds of case studies (discussed below), we modified the case study selection criteria for the third and fourth round of case studies to define more precisely the types of cases in which we were now interested. Through this redefinition, we identified some key indicators of systems that create records or that have the potential to create records [see Table 2].

Case Study Data Collection

A drawback of any research that employs multiple selective case studies is the limited degree to which it is possible to compare across or generalize from individual case studies. Each case is highly sensitive to its own national, juridical, institutional, and technological contexts. Moreover, InterPARES case studies were conducted under a range of different conditions by different investi-

Table 2 Key Indicators of Systems that Create or that have the Potential to Create Records

- If the action in which the system participates is juridically required;
- If there is a business procedure in place to carry out that action; and
- If the system operates within the management or strategic decision-making levels of the organization.

gators. As a result, each case study had to be selected and analyzed on its own merits for how it might inform our theory development. It was necessary, therefore, to be cautious about the extent to which we could look for patterns emerging across case studies in similar institutional settings or in systems performing similar functions in different settings. In an effort to control the individual differences between case studies and within case study rounds as much as possible, we developed a *Case Study Interview Protocol (CSIP)* to standardize the interview process for the case studies, as well as to provide data for populating the *Template Element Data Gathering Instrument (TEDGI).*[17]

The *CSIP* (essentially the interview script) was divided into five sections that reflect areas delineated in the *Template for Analysis:* Context (juridical-administrative, provenancial, procedural, and documentary), Intrinsic Elements of Form, Extrinsic Elements of Form, Annotations, and Medium and Technological Context. A range of standardized questions was asked to elucidate each aspect. The same question was sometimes asked in different ways within the same section to check for consistency in responses. The same question was also sometimes asked in a different way in more than one section to identify any alternate perspectives of respondents with different backgrounds (for example, records managers and systems personnel). Interviewers, predominantly institutional archivists or archival science students participating in InterPARES, sought out respondents who were the records creators, records managers, and systems personnel primarily responsible for working with the electronic systems under study. Because of local requirements and practicalities, some interviews were with individuals, and some with groups of individuals. In some case studies, multiple interviews with different individuals were held. Interviewers also collected supporting materials such as technical documentation, organization charts, and workflow rules; and sometimes followed up with interviewees when further information was required. The interviewers were then responsible for translating the data they had gathered through the CSIP and supporting documentation into the TEDGI, and for transmitting copies of all the case study data to both the University of British Columbia (UBC) and the University of California, Los Angeles (UCLA) for analysis. In the third and fourth rounds of case studies, researchers at UBC were responsible for compiling the TEDGI.

Case Study Data Analysis

Each round of case studies was described and analyzed not only from the perspective of contemporary archival diplomatics but also through the appli-

[17] Copies of all instruments used are available on the InterPARES website at <http://www.interpares.org/reports.htm.> (December 4, 2002).

cation of analytical methods drawn from the social sciences.[18] The rest of this paper will primarily discuss the analysis from the latter, deductive perspective, although it will contrast this with the outcomes of the diplomatic analysis. This analysis addressed the following questions that expanded upon those posed at the outset of the research:

1. What are the elements that are most commonly present across case studies?
2. What are the elements that are most commonly absent, or that cannot be discerned across case studies?
3. What business functions are being supported by the electronic systems studied?
4. What are the activities and transactions performed by the electronic systems in support of the business functions?
5. At which level within the organization do the electronic systems exist?
6. What are the relationships between paper and electronic components of record-keeping systems?
7. In what ways do records creators, custodians, and systems personnel conceptualize the nature and role of the electronic records and/or record-keeping system being studied?
8. What are the variances in language used to describe records by records creators, custodians, and systems personnel?
9. To what extent should or could the conceptualizations and language of such personnel be factored into the design of a method to identify and ensure the preservation of authentic electronic records?

The first type of analysis examined how and to what degree the identity and integrity of electronic records is supported within and across case studies. In undertaking the diplomatic analysis of the case studies, we had begun with an assumption that the diplomatic elements of electronic records would be the same (or at least the fundamental elements would be similar) as those of traditional records. However, we began to realize that these elements are less explicit in electronic records, and that more of the record's identifying elements are found in its context, instead of on the face of the records, as was the case for traditional records. As a result, the diplomatic analysis often focused on what was apparently absent or inadequate in the systems that were studied when held up against the ideal record represented by the *Template,* rather than effectively identifying alternative, new, or unanticipated ways in which authenticity requirements were being met in these systems.

Each case study was reanalyzed, therefore, in order to determine which, if any, aspects of the systems examined corresponded to, or supported elements

[18] The diplomatic analysis was conducted by InterPARES researchers at the University of British Columbia; the other analyses were conducted by researchers at the University of California, Los Angeles and at the University at Albany, State University of New York.

establishing the identity and integrity of electronic records (the key concerns of authenticity), as delineated in the *Template for Analysis*. The *identity* of a record refers to how its attributes uniquely characterize it and distinguish it from other records. These attributes might include the names of the persons involved in its formation; the date(s) of creation and of transmission; any indication of the action or matter in which the record participates; any means by which a record is linked to other records participating in the same action (for example, through a classification code or other unique identifier); and indications of any attachment. The *integrity* of a record refers to its wholeness and soundness in all essential respects (that is, that the message that the record is meant to communicate in order to achieve its purpose is unaltered).

The analysis of elements of identity and integrity examined not only specific elements, but also a variety of contexts, sources, and techniques through which elements might be manifested or their purposes achieved. The case study data was coded to see whether any patterns were discernible across all case studies, or across those that seem likely to contain similar types of records. The key findings of this analysis are indicated below.

Key findings of the analysis of elements of identity and integrity:

- Authenticity in all cases is assured mainly through procedural means. The most common means identified were access privileges (including use of passwords, user IDs, and user profiles), followed by the use of audit trails and backup procedures.
- While few explicit measures are in place to ensure the authenticity of individual electronic records, record creators implicitly address this issue through the management of the electronic system as a whole.
- Authentication technologies only address the authenticity of records over space and in time, and not across time, as would be required for archival records.
- Traditional diplomatic elements such as the physical or juridical persons involved in the formation of the record are less explicitly expressed in electronic than in paper records, and are frequently implicit, inferred, or inherited from the context of the system or some other aspect of the system.

The second analytical approach examined characteristics of case studies by type of business information system. This analysis applied a model commonly used in business administration to identify types of information systems developed and used in an organization to support business processes and to fulfill the mission of the organization.[19] The model provided one way to describe the nature of systems found in an organization, and, thereby, potentially a method to help discern systems that are likely to create records, and whether those records are likely to be dispositive, probative, supporting, or narrative.

[19] Kenneth C. Laudon and Jane P. Laudon, *Management Information Systems: New Approaches to Organization and Technology* (Upper Saddle River, N.J.: Prentice Hall, 1996).

Diplomatics traditionally identifies two categories of records based on the relationship between a record and the action in which it participates. *Dispositive* records are records whose written form is required by the juridical system and which can effect an action (for example, a contract); *probative* records are records whose written form is required by the juridical system as proof that an action has taken place prior to its documentation (for example, a receipt). Recent work in contemporary archival diplomatics, notably that of Luciana Duranti, identifies two additional categories of records. *Supporting* records are records whose written form is discretionary; they are created to provide support for, and are procedurally linked to, a legally relevant action. They do not in themselves constitute the action and are not used to prove the action, but they assist in decision-making. *Narrative* records are also records whose written form is also discretionary; however they do not participate procedurally in any legally relevant action but are created as part of routine work processes.[20]

In the model used, an organization is divided into four levels [Table 3], and organizational functions are supported by six major types of systems [Table 4]. Operational level systems, such as transaction processing systems, help operational managers keep track of the organization's everyday activities. Knowledge level systems, such as office automation systems and knowledge work systems, help knowledge and data workers design products, distribute information, and manage paperwork. Management level systems, such as management information systems and decision support systems, help middle managers monitor and control business activities. Strategic level systems, such as executive support systems, help senior managers with long-term planning. The model also delineates the information inputs, processes, and outputs that serve as indicators of the type of system being examined.

This business information systems analysis examined the organizational level and information inputs, processes, and outputs associated with each case study in order to try to identify the type and nature of each system and the likelihood that it generates, or should generate, records. In recognition of the "mixed" nature of most of the systems studied, we also extended the model to identify more closely not only electronic but also paper outputs. Because stable content is considered to be an identifying characteristic of authentic records, we further categorized the status of system outputs in order to understand the degree to which they were stable [Table 5].[21]

[20] See Duranti, *Diplomatics: New Uses for an Old Science,* 67–69.

[21] According to the Authenticity Task Force's Research Methodology Statement, a fixed form "means that (1) the binary content of the record, including indicators of its documentary form, are stored in a manner that ensures it remains complete and unaltered; and (2) technology has been maintained and procedures defined and enforced to ensure that the content is presented or rendered with the same documentary form it had when it was set aside." The Statement is available on the project website.

Table 3 Levels at Which Information Systems are Implemented within Organizations[26]

1) *Operational level:* information systems monitor the elementary activities and transactions of the organization.
2) *Knowledge level:* information systems support knowledge and data workers in an organization.
3) *Management level:* information systems support the monitoring, controlling, decision-making, and administrative activities of middle managers.
4) *Strategic level:* information systems support the long-range planning activities of senior management.

Table 4 Major Organizational System Types[27]

1) *Transaction processing system (TPS):* computerized system that performs and records from daily routine transactions necessary to conduct the business; these systems serve the operational level of the organization.
2) *Knowledge work system (KWS):* information system that aids knowledge workers in the creation and integration of new knowledge in the organization.
3) *Office automation system (OAS):* computer system, such as word processing, electronic mail system, and scheduling system, that is designed to increase the productivity of data workers in the office.
4) *Management information system (MIS):* information system at the management level of an organization that serves the functions of planning, controlling, and decision making by providing routine summary and exception reports.
5) *Decision-support system (DSS):* information system at the management level of an organization that combines data and sophisticated analytical models to support semi-structured decision-making.
6) *Executive Support System (ESS):* information system at the strategic level of an organization designed to address unstructured decision making through advanced graphics and communications.

Table 5 Stability of System Output

- Fixed: Once output is created, it is immutable. If it needs to be changed, either an update must be appended, or a new version must be created.
- Transient: Output is created for temporary use only: for example, a screen display providing the results of an information query.
- Dynamic: Output is stored on the system but can be changed, updated, annotated, and overwritten.

The majority of the case studies focused on systems that function at the operational and knowledge levels within the organization, and less frequently at the management level. In comparing this analysis with the diplomatic analysis of the same case studies, we can see that those systems identified through the diplomatic analysis as containing (or that should contain) dispositive or probative records for the most part carry out at least some management as well as operational and knowledge-level functions. One could speculate, therefore, that systems addressing functions at the management level and above would be more likely to contain records and less transactional data.

[26] Taken from Laudon and Laudon, *Management Information Systems: New Approaches to Organization and Technology*.

[27] Taken from Laudon and Laudon, *Management Information Systems: New Approaches to Organization and Technology*.

Key findings of the business information systems analysis:

- The analysis indicated the complexity of the systems studied—almost no system exists independent of a wider record-keeping system, and most relate to more than one organizational level and perform a range of functions rather than conforming to one of the discrete types contained in the business model.
- Most record-keeping systems are a hybrid of electronic and paper records. This "mixed" environment must be taken into account when understanding the nature of any potential record generated by the system.
- Many of the systems studied contained primarily transactional data, and most of them generated primarily transient or dynamic output.
- Systems addressing functions at the management level and above may be more likely to contain records and less transactional data.

The third analytical approach employed was a functional analysis of case studies. As the research progressed, it became increasingly clear that understanding the nature and boundaries of electronic records required a detailed understanding of the business functions and activities of the record-keeping systems being studied. This outcome supports the findings of several other electronic records research projects, notably the Pittsburgh Project and the Indiana University Electronic Records Project, and should not, therefore be considered surprising. We selected the National Archives of Australia's *DIRKS (Designing and Implementing Recordkeeping Systems) Manual* as a robust and replicable method of functional analysis.[22] The purpose of conducting this functional analysis was to describe, unambiguously for non-archivists, and systems designers in particular, the nature of the record-keeping function performed by the system. We concluded, however, that it was not possible to render an accurate functional decomposition of each case study from the case study data that had been collected. We reached this conclusion after attempting both a narrative and graphical representation of the major functions of the systems being studied, and a break-down of the actions and transactions that support those functions, and then receiving feedback from interviewers and respondents upon the draft decompositions. The major reason for our inability to produce an accurate functional decomposition was that the *CSIP*, developed from the diplomatic perspective of analyzing individual documents, had not been designed to capture the appropriate depth of functional detail about the record-keeping system as a whole.

[22] National Archives of Australia, *Designing and Implementing Recordkeeping Systems: Manual for Commonwealth Agencies*. Available at <http://www.naa.gov.au/recordkeeping/dirks/dirksman/dirks.html> (December 4, 2002).

Key findings of the functional analysis:

- Understanding the nature and boundaries functions and activities of the record-keeping systems being studied.
- The diplomatic approach, with its focus on the individual record or document, and the functional analysis approach with its focus on business processes, can both yield valuable insights but are incompatible as complementary research approaches.

Our fourth and final approach consisted of a content analysis of transcribed case study interview data. One or our concerns throughout this research was that the understanding of the nature of electronic records and the concept of authenticity that we had built up through this research, as well as the ways in which that understanding was expressed through the terminology used in the *Case Study Interview Protocol* and other InterPARES products, would not match that of, or be understandable by, record-keepers and systems personnel. Although the case study interviews were heavily scripted to ensure some level of consistency across cases, some interviews were recorded and transcribed. Of these interviews, some contained additional discussion about the nature and functionality of the electronic recordkeeping in which the respondents were engaged. We examined selected case study transcripts to gain a closer understanding of respondent perspectives and terminology. A complete content analysis was conducted of one case study (of a DataCAD™ system of a small U.S. architectural firm) that demonstrated the value of such an approach for future research.[23] It should be noted, however, that the case studies were not originally intended to be subjected to content analysis, but were created for the purposes of grounded theory development. Had we decided *a priori* to use content analysis as a method, the interviews, or certain components of the interviews, would need to have been conducted in a more free-form or conversational manner which would have allowed respondents to expand their commentary and would have not asked the interviewees questions that used InterPARES' own terminology and rhetorical tropes.

The content analysis was guided by several specific questions:

1. To what factors besides authenticity do records creators, custodians and systems personnel give weight when addressing the permanent preservation of electronic records?
2. Currently, when the long-term authenticity of records cannot be assured, what metrics and/or heuristics are put in place, why, and by whom?

[23] For a full discussion of the analysis, see Ciaran Trace, "Applying Content Analysis to Case Study Data: A Preliminary Report," June 2001. Available at <http://www.interpares.org/reports.htm.> (December 4, 2002).

3. In what ways do records creators, custodians, and systems personnel conceptualize the nature and role of the electronic records and/or record-keeping system being studied?
4. What are the variations in language used to describe records by records creators, custodians, and systems personnel?
5. What is the relative importance of business policy, pragmatism, and institutional culture of the organization in determining the fate of electronic records and record-keeping systems?

Key findings of the content analysis:

- Content analysis appears to provide an additional means of gaining insight into transparent or implicit aspects of the nature of records and authenticity as understood in institutional settings. In particular, it may assist in developing an understanding of how organizations view their own records; the factors that influence how records are created and formatted; and how, and to what extent, electronic and paper records are similar and different.
- The most effective way to elicit rich data for content analysis is through the use of semi-structured or open-ended interviews with individuals and focus groups, thus facilitating variance of language and extended expression of opinions and examples on the part of interviewees.

Conclusion and Areas for Further Research

Electronic records are complex physical objects and highly contingent intellectual constructs. Because of these characteristics, it is often difficult for archivists, whether researching or managing electronic records, to identify a single, appropriate unit of analysis. As may be seen from the discussion in this paper, diplomatics approaches the issue with a focus on the individual record, archival science from the perspective of the record aggregate, systems analysis from that of the automated information or record-keeping system, and content analysis from the perspective of the interviewee (i.e., the records creator, custodian, or systems person). Each of these perspectives contributes to our understanding of the nature of the record and its long-term authenticity and preservation needs. Employing both the deductive and the inductive approaches allowed the Authenticity Task Force to construct a detailed profile of the complexity of contemporary electronic records and their embeddedness in their juridical-administrative, provenancial, procedural, documentary, and technological contexts.

This profile and the other findings discussed in this paper provided input for and raised substantive issues about "truth tests" for authenticity that were factored into the overall outcome of the research conducted within the

Authenticity Domain of InterPARES 1—a set of *Benchmark and Baseline Requirements for Assessing and Maintaining the Authenticity of Electronic Records* [see APPENDIX].[24] The *Benchmark Requirements* presume that establishment of the authenticity of records while they are being generated and maintained by the records creator will be based upon the *number* of requirements that have been met and the *degree* to which each has been met. The requirements are, therefore, cumulative: the higher the number of satisfied requirements, and the greater the degree to which an individual requirement has been satisfied, the stronger the presumption of authenticity. When it is not possible to establish a presumption of authenticity through the *Benchmark Requirements,* then several other methods of verification of authenticity are possible. These methods include, but are not limited to, a comparison of the records in question with copies that have been preserved elsewhere or with backup tapes; comparison of the records in question with entries in a register of incoming and outgoing records; textual analysis of the record's content; forensic analysis of aspects such as medium and script; a study of audit trails; and the testimony of a trusted third party. The *Baseline Requirements* are intended for those who are responsible for preserving records, generally the archives. These requirements apply to the maintenance of records, as well as to producing copies according to procedures that also maintain their authenticity. Unlike the *Benchmark Requirements,* all of the requirements included in the *Baseline Requirements* must be met before the preserver can attest to the authenticity of the electronic copies in its custody.

In terms of methodological outcomes, this work raised several questions that would be fruitful to examine through further research. Firstly, is it possible to develop an analytical framework integrating aspects of contemporary diplomatics and archival theory that addresses both the document and record aggregates and identifies and elucidates the role of their different contexts? Secondly, can we provide a more detailed analysis of the various contexts of the records and the ways in which the archival bond might be expressed within those contexts?[25] And finally, can we develop more finely grained instruments that could extract specific aspects of different contexts and tie them to the record in ways that establish the archival bond?

[24] These are contained in Appendix 2, "Requirements for Assessing and Maintaining the Authenticity of ElectronicRecords," of the final report of the InterPARES Project. <http://www.interpares.org> (December 4, 2002).

[25] The archival bond is defined in the InterPARES *Glossary* as "the relationship that links each record, incrementally, to the previous and subsequent ones and to all those which participate in the same activity. It is originary (i.e., it comes into existence when a record is made or received and set aside), necessary (i.e., it exists for every record), and determined (i.e., it is characterized by the purpose of the record).

Appendix: Benchmark and Baseline Requirements

Benchmark Requirements Supporting the Presumption of Authenticity of Electronic Records (Requirement Set A)

	To support a presumption of authenticity the preserver must obtain evidence that:
REQUIREMENT A.1: Expression of Record Attributes and Linkage to Record	the value of the following attributes are explicitly expressed and inextricably linked to every record. These attributes can be distinguished into categories, the first concerning the identity of records, and the second concerning the integrity of records.

A.1.a Identity of the record:

- ***A.1.a.i*** Names of the persons concurring in the formation of the record, that is:
  - name of author[1]
  - name of writer[2] (if different from the author)
  - name of originator[3] (if different from name of author or writer)
  - name of addressee[4]
- ***A.1.a.ii*** Name of action or matter
- ***A.1.a.iii*** Date(s) of creation and transmission, that is:
  - chronological date[5]
  - received date[6]
  - archival date[7]
  - transmission date(s)[8]
- ***A.1.a.iv*** Expression of archival bond[9] (e.g., classification code, file identifier)
- ***A.1.a.v*** Indication of attachments

[1] The name of the physical or juridical person having the authority and capacity to issue the record or in whose name or by whose command the record has been issued.

[2] The name of the physical or juridical person having the authority and capacity to articulate the content of the record.

[3] The name of the physical or juridical person assigned the electronic address in which the record has been generated and/or sent.

[4] The name of the physical or juridical person(s) to whom the record is directed or for whom the record is intended.

[5] The date, and possibly the time, of compilation of a record included in the record by the author or the electronic system on the author's behalf.

[6] The date, and possibly the time, when a record is received by the addressee.

[7] The date, and possibly the time, when a record is officially incorporated into the creator's records.

[8] The date and time when a record leaves the space in which it was generated.

[9] The archival bond is the relationship that links each record, incrementally, to the previous and subsequent ones and to all those participate in the same activity. It is originary (i.e., it comes into existence when a record is made or received and set aside), necessary (i.e., it exists for every record), and determined (i.e., it is characterized by the purpose of the record).

	A.1.b Integrity of the record: ***A.1.b.i*** Name of handling office[10] ***A.1.b.ii*** Name of office of primary responsibility[11] (if different from handling office) ***A.1.b.iii*** Indication of types of annotations added to the record[12] ***A.1.b.iv*** Indication of technical modifications;[13]
REQUIREMENT A.2: Access Privileges	the creator has defined and effectively implemented access privileges concerning the creation, modification, annotation, relocation, and destruction of records;
REQUIREMENT A.3: Protective Procedures: Loss and Corruption of Records	the creator has established and effectively implemented procedures to prevent, discover, and correct loss or corruption of records;
REQUIREMENT A.4: Protective Procedures: Media and Technology	the creator has established and effectively implemented procedures to guarantee the continuing identity and integrity of records against media deterioration and across technological change;
REQUIREMENT A.5: Establishment of Documentary Forms	the creator has established the documentary forms of records associated with each procedure either according to the requirements of the juridical system or those of the creator;
REQUIREMENT A.6: Authentication of Records	if authentication is required by the juridical system or the needs of the organization, the creator has established specific rules regarding which records must be authenticated, by whom, and the means of authentication;
REQUIREMENT A.7: Identification of Authoritative Record	if multiple copies of the same record exist, the creator has established procedures that identify which record is authoritative;
REQUIREMENT A.8: Removal and Transfer of Relevant Documentation	if there is a transition of records from active status to semi-active and inactive status, which involves the removal of records from the electronic system, the creator has established and effectively implemented procedures determining what documentation has to be removed and transferred to the preserver along with the records.

[10] The office (or officer) formally competent for carrying out the action to which the record relates or for the matter to which the record pertains.

[11] The office (or officer) given the formal competence for maintaining the authoritative record, that is, the record considered by the creator to be its official record.

[12] Annotations are additions made to a record after it has been completed. Therefore, they are not considered elements of the record's documentary form.

[13] Technical modifications are any changes in the digital components of the record as defined by the Preservation Task Force. Such modifications would include any changes in the way any elements of the record are digitally encoded and changes in the methods (software) applied to reproduce the record from the stored digital components; that is, any changes that might raise questions as to whether the reproduced record is the same as it would have been before the technical modification. The indication of modifications might refer to additional documentation external to the record that explains in more detail the nature of those modifications.

Baseline Requirements Supporting the Production of Authentic Copies of Electronic Records (Requirement Set B)

	The preserver should be able to demonstrate that:
REQUIREMENT B.1: Controls over Records Transfer, Maintenance, and Reproduction	the procedures and system(s) used to transfer records to the archival institution or program; maintain them; and reproduce them embody adequate and effective controls to guarantee the records' identity and integrity, and specifically that **B.1.a** Unbroken custody of the records is maintained; **B.1.b** Security and control procedures are implemented and monitored; and **B.1.c** The content of the record and any required annotations and elements of documentary form remain unchanged after reproduction.
REQUIREMENT B.2: Documentation of Reproduction Process and its Effects	the activity of reproduction has been documented, and this documentation includes **B.2.a** The date of the records' reproduction and the name of the responsible person; **B.2.b** The relationship between the records acquired from the creator and the copies produced by the preserver; **B.2.c** The impact of the reproduction process on their form, content, accessibility and use; and **B.2.d** In those cases where a copy of a record is known not to fully and faithfully reproduce the elements expressing its identity and integrity, such information has been documented by the preserver, and this documentation is readily accessible to the user;
REQUIREMENT B.3: Archival Description	the archival description of the fonds containing the electronic records includes—in addition to information about the records' juridical-administrative, provenancial, procedural, and documentary contexts—information about changes the electronic records of the creator have undergone since they were first created.