Providing Grounds for Trust II: The Findings of the Authenticity Task Force of InterPARES

HEATHER MACNEIL

RÉSUMÉ Entre 1999 et 2002, le projet InterPARES (International Research in Permanent Authentic Records in Electronic Systems) s'est penché sur les problèmes reliés à la conservation à long terme de documents électroniques authentiques. Dans le cadre du projet, c'est le groupe de travail sur l'authenticité qui fut chargé d'élaborer les exigences conceptuelles nécessaires à l'évaluation et au maintien de l'authenticité des documents électroniques. Cet article présente certains résultats de la recherche du groupe de travail, dans la suite d'un article pécédent paru dans *Archivaria* 50. Il examine les resultats de analyse de systèmes de documents électroniques actifs dans la perspective de la diplomatique archivistique contemporaine et présente la formulation définitive des exigences nécessaires à l'évaluation et au maintien de l'authenticité des documents électroniques.

ABSTRACT Between 1999 and 2002, the International Research in Permanent Authentic Records in Electronic Systems (InterPARES) Project investigated the issues associated with the long-term preservation of authentic electronic records. Within the InterPARES Project, the Authenticity Task Force was given the task of developing conceptual requirements for assessing and maintaining the authenticity of electronic records. This article presents some of the results of the research undertaken by the Task Force following up on a previous article published in *Archivaria* 50. It examines the results of analyzing the case studies of live electronic systems from the perspective of contemporary archival diplomatics and presents the final version of the requirements for assessing and maintaining the authenticity of electronic records.

This article reports on the findings of the Authenticity Task Force of the *International Research in Permanent Authentic Records in Electronic Systems (InterPARES) Project*.[1] The goal of the InterPARES Project was to formulate

1 The InterPARES Project was a three-year research initiative, which began in January 1999 and concluded in December 2002. The InterPARES research team comprised an international and multidisciplinary group of scholars, specialists, and practitioners drawn from archival studies, the humanities and social sciences, and the computer, mathematical, and chemical sciences. The project director was Luciana Duranti, an archival studies professor in the School of Library, Archival and Information Studies at the University of British Columbia. The collaborators were a consortium of eight national and multi-national research teams representing Australia, Canada, China, France, Ireland, Italy, the Netherlands, Sweden, the United King-

principles and methods for ensuring the long-term preservation of authentic electronic records. The Authenticity Task Force[2] of InterPARES was given the specific charge of identifying conceptual requirements for assessing and maintaining the authenticity of electronic records.

To fulfill this charge, the Task Force adopted two complementary analytical approaches. The first approach was a theoretical and deductive one, based on contemporary archival diplomatics. The second approach was an inductive and empirical one that employed selected case studies of live electronic systems. A preliminary report of the work accomplished by the Task Force at the project's mid-point was published in a previous article in *Archivaria*.[3] That article explained the premises underpinning the Task Force's research, examined the contemporary archival diplomatic model of an ideal electronic record developed by the Task Force, and summarized the preliminary draft of the requirements for assessing and maintaining the authenticity of electronic records. The present article provides a brief overview of the research's premises, examines the results of analyzing the case studies of live electronic systems from the perspective of contemporary archival diplomatics, and presents the final version of the requirements developed by the Task Force for assessing and maintaining the authenticity of electronic records.[4]

---

dom, and the United States. Included on the national teams were representatives from the national archives of Canada, China, France, Ireland, Italy, the Netherlands, Sweden, the United Kingdom, and the United States. InterPARES also counted among its participants a range of industries, including the pharmaceutical, chemical, biotechnology, computer software, and high technology industries, all of which were represented through the participation of the Collaborative Electronic Notebooks System Association (CENSA). Funding for the international direction of the project and for the Canadian research team was provided by the Social Sciences and Humanities Research Council of Canada. For a detailed description of the project as a whole and its findings see InterPARES Project, *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPares Project* (August 2002), available at <www.interpares.org>.

2 The members of the Authenticity Task Force were: Heather MacNeil, University of British Columbia (Chair), Chen Wei, Beijing Municipal Archives, Luciana Duranti, University of British Columbia, 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.

3 Heather MacNeil, "Providing Grounds for Trust: Developing Conceptual Requirements for the Long-Term Preservation of Authentic Electronic Records." *Archivaria* 50 (Fall 2000), pp. 52–78.

4 For a full account of the work of the Authenticity Task Force, including the original research questions, the collection and analysis of case study data, as well as the Task Force's findings and recommendations, see Authenticity Task Force, "Establishing and Maintaining Trust in Electronic Records: The Final Report of the Authenticity Task Force," in *The Long Term Preservation of Authentic Electronic Records*, pp. 1–33 and appendices 1 and 2. Available at <www.interpares.org>.

An authentic record is one that can be proven to be (a) what it claims to be and (b) free of falsification or inappropriate modification. To assess the authenticity of an electronic record and to maintain it over time, the preserver must be able to establish its *identity* and demonstrate its *integrity.* The *identity* of a record refers to the attributes of a record that uniquely characterize it and distinguish it from other records. The *integrity* of a record refers to its wholeness and soundness: a record has integrity when it is complete and uncorrupted in all its essential respects. Accordingly, in developing the requirements for authenticity, the Authenticity Task Force focussed on identifying the most effective means of protecting the identity and integrity of electronic records over time and across technologies.

To adhere to that focus, the Task Force found it necessary to draw a distinction between *authenticity* and *authentication*. In its simplest terms, authenticate means "to prove or serve to prove the authenticity of."[5] In legal terms, authentication is "the act or mode of giving authority or legal authenticity to a statute, record, or other written instrument, or a certified copy thereof, so as to render it legally admissible in evidence.… An attestation made by a proper officer by which he certifies that a record is in due form of law, and that the person who certifies it is the officer appointed to do so."[6] The diplomatic notion of authentication is consistent with the legal meaning of the word. In diplomatic terms, authentication "is the legal recognition that a signature is affixed by and belongs to the person whose name it expresses, that a document is what it purports to be, or that a copy conforms to an original."[7]

The legal and diplomatic understanding of authentication shaped the Authenticity Task Force's definition of the term, which is, "a declaration of a record's authenticity at a specific point in time by a juridical person entrusted with the authority to make such declaration."[8] Authentication typically takes the form of an authoritative statement (which may be in the form of words or symbols) that is added to or inserted in the record attesting that the record is authentic. Digital signatures are a specific example of an authentication technology that has been developed to address the need for secure electronic communication across open networks such as the Internet. While such signatures – which identify the sender of a data object and verify that it has not been altered in transmission – can support the authentication of electronic records,[9]

5 Merriam-Webster Online Dictionary, <www.m-w.com/cgi-bin/dictionary>, s.v. "authenticate."

6 *Black's Law Dictionary*, 6th ed. (St. Paul, Minn., 1990), s.v. "authentication."

7 Luciana Duranti, *Diplomatics: New Uses for an Old Science* (Lanham, Maryland, 1998), p. 139.

8 "The InterPARES Glossary," *The Long Term Preservation of Authentic Electronic Records*, s.v. "authentication."

9 In its diplomatic analysis of the elements of an electronic record, the Authenticity Task Force treated the digital signature as functionally analogous to the medieval sovereign's seal. See MacNeil, "Providing Grounds for Trust," pp. 60–63.

they are not sufficient to assess and maintain their authenticity over time and across technologies. Further research is needed to determine the specific impact of digital signatures on the long-term preservation of authentic electronic records.[10]

Preserving a record's authenticity is predicated on its endurance and stability over time. Preserving the authenticity of a record in the digital world is complicated by the fact that there are no stable and enduring physical objects in that world. The Preservation Task Force of InterPARES found that:

> Empirically, it is not possible to preserve an electronic record: it is only possible to preserve the ability to reproduce the record. That is because it is not possible to store an electronic record in the documentary form in which it is capable of serving as a record. There is, inevitably, a substantial difference between the digital representation of the record in storage and the form in which it is presented for use. It is always necessary to use some software to translate the stored digital bits into the documentary form of the record. This entails an inevitable risk that, regardless of how well the digital data were protected in storage, the record may be inappropriately altered when the stored bits are retrieved and presented for use as a record.[11]

This finding requires that we re-think our reliance on the notion of an unbroken chain of custody as a guarantor of record authenticity. With non-digital forms of records, continuous custody has been considered sufficient grounds for asserting their authenticity. However, as the Preservation Task Force points out: "Given that the storage and retrieval processes for electronic records inevitably entail physical and representational transformations, the traditional concept of an unbroken chain of custody needs to be expanded to encompass the processes that are necessary to ensure that an electronic record is transmitted over time without inappropriate alteration."[12] The Preservation Task Force calls this expanded principle the unbroken "chain of preservation." The principle asserts that: "the entire process of committing an electronic record to storage, maintaining it in storage, retrieving, and presenting it must adequately preserve all its essential attributes in order to support a credible claim that the retrieved electronic record is authentic."[13]

10 The impact of (1) digital signature technologies and (2) the infrastructure supporting them on the management of authentic electronic records over the long term has been identified by the Authenticity Task Force as an area requiring further research. See Authenticity Task Force, "Establishing and Maintaining Trust," p. 33.

11 Preservation Task Force, "Trusting to Time: Preserving Authentic Records in the Long Term: Preservation Task Force Report," in *The Long Term Preservation of Authentic Electronic Records*, p. 5. Available at <www.interpares.org>.

12 Preservation Task Force, "How to Preserve Authentic Electronic Records," in *The Long Term Preservation of Authentic Electronic Records*, Appendix 6, p. 8.

13 Ibid.

Authenticity is particularly at risk when records are transmitted across space (that is, when they are sent between persons, systems, or applications) or time (that is, when they are stored offline, or when the hardware or software used to process, communicate, or maintain them is upgraded or replaced).[14] Therefore, in the case of records maintained in electronic systems, the traditional presumption of authenticity must be supported by evidence that a record is what it claims to be and has not been inappropriately modified or corrupted. The requirements for assessing and maintaining the authenticity of electronic records developed by the Authenticity Task Force concern that evidence.

The theoretical framework that shaped the work of the Authenticity Task Force was provided by *contemporary archival diplomatics*, a hybrid discourse consisting of an adaptation of traditional diplomatic concepts and methods to contemporary record-keeping environments, and an integration of those concepts and methods with those drawn from archival science.[15]

Viewed from that perspective, an electronic record, like its non-electronic

14 As the Preservation Task Force explains: "A technological boundary exists between any two states of a system or of interoperating systems when the transition from one state to another does, or can, entail significant changes in the attributes or methods of a digital object. For records, significant changes are those that affect identity or integrity. Technological boundaries exist at macro and micro levels. Macro level boundaries occur at the interfaces between systems, subsystems or applications, such as during system, media, or data format migrations or in transfers between the "live" systems in which the records are created, and other applications in which they are transmitted over space or stored over time. Micro level boundaries occur when a record is decomposed into separate digital components or is reconstituted from its components, and when different methods are invoked to process distinct components. Transitions from storage representation to presentation for use can involve both macro and micro boundaries.

Preservation control is critical in transitions across technological boundaries. Preservation control consists of actions, conditions, and constraints designed to ensure the preservation of records and their continued authenticity. While preservation controls during maintenance of the records in storage must be adequate and effective, the risks of corruption or loss of records are more frequent and complex during transitions across technological boundaries. Thus preservation controls can be divided into two types: *systemic controls* are those that ensure records remain unchanged over time within a given system or subsystem; *dynamic controls* are those that ensure records remain authentic across technological boundaries." Ibid., pp. 5–6.

15 The concepts and methods of traditional diplomatic analysis and their adaptation to contemporary record-keeping practices were first introduced to North American archivists by Luciana Duranti in a series of articles written between 1989 and 1992. See Luciana Duranti, "Diplomatics: New Uses for An Old Science," *Archivaria* 28 (Summer 1989), pp. 7–27; "Diplomatics ... (Part II)," *Archivaria* 29 (Winter 1989–90), pp. 4–17; " Diplomatics ... (Part III)," *Archivaria* 30 (Summer 1990), pp. 4–20; "Diplomatics ... (Part IV)," Archivaria 31 (Winter 1990–91), pp. 10–25; "Diplomatics ... (Part V)," *Archivaria* 32 (Summer 1991), pp. 6–24; "Diplomatics ... (Part VI)," *Archivaria* 33 (Winter 1991–92), pp. 6–24. The articles were subsequently published in a single volume as Duranti, *Diplomatics: New Uses for an Old Science*. For a detailed examination of the evolution of contemporary archival diplomatics see Heather MacNeil, *Trusting Records: Legal, Historical and Diplomatic Perspectives* (Dordrecht, 2000), pp. 86–112.

counterpart, is a complex of elements and their relationships. It possesses a number of characteristics, including a *fixed documentary form*, a *stable content*, an *archival bond with other records*, and an identifiable juridical-administrative, provenancial, administrative, procedural, documentary and technological *context*. It participates in or supports an *action*, and at least three *persons* (author, writer, and addressee) are involved in its creation.[16]

These characteristics manifest themselves both explicitly and implicitly through a range of elements that are found both inside the record and outside of it as part of the larger documentary and administrative framework in which the record is created and maintained. For example, the archival bond may manifest itself as a classification code or other record identifier that appears on the face of the record or in its profile; the name of the author may take the form of letterhead or an electronic signature; the record's documentary context may reveal itself in a classification scheme, and so on. The purpose served by these elements necessarily will vary, depending on their specific form of manifestation. For example, the identification of the name of the author that appears as letterhead serves the purpose of identifying the record's provenancial or administrative context; when the name of the author takes the form of an electronic signature, it serves the purpose of attesting to the validity of the record itself, or its content, or both.

In traditional diplomatic theory, the elements most relevant to a consideration of a record's authenticity are typically found in the record's documentary form and in annotations. The documentary form is the primary means by which the content of a record, its immediate administrative and documentary context, and its authority are communicated. In contemporary record-keeping environments the elements of documentary form might include discursive elements such as the name of the author and addressee, the date, the description of the action, and the attestation; as well as non-discursive elements such as its mode of representation (e.g., textual, graphic, moving image), specific presentation features (e.g., deliberately employed colors, special layouts, hyperlinks, sample rates of sound files, resolution of image files), as well as seals and spe-

16 The list of identifying characteristics of an electronic record was formulated by the Authenticity Task Force early in its research. It represents a substantially revised version of the list of components of an electronic record identified in a previous research project carried out between 1994 and 1997 by archival researchers at the University of British Columbia. The goal of that project, which was entitled "The Preservation of the Integrity of Electronic Records" (commonly known as "The UBC Project"), was to identify and define conceptually the nature of an electronic record and the conditions necessary to ensure its reliability and authenticity during its active and semi-active life, based on the concepts and methods of diplomatics and archival science. 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), pp. 46–67. In the UBC Project, the list of components of an electronic record included a *medium*, a *content*, *a physical and intellectual form*, an *action*, *persons*, an *archival bond*, and a *context*.

cial signs (e.g., digital signatures and watermarks). Annotations are additions made to the record in the course of its execution, handling, and management. Examples include the record's date of transmission and receipt, its classification code, registration number, version numbers, and written comments made by persons handling the matter to which the record relates.

The Task Force hypothesized that an understanding of the ways in which the characteristics of a record identified by diplomatics manifest themselves in an electronic environment, and the specific role played by the individual elements of an electronic record in asserting that record's identity and its integrity, would provide a logical foundation for the formulation of conceptual requirements for assessing and maintaining authenticity. To test that hypothesis, the Task Force developed a *Template for Analysis*, which was a model of an ideal electronic record based on all its possible known elements. The elements of an electronic record included in the *Template for Analysis* were organized into four main categories: (1) *documentary form,* which was subdivided into *intrinsic elements of form* and *extrinsic elements of form*, (2) *annotations,* (3) *context*, and (4) *medium*.[17] The *Template* decomposed each of these categories into its constituent elements, defined each element, explained its purpose, and indicated whether, and to what extent, that element was instrumental in establishing a record's authenticity.[18]

The *Template* subsequently provided the basis for diplomatic analyses of a wide range of live electronic systems that contained, generated, or had the potential to create electronic records.[19] The analyses were carried out through four rounds of case studies.[20] Our expectation was that this analysis would

17 At the beginning of the research, medium was viewed as a distinct part of the record itself. At the end of the research, however, the Task Force concluded that medium should be considered part of the record's context (specifically, its technological context). See "Establishing and Maintaining Trust," pp. 6–7.

18 It is important to point out that certain elements may be found in more than one category. For example, a digital signature is considered both an extrinsic element of form and an annotation.

19 The basis for the selection of case studies is explained in the final report of the Authenticity Task Force. See "Establishing and Maintaining Trust," pp. 8–9. Overviews of case studies conducted by research collaborators at the National Archives of Canada are available on the project Web site at <www.interpares.org/reports>. The diplomatic analyses of the systems were carried out by research assistants at the University of British Columbia. The research assistants were: Lisa Beitel, Robert Edwards, Anna Gibson, Prisca Giordani, Elaine Goh, Erica Hernandez, Robyn Hulley, Ian McAndrew, April Miller, Claire Vesseirre, Lara Wilson, and Jane Zhang.

20 Although contemporary archival diplomatics was the primary type of analysis carried out on the electronic systems, it was by no means the only one. Four other types of analysis were carried out by InterPARES researchers at the University of California, Los Angeles and the University of Albany. These included (1) an analysis of how and to what degree the identity and integrity of electronic records were supported within and across case studies; (2) an analysis of the characteristics of case studies by type of information system; (3) a functional analysis of case studies; and (4) a narrative analysis of transcribed case study interview data. See "Establishing and Maintaining Trust," pp. 16–20.

enable us to identify general requirements for authenticity and provide the foundation for the development of a typology of electronic records based on authenticity requirements for specific types of records.

In analyzing the live systems, we were specifically concerned with (1) establishing the status of the digital entities contained within them as records and (2) identifying the elements of such records specifically associated with identity and integrity. With respect to (1), we found that a surprisingly large number of the systems examined in early case study rounds did not appear to contain records when measured against the evaluation criteria established by contemporary archival diplomatics.[21] This was due largely to the fact that many of those systems were dynamic; as a result, the digital entities contained within them tended to lack a fixed documentary form or a stable content. With respect to (2), we had expected that elements falling within the categories of *documentary form* and *annotations* would play key roles in establishing the identity and demonstrating the integrity of electronic records. In fact, in the case studies analyzed, it was often difficult to determine the significance of the absence or presence of specific elements of documentary form or annotations to a consideration of a record's authenticity. The Task Force found that the determination of documentary forms in general and the establishment of required elements of form and annotations in particular were deeply embedded within specific institutional and procedural contexts and were resistant to any easy generalizations. As a consequence, the Task Force's efforts to construct a typology that would provide a meaningful differentiation and specification of requirements for authenticity according to types of records failed. In the end, we were simply unable to establish a meaningful correlation between authenticity and the presence or absence of specific documentary elements or annotations capable of generalization into a single, comprehensive typology.[22]

Our experience with analyzing electronic systems from the perspective of contemporary archival diplomatics taught us much about the limits of diplomatic analysis in general and the limits of the diplomatic model of an ideal electronic record (as embodied in the *Template for Analysis*) in particular. As

21 It is important to emphasize the qualification of this finding implicit in the phrase "did not *appear* to contain records." The qualification is necessary because in some cases, the information provided about the system was not sufficient to permit a detailed analysis of the system and the entities contained within it. For a discussion of the challenges faced by the UBC research assistants in analyzing the electronic systems in diplomatic terms see "Establishing and Maintaining Trust," pp. 11–14. For a more general discussion of the limits of the case study design and instrumentation, see Ibid., pp. 25–27.

22 For a more detailed discussion of the Task Force's efforts to construct a typology, see "Establishing and Maintaining Trust," pp. 14–16. A more effective approach might be to construct specific typologies based upon the functions and procedures of individual creators. Such typologies might be translatable into similar settings, but they would still be limited because each creator works within a specific juridical context and may interpret and implement its functions and procedures differently.

it is currently articulated, contemporary diplomatics remains rooted in a very traditional conception of what a record is. Its capacity to extend the range of understanding about the nature of different kinds of electronic systems and the variety of entities contained within them is thereby limited. While it is quite effective in decomposing electronic systems containing digital objects that behave like traditional records, i.e., in systems in which the digital objects are fixed and circumscribable, it is less helpful in decomposing electronic systems containing digital objects that behave differently, i.e., in systems in which the entities are fluid and less easy to circumscribe.

To increase the utility of diplomatic analysis as a tool for understanding and analyzing diverse electronic systems, a reorientation of its concepts and principles is necessary, one that will accommodate a broader interpretation of the characteristics of electronic records and the manner in which they manifest themselves in different electronic environments. In specific terms this might mean focussing less attention on establishing whether the record is complete, stable, and unchangeable, and more attention on determining whether and to what extent the system is capable of tracking changes and how that tracking function might be managed over time. An implication of this reorientation is that, inevitably, we will be forced to make difficult decisions about the nature and extent of the changes that will and will not be captured and preserved over time. While it is neither feasible nor desirable to capture and preserve every change, it is essential that we provide logical and defensible reasons for the changes we choose to include and exclude.

An increased attention to the characteristics and behaviours of fluid systems does not imply an abandonment of fixity as a desirable characteristic of electronic systems. One of the things the diplomatic analysis highlighted was the extent to which electronic systems are still being designed to manage data rather than records. This seems to be the case even when the purpose for which the system is designed would appear to require the creation and maintenance of fixed records rather than fluid data. What is needed is a deeper analysis of the nature and purpose of different kinds of electronic systems that would enable us to specify the degree of fixity and stability necessary to protect the authenticity of certain types of records; and to stipulate, in the absence of fixity, alternative means for protecting it.

The limits of the diplomatic model of an ideal electronic record in particular may be attributed to two factors. The first factor is that the model was built on the premises of *general diplomatics*. *General diplomatics* seeks to decontextualize records, to eliminate their particularities and anomalies in the interest of identifying the common, shared elements of records that cut across juridical, provenancial, and technological boundaries. The case studies of electronic systems revealed that we are living in an era that is analogous to the age of medieval manuscripts where documentary variation was the norm rather than the exception. Our frustrated efforts to impose general diplomatics on this

reality may be read, in retrospect, as a cautionary tale about the dangers of premature universalization. Given the variety and complexity of current electronic systems, a more useful approach might be to adapt the approach of *special diplomatics*, which, traditionally, has focussed on analyzing individual chanceries and specific juridical systems. For electronic records, this means beginning with an analysis of the various features of individual electronic systems and record-keeping environments in their own terms, with all their particularities and anomalies; and, on the basis of that analysis begin to build a more general framework. In this way we can strike a more equitable balance between ideal and local features of electronic records.

Recognizing the need to accommodate local variation does not, however, invalidate the ideal elements defined by contemporary archival diplomatics. While they are incomplete, the elements are far from irrelevant to a consideration of authenticity and provide a necessary context of authority within which variation may be accommodated.[23] The case studies revealed little consistency in the way the attributes that specifically establish the identity of a record (e.g., the names of the author and addressee, the indication of the action or matter, the manifestation of the bond linking the record to others participating in the same action) are captured and expressed from one electronic system to another. In many cases, certain attributes (for example, the expression of the archival bond) were not captured at all. This finding underlines the need to make certain of those elements explicit to ensure that knowledge of key indicators of identity is not lost when the records are removed from the specific electronic system and record-keeping environment in which they have been created and actively used.

The second limiting factor of the diplomatic model of an ideal electronic record is that it lacks a sufficiently detailed vocabulary for describing and analyzing the various contexts in which electronic records are created, maintained, and used. The elements of context identified in the *Template for Analysis* corresponded to a hierarchy of frameworks ranging from the general to the specific. They included the record's *juridical-administrative context* (i.e., the legal and organizational system in which the record creator is situated); its *provenancial context* (i.e., the mandate, structure, and functions of the record creator); its *procedural context* (i.e., the business procedure in the course of which the record is created); its *documentary context* (i.e., the broader aggregation to which the record belongs and its internal structure); and its *technological context* (i.e., the technological environment surrounding the record). A detailed knowledge of these elements is critical to an understanding of the business pro-

23 The notion of variation within a context of authority is discussed in Philip E. Doss, "Traditional Theory and Innovative Practice: The Electronic Editor as Poststructuralist Reader," in Richard J. Finneran (ed.), *The Literary Text in the Digital Age* (Ann Arbor, 1996), p. 221. Doss is speaking in the context of creating hyperlinked scholarly editions of literary texts but there are obvious parallels with other kinds of electronic records.

cesses in the course of which electronic records are created, maintained, and used, the types of records generated from these processes, and the connection between and among those processes, the electronic system that supports them, and the creator's broader functions and mandate.

The case studies revealed that the elements relating to context, in particular to procedural and technological context, were most relevant to an understanding of the electronic record-keeping environment and appeared to provide the main grounds on which creators based their presumption of the records' authenticity. However, while the elements of context were identified in the *Template*, none of them was sufficiently decomposed to permit an in-depth analysis. For example, in several case studies, audit trails, which are considered part of the records' technological context, were identified by the creator as a significant means of ensuring the authenticity of electronic records. In the *Template*, technological context was decomposed into five sub-elements: *hardware* (the storage, microprocessor, network, peripheral devices, and architecture); *software* (the operating system, system software, network software, and application software); *data* (the file structure and file format); *system models* (i.e., abstract representations of the entities, activities, and/or concepts in the system as well as their attributes, characteristics, and the functional relationship between them); and *system administration* (i.e., the set of procedures that ensure correct, secure, reliable, and persistent operation of the system).

System administration covers a broad range of procedures, including the maintenance of audit trails, but because we did not decompose this sub-element any further in order to identify and elaborate the specific procedures falling within system administration, we were only able to obtain fairly general information about those procedures. With respect to audit trails, the data collected in individual case studies left unanswered at least two important questions concerning the way in which an audit trail functioned in a particular environment: Firstly, what actions taken on an electronic record are recorded and stored in the audit trail?[24] Secondly, what types of information are captured about each action?[25] Because we were unable to answer these questions in any definitive way, it was difficult to assess the extent to which an audit trail supported the creator's presumption of authenticity in particular cases.[26]

24 For example, is the audit trail unalterable? Does it record the date and time of capture of all electronic records? Does it record any changes made to the metadata associated with records? Does it record the date and time of creation, amendment, and deletion of metadata? Does it record any changes made to the access privileges affecting a record? Does it record any deletion or destruction action on an electronic record?

25 For example, does the audit trail identify the individual initiating and carrying out the action? Does it capture the date and time of the event?

26 The lack of decomposition of the element *procedural context* also hindered efforts by InterPARES researchers at UCLA to conduct a functional analysis of each case study. See Authenticity Task Force, "Establishing and Maintaining Trust," p. 19.

Our failure to decompose the elements of context in sufficient depth also hindered our ability to understand and analyze the specific nature of record aggregations in electronic systems. One significant distinction between a traditional diplomatic approach to analyzing records and a contemporary archival diplomatic approach is that, while the traditional approach focuses exclusively on individual records, the contemporary approach takes into account record aggregations (e.g., files, series, fonds). This distinction, however, was not as evident as it should have been in the model of an ideal electronic record. The majority of elements included in the *Template* fell into the category of *documentary form*, meaning that they were only relevant at the level of individual records. While this did not pose a problem when we were examining electronic systems containing homogenous aggregations of records (i.e., systems containing records that all share the same documentary form), many of the systems examined in the case studies contained heterogeneous aggregations of records (e.g., systems containing records that have a variety of documentary forms). In the *Template*, *documentary context* was the element designed to capture information about record aggregations. However, like the other elements of context, it was not decomposed sufficiently to permit a detailed analysis of the various kinds of aggregations found in electronic systems.

At first glance, the impoverished representation of context in the *Template* seems a surprising oversight, given the central importance archivists attach to it. On reflection, however, it is likely that the very taken-for-grantedness of the centrality of context to an understanding of records blinded us to the fact that the "elements" we identified were little more than general categories, each of which required decomposition into more meaningful units of analysis. What we failed to do, in other words, was to translate our implicit understanding of the nature and complexity of context into explicit terms of reference by naming and localizing its various aspects.

Although contemporary archival diplomatics did not prove in the end to be as explanatory or predictive a model for analyzing electronic records as we might have wished, it was, nevertheless, enormously productive because it inspired argument and debate and opened up new lines of inquiry. Some of the new lines of inquiry that emerged in the course of our research are:

1. Is it possible to develop an analytical framework that addresses both the document and record aggregates and identifies and elucidates the role of the different contexts of the records in relation to both individual records and record aggregates?

2. Can we provide a more detailed analysis of the various contexts in which records are created, maintained, and used, and the ways in which the archival bond might be expressed within those contexts? Can we develop more

finely grained instruments that could extract specific aspects of different contexts and tie them more closely to the records?

3. Is it possible to develop meaningful typologies of records of specific creators or specific functions and procedures?[27]

Moreover, we should not underestimate the critical importance of failure in any research activity.[28] Understanding and rigorously documenting what did not work and why are as critical to the advancement of knowledge as understanding and documenting what did work.

Although we failed to develop a comprehensive typology of electronic records, we were successful in developing general requirements for authenticity. A preliminary set of those requirements was drafted in October 2000. The requirements were identified on the basis of data gathered in the first two rounds of case studies concerning the methods used by record creators to support their presumption of the authenticity of their electronic records. The analyses of case studies revealed that record creators tend to rely primarily on generic technological and procedural controls to ensure the authenticity of their records and to treat authenticity as part of the management of the electronic system as a whole rather than as part of the management of individual records within the system.[29] The commonest means of protecting the *integrity*

27 Authenticity Task Force, "Establishing and Maintaining Trust," pp. 32–33.

28 The importance of failure is staunchly defended by John Unsworth who discusses it in the context of evaluating electronic scholarly research projects. See John Unsworth, "The Importance of Failure," *Journal of Electronic Publishing* 3, no. 2 (December 1997), available online at <www.press.umich.edu/jep/03–02/unsworth.html>. In making his case, Unsworth draws on two complementary theses articulated by the scientific philosopher Karl Popper. The first thesis states: "We know a great deal. And we know not only many details of doubtful intellectual interest, but also things which are of considerable practical significance and, what is even more important, which provide us with deep theoretical insight, and with a surprising understanding of the world." This first thesis is qualified by a second one, reminding us that: "Our ignorance is sobering and boundless. …With each step forward, with each problem which we solve, we not only discover new and unsolved problems, but we also discover that where we believed we were standing on firm and safe ground, all things are, in truth, insecure and in a state of flux." For Unsworth, these two theses eloquently express, "the importance – the utility – of what we do know and, on the other hand, the ephemeral, contingent, transitional character of that knowledge – and, therefore, the need for experiment, the indispensability of mistakes, and the necessity of recognizing, documenting, and analyzing our failures." Ibid, p. 2.

29 Case study interviewers reported that many record creators did not appear overly concerned about the authenticity of records contained within their electronic systems and seemed confident that generic technological controls over those systems were sufficient to protect the authenticity of the records contained within them. In many cases, given the paucity of controls identified, such confidence was clearly misplaced. It may also be the case, however, that the record creators had a different understanding of the meaning of authenticity and the methods necessary to protect it. For an empirical study of how practioners in records and information management understand the concept of authenticity, see Eun G. Park, "Understanding 'Authenticity' in Records and Information Management: Analyzing Practitioner Constructs," *American Archivist* 64 (Fall/Winter 2001), pp. 270–91.

of records in the case studies examined were access privileges – including passwords, user identifications, and user profiles – followed by various procedures that either prohibit or discourage modification of records once they are considered complete, the use of audit trails, and system backup procedures. The commonest means of establishing the *identity* of records were classification codes, the linking of related electronic and non-electronic records, and record profiling. On the whole, record creators appeared to be more concerned with protecting the integrity of records than with establishing their identity.[30]

The emphasis record creators place on controls aimed at protecting the integrity of the electronic system is consistent with recent amendments to the *Canada Evidence Act* dealing with the application of the best evidence rule to electronic records. The function of the best evidence rule is to ensure the integrity of records admitted in litigation. The application of that rule in a traditional record-keeping environment requires that the proponent of evidence produces, whenever possible, the original document since alterations are more likely to be detected on the original. The amendments to the rule included in Part 3 of the *Personal Information and Electronic Documents Act*[31] provide an alternative means of addressing the need to expose the particular vulnerability of electronic documents to undetectable change. They do so by shifting the focus of the rule from a dependence on proof of the integrity of the electronic document to a dependence "on proof of the integrity of the electronic documents system by or in which the electronic document was stored."[32] The principle that the integrity of an electronic records system as a whole, including the procedures used to maintain the system, creates a presumption of the integrity of the records within it is an important one and is implicit in many of the requirements developed by the Authenticity Task Force.[33]

The data compiled from the case studies were then compared to the recommended methods for maintaining authentic records identified in a previous

30 The analysis of the final two rounds of case studies reinforced this initial finding.

31 *Personal Information Protection and Electronic Documents Act*, R.S.C. 2000, C. 5, p. 3, sec. 56, amend. sec. 31.2, 31.3. The amendments to the *Act* incorporate recommendations made by the Uniform Law Conference of Canada (ULCC). For further discussion of the amendment to the best evidence rule see Uniform Law Conference of Canada, "Uniform Electronic Evidence Act Consultation Paper," March 1997, para. 3, at <www.law.ualberta.ca/alri/ulc/>, (March 1997), para. 24–26, 27.

32 An electronic documents system "includes a computer system or other similar device by or in which data is recorded or stored, and any procedures related to the recording or storage of electronic documents." Ibid., amend. sec. 31.8.

33 In the final version of the requirements for assessing and maintaining authenticity, the principle is embodied specifically in the benchmark requirements, which identify the salient features of a trusted record-keeping system. Further investigation is needed to determine whether this principle, which is well accepted in common law jurisdictions, is equally acceptable in civil law jurisdictions.

research project carried out at the University of British Columbia (commonly known as the UBC project).[34] On the basis of this comparison, the research assistants identified additional methods for supporting and strengthening such presumption, based on the ones recommended in the UBC project. These additional methods included: instituting procedures for the systematic removal of records from the live electronic system for preservation purposes; instituting procedures to prevent loss or corruption of records due to unauthorized alteration, technological obsolescence, or media fragility; establishing rules for transmitting and copying electronic records; and instituting procedures for profiling records to capture essential identifying metadata (e.g., names of author and addressee, date, indication of action or matter, version number). A preliminary draft of the requirements, identifying the most effective means of establishing the identity and demonstrating the integrity[35] of electronic records, emerged from this process.

In April 2001, representatives of the three main InterPARES task forces – the Authenticity, Appraisal, and Preservation task forces – met to review the draft requirements in light of new case study data and to reconcile them with the analytical work being carried out by the Appraisal and Preservation Task Forces.[36] During that meeting, the draft requirements were substantially revised. The revised draft was submitted to the InterPARES investigators and collaborators for their comments after which it was posted to the InterPARES Project Web site for comment from the archival community. *Requirements for*

34 For a brief description of that project see footnote 16 above. The full set of procedural rules for creating and maintaining reliable and authentic electronic records during their active and semi-active life may be found on the UBC Project Web site at <www.slais.ubc.ca/duranti>.

35 It is necessary to explain the interpretation of integrity adopted by the Authenticity Task Force in the course of developing the requirements. As mentioned at the beginning of this article, the integrity of a record refers to its wholeness and soundness: a record has integrity when it is complete and uncorrupted in all its essential respects. This does not mean that the record must be precisely the same as it was when first created for its integrity to exist and be demonstrated. When we refer to an electronic record, we consider it essentially complete and uncorrupted if the message that it is meant to communicate in order to achieve its purpose is unaltered. This implies that its physical integrity may be compromised, provided that the articulation of the content and any required annotations and elements of documentary form (meaning required by the creator) remain the same. This interpretation of integrity emphasizes preserving the authority of the record as an instrument rather than preserving its "purity" (or what some might call its original "look and feel") as a text. It is understood of course that the two may coincide with certain types of electronic records, i.e., that its authority as an instrument may depend on preserving its purity as a text.

36 The main analytical work of both the Preservation and Appraisal Task Forces consisted of the preparation of a model of the activities involved in the appraisal and preservation of authentic electronic records. The task forces viewed such model as the principal means of identifying the various theoretical and methodological questions that arise in the course of appraising and preserving authentic electronic records. The final reports of the Appraisal and Preservation Task Forces are available on the InterPARES Web site at <www.interpares.org>.

Assessing and Maintaining the Authenticity of Electronic Records is the final product of that process.[37]

The requirements are directed toward the preserver of electronic records, i.e., "the juridical person whose primary responsibility is the long-term preservation of authentic records."[38] Such person might be an archival institution, such as a national or provincial archives, or an office of an organization, such as the archives division of churches, businesses, and universities. Although the requirements are directed toward the preserver, they have obvious relevance to record creators for three reasons: firstly, the requirements identify criteria for determining whether a creator's electronic records may be presumed authentic; secondly, if we consider the principle of an unbroken chain of preservation enunciated by the Preservation Task Force of InterPARES, it is clear that the preservation of authentic electronic records is a responsibility shared by record creators and preservers and, in many cases, the creator and preserver will be the same person; and, thirdly, record creators that are obliged to retain electronic records for many decades for business purposes must concern themselves with maintaining their authenticity over the long-term.

The first set of requirements – the benchmark requirements – deal with the assessment of authenticity. They establish whether and to what extent the records have been maintained by the creator using technologies and administrative procedures that either ensure their authenticity or at least minimize risks of change from the time the records were first set aside to the point at which they are subsequently accessed. In other words, they define the evidence that demonstrates how the record creator established and maintained the chain of preservation while the records remained in its custody.

The second set of requirements – the baseline requirements – deal with the maintenance of authenticity. After a body of electronic records has been transferred from the creator to the preserver, their authenticity needs to be maintained over the long term. To do so, the preserver must manage those records in accordance with procedures that ensure their continuing authenticity and produce copies of those records in a manner that ensures their authenticity is not compromised by the reproduction process. In other words, the baseline requirements articulate what the preserver must do to ensure that the chain of preservation remains unbroken from the moment the records are transferred to the archival institution or program. Both the benchmark and baseline requirements are based on the notion of trust in record-keeping and record preservation from the moment of the records' creation. The standard of trust to which they aspire is measured in terms of circumstantial probability rather than certainty.

37 *Requirements for Assessing and Maintaining the Authenticity of Electronic Records* is included as an appendix to this article.

38 "The InterPARES Glossary," *The Long Term Preservation of Authentic Electronic Records*, s.v. "preserver."

The benchmark requirements focus on active and semi-active records and enumerate the salient characteristics of *a trusted record-keeping system*. A trusted record-keeping system is the whole of the rules that control the creation, maintenance, and use of the creator's records, and that support a presumption of the authenticity of the records within the system. A presumption of authenticity 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.

The baseline requirements focus on inactive records and enumerate the procedures necessary to enable record preservers to attest to the authenticity of electronic records after they have been transferred to their custody. The requirements are predicated on the role of the preserver as a *trusted custodian*. To be considered a trusted custodian, the preserver must demonstrate that it has no reason to alter the preserved records, or to allow others to alter them, and that it is capable of implementing procedures that ensure that any loss or change to records over time is avoided or at least minimized. Unlike the benchmark requirements, all of the baseline requirements must be met before the preserver can attest to the authenticity of the electronic records in its custody.

To place the requirements in context, the Authenticity Task Force conducted comparative analyses of the benchmark and baseline requirements against three prominent records management standards: the International Organization for Standardization's (ISO) *Draft International Standard on Records Management*, the U.S. Department of Defense's (DoD) *5015.2 Records Management Standard*, and the European Commission's *Model Requirements Specification* (MoReq).[39] The ISO and MoReq standards contain provisions that can be considered as counterparts to the individual benchmark requirements, while the DoD standard identifies provisions that function parallel to the stipulations contained in both the benchmark and the baseline requirements. Generally speaking, the comparative analyses revealed that, although they are sometimes expressed differently, the provisions of the three standards examined demonstrate a substantial degree of consistency with the benchmark and baseline

39 International Organization for Standardization, Technical Committee ISO/TC 46 Information and Documentation, Subcommittee 11, Archives/Records Management, *International Standards Organization Draft International Standard (ISO/DIS 15489) Information and Documentation – Records Management* (Geneva, 2000); United States, Department of Defense, Assistant Secretary of Defense for Command, Control, Communications, and Intelligence Design Criteria, *Standard for Electronic Records Management Software Applications* (DoD 5015.2-STD) (Washington, 2001); European Commission, *Requirements for the Management of Electronic Records* (MoReq Specification), prepared by Cornwell Affiliates plc. (Bruxelles-Luxembourg, 2001).

requirements, meaning that for most of the requirements we were able to identify a counterpart provision in one or more of the three standards.[40]

The benchmark and baseline requirements are also consistent with the standards of professional conduct outlined in the International Council on Archives' Code of Ethics. Among other things, the Code establishes the obligation of archivists to work with record creators to ensure good record-keeping practices from the moment of a record's creation; to protect the authenticity of the records in their care; and to record and justify any actions they take with respect to their long-term management.[41] The requirements are a concrete embodiment of these ethical obligations in the specific context of electronic records.

The work accomplished by the InterPARES Project constitutes a milestone in the search for logical and defensible standards and methods for protecting the authenticity of electronic records over the long term and complements the work undertaken by a number of other research initiatives related to the preservation of digital information in general and electronic records in particular.[42]

At the same time, much work remains to be done.[43] Based on the knowledge the Authenticity Task Force of InterPARES has gained over the past three years, there are three simple but important lessons that might usefully inform future research in this area. Firstly, we still have some distance to travel in our search for ways and means of capturing and preserving the

40 The comparative analyses may be found on the InterPARES Web site at <www.interpares.org>.

41 International Council on Archives, "Code of Ethics" (Beijing, China, 6 September 1996), principles 3 and 5 and accompanying commentary.

42 The current research initiatives most closely related to the work of the InterPARES Project in general and that of the Authenticity Task Force in particular are summarized in the final report of the Authenticity Task Force. See "Establishing and Maintaining Trust," pp. 30–31. For a comprehensive and up-to-date bibliography of the work and research accomplished in the area of managing and preserving electronic records over the past decade, see Minnesota Historical Society, "NHPRC Electronic Records Research Agenda: 1991 Research Issues and Related References," (Draft) November 2002, at <www.mnhs.org/preserve/records/erbibliography.pdf>. Though its primary focus is on research initiatives funded by the U.S. National Historical Publications and Records Commission (NHPRC) since it first articulated an electronic records research agenda in 1991, the bibliography identifies many other local, national, and international initiatives related to the management and preservation of electronic records, and situates each of them in the context of the original ten research questions identified in the NHPRC's 1991 agenda. See National Historical Publications and Records Commission, *Research Issues in Electronic Records* (St. Paul, Minn., 1991).

43 Some of that work will take place as part of the InterPARES 2 Project, a new five-year research initiative, which began in January 2002 and which will build on the findings of what is now known as InterPARES 1. InterPARES 2 will address issues relating to the reliability, accuracy, and authenticity of electronic records throughout their lifecycle. It will focus on records produced in experiential, dynamic, and interactive digital environments and records resulting from artistic, scientific, and governmental activities. For a description of the InterPARES 2 Project see <www.interpares.org>.

unique characteristics of electronic records and electronic systems that will enable us, in turn, to assess and maintain their authenticity over time. Secondly, in the course of that search, we need to develop a more robust interpretation of the necessary and desirable characteristics of electronic records. Such interpretation should take into account the fluid, localized, and process-oriented nature of electronic systems, without, however, losing sight of the record itself, and the desirability of building fixity and standardization of documentary forms into the design of electronic records systems. Finally, the authority and legitimacy of the claims we make for the authenticity of electronic records derive entirely from the integrity and internal coherence of the procedures we adopt to manage them. We are bound, therefore, not only to design and implement procedures that provide a strong circumstantial probability of record trustworthiness but also to provide an honest and adequate account of our choices and decisions, including the compromises we have made, in the course of our stewardship.

In the meantime the findings and products of the research undertaken by all the task forces of the InterPARES Project are available on the Project's Web site. It is time now for the international community of record practitioners and scholars and information technology specialists to examine the findings in order to critique and test them, to conceptualize them within particular organizational structures and technological environments, to situate them in relation to the findings of related research initiatives, and to identify creative and practical ways of interpreting, implementing, and improving upon them.

44 The preserver is the juridical person whose primary responsibility is the long-term preservation of authentic records. The preserver's responsibilities include appraisal.

Appendix

Authenticity Task Force Requirements for Assessing and Maintaining the Authenticity of Electronic Records

The requirements for assessing and maintaining the authenticity of electronic records that are identified in this document fall into two groups: the first group includes requirements that support the presumption of the authenticity of electronic records before they are transferred to the custody of the preserver,[44] while the second group includes requirements that support the production of authentic copies of electronic records that have been transferred to the custody of the preserver. The report is organized into the following sections:

1. Conceptual Framework for the Requirements for Authenticity
2. Benchmark Requirements Supporting the Presumption of Authenticity of Electronic Records
3. Baseline Requirements for the Production of Authentic Copies of Electronic Records
4. Commentary on the Benchmark Requirements Supporting the Presumption of Authenticity of Electronic Records
5. Commentary on the Baseline Requirements for the Production of Authentic Copies of Electronic Records

1. Conceptual Framework for the Requirements for Authenticity

1.1 *Introduction*

Authenticity is defined as "the quality of being authentic, or entitled to acceptance."[45] *Authentic* means "worthy of acceptance or belief as conforming to or based on fact" and is synonymous with the terms *genuine* and *bona fide*. *Genuine* "implies actual character not counterfeited, imitated, or adulterated [and] connotes definite origin from a source." *Bona fide* "implies good faith and sincerity of intention."[46] From these definitions it follows that an *authentic record* is a record that is what it purports to be and is free from tampering or corruption.

In both archival theory and jurisprudence, records that the creator[47] relies on

45 *Oxford English Dictionary,* 2nd ed., s.v. "authenticity."

46 *Merriam-Webster Online Dictionary*, s.v. "authentic."

47 The creator is the physical or juridical person in whose archival fonds the record exists. The fonds is the whole of the records created (meaning made or received and set aside for action or reference) by a physical or juridical person in the course of carrying out its activities.

in the usual and ordinary course of business are presumed authentic. However, digital information technology creates significant risks that electronic records may be altered, either inadvertently or intentionally. Therefore, in the case of records maintained in electronic systems, the presumption of authenticity must be supported by evidence that a record is what it purports to be and has not been modified or corrupted in essential respects. To assess the authenticity of an electronic record, the preserver must be able to establish its *identity* and demonstrate its *integrity.*

The identity of a record refers to the distinguishing character of a record, that is, the attributes of a record that uniquely characterize it and distinguish it from other records. From an archival-diplomatic perspective, such attributes include: the names of the persons concurring in its formation (that is, its author, addressee, writer, and originator); its date(s) of creation (that is, the date it was made, received, and set aside) and its date(s) of transmission; an indication of the action or matter in which it participates; the expression of its archival bond, which links it to other records participating in the same action (for example, a classification code or other unique identifier); as well as an indication of any attachment(s) since an attachment is considered an integral part of a record.[48] The attributes[49] that establish the identity of a record may be explicitly expressed in an element of the record, in metadata related to the record, or they may be implicit in its various contexts. Those contexts include: its *documentary context*, that is, the archival fonds to which a record belongs, and its internal structure; its *procedural context*, that is, the business process in the course of which the record is created; its *technological context*, that is, the characteristics of the technical components of an electronic computing system in which records are created; its *provenancial context*, that is, the creating body, its mandate, structure, and functions; and its *juridical-administrative* context, that is, the legal and organizational system in which the creating body belongs.

The *integrity* of a record refers to its wholeness and soundness: a record has

48 An attachment is a document that constitutes an integral part of the whole record, notwithstanding the fact that it exists as a linked, but physically separate, entity.

49 The use of the terms *attribute* and *element* in this report should not be confused with the way the terms are used in other contexts, such as the various Standard Generalized Markup Languages (SGML). In this report, a *record attribute* is a defining characteristic of a record or of a record element. A *record element* is a constituent part of the record's documentary form and may be either extrinsic or intrinsic. An attribute may manifest itself in one or more elements of a record's documentary form. For example, the name of the author of a record is an attribute, which may be expressed as a superscription or a signature, both of which are intrinsic elements of documentary form. For a more detailed explanation of the extrinsic and intrinsic elements of documentary form see the Authenticity Task Force's *Template for Analysis* which is available on the InterPARES Web site. An attribute may also manifest itself in the form of an annotation(s) to a record, in metadata linked to it, or in one or more of its various contexts.

integrity when it is complete and uncorrupted in all its essential respects. This does not mean that the record must be precisely the same as it was when first created for its integrity to exist and be demonstrated. Even in the paper world, with the passage of time, records are subject to deterioration, alteration and/or loss. In the electronic world, the fragility of the media, the obsolescence of technology and the idiosyncrasies of systems likewise affect the integrity of records. When we refer to an electronic record, we consider it essentially complete and uncorrupted if the message that it is meant to communicate in order to achieve its purpose is unaltered. This implies that its physical integrity, such as the proper number of bit strings, may be compromised, provided that the articulation of the content and any required annotations and elements of documentary form remain the same.[50] The integrity of a record may be demonstrated by evidence found on the face of the record, in metadata related to the record, or in one or more of its various contexts.

1.2 *Benchmark Requirements for Assessing the Authenticity of Electronic Records*

The records of the creator belong to one of two categories. The first category comprises those records that exist as created. They are considered authentic because they are the same as they were in their first instantiation. The second category comprises those records that have undergone some change and therefore cannot be said to exist as first created; they are considered authentic because the creator treats them as such by relying on them for action or reference in the regular conduct of business. However, the authenticity of electronic records is threatened whenever they are transmitted across space (that is, when sent to an addressee or between systems or applications) or time (that is, either when they are in storage, or when the hardware or software used to store, process, or communicate them is updated or replaced). Given that the acts of setting aside an electronic record for future action or reference and of retrieving it inevitably entail moving it across significant technological boundaries (from display to storage subsystems and vice versa), virtually all electronic records belong to the second category. Therefore, the preserver's inference of the authenticity of electronic records must be further supported by evidence – provided in association with the records – that they have been maintained using technologies and administrative procedures that either guarantee their continuing identity and integrity or at least minimize risks of

50 For example, for an electronic mail message, an authentic copy of a complete message may include only the text. Provided it clearly indicated the author, addressee, receivers, and date as well as the content, it would not need to appear in the same way in which it was seen by the author or addressee. By contrast, an authentic copy of a map would have to retain its original presentation features, including color and feature presentation. Provided these requirements were met, an authentic copy could be produced in GIF, JPEG, or GML format.

change from the time the records were first set aside to the point at which they are subsequently accessed. The requirements for assessing the authenticity of the creator's electronic records concern this evidence.

1.2.1 *The Presumption of Authenticity*

A presumption of authenticity is an inference that is drawn from known facts about the manner in which a record has been created and maintained. The evidence that supports the presumption that the record creator created and maintained them authentic are enumerated in the *Benchmark Requirements Supporting the Presumption of Authenticity of Electronic Records* (Requirement Set A). A presumption of authenticity 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. This is why these requirements are termed "benchmark" requirements.

1.2.2 *The Verification of Authenticity*

In any given case, there may be an insufficient basis for a presumption of authenticity, or the presumption may be extremely weak. In such cases, further analysis may be necessary to verify the authenticity of the records. A verification of authenticity is the act or process of establishing a correspondence between known facts about the record and the various contexts in which it has been created and maintained, and the proposed fact of the record's authenticity.[51] In the verification process, the known facts about the record and its contexts provide the grounds for supporting or refuting the contention that the record is authentic. Unlike the presumption of authenticity, which is established on the basis of the benchmark requirements, this verification involves a detailed examination of the records themselves and reliable information available from other sources about the records and the various contexts in which they have been created and maintained. Methods of verification 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 the

51 In common usage, *verify* is synonymous with the terms *validate*, *confirm*, *corroborate*, and *substantiate*. According to *Webster's Online Dictionary*, "*validate* means to attest to the truth or validity of something; *confirm* implies the removing of doubts by an authoritative affirmation or by factual proof; *corroborate* suggests the strengthening of something that is already partly established; *substantiate* implies the offering of evidence that sustains the contention."

medium, script, and so on; a study of audit trails; and the testimony of a trusted third party.

1.3 *Baseline Requirements Supporting the Production of Authentic Copies of Electronic Records*

After the records have been presumed or verified authentic in the appraisal process, and have been transferred from the creator to the preserver, their authenticity needs to be maintained by the preserver. In order to do so, the preserver must carry forward the records in accordance with the baseline requirements that apply to the maintenance of records, producing copies according to procedures that also maintain authenticity.[52] The production of authentic copies is regulated by the *Baseline Requirements for the Production of Authentic Copies of Electronic Records* (Requirement Set B). 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. This is why the requirements for the production of authentic electronic copies are termed "baseline" requirements.

Satisfaction of these baseline requirements will enable the preserver to certify that copies of electronic records are authentic. Traditionally, the official preserver of the records has been the person entrusted with issuing authentic copies of such records. To fulfill that role, the preserver needed simply to attest that the copy conformed to the record being reproduced. With electronic records, the difficulties related to preservation make it prudent for the preserver to produce and maintain documentation relating to the manner in which it has maintained the records over time as well as the manner in which it has reproduced them to support its attestation of authenticity.

A copy is the result of a reproduction process. A copy can be made from an original or from a copy of either an original or another copy.[53] There are several types of copy. The most reliable copy is a copy in the form of an original, which is identical to the original although generated subsequently. An imitative copy is a copy that reproduces both the content and form of the record, but in such a way that it is always possible to tell the copy from the original. A simple copy is a copy that only reproduces the content of the original.

Any of these types of copy is authentic if attested to be so by the official

52 It is understood that the records that are maintained by the preserver only exist as copies of the creator's records.

53 In common language, *copy* and *reproduction* are synonyms. For the purposes of this research, the term reproduction is used to refer to the process of generating a copy, while the term copy is used to refer to the result of such a process, that is, to any entity which resembles and is generated from the records of the creator. An original record is the first, complete record, which is capable of achieving its purposes (that is, it is effective). A record may also take the form of a draft, which is a temporary compilation made for purposes of correction.

preserver. By virtue of this attestation, the copy is deemed to conform to the record it reproduces until proof to the contrary is shown. Such attestation is supported by the preserver's ability to demonstrate that it has satisfied the applicable baseline requirements for maintenance and all of the requirements for the production of authentic copies.

2. Benchmark Requirements Supporting the Presumption of Authenticity of Electronic Records

2.1 *Preamble*

The benchmark requirements are the conditions that serve as a basis for the preserver's assessment of the authenticity of the creator's electronic records. Satisfaction of these benchmark requirements will enable the preserver to infer a record's authenticity on the basis of the manner in which the records have been created, handled, and maintained by the creator.

Within the benchmark requirements, Requirement A.1 identifies the core information about an electronic record – the immediate context of its creation and the manner in which it has been handled and maintained – that establishes the record's identity and lays a foundation for demonstrating its integrity. Requirements A.2-A.8 identify the kinds of procedural controls over the record's creation, handling, and maintenance that support a presumption of its integrity.

2.2 *Benchmark Requirements (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[54]

54 The name of the author is 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.

	• name of writer[55](if different from the author) • name of originator[56] (if different from name of author or writer) • name of addressee[57]
A.1.a.ii	Name of action or matter
A.1.a.iii	Date(s) of creation and transmission, that is: • chronological date[58] • received date[59] • archival date[60] • transmission date(s)[61]
A.1.a.iv	Expression of archival bond[62] (for example, classification code, file identifier)
A.1.a.v	Indication of attachments
A.1.b	Integrity of the record:
A.1.b.i	Name of handling office[63]
A.1.b.ii	Name of office of primary responsibility[64] (if different from handling office)

55 The name of the writer is the name of the physical or juridical person having the authority and capacity to articulate the content of the record.

56 The name of the originator is the name of the physical or juridical person assigned the electronic address in which the record has been generated and/or sent.

57 The name of the addressee is the name of the physical or juridical person(s) to whom the record is directed or for whom the record is intended.

58 The chronological date is the date, and possibly the time, of a record included in the record by the author or the electronic system on the author's behalf in the course of its compilation.

59 The received date is the date, and possibly the time, when a record is received by the addressee.

60 The archival date is the date, and possibly the time, when a record is officially incorporated into the creator's records.

61 The transmission date(s) is the date and time when a record leaves the space in which it was generated.

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

63 The handling office is 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.

64 The office of primary responsibility is the office (or officer) given the formal competence for

A.1.b.iii	Indication of types of annotations added to the record[65]
A.1.b.iv	Indication of technical modifications;[66]

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;

maintaining the authoritative record, that is, the record considered by the creator to be its official record.

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

66 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

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.

3. Baseline Requirements Supporting the Production of Authentic Copies of Electronic Records

3.1 *Preamble*

The baseline requirements outline the minimum conditions necessary to enable the preserver to attest to the authenticity of copies of inactive electronic records.

3.2 *Baseline Requirements (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 that 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.

4. Commentary on the Benchmark Requirements Supporting the Presumption of Authenticity of Electronic Records

The assessment of the authenticity of the creator's records takes place as part of the appraisal process. That process and the role of the Benchmark Require-

ments within it are described in more detail in the *Final Report of the Appraisal Task Force*. This assessment should be verified when the records are transferred to the preserver's custody.

A.1: Expression of Record Attributes and Linkage to Record

The presumption of a record's authenticity is strengthened by knowledge of certain basic facts about it. The attributes identified in this requirement embody those facts. The requirement that the attributes be expressed explicitly and linked inextricably[67] to the record during its life, and carried forward with it over time and space, reflects the Task Force's belief that such expression and linkage provide a strong foundation on which to establish a record's identity and demonstrate its integrity. The case studies undertaken as part of the work of the Task Force revealed very little consistency in the way the attributes that specifically establish the identity of a record are captured and expressed from one electronic system to another. In certain systems, some attributes were explicitly mentioned on the face of the record, in others they could be found in a wide range of metadata linked to the record or they were simply implicit in one or more of the record's contexts. In many cases, certain attributes (for example, the expression of the archival bond) were not captured at all. The Task Force's concern is that, in the absence of a precise and explicit statement of the basic facts concerning a record's identity and integrity, it will be necessary for the preserver to acquire enormous, and otherwise unnecessary, quantities of data and documentation simply to establish those facts.

The link between the record and the attributes listed in Requirement A.1 is viewed by the Task Force as a conceptual rather than a physical one, and the requirement could be satisfied in different ways, depending on the nature of the electronic system in which the record resides. For example, in electronic records management systems, this requirement is usually met through the creation of a record profile.[68] In other types of systems, the requirement could be fulfilled through a topic map. A topic map expresses the characteristics (that is, *topics*) of subjects (for example, records or record attributes) and the relationships between and among them.

When a record is exported from the live system, migrated in a system update, or transferred to the preserver, the attributes should be linked to the record and available to the user. When pulling together the data prior to

67 For the purposes of this requirement, *inextricable* means incapable of being disentangled or untied, and *link* means a connecting structure.

68 If the attribute values contained in the profile are also expressed independently as entries in a register of all records made or received by the creator, then, in addition to establishing the identity and supporting the inference of the integrity of the record, they would also corroborate such identity and strengthen the inference of integrity.

export, the creator should also ensure that the data captured are the right data. For example, in the case of distribution lists, the creator must ensure that if the recipients specified on "List A" were changed at some point in the active life of records, the accurate "List A: Version 1" is exported with the records associated with the first version, and that the second version is sent forward with those records sent to recipients on "List A: Version 2."

A.2 Access Privileges

Defining access privileges means assigning responsibility for the creation, modification, annotation, relocation, and destruction of records on the basis of competence, which is the authority and capacity to carry out an administrative action. Implementing access privileges means conferring exclusive capability to exercise such responsibility. In electronic systems, access privileges are usually articulated in tables of user profiles. Effective implementation of access privileges involves the monitoring of access through an audit trail that records every interaction that an officer has with each record (with the possible exception of viewing the record). If the access privileges are not embedded within the electronic system but are based on an external security system (such as the exclusive assignment of keys to a location), the effective implementation of access privileges will involve monitoring the security system.

A.3 Protective Procedures: Loss and Corruption of Records

Procedures to protect records against loss or corruption include: prescribing regular backup copies of records and their attributes; maintaining a system backup that includes system programs, operating system files, etc.; maintaining an audit trail of additions and changes to records since the last periodic backup; ensuring that, following any system failure, the backup and recovery procedures will automatically guarantee that all complete updates (records and any control information such as indexes required to access the records) contained in the audit trail are reflected in the rebuilt files and also guarantee that any incomplete operation is backed up. The capability should be provided to rebuild forward from any backup copy, using the backup copy and all subsequent audit trails.

A.4 Protective Procedures: Media and Technology

Procedures to counteract media fragility and technological obsolescence include: planning upgrades to the organization's technology base; ensuring the ability to retrieve, access, and use stored records when components of the electronic system are changed; refreshing the records by regularly moving them from one storage medium to another; and migrating records from an obsolescent technology to a new technology.

A.5 Establishment of Documentary Forms

The documentary form of a record may be determined in connection to a specific administrative procedure, or in connection to a specific phase(s) within a procedure. The documentary form may be prescribed by business process and workflow control technology, where each step in an administrative procedure is identified by specific record forms. If a creator customizes a specific application, such as an electronic mail application, to carry certain fields the customized form becomes, by default, the required documentary form. It is understood that the creator, either acting on the basis of its own needs, or the requirements of the juridical system, not an individual officer, establishes the required documentary form(s) of records.

When the creator establishes the documentary form in connection to a procedure, or to specific phases of a procedure, it is understood that this includes the determination of the intrinsic and extrinsic elements of form[69] that will allow for the maintenance of the authenticity of the record. Because, generally speaking, that determination will vary from one form of a record to another, and from one creator to another, it is not possible to predetermine or generalize the relevance of specific intrinsic and extrinsic elements of documentary form in relation to authenticity.

A.6 Authentication of Records

In common usage, to authenticate means to prove or serve to prove the authenticity of something. More specifically, the term implies establishing genuineness by adducing legal or official documents or expert opinion. For the purposes of the benchmark requirements, authentication is understood to be a declaration of a record's authenticity at a specific point in time by a juridical person entrusted with the authority to make such declaration. It takes the form of an authoritative statement (which may be in the form of words or symbols) that is added to or inserted in the record attesting that the record is authentic.[70] The requirement may be met by linking the authentication of specific types of records to business procedures and assigning responsibility to a specific office or officer for authentication.

The authentication of copies differs from the validation of the process of reproduction of the digital components of the records. The latter process

69 The extrinsic and intrinsic elements of form are defined and explained in the Authenticity Task Force's Template for Analysis, which is available at <http://www.interpares.org/reports.htm>.

70 The meaning of authentication as it is used by the Authenticity Task Force in this report is broader than its meaning in Public Key Infrastructure (PKI) applications. In such applications, authentication is restricted to proving identity and public key ownership over a communication network.

occurs every time the records of the creator are moved from one medium to another or migrated from one technology to another.

A.7 Identification of Authoritative Record

An authoritative record is a record that is considered by the creator to be its official record and is usually subject to procedural controls that are not required for other copies. The identification of authoritative records corresponds to the designation of an office of primary responsibility as one of the components of a records retention schedule. The Office of Primary Responsibility is the office given the formal competence for maintaining the authoritative (that is, official) records belonging to a given class within an integrated classification scheme and retention schedule. The purpose of designating an Office of Primary Responsibility for each class of record is to reduce duplication and to designate accountability for records.

It is understood that in certain circumstances there may be multiple authoritative copies of records, depending on the purpose for which the record is created.

A.8 Removal and Transfer of Relevant Documentation

This requirement implies that the creator needs to carry forward with the removed records all the information that is necessary to establish the identity and demonstrate the integrity of those records, as well as the information necessary to place the records in their relevant contexts.

5. Commentary on the Baseline Requirements for the Production of Authentic Copies of Electronic Records

The establishment and implementation of the baseline requirements take place as part of the function of managing preservation. The preservation function and the role of the Baseline Requirements within it are described in more detail in the *Final Report of the Preservation Task Force*.

B.1 Controls over Records Transfer, Maintenance, and Reproduction

The controls over the transfer of electronic records to archival custody include establishing, implementing, and monitoring procedures for registering the records' transfer; verifying the authority for transfer; examining the records to determine whether they correspond to the records that are designated in the terms and conditions governing their transfer; and accessioning the records.

As part of the transfer process, the assessment of the authenticity of the creator's records, which has taken place as part of the appraisal process, should

be verified. This includes verifying that the attributes relating to the records' identity and integrity have been carried forward with them (Requirement A.1), along with any relevant documentation (Requirement A.8).

The controls over the maintenance of electronic records once they have been transferred to archival custody are similar to several of the ones enumerated in the benchmark requirements. For example, the preserver should establish access privileges concerning the access, use, and reproduction of records (Requirement A.2); establish procedures to prevent, discover, and correct loss or corruption of records (Requirement A.3), as well as procedures to guarantee the continuing identity and integrity of records against media deterioration and across technological change (Requirement A.4). Once established, the privileges and procedures should be effectively implemented and regularly monitored. If authentication of the records is required, the preserver should establish specific rules regarding who is authorized to authenticate them and the means of authentication that will be used (Requirement A.6).

The controls over the reproduction of records include establishing, implementing, and monitoring reproduction procedures that are capable of ensuring that the content of the record is not changed in the course of reproduction.

B.2 Documentation of Reproduction Process and its Effects

Documenting the reproduction process and its effects is an essential means of demonstrating that the reproduction process is transparent (that is, free from pretence or deceit). Such transparency is necessary to the effective fulfillment of the preserver's role as a trusted custodian of the records. Documenting the reproduction process and its effects is also important for the users of records since the history of reproduction is an essential part of the history of the record itself. Documentation of the process and its effects provides users of the records with a critical tool for assessing and interpreting the records.

B.3 Archival Description

Traditionally it has been a function of archival description to authenticate the records and perpetuate their administrative and documentary relationships. With electronic records, this function becomes critical. Once the records no longer exist except as authentic copies, the archival description is the primary source of information about the history of the record, that is, its various reproductions and the changes to the record that have resulted from them. While it is true that the documentation of each reproduction of the record copies[71] may

71 Although, technically, every reproduction of a record that follows its acquisition by the preserver is an authentic copy, it is the only record that exists and, therefore, should normally be referred to as "the record" rather than as "the copy."

be preserved, the archival description summarizes the history of all the reproductions, thereby obviating the need to preserve all the documentation for each and every reproduction. In this respect, the description constitutes a collective attestation of the authenticity of the records and their relationships in the context of the fonds to which the records belong. This is different from a certificate of authenticity, which attests to the authenticity of individual records. The importance of this collective attestation is that it authenticates and perpetuates the relationships between and among records within the same fonds.