Structural and Formal Analysis: The Contribution of Diplomatics to Archival Appraisal in the Digital Environment

Luciana Duranti

"Analysis is the essence of archival appraisal" (Schellenberg, 1956a, 45 [277]). All those who have written about appraisal, regardless of their perspective, beliefs, and context, agree that the key to the accurate assessment of the value of records is a systematic and rigorous analysis of their context, inter-relationships, form, content, and/or use. They may disagree on the methodology for or on the object of analysis but, since the mid-nineteenth century, the idea that appraisal could be based on intuition has all but disappeared and has been replaced by the conviction that appraisal can only result from a scientific process of analysis, regardless of the interest being served and the criteria being followed.

Structural analysis was introduced in the discourse on appraisal in the twentieth century by German theorists. Although many archivists in Germany still supported the primacy of content analysis aimed at determining the usefulness of records for future historical research (Zimmerman, 1959), structural analysis began to dominate appraisal methodology, mostly as a consequence of the widespread international acceptance of the principle of provenance as the theoretical basis of archival arrangement. If meaning is derived from context, then an understanding of the administrative structure of a records creator should be able to guide not only arrangement, but also appraisal (Heredia Herrera, 1987, 123). To German archivists, the destruction of copies and transitory records was still the proper thing to do, because they were extraneous to the understanding of context and structure (Doehaerd, 1950, 325), until, in 1939, Hans O. Meissner re-issued and developed the systematic appraisal standards formulated in 1901

by Georg Hille. His primary contribution to appraisal methodology was the use of structural analysis to gather an understanding of the organization, functions and activities of the records creating body. However, he believed that such analysis had to be combined with that of subject content in order to be able to identify records of value (Klumpenhouer, 1988, 52). In 1940, Hermann Meinert endorsed Meissner's standards arguing though that the value of records depends primarily on the significance of a records creator within an administrative hierarchy, which can be determined through an analysis of its position in such structure, of the nature of its activities, and of their relationship with those of superior and subordinate administrative units (Schellenberg, 1956b, 137). This was the first articulation of the now generally accepted proposition that records must be appraised in their administrative context. In 1957, Georg Wilhelm Sante stated that the process of appraisal must begin with the functional analysis of the creator and then proceed to an assessment of the significance of each function and of the administrative body carrying it out (Sante, 1958, 93).

This German confidence in structural analysis as the best means of establishing value continued until the 1980s. However, in the 1960s, its object began to shift from the records creator to the records body. Johannes Papritz expressed his belief that scientific principles should guide appraisal and that knowledge of the structural form of the record body would serve as a precondition because, by creation and meaning, the record body constitutes a logical unit in which each document exists in relationship to the entire body of records (Papritz, 1964, 220). Several years later, Hans Booms, in his seminal article "Society and the Formation of the Documentary Heritage," stated that structural analysis derives from an implicit and excessively ideological assumption of the ultimate value of the public realm over the private (Booms, 1987, 90) and dismissed the validity of Papritz's assertions saying that value cannot logically flow

from the structural form of the record body, in contrast with the widespread belief in the rest of Europe that there is a direct link between structure and value. Champion of such belief was the Italian archival theorist Elio Lodolini, who insisted that the importance of structural analysis resides in the fact that it reflects a theoretical understanding of the organic nature of the archival *fonds* and therefore supports the maintenance of the integrity of series through the empirical process of appraisal (Lodolini, 1987, 214).

This idea that the protection of the nature of the record had to be at the heart of any appraisal process was only implicit in the concept of structural analysis as it developed in Germany and was received in Italy, while it was central to the British views on appraisal and the consequent procedures. While the principle of provenance was never explicitly referred to in connection with appraisal, the use of structural analysis was linked to the consideration of administrative use as the key to appraisal and its equation with historical value. In 1954, the Grigg report accepted Jenkinson's principle that the authenticity, impartiality and interrelatedness of the records must be protected in the process of appraisal, but took the view that a structural analysis of the body of records using the functions and activities of the records creator as point of reference would ensure that the appraiser maintain an objective stance (Grigg Report, 1954, 30). Michael Cook noted in 1987 that the broad correspondence between administrative and research value established by the Grigg report had not been seriously challenged yet since it was advocated (M. Cook, 1987, 52). Indeed, the use of the principle of provenance in appraisal through structural analysis not only preserves the archival and evidentiary nature of the documents, but subsumes the content analysis inherent in the principle of pertinence, thereby resolving the conflict between the two principles.

The endurance of the concept of structural analysis was also due to the fact that it relied on an argument hard to contest, that records do not merely refer to their creator's activities but are material parts of them and directly connected to them (Jenkinson, 1937, 3). This explains why the concept is also found in American writings, which are not concerned as the European ones with the protection of the nature of the records. The first to discuss the use of structural analysis in appraisal was Philip Brooks, who linked it to proper records management strategies: "The whole appraisal function...can best be performed with a complete understanding of the records of an agency in their relationships to each other as they are created rather than after they have lain forgotten and deteriorating for twenty years" (Brooks, 1940, 226). Brooks advocates a two-fold analysis of the administrative structure and functions of the records creator and of the relationships of the records to each other. He focuses on the relationships that exist between records as written evidence of functions and activities of the creating agency, and uses provenancial information to create an objective framework in which value is related to the accuracy with which the records represent the records creator (Brooks, 1940, 231). Brooks methodology of appraisal has had quite an impact in North America, filtered as it was through the writings of Schellenberg.

Arguing from the perspective of the principle of provenance as understood in the German tradition, Schellenberg identifies structural analysis of the administrative context of records creation as the primary means for ascertaining evidential value, which is directly related to the hierarchical position of each office in the administration, the functions performed by each office, and its activities in the execution of each function. "The archivist must know how records came into being if he is to judge their value for any purpose"—he states. But the aim is to preserve the records that most effectively document the substantive functions of the organization

(Schellenberg, 1956, 243-253). In fact, when it comes to informational value, Schellenberg is no longer so much interested in structural analysis as he is in formal analysis. He identifies form among the three tests by which informational value may be judged and states that formal analysis is meant to identify records that are in the most complete, usable and concentrated form available (Schellenberg, 1956, 256-257). Schellenberg's emphasis on the importance of structural and formal analysis was not challenged for at least three decades, and subsequent writers on appraisal have considered a given that records must be appraised in context. In the mid-eighties, however, "functional analysis" became quite popular, and in the early 1990s Samuels explicitly contrasted it to structural analysis, stating that the latter was made obsolete by the fluidity of organizational structures and that functions had to be examined independently of where they occur (Samuels, 1992, 20-24). Regardless, the methodology of such analysis, as well as that of structural and formal analysis, has not been further discussed in the United States.

By contrast, in Canada, structural and formal analyses have been the focus of attention of most writers on appraisal. "Archivists must...look at the processes and functions behind records creation. In this first and most important phase of appraisal, they must understand why records were created rather than what they contain, how they were created and used...and what formal functions and mandates they supported" (Cook,T., 1991, 38). And, "the interaction of structure and function together articulates the corporate mind (or programme) of the records creator" (Cook, T., 1992, 46). Cook's emphasis on relating the processes and functions of administrative structures to the circumstances of creation, supporting the centrality of provenancial information in appraisal, was complemented by Craig's focus on the analysis of the records: "The reality of the record base must be an indispensable component of all acts of appraisal. Without an understanding of documents and records, of their forms and of their functions, and of how they

were created and used, a plan can be so easily upset by the attractiveness of concentrating on information divorced from the realities of its documentary expression...., it is the record which is our special area of knowledge; it will be a sad day and a dangerous step when faith in planning replaces the study and knowledge of records." (Craig, 1992, 179). Finally, in 1998, Lemieux suggested that, on the basis of the organizational configurations identified by Mintzberg, one could determine which functions are "organizationally significant" without the need to analyze the actual functions, and that would lead to the identification of the "sites of archivally significant records" (Lemieux, 1998, 32-85).

Yet, none of the Canadian writers—similarly to European and American writers—discussed how to gain that understanding of the records context (i.e., structure and/or functions and activities) and of the records themselves, an understanding that is identified by all of them as the essential pre-requisite to sound appraisal. At least, not explicitly, until this author identified in diplomatic criticism the most appropriate methodology for conducting structural and formal analysis of records creators and the related bodies of records for the purposes of appraisal: "The relationship between the records and the actions from which they derive, as embedded in the records intellectual forms and in their forms of aggregation, which tend to be very repetitive, will enable us to identify which functions and activities generated them, and their relative significance. Record forms will guide us to meaning, context and value, and so will the processes and procedures, the functions and activities of records creators" (Duranti, 1991, 26). The only other writer who considered form the key to a full understanding of the record was Peter Sigmond, the leader of a project called "Commentaries on Sources," undertaken in the Netherlands in the mid-eighties. This project had the objective of identifying and describing the procedures and record types used by Dutch government bodies during the 19th century to carry

out their mandates. The researchers focused on the identification of recognizable patterns of action starting from a diplomatic analysis of the records, rather than relying on the logical breakdown of functions and activities resulting from the study of the organizations' mandates without examining any existing aggregations of records (Sigmond, 1991-92, 141-147). This work was very successful within the limits that it had defined and provided evidence of the continuing validity of diplomatic criticism. However, it did not have much resonance because the research was carried out on records that were a century old—therefore, relatively few in number—and on paper—therefore, fixed in form and content, structured, and manifestly interrelated. By contrast, the proposal made by this author that diplomatic criticism be used on contemporary and even yet to be created electronic records, that is, in a prospective way rather than only in a retrospective way (Duranti, 1990), attracted much attention, particularly as it came in response to a clear emergency call on the part of electronic records experts.

With few exceptions, electronic records experts have generally appreciated the use of structural analysis. In 1984, Naugler proposed a two-tiered approach to the appraisal of electronic records that combined structural and content analysis with technical analysis (Naugler, 1984). In 1990, the United Nations Advisory Committee for the Co-ordination of Information Systems (ACCIS) recommended an analytical process that begins at the design phase of the information system that is expected to make or receive and maintain a creator's record, thereby endorsing both Brook's position that the appraisal process must begin as early as possible in the records life-cycle, and the concept of structural and formal analysis as it was developed by Western archival theorists in the course of the twentieth century. But the ACCIS report went much further in the explicit concern that it manifested for the protection of the nature of the record as, to that purpose, it recommended interference with the records creation process. In

fact, it stated that design decisions must determine, for all records to be produced in the course of significant transactions, record layout, the linkage of records to each other and to those in other systems, and the function that each type of record fulfills. Although this author appreciated at the time the emphasis put by the ACCIS report on the identification of the entities in the system that are records in order to conduct an appraisal that is effective and efficient, she also felt that the report went too far in its recommendations and that the structural and formal analysis should be used differently in support of the appraisal function and should be guided by the methodology of diplomatics (Duranti, 1990, 12; Duranti, 1997).

Diplomatics was developed as a science for the purpose of determining the authenticity of records of unproven origin. Thus, it comprises a body of concepts about the nature of records, their characteristics, components, effects, and relationships, and the requirements for their trustworthiness and genuineness, and a methodology for identifying the entity record and assessing its authenticity in any context. The methodology involves detailed sub-methods for analysing the juridical-administrative system in which the records are created, the records creators, their functions and activities, their procedures and processes, their record-making and recordkeeping systems, and their records. These methods are of a comparative nature in that they create models and templates representing the ideal record, the ideal procedure, the ideal structure of a record system, etc., based on past knowledge, and bounce unknown or new situations against them. In the course of this analytical process the researchers learn about the situation under study by recognizing what is known and describing what is new, and develop new models and templates on the basis of the acquired new knowledge. While diplomatics originated from the need of discovering, understanding and assessing what exists, and it is still used that way by professional diplomatists, the body of knowledge that it has accumulated over

the centuries can be easily used for determining the features of what will exist, by designing records forms, structuring procedures, developing records systems, and for supporting several archival activities, including appraisal, in ways that no other discipline can offer.

Twenty years after the ACCIS report, the need for a structural and formal analysis supported by the methodology of diplomatics has become evident and urgent. Its importance has been demonstrated by the research conducted in the past ten years in the context of the InterPARES project, an international multidisciplinary collaborative project aimed at the development of theory, methods and practices for the long term preservation of electronic records (www.interpares.org). Some of the findings of InterPARES regard appraisal, and specifically its methodology. With digital records, appraisal has changed in four fundamental ways: first, the preserver must assess the authenticity of the records considered of continuing value; second, the preserver must determine the feasibility of the preservation of authentic records; third, the disposal decision must be made very early in the life of the records; and, fourth, the preserver must constantly monitor the records of the creator and, if warranted by the changes that they have undergone through time, revise the disposal decision (Eastwood, 2004, 202-208; InterPARES Project, 2001b). Of these new activities involved in appraisal of digital records, two in particular are based on formal and structural analysis: the determination of the authenticity of the records and the monitoring of the records from the moment they are created to the time they become inactive. However, before discussing these activities and the methods for carrying them out, the most significant issue presented by appraisal of digital records needs to be discussed: the identification, in each given case, of the object of appraisal, the records. Increasingly, archivists feel unable to find records in digital systems, as they appear to contain only data. When such a situation occurs, archivists must determine, on the basis of the creator's

functions and activities, whether a record should exist, and if so, must help the creator to redesign the system to enable it to create records that can be preserved and can serve either a memorial or an evidential function. This is quite difficult when, for reasons primarily of social and professional responsibility, but also of transparency and accountability, records must be created that contain a certain type of data and must be kept for the same use for which they are created, probably for a very long time. In such cases, the assessment of the value of specific data sets is conducive to the definition of the form of the records that should contain them and of the digital presentation that will allow for their long term use, accessibility and preservation. Although appraisal ends up serving a creation purpose and is followed by the actual creation of new documentary forms, it does not affect the impartial, involuntary nature of the resulting records, as the data sets already exist and the records that come to contain them are used in the usual and ordinary course of business by the creator for its own purposes, rather than being generated for research purpose.

In order to demonstrate the problems presented by digital systems in terms of the identification of the entities to appraise, the concept of digital record as formulated by diplomatics should be presented. The diplomatic concept of digital record identifies the following necessary characteristics: 1) a fixed form; 2) an unchangeable content; 3) explicit linkages to other records within or outside the digital system, through a classification code or other unique identifier; 4) an identifiable context of creation; 5) the involvement of five identifiable persons, an author—the person responsible for issuing the record, an addressee—the person for whom the record is intended, a writer—the person responsible for the articulation of content, an originator—the person responsible for the space from which a record is sent or in which it is generated and saved; and a creator—the person in whose *fonds* or archive the record

exists; 6) an action, in which the record participates or which the record supports either procedurally or as part of the decision making process (Duranti, 2009). With complex digital systems, the characteristics that create the most problems are the first two: fixed form and stable content. We can say that a digital record has a fixed form if its binary content is stored so that the message it conveys can be rendered with the same documentary presentation it had on the screen when first saved, even if its digital presentation has been changed, for example, from Word to .pdf. We can also say that a digital record has a fixed form if the same content can be presented on the screen in several different ways but in a limited series of pre-determined possibilities: in such a case we would have different documentary presentations of the same record (e.g. statistical data viewed as a pie chart, a bar chart, or a table). The latter situation brings forth the issue of "stored record" versus "manifested record".

A stored record is constituted of the linked digital component(s) that are used in re-producing the record, which comprise the data to be processed in order to manifest the record (content data and form data) and the rules for processing the data, including those enabling variations (composition data). A manifested record is the visualization or materialization of the record in a form suitable for presentation to a person or system. Sometimes, it does not have a corresponding stored record, but is re-created from fixed content data when a user's action associates them with specific form data and composition data (e.g. a record produced from a relational database). If the same user's action produces always the same documentary presentation with the same content, we can say that the manifested entity, even when it does not have a corresponding stored record, has fixed form and stable content and, if all other records characteristics are present, is a record. In contrast, when one stored record may results in several documentary presentations, as mentioned above, it is an appraisal decision to determine whether

the entity to keep as the record of an activity is the stored one or one or more of the manifested ones, and such a decision is made when the entity identified as the record is assigned a retention period. There might also be situations in which a stored record is never manifested, as is the case with software patches that enable the playing of electroacustic music, or with interacting business applications, workflow generated and used to carry out experiments, analyses of observational data carried out by interpreting software, etc. Also in this case, it is an appraisal decision to determine which enabling entities should be retained with other records of the same activity, manifested or not. Clearly, these decisions require both structural and formal analyses of functions, activities and records, as aggregates and as individual entities. These analyses necessitate the use of diplomatic theory and methodology, also when the matter to be studied is not only contextual and formal, but also content-related. And this takes us to the concept of stable content.

A digital entity has stable content and can be considered a record—if all other conditions are satisfied, if the data and the message in it are unchanged and unchangeable, meaning that data cannot be overwritten, altered, deleted or added to. However, there are cases in which we consider as having stable content entities that demonstrate bounded variability. A digital entity has bounded variability when changes to its form are limited and controlled by fixed rules, so that the same query or interaction always generates the same result, and when the user can have different views of different subsets of content, due to the intention of the author or to different operating systems or applications. While the former definition of stable content applies to static digital entities, the latter is significant when the entities we are looking at are interactive.

A static digital entity is one that does not provide possibilities for changing its manifest content or form beyond opening, closing and navigating, for example e-mail, reports, sound recordings, motion video, and snapshots of web pages. These entities, if all other conditions are satisfied, are records because they have fixed form and stable content. An interactive digital entity, instead, presents variable content, form, or both, and the rules governing the content and form of presentation may be either fixed or variable. Interactive entities may or may not be records, depending on whether they are non-dynamic or dynamic. Non-dynamic entities are those for which the rules governing the presentation of content and form do not vary, and the content presented each time is selected from a fixed store of data. Examples are interactive web pages, online catalogs, and entities enabling performances: if the other conditions exist, they are records. Dynamic entities are those for which the rules governing the presentation of content and form may vary: they are potential records, in that they become records if either the digital system in which they exist is redesigned in such a way that the rules do not vary any longer, or they are moved to another system that only maintains digital records (i.e. static or non-dynamic entities). Examples of dynamic entities are: entities whose variation is due to data that change frequently (e.g., the design permits updating, replacement or alterations; it allows data collection from users or about user interactions or actions; or it uses these data to determine subsequent presentations); entities whose variation is due to data continually received from external sources and not stored within the system; entities produced in dynamic computing applications that select different sets of rules to produce documents, depending on user input, sources of content data, and characteristic of content (e.g. weather sites); entities produced by evolutionary computing where the software generating them can change autonomously (e.g. scheduling and modeling of financial markets; edutainment sites), etc. (Duranti L. and Thibodeau K., 2006) .

The question that immediately comes to mind is: "why cannot we simply re-design in some standardised way each type of system to enable it to produce and keep static or interactive but non-dynamic entities?" The answer is simple: because the way in which each system is designed entirely depends on the function that the system fulfills in the context of the records creators and/or uses over time. The InterPARES research project has proven through many case studies that solutions to these issues are specific, in addition to being dynamic. For example, a Geographic Information System (GIS) used by a public body to make decisions on the basis of the data available at the time of the decision itself, probably would require that the data sets are time stamped so that an auditor can see not necessarily what a decision maker *saw* at a given point in time but what s/he *would have been able to see* if s/he had followed the required procedure, and hold him/her accountable on that basis. In contrast, a GIS used by a private research team to carry out testing might need to be able to stabilize and fix all the users' interactions and their effects, not by time, but by type of action or by type of material on which the action was carried out. In another example, a relational database of a public office having high interoperability with the digital systems of functionaries in other public offices, and incorporating a PKI for continuous authentication and extreme assessment of integrity of data, may need to provide access the material attesting public transactions and to periodically transfer them to a public archives. However, the database does not contain any records, even if it is non-dynamic (i.e. the database itself is a record), thus the only way of making the material available, given the reasons for doing so, might be to define an XML schema which may serve as a translation device between the complex data model used by the database, and a less complex model, to be defined, sufficient to satisfy the needs of users. The content of each data set could then be exported to a file according to the XML schema and imported into a parallel relational

database sufficiently simple to be maintained for purposes of access and preservation (e.g., Microsoft Access). All these decisions are indeed appraisal decisions about what to select, why, and in which form, and need to be based on a structural analysis of the functions and activities of the creating body as well as on a structural and formal analysis of the system containing the material in question and of each digital entity in the context of the entire records system (digital and otherwise) in which they belong. Once again, the knowledge necessary to carry out such analyses derives from diplomatics.

However, as mentioned earlier, the need for diplomatics knowledge in conducting appraisal of digital records is not limited to the identification of the records to be appraised and to a consequent decision based on their characteristics, purpose and use, and to the re-design of digital systems when it is needed to have records where data sets should be embedded in records, but also to assess the authenticity of the records to be appraised for continuing preservation and to monitor the records and the system in which they exist during the period between the initial assessment of value and the implementation of the final appraisal decision (e.g. the transfer to an archival institution).

Traditional appraisal literature has never concerned itself with the authenticity of records. Yet, authenticity does represent a great challenge for the appraiser of digital records. In 1922, Sir Hilary Jenkinson identified authenticity as one of the characteristics of archival material, and linked it to the procedures of creation, maintenance and preservation. Archival documents (i.e. records) are trustworthy as statements of fact because those who generate them need to rely on them for action or reference, and they are trustworthy as records because that same need of the creator and its legitimate successor ensures that proper guarantees are put into place to keep them

intact over time, both in the short and the long term (Jenkinson, 1922, 8-9, 39). With digital records, a presumption of records authenticity based on the reliance on them by their creator and on a legitimate chain of unbroken custody is no longer possible. Because of their manipulability, vulnerability and fragility, the authenticity of digital records is constantly at risk, especially when they are transmitted across space and time and when they are migrated from an obsolescent system to a new one. Thus, authenticity cannot be considered a characteristic of all digital records, but only of those whose processes of creation, maintenance and preservation respect certain pre-established authenticity requirements. In all other cases, authenticity must be verified.

An authentic record is one that is what it purports to be. According to diplomatics, record authenticity comprises identity and integrity. "Identity" refers to the attributes of a record that, together, uniquely characterize it and distinguish it from other records. These attributes include: the names of the persons concurring to its formation; its date(s) of creation and transmission; an indication of the matter or action in which it participates; the expression of its archival bond, that is, of its relationships to the other records within the same *fonds*; as well as an indication of any enclosure(s). These attributes may be explicitly provided by formal elements of the record (e.g., a signature, a subject line, or a classification code) or by information linked to the records (e.g., a register entry), or may be implicit in the various contexts of the record (i.e., administrative, provenancial, procedural, documentary, or technological). "Integrity" refers to the wholeness and soundness of a record. A record has integrity if it is intact and uncorrupted, that is, if the message that it is meant to communicate in order to achieve its purpose is unaltered. Its physical integrity (e.g., in a digital record, the proper number of bit strings) may be compromised, provided that the articulation of the content and its required elements of form remain the same. Integrity may

be demonstrated by evidence found on the face of the record, in information related to the record, or in one or more of its contexts.

In light of these definitions, in order to assess the authenticity of a record, one has to establish its identity and demonstrate its integrity. However, the question arises of who should be responsible for such an assessment, the archivist or the researcher. Traditionally, archivists have rejected such responsibility. In cases of demonstrable legitimate unbroken custody, archivists have presumed authenticity, but even so, if asked to declare a record authentic, they would only go as far as to declare that a record in their custody was as authentic as when transferred to the archives. In a 1949 article, Herman Kann wrote that, if a researcher were to request a declaration that a record is authentic, the archivist should firmly refuse to issue it, because this kind of interpretation is not part of its responsibilities (Kann, 1949, 363). Thus, the assessment of records authenticity is traditionally a responsibility of the researcher. This is perfectly consistent with archival theory and the impartiality that it accords to the professional archivist. The only important role archivists have with respect to authenticity is to describe the records in their custody in context, by making explicit, stabilizing and perpetuating their relationships with their creator and among themselves: this elucidation of the records in their various contexts is one of the primary instruments of any researcher who wishes to assess the records' authenticity. Indeed, archival description is a collective attestation of the authenticity of the records of a *fonds* and of all their interrelationships, as made explicit by their administrative, custodial and technological history, the illustration of their scope and content, and the hierarchical representation of the records aggregates. The unique function of archival description is to provide an historical view of the records and of their becoming while presenting them as a universe in which the individuality of each member is subject to the bond of a common provenance and destination.

Beyond archival description and certification of the authenticity of copies of records in their custody, archivists have traditionally held no other role with respect to the authenticity of records, especially when it came to appraisal.

Archival theory accepts appraisal only in the measure in which it respects the characteristics of the records, that is, their naturalness, impartiality, authenticity, interrelatedness, and uniqueness, by not interfering with them in the process of assigning value. In other words, as long as the archivist does not import into his function elements of personal judgment, the selection of records for permanent preservation is a legitimate archival endeavor. In this context, the assessment of the authenticity of records is an activity that risks compromising the impartiality of the records by alerting the creators to their inherent value, interpreting the records formal elements, and evaluating their processes of creation and maintenance. Also archival practice has traditionally rejected the assessment of the authenticity of the records as part of appraisal, on the grounds that it would make appraisal far too laborious and time consuming. However, this common stance of archival theorists and practitioners could be held only because, with traditional records, the documents entering an archival institution or program were the same that were made or received and set aside by their creator or legitimate successor and evaluated by the archivist in the appraisal process. Thus, the assessment of authenticity could be easily delegated to future researchers, who would be able to analyze the records under scrutiny in their original instantiation, that is, in the same form and status of transmission (i.e. degree of perfection: draft, original or copy) they had when first made or received and set aside. This is no longer the case.

Digital records undergo several changes from the moment they are generated to the moment they become inactive and are ready for the implementation of the appraisal decision made while they were still used by the creator. Some of those changes are intentional. Information technology is in a constant state of development. Records creators continually update their systems and the live digital entities contained in them, at times with minimal consequences for the form, functionality, organization and metadata of the records, other times with dramatic consequences. The latter situation is more likely to occur when records generated in an obsolete system are migrated to a new one. In addition to intentional changes, inadvertent changes occur, simply because of the fact that it is impossible to keep a digital record; it is only possible to store its digital components in a way that it can be reproduced when needed.

A digital component is a digital object that contains all or part of the content of an electronic record, and/or data or metadata necessary to order, structure, or manifest the content, and that requires specific methods for storage, maintenance and preservation. It is distinguished from the extrinsic and intrinsic elements of form, which are those parts of a record that constitute its external appearance and convey the action in which it participates and the immediate context in which it was created, because a digital component is simply a unit of storage, not meaningful *per se*. Every time a digital record is reproduced from its digital components, it is slightly different from the previous time. This happens because there are three steps in the reproduction of a digital record. The first step is to reassemble all the record's digital components in the correct order. The second step is to render the components, individually and collectively, in the correct documentary form or presentation (if we are dealing with a manifested record). The third step is to reestablish the relationships between the digital components of the record in question and all the other records that belong in the same archival aggregate (e.g., series, file). This

requires, first, to recreate the structure of the archival aggregate, and then to fill it with the records that belong to it. Each step involves a margin of error. Considering that the processes of storage and retrieval by re-production imply transformations, both physical and of presentation, the traditional concept of unbroken custody must be extended to include the processes necessary to ensure the unaltered transmission of the record through time, and must therefore become an "unbroken chain of preservation," which begins when the records are created respecting established authenticity requirements, and continues with the documentation of all the changes to the records and of the processes of appraisal, transfer, reproduction and preservation. However, the most important consequence of this situation is that the appraisal function must include appropriate activities aimed at ascertaining the authenticity of the records considered for selection, monitoring it, and attesting it.

The appraisal of digital records, therefore, more then ever, must rely on a diplomatics based structural and formal analysis. This analysis is also necessary to establish whether the digital components embodying the essential elements that confer identity to and ensure the integrity of the records can be preserved, given the current and anticipated technological capabilities of the archives. This determination process comprises three steps. The appraiser should identify the record formal elements that provide informational content and those that need to be preserved according to the authenticity requirements that constitute the specific terms of reference for the designated preserver. Then, the appraiser identifies where these crucial record elements are manifested in the digital components of the record. Finally, the appraiser reconciles these preservation requirements with the preservation capabilities of the organization that is responsible for the continuing preservation of the body of records being appraised. The appraisal decision comprises two parts: a determination of what must be transferred to the archives,

including the list of the digital components of the records, and a determination of how and when this should happen, including the identification of acceptable digital presentations or formats (especially for stored records that do not have corresponding manifested records) and methods of transmission to the archives (InterPARES Project, 2001b). One key component that must remain inextricably linked to the record is the metadata, the presence of which constitutes the key authenticity requirement for any digital records.

Diplomatically, records metadata can be divided in identity metadata and integrity metadata. The former include, at a minimum, the identity attributes identified earlier: the names of the five persons concurring in the record creation; the date(s) and time(s) of issuing, creation and transmission; the matter or action in which the record participates; the expression of its archival bond; the record documentary presentation, or form; its digital presentation, or format; the indication of any attachment(s); the presence of a digital signature. The latter include data related to responsibility for the record and to changes made to the record and, at a minimum: the name of the juridical person responsible for the record; the name(s) of the person(s) handling the record over time; the name of the person responsible for keeping the record; an indication of annotations; an indication of technical changes; an indication of the presence or removal of a digital signature other than with which the record was received; the time of planned removal from the system; the time of transfer to a custodian; the time of planned deletion; and the existence and location of duplicates outside the system. Metadata that need to be added by the appraiser and the preserver for the purpose of providing additional grounds for the authentication of the records over time can be determined on the basis of the structural and formal analysis of the records. The presence of identity and integrity metadata is just one of the factors the appraiser needs to assess in order to declare a presumption of authenticity for the digital records

under examination. The others are that the digital materials being appraised are protected from unauthorized action using physical security, access privileges and blocks on modifying records once filed pursuant to a classification plan; that they are protected from accidental loss and corruption using daily backups; and that steps had been taken by the creator to prevent hardware and software obsolescence by upgrading and migrating to new technology and retaining relevant documentation for long-term preservation (InterPARES Project, 2001a). A protection of the records entrusted to digital authentication technology is not sufficient to ensure their authenticity and indeed the appraiser should advocate the removal of any type of digital signature to guarantee the records continuing preservation.

It is one of the most important contributions of diplomatics to the trustworthiness of the records to have differentiated authenticity and authentication, which is defined as a declaration of authenticity, resulting either by the insertion or the addition of an element or a statement to a record. In a diplomatic analysis of the extrinsic elements of electronic records, digital signatures are identified as examples of electronic seals, being functionally equivalent to medieval seals, which were not only a means of verifying the origin of the record and the fact that it was intact, but also made the record indisputable and incontestable, that is, had a non-repudiation function. The analogy is not perfect, because the medieval seal was associated exclusively with a person, while the digital signature is associated with a given person and a specific record, and because the former was an expression of authority, while the latter is only a mathematical expression. However, it is essential to remember that authenticity is a property of the record that accompanies it for as long as it exists; while authentication is a means of proving that a record is what it purports to be at a given moment in time.

Once the appraiser has concluded his/her structural and formal analysis, a presumption of authenticity can be issued, based upon the number of requirements that have been met and the degree to which each has been met. If the presumption of authenticity is too weak, a verification of authenticity will be necessary. A verification of authenticity is the act or process of establishing a correspondence between known facts about the records and the various contexts in which they have been created and maintained, and the proposed fact of the records authenticity. It involves a detailed diplomatic examination of the records in all their contexts, and of reliable information available from other sources (e.g., audit trails, backups, copies preserved elsewhere), and, if needed, even a textual analysis. The resulting assessment may affect the determination of the records value. This information is also crucial to understanding and using the records once they have been transferred to the preserver. Future users of the records must know how well founded a declaration of authenticity of the records is and what information that declaration is based on in order to make their own assessment, long after the fact, at a time when accumulating relevant information will be difficult, if not impossible. If the appraiser has good reason to suspect that the records no longer reflect what they were at the time of their creation and primary use, he or she may decide not to preserve them (InterPARES Project, 2001b; InterPARES 2 Project, 2008a).

Once the initial appraisal is concluded, the records selected for long-term preservation must be continually monitored till the day of the transfer to the designated preserver, mostly for changes in their technological context and characteristics, but also for changes to their function, use, and consequently value. In most cases, monitoring produces minor revisions to the documentation on their appraisal and to the terms and conditions of transfer. In some cases, however, it may be necessary to repeat the appraisal because of changes that can affect the

feasibility of preservation of the selected records in the archival environment, or because of changes that can affect their identity.

The InterPARES 2 report Modeling Digital Records Creation, Maintenance and Preservation states:

> Because there may be changes in the way records are generated or organized, in the technology the creator uses to create them, or in the preserver's preservation capabilities, part of appraising digital records involves monitoring records that have already been appraised to identify any necessary changes to appraisal decisions over time. As well, because the creator's organizational mandates and responsibilities may change over time, as might the way those responsibilities are carried out, such that data accumulated in formerly appraised systems may be put to new uses, it is possible that systems that did not initially contain records may be upgraded to do so, especially in organizations with hybrid paper and electronic recordkeeping systems. Likewise, it is likely that the preserver's preservation capabilities will change over time, as might its organizational mandates and responsibilities. Therefore, in addition to monitoring changes to the creator's appraised records, it is also necessary for the preserver to keep track of appraisal decisions in relation to subsequent developments within the creator's and/or preserver's operations that might make it necessary to adjust or redo an appraisal, such as substantial changes to: (1) the creator's organizational mandate and responsibilities, (2) the creator's recordmaking or recordkeeping activities or systems, (3) the preserver's records preservation activities or systems and/or (4) the preserver's organizational mandate and responsibilities. (InterPARES Project, 2008b, 35)

Several InterPARES case studies have demonstrated such necessity. For example, a university student registration system underwent three migrations to new and more complex technologies. The first system had a current and an historical part, the latter containing the records of *alumni*, which would be accessible but impossible to modify, delete or add to. The designers of the second system eliminated the historical component so that all information on students and *alumni* was kept live. Both systems contained, about each individual, the data sets required by the registrar office to carry out its specific functions. The third system was much more complex and allowed for manipulating data in ways that could support activities that were

not the responsibility of the registrar, but of other university offices, such as recruitment of new students, and planning of university events, which could be carried out using the data sets in the registration system if they were enriched with additional data and manifested in several additional ways. Thus, the registrar's office began asking the students to provide upon registration more information about themselves—as detailed by the interested university offices which were going to use the system—and the system was configured to produce manifested records pre-defined by those offices in order to exhibit the relevant data sets in such a way that they could be used to carry out their functions.

Clearly, the retention and disposition decision made for the first iteration of the university registration system was no longer valid for its second and third iteration. Furthermore, the methodology for conducting appraisal now required beforehand a structural analysis of the university offices using the system and the functions and activities that they so fulfilled. On that basis, the data sets used in the course of each activity as content of each manifested record had to be identified by formal analysis and put into relation with the other records produced by the same activity and residing outside the registration system. Following this analytical process, a determination had to be made as to the relative value of each data set contained within the registration system with respect to each function carried out by each user office and to the other records/data sets created in the process, and then an appraisal decision had to be linked to each data set. If a data set was linked to multiple appraisal decisions, the decision requiring the longest retention period would override any other. Certainly, a redesign of the system based on the structural analysis of its users and the formal analysis of the records aggregates used and added to by each would allow for a proper control on the records and a better appraisal. Such system would have embedded into it integrated business and documentary procedures for each user

office, an integrated classification system and retention schedule consistent with that of the office of primary responsibility for each type of manifested records, and appropriate identity and integrity metadata sets.

However, redesigning digital systems is a very expensive endeavor and not one likely to happen till such time when their users will encounter some serious legal problem requiring it. And a long time will pass before knowledgeable records professionals will be asked to participate in the initial determination of requirements and specifications for the design of record-making and recordkeeping systems, and this means that our knowledge of diplomatics will have to continue to be used retrospectively—analyzing the characteristics and behavior of what records/data exist, rather that prospectively—analyzing records creators functions and activities to determine what form the records yet to be generated should take, be they manifested or stored, static or interactive. While we wait for that time to come, we need to continue to develop the science of diplomatics and its methodological tools so that the appraisal of the next complex forms of digital records (e.g. holographic) in the context of the new activities that will use them will not find us unprepared.

List of References

Advisory Committee for the Co-ordination of Information Systems (ACCIS) (1990). *Management of Electronic Records: Issues and Guidelines.* New York: United Nations.

Booms, H. (1987) Society and the Formation of a Documentary Heritage: Issues in the Appraisal of Archival Sources. Trans. Hermina Joldersam and Richard Klumpenhouer, *Archivaria* 24: 69-107.

Brooks, P. (1940) The Selection of Records for Preservation. *American Archivist*: 221-234.

Cook, M. (1986) *The Management of Information From Archives*. Hants, England: Gower Publishing Co.

Cook, T. (1991) *The Archival Appraisal of Records Containing Personal Information: a RAMP Study with Guidelines*. Paris: UNESCO General Information Program and UNISIST. PGI 91/WS/3.

Cook, T. (1992) Mind Over Matter: Towards a New Theory of Archival Appraisal, in Barbara L. Craig, ed., *The Archival Imagination: Essays in Honour of Hugh A. Taylor.* Ottawa: Association of Canadian Archivists.

Craig, B. (1992) The Acts of the Appraisers: The Context, the Plan and the Record, *Archivaria* 34: 175-180.

Doehaerd, R. (1950) Remarks on Contemporary Archives, *American Archivist:* 14, 325.

Duranti, L. (1990) Diplomatics: New Uses for an Old Science (Part II), *Archivaria* 29: 4-17.

Duranti, L. (1991) ACA 1991 Conference Overview. *ACA Bulletin,* July 1991.

Duranti, L. 1997. The Thinking About Appraisal of Electronic Records: Where We Have Been and Where We Are Going. *Janus,* International Council on Archives ed. Paris: International Council on Archives, 47-67

Duranti, L. and Thibodeau, K. (2006) The Concept of Record in Interactive, Experiential and Dynamic Environments: the View of InterPARES. *Archival Science,* 6, 1: 13-68. Also online: http://dx.doi.org/10.1007/s10502-006-9021-7

Duranti, L. (2009) Diplomatics. *Encyclopedia of Library and Information Science*. Marcia Bates, Mary Niles Maack, Miriam Drake eds. New York, Basel, Hong Kong: Marcel Dekker, INC.

Eastwood, T. (2004) Appraising Digital Records for Long-term Preservation, *Data Science Journal* 3: 202-208. http://www.jstage.jst.go.jp/article/dsj/3/0/202/_pdf

Heredia Herrera, A. (1987) *Archivistica General: Teoria y Practica*. Sevilla: Deputacion Provincial de Sevilla.

InterPARES Project. (2001a) Authenticity Task Force Report. Available at http://www.interpares.org/book/interpares_book_d_part1.pdf

InterPARES Project. (2001b) Appraisal Task Force Final Report. Available at http://www.interpares.org/book/interpares_book_e_part2.pdf

InterPARES 2 Project. (2008a) Methods of Appraisal and Preservation. Available at http://www.interpares.org/ip2/display_file.cfm?doc=ip2_book_part_4_domain3_task_force.pdf

InterPARES 2 Project. (2008b) Modeling Digital Records Creation, Maintenance and Preservation. Available at http://www.interpares.org/ip2/display_file.cfm?doc=ip2_book_part_5_modeling_task_force.pdf

Jenkinson, H. (1922) *A Manual of Archive Administration*. 2nd edition. Reprint. London: Percy Lund, Humphries & Co. Ltd. 1965.

Kann, H. (1949) A Note on the Authentication of Documents, *The American Archivist*, vol. XII, no. 4:363.

Klumpenhouer, R. (1988) *Concepts of Value in the Archival Appraisal Literature: An Historical and Critical Analysis*. Unpublished MAS thesis, University of British Columbia.

Lemieux, V. (1998) Applying Mintzberg's Theories on Organizational Configuration to Archival Appraisal, *Archivaria* 46: 32-85.

Lodolini, E. (1987) *Archivistica: Principi e Problemi.* Milano: Franco Angeli.

Naugler, H. (1984) *The Archival Appraisal of Machine-Readable Records: A RAMP Study with Guidelines.* Paris: UNESCO.

Papritz, J. (1964) Zum Massenproblem der Archive, *Der Archivar* 17, 220. Quoted in Booms, 94.

Samuels, H. (1992) *Varsity Letters*. *Documenting Modern Colleges and Universities.* Metuchen, NJ and London: Society of American Archivists and the Scarecrow Press, Inc.

Sante G. W. (1958) Behorden—Aktn—Archive. Alte Taktik und Neue Strategie, *Archivalische Zeitschrift* 54: 93, quoted in Klumpenhouer, 134.

Schellenberg T. R. (1956a) The Appraisal of Modern Public Records, *Bulletins of the National Archives* 8, October, 5—46 [237—278].

Schellenberg T. R. (1956b) *Modern Archives: Principles and Techniques*. Chicago: University of Chicago Press.

Sigmond, P. (1991-92) Form, Function and Archival Value, *Archivaria* 33: 141-47.

United Kingdom (1954). *Committee on Departmental Records Report.* Sir James Grigg, Chair, Cmnd. 9163. Referred to in the text as "Grigg Report."

Zimmerman, F. (1959) Wesen und Ermittlung des Archivwertes: Zur Theorie einer Archivalischen Wertlehres, *Archivalische Zietschrift* 54, 104—107.