

The Concept of Record in the Digital World: The View of InterPARES

Luciana Duranti

The InterPARES Project is an international research endeavor that aims at developing the theoretical and methodological knowledge essential to the long-term preservation of authentic records created and/or maintained in digital form. This knowledge should provide the basis from which to formulate model policies, strategies and standards capable of ensuring the longevity of such material and the ability of its users to trust its authenticity. InterPARES has developed in two phases. InterPARES 1 dealt with records mandated for accountability and administrative needs that are created in databases and document management systems. InterPARES 2 has focused on the portion of society's recorded memory that is digitally produced in interactive, dynamic and experiential systems in the course, and as a byproduct, of artistic, scientific and electronic government activities.

The InterPARES research team determined at the very beginning of the first phase of the project that, in order to be able to identify among the various types of digital information contained in existing systems what corresponded to a record, it was necessary to agree on the definition and concept of record and on how such entity differed from document, information and data. Thus, the team called a **record** any document created (i.e., made or received and set aside for action or reference) by a physical or juridical person in the course of a practical activity as an instrument and by-product of such activity, thereby adopting the traditional archival definition. The team then proceeded to define **document** as recorded information, **information** as a message

intended for communication across space or time, and **data** as the smallest meaningful piece of information. Finally, an **electronic record** was defined as a record created in electronic form, meaning that a document received in electronic form but set aside for action in paper form is a paper record, while a document received on paper but scanned in the computer and only used as a digital file is an electronic record. This definition is fully consistent with the archival principle that whatever the creator treats as a record in the course of any given action is indeed a record of the action in question.

After having chosen a definition of electronic record based on traditional archival theory, the team considered essential to delve into the concept of record underlying and implied by such definition by determining what the necessary characteristics of an electronic record are on the basis of both archival theory and diplomatic theory. The support of diplomatic theory was considered necessary because diplomatics is a discipline that examines records as items rather than as part of aggregations and is therefore able to complement the knowledge provided by archival theory for the purpose of identifying the characteristics embedded in the record itself rather than in its relationships. The following necessary characteristics were identified: 1) a fixed form, meaning that the entity's binary content must be stored so that it remains complete¹ and unaltered, and its message can be rendered with the same documentary form it had when first set aside; 2) an unchangeable content;² 3) explicit linkages to other records within or

1 Completeness here is not mentioned as a characteristic of the record, because an incomplete record is still a record, albeit a bad one, but as a characteristic of a fixed form, according to which a form that is fixed is one that does not lose any of its original elements in the process of being stored and retrieved.

2 The stability of the record, as determined by its fixed form and its unchangeable content, is only implied in the part of the archival definition that reads that a record is a document (i.e. rather than just data or information), but it is explicitly stated in the diplomatic definition and concept of record (see Luciana Duranti, *Diplomatics. New Uses for an Old Science* (Lanham, Maryland, and London: The Scarecrow Press, Inc., The Society of American Archivists and the Association of Canadian Archivists, 1998), pp. 41-58.

outside the digital system, through a classification code or other unique identifier;³ 4) an identifiable administrative context; 5) an author, an addressee, and a writer; and 6) an action, in which the record participates or which the record supports either procedurally or as part of the decision making process.⁴

Having specified the necessary characteristics of an electronic record, the research team decided to accept as a working hypothesis the fundamental assumption of diplomatics that, regardless of differences in nature, provenance or date, from a formal point of view, all records are similar enough to make it possible to conceive of one typical, ideal documentary form containing all possible elements of a record. On the basis of this assumption, the team itself hypothesized that, while they may manifest themselves in different ways, the same formal elements that are present in traditional records exist either explicitly or implicitly in electronic records, and that all electronic records share the same formal elements. Thus, it created a template, that is, a decomposition of the ideal electronic record, first, into its constituent parts, and then, within the part “form,” into its elements.⁵ In the template, the parts and elements are defined and their purpose is explained. The research team used the template as an instrument for the systematic

³ This characteristic corresponds to the archival bond, which is implied in the archival definition when records creation is linked to an activity, but it is made explicit by archival theorists of all times and cultures. See Luciana Duranti, “The Archival Bond,” *Archives and Museum Informatics* 11, nos.3-4 (1997): 213-218.

⁴ While characteristics 4 and 6 can be deduced from the archival definition, characteristic 5 derives from the diplomatic concept of record: it was considered important in order to distinguish records from digital entities resulting from simply querying a database. The author is the person issuing the record, the writer is the person determining the articulation of the discourse in the record, and the addressee is the person for whom the record is intended. As a record must participate in an action and any action must fall on somebody, the addressee is necessary to the existence of the record. See the Appendix 2 of the book *The Long-term Preservation of Electronic Records: the InterPARES Project* on the InterPARES web site www.interpares.org [<http://www.interpares.org/book/index.cfm>].

⁵ The reason why the constituent parts of the record ended up in the template that is supposed to represent the ideal form of a record is that all identified constituent parts used to be regarded as necessary extrinsic elements of form by traditional diplomatists. It was important to show their presence, definition and purpose, and the fact that they are now independent of form.

analysis of the electronic entities contained in several different systems, for the purpose of establishing which ones are records.

The template is composed of four sections corresponding to the four necessary constituent parts of every record: documentary form, annotations, context, and medium.⁶ The documentary form⁷ includes, among the intrinsic elements,⁸ the names of the persons concurring to the creation of the record, the chronological date, the place of origin of the record, the indication and description of the action or matter, the attestation, and a statement of validation; and, among the extrinsic elements,⁹ overall presentation features (e.g. text, image, sound, graphic), specific presentation features (e.g. layouts, hyperlinks, colors, sample rate of sound files, resolution of image files, scales of maps), electronic signatures and seals (e.g. digital signature), digital time stamps, and special signs (e.g. digital watermarks, organization crest, personal logo).¹⁰

The annotations¹¹ fall into three fundamental groups: 1) additions made to the record after its creation as part of its transmission (e.g. priority of transmission, date of compilation and date of transmission in an e-mail record, the indication of attachments),

6 In a previous research endeavour commonly known as the UBC-DoD project, the parts constituting the records were identified as: medium, form, action, persons, archival bond, content and context. See Luciana Duranti and Heather MacNeil, "The Preservation of the Integrity of Electronic Records: An Overview of the UBC-MAS Research Project," *Archivaria* 42 (Spring 1997): 46-67; and Luciana Duranti, Terry Eastwood and Heather MacNeil, *Preservation of the Integrity of Electronic Records* (Dordrecht, Kluwer Academic Publishing, 2002: Chapter 1. In the context of InterPARES, it was decided that action, persons, archival bond and content, contrary to the other constituent parts, continue to manifest themselves in formal elements and are inextricable from them, so they do not have to be identified separately from the form. As it regards the annotations, which were not among the parts identified in the MAS-DoD project, they were added to the constituent parts because they are often linked to the record rather than embedded in it, and need therefore to be looked at separately from the record form.

7 Defined as "The rules of representation according to which the content of a record, its administrative and documentary context, and its authority are communicated.

8 Defined as "The elements of a record that convey the action in which the record participates and its immediate context.

9 Defined as "The elements of a record that constitute its external appearance."

10 See Authenticity Task Force, "Template for Analysis," in *The Long-term Preservation of Electronic Records: the InterPARES Project*, cit. <<http://www.interpares.org/book/index.cfm>>.

11 Defined as "Additions made to a record after it has been created."

2) additions made to the record in the course of handling the business matter in which the record participates (e.g. date and time of receipt, action taken, name of handling office), and 3) additions made to the record in the course of managing it as a record (e.g. filing date, class code, registration number). The categorization of the contexts of the record¹² and the list of what would reveal them correspond to an hierarchy of frameworks that goes from the general to the specific: 1) juridical-administrative context (manifested in, for example, laws and regulations), 2) provenancial context (manifested in, for example, organizational charts, annual reports, tables of users in a database), 3) procedural context (manifested in, for example, workflow rules, codes of administrative procedure), 4) documentary context (manifested in, for example, classification schemes, records inventories, indexes, registers), and 5) technological context (manifested in, for example, hardware, software, system models, system administration).¹³

The medium¹⁴ was difficult to place within the template, because, although it is still necessary for an electronic record to exist, it is not inextricably linked to the message, does not store the record as such, but a bit-stream, and its choice by the record-maker or keeper can be either arbitrary or based on reasons related to preservation rather than to the function of the record. In addition, the medium is not a relevant factor in assessing a record's authenticity—one of the primary purposes of InterPARES—at least from the perspectives of the creator and of the record preserver.¹⁵ This was confirmed by the case studies undertaken by the research team, by the end of which the team was

¹² Defined as “The framework of action in which the record participates.”

¹³ For details related to annotations and contexts, see the Template for Analysis referenced above.

¹⁴ Defined as “The physical carrier of the message.”

¹⁵ An additional reason for the InterPARES team to dissect the concept of record was to identify what parts or elements contribute to the authenticity of the record and to the ability to verify it.

convinced that, with electronic records, the medium should not be considered a constituent part of the record but a part of the record technological context.

The analyses of the case studies conducted using the template indicated that only a half of the examined systems contained records (twelve out of twenty-two), primarily because the entities identified within the other half did not appear to possess either a fixed documentary form or a stable content. When systems did contain records, these could rarely be compared with the model represented by the template, because, although they were able to achieve their purposes, they were not good records. For example, in most systems, there was no explicit manifestation of the relationship among the records participating in the same affair or matter, and, although it was easy to identify the business processes supported by the system, it was not always possible to determine how the records participated in or supported specific actions. In addition, it was often difficult to determine the significance of the presence or absence of given elements of documentary form or of annotations.

More importantly, the case studies showed that, with digital records, a key concept to consider is that of *records attributes*, which are the defining characteristics of each given record or of a record element in it. A *record element* is a constituent part of the record's documentary form and, as seen earlier, may be either extrinsic, like a seal, or intrinsic, like the salutation.¹⁶ An attribute may manifest itself as one or more record elements. For example, the name of the author of a record is an attribute, which may be expressed as a letterhead or a signature, both of which are intrinsic elements of documentary form, that is, record elements. In addition to attributes that manifest themselves in the form of the record, that is, on the face of the record, as record elements,

¹⁶ A defining characteristic, or attribute, of the record element "seal" may be its legend.

every record has attributes that are implicit in other parts of the record, such as the name of the creator or of the medium, but in digital records they are also expressed, albeit outside the documentary form. Because of this, they are mostly transparent to the user, and manifest themselves as metadata included in either a record profile,¹⁷ a topic map, or other digital entity linked to the record. Attributes made explicit outside the record as metadata demonstrating its identity are important to uniquely identify any digital record, but they become the primary means for the identification of digital objects that do not have—or at least for as long as they lack—a stable content or a fixed form. This will become clearer later on.

As if the distinction between record elements and attributes were not sufficiently complex, with electronic records we also have to differentiate elements and attributes from the record's digital components. A *digital component* is a digital object that may contain all or part of a record, and/or the related metadata, or more than one record, and that requires specific methods for preservation.¹⁸ For example, an e-mail containing a textual message, a picture and a digital signature has at least four digital components that are stored in different part of the system, although they are linked among themselves, and require different protection measures: the header, the text of the message, the picture and the digital signature. In contrast, a report with textual attachments may consist of only one digital component. In other words, a digital component is a unit of storage, but one that needs to be identified when the concept of digital record is dissected.

¹⁷ A record profile is an annotation inextricably linked to the record that includes several fields, which are either automatically or manually filled in with the record's metadata.

¹⁸ See Preservation Task Force Report in *The Long-term Preservation of Electronic Records: the InterPARES Project*, cit. <<http://www.interpares.org/book/index.cfm>>.

Finally, the InterPARES 1 team felt the need to point out that the relation between a digital record and a computer file can be one-to-one, one-to-many, many-to-one, or many to many, thus we should never use the terms record and file interchangeably; that the same presentation of a record can be created by a variety of digital presentations and, vice-versa, from one digital presentation a variety of record presentations can derive, thus fixed form does not imply that the bit streams must remain intact overtime; and that it is possible to change the way in which a record is contained in a computer file without changing the record, thus the name of a record form does not necessarily indicate what digital object we are dealing with.¹⁹

The concept of digital record presented above, with all its characteristics, parts, formal elements, attributes and digital components, has worked quite well with databases and document management systems. However, it may appear problematic when applied to the entities examined by InterPARES 2, the most salient characteristic of which seems to be, as mentioned earlier, the lack of a stable content and/or fixed form, not because they are bad records, as it was often the case with the digital entities examined during InterPARES 1, but because fluidity is part of their nature and contributes to the accomplishment of their purpose as instruments of or support for action. They are experiential, interactive and dynamic records.

Experiential records are electronic objects the essence of which goes beyond the bits that constitute the object to incorporate the behavior of the rendering system, or at least the interaction between the object and the rendering system. Defining the characteristics, parts, elements, attributes and components of such objects is much more complex than with traditional electronic records, because it is dependent not only on the

¹⁹ Ibidem.

object per se, but on the environment in which the object is experienced. Examples of experiential digital objects range from audio and moving images embedded in a web page to virtual reality systems.

Interactive records are records made and maintained in interactive systems, where each user's entry causes a response from or an action by the system. To determine the boundaries of such records (i.e., where one record ends and another begins), when they can be considered finished rather "in progress", when they are complete rather than partial or incomplete, etc., one needs to ascertain 1) how user input affects the creation and form of each record (as is the case with much on-line commerce); and 2) if and when the interactive system and its inherent functionality are to be regarded as meaningful parts of the record. Examples of interactive systems range from web pages delivering government services online to musical performances based on human-computer interaction to commercial video games.

Dynamic records are documents whose content is dependent upon data that vary continuously and are held in several databases and spreadsheets. Examples range from simple web pages with embedded links to complex systems where information is stored and updated in order to be shared via wireless transmission by multiple mobile users in diverse ways. The increasing reliance on such documents by individuals and institutions will necessitate understanding how the information they contain is captured and saved.

Whether experiential, interactive, and dynamic digital objects are records primarily depends on their relationship to the activity of their creator. It is out of question that these objects must be subjected to same kind of scrutiny which any document undergoes in relation to the action and procedure in which they participate, the archival bond they have

with other records of the same creator, etc., in order to establish whether they are records or not. However, even if they pass such scrutiny with respect to these fundamental requirements, the question of form looms large. Is it possible to have a record in fluid form and with undetermined boundaries? If not, should an entity with fixed form be generated for the purpose of making a record to be kept in a trusted recordkeeping system and perhaps preserved over the long term? If yes, who should make it? On the basis of which criteria? When in the entity's lifecycle?

The InterPARES 1 Authenticity Task Force's report suggests the possibility of trading the record characteristics of stability of content and fixity of form (including completeness of content and form with respect to the first and to any subsequent instantiations of the record) with the ability of the system containing it to track and preserve any change to the record. In other words, the Task Force was inclined to shift the requirements of stability and fixity from the record to the log of the changes to the record once the record was no longer active;²⁰ in this context, the entity identified as the record and to be kept intact would then be the last instantiation of the fluid entity, plus the complete log of changes, and the metadata of both. This option is conceptually sound only if the creator uses this set of entities as its record, but this scenario is very unlikely because it would be highly impractical. Alternatively, one might look at each digital object participating in the creator's activity as an instrument and by-product of it as existing at any given time in one of two modes, as a record *in fieri*, that is, in becoming, when its process of creation is ongoing, that is, when the entity is accessed to add data or information, and as a record when the entity is accessed for use. This would imply the stabilization and maintenance of every instantiation accessed for use and its metadata.

²⁰ Here, active is used to mean "subject to changes or additions."

Conceptually, this option is as sound as the previous one, but it appears to be equally impractical.

One other option can be developed from the findings of the case studies carried out in the context of the artistic and e-government activities, results that are remarkably similar. The records of each individual activity that has been examined both comprise and are each composed of a mix of analogue and digital entities interacting among themselves, often with the mediation of human beings, instruments and/or computer technology. This situation presents issues of three types: 1) issues related to the maintenance of each digital object, be it larger, smaller or equal to a record, in a way that its accuracy/reliability and authenticity can be ensured; 2) issues related to the maintenance of the relationships among analogue and digital entities, and of the ability of the various digital objects to interact with each other, with or without human or technological mediation, both within a record and between records, in precisely the same way in which they were meant to interact when generated; and, most important in the context of this paper, 3) issues related to the identification of the boundaries of the entity record.

Although it is theoretically improper to base the identification of an entity on the requirements for its preservation beyond its active state, it is methodologically justifiable in the context of InterPARES, which 1) aims at finding solutions to the long-term preservation of digital records (hence, the entity we identify as a record has to be preservable), 2) has determined that the only way of preserving digital records is to produce authentic copies of them (hence, the entity we identify as a record must have a definitive complete instantiation—a state of stability beyond which no change will occur-

- which is the entity of which authentic copies would be made), and 3) has stated that the chain of preservation begins at creation (hence, the entity we identify as a record at creation should be the one that we can preserve). In this context, it is important to remember that InterPARES 1 has clearly formulated the difference between the copies of record made by the creator in the course of business and for the purposes of its business (be they the reproductions generated every time a record is retrieved or the copies made as a result of system upgrade or record migration), which it has called “the records of the creator,” and the copies of records made by the preserver in the course and for the purposes of archival functions, which it has called “the authentic copies of the records of the creator.” This distinction is vital for identifying the entity record, because it means that, if the creator alters the form of the record in order to be able to keep it for future action or reference, the result of such alteration is a record of the creator, and the preserver will have to carry forward an authentic copy of the last instantiation of that record as produced by the creator, including the metadata attesting to the changes. Finally, to arrive at the articulation of a concept of record on the basis of conceptual requirements for preservation is justifiable because we would not be defining records, something that the archival and diplomatic sciences have already done quite satisfactorily, as aggregates the one, and as items the other:²¹ What we would be doing is

²¹ The archival and diplomatic traditional definitions have served us quite well through the centuries because they have enough rigour to show a clear demarcation between an entity that is a record and one that is not a record, and enough flexibility to be applicable to entities produced in all the technological, administrative and cultural environments that have existed to date. All attempts to refresh those definitions by making them more specific have in my opinion miserably failed. For example, the definition coined by the *Guide for managing electronic records from an archival perspective*, issued by the International Council on Archives Committee on Electronic Records (ICA Studies, February 1997), which, at p. 22, reads “A record is recorded information produced or received in the initiation, conduct or completion of an institutional or individual activity and that comprises content, context and structure sufficient to provide evidence of the activity,” is at the same

to develop a description of what entity the archival and diplomatic definitions concretize themselves into in the context of a dynamic, interactive or experiential digital environment. With this understanding, we can proceed to discuss the three types of issues identified earlier.

The first type of issues, related to the maintenance of each digital object, may appear easy to solve by using migration. However, migration of digital interacting entities existing in different formats often makes their interaction impossible, alters the functionality of the entities, and results in partial, inaccurate, unreliable and inauthentic reproductions.²² The difficulty of the second type of issues, related to the maintenance of the relationships among analogue and digital entities, derives from the fact that those relationships, as well as the interaction among the digital entities in question are usually not documented in a way that makes it possible to re-enact them in a different environment or when one of more of the digital entities is upgraded. To overcome this problem requires the development of a special kind of notation for arts material and of metadata schemata for e-government material that are capable of describing in an objective, detailed and standardised way the interaction between the record's digital and analog components, a digital record and another, and the record's components or the records themselves with the mediating entity, so that such interaction can be accurately reproduced. The third type of issue, related to the identification of the boundaries of the entity record, is to be solved case by case, but on the basis of a renewed understanding of

time ambiguous and inflexible, and would certainly exclude from the record category all dynamic, experiential and interactive records.

²² Migration as a method of maintenance and preservation will present us with this sort of problem for a long time, at least until we have developed a sense of what change in a record is acceptable to the point that we can still say that, regardless of it, the record has preserved its identity and integrity. With paper we know, on the basis of centuries of experience, how much damage a record can tolerate to be considered intact, or how different a copy can be from the item it reproduces to be considered authentic, but with the digital medium, we have to define parameters and develop standards.

what is implied in the definition of record, an understanding that must be linked to the answer to the previous two issues.

The case studies completed in the area of the performing arts are helping to find such answer. With music, the work is considered to be the performance, while the score is regarded a set of instructions that allows performers in different times and places to reproduce the same work.²³ Each performance is a little different, depending on how detailed the score is, and therefore on how much discretion is left to the performer in interpreting the piece, on the ability of the performer, on how the related musical instruments have changed over time, on the acoustics of the place where the performance occurs, etc., but it is close enough to the original work to be easily identified by the audience for what it is. In other words, although the original performance cannot exist as a live work of art longer than its manifestation, the presence of a score ensures the accuracy and authenticity of the live performances that will follow. However, with electronic music, it is becoming quite clear that the set of instructions included in the score, when it exists, is not sufficient to reproduce the piece: one needs to have also the computer codes, the patches, a synthesiser when used, and the interaction between the performer(s) and all of the above, an interaction that so far has never been described as a reproducible set of instructions. Increasingly, both composers and InterPARES researchers are arriving at the conclusion that the only way of keeping digital music as a record is to describe each component of it and the interactions among them, that is, to produce a set of instructions for re-creating each part of the piece and the piece as a whole.

²³ There is some disagreement on whether the score is also a work in its own right, but this does not invalidate the option I intend to present. If anything, it supports this option.

Furthermore, through case studies of visual art, InterPARES is beginning to advance the proposition that, in the digital world, every art form is becoming performing art in character, in that it can only be manifested over time by re-creating it on the basis of a record made up of a set of instructions, rather than by migrating or even emulating its components and hoping that they will be able to behave as their first instantiations. This proposition is supported by several initiatives, separate from InterPARES, which have tried to solve the preservation problems presented by digital objects that are characterized by their capacity of evolution, their interactivity, their dependence for form and content from external factors, and the centrality of movement and multimediality to their meaning. The fundamental issue these initiatives are confronted with is whether to preserve these objects or maintain them live. To preserve them means to choose between fixing them in a definitive form with one final act of interpretation and representation, and treating them like a musical score, by creating a description of them, a document that opens to the possibility of generating future reiterations of the work. In the former case, it would be necessary to identify as the work a collection of, for example, installations, machines, software, Internet screens, videos of interactions among the parts, etc. This would imply loss of causality, dynamicity and artistic experience. In the latter case, the act of interpretation would be left to the performer or the user in general, accepting the fact that each act of interpretation will have a different result, although always identifiable as the same work. A project that tries to find a compromise between the two preservation alternatives is the Rhizome ArtBase. This project has developed two concepts, that of “connected art object” and that of “cloned art object.” The former comprises the description of the work and its components, a thumbnail of the work,

keywords and metadata, a link to the URL of the work, the biography of the artist, and his/her certification that this aggregate of parts corresponds to the work and constitutes an adequate representation of it. The latter includes, in addition, an authentic copy of the work preserved in the server of the project.²⁴

The project Rhizome ArtBase brings to light the most important issue presented by the identification of the work to be preserved, that of artistic intention. Lacking the possibility of carrying forward into the future an intact work of electronic art, one is left with the option of capturing the essence of the work, but this implies that the author and/or creator become active participant in the preservation act and state in explicit and authentically preservable way that the entity being preserved is the substance of the work in question. The authenticity of the work is ensured by the personal involvement of its author/creator in the decision of how it will be re-created in the future. Its nature of record is ensured by the fact that the author generates this surrogate, or potential work, as a regular part of his/her creative activity and for its purposes: it is a very fine line to walk, but, as long as the creator does not produce this entity for the preserver, but for his/her own benefit, it is one consistent with the concept of record. Certainly, examining the list of the entities comprising the connected and the cloned object, it appears evident that, from their sum it is not possible to generate a work identical to the original one, but its essence would be captured. In InterPARES terms, to equate the concept of work of art with the objects described above means to give preference to continuing accessibility over accuracy (obviously, reliability is maintained by the control of the author/creator on

²⁴ Alena Williams, "Rhizome.org," in Alain Depocas, Jon Ippolito, and Caitlin Jones eds., *Permanence Through Change. The Variable Media Approach*, New York and Montreal: Guggenheim Museum Pub. And The Daniel Langlois Foundation for Art, Science and Technology, 2003, p. 39.

the process) without compromising the spirit of authenticity, given that the identity of the work is ensured and its integrity²⁵ can be (potentially at least) reconstituted.

And here is a hypothesis that I wish to propose: With regard to the records resulting from e-government (and possibly e-science) activities, we might be dealing with the same kind of scenario presented by the digital works of art. With e-government records, a few recurring features are likely to occur. For each service digitally delivered by a government to a citizen in an interactive mode, there will be a record spread across several interacting technologies, a record that has no clear boundaries, and changes continuously on the basis of the input of the user (either the government officer or the citizen) and/or of the reaction of the system to such input, and a record that rarely corresponds to one action and more often includes the whole interaction between a government office and a citizen with respect to one matter (i.e., what used to be a paper file, a dossier that is). Thus, one will need to identify 1) the boundaries of the digital entity that constitutes the record which, once made or received, and repeatedly set aside in different instantiations, is to be kept for future action or reference in a trusted recordkeeping system, 2) the essence of such entity, i.e., keeping in mind the acceptable degree of change from the moment the record has reached its final instantiation, its constituent parts and digital components to be kept stable as content and fixed as form and linked among themselves, 3) its attributes to be manifested in metadata permanently attached to the record as an annotation (thereby becoming a constitutive part of the record), and 4) the necessary accompanying documentation of what is not fully preservable, that is, interactivity, connectivity, and functionality. Once this identification

²⁵ In the Authenticity Task Force Report, cit., integrity is taken to mean that the substance of the message is conveyed in the same intellectual form as its first instantiation.

has occurred, one could assemble the stabilized essence of the record, its metadata and system documentation and treat this entity as the record. Would such procedure be acceptable if it were the interest and responsibility of the creator to carry it out? Is such an idea contrary to the theoretical understanding of what is a record?

Although, as stated in the InterPARES 1 Authenticity Task Force report, the study of new record types seems to indicate that not always what is known can guide to the understanding of what is unknown, I believe that we should keep looking for past situations that can be related to each present situation that one is observing. Certainly, there have never been in the past interactive records such as those resulting from e-government activities, but in Medieval times offices have kept records “attributes” in such a way that, when a finished record did not exist, complete and effective records could be created at will years later. I am not thinking of record metadata, which have also been generated for centuries in form of “*regestum*”, because they existed either in addition to the record or as its surrogate once the record had been destroyed, in either case for the purpose of proving the existence of the record, not of producing it when it was needed some time in the future. I am rather referring to the “*imbreviaturae*” of medieval notaries. When the notaries became so powerful as a profession that most transactions had to be recorded and preserved by them, they did no longer go through the trouble of writing out the records of the transactions that they witnessed. They would take a parchment, fold one corner forward, and write on it the transaction type, the names of the parties, the date, the description of the transacted property or matter, and any other data specific to that transaction. Then, they would file away the blank parchment with the

annotated corner²⁶ and, at the end of each year, bind all the *imbreviaturae* of the year in the same volume, and index the volume and or keep a separate registration of the occurred transaction in a book of *regesta*. If, later on, one or more of the parties to that transaction or their descendants wanted the complete record of the transaction, the notary would find the volume containing the *imbreviatura* in question by date, retrieve the document in it through the index or the register, take a new piece of parchment (or paper, if appropriate), and write out the record with the data written on the *imbreviatura* corner and the formulas contained in a special book, called *formularium*, which provided clear instructions for writing out a record for each type of transaction that occurred in a specific range of years. Thus, what the notaries really maintained was not the complete record of each transaction, but a record of the content of the transaction and of the documentary form in which it had to be manifested, and the ability to produce a complete record upon request by integrating content and form. In other words, they kept a record of the fact that a transaction had occurred (register and/or index), the data of the transaction (*imbreviatura*), and sets of instructions guiding them to make the accurate and authentic record of the transaction when needed (*formularium*), even centuries later, as each notary left its archives to its legitimate successor. However, precisely because of this trust, almost never a party or its successors requested that a complete record be issued: the existence of the *imbreviatura* in a notary archives was sufficient evidence of the transaction.

From the observation of the *imbreviaturae*, one can imagine a similar way of keeping the interactive records of e-government activities across technologies: at the completion of each transaction, the handling office/officer would, as a matter of routine,

²⁶ Sometimes, rather than on a corner, they would write the data on the back of the medium.

separate the data of the record from its form and technological environment, stabilize the former and the metadata of the original record, and link them to the latter by means of a description of the original form and functionality. As with the *imbreviaturae*, most times, this set of documents of a transaction, properly registered, would constitute a record adequate to serve both administrative and historical accountability, as well as legal purposes. In the few cases in which a complete and finished record of the transaction would have to be re-produced, it would likely be sufficient to embed the data in the appropriate record form, and accompany this record with the description of the functionality of the original digital environment. The fundamental difference between e-government records and the *imbreviaturae* is that interactive records come into existence as complete and effective records at the end of the interaction between the government and the citizens, however abstract the concept of complete can be with a record that is live with its original functionality, while the *imbreviaturae* only exist as initial rough drafts of a potential record. Thus, while the record generated from an *imbreviatura* would be created as a first instantiation of a record, its original, that is, the record generated by re-producing the e-government record after it had been taken apart in order to set its components aside as fixed entities, would be created as an authentic copy of the original interactive record. However, because the creator would use it in the usual and ordinary course of business, for further action or reference, this authentic copy would be considered a record of the creator, as discussed earlier in this paper.

This hypothesis is, however, workable only on the assumption that, upon completion of the interaction between the parties, the finished entity will be the exclusive responsibility of a trusted custodian similar to a notary, that is a person who has no stake

in the content of the record and can therefore fulfil the role of a third neutral party, and a person who is formally recognized competent to maintain the record because of his/her professional qualifications—a records officer, that is. This person would be the one making the surrogate of the interactive record, keeping it accessible to the competent offices in a trusted recordkeeping system, and generating the complete re-production upon request.

Thus, in the scenario depicted above, and keeping in mind the accepted definition of record, to what entity does a record correspond? I would suggest that, while the business procedure is active, the interactive digital entity constitutes the overall record of the transaction. If instantiations accessed for use by the parties at each given time are set aside with a fixed form and a stable content and linked to other records of the same transaction, they are also records of the transaction. Once the business procedure is concluded, the final record of the transaction will consist of the data contained in the last instantiation of the interactive entity²⁷ and its metadata, properly linked to an exemplary of its form and a description of its digital environment (i.e. record functionality and system documentation) that would already be maintained in the recordkeeping system to which such record will be transferred. The key to the reliability, accuracy and authenticity of such record will be the fact that the responsibility for generating and maintaining it as the source of future re-productions of the original interactive entity in its active state will reside with the creator (and, within the creator, with the record office):

²⁷ Assuming that no data would be deleted in the course of the transaction, as good record making would recommend. If data were deleted as a matter of course during the transaction, a log of changes would have to accompany the data contained in the last instantiation. If it were part of the formal procedure to delete data in the course of the transaction, the record of the transaction would be complete without the log, but the office would have to keep a description of the procedure linked to the series of records subject to it in order to account for the deletions.

the record-source will be the record of the creator just like the interactive entity was. If a re-production of the original interactive entity from the record-source will be made by the creator for its purposes, such re-production will still be the record of the creator, while, if it will be issued to an external user for other purposes either by the creator or by the preserver (if the record-source were transferred to an archives), such reproduction will be an authentic imitative copy of the record of the creator.

Of course, this option only addresses situations in which e-government creates dossier-like digital interactive records. But InterPARES 2 case studies have dealt with more complex situations. One example will suffice to show similarities and differences: the VanMap case study. VanMap is a GIS system that allows the City of Vancouver to meet the needs of internal users in providing services to Vancouver's citizens and businesses. VanMap supports the functions and activities of the following departments: Community Services Group, Engineering Services, Corporate Services Group, Board of Parks and Recreation, Vancouver Police Department, Fire and Rescue Services. The decisions on the layers and the set of data they should contain are made collectively by the departments and the VanMap Technical Team. Data are uploaded by each department directly in Oracle Spatial or taken as extracts from external offices databases (for example, permit and license data stored in PRISM or License+ are extracted to a SQL server; property tax data are extracted from the SQL Property Tax System, etc.) for inclusion in VanMap by the Technical Team, which is responsible for its administration. Engineering and CSG graphics are created in the form of CAD drawings in AutoDesk, or keyed or drawn in the Oracle Spatial database. VanMap data are overwritten at each

update and, every once in a while, existing layers are modified to receive different kinds of data sets and new layers are added.

The VanMap, as an indivisible whole and in the context of each of the numerous business processes in which it participates, perfectly corresponds to the archival concept of record in that it is treated by the creator as a record, it is linked to the other records participating in each business process by a documentary and procedural relationship, and it is the byproduct and residue of the transaction of affairs. As such, besides exhibiting all parts and characteristics identified by InterPARES 1, except of course stable content and fixed form,²⁸ VanMap manifests the traditional archival characteristics of naturalness, impartiality, interrelationship, authenticity, and uniqueness in context.²⁹ However, this conceptual recognition is logical, clear and of no consequence while one focuses on the VanMap as a live system, but it becomes very much clouded when one focuses on each business process and tries to identify the records participating in it. Would the whole VanMap be a record in each of the very numerous business processes carried out by the city departments that use it? Unless the instantiation of VanMap and, within it, the specific data layers and data sets participating in each given business process have been extracted from the system, frozen and, as a matter of course, attached in such form to the

²⁸Following the VanMap diplomatic analysis, it might be argued that, although its digital components may undergo dramatic changes every so many years, the documentary form of VanMap is quite stable, because its intrinsic and extrinsic elements do not change other than in their content.

²⁹A discussion of these characteristics can be found in Terry Eastwood, "What is Archival Theory and Why is it Important?" *Archivaria* 37 (Spring 1994): 122-130. As it regards uniqueness, it should be noted that, in the case of VanMap, it is also present with regard to content, as the information provided by VanMap and resulting from the combination of data sets originating in different departments and offices, does not exist anywhere else, although many of its data sets exist either as such, as part of records, or as isolated data in individual offices.

A reduced copy of VanMap accessible to the public at large is kept outside the city firewall to ensure that users do not either accidentally or maliciously see layers that are confidential within the VanMap for city staff. This digital entity is to be regarded as a publication and does not have any of the archival characteristics of the VanMap used by the city employees. Of course, if a user were to download the public VanMap to its system and use it in the course and for the purposes of its business, that specific downloaded entity could be a record in the user's fonds.

related records—a routine that nobody uses—the observed entity, which existed for the time necessary to carry out a transaction with a citizen, or to inform somebody of a given situation, or to make decisions or plans, disappears. Each instantiation was “made” at each update of the data, and “received” at each use, but never “set aside,” thus, it was only a potential record that never materialized. As a consequence, while we can still regard the whole VanMap as a record of the City of Vancouver collectively generated by its staff, it is not a record in the context of any given business transaction.

This situation should be rather disturbing to the City Council though, because its ability to account to the citizens for the actions of the city staff that affect them, both individually and collectively, is greatly diminished by the inability to demonstrate the factual grounds of city decisions.³⁰ Thus, as archivists responsible for advising creators about proper recordkeeping practices, we could imagine a solution capable of supporting both current and historical accountability: we would recommend the VanMap Technical Team, which includes representatives of the city departments, to develop a detailed description of each business process in which VanMap is involved and of the way in which VanMap is used in each of them, thereby revealing the archival bond between the records of each business process and VanMap. It is a centuries-long tradition to embed in a code of administrative procedure the function of a record that serves multiple activities and procedures, but of which only one original exist (see for example the series of the maps of the cadastre, which were and are used as records in several procedures having different purposes). Thus, this solution, per se, is not problematic as a principle or as a practice. What is problematic is that VanMap does not contain historic data, as it is

³⁰ This is where the uniqueness of the VanMap content comes into play because it is the coexistence and interaction of data of different, both internal and external, origin that makes VanMap a vital source for making decisions, defending and proving their factual basis.

constantly updated by overwriting superseded information. This means that, if a citizen who has been denied a building permit, for example, knowing the procedure and the part that VanMap has in it, requests access to the information contained in VanMap when his application was submitted, his request cannot be satisfied. Thus, this proposed solution would have to be accompanied by some additional procedure. One could recommend that each staff member using VanMap in the course of a transaction freeze the view related to each decision and save it as an attachment to the records of the business process in question. But this procedure is laborious without supporting the work of the staff members in any other way, therefore is not likely to be followed. Another, more effective procedure that would indeed make the work of some, if not all, departments more accurate, reliable and efficient, would be to configure the system in such a way that every day, at the closing of business, a complete image of VanMap be preserved live and fully functional within it, with of course the related attributes attached as metadata.³¹ Of course, as these images would not be explicitly linked to any individual concrete business transaction, they would not be records. The record would be still the VanMap as a whole, but the presence in it of this historical stratification of data would make two things possible: first, accountability would be served, and second, the VanMap would become preservable as a record. The first outcome is quite obvious, while the second requires some explanation.

³¹ The VanMap manager, having discussed this option with the case study team, found it not only technically feasible but also presumably exciting for several departments who would make good use of an historical stratification of data, such as the planning department. The idea of this option came to members of the team working on a science related case study, where the records creator maintains in a live system not only all the data of astronomic observations, and the images putting them into the needed relations, but also a stratification of the images taken at the end of each day representing all data accumulated in the previous twenty four hours.

VanMap, in its present configuration, constitutes a record that is constantly in a state of becoming and cannot be preserved. However, if the creator, in the usual and ordinary course of its business and for the purposes of its business, decided not only to have an historical stratification of daily images, but to remove them on a regular basis, yearly for example, from the active system into a live and fully functional, but inactive system, separate from the first by a firewall, thereby setting them aside as an indivisible whole related by procedure to all the city business which needs to use it either for reference or for accountability purposes, then each year of the VanMap could be regarded as a record with stable content and fixed form, that is a finished record, fulfilling all the requisites implied in the archival definition of record.³² I

In conclusion, all InterPARES case studies have pointed to the fact that dynamic, interactive and experiential digital objects can only be records in becoming, potential records, that is. If the creator treats them as records, associates them with entities that are undeniably records, and do so in the course of activity and for its purposes, these objects only need a stable content and a fixed form to materialize themselves as fully finished records. If the acquisition of these two characteristics occurs as a matter of course at the hand of the creator, because the creator will need either to have records of the actions for which the written form is required by the legal system, to provide an account to itself and

³² Some participants in the professional discourse on the concept of record have reduced it to a polarization between those who think that data or recorded information are records and those who think that data files can never be regarded as records. This is indeed a false dichotomy, because every entity that becomes associated in the usual and ordinary course of affairs with an archival aggregation (i.e., dossier, series, fonds), be it digital or not, regardless of its original nature, is subject to the law that governs every universality, according to which every individual entity that becomes a member of a collectivity, subject its individuality to the nature of the collectivity and shares the attributes of all other members of it (Archives as *universitas* is a very old concept. *Universitas* is the term from which the word university derives, a word referring to an institution whose members share the same nature, rights and duties with respect to the *raison d'être* and the purposes of the institution). This is the reason why the VanMap used by the city staff is a record of the City, regardless of the fact that it is a collection of data, while the VanMap used by the public it is not.

to its stakeholders of the activities carried out, or to use the records itself for further action, future activities or reference, then the results will be records in all respects, by any analysis and standard (certainly from an archival, diplomatic, administrative, and legal point of view). A proviso is however required in the context of e-government and of business organizations: the stabilization of the content and the fixing of the form must be carried out by a neutral third party having no stake in the content of the records of any type of business transaction carried out by the creator, and who is competent in the science of the records. This is of course the definition of a record officer. On the contrary, as it regards the sciences and the arts, the author (be it an individual or a collective person), that is, the person who has the highest stake in the content of the record, would also be the most reliable person to make the record because of the nature and purposes of the scientific and artistic activities.

Meeting the challenge of identifying the record in the digital world is much more than establishing policies and procedures or developing metadata schemata. Every time a solution is proposed, more questions come up. Thus, I do not believe that a definitive answer that is valid in the context of all future technological environments can be provided, but what InterPARES 2 can do is to establish the conceptual underpinning, the parameters, and the method of analysis that will determine the answer to the question of what entity corresponds in a known given environment to the archival definition of the record, a definition that has survived the administrative and technological changes brought about by centuries of human activities and is likely to remain valid for the foreseeable future.

Committing the Web to Memory: Transmitting Web-based Records over Time

Jim Suderman

This paper will draw upon the findings of an InterPARES 2 case study on website exhibits. , , are at risk due to The interaction between the creator's platform and the rendering of the record by the user's platform results in a risk to the meaning and intent of web-based records, perhaps even more than to other electronic records The paper will examine processes that result in the creation and maintenance of web-based records in terms of record values appropriate to their context. These values define the characteristics of record authenticity, reliability and accuracy. From this base the paper will conclude with an exploration of the usefulness of web-based records as evidence over time.

The title of this paper, “Committing the web to memory,” plays on a phrase that represents what for me was a school activity – memorizing information, like Hamlet’s soliloquy, to produce later in a performance or a test – and something we probably do every day: save a document to computer memory to access it later. The inventor of the world wide web, Tim Berners-Lee, defines the web as “The set of all information accessible using computers and networking, each unit of information identified by a U[niversal] R[esource] I[ndicator].”¹ This definition focuses almost exclusively on the technology: computers, networks, universal resource indicators. The exception is the reference to units of information. Identifying Hamlet’s soliloquy by a URI may be effective for moving it around networked computers, but all other aspects of its cultural context must be supplied by the viewer. It is safe to say that Shakespeare did not compose the speech only for it to be accessed by networked computers. For this reason I suggest expanding the preceding definition of the web to qualify “units of information” by the phrase “developed for a specific purpose” and made accessible using computers and networking.

The principles that guide archival practice are designed to preserve the intent of the creator over time for the benefit of subsequent users of the creator’s records. The code of ethics

¹ From the glossary of Tim Berners-Lee with Mark Fischetti. *Weaving the Web. The Original Design and Ultimate Destiny of the World Wide Web by its Inventor.* (San Francisco: Harper, 2000). Definition available at <http://www.w3.org/People/Berners-Lee/Weaving/glossary.html> (checked 14 March 2005).

adopted by the International Council of Archives requires archivists to preserve and provide access to records in such a manner that protects “the integrity of archival material and thus guarantee that it continues to be reliable evidence of the past.”² This statement is consistent with a 1991 resolution of the European Council recognizing archives as having a twofold purpose, as “a basis for decision-making in the public sector on the one hand, and as a vital component of a nation’s cultural heritage on the other.”³ It is valid, then, to examine how archival principles operate in relation to the preservation of web-based records – governmental, scientific, or artistic.

The word “archives” in the European resolution is a reference to records: units of information with a specific purpose. To be properly understood records must be considered in the context in which they were created. For decision-makers to be accountable the records they use must be accurate and reliable. In terms of cultural heritage, de-contextualized information cannot act as a vital component because at the very least the cultural context must be established for records to act as a vital component of a nation’s cultural heritage. For these reasons the two purposes in the European resolution are useful starting points when approaching web-based records for the purpose of preserving and reproducing them in the future. For web-based information transmitted across space and through time to stand as evidence of the intentions and purposes of its creators requires preservation activities that identify the web-based content with its creators and their purposes and intentions.

² International Council of Archives. *Code of Ethics* (1996). Available at http://www.ica.org/biblio/code_ethics_eng.html. (Checked 14 March 2005.)

³ “Resolution of the Council and the Ministers of Culture, meeting within the Council of 14 November 1991 on arrangements concerning archives.” *Official Journal C 314, 05/12/1991 p. 0002-0002*. Available at [http://europa.eu.int/smartapi/cgi/sga_doc?smartapi!celexapi!prod!CELEXnumdoc&lg=EN&numdoc=41991X1205\(01\)&model=guichett](http://europa.eu.int/smartapi/cgi/sga_doc?smartapi!celexapi!prod!CELEXnumdoc&lg=EN&numdoc=41991X1205(01)&model=guichett). (Checked 14 March 2005.)

This paper examines some of the means by which the creator's intent and purpose can be ascertained and represented and how some current initiatives, in particular the Internet Archive, approach the preservation of web-based information. The first part draws heavily on the findings of a case study on the creation, maintenance and preservation of web exhibits conducted within the InterPARES 2 Project. This research project is examining authenticity, reliability and accuracy of records created within emerging digital systems.⁴ An envisioned outcome of the Project is a principle-based policy framework that will provide preservation guidance for creators and memory institutions responsible for preserving the accessibility of accurate, reliable and authentic records over time.

The case study focussed on web exhibits created by two Canadian archival institutions. Individuals fulfilling the roles of manager, website coordinator, scanning technician and exhibit curator were interviewed to highlight the purpose and process of creation, the technology involved, and the on-going record keeping requirements.

Both subject institutions are required to make their archival holdings known to as wide an audience as possible. Web exhibits are seen as an important and cost effective means to fulfil this responsibility and both institutions devote significant resources to the development of web content in general and web exhibits in particular. Although web exhibits are not transactional

⁴ This research project is examining authenticity, reliability and accuracy of records created within emerging digital systems and is directed by Dr. Luciana Duranti, University of British Columbia. See www.interpares.org for general information on the InterPARES 2 Project and its predecessor project. The InterPARES 2 Project proposal defines accuracy as "the truthfulness of the content of the record and can only be established through content analysis. With administrative and legal records, it is usually inferred on the basis of the degree of the records' reliability and is only verified when such degree is very low." Authenticity is defined as "the trustworthiness of a record as a record. An authentic record is one that is what it purports to be and has not been tampered with or otherwise corrupted. Authenticity is established by assessing the identity and the integrity of the record. It must be possible to ascertain at all times what a record is, when it was created, by whom, what action or matter it participated in, and what its juridical/administrative, cultural, and documentary contexts were. It must also be possible to ascertain the wholeness and soundness of the record: whether it is intact or, if not, what is missing." Reliability is defined as "the trustworthiness of a record as a statement of fact. It exists when a record can stand for the fact it is about, and is established by examining the completeness of the record's form and the amount of control exercised on the process

records in the same way as contracts or registrations, the study showed that considerable efforts are made to ensure that exhibits are accurate, authentic and reliable. The study identified three more or less concurrent processes in the creation of web exhibits: research, administrative, and technological. Each contributes to some aspect of ensuring that the exhibits were trustworthy records of the institution.

Research process

The research process for creating web exhibits is based primarily on scholarly research practices. Once an exhibit topic or subject is determined and approved, the curator reviews existing secondary literature. This review provides a basis for a preliminary narrative for the exhibit topic and assists in identifying items for inclusion in the exhibit. Items for inclusion are predominantly digital images of conventional, that is paper-based, records from the holdings of the subject institutions. Items are chosen on the basis of

- their authenticity
- their ability to represent a variety of narrative viewpoints
- their value for quotations that contribute immediacy and provide context to the narrative
- and visual appeal (e.g., maps or technical drawings).

The initial identification of items for use as sources includes compiling citations to assist with their review and final selection. At the same time as sources are being selected the narrative is refined and images of source documents, quotations and citations begin to be integrated.

Administrative process

The study identified two key managerial approvals. The first was the initial approval for the exhibit concept or topic. This approval clearly establishes institutional ownership of the

of its creation.” See Luciana Duranti, *InterPARES 2 Project Detailed Proposal*, p. 1.1-11, available at http://www.interpares.org/display_file.cfm?doc=ip2_detailed_proposal.pdf. (Checked 14 March 2005).

exhibit process and allows the research and technological processes to begin. The final approval, by the institutional head, marks the end of the process of creation; it thus marks the completion of the record, the institution's satisfaction with it and the acceptance of the institution's ownership and responsibility for it. Because exhibits often portray historical events, one consideration in the final approval is to ensure that events are not portrayed in a partisan or otherwise inappropriate fashion.

There are also approval steps during the process of creation. These govern factors such as the focus of the narrative, the selection of images and text, the allocation of resources – for example, approval is needed if identified source materials require extensive conservation work prior to scanning – and the 'look and feel' of the technological components of the exhibit. Interim approvals may also involve a review by other staff to help ensure topical accuracy or clarity of presentation.

Technological process

The technological process involves two sub-processes. The first is the creation of the digital components used within the web exhibit and is a process that is shared by the website coordinator and the scanning technician. The second sub-process is the creation of the web exhibit itself, normally the responsibility of the website coordinator.

Creating the digital components predominantly means scanning or digitising source materials into image files, but may also include the digitisation of video files or recording of sound files. The scanning process was developed independently of web exhibit creation as an accessibility and preservation initiative for photographic holdings, including negatives and rare formats, in both institutions. The scanning technician creates high quality tiff format⁵ images

⁵ Tiff stands for Tagged Image File Format, a popular, flexible and public domain raster file format. The tiff specification is available at <http://partners.adobe.com/public/developer/tiff/index.html>. (Checked 9 March 2005).

from which images in jpeg format⁶ are derived for use as web exhibit components. The technician's visual skills are supported by high quality scanning software and hardware, including calibrated monitors for correct display. In the creation of the digital components for web exhibits the same high degree of accuracy is sought as for the reproduction of photographic records for access and preservation purposes. The relationship of the scanned image to the source item and any other records it may be related to is maintained by the contextual data added to each image. Details of the scanning process, including technical settings, are likewise maintained as attached data elements. The contribution of the rigour of this process to exhibit components is analagous to that which the scholarly research process brings to the development of the exhibit narrative.

The second technological sub-process is the assembly of the digital components comprising the exhibit content into the actual exhibit. Exhibits consist of a number of linked, HTML-encoded pages; each page is comprised of a number of digital files including the HTML code file and image (and other) files needed to complete its content. A style sheet governing type style, font, color and other visual and structural aspects is used for all web pages comprising the website.⁷ Web pages are constructed according to a standard corporate web page template that contributes considerably to the identity of the record through logos, navigational links and copyright statements. Another corporate standard requires that all web content, including exhibits, must meet accessibility guidelines specified by the World Wide Web Consortium.⁸

⁶ The jpeg [Joint Photographic Experts Group] format is a compression format for images and is an ISO standard *Digital Compression and Coding of Continuous-tone Still Images, Part 1: Requirements and Guidelines*. (ISO/IEC IS 10918-1). The jpeg format appears to have a baseline specification, from which additional extensions are added, which may or may not be supported by supporting applications. Specification information is available at "JPEG File Interchange Format" at <http://netghost.narod.ru/gff/graphics/summary/jfif.htm>. (Checked 9 March 2005.)

⁷ Definition of cascading style sheets and specifications are available from the World Wide Web Consortium, at <http://www.w3.org/Style/CSS/>. (Checked 9 March 2005.)

⁸ World Wide Web Consortium, "Web Content Accessibility Guidelines 1.0", available at <http://www.w3.org/TR/WCAG10/> (Checked 9 March 2005.)

These requirements include accurate text representations of images so that browser software can “read” the text describing an image to blind visitors to the exhibit, for example. This process not only requires the correct assembly of all the exhibit components but also checking to ensure that they will display correctly on the most commonly used software and hardware platforms.⁹

The exhibit is finalized on a development server that emulates the actual environment encountered by an exhibit visitor. While residing on the development server exhibits are checked using the designated browsers to ensure proper display. Once the exhibit receives final approval the website coordinator forwards it to the corporate service provider where it is uploaded to the production server. It is from this server that visitors, including staff of the creating institutions, actually access the exhibits. For reasons of security and efficiency, the production server is maintained centrally for all Government offices and is therefore not in the control of the institutions developing the exhibits.

Each of the three processes¹⁰ contributes to the accuracy, authenticity, and reliability of the web exhibits. In terms of accuracy, scholarly research practices contribute directly to the “truthfulness” of exhibit content through the critical use of sources and source citations to develop the exhibit narrative, i.e., the critical use of sources provides independently verifiable facts on which the narrative is based. To gain managerial confirmation that accessibility requirements have been met, textual representation of non-textual content must be accurate. Skilled staff, high quality tools, and an established procedure for digitizing images contribute to the accuracy of the image components of web exhibits. Testing how the web exhibits display on several common platforms helps ensure that the exhibit will be accurately rendered for visitors.

⁹ Corporate web log statistics are used by the two institutions in the study to determine the most commonly used software and hardware.

The curatorial process enhances authenticity through the selection of authentic sources to include in the exhibit. This means sources whose identity and integrity can be shown to be intact. Executive level approval of the finished exhibit provides a clear indication that the institution accepts the record as being what it purports to be.¹¹ The procedures and tools for scanning images transfers the authenticity of the selected source records to the digitized copies. Usage of the corporate web template on each page is significant because the template elements identify the exhibit as belonging to the creating institution.

Control over the entire process of exhibit creation is evident from the initial and final managerial approvals. The final approval is also an indicator that the record is complete. On-going managerial review during the exhibit process serves as a check that scholarly research and other required practices are being followed. The established procedure for scanning results in the production of digitized exhibit components that are complete and reliable. Finally, there is the testing of the assembly of digital components into web pages, and pages into exhibits by the website coordinator within a development or test environment. These elements of the three processes contribute to the reliability of the web exhibits as records.

The study also revealed noteworthy gaps in accuracy, authenticity and reliability both in the rendering of the record, and in relation to institutional record-keeping and security. Web exhibits are not integrated into any existing record keeping systems. While the website itself could be developed as a record keeping system, it currently lacks key characteristics that would make this possible. Notably absent are a file classification plan or other tool that establishes the relationship between web exhibits and the other records of the institution, and defined record

¹⁰ The three processes outlined here are consistent with those defined by Martin R. Kalfatovic, *Creating a Winning Online Exhibition. A Guide for Libraries, Archives, and Museums* (Chicago and London: American Library Association, 2002), p. 20.

retention requirements that set out how long the exhibits must be maintained. If the web exhibits were to be removed from the context provided by the institutional web site, their trustworthiness as institutional records would suffer. The degree of loss of trustworthiness might be mitigated at least partially by compensatory mechanisms such as a preservation plan for the records.

Web standards and security are outside the ambit of the subject institutions in the study. Web standards, such as the requirement to use the web page template, and security, including the maintenance of the live or production web servers, are both centralized within the corporation. There was no evidence to suggest that security procedures and policies had been developed in relation to specific institutional activities or record requirements. The institutions under study followed no special procedures to ensure that web exhibits were not tampered with or modified after being up-loaded to the production server.

Since exhibits are developed specifically for users or clients external to the institutions, it is worthwhile to note deficiencies in terms of rendering the exhibits. Corporate standards require only that web content display properly on the most common browsers with minimal display requirements: that images display with a 256 color palette, for example.¹² Emerging technologies such as cell phones can access web content with non-standard browsers and have very limited display capabilities. A cell phone tested within the study failed to render several components of web exhibits. Generally speaking, it successfully displayed text and the required navigation links. However, neither the graphic components nor their alternative text description were rendered. As a result, all illustrations, including the institution's name and other identity

¹¹ This reflects authenticity as it is defined within the ISO 15489 standard: *Information and documentation – Records management – Part 1: General*, (2001), p. 7, section 7.2.2 “Authenticity.”

¹² Government of Ontario. *GO-ITS 23.1 – Internet Public Access – Product Design*. Sections 1.7.8 and 1.7.9 (2002), available at restricted intranet site http://www.gov.on.ca/MBS/techstan/GOITS_23_1_Internet_Public_Access_Product_Design.htm. (Checked 14 March 2005.)

components, were not displayed because this information was all formatted as graphics. When conventional monitors and browsers were used to access the exhibits off-line, one browser failed to locate or apply the style sheet. The result was that some components, such as the navigation links, appeared but did not display correctly. For the viewer, improper rendering of exhibits will negatively affect the accuracy, authenticity and reliability of records. As technology continues to evolve rapidly this may negatively affect the accuracy, authenticity and reliability of the exhibits *even for the creator* unless information such as an optimum viewing platform is identified or some other means of illustrating what should appear is provided.

It is important to analyze what makes web-based records accurate, authentic, and reliable especially as web-archiving initiatives are already underway. Brewster Kahle's founding vision for the Internet Archive is to provide universal access to all human knowledge. Accessibility is, indeed, a key component in the mandates of most "memory" institutions like archives and libraries. And while it is worth noting that nowhere in the Internet Archive's *Terms of Use* do the terms "preserve" or "preservation" appear, preservation is, nevertheless, implicit. For example, the proviso that "according to standard academic practice, if you use the Archive's Collections for any research that results in an article, a book, or other publication, you list the Archive as a resource in your bibliography" implies that the cited source will be preserved so that others may check it for themselves. The *Terms of Use* go on to say that "the Archive makes no warranty or representation regarding the accuracy, currency, completeness, reliability, or usefulness of the content in the Collections" nor "that access to the Collections will be uninterrupted, timely, secure, or error free, or that defects, if any, will be corrected."¹³ These

¹³ Internet Archive, *Terms of Use* (10 March 2001), available at <http://www.archive.org/about/terms.php>, accessed 20 October 2004.

statements absolve the preserver of any responsibility to creators and users alike for maintaining and providing access to trustworthy records.

My purpose here is not to belittle or defame Mr. Kahle's laudable initiative but rather to contrast the idea of records as units of information created for a specific purpose with those which are simply identified by a URI and accessed using networked computers. The Berners-Lee definition of the web is very much reflected in the Internet Archive, which is designed first and foremost to make accessible whatever it ingests. Information that is not restricted is harvested and maintained only in the context of date and URI. The Internet Archive's acquisition process does not include collecting information on website creators identities or intentions and so if this information has not been made explicit in the web content they will not be evident in the Internet Archive.

Contrast the Internet Archive with author- or creator-focussed preservation initiatives that have existed for some time, especially in the library field where national libraries have sought to extend depository programs to include digital, including web, publications. Library and Archives Canada asks depositors to send them an email if they have something to deposit and promises to maintain the "integrity and authenticity of their online publications for future generations."¹⁴ A more ambitious project is the PANDORA initiative of the National Library of Australia which takes responsibility for capturing, archiving and providing long-term access to selected online electronic Australian publications.¹⁵ The purpose of this brief comparison of two approaches to archiving web-based records is to highlight how the fundamental differences in the two general approaches affect how what is gathered is maintained for preservation.

¹⁴ Library and Archives Canada website, *Electronic Collection*, "Depositing to the Electronic Collection", available at <http://www.collectionscanada.ca/electroniccollection/003008-200-e.html>. (Checked 14 March 2005.)

¹⁵ See "Pandora Project" at the National Library of Australia website, <http://www.nla.gov.au/policy/plan/pandora.html>, accessed 10 March 2005.

The web is, by definition, a means of transmitting of information across space.

Committing the web “to memory” is a commitment to transmitting information through time as well. For web content to move from being simple units of information to records that can support accountable decision-making or contribute to a nation’s cultural heritage it is critical that such information be identified in terms of its context of creation. The role of creators in the preservation of meaningful web content is as important as that of preservers. However, the case study suggests that while considerable effort is made in the creation of web-based records, much of the value achieved thereby is attenuated by the absence of record keeping and preservation processes to maintain the records.

Preservation activities must include providing a description the creator’s purpose and intent both in the process of creating web-based records. The preserver must also document how records are maintained and reconstituted to counteract erosion of the accuracy, authenticity and reliability of those records over time.

I began this paper with a play on the words “committing to memory.” A technology based preservation initiative, such as the Internet Archive, would treat Hamlet’s soliloquy as an “information unit” beginning with “To be or not to be” and ending with “Be all my sins remember’d.” Its identification would simply be a URI, e.g., <http://www.wowzone.com/hamlet.htm>. This is sufficient for the purpose of moving it across space using networked computers but is insufficient for accountable decision-making or contributing to a nation’s cultural heritage. An author or creator-oriented preservation of the same text would preserve its relationship to a character named Hamlet in an eponymous play by a sixteenth century English playwright named William Shakespeare. This approach supports decision-making, by establishing that the soliloquy is part of a larger work, for example. This

approach also makes clear to whose cultural heritage the speech contributes and is key therefore in moving information through time.



PROPUESTA DE DISEÑO DEL SERVICIO DE INFORMACIÓN EN LÍNEA PARA LA DIVISIÓN DE ARCHIVO CENTRAL DEL INSTITUTO AUTÓNOMO BIBLIOTECA NACIONAL Y DE SERVICIOS DE BIBLIOTECA

Autor: Isneida Thais Riverol Burgos
neidita@cantv.net.ve

Caracas, marzo 2005

PONENCIA

El objeto de este trabajo es aportar a los profesionales de la Archivología, encargados de proyectar servicios de información en línea, las pautas necesarias y elementos que deben considerarse para el diseño de sistemas de información; con el fin de aportar soluciones al problema de carencia de sistemas de información gerencial, y mejorar los procesos, el entorno y recursos de la organización, que permita maximizar su rendimiento en el ejercicio y estimule el uso de la información como defensa estratégica. En este sentido el trabajo esta dirigido a la elaboración de una propuesta específica, para la División de Archivo Central del Instituto Autónomo Biblioteca Nacional y de Servicios de Biblioteca, organización pública destinada a dar servicios de información.

Como producto final se obtuvo la presente propuesta la cual está conformada en base a los siguientes diez fundamentos conceptuales: objetivo, estructura, funciones, reglas, ambiente, insumo, proceso, producto, retroinformación y resultado planteados por Milán Juranovic, (Juranovic, M., 1988), que son elementos presentes en todo sistema viviente:

1 Objetivo

El servicio de información en línea tendrá como objetivo proporcionar información para la toma de decisiones y solución de problemas, que permita dirigir y controlar todas las operaciones de la organización.

2 Estructura

El sistema de información gerencial, debe estar estructurado en una red como médula de información asociado a las estaciones de trabajo de la organización. Deben estar interconectados por medio de canales donde fluyan datos e información. Cada estación de trabajo captura y procesa los datos originados por las transacciones o eventos de la red.

Asimismo, que los datos provenientes de la red, el centro de información de una estación de trabajo recibe datos de la Dirección o de otra unidad administrativa, igualmente, envía información hacia otros departamentos o estaciones de trabajo o al medio ambiente de la organización.

El sistema de información gerencial integra los subsistemas de información determinados por las funciones organizacionales, que radica en las características esenciales a cada una de tales funciones; poseen operaciones, datos e información diferentes que unidos imposibilitan la concepción de un solo sistema de información de carácter general. Debido a esto, la estructura de un sistema de información está definida por las funciones de la organización.

El logro de los objetivos del sistema de información gerencial depende de los recursos humanos, tecnológicos, físico y material que disponga. Por lo tanto, para que la aplicación de estas tecnologías se convierta en aumentos de la productividad y consecuente mejora de la competitividad, son precisas mejoras en la formación del personal, así como en la gestión de la organización.

3 Funciones

Las funciones de los recursos humanos, deben estar dirigidas inicialmente a la capacitación del personal de la División de Archivo Central en el uso de la herramienta sistemas de computación de archivos vigentes SCAV, el personal de la organización, usuario del sistema debe capacitarse en la consulta del sistema y elaboración de documentos controlados.

El profesional del área de la información debe tener la función de administrar el sistema de información, los asistentes de archivo, deben ejecutar funciones de transcritores de datos y el grupo, conformado por el ingeniero en sistemas, programadores y analistas se encargaran de dar soporte técnico.

Las funciones del sistema de información propuesta, son funciones comunes, que a continuación se exponen según Senn:

Procesamiento de Transacciones. El sistema debe capturar, ordenar, clasificar, calcular, sintetizar y almacenar los datos producidos por las transacciones ordinarias de la organización.

Definición de archivos. El sistema debe acumular los datos capturados, por el procesamiento de transacciones, en base a (1) una estructura de almacenamiento apropiada, que

en este caso deben ser base de datos o archivos; (2) una técnica que facilite su almacenamiento, acceso y actualización; y (3) un emisor-receptor de datos apropiado de almacenamiento (discos, cintas, diskettes, otros.).

Mantenimiento de Archivos. Los archivos o bases de datos del sistema de información, deben mantenerse actualizados. Las operaciones básicas de mantenimiento son la inclusión, la transformación y la exclusión de datos en los medios de almacenamiento.

Generación de Reportes. El sistema debe realizar esta función fundamental, cuyo propósito es originar la información demandada y transferirla a las estaciones de trabajo que la soliciten.

Procesamiento de Consultas. El sistema debe permitir las consultas interactivas de los usuarios, las cuales son interrogantes no predefinidas y cuyas respuestas son cortas, por lo que no requieren reporte. Esta función es establecida por los subsistemas de administración de datos, que suministra el acceso a los datos, y de procesamiento de información, que convierte los datos almacenados en información.

Mantenimiento de la Integridad de los Datos. El sistema de información debe mantener datos confiables y veraces. Esta función debe avalar la integridad de los datos y resguardarlos contra accesos no autorizados y reformas mal intencionado.

4 Reglas

Se determina la filosofía de gestión del servicio de información en línea, se exponen su naturaleza, propósito, valores, misión, visión objetivos, estrategias, acciones a tomar, un análisis FODA y luego, el conjunto de normas generales que se establecerán, para que los componentes activos del mismo dirijan su actividad hacia el objetivo propuesto.

Filosofía de gestión

Naturaleza: La naturaleza es de servicio.

Propósito: Proponer un servicio de información eficaz y efectivo.

Misión

El servicio de información en línea, está adscrita a la División de Archivo Central, tiene por misión mantener registro del fondo documental, archivo automatizado de la documentación que

reposa en la División de Archivo y circula en la Institución, en forma organizada y pertinente, posibilitando una respuesta expedita a las consultas realizadas. Proporcionar la información en el menor tiempo posible, asegurando la satisfacción del cliente / usuario.

Visión

Ser vista como servicio con elevada capacidad de respuesta al usuario, satisfaciendo sus necesidades de información y documentación a través de las tecnologías de información y comunicación.

Objetivos

1. Satisfacer las necesidades de información de los usuarios.
2. Preservar y conservar el fondo documental.
3. Sistematizar y simplificar las funciones de manejo.
4. Automatizar la documentación administrativa jurídica, producida y recibida en la organización.
5. Proporcionar un servicio de información eficaz y efectivo.

5 Ambiente

La sociedad venezolana demanda del Instituto Autónomo Biblioteca Nacional y de Servicios de Biblioteca, servicio de información en todas las áreas del saber. Dentro de este contexto social, se encuentran Sociedades Civiles, Compañías Anónimas, entes públicos y comunidad en general.

El Instituto Autónomo Biblioteca Nacional, es una institución pública, orientada a satisfacer necesidades de información en diversas áreas.

Para desempeñar esta actividad asertivamente, el sistema de información gerencial será subsistema del sistema mayor, es decir, de la red del Instituto, porque este sistema funcionará desde la red y estará instalado a través del módulo consulta en todas las estaciones de trabajo, permitiendo una interacción entre el sistema y el ambiente.

Esta interrelación permitirá que el sistema se vaya adaptando al ambiente, en respuesta a las necesidades de información de los usuarios y clientes, además de las imposiciones de la sociedad de que forma parte.

Los usuarios internos del sistema de información estará conformado por el personal que labora en la División de Archivo Central, que se encargarán de administrar y alimentar el sistema. Los usuarios externos estarán conformados por los empleados de la organización, autorizado para consultar el sistema, clientes, visitantes, investigados y estudiantes. Los productos de los sistemas serán aprovechados por los empleados del Instituto, que deseen consultar cualquier información concerniente a la actividad de la organización.

El ambiente en que debe ser desarrollado el sistema, dependerá en gran medida de la interacción con el ambiente, que obligará a estar enfocado hacia la confianza en la gerencia y en la generación de información confiable a través del servicio de información, para la toma de decisiones.

6 Insumo

El sistema de información se alimentará del ambiente externo e interno que le proporcionará el insumo (input) necesario, para procesar el producto (output). Esta actividad de información proporcionará al sistema la adaptación necesaria para sobrevivir.

El input requerido es información de cada documento generado y recibido en la organización. Cada dato necesario será extraído de acuerdo a los campos creados en cada base de datos, que estará diseñada de acuerdo a la necesidad de los usuarios.

De acuerdo a la tipología documental a continuación se exponen las bases de datos que debe contener el sistema y los campos requeridos por los usuarios en cada tipo documental:

Asambleas: estructurada de manera que el usuario obtenga información detallada de las asambleas ordinarias y extraordinarias de las empresas de los clientes y organización. Está diseñada con los siguientes campos: número de referencia, nombre de la empresa, fecha, quórum 1ra (Asamblea ordinaria y extraordinaria), quórum 2da (Asamblea ordinaria y extraordinaria), quórum 3ra (Asamblea ordinaria y extraordinaria), fecha de celebración de asamblea, lapso para convocar asamblea, ordinaria (hábiles o continuos) palabras claves.

Constancia de calificación SIEX: estructurada de manera que el usuario obtenga una información detallada de la constancia de calificación SIEX. Está diseñada con los siguientes campos: número de referencia, fecha, fecha de solicitud, fecha de calificación, número de

calificación, tipo de calificación, porcentaje nacional, porcentaje extranjero, porcentaje neutro, palabras claves.

Contratos: estructurada de manera que el usuario obtenga información detallada de los convenios de tecnología de los clientes. Está diseñada con los siguientes campos: Número de referencia, fecha, nombre de la empresa, clase de contrato, fecha de solicitud y número, fecha de autorización y número, fecha de registro y número, fecha de entrada en vigencia, fecha de vencimiento, regalía autorizada, monto, moneda de pago, forma de pago, palabras claves.

Créditos / préstamos: estructurada de manera que el usuario obtenga una información detallada del crédito o préstamo. Está diseñada con los siguientes campos: Número referencia, Empresa, Asunto, Parcela, Fecha documento, Fecha registro, Tradición documental, Fecha notaria, Código, Palabras claves, Acreditado por.¹

Documentos: dicha base de datos está montada en el sistema, diseñada para ingresar documentos recibidos y generados. No se modificó ningún campo. Los campos que contiene son los siguientes: número de referencia, fecha del documento, título, receptor, emisor, organismo receptor, organismo emisor, palabras claves.

Ficha de personal: estructurada de manera que el usuario (Administración) obtenga una información general del personal de la empresa. Está diseñada con los campos: Nombre del trabajador, Cédula de identidad, Fecha de nacimiento, Fecha de ingreso, Salario inicial, Cargo, Motivo de ingreso, Cursos realizados dentro de la empresa, Salario final, Motivo de egreso y palabras claves.²

Funcionarios: estructurada de manera que el usuario obtenga información detallada de los funcionarios que representan a cada empresa de los clientes. Está diseñada con los siguientes campos: número de referencia, fecha, nombre de la empresa, presidente, vicepresidente, gerente general, secretario, tesorero, representante judicial, comisario principal, comisario suplente, otros funcionarios, cargo, nombre, palabras claves, número de expediente.

Legales: estructurada de manera que el usuario obtenga una información detallada del contenido de documentos legales, su registro y notaria. Está diseñada con los campos siguientes: Número de referencia, Empresa, Asunto, Inmueble, Fecha documento, Tradición documental,

¹ RIVEROL Isneida. **Informe SCAV**, 1999. (Comunicación personal). - - h. 2

² Íbidem, h. 3

Datos de registro: fecha, número documentos, folios, Protocolo, Tomo, Datos de notaria: fecha, número de documento, folio, protocolo, número de tomo. Palabras claves, Código.³

Minuta de Comité: estructurada de manera que el usuario (Accionistas) obtenga una información detallada de los asuntos tratados en cada comité de trabajo. Está diseñada con los siguientes campos: Comité n°, Presentes, Puntos a tratar, Acuerdos principales, palabras claves.⁴

Permisología: estructurada de manera que el usuario obtenga una información detallada de los permisos necesarios de su cliente. Está diseñada con los siguientes campos: Número referencia, Fecha documento, Proyecto, Organismo, Dirección, Estado, Municipio, Parcela, Tipo permiso, Número permiso, Tradición documental, Código, Observaciones, Palabras claves.⁵

7 Proceso

Después de establecida la filosofía de gestión, a continuación se establecen los procesos y procedimientos que guíe al personal en el uso del sistema y en el servicio de información a prestar:

- Archivo de documento generado
- Consulta información a través del sistema
- Creación de usuarios
- Desincorporación de archivo
- Diseminación selectiva de información
- Emisión de documento electrónico generado
- Informe mensual de actividades del sistema
- Ingreso al sistema para administrarlo
- Ingreso al sistema para consulta
- Ingreso de documentos al sistema
- Ingreso de préstamo de expediente o documento al sistema
- Reporte diario
- Respaldo de bases de datos
- Revisión base de datos

8 Producto

³ Íbidem, h. 2

⁴ Íbidem, h. 3

⁵ Íbidem, h. 2

El servicio de información en línea, ofrecerá diversos servicios y productos de información que serán fundamentales al usuario, para satisfacer y dar respuestas a sus requerimientos. Los servicios son bienes de consumo o producción cometidos por la organización, destinados a satisfacer las necesidades del usuario u organización. Los servicios son intangibles y regularmente consumidos en el momento en que son producidos. Los productos son bienes de producción que comprenden servicios y bienes tangibles que son comprados o usados por los usuarios. A continuación se relacionan los servicios y productos que ofrecerá el servicio de información:

Servicio de alerta: A través del correo interno se informará a los usuarios sobre los documentos generados y recibidos correspondientes a su área: documentos relacionado con su cliente, comunicación personal, información de interés, entre otros.

Servicio de consulta a SCAV: Es un sistema con bases de datos que contendrá los registros de los distintos documentos existentes en la organización y su objetivo será ofrecer al usuario, la posibilidad de conocer y tener acceso al fondo documental. Este servicio estará disponible al usuario en cada estación de trabajo, que le facilitará la consulta documental, imprimir documentos, enviar documentos y editar documentos.

Servicio de difusión de información: El servicio estará destinado a difundir a través del correo electrónico, información interna generada en la organización sobre intereses comunes a todos los trabajadores.

Servicio de encuadernación: Este servicio colaborará con la conservación documental, en virtud de que existen documentos voluminosos y la encuadernación evitará su deterioro y facilitará al usuario su traslado.

Servicio de fotocopiado: Este servicio permitirá suministrar la información requerida por los usuarios a través de copias digitalizadas, fotocopias, soportes magnéticos (Diskette, CD), correo electrónico o fax. El servicio tiene su base en la necesidad del usuario de obtener reproducciones de un material original.

Servicio de Internet: El servicio de información ofrecerá acceso a fuentes de información, permitiendo a los usuarios buscar y compartir información en forma rápida. El acceso a Internet

ofrecerá la posibilidad de realizar búsquedas en diferentes direcciones y bases de datos, además de enviar y recibir documentos a clientes y asociados.

Servicio de investigación: Este servicio consistirá en la preparación y entrega de búsquedas de documentos, en las que se recopila información sobre temas específicos.

Servicio de préstamo: Servicio destinado a facilitar el expediente o documento original al usuario. A fin de satisfacer las demandas de información.

Servicio de referencia: Este servicio estará destinado a dar respuesta a las diversas preguntas que hacen los usuarios que consultan el sistema. Se le recomendará forma de búsqueda y tipo de documento, según su necesidad, dónde localizarlo y como hacer uso de los servicios que se prestan. El servicio puede ser en persona, vía telefónica, correo electrónico o mensajería interna.

Servicio de reprografía: Este servicio optimizará los recursos de la organización, permitirá suministrar formatos y cantidad requerida por los usuarios, tacos de papel para tomar notas y tarjetas de presentación del personal. El servicio tiene su base en la necesidad de contar con formatos estandarizados dentro de la organización.

El sistema de información proporcionará los siguientes reportes: reporte de tesoro de organizaciones, conceptos y personas. Reporte de préstamos y listado de documentos desincorporados. Igualmente ofrece como producto documentos de reportes de envíos de correspondencia, copias de documentos digitalizados y cualquier otra información requerida.

9 Retroinformación

El Jefe de la División de Archivo Central emitirá un informe mensual a la Dirección Ejecutiva, que contendrá estadísticas de las actividades rutinarias de la División, los logros alcanzados, servicios prestados, documentos ingresados y consultas al sistema en el mes anterior a la fecha del informe, control de préstamos y cualquier otra actividad relevante.

Entre los empleados y el jefe de la División de Archivo Central se mantendrá información constante, sobre las irregularidades y ventajas del sistema. El Jefe de División mantendrá comunicación con la Oficina Tecnología de Información, sobre las necesidades de mantenimiento

de los sistemas. Habrá continua comunicación con el usuario que irá dando las pautas del servicio a prestar de acuerdo a sus necesidades.

10 Resultado

Para lograr los objetivos del servicio de información en línea, el Jefe de División tomará indicadores de evaluación de resultados, porcentajes de metas alcanzadas por el servicio y la División de Archivo Central, que le permitirán evaluar la eficacia de gestión y operación de la División. Dentro de las metas incluirá eficiencia, organización y rapidez en el ingreso, almacenamiento, búsqueda y recuperación de información.

Estos indicadores permitirán tomar las acciones precisas, para mantener los sistemas y el servicio de información en línea, orientados hacia el logro de los objetivos planteados.

CONCLUSIONES

Una vez finalizado este trabajo podemos concluir en principio, que la implantación y ejercicio de un sistema de información automatizado dentro de una organización, origina un gran impacto positivo o negativo y de acuerdo a ello, su aceptación o rechazo por parte de los usuarios y de la organización.

Un sistema de información es un sistema, porque posee las características siguientes: abierto, porque debe interactuar con su ambiente mediante el intercambio de información y adaptarse a sus necesidades.

Sistema hombre máquina, para admitir el uso de ordenadores que sirvan de herramientas, para automatizar los procesos habituales de transformación, los cuales son conducidos e inspeccionados por el hombre.

Su entrada ha de estar constituida por datos y su salida por información, el proceso de transformación de datos en información, constituye la función esencial del sistema de información.

La elaboración de un proyecto de servicio de información en línea, debe ser elaborada tomando en cuenta los elementos conceptuales de todo sistema viviente, para asegurar su permanencia en el tiempo y su éxito dentro de la organización.

El sistema de información gerencial debe facilitar la retroalimentación (feedback) y la promoción de una acción proactiva, avanzando hacia la posible evolución de la conducta o comportamiento del proyecto o iniciativa.

RECOMENDACIONES

La propuesta de diseño del servicio de información en línea, proporciona alternativas factibles según los recursos que puedan presentarse y que acondicionen al Instituto en el momento de su puesta en marcha. Por tal sentido se plantean las siguientes recomendaciones:

Se recomienda la implantación del servicio de información en línea, a fin de optimizar el uso de la información como herramienta vital para el logro de sus objetivos, satisfacer sus insuficiencias y la de sus usuarios.

Se recomienda la elaboración de un proyecto de servicio de información en línea, con las especificaciones presentadas en esta propuesta, con la finalidad de que la organización cuente con un sistema de información gerencial, que le permita una toma de decisiones exitosa.

La implantación debe hacerse tomando en cuenta todos los recursos necesarios, recursos humanos, recursos tecnológicos, recursos físicos y materiales, a fin de garantizar un servicio eficaz.

BIBLIOGRAFÍA

Administración y uso de la información [en línea]. <http://erc.msh.org/fpmh_spanish/chp7/pl.html>
[Consulta: 22 de octubre 2002]

ANDER-EGG, Ezequiel. **Introducción a la planificación** / Ezequiel Ander-Egg. - - Buenos Aires : El Cid Editores, 1976. - - 250 p.

Añez, Luis. Currículo Vitae [en línea].
<http://seleccion.net/online/finanzas/financemanagers/Luis_Enrique_Anez_Castillo.htm> [Consulta: 22 de marzo 2004]

Archivística : estudios básicos. - - Sevilla : Diputación Provincial de Sevilla, 1981. - - 256 p.

BALESTRINI ACUÑA, Miriam. **Como se elabora el proyecto de investigación** / Miriam Balestrini Acuña. - - 6ª. ed. - - Caracas : BL Consultores Asociados. Servicio Editorial, 2002. - - 248 p.

BLANCO, Luis. **Manual para la elaboración, presentación y evaluación del trabajo de licenciatura** / Luis Blanco, Josefina Cáceres, Gavidia Wiston, Carlos Ortuño. - - Caracas : Universidad Central de Venezuela, Facultad de Humanidades y Educación, Escuela de Bibliotecología y Archivología, 1993. - - 46 p.

Capítulo 4. Teoría sobre sistemas de información [en línea].

<http://mailweb.udlap.mx/tesis/lis/teodoro_v_jc/capitulo4.html> [Consulta: 22 de octubre 2002]

Capítulo 1. El lenguaje visual FOXPRO [en línea].
<<http://www.politec.une.edu.py/taller3/docs/guia1/ufp.htm-6k>> [Consulta: 11 de julio 2003]

CELIS, Alejandro. **Requerimientos mínimos de hardware que se recomienda para el funcionamiento del SCAV**, 2002. (Comunicación personal). - - 2 h.

CHAVEZ, **Ernesto**. **Filosofía de Gestión de la Dirección de Recursos Humanos**, 2003. (Comunicación personal). - - h. 1-4

Cibanal, Luis. Teoría general de sistema [en línea].

<http://perso.wanadoo.es/aniorte_nic/apunt_terap_famil_2.htm> [Consulta: 15 de julio 2003]

COLINA U., Milagro D. **Diagnóstico de las necesidades de información de los ingenieros mecánicos que ejercen sus actividades en empresas privadas del área metropolitana de Caracas, como un aporte a la red de ingeniería, arquitectura y afines (REDINARA)** / Milagro D. Colina U., Carmen R. Gómez P.; Luis Enrique Díaz. - - Caracas. - - 1989. - - 183 h. - - Tesis (Licenciatura en Bibliotecología). - - Universidad Central de Venezuela, Facultad de Humanidades y Educación, Escuela de Bibliotecología y Archivología.

CRUZ MUNDET, José R. **Manual de archivística** / José Ramón Cruz Mundet. - - Madrid : Fundación Germán Sánchez Ruipérez, 1994. - - 408 p.

Determinar como el uso de manuales de procedimientos podrán contribuir a [en línea].

<...Determinar%20como%20el%20uso%20de%20manuales%20de%20procedimientos%20podra>

[Consulta: 25 de junio 2002]

Diccionario de la lengua española / José Alemany y Bolufer, Ramón Sopena (editor). - - Barcelona (España) : Limpida Fons, (1910). - - 1752 p.

Estivill, Assumpció. Cómo citar recursos electrónicos [en línea]. Cristóbal Urbano. Mayo 1997. Ver. 1.0. <<http://www.geocntes.com/Atnens/itnaca/0/2/recuerec.num>> [Consulta: 20 de septiembre 1999]

FERRER DE HERNÁNDEZ, Adriana. **Diccionario básico del proceso investigativo** / Adriana Ferrer de Hernández. - - Caracas : Ediciones CIAR, 1993. - - 144 p.

GARCÍA, Ramón. **Pequeño Larousse ilustrado** / Ramón García. - - Buenos Aires : Larousse Argentina, 1984. - - 1.663 p.

HERNÁNDEZ S., Roberto. **Metodología de la investigación** / Roberto Hernández S. ; Carlos Fernández C. ; Pilar Baptista L. - - México : McGraw-Hill, 1991. - - 505 p.

HOCHMAN, Elena. **Técnicas de investigación documental** / Elena Hochman, Maritza Montero. - 6ª. ed. - - México : Trillas, 1978. - - 88 p.

JOHNSON, Richard A. **Teoría, integración y administración de sistemas** / Richard A. Johnson. - - México : Editorial Limusa - Wiley, 1966. - - 399 p.

JURANOVIC, Milán R.: **Fundamentos de sistemas y sistemas de información gerencial** / Milán R. Juranovic. - - Caracas : Contemporánea de Ediciones, 1998, 194 p.

MONTEVERDE, I: **Manual de inducción**, Caracas : UICA, 1996.

MONTILVA C., Jonás A.: **Desarrollo de Sistemas de Información** / Jonás Montilva. - - 2ª. ed. - - Mérida : Universidad de los Andes-Consejo de Publicaciones, 1995, 262 p.

MORLES, Víctor. **Planeamiento y análisis de investigaciones** / Víctor Morles. - - Caracas : Ediciones El Dorado, 1992. - - 89 p.

Olazábal, Angel. Teoría General de Sistemas. (Pensamiento de sistemas), [en línea]. <<http://www.udabol.ed.bo/biblioteca/sistemas/8tgs/14gpensamsis/tgs.htm>> [Consulta: 15 de julio 2003]

Paños Antonio. Reflexiones sobre el papel de la información como recurso competitivo en la empresa, Anales de Documentación [en línea]. No. 2 (1999) <<http://www.um.es/fccd/anales/ad02/AD02-1999.PDF>> [Consulta: 15 de julio 2003]

Propuesta de un modelo de sistema de información gerencial para mejorar el proceso de selección de personal en la Dirección de Relaciones de Trabajo de la Universidad de Carabobo [en línea]. Octubre 2002. <<http://www.gestiopolis.com/recursos/documentos/fuldocs/rrhh/sigselecc.html>> [Consulta: 22 de octubre 2002]

RAMÍREZ, Tulio. **Como hacer un proyecto de investigación** / Tulio Ramírez. - - Caracas : Carhel, 1992. - - 83 p.

RIVEROL, Isneida. **Diagnóstico. Departamento de Archivo. Benson, Pérez Matos, Antakly & Watts**, 2003. (Comunicación personal). - - 14 h.

_____. **Informe SCAV**, 1999. (Comunicación personal). - - 4 h.

Ros, García. Auge de los sistemas de información y documentación en las organizaciones [en línea]. <<http://www.ucm.es/info/multidoc/multidoc/revista/num2/jros.html>> [Consulta: 22 de octubre 2002]

ROTUNDO, Emiro. **Introducción a la teoría general de los sistemas** / Emiro Rotundo. - - 2ª. ed. - - Caracas : Universidad Central de Venezuela, Facultad de Ciencias Económicas y Sociales, División de Publicaciones, 1973. - - 96 p.

SABINO, Carlos. **Como hacer una tesis : guía para la elaboración y redacción de trabajos científicos** / Carlos Sabino. - - Caracas : Editorial Panapo, 1993. - - 212 p.

_____. **El proceso de investigación : una introducción teórico - práctica** / Carlos Sabino. - - Caracas : Editorial Panapo, 1992. - - 210 p.

Sánchez, Bárbara. La Información como Recurso en el Desarrollo de las Organizaciones de las Administraciones Públicas [en línea]. <<http://www.um.es/fccd/anales/ad03/AD10-2000.PDF>> [Consulta: 20 de julio 2003]

SCAV para Windows. Manual del usuario, versión 2.0. - - Caracas : Archicentro Sistemas, 1997. - - 360 p.

SCHMELKES, Corina. Manual para la presentación de anteproyectos e informes de investigación (tesis) / Corina Schmelkes. - - México : Harla, 1988. - - 212 p.

SOTO, Leonardo. **Uso de FoxPro pasos a seguir**, 2003. (Comunicación personal). - - 4 h.

_____. **Proyecto administración y control de expedientes**, 2001. (Comunicación personal). - - 91 h.

TEKHAUS, Ernesto. **Componentes mínimos para una red**, 2002. (Comunicación personal). - - 1 h.

Torres, Juan. Cadena de valor y su aplicación a los sistemas de información [en línea]. Diciembre 2001.

<<http://www.timagazine.net/magazine/1201/valor.cfm?id=1201/valor&clave=ti%20OR%20informacion%20OR%20empresa%20OR%20business>> [Consulta: 22 de octubre 2002]

VAN, John. **Teoría general de sistemas** / John Van. - - México : Editorial Trillas, 1989. - - 607 p.

La Archivística en la Filosofía de la Información (o viceversa): Inserción de un paradigma teórico.

MSc. Ania R. Hernández Quintana

Departamento de Bibliotecología y Ciencia de la Información. Facultad de Comunicación. Universidad de La Habana. Cuba

Para explicar la paradoja del sentido, Deleuze (Cf. Drumond 2003) recuerda a Alicia en el país de las maravillas. Ella, como Edipo, enfrenta siempre dos caminos, por ejemplo: crecer y dejar de ser menor. Pero, a la vez, Alicia no es mayor y menor. Es mayor ahora y menor antes, pues el hecho de crecer se hace posible solamente dejando de ser menor, y a la inversa.

La paradoja del sentido se explica cuando se puede ir en dos direcciones al mismo tiempo pero, a la vez, ser imposible una identificación única, pues puede enfatizarse con igual éxito uno u otro de los efectos. Afirmaremos que la mayor paradoja del sentido para la especialidad en los tiempos que corren consiste en ser imposible la Archivística sin la Gestión Documental (y viceversa), para lo cual deben confluír en una empresa común el saber hacer de cada una para el tratamiento de la información consignada de forma electrónica o digital.

El territorio continuo que comprende la integridad y el acceso a los documentos les concierne a ambas para atender la proliferación de información desprovista de sentido, anónima y descontextualizada. La unicidad, autenticidad y preservación documental, como principios de la nueva Archivística, enfrentan los retos de la ubicuidad digital, de la manipulación dolosa y la obsolescencia tecnológica, pero son, a no dudarlo, principios esenciales para diseñar el “Web de la verdad”.

Archiveros y gerentes documentales están compulsados a enfrentar, junto con juristas y otros especialistas y técnicos, procedimientos que arruinan la nueva

ética informacional, fecunda en prácticas ingenuas y malévolas. Albuquerque (2005) nombra este fenómeno como la “zona ceniza”, aquella donde impera el oportunismo en la creación documental, porque no existen límites claros en el ciberespacio entre el bien y el mal, lo cierto y lo errado, lo legal y lo ilegal (otra vez la paradoja del sentido), bajo la apariencia de la legalidad o la lealtad personal e institucional. Los involucrados en la zona ceniza pueden cambiar la historia o reinventar la realidad, pueden desconocer la tradición y falsear los recursos.

Gutiérrez (2000) incluso es más apocalíptico, cuando afirma que “gran parte de esa información es errónea o tendenciosa, orientada a defender una postura o intereses”. Cazau (2002) se preocupa sobre quién controla la calidad de la información en red, “problema que adquiere dimensiones importantes particularmente cuando se trata de recuperar la información dentro de una masa de datos frecuentemente sepultada por una inmensísima cantidad de documentos irrelevantes, sin ningún valor científico o instrumental”.

Si abordamos el problema paradójicamente, “solo habrá que esperar que se resuelva una ecuación”, apunta Dollar (1994), recomendando a los archiveros que nos concentremos en cuestiones intelectuales, pues las tecnológicas serán inevitablemente resueltas en su momento.

La paradoja del sentido está en la esencia de los hipertextos y del ciberespacio, por lo que organizar documentos, información y conocimientos en esta nueva zona mediática, gnoseológica, documental y cultural, entraña, primero, reconstruir las definiciones que necesitamos para entendernos y reconocer los nuevos espacios profesionales, para luego confirmar que las implicaciones sobre la identidad, la localización y la clasificación de los acervos físicamente codificados no son exactamente las mismas en un entorno donde “se descolocan definiciones, determinaciones, exclusiones e inclusiones” (Levy, 1996), donde se hacen inestables los dogmas, el sentido único y el sentido común, donde se disemina en

flujos el conocimiento y se universaliza a través de una distribución y una conexión de “todos con todos”. (Drumond, 2003)

Pero, ¿cómo se ventilan acaso los principios fundamentales de la Archivística en este espacio?

1. La profesora Duranti ya manifestaba en este mismo escenario que la unicidad de los documentos archivísticos estaba marcada por las relaciones contextuales de los creadores, pero que los contextos de los registros digitales no son siempre evidentes o explícitos. Los creadores, en virtud de sus propias funciones, son los más interesados en incluir esa función contextual en los actuales sistemas de información, pero los archiveros y los gerentes debemos comprender que la valorización como profesionales de la información de estos tiempos radica en la construcción de una nueva visibilidad, en tanto seamos capaces de enfrentar estudios arqueológicos (archimaining, si se me permite el término, en clara alusión al bibliomaining de Scott, 2005) para derivar en instrumentos de ayuda a los gerentes de sistemas documentales digitales.

2. La autenticidad se presume mientras no se demuestre lo contrario, de contexto y de contenido. Para garantizarla, los documentos archivísticos tradicionales se han mantenido bajo custodia en entidades específicas. Ahora, la arquitectura cliente-servidor homologa esta función, pero la información archivística no es solo información registrada, sino información orgánica, por lo que nuestra participación en la construcción de sistemas dinámicos de metadatos es impostergable, para poder garantizar el reconocimiento de este principio. El perfeccionamiento de normas archivísticas de descripción cada vez más relevantes, que vinculen coherentemente a formadores, productos y proveedores contribuirá a advertir los vestigios de una falsificación.

3. La preservación de la legibilidad, recuperabilidad e inteligibilidad de los documentos, se concreta en la planeación estratégica de las migraciones, donde

la valoración archivística tiene que jugar un papel crucial para que esos planes sean realistas y factibles.

Los registros convencionales (todos los documentos antes de la era de la computación) se caracterizaban por una semejanza entre la forma del registro y el material registrado, amén de una extensa variedad tipológica, que permitían tratar con una continuidad espacio-temporal. Este carácter permitió distinguir a los documentos analógicos en virtud de su forma, de sus signos portadores de sentido y como medio para evidenciar relaciones sociales. (Pédauque, 2003)

En virtud de la forma, el saber hacer profesional privilegiaba numerosos puntos de vista, algunos bien antiguos, como la caligrafía, la tipografía u otras formas de representación, y así se privilegiaban, por ejemplo, la catalogación o la clasificación y la gestión de documentos e, incluso, la diplomática archivística.

La digitalización cambió drásticamente esta situación al descomponer todo el material registrado en cálculos matemáticos discretos representados en ceros y unos. Deleuze y Guattari (1995) proponen comprender las formas simbólicas del ciberespacio como metáforas, pues no se confinan al hecho físico de la realización, a la rigidez temporal resultante del registro material, y sobre todo, porque expanden esencias semánticas.

Este movimiento de las formas tiene una estrecha relación con los contenidos y con el sentido, especialmente cuando se ha discontinuado el límite entre las estructuras lógicas y las físicas, o mejor, cuando se distribuyen ambas en el mismo momento.

La información de naturaleza archivística “todavía está escondida y es poco comprendida” (Lopes, 1998), por lo que rozar algunos aspectos filosóficos se va haciendo notable en estos tiempos. Con el advenimiento de la postmodernidad (pura paradoja del sentido) ninguna disciplina es una isla. Tampoco la Archivística y, menos, su gremio. Los problemas que enfrentamos, desde el punto de vista

práctico, están relacionados con el paso del dominio análogo al digital y con la amplitud pragmático-funcional de la entidad archivística. Desde el punto de vista teórico, la reciente gestación de un marco epistémico que examina problemas de la dinámica social de la información, obliga a compartir espacios de reflexión desde la convergencia de disciplinas sociales e informáticas en virtud de una finalidad: la recuperación oportuna, eficiente, fidedigna e inteligente.

Estas dos dimensiones, la práctica y la teórica, se afectan por igual de los paradigmas dominantes que emplean e identifican las actividades, organizaciones y colectividades, y por supuesto, irrumpen sobre cualquier consideración de ciencia.

Para la Archivística, la gestión documental digital debe suponer, como se ha visto, una revisión y un ajuste de principios críticos, en especial porque ha rebasado una función testimonial autónoma para ampliar su alcance a la gestión de inteligencia en red, lo cual expande tanto su investigación y como su propio objeto de estudio.

La Archivística y la Gestión Documental deben colaborar, en definitiva, “con la concepción de un sistema matainformacional para asegurar que estos contengan la información contextual necesaria para la comprensión del conjunto de documentos y sistemas documentales”. (Dolar, 1994) Esta es la estrategia inmediata y fundamental de los archiveros encargados de la gestión documental electrónica. Pero para que sea fructífera, se necesita del dinamismo profesional que desarrolle nuevas aptitudes tecnológicas y multidisciplinarias.

El mundo virtual, como dispositivo informacional, se define como un espacio de interacción dentro del cual quien explora puede controlar directamente a los representantes de sí mismo (Levy, 1996) a través de sus actos. Tanto en la producción de documentos (hipertextos), como en la creación de ideas, en el ciberespacio se potencian significaciones más próximas al propio movimiento multidireccional del pensamiento que a la tradicional imagen de una forma

documental. Esto quiere decir, que la práctica profesional archivística tiene y comparte un nuevo contexto social, introducido por la digitalización, para gestionar la memoria humana objetivada.

Los archiveros no estamos solos, no somos los únicos y mucho menos somos huérfanos. No ya. Por primera vez en la historia, todas las disciplinas que se ocupan por separado de atender la información documentada, enfrentamos, al unísono, “una forma unitaria de registro que resulta idónea para todas las formas de expresión humana” (Bizjak, 2000). Por primera vez en la historia, un “progenitor común” (Floridi, 2004) se nos está presentado y propone disolver la crisis de identidad que nos afecta como “gueto intelectual” (Pierce, 1992), como trincheras del aislacionismo documental-funcional: la Filosofía de la Información.

Por tanto, así como no se puede desconocer la naturaleza de los nuevos documentos, los riesgos y las ventajas que representan, lo cual ya hace tiempo es obvio, no debemos tampoco seguir tratando el análisis del fenómeno esencial (los registros de información y conocimiento) por separado, desde las remotas parcelas técnico-disciplinares de bibliotecarios, documentalistas, científicos de la información, arquitectos o archiveros, excluyendo innecesariamente, desde estos presupuestos, importantes y comunes referentes teóricos y filosóficos.

Nuestra misión en el segundo milenio y en la “segunda generación de estudios relativos a los documentos” (Wallace, 1993) es vital: contribuir a que desaparezcan las limitaciones espacio-temporales para la transferencia informacional, construir mediaciones auténticas cargadas de valor semántico e insertarnos absolutamente en una nueva dinámica cultural con implicaciones prácticas muy cercanas a todos.

Esto sobrepasa la función de garantes de un acceso permanente a recursos auténticos pues implica no solo identificar, sino comprender significaciones e interpretar contenidos. Hedstrom (1993) lo representaba de manera inmejorable:

“decidir qué es esencial e imaginar qué es posible”. Así, la gestión metainformativa se convierte en una estrategia alternativa a las prácticas descriptivas actuales, y tiene consecuencias sobre el tratamiento archivístico, pero para ello debemos reevaluar nuestros objetivos y trazar soluciones novedosas desde la informática, la gestión y los principios archivísticos.

Para congeniar sobre principios fundamentales, quedaba por construir una filosofía que incorporara elementos de la información en términos de materia de estudio, que contribuyera a disipar el minimalismo de fondo de cada una de estas disciplinas y que reforzara nuestra identidad como comunidad científica.

La crisis de credibilidad científica de la Archivística, lo mismo que de la Bibliotecología, se apoya en sendos y centenarios enfoques pragmáticos que con la Filosofía de la Información, entre otras dimensiones, puede terminar (o reformularse) en la medida en que satisfaga su rol de fundamento teórico, al proporcionar una comprensión sistémica de sus conceptos básicos, estudiando su naturaleza, valor y fines de las prácticas específicas.

Cada una por separado ha buscado apoyos teóricos en la Historia, en la Comunicación, en la Psicología, en la Matemática, en fin, en fuentes externas a las propias disciplinas. Recordemos que para el común de los mortales, la Archivística, por ejemplo, sigue siendo considerada una ciencia auxiliar, no una ciencia en sí misma. Colaborando con la construcción de la Filosofía de la Información desarrollaremos un fundamento propio sin renunciar a especificidades y de la mano de la interdisciplinariedad, para “para explicar y dirigir la construcción intencionada de nuestro entorno intelectual y ofrecer el tratamiento sistemático de los fundamentos conceptuales de la sociedad contemporánea”. (Floridi, 2004)

Con la revolución digital se alcanzó un nuevo nivel de sinergia entre los saberes y la tecnología que dominó todo el discurso profesional de los años 80 y 90 del siglo

XX. Las necesidades de la recurrente Sociedad de la Información compulsaron nuevas estrategias perceptivas y organizativas que colocaron el fenómeno “qué es información” en el centro de un debate filosófico, que no solo revisó viejas cuestiones, sino que identificó otras de gran novedad. Las tecnologías no solo se insertan en las disciplinas científicas, técnicas e ingenieriles, sino que fecundan el procesamiento, almacenamiento, distribución y aplicaciones de información.

La relación información-informática fue tan controversial y productiva en esas décadas, que Bynum y Moor (1998) la calificaron como una nueva fuerza en el escenario filosófico, con profundas implicaciones para la comprensión de muchos conceptos emblemáticos como mente, conciencia, experiencia o verdad, razón por la cual compilaron importantes trabajos que resumían una época de presentismo filosófico, de seducción tecnológica, pero con miradas diversas sobre la esencia informacional.

Es en Floridi (2003) y en su Filosofía de la Información, que se define un nuevo espacio de investigación que relaciona con mucha más justicia ambos fenómenos: “la naturaleza conceptual y los principios básicos de la información, incluyendo su dinámica, su utilización y sus ciencias, y la elaboración y aplicación de metodologías teórico-informativas e informáticas y su aplicación a problemas filosóficos”. Este es el planteamiento original, pero realmente la Filosofía de la Información tiene como tema esencial a la información y no a la computación, puesto que analiza que la última presupone la primera.

“La Filosofía de la Información no se plantea desarrollar una teoría unificada de la información, sino más bien una familia integrada de teorías que analicen, evalúen y expliquen los diversos principios y conceptos de información, su dinámica y utilización, prestando especial atención a las cuestiones sistémicas que provienen de diferentes contextos de aplicación y las interconexiones con otros conceptos clave de la Filosofía, tales como el ser, el conocimiento, la verdad, la vida y el significado”. (Floridi, 2004)

El enfoque metateórico de Floridi persigue esclarecer si un problema o una explicación pueden ser legítima y plenamente reducidos a un problema o explicación informacional, implica preguntarse cómo sería un sistema sin tener en lo absoluto una naturaleza informativa.

Como se podrá inferir de lo anterior, la inserción de la tríada gestión documental, de información y del conocimiento, en no importa cual ambiente, como estrategia para el éxito organizacional (Ponjuán, 2004), tiene aquí un importante fundamento, ciertamente diacrónico, que solo demuestra una vez más la distancia temporal entre filosofía, teoría y práctica.

Pero también, bajo esta égida, se evidencia la necesidad de construir los encuadres filosóficos que ayuden a formular una Ciencia de la Información global (ya que la recuperación es un problema que afecta a todos los sistemas y técnicas que registren, documenten o archiven información), y no una ciencia exclusiva del norte anglosajón, pragmático, rico, digitalizado y conectado.

Desde la Filosofía de la Información se ofrecen razones y presupuestos para que todas las culturas informacionales, generadas desde distintas prácticas y objetivos, confluyan en el *continuum* desde sus diferencias, pero con igual reconocimiento científico. Se trata de encontrar soluciones pensáticas múltiples para los variados sentidos con que puede comprenderse la propia noción de información, que sin un adjetivo para acompañarla, en nuestro caso más de uno (registrada, auténtica y orgánica), conduce a una explicación simplista y a resultados incompletos.

La Ciencia de la Información (y los saberes archivísticos ahí contemplados), comparte con la Filosofía de la Información los dos aspectos de su definición, no se reduce a uno u otro (recordemos la paradoja del sentido): su objeto de investigación es la información documentada, sus ciclos y los procedimientos, las

técnicas y los instrumentos por lo que estos se implementan, se manejan y se regulan.

De esta manera, puede calificarse a la Archivística como Filosofía Aplicada de la Información, porque tanto ella como la propia Ciencia de la Información pueden emplear principios fundamentales para resolver problemas prácticos definidos y tratar con fenómenos específicos concretos, realizar investigaciones empíricas con propósitos prácticos de servicio y contribuir al desarrollo de su investigación básica. (Floridi, 2002)

Shannon (1993) advertía que no se podía esperar un concepto único de información, y es así porque efectivamente, la información tiene una naturaleza fragmentada, que adquiere su sentido último en un contexto específico. En cada noción de información se encuentran fragmentos de conocimiento, cuyos contenidos transitan por las distintas dimensiones, orales, escritas, visuales, o integradas, que permiten una comprensión dependiente de la perspectiva gnoseológica, filosófica o pragmática que le sirva de partida a los sujetos.

Como mediadores culturales, los archiveros y todos los profesionales de la información, hemos transitado de la cultura oral a la materializada en la escritura y de ahí a la digital. El paradigma más notorio que envuelve a las formas culturales de transmisión de información se ha relacionado con su acceso, pero hoy día, justamente por las paradojas del sentido, el cambio paradigmático estriba en colaborar con la comprensión de los acervos.

El texto tradicional, entendido como flujo de información direccionado y sucesivo, en los nuevos sistemas de información en red se vuelve multidimensional, sin una ordenación lineal obligatoria, que compromete aún más a los significados. La Filosofía de la Información reconoce esta “nueva etapa en la semantización del ser” (Floridi, 2004) al plantear su problema fundamental y analizar la naturaleza de la información.

Para ello, Floridi propone tres dimensiones de estudio:

1. información como realidad o información ecológica: patrones de señales físicas, tal como lo ve la Teoría Matemática de la Comunicación, interesada no en la utilidad, relevancia, significado, interpretación o pertinencia de los datos, sino en el nivel de detalle, disposición, localización y frecuencia de los datos no interpretados, como los registros que se replican o redistribuyen en el entorno digital sobre un documento,
2. información para la realidad o información genérica: información de instrucción, por ejemplo, el gran complejo de asistentes para ayudas e interacciones, e
3. información acerca de la realidad o información semántica: “información interesada”, en la medida que reduce el grado de incertidumbre o sorpresa, dado un estado de conocimiento del informado. Esta información se encuentra en las estructuras profundas (Van Dijk, 1992) de los documentos y es esencialmente contextual. La información sobre la realidad puede tener enfoques probabilísticos, modales, sistémicos, inferenciales o significativos, pero todos buscan exponer el contenido semántico adoptando una orientación proposicional. (Floridi, 2004)

A la Archivística le incumben las tres dimensiones. En virtud del tiempo que generosamente han dispuesto, apenas esbozaré algunos elementos de la información semántica y de su relación con nuestra área, pues sigue siendo esta, en la práctica, un trabajo sobre datos e información registrados en soportes convencionales y/o con datos e información derivados de sistemas de software, pero aún tenemos limitaciones para comprender la importancia capital de generar instrumentos que evidencien los nexos entre estas fuentes y el conocimiento desde los documentos archivísticos.

El poder semántico de la información llevó, tiempo atrás, a reconocerla como “diferencias que importan” Batenson (1979). La relación de reciprocidad de los documentos y sus contenidos, así como la complejidad de estas relaciones, aún no tiene suficientes representaciones en los sistemas de recuperación de la información en línea. Así como no existe todavía la Web de la verdad que comentábamos antes, tampoco existe la Web semántica.

Ambas constituyen un ideal en virtud de necesidades muy concretas en el campo, téngase por caso, del gobierno electrónico o la gestión de contenidos, pero su concreción no es obligación exclusiva de los programadores, es especialmente, una función nuestra, pues es el contenido documental quien determina el proceso que enlaza creador y uso, es el que forma un mejor conocimiento de la realidad y que el permite actuar sobre ella.

Los contenidos y los significados son un problema estratégico para los sistemas de información, son el punto de giro para un tratamiento no solo técnico, sino científico de la información, y no es solo un problema informacional o comunicacional, es también un problema filosófico. El mejor caso de estudio es el fenómeno de las ontologías.

La ontología es una rama de la Filosofía que se ocupa del ser en general y de sus propiedades trascendentales. La ontología filosófica no se reduce a un lenguaje o a una forma, aunque a través del lenguaje describa las cosas en sí y las relaciones entre las cosas. Su objetivo es describir la “verdad”, en tanto conocimiento sobre determinado objeto, sujeto o suceso, y poder distinguirla de falsas creencias, bajo el principio de que solo el conocimiento verídico es útil.

Primero fue la ontología filosófica, luego se transformó interdisciplinariamente y llegó a los sistemas de información, en los años 90, con igual denominación pero con otras connotaciones y funciones. Su irrupción inició una auténtica revolución en el

área de los lenguajes documentales controlados (Quin y Paling, 2000-01) y en especial significó una confrontación directa con los tesauros.

Para los sistemas de información, las ontologías son un lenguaje formal diseñado para representar un campo determinado de conocimiento con uno o más propósitos específicos en entornos informáticos y requiere, primero, la determinación de los conceptos para, después, constituirse en un nivel de explicación y relación. La definición más generalizada es de Gruber (1993): la ontología es una especificación explícita de una conceptualización.

Existen dos niveles de análisis sobre las ontologías en los sistemas de información: como esfera de investigación general y como producto particular de descripción, es decir, como fenómeno formal, la primera, y como ontología material, sobre algo o para algo, la segunda.

En las ontologías formales, la descripción que se deriva de la investigación es general, y por tanto, aplicable a todas las esferas materiales o campos de la realidad, esto es: entidades naturales físicas (lugares, animales, plantas), entidades subjetivas (las que dependen de los sentidos, como el olor o el dolor), artefactos físicos (objetos materiales), objetos sociales (construidos en virtud de roles, como el dinero) y entidades abstractas (como los colores o los números). Estas clases no nos son desconocidas, es una descomposición similar a la que presentan los tesauros.

Para Gruber y para los sistemas de información, “lo que existe es lo que puede ser representado”, pero ¿realmente todo lo describible filosóficamente hablando, puede representarse en los sistemas de información? Zúñiga (2001) se pregunta cómo representar la pérdida de dignidad, la angustia, la felicidad, a modo de ejemplos que todavía no tienen respuesta cabal en ningún lenguaje, ni en los tesauros, ni en las ontologías.

Las ontologías formales son clasificaciones, demarcaciones, taxonomías del mundo, que además deben conceptualizarse, en los sistemas de información, para explicar la visión que se tiene de ese mundo. Deben ser explícitas, dichas, escritas, representadas, para que pueda ser expresada, reconocida y compartida por los distintos niveles de conocimiento de los grupos de actores, porque cada uno en su dominio tiene su particular manera de reconocer la realidad. Esta es la diferencia esencial que tienen con los tesauros, que ni explican ni justifican el orden ni los por qué de los tipos de relaciones que establecen.

Lo que marca la diferencia con los actuales tesauros es la presencia de una mayor variedad de relaciones entre conceptos de dominios formalizados y su función más importante es la de almacenar conocimiento de forma que pueda ser utilizada por sistemas automáticos capaces de realizar deducciones a partir de la variedad de relaciones entre los conceptos. (García, 2004)

Las ontologías son un instrumento de descripción, son una lista jerárquica de términos de un área específica, son vocabularios categorizados que incluyen un análisis semántico de palabras para lograr zonas de significación común y poder enlazarla con otras. Son “artefactos técnicos constituidos por un vocabulario específico utilizado para describir una cierta realidad, más un conjunto de hipótesis explícitas en relación con el previsto significado de las palabras del vocabulario”. (Guarino Cf. Zúñiga, 2001)

Las relaciones significan determinado estado de las cosas, por eso es que son importantes para representar contextos, y como tal indicadas para la gestión documental archivística, pues pueden describir más puntualmente las actividades de las organizaciones, las tareas concretas de campos determinados y, además, la estructura de almacenamiento de las bases de datos.

Mientras la ontología filosófica crea un sistema de categorías para desmenuzar la realidad conceptualmente y ofrecer lo indudable sobre su naturaleza, su campo de

acción y sus límites, la ontología informacional parte de esa clasificación para dotar de sentido, de significado, de semántica, de relaciones, a los conceptos en virtud de un contexto gnosoelógico.

Silla y trono, son dos artefactos físicos, cada uno tiene su conceptualización, sus definiciones son traducibles en un lenguaje, una se distingue de la otra en virtud, por ejemplo, de las relaciones de poder. Pero solo reconociendo el contexto de ciertos porteros y recepcionistas, que descansan en el trono que otorga el poder de sus pequeñas y efímeras decisiones cotidianas, se alcanza a comprender cabalmente un discurso registrado en un documento de archivo para poderlo representar significativamente.

Las ontologías de los sistemas de información, entonces, tienen un componente semántico importante para disminuir la confusión terminológica y conceptual, y son un ejercicio de interpretación en ambiente Web. Las ontologías constituyen el inicio de una Web semántica (Berners-Lee, 2001) que entienda y distribuya en virtud de las diferencias contextuales y también son el comienzo de la representación de la verdad conocida ¿Alguna vez han estado más cerca Filosofía e información?

Para los sistemas integrados de gestión de información los principios fundamentales de la Archivística no pueden pasar inadvertidos. Sin embargo, en estos tiempos de convergencia, no son los únicos a través de los cuales se garantiza la autenticidad. La memoria documental es fidedigna, especialmente mientras más y mejor se conozca y comprenda, mientras más se comparta y sirva, y para ello el duro camino que queda estriba en ofrecer cada vez mejores modelos de representación de esos conocimientos.

La organicidad documental no es solo formal, es un *continuum* intencionado, y por tanto, profundamente ligado también a los significados, siempre paradójicos, que por lo menos ayudan a concluir a esta ponente que Alicia huía buscando regresar.

Referencias:

BARRETO, A. <abarreto-l@aldoibct.bighost.com.br> "A zona cinza". Lista de divulgação, discussão e crônicas do IASI - Instituto para Inserção na Sociedade da Informação. Inclusão social - Informação e Desenvolvimento humano. Mensaje 13-03-05. [Consultado el 13-03-05]. Archivo en <http://www.e-iasi.org>

BATENSON, G. (1980) *Mind and nature: a necessary unity*. New York: Bantam.

BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. The Semantic Web, *Scientific American*, 2001.

Disponible en:

<http://www.scientificamerican.com/2001/0501issue/0501berners-ee.html> (Consulta: enero de 2005)

BYNUM, T.; MOOR, J. (1998) *The digital Phoenix: How computers are changing Philosophy*. Oxford: Blackwell.

BIZJAK, P. *Memory Managers: A New Paradigm of Library Science*. Disponible en:

<http://www.webpages.uidaho.edu/~mbolin/bizjak.html> (Consulta: octubre de 2004)

CAZAU, P. (2002) *Las publicaciones científicas en Internet*. Disponible en:

http://ucm.es/BUCM/psi/guia_red_pub.htm (Consulta: septiembre de 2004)

DELEUZE, G. (1998) *Lógica do sentido*. São Paulo: Perspective.

DELEUZE, G.; GUATTARI, F. (1995) *Mil platôs: capitalismo e esquizofrenia*. São Paulo: Editora 34.

DOLLAR, CH. (1994) *Tecnologias da informação digitalizada e pesquisa acadêmica nas ciências sociais e humanas: o papel crucial da arquivologia*. Disponible en:

<http://www.cpdoc.fgv.br/revista/arg/137.pdf> (Consulta: enero de 2002)

DRUMOND, S. (2003) *Virtual organization of knowledge in cyberspace*. Disponible en:

http://www.dgz.org.br/dez03/Art_05.htm (Consulta: octubre de 2004)

FLORIDI, L. (2002) *On defining library and information science as applied philosophy of information*. *Social Epistemology*, v.16, n.1, 37-49.

FLORIDI, L. (2003) *Two approaches to the Philosophy of Information*. *Minds and Machines*, 13(4), 459-469.

FLORIDI, L. (2004) *Open problems in the Philosophy of Information*. *Metaphilosophy*. Disponible en: <http://www.ethics.acusd.edu/video/CAP/CMU2001/Floridi/index.html>. (Consulta: octubre 2004)

GARCÍA, A. (2004) *Instrumentos de representación del conocimiento: tesauros versus ontologías*. *Anales de Documentación*, n.7, pp. 79-95.

GUTIÉRREZ, A. (2000) *Alfabetización multimedia*. Disponible en:

<http://www.doe.uva.es/alfonso/web/AlfMultIndice.htm> (Consulta: septiembre de 2003)

GRUBER, T. (1993) *A translating approach to portable ontology specifications*. *Knowledge Acquisition*, 5, pp.199-220.

HEDSTROM, M. (1993) *Descriptive Practices for Electronic Records: Deciding What is Essential and Imagining What is Possible*. *Archivaria*

LÉVY, P. (1996) O que é o virtual? São Paulo: Editora 34.

LOPEZ, L. (1998) A imagem e a sombra da arquivística. Rio de Janeiro: Arquivo do estado do Rio de Janeiro.

PÉDAUQUE, R. (2003) Document: forme, signe et médium, les re-formulations du numérique. Disponible en: http://archivesic.ccsd.cnrs.fr/sic_00000511.html (Consulta: marzo de 2004)

PIERCE, S. (1992) Dead Germans and the theory of librarianship. *American Libraries*, 23, 641-643.

PONJUÁN, G. (2004) Gestión de información. Dimensiones e implementación para el éxito organizacional. Nuevo Paradigma.

QIN, J.; PALING, S. (2000-01) Converting a controlled vocabulary into an ontology: the case of GEM. Disponible en: <http://informationr.net/ir/6-2/paper94.html> (Consulta: septiembre de 2003)

SCOTT, N. (2005) Uma estrutura de estudo para Bibliotecas virtuais baseada na Arqueologia (A Framework for Internet Archeology: Discovering Use Patterns in Digital Library and Web-Based Information, Resources) Disponible en: http://www.firstmonday.org/issues/issue10_2/nicholson/ (Consulta en: febrero de 2004)

SHANNON, C. (1993) *Collected Papers.*, N.J.A. Sloane and A.D. Wyner (eds.)

WALLACE, D. (1993) Metadata and the Archival Management of Electronic Records: A Review. *Archivaria*.

ZÚÑIGA, G. (2001) Ontology: Its transformation from Philosophy to Information Systems. Disponible en: http://www.acton.org/about/staff/zuniga_PDFs/Ontology_Info_Systems.pdf (Consulta: octubre de 2004)