

Critical Factors for Digital Records Preservation

Katia P. Thomaz

Tech-in Gestão, Projetos e Tecnologia, Brazil

katia.thomaz@uol.com.br

Abstract

The research on the preservation of digital records, developed in the Information Science Post-Graduation Program at Universidade Federal de Minas Gerais – UFMG (Brazil) using a transdisciplinary approach, has investigated the computer environment of medium and large Brazilian public organizations, in order to identify factors that play a key role in ensuring the long term preservation of their records. This paper presents an overview of the research, outlining the objectives and methodology, summarizing the conceptual analysis, and discussing its main findings, including 160 critical factors for long term preservation of digital records, and a digital preservation information model showing relationships between objects, actors and actions.

Keywords: archival science, records management, electronic records, digital preservation.

Introduction

The increasing production of digital documents, mainly documents originally created in the digital environment, has threatened the human capacity of using these records as reliable information sources, due to the new challenge imposed by their preservation. As stated by our Spanish colleague Vasquez de Parga on UNESCO's 1998 World Panel on Communication and Information, digital preservation is probably "the greatest challenge facing the archive community throughout the world - the challenge of accepting its full responsibility for ensuring adequate archival processing of these new records and incorporating them fully into their countries' archival systems so that historical memory can be permanently preserved, administrative agencies can function properly and citizens' rights, based on the evidential and legal value of such records, can be protected".

While valuable studies have been undertaken in many countries to preserve digital records, as well as its challenging issues, preservation requirements and archival functions, an academic research has developed in the Information Science Post-Graduation Program at Universidade Federal de Minas Gerais – UFMG (Brazil) investigated the computer environment of medium and large Brazilian public organizations, in order to identify factors that play a key role in ensuring the long term preservation of their digital records.

Material published as part of this journal, either on-line or in print, is copyrighted by the Informing Science Institute. Permission to make digital or paper copy of part or all of these works for personal or classroom use is granted without fee provided that the copies are not made or distributed for profit or commercial advantage AND that copies 1) bear this notice in full and 2) give the full citation on the first page. It is permissible to abstract these works so long as credit is given. To copy in all other cases or to republish or to post on a server or to redistribute to lists requires specific permission and payment of a fee. Contact Publisher@InformingScience.org to request redistribution permission.

This paper presents an overview of the research, outlining the objectives and methodology, summarizing the conceptual analysis, and discussing its main findings.

Literature Review

The literature on preserving digital records deals with challenging issues, preservation requirements and archival

functions. Challenging issues have been studied by a number of authors (Beagrie & Jones, 2002; Brand, 1999; Bullock, 1999; Conway, 1996; European Commission, Interchange of Data between Administrations [EC/IDA], 2001; Hedstrom, 1997/1998; International Council on Archives, Committee on Electronic Records [ICA/CER], 1997; Lusenet, 2002; National Library of Australia, 2003; Public Records Office [PRO], 1999; Rothenberg, 1999; Thibodeau, 2001; Watters & Garret, 1996). This literature allows grouping the challenging issues into the following categories:

I - Lack of evaluation policies. The impact caused by the definition of selection criteria is very crucial in the digital environment. The digital document which is not selected in the initial phases of its lifecycle, is likely to be lost or become useless in the future.

II - Lack of descriptive policies. The complex nature of technology demands a rather detailed descriptive approach, i.e. metadata to digital objects for maintenance. Metadata for both internal and external document elements became crucial.

III - Physical vulnerability. Hardware and storage media are inherently unstable and, without the appropriate installation and maintenance, they can deteriorate very quickly, even if they do not seem to be damaged externally.

IV - Logical vulnerability. The digital environment is sensitive to changes (some emerge from the management needs itself) which can compromise the integrity, authenticity, and the history of digital objects.

V - High technological obsolescence. The technology renewal cycle is short (3-5 years) as opposed to decades and centuries associated with the preservation of physical objects.

VI - High technological dependence. All digital objects require specific hardware and software to be accessed and each of these elements generally requires contract agreements, which are commonly difficult to negotiate.

VII - Difficulty in recruiting properly qualified staff. The technology involved in accessing digital objects requires a significant diversity of activities that can be performed by scarce specialists.

Another part of the preservation literature focuses on preservation requirements. Based on the reference model Open Archival Information System – OAIS, Bullock (1999) identified actions (“requisites”) to be observed step-by-step in the preservation of electronic records. They are:

I - Fixing the object as a discrete whole. The boundaries of a digital object are not clear, especially if it is compound object created by assembling different media or by linking to resources from around a network.

II – Preserving the physical presence. It refers to keeping the computer file, the series of 1s and 0s that are the basis of a digital object.

III - Preserving the content. It refers to keeping the ability to access the content at its lowest level (e.g., ASCII text) without the embellishments of font variations and layout features.

IV - Preserving the presentation. It refers to keeping the original look of a digital object. The layout specifications must also be preserved, especially when they contribute significantly to the understanding and interpretation of the content. Layout specifications include different font faces and sizes, the use of white space, columns, marginalia, headers, footers, pagination, and so on, sometimes separated from the content.

V - Preserving the functionality. It refers to keeping the dynamic aspects of a digital object (e.g., multimedia components, the hypertext format, the capability of generating dynamic content automatically from data stores, navigation functions, and interactive tables of contents).

VI - Preserving the authenticity. It refers to securing digital object against unauthorized changes and monitoring digital object through multiple 'copying' cycles to ensure that each copy is an acceptable rendition of the original.

VII - Locating and referring to the original object over time. It refers to being able to match a citation to a digital object, and to distinguish it from other versions or editions.

VIII - Preserving provenance. It refers to asserting the origin and chain of custody of an object and contributes to define it as a whole. This helps to confirm that the work is authentic and its content is intact.

IX - Preserving context. It refers to describing the hardware and software dependencies of a digital object, and its mode of distribution and linkages to other digital objects.

The third segment of the preservation literature concerns archival functions. Open Archival Information System (OAIS) has identified particular archival functions (see Consultative Committee for Space Data Systems [CCSDS], 2002). They are:

I - Ingest. The OAIS entity that contains the services and functions for accepting Submission Information Packages from Producers, preparing Archival Information Packages for storage, and ensuring that Archival Information Packages and their supporting Descriptive Information become established within the OAIS.

II - Archival storage. The OAIS entity that contains the services and functions for storing and retrieving Archival Information Packages.

III - Data management. The OAIS entity that contains the services and functions for populating, maintaining, and accessing a wide variety of information. Some examples of this information are catalogs and inventories on what may be retrieved from archival storage, processing algorithms that may be run on retrieved data, consumer access statistics, consumer billing, event based orders, security controls, and OAIS schedules, policies, and procedures.

IV - Administration. The OAIS entity that contains the services and functions for controlling the operation of the other OAIS functional entities on a day-by-day basis.

V - Preservation planning. The OAIS entity that contains the services and functions for monitoring the OAIS environment and for providing recommendations to ensure that the stored information remains accessible to the designated community over the long term, even if the original computational environment becomes obsolete.

VI - Access. The OAIS entity that contains the services and functions for accessing the archival information holdings and related services.

Methodology

The lack of a standard research methodology in the digital preservation field required the development of a specific process of investigation. Initially, since digital records exist in a complex context, a key instrument to identify digital preservation factors during the data collection phase was created. According to Chiavenato (1987), the organization can be seen as a system that operates through interrelated and interdependent variables summered as activities, management, people, environment and technology. These five organization variables, extended and enriched by OAIS reference model (Consultative Committee for Space Data Systems [CCSDS], 2002) and other General Management theories (Blau & Scott, 1970; Bowditch & Buono, 1997; Champion, 1985; Hall, 1984; Miranda & Inácio, 1978) resulted in the Digital Object Context Model, a simple diagram shown in Figure 1.

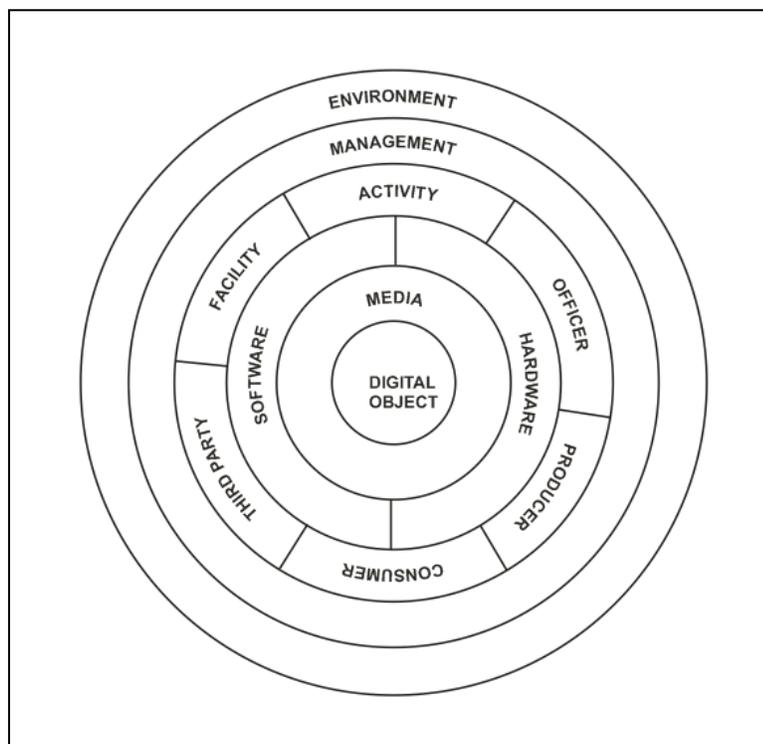


Figure 1: Digital Object Context Model
Source: Thomaz, 2004

The Context Model includes 13 organization variables, as follows:

I Archival Digital Object: digital object created in the course of an activity, as a mean for and byproduct of it, and kept for future access;

II Storage Media: different types of physical materials in which the archival digital object is recorded and stored, like floppy disk, hard disk, tape and optic disk;

III Presentation Software: software which is needed to present all or part of an archival digital object, in a way that people can understand;

IV Processing Hardware: hardware which is needed to process the storage media and run the presentation software;

V Maintenance Activity: organization action which aims at properly keeping – in order, running, updated, etc. – the archival digital object, the storage media, the presentation software, the processing hardware and the facility;

VI Business Activity: an organization action which aims at carrying on the organization business;

VII Officer: a person, often a technician or a lower employee, who performs the maintenance activity;

VIII Producer: a person or client-system that produces the archival digital object to be stored;

IX Consumer: a person or client-system interested in the archival digital object;

X Third Party: a person or organization that makes products, provides services, or certifies – to declare true, accurate, certain by formal statement – a digital object, technology and facility;

XI Facility: a physical space holding processing hardware or storage media;

XII Management: an organization structure that is needed to perform the maintenance and business activities;

XIII Environment: all the conditions, circumstances, and influences surrounding, and affecting the development of the organization.

The relationships between the organization variables show how the center depends on the periphery. In fact, the Archival Digital Object depends on the Storage Media where it is recorded, on the Presentation Software that interprets it, on the Processing Hardware that reads and processes it, on the interaction between Maintenance Activity, Business Activity, Officer, Producer, Consumer, Third Party and Facility that manages it, on the Management to establish preservation policies, and finally, on the Environment in which it is kept.

Due to the difficulty of studying actual cases of long-term digital preservation in Brazil, as demonstrated by Santos (2002), it was necessary to establish the technology change project, a process involving the substitution of any hardware/software components in a computer environment, as an initial object of analysis. In this sense, a technology change project could be considered as induced obsolescence. It was presumed that an investigation of the different elements of this changing process, recently promoted by any organization, could reveal a set of essential factors related to digital preservation. The reason was that this changing process represents a high level risk for preservation, since the data must go through a migration procedure from the environment in which they are controlled and stable, to a new environment, without losing any of their original characteristics.

The primary objectives of the research included increasing knowledge in the digital records preservation field. This general objective, if reached, could stimulate the emergence of new works and research projects that would complement this initiative. As a specific objective, the research looked for digital preservation factors, i.e. those elements that must be controlled to ensure the long term preservation of digital records. In other words, it aimed at identifying the constraints that should be taken into account to reduce the risk of losing digital records.

The definition of scope and primary objectives were conducive to a qualitative approach. In this kind of investigation, the researcher acts as an interpreter of the reality, developing concepts, ideas, and understandings of the object of study based on the patterns found in the data. As research methodology, a mix of case studies, grounded theory, interviews, documents analysis, focus groups and peer reviews were chosen, in which one method complemented the other.

The following companies took part in the research: Companhia Energética de Minas Gerais – CEMIG (Power Company of the State of Minas Gerais), Companhia de Saneamento Básico de Minas Gerais – COPASA (Water Company of the State of Minas Gerais), Empresa de Informática e Informação do Município de Belo Horizonte – PRODABEL (Computer and Information Company of the City of Belo Horizonte), and Companhia de Tecnologia da Informação do Estado de Minas Gerais – PRODEMGE (Information Technology Company of the State of Minas Gerais). These companies hold many of the digital records of the State of Minas Gerais and of the City of Belo Horizonte, and are responsible for the safekeeping and administration of these records. The potential interest in the results of this study on the part of other Brazilian utilities companies was another reason for the choice of these case studies.

From all the technology change projects successfully promoted during the period 2000-2002 by the companies that took part in the research, sixteen were selected to have their technical records analyzed. These records were 71 documents totaling 522 pages (e.g., minutes, memoranda, correspondences, procurement documents and technical studies). The documents were generated during the projects using a variety of methods, both conventional and digital. This constituted the

first object of study in the data collection phase. The period of three years was considered ideal for the investigation, because of the average technology renewal cycle identified by the relevant literature (Brand, 1999; Hedstrom, 1997/1998). Figure 2 shows the distribution of the investigated projects for each company as well as the type of component involved in the technology change.

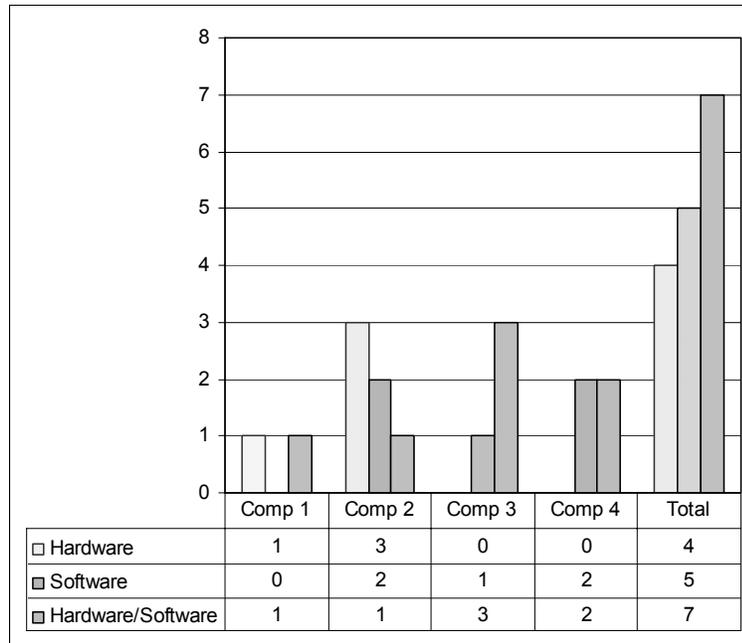


Figure 2: Number of investigated projects per company and type of element involved in the technology change
Source: Thomaz, 2004

Two categories of respondents were interviewed in the investigated companies. The first category involved four technical support managers who were interviewed at the beginning of the data collection phase. A technical support manager is a middle level manager associated with the computer industry who manages, plans and supervises technical support activities, including hardware and software studies, controls the performance of systems and technical resources in order to correct irregularities, and suggests modifications, improvements, upgrades and adjustments of information systems in order to attend to new user requirements. The second category of respondents covered 40 computer specialists, including system analysts and technicians (programmers) that had at least five years of work experience and participation in one technology change project. These respondents volunteered to take part in the evaluation process of the digital preservation factors.

Overall, the data collection phase was performed through four steps:

I – Interview with a technical support manager of each company to describe the company itself and list the technology change projects concluded during the period 2000 until 2003.

II – Documents analysis of projects to identify the first candidates to digital preservation factor

III – First interaction with computer specialists, gathered into eight focus groups, to validate and refine the set of digital preservation factors

IV – Second interaction with computer specialists through e-mail to finalize the set of digital preservation factors.

All these activities required several research tools, 14 in total, gradually refined step-by-step. Through an interactive process of data collecting, analyzing, evaluating, consolidating and arranging, a repetitive cycle seeking a better specification and quality of the information, the research evolved from the raw data, found in the project records, to the last interaction with the computer specialists to evaluate the relevance of a group of digital preservation factors and the associated Glossary.

It is interesting to note that during the research process, the professionals of the investigated organizations realized the relevance of the topic of digital preservation, even though there was no previous actual concern towards it. The involvement of those professionals, extremely competent workers that embraced the investigation was another strong element of the study, contributing to a good researcher-respondent relationship and the study's results.

Findings

The study identified 160 digital preservation factors (see Appendix). A Glossary with 366 entries detailing the meaning of organization variables, digital preservation factors, and associated terms was also produced. In addition, Figure 3 shows the distribution of the digital preservation factors according to the organization variables in the model depicted in Figure 1. The frequencies of the factors indicated that the organization variables Environment, Maintenance Activity, Archival Digital Object, and Storage Media, should demand greater preservation efforts since most of the factors reside in these classes.

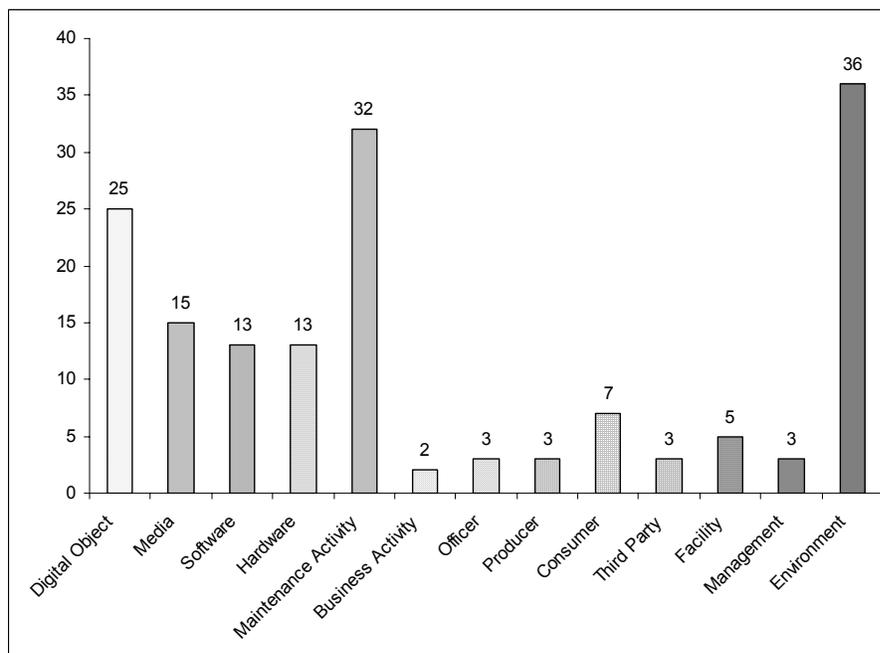


Figure 3: Distribution of Digital Preservation Factors per Class of Objects
Source: Thomaz, 2004

In order to validate and consolidate the research findings, the digital preservation factors were analyzed in light of the key aspects discussed in the literature review section: challenging issues, preservation requirements and archival functions. A table was constructed with the factors as rows and these aspects, as columns. Next, each pair factor/aspect was tested on the following

Critical Factors for Digital Records Preservation

question: “Does this factor contribute in any way to this aspect?” For example, “Does digital object identification contribute in any way to lack of evaluation policies?”, “Does digital object identification contribute to lack of descriptive policies?” and so on. The answer for each pair was based upon their definitions. Each cell was filled with an “X”, in case of a positive answer, or left blank, in case of a negative answer. Then, a histogram was build with the occurrences of factors by aspect. The construction process of these tables involved discussion with and revision by peers.

Regarding seven challenging issues, the results, shown in Figure 4, indicated that the highest concentrations of digital preservation factors are on Physical Vulnerability, with 61 factors, Logical Vulnerability, with 49 factors, and Technological Obsolescence, with 27 factors. Hence, physical and logical vulnerability should constitute the main digital preservation concerns, and not technological obsolescence.

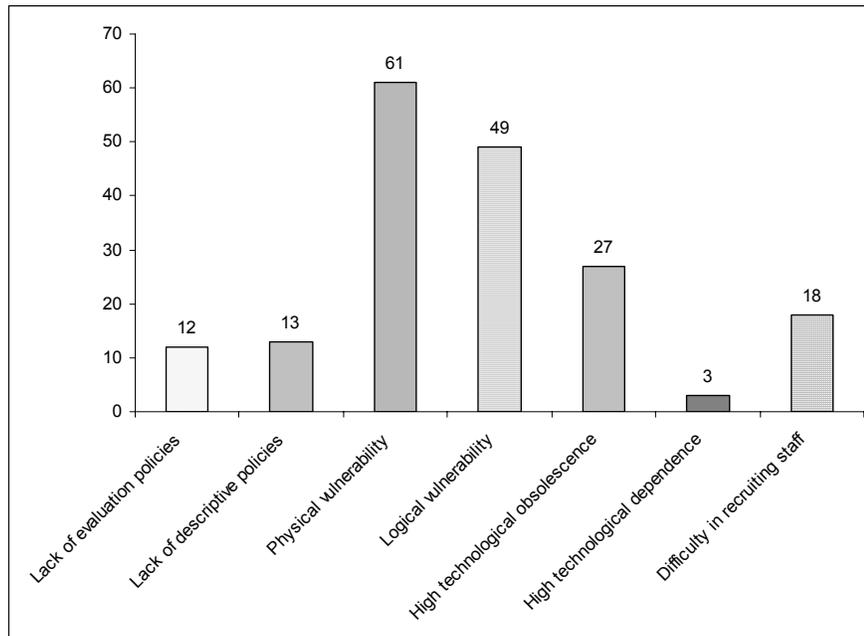


Figure 4: Distribution of Digital Preservation Factors per Challenging Issue
Source: Thomaz. 2004

Regarding nine preservation requirements, the results, shown in Figure 5, indicated that the highest concentration of digital preservation factors are on Preserve the Content (75 factors), Preserve the Physical Presence (71), Preserve the Presentation (41), and Preserve the Functionality (40).

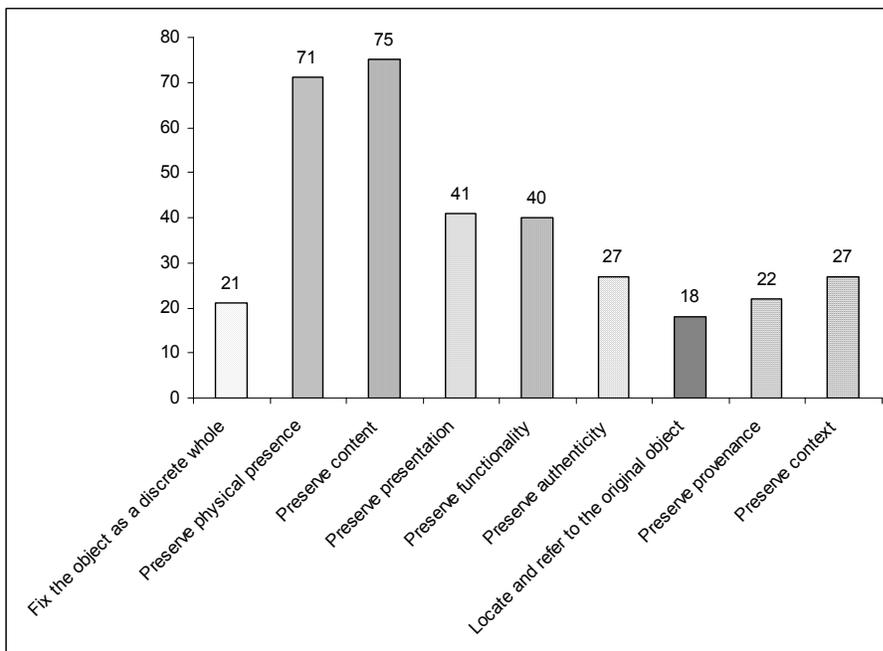


Figure 5: Distribution of Digital Preservation Factors per Preservation Requirement
Source: Thomaz, 2004

Also, regarding six archival functions, the results, shown in Figure 6, indicated 75 factors for Administration, 50 for Preservation Planning, 17 for Archival Storage, 9 for Access, and 6 for both Ingest and Data Management.

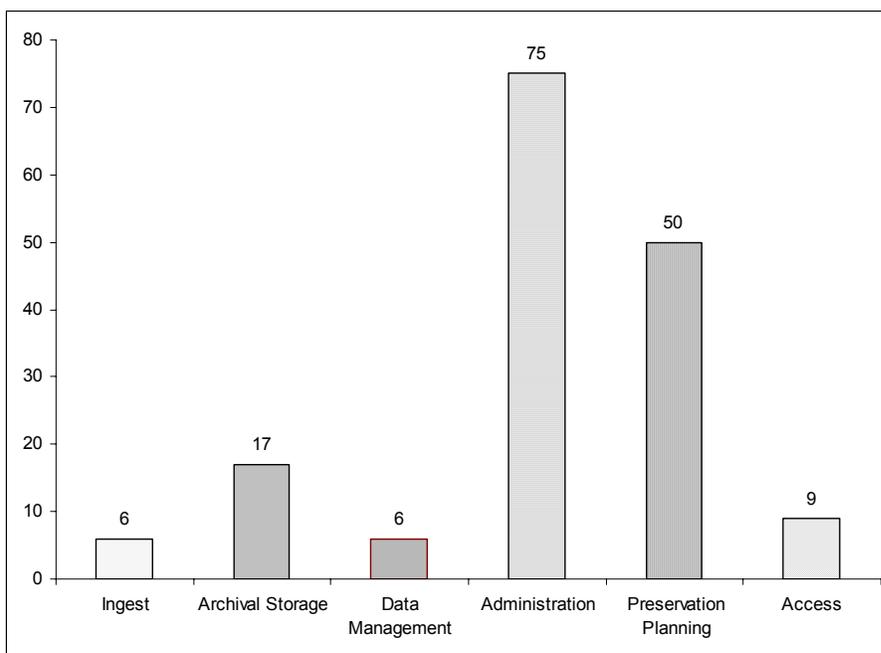


Figure 6: Distribution of Digital Preservation Factors per Archival Function
Source: Thomaz, 2004

During the analytic process, based on the comparison among organization variables, digital preservation factors, and literature, it was noticed that new findings could contribute to extend or detail Figure 1. In fact, new approaches were developed, leading to the Digital Preservation Information Model, elaborated using the Unified Modeling Language – UML’s class diagram, shown in Figures 7a to 7b in the Appendix. In addition to the relationships between the organization variables identified in the Context Model, the developments, which define the different approaches and roles, are also depicted. The Information Model starts with the Management handling the other internal organization variables. Producer performs one or more Business Activity that generates zero or more Groups of Records, and some of them are designated as Long-Preservation Group. A Long-Preservation Group contains one or more Records, and some of them are Digital Records. A Digital Record is composed of one or more Digital Objects that are recorded on certain Storage Media and presented by specific Presentation Software, both processed by compatible Processing Hardware. This technological complex – Storage Media, Presentation Software and Processing Hardware – is maintained by one or more Maintenance Activities performed by one or more Officers. It also supports one or more Business Activities that support one or more Consumers. At the end of the figure it is possible to see the influence of the Environment over all internal organization variables in terms of change/evolution, including the Management itself.

There are some aspects to be highlighted in the Information Model: the transition of Group of Records to Long-term Group (Figure 7b) that would involve a complex decision-making; the multiplicity of the Record and the delimitation of the Digital Objects which compose the Record (Figure 7b); the Maintenance Activities, if they do not assemble enough information on the technological components, can lead the whole model into a collapse (Figure 7c); and the difficulty in evaluating influences of the Environment over the various organization variables (Figure 7d; see Appendix).

Also, the association of Information Model and digital preservation factors provides a broad understanding of what is necessary to ensure the careful preservation of digital records over a long term, taking into account the three functional categories of metadata identified by the Working Group on Preservation Metadata [PMWG] (2001):

I Descriptive: facilitating source discovery and identification;

II Administrative: supporting source management within a collection;

III Structural: binding together the components of complex information objects.

Conclusion

The results of this study are relevant for different fields, including archival science, computer science, and information science. The methodology used has proven to be extremely effective, and provided the desired precision.

The Digital Preservation Information Model and associated factors offer a holistic view of the digital environment and provide the information professional and/or researcher with a more accurate analysis of its components. Any particular instance of a digital archive finds itself mirrored in the Information Model, allowing each mentioned aspect to be the object of implementation and detailed study, so that the professional or researcher does not forget any component or lose the sense of the whole.

Our Information Model recognizes the highly distributive nature of digital records in the organization and the need for local implementation of effective policies and procedures to support its preservation. Among the classes involved in long-term digital preservation, besides the conventional hardware-software-data triad, it also considers the Officer, the Supplier, the Manufacturer,

the Certifier, the Storage Media, the Computer Room, the Storage Room, the Management and the Environment. There is also a great amount of digital preservation factors inside the classes, which will have to be taken into account no matter what position the class takes in the Information Model.

Overall, the Digital Preservation Information Model and associated factors offer a guidance for:

- gaining a broad understanding of the organization variables involved in the preservation of and access to digital records over a long term;
- describing and comparing current and future digital preservation strategies;
- providing a foundation that may be expanded by other efforts to address long term preservation of records that are not in digital form (e.g., paper, microfilm etc.); and
- guiding the identification and generation of new standards.

The major research findings – Digital Preservation Information Model, associated factors, and comparisons with other key aspects of digital preservation literature – represent innovative and original approaches, and may constitute useful tools for digital preservation management, since they allow the evaluation of risks associated with the necessary decision-making. For example, once a digital preservation program is planned, the overlooked factors would automatically reveal the related risks, that is, the factors, challenging issues, preservation requirements and/or archival functions which were weakly or not at all considered. Uncertainty reigns in the absence of information (Giddens, 1991). In contrast, information transforms uncertainty into a risk, that is, in the probability that a certain negative effect may occur. The risk can be better predicted and managed when there is sufficient and precise information. These tools, once they represent the current state-of-the-art, should be constantly revisited and updated so that they can be adapted to new solutions generated by the evolution of technologies and processes.

Finally, the research progressed from theory to praxis, since this path provides instruments of effective concreteness. This fact is relevant because of the urgency of researches, discussions and practices related to the subjects of heritage and memory.

References

- Beagrie, N., & Jones, M. (2002). *Preservation management of digital materials: A handbook*. London: Digital Preservation Coalition. Retrieved November 6, 2003, from <http://www.dpconline.org/graphics/handbook/>
- Blau, P. M., & Scott, R. (1970). *Organizações formais* [Formal organizations]. São Paulo, SP: Atlas.
- Bowditch, J. L., & Buono, A. F. (1997). *Elementos de comportamento organizacional* [A primer on organizational behavior]. São Paulo, SP: Pioneira.
- Brand, S. (1999, February). Scaping the digital dark. *Library Journal*, 124 (2), 46-48.
- Bullock, A. (1999). *Preservation of digital information: Issues and current status*. Ottawa: National Library of Canada. Retrieved November 25, 2003, from <http://www.nlc-bnc.ca/publications/1/p1-259-e.html>
- Champion, D. J. (1985). *A sociologia das organizações* [The sociology of organizations]. São Paulo, SP: Saraiva.
- Chiavenato, I. (1987). *Teoria geral da administração* (Vol. 1, 3rd ed.) [General management theory]. São Paulo, SP: McGraw Hill.
- Consultative Committee for Space Data Systems. (2002). *Reference model for an Open Archival Information System (OAIS)*. Washington, DC. Retrieved November 14, 2002, from <http://www.ccsds.org/documents/650x0b1.pdf>

Critical Factors for Digital Records Preservation

- Conway, P. (1996). *Preservation in the digital world*. Washington, DC: Commission on Preservation and Access. Retrieved November 23, 2003, from <http://www.clir.org/pubs/reports/conway2/>
- European Commission, Interchange of Data Between Administrations. (2001). *Model Requirements for the management of electronic records (MoReq)*. Luxembourg. Retrieved October 23, 2003, from <http://europa.eu.int/ISPO/ida/export/files/en/635.pdf>
- Giddens, A. (1991). *As conseqüências da modernidade* [The consequences of modernity]. São Paulo, SP: UNESP.
- Hall, R. H. (1984). *Organizações, estruturas e processos* [Organizations: structure and process]. Rio de Janeiro, RJ: Prentice Hall do Brasil.
- Hedstrom, M. (1997/1998). Digital preservation: A time bomb for digital libraries. *Computer and the Humanities*, 31(3), 189-202. Retrieved November 25, 2003, from <http://www.uky.edu/~kiernan/DL/hedstrom.html>
- International Council on Archives, Committee on Electronic Records. (1997). Guide for managing electronic records from an archival perspective. Paris. Retrieved February 20, 2004, from <http://www.ica.org/biblio.php?pdocid=3>
- Lusenet, Y. (2002). Digital heritage for the future. *Cadernos BAD*, 2, 15-27.
- Miranda, G., & Inácio M. P. (1978). *Organização e métodos* (4th ed.) [Organization and methods]. São Paulo, SP: Atlas.
- National Library of Australia. (2003). *Guidelines for the preservation of digital heritage*. Paris: UNESCO. Retrieved November 25, 2003, from <http://unesdoc.unesco.org/images/0013/001300/130071e.pdf>
- Public Record Office. (1999). *Management, appraisal and preservation of electronic records: Principles* (2nd ed.). Kew. Retrieved February 19, 2004, from <http://www.pro.gov.uk/recordsmanagement/erecords/guidelines/principles.pdf>
- Rothenberg, J. (1999). *Ensuring the longevity of digital information*. Washington: Commission on Library and Information Resources. Retrieved November 26, 2003, from <http://www.clir.org/pubs/archives/ensuring.pdf>
- Santos, V. B. (2002). *Gestão de documentos eletrônicos: Uma visão arquivística* [Electronic records management: An archival perspective]. Brasília: Associação Brasileira de Arquivologia.
- Thibodeau, K. (2001, February). Building the archives of the future. *D-Lib Magazine*, 7(2), 1-13.
- Thomaz, K. P. (2004). *A preservação de documentos eletrônicos de caráter arquivístico: Novos desafios, velhos problemas* [Electronic records preservation: New challenges, old problems]. (Doctoral dissertation, Universidade Federal de Minas Gerais, 2004).
- Waters, D., & Garrett, J. (1996). *Preserving digital information: Report of the Task Force on Archiving of Digital Information commissioned by the CPA and the Research Libraries Group*. Washington, DC: Commission on Preservation and Access. Retrieved November 7, 2003, from <http://www.rlg.org/ArchTF>
- Working Group on Preservation Metadata. (2001). *Preservation metadata for digital objects: A review of the state of the art*. Retrieved December 6, 2001, from Online Computer Library Center Web site http://www.oclc.org/research/projects/pmwg/presmeta_wp.pdf

Appendix

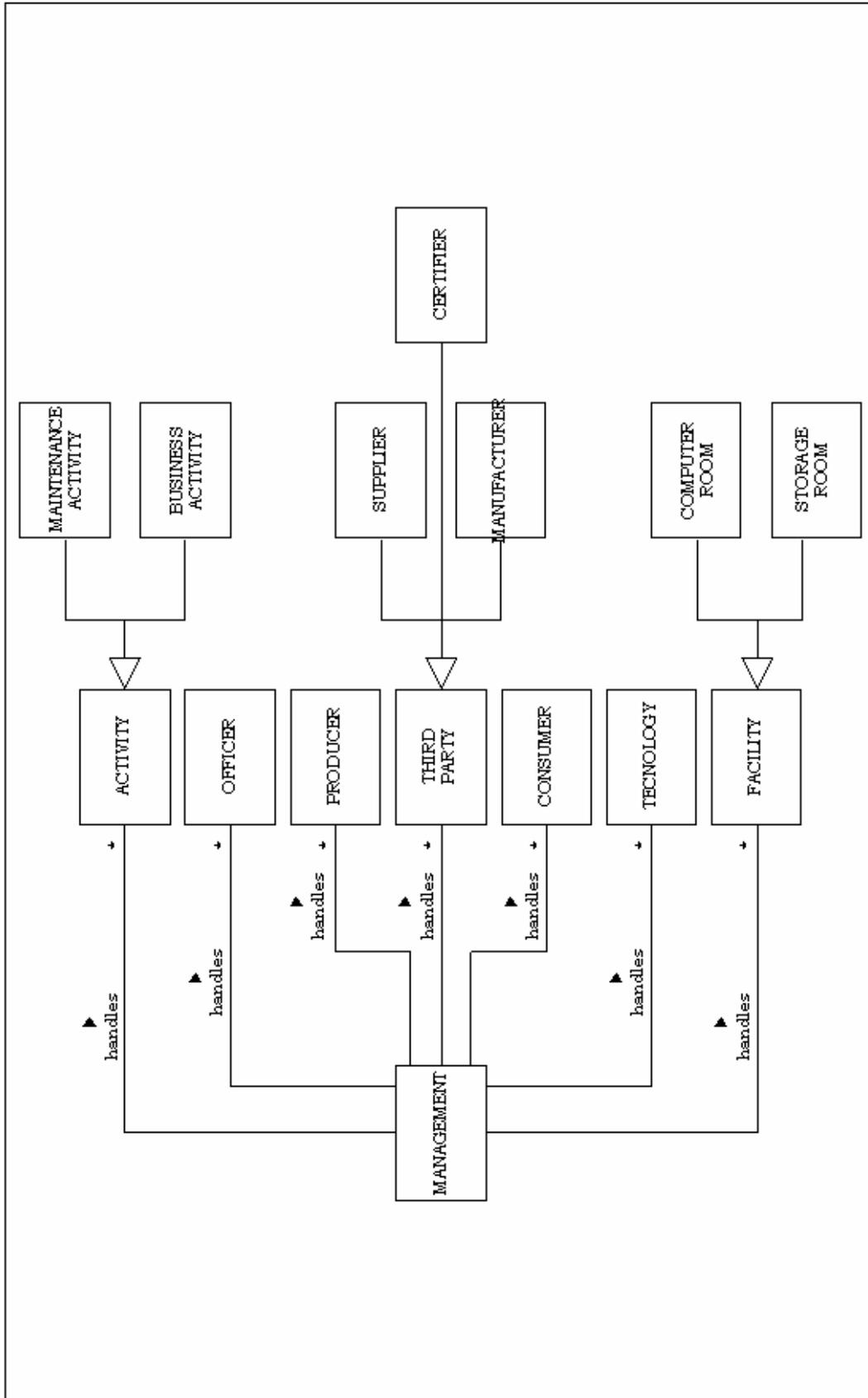


FIGURE 7 a: Digital Preservation Information Model

Source: THOMAZ, 2004

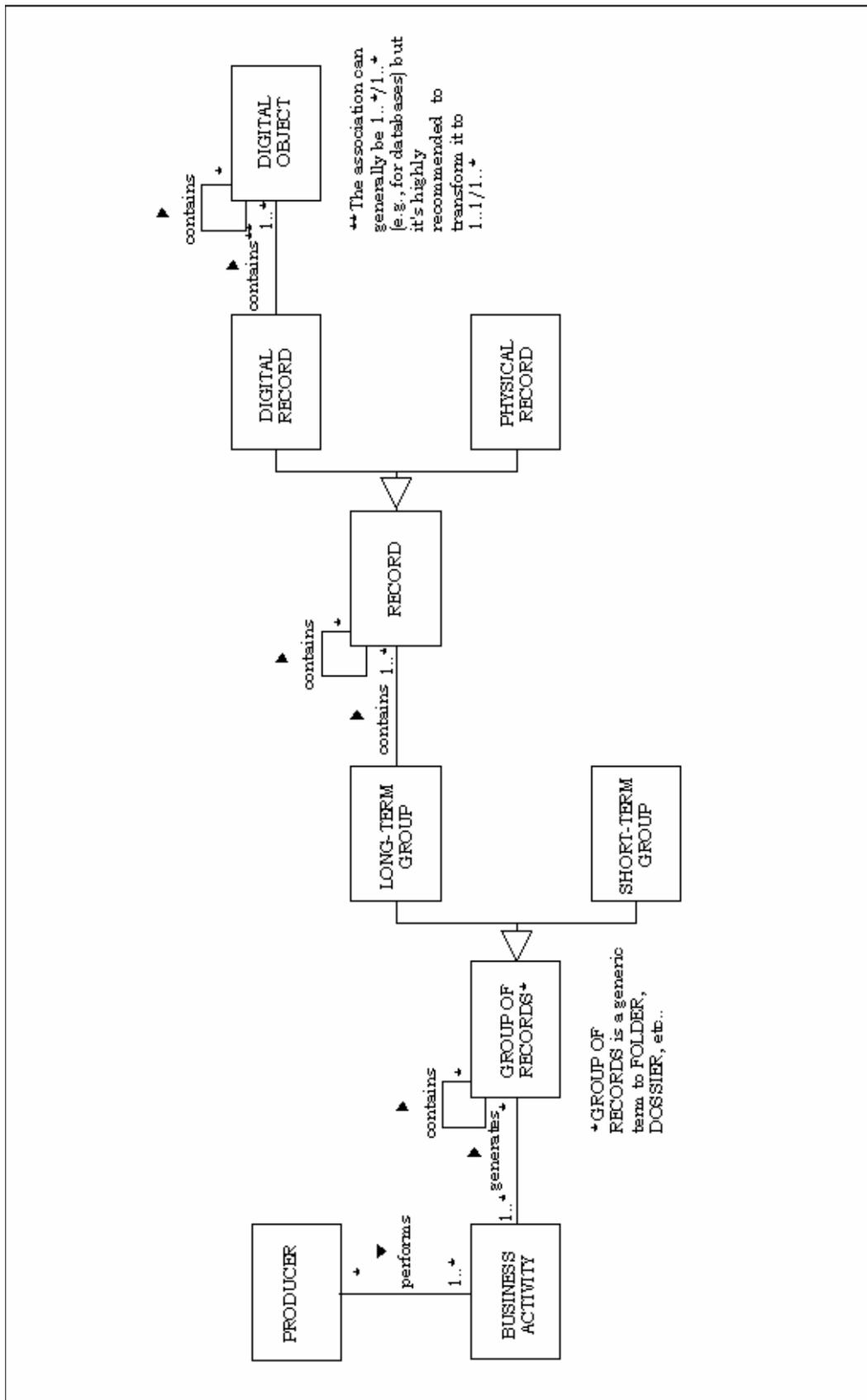


FIGURE 7b: Digital Preservation Information Model (cont.)

Source: THOMAZ, 2004

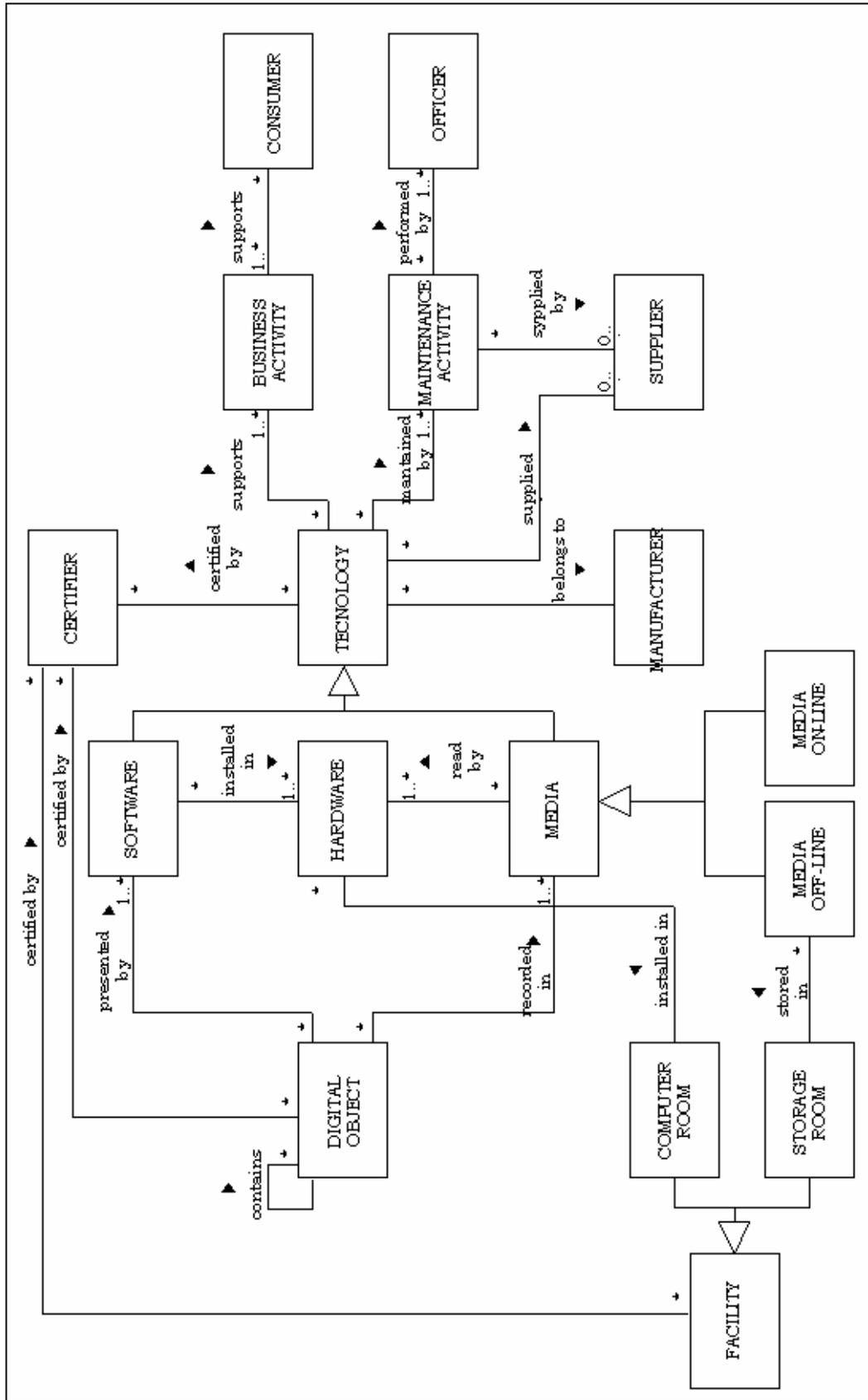


FIGURE 7c: Digital Preservation Information Model (cont.)
 Source: THOMAZ, 2004

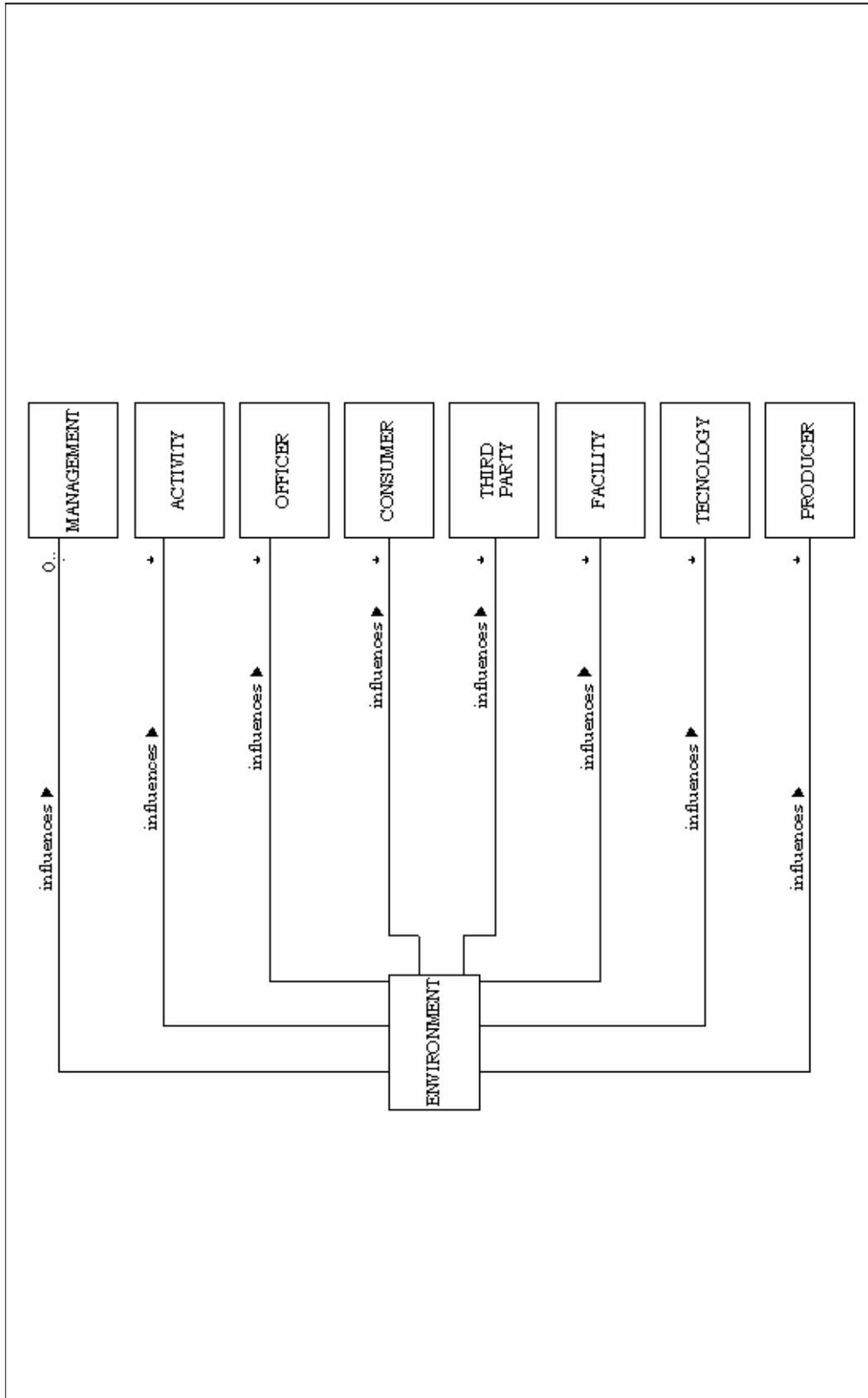


FIGURE 7d. Digital Preservation Information Model (cont.)
Source: THOMAZ, 2004

Digital Preservation Factors

ID	NAME
01 – Archival Digital Object	
01.01	digital object identification
01.02	digital object type
01.03	digital object address
01.04	digital object function
01.05	digital object justification
01.06	digital object security
01.07	digital object expiration
01.08	digital object context
01.09	digital object agreement
01.10	digital object representation
01.11	digital object compression
01.12	digital object prevention
01.13	digital object certification
01.14	digital object logical protection
01.15	digital object description
01.16	digital object presentation
01.17	digital object semantic
01.18	digital object contingency
01.19	digital object responsibility
01.20	digital object area
01.21	digital object volume
01.22	digital object generation
01.23	digital object access
01.24	digital object provenance
01.25	digital object history
02 – Storage Media	
02.01	storage media identification
02.02	storage media type
02.03	storage media format
02.04	storage media address
02.05	storage media certification
02.06	storage media expiration
02.07	storage media requirement
02.08	storage media constraint
02.09	storage media agreement
02.10	storage media description
02.11	storage media prevention
02.12	storage media physical protection
02.13	storage media logical protection
02.14	storage media contingency
02.15	storage media disposition
03 – Presentation Software	
03.01	presentation software identification
03.02	presentation software type
03.03	presentation software certification

Critical Factors for Digital Records Preservation

ID	NAME
03.04	presentation software expiration
03.05	presentation software requirement
03.06	presentation software constraints
03.07	presentation software agreement
03.08	presentation software user interface
03.09	presentation software interface
03.10	presentation software description
03.11	presentation software logical protection
03.12	presentation software prevention
03.13	presentation software contingency
04 – Processing Hardware	
04.01	processing hardware identification
04.02	processing hardware type
04.03	processing hardware certification
04.04	processing hardware expiration
04.05	processing hardware requirement
04.06	processing hardware constraint
04.07	processing hardware interface
04.08	processing hardware agreement
04.09	processing hardware description
04.10	processing hardware prevention
04.11	processing hardware physical protection
04.12	processing hardware logical protection
04.13	processing hardware contingency
05 – Maintenance Activity	
05.01	maintenance activity identification
05.02	maintenance activity description
05.03	maintenance activity agreement
05.04	digital object efficacy
05.05	digital object prevention efficacy
05.06	digital object certification efficacy
05.07	digital object logical protection efficacy
05.08	digital object contingency efficacy
05.09	storage media efficacy
05.10	storage media occupation
05.11	storage media prevention efficacy
05.12	storage media physical protection efficacy
05.13	storage media logical protection efficacy
05.14	storage media contingency efficacy
05.15	storage media disposition efficacy
05.16	presentation software efficacy
05.17	presentation software prevention efficacy
05.18	presentation software logical protection efficacy
05.19	presentation software contingency efficacy
05.20	processing hardware efficacy
05.21	processing hardware occupation
05.22	processing hardware prevention efficacy

ID	NAME
05.23	processing hardware physical protection efficacy
05.24	processing hardware logical protection efficacy
05.25	processing hardware contingency efficacy
05.26	officer efficacy
05.27	third-party efficacy
05.28	facility efficacy
05.29	facility efficacy
05.30	facility prevention efficacy
05.31	facility physical protection efficacy
05.32	maintenance cost
06 – Business Activity	
06.01	business activity identification
06.02	business activity requirement
07 – Officer	
07.01	officer identification
07.02	officer qualification
07.03	officer qualification expiration
08 – Producer	
08.01	producer identification
08.02	producer qualification
08.03	producer qualification expiration
09 – Consumer	
09.01	consumer identification
09.02	consumer qualification
09.03	consumer qualification expiration
09.04	consumer security credential
09.05	consumer security credential expiration
09.06	consumer access right
09.07	consumer access right expiration
10 – Third-Party	
10.01	third-party identification
10.02	third-party qualification
10.03	third-party qualification expiration
11 – Facility	
11.01	facility identification
11.02	facility certification
11.03	facility
11.04	facility prevention
11.05	facility physical protection
12 – Management	
12.01	organization policies and strategies
12.02	organization structure
12.03	budget
13 – Environment	
13.01	digital object relevance
13.02	digital object representation reliability
13.03	digital object compression reliability

Critical Factors for Digital Records Preservation

ID	NAME
13.04	digital object prevention reliability
13.05	digital object certification reliability
13.06	digital object logical protection reliability
13.07	digital object contingency reliability
13.08	storage media reliability
13.09	storage media format reliability
13.10	storage media prevention reliability
13.11	storage media physical protection reliability
13.12	storage media logical protection reliability
13.13	storage media contingency reliability
13.14	storage media disposition reliability
13.15	presentation software reliability
13.16	presentation software user interface reliability
13.17	presentation software interface reliability
13.18	presentation software prevention reliability
13.19	presentation software logical protection reliability
13.20	presentation software contingency reliability
13.21	operational system reliability
13.22	database reliability
13.23	processing hardware reliability
13.24	processing hardware interface reliability
13.25	processing hardware prevention reliability
13.26	processing hardware physical protection reliability
13.27	processing hardware logical protection reliability
13.28	processing hardware contingency reliability
13.29	digital object readability
13.30	digital object presentation reliability
13.31	maintenance specialist availability
13.32	third-party reliability
13.33	facility reliability
13.34	facility prevention reliability
13.35	facility physical protection reliability
13.36	technology alert

Biography



Katia Thomaz is a Director of Tech-in - Gestão, Projetos e Tecnologia Soc. Civil Ltda (<http://www.techin.uaivip.com.br>). She holds a Ph.D. in Information Science from Universidade Federal de Minas Gerais in Brazil. Her research has received the National Information Science Research Award from Associação Nacional de Pesquisa e Pós-graduação em Ciência da Informação (ANCIB) (<http://www.ancib.org.br>), and she has published many articles on digital preservation describing its challenging issues, preservation requirements and archival functions. Her experience includes diverse teaching, managerial and technical skills in system analysis and design, information management, records management, knowledge management, project management, quality management, and organizational strategic planning. Currently she teaches records management and project management to both undergraduate and graduate students.