

Digital Diplomatics and Measurement of Electronic Public Data Qualities: What lessons should be learned?

Introduction

In archival science, the interest for studying the concept of quality is not well-developed. This paper presents the main results of QADEPs project (Qualités des documents et des données électroniques publics) realized at the Geneva Business School with the collaboration of several public and private swiss partners, which are themselves dealing with the definition and the measurement of the important qualities applied to electronic records and data in public institution. It specifies the measurement of three dimensions: the Trustworthiness, the Exploitability and the Representativeness. For the purpose of this paper, we focus in particular on the main indicators and variables related to the Trustworthiness dimension. Classic and digital diplomatics literature represents one of the main resources in this field.

The goal of the project is, in a conceptual phase, to define the variables to appraise the quality of public electronic archives in the perspective of a sustainable conservation in an OAIS system type. In a second empirical phase, the model is tested on three public institutions (Swiss Federal Archives, Archives of the state of Geneva, Archives of the state of Wallis) in order to verify its applicability. This research promises to bring significant benefits: theoretical results as well as empirical results. Regarding the theoretical results, the conceptual framework will be presented as a general mapping, which will summarize the dimensions and indicators of the qualities of electronic records. This structure will represent the various aspects of the quality of public electronic archives. A sample of variables related to authenticity will be then presented to enable an accurate understanding of how we can define and quantify the qualities of a public electronic document to ensure its long term preservation. Moreover, this paper will focus the trustworthiness dimension of electronic records studied as part of the QADEPs project. In short, it will highlight in particular the measurement of authenticity, and will demonstrate a relevant use case developed on the basis of rich diplomatics knowledge.

Relationship between diplomatics and the study of records qualities

Diplomatics is recognized as a medieval discipline that consist to study exclusive attributes of records and manuscripts. It focuses not only on the analysis of their nature but also on their broad context as well as their creation, and their transmission conditions (Duranti, 1998). What is studied in diplomatics in particular is the relationship between the archives (content, container, and context) and the fact/act with which those archives are dealing. Digital diplomatics has the same interest except that its scope is slanted towards an electronic environment and more specifically towards digital documents or those which were converted into a digital object from a paper document.

The main goal of diplomatics and digital diplomatics is to examine the intrinsic and extrinsic characters of records in order to demonstrate their authenticity. In our study, authenticity, reliability and others internal and contextual records attributes will be studied as important characteristics that should be considered when establishing the quality levels of those records. Thus, from this perspective, we will take advantage of this old discipline to conduct an emergent and recent research which is focusing on defining and measuring qualities of digital records.

QADEPs Project: Main lines

This section presents the main lines of the study. The context and background will be presented first to allow a full understanding of the current issues linked to the assessment of the quality of electronic documents and therefore the need of a practical implementation of our project. Then, the three objectives of this research will be specified to give an insight on the guiding principles that were followed throughout the project.

Context and Background

Whether in public or private companies, any good activity process contains one or more stages dedicated to the control of its quality, service and the product that results of this process. This phase is also considered as capital, as shown by two series of standards: ISO 9000 and ISO 14000. The International Organization for Standardization published in September 2011 a third series of ISO 30000, on records management activities. This demonstrates the importance of integrating quality principles in the implementation of information governance processes.

Documents produced or received by each employee as part of his/her duties at each stage of the activity processes should be well documented. Therefore, it is necessary firstly to identify and define clearly and systematically the whole lifecycle of each activity and its related processes, and secondly to define the quality criteria that should be applied for selecting relevant information derived from all these institutional processes and activities.

The application of quality criteria will enable an adequate reduction of the mass of documents to scan, since only a small volume of them contains the key information that needs to be saved and stored. Digitizing services will thus greatly be facilitated and the institutions will have more savings opportunities. This offers a pragmatic and methodological way to ensure the setting and control of the quality of digital information. The organization will therefore significantly increase the excellence of its information service. These arguments show the relevance of this project in the current context. Providing a conceptual framework and operational measuring instruments correspond to the process of the information quality assessment by which an institution would ensure the rationality of its information, record and data management practices.

Objectives

This project has three objectives:

1. Identify and define key dimensions of the quality of electronic and public documents and their related indicators;
2. Identify the variables and develop the appropriate and operational tools for assessing the quality of public and electronic documents
3. Test the applicability of the measurement of each variable on real cases to verify its relevance, feasibility and its automation.

The realization of these objectives requires an exploratory approach which will be exposed in the next section.

Methodology

This section is structured in two parts: the first one describes the conceptual phase of the study, which refers essentially to the definition of the conceptual framework. The findings of

this particular phase have played a crucial role as they influenced the overall structure of the study. The second section highlights the main stages of the development of the empirical phase in which we developed the measuring process.

Conceptual phase: Definition of a conceptual framework

The conceptual framework describes the dimensions and key indicators of electronic records quality in the context of public institutions. Through an extensive literature review, the goal was to provide a consistent model based on reliable resources (including international standards), to give a clear representation of what is encompassed by the notion of quality for electronic and public records. Developing concepts definitely helps reaching an adequate understanding of various situations but it does not solve or control them. For this, we need to develop more operationalization.

In order to manage the flow of electronic documents and to preserve their legal value, various projects (InterPARES, KEEP Project, etc.) and a number of standards (ICA-Req, ISO 14'641, ISO 14'721, Moreq, 2010) have been created and carried out since the 90s. However, none of these projects or texts has suggested a suitable solution to ensure the quality of the documents (Makhlouf Shabou, 2011).

The present study takes advantage of the results of our recent research (Makhlouf Shabou, 2011) and adapts them to the study of the qualities of electronic records. To do so, an in-depth literature review was conducted focusing on the international standards related to information processing as in the ISO 15489, 23081, 26324, 30300, 30301 or the Afnor NF Z 42-013 standard. Those texts are valuable resources since even if they do not always specify the processing of the quality of records, they describe processes, acknowledged management practices that allow to ensure a high level of quality. On this basis, it was then possible to define the variables in order to confirm if the given recommendations in those standards have been applied or not.

Amongst the other essential references, useful to the development of our analyzing tool, there was the practices of a number of institutions as the National Archives of various countries (especially Switzerland, Canada, Germany and Australia). The practices of the latest institutions provide a point of comparison with the standards by describing how document processing is conducted under a perspective of long term archiving. Indeed, most of the standards considered here assume rather a records management perspective (current and intermediary records) and once the lifecycle period of the documents ends the priorities in terms of archiving also evolve. For example, the use of encryption guarantees security during the life of the document, nevertheless, it becomes a significant risk when it comes to long-term archiving. Most institutions, like Library and Archives Canada, refuse to take responsibility for such objects¹.

We explored also the scientific literature and the results of certain research groups like InterPARES. The issues addressed were often very specific and therefore, it required an important effort of adaptation to draw operational variables. The bibliography at the end of this article will show the various references used.

¹ <http://www.collectionscanada.gc.ca/gouvernement/produits-services/007002-3015-f.html>

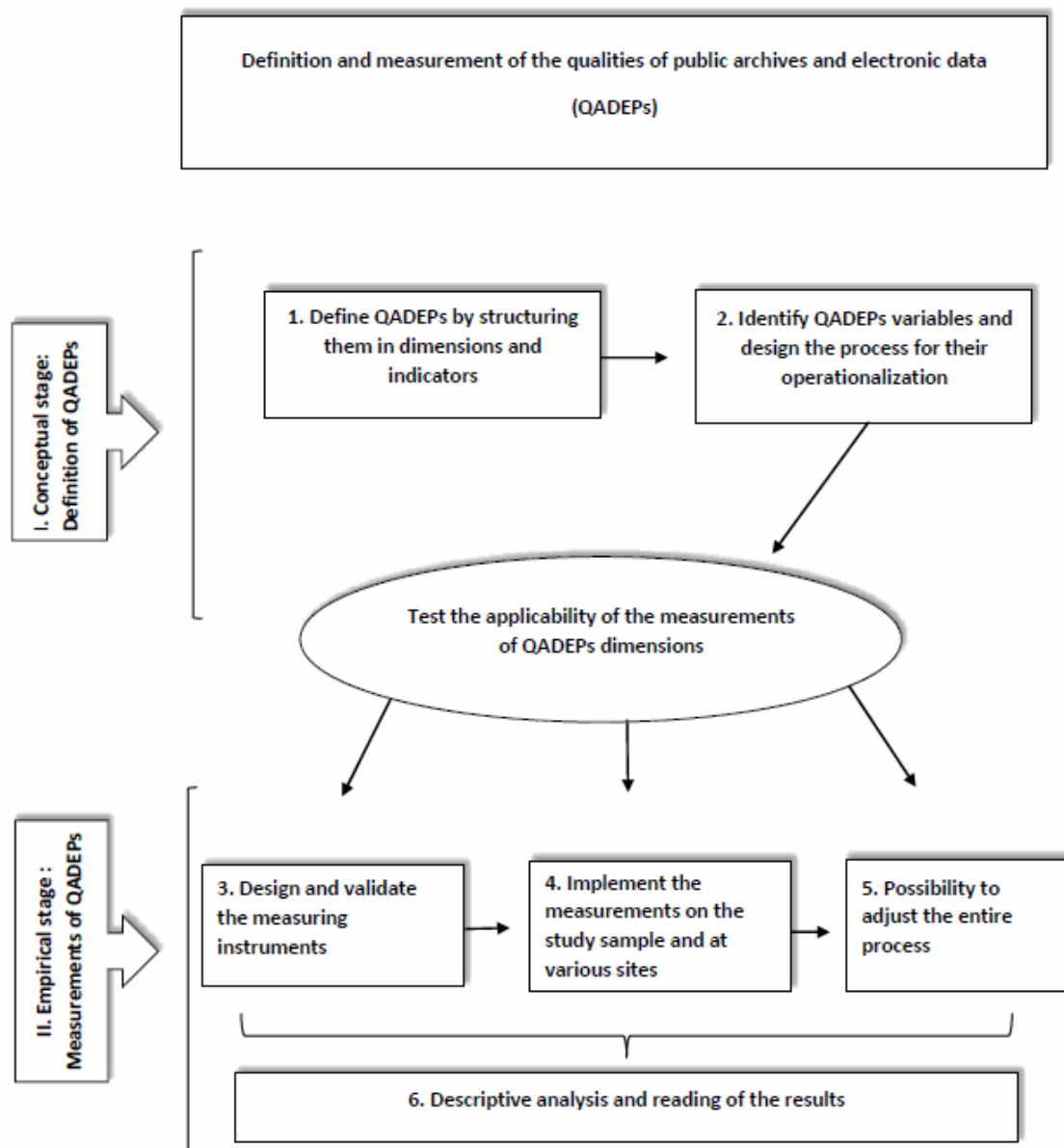


Figure 1 - Methodology

By combining those various references, it allowed us to develop a reliable tool based on stable and recognized foundations (international standards), but also close to the institution practices and in compliance with existing research.

Empirical Phase: Development of the measurement process

Fortin (Fortin, 2006, p. 274) defines the operationalization, mainly, by a four-step process: 1) the development of the theoretical definition of the concept, 2) determining the dimensions of the concept, 3) the choice of empirical indicators, and 4) selection or development of the measuring instrument. Drawing on this approach and based on the definition of intrinsic and contextual dimensions of QADEPs in terms of indicators, we tested the operationalization of the model established upon the measurement of a specific set of variables on a reduced sample of public and digital records.

The testing of the applicability of the QADEPs measurements was carried out in three steps:

1. designing of the QADEPs measuring instruments,
2. validation of instruments with different partners,
3. implementation of the QADEPs measures on a sample.

For testing the identified metrics of QADEPs project, we selected three institutions: the Swiss Federal Archives, the Archives of the State of Wallis and the Archives of the State of Geneva. These tests were primarily intended to check:

- the relevance : do these metrics reflect a high level of quality for records and data?
- the applicability of the variables : are they ambiguous ? Is it easy to make a choice amongst the different quality levels?
- the automation possibilities: would it be possible for a computer system to perform an automated quality measurement?

Once the conceptual framework was defined, the next goal was to develop variables to assess on measuring scales the dimensions and indicators defined above. Four sub-objectives have guided the development and the verification of these measurements:

1. to enable automated calculation, which would offer an undeniable advantage for quality assessment of substantial quantities of data;
2. to limit as much as possible subjective assessments;
3. to document the application of each measurements to facilitate their reusability and their adaptability to other contexts;
4. to confirm the applicability and relevance of the variables through tests in several institutions.

Overview of Findings

This section explains the findings resulting of the two main phases of the project being the conceptual phase and the empirical phase. As an example, we have chosen to develop the measurement of authenticity particularly in order to illustrate how we worked on the automation question and to provide an example of the application tool.

Conceptual framework

The conceptual framework seeks to define the quality of electronic records according to several dimensions, sub-dimensions and indicators. It (see figure 2) has three dimensions and eight sub-dimensions which contain fifteen indicators.

Dimensions and sub-dimensions

By dimension, we mean a particular facet of quality applied to electronic archives. For this study, we selected three of them: the Trustworthiness, the Exploitability and the Representativeness.

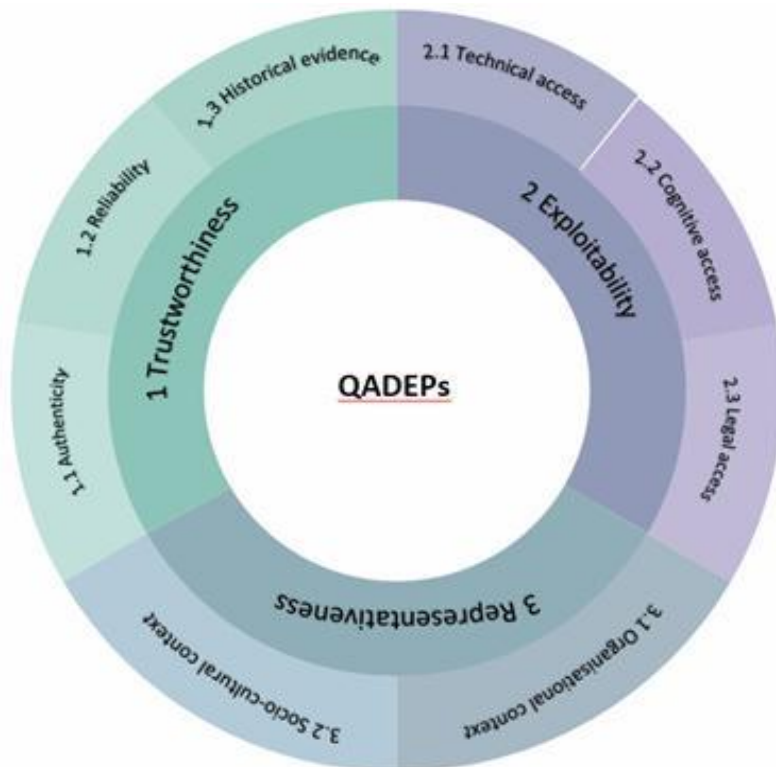


Figure 2 - Overview of the conceptual framework

Globally the main dimensions proposed in this conceptual framework could be defined as follows :

Trustworthiness

This dimension refers to the ability of a document to gain the trust of the user as the preferred supporting facts source. This quality depends on the authenticity and the reliability and the durability of these qualities over time (Makhlouf Shabou, 2011, p. 115; InterPARES 2, 2013).

Exploitability

This dimension refers to the ease of use a document thanks to its location, retrievability, diffusion and interpretability. In other words, the exploitability depends on three types of document accessibilities: 1) technical accessibility including physical material needed for reading; 2) legal accessibility including regulatory and administrative environments required for the diffusion of document; and 3) cognitive accessibility that guarantees an adequate comprehension and interpretation of document contents (ISO 15489, 2001; Makhlouf Shabou, 2011, p. 120-122).

Representativeness

This dimension refers to the capacity of the documents to provide a significant testimony of the institutional context in which they were created. This quality depends on two essential elements: the completeness of testimony and the representativeness of the socio-cultural context in which these documents were created (Makhlouf Shabou, 2011, p. 123).

Empirical phase: an overview of the main results with an illustrative case

The development of the conceptual framework allowed an accurate definition of each concept deriving from the main quality dimensions. This was essential to identify adequately the variables and the metrics that will enable the measurement of these three dimensions. The realization of this last empirical phase allowed to define eight sub-dimensions detailed by seventeen indicators, which are organized as shown in Figure 3:

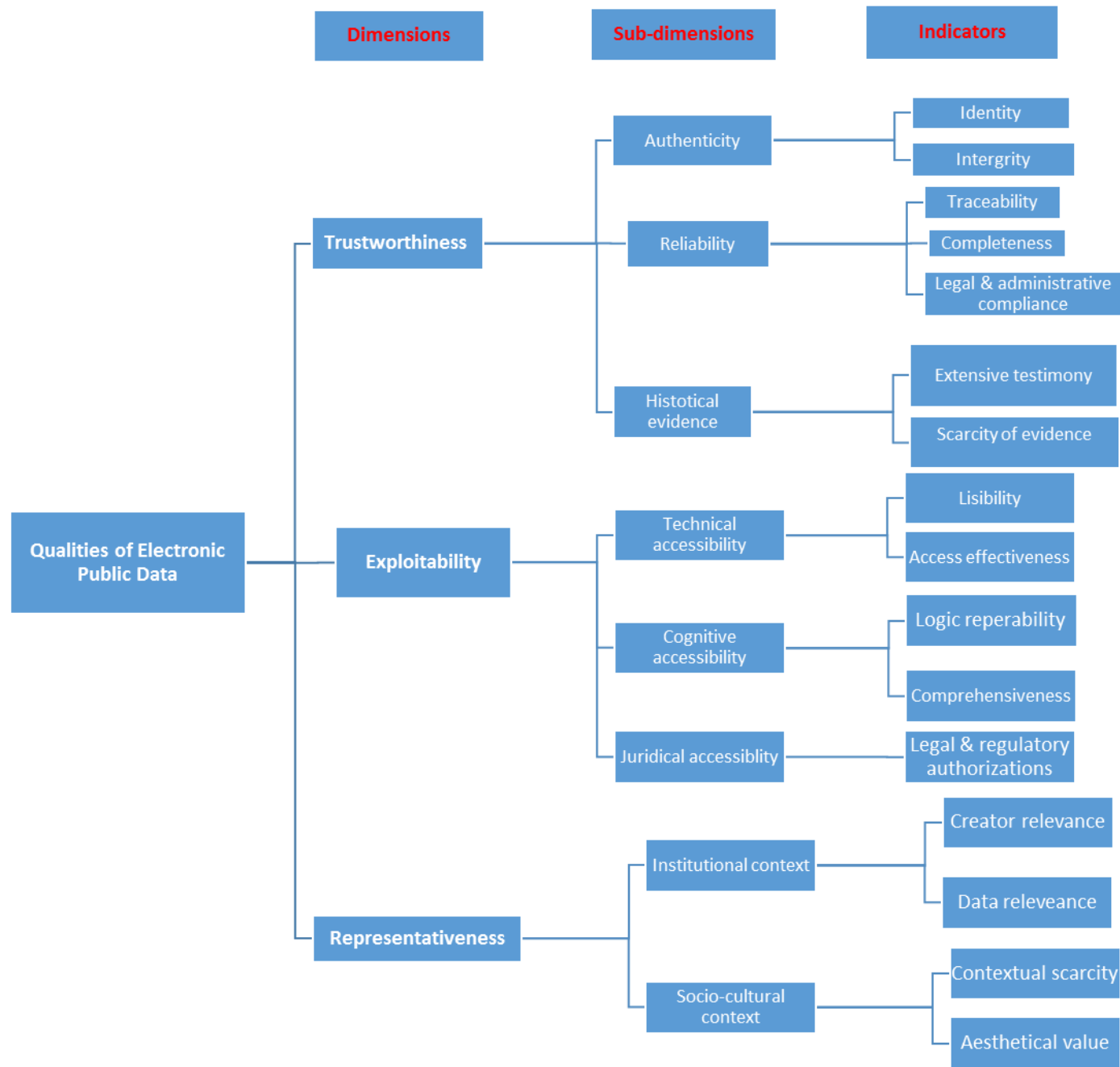


Figure 3 - Conceptual Framework: Structure of the Dimension Qualities

The different qualities studied and tested in the context of electronic public data are structured as follow (Figure 3) :

- Trustworthiness (dimension)
 - Authenticity (sub-dimensions)
 - Identity (indicator)
 - Integrity
 - Reliability
 - Traceability
 - Completeness
 - Legal and administrative compliance
 - Historical evidence
 - Extensive testimony
 - Scarcity of evidence
- Exploitability
 - Technical accessibility
 - Lisibility
 - Access effectiveness
 - Cognitive accessibility
 - Logic reparability
 - Comprehensiveness
 - Juridical accessibility
 - Legal and regulatory authorizations
- Representativeness
 - Institutional context
 - Creator relevance
 - Data relevance
 - Socio-cultural context
 - Contextual scarcity
 - Aesthetical value

We will not detail the measurement of all dimensions presented here with their different hierarchical levels in this paper, but we will focus more on the authenticity concept which represent may be the most discussed concept in the literature, both in Digital Diplomatic and also in different international standards (see section 4.2.3).

On the basis of the last conceptual level which is the indicator level, variables were developed and tested in order to verify 1) their relevance, 2) the feasibility of their empirical measurements and 3) the possibility of their automation.

Assessment method: quality levels and data scoring

The forty-six variables were identified for three quality dimensions and their related sub-dimensions and indicators. At least, for each indicator we provided a minimum of one variable or more in order to ensure its measurement. The measurement of each variable is based on a ranking system that provides specific scores reflecting the quality level of each evaluated data. The attribution of those scores was based on the verification of a set of specific conditions.

For example, to evaluate the integrity indicator of electronic data, we looked at two variables:

- 1) Access control which refers to the degrees of control, security and traceability of data access. Specifically, four quality levels were proposed:
 - Level 1: lowest one, no access control during the lifecycle
 - Level 2: existing access control during the lifecycle by a user authentication
 - Level 3: existing access control during the lifecycle by a reporting of events
 - Level 4: highest one, existing access control during the lifecycle by an algorithmic system (checksums)

- 2) Fixity of electronic data which is based on the stability of its carrier and its format. For measuring this stability, three assessment levels were proposed:
 - Level 1: lowest one, applied when the fixity is not controlled
 - Level 2: medium one, applied when the fixity is controlled by, at least one checksum
 - Level 3: highest one, applied when the fixity is controlled, periodically, by planned checksums.

The number of quality levels for each metric is different, it varies between two and six levels.

It depends on the complexity of each variable. The rank corresponding to quality levels obtained after each completed measurement was considered as the official score and final result of the measurement. This scoring method starts with level 1, which is the lowest one. This method does not provide a level 0, as it seems difficult to assume, categorically, that the quality of information is inexistent.

Application units

There were three main application units during the measuring process: items, series and fonds. These units have been used according to the following definitions:

- 1) **Items**: “a digital objects that represent a collection of data with defined boundaries that is treated as a single entity” (Pearce-Moses, 2014)².
- 2) **Series**: “group of similar records that are arranged according to a filing system and that are related as the result of being created, received, or used in the same activity; a file group; a record series” (Pearce-Moses, 2014)³.
- 3) **Fonds**: “The entire body of records of an organization, family, or individual that have been created and accumulated as the result of an organic process reflecting the functions of the creator” (Pearce-Moses, 2014)⁴.

The series and fonds were considered, especially, in the measurement of the totality of the seven variables deriving from the Representativeness dimension, for which a

² Definition available at <http://www2.archivists.org/glossary/terms/o/object>

³ Definition available at <http://www2.archivists.org/glossary/terms/s/series>

⁴ Definition available at <http://www2.archivists.org/glossary/terms/f/fonds>

large set of data was needed. These two units were also applied on the measurement of the five variables related to the sub-dimension of Historical evidence into the Trustworthiness dimension (see Fig.3). For the rest of the variables, we applied the items as measurement unit.

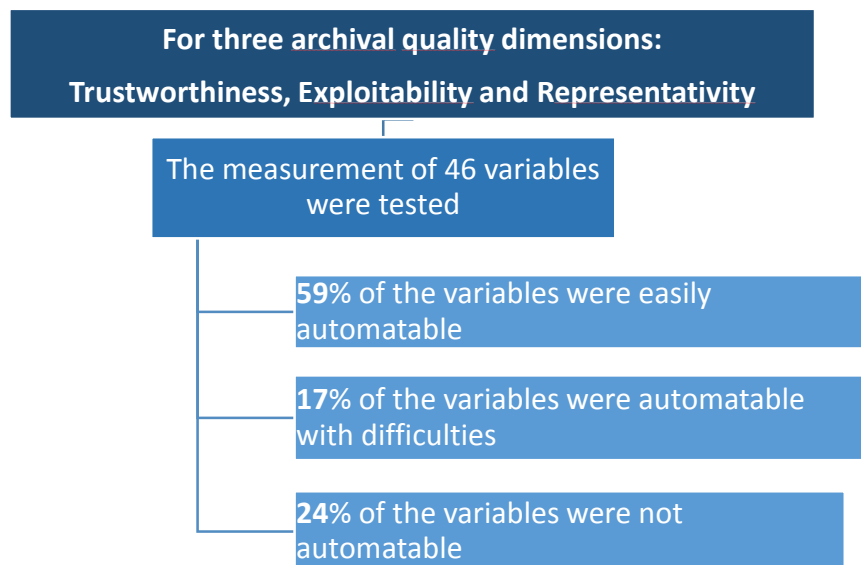


Figure 4 -Measurement of Trustworthiness, Exploitability and Representativeness

Each variable was measured by a specific metric. As shown in figure 4, 59% of metrics were easily measured and could completely be automated. And also, 17% of metrics were computable, but would require the creation of specific metadata which would be expensive. However, 24% of the metrics were not computable due to the subjectivity of some criteria on which the assessment is based, since these are founded on manual and human evaluation. This reduces the systematic reproducibility of those metrics.

A focus on authenticity and its measurement

As mentioned above, authenticity represents an important sub-dimension in the Trustworthiness dimension. Authentic records are the ones that can prove they are what they are supposed to be, that they are really produced or received by the person who claims to have produced or received them. (ISO 15489, 2001). At least, record authenticity is based on multiple types of matching metadata: the matching between its declared nature and its real nature, the matching of the declared producer/receiver and the real producer/receiver and finally the matching between the declared creation/reception date and the real creation/reception date. The efficient assessment of those previous conditions seems to be difficult, nevertheless it ought to be possible if we try to examine some observable aspects related to records identity and integrity.

The identity is "The whole of the characteristics of a document or a record that uniquely identify it and distinguish it from any other document or record" (InterPARES 2, 2013). The measurement of records identity was based, according to our study, on the examination of 10 variables: 1) existing identifier, 2) existing title, 3) name format of records file, 4) history/journal of previous identifiers for a record, 5) history/journal of previous records title, 6) history/journal of previous file names, 7) knowledge of

creator, 8) information about the file format, 9) information about timestamp, 10) information about file size.

The integrity is “property of safeguarding the accuracy and completeness of assets. Integrity demonstrates that the record is complete and has been unaltered. It is necessary that a record be protected against unauthorized alteration. Records management policies and procedures should specify what additions or annotations may be made to the record after it is created, under what circumstances additions or annotations may be authorized, and who is authorized to make them. Any authorized annotation, addition or deletion to a record should be explicitly indicated and traceable” (ISO 15489:2001; ISO 13335-1:2004)⁵. The measurement of records integrity, was based, according to our study, on the examination of 3 variables: 1) the rigor of the files and records access control, 2) fixity of the byte string, 3) the recordness of records in a specific system.

The measurement of authenticity, as one of eight sub-dimensions developed in our study, represents 13 variables on 46 which cover almost 30% of the total number of variables. This demonstrates the importance of the place of authenticity in the electronic documents qualities assessment. While authenticity is not an exclusive interest to electronic documents but rather a quality that has been extensively studied in the context of diplomatics, it should definitely be recognized here as the real root of this research field. As shown in Annex 1, many sources proposes definition to this concept. They globally propose a convergent definitions. However, they do not propose an accurate methods to evaluate this quality. They also do not specify the variables that could be measured this evaluation. Nevertheless, diplomatics sources, specially the InterPares researches dad briefly discussed the concept of authenticity assessment (table in Annex 1) without a deep description of its methods.

Authenticity Measurement: possibility of automation

As mentioned above, once the relevance and the feasibility of the measure of authenticity was tested and validated, our interest was directed to verifying the automation of those measures.

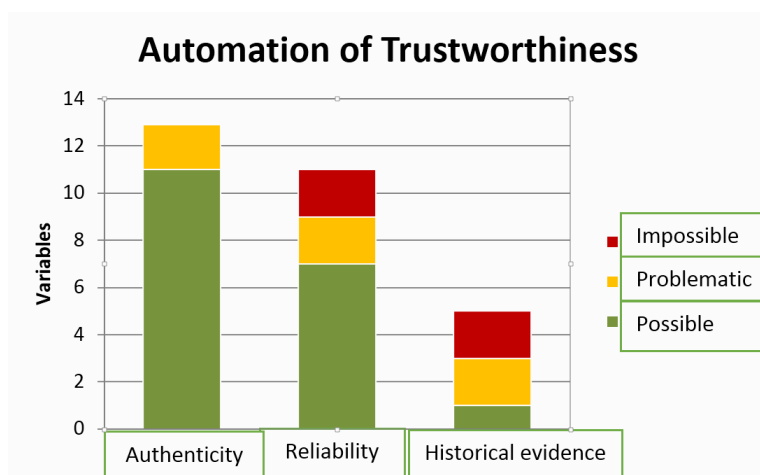


Figure 5 - Automation of Trustworthiness

⁵ Available at <http://www.ocio.gov.nl.ca/ocio/im/glossary.html#Integrity>.

The measurement of variables related to the sub-dimension of authenticity were the most feasible. The testing of their automation was the easiest one, because the metrics are based on the intrinsic qualities which generally are given in the basic metadata made available by operating systems used for the basic processing of information.

Authenticity Measurement: the application guide

The aim of the application guide is 1) to provide a definition for the concepts that were used throughout the study, 2) to offer a complete description for each dimensions, sub-dimensions, indicators and variables used and therefore, to support the reproducibility of each measurement, 3) to justify the choice of different quality levels and to present the references on which our study was based.

B. Detailed description of variables

1. Trustworthiness

Definition: capacity of a document to gain his user confidence as being the privileged source in support of facts. This quality rests upon authenticity and reliability. (MAKHLOUF SHABOU, 2010, p. 115 ; INTERPARES 2, 2013)

1.1. Sub-dimension : Authenticity

Definition : « an authentic document is a document whose can prove that it is what it pretends to be, that it was produced or received by the person who pretends to have produced or received, and it was produced or received at the moment it pretends to have been. » (ISO 15489-1:2001, p. 7)

1.1.1. Indicator : Identity

Definition : « identity :The whole of the characteristics of a document or a record that uniquely identify it and distinguish it from any other document or record. With integrity, a component of authenticity. » (INTERPARES 2, 2013)

v1. Existence of an identifier	
1.	no existing identifier
2.	existing identifier but without standards
3.	existing and standardized identifier inside the institution
4.	existing and standardized identifier inside and outside the institution

Application note : this variable examines the existence, the uniqueness, the standardization and the durability of the identifiers.

Level 1 : is applied when the object doesn't possess an indicated identifier in its current form.

Level 2 : is applied when an identifier exists, but there is no explicit rules about his establishment. This level takes, in most cases, the URI as their durability is not always guaranteed.

Level 3 : is applied when the identifiers are defined by the rules of the institution. The unicity is guaranteed, but only inside the institution. On the other hand we could find the same identifier attributed to another object in a different institution. It should be noted that the identifiers appealed to the ISIL codes attributed by a national center generally belong to this level because the ISIL code registers only the institution. The remainder of the identifier is managed and recorded internally.

Level 4 : is applied when the identifiers are defined according to norms or standards and are recorded in external registers of the institution. These identifiers must be definitively attributed and their unicity guaranteed. Their durability must be guaranteed, even if there are modifications or moving. Example: *Digital object identifier (DOI)*¹, *Persistent URLs (PURL)*², etc.

Observations : the same object can have several identifiers : for example, one for the internal functioning of the institution and another external DOI type. The assessment of the variable will only consider the identifier which has the highest value.

Sources :

ORGANISATION INTERNATIONALE DE NORMALISATION. *ISO 15511 : Information et documentation - Identifiant international normalisé pour les bibliothèques et les organismes apparentés (ISIL)*. Genève. 2011.

¹ DOI are made of a prefix and a suffix like the following : 10.xxxxx/xxxxxx. The 10 is own to the DOI. The prefix indicate the agency which records the DOI, as for the suffix, it identifies the object it-self inside a register. The suffix can take numerous forms like an issn or any chain of characters.

²cf. PURL'S CONTRIBUTORS, 2011 : « PURLs are Web addresses or Uniform Resource Locators (URLs) that act as permanent identifiers in the face of a dynamic and changing Web infrastructure. Instead of resolving directly to Web resources (documents, data, services, people, etc.) PURLs provide a level of indirection that allows the underlying Web addresses of resources to change over time without negatively affecting systems that depend on them. This capability provides continuity of references to network resources that may migrate from machine to machine for business, social or technical reasons. »

Figure 6 – The Presentation of the application guide: sample⁶

⁶ This is just one example of the contents presented in the guide. The English version of this guide with its exhaustive section will be available soon. Contact : basma.shabou@gmail.com or basma.makhlouf-shabou@hesge.ch

As shown in the sample presented in Figure 6, the overall content of the application guide follows the same structure as the general framework observed throughout the study. It begins with a definition of the dimension (in this study being either trustworthiness, exploitability or representativeness), it continues with a definition of each sub-dimension and indicator selected, then one can find a description of the variables with their detailed application notes containing the explanation for each quality levels chosen and to conclude, the guide mentions the references on which the application notes were based. The application notes offer further information in the form of remarks or references mainly to specify certain application processes or to define a concept more accurately.

Conclusion

As mentioned above, our research intended to define different dimensions that specify the qualities of data and electronic records in the context of public institutions and their related sub-dimensions, indicators and variables. On the basis of those quality dimensions, a specific and rigorous measurement method was designed and tested.

Three main lessons should be learned from this study. First, the measurement of electronic records and data was possible and its process was also verifiable and largely computable. Second, the method proposed in this study is reproducible. It is clearly presented. This study provides a significant results which refers to a specific level of quality from the point of view of a particular aspect and cannot be understood as a general and absolute evaluation. For example we can evaluate the exploitability of a given set of data as high quality only from the point of view of specific sub dimensions. Third, the realization of this research demonstrates the multidimensional nature of the definition and measurement of records and data qualities. Many elements should be studied and considered altogether to enable adequate definition and an accurate measure of specific records or data quality. That is why we noticed also the relative nature of obtained results specifically at the sub-dimensions and indicators levels which were based on the compilation and the interdependence between many variables.

In fact, our research which is defining and measuring the quality of data and electronic records and archives promises four significant benefits. In theory, it develops a conceptual framework that defines an accurate architecture of main qualities applied on electronic and public records and data. At the methodological level, it offers a method for measuring the historical qualities already tested in a real context of public institutions. At the professional level, firstly, it allows the assessment of the relevance of archival material that has been chosen for the appraisal; furthermore, it provides professionals a measurement grid for measuring archival qualities already tested and also the guide that facilitates its application.

At the academic level, this research has provided opportunities for collaboration with another department of our school. A group of five students from the Business Information Systems Department of the Geneva Business School are developing software that should enable the automation of the metrics related to authenticity and reliability of the sub-dimensions. The results are expected for the end of 2014. At the international level, the QADEPs results are shared with the InterPARES Trust research, and further collaborations are planned to develop this research area.

This study is, to our knowledge, the first reflection conducted in the domain of archival quality assessment. At this exploratory step, we could not pretend to offer the way to measure exhaustively the quality of institutional archives. To do that, the research should be

reproduced in different contexts in order to be confirmed and reinforced. The completion and the development of those measurements are necessary for the validity of this research's conceptual framework and the relevance of its results.

Acknowledgements

This study was supported by the University of Applied sciences and Art of Western Switzerland (HES.so Genève). It has been carried out by Dr Basma Makhoulf Shabou in collaboration with Mrs Mellifluo, Mr Rey, Dr. Schneider and Mr. Boder from HEG-Geneva; Mr. Fluckiger and Ducry from Archives of State of Geneva; Mr. Dubois from Archives of State of Wallis; Mr. Kellerhals from Swiss Federal Archives and Dr. Christophe Graf from the University of Bern; and Dr Wildi from Docuteam Sàrl.

Bibliography

The Federal Assembly of the Swiss Confederation, (2011). Federal Act on Data Protection (FADP) of 19 June 1992 (Loi fédérale du 19 juin 1992 sur la protection des données (LPD)), available at http://www.admin.ch/ch/e/rs/235_1/

French Association for Standardization (2009). *NF Z 42-013 - Electronic archival storage - Specifications relative to the design and operation of information processing systems in view of ensuring the storage and integrity of the recordings stored in these systems.* Paris

Banat-Berger, Françoise et Huc Claude (2011). « Section 11 - Gestion d'un projet d'archivage électronique » (Management of an electronic archiving project) in *PIAF 2.0 - Espace formation, Module 7. 22 novembre 2011.* <http://www.piaf-archives.org/espace-formation/mod/resource/view.php?id=101>

Library and Archives Canada (2006). Guidelines For Records Created Under a Public Key Infrastructure Using Encryption And Digital Signatures. <http://www.bac-lac.gc.ca/eng/services/government-information-resources/disposition/records-appraisal-disposition-program/Pages/guidelines-records-public-key-infrastructure-encryption-digital-signatures.aspx>

Boles, Frank et al. (1991). *Archival appraisal.* Neal-Schuman Publishers, New York.

Centre de coordination pour l'archivage à long terme de documents électroniques (2013). SIP-Val: guide d'application (Coordination center for long-term archiving of electronic documents (2013). SIP -Val: application guide), available at : <http://kost-ceco.ch/cms/download.php?abacbd41cb554aa9a1d3c154c2675b20>

Centre de coordination pour l'archivage à long terme de documents électroniques, [s.d]. *Catalogue des formats de données d'archivage* (Coordination center for long-term archiving of electronic documents (2013). SIP -Val: application guide), KOST-CECO, available at : <http://www.kost-ceco.ch/wiki/whelp/Cfa/index.php>

Confédération Suisse (2012). *Bibliothèque nationale suisse BN : Centre ISIL Suisse*. In : *Département fédéral de l'intérieur, 5 juillet 2012*, available at : http://www.nb.admin.ch/nb_professionnel/01540/index.html?lang=fr

Conférence des recteurs et des principaux des universités du Québec (2009). *Mesures transitoires et bonnes pratiques de gestion des documents numériques*, available at : http://www.crepuq.qc.ca/IMG/pdf/mesures_transitoires_bonnes_pratiques_GDN-10juin.pdf

Consultative committee for space data systems, 2011. "CCSDS 652.0-M-1: Audit and certification of trustworthy digital repositories" in *Magenta book*, Washington, DC, available at : <http://public.ccsds.org/publications/archive/652x0m1.pdf>

Cook, Terry (2006). "Macroappraisal in Theory and Practice : Origins, Characteristics, and Implementation in Canada, 1950-2000", *Archival Sciences*, Vol. printemps No. 5, pp. 101-161.

Cox, Richard J (2011). "Appraisal and the future of archives in the digital era" in : Hill, Jenny (ed.), *The future of archives and recordkeeping: a reader*, Facet, London, pp. 213-237.
DLM Forum Foundation (2011). *Version 1.1 : MoReq2010®: modular requirements for records systems - volume 1: core services & plug-in modules*, available at : http://moreq2010.eu/pdf/moreq2010_vol1_v1_1_en.pdf

Drümmer, Olaf, Oettler, Alexandra and Seggern, Dietrich von (2008). *PDF/A : l'essentiel : archivage numérique à long terme avec PDF*, Association for Digital Document Standards, Berlin.

Dubourg, Christian and Thomas, Michel (2010). "Archivage électronique et conformité réglementaire" in *Livre blanc*, Ever Team, available at : <http://www.bivi.fonctions-documentaires.afnor.org/livres-blancs/livre-blanc-ever-team-archivage-electronique-et-conformite-reglementaire>

Duranti, Luciana, Society of American Archivists and Association of Canadian Archivists (1998). *diplomatics: new uses for an old science*, Scarecrow, Lanham.

Eastwood, Terry (ed.) (1992). *The archival fonds: from theory to practice = Le fonds d'archives: de la théorie à la pratique*, Bureau of Canadian Archivists, Planning Committee on Descriptive Standards, Ottawa.

Fortin, Marie-Fabienne (2006), *Fondements et étapes du processus de recherche* (Foundations and steps of the research process), Chenelière Education, Montreal.

Hollingsworth, David (1995). *TC00-1003 : The Workflow Reference Model [en ligne]. Winchester. The Workflow Management Coalition*, available at :
<http://www.wfmc.org/Download-document/TC00-1003-The-Workflow-Reference-Model.html>.

International Council on Archives (1995). *ISAAR(CPF) : International Archival Authority Record for Corporate Bodies, Persons and Families*, Ottawa, available at :
[http://www.icacds.org.uk/eng/ISAAR\(CPF\).pdf](http://www.icacds.org.uk/eng/ISAAR(CPF).pdf).

International Council on Archives (2000). *ISAD(G): General International Standard Archival Description - Second edition*, Ottawa, available at :
<http://www.ica.org/10207/standards/isadg-general-international-standard-archival-description-second-edition.html>

International Council on Archives (2007). *ISDF: International Standard for Describing Functions - Première édition*, Dresde, available at :
<http://www.ica.org/10226/normes/isdf-norme-internationale-pour-la-description-des-fonctions.html>

International Organization for Standardization (2001). *Information and documentation – Records management -- Part 1: General*, ISO 15489-1:2001, ISO, Geneva.

International Organization for Standardization (2004). *ISO/TR 15801 : Electronic imaging – Information stored electronically - Recommendations for trustworthiness and reliability*, ISO, Geneva.

International Organization for Standardization (2008). *ISO 26122 : Information and documentation - Work process analysis for records*, ISO, Geneva.

International Organization for Standardization (2011). *ISO 15511 : Information et documentation - Identifiant international normalisé pour les bibliothèques et les organismes apparentés (ISIL)*, ISO, Genève.

International Organization for Standardization (2012a). *ISO 13008 : Information and documentation - Digital records conversion and migration process* ISO, Geneva.

International Organization for Standardization (2012b). *ISO 23081-1 : Information and documentation -- Records management processes -- Metadata for records -- Part 1 principles*, ISO, Geneva.

International Organization for Standardization (2012c). *ISO 23081-2 : Information and documentation -- Records management processes -- Metadata for records -- Part 2 conceptual and implementation issue*, ISO, Geneva.

International Organization for Standardization (2012d). *ISO 26324 : Information et documentation - Système d'identifiant numérique d'objet*, ISO, Genève.

INTERPARES 2 (2013). "Terminology Database" in *INTERPARES Project*, available at http://www.interpares.org/ip2/ip2_terminology_db.cfm

Library of Congress (2011). *METS : vue d'ensemble et tutoriel*, available at : http://www.loc.gov/standards/mets/METSOverview.v2_fr.html

Library of Congress (2012a). *EAD: Encoded Archival Description: Version 2002 Official Site*, available at : <http://www.loc.gov/ead/>

Library of Congress (2012b). *Metadata Encoding and Transmission Standard (METS) Official Web Site*, available at : <http://www.loc.gov/standards/mets/>

Makhlouf Shabou, Basma (2011). *Étude sur la définition et la mesure des qualités des archives définitives issues d'une évaluation*. Ph.D. thesis, Montreal, Université de Montréal - École de bibliothéconomie et des sciences de l'information, available at : https://papyrus.bib.umontreal.ca/jspui/bitstream/1866/4955/5/Makhouf-Shabou_Basma_2011_these.pdf

Maltais, Corinne and Lemieux, Nadine (2009). "1001 stratégies pour nommer ses fichiers et dossiers numériques", *Archives au présent : le magazine en ligne des archivistes du Québec*, 16 mai 2009, available at :

http://www.archivistes.qc.ca/spip.php?page=article-24&id_article=278

Menne-Haritz (1994). "The Principle of Provenance : Report from the First Stockholm Conference on Archival Theory and the Principle of Provenance, 2-3 September 1993" in *The Principle of Provenance*, Swedish National Archives, Stockholm, pp. 103-131.

Microsoft (2013). "Journaux des événements dans Windows SharePoint Services", *MSDN Library*, available at : [http://msdn.microsoft.com/fr-fr/library/bb861918\(v=office.12\).aspx](http://msdn.microsoft.com/fr-fr/library/bb861918(v=office.12).aspx)

Office of the chief information officer (2012). *Information Management and Information Protection Glossary of Terms*, Office of the Chief Information Officer, available at : http://www.ocio.gov.nl.ca/ocio/im/glossary.html#Information_Asset

Paskin, Norman (2010). *Digital Object Identifier (DOI®) System*, Third edition, CRC Press, Boca Raton Fla., available at : <http://www.doi.org/overview/080625DOI-ELIS-Paskin.pdf>

PCR DIGITAL FORENSICS (2013). "Glossary of Computer Forensics Terms", *PCR Digital Forensics*, available at : http://www.pcrforensics.com/index.php?option=com_glossary&Itemid=132

Pearce-Moses, Richard (2014). "A Glossary of Archival and Records Terminology" in *The Society of American Archivists*, available at : <http://www2.archivists.org/glossary>

Portail International Archivistique Francophone (2009). *Les critères d'authenticité d'une archive électronique*, available at : <http://www.piaf-archives.org/content/les-criteres-dauthenticite-dune-archive-electronique>

PREMIS Editorial Committee (2012). *PREMIS - Data Dictionary for Preservation Metadata – version 2.2*, available at : <http://www.loc.gov/standards/premis/v2/premis-2-2.pdf>

PURL'S CONTRIBUTORS (2011). "Persistent URLs" in *Persistent URLs*, available at : <http://www.purlz.org/>

République du Canton de Genève, Archives d'Etat (2011). *Formats de fichiers adaptés à*

l'archivage électronique à moyen et long terme, 19 octobre 2011, République et Canton de Genève, Genève, available at :

http://etat.geneve.ch/dt/SilverpeasWebFileServer/20111019_formats_archivage_etatge.pdf?ComponentId=kmelia66&SourceFile=1319206910719.pdf&MimeType=application/pdf&Directory=Attachment/Images/

Rouchon, Olivier (2009). "PAC, la plateforme d'archivage pérenne de documents électroniques au CINES" in *Archivage et stockage pérennes : enjeux et réalisations*, Hermes science, Paris, pp. 21-40.

Rudersdorf, Amy (2012). "Digital preservation ingest can be a « CINCH »", *Library Hi Tech*. 31 août 2012, Vol. 30 No 3, pp. 449-456. DOI 10.1108/07378831211266591.

Samuels, Helen Willa (1992). *Varsity letters documenting modern colleges and universities*, Scarecrow Press, Chicago.

Stanescu, Andreas (2005). "Assessing the durability of formats in a digital preservation environment: The INFORM methodology", *OCLC Systems & Services*, Vol. 21 No 1, pp. 61-81. DOI 10.1108/10650750510578163.

State University Library Services (2013). *PREMIS in METS Toolbox*, available at : <http://pim.fcla.edu/>

The Consultative Committee for space data systems (2012). *ccsds 650.0-M-2 : Reference model for an open archival information system (OAIS): recommended practice*, The Consultative Commit, Washington, DC., available at : <http://public.ccsds.org/publications/archive/650x0m2.pdf>

The Dublin Core Metadata Initiative (2012). *Using Dublin Core - The Elements*, available at : <http://dublincore.org/documents/usageguide/elements.shtml>

The Dublin Core Metadata Initiative (2013). "XML Schemas to support the Guidelines for implementing Dublin Core Metadata in XML", in *DCMI Home: Dublin Core Metadata Initiative (DCMI)*, available at : <http://dublincore.org/schemas/xmls/>

The National Archives [s.d]. "Download DROID: file format identification tool" in *The National*

Archives, available at :

<http://www.nationalarchives.gov.uk/information-management/projects-and-work/droid.htm>

The National Archives [s.d]. *PRONOM* in *The National Archives*, available at :

<http://www.nationalarchives.gov.uk/PRONOM/Default.aspx>

Todd, Malcolm (2009). "09-02 : File formats for preservation" in *Technology Watch Report. S.I. Digital Preservation Coalition*, available at :

http://www.dpconline.org/component/docman/doc_download/375-file-formats-for-preservation

Workflow Management Coalition (1998a). *WFMC-TC-1015 : Audit Data Specification*, The Workflow Management Coalition, Winchester, available at :

<http://www.wfmc.org/Download-document/WFMC-TC-1015-Ver-1.1-Audit-Data-Specification.html>

Workflow Management Coalition (1998b). *WFMC-TC-1019 : Workflow Security Considerations - White Paper*, The Workflow Management Coalition, Winchester, available at :

<http://www.wfmc.org/Download-document/WFMC-TC-1019-Ver-1-Workflow-Security-Considerations.html>

Zierau, Eld Maj-Britt Olmütz (2012). "A holistic approach to bit preservation", *Library Hi Tech*, 31 août 2012, Vol. 30 No 3, pp. 472-489.

Annex 1 Authenticity : Theoretical Basis

Sources	Authenticity : definition and measurement
QADEPs	Indicator : Identity
	<ul style="list-style-type: none"> Existing identifiers for default format
	<ul style="list-style-type: none"> Existing title for default format
	<ul style="list-style-type: none"> Existing naming format for default format
	<ul style="list-style-type: none"> Audit trail of existing identifiers for changes and / or changes in compliance with default format
	<ul style="list-style-type: none"> Audit trail of existing titles for changes and / or changes in compliance with default format
	<ul style="list-style-type: none"> Audit trail of existing naming formats for changes and / or changes in compliance with default format
	<ul style="list-style-type: none"> Existence of metadata informing on the producer
	<ul style="list-style-type: none"> Existence of metadata indicating the format
	<ul style="list-style-type: none"> Existence of timestamp
	<ul style="list-style-type: none"> Compare size of existing elements with size of previous elements
	Indicator : Integrity
	<ul style="list-style-type: none"> Level of access control
<ul style="list-style-type: none"> Checksums control 	
Theoretical Basis	
Diplomatics : InterPARES (2015)	<p>Authenticity <i>authenticity</i> n., The trustworthiness of a record as a record; i.e., the quality of a record that is what it purports to be and that is free from tampering or corruption. [Archives]</p> <p>assessments of authenticity n., The determination of whether a document has all the formal elements that it was supposed to present when first made or received and set aside. [Archives]</p>

	<p>Identity Definition: The whole of the characteristics of a document or a record that uniquely identify it and distinguish it from any other document or record. With integrity, a component of authenticity. »</p> <p>Source : INTERPARES, PROJECT, 2013. Terminology database. <i>InterPares Project</i> [online]. 2013. [Consulted the 14th January 2015]. Available at : http://www.interpares.org/ip3/ip3_terminology_db.cfm</p>
ISO 14641	<p>Integrity Definition : attribute of a document whose content is completed and unaltered Source : INTERNATIONAL ORGANISATION FOR STANDARIZATION, 2012. <i>Electronic archiving -- Part 1: Specifications concerning the design and the operation of an information system for electronic information preservation</i>. Geneva : ISO, 24th January 2012. ISO 14641-1</p>
OAIS : ISO 14721	<p>Fixity information : This information provides the Data Integrity checks or Validation/Verification keys used to ensure that the particular Content Information object has not been altered in an undocumented manner. Fixity information includes special encoding and error detection schemes that are specific to instances of Content Objects. Fixity Information does not include the integrity preserving mechanisms provided by the OAIS underlying services, error protection supplied by the media and device drivers used by Archival Storage. The Fixity Information may specify minimum quality of service requirements for these mechanisms.</p> <p>Source : INTERNATIONAL ORGANISATION FOR STANDARIZATION, 2012. <i>Space data and information transfer systems -- Open archival information system (OAIS) -- Reference model</i>. Geneva : ISO, 21st August 2012. ISO 14721</p>
ISO 15489 (2001)	<p>Authenticité Definition : Quality of a document can be proven to be what it claims to be, it was effectively produced or received by the person who claims to have produced or received, and was produced or received by the time he claims to have been. (<i>« un document authentique est un document dont on peut prouver qu'il est bien ce qu'il prétend être, qu'il a été effectivement produit ou reçu par la personne qui prétend l'avoir produit ou reçu, et qu'il a été produit ou reçu au moment où il prétend l'avoir été. »</i>)</p> <p>Integrity Definition : « Integrity is the property of safeguarding the accuracy and completeness of assets. Integrity demonstrates that the record is complete and has been unaltered. It is necessary that a record be protected against unauthorized alteration. Records management policies and procedures should specify what additions or annotations may be made to the record after it is created, under what circumstances additions or annotations may be authorized, and who is authorized to make them. Any authorized annotation, addition or deletion to a record should be explicitly indicated and traceable »</p> <p>Source : INTERNATIONAL ORGANISATION FOR STANDARIZATION, 2001 <i>Information and documentation -- Records management -- Part 1: General</i>. Geneva : ISO, 2001. ISO 15489.</p>
ICA-Req	<p>Authenticity Definition : the record can be proven to be what it purports to be, to have been</p>

	<p>created or sent by the person that created or sent it, and to have been created or sent at the time it is purported to have occurred</p> <p>Source : INTERNATIONAL COUNCIL ON ARCHIVES, 2008. <i>Principles and functional requirements for records in electronic office environments.</i></p>
<p>MoReq Specifications</p>	<p>Authenticity, principle of –</p> <p>Definition : Along with integrity, reliability and usability, one of the central characteristics of a record according to ISO 15489.</p> <p>An authentic record is one that can be proven to be what it purports to be.</p> <p>Source : DLM FORUM FOUNDATION, 2011. Version 1.1 : MoReq2010®: modular requirements for records systems - volume 1 : core services & plug-in modules [en ligne]. S.I. [Consulted the 14th January 2015]. Available at : http://moreq2010.eu/pdf/moreq2010_vol1_v1_1_en.pdf.</p>