

ICA-Req as a tool for audit software functional requirements: evaluating data and metadata to ensure long-term preservation

Lourdes FUENTES HASHIMOTO

Archives départementales de la Seine Saint-Denis
18 avenue du Président Salvador Allende 93 000 Bobigny, France
E-Mail: lfuenteshashimoto@cg93.fr

ABSTRACT

Usually, we create data without considering its quality (in terms of authenticity/reliability), its life cycle and its potential preservation. Are we sure that our systems are creating reliable data? How long do we want to use the data we create? Have we established rules to manage its life cycle? Those are questions we rarely think about because we are more concerned about the immediate use of data in a specific area or business. Besides, systems are rarely designed for this purpose. Preservation commonly becomes an issue after we have generated data and we realize that we need to plan its preservation. But data and metadata long-term preservation highly depends on their quality and quality needs to be prepared, ideally before data creation. Since IT systems are a determining factor, we should evaluate every system that produces the data we intend to preserve. ICA-Req (ISO 16175) provides a set of functionality requirements to consider when specifying a new system or auditing an existing one. Firstly, this presentation will explain how ICA-Req can be used as a tool for audit system's functional requirements to ensure long-term preservation. Secondly, it will present an example of its implementation at the Archives of Seine Saint-Denis (French local administration) and the results in terms of methodology and practice.

Keywords: international standards, records management, information systems, data, metadata, audit, archives, digital preservation.

INTRODUCTION

Data and metadata long-term preservation highly depends on their quality. The best way to be sure of their quality is to apply specific rules at the very moment we are creating them. Those rules should be the result of a previous analysis and organization in order to respect work processes. Quality is related to security (material security and access control) but also to the idea of “integrity” (which refers also to “authenticity” and “readability”). Digital preservation implies the preservation of data integrity, that is to say we are preserving data authenticity and readability over time. With this aim in view, we have created sophisticated systems that we call “digital repositories” or “digital preservation systems”. Nevertheless, it is absurd to dump data into these systems if we are not sure of its quality (in terms of integrity). That is the reason why we need to be particularly demanding when we are implementing new systems that will create the data we intend to preserve. These systems have a key role in digital preservation. Since there is undoubtedly a link between data creation and its preservation, the evaluation of the systems we use to create data should be the basis of any preservation policy. Thus, digital preservation is rather preventive than curative.

How should we evaluate our systems? There is an international standard, ICA-Req [1], which has been designed to improve electronic records (data and metadata) management practices and to contribute to establish best practices for planning long-term preservation and access. It has been developed in 2008 by the National Archives of Australia in conjunction with the International Council on Archives (ICA) and members of the Australasian Digital Recordkeeping Initiative. The French translation was launched in December 2010. A few archival institutions in France, such as the National Archives (Digital Archives

office in Fontainebleau), the Ministry of Culture, the Ministry of the Interior and the Archives of Seine Saint-Denis have started using this standard in 2011 and developing tools based on ICA-Req's guidelines. ICA-Req consists of three modules: the first module presents ICA-Req's generic principles, the second one focuses on functional requirements for ERMS (Electronic Records Management Systems), the third one describes the most important functional requirements for the systems that we use to produce data. The three modules can be used separately but it is also possible to combine them so that they become the basis of a digital archival policy. The main aims of the project were to produce globally harmonised generic principles and functional requirements for software used to create and manage electronic data and to provide guidelines to meet the needs of the international archival community.

ICA-Req's module 3 sets up the most important requirements to identify work processes and to evaluate both software and data that is created. The main requirements can be used to establish a list to consider when specifying a new system or auditing an existing one. ICA has also published a workbook for archivists and records management to explain how to use the module 3, it will be soon available in French [2]. The Archives of Seine Saint-Denis have implemented ICA Req's module 3. We will explain how we have built an audit tool and describe the results in terms of methodology.

ICA-REQ'S MODULE 3

Digital preservation starts when we create data: we must consider it as a part of each work process and not as an independent issue. The electronic information generated by a system increasingly serves as the only evidence of a work process, even if the system is not being designed for this purpose. Without reliable electronic information, organizations are exposed to different kinds of risk (inefficiency and legal issues in particular). Scientific and technical sectors are also concerned because it is fundamental to produce reliable data, to maintain its integrity over time and to preserve it for further scientific work, including historical research. How to evaluate a system? Which are the criteria to decide either a system is or not preserving data integrity? How to improve data and metadata management? ICA-Req's module 3 provides a set of functionality requirements that can be used as an audit tool in order to evaluate systems.

Guidelines

ICA-Req will help organizations to understand processes and requirements for identifying and managing electronic records and to develop requirements to be included in a design specification when building, upgrading or purchasing software. It provides a set of functionalities with recommended levels of obligation. However, ICA-Req does not provide a complete specification but rather outlines a number of key requirements. It does not focus on general system management. Therefore, design requirements such as usability, searching, system administration and performance are beyond the scope of this standard. It focuses mostly on the implementation of recordkeeping functionalities including the management of the life cycle of electronic data and the evaluation of metadata quality. It also highlights the importance of analyzing work processes since recordkeeping integration is challenged in an environment where processes are poorly defined. Thus, data and metadata quality depends on a prior analysis of work processes 1) to define which data is required to be kept and 2) to improve data management identifying strategic functionalities and user's profiles. ICA-Req describes what an appropriately managed electronic record is, it will:

- Provide an information resource that can be trusted and that can be used to demonstrate and account for organizational activities.
- Aid transparent, informed and quality decision making and planning.
- Enable consistency, continuity and efficiency.

The term "system" refers to an IT system. This is in contrast with records management principles that usually include people, policies and procedures. Nevertheless, it is important to consider these aspects to ensure that data is appropriately managed. There are other specific terms which have been defined in a glossary to avoid misunderstanding.

Functional requirements

ICA-Req describes 125 functionalities that can be used to develop a software design specification for a system with records management functionalities or to review, assess and audit existing systems. Functional requirements are divided into four different sections according to key records management concepts:

- Creating electronic records in context.
- Managing and maintaining electronic records. Records must be managed to ensure they have the following characteristics: authenticity, reliability, integrity, usability.
- Supporting import, export and interoperability.
- Retaining and disposing electronic records as required.

Metadata is essential to the appropriate management of electronic records. For this reason, requirements for electronic records metadata are incorporated into the four sections. There is also a specific category with records management metadata.

There are three different levels which indicate the relative importance of each of the functional requirements:

- “Must”- requirements that use “must” are absolutely necessary for compliance with the specification.
- “Should” – requirements that use “should” may be ignored if a valid reason exists but the full implications of ignoring must be understood before choosing a different course.
- “May” – requirements that use “may” are optional.

It is acknowledged that, depending on the model chosen, the requirements do not have to be met purely within the system but may be met through the use of additional tools. Therefore, integration with other systems must be taken into account.

IMPLEMENTING ICA-REQ’S MODULE 3: A CASE STUDY OF THE FRENCH LOCAL ADMINISTRATION

ICA-Req does not prescribe any specific implementation approach. It provides a generic set of requirements with recommended levels of obligation. Consequently, an adaptation is necessary to analyze systems within a specific context. Moreover, additional analysis is required to understand processes, risks, legal issues, and specific needs in each environment/organization. We will present here a methodology drawn up by the Archives of Seine Saint-Denis.

Context

The French public administration has taken a series of decisions to develop e-governance and e-services over the last ten years. Thus, many procedures have been automated. In parallel, the context of production of strategic documents has evolved with the development of information technologies. The electronic information that is generated is the only evidence of the administration’s activities. This information is considered as public records and must be preserved in accordance with archival rules and legislation. We are facing new challenges in terms of IT security, records management and digital preservation. In this very specific context, the Archives of Seine Saint-Denis, which have the responsibility of preserving the archives produced by the local administration in the county (*département* in French) of Seine Saint-Denis, have considered to implement ICA-Req’s module 3 to audit every system which produce important/strategic data. Two main ideas are the basis of our approach. We decided to:

- Evaluate software functionalities and their ability to produce and/or manage reliable data and metadata without compromising their integrity. This is intended to prevent the transfer of corrupted data to a digital preservation system.
- Evaluate software ability to select data and metadata in order to start an archival procedure or in order to properly purge a system.

Methodology: how to build an audit tool and how to use it?

First, it was necessary to understand ICA-Req's guidelines. Then we decided to consider only the most important requirements to establish a list of functionalities that are tenable and realistic, so we made a compromise. However, it should be noted that some functionalities that were not deemed mandatory by ICA-Req were essential in the context of the Seine Saint-Denis' administration. A total of sixty requirements were selected. We tried to understand each requirement and its impact in electronic records management. It was decided to rephrase a few requirements to ensure a better understanding of this tool because the checklist would be sent to persons engaged in different professions and different backgrounds such as IT professionals, software editors, administration officials and archivists.

The next step was to adapt the checklist to the system that we intended to evaluate. The Archives of Seine Saint-Denis have evaluated two systems between March and August 2011. We will present in this paper only one case. The administration wanted to adopt an ECM system (Enterprise Content Management) to manage born digital official records. Since they are very strategic for the administration and have a strong legal value, it was important to audit the system.

The checklist became an audit tool. It has three columns:

Functionality/Requirement	Yes/No	Notes
(Example) The system must allow the manual or automatic updating of all metadata.		

Table 1: The structure of the checklist

Notes are fundamental in order to describe functionalities and procedures since the aim of the checklist is to allow archivists to understand the system. We also use this column to make recommendations for further improvement and to describe a possible risk when it is necessary. Notes were used during the evaluation of the ECM system in order to write a full report which draws attention to some specific points:

- Data integrity and security could be compromised by a lack of specific features.
- The life cycle of the electronic records has not been sufficiently taken into account.
- User profiles were not well defined.

CONCLUSION

The Archives of Seine Saint-Denis have noticed two benefits at the end of this work:

- The implementation of ICA-Req has had a positive impact in our archival practice.
- This audit tool allows the archivists to communicate better with all the people involved in the implementation of a new system. Archivists are able to explain the link between data creation and its preservation.

The quality of data and metadata is the cornerstone of a digital preservation policy. We should consider data creation as the first step in digital preservation. Producing data without establishing life cycle rules and retention schedules jeopardizes the success of any digital archival policy.

REFERENCES

- [1] – International Council on Archives, ICA-Req, Principles and functional requirements for records in electronic office environments (2008).
- [2] – Margaret Crockett, Principles and Functional Requirements for Records in Electronic Office Environments: Module 3 Guidelines and Functional Requirements for Records in Business Systems. Workbook for Archivists and Records Managers (2010).

Lourdes Fuentes Hashimoto is a French-Mexican records manager and digital archivist. She is currently in charge of electronic records management and digital preservation at the Archives départementales of Seine Saint-Denis (French local administration). She also provides support related to data migration. She is a member of the International Council on Archives where she is currently working with other archivists on the development of the ICA-Req French toolkit (a set of tools to adapt ICA-Req to the French archival practice). She holds a master's degree in modern history from the University of Bordeaux and a master's degree in digital technologies and archival studies from the Ecole nationale des chartes.