

Functional and System Appraisal of Electronic Records

Jim Suderman

Presented to the SARBICA Conference, Hanoi, 3 May 2004

The InterPARES 2 Project defines appraisal as “the process of assessing the value of records for the purpose of determining the length and conditions of their preservation.”¹ This presentation will use this definition but will focus less on the records and more on the functions of the systems containing the records – part of the process of assessing value through compiling information about the electronic records.

In this session I will introduce a typology of systems and describe one method of system design that has been useful, in my experience, for appraising electronic records. I will conclude with some of my own experience using system analysis to support the appraisal of structured and semi-structured records. The underlying assumption here is based on an assumption that not all electronic records are the same. The kind of records information systems contain and how those records are structured is, to a large extent, determined by the function of the system within the overall business context of the creating organization. The purpose of this session is to suggest that systems fall into general functional categories, on which further specific analysis of the function(s) of the system can materially assist in defining records and their value.

A general typology of information systems

It is generally presumed that archival electronic records, i.e., documents “made or received in the course of a practical activity as an instrument or a by-product of such activity, and set aside for action or reference,”² exist within some kind of information system. There is a typology of systems taken from a 1996 systems textbook that is summarized within the Authenticity Task Force Report.³ Despite its age, I find this typology remains helpful because it is based on general functions that remain constant or at least do not change as frequently as the underlying technologies. There are six major types of systems, of which I will only list four:

- Transaction Processing
- Management Information
- Office Automation
- Decision Support

Transaction Processing systems manage transactions at their most basic level. So, for example, transaction systems would manage our bank deposits and withdrawals, or documenting purchases at a grocery store. These kinds of systems frequently deal with

¹ From the InterPARES 2 Terminology database, checked 11 March 2004. Note that access to this database is currently limited to InterPARES 2 researchers.

² From the InterPARES 2 Terminology database, checked 12 March 2004. Currently this database is only available to members of the InterPARES 2 research teams.

³ InterPARES “Authenticity Task Force Report,” p. 18, available at http://www.interpares.org/book/interpares_book_d_part1.pdf, checked 26 March 2004. The textbook is by Kenneth C. and Jane P. Laudon and is entitled *Management Information Systems: New Approaches to Organization and Technology*. (New Jersey: Prentice Hall, 1996).

a high volume of transactions and, necessarily, these transactions are well defined in terms of their information content and structure. The information they contain can obviously be very important, but in most instances probably not for very long.

Management Information systems monitor transaction systems. It is their business to report exceptions and provide summaries of transaction activities. These reports assist an organization's management to effectively manage the daily business of the organization. Because these systems monitor transaction systems, their scope and purpose is generally clearly defined. For example, a web log application monitors all transactions, i.e., hits, downloads, etc., on the web server. This data is summarized for monitoring and on-going management of the website.

Office automation systems are those with which we're probably all familiar. These systems are built around office productivity applications like those comprising Microsoft Office. The purpose of this kind of system is often not well-defined. Users of the system have a great deal of individual flexibility in terms of what information to take into and maintain within the system. It is not surprising therefore that some of the most galling archival problems, e.g., email, in terms of defining records in context arise from this type of system.

The decision support systems I have experience with have been large and costly systems, and involve a data warehouse. Fundamentally these are reporting systems based on one or more databases. They are reporting systems because the systems do not make decisions; they make reports based on pre-defined data sets and within defined parameters. The purpose of these reports is to help decision-makers make decisions. These reports may or may not be central to the decision-making process within an organization. For example, political factors are often paramount for senior decision-makers in government. The database component is historical, rather than transactional – meaning that the data, which is often drawn from transactional systems, is transformed so that data from dissimilar source systems can be effectively compared and used together.

A typology such as this is helpful for the purposes of appraisal because it can provide us with a general idea of what types of records to expect from a system, and what some of the characteristics might be. A generalized typology like this allows one to categorize electronic records very generally in terms of the broad function or purpose of the system in which they exist. It does not provide enough information to support appraisal of records within a specific business. Additional information, specific to the creator, must be sought. This is consistent with the general archival approach to description which emphasizes describing records from the general to the specific.⁴

⁴ "Rule 2.1: Description from the General to the Specific." ICA Committee on Descriptive Standards, *ISAD(G): General International Standard Archival Description*, 2nd ed., (2000), available at http://www.ica.org/biblio/isad_g_2e.pdf, checked 26 March 2004.

Any method of system analysis and design would inevitably provide additional information to support archival appraisal. The use case methodology is a common method and is introduced here because it builds on an existing archival and records management strength, which is knowledge of function and activity. Use case methodology is also described as an effective way to communicate the essentials of a system to a non-technical audience, frequently the users of the system. This method supports archival appraisal in two ways: 1) the system is defined in terms of its functions, activities, and transactions which provides essential context for the records; and 2) it can be an effective means of representing the system in an appraisal report.

The use case methodology describes a system in terms of its functions. A “use case” is a discrete interaction between a user and the system. Use cases must be linked to specific business goals, which means that they must identify business decisions and actions.⁵ For effective system design, i.e., defining system requirements, use cases would ultimately be defined at a very detailed level. An example of a very specific use case would be the interaction that occurs when a user logs into a system. For most archival appraisal requirements it would not be necessary to break the functionality down to such a detailed level. Defining a system in general terms of who is interacting with it and for what purposes at whatever level the appraisal requires can be an enormously effective communication tool (e.g., for archival description or records retention scheduling). It can also be immediately helpful in terms of understanding responsibilities for record creation, use, maintenance, and preservation.

This is very similar to the work process analysis for recordkeeping purposes standard recently developed in Australia. “Work process” is defined as “the sequence of steps or actions to produce a business outcome that complies with governing rules.”⁶ Two types of analysis are described in this standard. Functional analysis involves a top down approach starting with the broad goals and objectives of an organization and breaking these down, layer by layer, down to the level of recurring transactions. Sequential analysis identifies the sequence of actions or transactions of a work process and their relationship to other processes.

The following examples are intended to illustrate how general and specific knowledge of system functions have been utilized to support archival appraisal. The first concerns a data warehouse – a project with multi-phased roll out. The appraisal was undertaken relatively early in the project, when the warehouse was in its first of three phases. The appraisal was undertaken to establish an approved disposal authority, i.e., record retention schedule, for the warehouse contents.

⁵ Sparx Systems *Enterprise Architect: UML Modelling and Design Tool* “Use Case Model”, available at http://www.sparxsystems.com.au/use_case_model.htm, (checked 12 March 2004). See also Richard Woolridge, CBD/e White Papers, “Introduction to Use Case Analysis” (Castek, 1999), accessible at http://www.cbd-hq.com/PDFs/cbdhq_991115rw_usecaseanalysis.pdf, checked 12 March 2004.

⁶ This is taken from the “Introduction” to the draft for public comment circulated in 2002, in the author’s possession. The draft standard is no longer accessible on the web.

There was no previous institutional experience to draw on for the appraisal of records from this kind of information system. In spite of this the appraisal was begun conventionally enough by trying to define the system, initially described simply as a database, in terms of its inputs, outputs and the data processing or functionality of the database.⁷ It quickly became clear that this database had inputs from many databases and from different jurisdictions. A similar finding was discovered with the outputs, which were used by a number of different offices, especially executive and policy groups. Descriptions of the data processing used words like “transformation” and involved individuals called “power users”. All this evidence suggested that this database was unusual. Further inquiries confirmed this and its true nature as a data warehouse became visible to us. Its purpose was to support decision-making and policy development.

Researching into exactly what a data warehouse / decision support system was led to an understanding that these systems take in legacy or “historical” data from disparate systems and transform it so that data from one source can be meaningfully related with data from other sources. Specialists, or “power users”, used specialized hardware and software to create (involving a second transformation) and query ‘data marts’ within the warehouse. Pre-programmed and *ad hoc* reports were generated for use by other offices. System documentation was plentiful and voluminous as these systems are costly, complex, and must be trustworthy. We were at a loss as to how to appraise datasets acquired from the source systems, the transformed data, the data marts, and the *ad hoc* reports. (The pre-programmed ones were not a problem). Furthermore, because of the specialized hardware and software used within the data warehouse we began to have doubts about the feasibility of being able to preserve anything from this source e.g., reports that existed simply as views of data on a screen. This was a concern because, with a fledgling electronic records program, there was no indication as to how much weight should be given to cost considerations in the appraisal. Having learned what specific software was being used, we researched that to try and get a handle on where records, if any, might be found. In general we got caught up in too detailed a view of the activities taking place within the warehouse. The only headway we had made was in eliminating the data from the source systems from the appraisal simply because the records in those source systems had already been appraised.

The appraisal pieces started falling into place when we stepped back from an increasingly complex description of the contents and technology of the system and looked again at the overall function of the system: that of supporting decision-making and policy development. This refocused the appraisal on the function that the system, and the information “products” it produced, was designed to support. The emphasis now was on determining how the products of the warehouse supported decision-making and policy development. In this case, it soon became evident that there were many supports for these two functions other than the data warehouse. In effect the data warehouse provided only one facet of a multi-faceted decision-making and policy

⁷ This is consistent with the InterPARES Authenticity Task Force conclusion that “understanding the nature and boundaries of electronic records required a detailed understanding of the business functions and activities of the record-keeping systems.” InterPARES Authenticity Task Force Report, p. 19.

development process. We perceived the importance of the transformed data and reporting capabilities of the warehouse to be high when our focus was on the specific technology and cost of the system. Once we reoriented our viewpoint to a functional one in which the system was placed in the larger context of the business process it supported, these values were correspondingly adjusted. In this way, by determining what type of system we were looking at and by examining how the principle clients at the highest level interacted with the system, we were able to reach an appraisal decision. In this case, the decision was that the system maintained no records with archival value. Had the appraisal concluded that some portion of the warehouse warranted preservation, then further analysis of the interactions between the users and the system, at the appropriate level, would have been necessary.

The previous example involved a complex system with highly structured information and well defined points of intersection with the business process to which it was connected. Far more frequently we are called to appraise records which are much closer to 'traditional' (meaning paper-based) records emerging from office automation systems. Appraising records emerging from office software, including word processing, email, spreadsheet, and presentation applications, etc., can be difficult because these are often poorly defined systems in terms of their connection with organizational business processes.

While this type of system may provide very little information of direct relevance to an appraisal, because office automation systems are primarily for enhancing productivity rather than supporting well-defined transactions, we can expect to deal with semi-or unstructured information. Frequently office automation tools are put in place primarily to enhance personal productivity rather than effective management of records. This is especially true for email.

Knowing the general type of system from which the records emerge, and examining systems from the perspective of their functions is an effective approach to appraising electronic records. The use case method of system analysis provides a common link with the archival and records management reliance on functions or business activities. A general and specific analysis of systems contributes meaningfully to appraisal both in terms of assessing the value of the records (i.e., through their relationship to the function they support) and also to help determine the record elements required for preserving the authenticity of the records.⁸ The analysis of the system in terms of its function will also contribute significantly to the archival description of the records, an essential preservation requirement.

⁸ These are specific appraisal activities identified in the InterPARES "Model of the Selection Function", (items A221, A222, A231, and A 232) available at http://www.interpares.org/book/interpares_book_m_app04i.pdf, checked 26 March 2004.