

Association for Information Systems AIS Electronic Library (AISeL)

AMCIS 2007 Proceedings

Americas Conference on Information Systems
(AMCIS)

December 2007

A Descriptive Case Study of Electronic Records Taxonomy Development at the Central Intelligence Agency

David Balda

Summer Bartczak
AFIT

Rhonda Syler
University of Arkansas at Little Rock

Alan Heminger
AFIT

Follow this and additional works at: <http://aisel.aisnet.org/amcis2007>

Recommended Citation

Balda, David; Bartczak, Summer; Syler, Rhonda; and Heminger, Alan, "A Descriptive Case Study of Electronic Records Taxonomy Development at the Central Intelligence Agency" (2007). *AMCIS 2007 Proceedings*. 98.
<http://aisel.aisnet.org/amcis2007/98>

This material is brought to you by the Americas Conference on Information Systems (AMCIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in AMCIS 2007 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

A DESCRIPTIVE CASE STUDY OF ELECTRONIC RECORDS TAXONOMY DEVELOPMENT AT THE CENTRAL INTELLIGENCE AGENCY

David A. Balda, david.balda@hickam.af.mil
Summer E. Bartczak, Air Force Institute of Technology, Summer.Bartczak@afit.edu
Rhonda A. Syler, University of Arkansas at Little Rock, rasyler@ualr.edu
Alan R. Heminger, Air Force Institute of Technology, Alan.Heminger@afit.edu

Abstract

Electronic records are growing at a rapid pace and the management of these records is evolving into a large, complex problem facing all organizations, including the military. The Air Force, specifically, is experiencing difficulty transitioning from a paper-based records management methodology to an electronically-based one. In contrast, the Central Intelligence Agency (CIA) has developed and implemented an electronic records management program that has been quite successful and that has gained notoriety inside and outside of governmental channels. This research uses a descriptive case study approach to examine the CIA's experience in developing an electronic records management program. Specific attention is given to the process of developing an overall organizing electronic records taxonomy. This study provides insight into an overall e-records implementation and strategy approach leading to specific recommendations for developing an enterprise-wide electronic records taxonomy. Additional benefits of the research include lessons learned from the CIA regarding the taxonomy development process and implementation.

Keywords: electronic records taxonomy, records management, electronic records

The views expressed in this paper are those of the authors and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the United States Government.

Introduction

Electronic recordkeeping is becoming a big problem and challenge for all organizations. What constitutes an electronic record, how it is stored, archived, superseded, and ultimately destroyed are some of the tough decisions facing the leadership and management of all organizations. The problem of managing this electronic flood of records is of great concern to many organizations as they struggle to develop ways to manage all aspects of electronic records and establish electronic recordkeeping processes. Electronic recordkeeping requires many activities to include analyzing the records that flow through an organization, crafting an electronic records management process to store and retain the records for specific time periods, and creating a retrieval scheme that makes it easier for the user to gain access to the records. Yet, perhaps the most difficult activity is the development of a classification method, or taxonomy, with which to organize the electronic records. In the context of this research, a taxonomy is defined as a necessary guiding framework that allows organizations to

Balda, Bartzak, Syler, & Heminger, *Electronic Records Taxonomy* sort their electronic records into logical subsets which enables a user's ability to retrieve it once it has been stored. As such, a taxonomy becomes a cornerstone of an organization's information and knowledge management approach because it is necessary to accurately file, store, and retrieve each electronic record that is processed. Regardless of business sector, the existence of a taxonomy enables an organization to achieve more efficient and effective information retrieval (Bruno & Richmond, 2003).

The objective of this research is to explore how the CIA devised and implemented its organizational electronic records taxonomy. Because the CIA is a federal entity with a history of success with taxonomy development, it is an ideal organization from which to learn, especially for the Air Force who is in the infancy of its efforts. To do this the following general research questions were investigated:

- (1) How did the CIA approach the development and implementation of its electronic records taxonomy?
- (2) What are key lessons learned that can be identified from the CIA's experience?

Literature Review

According to Menkus (1996), electronic records are "units of information that are comprehensive, authentic, accurate, reliable and contemporaneous with the events and decisions that they purported to reflect" (p. 39). His statement simply translated means that an electronic record takes a snapshot in time to capture an event or decision. A record is a static event that is retained for historical documentation and potential future use. Much like the Dewey decimal system used by any library to organize its collection, a well-defined electronic records taxonomy enables both the contributors and retrievers to quickly and easily locate appropriate electronic records.

There are a number of reasons why a taxonomy is a necessary foundation for an electronic records management system. A good taxonomy can provide 1) identification to control where information should be stored, categorized, and filtered, 2) discovery by providing additional information that will guide users on where to place the records, and 3) improvement in the speed of retrieving records (Bruno & Richmond, 2003). A robust taxonomy can also help to build trust within all the people that use it (Bruno & Richmond, 2003) in that they can retrieve any piece of information, even electronic records, with ease.

There are different ways to approach building an electronic records taxonomy. It can be built from scratch, modified from an existing file plan structure (internal or external to the organization) (Olsen, 2001), or developed through use of specialized software applications. Although building a taxonomy from scratch entails an enormous amount of time and effort, it can prove to be the most valuable approach for an organization as it leads to a more specialized taxonomy that is developed around the specific business. An approach involving modifying an organization's existing file plan (which is often in existence to guide paper records management) can also provide an organization with a starting point. Organizations can also choose to adopt a taxonomy from another organization and adjust it to fit their needs--a viable option when the

Balda, Bartczak, Syler, & Heminger, *Electronic Records Taxonomy* organizations have clear similarities. Finally, organizations may choose to a method called “autocategorization” which makes use of a software program consisting of a programmed rule set or learning application that automatically sorts e-records into pre-defined categories based upon a pre-designed taxonomy. This approach limits the need for human intervention but is less accurate.

Methodology

This research uses a case study approach. Prior case study research on electronic records management issues was found to be sparse at best. Of the literature that was uncovered, the themes were varied. For instance, one case study reported on an initiative at Indiana University that concerned its efforts to alter traditional records management strategies for paper records to accommodate electronic records (Bantin, 2001). Another case study revolved around electronic records management within the Australian government and how each Agency was managing electronic records differently (Sletten, 1999). Neither of these case studies, however, explored the issues of taxonomy development within the organizations.

This study extends the literature by using a descriptive case study method to investigate the CIA’s development and implementation of an electronic records taxonomy. This single case study is based on the premise that the problem of managing electronic records is representative of the situation many organizations find themselves in, and, as such, correlates well to one of Yin’s rationale for using a single case study approach – the representative, or typical case (Yin, 2003). To improve the overall quality of the research, case study tactics (Yin, 2003) were applied to address the construct validity, internal validity, external validity, and reliability of this effort although not described in detail here due to space constraints.

Specifically, the study focuses on two embedded units of analysis within the CIA--the CIA Records and Classification Management Group and a collection of key individuals directly involved in the electronic records taxonomy development and implementation process (but not direct members of the Records and Classification Management Group). Semi-structured interviews were conducted with a sampling of four members of the original Records and Classification Management Group who were internal to the taxonomy development process and three external staff members-- the Chief Information Officer (CIO) for the Office of Information Management and two members from the project management staff. In addition, supporting documentation, such as charts, policy documents, and project management documents were collected to provide additional data relevant to the research questions.

Analysis of the data collected was performed through triangulation for each research question. Consistent themes were identified through pattern-matching and cross-checked against the literature to identify links to theory and previous research findings.

Results

CIA's Four-Pronged Approach to Managing Electronic Records

Development of a sound electronic records taxonomy was to be an integral part of a larger plan for handling electronic records across the CIA enterprise. Previously, any decision to file a record was made immediately after a user created a record (post-creation). According to the Agency, handling records post-creation as a future strategy would require the directorates to meet every few years to upgrade systems and applications to meet the changing needs of the Agency and its associated business processes. Furthermore, managing records post-creation would cause increased spending due to technological fixes needed to correct and update network infrastructure and software applications. As such, the CIA decided it would manage electronic records focusing on the pre-creation process. To guide the process, it developed a four-pronged approach (Figure 1) for managing the Agency's electronic records to better support Agency knowledge.

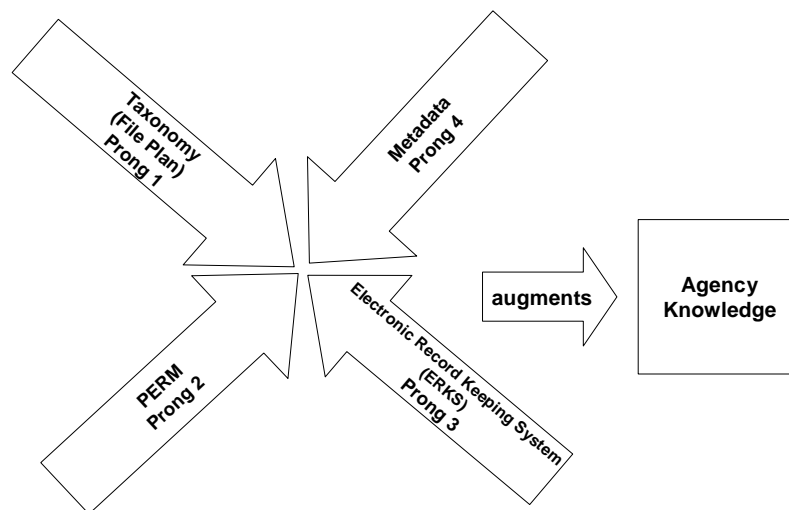


Figure 1. CIA Four-pronged Approach for Managing Electronic Records (Olsen, 2001)

The focus of the first prong of the approach was to devise a taxonomy. The Agency believed that a solid electronic records taxonomy would ensure homogeneity in the structure and composition of information assets which would enable immediate accessibility of the available information and expedite retrieval (Olsen, 2001). The focus of prong two of the approach called for the creation of a Proactive Electronic Records Management (PERM) software solution (Johnson, 2003). This initiative was further divided into four phases. Phase I consisted of developing and implementing an electronic software platform that integrated multiple software packages used to manage different electronic records into a single user-interface program. Phase II began with the population of the electronic record taxonomy levels into the software. Phase III addressed the addition of tools for users for managing the electronic records data. Finally, in Phase IV, a schema, or organization method, was applied to the electronic records storage to ensure long-term accessibility and retention. Prong three of the approach focused on implementation of an Agency-wide standard for the creation, storage, and deletion of electronic records

Balda, Bartczak, Syler, & Heminger, *Electronic Records Taxonomy* within any new software application. The result was the Electronic Record Keeping certification standard, or ERKS, which established a baseline electronic record format for any record created from within an application software system. When ERKS was applied to all newly created Agency software systems, it provided a method to automatically capture and retain records produced in accordance with Agency-specified record keeping guidelines without human intervention. Finally, prong four of the approach called for establishing standard metadata requirements (i.e. dates, author information, record characteristics, record disposition information, and classification) for electronic records.

CIA Taxonomy Development Methodology

In general, the CIA approached developing an electronic records taxonomy by first examining its own business rules. These business rules dictate how the Agency handles record information generated from the business processes and what must be retained in accordance with federal guidelines. The records generated from a directorate's business processes were being altered from paper to electronic; however, the method to handle them needed to remain similar--records still needed to be filed, maintained, and retained in accordance with regulatory guidelines

As one of the early organizations to embark on such an effort, the CIA found no specific methodology to follow in creating its taxonomy—it evolved over time. As such, the Agency chose to begin its effort using its current paper-based file plans and processes as a foundation. One of the key challenges was determining retention periods for each record type as certain paper-based records were already being flagged with a record control schedule timestamp to ensure preservation. Another challenge involved reconciling different directorates and their unique ways of doing business. Each directorate had its own unique business process and, therefore, its own unique paper-based file plan. The implementation team started by combining all the directorate file plans into one master Agency file plan focusing specifically on those record items that had a retention period of any kind. Records that did not have a retention timestamp were eliminated from the master file plan. Next, the team grouped the records into logical main categories and subjugated associated files underneath each of the main categories. After the master file plan was drafted, it had to be verified and reviewed by each individual directorate. The implementation team was supported by subject matter experts from within each directorate. It is important to note that even before the electronic record management taxonomy project was conceived, the Agency had already placed a high-level Information Management Officer (IMO) within each directorate. The philosophy behind the strategic placement of the IMOs was to add expertise in managing the records and information that were specific to each directorate. As such, IMOs became key figures in the taxonomy creation and implementation process.. These individuals also had an understanding of records management guidelines and became an integral part of controlling and maintaining the integrity of records created in their respective directorate as well as ensuring compliance with federal record keeping guidelines.

The implementation team, IMO, and selected directorate subject matter experts, went through each record category of the master file plan individually. Every record item and category was scrutinized to ensure it had direct ties to the business

Balda, Bartzak, Syler, & Heminger, *Electronic Records Taxonomy*
processes within the directorate—if not, it was deleted. This process was continued until each directorate was satisfied with its own unique file plan and then all were merged together to form one large enterprise file taxonomy.

CIA’s Conceptual/Logical Electronic Records Taxonomy

The final version of the Agency’s file plan is comprised of three levels. The top level consists of 13 main categories that directly correspond to Agency functions such as Administration, Finance, Health Safety & Environment, Human Resources, Information Management, Information Release, Information Technology, Intelligence Activities, Logistics, Operations, Policy, Projects & Programs, and Security. (Figure 2)

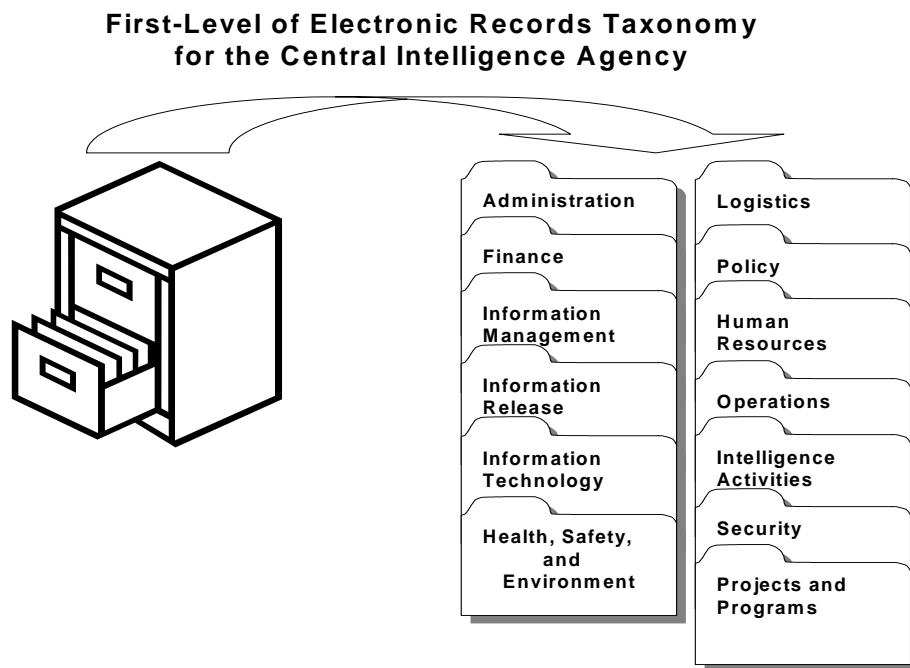


Figure 2. First level of the CIA Electronic Records Management Taxonomy

The first (or top) level covers all the main business process areas that comprise the Agency. At the second level, the various business areas are further broken down into categories that are more specific. For example, the Administration area is broken down into 12 categories to include: acquisition, supply and finance, activities reports/weeklies, budget, calendars, general administration, health safety and environment, information management, information release, information technology, logistics, security, and travel and transportation. The third and final level is specifically reserved for the end users. At this level, end users can make decisions about where to place a document into the taxonomy. As an example, within the main category of “Administration” and the second-level of “Acquisition” there are the options of billing records, charge-back calculations, component financial accounting files, contracting officer contract files, inventory, purchase orders/loan issue receipts, and support documentation. The end-user can select the appropriate third-level category and the electronic records management system will place the document in the appropriate position within the repository. The record

retentions are also affixed to the document based upon the third level of the electronic record taxonomy where the record is maintained for the proper length of time (Figure 3).

Example of Second and Third Levels of the Electronic Records Taxonomy for the Central Intelligence Agency

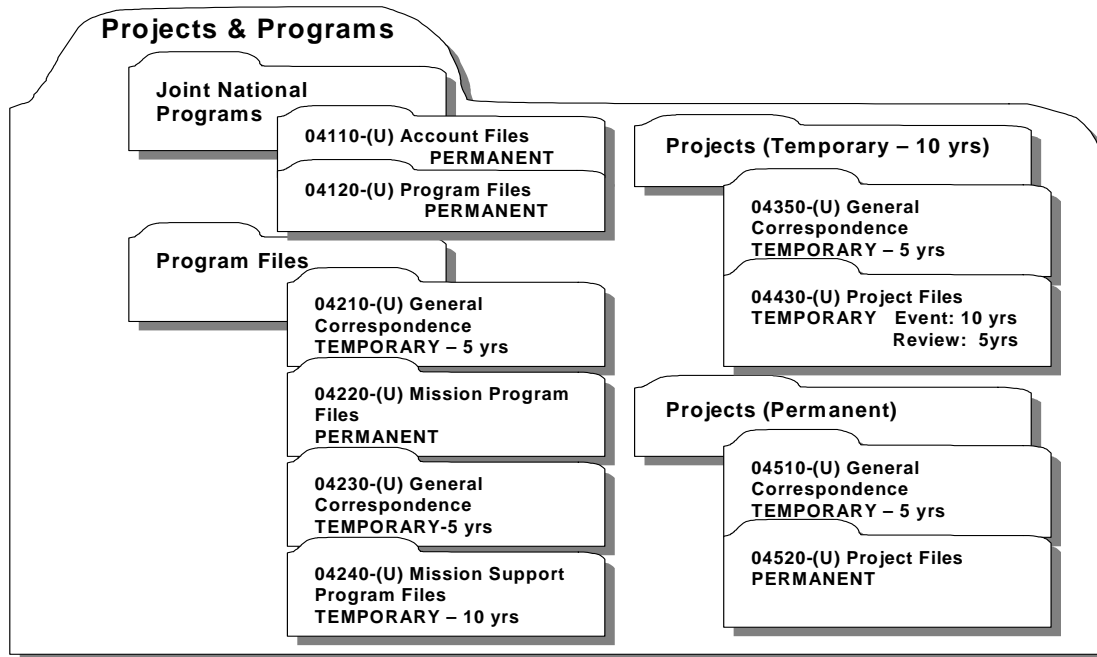


Figure 3. Depiction of Second & Third Levels of the Electronic Records Taxonomy

The primary reason that the Agency decided upon only three levels for the taxonomy was to simplify. The Agency's Records and Classification Group believed that if the taxonomy were to expand to more than 13 different categories and progress deeper than three levels that the increased burden placed upon the end users would result in a serious reduction in the use of the taxonomy and ultimately the electronic records management system. Therefore, the designers and implementers of the taxonomy spent individual time with each directorate to identify data requirements and specific records that related directly to its business rules.

CIA's Usage of the Taxonomy within an Electronic Records Management System

As mentioned previously, the Agency incorporated a number of different commercial off-the-shelf (COTS) software products and integrated them into a single user interface named the Permanent Electronic Records Management (PERM) system. The software included: Foremost, DocsOpen, Lotus Domino, Convera Retrievalware, Oracle, and Microsoft Internet Information Service (IIS). Foremost was selected for its records management capabilities. DocsOpen was able to store any Microsoft Office document. As the Agency uses Lotus Notes for its e-mail service, Lotus Domino was a natural choice. Convera Retrievalware was purchased because it had an index capability and allowed full-text searching of textual

Balda, Bartczak, Syler, & Heminger, Electronic Records Taxonomy documents and attachments. Oracle was used for housing document metadata and Microsoft IIS enabled web-based access to Convera Retrievalware.

The Agency's electronic record taxonomy was built directly into PERM system thereby eliminating the need for the end-user to search for a desired filing location for the record. Anytime a user creates an electronic record, the system prompts him/her for the type of record he/she will be creating, either permanent or temporary. Once the appropriate type of record is selected, the PERM system presents the user with a listing of possible record types available. These record types are tailored to meet the needs of the end user. Once the user selects the record type the software then attaches file tags to the record for filing and retrieval. Since the taxonomy has the appropriate information regarding the file type, retention period and specific handling information, the software automatically manages the record in a predetermined way. At the time of this case study, the Agency was in the final stages of developing an automatic electronic record classification system using autocategorization software in order to achieve some additional efficiencies.

Some Lessons Learned from the CIA's Experience

There were many growing pains associated with developing an Agency-wide electronic records taxonomy; however, the CIA has gained some notoriety for its efforts. General Motors, the U.S. Coast Guard, and, now, the U.S. Air Force have examined the CIA's efforts for key lessons learned.

Lessons Learned when Building a Taxonomy

Taxonomy as a Foundation. The CIA realized that a taxonomy was needed as the cornerstone for constructing a viable electronic records management system. The Agency used the Federal records control schedule (RCS) to guide the building of its taxonomy. The RCS identifies records and the specified periods of time that each are required to be maintained by organizations. As such, the retention period of records became the main foundation to building the electronic records taxonomy. The RCS has, in fact, defined the business process (by its records), maintains access controls for each file series, and retention periods on when the material can be destroyed.

Involve the Right People. The Agency also built a team of highly competent personnel to assist in building its taxonomy. Personnel chosen to participate in the electronic records taxonomy development process were credentialed with Masters of Library Science, Masters in Information Systems, and Masters in Computer Science. Each member of the team brought his/her own specific expertise (records management or otherwise) to the table and offered valuable suggestions to create a viable, end-user focused electronic records management taxonomy.

One key action in involving the right people was the strategic placement of IMOs within each directorate. This provided solid, foundational advice to the organizational leadership. Because the IMOs understood the internal processes of their specific areas, they were critical in identifying the records and information each directorate created, received, or exchanged with other directorates.

It was also recognized that managers or individuals who became convinced of the poor state of recordkeeping and its impact on mission accomplishment become strong advocates for change. It was found that development of a working taxonomy with visible benefits went along way in creating vocal proponents (Olsen, 2001)

Governance is Important. Another key element involved the way the CIA chose to have a centralized governance for establishing and executing mandatory records and information management policies for the entire enterprise. This action was critical to successful generation and promulgation of policy and procedures and in achieving subsequent compliance.

Be General but Specific. The Agency realized it could not achieve perfection in accurately defining every record category. Instead, it decided to maintain a broad perspective and looked to identify general categories. It was determined that if the record category areas were defined too restrictively, it would hamper use and reliance upon the taxonomy by users. At the same time it was also learned that taxonomy names or file tags had to be unambiguous and organization specific to promote full user understanding and accuracy of filing and retrieval .

Lessons Learned in Implementing the Taxonomy

Mandate Compliance at the Right Levels. The Agency realized that once the file tags were completed that the taxonomy should be baselined and placed under a configuration control process. Such action ensured that everyone within the organization used the same taxonomy. Beyond the baseline, it was determined that flexibility could be allowed at the lower levels of the taxonomy in order that each component of business could further tailor the file plan to meet its needs.

Start Small and Grow. Finally, it was learned that the taxonomy should be tested beginning at the lower levels of the organization, such as a small office. It was found that testing a taxonomy on a lower level office's paper records was an excellent start to assessing usefulness across the enterprise.

Seek Buy-in. After the Agency developed a solid taxonomy, it was incorporated into a software application that allowed end users to file records directly from desktops and applications. Key lessons were also learned in bringing the application on-line such as the need to perform a detailed pre-installation briefing with each directorate leader to explain the process and seek buy in. To further enhance buy-in and generate interest in the file plan, a marketing campaign was developed to reach end users which made use of newsletters, posters, and briefings to explain the benefits of implementing the taxonomy and electronic records management software..

Plan for Incremental Roll-out. The Agency also developed an aggressive rollout plan combined with a technical assistance program. The Agency used a 32-day installation plan that was quite aggressive to convert everyone within a directorate from a paper-based records management program to an electronically-based system. However, this rollout plan was revised after first converting the Records and Classification Group to the new electronic records management system. The rollout plan was perfected before implementing it within other directorates.

Educate and Train. The Agency also discovered the importance of user education and training. They made sure user training occurred immediately after the director pre-installation briefing. This laid the foundation for the users to begin thinking differently about records management and to get accustomed to their new business process. The seven-day technical assistance program was created to allow directorate personnel an opportunity to ask questions specifically pertaining to the new system and its operations before the implementation team departed.

Conclusion

The CIA's success in devising an electronic records taxonomy involved a number of different factors that facilitated its development and its implementation throughout the Agency. The research revealed that the development of an Agency-wide electronic records taxonomy was a critical precursor to the CIA successfully managing its electronic records. Furthermore, the research identified that although taxonomy development was a critical and difficult first step, that creating buy-in for implementation, execution of implementation, and eventually maintaining the taxonomy and electronic records management system solution were equally difficult. Beyond the specific lessons learned offered, the research highlighted the struggles that all organizations are having in the realm of e-records management and how few resources there are to help them through the struggle. The fact that the CIA had to develop its own methodology for approaching the taxonomy development problem and devise its own "system" solutions for the most part indicate that this is a problem that organizations are finding hard to get their arms around.

Limitations & Future Research

The limitations of this study are in line with those generally stated for case study research. To begin, this was a single case study so the generalizability of the results is somewhat limited. Although there are many aspects of the CIA that make it like any other organization, it is definitely different with respect to many others. Additionally, a limited number of individuals were interviewed. Furthermore, the individuals interviewed had specific experience in the development of the CIA's electronic records taxonomy but may not have had the full and complete knowledge the broader of records/electronic management. Finally, although research design considerations were put in place, this study was conducted by a single researcher which may have introduced bias and impacted interpretation of the data and results.

Despite the limitations of the study, the research revealed how little is published in practitioner or academic literature on the specifics of dealing with electronic records management and taxonomy development issues in organizations. As such, the findings from this study are useful and open the door to other lines of inquiry. Generally, we feel additional case study investigations of other electronic record management and taxonomy development initiatives would be most informative at this point. More specifically, an effort to compare electronic records management strategies and technologies used across the Federal sector would provide an interesting extension to this study. Finally, an empirical look at benefits of taxonomy

Balda, Bartczak, Syler, & Heminger, Electronic Records Taxonomy development and implementation versus the emerging focus on information and knowledge discovery through search mechanisms might provide indication of the future of electronic records management in organizations.

Bibliography

- Bantin, P. C. (2001). "The Indiana University Project: Lessons Learned". *The Information Management Journal*, January 2001, 16-24.
- Bruno, D., & Richmond, H. (2003). "The Truth About Taxonomies". *The Information Management Journal*, March/April 2003, 44-53.
- Creswell, J. W. (2003). *Research Design - Qualitative, Quantitative, and Mixed Methods Approaches* (2nd ed.). Thousand Oaks: Sage Publications.
- Johnson, C. (2003). *Proactive Electronic Records Management (PERM) Informational Briefing for the U.S. Air Force*". Powerpoint Presentation.
- Menkus, B. (1996). "Defining Electronic Records Management". *Records Management Quarterly*, 30(1), 38-42.
- Olsen, C. J. (2001). *Build It, Buy It, Steal It - But Your Organization Needs a Taxonomy for Its Information Survival*". Paper presented at the ARMA, Montreal, Canada.
- Sletten, L. (1999). "Lessons from Down Under: Records Management in Australia". *The Information Management Journal*, January 1999, 26-32.
- Yin, R. K. (2003). *Case Study Research: Design and Methods* (3rd ed. Vol. 5). Thousand Oaks: Sage Publications.