

Inge Alberts
University of Ottawa, Ottawa, ON, Canada

An Information Governance Methodology to Tackle Digital Recordkeeping Challenges: The Convergence of Artificial Intelligence, Business Analysis and Information Architecture

Abstract

Information Governance (IG) is defined as a “strategic, cross-disciplinary framework composed of standards, processes, roles, and metrics that hold organizations and individuals accountable for the proper handling of information assets” (ARMA, 2018, 1). Within this framework, sound recordkeeping practices are paramount as they ensure information security and privacy, efficiency in the delivery of programs and services, and reliable methods for consistently managing data, e-documents, and records along the same continuum (Desrochers, 2012; Hagemann, 2013; Smallwood, 2014). Faced with the exponential growth of information sources, the instability of new digital forms and the convergence of archival and content management systems, organizations are faced with numerous challenges to adopt an effective and all-encompassing approach to exploit their information assets.

As a “super-discipline” that includes components of several key fields such as law, records management, information technology, security management and business analysis, Information Governance calls for a “new breed of information professional who is competent across these established and quite complex fields” (Smallwood, 2014, 4-7). Today the profession is witnessing an evolution of the traditional role of the recordkeeping professional, whether a recordkeeper or an archivist, whose tasks now focus on the entire information continuum – from the point of creation to the end of the lifecycle – while also including strategic planning, business analysis and risk management. As a result, these professionals must acquire additional skills so that the management of information effectively contributes to the achievement of one organization’s business objectives. To that aim, the optimization of business processes and the valorization of information constitute two sides of the same coin.

Despite the significant pressures to manage information as a strategic resource, organizations are still struggling to implement effective digital solutions to support the Information Governance requirements throughout the information lifecycle. This situation is truly symptomatic of profound “paradigm shifts” (Joseph, Debowski & Goldschmidt, 2012) within the recordkeeping discipline in search of a new “orthodoxy” (Lappin, 2010). The discipline needs to replace the EDRMS (Electronic Document and Records Management System) conceptual approach referred to as the “Central Registry Model” (Carnahan, 2014) where digital records are being saved in a virtual folder structure representing the corporate business classification with associated retention and access rules. Recent studies have revealed that this strategy is largely insufficient to meet the legal and operational recordkeeping requirements (McDonald, 2010; Joseph et al., 2012; Nguyen et al., 2014) as “only on average 7-9 percent of enterprise content is managed as “official” or scheduled records” (Hagemann, 2013, 229).

The main reason for this lack of success lies in the transfer of the recordkeeping responsibilities from the specialists to individual employees who are generally unskilled in records management practices (Joseph *et al.*, 2012, 63). With the EDRMS approach, staff is expected to create a new folder at the start of a project, decides which piece of information has value for the organization and manually classifies each record against the business classification scheme (Lappin, 2010, 253). A report from the U.S. National Archives and Records Administration concludes that depending on busy employees who are focused on achieving their organization mission leads to inconsistent recordkeeping across the government (NARA, 2014, 5). These conclusions are reflected in different studies that address the lack of employee motivation with respect to tasks related to the appraisal and classification of records (Bailey, 2009; Mäkinen & Henttonen, 2011, Goldschmidt, Joseph *et al.*, 2012; Jordan & deStricker, 2013; McKemmish & Piggott, 2013; Oliver & Foscarini, 2014).

In 2009, Steve Bailey published a paper entitled “Forget electronic records management, it’s automated records management that we desperately need” which explores innovative ideas to automate manual recordkeeping tasks. However, despite significant developments in natural language processing and machine learning techniques, the recordkeeping discipline has been reluctant to apply automation to better deal with the growing volumes of digital information (Vellino & Alberts, 2016; Rolan *et al.*, 2019). The main reason is that a valid recordkeeping system (the policies, the standards, the methodologies, the processes and the technologies) should embed the Information Governance compliance requirements insuring accountability, transparency, integrity, protection, compliance, availability, retention and legally defensible disposition practices (ARMA, 2018). As the field of artificial intelligence is becoming a critical requirement for the recordkeeping profession, experts need an approach which embeds the compliance requirements and ensure the traceability of the decisions over the automated processes (Alberts & Eby, 2019).

In response to this requirement, this paper presents a practical Information Governance methodology that borrows concepts, tools and techniques from the fields of Artificial Intelligence, Business Analysis and Information Architecture. This methodology which was developed collaboratively (see Alberts, Schellinck, Eby & Marleau, 2010; Alberts, Marleau, & Eby, 2015; Alberts & Eby, 2019) is based on the available literature, standards and methodologies in records management (such as Smallwood (2014), DIRKS (NSW, 2018) and ISO 15489), several action research projects, as well as interviews and cognitive inquiries with expert consultants in Information Management and Digital Recordkeeping (N=24). The methodology focuses on business analysis for identifying information of “value” for an organization (this value being specific to each organization) and on information architecture (describes here as a common set of terms with their semantic associations).

The methodology is based on the records management concept of “macro-appraisal”, which is used to assess the value of information in context (Alberts *et al.*, 2010). To determine the value of information, a functional model is developed, and information resources are inventoried using process mapping. Because business processes clearly represent how, where, and which information is produced to fulfill the organization’s mandates, they allow valuable resources to be targeted. Process maps provide rich insights on the form and format of a valuable record –

whether a source of data, unstructured information, or knowledge - where they acquire value during business, how they circulate within the organization, and who is responsible for them. These business-driven inventories are an incredible source of contextual knowledge to develop records management automation strategies. Leveraging an information architecture built using business principles is paramount to the success of any Artificial Intelligence initiative that seeks to automate the records and information management process (Alberts & Eby, 2019).

Moreover, the analysis of which information resource has value as well as the development of the information architecture serve as a stepping-stone for the implementation of the information governance strategies; each area of expertise (such as Information Management, Information Technologies, Security, Business Performance, etc.) will have in hand a common representation of information flows, the creation of value in the organization, the business risks associate with each information resource as well as the technologies required to fulfill the organizational mandate. For the specialists in these different areas of expertise, having a common representation of business concepts will allow the development of targeted information management strategies and tools, such as ontologies, glossaries, thesauri, procedures, retention schedules, metadata schemas, hard drive permission structures, etc.

In the paper, the five-step methodology comprising 1) IM Need and Capacity Analysis; 2) Functional Analysis; 3) Process Analysis; 4) Information Architecture Development; 5) NLP Requirement Specifications and Iteration will be presented. This presentation will be followed by a discussion demonstrating how the methodology fulfills the Information Governance compliance requirements while promoting a better coordination of information management strategies, with both IT, security and performance measurement strategies. The methodology also lays the foundations to integrate recordkeeping automation to current recordkeeping practices based on techniques derived from research in artificial intelligence, and more specifically Natural Language Processing (NLP) (Albert and Eby, 2019).

As reflected by the dynamism of the current research on recordkeeping automation, there is a tremendous opportunity to build scalable recordkeeping practices to handle 21st century digital records volume. For private and public organizations, leveraging technologies to tackle the current information management crisis will enable agencies to meet their objectives of operational efficiency, transparency and accountability. For recordkeepers, it will initiate a long-awaited modernization “where specialists manage the semantic engines that create meaning from the mountains of information [...], a profession of finesse and tuning to best create knowledge from data and information” (Carnahan, 2014, p. 16).

References

ARMA (2018). *Information Governance Body of Knowledge (iGBOK)*. Overland Park: Arma International. 274 p.

Alberts, I., & Eby, C. (2019). Enterprise Knowledge Graphs for Intelligent Information Management. *ARMA NCR 2019: Fall Information Management Conference*, Ottawa, Canada.

Alberts, I., Marleau, Y., & Eby, C. (2015). Information Value as a Dynamic Concept: Moving Towards More Agile Records Management, Ottawa, Canada, 24-26 November 2015, *ARMA NCR 2015: Fall Information Management Conference*.

Alberts, I., Schellinck, J., Eby, C., & Marleau, Y. (2010). Bridging Functions and Processes for Records Management/Comblent le fossé entre les fonctions et les processus dans la gestion documentaire. *Canadian Journal of Information and Library Science*, 34(4), 365-390.

Bailey, S. (2009). Forget Electronic Records Management, It's Automated Records Management that we Desperately Need. *Records Management Journal*, 19(2), 91-97.

Carnahan, A. (2014). Big Data Needs Big Classification. *IQ: The RIM Quarterly*, 30(1), 16-19.

Desrochers, P. (2012). Recordkeeping and Linking Government Data in Canada. *IEEE Intelligent Systems*, (3), 50-53.

Hagmann, J. (2013). Information Governance—Beyond the Buzz. *Records Management Journal*, 23(3), 228-240.

International Standard Organization (2016). ISO 15489-1:2016 Information and documentation — Records management — Part 1: Concepts and principles.

Jordan, I., & de Stricker, U. (2013). Information Management in the Canadian Federal Government: Principles, Practices and the Role of Information Professionals. URL: <https://curve.carleton.ca/39509abe-6e49-483b-8620-bc408a4e0819>. Last visited on 29/12/2015.

Joseph, P., Debowski, S., & Goldschmidt, P. (2012). Paradigm Shifts in Recordkeeping Responsibilities: Implications for ISO 15489's Implementation. *Records Management Journal*, 22(1), 57-75.

Lappin, J. (2010). What will be the Next Records Management Orthodoxy? *Records Management Journal*, 20(3), 252-264.

Mäkinen, S., & Henttonen, P. (2011). Motivations for Records Management in Mobile Work. *Records Management Journal*, 21(3), 188-204.

McDonald, J. (2010). Records Management and Data Management: Closing the Gap. *Records Management Journal*, 20(1), 53-60.

McKemmish, S., & Piggott, M. (2013). Toward the Archival Multiverse: Challenging the Binary Opposition of the Personal and Corporate Archive in Modern Archival Theory and Practice. *Archivaria*, 76, 111– 144.

National Archives and Records Administration (2014). Automated Electronic Records Management Report/Plan. Office of the Chief Records Officer for the U.S. Government. URL:

<https://www.archives.gov/files/records-mgmt/prmd/A31report-9-19-14.pdf>. Last visited on 25/07/2020.

Nguyen, C., Stockdale, R., Scheepers, H., & Sargent, J. (2014). Electronic Records Management-An Old Solution to a New Problem: Governments Providing Usable Information to Stakeholders. *International Journal of Electronic Government Research*, 10(4), 94-116.

NSW State Archives & Records. (2018). Strategies for Documenting Government Business: the DIRKS Manual. URL: https://www.records.nsw.gov.au/sites/default/files/Recordkeeping/DIRKS_Manual.pdf. Last visited on 10/01/2020.

Oliver, G., & Foscarini, F. (2014). *Records Management and Information Culture: Tackling the People Problem*. Facet Publishing.

Rolan, G., Humphries, G., Jeffrey, L., Samaras, E., Antsouпова, T., & Stuart, K. (2019). More human than human? Artificial intelligence in the archive. *Archives and Manuscripts*, 47(2), 179-203.

Smallwood, R. (2014). *Information Governance: Concepts, Strategies and Best Practices*. Hoboken, New Jersey: Wiley.

Vellino, A. & Alberts, I. (2016). Assisting the Appraisal of E-mail Records with Automatic Classification. *Records Management Journal*, 26(3), 293-313.