

Development of an ontology of functional activities for archival systems

Georg Gaenser, School of Library, Archival and Information Studies, University of British Columbia

Abstract

The project “development of an ontology of functional activities for archival systems” aimed at creating an overarching ontology for archival and records management environments in which current, semi-current and noncurrent records are managed in a coherent system. The basis of the ontological representation of recordkeeping functions is the analysis of the contents of the most relevant international standards. An ontological representation of the functional activities provides records professionals, stakeholders and clients with a clear image of the functions that need to be implemented and performed in records management and archives.

Project Aim

The development of an ontology of functional activities for archival systems is an InterPARES Trust project. The research was conducted by Giovanni Michetti and Georg Gaenser. The aim of this project was to develop an ontology representing the main functions and activities performed in records management and archival environments. An ontology is a – usually graphical – classification of concepts, that shows the relationships of these concepts and describes them. The ontology of functional activities for archival systems is based on an analysis of the most relevant international standards. Since there are over 100 standards that deal with aspects of records management and archives, the project aims at combining and connecting the information of the 20 most widely used standards in an overarching ontology. The project aims to provide a comprehensive and systematic image of the functions and activities that are performed throughout a record’s life-cycle. Thus, the ontology helps practitioners, stakeholders and software developers to understand which functions and activities need to be performed by records professionals, how these functions are related and what they involve.

Research Method

The research-project for the development of the ontology was divided into three interconnected steps. The steps followed each other sequentially, however, since all three steps were interrelated, analysis of the sources and adding data could happen throughout the project duration.

The first step involved the identification and selection of the sources that subsequently form the basis of the ontology. The aim was to choose the relevant and most commonly used standards to form a

solid foundation for the ontology. The project team identified widely used and accepted standards such as the ISO 15489 series (Records management), the ISO 16175 series (aka. ICA requirements, Principles and functional requirements for records in electronic office environments), the OAIS Reference Model (ISO ISO 14271), the ISO 27000 series (Security techniques Information security management systems) or the ISO 30300 series (Management systems for records), and several other standards and technical reports¹ to cover every aspect from creation to preservation. In total 20 standards² and technical reports were chosen in this process, i.e. identified to contain relevant information and subsequently analyzed. The assumption of relevance of the chosen standards and technical reports is supported by the literature.³

Except for two de-facto standards, the Core Archival Functions (CAF) by the SAA and the General Accepted Recordkeeping Principles (GARP) by ARMA, all standards used for the ontology were issued by the International Organization for Standardization and are thus de-jure standards.⁴ In addition, three ISO technical reports were analysed – ISO TR 15489-2 (Records Management), ISO TR 26122 (Work process analysis for records) and ISO TR 15081 (Electronically stored information). The analysis started with ISO 15489-1, a widely accepted and endorsed standard for Records Management, that is used world-wide in all kinds of organisations and administrative contexts.⁵ Since ISO 15489-1 is considered to be a high-level guideline to set up principles and practices for a Records Management program, a range of standards was subsequently developed. These standards try to fill gaps and elaborate or develop concepts and principles brought up in ISO 15489-1.⁶ Since there exist more than 100 standards that have some relationship to or

¹ See “ISO deliverables”, accessed February 6, 2018, <https://www.iso.org/deliverables-all.html> – “A Technical Report [...] may include data obtained from a survey, for example, or from an informative report, or information of the perceived ‘state of the art’.” Whereas, according to ISO an international standard “[...] provides rules, guidelines or characteristics for activities or for their results, aimed at achieving the optimum degree of order in a given context.”

² See full list of used sources under used sources at the end of this article.

³ E.g. Robert Smallwood, *Information Governance: Concepts, Strategies and Best Practices* (Hoboken, NJ: John Wiley and Sons, 2014), 77-81 and 85-86. See also, Shadrack Katuu, “Managing digital records in a global environment: A review of the landscape of international standards and good practice guidelines”, *The Electronic Library* 34, no. 5 (2016): 873-876.

⁴ See M. Pember, “Sorting out the standards: what every records and information professional should know”, *Records Management Journal* 16, no. 1 (2006): 22-23. About the difference between de-jure and de-facto standards see also Smallwood, *Information Governance*, 76.

⁵ See P. Joseph, Debowski, S. and Goldschmidt, P., “Paradigm shifts in recordkeeping responsibilities: implications for ISO 15489's implementation”, *Records Management Journal* 22, no. 1, (2012): 59 and Katuu, *Managing digital records in a global environment*, 873-874.

⁶ See Katuu, *Managing digital records in a global environment*, 873; See also R. Frost, “New ISO management system standards for records facilitate transparency in corporate governance” 2011, accessed February 6, 2018, <https://www.iso.org/news/2011/11/Ref1487.html> and M. Cottin et al., “Méthodes techniques et outils”, *Documentaliste-Sciences de l'Information* 48, no. 4 (2011): 10-11.

cover aspects of records management or archival functions⁷, it was necessary to identify the most widely accepted and used standards.

The second step comprised a systematic analysis of the selected sources and the modelling of statements drawing from the content of the sources. The contents of the selected standards were condensed into “semantic triples” – subject, predicate, object – that are the foundation of the ontology. All diagrams of the ontology are based on these statements. In modelling the statements, the project-team tried to stay as close as possible to the sources. However, to achieve and enhance consistency in modelling the “triples”, the utilised predicates were limited to a few.

Subject	Predicate	Object
Records Management	<i>Involves</i>	Metadata Management [taken from ISO 16175-2, p. 9 and ISO 23081-1, pp. 2 and 4]
Records Management	<i>Involves</i>	basing decisions on analysis and risk assessment of business activities [taken from ISO 15489-1, p. 3]
Disposition	May involve	destruction of records and metadata [taken from ISO 15489-1, p. 18]
Appraisal	Aims at	identifying business, legal and other requirements for records, used to specify records creation when work processes and records systems are designed or redesigned [taken from ISO 15489-1, p. 16]

Fig. 01 – Examples of the modelled statements

Thus, solely descriptive statements use “=”, statements that describe an outcome, a goal or a consequence use “aims at” and statements, that describe that an object is included, implied or encompassed, use “Involves”. In cases where the latter is conditional, “May involve” is used as predicate (Fig. 01). To clarify, that some statements from certain standards apply only under specific circumstances, a “qualification” is inserted in square brackets (Fig. 02).

Subject	Predicate	Object
Creation	<i>Involves</i>	[in records management systems] capturing evidence of a business activity [taken from 16175-2, p. 13]

Fig. 02 – Example of a statement with “qualification”

A “qualification” in this ontology is usually based on the scope and purpose notes provided by the standards in their introductions. However, whenever a statement in this ontology is the result of a merger of similar or identical statements from two or more standards, no qualification is added. It can be assumed, that a statement, occurring in several sources, is generally applicable.

⁷ See InterPARES 3 Project: TEAM Canada. “General Study 04 – International Standards Relevant to the InterPARES 3 Project” Vancouver: InterPARES 3 Project (2012), accessed February 6, 2018, www.interpares.org/ip3/display_file.cfm?docip3_canada_gs04_international_standards.pdf; See also Katuu, *Managing digital records in a global environment*, 893-894.

This analytic procedure involved adapting and refining statements to make them fit these requirements. Nevertheless, the statements were modelled as close as possible to the sources where it was feasible and did not distort readability. The described method was chosen for its greater flexibility. It is adaptive enough to account for the inconsistencies of the sources.

In some cases, statements can be the result of merging two or more similar statements (Fig. 03). This is recognizable and traceable since the sources, including page numbers, are noted in a coded form for each statement in square brackets – e.g. [A3]. (“A” stands for ISO 15489-1, “B” for GARP, “C” for ISO 23081-2 (Metadata for records), etc., the number indicates the page.) Statements from different standards, that are – essentially – the same or similar were merged. Therefore, the final statements used as the basis for the ontology are the result of controlled and consistent interpretation.

Subject	Predicate	Object
Disposition	<i>Involves</i>	authorizing retention decisions in accordance with jurisdictional requirements [A18] [H15] [J52]

Fig. 03 – Example of a merged statement

The third step was the development of a visual representation of the results of the analysis by means of a mind-mapping approach. Mind maps are a common and simple tool to describe a concept and organise and connect information about a concept in a graphical way. Usually the concept to be described is in the center of the mind map and the information and descriptions, as well as related concepts are arranged around the central concept and connected with lines.

Again, this approach was used to maintain the flexibility of the ontological representation. The ontology is expressed in the form of diagrams where functions and activities are represented as classes. The classes are defined and specified by the statements. These are connected through the predicates as outlined before. In this way, a graphical representation of the network of relationships was achieved, that visualizes the complex functional perspective on archives and records management.

Currently, the research team develops a web-based dissemination of the ontology. The diagrams seen on the following pages are the result of this development.

Results and Findings

The ontology consists of an overview-map (Fig. 04) featuring an array of nine main recordkeeping functions and 105 sub-maps, representing 105 (sub-)functions in total. These sub-maps contain the statements created in the analysis of the sources.

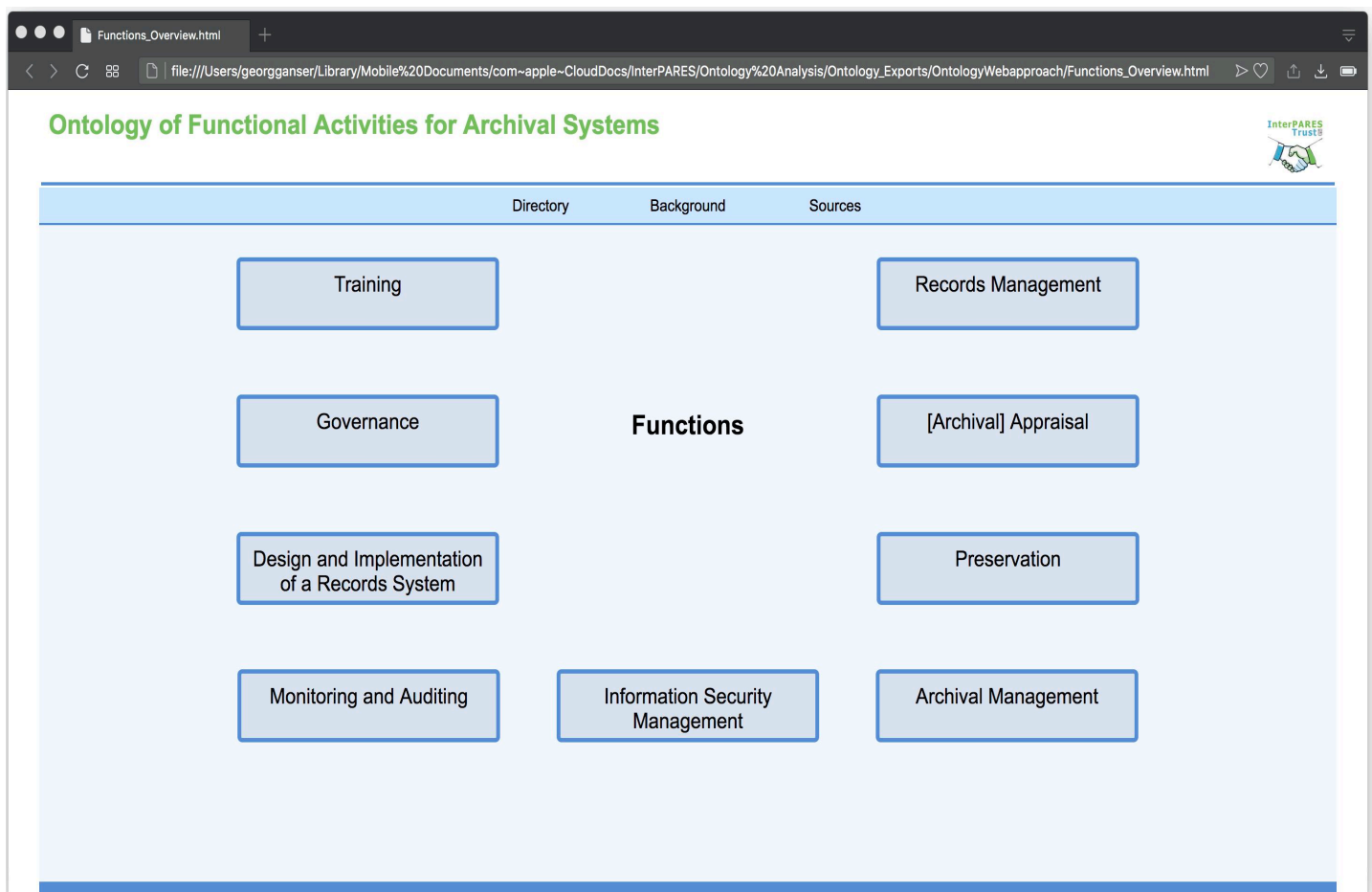


Fig. 04 – The Overview-map of the ontology

Relations in the form of sub-functions are derived from the statements that result in visually represented relationships and links to the corresponding sub-maps. Beside these sub-function-relationships the ontology employs “Related to”-relationships, that are derived from the statements, too. This type of relationship is used when a function is mentioned as being connected to another function in a statement (Fig. 05).

A sub-map can have multiple relationships with other sub-maps – these are explicitly derived from the utilized statements. In addition, the ontology makes use of another “Related to”-relationship, which is represented by a connection with a dashed box and green font (Fig. 05). This means that this relationship is an implication by the authors made in the course of the analysis of the sources. These relationships are not explicitly found in statements, but can still be derived from the analysis. The graphical representation of the ontology is a multilevel representation of the functions involved in recordkeeping.

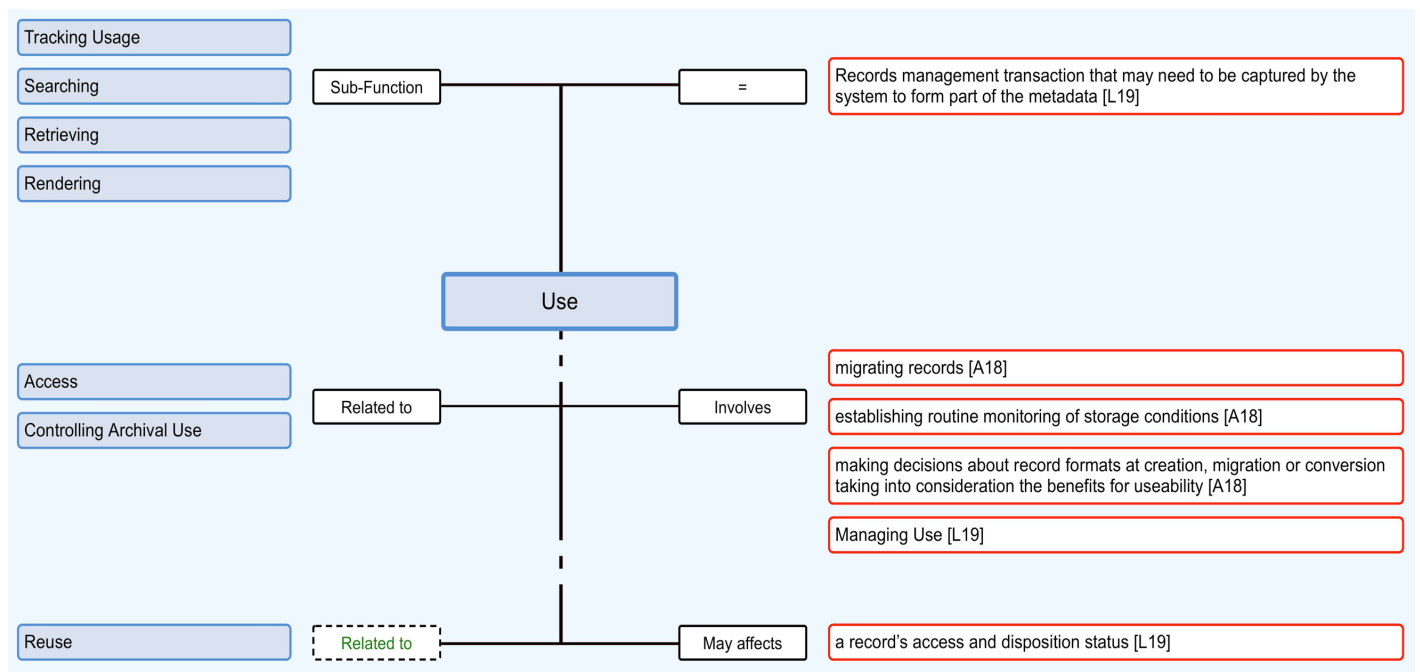


Fig. 05 – Ontology-diagram for “Use” with three relationship-types (extracted version)

A (Sub-)Function can be a Sub-Function of several other (Sub-)Functions. This represents the multiple relationships between functions of all levels.

This is due to certain inconsistencies of the sources.⁸ The ontology does not aim at resolving these inconsistencies or incoherencies but to making them visible. The ontology tries to represent all approaches to recordkeeping from the analyzed sources, to provide an overall image of the functions and activities and their (possible) relationships. Due to the graphical representation in the form of the mind-maps, these relationships and connections are clearly visible and understandable. It shows the interrelations and connections between records management and archival functions and the need for cooperative and integrated approaches.

The ontology highlights redundancies in some areas, whereas other functions – especially core archival functions – are underrepresented in standardization. This is probably due to the standardization procedures which are sometimes described as “controversial” and due to the promotion of certain interests, such as the continuum model or the notion of appraising which records need to be created in ISO 15489 (Records management).⁹ In addition, most archival standards focus on either description or

⁸ See e.g. J. Furner and A. Gilliland, “Archival IR. Applying and Adapting Information Retrieval Approaches in Archives and Recordkeeping Research”, in *Research in the Archival Multiverse*, ed. Gilliland A., McKemmish, S., and Lau A. (Melbourne: Monash Publishing, 2017), 595-598, on differences in the terminology used in international archival and records management standards.

⁹ See Gillian Oliver, “International records management standards: the challenges of achieving consensus”, *Records Management Journal*, 24, no. 1 (2014): 22-13; See also Giovanni Michetti, “Unneutrality of archival standards and

digital preservation – other archival core functions¹⁰ seem to be less standardized. Most standards deal with aspects of records management and (digital) preservation.¹¹ Archival functions such as Arrangement are solely covered by the SAA's Core Archival Functions whereas Outreach is not covered by any standard. Hoffman sees "two main clusters of archival standards" – "archival description standards", such as EAD or ISAD(G) and "digital preservation standards", such as OAIS.¹²

Finally, the ontology covers a wide range of functions and activities showing how the profession develops and adapts to arising challenges, especially in digital and networked environments. Functions like Governance and Information and Security Management indicate emerging responsibilities of records professionals.

Graphical Representation

Most of the analyzed standards deal with the current phase of a records' life. Therefore, the diagram for Records Management (Fig. 06) is the densest diagram of the ontology. This diagram contains the largest number of statements and thus relationships to other sub-functions. The sub-functions of Records Management are derived from the detailed statements, and feature classic functions and activities such as Capture, Registration, Classification, Access Management and Disposition.¹³

However, due to the changing business environments – especially considering requirements for digital records and digitalization¹⁴ in general – functions like Metadata Management and Migration are included in the Records Management's sub-functions as well. This development can be observed on all levels throughout the diagrams of the ontology. Many statements represented in the ontology deal with specific issues of the digital environment – they reflect the technological shifts in the records professions.¹⁵

processes", in *Re:inventing Information Science in the Networked Society. Proceedings of the 14th International Symposium on Information Science (ISI 2015). Zadar, Croatia, 19th-21th May 2015*, ed. F. Pehar, C. Schlögl, C. Wolff (Glückstadt: Verlag Werner Hülsbusch, 2015), 144-159.

¹⁰ Duranti and Michetti list "Appraisal and Acquisition, Arrangement and Description, Retention and Preservation, Management and Administration, and Reference and Access" as archival core functions, see Luciana Duranti and Giovanni Michetti (2017), "The Archival Method", in *Research in the Archival Multiverse*, ed. Gilliland A., McKemmish, S., and Lau A. (Melbourne: Monash Publishing, 2017), 77.

¹¹ See Hans Hoffman, "Archival Standards", in *Encyclopedia of Archival Science*, ed. Duranti L. and P. C. Franks (Rowman and Littlefield, 2015), 87-89.

¹² Hoffmann, *Archival Standards*, 87.

¹³ See e.g. Franks, P., *Records and Information Management* (American Library Association, Neal-Schuman, Chicago, IL, 2013) or Elizabeth Shepherd and Geoffrey Yeo, *Managing Records: A Handbook of Principles and Practice*. (Facet, London, UK, 2003).

¹⁴ Digitalization describes the shift from "analogue" to "digital" – "digital workplace".

¹⁵ See e.g. Joseph, Debowski, and Goldschmidt, *Paradigm shifts in recordkeeping responsibilities*, 58 and 69-70.

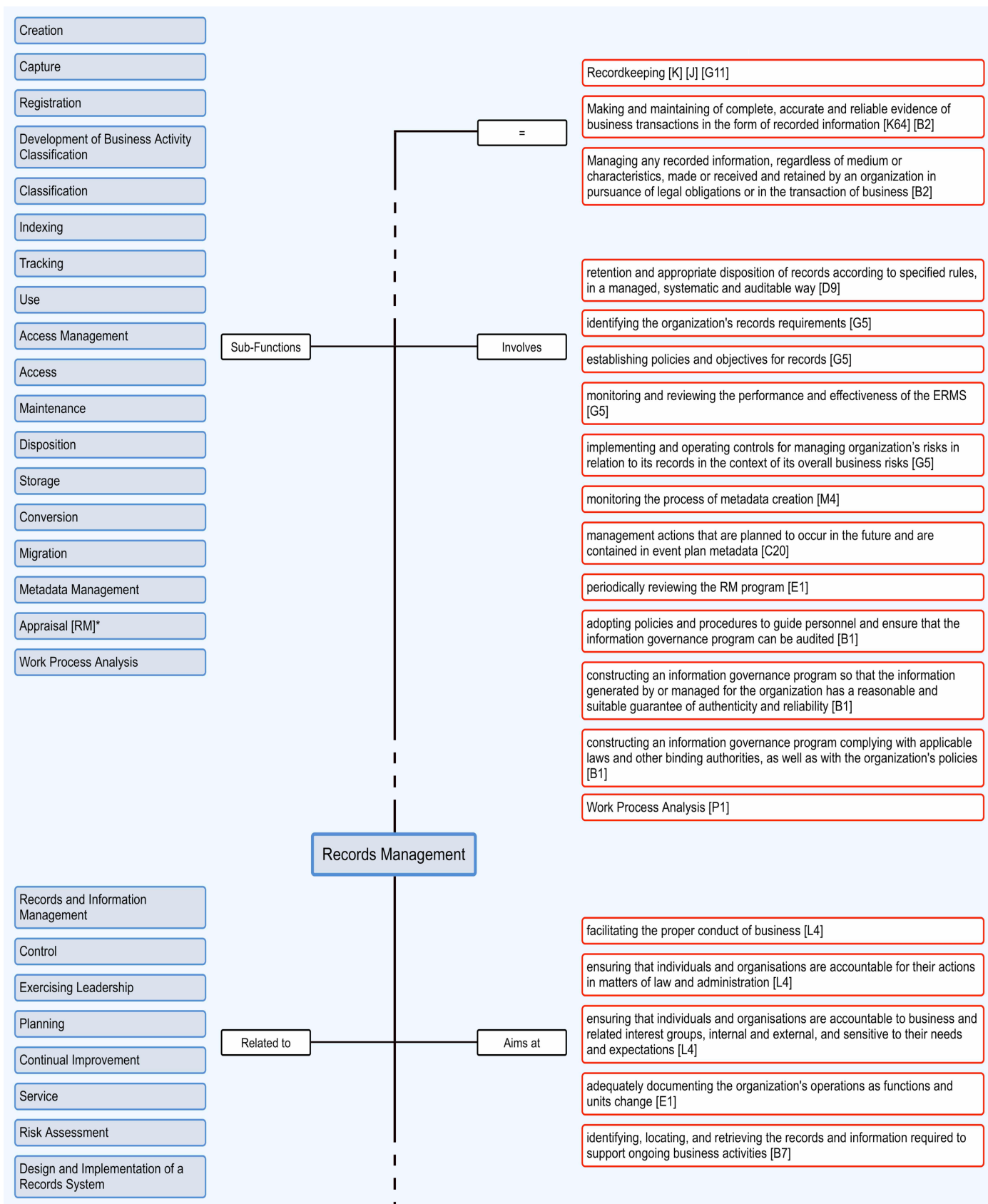


Fig. 06 – Ontology-diagram for “Records Management” (extracted version)

The ontological representation clearly shows that Records Management is not a stand-alone endeavour. This is reflected by many statements and related functions such as Control, Risk Assessment, and Design and Implementation of a Records System.

The Records Management function is related to some of the other high-level functions either directly or through one or more sub-functions. For example, the sub-function Access (Fig. 07) is shared by the main functions Records Management, Archival Management, and Preservation—this makes sense, because Access, along with its associated activities and tasks, assumes a relevant role throughout the whole records' life.

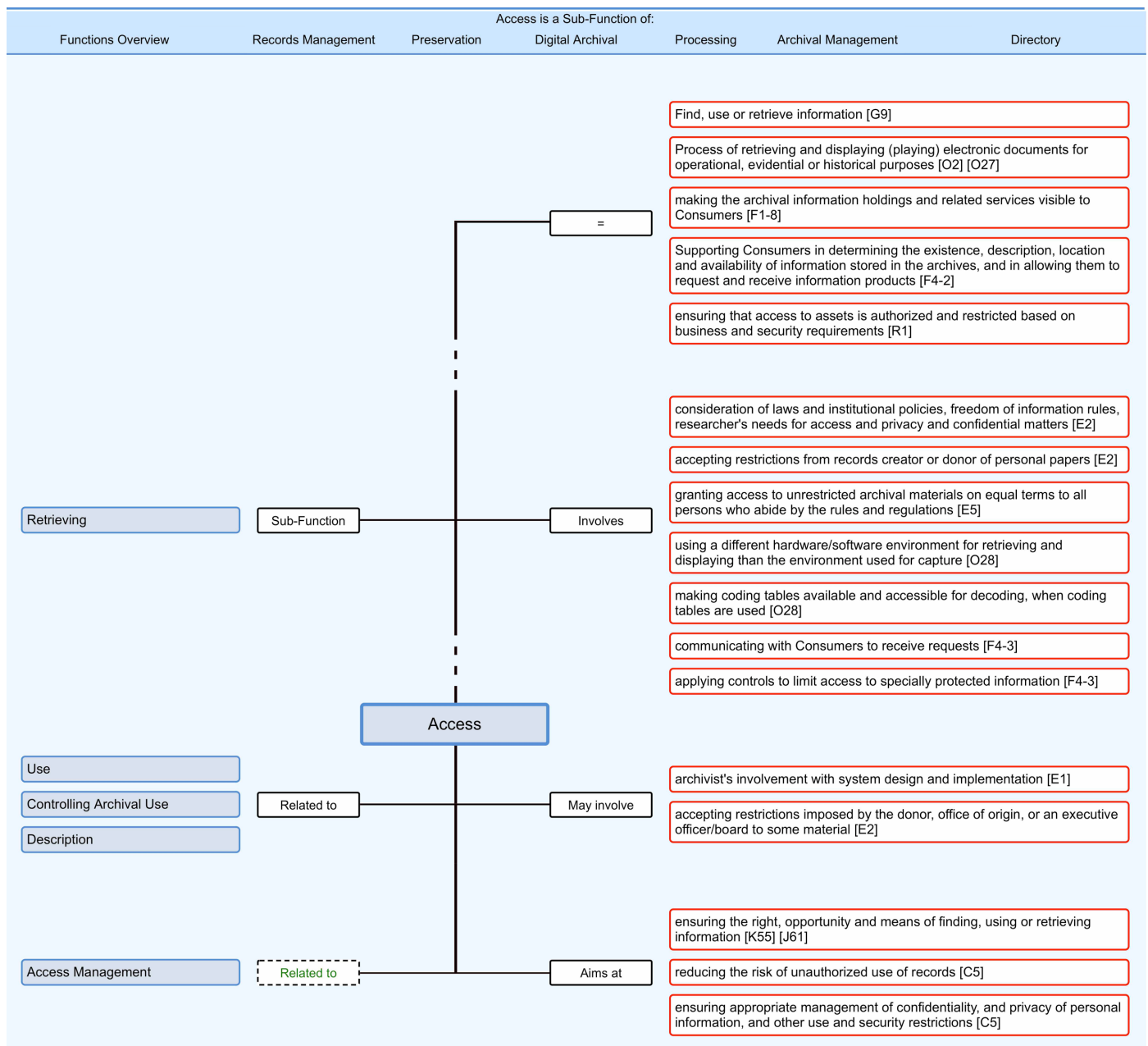


Fig. 07 – Ontology-diagram for “Access” (extracted version)

Therefore, the diagrams in this ontology can be used to point out shared responsibilities of all records professionals whether in the records management phase or in the archives. The diagrams and the related (sub-)functions contain several links to Information Governance and aspects of Information Governance programs, such as enterprise architecture and Information and Security Management. The Information and Security Management function in combination with the high-level functions Governance and Records Management, their relationships and connections found in the international standards, indicate the rising significance of Information Governance for the profession.¹⁶

Conclusion

An ontological representation of the functional activities provides records professionals, stakeholders and clients with a clear image of the functions and activities that need to be implemented and performed in records management and archives. By making inconsistencies and redundancies in the existing standards visible the ontology shows a certain need for systematization and precision in standardization. Furthermore, the ontology might prove useful as a guideline for software and system development, as well, as it visualizes connections and interrelations of functions. In addition, it can be used to clarify the responsibilities of archivists and records managers, to support their role in ensuring transparency of records processes, and to promote the active and direct involvement of records professionals in digital preservation. The ontology can be used to facilitate a better understanding of records management and archival processes and it can help in the drafting of procedures and policies. Finally, the ontology can improve the understanding of connections to other professions in the information management field. It shows that recordkeeping is a continuous effort that needs collaboration and interoperability of functions. There is indeed a tendency to “blur the distinction” between records management standards and archival standards. Standardization is moving towards developing “integrated standards that regard records during their entire existence, from their creation through use, retention, and preservation, to their final disposition.”¹⁷ The relationships represented in the ontology show that the functions and activities involved in recordkeeping are not isolated endeavours and that their nature is continuous.¹⁸

The ontology shows a certain need for records professionals to fill standardization-gaps in Information Governance and to get involved in this field, since Information Governance has great influence

¹⁶ See Smallwood, *Information Governance* and Franks, *Records and Information Management*, 311 and 329-330.

¹⁷ Hoffmann, *Archival Standards*, 87.

¹⁸ See Frank Upward, “Records Continuum”, in *Encyclopedia of Archival Science*, ed. Duranti L. and P. C. Franks (Rowman and Littlefield, 2015), 334-338.

on recordkeeping.¹⁹ The ontological representation of recordkeeping functions as they appear in the most relevant standards clearly indicates the growing importance of Information Governance and the need for a broadened view of the responsibilities of the records profession.²⁰ Instead of focusing on one stage in the record's life-cycle archivists and records managers have to develop a holistic understanding of their professions and embrace cooperation and overlapping responsibilities.²¹ The profession needs to see the management of records and information throughout the life-cycle as a common aim, that requires collaboration. The ontology clearly illustrates the connections and dependencies between records management and archival functions.

The ontology can serve as a useful tool for records professionals to support the cooperation within the records profession, their involvement with other information professions and to guide holistic approaches for the recent developments of the profession towards digitalization, increasingly networked environments, and cloud services. The ontological representation of records management and archival functions shows stakeholders and records professionals that our profession is well equipped with standardized means and methods and that the responsibilities span indeed to digital and networked environments and records in clouds as well. The author believes that a graphical representation of the functions and activities of records professionals can facilitate these aims.

Used sources:

- ISO 15489-1:2016 Information and documentation – Records management – Part 1: Concepts and principles (International Standards Organization, 2016);
- ISO/TR 15489-2:2001 Information and documentation—Records management—Part 2: Guidelines. (International Organization for Standardization, 2001);
- Generally Accepted Recordkeeping Principles. (ARMA International, 2014);
- ISO 23081-1:2006 Information and documentation—Records management processes—Metadata for records—Part 1: Principles. (International Organization for Standardization, 2006);

¹⁹ So far only the Australian Information Management Standard, that covers some aspects of Information Governance and the ARMA Information Governance Maturity Model are concerned with Information Governance. Both standards were not considered for the ontology, since these publications cannot be considered as widely in use, yet. A good example that illustrates the influence of Information Governance on recordkeeping is the Information Governance Reference Model. See "Information Governance Reference Model" 2018, accessed July 15, 2018, <https://www.edrm.net/frameworks-and-standards/information-governance-reference-model/>.

²⁰ See Smallwood, *Information Governance* and Franks, *Records and Information Management*, 329-330; See also the information about ARMA's Information Governance Professional (IGP) certification, "IGP CERTIFICATION" 2017, accessed July 15, 2018, <https://www.arma.org/page/Certifications>.

²¹ See also Franks, *Records and Information Management*, 311.

- ISO 23081-2:2007 Information and documentation—Records Management Processes—Metadata for Records—Part 2: Conceptual and implementation issues. (International Standards Organization, 2007);
- ISO 16175-1:2010 Information and documentation—Principles and functional requirements for records in electronic office environments—Part 1: Overview and statement of principles. (International Organization for Standardization, 2010);
- ISO 16175-2:2011 Information and documentation—Principles and functional requirements for records in electronic office environments—Part 2: Guidelines and functional requirements for digital records management systems. (International Organization for Standardization, 2011);
- ISO 16175-3:2010 Information and documentation—Principles and functional requirements for records in electronic office environments—Part 3: Guidelines and functional requirements for records in business systems. (Organization for Standardization, 2010);
- Core Archival Functions. (Society of American Archivists, 2016);
- ISO 14721:2012 Space Data and Information Transfer Systems—Open Archival Information System (OAIS)—Reference Model. (International Organization for Standardization, 2012);
- ISO 30300:2011 Information and documentation—Management systems for records—Fundamentals and vocabulary. (International Organization for Standardization, 2011);
- ISO 30301:2011 Information and documentation—Management systems for records—Requirements. (International Organization for Standardization, 2011);
- ISO 30302:2011 Information and documentation—Management systems for records—Guidelines for implementation. (International Organization for Standardization, 2015);
- ISO 16363:2012 Space data and information transfer systems—Audit and certification of trustworthy digital repositories. (International Organization for Standardization, 2013);
- ISO 14641-1:2012 Electronic archiving—Part 1: Specifications concerning the design and the operation of an information system for electronic information preservation. (International Organization for Standardization, 2012);
- ISO/TR 26122:2008 Information and documentation—Work process analysis for records. (International Organization for Standardization, 2008);
- [ISO/TR 15801:2017 Document management—Electronically stored information—Recommendations for trustworthiness and reliability. (International Organization for Standardization, 2017)];

- ISO/IEC 27000:2016 Information technology—Security techniques Information security management systems—Overview and vocabulary. (International Organization for Standardization, 2016);
- ISO/IEC 27001:2013 Information technology—Security techniques—Information security management systems—Requirements. (International Organization for Standardization, 2013);
- ISO/IEC 27002:2013 Information technology—Security techniques—Code of practice for information security controls. (International Organization for Standardization, 2013).

References

- Cottin, M. et al. “Méthodes techniques et outils”, *Documentaliste-Sciences de l'Information* 48, no. 4 (2011): 10-15;
- Duranti, Luciana and Michetti, Giovanni. “The Archival Method”, in *Research in the Archival Multiverse*, ed. Gilliland A., McKemmish, S., and Lau A. (Melbourne: Monash Publishing, 2017), 75-95;
- Franks, P. *Records and Information Management*. American Library Association, Neal-Schuman, Chicago, IL, 2013;
- Frost, R. “New ISO management system standards for records facilitate transparency in corporate governance” 2011. Accessed February 6, 2018, <https://www.iso.org/news/2011/11/Ref1487.html>;
- Furner, J. and Gilliland, A. “Archival IR. Applying and Adapting Information Retrieval Approaches in Archives and Recordkeeping Research”, in *Research in the Archival Multiverse*, ed. Gilliland A., McKemmish, S., and Lau A. (Melbourne: Monash Publishing, 2017), 580-631;
- Hoffmann, Hans. “Archival Standards”, in *Encyclopedia of Archival Science*, ed. Duranti L. and P. C. Franks (Rowman and Littlefield, 2015), 86-89;
- Joseph, P., Debowski, S. and Goldschmidt, P. “Paradigm shifts in recordkeeping responsibilities: implications for ISO 15489's implementation”, *Records Management Journal* 22, no. 1, (2012): 57-75;
- InterPARES 3 Project, TEAM Canada. “General Study 04 – International Standards Relevant to the InterPARES 3 Project” Vancouver: InterPARES 3 Project (2012). Accessed February 6, 2018, www.interpares.org/ip3/display_file.cfm?docip3_canada_gs04_international_standards.pdf;
- Katuu, Shadrack. “Managing digital records in a global environment: A review of the landscape of international standards and good practice guidelines”, *The Electronic Library* 34, no. 5 (2016): 869-894;

- Michetti, Giovanni. "Unneutrality of archival standards and processes", in *Re:inventing Information Science in the Networked Society. Proceedings of the 14th International Symposium on Information Science (ISI 2015). Zadar, Croatia, 19th-21th May 2015*, ed. F. Pehar, C. Schlögl, C. Wolff (Glückstadt: Verlag Werner Hülsbusch, 2015), 144-159;
- Oliver, Gillian. "International records management standards: the challenges of achieving consensus", *Records Management Journal*, 24, no. 1 (2014): 22-31;
- Pember, M. "Sorting out the standards: what every records and information professional should know", *Records Management Journal* 16, no. 1 (2006): 21-33;
- Shepherd, Elizabeth and Yeo, Geoffrey. *Managing Records: A Handbook of Principles and Practice*. Facet, London, UK, 2003;
- Smallwood, Robert. *Information Governance: Concepts, Strategies and Best Practices*. Hoboken, NJ: John Wiley and Sons, 2014.
- "IGP CERTIFICATION" 2017. Accessed February 6, 2018, <https://www.arma.org/page/Certifications>.