Toward an International Conceptual Model for Archival Description: A Preliminary Report from the International Council on Archives' Experts Group on Archival Description

Gretchen Gueguen, Vitor Manoel Marques da Fonseca, Daniel V. Pitti, and Claire Sibille-de Grimoüard

ABSTRACT

The International Council on Archives (ICA) appointed the Experts Group on Archival Description (EGAD) in late 2012. ICA charged this group with developing a conceptual model for archival description that will reconcile and integrate the four existing ICA descriptive standards. The EGAD will build on more than twenty years of ICA standards development, national or project-based modeling work in the archival community, and the modeling work of allied professional communities. This work will have as its core objective developing a conceptual model that reflects an international professional consensus and that positions the archival community to take full advantage of opportunities presented by current and emerging communication technologies, including the opportunities to work cooperatively within and without the archival community in a shared quest to provide enhanced access to and understanding of the human record.

KEY WORDS

Archival Records, Description, Technology

In 2012, the International Council on Archives (ICA) formed the Experts Group on Archival Description (EGAD) as the partial successor to the Committee on Best Practices and Standards (CBPS). As with its predecessor, ICA charged EGAD with developing standards for the description of records based on archival principles. For the 2012–2016 term, the EGAD is specifically charged with developing a comprehensive descriptive standard that reconciles, integrates, and builds on the four existing standards: General International Standard Archival Description (ISAD(G)); International Standard Archival Authority Records—Corporate Bodies, Persons, and Families (ISAAR(CPF)); International Standard Description of Functions (ISDF); and International Standard Description of Institutions with Archival Holdings (ISDIAH). The ICA further mandates that the EGAD develop the standard using "conceptual modeling" methods.

Understanding of archival description has continued to evolve since ICA formally embraced standards development as part of its mission in 1989. To keep current with the changing understanding and the ongoing transformation of communication technologies that provide the foundation for archival description systems, ICA has periodically reviewed the standards (see Table 1). Of the four standards, only the first two, ISAD(G) and ISAAR(CPF) have been reviewed and revised, resulting in second editions of each. Both ISDF and ISDIAH were developed recently and would only now be considered for review.

At the beginning of the development of what became four descriptive standards, ICA released the *Statement of Principles Regarding Archival Description*.[1] In section 1.3, the rationales for descriptive standards are given, including ensuring consistent, relevant, and explicit descriptions; facilitating the retrieval and exchange of information about archival materials; and making it possible to integrate descriptions from different repositories into a unified information system. These objectives endure.

To be effective in realizing the objectives, though, the standards need to be widely embraced and employed by the archival community. Unfortunately, of the four standards, only ISAD(G) has been widely accepted and used in the development of archival descriptive systems. In large part this is because ISAD(G) reflects the long-standing and still predominant descriptive practice of using a *single*, record-focused, provenance-based description that includes all of the various facets of description. The emergence of three additional standards that focus on components of the description found in ISAD(G) encourages the development of systems in which the components of descriptions are separated and interrelated to form complete archival description as it has traditionally been understood and to enable producing new and potentially powerful perspectives that facilitate the use and understanding of records. The ICA standards, however, do not provide guidance on how the four standards can be applied together in such systems. This has hindered understanding of the standards as

well as the economic and functional benefits of developing systems in which the major components of description are maintained separately.

The objectives of the EGAD therefore will be to develop a conceptual model for archival description that will address the shortcomings of the existing ICA standards and to provide guidance for developing archival description systems that both respect traditional understandings of description and pave the way for creating new paths into and novel perspectives on records.

History of ICA Standards

In 1988, the National Archives of Canada, in cooperation with ICA, convened a meeting of experts in archival descriptions. One of the resolutions passed at this meeting called for ICA to establish a working group "consisting of archivists knowledgeable in descriptive theory and practice to develop international standards for the description of archives."[2] ICA responded to this resolution by organizing a meeting in Paris the following year. This set in motion activities that would, over the next twenty years, lead to the development of a *Statement of Principles Regarding Archival Description* and four descriptive standards: General International Standard Archival Description (ISAD(G); International Standard Archival Authority Records—Corporate Bodies, Persons, and Families (ISAAR(CPF)); International Standard Description of Functions (ISDF); and International Standard Description of Institutions with Archival Holdings (ISDIAH).

Table 1. Development of ICA Standards

Standard	Edition	Development Dates	Publication Date
Statement of Principles		(1988) 1989–1992	1992
ISAD	1st	1990–1993	1994
ISAAR	1st	1993–1995	1996
ISAD	2nd	1996–2000	1999
ISAAR	2nd	2000–2004	2004
ISDF	1st	2005–2007	2007
ISDIAH	1st	2005–2008	2008

The four standards were developed in succession. ISAD(G) and ISAAR(CPF) were released, and then revised, before ISDF and ISDIAH were developed. Though the individual archivists involved in the development work have changed since the first standard, considerable overlap in working group membership as the focus shifted from one standard to the next ensured continuity along the way.

It is clear that the archivists who met in Ottawa in 1988 did not envision the eventual creation of four standards. The prevailing descriptive model at the time was to use a single description for one archival *fonds*, beginning with the most general description and, in increasing specificity, the parts of the *fonds*, the parts of the parts, and so on. All attributes of the records in a *fonds*, and the context of the records' creation and use, were combined in one description, with the different descriptive elements interwoven to form a complete, comprehensive description. The description thus covered the records as such, as well as what we have come to understand over time as the record context: the creator, functions and activities, and holding repository. ISAD(G) embodies this understanding in both the first and second editions.

While combining the description elements together dominated ICA standards work at the beginning, it is clear that early on some of the archivists involved began to see advantages in separating particular components of the description. ISAD(G) itself suggested the possibility of maintaining "access points" using the methods of authority control (that is, maintaining them separately from the remainder of the description).[3] In this regard, mention is made in ISAD(G) of the intention to develop the standard that would be named ISAAR(CPF) and released two years after it.[4] The working group that developed ISAAR(CPF) was in fact appointed in 1993, before the release of ISAD(G) in 1994. This working group focused on separating the "access points" in the archival description from the remainder of the description. More specifically, it focused on the names of record creators: once separated, the record descriptions and the name descriptions could be *maintained* independently and *interrelated* as appropriate.[5] For users, the interrelated pieces of the description would be brought together to form a complete description.

Separating authority control from record description was modeled on the computer-based "linked authorities systems" in libraries that emerged in the 1980s. These systems were based on the authority files used in maintaining access points in library card catalogs. A central argument used for developing an archival authority record standard was the same used in the library community: it was more economical.[6] Archives holding parts or fragments of the same *fonds* could share the creator contextual information. Colonial, immigration, and trade records were cited as common instances where exchanging the descriptions of creators would be useful.

While the immediate inspiration for the separation of the names from the other elements of description was the "linked authority systems" in libraries, the argument for separation has a long history. Charles Jewett, the librarian of the Smithsonian Institution in the mid-nineteenth century, envisioned a universal catalog of the world's knowledge. The predominant form of the library catalog at that time was the book catalog, which was expensive to produce and

out of date by the time of publication. Jewett did not advocate for a new form of catalog, but proposed radically changing the way in which cataloging data were created and maintained. His methods would enable use and reuse of the descriptions to produce a wide variety of book catalogs for individual libraries and a "universal catalog" that included all libraries.[7] The core of Jewett's argument for standards and new methods was that maintaining key components of description separately would enable flexible use of the components in various combinations to produce a wide variety of products, each of which would provide a different perspective and support a different strategy in locating and using the described resources. Arguably, this approach continues to motivate innovation in the development of information description and access systems, including archival systems.

Though the ICA working group that developed ISAAR(CPF) was driven by library methods and systems, the argument for separation in archival description had been made more than thirty years earlier. Peter Scott, working with others at the Australian Commonwealth Archives Office, argued in 1966 for the end of the "record group concept." The primary motivation was that the monolithic *fonds*-level description made it difficult (if not impossible) to maintain the integrity of record series with multiple provenances, which many series had due to the frequent transfer of functional responsibility from one agency to another. Separating the description of the record creators from the record description and shifting the descriptive focus to the series enabled maintaining the integrity of a series, *without* abandoning the provenance as such.[8] Various American archivists resumed making Scott's argument some years later.[9] While the "series system" did not have an ostensible impact on ICA standards development, the ongoing separation of the description components increasingly reflected the descriptive methods advocated by Scott.[10]

The ICA standards appear to have been motivated by the broader and more persistent objective of separating components of description, so as to have a more economical and flexible method of creating and maintaining the data that comprises description. Once separated, the components can be exploited to create a wide variety of discovery and access systems that multiply the use perspectives, all the while maintaining the essential archival *respect des fonds.* Arguably, though, the single *fonds*-level description continued to dominate the imaginations of the archivists involved. While laying the foundation for separation, the standards failed to articulate an encompassing conceptual framework that would provide guidance on how the standards were intended to work together to form a complete system of description. It can be further argued that the lack of a conceptual framework has impeded the embrace of the power of separation, and, as a result, the international archival community has not

widely accepted and implemented systems that employ fully the ICA standards. Of the four standards, the only to be widely embraced is ISAD(G).

In 2008, in response to this "missing piece," ICA made the decision to develop a compendium of the four existing standards. Two documents, both released in July 2012, resulted from the subsequent deliberations: a common chapter on the interrelations of the four standards,[11] and a "Progress report for revising and harmonising ICA descriptive standards."[12] After analyzing the four standards, in particular looking for overlap, inconsistencies, and ways to reconcile them into a coherent whole, it was recommended that a "conceptual model" would first need to be developed and that this conceptual model would serve as a reference for revising the four standards into a complementary, consistent set.[13] Following this recommendation, ICA appointed the EGAD with the charge of developing a conceptual model, to be released in 2016.

Information Technologies Context

Resource description and access relies on communication technologies. As new research media and methods have emerged, the communities for whom description is a primary activity have sought to exploit the emerging technologies in reimagining description to make it more economical and to enhance discovery, access, use, and understanding, thereby more fully realizing core principles and objectives. For Jewett, the technology that made possible his reimagining of cataloging was stereotype printing. It was more economical and flexible, and provided the means to reuse data in multiple reformulations. Later, many of Jewett's ideas were to be realized via written and typed cards and card catalogs (which were then further used, much as Jewett imagined, to produce universal catalogs as printed books).

In 1986, two years before ICA began its standards development activities, two standards were released that have come to dominate the technology landscape: Structured Query Language (SQL: ANSI X3.135-1986; ISO/IEC 9075) and Standard Generalized Markup Language (SGML: ISO 8879:1986). Of the two, SQL had immediate widespread impact on government, business, academic, and other users. With the emergence of eXtensible Markup Language (XML: W3C) as a subset of SGML in 1998, the two technologies were both broadly supported and, when not in competition with one another, were used together in a variety of configurations.

The two technologies, database and markup, take different approaches to representation of data. Each is best suited to particular uses. Users thus have needed to analyze their data, to ascertain their features, while also bearing in mind how they want to use the data in order to decide which of the two technologies to use.

In general, databases efficiently support information that can be represented according to the following criteria: regular number of fields; order of the fields generally not significant; each field restricted to data; interrelated fields have a fixed or shallow hierarchy; and the data in each field are controlled with respect to form and structure. Information that is well accommodated in databases might be characterized as the data one finds in the forms that we must complete, for example, when applying for a driver's license. Database technologies excel not only in separating various data fields but also in reliably managing the interrelating of the fields, allowing users to bring them together in various configurations that serve a wide variety of perspectives and uses. This approach to data representation is frequently referred to as "data-centric."

Markup technologies, on the other hand, were intended to model traditional documents (texts of various sorts such as essays, letters, articles, books, and so on). In general, all of these types of data objects share many of the following features: irregular number of components (e.g., chapters or paragraphs), sequence or order is important (e.g., order of paragraphs matters); semiregular structure and unbounded hierarchy; arbitrary intermixing of data with markup; and arbitrary number of interrelations within and among documents. This approach to data has commonly been called "document-centric."

As powerful as the two technologies have been, much and perhaps most real world information is not represented well in either one or the other. Archival description, particularly in the single *fonds*-level description (its most widespread form), is adequately but not perfectly accommodated by database technologies in some parts and in other parts by markup technologies. That neither technology clearly dominates the archival implementation landscape reflects the "betwixt and between" nature of the traditional single description. Many description systems use one or the other or a carefully crafted combination of the two, with database technologies used for the creation and maintenance of the description and markup technologies used to communicate from machine to machine and from machine to end users. Technological developments within and between the two technologies ameliorate if not eliminate the weakness of each and thus help sustain their dominance over the representation landscape.

A year after XML was released, the W3C released the Resource Description Framework (RDF). RDF introduced a very different data representation model: graph technologies. Graph technologies introduce data representation as statements, typically characterized as subject-predicate-object, with each statement called a "triple." While XML supports a specific form of graphs, the hierarchy (or tree) triples enable unbounded representation of networks of interconnected data objects as well as real world objects (represented by data). Given that the real world within which we live and work may be understood as a vast, dynamically

interrelated network of people and objects situated in space and time, graph technologies offer new and more expressive forms of representation.

RDF has been combined with related standards and methods to formalize and implement semantic technologies. Semantic technologies propose to create a more semantically precise and logically structured network of interrelated information that can be computationally exploited to reveal new knowledge, new patterns, and thus an information universe that is more than the sum of its parts. Semantic technologies are increasingly used in various domains to share and interconnect data that are created and maintained in disparate contexts, and thus overcome the separation between different descriptive systems and the things they describe. An especially important initiative, the Semantic Web, grandly envisions the World Wide Web as a vast semantic network of interconnected statements about the world, objects in the world, and specific relations of the objects.[14] To realize this vision, data creators are encouraged to make data freely available on the Internet in a machine-readable form known as Linked Open Data (LOD) that can be used in systems other than the system in which the data are maintained. Many people, quite reasonably, doubt that this vision will ever be fully realized, but semantic technologies are increasingly demonstrating that grand if not grandiose objectives are within reach. In particular, the cultural heritage communities are increasingly sharing a vision, to paraphrase Jewett, of "the universal catalog of cultural heritage."

The cultural heritage communities that have long worked largely in isolation from one another, each pursuing its own descriptive practices and developing systems based on these practices, are now increasingly using semantic technologies to explore interconnecting disparate description systems and realize integrated access to cultural heritage. Europeana endeavors to provide union access to digital archives, library, and museum resources in Europe.[15] Based on the data model (or ontology) developed by Europeana, the Digital Public Library of America (DPLA) has embarked on a similar initiative.[16] The Library of Congress is making many of its controlled vocabularies available as LOD, with work underway to expose the bibliographic descriptions as well.[17] The OCLC Virtual International Authority File (VIAF) is making millions of authority file "clusters" available as LOD.[18] *Wikipedia* articles about people are systematically being linked to VIAF records for the same people, and a subset of *Wikipedia* is also being exposed as LOD in DBpedia.[19] OCLC is also experimenting with making a subset of WorldCat records available as LOD.[20] The Bibliothèque Nationale de France is using semantic technologies to provide unified Internet access to its different curatorial-based systems of description (including MARC- and EAD-based descriptions) and to expose the data as LOD.[21] The British Library has released the British National Bibliography (BNB) as LOD.[22]

While the archival community is participating in these and other initiatives and projects, it has been hindered (with notable exceptions) by assorted factors: descriptive practices less standardized than those of other communities, problems with the overall complexity of its descriptive practices (hierarchy and ordered lists being a particular challenge), and lack of a clear and comprehensive conceptual framework for archival description. Still, with the guidance provided by the ICA standards, in particular ISAD(G), the archival community has made great strides in standardizing descriptive practice.

Nevertheless, much remains to be done before the community can fully participate in and fully realize the opportunities that advanced and emergent technologies offer for making archival description significantly more effective in the discovery, use, and understanding of archival resources to better serve the communities that use archives and to attract new communities of users.

Conceptual Models

Conceptual modeling is a formal technique for representing the principal concepts and relations among them for a given knowledge domain. As different data representation methods (e.g., database, markup, and graph technologies) have emerged, specific methods for modeling the data that are to be represented and exploited in the systems have been developed. A popular modeling method that emerged for database technologies is the entity-relationship model (ER). With the emergence of object relational databases, the ER approach was adapted to reflect object-oriented data representation methods. With the emergence of graph-based semantic technologies, more modeling methods and tools have emerged to support them. The W3C, based on RDF, developed the Web Ontology Language (OWL) as a method for formal "knowledge representation" of domain-specific ontologies. Ontologies expressed in formal machine-readable forms can be processed by "reasoning" software that not only tests the logical integrity of the model, but also uses the ontology to inform processing of the data assembled in compliance with the ontology.

From 1992 to 1995, the International Federation of Library Associations (IFLA) Study Group on Functional Requirements for Bibliographic Records (FRBR) developed a conceptual model for bibliographic description. The Study Group employed ER conceptual modeling techniques in developing the model.[23] In 1996, the International Council of Museums (ICOM) International Committee for Documentation (CIDOC) began the development of a conceptual model for the description of museum objects.[24] The Conceptual Reference Model (CRM), though initially focused on museums, came to be conceived as a reference model that could serve the broader ambition of enabling integrated access to cultural heritage, thus encompassing archives, library, and museum access. In this

regard, the International Working Group on FRBR/CIDOC CRM Harmonisation was formed in 2003. The working group has focused on mapping FRBR concepts to CRM concepts and, where necessary, enhancing and refining CRM concepts to facilitate the mapping, thereby making the CRM a single, overarching semantic model. The extension of the CRM that accommodates the FRBR entity-relationship model into the objected-oriented CRM is called FRBRoo and was released in its first version in 2009.[25]

The archivists and museum specialists involved in the development of the CRM and the FRBRoo extension have expressed interest in working with the archival community to accommodate archival description and enable the model to fully incorporate the archives, library, and museum communities. While many in the archival community would very much like to collaborate in this endeavor, the lack of a conceptual model that comprehensively and accurately addresses the archives domain concepts (as was done with FRBR for the bibliographic domain) has been an impediment. Before collaborating, the archival community must develop a conceptual model that first and foremost addresses the principles and needs of archivists. Once this is accomplished, the community will be well positioned to join with allied professional communities in a common quest to provide integrated access to human expression in all of its forms.

Archival Community Conceptual Models

While development of an international archival conceptual model is just beginning, the EGAD's work will be greatly facilitated by one binational and two national modeling initiatives. Of the three initiatives, the most long-standing of the models is represented by the closely aligned *Australian Government Recordkeeping Metadata Standard Version 2.0* (2008) *(AGRkMS)* and the Archives New Zealand's *Technical Specifications for the Electronic Recordkeeping Metadata Standard Version 1.0* (2008).[26] Both are based on thorough revisions of the Australian *Recordkeeping Metadata Standard for Commonwealth Agencies Version 1.0* published in 1999.[27] In Spain, the Comisión de Normas Españolas de Descripción Archivística (CNEDA) began work in 2007 and published the *Modelo Conceptual de Descripción Archivística y Requisitos de Datos Básicos de las Descripciones de Documentos de Archivo, Agentes y Funciones* in 2012.[28] Recently, in Finland, the Arkistolaitos (National Archives) appointed a working group to develop a conceptual model, and the working group released Draft Version 0.1 of the *Finnish Conceptual Model for Archival Description*.[29] Also worthy of mention are efforts in the United Kingdom to develop an ontology based on the Encoded Archival Description (EAD) communication standard (LOCAH)[30] and work on the ICA-sponsored AtoM system that is developing a platform supporting the ICA descriptive standards and that enables LOD exposure of the data in compliance with existing ontologies.[31]

The *AGRkMS* is grounded in two ISO standards, the *Australian Standard on Records Management* (*AS ISO 15489*) and the *Metadata for Records* (*AS ISO 23081*) and the work of the Monash University SPIRT Research Team. The standard fully embraces a "multiple-entity" model of archival description, distinguished from the "single-entity" approach to archival description that has represented and continues to represent the most common approach to archival description. The multiple-entity approach separates and interrelates the major entities that comprise the single-entity approach. The model identifies five entities: *Record, Agent, Business, Mandate,* and *Relationship*. The Relationship entity plays the role of interrelating the other four entities to form a complete description. Relationship is event based, which is to say that relations document specific human activities or events. While the developers of the standard clearly encourage the separation of the entities, they carefully describe how the standard can be used in single-entity systems where the single entity is the record and the other entities are treated as attributes of the record, and other systems that only partially separate the entities, such as record and agent.

CNEDA employs ER data modeling techniques and takes into consideration the four ICA standards as well as the ICA-AtoM, various ISO standards (including 15489 and 23081 mentioned above), and many others, including the *AGRkMS*. As with the *AGRkMS*, CNEDA recognizes that the archival community is in transition, "from a one-dimensional archival description (focusing almost exclusively on the representations of records), to another multidimensional description, aimed at creating and maintenance [*sic*] of representations of entities of different type (records, agents, business, etc.) and their interrelationships."[32] The model identifies four entities that are the same as in the *AGRkMS*: *Record*, *Agent*, *Business*, and *Mandate*. It identifies two additional entities, *Concept, object, or event* and *Place*.[33] The relationships are not designated an entity, as in the *AGRkMS*, though they play much the same essential role.

The CNEDA model makes a clear distinction among three "dimensions" of the domain within which archivists work. First, there is the "real world," the world shared by us all, which can be conceptualized and understood from multiple perspectives, formal as well as informal. Second is a conceptual model of the real world based on a specific perspective, the archival perspective. This conceptual model of the world is represented in the CNEDA model. Third, there are standards that specify the data requirements of systems that represent real world phenomena based on the conceptual model.[34] Perhaps the latter may be characterized as "implementation guidelines."

Draft Version 0.1 of the *Finnish Conceptual Model for Archival Description* was released in January 2013 and is thus in the very earliest stage of development. Unlike the other two models, it is being developed within the context of integrated access to cultural heritage as represented in the National Digital Library

(hereafter FNDL). The FNDL recommends that all cultural heritage use *Resource Description and Access* (RDA) as a starting point.[35] RDA is a library standard based on FRBR and thus, by extension, is related also to FRBRoo and, by further extension, to CRM. The early draft of the Finnish model thus reflects the influence of RDA, FRBR, FRBRoo, and CRM. At the same time, the model is mindful of the ICA standards and places particular focus on the principle of provenance. Like the *AGRkMS* and the CNEDA model, the Finnish model advocates separation and interrelation of the core description entities. It recognizes *Function*, *Agent*, *Information Resource*, *LifeCycle Event*, *Mandate*, *Place*, *Temporal Event*, and *Subject*.[36] LifeCycle Event is described as Recordkeeping Business. Again, it overlaps substantially with the *AGRkMS* and the CNEDA model.

In addition to these national standards, the LOCAH project in the United Kingdom and Artefactual Systems, the developers of ICA-AtoM, have also developed models. A primary motivation for both is to enable exposing description on the Internet as LOD. Artefactual Systems has developed a model of archival description based on the existing ICA standards, though mapped to existing, widely used, if not specifically archival, ontologies. The AtoM model prominently features archival materials (*Records*), *Agents*, *Archival Agency*, and *Events*. The LOCAH project has taken a different approach by focusing on EAD as the basis for the model. This is motivated by the desire to expose existing description represented in large aggregations of EAD instances (such as ArchivesHub) as LOD. The LOCAH model includes *Record*, *Agent*, *Event*, *Time*, *Place*, and *Subject* entities. [37] It is similar in approach to models developed in the library community to expose existing authority and cataloging records (for example, the Library of Congress Linked Data Service and the emerging BIBFRAME model).

The ICA Experts Group on Archival Description

The ICA Programme Commission formed the Experts Group on Archival Description late in 2012. The EGAD is charged with the harmonization of the four existing ICA standards, ISAD(G), ISAAR(CPF), ISDF, and ISDIAH, based on a formal archival description conceptual model. The EGAD's members are drawn from the international professional community and have demonstrated expertise in archival description and standards. There are twelve full members and an additional nine corresponding members (full members are required to attend EGAD meetings but corresponding members are not). All twenty-one members will be responsible for developing the conceptual model and related documentation, and all are also responsible for representing, informing, and gathering input from their respective communities.

The work of the EGAD began in late 2012 and will be completed in 2016. Over the course of the development, the EGAD will consult widely with the

international archival community and related professional organizations. As milestones in development are reached, the EGAD will disseminate drafts of the model and documentation and gather community input. The EGAD will endeavor to ensure that the resulting international standard reflects a community-wide consensus and that it can be applied in all cultures, languages, and scripts.

The methodology for developing the archival conceptual model will initially focus on defining the scope (or domain) of the model. The model must first and foremost address the principles and needs of the archival community. In particular, the model will be grounded in the enduring principle of provenance. The model will focus on the separation and interrelation of the primary components of archival description to provide a foundation for the development of archival description systems that will support a variety of perspectives on archival resources, including both the perspective represented in traditional *fonds*-level finding aids and guides, and the perspective represented in the series system first advocated by Peter Scott in Australia. Systems based on the model will support multiple paths into archival resources, making it possible, for example, to find all bodies that have performed a specific function, all records created by the same body, and all records documenting the same function, and each of these in relation to specific times and places. Our understanding of archival description has evolved, as have the information technologies that are needed to represent and communicate it, and it is clear that an international standard based firmly on the principle of provenance while at the same time supporting different emphases and approaches to the description of the records is within reach.

Other professional communities will be interested in the model, in particular the library, museum, and scholarly research communities and more generally the interdisciplinary community that is interested in semantic technologies, particularly as applied to integrated cultural heritage access initiatives such as Europeana and DPLA. Thus, the scope of the model will take into consideration the interests of these communities, as well as interest within the archival community for cooperation with these other communities on a shared and expansive vision of description and access to cultural resources.

In addition to the four ICA standards, the work of the EGAD will be greatly facilitated by the substantial work that has been done, or is underway, by members of the archival community. The *AGRkMS* and the closely aligned New Zealand model, the CNEDA model, and the model just underway in Finland will be particularly important. While the conceptual landscapes articulated in each of these models overlap substantially, there are also substantial differences in approach and in the conceptualizations themselves. The differences may only be a matter of terminology, but in some cases the differences are more substantive

and will call for goodwill and flexibility. Taking advantage of the work already accomplished, once the scope of the ICA model is defined, attention will shift to comparing these three models, looking for common ground, differentiating apparent conflicts from substantial ones, and reconciling terminology and understandings.

The first and perhaps most challenging task will be to identify high-level entities: records, agents, business, and so on. Focus will then subsequently move to identifying each component's attributes and, critically, the relations among them that enable variously assembling the components to support different descriptive perspectives. From this development process, it is anticipated that a conceptual model that reflects international archival consensus will reposition archival description to more effectively and efficiently identify, authenticate, and manage records through the description of their context, content, and structure. Furthermore, the description will enhance user discovery, location, and understanding of records.

While the focus will initially be on an archival conceptual understanding of the world, both the library and museum conceptual models will be consulted, possibly along with other conceptual models that come to the attention of the EGAD as the work unfolds. Considerable overlap exists in the descriptive practices of the allied cultural heritage communities, as all have in common a focus on human activity and the artifacts (broadly defined) this activity produces. Both the library and museum communities are much further along in developing international standard conceptual models, and the EGAD can greatly benefit from these allied standards. Further, such consulting will help ensure that archivists will have a firm foundation for collaborating with the allied professions in the integration of access.

In 2016, the final archival conceptual model will be released as a formal document, including text and diagrams, as well as be formally represented in OWL. Using OWL to represent the model will enable using development and testing tools that ensure the model is formally correct. It will also facilitate its use immediately in a variety of semantic technologies, mapping to FRBRoo and CRM, and exposing archival description using LOD techniques. Drafts of the documentation and formal OWL model will be released for testing and comment before final publication. Once revision is completed, the model will be published freely online and will be disseminated before the ICA Eighteenth International Congress.

Conclusion

Since at least the middle of the nineteenth century, the methods used in resource description have progressively emphasized separating and interrelating

key components of description to accommodate the production of familiar and proven modes of access and at the same time open new paths into and perspectives on described resources. The ongoing emergence of new communication technologies has provided the foundation for this recurring re-envisioning of resource description. The two interdependent motivations for the separation have remained constant: *improving the economy and accuracy of description* and *enhancing access to and understanding of the described resources.*

The emergence of the series system in Australia in the 1960s, the much later development of the four ICA standards, and the descriptive models that have emerged in Australia, Spain, and Finland reflect this trend in the archival community. The dramatic emergence of computer networks and computer-based data representation technologies in the last two decades of the twentieth century and continuing into the twenty-first century has transformed and is transforming communication. These advancements present new opportunities for re-envisioning (and modeling) archival description and also for lowering the technological challenge of realizing the opportunities. The four ICA standards were successively developed in parallel with the emerging technologies, though they only have faintly reflected the possibilities presented by them. The ICA standards, in particular ISAD(G), have helped promote and ensure consistency in descriptive practice, and have thereby facilitated the development of many significant multi-institutional collaborations. Despite the significant impact of the standards, they have failed to provide adequate guidance for developing the next generation of archival description systems.

For a variety of reasons, now is an excellent time for archival experts from around the world to consider all of the ICA archival descriptive standards and the exceptional modeling work completed or underway in the archival community and among allied professional communities. All the ongoing initiatives within the archival community for developing conceptual models at a national level or in the framework of specific projects demonstrate that the profession continues to explore its principles and re-imagine its practices as new communication technologies present unprecedented opportunities to more effectively fulfill the archival mission. Local and national initiatives highlight the need for an international standard archival conceptual model for archival description. Based on a review of the ICA standards and the archival and allied professional modeling work, the EGAD will develop a model that serves the professional principles and objectives of the international archival community, providing guidance for developing archival descriptive systems and a solid intellectual and technological foundation for collaborations within and outside the community.

Notes

1   Ad Hoc Commission on Descriptive Standards, *Statement of Principles Regarding Archival Description,* (Ottawa: International Council on Archives, 1988), 6.

2   Ad Hoc Commission on Descriptive Standards, *Statement of Principles Regarding Archival Description*, 1.

3   Ad Hoc Commission on Descriptive Standards, *ISAD(G): General International Standard Archival Description* (Ottawa: International Council on Archives, 1994), 5.

4   Ad Hoc Commission on Descriptive Standards, *ISAD(G)*, 3.

5   Committee on Descriptive Standards, *ISAAR (CPF): International Standard Archival Authority Record for Corporate Bodies, Persons and Families* (Ottawa: International Council on Archives, 1994), 7.

6   Committee on Descriptive Standards, *ISAAR (CPF)*, 7.

7   Charles C. Jewett, *On the Construction of Catalogues of Libraries and Their Publication by Means of Separate, Stereotyped Titles* (Washington, D.C.: Smithsonian Institution, 1853), 23. "The headings (if they be names) are to stand on plates distinct from the titles [descriptions of the books]. This is required, in order to avoid repeating them for each title. They must be separate from the titles, that other titles may, if occasion require, be interposed." Jewett's argument for separation included more than names and titles.

8   Peter Scott, "The Record Group Concept: A Case for Abandonment," *The American Archivist* 29 (October 1966): 493–504

9   In the 1980s, several articles appeared in support of separation: Richard H. Lytle, "Intellectual Access to Archives," *The American Archivist* 43 (Winter/Spring 1980); Lytle and David A. Bearman, "The Power of the Principle of Provenance," *Archivaria* 21 (Winter 1985–1986); Max J. Evans, "Authority Control: An Alternative to the Record Group Concept," *The American Archivist* 49 (Summer 1986): 240–61; Bearman and Richard Szary, "Beyond Authorized Headings, Authorities as Reference Files in a Multi-disciplinary Setting," in *Occasional Papers of the Art Library Society of North America*, no. 6 (Tucson, Ariz.: Art Library Society of North America, 1987).

10   While the impact of the Australian "series system" is not explicit in the ICA standards, the ongoing participation of Australian archivists in the development of the standards since 1993, when Chris Hurley was appointed to the Ad Hoc Commission on Descriptive Standards, has no doubt had significant influence.

11   Committee on Best Practices and Standards, *Relationship in Archival Descriptive Systems* July 4, 2012), ica.org, "Standards," http://www.ica.org/13149/standards/cbps-relationship-in-archival-descriptive-systems.html.

12   Committee on Best Practices and Standards, *Progress Report for Revising and Harmonizing ICA Descriptive Standards* (July 4, 2012), ica.org., "Standards," http://www.ica.org/13155/standards/cbps-progress-report-for-revising-and-harmonising-ica-descriptive-standards.html.

13   Committee on Best Practices and Standards, *Progress Report for Revising and Harmonizing ICA Descriptive Standards*, 15.

14   W3C, "Semantic Web," http://www.w3.org/standards/semanticweb/.

15   Europeana, "About Us," http://www.europeana.eu/portal/aboutus.html.

16   Digital Public Library of America, "Metadata Application Profile," http://dp.la/info/map/.

17   Library of Congress Linked Data Service, "About," http://id.loc.gov/about/.

18   OCLC, "VIAF," http://www.oclc.org/viaf.en.html.

19   DBpedia, "About," http://dbpedia.org/About.

20   OCLC, "OCLC Provides Downloadable Linked Data File for the One Million Most Widely Held Works in WorldCat" (August 14, 2012), https://www.oclc.org/news/releases/2012/201252.en.html.

21   Bibliothèque Nationale de France, "Semantic Web and data model," http://data.bnf.fr/semanticweb-en.

22   British Library Metadata Services, "Free Data Services," http://www.bl.uk/bibliographic/datafree.html.

23 International Federation of Library Associations and Institutions Study Group on the Functional Requirements for Bibliographic Records, *Functional Requirements for Bibliographic Records: Final Report* (Munich: K.G. Saur, 1998), 9.

24 International Council of Museums, The CIDOC Conceptual Reference Model, "Who We Are," http://www.cidoc-crm.org/who_we_are.html.

25 International Council of Museums, The CIDOC Conceptual Reference Model, "FRBRoo Introduction," http://www.cidoc-crm.org/frbr_inro.html.

26 National Archives of Australia, *Australian Government Recordkeeping Metadata Standard Version 2.0*, (2008), http://www.naa.gov.au/Images/AGRkMS_Final%20Edit_16%2007%2008_Revised_tcm16-47131 .pdf; and Archives New Zealand**,** *Technical Specifications for the Electronic Recordkeeping Metadata Standard Version 1.0* (2008), http://archives.govt.nz/sites/default/files/TS_4.pdf.

27 National Archives of Australia, *Recordkeeping Metadata Standard for Commonwealth Agencies Version 1.0* (Canberra: National Archives of Australia, 1999).

28 Comisión de Normas Españolas de Descripción Archivística, *Modelo Conceptual de Descripción Archivística y Requisitos de Datos Básicos de las Descripciones de Documentos de Archivo, Agentes y Funciones* (Ministerio de Educación, Cultura, y Deporte, June 18, 2012), http://www.mcu.es/archivos/docs/ NEDA_MCDA_P1_P2_20120618.pdf. An English language summary is also available as Comisión de Normas Españolas de Descripción Archivística, *Report on the Work of CNEDA (2007–2012): Toward a Conceptual Model for Archival Description in Spain* (Ministerio de Educación, Cultura, y Deporte, July 11, 2012), http://www.mcu.es/archivos/docs/MC/CNEDA/ReportCNEDA_11_07_2012.pdf.

29 Arkistolaitos, *Finnish Conceptual Model for Archival Description Draft Version 0.1* (March 1, 2013), http:// www.arkisto.fi/uploads/Arkistolaitos/Teht%C3%A4v%C3%A4t%20ja%20toiminta/Hankkeet/AHAA /The%20Finnish%20Conceptual%20Model%20of%20Archival%20Description_text.pdf.

30 LOCAH Project, http://archiveshub.ac.uk/locah/.

31 ICA AtoM, "Entity Types," https://www.ica-atom.org/doc/UM-1.3.

32 Comisión de Normas Españolas de Descripción Archivística, *Report on the Work of CNEDA (2007– 2012)*, 7.

33 Comisión de Normas Españolas de Descripción Archivística, *Report on the Work of CNEDA (2007– 2012)*, 14.

34 Comisión de Normas Españolas de Descripción Archivística, *Report on the Work of CNEDA (2007– 2012)*, 7–9.

35 Arkistolaitos, *Finnish Conceptual Model for Archival Description Draft Version 0.1*, 6.

36 Arkistolaitos, *Finnish Conceptual Model for Archival Description Draft Version 0.1*, 16–34.

37 Pete Johnston, "Two Changes to the Model and Some Definitions," LOCAH Project (February 16, 2011), http://archiveshub.ac.uk/locah/2011/02/two-changes-to-the-model-and-some-definitions/.

ABOUT THE AUTHORS

Gretchen Gueguen is a research assistant to the ICA Experts Group on Archival Description (EGAD).

Vitor Manoel Marques da Fonseca is an executive member of EGAD, a senior archivist at the Arquivo Nacional (Brazil), and a professor at the Departamento de Ciência da Informação (Information Science Department) of Universidade Federal Fluminense.

Daniel V. Pitti is chair of the EGAD and associate director of the Institute for Advanced Technologies in the Humanities at the University of Virginia.

Claire Sibille-de Grimoüard is an executive member of the EGAD, archive curator and head of the Archival Processing and Preservation Office, Direction Générale du Patrimoine (General Directorate of Cultural Heritage), Service Interministériel des Archives de France (Interdepartmental Service of French Archives).