Archival Science (2023) 23:629–673
https://doi.org/10.1007/s10502-023-09422-w

ORIGINAL PAPER

A metadata model for authenticity in digital archival descriptions

André Pacheco[1] · **Carlos Guardado Da Silva**[1] · **Maria Cristina Vieira De Freitas**[2]

Accepted: 4 July 2023 / Published online: 25 July 2023

Abstract

The advent of the information paradigm has shaken many of the principles of archival theory and practice. One key issue is knowing to which extent can digital information be trusted. Digital resources are represented by metadata, and trust consists in demonstrating their authenticity. Since the traditional elements used to verify the authenticity of analog records are not suitable in the digital world, the field faces a major challenge. The use of abundant, pertinent and constantly captured metadata seems to one of the most relevant solutions. This article aims to contribute to tackle this issue by setting the goal of proposing a model that attempts to include the most relevant metadata elements to capture the information that contributes for ascertaining the authenticity of digital archival descriptions. To that end, mixed methods methodology are employed. A qualitative documentary research is used to collect, analyze and interpret a corpus of scientific literature. As a complement, the quantitative technique requirements engineering is used to extract from international description standards the metadata requirements that can assist in the presumption of authenticity. Both approaches are then combined through a critical lens into a single unifying model for authenticity that is deemed as complete as necessary but as simple as possible. The model can be used by organizations or as a contribution to the discussion of authenticity and trustworthiness in digital archival descriptions.

Keywords Information representation · Archival description · Metadata · Model · Authenticity · Crosswalk

This article summarizes the PhD thesis of André Pacheco (2022), defended in Portuguese at Coimbra University, Portugal, and supervised by M.C.V.F. and C.G.S.

Extended author information available on the last page of the article

Introduction

As technologies evolve and transform, so do the communication processes enabled by them. A document is information in a medium, so a change in the medium changes what and how information is captured and transmitted. Ultimately, this changes the message. Technology is not neutral since "we shape our tools and thereafter they shape us" (Culkin 1967, p. 70).

A change in information processes —from creation to transmission — requires an equal change in archival theory and practice, at least by constantly questioning its principles and methods in order to assess if they still remain relevant in the new contexts. The transition from an analog to a digital paradigm, *i.e.*, from the management of records as physical entities to digital information management, has scrambled several core notions of the field, such as the principle of provenance and *respect des fonds*. Another aspect in which cracks have appeared is the presumption of authenticity, which traditionally relied on demonstrating a record's identity and physical integrity. In the digital environment, integrity is not a stable concept but a property of records that is verified by the preservation of a change log over time, since it is inevitable that a digital document is constantly changed and reinterpreted.

As a result, digital information management requires a constant stream of metadata to provide contextual information. Context, as expressed by metadata, can become more important than content, since it allows for content to be reconstructed or, from the perspective of a user query, it allows users to refine and filter search parameters to obtain a relevant corpus, even before viewing a single resource, known as positive search abandonment (Stamou & Efthimiadis 2010, p. 1). Since in the archival domain metadata emerges as a way of documenting context, it thus assists in preserving the identity and integrity of records, demonstrating their authenticity. This was one of the findings of the Inter PARES 1 project (Duranti & Blanchette 2004) that implies that archival description should not be seen as static, but rather as a "fluid, evolving and socially constructed practice" (Yakel 2003, p. 2).

The fluidity and social shaping of archival descriptions becomes particularly relevant the further we transition deeper into an interconnected web of resources, services and agents. The current vision of the Semantic Web aims to transform the web in a global database that enables links to be automatically created between resources as per the request of users (Machado et al. 2019). This intention relies on the resharing of metadata, which in turn requires interoperability between the various representation systems (or archival descriptions), understood as the seamless sharing of information across different systems. However, it is usually noted in the literature that interoperability in the web is hard, and as a result archival descriptions remain isolated (Rolan 2015). Additionally, since representation is subjective, how each community chooses to represent resources varies, further enhancing the problem of interoperability.

In order to facilitate the presumption of authenticity of resources in a shared web environment, and thus to ascertain their trustworthiness, an effort is

necessary to bridge these concepts across different communities. This work attempts to contribute to this gap in three different steps: 1) by performing a systematic review of the archival literature pertaining to archival description and metadata over a 10 year period; 2) to conduct a metadata mapping and comparison exercise drawing on ten relevant data structure standards; and 3) to merge both meta-analysis into a consolidated metadata model that can serve to represent the best practices in providing information about authenticity. As a result, it is a pioneer work that integrates the literature review as a critical theory for the analysis of the standards, that contextualizes the metadata elements featured in the standards in light of the previous literature review, documenting how that theory has been applied in practice, and consolidates those two analyses in a proposal of a metadata model that can be treated as a starting point for a reflection of the concepts of authenticity and trustworthiness, while at the same time serving as a representation of the most common metadata elements used or argued for in the field. By drawing input from a broad range of authors and publication venues over a large period of time, as well as from metadata standards from various practice communities and regions, the authors hope that it can be used both in practical application and/or as a starting point for theoretical thinking in how to increase the trustworthiness of descriptions and enhancing their interpretative context.

In this context, it is pertinent to revisit the principles on which the presumption of authenticity traditionally rested, to discuss if they remain relevant in the digital environment and, if necessary, to suggest an alternative. A mixed methods methodology is used pairing documentary research and requirements engineering complementarily in an attempt to rely on the scientific (articles in peer reviewed journals) and technical (metadata standards) literatures as input for the elaboration of the model. The first section of the article discusses the role of metadata in archives. The second section debates the constraints of representations and how authenticity can be conveyed. The third section presents the results of the analysis of the scientific and technical literatures, extracting as requirements the main ideas observed for the presumption of authenticity. The fourth and final section analyses and systematizes the requirements extracted previously, and consolidates them critically under a single model.

Metadata in archives

Definition and interpretation

Metadata are a form of archival description, which in turn is a form of information representation. Before delving into that discussion, it is important to reach an understanding of what metadata are, since it is a concept used across multiple fields with disparate interpretations. For example, Furner (2020) identified 96 separate ISO standards that provided 46 different definitions. The term ‘metadata’ was first used in 1968 by Philip Bagley, a computer scientist, who argued that one of the essential components of a programming language is “the ability to associate explicitly with a data element a second data element which represents data

'about' the first data element. This second data element we might term a 'metadata element'" (Bagley 1968). There are differences of opinion in the archival literature as to what can be metadata. For example, Zeng and Qin (2016) argue that the nutritional label of foods should be considered metadata, since they state a property (calories) that explains the meaning of a data value (250). However, Haynes (2018) disagrees by arguing that metadata can only be applied to informational resources traditionally found in archives, libraries or museums.

At the heart of the question is the scope of the represented object. We tend to agree with the broader interpretation of Zeng and Qin, which encompasses the notion that metadata can be generally used to describe any data that expresses a property about other data. Furthermore, they are "structured data about other data" (Gladney 2007, p. 7). The notion of structure is key and broadly found in the archival literature (Clobridge 2010; Cron 2016; Daniel & Daniel 2012; Gilliland 2016; Gueguen et al. 2013; ISO/IEC TR 19583–1 2019; NISO 2017; Witten et al. 2010). The Digital Preservation Coalition further adds that metadata are "data about a digital resource that is stored in a structured form suitable for machine processing" (Digital Preservation Coalition 2015, p. 56), thus inputting the notion of machine readability. Therefore, metadata can be understood as enriching the data with context and structure, in a manner that is machine processable. In fact, context and structure are essential characteristics of records (Duranti 1997; International Council on Archives 1997; Janes 2012), so it is no surprise that they also manifest in the metadata that represent them (Apostolou 2009; Cron 2016; Gilliland 2016; The National Archives of the UK 2002). According to Gilliland (2016, p. 2), content "relates to what the object contains or is about and is intrinsic to an information object", context "indicates the who, what, why, where, and how aspects associated with the object's creation and subsequent life and is extrinsic to an information object", while structure "relates to the formal set of associations within or among individual information objects and can be intrinsic, extrinsic, or both".

Metadata themselves require a structure in which they can be recorded. Such a structure is a metadata record, generally understood as a set of metadata that represent a resource, including their semantic, syntax and structure (Lubas et al. 2013). Jeffrey Pomerantz highlights that, the same way a book is a "container for data but is not data itself … metadata is data, but metadata cannot exist outside of a container: a metadata record must exist in some format, be it physical or digital. Likewise a metadata record is itself a container for data about an object" (Pomerantz 2015, p. 12). The metadata record thus acts as the internal context for a metadata set, whereas the external context provided by metadata is the network of relationships that a resource establishes with any other resource with which it can be linked. These relationships overlap with the archival bond for resources produced in the course of the same activity by the same person, but can be expanded to any other resource within the same representation system through means of the informational bond. This view is reinforced by the American National Archives and Record Administration (NARA), who write that "metadata elements also provide contextual information that explains how electronic records were created, used, managed and maintained and how they are related to other records" (NARA 2015, Metadata in Archives).

However, the interpretation of resources changes because context is not fixed. Metadata are created by certain people at a certain point in time, therefore with some kind of bias. Representations are subjective and metadata, as a form of information representation, do not escape this fate. As Gartner (2016, p. 4) puts it, "there is nothing objective about metadata: it always makes a statement about the world, and this statement is subjective in what it includes, what it omits, where it draws its boundaries and in the terms it uses to describe it". The notion of metadata as statements is not only common in the archival literature, but also necessary to understand that metadata are not objective because they are a human construct and everything that is human is not neutral. Yakel states that "through the process of selection of information for inclusion and choice of access points, archivists reveal and conceal, making finding aids political statements" (Yakel 2003, p. 19), imparting the subjectivity on the finding aids themselves, *i.e.*, on the archival descriptions.

In summary, "metadata is best understood as 'any statement about an information resource', regardless of what it is being used for, which metadata vocabulary is being used, and how the metadata is represented" (Garshol 2004, p. 379). The word 'statement' accounts for the subjectivity of the representation process. The 'information resource' can both refer to records, in the traditional notion of "recorded information produced or received in the initiation, conduct or completion of an institutional or individual activity and that comprises content, context and structure sufficient to provide evidence of the activity" (International Council on Archives 1997, p. 21), as well as any other object that can provide information, such as any entity that establishes a relationship with a record and is captured in the representation system, whether it is a person, an occupation, or a property.

Metadata interoperability

There are several forms in which metadata can be expressed. The most commonly used data models in archives are the relational model, which includes both tabular (e.g., an Excel sheet) and relational data (e.g., a relational database), the hierarchical model (e.g., XML) and the graph model (e.g., Resource Description Framework). These all have advantages and disadvantages, depending on the goals and constraints of the data curators. Despite emerging across the last decades, each still has a place in archival description, although serving different purposes. Data models can even be mixed. For example, data can be structured in XML but their relationships enhanced by RDF graph representations in the form of triples.

Regardless of which data model is used, it is key that data remains interoperable across different representation systems. Interoperability is one of the five key principles for the construction of metadata suggested by Marcia Zeng and Jian Qin (2016, p. 28), the others being modularity, extensibility, refinement and multilingualism. Interoperability can be understood as "the ability of multiple systems with different hardware and software platforms, data structures, and interfaces to exchange data with minimal loss of content and functionality" (NISO 2004, p. 2). Communicating across different systems so that data is minimally lost is not only a recurring problem for data curators, but it is also one of

the greatest hurdles in metadata management. The proliferation of options and standards for describing resources in recent decades exacerbates this problem by multiplying the number of possibilities available. Since these standards are often developed in isolation, seamless compatibility is not automatic. It requires effort for information to be shared efficiently and automatically. In an ideal situation, users "should be able to discover through one search what digital objects are freely available from a variety of collections, rather than having to search each collection individually" (Tennant 2001, p. 118). Databases should allow for cross-queries in order to obtain as many pertinent results as possible from different contexts, thus adding value and meaning to the resources retrieved. In order to achieve this, metadata are "governed by community-developed and community-fostered standards and best practices in order to ensure quality, consistency and interoperability" (Gilliland 2016, p. 2). However, in order for an efficient sharing of metadata, semantic interoperability is required in addition to syntactic interoperability. Not only it is necessary to communicate, it is also crucial to know how to interpret the communicated data.

As a result, there are at least two levels of interoperability. Syntactic interoperability refers to the standardization of communications between both systems (Schaeffer et al. 2012; Sudmanns et al. 2018) and is made possible by a "strict adherence to standard metadata formats and protocols while creating metadata structures and possible only through metadata deployment" (Ramesh et al. 2015, p. 197). Without syntactic interoperability, "data and information cannot be handled properly with regard to formats, encodings, properties, values, and data types; and, therefore, they can neither be merged nor exchanged" (Zeng and Qin 2016, p. 348). In turn, semantic interoperability can be defined as the "capacity for different agents, services, and applications to communicate data, information, and knowledge while ensuring accuracy and preserving the meaning of that data, information, and knowledge" (Zeng and Chan 2010, p. 4646). Without semantic interoperability, "the meaning of the used language, terminology and metadata values cannot be negotiated or correctly understood" (Koch 2006, Interoperability).

Furthermore, interoperability can be observed at various levels, namely at a schema, record or repository level. Since this paper aims to suggest a metadata element set that can be used to represent the authenticity of resources, it focuses on semantic interoperability at the schema level, which comprises the scope of metadata standards. Several types of standards exist that accomplish different goals. Data structure standards govern the elements used in the metadata record. Data content and data value standards clarify semantics, whereas data exchange standards refer to syntax. Amongst these types, this work focuses on structure standards that elucidate about the metadata elements to be used, regardless of how those metadata are expressed by syntax or semantics. Thus, data structure standards (sometimes also called metadata element sets, element sets, schemas or metadata vocabularies) can be seen as "'categories' or 'containers' of data that make up a record or other information object" (Gilliland 2016, p. 3). Those categories are the metadata elements, also known as data elements or data fields. They are the units of information that a standard dictates that should be captured, in the hierarchical order defined by the data structure (Elings & Waibel 2007).

One of the goals of this paper is to suggest a set of elements that can be used to reflect the authenticity of resources across multiple digital contexts, as informed by the meta-analysis of the scientific and technical literatures. Therefore, in the context of this work, the resulting metadata element set is posited as a model for authenticity, in the sense that a model is an abstract representation of an object or a state for analytic ends. Being "an abstractive representation of some object or state of affairs" (Wartofsky 1979, p. 4), the model individualizes a property of an object. It does not require a complete and reliable representation of the object, but merely a similarity to some of its properties. In this case, the object is archival description. The particular property to be studied is authenticity, not all the other characteristics of records. Thus, one of the aims is to create a model for authenticity in digital archival descriptions, manifested in the proposal of a metadata element set. It does not aim to be exhaustive, but rather to be as complete as necessary and as simple as possible. However, before attempting to do so, it is important to understand what authenticity encompasses in order to be able to establish criteria for populating the content of the model.

Authenticity and trustworthiness in archival representations

Information Representation

First, it is important to understand that all descriptions in archives are information representation, which in turn are a form of representation. Rosenberg argues that "the essential and characteristic human activity is representation—that is, the production and manipulation of representations (Rosenberg 1981, p. 1). Representations have certain philosophical constraints that emanate throughout the represented objects. One is that a representation is, at its fundamental level, "something that stands, or is believed to stand, for something else" (Yeo 2018, p. 129). The solar system is studied in a classroom using balls that represent planets, with the essential condition that students interpret those balls as an accurate representation of the actual planets. Representations serve a purpose of bringing to attention a distant object, acting as "surrogates for memory" (Jimerson 2015, p. 99). They make present something that is absent. In the case of information representation in archives, archival description enables us to discover and learn about records that are held elsewhere. This aspect is particularly important in the digital realm, in which we interact exclusively with representations in our screens. Resources are searched, aggregated and contextualized using metadata and, if considered pertinent for the objectives of the user, they may consult a digitalized version of the resource (such as a PDF). In archives, records can also be considered a representation of the actions they describe. However, this study focuses merely on archival description as representations of resources, and as metadata as statements about those records.

Additionally, one of the most critical axioms is that a representation is a simplification of an object. Therefore, they are necessarily incomplete. That is a consequence of the reason why we represent — to reduce the complexity of an object so that it can more easily be studied and manipulated. In the classroom example, the

balls might have the same shape and relative distance to one another as the actual planets, but do not possess their mass. If some properties of the object are represented but others are not, there is a choice. Such conscious and deliberate choice is made by the creator of the representation. Consequently, no representation is neutral.

The consciousness about the subjectivity of representations has been introduced in archival science by several postmodernist thinkers (Brothman 1993; R. Brown 1991; Harris 1997; Hedstrom 1993; Nesmith 1999; Upward 2000). Modernists, inspired by the French philosopher August Comte (1798–1857), believed that knowledge was limited by what could be experienced, or positively determined. It was believed that science was above any cultural, historical, political or social context, encapsulating the truth within itself, as if it were an objective truth. Post-modernists challenged this notion, arguing that the human perspective conditions the interpretation of phenomena. A consequence is that archivists are no longer seen as "passive guardians of an inherited legacy" but instead as "actively shaping collective (or social) memory" (Cook 2001, p. 4). Archivists shape memory by selecting what to preserve and then by building representations of what was kept. How records are represented by archival description has also been a contentious topic in archival science. The traditional approach has been to the *fonds*, according to the principles of provenance and original order. The several international standards for archival description published by the International Council on Archives (ICA) throughout the 1990s have prioritized collection-level descriptions of records, according to a hierarchical structure from the general to the particular, from the *fonds* to the item. The ICA has since then changed its shift and, in its latest standard *Records in Contexts*, it acknowledges that this focus on the person or group that has accumulated a body of records "often does not reflect the social and material complexity of the origins of records" (International Council on Archives 2016, p. 5).

Bailey also protested against the idea of the *fonds* as the single method of archival description, since "in a database, objects are related but not ordered. The database logic is nonlinear and there is no original order because order is dependent upon query. … Digital objects will have an identifier, yes, but where they 'rest' in intellectual space is contingent, mutable. The key point is that, even at the level of representation, arrangement is dynamic: access and representation need not depend on the fonds" (Bailey 2013). Similarly, Anne Gilliland reinforces the criticism of using exclusively collection-level, hierarchical metadata arguing that, while they remain valuable for retaining context and original order, it "represents an oversimplified view of the actual complexities of records-creation process and provenance … and necessarily perpetuates a paper-based descriptive paradigm" (Gilliland 2016, pp. 5–6).

As a result, in addition to collection-level metadata, it becomes a requirement for digital archival representation to also include item-level metadata, so that resources can be easily recontextualized in novel aggregations upon query by users. Despite some authors believing that "a relentless focus on the aggregate part is what sets us [archivists] apart from librarians and museum curators " (Greene 2009, p. 24), David Bearman suggests that "item level information is fundamentally more valuable because it can generate more valid collective level data in addition to serving the needs of item documentation" (Bearman 1996, p. 205). Item-level description seems

to be a necessary complement to collection-level descriptions. Both serve distinct but crucial roles. While collection-level description communicates the context of a records producers and provenance, item-level metadata facilitates the discovery of resources and their recombination in different aggregates, multiplying their contexts as well as their potential interpretations and meanings. As Terry Cook summarizes:

> *Nothing is neutral. Nothing is impartial. Nothing is objective. Everything is shaped, presented, represented, re-presented, symbolized, signified, signed, constructed by the speaker, photographer, writer, for a set purpose. No text is a mere innocent by-product of action … but rather a consciously constructed product … Texts are all a form of narration more concerned with building consistency and harmony for the author … than they are evidence of acts and facts, or juridical or legal frameworks. And there is not one narrative in a series or collection of records, but many narratives, many stories, serving many purposes for many audiences, across time and space* (Cook 2001, p. 7).

It is not the role of archivists and similar data stewards to limit which queries can be made beyond the legal restrictions (such as embargos), but instead to facilitate discovery and reuse of resources through adequate and complete enough archival descriptions that maximize the relationships established by resources, enabling users to construct their queries in a manner that is not pre-determined. This notion reinforces the importance of developing a model for authenticity in archival description that can serve as a guideline for these professionals.

Authenticity, trustworthiness and reliability

The importance of authenticity for the archival profession is attested by its presence as one of the fundamental purposes of an archivist's ethical code: to ensure the authenticity of records under custody (International Council on Archives 1996; Society of American Archivists 2012). Authenticity refers to "trustworthiness of the record as a record, i.e., the quality of a record that is what it purports to be and is free from tampering or corruption" (InterPARES 2 2008, p. 8). It requires establishing the identity and demonstrate the integrity of records over time. Identity is understood as "the whole of the characteristics of a document or a record that uniquely identify it and distinguish it from any other document and record" whereas integrity refers to "quality of being complete and unaltered in all essential aspects" (InterPARES 2, 2008, pp. 25–26). To question identity is to ask "was it written by who purports to have written it?", while integrity attempts to answer the question "has it been altered in any way since it was first created and, if so, has such alteration changed its essential character?" (MacNeil 2000, p. 53).

The ICA also defines reliability and usability as two other essential characteristics of records. Reliability means that "a record can be trusted as a full and accurate representation of the transaction(s) to which they attest, and can be depended on in the course of subsequent transaction", whereas usability implies that records can be located, retrieved, preserved and interpreted (International Council on Archives 2008, p. 13). The relevance in establishing that a record is

authentic and reliable is related to the ability to demonstrate its trustworthiness as a representation of the actions it manifests. Trustworthiness is "the quality of being dependable and reliable" (Society of American Archivists 2005, p. 388), and is generally defined as hinging on reliability and authenticity.

In the analog paradigm, authenticity was demonstrated by a tight control on recordkeeping processes. In the digital realm, integrity is constantly being challenged as information flows across systems and mediums. Therefore, it has become more consensual that, when managing digital resources, the elements that contribute to the presumption of authenticity "are observable not in the document itself but rather in the procedures" (Eastwood 1994, p. 127) of record creation, maintenance and preservation. Since archival descriptions are basically a metadata element set, they feature as one of the key procedures for capturing the authenticity of records (InterPARES 1 2002), in the sense that they clarify the arrangement of records and the circumstances (the contexts) in which they were created (International Council on Archives 2016; MacNeil 2005). The link between authenticity and archival description is further documented by the *Canadian Rules for Archival Description*, when stating that one of the goals of archival description is "to establish grounds for presuming the authenticity of archival material by documenting its chain of custody, arrangement, and circumstances of creation and use" (Bureau of Canadian Archivists 2008, p. xxii). Given the evolution in importance of integrated digital descriptions, it becomes increasingly important that metadata is interoperable across information systems and archival collections, so that the chain of custody, arrangement and custody of records are preserved, and their authenticity can be reinforced.

It is also relevant to highlight that authenticity is not a binary question, so that something would be simply authentic or not (Factor et al. 2009). First, authenticity can only be understood in reference to the goals of a user. As Bonnie Mak explains, "the purpose of establishing the authenticity of a record is to position it as trustworthy for a particular purpose, and authoritative within a particular framework" (Mak 2015, p. 122). Furthermore, "the level of confidence one can have in authentication is related to the strength of the evidence, and should not be considered absolute" (Suvak 2015, p. 117). As a consequence, it is observed that there is a gradient in authenticity (Carta 2017, p. 195), a spectrum in which one end is 'more authentic' and the other 'less authentic'. The middle of this spectrum is filled with a grey area of nuance. In practice, this means that archivists cannot claim that a resource or a description is authentic. At most, they can provide the maximum amount of evidence possible so that a user determines if, in their perspective, the object represented is authentic and, therefore, trustworthy. Authenticity is itself a subjective and personal notion, since what might be enough for a user may not be sufficient for other. As a result, the model discussed in this paper cannot have the pretentious goal of ensuring the authenticity of archival descriptions, but merely to provide a guideline to capture certain information about resources so that, when present, they may be complete but simple enough so that users can form their own judgement, and so that such judgement may with greater likelihood conclude that the descriptions are authentic, reliable and trustworthy in as many contexts as possible.

Methodology

The plurality of perspectives, backgrounds in authors of the field, as well as the different scopes of several metadata standards, calls for a consolidated rethinking of how to represent authenticity in the digital web, a problem acknowledged by the International Council on Archives (2019) when developing *Record(s) in Context(s)*, the latest international archival standard. The main goal of the study is to propose a model that can help bolster the presumption of authenticity in digital archival descriptions. In order to do so, it commits to an extensive review of the scientific and technical literatures as data points for the construction of the model. Documentary research (C. G. da Silva 2021) is employed as a qualitative method, determining the need for assembling a literary corpus and positioning the researchers as the interpreters of the meaning of the texts. The first step in selecting the corpus of scientific works that comprises the scientific literature review was selecting the database Library, Information Science and Technology Abstracts (LISTA) as the source for data collection, due to the relevance of publications it indexes and availability of the texts. Within LISTA, there was an initial query with the terms "Title = 'archiv*', 'arqª' or 'records management'". The goal was to identify the most relevant (high *Índice Compuesto de Difusión Secundaria*, or ICDS) journals that versed directly about archives or records management, as defined by their presence in the title.

This query yielded on May 2019 a total of 25 results, which were then further filtered by excluding the journals with an ICDS < 6.5, not being peer-reviewed, not having descriptive information, not being active, not being in English, Spanish, French or Portuguese, or being off-topic for the purposes of the research. The result of the application of these filters was the selection of 12 journals deemed appropriate for the research goals: *American Archivist*, *Archival Issues: Journal of the Midwest Archives Conference*, *Archival Science*, *Archivaria*, *Archives [London]*, *Archives [Quebec]*, *Archives & Manuscripts*, *Archives & Records*, *Cadernos de Biblioteconomia*, *Arquivística e Documentação*, *Journal of Archival Organization*, *Provenance: the Journal of the Society of Georgia Archivists* and *Records Management Journal*. Once the journals were identified, all articles published between 1 January 2009 and 31 December 2019 were retrieved. Their pertinence and inclusion in the study was analyzed through reading the title and abstract. A total of 213 pertinent articles were extracted from all journals.

As a complement to the scientific literature review, it was considered important to also include in the model the insights from data structure standards for archival description, since these instruments play a key role in influencing archival practice. Since no single standard can adequately capture every context of representation of the several types of collections or resources across different communities (Apostolou 2009; Gilliland 2016), it was decided that the sample should include data structure standards from multiple fields (archives, libraries and museums), geographies (Europe, Americas and Australia) and languages (Portuguese, English or Spanish), and available for free on the Internet. Based on these criteria, the following ten standards were chosen: *General Standard for*

Archival Description (ISAD-[G]), *Australian Recordkeeping Metadata Schema* (AGRkMS), *Encoded Archival Description* (EAD), *Esquema de Metadados para la Gestión del Documento Electrónico* (e-EMGDE), *Describing Archives: A Content Standard* (DACS), *Dublin Core Metadata Element Set* (DCMES), *Visual Resources Association Core* (VRA Core), *Metadata Object Description Schema* (MODS), *Categories for the Description of Works of Art* (CDWA) and *Records in Contexts* (RiC).

ISAD-(G) and DACS were selected for being international archival standards; AGRkMS for being an archival standard used in Australia; EAD for being an international standard focused on finding aids; e-EMGDE for capturing the Spanish-speaking worldview on public recordkeeping; DCMES for being a multi-purpose international standard with broad application in representing content on the web; VRA Core to include in the analysis an international perspective on description of visual resources and images; MODS to include the bibliographic approach, as it is used internationally; CDWA for its focus in the representation of artworks in museum contexts; and finally, RiC v0.2 was included, albeit being still in draft version, due to its anticipated importance in shaping archival description worldwide, as it is the latest ICA international archival standard, and has been designed specifically with a concern of how to capture the multiple contexts in which resources circulate. The element set should not change significantly in the final version, thus it is expected that the current analysis remains relevant. Thus, the metadata mapping would be incomplete without its mention. Other standards could have been included, but it was deemed that these ten, given their diversity in scope and application contexts, were sufficient for the goals of the study.

These standards were analyzed using the quantitative technique of requirements engineering (Bennaceur et al. 2019). It is a technique borrowed from computer science, in which it is used for software development, since it is useful for defining the requirements that a certain system must have in order to function. Within the scope of this study, the goal is to use this technique to elucidate about the requirements of authenticity, as expressed in the technical standards. Those requirements are first derived from archival theory, according to the scientific literature review, and then critically analyzed in the standards. The expectation is that requirements engineering can assist in identifying the metadata elements employed by the standards most directly correlated with the demonstration of authenticity and, thus, trustworthiness of descriptions. As a result of the combination of qualitative and quantitative techniques, this research follows a mixed methods approach (Creswell & Clark 2013).

The use of mixed methods is particularly useful for this study since quantitative methods allow for a detailed understanding of a problem, although usually lack the ability to extrapolate those results to a broader social context. In contrast, qualitative methods are usually limited in their capacity to analyze large volumes of data, and are more heavily biased by the subjectivities of researchers. Therefore, they are used in a complementary manner, so that the strengths of the qualitative approach compensates for the weaknesses of the quantitative, and vice-versa (Brewer & Hunter 2012; Creswell & Clark 2018; Jick 1979). In the context of this study, it is expected that documentary research illustrates how to represent resources' authenticity through archival descriptions, but by itself would lack a broader context on how

Table 1 Analytical categories emerging from the literature review

Category	Author and work
Information representation	(Duff et al. 2013; Niu 2015; Serewicz 2010; Silva 2012; Verborgh et al. 2015; Yakel 2003; Yeo 2012; Zhang 2012a, 2012b)
Archival description	(Anchor 2013; Douglas 2016; Gracy and Lambert, 2014; Hedstrom 1993; MacNeil 2009, 2012; Meehan 2009; Moyano Collado 2013; Niu 2013; Padrón and Cabero 2019; Zhang, 2012c)
Information management	(Acker 2017; Bailey 2013; Brown, 2019; Cumming 2010; Duranti and Franks, 2015; Kallberg, 2012; MacNeil 2017; McLeod 2014; Ridener 2009; Tough 2016; Wright 2014; Yeo 2011)
Metadata	(Apostolou 2009; Baca, 2016; Beyene and Godwin 2018; Chen et al. 2011; Gartner 2016; Gilliland 2016; Gladney 2009; Haynes 2018; Li and Sugimoto 2017; NISO 2017; Pomerantz 2015; Yeo 2018; Zeng and Qin 2016)
Metadata standards	(Andrade et al. 2014; Baños-Moreno et al., 2019; Botão, 2011; Bunn 2013; Dow, 2009; Dryden 2009; Ducheva and Pennington, 2019; Henttonen, 2009; Linden, 2017; Llanes-Padrón and Moro-Cabero 2017; Moro-Cabero et al., 2011; Pastor-Sánchez and Llanes-Padrón, 2017; Riley, 2010; Rolan, 2017; Woodley 2016; Youn 2015)
Authenticity	(Bhatia and Wright de Hernandez 2019; Bountouri et al. 2017; Dryden, 2011; Duncan 2009; Duranti et al. 2019; Engvall 2019; Gladney 2009; Hofman et al. 2019; Jansen 2015; Lemieux 2016; Mak, 2012; McLeod and Gormly 2017; Price and Smith 2011; Rogers 2015; Rogers and Tennis, 2013; Yeo 2013)
Linked data	(Gartner 2015; Hooland and Verborgh 2014; Jones 2018; Machado et al. 2019; Niu 2016; Rolan 2015; Rolan et al. 2019; Samouelian 2009)

authenticity is captured not only by archival standards, but also in other communities of practice. On the other hand, the exclusive use of requirements engineering would provide a detailed, comparative list of each metadata element used, but would be void of the critical thinking to analyze, debate and frame them within the theoretical discourse. Their complementary use as part of a mixed methods approach reinforces the ability to conduct the meta-analysis of the scientific literature, to be able to translate that learning into the specific metadata elements used by the standards and, in turn, to interpret and contextualize those elements in light of the literature review.

Results

Scientific requirements

In order to facilitate the analysis of the vast scientific literature, it became necessary to group the articles thematically. As a result, the following emergent categories were observed based on data analysis: (1) information representation (IR); (2) archival description (AD); (3) epistemology of information management (IM), sometimes referred to as records management; (4) metadata standards (MDS); (5) epistemology of metadata (MD); (6) authenticity and trustworthiness (AUT); (7) and linked data

(LD). Table 1 illustrates some of the most relevant works and authors identified in each category.

These categories form the thematic pillar of the literature review, in which it is possible to categorize the majority of works identified in the documentary corpus. Articles were aggregated under 'Information Representation' as it was noted that they frequently focused on the information mediation between archive users and how archivists organize collections. The category 'Archival Description', albeit a part of information representation, was chosen to adopt terminological consistency with standards for archival description. These texts tend to question the active and subjective influence of archivists in description of resources and in the recreation of context, revisit the essential functions of archival description or explore the relationship between archival description and digital curation. 'Information Management' articles tend to focus specifically on procedures for information management and organization in archives.

The category 'Metadata' was chosen to represent the scientific discussions closest to the concept of metadata, particularly from an archival perspective, although some of the authors included are not archivists, which brings a welcome diversity into the analysis. The category 'Metadata Standards' captures research that focuses on the topics of mapping between metadata standards, difficulties in implementing standards, guidelines for the development of standards and guidelines for adapting metadata standards to the Semantic Web. The category 'Authenticity' captures the works which explicitly aim to debate the authenticity and/or trustworthiness of resources, documents or information. Finally, under 'Linked Data' are included articles that problematize archives from the perspective of concepts such as the Semantic Web and linked data, therefore usually focused on interoperability, as well as in the sharing and reuse of digital archival descriptions.

A detailed list of the bibliographic references of articles that were consulted but do not featured as direct citations is included in Appendix 1. Each single article in Table 1 was read and analyzed under the lens of which criteria it identifies as a precondition for authenticity to be inferred from archival descriptions. A summary of the emerging requirements found in each category is provided in Appendix 2.

Technical requirements

For the ten standards that comprise the technical corpus, previously identified in the methodology, a detailed analysis of their data structure was performed. Each metadata element was considered and, based on the conceptual notions of authenticity as seen in the literature review, a decision was made whether or not to consider them relevant in establishing the authenticity of resources. Since authenticity is comprised of the demonstration of identity and the verification of integrity, the analysis focused on these two parameters. It is noteworthy to observe that authenticity was found in relation not only to resources, but also to the descriptive practice itself. Metadata elements that identify the creator of the archival description, as well as notes or rules followed, were also considered as key for authenticity,

which in this case can be seen as meta-metadata, or elements for the control of descriptions.

Thus, a summary of all metadata elements considered relevant for informing about the authenticity of both resources and descriptions themselves is presented in Table 2. The elements are grouped according to their focus: identity, integrity, and description's control, and are separated vertically by a blank space. It should also be noted that, due to time and analytical capacity constraints, not every single metadata element of the standards was considered, but only the first two levels of representation — categories and sub-categories. From the 26 elements of ISAD(G), 16 were identified as potentially relevant to asserting authenticity, among which six were considered mandatory by the standard creators; from AGRkMS's 27 metadata elements and 44 sub-properties, 12 were chosen; regarding EAD, 29 metadata were selected amongst the 165 metadata elements and 85 attributes; in the case of e-EMGDE, amidst its 30 elements, 21 were selected for the mapping; from DACS' 25 elements, 16 were considered related to authenticity; from DCMES' essential set of 15 elements, ten elements and four refinements were selected; for VRA Core 4.0, 17 metadata terms were chosen amongst the 19 categories and their sub-elements; from MODS' 20 elements, 12 metadata elements and four sub-elements were selected; from CDWA's abundant 540 categories and subcategories were chosen 51 elements and sub-elements; and finally, from RiC were extracted 18 entities and attributes based on the drafted 22 entities and 44 attribute.

Construction of the model

Analysis and discussion of the scientific requirements

The identification of the scientific requirements in Appendix 2 enabled an overview of the discussions on the concept of authenticity across several perspectives and authors within the field. In order to further facilitate the analysis, the unique requirements were traced and systematized in Appendix 3. A total of 32 requirements were identified. The most common ones are requirements 1, 7, 2, 4, 10 and 17. This seems to indicate that a need for greater transparency and accountability by the creators of the descriptions (requirement 1), as consequence of the growing consciousness of the subjectivity of the representation process (requirement 10) is generally recognized in the archival literature. Therefore, the intellectual control that archivists elaborate can and should be questioned, enabling the possibility that resources belong simultaneously to multiple contexts and collections, depending on the interpretation perspective (requirement 4). Furthermore, it is also generally agreed upon that metadata must be collected across the resources' lifecycle (requirement 7) and must include abundant and detailed contextual information as a way of documenting provenance (requirement 2) and strengthening the verification of their integrity (requirement 17). The fact that these requirements feature in the majority of categories (4 or more) does not mean that less frequent ones are less

Table 2 Summary of the metadata elements for authenticity identified in each standard

ISAD(G)	AGRkMS	EAD	e-EMGDE	DACS	DCMES	VRA	MODS	CDWA	RiC
Reference code	Name	Agencycode	Identificador	Reference code	Coverage	Date	Dateissued	Creation Date	Identifier
Dates	Identifier	Recorded	Nombre	Title	Date	Material	DateCreated	Measurements	Name
Extent and medium	Date range	Unitid	Fechas	Date	Format	Measurements	DateCaptured	Dimensions description	Date
Level of description	Extent	unitdate	Tamaño	Extent	Identifier	Description	PhysicalDescription	Materials/techniques description	Record Resource Ext
System of arrangement	Medium	Unitdatestructured	Soporte	Scope and content	Subject	Subject	Identifier	Physical description	Documentary Form
Scope and content	Document Form	Physdesc	Tipo documental	System of arrangement	Creator	CulturalContext	Classification	Place/location authority	Record Set Type
Title	Category	Physdescstructured	Categoría	Name of creator(s)	Contributor	StylePeriod	TargetAudience	Catalog level	Classification
Name of Creator	Description	Archdesc	Descripción	Name and location of repository	Provenance	Title	Abstract	Object/Work Type	Scope and Content
Administrative/biographical history	Coverage	c	Ámbito	Administrative/biographical history	Title	WorkType	Genre	Classification term	Structure
	Related entity (relationship)	Scopecontent	Entidad Relacionada			Agent	Subject	Styles/Periods…	Agent
Existence and location of originals	Change history	Unittitle	Clasificación	Custodian history	AccrualMethod	Source	TitleInfo	Orientation/Arrangement	History

Table 2 (continued)

ISAD(G)	AGRkMS	EAD	e-EMGDE	DACS	DCMES	VRA	MODS	CDWA	RiC
Existence and location of copies	Name	Origination	Historia del cambio	Immediate source of Acquisition	AccrualPeriodicity		Name	Subject matter	
Physical characteristics and technical requirements	Identifier	Bioghist	Nombre (Agente)	Appraisal, destruction, and scheduling information	AcccrualPolicy	Location	TypeofResource	General subject terms	Integrity
Related unit of description			Órgano	Accruals	Relation	Inscription		Context descriptive note	State
Source of acquisition	Disposal	Custodhist	Origen del documento	Existence and location of copies	Source	Relation	DateModified	Title text	Accrual
Archival history	Location	Acqinfo		Existence and location of originals		StateEdition	RelatedItem	Creator identity	Physical characteris
Appraisal, scheduling and destruction information	Integrity check	Accruals	Dictamen	Related archival materials	Publisher	Technique	Location	Creator role	Production technique
		Appraisal	Características técnicas			Textref		Creator description	Place
Archivist's note	Contact (Agent)	Phystech	Localización	Notes			RecordInfo	Person/corporate body authority	Authenticity Note

Table 2 (continued)

ISAD(G)	AGRkMS	EAD	e-EMGDE	DACS	DCMES	VRA	MODS	CDWA	RiC
Rules or conventions		Originalsloc	Verificación de integridad	Description control		–			
Dates of descriptions		Altformavail	Firma					Materials/techniques	–
		Relatedmaterial	Estado de elaboración					Inscriptions/marks	
		Separatedmaterial						State	
			Contacto					Facture	
		Processinfo						Condition/examination history	
		Conventiondeclaration						Conservation/treatment history	
		Maintenanceevent						Related works	
		Eventdatetime						Related visual documentation	
		Maintenanceagency						Related textual references	
		Maintenance history						Current location	
		Maintenance status						Ownership/collecting history	
								Cataloging history	

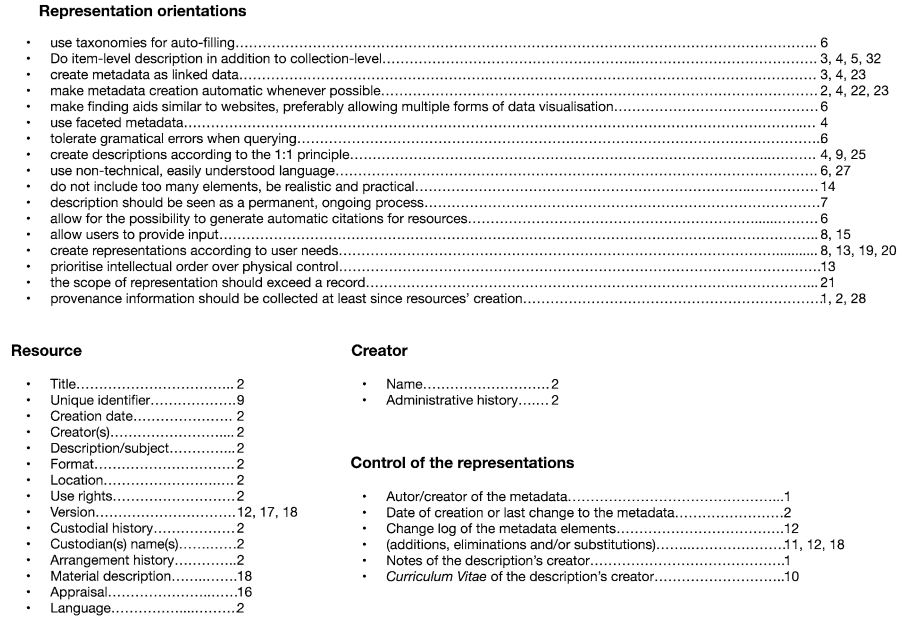

Fig. 1 Metadata elements and orientations derived from the scientific literature. The number indicates the requirement fulfilled, according to Appendix 3

important. An argument could be made that those are even more relevant, since they are more frequently overlooked. As such, the model considers all requirements equally.

The requirements seem to be structured across four dimensions: (1) theoretical insights on representation; (2) metadata about the resource, (3) about the creator of the resource and (4) about the control of representations. A systematization of this interpretation is summarized in Fig. 1, which serves as the final analysis of the information extracted from the scientific literature.

Analysis and discussion of the technical requirements

The data extracted from the international standards was analyzed in a similar way to the articles. The metadata elements identified across every standard, as shown in Table 2, were standardized and the unique requirements were identified. Appendix 4 provides a list of the unique requirements, serving also as a crosswalk between the 10 standards that were scrutinized. The distribution of the metadata elements also sheds some light on the differences between the standards. ISAD(G), as an ICA international standard for archives, includes several elements directly related to the archival perspective. It is also a reflection of the time of its creation, in the 1990s, revealing a documentary inclination for the management of records as physical

entities and for a multilevel, hierarchical structure. As for AGRkMS, a standard for records management in public administration, it shows some similarity with archival practices, but is more focused on digital information when compared with ISAD(G). It is also notable for the absence of metadata elements for control of the descriptions.

The nature of EAD as a standard for the codification of archival descriptions is clearly made evident by its metadata structure. Some of its elements, like archival description, biographical history, custodial history and appraisal are characteristic of elements used in information management in archives. As such, it is closely related to ISAD(G) and AGRkMS, but distinguishes itself by also focusing on the transparency of the descriptions. The standard e-EMGDE is a product of the juridical context that governs record creation in Spanish public administration. In addition to the traditional elements used in archival description, it distinguishes between public and private documents, and shows some characteristics of a focus on the medium. Here it is also found the absence of metadata for the control of descriptions.

The simplistic approach of DCMES is reflected upon its vaguer designations and omission of some of the most usual elements used in archives, such as administrative history. Instead, it seems to be an essential set of elements necessary for the basic identification of resources. VRA Core is another standard that is not from the archival community, but is instead used for the representations of artwork. Therefore, it includes some elements important for this community, such as cultural context. There is also an almost absolute absence of control metadata. Likewise, MODS is a standard for the description of bibliographic resources commonly used in libraries. It also omits certain elements of the archival domain but, contrarily to VRA Core, emphasizes the control of descriptions by documenting their author, language, sources and footnotes. Another standard used in artwork is CDWA, which focuses on artistic objects such as architecture. It dedicates great detail to the physical characteristics of works, such as techniques and construction methods, but seems to disregard metadata related to the digital management of resources. Lastly, RiC v0.2, although still in consultation draft, as the latest international standard for archival description, seems to consolidate the representation of information communities that are close but sometimes distant, such as archives, libraries and museums. As a result, the metadata elements do not seem to be exclusive to any of these communities, but rather serve as a flexible conceptual model that each information manager may tailor to their needs. However, there is a lack of explicit metadata for the control of descriptions, which could be a useful addition.

Despite the differences between some standards, some requirements stand out as most frequent. Considering the universe of 10 standards, the creation date of resources (T4 in Appendix 4) features in every single one. The resource's title (T2), creator (T3), physical dimensions (T5), administrative history (T16), current location (T19) and related resources (T20) feature in 90% of standards. The use of an identifier for resource description (T1) and the custodial history occur in 80% of standards. With less frequency albeit in the majority of standards is also observed

Resource

- Identifier.....T1
- Title.....T2
- Date.....T4
- Extension.....T5
- Medium.....T6
- Format.....T7
- Physical characteristics.....T8
- Target-audience.....T9
- Name and location of the repository..T10
- Documentary form.....T11
- Orientation/arrangement.....T12
- Description level.....T13
- Arrangement system.....T14
- Scope and content.....T15
- Administrative history.....T16
- Origin of the resource.....T17
- Elaboration state.....T18
- Location.....T19
- Related resources.....T20
- Custodial history.....T21
- Appraisal.....T22
- Acquisition source.....T24
- Integrity declarations.....T25
- History of physical interventions.....T26

Aggregate

- Accruals.....T23

Creator

- Name.....T3
- Biographical history.....T16

Control of the representations

- Creator of the description.....T27
- History of the description.....T29
- Rules or conventions.....T30
- Sources of the description.....T31
- Language of the description.....T32
- Notes.....T33
- Contact.....T34

Fig. 2 Metadata elements identified in the technical literature, with the indication of the requirement fulfilled, according to Appendix 4

in 60% of the sample the description of the physical characteristics of resources, such as documentary form or type of resource (T11), description level (T13) and the classification system (T14). In 50% of the standards are found the physical medium of the resources (T6), information about appraisal, including destruction (T22), additions to the collection (T23), the author of the descriptions (T27) and their date (T28).

As with the scientific requirements, all requirements were considered equally when constructing the model, even the least frequent ones. The final assessment of the metadata elements to include is presented in Fig. 2, which seems to orbit around four dimensions: metadata about the resource, the aggregate, the creator, and the control of representations.

Proposed model

A comparison between the requirements extracted from each approach reveals some similarities. For example, both perspectives include arguments for the description of resources, creators, or control of the representations. Regarding the creators, both technical and scientific literatures suggest the indication of their name and biographical history. As for meta-metadata, there is also a juxtaposition in recommending the creator of the description, the data, a changelog and creator notes. However, the

1. Resource

1.1 Identifier
1.2 Title
1.3 Date
1.3.1 Creation date
1.3.2 Date of the last change
1.3.3 Version
1.3.4 Elaboration state
1.4 Creator
1.5 Physical characteristics
1.5.1 Dimensions and medium
1.5.2 Format
1.5.3 Software and hardware
1.6 Description level
1.7 Subject
1.8 Language
1.9 Location
1.9.1 Repository's name
1.9.2 Custodian's name
1.9.3 Address
1.10 Representation form
1.11 Custodial history
1.12 Arrangement history
1.13 Appraisal
1.13.1 Elimination
1.13.2 Retention plans
1.14 Related resources
1.15 Integrity declarations
1.16 Log of physical interventions

2. Creator

2.1 Identifier
2.2 Name
2.3 Biographical history

3. Aggregate

3.1 Accruals

4. Meta-metadata

4.1 Metadata creator
4.2 Metadata date
4.2.1 Metadata creation date
4.2.2 Metadata last change date
4.3 History of changes to the metadata
4.4 Rules/conventions
4.5 Sources
4.6 Metadata language
4.7 Notes
4.8 Contact

Fig. 3 Proposal of the metadata model for the authenticity of archival descriptions

scientific literature is usually more summary, repeating a smaller number of essential requirements. As a result, as far as control metadata are concerned, the technical literature includes all these and a few more, so is more complete. To counterbalance, the scientific literature details more completely some considerations for the representation process, including theorizing the objectives of description and some implementation guidelines.

Once the contributions of each portion of the corpus have been clearly articulated, a comprehensive analysis that uniformizes them is now possible. It is worth remembering that the ultimate goal of the model is for the resources it describes to be considered trustworthy, which is made possible when authenticity is established. The analysis is driven by an overarching concern to capture not only the identity and integrity of the descriptions of resources, but also the identity and integrity of the descriptions themselves. Based on these premises, the interpretation of the data yielded 42 metadata elements that comprise the model proposed, reproduced in Fig. 3. It is believed that these elements capture the most relevant dimensions of authenticity. It does not simply combine all metadata elements from both data sources, since that would result in a long list that could be hard to implement. Instead, it aims to be a trade-off between simplicity and completeness, in the sense that it intends to be simple enough to be easily implemented, but also complete

enough to not overlook any element that might be important for users when appraising authenticity.

As such, concerning the resources, the indication of an identifier, title, dates, creator name and physical characteristics are crucial to manage resources as physical objects. The description level, subject and representation form contribute to describe the structure of the aggregate, the activities that are reflected in the resources, and the nature of the resources in that aggregate. Furthermore, the description of the resources' location and custodial history is essential to inform about the chain of custody over time. The declarations of integrity also assist in clarifying the history of transmission of records, highlighting the physical manipulations that they underwent. A description of the interventions of archivists in the management of resources is enabled by the elements of classification history, appraisal, and physical interventions. Regarding provenance, language can also be an important information. Metadata about the creator contribute to strengthen provenance by indicating an identifier, a name and a biographical history. Still referring to resources, the inclusion of information about related resources is important to establish bonds with resources from other aggregates, contextualizing the interpretation and meaning of resources.

As for the aggregation, a description of new additions may aid in clarifying interventions by the curators. Other information may be provided at the aggregation-level, such as physical characteristics, description level, subject, location, amongst others, and automatically inherited by lower description levels. However, they feature in the model underneath the resources as a reminder that this information should be available at an item-level in order to facilitate retrieval, enhancing discovery, interpretation and contextualization. Lastly, it is believed that all metadata elements suggested for the control of descriptions — or meta-metadata — can be important to attest the authenticity of representations by elucidating about their provenance, promoting transparency of description procedures and holding the creators of the descriptions accountable. A more detailed list of every element that includes their definition is included in Appendix 5.

Reference must be also made to the elements that were chosen not to be included in the model. Regarding the requirements identified in the scientific literature, the suggestions about the inclusion of use rights for records and the *Curriculum Vitae* (CV) of curators were excluded.[1] As for the technical literature, it was decided not to include the elements referring to the target audience, the documentary form, the orientation/disposition of visualization and the origin of resources. Indicating a target-audience seems like a subjective and unnecessary prediction that, although not forgetting that many organizations serve a primary community, it is not absolute in the digital realm where resources may be retrieved in unexpected ways. It is also admitted that documentary form could be useful for the description of analog records, but harder to justify in the digital. Information about the orientation of resources

[1] For some metadata practitioners, the future of descriptive metadata for digital collections will be automated and machine actionable so the idea of focusing on individuals' professional backgrounds appears dated and unnecessary.

seem more appropriate for artworks, therefore questionable for archival description. Finally, the origin of the resource originates in a Spanish standard for recordkeeping in public administration, due to the need to distinguish between private and public records. Although sometimes useful, it was considered that this requirement was not essential for the demonstration of authenticity.

Conclusion

Metadata is understood as data about data that adds context and structure in human and machine readable form, so that it can be meaningfully interpreted and used by both humans and machines. It can elucidate about the identity and integrity of resources, thus strengthening the presumption of authenticity and consequent trustworthiness of archival descriptions. However, as a form of representation, archival description and, by extension, metadata, are not neutral. Every representation implies a simplification of the represented object, which means a choice of which elements to include or exclude from the representation. Thus, metadata is not objective, but rather a subjective human construct shaped by the context of its creator. Furthermore, authenticity is not an absolute concept, but equally subjective in the sense that it depends on the goals and the reference points of the user. What is authentic for one might not be enough for other, so authenticity should be seen as a nuanced concept that plays out across a grey scale. It is not the role of archivists to ensure authenticity, but merely to provide as much information as necessary so that users can form their judgement.

In a digital context in which information can be retrieved by multiple representation systems, it is key that no metadata is lost in translation, in order to preserve resources' context and meaning, as well as to maintain their trustworthiness. This requires a semantic interoperability of the metadata elements used to describe resources that is facilitated if information management communities articulate together to standardize description procedures. The sheer diversity of approaches, standards, and professional backgrounds in the information management field may create confusion, compromising the ability for researchers and practitioners to dialog, and for information systems to communicate. This article's proposal of a model for authenticity aimed to contribute to filling that gap, not only performing a thorough meta-analysis of the literature of the last decade, but also a mapping of how metadata standards have expressed authenticity, combining both inputs in a suggested a set of metadata elements firmly anchored in relevant literature. The model strived to capture the authenticity not only of resources, but also of the descriptions themselves as a postmodern acknowledgement of the archivist's subjectivity in the representation process. A compromise was made between simplicity and complexity, so that the elements included in the model followed a rationale of being as complete as necessary while as simple as possible. It can be used by data practitioners as a benchmark for application, as well as by archival scholars or standards researchers as a starting point for the discussion of authenticity and trustworthiness in digital descriptions.

Appendix 1: Bibliographic references of Table 1

These references are mentioned in Table 1 as input for the literature review, but are not cited in the main text.

Baños-Moreno, M., Valentín Ruiz, F., & Blázquez Martín de las Mulas, A. (2019). e-EMGDE, RIC, NEDA y las normas de codificación: más allá de un perfil de aplicación en archivos. *Atas Do IV Congresso ISKO Espanha-Portugal*

Botão, A. (2011). *Recuperação da informação digital: a Norma Brasileira de Descrição Arquivística (NOBRADE) na descrição de material imagético*. Universidade Federal do Rio de Janeiro

Brown, C. (Ed.). (2019). *Archival futures*. Facet Publishing

Dow, E. (2009). Encoded archival description as a halfway technology. *Journal of Archival Organization*, *7*(3), 108–115. https://doi.org/10.1080/15332740903117701

Dryden, J. (2011). Measuring Trust: Standards for Trusted Digital Repositories. *Journal of Archival Organization*, *9*(2), 127–130. https://doi.org/10.1080/15332748.2011.590744

Ducheva, D., & Pennington, D. (2019). RDA in Europe: Implementations and perceptions. *Journal of Librarianship and Information Science*, *51*(2), 387–402. https://doi.org/10.1177/0961000617709060

Duranti, L., & Franks, P. (Eds.). (2015). *Encyclopedia of archival science*. Rowman & Littlefield

Gracy, K., & Lambert, F. (2014). Who's ready to surf the next wave? A study of perceived challenges to implementing new and revised standards for archival description. *American Archivist*, *77*(1), 96–132. https://doi.org/10.17723/aarc.77.1.b241071w5r252612

Henttonen, P. (2009). A comparison of MoReq and SÄHKE metadata and functional requirements. *Records Management Journal*, *19*(1), 26–36. https://doi.org/10.1108/09565690910937227

Kallberg, M. (2012). Archivists 2.0: Redefining the archivist's profession in the digital age. *Records Management Journal*, *22*(2), 98–115. https://doi.org/10.1108/09565691211268162

Linden, L. (2017). *O tratamento temático da informação em instrumentos normativos de descrição arquivística,.96-124.* [Universidade Federal de Santa Catarina]. https://doi.org/10.19132/1808-5245243

MacNeil, H. (2012). What finding aids do: Archival description as rhetorical genre in traditional and web-based environments. *Archival Science*, *12*(4), 485–500. https://doi.org/10.1007/s10502-012-9175-4

Mak, B. (2012). On the uses of authenticity. *Archivaria*, *73*(Spring 2012), 1–17

Moro-Cabero, M., Martín-Pozuelo, M., & Zazo, J. (2011). ISO 15489 and other standardized management systems: Analogies and synergies. *Records Management Journal*, *21*(2), 104–121. https://doi.org/10.1108/09565691111152044

Riley, J. (2010). *Seeing Standards: A Visualization of the Metadata Universe*. https://doi.org/10.1080/07317131.2011.598072

Rogers, C., & Tennis, J. (2013). Authenticity as a Social Contract—We Are Our Records. *IConference 2013 Proceedings*, 801–803. https://doi.org/10.9776/13393

Rolan, G. (2017). Towards interoperable recordkeeping systems: A meta-model for recordkeeping metadata. *Records Management Journal*, *27*(2), 125–148. https://doi.org/RMJ-09-2016-0027

Zhang, J. (2012a). Archival context, digital content, and the ethics of digital archival representation. *Knowledge Organization*, *39*(5), 332–339. https://doi.org/10.5771/0943-7444-2012-5-332

Appendix 2: Summary of the requirements identified in the scientific literature, organized by theme

Category	Requirements
Information representation	The archivist should be able to explain archival description's rules and procedures (Duff et al. 2013), in an effort to be more transparent (Serewicz 2010)
	The need to include contextual information, namely about the records' creator, history and provenance (Duff et al. 2013)
	Records establish links not only with members of their aggregate but also to records in other collections and repositories (Duff et al. 2013)
	A good relationship between users and archivists should be nourished in order to facilitate retrieval (Duff et al. 2013)
	Archival description should occur not only at collection-level, but also item-level (Niu 2015a). Some possible metadata are 'title', 'subject', 'description', 'author', 'creator', 'receiver', 'contributor', 'date', 'placer' e 'material type' (Zhang 2012b)
	Intellectual control can be applied to content-level through the description of *datasets*, figures, tables or text annotations (Niu 2015a; Zhang 2012b)
	The digital environment enables resources to belong to more than a collection, and a collection to belong to multiple communities (Niu 2015a)
	Finding aids may be independent but interconnected in a single aggregator (Yeo 2012). Relationship metadata that indicate the origin of descriptions, such as 'contains' and 'derived from' should be included (Niu 2015a), as well as metadata regarding provenance at a creator level, elaborating on the 'administrative history' (Zhang 2012b)
	Finding aids must establish links to other resources through metadata structured as linked data (Serewicz 2010)
	Finding aids must be easy to use, preferably as websites, and it is recommended that they are built in a manner that is familiar for their users (Niu 2015a)
	Metadata should be captured and accumulated throughout the resources' lifecycle, either by the creator or by other organizations (Niu 2015a)
	Representation should be bottom-up, built upon the users' (Serewicz 2010)
	Representation systems should become more flexible and transversal, as a manifestation of the complexity of organisations (I. Silva 2012)
	Resources should have a unique, machine-readable identifier in order to facilitate metadata's automatic processing and to improve discovery and access to resources (Verborgh et al. 2015)
	Archival representations are subjective, socially-constructed practices that should be questioned. They are also a process that evolves over time, requiring a revisit and reinterpretation as context changes (Yakel 2003)
	Past archival descriptions should be preserved as a historical account and as an example of the evolution in representation narratives (Yakel 2003)
	The authors and the version of each finding aid should be identified (Yakel 2003)
	It is increasingly important to prioritize intellectual order over physical order (Yakel 2003; Zhang 2012b)

Category	Requirements
Archival description	According to the More Product Less Process methodology, it is suggested that description processes should be simplified with essential metadata in order to lower the description costs of large document volumes (Anchor 2013)
	Description must be centered on user needs (Anchor 2013)
	Description should be more honest, acknowledging the transformative processes that influence archive formation over time. Notes and/or citations of the creator of the description should be included, indicating creatorship (Douglas 2016; Meehan 2009) and even their CV (MacNeil 2009)
	Descriptions must be more transparent, possibly adding footnotes that document archivists' decisions, in addition to explaining description (Douglas 2016; MacNeil 2009; Meehan 2009)
	A significant focus should be placed on the custodial history of records. Suggestion of an autonomous description field, subdivisible in three elements: 'name(s) of the custodian(s)', 'custodial history', and 'Arrangement history of records' and related finding aids' (Douglas 2016). These should include a detailed log of the records' history until acquisition by the archival institution, including 'conversation history' (MacNeil 2009)
	The possibility of creating an independent field for creator's history with 'creator's name' and 'administrative history', as well as a field for archival history with the fields 'custodian name', 'custodial history', 'history of the arrangement of records' and 'history of the arrangement of finding aids' (MacNeil 2009)
	Archival descriptions are subjective constructs created by archivists in a given context (Douglas 2016; MacNeil 2009; 2012)
	The suggestion for users to contribute to archival description through comments, tags, annotations or reviews (Douglas 2016)
	Archival descriptions should include:
	File name, location in the storage system and technical hardware and software specifications that would assist in information retrieval (Hedstrom 1993). As for metadata, 'codification format' and 'physical material' should be included (Llanes-Padrón 2019). Physical changes to records should also be documented (MacNeil 2009)
	Information that documents the relationship between records' context and content (Hedstrom 1993; Meehan 2009)
	A log about records' destruction. This information may be included in the elements 'administrative history', 'appraisal' and 'planned destruction and eliminations' (Niu 2013)
	Information about records management throughout their life-cycle, including audit trails (Hedstrom 1993; Niu 2013)
	The inability of a single hierarchical structure to capture sufficiently the complexity of relationships that records establish during their use, as they may be interpretable across different contexts (Hedstrom 1993; Llanes-Padrón 2019; Meehan 2009; Zhang 2012c)
	Resources should be described independently, at item-level, in order to enable the elaboration of several aggregates by web users (Moyano Collado 2013; Zhang 2012c)
	Metadata should be added gradually and progressively throughout the records' life-cycle (Niu 2013)

Category	Requirements
Information management	Representation should occur at item-level in order to enable users to recontextualize resources during their individualized queries (Bailey 2013)
	Records can be interpreted in multiple ways since their meaning and context vary over time (Bailey 2013; Cumming 2010; MacNeil 2017)
	Archival description should be a permanent process throughout records' life-cycle (Cumming 2010)
	The physical conditions of materials should be documented as a critical part in the perception of their meaning, context and authenticity (Acker 2017)
	Recordkeeping challenges need to be tailored to organizational needs, and there is no single, universal and magical solution (McLeod 2014)
	Metadata should reflect, above all, the context of the activities of which records are a subproduct (McLeod 2014; Tough 2016; Wright 2014)
	The creator of the representation should be identified (Wright 2014)
	The object of representation in archives should be (Yeo 2011), to allow semantic relationships to be established with resources from other collections (Wright 2014)
	Archivists are active agents in the construction of the representation and their meanings (Ridener 2009)
Metadata	The identification of who owns the metadata, as well as a log of changes over time, are part of a basis for trustworthiness (Apostolou 2009)
	User annotations, comments and evaluations may be included to boost content and context (Apostolou 2009)
	Acknowledging that each community has different requirements and needs, therefore there is no universal metadata structure, nor there is a common ground on which to define 'essential' metadata (Apostolou 2009; Gladney 2007; Haynes 2018)
	Metadata creation should be seen as an incremental process with shared responsibilities (Baca 2016)
	In order to simplify and accelerate metadata management, their creation must be as automatic as possible (Baca 2016)
	Metadata should be created by organisations so that they are easily sharable and reusable (Baca 2016; Gartner 2016)
	Faceted metadata can be used to improve access (Gartner 2016) and to fulfill the query goal by simply looking at the results retrieved (Beyene and Godwin 2018)
	The relationship between provenance and authority of the creator, according to the demonstration of its identity, is highlighted:
	For the explanation of provenance, the following metadata elements are considered important: 'Title', 'Identifier', 'Format', 'Subject', 'Rights', 'Editor' (Chen et al. 2011), as well as 'Creation date e 'Description' (Haynes 2018)
	For descriptions to be complete, as many elements as possible should be included, namely 'Creator', 'Contributor', 'Format', 'Editor', 'Source' and 'Coverage' (Chen et al. 2011), in addition to 'responsible entity' and 'update frequency (Zeng and Qin 2016)
	For the assessment of the resources' integrity, it is suggested to include metadata such as 'source', 'relationship', 'version/edition' and 'digital signature' (Zeng and Qin 2016)
	Metadata that can serve audit trails should be included to encourage the verification of resources' provenance and integrity, ultimately reinforcing trust in institutions (Gladney 2007; Li and Sugimoto 2017)
	Provenance may also be strengthened with metadata that answer the following questions: Who created the metadata? When were they created or modified? What are the circumstances that led to their creation? (Haynes 2018). Accruals, elimination and substitution of metadata elements and their definitions should be included in the answers. (Li and Sugimoto 2017)
	Metadata about retention/elimination policies should be included, such as 'Elimination act', 'Temporal interval of the elimination', 'Expected date for the elimination' and 'Elimination authorized by' (Haynes 2018)
	Each resource should have one and only description, according to the one-to-one principle (Zeng and Qin 2016)

Category	Requirements
Metadata standards	The mapping of metadata across standards is not linear. There are several inconsistencies and imperfect or non-existent matches (Andrade et al. 2014; Woodley 2016)
	There is a distinction between archival description standards and information exchange standards. The former attempt to ensure records' authenticity across time, contexts and domains, whereas the latter serve as safety vests that dictate how information is shared (Bunn 2013)
	The most successful models tends to be free and are written clearly, ensuring that they are easily read and understood (Dryden 2009)
	The creation context of records should be captured, and descriptions should be shareable and reusable between several communities in the web (Llanes-Padrón and Moro-Cabero 2017; Llanes-Padrón and Pastor-Sánchez 2017)
	Metadata creation is seen as n iterative process throughout the life-cycle of the records. Metadata for appraisal, control, preservation, retrieval, access, use and an audit trail should be included as part of intellectual control (Mckemmish et al. 1999)
	If a standard includes too many metadata elements, it becomes harder to implement, generating costs in time and money (Youn 2015)
Authenticity	Mechanisms should be included to document the variability of digital objects through provenance metadata, fixating their physical integrity. Some possibilities are checksums or file size verification (Bountouri et al. 2017). An object's integrity may also be ascertained using cryptographic authentication (Gladney 2009; Rogers 2015)
	Provenance information, which are essential to document a record's life-cycle, should be collected since objects are created (Bountouri et al. 2017)
	The assertion of authenticity and trust are subjective acts shaped by various interpretations (Duncan 2009; Duranti et al. 2019; Engvall 2019; McLeod and Gormly 2017; Rogers 2015; Yeo 2013). A starting point may be to anchor the reference point of the "other" according to whom the authenticity degree is measured (Duncan 2009)
	The criteria used for trust in digital records are necessarily distinct from the principles that govern analog record creation. In the digital realm, enhanced transparency and accountability of records' creation and management processes, as well of their metadata, is required (Duranti et al. 2019; Engvall 2019; Rogers 2015; Yeo 2013)
	The explanation of the origin, provenance, chain of custody and a log of changes and audit are suggested to be indispensable aspects for transparency and to establish the grounds for trust (Duranti et al. 2019; Gladney 2009; Price and Smith 2011)
	Provenance information should be associated to the resources so that descriptions are self-explanatory (Gladney 2009). Some of the key metadata elements of an abstract concept of record, according to a Diplomatic perspective, might be: 'action', 'bond', 'addressee', 'writer', 'author', 'adminContext', 'digitalObject', 'ingestDate' (Jansen 2015)
	Each record, or even each version of each record, should have a unique identifier (Gladney 2009)
	Collections from various institutions should be aggregated together in order to enrich context and better reflect society (Price and Smith 2011)
	Metadata creation should be made as automatic as possible in order to save costs and make description viable (Yeo 2013)
	As for authenticity indicators, description should include information about the software that generated the records, about actions carried out on the records, retention and elimination schedules, audit traits and cryptographic validation techniques (Rogers 2015)
	Blockchain technology is pertinent to recordkeeping as a safeguard for the integrity and identity of records, since it documents provenance permanently (Bhatia and Wright de Hernandez 2019; Hofman et al. 2019; Lemieux 2016)

Category	Requirements
Linked data	Interoperability is facilitated by a higher granularity in descriptions (Gartner 2015; Jones 2018; Niu 2016)
	It is important to use unique identifiers to define entities and their relationships (Gartner 2015)
	Some metadata elements may be useful for aggregators of multiple repositories, such as those that describe the collection (identifiers, titles), the components ('physical location', 'material description', 'content description'), holding institutions, dates, relationships, rights, use impediments, language and subject (Gartner 2015)
	Each representation model (tabular, relational, hierarchical and graph) has intrinsic advantages and disadvantages, so that each has its own value in being implemented, according to context (Hooland and Verborgh 2014):
	The use of RDF as a metadata sharing format has the disadvantage of requiring a high number of triples for description, as well the problem of intellectual control and curation rights (Gartner 2015)
	Nevertheless, it is still advantageous to make descriptions more flexible by expressing metadata in a machine processable manner, for example through linked data that facilitates the integration of collections (Gartner 2015; Hooland and Verborgh 2014; Jones 2018; Machado et al. 2019; Niu 2016)
	The focus of description should be in the relationships built within networks, not in rigid hierarchies built upon documents, collections or organizations (Jones 2018)
	Knowledge should be made explicit in metadata that facilitates users to apprehend the meaning of resources, who should not be required to ask the archival staff for further information (Jones 2018). Users should also be able to build their own narratives and interpretations (Rolan 2015; Samouelian 2009)
	Descriptions are historical products of a certain socio-temporal context (Jones 2018; Rolan 2015)
	Archives should integrate tools that invite user interaction, such as blogs, wikis, ratings, reviews, podcasts and bookmarks (Samouelian 2009)
	Artificial intelligence might become a precious ally in information management (Rolan et al. 2019)

Appendix 3: Unique requirements identified in the scientific literature

#	Requirements	Topic
1	More transparency and accountability in description procedures. The creator of the metadata should be identified and provide explanations when necessary (Douglas 2016; Duff et al. 2013; Duranti et al. 2019; Engvall 2019; Haynes 2018; Li and Sugimoto 2017; MacNeil 2009; Meehan 2009; Rogers 2015; Serewicz 2010; Wright 2014; Yeo 2013)	IR, AD, IM, MD, AUT
2	To include contextual information about the records' creator, history and provenance, clarifying about custodial history (Chen et al. 2011; Douglas 2016; Duff et al. 2013; Duranti et al. 2019; Gladney 2009; Haynes 2018; MacNeil 2009; Niu 2015a; Price and Smith 2011; Zeng and Qin 2016; Zhang 2012b)	IR, AD, MD, AUT
3	Records establish relationships not only with members within their aggregate but also with records from other collections and repositories (Duff et al. 2013; Jones 2018; Price and Smith 2011; Serewicz 2010; I. Silva 2012; Yeo 2012)	IR, AUT, LD

#	Requirements	Topic
4	Records may belong to multiple aggregations and be interpreted in various contexts (Bailey 2013; Cumming 2010; Jones 2018; MacNeil 2017; Moyano Collado 2013; Niu 2015a; Rolan 2015; Samouelian 2009; Zhang 2012c)	IR, AD, IM, LD
5	Archival description should occur not only at collection but also item-level (Bailey 2013; Moyano Collado 2013; Niu 2015a; Zhang 2012b, 2012c)	IR, AD, IM
6	Finding aids should be user-friendly, built as websites (Niu 2015a)	IR
7	Metadata should be captured across the resources' life cycle. Description is a permanent, incremental process that requires constant updating. (Baca 2016; Cumming 2010; Mckemmish et al. 1999; Niu 2013, 2015a; Yakel 2003)	IR, AD, IM, MD, MDS
8	Representations must be built according to user needs (Anchor 2013; Serewicz 2010)	IR, AD
9	Unique identifiers should be used (Gartner 2015; Gladney 2009; Verborgh et al. 2015)	IR, AUT, LD
10	To acknowledge that archival representations are subjective, socially-constructed practices that can be questioned (Douglas 2016; Duncan 2009; Duranti et al. 2019; Engvall 2019; Jones 2018; MacNeil 2009, 2012; McLeod and Gormly 2017; Ridener 2009; Rogers 2015; Rolan 2015; Yakel 2003; Yeo 2013)	IR, AD, AUT, LD
11	To preserve past descriptions as a historic record of narratives (Yakel 2003)	IR
12	To identify the authors and the version of each finding aid, as well as the changes that metadata suffered over time, such as new elements, modifications or deletion. This includes changes to their definitions (Apostolou 2009; Li and Sugimoto 2017; Yakel 2003)	IR, MD
13	In the digital context, intellectual order should be prioritized over physical order (Yakel 2003; Zhang 2012b)	IR
14	Description should be complete but simple enough to be viable (Anchor 2013; Youn 2015)	AD, MDS
15	Users should be allowed to contribute to archival description through mechanisms such as comments, tags, notes or reviews (Apostolou 2009; Douglas 2016; Samouelian 2009)	AD, MD, LD
16	To include information about the appraisal of records, including elimination schedules (Haynes 2018; Niu 2013; Rogers 2015)	AD, MD
17	Mechanisms to verify integrity and provenance, such as audit trails, version log, digital signatures or checksums (Bountouri et al. 2017; Duranti et al. 2019; Gladney 2007, 2009; Hedstrom 1993; Li and Sugimoto 2017; Mckemmish et al. 1999; Niu 2013; Price and Smith 2011; Rogers 2015; Zeng and Qin 2016)	AD, MD, MDS, AUT
18	Archival description should elucidate about the physical characteristics and changes to records (Acker 2017; Hedstrom 1993; Llanes-Padrón 2019; MacNeil 2009)	AD, IM
19	To acknowledge that there is not a single universal structure for the representation of resources, and that a solution depends on the implementation context and the needs of each community (Apostolou 2009; Gladney 2007; Haynes 2018; McLeod 2014)	IM, MD
20	The main goal of metadata is to reflect the context of the activities of which the records are a sub-product (McLeod 2014; Tough 2016; Wright 2014)	IM

#	Requirements	Topic
21	The object of representation should be expanded beyond the traditional understanding of record (Wright 2014; Yeo 2011)	IM
22	Metadata creation should be as automated as possible (Baca 2016; Yeo 2013)	MD, AUT
23	Metadata should be created so that they can be easily sharable and reusable (Baca 2016; Gartner 2015, 2016; Hooland and Verborgh 2014; Jones 2018; Llanes-Padrón and Moro-Cabero 2017; Llanes-Padrón and Pastor-Sánchez 2017; Machado et al. 2019; Niu 2016)	MD, MDS, LD
24	It is advantageous to use facets in queries (Beyene and Godwin 2018; Gartner 2016)	MD
25	Each resource should have a single description, according to the 1:1 principle (Zeng and Qin 2016)	MD
26	Crosswalks between standards are often inconsistent (Andrade et al. 2014; Woodley 2016)	MDS
27	It is beneficial that models are free to use and are written simply, in an easy to understand language (Dryden 2009)	MDS
28	Provenance information should be collected since resources are created (Bountouri et al. 2017)	AUT
29	Resources should have enough associated information to become self-descriptive (Gladney 2009; Jansen 2015)	AUT
30	Blockchain technology has potential to aid in archival description (Bhatia and Wright de Hernandez 2019; Hofman et al. 2019; Lemieux 2016)	AUT
31	Artificial intelligence may be helpful (Rolan et al. 2019; Yeo 2013)	AUT, LD
32	The more granular descriptions are, the more interoperability is facilitated (Gartner 2015; Jones 2018; Niu 2016)	LD

Appendix 4: Unique requirements identified in the analysis of the technical literature

#	Requirement	ISAD(G)	AGRkMS	EAD	e-EMGDE	DACS	DCMES	VRA	MODS	CDWA	RiC
T1	Identifier	Reference code	Identifier	Unitid, agencycode, recordid	Identificador	Reference code	Identifier	–	Identifier	–	Identifier
T2	Title	Title	–	Unittitle	Nombre	Title	Title	Title	TitleInfo	Title text	Name
T3	Creator	Name of creator	Name	Origination	Órgano	Name of creator(s)	Creator, Contributor	Agent	–	Creator identity Creator role Creator description Person/corporate body authority	Agent
T4	Date	Dates	Date range	Unitdate, unitdate-structured	Fechas	Date	Date	Date	Dateissued DateCreated DateCaptured	Creation date	Date
T5	Extent	Extent	Extent	Physdesc, physdesc-structured	Tamaño	Extent	Format	Measurements	–	Measurements Dimensions description	Record resource extent
T6	Medium	Medium	Medium	–	Soporte	–	Format	–	–	Materials/techniques description	–
T7	Format	–	–	–	Formato	–	Format	–	–		–

#	Requirement	ISAD(G)	AGRkMS	EAD	e-EMGDE	DACS	DCMES	VRA	MODS	CDWA	RiC
T8	Physical characteristics	Physical characteristics and technical requirements	–	Phystech	Características técnicas	–	–	Material Technique	physicalDescription	Physical description Facture	Physical characteristics Production technique
T9	Target audience	–	–	–	–	–	–	–	TargetAudience		–
T10	Name and location of repository	–	–	–	–	Name and location of repository	Coverage	–	Name	Place/location authority	–
T11	Documentary form	–	Document form	–	Tipo documental	–	–	WorkType	TypeofResource	Object/Work Type	Documentary form type
T12	Orientation/Arrangement	–	–	–	–	–	–	–	–	Orientation/Arrangement	Structure
T13	Description level	Level of description	Category	Archdesc, c	Categoría	–	–	–	–	Catalog level	Record set type
T14	System of arrangement	System of arrangement	–	–	Clasificación	System of arrangement	–	–	Classification	Classification term	Classification
T15	Scope and content	Scope and content	Description	Scopecontent	Descripción	Scope and content	Subject	Subject Description	Abstract Subject	Subject matter General subject terms Context Descriptive note	Scope and content

#	Requirement	ISAD(G)	AGRkMS	EAD	e-EMGDE	DACS	DCMES	VRA	MODS	CDWA	RiC
T16	Administrative/bibliographic history	Administrative/biographical history	Coverage	Bioghist	Ámbito	Administrative/biographical history	–	Cultural-Context StylePeriod	Genre	Styles/Periods…	History
T17	Origin of the resource	–	–	–	Origen del documento	–	–	–	–	–	–
T18	Elaboration status	–	–	–	Estado de elaboración	–	–	–	–	State	State
T19	Location	Existence and location of originals, … of copies	Location	Originalsloc, altformavail	Localización	Existence and location of copies, existence and location of originals	–	Location	Location	Current location	Place
T20	Related resources	Related unit of description	Related entity	Relatedmaterial separatedmaterial	Entidad Relacionada	Related archival materials	Relation	Relation StateEdition textref	relatedItem	Related works Related visual documentation Related textual references	–
T21	Custodial history	Archival history	Change history	custodhist	Historia del cambio	Custodian history	Provenance	–	DateModified	Ownership/collecting history	–

#	Requirement	ISAD(G)	AGRkMS	EAD	e-EMGDE	DACS	DCMES	VRA	MODS	CDWA	RiC
T22	Appraisal	Appraisal, destruction and scheduling information	Disposal	Appraisal	Dictamen	Appraisal, destruction, and scheduling information	–	–	–	–	–
T23	Accruals	Accruals	–	Accruals	–	Accruals	accrualMethod, accrualPeriodicity, acccrualPolicy	–	–	–	Accrual
T24	Acquisition source	–	–	Acqinfo	–	Immediate source of Acquisition	–	–	–	–	–
T25	Integrity declarations	–	Integrity check	–	Verificación de integridad, Firma	–	–	Inscription	–	Inscriptions/marks	Integrity Authenticity note
T26	History of physical interventions	–	–	–	–	–	–	–	–	Condition/examination history Conservation/treatment history	–
T27	Creator of the description	–	–	Maintenancea-gency	–	Description control	Publisher	–	recordInfo	Cataloging history	–

#	Requirement	ISAD(G)	AGRkMS	EAD	e-EMGDE	DACS	DCMES	VRA	MODS	CDWA	RiC
T28	Date of the description	Dates of description	–	maintenanceevent, eventdatetime	–	Description control	–	–	RecordInfo	Cataloging history	–
T29	History of the description	–	–	Maintenance history, maintenance status	–	–	–	–	–	–	–
T30	Rules or conventions	Rules or conventions	–	Conventiondeclaration	–	Description control	–	–	–	–	–
T31	Sources of the description	–	–	–	–	Description control	Source	Source	RecordInfo	–	–
T32	Language of the description	–	–	–	–	–	–	–	RecordInfo	–	–
T33	Notes	Archivist's note	–	Processinfo	Notes	–	–	–	–	Cataloging history	–
T34	Contact	–	Contact (Agent)	–	Contacto	–	–	–	–	–	–

Appendix 5: Definitions of the metadata elements of the model.

Referring to	Element	Definition
1. Resource	1.1 Identifier	A unique identifier of the Resource, such as an URI
	1.2 Title	The name of the resource
	1.3 Date	Chronological information for context
	1.3.1 Creation date	The creation date of resources
	1.3.2 Date of the last change	The date of the last change to the resources
	1.3.3 Version	Information about the version of the document
	1.3.4 Elaboration state	Information about the transmission state (copy, original, authenticated copy)
	1.4 Creator	Name of the one responsible for the creation of resources
	1.5 Physical characteristics	Information relevant for physical handling
	1.5.1 Dimensions and medium	Information about the dimensions and materials of the documents
	1.5.2 Format	Information about file extensions
	1.5.3 Software and hardware	Information about technical characteristics that may assist in interpreting a file
	1.6 Description level	Identification of the description level
	1.7 Subject	A description of the content
	1.8 Language	Information about the language(s) used
	1.9 Location	Geographical information about the custody
	1.9.1 Repository's name	Official naming of the repository
	1.9.2 Custodian's name	Formal name of the custodian individual or organization
	1.9.3 Address	Geographical or digital (URI) location of the repository
	1.10 Representation form	Information about the internal structure and order, the arrangement system or about how to visualize information
	1.11 Custodial history	Information about custody transfers
	1.12 Arrangement history	Clarification of the several arrangements adopted over time
	1.13 Appraisal	Information about the interventions of archivists
	1.13.1 Elimination	Indications of which documents were deleted, and when
	1.13.2 Retention plans	Information about expected deletions or custody transfers
	1.14 Related resources	Identification of related resources
	1.15 Integrity declarations	Integrity verification mechanisms such as checksums, signatures, seals, stamps or marks
	1.16 Log of physical interventions	Indication of physical changes to the resources by custodians, such as conservation and restoration or digital preservation actions, such as migrations

Referring to	Element	Definition
2. Creator	2.1 Identifier	A unique identifier of a resource's creator
	2.2 Name	The official name of a resource's creator
	2.3 Biographical history	Biographical details of the creator that may contextualize resources
3. Aggregate	3.1 Accruals	Information about expected accruals
4. Meta-metadata	4.1 Metadata creator	Name of the creator of the metadata
	4.2 Metadata date	Useful chronological information for contextualizing representations
	4.2.1 Metadata creation date	Indication of the creation date of the description
	4.2.2 Metadata last change date	Indication of the date the description structure was last changed
	4.3 History of changes to the metadata	A record of changes to the description structure
	4.4 Rules/conventions	To make explicit the principles on which the description is based
	4.5 Sources	Information sources used in the description
	4.6 Metadata language	Language of the description
	4.7 Notes	Any further information considered pertinent by the creators of the description
	4.8 Contact	A contact of those responsible for the description

Author contributions The study was carried out by A.P. under the methodological guidance of C.G.S. and M.C.V.F. Data collection, interpretation and draft of the first version of the manuscript were performed by A.P. Subsequent versions were reviewed and improved by C.G.S and M.C.V.F. All authors have read and agreed to the published version of the manuscript.

Funding Open access funding provided by FCT|FCCN (b-on). The PhD thesis that preceded this article was funded by Fundação para a Ciência e Tecnologia (FCT) under the research grant SFRH/BD/131004/2017.

Declarations

Conflicts of interest The authors declare no conflict of interest.

References

Acker A (2017) When is a record?: a research framework for locating electronic records in infrastructure. In: Gilliland A, McKemmish S, Lau A (eds) Research in the Archival Multiverse. Monash University Publishing, pp 286–323

Anchor R (2013) 'More product, less process': method, madness or practice? Arch & Rec 34(2):156–174. https://doi.org/10.1080/23257962.2013.818937

Andrade M, Berti Júnior D, Pereira G, Baptista A (2014) Esquemas de metadados utilizados por repositórios digitais científicos e de objetos educacionais : estudo preliminar. Cadernos BAD 2:219–222

Apostolou C (2009) The role of metadata in managing knowledge. Loughborough University, UK

Baca M (ed) (2016) *Introduction to metadata* (3°). Getty Research Institute., California, USA

Bagley P (1968) Extension of programming language concepts. University City Science Center, Philadelphia, p 0223

Bailey J (2013) Disrespect des fonds: rethinking arrangement and description in born-digital archives. Arch Jour 3:201–212

Bearman D (1996) Item level control and electronic recordkeeping. Arch Mus Inf 10(3):195–245. https://doi.org/10.1007/bf02802369

Bennaceur A, Tun T, Yu Y, Nuseibeh B (2019) Requirements engineering. In: Cha S, Taylor R, Kang K (eds) Handbook of software engineering. Springer International Publishing. https://doi.org/10.1007/bf02802369

Bhatia S, Wright de Hernandez AD (2019) Blockchain is already here. what does that mean for records management and archives? J Arch Org 16(1):75–84. https://doi.org/10.1080/15332748.2019.1655614

Bountouri L, Gratz P & Sanmartin F (2017). Digital preservation: how to be trustworthy. *Conference of the Marie Sklodowska-Curie Initial Training Network for Digital Cultural Heritage (ITN-DCH)*, *10605*, 364–374. https://doi.org/10.1007/978-3-319-75826-8

Brewer J, Hunter A (2012) The multimethod approach and its promise. Found Multimet Res. https://doi.org/10.4135/9781412984294.n1

Brothman B (1993) The limits of limits: derridean deconstruction and the archival institution. Archivaria 36:205–220

Brown R (1991) The value of "narrativity" in the appraisal of historical documents: foundation for a theory of archival hermeneutics. Archivaria 32:152–156

Bunn J (2013) Developing descriptive standards: a renewed call to action. Arch Rec 34(2):235–247. https://doi.org/10.1080/23257962.2013.830066

Bureau of Canadian Archivists. (2008). *Rules for archival description*. Bureau of canadian archivists. http://www.cdncouncilarchives.ca/RAD/RADComplete_July2008.pdf (en);

Carta G (2017) Metadata and video games emulation: an effective bond to achieve authentic preservation? Rec Man Jour 27(2):192–204. https://doi.org/10.1108/RMJ-10-2016-0037Permanent

Chen Y-N, Wen C-Y, Chen H-P, Lin Y-H & Sum H-C (2011). Metrics for metadata quality assurance and their implications for digital libraries. In C. Xing, F. Crestani, & A. Rauber (Eds.), *digital libraries: for cultural heritage, knowledge dissemination, and future creation. ICADL 2011. lecture notes in computer science, vol 7008* (pp 138–147). Springer. https://doi.org/10.1007/978-3-642-24826-9_19

Clobridge A (2010) Building a digital repository program with limited resources. Chandos Publishing

Cook T (2001) Archival science and postmodernism: new formulations for old concepts. Arch Sci 1:3–24

Creswell J & Clark V (2013). *Pesquisa de métodos mistos* (2nd ed). Penso

Creswell J, Clark V (2018) *Designing and conducting mixed methods research* (3°). SAGE Publications, California, USA

Cron B (2016). Metadata in electronic records management. *Records Express*. https://records-express.blogs.archives.gov/2016/11/21/metadata-in-electronic-records-management/

Culkin JM (1967) A schoolman's guide to Marshall McLuhan. Saturday Rev 51–53:70–72

Cumming K (2010) Ways of seeing: contextualising the continuum. Rec Man Jour 20(1):41–52. https://doi.org/10.1108/09565691011036224

da Silva CG (2021) Investigação documental. In: Gonçalves S, Gonçalves J, Marques C (eds) Manual de investigação qualitativa: conceção, análise e aplicações. Pactor, pp 103–123

Daniel L & Daniel L (2012). *digital forensics for legal professionals: understanding digital evidence from the warrant to the courtroom*. Syngress.
Douglas J (2016) Toward more honest description. Am Arch 79(1):26–55. https://doi.org/10.17723/0360-9081.79.1.26
Dryden J (2009) The open archival information system reference model. J Arch Org 7(4):214–217. https://doi.org/10.1080/15332740903334116
Duff W, Yakel E, Tibbo H (2013) Archival reference knowledge. Am Arch 76(1):68–94
Duncan C (2009) Counterpoint: authenticity or bust. Archivaria 68:97–118
Duranti L (1997) The Archival Bond. Arch Museum Inform 11(3):213–218. https://doi.org/10.1023/A:1009025127463
Duranti L & Blanchette J (2004). *The authenticity of electronic records: the InterPARES approach*. http://www.interpares.org/display_file.cfm?doc=ip1-2_dissemination_cpr_duranti~blanchette_ist_2004.pdf
Duranti L, Rogers C, Mcleod J, Shepherd E, Foscarini F, Shabou B, Franks P, Stančić H, Michetti G, Cunningham A, Oliver G, Engvall T, LemieuxV, Hofman D, & Pearce-Moses R (2019). *trusting records and data in the cloud: the creation, management, and preservation of trustworthy digital content*. https://doi.org/10.29085/9781783304042
Eastwood T (1994) What is archival theory and why is it important? Archivaria 1(37):122–130. https://doi.org/10.1590/198053142851
Elings M, Waibel G (2007) Metadata for all: descriptive standards and metadata sharing across libraries archives and museums. First Monday 12(3):23. https://doi.org/10.5210/fm.v12i3.1628
Engvall T (2019) Situating trust challenges of online trade. Rec Man Jour 29(1/2):272–286
Factor, M., Henis, E., Naor, D., Rabinovici-Cohen, S., Reshef, P., & Ronen, S. (2009). Authenticity and provenance in long term digital preservation: modeling and implementation in preservation aware storage. *Workshop on Theory and Practice of Provenance*. https://www.usenix.org/legacy/event/tapp09/tech/full_papers/factor/factor.pdf
Furner J (2020) Definitions of metadata: a brief survey of international standards. J Asso Inf Sci Tech 71(6):E33–E42. https://doi.org/10.1002/asi.24295
Garshol LM (2004) Metadata? thesauri? taxonomies? topic maps! making sense of it all. J Inf Sci 30(4):378–391. https://doi.org/10.1177/0165551504045856
Gartner R (2015) An XML schema for enhancing the semantic interoperability of archival description. Arch Sci 15(3):295–313. https://doi.org/10.1007/s10502-014-9225-1
Gartner R (2016) Metadata: shaping knowledge from antiquity to the semantic web. Springer
Gilliland A (2016). Setting the stage. In M. Baca (Ed.), *Introduction to Metadata* (3rd ed, pp. 1–20). Getty Research Institute.
Gladney H (2007) Preserving digital information. Springer
Gladney HM (2009) Long-term preservation of digital records: trustworthy digital objects. Am Arch 72(2):401–435
Greene M (2009) The power of archives: archivists' values and value in the postmodern age. Am Arch 72(1):13–41. https://doi.org/10.17723/aarc.72.1.k0322x0p38v44l53
Gueguen G, Marques M, Fonseca D, Pitti D, Sibille-De Grimoüard C (2013) Toward an international conceptual model for archival description: a preliminary report from the international council on archives' experts group on archival description. Am Arch 76(2):566–583
Harris V (1997) Claiming less, delivering more: a critique of positivist formulations on archives in South Africa. Archivaria 44:132–141
Haynes D (2018) *Metadata for information management and retrieval* (2°). Facet Publishing, UK
Hedstrom M (1993) Descriptive practices for electronic records: deciding what is essential and imagining what is possible. Archivaria 36:23
Hofman D, Lemieux V, Joo A, Batista DA (2019) The margin between the edge of the world and infinite possibility: blockchain, GDPR and information governance. Rec Man Jourl 29(1–2):240–257. https://doi.org/10.1108/RMJ-12-2018-0045
InterPARES 2 (2008). The InterPARES 2 Project Glossary. In L. Duranti & R. Preston (Eds.), *International Research on Permanent Authentic Records in Electronic Systems (InterPARES) 2: Experiential, Interactive and Dynamic Records*. Associazione Nazionale Archivistica Italiana.
InterPARES 1 (2002). *Requirements for assessing and maintaining the authenticity of electronic records*.
International Council on Archives. (1996). *Code of ethics*.

International Council on Archives. (1997). *Guide for managing electronic records from an archival perspective* (Issue February). ICA Studies. http://www.ica.org/sites/default/files/ICA Study 8 guide_eng.pdf

International Council on Archives. (2008). *Principles and functional requirements guidelines and functional requirements for records in electronic office environments - module 2: guidelines and functional requirements for electronic records management systems*. www.ica.org

International Council on Archives. (2016). *Records in contexts. A conceptual model for archival description. Consultation draft v0.1*. https://doi.org/10.1017/S0940739197000234

International Council on Archives (2019). *Records in contexts: a conceptual model for archival description (Consulation Draft v0.2)*. https://www.ica.org/sites/default/files/ric-cm-0.2_preview.pdf

ISO/IEC TR 19583–1 (2019). *Information technology—concepts and usage of metadata—part 1: metadata concepts*. ISO. https://www.iso.org/obp/ui#iso:std:iso-iec:tr:19583:-1:ed-1:v1:en

Janes A (2012) Of maps and meta-records: eighty- five years of map cataloguing at the national archives of the United Kingdom. Archivaria 74:119–165

Jansen A (2015) Object-oriented diplomatics: using archival diplomatics in software application development to support authenticity of digital records. Rec Man Jour 25(1):45–55. https://doi.org/10.1108/RMJ-04-2014-0022

Jick T (1979) Mixing qualitative and quantitative methods: triangulation in action. Adm Sci Quar 24(4):602–611

Jimerson R (2015) Archives and memory. In: Duranti L, Franks P (eds) Enc of Arch Sci. Rowman & Littlefield, pp 99–101

Jones M (2018) From catalogues to contextual networks: reconfiguring collection documentation in museums. Arch Rec 39(1):4–20. https://doi.org/10.1080/23257962.2017.1407750

Koch, T. (2006). Electronic thesis and dissertation services: semantic interoperability, subject access, multilinguality. In *E-Thesis Workshop*. http://www.ukoln.ac.uk/ukoln/staff/t.koch/publ/e-thesis-200601.html

Lemieux V (2016) Trusting records: is Blockchain technology the answer? Rec Man Jour 26(2):110–139. https://doi.org/10.1108/RMJ-12-2015-0042

Li C, Sugimoto S (2017) Provenance Description of Metadata Vocabularies for the Long-term Maintenance of Metadata. J Data Inf Sci 2(2):41–55. https://doi.org/10.1515/jdis-2017-0007

Llanes-Padrón D (2019) La descripción archivística y la curaduría digital: relaciones y perspectivas. Atas Do IV Congresso ISKO Espanha-Portugal. https://doi.org/10.31229/osf.io/265vr

Llanes-Padrón D, Moro-Cabero M (2017) Records in contexts: A new model for the representation of archival information in semantic web scenarios. Prof De La Infn 26(3):525–533. https://doi.org/10.3145/epi.2017.may.19

Lubas R, Jackson A, Schneider I (2013) The metadata manual: a practical handbook. Chandos Publishing, UK

Machado LMO, Souza RR, Simões M (2019) Semantic web or web of data? a diachronic study (1999 to 2017) of the publications of tim berners-lee and the world wide web consortium. J Asso Inf Sci Tech 70(7):701–714. https://doi.org/10.1002/asi.24111

MacNeil H (2000) Trusting records: legal, historical, and diplomatic perspectives. Kluwer Academic Publishers

MacNeil H (2005) Picking our text: archival description, authenticity, and the archivist as editor. Am Arch 68:264–278

MacNeil H (2009) Trusting description: authenticity, accountability, and archival description standards. J Arch Organ 7(3):89–107. https://doi.org/10.1080/15332740903117693

MacNeil H (2017) Deciphering and Interpreting an Archival Fonds and Its parts. In: Gilliland A, McKemmish S, Lau A (eds) Research in the Archival Multiverse. Monash University Publishing, pp 161–197

Mak B (2015) Authenticity. In: Duranti L, Franks P (eds) Enc of Arch Sci. Rowman & Littlefield, pp 119–122

McKemmish S, Acland G, Ward N, Reed B (1999) Describing records in context in the continuum: The Australian Recordkeeping Metadata Schema. Archivaria 48(1):3–37. https://doi.org/10.4225/03/580562c352056

McLeod J (2014) Reinventing archival methods: reconceptualising electronic records management as a wicked problem. Arch Man 42(2):193–196. https://doi.org/10.1080/01576895.2014.911687

McLeod J, Gormly B (2017) Using the cloud for records storage: issues of trust. Arch Sci 17(4):349–370. https://doi.org/10.1007/s10502-017-9280-5
Meehan J (2009) Making the leap from parts to whole: Evidence and inference in archival arrangement and description. Am Arch 72(1):72–90. https://doi.org/10.17723/aarc.72.1.kj672v4907m11x66
Moyano Collado J (2013). La descripción archivística. De los instrumentos de descripción hacia la web semántica. *Anales de Doc*, *16*(2). https://doi.org/10.6018/analesdoc.16.2.171841
NARA (2015). *Bulletin 2015–04: Metadata guidance for the transfer of permanent electronic records*. https://www.archives.gov/records-mgmt/bulletins/2015/2015-04.html
Nesmith T (1999) Still Fuzzy, But More Accurate: Some Thoughts on the "Ghosts" of Archival Theory. Archivaria 47:136–150
NISO (2004). *Understanding metadata*. NISO Press.
NISO. (2017). *Understanding metadata: what is metadata, and what is it for?* NISO Press. http://www.niso.org/apps/group_public/download.php/17446/Understanding Metadata.pdf
Niu J (2013) Recordkeeping metadata and archival description: A revisit. Arch Man 41(3):203–215. https://doi.org/10.1080/01576895.2013.829751
Niu J (2015) Archival intellectual control in the digital age. J Arch Org 12(3–4):186–197. https://doi.org/10.1080/15332748.2015.1154747
Niu J (2016) Linked data for archives. Archivaria 82:83–110
Padrón D L & Cabero M M M (2019). *La descripción archivística y la curaduría digital: relaciones y perspectivas.* https://doi.org/10.31229/osf.io/265vr
Pastor-Sánchez J-A, Llanes-Padrón D (2017) Records in context: el camino de los archivos hacia la interoperabilidad semántica. Anuario ThinkEPI 11:297. https://doi.org/10.3145/thinkepi.2017.56
Pomerantz J (2015) Metadata. MIT Press
Price DM, Smith JJ (2011) The trust continuum in the information age: a canadian perspective. Arch Sci 11(3–4):253–276. https://doi.org/10.1007/s10502-011-9148-z
Ramesh P, Vivekavardhan J, Bharathi K (2015) Metadata diversity interoperability and resource discovery issues and challenges. DESIDOC J Lib Inf Technol 35(3):193–199. https://doi.org/10.14429/djlit.35.3.8074
Ridener J (2009) From polders to postmodernism: a concise history of archival theory. Litwin Books
Rogers C (2015). *Virtual authenticity: authenticity of digital records from theory to practice* (Issue April) [University of British Columbia]. https://doi.org/10.14288/1.0166169
Rolan G (2015) Towards archive 20: Issues in archival systems interoperability. Arch Man 43(1):42–60. https://doi.org/10.1080/01576895.2014.959535
Rolan G, Humphries G, Jeffrey L, Samaras E, Antsoupova T, Stuart K (2019) More human than human? Artificial intelligence in the archive. Arch Man 47(2):179–203. https://doi.org/10.1080/01576895.2018.1502088
Rosenberg J (1981) Linguistic representation. D. Reidel Publishing Company, Netherland
Samouelian M (2009) Embracing web 2.0: Archives and the newest generation of web applications. Am Arch 72(1):42–71
Schaeffer B, Baranski B, Foerster T, Brauner J (2012) A service-oriented framework for real-time and distributed geoprocessing. Lect Not Geoinf Cart 199579:3–20. https://doi.org/10.1007/978-3-642-10595-1_1
Serewicz L (2010) Do we need bigger buckets or better search engines?: the challenge of unlimited storage and semantic web search for records management. Rec Man Jour 20(2):172–181. https://doi.org/10.1108/09565691011064313
Silva I (2012). *A organização e a representação do conhecimento no domínio da arquivística*. Universidade Estadual Paulista.
Society of American Archivists (2005). *Glossary of archival and records terminology*. http://www2.archivists.org/glossary/terms/f/fonds
Society of American Archivists (2012). *Society of American Archivists core values statement and code of ethics*. https://www2.archivists.org/statements/saa-core-values-statement-and-code-of-ethics
Stamou S & Efthimiadis E (2010). Interpreting User inactivity on search results. *advances in information retrieval. ECIR 2010. Lect Not in Comp Sci,l 5993*. https://doi.org/10.1007/978-3-642-12275-0_12
Sudmanns M, Tiede D, Lang S, Baraldi A (2018) Semantic and syntactic interoperability in online processing of big Earth observation data. Int J Dig Earth 11(1):95–112. https://doi.org/10.1080/17538947.2017.1332112
Suvak W (2015) Authentication. In: Duranti L, Franks P (eds) Enc of Arch Sci. Rowman & Littlefield, pp 116–119

Tennant R (2001) Different Paths to Interoperability. Lib Jour 126(3):118–119
The National Archives of the UK (2002). *Functional requirements for electronic records management systems-reference document*.
Tough A (2016) Thinking about and working with archives and records: a personal reflection on theory and practice. Arch Rec 37(2):225–238. https://doi.org/10.1080/23257962.2016.1147343
Upward F (2000) Modelling the continuum as paradigm shift in recordkeeping and archiving processes, and beyond—a personal reflection. Rec Man Jour 10(3):115–139. https://doi.org/10.1108/EUM0000000007259
van Hooland S, Verborgh R (2014) Linked data for libraries, archives and museums. Facet Publishing Production, UK
Verborgh R, Van Hooland S, Cope AS, Chan S, Mannens E, Van De Walle R (2015) The fallacy of the multi-API culture conceptual and practical benefits of representational state transfer. J Doc 71(2):233–252. https://doi.org/10.1108/JD-07-2013-0098
Wartofsky M (1979) Models: representation and the scientific understanding, 129. D. Reidel Publishing Company, Dordrecht, Netherlands
Witten I, Bainbridge D, Nichols D (2010) How to build a digital library. Elsevier, Morgan Kaufmann
Woodley M (2016). Metadata matters: connection people and information. In M. Baca (Ed.), *Introduction to Metadata* (3° ed,). Getty Research Institute.
Wright K (2014) Broadening the record and expanding the archives. Arch Man 42(2):219–221. https://doi.org/10.1080/01576895.2014.911693
Yakel E (2003) Archival representation. Arch Sci 3(1):1–25. https://doi.org/10.1007/BF02438926
Yeo G (2011) Rising to the level of a record? some thoughts on records and documents. Rec Man Jour 21(1):8–27. https://doi.org/10.1108/09565691111125071
Yeo G (2012) Bringing things together: aggregate records in a digital age. Archivaria 74:43–91
Yeo G (2018) Records, information and data. Facet Publishing, UK
Yeo G (2013). Trust and context in cyberspace. In *Arch & Rec* (Vol. 34, Issue 2, pp. 214–234). Taylor & Francis. https://doi.org/10.1080/23257962.2013.825207
Youn E (2015) Adoption of ISAD(G) in practice: a close look at the standardization process of ISAD(G) in a manuscript archives of Korea. Arch Rec 36(2):128–145. https://doi.org/10.1080/23257962.2015.1029892
Zeng M L & Chan L M (2010). Semantic Interoperability. In M. Bates & M. Maack (Eds.), *Enc of Lib and Inf Sci. Volume I* (3rd ed., pp. 4645–4662). Dekker Encyclopedias, Taylor and Francis Group.
Zeng ML, Qin J (2016) *Metadata* (2°). Facet Publishing, UK
Zhang J (2012b) Archival representation in the digital age. J Arch Org 10(1):45–68. https://doi.org/10.1080/15332748.2012.677671
Zhang J (2012c) Original order in digital archives. Archivaria 74:167–193

André Pacheco is a Ph.D. in Information Science (2022) from the University of Coimbra. He is a data curator at Minho University and a researcher at Center for Classical Studies of the University of Lisboa. He is particularly interested in metadata, archival description, trustworthiness and authenticity. He recently authored a book on archival description intitled "Arquivos Digitais: Metadados e Autenticidade".

Carlos Guardado Da Silva is a Ph.D. in Medieval History (2004), from the University of Lisbon, and Aggregate in Information Science (2016) by the University of Coimbra. He's researcher at Center for Classical Studies of the University of Lisboa, professor and director of the Master Degree in Documentation and Information Sciences, at School of Arts and Humanities of the University of Lisbon. He's also member of the ISKO (International Society for Knowledge Organization)—Iberian Chapter. His research interests are Information Science, History and heritage.

Maria Cristina Vieira De Freitas is a Ph.D. in Library Science and Documentation (2010) from the University of Salamanca. She is an assistant professor at University of Coimbra, as well as Director of the Archive of the University of Coimbra (2019—present), integrated member of the CEIS20 Research

Center (University of Coimbra), and collaborating member of the ARBIDOC Group (University of León), and the Teresa Andrés Group (University of Salamanca). Her primary research interests are archival theories and methods, archival information organization, knowledge management and qualitative research (with emphasis on software for data collection and analysis).

Authors and Affiliations

André Pacheco[1] · Carlos Guardado Da Silva[1] · Maria Cristina Vieira De Freitas[2]

✉ André Pacheco
a.pacheco@campus.ul.pt

Carlos Guardado Da Silva
carlosguardado@campus.ul.pt

Maria Cristina Vieira De Freitas
cristina.freitas@fl.uc.pt

[1] Faculdade de Letras, Centro de Estudos Clássicos, Universidade de Lisboa, Lisbon, Portugal

[2] Faculdade de Letras, Centro de Estudos Interdisciplinares Do Século XX, Universidade de Coimbra, Coimbra, Portugal