

# **Pacheco, A. (2022). *Arquivos Digitais – Metadados e Autenticidade*. Edições Colibri.**

MADALENA LOPES DAMIÃO RODRIGUES

Mestranda em Ciências da Informação e Documentação

Faculdade de Letras, Universidade de Lisboa

[mdamiao@campus.ul.pt](mailto:mdamiao@campus.ul.pt)

ORCID: <https://orcid.org/0009-0001-9208-9713>

No mundo atual, a transição para a utilização de tecnologias e o digital expandiu-se a nível global. As instituições da informação acompanharam este fenómeno, em particular para conseguir responder às novas necessidades de comunicar e informar dos seus utilizadores. No âmbito arquivístico, isto também foi uma realidade, surgindo, neste quadro, as bases de dados e os arquivos digitais. No entanto, esta transição e as mudanças que implicam colocam novos problemas ou necessidades que precisam de uma resposta diferente e que exigem da Arquivística uma mudança de paradigma de forma que possa continuar a acompanhar aquilo que tanto os presentes como os futuros utilizadores possam exigir, nomeadamente no que respeita à autenticidade dos dados e textos digitais.

No mundo que vive intimamente ligado ao digital, surge a preocupação da preservação da informação digital, algo que não se pode dissociar dos metadados (Formenton & Gracioso, 2022), elementos-base para a garantia do acesso ao conteúdo do documento de arquivo, permitindo encontrá-lo, manejá-lo e rastreá-lo, inclusive quando da criação de novas versões. Ora, a forma de garantir a preservação e a confiabilidade ou integridade dos documentos de arquivo em formato eletrónico é a autenticidade (Lee et al., 2002; Gracy & Kahn, 2012), sendo, por isso, esta uma das maiores preocupações dos arquivistas na era digital. É relevante que se refiram ambas as coisas em paralelo, dado o seu papel no futuro dos arquivos digitais. Esse futuro não pode descurar princípios como o princípio da interoperabilidade, tanto de documentos como de sistemas, isto é, semântica e técnica, algo

que se pode obter através dos metadados com aquilo que comumente se designa por “vínculo arquivístico”, que ainda é considerado o foco da arquivística (Stančić & Bralić, 2021). Além disso, o próprio princípio paradigmático da arquivística em termos de acesso à informação deverá ser mantido no foco da disciplina, se se pretender que os utilizadores mantenham satisfeitas as suas necessidades de informação, bem como para as instituições tutelares dos documentos cumprirem a sua função, do ponto de vista social, enquanto instituições de informação e de memória.

É neste contexto que podemos situar o livro em apreciação, que tem como base a tese de doutoramento do autor, André Pacheco, apresentada originalmente à Faculdade de Letras da Universidade de Coimbra, em 2021, com o título *Metadados para a descrição arquivística digital: Proposta de um modelo para a autenticidade*. O seu trabalho, adaptado para uma versão em livro, editado na coleção “Ciência da Informação”, que é o objeto desta recensão, foca-se, precisamente, no problema da garantia da presunção de autenticidade documental nos arquivos digitais através dos metadados. André Pacheco é doutorado em Ciência da Informação pela Universidade de Coimbra e o tema que apresenta neste trabalho é, precisamente, o foco da sua especialização enquanto cientista da informação. É também investigador no centro de Estudos Clássicos da Universidade de Lisboa, com experiência em investigação académica, e colabora em projetos europeus e no desenvolvimento de soluções em contexto empresarial.

O livro estrutura-se em cinco capítulos. No primeiro, discute-se o conceito de “metadados” desde o seu aparecimento em 1968 e apresenta-se uma síntese relativa aos tipos de metadados existentes na literatura científica e à terminologia das normas de metadados. Identificam-se, desde logo, quatro tipos principais de normas. Este é um capítulo essencial na medida em que toca o problema da gestão da informação no âmbito da interoperabilidade de sistemas, uma questão intimamente relacionada com a normalização e a acessibilidade dos documentos e dos sistemas de arquivo.

No segundo capítulo, o foco está na forma como se representa a informação nos arquivos, sendo que essa representação abrange tanto a classificação como a descrição documental. Após apresentar a evolução da representação da informação desde a Antiguidade até ao final do século XX, o autor aborda a passagem de um paradigma moderno para um pós-moderno de forma que a Arquivística possa responder aos desafios colocados pela Era digital. Aqui encontra-se também uma reflexão sobre os objetos digitais e a descrição arquivística, na medida em que o princípio da proveniência não é suficiente, devendo-se também considerar os contextos. Neste segun-

do capítulo, destaca-se aquilo que será um dos dois pontos fundamentais da obra quando, no lugar de “vínculo arquivístico” ou como uma ampliação do conceito, ao autor propõe o termo “vínculo informacional” (*informational bond*) (pp. 138-142).

No terceiro capítulo — e aquele de maior interesse para efeitos desta recensão —, o da metodologia, o autor explica os métodos e procedimentos que seguiu na sua investigação, regressando ao ponto de partida das “questões de investigação” e aos seus objetivos. O objetivo geral do autor consiste na exploração da literatura científica e técnica de forma a demarcar não só os conceitos basilares da presunção da autenticidade dos documentos de arquivo, como também que elementos dos metadatos devem estar incluídos nas normas de metadatos que representam a dinâmica da informação arquivística na *web*, recorrendo a um modelo de dados em grafo. Delimitados os objetivos, os resultados esperados seriam: compreender qual o papel da descrição arquivística na verificação da (presunção da) autenticidade dos documentos de arquivo, ou das suas representações; e desenvolver um modelo de metadatos com aqueles que foram identificados na literatura como relevantes para uma representação arquivística autêntica na *web*.

Com este objetivo geral e estes resultados em mente, André Pacheco assume o paradigma interpretativo, também designado qualitativo ou construtivista. Este paradigma considera a existência de uma realidade múltipla, no sentido de a “realidade social” ser uma construção própria dos seus atores, sendo estes agentes não neutros. De acordo com esta perspetiva, os documentos são também construções sociais, exigindo uma interpretação e desconstrução por parte de quem os investiga, não podendo ser considerados de forma “isolada”.

Em seguida, são apresentados os métodos de investigação e de recolha de dados. O autor recorre à investigação documental e à engenharia de requisitos, por considerar que estes são os métodos que concedem o carácter qualitativo à metodologia de investigação, apesar de também apresentar componentes quantitativos. Assim, a metodologia utilizada é, na verdade, mista, com ambas as partes a complementarem-se e interligarem-se para cumprir o objetivo enunciado. Além disso, importa referir que a abordagem utilizada na revisão da literatura científica e técnica foi sistêmica e indutiva, seguindo o paradigma qualitativo. Sumariando: André Pacheco partiu dos dados obtidos através da investigação documental e inferiu uma teoria a partir desses dados; após a formulação dessa teoria, utilizou a engenharia de requisitos para a desconstruir ao analisar normas de metadatos; esses dados, por sua vez, foram utilizados para criticar e refletir sobre a teoria inicialmente

proposta, e, no fim, para a construção de um modelo de metadados que permite reforçar a presunção de autenticidade na descrição arquivística.

Em relação à investigação documental, termo cunhado por Carlos Guardado da Silva na língua portuguesa, também designada análise documental (Bowen, 2009), é um método ou processo de investigação que se baseia na análise de documentos com informação relevante para a investigação em curso, sendo os documentos a fonte de informação. Pode incluir uma abordagem qualitativa, quantitativa ou mista, e sendo um método abrangente para as investigações. No entanto, apresenta as suas vantagens e desvantagens (Bowen, 2009; Ahmed, 2010), algo reconhecido pelo autor que se mostra bastante ciente do método que utiliza. Como vantagem, por exemplo, tem a seu favor a não-reatividade dos documentos (não são feitos de propósito para a investigação, não sendo, assim, parciais); já como desvantagem apresenta, por exemplo, a dificuldade na recuperação da informação que pode estar pouco acessível. Relewa também dizer que a investigação documental presume a heurística e a hermenêutica dos documentos, enaltecendo a natureza interpretativa das fontes da informação, e também em conformidade com a perspetiva oferecida pelo paradigma construtivista que o autor adotou. No caso concreto do autor, André Pacheco reuniu literatura científica e técnica onde recolheu os seus dados, realizando, desta forma, uma revisão da literatura, que antecedeu a formulação da sua teoria, a sua desconstrução, e a proposta de um modelo de metadados.

Uma vez que a investigação documental tem, no seu âmago, a recolha e a seleção de determinadas fontes documentais, importa destacar o *corpus documental* usado na investigação. André Pacheco não se esquece de focar de forma bastante exaustiva as fontes que usou: doze revistas científicas, artigos científicos publicados entre 1 de janeiro de 2009 e 21 de dezembro de 2019, e ainda revisões de monografias. Em relação às primeiras fontes, as revistas científicas foram identificadas na base de dados LISA (*Library & Information Science Abstracts*) por considerar que esta estava mais completa face a outras, tendo sido realizada uma pesquisa nesta base com recurso a filtros qualitativos (seleção de títulos, o *Índice Compuesto de Difusión Secundaria* ser, no mínimo, 6.5, ser *peer-reviewed*, estar em atividade, ser numa língua que o investigador possa ler, estar enquadrada na investigação, disponibilizar informação descritiva) que levaram, então, aos 12 títulos utilizados. Já os artigos pertenciam a essas revistas que foram recuperadas, no período temporal enunciado, e relevantes para a investigação segundo o seu título e o *abstract*. Por fim, as monografias também foram selecionadas por estarem indicadas nestas revistas que foram recuperadas. De seguida, o autor

analisou os dados recolhidos, elaborando categorias de análise relacionadas com a investigação em curso. Relembra-se que a investigação documental foi, na investigação, utilizada de forma mista: num primeiro momento, de natureza mais qualitativa; num segundo momento, de âmbito quantitativo. Assim, a investigação documental acabou por ser uma grande base da investigação de André Pacheco, permitindo a elaboração de um *corpus documental* e a análise de dados.

Ainda que se possa considerar que a investigação documental possa não bastar por si só, pois implica uma base mais teórica, o autor não descarta a parte empírica, e completa o seu aparelho metodológico com a engenharia de requisitos. Este segundo método implica a combinação de um formalismo para representar e analisar requisitos e processos que apoiam e orientam o utilizador desse formalismo através de, por exemplo, a identificação de requisitos, e a sua verificação (Jureta, 2012). Ao empregar este método, o autor pretende conseguir identificar os elementos de metadados que são necessários à verificação de autenticidade, de forma a, mais tarde, na investigação e após a análise desses elementos e das normas sobre os mesmos, formular um modelo de metadados.

Na continuação do seu trabalho, e passando ao quarto capítulo do livro, encontra-se aqui, então, a identificação dos requisitos de metadados a partir da revisão da literatura científica e técnica. André Pacheco analisa nove categorias — representação da informação, descrição arquivística, gestão da informação, metadatos, normas de metadados, autenticidade, *linked data* —, indicando para cada uma delas o conjunto de requisitos que considera essenciais para o modelo que depois propõe. Relativamente à literatura técnica, complementando a revisão de literatura científica que foi realizada e já referida sumariamente, que é discriminada neste momento, o autor examinou dez normas, locais e internacionais, com o intuito de respeitar a necessidade de a investigação ser o mais abrangente possível ao mesmo tempo que se mantém exequível, de descrição arquivística: ISAD(G), AGRkMS, EAD, e-EMGDE, DACS, DCMES, VRA Core, MODS, CDWA, RiC. Cada uma destas normas é contextualizada historicamente, identificando-se a estrutura dos elementos de metadados relativos a cada uma delas, e ainda se indica quais desses elementos é que são considerados obrigatórios pelos autores das normas. Em seguida, o autor identifica, para cada uma das normas apresentadas, os metadados potencialmente pertinentes para proceder à verificação de autenticidade.

Chegando ao último capítulo do livro de André Pacheco, é aqui sistematizada a apresentação dos requisitos previamente identificados. O autor expõe em tabela (Tabela 12, pp. 250-252) os trinta e dois requisitos científicos que

foram identificados, seguindo-se depois os requisitos técnicos (Tabela 14, pp. 260-262). Com bases nestes requisitos, que são, finalmente, apresentados, André Pacheco propõe um modelo de averiguação da autenticidade dos metadados, que ilustra (p. 270). É esta proposta que aparece no final da obra que pretende ser uma possibilidade de verificação da autenticidade dos documentos de arquivo, de forma a garantir aos arquivistas ou utilizadores de arquivos a confiabilidade da informação a que acedem nas instituições de informação.

O autor mostra-se atento a grandes questões no âmbito da disciplina da arquivística, elaborando uma investigação pertinente que tenta dar resposta à questão central da autenticidade dos documentos de arquivo utilizando os metadados como base da sua investigação. Utiliza uma metodologia que se mostra adequada, uma vez que junta métodos mistos e se baseia na literatura científica e técnica atual, de forma abrangente e sustentada na ideia de interpretação. Esta questão é essencial e vem no seguimento da ideia de que os documentos de arquivo não são neutros, resultando da subjetividade dos seus produtores e, por isso, necessitando de ser contextualizados nas suas múltiplas dimensões, e ainda do próprio princípio da interoperabilidade, ou seja, da comunicabilidade entre documentos e sistemas. O próprio paradigma construtivista, de resto, no qual o autor se baseia, sustenta que a sociedade e as instituições são uma construção resultante da interação e das interpretações dos atores. Neste sentido, a chamada de atenção para a subjetividade procura consciencializar os participantes no processo de construção e gestão da informação para a necessidade de se relacionarem com a comunidade e considerarem os vários pontos de vista, algo notório ao longo da investigação e na relevância dos resultados obtidos.

Em suma, a investigação culmina numa proposta de modelo de requisitos para autenticidade de metadados que poderá ser uma opção para apaziguar os arquivistas no que toca à preservação de documentos de arquivo digitais, tanto a nível nacional como internacional. É aqui, e no cunhar da expressão “vínculo informacional” que abrirá também novas portas para entender a interoperabilidade dos documentos e sistemas de arquivo, que irá assentar o interesse e a relevância da obra de André Pacheco para a comunidade arquivística.

## Referências bibliográficas

Ahmed, J. U. (2010). Documentary Research Method: New Dimensions. *Indus Journal of Management & Social Science (IJMSS)*, 4(1), 1-14. <http://ideas.repec.org/s/iijh/journal.html>

- Bowen, G. A. (2009). Document Analysis as a Qualitative Research Method. *Qualitative Research Journal*, 9(2), 27-40. <https://doi.org/10.3316/QRJ0902027>
- Bustelo-Ruesta, C. (2011). Los grandes temas relacionados con la gestión de documentos: desafíos y oportunidades. *Profesional De La información*, 20(2), 129-134. <https://doi.org/10.3145/epi.2011.mar.01>
- Formenton, D., & Gracioso, L. de S. (2022). Padrões de metadados no arquivamento da Web: recursos tecnológicos para a garantia da preservação digital de websites arquivados. *RDBCI: Revista Digital de Biblioteconomia e Ciência da Informação*, 20(00), e022001. <https://doi.org/10.20396/rdbci.v20i00.8666263>
- Gracy, K. F., & Kahn, M. (2012). Preservation in the Digital Age. *Library Resources & Technical Services*, 56(1), 25-43. <https://doi.org/10.5860/lrts.56n1.25>
- Jureta, I. (2012). Requirements Engineering Methods: A Classification Framework and Research Challenges. *arXiv*. <https://doi.org/10.48550/arXiv.1203.1717>
- Lee, K.-H., Slaterry, O., Lu, R., Tang, X., & McCrary, V. (2002). The State of the Art and Practice in Digital Preservation. *Journal of Research of the National Institute of Standards and Technology*, 107(1), 93-106. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4865277/>
- Stančić, H., & Bralić, V. (2021). Digital Archives Relying on Blockchain: Overcoming the Limitations of Data Immutability. *Computers*, 10(8), 91. <https://doi.org/10.3390/computers10080091>

**BAUC VOL. XXXVI, N.º 1**

**NOTA DE APRESENTAÇÃO**

**ESTUDOS**

Los archivos municipales en Extremadura (España)  
a finales del s. XVIII a través del interrogatorio de la real audiencia  
Carmen Solano Macías; Agustín Vivas Moreno

Arquivística Musical Histórica aplicada a arquivos de Bandas de Música:  
um olhar sobre o estado da questão  
Ana Raquel Coelho

Impactos do Efeito Filtro Bolha no Engajamento de *Fake News*  
Luis Yago Santos Pessoa; Clara Vasconcelos Gusmão; Lucas Daniel Anselmo Tabosa de Andrade;  
Letícia Ferreira Neves; Walter de Macedo Rodrigues; Maria Amália Arruda Camara

Morte e glorificação de D. Miguel da Anunciação (1703-1779),  
Bispo de Coimbra  
Guilhermina Mota

Organização e Representação da Informação do Projeto “Mesas do Castelinho”  
no Centro de Arqueologia da Universidade de Lisboa  
Matilde Seca

Paleografia e ciência da informação:  
reflexões em torno de um diálogo intercientífico  
Carlos Guardado da Silva; Alexandre Faben

A produção e conservação informacional  
de Maria do Carmo Barros Leite (1841-1911)  
Joana M. Couto

**RECENSÕES CRÍTICAS**

Comerford, K. M. (2022). *Jesuit Libraries*. Brill.  
Sofia Bettencourt da Silva

Edmond, J., Horsley, N., Lehmann, J., & Priddy, M. (2022).  
*The Trouble With Big Data: How Datafication Displaces Cultural Practices*.  
Bloomsbury Academic.  
Anabela Pires Duarte

Pacheco, A. (2017). *Informação Digital: O vértice comum  
entre a diplomática e a ciência da informação*. Edições Húmus.  
Maria Beatriz Merêncio

Pacheco, A. (2022). *Arquivos Digitais –  
Metadados e Autenticidade*. Edições Colibri.  
Madalena Lopes Damião Rodrigues

**ISSN**

0872-5632  
2182-7974

**MORADA PARA CORRESPONDÊNCIA**

Boletim do Arquivo da Universidade de Coimbra  
Arquivo da Universidade de Coimbra  
Rua de S. Pedro, 2, 3000-370 Coimbra, Portugal  
URL: <http://www.uc.pt/auc>