



UNIVERSIDADE FEDERAL DA PARAÍBA
CENTRO DE CIÊNCIAS SOCIAIS APLICADAS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO
MESTRADO EM CIÊNCIA DA INFORMAÇÃO

**OS REGISTROS DIGITAIS ARQUIVÍSTICOS DO CENTRO DE CIÊNCIAS
JURÍDICAS DA UNIVERSIDADE FEDERAL DA PARAÍBA:**
descrição, categorização e metadados de preservação digital

PEDRO FELIPY CUNHA DA SILVA

JOÃO PESSOA

2020

PEDRO FELIPY CUNHA DA SILVA

**OS REGISTROS DIGITAIS ARQUIVÍSTICOS DO CENTRO DE CIÊNCIAS
JURÍDICAS DA UNIVERSIDADE FEDERAL DA PARAÍBA:**
descrição, categorização e metadados de preservação digital

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação da Universidade Federal da Paraíba, na linha de pesquisa Ética, Gestão e Políticas de Informação, em cumprimento às exigências para a conclusão do Curso de Mestrado.

Área de concentração: Ética, Gestão e Políticas de Informação

Orientador: Prof. Dr. Wagner Junqueira de Araújo

JOÃO PESSOA

2020

Catálogo na publicação
Seção de Catalogação e Classificação

S586r Silva, Pedro Felipy Cunha da.

Os registros digitais arquivísticos do Centro de Ciências Jurídicas da Universidade Federal da Paraíba: descrição, categorização e metadados de preservação digital / Pedro Felipy Cunha da Silva. - João Pessoa, 2020.

180 f. : il.

Orientação: Wagner Junqueira de Araújo.
Dissertação (Mestrado) - UFPB/CCSA.

1. Ciência da Informação. 2. Arquivo digital. 3. Modelos de categorização de registros digitais. 4. Metadados de preservação digital. 5. Registros digitais arquivísticos. 6. Arquivologia digital. I. Araújo, Wagner Junqueira de. II. Título.

UFPB/BC

CDU 02(043)

PEDRO FELIPY CUNHA DA SILVA

**OS REGISTROS DIGITAIS ARQUIVÍSTICOS DO CENTRO DE CIÊNCIAS
JURÍDICAS DA UNIVERSIDADE FEDERAL DA PARAÍBA:**
descrição, categorização e metadados de preservação digital

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação da Universidade Federal da Paraíba, na linha de pesquisa Ética, Gestão e Políticas de Informação, sob a orientação do Prof. Dr. Wagner Junqueira de Araújo, como parte dos requisitos necessários à obtenção do Título de Mestre em Ciência da Informação.

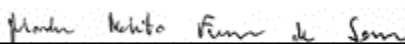
DATA DA AVALIAÇÃO: 02/12/2020

CONCEITO: **APROVADO**

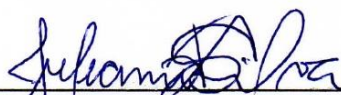
BANCA EXAMINADORA



Prof. Dr. Edvaldo Carvalho Alves
Presidente – PPGCI/UFPB



Prof. Dr. Marckson Roberto Ferreira de Souza
Membro Interno – PPGCI/UFPB



Profa. Dra. Julianne Teixeira e Silva
Membro Externo – DCI/UFPB

JOÃO PESSOA
2020

RESUMO

O crescimento do uso da tecnologia como mediadora da ação das pessoas e instituições ocasionou o aumento do número de registros de atividades em ambientes digitais. Gerir e armazenar tais registros, conforme critérios que permitam a sua organização e acesso no presente e no futuro, constituem-se em um desafio para a comunidade científica e profissional da área de informação. No contexto da transformação digital do serviço público, as universidades federais concentram várias demandas relacionadas à criação e armazenamento desses registros digitais. Para as finalidades deste trabalho, entende-se como registro digital o conjunto de atributos que permitem que os dados constantes de um documento em ambiente digital possam se tornar um documento manifestado e acessível ao usuário no decorrer do tempo, associado ao conteúdo que o fundamenta, e às transações e requisitos que o tornam válido, e que, quando criado para evidenciar e servir de prova das atividades de uma pessoa ou instituição, é um registro digital arquivístico. Nesse sentido, o presente trabalho teve como objetivo estudar como se constituem os registros digitais arquivísticos produzidos pelo Centro de Ciências Jurídicas da Universidade Federal da Paraíba (CCJ/UFPB), por meio da sua descrição, categorização e identificação dos metadados necessários à sua preservação digital. Para tal, buscou conhecer na literatura modelos que permitam descrever e categorizar os registros digitais arquivísticos criados pelo CCJ/UFPB; propor a categorização de uma amostra dos registros digitais arquivísticos produzidos pelo CCJ/UFPB baseada nos modelos disponíveis; e identificar os metadados necessários para a preservação digital dos registros digitais arquivísticos criados pelo CCJ/UFPB. Metodologicamente, trata-se de um estudo quantitativo e qualitativo, composto de pesquisa bibliográfica, descritiva e aplicada. A revisão sistemática e o estudo de caso nortearam os procedimentos utilizados para a coleta e análise dos dados. Os resultados apontam e descrevem duzentos e trinta e quatro tipos de registros digitais arquivísticos criados pelo CCJ/UFPB, categorizados de acordo com cinco modelos de registros digitais. Apresenta metadados de identificação da origem e localização, distinção entre nato digitais e digitalizados, descrição do dispositivo e sistemas utilizados, descrição automática e manual, demonstrados como os metadados para a preservação digital dos registros digitais arquivísticos. E oferece uma reflexão sobre a complexidade da natureza desses registros, os desafios teóricos e práticos decorrentes, dentre eles a necessidade de capacitação profissional dos arquivistas, com destaque para a definição de normas e padrões para os metadados de preservação digital específicos dos registros digitais arquivísticos.

Palavras-chave: Arquivo digital. Modelos de categorização de registros digitais. Metadados de preservação digital. Registros digitais arquivísticos. Arquivologia digital.

ABSTRACT

The use of technology as a mediator of the action of people and institutions has been growing, what has led to an increase in the number of records of activities in digital environments. The task of managing and storing such records, according to criteria that offer an operative arrangement and access now and in the future, has become a challenge for the scientific and professional community of Information. In the context of the digital transformation of the public service, federal universities have been holding several demands related to the creation and storage of these digital records. For the purposes of this work, a digital record is defined as a set of attributes that allow the data included in a document within a digital environment to become an unfold and accessible document to any user over time along with the content on which it is based, the transactions and requirements that validate it, and, besides all this, whenever it is produced to evidence and to serve as a proof of the activities of a person or institution, is labeled as an archival digital record. In this sense, this work focused on studying how the digital archival records produced by the Center of Legal Sciences of the Federal University of Paraíba (CCJ/UFPB) are constituted, through their description, categorization and identification of the metadata necessary for their digital preservation. In order to do so, it was necessary a survey into technical literature models that enabled describing and categorizing the archival digital records created by the CCJ/UFPB; a categorization of a sample of the archival digital records produced by the CCJ/UFPB based on those available models is then proposed; as well as identifying the necessary metadata for a digital preservation of such archival digital records. Methodologically, it is a both quantitative and qualitative study, accomplished through bibliographic, descriptive and applied research. Systematic review and case study guided the procedures used for data collection and analysis. The final result plot and describe two hundred thirty-four types of digital archival records created by CCJ/UFPB, categorized as five models of digital records. It presents the metadata identifying origin and location, distinction between digital and digitized, description of the device and systems used, automatic and manual description, demonstrated as the metadata for the digital preservation of archival digital records. It also offers a reflection on the complexity of the nature of these records and eventual emerged theoretical and practical challenges, one if these is the archivists professional training need, with a special attention to the definition of standards and rules for digital preservation metadata specific to archival digital records.

Keywords: Digital Archive. Digital record categorization models. Digital preservation metadata. Archival digital records. Digital Archivology.

LISTA DE ABREVIATURAS E SIGLAS

AFD – Assentamento Funcional Digital
BASE – *Bienlefeld Academic Search Engine*
CCJ – Centro de Ciências Jurídicas
CCSDS – Comitê Consultivo para Sistemas de Informação Espacial
CTDE – Câmara Técnica de Documentos Eletrônicos
DBTA – Dicionário Brasileiro de Terminologia Arquivística
DNB – *Deutsche Nationalbibliothek*
eSIC – Sistema Eletrônico de Informação ao Cidadão
FalaBR – Plataforma Integrada de Ouvidoria e Acesso à Informação
IBICT – Instituto Brasileiro de Informação em Ciência e Tecnologia
IFES – Instituições Federais de Ensino Superior
MDA – Massa Documental Acumulada
NARSSA – Arquivo Nacional da África do Sul
NUP – Número Único de Protocolo
ORCID – *Open Researcher & Contributor ID*
PEN – Processo Eletrônico Nacional
PRISMA – Preferred Reporting Items for Systematic reviews and Meta-Analyses
RSL – Revisão Sistemática de Literatura
SCDP – Sistema de Concessão de Diárias e Passagens
SIAFI – Sistema Integrado de Administração Financeira
SIG – Sistemas Integrados de Gestão
SIGAA – Sistema Integrado de Gestão das Atividades Acadêmicas
SIGEP – Sistema de Gestão de Pessoas
SIPAC – Sistema Integrado de Patrimônio, Administração e Contratos
TDIC – Tecnologias Digitais de Informação e Comunicação
UFGD – Universidade Federal da Grande Dourados
UFS – Universidade Federal de Sergipe
UFSJ – Universidade Federal de São João Del Rei
UFABC – Universidade Federal do ABC
UFMA – Universidade Federal do Maranhão
UFPB – Universidade Federal da Paraíba
UFLA – Universidade Federal de Lavras
UFOB – Universidade Federal do Oeste da Bahia
UFRB – Universidade Federal do Recôncavo da Bahia
UFRN – Universidade Federal do Rio Grande do Norte
UFRA – Universidade Federal Rural da Amazônia
UFRRJ – Universidade Federal Rural do Rio de Janeiro
UFERSA – Universidade Federal Rural do Semiárido
UNIVASF – Universidade Federal do Vale do São Francisco
UNILA – Universidade Federal da Integração Latino-Americana
UNIFEI – Universidade Federal de Itajubá

LISTA DE FIGURAS

Figura 1 – Percurso metodológico da pesquisa	27
Figura 2 – Publicações por área de conhecimento	43
Figura 3 – As subáreas da ciência da informação no século XXI relacionadas com os registros digitais arquivísticos	44
Figura 4 – Palavras-chave da RSL	49
Figura 5 – Rede de coautorias	51

LISTA DE GRÁFICOS

Gráfico 1 – Publicações por ano	42
Gráfico 2 – Registros digitais arquivísticos do CCJ/UFPB por setor	111
Gráfico 3 – Registros digitais arquivísticos das subunidades da Direção do CCJ/UFPB	112
Gráfico 4 – Registros digitais do CCJ/UFPB por gênero e por setor	114
Gráfico 5 – Registros digitais do CCJ/UFPB divididos por espécie documental	115
Gráfico 6 – Principais espécies documentais dos registros digitais do CCJ/UFPB	116
Gráfico 7 – Espécies documentais x finalidades dos registros digitais do CCJ/UFPB	117
Gráfico 8 – Registros digitais do CCJ/UFPB por atividade	118
Gráfico 9 – Registros digitais arquivísticos do CCJ/UFPB – Finalidades x Plataforma	119
Gráfico 10 – Registros digitais do CCJ/UFPB – espécie documental x modelo de Rondinelli (2011)	121
Gráfico 11 – Registros digitais do CCJ/UFPB – gênero documental x modelo de categorias de Katuu (2016)	122
Gráfico 12 – Registros digitais do CCJ/UFPB por atividade, com base em Zhao <i>et al</i> (2019)	123
Gráfico 13 – Registros digitais do CCJ/UFPB com base em Garmendia (2020)	124
Gráfico 14 – A espécie documental “registro” e os gêneros documentais no caso do CCJ/UFPB	126
Gráfico 15 – Setor responsável pela resposta sobre metadados de preservação digital nas universidades federais que utilizam o SIPAC	130
Gráfico 16 – Metadados de preservação digital nas IFES que utilizam o SIPAC	131
Gráfico 17 – Metadados de preservação digital de registros digitais arquivísticos das IFES que utilizam o SIPAC/Módulo Protocolo	132
Gráfico 18 – Metadados armazenados pelas IFES para identificação de origem e localização dos usuários do SIPAC	140
Gráfico 19 – Metadados armazenados pelas IFES para diferenciar registros nato digitais daqueles digitalizados	144
Gráfico 20 – Metadados de descrição do dispositivo e sistema operacional armazenados pelas IFES que utilizam o SIPAC	147
Gráfico 21 – Metadados descritivos de dados armazenados pelas universidades federais que utilizam o SIPAC	149
Gráfico 22 – Universidades federais que utilizam o SIPAC em relação à quantidade de Metadados descritivos de dados	151
Gráfico 23 – Metadados descritivos incluídos de forma manual pelo usuário nas universidades federais que utilizam o SIPAC	153
Gráfico 24 – Metadados descritivos incluídos de forma manual pelo usuário no SIPAC da Universidade Federal da Paraíba	154

LISTA DE QUADROS

Quadro 1 – Universidades Federais que utilizam o SIPAC	38
Quadro 2 – Especificidades dos registros digitais	60
Quadro 3 – Dificuldades para a gestão de registros digitais	64
Quadro 4 – Elementos intrínsecos do documento arquivístico	77
Quadro 5 – Classificação dos documentos digitais	80
Quadro 6 – Classificação dos documentos digitais interativos	80
Quadro 7 – Categorização de trabalhos no ORCID	88
Quadro 8 – Categorização de objetos no <i>Figshare</i>	89
Quadro 9 – Metadados para preservação digital de registros digitais arquivísticos	128
Quadro 10 – Questionário sobre metadados de preservação digital para as universidades federais que utilizam o SIPAC	129
Quadro 11 – Metadados de preservação digital de registros digitais arquivísticos das IFES que utilizam o SIPAC/Módulo Protocolo.....	133
Quadro 12 – Metadados de diferenciação entre nato digital e digitalizado nas IFES que utilizam o SIPAC.....	145
Quadro 13 – Metadados acerca do tipo de conferência para os documentos digitalizados nas IFES que utilizam o SIPAC	145
Quadro 14 – Metadados de descrição de dados de universidades federais que utilizam o SIPAC.....	150

LISTA DE TABELAS

Tabela 1 – Critérios de pesquisa	33
Tabela 2 – Resumo da aplicação das fases do PRISMA nesta RSL	34
Tabela 3 – Palavras-chave mais utilizadas na RSL sobre registros digitais para arquivos	47
Tabela 4 – Produção por autor	50
Tabela 5 – Citações por autor no <i>Research Gate</i>	50
Tabela 6 – Altimétricas dos autores	52
Tabela 7 – Títulos, autores, periódicos e ano dos artigos mais citados	53
Tabela 8 – Artigos mais baixados	56

SUMÁRIO

1. CONSIDERAÇÕES INICIAIS.....	13
1.1 PROBLEMATIZAÇÃO E JUSTIFICATIVA.....	14
1.2 OBJETIVOS	21
1.2.1 Objetivo geral	221
1.2.2 Objetivos específicos.....	22
2. PERCURSO METODOLÓGICO	23
2.1 CARACTERIZAÇÃO DA PESQUISA	23
2.2 CAMPO EMPÍRICO	24
2.3 SUJEITOS DA PESQUISA.....	26
2.4 PROCEDIMENTO PARA O ESTUDO DE CASO.....	288
2.5 PROCEDIMENTO PARA REVISÃO DE LITERATURA SOBRE GESTÃO DE REGISTROS DIGITAIS PARA ARQUIVOS	332
2.6 PROCEDIMENTO PARA REVISÃO DE LITERATURA SOBRE REGISTROS DIGITAIS ARQUIVÍSTICOS.....	35
2.7 PROCEDIMENTO PARA REVISÃO DE LITERATURA SOBRE METADADOS DE PRESERVAÇÃO DE REGISTROS DIGITAIS	36
2.8 PROCEDIMENTO PARA COLETA DE DADOS SOBRE METADADOS DE PRESERVAÇÃO DIGITAL NAS IFES.....	37
2.9 PROCEDIMENTO DE ANÁLISE DOS DADOS.....	399
3. FUNDAMENTAÇÃO TEÓRICA	41
3.1 GESTÃO DE REGISTROS DIGITAIS PARA ARQUIVOS: REVISÃO SISTEMÁTICA DE LITERATURA	41
3.1.1 Revisão sistemática de literatura – dados quantitativos	42
3.1.2 Revisão sistemática de literatura – dados qualitativos	56
3.1.3 Considerações sobre os registros digitais arquivísticos	66
3.2 REGISTROS DIGITAIS ARQUIVÍSTICOS: CONCEITOS, MODELOS DE CATEGORIZAÇÃO E PERSPECTIVAS	69
3.2.1 Conceitos relacionados aos registros digitais arquivísticos: um olhar para a literatura nacional	69
3.2.2 Modelos de categorização dos registros digitais	83
3.2.3 Perspectivas sobre os registros digitais arquivísticos	90
3.3 METADADOS DE PRESERVAÇÃO DIGITAL PARA REGISTROS DIGITAIS ARQUIVÍSTICOS.....	102
4. ANÁLISE DOS REGISTROS DIGITAIS ARQUIVÍSTICOS DO CCJ/UFPB	110
4.1 DESCRIÇÃO DOS REGISTROS DIGITAIS	110
4.2 CATEGORIZAÇÃO DOS REGISTROS DIGITAIS	120
4.3 ANÁLISE DOS METADADOS DE PRESERVAÇÃO DIGITAL.....	12727
5. CONSIDERAÇÕES FINAIS	160
REFERÊNCIAS.....	165
APÊNDICES	177
ANEXOS	182

1. CONSIDERAÇÕES INICIAIS

As tecnologias digitais de informação e comunicação (TDIC) já fazem parte do cotidiano das pessoas e das organizações. Sob o argumento da praticidade e da facilidade, de forma cada vez mais crescente, indivíduos e empresas realizam suas ações por intermédio de sites, *e-mails*, sistemas, aplicativos e outros recursos digitais de interação.

Contudo, a gestão, preservação e acesso às informações contidas nesses registros digitais ainda se configuram como um desafio para os profissionais e instituições, uma vez que suas características e possibilidades se modificam e ampliam a cada dia.

No caso das instituições federais de ensino superior (IFES), demandas como a necessidade de tramitação dos processos administrativos de forma eletrônica, e de disponibilização dos documentos acadêmicos em ambiente digital, corroboram a existência desse fenômeno.

Tais ações estão inseridas no contexto da transformação digital, proposta pelo Governo Federal para todos os órgãos da administração pública, dentre eles as IFES, que através das portarias do Ministério da Educação n.º 315/2018, 330/2018, 554/2019 e 332/2020, estabeleceu a necessidade de implantação do Diploma Digital e do Acervo Acadêmico Digital, questões com as quais os registros digitais das unidades acadêmicas estão relacionados.

Um estudo de tais registros digitais favorecerá a produção dessas informações em conformidade com os requisitos necessários para o seu acesso posterior, isto é, a proposta desta pesquisa poderá facilitar o trabalho de quem produz e de quem utiliza tais registros, além de contribuir na diminuição do uso do papel e com os consequentes benefícios ambientais decorrentes dessa ação.

Outro elemento relevante para a realização desta pesquisa é a aderência desse tema a outras ações de transformação digital relacionadas. É o caso do Processo Eletrônico Nacional (PEN), instituído pelo Decreto n.º 8.539/2015 e do Assentamento Funcional Digital (AFD), regido pela Portaria n.º 9/2018 do Ministério do Planejamento, Desenvolvimento e Gestão, atual Ministério da Economia; ambas de implantação obrigatória para todos os órgãos do poder executivo federal, incluídas as IFES.

Este trabalho pretende contribuir para a compreensão dos registros digitais, a partir do estudo de caso em que analisa os registros documentais produzidos nas atividades acadêmicas e administrativas, de um centro de ensino de uma universidade federal e os metadados utilizados para a preservação destes registros. Sua estrutura é composta de cinco capítulos.

No capítulo 1 é feita uma contextualização sobre os registros digitais arquivísticos, seguida de problematização sobre o tema e da justificativa para a realização da pesquisa, além de serem apresentados os objetivos geral e específicos do trabalho.

No capítulo 2 é apresentado o percurso metodológico, com caracterização da pesquisa, fixação do campo empírico, definição dos sujeitos da pesquisa, explicação sobre os tipos de registros utilizados como amostra, bem como dos procedimentos de realização da revisão de literatura, da coleta de dados sobre os metadados de preservação digital nas IFES e do estudo de caso ao qual o trabalho se propõe.

No capítulo 3 são apresentados os principais conceitos concernentes aos registros digitais arquivísticos, aspectos da sua gestão, modelos de categorização e um quadro conceitual sobre os metadados de preservação digital apresentados pela literatura.

No capítulo 4 são apresentados os resultados do trabalho, constituídos pela descrição e categorização dos registros digitais produzidos pelo CCJ/UFPB, e análise dos metadados de preservação digital com base nas respostas obtidas junto às universidades federais que compartilham o mesmo sistema de produção de documentos em ambiente digital.

No capítulo 5 são apresentadas as considerações acerca do estudo de caso realizado, limitações da pesquisa e sugestões de investigações futuras.

1.1 PROBLEMATIZAÇÃO E JUSTIFICATIVA

O avanço das TDIC tem provocado mudanças rápidas e constantes nas interações dos indivíduos. Conforme os estudos de Levy (1999), a vida social e cultural tornou-se uma extensão das novas redes de comunicação.

Para Le Coadic (1996), o crescimento no consumo de produtos e serviços informacionais, seguido do desenvolvimento de uma ciência e uma

indústria, também da informação, todas essencialmente influenciadas pela tecnologia, justifica a utilização do termo sociedade da informação.

A aproximação do ambiente social e dos ambientes digitais intermediados pela tecnologia relaciona-se ainda com o conceito de *cibercultura*, estabelecido por Levy (1999), que a define como

[...] o conjunto de técnicas (materiais e intelectuais), de práticas, de atitudes, de modos de pensamento e de valores que se desenvolvem com o crescimento do ciberespaço, definido por meio de comunicação, que surge da interconexão mundial dos computadores, abarcando não apenas a infraestrutura material da comunicação digital, mas também o universo oceânico de informações que ela abriga, assim como os seres humanos que navegam e alimentam esse universo (LEVY, 1999, p. 17).

É possível observar que as necessidades sociais passam a depender de forma crescente da utilização de informações, especialmente por meio de aplicações eletrônicas. Para Lastrez e Albagli (1999, p. 9), “esse novo papel da informação e do conhecimento nas economias vem provocando modificações substantivas nas relações, forma e conteúdo do trabalho, o qual assume um caráter cada vez mais informacional”.

Para Matterlart (2006), o valor do trabalho é substituído pelo valor do conhecimento, definido por Martins (2010, p. 7) como “resultado da aplicação e do uso produtivo da informação”.

Esse cenário, que alia mudanças sociais, uso das tecnologias digitais e necessidade crescente de informação e conhecimento, é contemplado na obra de Castells (2002), que cunhou o termo “sociedade em rede”. Para o autor, esta sociedade é caracterizada por mudanças na forma de organização social, graças ao surgimento das tecnologias da informação, num período de coincidência temporal com uma necessidade de mudança econômica e social. O autor define a tecnologia da informação como uma revolução com penetrabilidade em todas as esferas da vida humana. E acrescenta:

A revolução tecnológica não está na centralidade de conhecimentos e informações, mas na aplicação desses conhecimentos e dessa informação para a geração de conhecimentos e de dispositivos de processamento/comunicação da informação, em um ciclo de

realimentação cumulativo entre a inovação e seu uso (CASTELLS, 2002, p. 69).

Sobre essa revolução, Castells (2002, p. 67) afirma que “no final do século XX vivemos uma transformação da nossa ‘cultura material’, pelos mecanismos de um novo paradigma tecnológico que se organiza em torno da tecnologia da informação”. A tecnologia é compreendida, segundo o mesmo autor, como “o uso de conhecimentos científicos para especificar as vias de se fazerem as coisas de uma maneira ‘reproduzível’”. Castells (2002, p. 67) especifica ainda as tecnologias da informação como “o conjunto convergente de tecnologias em microeletrônica, computação (*hardware* e *software*), telecomunicações” entre outros.

O contexto de produção e utilização de registros digitais nas universidades federais está alinhado com o processo de transformação tecnológica descrito por Castells. As mudanças de comportamento dos usuários dos serviços das IFES influenciam e o ao mesmo tempo são influenciadas pelas novas ferramentas disponibilizadas pelo uso das tecnologias digitais.

Nesse sentido, ao escrever sobre as relações das transformações tecnológicas com as novas formas e processos sociais, Castells (2002, p. 43) aponta que “a tecnologia não determina a sociedade, nem a sociedade escreve o curso da transformação tecnológica [...] o resultado final depende de um complexo padrão interativo”.

Percebe-se que os usuários dos diferentes serviços desejam ser atendidos de forma cada vez mais rápida, com menos burocracia, e ambientes digitais são cada vez mais utilizados para a realização de ações de informação de diferentes tipos, ao passo que um estudo desenvolvido nesse cenário é oportuno.

Esse fenômeno, embora possa ser objeto de diversas ciências, encontra na Ciência da Informação um campo que pode oferecer contribuições na sua compreensão, no estabelecimento de relações de significado sobre ele, bem como na elaboração de propostas para a gestão dos problemas que dele decorrem.

Visto que as IFES estão produzindo documentos e registros diversos em ambiente digital, uma reflexão acerca dos registros criados, armazenados, e que precisarão ser preservados para poderem ser acessados nos ambientes digitais, constitui-se como possibilidade de estudo da ciência da informação.

Tal afirmação encontra sentido no trabalho de Araújo (2018) quando, ao analisar o estudo da informação, pontua

[...] a informação, objeto de estudo da ciência da informação, como um conceito, uma categoria abstrata, que só pode ser compreendida e estudada por meio do estudo de outros fenômenos, estes sim, com existência concreta, material, na esfera da realidade humana. Entre esses fenômenos estão a internet, as bibliotecas, as empresas, as escolas, os arquivos, os museus, a televisão, entre outros, que podem ser estudados em sua dimensão informacional - isto é, ligada à ação humana de produzir e reproduzir a realidade, a cultura e a si mesmos por meio dos registros de conhecimento (ARAÚJO, 2018, p. 104-105).

Ou seja, o fenômeno dos registros digitais, materializados pelos diversos ambientes virtuais de produção, tramitação e acesso a documentos e informações no âmbito das instituições, configura-se como elemento passível de contribuição na compreensão do fenômeno informacional.

Este trabalho pretende contribuir para a identificação, compreensão e preservação dos registros criados em ambientes digitais, no cumprimento das finalidades institucionais, portanto, registros digitais arquivísticos.

Visto que esta pesquisa trata dos registros digitais **arquivísticos**, consideram-se fundamentais para o início desta reflexão, os conceitos de arquivo, documento, documento digital e gestão de documentos. Para isso, utilizou-se o Dicionário Brasileiro de Terminologia Arquivística (DBTA), publicado pelo Arquivo Nacional (2005).

No dicionário citado, **documento** é definido como “unidade de registro de informações, qualquer que seja o suporte ou formato” (ARQUIVO NACIONAL, 2005, p. 73). Enquanto **documento digital**, é aquele “codificado em dígitos binários, acessível por meio de sistema computacional” (ARQUIVO NACIONAL, 2005, p. 75). A **gestão de documentos** é conceituada como o “conjunto de procedimentos e operações técnicas referentes à produção, tramitação, uso, avaliação e arquivamento de documentos em fase corrente e

intermediária, visando sua eliminação ou recolhimento” (ARQUIVO NACIONAL, 2005, p. 100). E para além das definições quanto à instituição, instalações e móvel de guarda, o termo **arquivo** é definido como o “conjunto de documentos produzidos e acumulados por uma entidade coletiva, pública ou privada, pessoa ou família, no desempenho de suas atividades, independentemente da natureza ou suporte” (ARQUIVO NACIONAL, 2005, p. 27).

Consoante tais concepções, é preciso compreender com clareza a diferença entre a definição de registro digital arquivístico proposta por Rogers (2015) – alicerce teórico desta pesquisa – e o conceito de documento arquivístico digital.

Registros digitais arquivísticos consistem em dados gerados pelo usuário (conteúdo), metadados gerados pelo sistema de identificação de origem e localização, metadados gerados pelo aplicativo que gerencia aparência e desempenho do registro (exemplo: formato de arquivo nativo), metadados gerados pelo aplicativo de descrição dos dados (exemplo: metadados do sistema operacional do sistema de arquivos) e metadados gerados pelo usuário que descrevem os dados (ROGERS, 2015, p. 16, p. 18, tradução nossa).¹

Ou seja, a condição arquivística dos registros digitais se dará, portanto, quando sua produção e acumulação forem realizadas como prova do desempenho das atividades das pessoas, físicas ou jurídicas, por eles responsáveis. E sua gestão, de maneira análoga à gestão de documentos, aos procedimentos relativos ao seu ciclo de vida.

Ao se combinar as definições do Arquivo Nacional (2005) com o conceito de Rogers (2015), o termo registro digital arquivístico é entendido, para fins deste trabalho, **como uma unidade de registro de informações e como um tipo de documento digital**. A aproximação com a definição de documento digital se dá por sua composição em dígitos binários e a necessidade de sistema computacional para sua leitura. Contudo, esse tipo de documento não se limita à forma clássica, uma vez que também considera

¹ No original: “Digital records consist of user-generated data (content), system-generated metadata identifying source and location, application-generated metadata managing the look and performance of the record (e.g., native file format), application-generated metadata describing the data (e.g., file system metadata OS), and user-generated metadata describing the data”.

outros elementos como parte da sua constituição, conforme define Rogers (2015) ao se referir aos metadados.

Embora o DBTA não traga um conceito para registros digitais arquivísticos, dado o seu caráter contemporâneo, o Conselho Nacional de Arquivos, por meio da Câmara Técnica de Documentos Eletrônicos (CTDE) apresenta, dentre outras coisas, a necessidade de gestão dos metadados. Presente na Resolução 43/2015, quando descreve as diretrizes para a implementação de repositórios digitais arquivísticos confiáveis, definidos como

[...] um ambiente de armazenamento e gerenciamento de materiais digitais. Esse ambiente constitui-se de uma solução informatizada em que os materiais são capturados, armazenados, preservados e acessados. Um repositório digital é, então, um complexo que apoia o gerenciamento dos materiais digitais, pelo tempo que for necessário, e é formado por elementos de *hardware*, *software* e metadados, bem como por uma infraestrutura organizacional e procedimentos normativos e técnicos (CONARQ, 2015, p. 9).

Ou seja, a gestão arquivística requer além do conhecimento relacionado aos documentos não digitais, compreender e estabelecer procedimentos pertinentes aos arquivos produzidos por meios digitais.

Ressalta-se que, para o termo registro, isoladamente, o DBTA apresenta duas definições. A primeira, relacionada a procedimentos de protocolo, como a “anotação sistemática em livro próprio”; a segunda, com maior proximidade ao tema que se deseja pesquisar, como “unidade de informação logicamente indivisível”.

A utilização de ambientes digitais e a consequente produção de registros digitais arquivísticos torna mais complexa a tarefa de gerenciar tais elementos (Rogers, 2015), dada a multiplicidade de possibilidades de interação e consequente significação pelo usuário (HENNTONEM, 2017).

A diversificação da natureza dos registros digitais é uma realidade e o aumento do seu volume tem sido exponencial nas últimas décadas (Katuu, 2016), de tal maneira que é necessário um esforço intelectual para a compreensão das novas demandas de gestão arquivística, necessárias para atuar com essa nova possibilidade de acervo. Portanto, a importância do arquivo não deve ser percebida apenas no amanhã (FENG, 2017). É preciso

reconhecer a contribuição da Arquivologia no presente para que o acesso aos registros digitais no futuro seja possível.

A quantidade de registros digitais gerados por meio de sistemas informatizados é cada vez maior, sendo necessário tratamento arquivístico específico, para que a informação possa ser preservada e acessada. Baron e Thurston (2016) definem a consequência da ausência desse tratamento como “*darckarchive*”, termo que teria a função de transmitir a noção de que o presente e o futuro volumes de informação digital, armazenados em arquivos, não estarão disponíveis a qualquer momento, sem que novas técnicas sejam empregadas por arquivistas. Seria como um “apagão arquivístico” do ambiente digital.

Há, portanto, a necessidade da existência de um “arquivista digital do futuro” (SCHWAITZER, 2018; HENTTONEM, 2017; BARON; THURSTON, 2016) como um profissional com a prerrogativa de enfrentamento dos novos métodos, altamente dependentes de meios automatizados de codificação, classificação e diversas outras tecnologias, de forma que o conteúdo eletrônico possa ser pesquisado e os arquivos digitais possam ser efetivamente acessados por aqueles que necessitarem.

Com base no exposto, esta pesquisa tem como elemento norteador a seguinte questão-problema: **como se constituem os registros digitais arquivísticos produzidos pelo Centro de Ciências Jurídicas da Universidade Federal da Paraíba?**

A importância desta pesquisa está na possibilidade de contribuir para que as informações produzidas em ambientes digitais sejam identificadas enquanto registros digitais arquivísticos com vistas à sua adequada preservação.

Não se pode falar em preservação sem que haja a gestão. Nesse sentido, a gestão dos registros digitais arquivísticos para Henntonem (2017), consiste, entre outras coisas, em um conjunto de processos que transferem e oferecem acesso à informação através de contextos espaciais e limites temporais.

Para Thurston (2016, p. 214) a falta de gestão pode permitir que os registros digitais sejam “alterados, apagados, fragmentados, corrompidos ou perdidos, seja através da interferência ou simplesmente através da

obsolescência tecnológica ou de um quadro de controle inadequado”. Isto é, para a preservação dos direitos ou da memória, a gestão dos registros digitais arquivísticos é premente, pois são evidentes os riscos que a falta dela traz para as instituições, para os cidadãos e para a sociedade como um todo.

Acrescenta-se como justificativa para esta pesquisa o fato de o autor trabalhar na Gestão Arquivística do Centro de Ciências Jurídicas da Universidade Federal da Paraíba, como ocupante do cargo de Arquivista junto à Direção do referido centro.

Este trabalho pretende contribuir ainda para a compreensão do fenômeno relacionado aos registros digitais arquivísticos nas unidades acadêmicas das IFES, com vistas a favorecer um melhor aproveitamento dos seus recursos informacionais, uma tomada de decisão mais segura por parte da administração e melhor atendimento às necessidades de informação do cidadão.

Sabe-se que, para além da gestão, a preservação dos registros digitais arquivísticos, e das informações que eles contêm, também deve ser uma preocupação, por esse motivo, este trabalho também se dedicará aos metadados de preservação digital.

É preciso garantir que o usuário possa acessar os seus registros não apenas no presente, mas também no médio e longo prazo, com vistas ao cumprimento da Lei de Acesso à Informação, Lei Federal nº 12.527/2011, e da garantia de direitos que podem estar subsidiados por eles, aspectos que podem ser considerados como parte da responsabilidade social da Ciência da Informação.

1.2 OBJETIVOS

Para a efetivação deste estudo foram definidos os seguintes objetivos:

1.2.1 Objetivo geral

De acordo com os conceitos apresentados de forma introdutória, o objetivo geral deste trabalho é **estudar a constituição dos registros digitais**

arquivísticos produzidos pelo Centro de Ciências Jurídicas da Universidade Federal da Paraíba.

1.2.2 Objetivos específicos

Para operacionalizar o objetivo geral, elencamos três objetivos específicos:

- I. **Descrever** os registros digitais arquivísticos produzidos pelo Centro de Ciências Jurídicas da Universidade Federal da Paraíba;
- II. **Categorizar** os registros digitais arquivísticos produzidos pelo Centro de Ciências Jurídicas da Universidade Federal da Paraíba quanto à sua natureza, conforme indicado na literatura da área;
- III. **Identificar** os metadados necessários para a preservação digital dos registros digitais arquivísticos produzidos pelo Centro de Ciências Jurídicas da Universidade Federal da Paraíba.

Apresentamos, a seguir, o percurso metodológico desta pesquisa.

2. PERCURSO METODOLÓGICO

Este capítulo apresenta os métodos utilizados para a consecução dos objetivos da pesquisa, sua caracterização e os procedimentos escolhidos para a análise dos dados obtidos.

2.1 CARACTERIZAÇÃO DA PESQUISA

A metodologia descreve como o trabalho pretende alcançar seus objetivos. No que se refere à obtenção de informações, este estudo se realizará por meio de pesquisa bibliográfica, que Rodrigues (2006, p. 89) define como aquela realizada a partir de fontes secundárias, ou seja, por meio de material já publicado, como livros, revistas e artigos científicos e publicações técnicas relacionadas com o objeto de pesquisa.

Será adotada a abordagem quali-quantitativa, definida por Marconi e Lakatos (2011, p. 269) como aquela em que os dados coletados são analisados sob a perspectiva estatística e também sob o ponto de vista da interpretação dos aspectos mais profundos e complexos que envolvem o ser humano, seus hábitos, atitudes e tendências de comportamento.

Quanto aos objetivos, este trabalho pretende ser uma pesquisa descritiva, definida por Prodanov e Freitas (2013, p. 52) como aquela que “visa a descrever as características de determinada população ou fenômeno ou o estabelecimento de relações entre variáveis”.

Quanto à natureza, será uma pesquisa aplicada, definida por Gil (2010, p. 27) como “voltada à aquisição de conhecimentos com vistas à aplicação numa situação específica”.

Quanto aos procedimentos técnicos, adotar-se-á a revisão sistemática de literatura (RSL) em bases de dados *online* de pesquisa científica que, segundo Miranda (2018, p. 24), são “estudos das produções científicas sobre determinado tema, em bases de dados referenciadas e em específico espaço de tempo”.

Do mesmo modo, trata-se de um estudo de caso que, de acordo com Gil (2010, p. 37), “é uma modalidade de pesquisa amplamente utilizada nas

ciências sociais. Consiste no estudo profundo e exaustivo de um ou poucos objetos, de maneira que permita seu amplo e detalhado conhecimento”.

A proposta é de realizar um estudo de caso único, que conforme Prodanov e Freitas (2013), não visa proporcionar o conhecimento preciso das características de uma população, mas, sim, expandir ou generalizar proposições teóricas a partir do caso estudado. Tal proposta foi utilizada para a consecução dos objetivos de descrição e categorização dos registros digitais, entretanto, para a identificação dos metadados de preservação digital, optou-se por ampliar o estudo para o conjunto de universidades federais que utilizam o mesmo sistema de produção de registros digitais.

A instituição escolhida para a pesquisa possui um trabalho de diagnóstico documental e implantação de um programa de gestão de documentos iniciado em 2018 e coordenado pelo autor desta pesquisa, que atua como servidor técnico-administrativo, ocupante do cargo de Arquivista. Tal contexto favoreceu a escolha da proposta de pesquisa ação que, conforme Thiollent (1998), define-se como

[...] pesquisa social com base empírica que é concebida em estreita associação com uma ação ou com a resolução de um problema coletivo e no qual os pesquisadores e os participantes representativos da situação ou do problema estão envolvidos de modo cooperativo ou participativo (THIOLLENT, 1998, p. 14).

A caracterização da instituição escolhida como local da pesquisa para os itens de descrição e categorização dos registros digitais arquivísticos, assim como as informações acerca da amostra de universidades participantes da pesquisa dos metadados de preservação digital, com as devidas justificativas, serão detalhadas no item a seguir.

2.2 CAMPO EMPÍRICO

O Centro de Ciências Jurídicas (CCJ), da Universidade Federal da Paraíba (UFPB), foi escolhido para ser o campo empírico deste trabalho por reunir as seguintes condições: i) possuir publicação técnica com dados acerca dos registros gerados em quantidade suficiente para subsidiar uma pesquisa

científica; ii) estar inserido em um cronograma de transformação digital que contempla a mudança de suporte na geração de registros e documentos, além de ser o local onde o autor da pesquisa exerce suas atividades laborais em função destinada a, dentre outras atribuições, supervisionar a produção, gestão e preservação dos registros em ambientes digitais.

Primeiramente, no ano de 2018, início das atividades do autor desse trabalho no CCJ/UFPB, após realização de entrevistas, visitas e levantamentos em todas as subunidades do centro de ensino, foi produzido um diagnóstico documental – instrumento de gestão arquivística constituído de dados sobre os registros gerados, que será utilizado como uma das fontes de dados para esta pesquisa.

Já no ano de 2019, em virtude das demandas do governo federal e da própria administração central da UFPB, diversas ações acadêmicas e administrativas passaram a realizar seus registros exclusivamente em ambiente digital. Além disso, os servidores do centro – tanto docentes como técnico-administrativos, bem como os discentes – dos cursos de graduação e pós-graduação, já estão habituados a realizarem pedidos, criarem e tramitarem diferentes demandas por meio de sites, *e-mails*, plataformas e sistemas de informação institucionais com diferentes finalidades.

Durante o ano de 2020, em virtude da pandemia global de coronavírus, a UFPB regulamentou o trabalho remoto e a realização de aulas por meio de plataformas digitais, cenário que torna ainda mais atual a necessidade de estudos acerca dos registros digitais.

O CCJ é também a unidade administrativa da UFPB em que o autor desta pesquisa realiza suas funções de técnico-administrativo, ocupante do cargo de Arquivista na Direção da unidade, já tendo, portanto, acesso a grande quantidade de dados referentes aos registros produzidos, suas características, finalidades, sua tramitação, condições de armazenamento e preservação, dentre outros elementos considerados relevantes para os objetivos propostos.

Para compor a análise dos metadados de preservação digital e favorecer a compreensão desse aspecto da investigação, optou-se por ampliar a pesquisa e buscar informações junto às outras instituições federais de ensino que utilizam o mesmo sistema de criação de registros digitais – o Sistema

Integrado de Patrimônio, Administração e Contratos (SIPAC), por meio de questionário enviado por meio digital, conforme descrito no item 2.8.

2.3 SUJEITOS DA PESQUISA

Serão considerados registros digitais produzidos pelo Centro de Ciências Jurídicas da UFPB, para os fins desta pesquisa, aqueles que tenham sido criados por servidores da instituição, utilizando os sistemas oficiais da universidade e as plataformas eletrônicas autorizadas, no cumprimento das funções dos setores e subunidades do CCJ visitados durante a realização do diagnóstico documental.

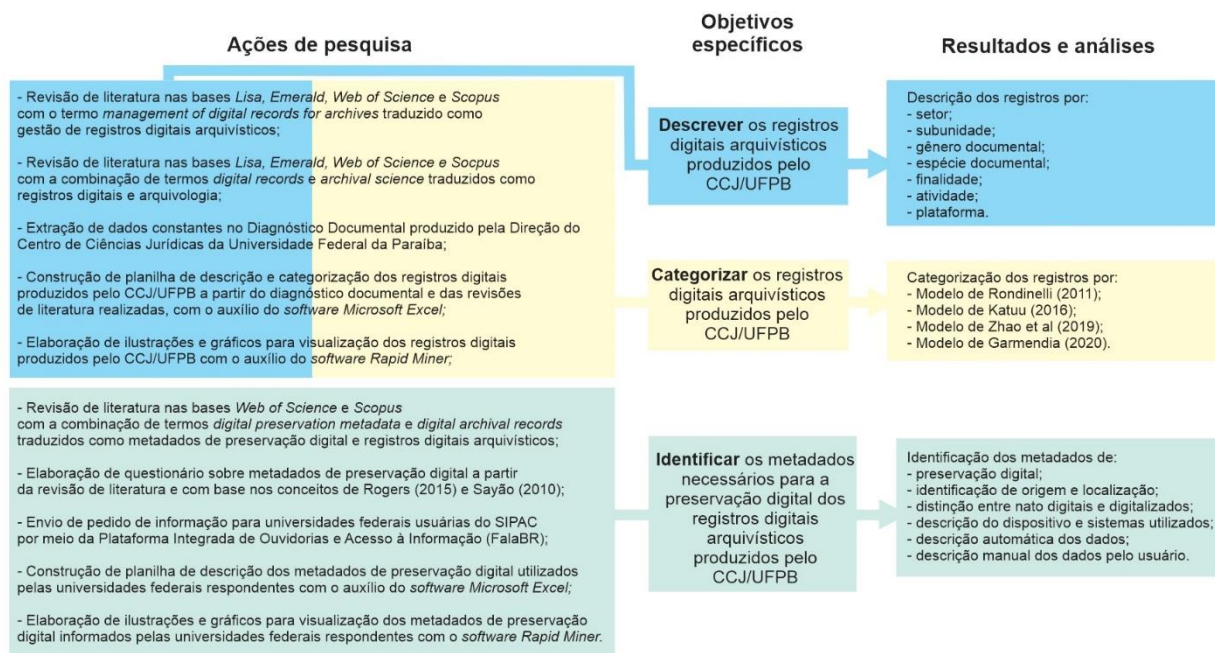
Sabe-se que discentes, servidores docentes e servidores técnico-administrativos do CCJ também produzem registros digitais em ambientes virtuais diversos, como aplicativos de trocas de mensagem, *e-mails* particulares, dentre outros. Entretanto, visando obter um resultado objetivo acerca dos registros que efetivamente correspondam à atividade do órgão, este trabalho não se ocupará dos registros digitais criados ou recebidos fora dos canais institucionais do CCJ e da UFPB, conforme legislação vigente durante a realização do levantamento.

A pesquisa buscou identificar quais registros digitais são criados pelo CCJ/UFPB, quais elementos são considerados preponderantes para a sua descrição. Ainda, permitiu relacionar tais elementos com os conceitos de classificação da Arquivologia e modelos de categorização de registros digitais arquivísticos encontrados na literatura.

O recorte utilizado foram os registros criados pelos setores e subunidades do CCJ da UFPB, como produtos da realização das competências institucionais dessas unidades, em ambiente digital.

Com vistas ao cumprimento do objetivo de identificar os metadados de preservação digital, entendeu-se ser prudente a inclusão das outras universidades federais que utilizam o mesmo sistema de produção de registros digitais para composição da análise.

A figura a seguir ilustra o percurso metodológico desta pesquisa, que será descrito de forma detalhada nos subitens dedicados a esta finalidade.

Figura 1 – Percurso metodológico da pesquisa

Fonte: elaborado pelo autor, 2020.

Inicialmente foi realizada uma pesquisa sobre a gestão dos registros digitais para arquivo. Posteriormente, sentiu-se a necessidade de ampliação da revisão de literatura, de modo que foi realizada nova busca, dessa vez sobre registros digitais e arquivologia. Em continuidade, pesquisou-se sobre metadados de preservação digital e registros digitais arquivísticos.

Com base nas leituras realizadas, passou-se ao mapeamento dos dados obtidos através do Diagnóstico Documental do CCJ/UFPB, obtendo-se a planilha de descrição e categorização dos registros digitais arquivísticos produzidos pelo referido centro (apêndice 1), cujo processamento foi realizado com o auxílio de *softwares* computacionais. As análises podem ser lidas nos itens 4.1 e 4.2.

Na sequência, foi produzido um questionário sobre os metadados de preservação digital (apêndice 2), enviado para as universidades federais usuárias do SIPAC, por meio da plataforma FalaBR. As respostas obtidas foram tabuladas, obtendo-se a planilha de identificação dos metadados de preservação digital (apêndice 3), e seu processamento também foi realizado por meio de *softwares*. A análise está no item 4.3.

2.4 PROCEDIMENTO PARA O ESTUDO DE CASO

De fevereiro a agosto do ano de 2018, no cumprimento das funções arquivísticas do Centro de Ciências Jurídicas da Universidade Federal da Paraíba, foi elaborado o Diagnóstico Documental do CCJ/UFPB (apêndice 4), que é o instrumento por meio do qual se procede a identificação dos documentos produzidos e acumulados por uma instituição com o objetivo de amparar o planejamento das ações de gestão de documentos.

De acordo com Silva (2018), a metodologia utilizada no diagnóstico foi dividida em duas etapas: a primeira, destinada aos depósitos de massa documental acumulada (MDA); a segunda, aos arquivos setoriais correntes. Os depósitos de MDA, para fins do diagnóstico, são os espaços para guarda de documentos que não são mais consultados com tanta frequência, cuja acumulação foi feita de forma não estruturada pelos servidores da instituição ao longo do tempo. Já os arquivos setoriais correntes são constituídos pelos documentos de acesso frequente para atender aos interesses administrativos, estes com maior potencial de contribuição para esta pesquisa.

Passaremos a descrever, conforme Silva (2018), os procedimentos de coleta de dados constantes do Diagnóstico Documental do CCJ/UFPB.

Na Sede do CCJ, localizada no Campus I, foi realizado o diagnóstico em 12 Arquivos Setoriais Correntes e em 1 Depósito de Massa Documental Acumulada, denominado como Depósito 1, a saber:

Arquivos Setoriais Correntes

- Gestão de Pessoas
- Gestão Administrativa
- Secretaria Administrativa, Direção e Conselho de Centro
- Coordenação de Monografias do Campus I
- Departamento de Direito Público
- Departamento de Direito Privado
- Departamento de Direito Processual e Prática Jurídica
- Coordenação da Graduação em Direito do Campus I
- Almoxarifado
- Setor de Informática e Redes
- Programa de Pós-Graduação em Ciências Jurídicas

- Biblioteca Setorial do Campus I

Depósito de Massa Documental Acumulada

- Depósito 1 - Campus I

Na Faculdade de Direito, localizada na Praça João Pessoa, foi realizado o diagnóstico em 1 Arquivo Setorial Corrente e em 2 Depósitos de Massa Documental Acumulada, denominados como Depósito 2 e 3, a saber:

Arquivos Setoriais Correntes

- Núcleo de Prática Jurídica

Depósitos de Massa Documental Acumulada

- Depósito 2 – Antiga Biblioteca
- Depósito 3 – Arquivo da Torre

Na Unidade Avançada do CCJ, localizada na cidade de Santa Rita, foi realizado o diagnóstico em 2 Arquivos Setoriais Correntes e em 1 Depósito de Massa Documental Acumulada, denominado como Depósito 4, a saber:

Arquivos Setoriais Correntes

- Departamento de Ciências Jurídicas
- Coordenação da Graduação em Direito do DCJ (Santa Rita)

Depósito de Massa Documental Acumulada

- Depósito 4 – DCJ (Santa Rita)

Antes de realizar a aplicação dos questionários, a Direção de Centro do CCJ, por meio da Gestão de Pessoas, enviou a todos os servidores do centro uma proposta de cronograma de visitas do Arquivista para a realização das entrevistas. O cronograma planejava o início das atividades em 12/03/2018 e o término em 12/04/2018, entretanto, a agenda de atividades do centro fez com que a sua concretização só fosse possível em 31/05/2018.

O planejamento do diagnóstico previa a aplicação do questionário com todos os setores e assessoriais do CCJ, contudo, a avaliação não pôde ser realizada com as seguintes unidades:

- Assessoria Acadêmico-Administrativa

- Assessorias Didático-Científico-Pedagógicas
 - Extensão
 - Pesquisa e Pós-Graduação
 - Graduação
 - Estágio
 - Convênios
 - Monitoria

A Assessoria Acadêmico-Administrativa não estava em funcionamento e as demais Assessorias não estavam disponíveis para a realização do Diagnóstico, seja pela ausência de servidor responsável ou pela ausência de acervo disponível para consulta. Apesar disso, nos demais setores, o trabalho de verificação e análise pôde ser realizado por meio de entrevistas com os gestores ou servidores por eles designados.

Para a obtenção dos dados necessários para a composição do diagnóstico, fez-se a utilização de um questionário (apêndice 5), utilizado como base para as entrevistas realizadas com os responsáveis pelos Arquivos Setoriais Correntes ou por servidores por eles designados.

No caso dos Depósitos de Massa Documental Acumulada, os dados foram obtidos por meio de observação pelo Arquivista, que utilizou os mesmos questionamentos presentes no formulário dos Arquivos Correntes para a composição das respostas, fazendo as adaptações necessárias ao contexto de desconhecimento do conteúdo do acervo, limitando-se às informações que poderiam ser extraídas das etiquetas e notações disponíveis. O questionário foi composto de 6 partes: 1) Identificação; 2) Produção Documental; 3) Arquivamento; 4) Acesso; 5) Condições de Armazenamento; e 6) Outros.

No campo de “Identificação”, foram colhidas informações a respeito do setor, sua vinculação hierárquica na estrutura do CCJ e os dados do responsável pelas respostas.

No item de “Produção Documental”, foram observadas as tipologias documentais produzidas, as espécies constantes do acervo e a existência ou não de procedimentos e fluxos documentais para normalização dos documentos.

No que se refere ao “Arquivamento”, a intenção era identificar que tipos documentais eram armazenados, que método de arquivamento era utilizado, se

era realizado algum tipo de classificação documental, quais materiais foram escolhidos para o acondicionamento e a relação entre o espaço disponível e o espaço utilizado para os arquivos dos setores – item com especial aderência ao objeto desta pesquisa.

No campo de “Acesso”, a proposta foi verificar quais elementos de recuperação da informação serão necessários após a padronização dos instrumentos de gestão de documentos nos setores e se já existia algum tipo de iniciativa que favorecesse a busca por documentos para atender aos interesses dos usuários internos, bem como avaliar questões de segurança de informação – item que também se insere com precisão no escopo deste trabalho.

No item das “Condições de armazenamento do acervo”, buscou-se avaliar a necessidade de adoção de medidas de conservação, preservação e restauração preventiva ou corretiva nos acervos correntes e a qualidade das instalações em que os arquivos se encontram para o cumprimento das suas finalidades de guarda.

Por fim, no item “Outros”, buscou-se identificar as ações já realizadas com a documentação do setor, na intenção de recuperar possíveis lacunas nos conjuntos documentais e melhor compreender a história administrativa do centro.

Este trabalho utilizou os dados colhidos durante a realização do Diagnóstico Documental do CCJ/UFPB para subsidiar a descrição detalhada dos registros digitais produzidos pelo referido centro.

Os dados obtidos no diagnóstico documental foram combinados ao conhecimento do pesquisador autor deste trabalho, que atua como servidor técnico-administrativo da unidade, acerca da produção de registros digitais pelo CCJ/UFPB, de modo que foi produzido um banco de dados contendo 234 tipos de registros digitais, materializado na planilha de descrição e categorização dos registros digitais arquivísticos do CCJ/UFPB (apêndice 1).

Esse banco de dados foi constituído das informações sobre setor de origem, gênero, espécie, finalidade, atividade e plataforma utilizada, além de contar com a informação acerca do enquadramento dos registros nos modelos e categorias propostas pelos autores Garmendia (2020), Zhao *et al* (2019), Katuu (2016) e Rondinelli (2011).

Em conformidade com os objetivos da pesquisa, e com base nos requisitos para metadados de preservação digital de registros digitais arquivísticos de Sayão (2018) e Rogers (2015), foi constituído um questionário (apêndice 2) com seis perguntas acerca dos metadados de preservação digital e enviado para as universidades federais que utilizam o mesmo sistema de produção de documentos da UFPB. Os detalhes do procedimento de coleta desses dados estão no item 2.8.

O conjunto dos dados foi analisado com o auxílio do *software* de visualização *Rapid Miner* e as análises foram detalhadas nos itens 4.1 Descrição dos registros digitais, 4.2 Categorização dos registros digitais e 4.3 Análise dos Metadados de Preservação Digital.

2.5 PROCEDIMENTO PARA REVISÃO DE LITERATURA SOBRE GESTÃO DE REGISTROS DIGITAIS PARA ARQUIVOS

No primeiro semestre do ano de 2019, durante a disciplina Tecnologia da Informação aplicada à Ciência da Informação, foi possível aprender a utilizar as principais ferramentas de gestão de referências, bem como os principais canais de acesso aos artigos científicos produzidos sobre registros digitais, de modo que foi produzida uma revisão sistemática de literatura (RSL) nas bases de periódicos nacionais e internacionais.

A RSL permite uma melhor visualização da pesquisa e a identificação de possíveis lacunas a serem preenchidas em investigações futuras (Miranda, 2018), além de ir ao encontro de artigos produzidos pela área de Ciência da Informação.

O modelo de RSL estabelecido foi o PRISMA – *Preferred Reporting Items for Systematic reviews and Meta-Analyses* –, que é composto das fases de identificação, seleção, elegibilidade e inclusão, conforme Moher *et al* (2009). Além disso, este trabalho se utilizou ainda de dados bibliométricos existentes nas bases para a composição de uma meta-análise sobre o tema da pesquisa.

Cabe destacar a singularidade do tema em questão na terminologia Arquivística. Conforme Pearce-Moses (2005, p. 13)² “o impacto mais

² [...] the most significant impact on archival language and professional boundaries resulted from the challenges of electronic records.

significativo na linguagem arquivística e nos seus limites profissionais resultaram dos desafios dos registros digitais”.

Verificou-se que não há consenso quanto ao termo utilizado para se referir à gestão de registros digitais arquivísticos, havendo trabalhos que utilizam *digital records management*, *recordkeeping*, *digital archival records management*, *e-records management*, *eBorn*, *digital archive material*, entre outros termos.

Considerando o objetivo desta pesquisa, a revisão optou por utilizar o termo *management of digital records for archives*, haja vista ter encontrado um número considerável de resultados, embora existam outras possibilidades de busca, como a pesquisa booleana com combinação de termos como “*digital records and archival science*”, que foi utilizada para a complementação desta pesquisa, na revisão descrita no item 2.6.

Sendo assim, o principal critério para inclusão de materiais nessa revisão baseada nas quatro fases do PRISMA foi identificar os conceitos e desafios na gestão de registros digitais para arquivos.

Considerando os fatores de velocidade nas transformações tecnológicas que alteram os registros digitais, o recorte temporal foi de um período de 5 anos, de janeiro de 2015 a julho de 2019. Os critérios de pesquisa são apresentados na tabela 1 abaixo:

Tabela 1 – Critérios de pesquisa

Base de dados	Tipo de publicação	Período	Registros
<i>Emerald</i>	<i>Research paper</i>	2010 a 2019	83
<i>Lisa</i>	Artigo	2010 a 2019	165
<i>Scopus</i>	Artigos de periódicos em inglês	2010 a 2019	106
<i>Web of science</i>	Artigos	2010 a 2019	88
Total			442

Fonte: dados da pesquisa, 2019.

Os metadados foram exportados para os *softwares Mendeley e Microsoft Excel*. Iniciou-se, portanto, a aplicação do método, conforme a seguir:

Tabela 2 – Resumo da aplicação das fases do PRISMA nesta RSL

Base de dados	Termo de busca	Resultado	Exclusão (duplicações)	Elegibilidade	Exclusão (metadados)	Elegibilidade	Exclusão 2010 a 2014	Elegibilidade	Exclusão (sem relevância)	Elegibilidade
<i>Emerald</i>	“management of digital records for archives”	83	29	54	38	16	14	2	1	1
<i>Lisa</i>		165	0	165	92	73	13	60	39	21
<i>Scopus</i>		106	12	94	13	81	37	44	23	21
<i>Web of science</i>		88	25	63	48	15	1	14	8	6
Total		442	66	376	191	185	65	120	71	49

Fonte: dados da pesquisa, 2019.

As buscas ocorreram entre as datas de 1 a 3 de junho de 2019 e se restringiram às bases *Emerald*, *Lisa*, *Scopus* e *Web of Science*. Tais bases foram escolhidas em virtude da relevância que os periódicos indexados possuem para a Ciência da Informação (CI) e por conterem alto número de trabalhos em seus acervos.

Foi recuperado um total de 442 documentos, que após importados para o *software Mendeley*, continha 66 repetições. Os artigos repetidos foram excluídos e posteriormente foi feita a análise de consistência dos metadados necessários para esta revisão. Constatou-se que 191 publicações não continham o conjunto de metadados desejados, quais sejam autores, palavras-chave, ano de produção, resumos. Excluídos os artigos sem os metadados, restaram 185 trabalhos a serem analisados, produzidos entre os anos de 2010 a 2019. Considerando que mais de 50% da produção estava concentrada no período de 2015 a 2019, sendo esse o intervalo mais recente, e a necessidade de diminuição da amostra para fins de eficiência da análise, foram excluídos 65 artigos com data de produção entre 2010 e 2014, restando 120 trabalhos a serem analisados. Após a leitura dos resumos, verificou-se que 71 artigos não estavam diretamente relacionados ao tema investigado, finalizando-se a seleção com 49 publicações consideradas adequadas ao interesse desta pesquisa.

Após a finalização das fases de elegibilidade, optou-se por trabalhar com os 49 artigos resultantes da última seleção, a partir dos quais as análises bibliométrica e de conteúdo foram realizadas.

Apoiado nos conceitos encontrados na revisão de literatura realizada, este trabalho busca ampliar a investigação acerca dos registros digitais, de forma a complementar o estudo já realizado e fundamentar a investigação a que se propõe.

A intenção é identificar nos artigos científicos, livros e demais publicações da Ciência da Informação, Arquivologia, Computação e demais áreas relacionadas, as propostas de classificação de registros digitais já existentes e os conceitos de documento arquivístico digital, com vistas a verificar a possibilidade da sua aplicação aos registros digitais produzidos pelo CCJ/UFPB.

De posse das informações obtidas pelo Diagnóstico Documental e do conhecimento extraído das leituras realizadas nesta RSL e na sua ampliação sobre Registros Digitais e Documentos Arquivísticos, pretende-se realizar um estudo comparativo de conceitos e das suas aplicações ao caso do CCJ/UFPB.

2.6 PROCEDIMENTO PARA REVISÃO DE LITERATURA SOBRE REGISTROS DIGITAIS ARQUIVÍSTICOS

Durante a escrita do projeto, sentiu-se a necessidade de ampliar a pesquisa para a inserção dos termos documento/registo digital e ciência arquivística/arquivologia. Tal ação cumpre o objetivo de ampliar a visão obtida após a primeira revisão de literatura sobre gestão de registros digitais para arquivos, descrita no item anterior.

Sendo assim, utilizando o portal de periódicos da CAPES, dessa vez optando pela combinação de termos “*digital records*” e “*archival science*”, foi escolhido o período entre 2015 e 2019, dada a sua relevância quanto ao objeto em questão, e foi feita a análise das publicações encontradas.

A elegibilidade levou em consideração os artigos que possuíam acesso aberto, a aderência do trabalho ao objeto da pesquisa com base na leitura dos metadados disponíveis, em particular o resumo e as palavras-chave, e ainda da vinculação da publicação, do autor e do periódico com a área da pesquisa.

De forma complementar ao material encontrado nas bases científicas referenciadas pela CAPES, outros trabalhos foram inseridos na análise mediante as leituras realizadas e a indicação de professores do Programa de Pós-graduação em Ciência da Informação da UFPB, mais as publicações técnicas consideradas relevantes após as análises (CONARQ, 2011; 2015; CTDE, 2008; ICA, 2005; PEARCE-MOSES, 2005).

2.7 PROCEDIMENTO PARA REVISÃO DE LITERATURA SOBRE METADADOS DE PRESERVAÇÃO DE REGISTROS DIGITAIS

No decorrer da pesquisa entendeu-se que o termo “metadados”, no contexto em questão, ainda necessitava de outros aportes teóricos. Sendo assim, em conformidade com o objetivo de identificar os metadados necessários para a preservação digital dos registros digitais arquivísticos do CCJ/UFPB, este trabalho buscou identificar, na literatura, elementos que fundamentassem a conceituação, tipificação e padrões existentes quanto aos metadados de preservação digital de registros digitais arquivísticos. Para isso, utilizou-se a metodologia do PRISMA – *Preferred Reporting Items for Systematic reviews and Meta-Analyses* –, já descrita nos itens anteriores.

As bases de dados escolhidas para esta revisão foram *Web of Science* e *Scopus*, por serem bases que contém resultados significativos para a área de Ciência da Informação.

A busca na base *Web of Science* foi realizada utilizando os termos “*digital preservation metadata*” e “*digital archival records*” de forma simultânea. Considerou-se, com base nas revisões descritas nos itens 2.5 e 2.6, que a utilização de operadores booleanos traria maior significância para os trabalhos a serem recuperados. A língua inglesa foi escolhida para a busca em virtude de já ter sido identificado por esta pesquisa que a maior parte dos trabalhos sobre registros digitais são publicados em periódicos internacionais.

Tendo em vista a contemporaneidade do tema, o período escolhido foi de 2015 a 2020, de modo que foram encontradas oito publicações. Após a leitura dos resumos e do conteúdo dos trabalhos encontrados, todos foram incluídos como elementos importantes para análise, porém outras referências

foram incorporadas ao estudo, de forma a complementar conceitos e análises sobre o tema.

A pesquisa na base de dados *Scopus* também foi realizada com os operadores booleanos indicados na base anterior. Entretanto, não foram recuperados trabalhos sobre o tema com esse filtro de busca. Desse modo, optou-se pela utilização do termo “*digital preservation metadata*”, no entendimento de que esse termo atenderia aos interesses deste trabalho. Sem limitação de período, para esse termo foram localizadas sete publicações, das quais apenas quatro artigos de periódicos estavam disponíveis para leitura. As outras três tratavam-se de *conference papers* e não estavam disponíveis para consulta em bases gratuitas.

A revisão se deu então por meio das doze publicações selecionadas, sendo oito delas extraídas da *Web of Science* e quatro da *Scopus*.

2.8 PROCEDIMENTO PARA COLETA DE DADOS SOBRE METADADOS DE PRESERVAÇÃO DIGITAL NAS IFES

Igualmente alinhado ao objetivo de identificar os metadados necessários para a preservação digital dos registros digitais arquivísticos, este trabalho buscou conhecer quais metadados dessa finalidade estão sendo utilizados pela UFPB e, de forma complementar, por outras instituições congêneres.

Para isso, foram enviados questionários para todas as universidades federais brasileiras que utilizam o mesmo sistema de produção de registros digitais que a UFPB, a saber, os Sistemas Integrados de Gestão (SIG), plataforma desenvolvida e distribuída pela Universidade Federal do Rio Grande do Norte (UFRN).

Importa esclarecer que tal sistema é constituído de diversos módulos, a exemplo do SIGAA, SIPAC, SIGRH, SIGPP, SIGED, SIGELEIÇÃO, entre outros. Contudo, considerando a amplitude de registros gerados pelos diferentes módulos e a necessidade de metadados específicos para cada um, este trabalho se limitou aos metadados gerados para/por um único módulo.

Com base nos dados obtidos no diagnóstico documental do CCJ, optou-se pela investigação acerca do Módulo Protocolo do SIPAC – Sistema

Integrado de Patrimônio, Administração e Contratos –, que é responsável, dentre outros aspectos, pela produção e tramitação de documentos e processos administrativos, e figura como o *software* com maior número de registros digitais produzidos no âmbito desta pesquisa.

Durante o Encontro Nacional de Arquivistas das Instituições Federais de Ensino Superior – ENARQUIFES (2019), obteve-se a informação de que, das cento e quatro IFES brasileiras, são dezesseis as universidades federais que utilizam o SIPAC para a produção de seus registros digitais.

Com o propósito de conhecer como essas instituições estão realizando a preservação digital por meio dos metadados, foi elaborado um questionário contendo seis perguntas, e esse instrumento de coleta de dados foi enviado para as dezesseis instituições por meio da Plataforma Integrada de Ouvidoria e Acesso à Informação (FalaBR), site criado pelo Governo Federal no ano de 2020, em substituição ao Sistema Eletrônico de Informação ao Cidadão (eSIC), como ferramenta de aplicação da Lei de Acesso à Informação.

O envio dos questionários foi realizado em 08/09/2020 para as seguintes universidades:

Quadro 1 – Universidades Federais que utilizam o SIPAC

Fundação Universidade Federal da Grande Dourados – UFGD
Fundação Universidade Federal de São João Del Rei – UFSJ
Fundação Universidade Federal de Sergipe – UFS
Fundação Universidade Federal do ABC – UFABC
Fundação Universidade Federal do Maranhão – UFMA
Fundação Universidade Federal do Vale do São Francisco – UNIVASF
Universidade Federal da Integração Latino-Americana – UNILA
Universidade Federal da Paraíba – UFPB
Universidade Federal de Itajubá – UNIFEI
Universidade Federal de Lavras – UFLA
Universidade Federal do Oeste da Bahia – UFOB
Universidade Federal do Recôncavo da Bahia – UFRB
Universidade Federal do Rio Grande do Norte – UFRN
Universidade Federal Rural da Amazônia – UFRA
Universidade Federal Rural do Rio de Janeiro – UFRRJ
Universidade Federal Rural do Semiárido – UFERSA

Fonte: dados da pesquisa, 2020.

Dentre as instituições pesquisadas, a UFRA ignorou a manifestação realizada para o seu Sistema de Acesso à Informação e não respondeu o

pedido de informação enviado pela plataforma FalaBR. Já a UNIFEI respondeu que não utiliza os sistemas SIG. As outras treze instituições enviaram respostas consideradas satisfatórias para os objetivos desta pesquisa, cujos resultados serão descritos no item 4.3 Análise dos metadados de preservação digital.

Convém elucidar que, após a realização dessa fase do trabalho, obteve-se a informação no portal de cooperação da UFRN que são 29 universidades federais que utilizam a plataforma SIG e não 16, como apresentado no ENARQUIFES (2019). Entretanto, considerando as dificuldades encontradas em obter as respostas das instituições dentro dos prazos legais, em virtude da adoção do regime de trabalho remoto por ocasião da pandemia de coronavírus, bem como as limitações de cronograma constantes no planejamento deste trabalho, optou-se por realizar a análise com as 16 instituições já pesquisadas, com a compreensão de que essa amostra já compreende mais do que 50% das universidades federais que utilizam o SIPAC.

2.9 PROCEDIMENTO DE ANÁLISE DOS DADOS

Inicialmente, detalhou-se quais são os registros digitais produzidos pelo CCJ/UFPB, primeiramente com a descrição para cada setor, departamento ou colegiado que produz registro digital no CCJ/UFPB. Na sequência, buscou-se descrever quais desses registros digitais são relacionados às atividades de ensino, pesquisa e extensão (atividades-fim) e quais são decorrentes de ações administrativas, financeiras, de controle de patrimônio, gestão de pessoas, dentre outras (atividades-meio), juntamente com a análise das características e elementos que constituem cada um dos registros digitais elencados.

Posteriormente, compõe-se uma proposta de categorização de registros digitais baseada nos modelos disponíveis nos artigos e publicações já existentes sobre o assunto. Para isso, uma vez identificados os registros digitais e suas finalidades, procurou-se relacioná-los com os modelos de categorização de registros digitais presentes na literatura acerca do tema e verificaram-se as possibilidades de utilização de algum ou de mais de um modelo para tais registros.

Em sequência, conhecidos os registros digitais produzidos pelo CCJ/UFPB, suas finalidades e possibilidades de categorização, apresentados nos itens 4.1 e 4.2, buscou-se identificar os metadados necessários para a sua preservação digital, tendo como base os conceitos presentes na literatura, especificamente a proposta de Rogers (2015) para definir registro digital arquivístico, e o trabalho de Sayão (2010) para definir o modelo conceitual para preservação digital de tais registros.

A partir dessa base teórica, constituiu-se um questionário enviado por meio da Plataforma Integrada de Ouvidorias e Acesso à Informação (FalaBR) para as instituições usuárias do SIPAC, e os resultados da pesquisa realizada com as universidades federais que utilizam esse sistema é apresentado no item 4.3.

As respostas obtidas das IFES quanto aos metadados de preservação digital foram tabuladas utilizando o *software Microsoft Excel* para a construção de uma base de dados, que foi analisada com o auxílio do *software* de visualização *Rapid Miner*, cujo detalhamento consta no item 4.3 Análise dos Metadados de Preservação Digital.

3. FUNDAMENTAÇÃO TEÓRICA

Neste capítulo, serão discutidos os conceitos relacionados aos registros digitais arquivísticos, aos documentos arquivísticos digitais, aos modelos de categorização e aos conceitos de metadados de preservação digital encontrados por esta pesquisa.

A seção 3.1, dedicada à revisão de literatura realizada no ano de 2019 sobre gestão de registros digitais para arquivos, se divide em três subseções: a primeira, 3.1.1, para tratar dos dados quantitativos; a segunda, 3.1.2, para tratar dos dados qualitativos; e a terceira, 3.1.3, para tratar da análise realizada.

Já a seção 3.2, dedicada à pesquisa realizada no ano de 2020 sobre documentos/registros arquivísticos digitais, se divide em: seção 3.2.1, sobre as contribuições da literatura nacional, com destaque para tese de Rondinelli (2011) sobre o conceito documento arquivístico frente à realidade digital; seção 3.2.2, sobre os cinco modelos de categorização de registros digitais encontrados durante a pesquisa bibliográfica; e 3.2.3, sobre as perspectivas a respeito dos registros digitais arquivísticos com base em diversos autores e instituições.

Por fim, a seção 3.3, apresenta definições de metadados, suas categorias, e os conceitos relacionados aos metadados de preservação digital, assim como suas relações com a preservação da informação dos registros digitais arquivísticos.

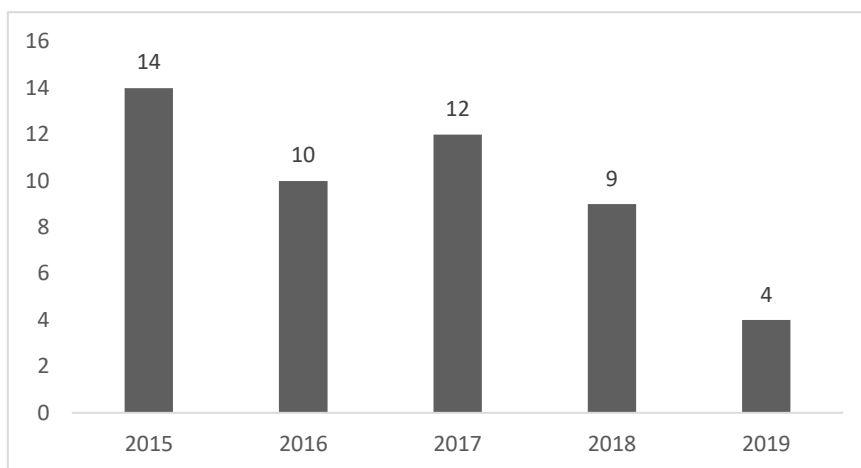
3.1 GESTÃO DE REGISTROS DIGITAIS PARA ARQUIVOS: REVISÃO SISTEMÁTICA DE LITERATURA

Nesta seção, serão apresentados os dados obtidos na revisão sistemática de literatura sobre o tema da gestão dos registros digitais arquivísticos.

3.1.1 Revisão sistemática de literatura – dados quantitativos

Dos 49 artigos selecionados, a maioria foi publicada em 2015, ano em que 14 trabalhos foram encontrados, o que representa 28,6% das publicações do período de 2015 a 2019. Os quantitativos de publicações por ano podem ser observados no gráfico 1:

Gráfico 1 – Publicações por ano



Fonte: dados da pesquisa, 2019.

Ressalta-se que, como o levantamento foi realizado até julho de 2019, é provável que outros textos tenham sido publicados até o final desta pesquisa. Nesta RSL também se buscou identificar em quais áreas do conhecimento os artigos foram produzidos, obtendo-se o resultado constante da figura 2 a seguir.

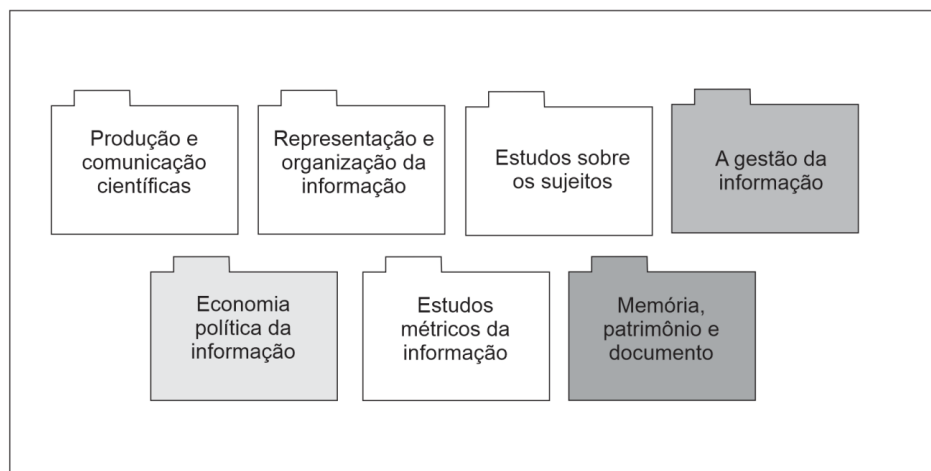
Figura 2 – Publicações por área de conhecimento

Fonte: dados da pesquisa, 2019.

Conforme é possível perceber na figura 2, que representa os temas dos periódicos nos quais os trabalhos foram divulgados, a área de publicação com maior relação com o tema pesquisado nas bases utilizadas e com os critérios definidos é *Information & Knowledge Management*, cuja tradução é Gestão da Informação e do Conhecimento, que possui 18 dos 49 artigos encontrados na busca, representando 36,7% dos trabalhos.

Entende-se que as áreas identificadas por essa revisão estão alinhadas à Ciência da Informação, com base em Araújo (2018). Para isso, é importante esclarecer que existem diferentes linhas de pensamento acerca do que é a Ciência da Informação, de qual é seu objeto, e quais as suas possibilidades de estudo. Essas questões não estão inseridas nos objetivos deste trabalho, entretanto, no âmbito desta pesquisa, utilizar-se-á como referência a divisão da ciência da informação do século XXI proposta por Araújo (2018), conforme figura 3, a seguir:

Figura 3 – As subáreas da ciência da informação no século XXI relacionadas com os registros digitais arquivísticos



Fonte: elaborado pelo autor com base em Araújo (2018).

Araújo (2018) propõe uma ciência da informação constituída de sete áreas, a saber: (a) produção e comunicação científicas; (b) representação e organização da informação; (c) estudos sobre os sujeitos; (d) a gestão da informação; (e) economia política da informação; (f) estudos métricos da informação; e (g) memória, patrimônio e documento.

No escopo desta pesquisa, consideramos que os registros digitais arquivísticos possuem relação com três dessas sete subáreas. Inicialmente, e em menor grau, com economia política da informação, pois a gestão dos registros digitais deverá proporcionar o seu armazenamento estável, sua preservação de acordo com as normas e padrões de segurança, de forma a garantir seu acesso posterior, possibilitando, assim, o direito de utilização dessa informação por parte da sociedade. Para Araújo (2018), a área de economia política de informação

[...] constitui-se com estudos voltados para a democratização da informação, [...] a criação de formas e sistemas alternativos de informação [...] bem como a identificação de uma responsabilidade social da ciência da informação como agente de inclusão e de fornecimento de informação como insumo para apoiar ações racionais específicas (WERSIG; NEVELLING, 1975 *apud* ARAÚJO, 2018, p. 44-45).

Em segundo lugar, em grau intermediário, o estudo dos registros digitais arquivísticos relaciona-se com a subárea da gestão da informação, uma

vez que a adoção de sistemas eletrônicos de produção de registros digitais visa, dentre outras coisas, a utilização da informação de maneira estratégica, a padronização de processos de trabalho, a virtualização de elementos da cultura organizacional, dentre outros. Segundo Araújo (2018),

[...] o entendimento sobre o significado de se estar numa sociedade pós-industrial (sociedade da informação ou ainda sociedade do conhecimento) foi se ampliando, de tal forma que foi sendo percebido que a informação que constitui um recurso importante para as organizações não é aquela que existe materialmente, mas aquela que ainda não existe como entidade física, que está na mente das pessoas que pertencem à organização. [...] Não bastava gerir os recursos informacionais, era preciso gerir também o conhecimento, criando condições próprias para transformá-lo em informação. (ARAÚJO, 2018, p. 43-44).

Dentre as subáreas da ciência da informação elencadas por Araújo (2008), aquela com a qual este trabalho possui mais identificação é Memória, patrimônio e documento. Para o autor, nessa subárea,

[...] a informação é entendida como o processo a partir do qual os indivíduos valorizam determinados registros e, nesse processo, participam do processo de construção da memória, portanto da cultura, e do real. Os serviços, produtos e sistemas de informação são estudados como atores na conformação de determinados quadros de sentido, modos de apreensão, inferências na valorização de determinadas fontes, documentos ou registros (ARAÚJO, 2018, p. 76).

Ou seja, uma pesquisa que busca descrever quais são os registros digitais produzidos por determinada instituição, suas relações com o conceito de documento arquivístico digital e quais metadados são necessários para a sua preservação digital, pode ser caracterizada como um trabalho de ciência da informação que se enquadra em três das sete subáreas da CI propostas por Araújo (2018).

Dando sequência à apresentação da RSL, percebe-se que as bases *Emerald*, *Lisa*, *Scopus* e *Web of Science* indexam periódicos de diversas áreas, o que corrobora o caráter interdisciplinar da Ciência da Informação. Nesse sentido, para aprimorar a compreensão das publicações, analisou-se as palavras-chave constantes nos artigos selecionados e verificou-se a existência

de 154 palavras-chave atribuídas aos 49 artigos analisados, das quais 132 (85,7%) palavras-chaves foram utilizadas apenas uma vez, e 22 (14,28%) mais de uma vez.

As palavras-chave com ocorrência superior a 2 trabalhos foram escolhidas para a composição da tabela 3, pois contribuem na explicação da multiplicidade de possibilidades de investigação do fenômeno da gestão de registros digitais para arquivos e dos conceitos relacionados ao objeto da pesquisa.

Tabela 3 – Palavras-chave mais utilizadas na RSL sobre registros digitais para arquivos

Palavras-chave	Tradução ³	Frequência	%	Acumulada
<i>Appraisal</i>	Avaliação	4	2,6	2,6
<i>Archival science</i>	Arquivologia ou ciência arquivística	2	1,3	3,9
<i>Archives</i>	Arquivos	9	5,8	9,7
<i>Archives management</i>	Gestão de arquivos	4	2,6	12,3
<i>Born-digital archives</i>	Arquivos nato digitais	2	1,3	13,6
<i>Digital archives</i>	Arquivos digitais	4	2,6	16,2
<i>Digital forensics</i>	Diplomática digital	2	1,3	17,5
<i>Digital humanities</i>	Humanidades digitais	3	1,9	19,5
<i>Digital information</i>	Informação digital	2	1,3	20,8
<i>Digital preservation</i>	Preservação digital	4	2,6	23,4
<i>Digital records</i>	Registros digitais arquivísticos	2	1,3	24,7
<i>Digital records management</i>	Gestão de registos digitais arquivísticos	2	1,3	26,0
<i>Digital storage</i>	Armazenamento digital	2	1,3	27,3
<i>Digitalization</i>	Digitalização	2	1,3	28,6
<i>Eletronic records</i>	Registros eletrônicos/digitais	3	1,9	30,5
<i>Eletronic records management</i>	Gestão de registros eletrônicos/digitais	6	3,9	34,4
<i>Information management</i>	Gestão da informação	6	3,9	38,3
<i>Information systems</i>	Sistemas de informação	2	1,3	39,6
<i>Preservation</i>	Preservação	2	1,3	40,9
<i>Provenance</i>	Proveniência	2	1,3	42,2
<i>Recordkeeping</i>	Manutenção de registros arquivísticos	6	3,9	46,1
<i>Records management</i>	Gestão de registros arquivísticos	14	9,1	55,2

Fonte: dados da pesquisa, 2019.

A análise feita a partir da tabela 3 permite destacar que a palavra “digital” é utilizada em oito contextos diferentes; o item lexical “gestão” é atribuído a cinco outros termos; a palavra “arquivo” tem três possibilidades de aplicação; e o conceito de “registro” é utilizado em dois termos distintos. Outros dois termos presentes, que são *appraisal* e *provenance*, se referem à função arquivística de avaliação e ao conceito clássico da Arquivologia de proveniência ou respeito aos fundos.

³ As traduções dos termos foram realizadas com base no Glossário Multilíngue de Terminologia Arquivística do Conselho Internacional de Arquivos (ICA) e no Glossário de Terminologia para Arquivos e Registros da Sociedade dos Arquivistas Americanos (SAA).

Tais achados revelam as diversas possibilidades de estudos acerca da gestão de registros digitais para arquivos e confirmam o entendimento de que essas publicações oferecem uma amplitude significativa de contribuições para a compreensão do objeto desta pesquisa.

Os dados apontam a predominância das palavras *management*, *records*, *digital* e *archives*, cuja tradução é, respectivamente, gestão, registros, digital e arquivos, nas palavras-chave indexadas nos artigos constantes desta revisão. Salienta-se que *archives* é comumente utilizado para referir-se aos documentos históricos ou aos arquivos históricos, enquanto *records* se refere aos documentos arquivísticos, em geral, mas seu uso mais comum na América do Norte (origem de 34,7% dos artigos contemplados na pesquisa) refere-se aos documentos arquivísticos em fase corrente e intermediária. No contexto dos arquivos, para Pearce-Moses (2005, p. 239), o termo *management* é definido como “a administração, organização, supervisão e condução de um negócio ou atividade”, enquanto que digital, para o mesmo autor, conceitua-se como “representação de informações através de uma sequência de unidades distintas, especialmente códigos binários”, definição semelhante àquela apresentada pelo Arquivo Nacional (BRASIL, 2005).

A figura 4, a seguir, foi produzida com o *software* NVivo e ilustra a nuvem das palavras-chave de todos os artigos selecionados.

Tabela 4 – Produção por autor

AUTOR	QTD
Klareld, Ann-Sofie	2
Ngoepe, Mpho	2
Niu, Jinfang	2
Katuu, Shadrack	2
Pan, Weimei	2

Fonte: dados da pesquisa, 2019.

Foram identificadas 201 citações referenciadas pelas bases de dados alusivas aos 49 artigos selecionados por esta revisão. Destas, 139 foram atinentes aos artigos produzidos por 10 autores, compreendendo 69% das citações realizadas. Esses dados foram sistematizados na tabela 5 a seguir:

Tabela 5 – Citações por autor no *Research Gate*

AUTOR	QTD
Conway, P.	32
Huvila, Isto	23
Dallas, Costis	20
Rogers, Corinne	20
Katuu, Shadrack	13
Baron, J.R.	7
Cohen, Frederick B.	7
Guo, W.	6
Henttonen, P.	6
Bhebhe, S.	5

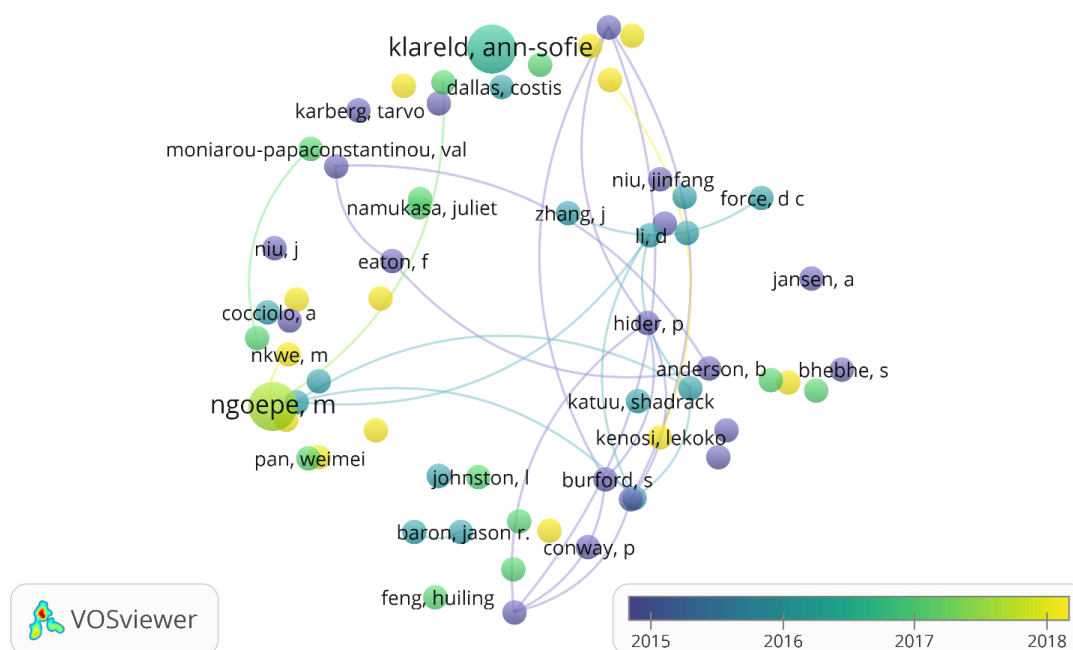
Fonte: dados da pesquisa, 2019.

Realizou-se uma análise dos artigos produzidos pelos 10 autores constantes na tabela 5 e verificou-se que apenas Huvila (2015) não tratava de forma preponderante da gestão de registros digitais para arquivos. Seu trabalho intitulado “*The unbearable lightness of participating? Revisiting the discourses of ‘participation’ in archival literature*”⁴, se dedica a uma análise do conceito de participação na literatura e nos discursos de arquivistas e gestores de registros e as implicações desse fenômeno nas suas práticas profissionais, e traz como uma de suas conclusões a existência de uma nova era nos arquivos, marcada por algumas revoluções, dentre elas a tecnologia e seus usos aplicados ao arquivo.

⁴ Que poderia ser traduzido como “A leveza insuportável da participação? Revisitando os discursos de ‘participação’ na literatura arquivística”.

A figura 5, a seguir, produzida com o *software VOS viewer* a partir dos dados bibliográficos, ilustra a rede de coautorias que se estabelece nos 49 artigos selecionados para essa revisão sistemática de literatura.

Figura 5 – Rede de coautorias



Fonte: dados da pesquisa, 2019.

Verificou-se que apenas 9 trabalhos foram realizados por mais de um autor, o que representa 18,4% dos trabalhos selecionados. Desses, apenas o autor Mpho Ngoepe, da Universidade da África do Sul, aparece em mais de uma publicação. Em coautoria com Marcia Nkwe, publicaram o artigo *Separating the wheat from the chaff with the winnowing fork: the eeny meeny miny mo appraisal approach of digital records in South Africa*⁵. Com Katuu Shadrack, Mpho Ngoepe publicou o artigo *Provision of Records Created in Networked Environments in the Curricula of Institutions of Higher Learning in Africa*⁶.

Na sequência buscou-se obter dados de altimetria dos responsáveis pelas publicações. Os autores escolhidos para a pesquisa das altimétricas

⁵ “Separando o joio do trigo com o garfo de joeiragem: a abordagem ‘uni-duni-tê’ da avaliação de registros digitais na África do Sul”, tradução nossa.

⁶ “Provisão de Registros Criados em Ambientes em Rede nos Currículos das Instituições de Ensino Superior na África”, tradução nossa.

foram aqueles que, de acordo com o Google Acadêmico, possuíam mais de 5 citações para os trabalhos constantes da amostra de 49 artigos desta revisão de literatura, que totalizam 10 autores.

A fonte para a pesquisa das altimétricas foi a rede social científica *Research Gate*. Os autores Jason R. Baron, Frederick B. Cohen, Wei Guo e Sindiso Bhebhe não possuem perfil nessa rede e por esse motivo suas altimétricas não puderam ser obtidas, de tal modo que a tabela 4, a seguir, apresenta as altimétricas de 6 autores, dentre os 10 selecionados.

Tabela 6 – Altimétricas dos autores

Nome do autor	<i>Research Interest</i>	<i>Citations</i>	<i>Reads</i>	<i>Recommendations</i>	<i>RG Score</i>
Shadrack Katuu	591,1	273	17671	1407	12,96
Isto Huvila	555,3	781	21737	27	27,72
Costis Dallas	130,7	192	2832	9	7,94
Paul Conway	86,6	13	812	9,1	9,51
Corinne Rogers	74,2	68	3659	29	5,69
Pekka Henttonen	44,3	47	7967	2	8,24

Fonte: dados da pesquisa, 2019.

Foram extraídos da rede *Research Gate* os índices denominados de *Research Interest*, *Citations*, *Reads* e *Recommendations*. De acordo com a rede social acadêmica, o índice *Research Interest* representa a soma do interesse de pesquisa de cada item de pesquisa que um autor adicionou ao seu perfil, em que cada item tem sua própria pontuação, com base nas suas citações, recomendações e leituras. Já o índice *Reads*, de acordo com o site, é contado sempre que alguém visualiza um resumo de publicação, clica em uma figura, exibe ou baixa o texto completo, clica em um projeto ou faz um questionamento ao autor.

Destacam-se os autores Katuu e Huvila, que juntos representam 72% das leituras dos autores constantes da seleção. Tal métrica pode se confirmar pelo indicador de interesse de pesquisa dos dois, de 591,1 e 555,3, respectivamente, que representam 77% do total. Esses dois autores também possuem os maiores índices de citações e recomendações pelos pares na rede social científica *Research Gate*.

Há ainda o indicador *RG Score*, que representa a métrica mais importante da rede social científica *Research Gate*, equivalente a uma noção de pontuação geral do pesquisador relacionada aos itens anteriores, para o qual Huvila está em primeiro lugar, seguido de Katuu, com menos da metade da pontuação do primeiro.

No período de 4 a 6 de junho de 2019, foram realizadas buscas no Google Acadêmico acerca da quantidade de citações por artigo para o conjunto dos 49 trabalhos desta revisão. Verificou-se que 71,4% (35 trabalhos) possuíam citações e 28,6% (14 trabalhos) não foram citados.

Dentre os 35 trabalhos com citações, a tabela 7 evidencia os 20 artigos mais citados. Os outros 15 distribuíram-se da seguinte forma: 2 trabalhos tiveram 3 citações cada; 5 trabalhos receberam 2 citações cada; e 8 trabalhos tiveram uma única citação.

Tabela 7 – Títulos, autores, periódicos e ano dos artigos mais citados

Título do artigo	Autor (es)	Periódico	Citações	Ano
<i>Digital transformations and the archival nature of surrogates</i>	Conway, P.	<i>Archival Science</i>	32	2015
<i>The unbearable lightness of participating? Revisiting the discourses of “participation” in archival literature</i>	Huvila, Isto	<i>Journal of Documentation</i>	23	2015
<i>Digital curation beyond the “wild frontier”: a pragmatic approach</i>	Dallas, Costis	<i>Archival Science</i>	20	2016
<i>Diplomatics of born digital documents – considering documentary form in a digital environment</i>	Rogers, Corinne	<i>Records Management Journal</i>	20	2015
<i>Managing digital records in a global environment: A review of the landscape of international standards and good practice guidelines</i>	Katuu, Shadrack	<i>The Electronic Library</i>	13	2016
<i>What lessons can be learned from the US archivist’s digital mandate for 2019 and is there potential for applying them in lower resource countries?</i>	Baron, J. R.; Thurston, A.	<i>Records Management Journal</i>	7	2016
<i>Digital diplomatics and forensics: going forward on a global basis</i>	Cohen, Frederick B.	<i>Records Management Journal</i>	7	2015
<i>Archives as a trusted third party in maintaining and preserving digital records in the cloud environment</i>	Guo, W.; Fang, Y.; Pan, W.; Li, D.	<i>Archival Science</i>	6	2016
<i>Privacy as an archival problem and a solution</i>	Henttonen, P.	<i>Records Management Journal</i>	6	2017

<i>Contemporary diplomatics of the civil and deceased estate case files found at the national archives of Zimbabwe</i>	<i>Bhebhe, S.</i>	<i>Archives and Records</i>	5	2015
<i>Event-based archival information organization</i>	<i>Niu, Jinfang</i>	<i>Archival Science</i>	5	2015
<i>Born-digital archives at the Wellcome Library: appraisal and sensitivity review of two hard drives</i>	<i>Sloyan, V.</i>	<i>Records Management Journal</i>	5	2016
<i>Identity and archives: return and expansion of the social value of archives</i>	<i>Feng, Huiling</i>	<i>Library Review</i>	4	2017
<i>Knowledge discovery from within: An examination of records management and electronic records management syllabi</i>	<i>Force, D. C.; Zhang, J.</i>	<i>Records Management Journal</i>	4	2016
<i>Isn't it information assets we're really talking about?' A discourse analysis of a panel discussion on digital archives</i>	<i>Klareld, Ann-Sophie</i>	<i>Archival Science</i>	4	2015
<i>Beyond preservation: investigating the roles of archivist</i>	<i>Vassilakaki, E.; Moniarou-Papaconstantinou, V.</i>	<i>Archives and Records</i>	4	2017
<i>Email as cultural heritage resource: appraisal solutions from an art museum context</i>	<i>Cocciolo, A.</i>	<i>Records Management Journal</i>	3	2016
<i>The future of access to public records? Public-private partnerships in US state and territorial archives</i>	<i>Kriesberg, Adam</i>	<i>Archival Science</i>	3	2017
<i>Separating the wheat from the chaff with the winnowing fork: The eenymeenyminyo appraisal approach of digital records in South Africa</i>	<i>Ngoepe, M.; Nkwe, M.</i>	<i>Records Management Journal</i>	3	2018
<i>Original order in the digital world</i>	<i>Niu, J.</i>	<i>Archives and Manuscripts</i>	3	2015

Fonte: dados da pesquisa, 2019.

Os trabalhos apresentados na tabela 7 foram identificados como relevantes tomando como base o número de citações. O número de citações de um trabalho acadêmico é um dos indicadores mais tradicionais existentes nas análises bibliométricas. Contudo, com o surgimento de novas formas de disseminação da produção acadêmica, nesta revisão optou-se por verificar a relação deste indicador com outros.

Os resultados das duas bases não é uniforme. É possível identificar, na análise do índice de citações do Google Acadêmico em comparação com o do *Research Gate*, que embora o autor com maior número de citações naquela base, Isto Huvila, com 723 citações, tem apenas 23 citações nesta. Contrariamente, Paul Conway tem apenas 13 citações indexadas pelo *Research Gate*, estando na sexta e última posição no *ranking*, e figura em

primeiro lugar no *ranking* da tabela 7, com 32 citações referenciadas pelo Google Acadêmico.

É importante destacar que o indicador de citação do *Research Gate* leva em consideração todos os artigos produzidos pelos autores e cadastrados naquela rede social acadêmica, enquanto que o número do Google Acadêmico é diretamente vinculado ao artigo que compõe esta revisão sistemática de literatura.

No que se refere à universidade de origem, dos 49 trabalhos escolhidos, foi possível obter a informação em 42 artigos, que representam 86% do total. Os 7 trabalhos, cuja universidade de origem não pôde ser identificada, são vinculados às seguintes instituições: Instituto de Tecnologia Educacional de Atenas; *Welcome Trust*, de Londres; Arquivo Nacional de Zimbábue; Instituto Pratt, de Nova Iorque; Administração Nacional de Documentos e Arquivos, dos Estados Unidos; Ministério da Cultura, da Itália; Instituto Nacional Francês de AudioVisual.

Dos 42 artigos com origem identificada em instituições de ensino superior, verificou-se que foram produzidos por 34 universidades, presentes nos cinco continentes.

Como último elemento de análise quantitativa, cada um dos artigos escolhidos para esta revisão de literatura teve a quantidade de *downloads* conferida nas bases Lisa, Emerald, *Web of Science*, *Scopus* e na plataforma Google Acadêmico. Desse modo, verificou-se que o conjunto dos 49 artigos escolhidos foi baixado 45.680 vezes, conforme ilustra a tabela 5, a seguir.

Para a construção da tabela 5 foram selecionados os dez artigos mais baixados, utilizando como critério o número mínimo de um mil e quinhentos *downloads*. É importante destacar que, dos 49 artigos, apenas 4 não possuíam *downloads* registrados no Google Acadêmico e nas bases selecionadas. Os 11 artigos com os maiores números de *downloads* representam mais de 50% do total de *downloads* registrados.

Tabela 8 – Artigos mais baixados

Título do artigo	Ano	Downloads	%	% Acumulado
<i>Born-digital archives at the Wellcome Library: appraisal and sensitivity review of two hard drives</i>	2016	3546	7,76	7,76
<i>Digital curation beyond the “wild frontier”: a pragmatic approach</i>	2016	2900	6,35	14,11
<i>Beyond preservation: investigating the roles of archivist</i>	2017	2550	5,58	19,69
<i>Digital transformations and the archival nature of surrogates</i>	2015	2200	4,82	24,51
<i>Knowledge discovery from within: An examination of records management and electronic records management syllabi</i>	2016	2042	4,47	28,98
<i>Methods, methodology and madness: Digital records management in the Australian government</i>	2017	1898	4,15	33,13
<i>Managing digital records in a global environment: A review of the landscape of international standards and good practice guidelines</i>	2016	1847	4,04	37,18
<i>The unbearable lightness of participating? Revisiting the discourses of “participation” in archival literature</i>	2015	1705	3,73	40,91
<i>Archival appraisal practice in U.S. state archives and records management programs</i>	2016	1700	3,72	44,63
<i>The implementation of electronic recordkeeping systems</i>	2017	1687	3,69	48,33

Fonte: dados da pesquisa, 2019.

Feita a análise quantitativa dos artigos, na seção posterior serão apresentados os conceitos encontrados na revisão de literatura sobre a gestão de registros digitais arquivísticos.

3.1.2 Revisão sistemática de literatura – dados qualitativos

Compreender a gestão dos registros digitais arquivísticos com base na literatura sobre o tema é uma tarefa que exige, inicialmente, a apreensão de alguns conceitos. O primeiro deles é o conceito de arquivo que, de acordo com Paes (2007), é a

[...] designação genérica de um conjunto de documentos produzidos e recebidos por uma pessoa física ou jurídica, pública ou privada, caracterizado pela natureza orgânica de sua acumulação e conservado por essas pessoas ou por seus sucessores, para fins de prova ou informação (PAES, 2007, p. 24).

Já o termo registro, de acordo com o Dicionário Brasileiro de Terminologia Arquivística, publicado pelo Arquivo Nacional (2005), é definido inicialmente como “anotação sistemática em livro próprio” ou “unidade de informação logicamente indivisível”.

Considerando que esta revisão utilizou para a constituição da amostra de trabalhos a serem analisados o termo *record*, em língua inglesa, verificou-se o conceito desse termo na literatura clássica da área de arquivos. *Records*, para Schellemborg (1956), são

Todos os livros, papéis, fotografias, ou outros materiais documentais, independentemente da forma física ou de características, elaborados ou recebidos por qualquer instituição pública ou privada no exercício de suas obrigações legais ou em conexão com a transação de seu próprio negócio e preservados ou destinados à preservação por aquela instituição ou seu legítimo sucessor como evidência de suas funções, políticas, decisões, procedimentos, operações ou outras atividades, ou por causa do valor informacional dos dados ali contidos (SCHELLENBERG, 1956, p. 16).

Buscou-se, ainda, a definição de *record* do Glossário de Terminologia para Arquivos e Registros da Sociedade dos Arquivistas Americanos. De acordo com Pearce-Moses (2005), existem as seguintes definições para o termo *record*:

1. Obra escrita ou impressa, de natureza legal ou oficial, que pode ser usada como evidência ou prova; um documento. - 2. Dados ou informações que foram corrigidos em algum meio; que tenha conteúdo, contexto e estrutura; e isso é usado como uma extensão da memória humana ou para demonstrar responsabilidade. - 3. Dados ou informações de forma fixa, criados ou recebidos no decurso de uma atividade individual ou institucional e retirados (preservados) como evidência dessa atividade para referência futura. - 4. Um instrumento arquivado para aviso público (aviso construtivo); - 5. Áudio · Um registro fonográfico. - 6. Computação · Uma coleção de elementos de dados relacionados tratados como uma unidade, como os campos em uma linha em uma tabela de banco de dados. - 7. Descrição · Uma entrada que descreve um trabalho em um catálogo; um registro de catálogo (PEARCE-MOSES, 2005, p. 326).

Com a definição de arquivo e de registro, e sabendo que o adjetivo digital diz respeito ao ambiente no qual o registro foi produzido ou armazenado, parte-se para a reflexão do que são os registros digitais arquivísticos. Com base nas referências utilizadas, não sendo a intensão desta revisão encerrar, mas apenas iniciar essa reflexão, o que são os registros digitais arquivísticos?

Inicialmente, utilizou-se a explicação de Rondinelli (2018, p. 142) para compreender a diferença entre documento de arquivo e documento arquivístico, sendo este o que se relaciona com as atividades do produtor/acumulador e aquele relacionado ao pertencimento ao local da sua guarda. De forma análoga, o termo “registro digital arquivístico” diz respeito ao registro criado em ambiente digital e que tenha natureza arquivística, ou seja, qualidade daquele que foi produzido ou recebido no decorrer das atividades de uma pessoa física ou jurídica. Essa reflexão será aprofundada nos itens seguintes deste capítulo. Partindo desses pressupostos, iniciam-se as contribuições dos autores presentes na revisão sistemática de literatura.

Os arquivos sempre tiveram uma função importante ao longo do tempo. Guo (2016) realiza uma contextualização histórica sobre o papel dos arquivos como instâncias responsáveis por dotar os registros de autenticidade quando avaliam os documentos e classificam-nos como elementos de prova ou evidência.

Um dos objetivos desse trabalho foi investigar, na literatura selecionada, a função do arquivo ou as possibilidades de contribuição do conhecimento arquivístico no contexto de produção de documentos em ambiente digital. Nesse sentido, Hoskins (2009) afirma que o arquivo, na perspectiva digital, é um novo lugar de memória, livre de espaço e restrições institucionais. Já Conway (2015) define os arquivos como construções sociais cujos significados dos seus rastros mudam com o tempo conforme são descritos, transformados e usados. Klareld (2015) complementa afirmando que o foco do arquivo não é apenas guardar e preservar, mas promover a reutilização da informação, ou seja, mudar a visão do arquivo para ser visto como um recurso, pois a informação já é vista assim.

O segundo conceito importante para esta pesquisa é o de registro. Os registros podem ser entendidos como documentos, informações registradas, produzidas ou recebidas no decurso da atividade prática, mas também como

instrumentos ou subprodutos dessa atividade, e há, de acordo com Rogers (2015), algo que os diferenciam: aos registros devem ser reservadas ações e recomendações adicionais.

Sem estabelecer tais condições, Henttonem (2017) concorda que registros são produzidos para documentar o desempenho de uma determinada tarefa organizacional. Já Klareld (2015) define registros como documentos mantidos como evidência.

Guo (2016), em contrapartida, apresenta como condições para que uma entidade eletrônica seja considerada um registro, que tenha forma fixa e conteúdo estável. Ou seja, os dados obtidos nos sistemas que fazem a gestão das atividades da empresa precisam ser convertidos em um formato não estruturado e imutável, como é o caso do PDF. Para o autor, a necessidade da geração de documentos em formato não estruturado tem como objetivo dar ao registro a possibilidade de servir como prova ou como ferramenta para as atividades de documentação e monitoramento, para o caso de ser necessária a reconstrução de quaisquer alterações possivelmente feitas nos bancos de dados. Contudo, essa abordagem não é uma unanimidade, pois se trata de uma discussão em construção na área.

As instituições, em sua maioria, não têm consciência de que os seus sistemas de tecnologia criam registros, que precisam ser capturados e protegidos quanto ao seu valor de prova. Para Baron e Thurston (2016, p. 214) a falta de gestão pode permitir que os registros digitais sejam “alterados, apagados, fragmentados, corrompidos ou perdidos, seja através da interferência ou simplesmente através da obsolescência tecnológica ou de um quadro de controle inadequado”.

Sobre registros digitais, Rogers (2015) estabeleceu o seguinte conceito:

Os registros digitais consistem em dados gerados pelo usuário (conteúdo), metadados gerados pelo sistema identificação de origem e localização, metadados gerados pelo aplicativo gerenciando a aparência e desempenho do registro (exemplo: formato de arquivo nativo), metadados gerados por aplicativo descrevendo os dados (exemplo: metadados do sistema operacional do sistema de arquivos) e metadados gerados pelo usuário descrevendo os dados (ROGERS, 2015, p. 16, p. 18 tradução nossa).

Com base em Rogers (2015), tem-se que, se as formas documentais não digitais necessitam de regras para verificação de autenticidade e confiabilidade, os registros digitais considerados semelhantes aos registros analógicos no que se refere à natureza arquivística (Conway, 2015; Katuu, 2016) também precisam de padrões para julgar seu valor de prova documental ou evidência.

Sendo assim, buscou-se a definição da gestão de registros digitais para arquivos. Para Henttonen (2017), o gerenciamento de registros em arquivos consiste, entre outras coisas, de um conjunto de processos que transferem e oferecem acesso à informação através de contextos espaciais e limites temporais.

Considerando serem os registros digitais objetos de diferentes análises e interpretações, o quadro 2 traz algumas especificidades dos registros digitais estabelecidos por autores desta RSL.

Quadro 2 – Especificidades dos registros digitais

Item	Descrição	Autor
Atributos	Forma fixa, conteúdo estável, ligações explícitas a outros registros, contextos identificáveis, identificação de pessoas envolvidas na formação do registro e uma ação na qual o registro participa.	Guo, 2016 e Rogers, 2015
Requisitos	Os registros digitais públicos devem ter presunção de confiabilidade, autenticidade e precisão.	Cohen, 2016
Crítérios	Transparência de preservação dos registros, segurança e estabilidade dos registros.	Guo, 2016
Categorias	(1) documentos criados usando aplicativos de escritório, como documentos de texto, planilhas e apresentações; (2) registros em ambientes <i>on-line</i> ou baseados na <i>Web</i> , como intranet, extranets e registros de transações <i>online</i> ; (3) registros gerados por sistemas de informações de negócios, como bancos de dados, sistemas de dados geoespaciais, sistemas de recursos humanos, sistemas financeiros, sistemas de fluxo de trabalho e sistemas de gerenciamento de conteúdo; e (4) sistemas de comunicação digital, tais como e-mail, serviços de mensagens curtas, serviços de mensagens multimídia, intercâmbio eletrônico de dados e fax eletrônico.	Katuu, 2016

Fonte: dados da pesquisa, 2019.

O item categorias será explorado na seção 3.2.2 desta pesquisa, porém, tais explicações sobre os registros digitais permitem visualizar a amplitude de abordagens possíveis para o tema e a necessidade de estudos sobre seus atributos, requisitos e critérios a serem observados nas possibilidades encontradas e em outras que possam existir.

Nesta revisão, além dos conceitos que já foram apresentados, o principal conceito abordado foi o de metadados, sem o qual, segundo os autores, não é possível realizar a gestão dos registros digitais para arquivos. Do mesmo modo, os termos “diplomática digital”, “curadoria digital”, “preservação digital” e “arquivamento digital” também foram considerados relevantes na análise dos trabalhos revisados. Logo, na sequência, serão apresentadas as definições dos autores revisados para cada um dos conceitos e termos em questão.

A gestão de metadados está intimamente ligada à gestão de registros digitais para arquivos (BARON; THURSTON, 2016; COHEN, 2015; ROGERS, 2015). Para autores como Niu (2015) e Conway (2015), os metadados são as formas de avaliação da proveniência dos registros digitais. Entretanto, nenhum dos autores traz um conceito sobre o termo: seus trabalhos concentram-se no estabelecimento de métodos para a utilização desses dados para verificação de autenticidade, confiabilidade e integridade – requisitos considerados fundamentais para a preservação e a gestão efetiva dos registros digitais.

Quando se trata de registros analógicos (não digitais), a diplomática pode contribuir na verificação da autenticidade. No caso dos registros digitais, Rogers (2015) escreve que a Diplomática digital, “fundamentada em uma base de princípios tradicionais da diplomática clássica, pode ajudar a identificar registros digitais através de seus metadados e determinar quais metadados precisam ser capturados, gerenciados e preservados” (ROGERS, 2015, p. 18).

A relação da diplomática com o objeto desta revisão também é identificada nos trabalhos de Cohen (2015), que analisa a necessidade de utilização dos metadados como elementos da análise diplomática, e de Bhebe (2015), quando aborda a diplomática como fonte para a construção de requisitos para verificação da autenticidade de recursos eletrônicos.

Com a compreensão de que os metadados constituem uma categoria de análise fundamental para o estudo dos registros digitais arquivísticos, o

tema foi aprofundado em uma revisão de literatura específica descrita na seção 3.3.

Dando continuidade, Dallas (2015) traz o conceito de curadoria digital como proposta de torná-lo um campo de investigação autônomo. Para o autor, o conceito tem sido utilizado de forma prescritiva, ou seja, para desenvolver regras normativas, procedimentos e sistemas, de modo a proporcionar a preservação digital.

De acordo com Guo (2016), a diferença básica entre curadoria digital e preservação digital é que esta assume um papel passivo, enquanto aquela enfatiza uma intervenção ativa e uso futuro de recursos digitais.

Cunningham (2008) ratifica o entendimento de Guo (2016) e vai ainda mais longe ao definir que a curadoria de registros digitais é uma atividade suficientemente distinta para justificar o uso do termo arquivamento digital, isto é, propõe que a Arquivologia, seus profissionais e pesquisadores, iniciem as reflexões sobre as possibilidades de ampliação da área ou até de constituição de uma subárea independente que trate dessa questão, com conhecimentos e aplicações específicas.

Para realizar a gestão dos registros digitais, Conway (2015) e Niu (2015) defendem a necessidade de ressignificação do conceito de proveniência.

Proveniência de registros originalmente significava a pessoa, família ou organização que tenha criado ou acumulado os registros. Durante as últimas décadas, os arquivistas perceberam a importância da função como proveniência de registros. A função tem sido considerada mais estável que a estrutura organizacional e tem sido usada como base para classificar e organizar registros (NIU, 2015, p. 319).

No contexto de digitalização de documentos, Conway (2015) afirma que o representante digital, que compreende o resultado da digitalização de um documento arquivístico, precisa ser analisado com um novo olhar. O termo “representante digital” se refere à tradução do termo “*digital surrogate*”, utilizada pelo autor e feita com base no trabalho de Luz e Flores (2017), autores brasileiros na área de documentos digitais, que se referem aos documentos provenientes de digitalização utilizando essa expressão.

Para Conway (2015), a natureza arquivística dos representantes digitais é uma oportunidade intelectual substancial para a Arquivologia e defende como proposta, para o acesso de longo prazo às coleções de representantes digitais, o seu tratamento como registros arquivísticos. De acordo com o autor, representantes digitais possuem os rastros próprios dos seus ciclos de vida, são fluidos e, portanto, merecem gerenciamento e preservação como arquivos.

A esse conceito, Conway (2015) atribui o termo “proveniência secundária”, apresentado como escolha para os casos de redigitalização, reingestão de múltiplas versões em repositórios arquivísticos digitais e substituição de fato das fontes originais.

A importância dos repositórios digitais também é identificada nesta revisão. No trabalho de Guo (2016), os repositórios digitais confiáveis são definidos como aqueles repositórios cuja missão é fornecer acesso confiável e de longo prazo a recursos digitais gerenciados à comunidade designada, agora e no futuro.

Sobre proveniência, Niu (2015) defende um novo tipo de abordagem para o termo (proveniência) composto das atividades de organizar e descrever registros digitais baseada em eventos de gerenciamento. Para o autor, tal método tem crescido e sido implementado com sucesso fora da comunidade arquivística e pode ser uma alternativa para a organização da informação e um redesenho dos metadados de descrição para registros de arquivo.

Nota-se, diante da quantidade de abordagens identificadas, quão árdua, permeada de interpretações acerca dos conceitos identificados e como é grande a necessidade de estudos mais aprofundados sobre cada um dos temas que lhes são pertinentes, com destaque para a “gestão de registros digitais arquivísticos”.

Sua complexidade é evidenciada em parte significativa dos trabalhos revisados, de maneira que o quadro 3 representa as principais dificuldades apontadas pelos autores para o tema em questão.

Quadro 3 – Dificuldades para a gestão de registros digitais

Descrição	Autor
Inviabilidade do gerenciamento com os recursos de gestão tradicionais em virtude das características específicas dos objetos digitais	Rogers (2015)
Ausência norma legal compatível para a gestão dos registros digitais	Katu (2016)
Perda de registros digitais em virtude de ausência de programa de preservação	
Falta de iniciativas colaborativas regionais e internacionais para preservação de registros digitais públicos	
Falta de monitoramento da evolução tecnológica e suas consequências para os registros digitais	Bhebhe (2015) e Vassilakaki e Moniarou-Papaconstantinou (2017)
Incapacidade dos profissionais envolvidos na gestão de registros eletrônicos por falta de treinamento adequado	
Custo de armazenamento, processamento e segurança dos registros em ambiente digital	Klareld (2015)
Atribuição de responsabilidade pela custódia e gestão dos registros públicos compartilhados entre diferentes órgãos de governo	
Ambiguidade na interpretação dos dados de mensuração dos registros digitais	
Ausência de convenção acerca do uso dos registros digitais com informações privativas	Henttonem (2017)
Restrições orçamentárias para questões de arquivo	Vassilakaki e Moniarou-Papaconstantinou (2017)
Altas taxas de entrada de material nos sistemas de gestão de registros digitais para arquivos	

Fonte: dados da pesquisa, 2019.

Além dos desafios evidenciados, há, na gestão de registros digitais, também uma possibilidade de contribuição para diversas áreas, seja na evolução da diplomática forense, quando os registros digitais passam a conter os elementos necessários para constituírem evidências e provas em casos judiciais, ou ainda na criação de repositórios de registros digitais que favoreçam o desenvolvimento sustentável global com a manutenção e preservação de informações estratégicas para a defesa dos interesses dos países (Baron e Thurston, 2016), bem como na geração de recursos potenciais para empresas privadas e empreendedores (KLARELD, 2015).

Em todos os casos, nos desafios e nas oportunidades, a revisão sistemática evidenciou o papel fundamental do Arquivista nesse processo e a necessidade de que esse profissional esteja capacitado para as demandas de gerenciamento de registros arquivísticos em ambiente digital.

Essa adequação na formação do profissional Arquivista, embora seja um estado de constante necessidade de aprimoramento, já vem sendo analisada desde a década de 1980 nos Estados Unidos, quando a gestão de registros eletrônicos já passa a ser reconhecida como parte integrante da educação arquivística norte-americana (FORCE; ZHANG, 2016).

Baron e Thurston (2016) descrevem a importância da atualização do Arquivista para a mudança de cenário e o aumento no volume e no formato das informações a serem processadas. Conway (2015) atribui ao Arquivista a responsabilidade por interferir na criação de produtos de informação a partir do acervo, registrar todos os eventos capazes de alterar a interpretação e o significado dos registros digitais, disponibilizar os metadados relacionados à proveniência digital dos registros e aprimorar a compreensão acerca das normas de confiabilidade, segurança e preservação de registros e repositórios digitais.

Para Guo (2016), as mudanças tecnológicas e sociais confirmam ainda mais importância do papel dos Arquivistas na garantia da confiabilidade dos registros, sendo necessário um reexame da base de conhecimento da Arquivologia que incorpore as mudanças tecnológicas e suas implicações na criação, manutenção e acesso aos registros digitais, incluindo todo seu ciclo de vida.

Além das questões de ordem tecnológica, o tratamento dos registros digitais, no que se refere à privacidade das informações neles contidos, também é abordado. A digitalização tem ressignificado o papel e o uso dos arquivos (HENTTONEM, 2017). Do mesmo modo, a privacidade tem sido motivo de preocupação para Arquivistas, dado o problema que a exposição de dados sigilosos pode causar.

Henttonem (2017) afirma que a gestão de registros e arquivos está estreitamente interligada com as questões de privacidade. De acordo com a autora, o profissional de arquivos que, no paradigma tradicional, era o mediador entre a informação e o estado, historiadores e a sociedade em geral; agora, no paradigma global, em que a tecnologia da informação por meio de seus profissionais fornece serviços de maneira global a todos os usuários da internet, precisa refletir acerca das implicações que as questões de privacidade no ambiente digital trazem para a sua atuação.

Embora a privacidade digital seja amplamente discutida pela literatura, para Henttonem (2017) são poucos os estudos que a abordam sob a perspectiva arquivística. Uma perspectiva trazida por Henttonem (2017) para a discussão da gestão dos registros para arquivos é a necessidade de harmonização entre as estratégias da sociedade para proteção de suas informações e a legislação que trata da transferência de registros para os arquivos. De acordo com a autora, a sociedade precisa estabelecer até que ponto é importante restringir acesso a determinadas informações por serem consideradas privadas. E, na sua visão, uma alternativa é a reflexão acerca do desenvolvimento de estratégias para controle do uso da informação como ferramenta mais eficiente do que o controle do acesso.

Baseado nas discussões de Henttonem (2017), é possível refletir que um dos fatores a ser considerado pelos profissionais responsáveis pela gestão de registros e arquivos é a potencialidade dos registros serem reidentificados e reutilizados em combinação com outros registros para a compreensão de fenômenos diversos que representam.

Esta revisão buscou contribuir na compreensão dessas novas demandas relacionadas à gestão de registros digitais arquivísticos. Pode-se observar que os registros digitais arquivísticos se configuram como um desafio, visto que têm características e requisitos próprios e, portanto, não contemplados em sua totalidade pelas teorias e princípios tradicionalmente utilizados pela Arquivologia. Do mesmo modo, o Arquivista, responsável pela gestão de documentos e registros, mas também pela custódia, preservação e acesso à informação que estes contêm, também se vê diante de novas demandas no que se refere à atualização dos conceitos e práticas da sua atuação profissional.

As considerações acerca dos resultados desta RSL serão descritos na subseção a seguir.

3.1.3 Considerações sobre os registros digitais arquivísticos

Esta revisão permitiu identificar que a gestão de registros digitais arquivísticos é um tema repleto de termos e possibilidades de investigação. As abordagens presentes nos artigos analisados ofereceram pistas para a

compreensão dos principais conceitos relacionados, seus desafios enquanto área de estudos e enquanto campo de atuação profissional.

A RSL elaborada ajudou a compreender o que são registros, quais as características dos registros digitais, como eles se relacionam ao contexto arquivístico e quais as referências para sua gestão. Embora apenas um dos autores tenha elaborado uma contribuição teórica mais consistente para o termo objeto desta análise, todos os demais ofereceram ricas respostas acerca dos elementos envolvidos, áreas com as quais existe relação e tarefas que impactam os aspectos teóricos e pragmáticos da questão.

Tamanha amplitude de conceitos e temas envolvidos na gestão dos registros digitais arquivísticos, traz consigo uma série de dificuldades e desafios detalhadamente descritos pelos autores estudados. Dentre as soluções, destaca-se o caso do arquivo municipal chinês, exemplificado por Guo (2016), que ao armazenar os registros digitais de empresas privadas, funciona como terceiro confiável em casos de necessidade de verificação da autenticidade, confiabilidade e integridade de tais registros.

Considerando as atribuições estabelecidas pela legislação brasileira ao Arquivo Nacional, entendemos ser esse órgão um propenso candidato para implementação de solução semelhante no país, servindo de repositório para os registros digitais arquivísticos dos órgãos do poder executivo federal, de modo que os requisitos de segurança possam ser preservados e utilizados quando necessário.

Entende-se que, enquanto não há um consenso acerca do que são efetivamente os registros digitais arquivísticos, quais as suas diferenças em relação aos documentos arquivísticos digitais, faz-se necessária a adoção de medidas por parte do poder público quanto à necessidade de se preservar o patrimônio documental digital brasileiro, independente da terminologia que desejemos adotar.

As diversas possibilidades de informação registradas que o governo produz (*sites*, aplicativos, *e-mails*, dados gerados por sistemas de negócios dos diferentes órgãos, áudios, vídeos e tantas outras formas de documentar a atividade governamental nos ambientes digitais) precisam ser reguladas antes da sua criação, e durante seu uso, para que se possa, efetivamente, realizar a sua gestão.

Quais são as diretrizes para a criação de informações registradas nas diversas plataformas digitais? Que metadados serão obrigatórios? Como elas devem ser classificadas? Quem irá realizar e como será a avaliação da sua temporalidade? Que critérios de armazenamento devem ser utilizados nas fases de gestão? Qual a política de preservação do material digital de valor permanente? Como garantir a interoperabilidade? Esses são alguns dos questionamentos que temos e podem servir para futuras pesquisas.

Alertamos para a proposta de Guo (2016) de que para ser considerado um registro digital, seu conteúdo seja convertido em um documento com forma fixa e conteúdo estável, pois tal argumentação vai na contramão da evolução tecnológica. Os registros digitais sempre irão possuir e assimilar as características que as tecnologias disponibilizarem. Portanto, é necessário que os profissionais arquivistas entendam tais características e busquem adequar a guarda documental às constantes mudanças, indicando e promovendo as eventuais adaptações necessárias nos sistemas de informações das organizações.

O contexto dos registros digitais requer que a prática profissional do arquivista seja repensada e se adapte ao processo de disrupção tecnológica que estamos vivendo e não o contrário. A informação é cada vez mais importante, o acesso aos arquivos cada vez mais necessário, mas as formas de organização e recuperação dessa informação precisam ser tão dinâmicas quanto as suas novas possibilidades de criação e uso.

Esta revisão também permitiu conhecer as soluções que a gestão dos registros digitais pode oferecer, as implicações e reflexões necessárias sobre as questões de privacidade dos registros, sobretudo os públicos.

Ficou evidenciado que os metadados são elementos fundamentais para que a gestão dos registros digitais aconteça, que o Arquivista precisa estar capacitado para lidar com as novas demandas impostas pelos recursos tecnológicos, dentre elas, a reidentificação, ressignificação e reutilização da informação.

3.2 REGISTROS DIGITAIS ARQUIVÍSTICOS: CONCEITOS, MODELOS DE CATEGORIZAÇÃO E PERSPECTIVAS

Nesta seção, serão apresentados os principais conceitos inerentes aos registros digitais arquivísticos, seus modelos de categorização e as perspectivas em relação aos registros digitais arquivísticos encontrados por esta pesquisa.

3.2.1 Conceitos relacionados aos registros digitais arquivísticos: um olhar para a literatura nacional

A revisão de literatura sobre a gestão de registros digitais para arquivos, realizada nas bases de periódicos científicos, possibilitou um reconhecimento da amplitude do tema. Contudo, considerando a baixa publicação de trabalhos em língua portuguesa nas bases consultadas, e que, após a aplicação dos critérios de seleção, nenhum trabalho de autores nacionais entrou no corpo da revisão, entendemos ser necessária a verificação de outras modalidades de comunicação científica sobre a questão, em que também autores brasileiros escrevem sobre o assunto, considerando o contexto nacional.

Nesse sentido, inicialmente encontramos um trabalho que traz uma contribuição significativa para essa temática. Trata-se da tese de doutoramento intitulada “O conceito de documento arquivístico frente a realidade digital: uma revisitação necessária”, defendida em 2011, junto ao Programa de Pós-Graduação em Ciência da Informação da Universidade Federal Fluminense (UFF), em parceria com o Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT).

Na referida tese, a autora Rosely Curi Rondinelli apresenta os conceitos de documento, documento arquivístico e documento arquivístico digital. Ela traz, ainda, uma proposta de categorização dos documentos digitais e indica as características e partes constituintes, segundo a Diplomática, dos documentos arquivísticos digitais. Passemos à apresentação dos conceitos encontrados nesse trabalho.

A primeira definição importante é a de documento. Ao analisar os estudos de Buckland (1991) sobre o paralelo da informação-como-conhecimento e informação-como-coisa, Rondinelli (2011, p. 38) estabelece que “[...] dados, textos, filmes, fotografias seriam o conhecimento representado de forma tangível, ou seja, seriam documentos”.

Tão amplo em possibilidades semânticas quanto o termo “registros digitais arquivísticos” é o trinômio “dado, informação e conhecimento” na CI. Para Meadow e Yuan,

[...] dado geralmente significa um conjunto de símbolos com pouco ou nenhum significado para o receptor. Informação é um conjunto de símbolos que têm um sentido ou significado para seu receptor. Conhecimento é a acumulação e integração da informação recebida e processada pelo receptor (MEADOW; YUAN, 1997, p. 701).

Sobre os dados, Rondinelli (2011, p. 39) explica que “num sentido lato, estes são entendidos como informação processada para uso. Já no sentido estrito, constituem-se em registros armazenados em computador”, enquanto que textos e documentos “seriam não somente documentos textuais e em papel, mas, também imagens e sons, tanto em meio convencional como eletrônico” (RONDINELLI, 2011, p. 39).

O conceito de documento é apresentado como

[...] evidência produzida num suporte (papel, fita magnética, disco, placa etc.) por meio de um instrumento de escrita (caneta, lápis, máquina de escrever, etc.) ou de um aparato para fixação de dados, imagens, e/ou vozes (DURANTI, 1998, p. 41).

Sobre o fato de o documento precisar ser “escrito”, esclarece que

O atributo “escrito” não é usado em diplomática como um ato isolado (desenhado, entalhado, tracejado, ou inscrito) mas sim no sentido de referir-se ao propósito e ao resultado intelectual do ato de escrever; isto é, à expressão de ideias numa forma concreta (documental) e sintática (regida por regras de disposição) (DURANTI, 1998, p. 41).

Convém registrar que, posteriormente, Duranti (2002, p. 11) conceitua documento apenas como “informação registrada”.

Rondinelli (2011) propõe convergências entre a Arquivologia e a Ciência da Informação quando da acepção do documento. O reconhecimento de que o documento possui materialidade ou fisicalidade é uma dessas possibilidades, pois

Um outro aspecto a ser considerado diz respeito à fisicalidade, ou materialidade do documento. Há que registrar que o tema, embora permeie os discursos dos autores da Ciência da Informação aqui estudados, é bem mais recorrente e mais explicitamente abordado na área arquivística. Ali a ideia de um conteúdo fixado num suporte aparece na maior parte dos conceitos de documento apresentados, o que, há que ressaltar, não significa o atrelamento da entidade documental a um determinado tipo de suporte mas sim à necessidade de estabilidade e permanência desse mesmo conteúdo (RONDINELLI, 2011, p. 55-56).

Convém pontuarmos a diferença entre arquivo administrativo e arquivo permanente, e as implicações terminológicas de adotar um ou outro como local de guarda de um documento. Tais termos se referem à aplicação da teoria das 3 idades ou teoria do ciclo vital dos documentos, construto importante para a ciência arquivística. Tal cenário “[...] oriunda dos estudos desenvolvidos nos Estados Unidos sobre gestão de documentos a partir dos anos de 1940, os documentos arquivísticos passam por três idades ou fases: corrente, intermediária e permanente” (ROUSSEAU; COUTURE, 1998).

O ciclo vital dos documentos é definido pelo DBTA como as “sucessivas fases por que passam os documentos de um arquivo, da sua produção à guarda permanente ou eliminação (ARQUIVO NACIONAL, 2005, p. 47). Enquanto isso, a teoria das 3 idades é definida pelo mesmo dicionário como a teoria segundo a qual “os arquivos são considerados arquivos correntes, intermediários ou permanentes, de acordo com a frequência de uso por suas entidades produtoras e a identificação de seus valores primário e secundário” (ARQUIVO NACIONAL, 2005, p.160).

Ou seja, a fase corrente corresponde aos documentos considerados administrativos e em fase de tramitação; a fase intermediária corresponde aos documentos tramitados com menor frequência e que aguardam destinação

(seja eliminação ou recolhimento). Já a fase permanente é o momento em que os documentos já foram avaliados como portadores de valor secundário ou histórico e, por esse motivo, serão preservados em caráter permanente.

Essa teoria é a base para uma discussão terminológica importante para essa pesquisa e para a Arquivologia quanto aos termos documento e registro, que está presente em diferentes idiomas.

O termo *records* é tratado de diferentes formas na literatura, a depender do país de origem da publicação e da vertente conceitual do autor. Para Theodore Schellemborg (1956), o conceito de *records* corresponde a

Todos os livros, papéis, fotografias, ou outros materiais documentais, independentemente da forma física ou de características, elaborados ou recebidos por qualquer instituição pública ou privada no exercício de suas obrigações legais ou em conexão com a transação de seu próprio negócio e preservados ou destinados à preservação por aquela instituição ou seu legítimo sucessor como evidência de suas funções, políticas, decisões, procedimentos, operações ou outras atividades, ou por causa do valor informacional dos dados ali contidos (SCHELLENBERG, 1956, p. 16).

Rondinelli (2011) apresenta ainda que,

[...] no contexto dos países de língua inglesa, registramos que a diferença entre os termos arquivo (*archives*) e documento arquivístico (*record*) não se restringe a considerações de ordem coletiva ou individual, conforme registrado anteriormente em relação aos países de língua latina, mas também ao fato do termo arquivo (*archives*) remeter sempre à preservação permanente dos documentos que o integram. Em outras palavras, na língua inglesa, o termo arquivo (*archives*) estaria para os documentos de valor permanente assim como o termo documento arquivístico (*record*) para os de valor corrente e intermediário (RONDINELLI, 2011, p. 201).

Antes de avançar para o próximo ponto, convém mencionar que existe uma corrente de pensamento que não corrobora com a teoria das 3 idades dos documentos. Trata-se da vertente australiana da Arquivologia que defende a corrente do *records continuum*, ou documentos/registros contínuos, que

[...] desafiam compreensões tradicionais que diferenciam 'archives' de 'records' com base na seleção para preservação permanente em custódia arquivística, e que focam na natureza

fixa dos documentos arquivísticos. Adotando uma visão pluralista de informação registrada, o pensamento *continuum* caracteriza os documentos arquivísticos como um gênero especial de documentos no que se refere à sua intenção e funcionalidade. Enfatiza sua natureza evidencial, transacional e contextual, rejeitando abordagens que focam no seu conteúdo e valor informacional (MCKEMMISH, 2001, p. 335).

É justamente no bojo do *records continuum* que existem importantes reflexões sobre a relação dos registros e documentos com os metadados. Sobre isso, Mckemmish (2001) diz que

[...] as ideias pós-modernas sobre documentos arquivísticos os vêem como objetos dinâmicos que são fixos em termos de conteúdo e elementos significativos da sua estrutura, mas vinculados a camadas cada vez mais amplas de metadados contextuais que gerenciam seus significados e permitem sua acessibilidade e usabilidade na medida em que eles se movem no tempo e no espaço (MCKEMMISH, 2001, p. 349).

Isto posto, iniciemos então o estudo do segundo conceito escolhido: documento arquivístico. Para os fins desta pesquisa, consideremos inicialmente a definição de documento arquivístico proposta por Yeo (2007). Para o autor,

Pode-se argumentar que definições que enfatizam apenas evidência ou informação são limitantes [...] e minimizam a complexidade dos documentos arquivísticos. Ênfase em evidência visa sempre vincular procedimentos de gestão de documentos ao mundo da lei e da governança corporativa; ênfase em informação sugere um alinhamento com biblioteconomia ou computação. Foco na memória talvez implique uma associação com a história ou identidade cultural. Todas essas perspectivas são válidas mas nenhuma é abrangente. A visão representacional de documentos arquivísticos é multidisciplinar e abrange uma ampla gama de interpretações (YEO, 2007, p. 343).

Em consonância com as intenções de pesquisa deste trabalho – os registros digitais arquivísticos, o conceito de Yeo (2007) permite que os produtores encontrem lugar também nas máquinas, quando estabelece que documentos arquivísticos podem ser: “[...] produzidos por pessoas ou equipamentos que participaram ou observaram a atividade representada ou por pessoas autorizadas a agir como seus representantes” (YEO, 2007, p. 338).

Com base em Yeo (2008), Rondinelli (2011) complementa:

Dentro dessa linha inovadora, o autor propõe ainda um questionamento sobre o que ele chama de documentos arquivísticos protótipos e documentos arquivísticos limítrofes. Entre os primeiros estariam os documentos escritos e com uma forma documental predeterminada, uma ata, por exemplo. Já entre os limítrofes estariam documentos com uma forma documental não convencional e passível de diferentes percepções como entidade documental por diferentes comunidades. Por exemplo, um *website* pode ser entendido como um documento arquivístico mas também como uma ferramenta de gestão corporativa ou mesmo uma manifestação de cultura contemporânea (RONDINELLI, 2011, p. 192-193).

Sem intenção de esgotar essa reflexão por aqui, nos indagamos: estariam os registros digitais arquivísticos contemplados na definição dos documentos arquivísticos limítrofes ou estes não conseguem abarcar a realidade dos metadados, defendida por Rogers (2015), como critério fundamental para caracterização dos registros digitais dos arquivos?

Para Duranti (2002, p. 11), um documento arquivístico é “[...] todo documento criado por uma pessoa física ou jurídica no decorrer de atividades práticas como instrumento ou subproduto dessas atividades [...]”.

De forma mais detalhada, Rondinelli (2011) fez a seguinte definição:

Os documentos arquivísticos surgem naturalmente no decorrer das atividades de uma pessoa física ou jurídica, como uma decorrência normal do ato de se registrar essas atividades. Desse registro espontâneo resulta uma **promessa** de imparcialidade dos documentos bem como um vínculo entre eles. Tal vínculo se traduz na relação orgânica que cada documento tem um com o outro e no fato do conjunto documental daí resultante ser dotado de organicidade. Como se trata de documentar atividades, cada documento é único no seu conjunto na medida em que ali desempenha uma função específica. Finalmente, a submissão desses documentos a procedimentos arquivísticos ininterruptos de gestão e de preservação, garante a sua autenticidade (RONDINELLI, 2011, p. 220, grifo nosso).

Com essa definição, a autora traz elementos considerados preponderantes para a constituição de um documento arquivístico, quais sejam a organicidade, a imparcialidade, a unicidade e a autenticidade.

A relação entre o conhecimento arquivístico e a diplomática também está presente em outros trabalhos. Para Tognoli *et al* (2019), embora atualmente a Arquivologia esteja dialogando mais proficuamente com as disciplinas que tem a informação e o conhecimento como objeto de estudo, durante anos esteve ligada às searas de história, administrativa e diplomática. Mas quando se estabelece essa associação entre Arquivologia e Diplomática, onde inicia e onde finaliza o escopo de cada área?

[...] a Diplomática olha os documentos arquivísticos como entidades individuais enquanto a Arquivologia os vê como agregações (DURANTI, MACNEIL, 1996). É esse componente orgânico dos arquivos que a Arquivologia oferece à Diplomática quando ambas se reassociam a partir do século XX. Nessa reassociação dá-se o complemento perfeito no qual a Diplomática cuida da gênese, dos elementos de forma e do status de transmissão (minuta, original e cópia) do documento e a Arquivologia trata de classificá-lo, temporalizá-lo, descrevê-lo e preservá-lo (RONDINELLI, 2011, p. 139).

Finalmente, passemos ao estudo do conceito de documento arquivístico digital. Rondinelli (2011) inicia o capítulo dedicado a esse conceito em sua tese com a seguinte constatação:

[...] o documento arquivístico pode se apresentar em qualquer suporte. Por essas assertivas inferimos que documentos digitais gerados no curso de atividades desempenhadas por pessoas físicas e jurídicas e em suportes tão diferentes como os magnéticos e óticos, também podem se constituir em documentos arquivísticos (RONDINELLI, 2011, p. 223).

De acordo com Rondinelli (2011), os documentos arquivísticos digitais, assim como os documentos arquivísticos não digitais, precisam apresentar as seguintes características:

[...] forma fixa, conteúdo estável, relação orgânica, contexto identificável, ação e o envolvimento de cinco pessoas: autor, redator, destinatário, originador e produtor. Há que ressaltar que dentre essas cinco pessoas, pelo menos as três primeiras têm que estar presentes num documento arquivístico (MACNEIL e colaboradores, 2005; DURANTI; THIBODEAU, 2008; DURANTI, 2009b; DURANTI, 2010 *apud* RONDINELLI, 2011, p. 227).

Quanto à “forma fixa e ao conteúdo estável, significam que o documento arquivístico digital tem que manter a mesma apresentação que tinha quando ‘salvo’ pela primeira vez” (RONDINELLI, 2011, p. 227).

No que se refere à relação orgânica,

[...] trata-se de mais uma característica eminentemente arquivística e que se encontra implícita no conceito de documento arquivístico, na medida em que, de acordo com esse conceito, os documentos se constituem em registros de atividades e, conseqüentemente, mantém um vínculo inextricável entre si. No caso do documento arquivístico digital, essa vinculação se dá entre documentos dentro e fora do sistema, isto é, nos chamados ambientes híbridos os quais se caracterizam por abranger documentos digitais e não digitais ao mesmo tempo (DURANTI; THIBODEAU, 2008 *apud* RONDINELLI, 2011, p. 228).

Ainda em relação aos elementos intrínsecos, considerando o prejuízo que a interpretação equivocada pode trazer para a compreensão do assunto, convém apresentar as seguintes definições:

Quadro 4 – Elementos intrínsecos do documento arquivístico

Termo	Definição
Autor	“pessoa física ou jurídica que tem autoridade e competência para emitir o documento arquivístico ou em cujo nome ou sob cujo comando o documento foi emitido” (DURANTI; PRESTON, 2008, p. 801);
Redator	“pessoa que tem autoridade e competência para articular o conteúdo do documento arquivístico” (DURANTI; PRESTON, 2008, p. 843);
Destinatário	“pessoa para quem o documento arquivístico é direcionado ou para quem se destina” (DURANTI; PRESTON, 2008, p. 796);
Originador	“pessoa designada no endereço eletrônico no qual o documento arquivístico foi gerado (isto é, do qual é enviado ou onde é compilado ou mantido)” (DURANTI; PRESTON, 2008, p. 827);
Produtor	“pessoa a cujo fundo ou arquivo o documento pertence” (DURANTI, 2010).
Data cronológica	“data (e, possivelmente a hora) de um documento arquivístico, incluída no documento por seu autor, ou pelo sistema eletrônico em nome do autor, no decorrer da sua elaboração” (DURANTI; PRESTON, 2008, p. 804);
Data tópica	“o lugar da elaboração de um documento arquivístico, incluído no documento por seu autor” (DURANTI; PRESTON, 2008, p. 841);
Indicação e descrição da ação ou assunto	“a identificação do assunto (linha de assunto) e o teor propriamente dito do documento” (DURANTI, 2005, p. 195);
Atestação	“[...] validação escrita de um documento arquivístico [...] por parte daqueles que participam da sua emissão (autor, redator, autenticador) bem como por testemunhas da ação ou da assinatura do documento” (DURANTI; PRESTON, 2008, p. 800-801). “Assim, em mensagens de correio eletrônico (emails), por exemplo, o nome do remetente que aparece no alto da mensagem, se constitui em uma atestação” (RONDINELLI, 2011, p. 230).

Fonte: Rondinelli, 2011, p. 229-230.

Anotações definem-se por “Acréscimos feitos ao documento arquivístico após a sua produção” (DURANTI; THIBODEAU, 2008, p. 409, nota 13) e podem ser:

1. acréscimos feitos ao documento arquivístico no âmbito da sua transmissão como, por exemplo, indicação de prioridade (urgente), data da elaboração e da transmissão, indicação de anexos;
2. acréscimos feitos no curso das atividades das quais o documento arquivístico participa como, por exemplo, data e hora do recebimento, providências tomadas e
3. acréscimos próprios da gestão arquivística como código de classificação, data do arquivamento (RONDINELLI, 2011, p. 231).

Já o contexto é conceituado como “ambiente em que ocorre a ação registrada no documento” (CÂMARA TÉCNICA DE DOCUMENTOS ELETRÔNICOS, 2011, p. 126), que pode ser identificado de acordo com os tipos abaixo:

Contexto jurídico-administrativo: refere-se a leis e normas externas à instituição produtora de documentos as quais controlam a condução das atividades dessa mesma instituição;

Contexto de proveniência: refere-se a organogramas, regimentos e regulamentos internos que identificam a instituição produtora de documentos;

Contexto de procedimentos: refere-se a normas internas que regulam a criação, tramitação, uso e arquivamento dos documentos da instituição;

Contexto documental: refere-se a código de classificação, guias, índices e outros instrumentos que situam o documento dentro do conjunto a que pertence, ou seja, ao fundo e

Contexto tecnológico: refere-se ao ambiente tecnológico (hardware, software e padrões) que envolve o documento. (CÂMARA TÉCNICA DE DOCUMENTOS ELETRÔNICOS, 2011, p.126-127).

O suporte é definido pela Câmara Técnica de Documentos Eletrônicos (2011, p. 132) como a “base física sobre a qual a informação é registrada”, e sobre os componentes digitais, para estes, são necessários alguns esclarecimentos de forma preliminar. Primeiramente, Rondinelli (2011, p. 233) explica que existem diferenças substanciais entre um documento digital quando este encontra-se exibido numa tela de computador (ou qualquer dispositivo) e quando ele é “salvo” e o “fechamos”, portanto, é armazenado em algum suporte. Segundo a autora, nesse momento o documento “deixa de ser compreensível aos olhos humanos para se transformar numa cadeia de bits”. Ao ser chamado à tela novamente, “um mecanismo é acionado no qual a cadeia de bits é processada por um software que a transforma no documento passível de leitura e compreensão”. Seguindo a explicação, a autora pontua que, por esse motivo,

[...] não é possível manter um documento digital exatamente como o vemos, a não ser que o imprimamos e, nesse caso, ele deixaria de ser digital. O que se pode manter é a capacidade de reproduzir o documento sempre que necessário (RONDINELLI, 2011, p. 234).

Por isso, um conceito de grande importância para o entendimento dos documentos digitais, e por conseguinte, dos registros digitais arquivísticos, é explicado pela autora: o documento digital manifestado.

Trata-se, pois, de uma realidade dinâmica inerente à própria natureza do documento digital pela qual este possui uma forma manifestada, isto é, aquela que é visível ao usuário, à qual corresponde uma codificação digital. Tal codificação se traduz em cadeias de bits inscritas num suporte magnético ou ótico. As cadeias de bits se constituem nos chamados componentes digitais os quais compreendem dados que determinam a forma do documento e os que definem seu conteúdo. Há ainda um terceiro grupo de dados que são aqueles que regulam o processamento dos dois primeiros, quando necessário, bem como os metadados relacionados ao documento. São os chamados dados de composição (RONDINELLI, 2011, p. 233).

Isto é, de acordo com Rondinelli (2011, p. 233), “um componente digital pode conter dados de forma e/ou de conteúdo e/ou de composição [...] e os metadados”. Os dados de forma relacionam-se à aparência do documento na tela em que for visualizado; os dados de conteúdo dizem respeito ao teor do documento; e os dados de composição são aqueles que contêm a forma como o documento é constituído (o *layout* de um *site*, por exemplo). Já os metadados, estes podem ter possibilidades das mais diversas (formato, direitos autorais, etc).

Em que pese todas as questões aqui apresentadas, Rondinelli (2011, p. 237) diz que “um documento não tem que conter todos os elementos da forma documental para ser arquivístico”. Ou seja, o documento arquivístico pode ser “incompleto no tocante a seus elementos intrínsecos e extrínsecos”, contanto que se adeque ao conceito de documento arquivístico.

Nesse sentido, e com base no trabalho de Duranti e Thibodeau (2008, p. 435-437), a autora apresenta uma proposta de classificação dos documentos digitais (arquivísticos ou não), em estáticos e interativos, conforme o quadro 5.

Quadro 5 – Classificação dos documentos digitais

Termo	Definição
Documento digital estático	Não permite alteração na forma e no conteúdo além das determinadas pela tecnologia como abrir, fechar, diminuir, aumentar etc. Exemplo: documentos com equivalentes em papel: carta em <i>Word</i> ; recibo de compra <i>on line</i> .
Documento digital interativo	Permite alteração de forma e/ou conteúdo por meio de regras fixas ou variáveis.

Fonte: Rondinelli, 2011, p. 241.

Já os documentos digitais interativos, subdividem-se em:

Quadro 6 – Classificação dos documentos digitais interativos

Termo	Definição
Documento digital interativo (não dinâmico)	As regras que gerenciam forma e conteúdo são fixas e o conteúdo é selecionado a partir de dados armazenados no sistema. Exemplo: catálogos de vendas <i>on line</i> .
Documento digital interativo (dinâmico)	As regras que gerenciam forma e conteúdo podem variar. Exemplo: serviços de previsão do tempo e de cotação de moedas cujos conteúdos estão sempre mudando (sistemas de informação e não documentos propriamente).

Fonte: Rondinelli, 2011, p. 241.

São apresentados, ainda, os documentos experienciais, que, embora não se tratem de uma categoria à parte, foram definidos de forma separada em virtude da “alta complexidade tecnológica” que envolvem.

O principal objetivo em classificar um documento como experiencial é enfatizar que este documento é claramente destinado a, não simplesmente, e, talvez, de forma alguma, comunicar uma informação específica e sim proporcionar uma experiência subjetiva (DURANTI; THIBODEAU, 2008, p. 431).

Sobre a forma fixa e conteúdo estável, no caso dos documentos arquivísticos digitais, Rondinelli (2011) conclui que os mesmos adquirem uma dinamicidade própria e coerente ao ambiente digital, e acrescenta:

Por esse dinamismo, uma certa variação na forma e no conteúdo do documento não compromete sua capacidade documental. Trata-se, na verdade, de uma nova concepção das características da forma fixa e do conteúdo estável, a qual

resulta no conceito de variabilidade limitada (RONDINELLI, 2011, p. 241).

Rondinelli (2011) coloca que, dentre as opções de interativo e estático, este – o documento estático – seria o documento digital que mais se adequaria ao status arquivístico, por ser detentor de forma fixa e conteúdo estável.

A função essencial de um documento arquivístico é servir como uma ponte no tempo, carregar informação sobre uma ação, um evento ou estado de negócios para quando for necessário em ações subsequentes ou como referência sobre o que aconteceu ou o que foi descrito ou dito no passado (DURANTI; THIBODEAU, 2008, p. 438).

Considerando Duranti e Thibodeau (2008), Rondinelli (2011) questiona como os documentos digitais interativos, que possuem variabilidade de forma e de conteúdo, poderiam ser essa “ponte no tempo”, já que “um documento arquivístico que não contenha uma mensagem fixa [...], não pode ser chamado de volta e não pode servir como um meio de lembrança” (DURANTI; THIBODEAU, 2008, p. 438).

Para responder a esse questionamento sobre como os documentos digitais interativos podem se constituir como arquivísticos é que, segundo a autora, se aplica o conceito de variabilidade limitada, que são “[...] mudanças na forma e/ou conteúdo de um documento arquivístico digital que são limitadas e controladas por meio de regras fixas, de maneira que a mesma consulta, pedido ou interações sempre geram o mesmo resultado” (DURANTI; PRESTON, 2008, p. 803).

Dito isso, Rondinelli (2011) conclui que documentos digitais que atendam aos critérios que atribuem a natureza arquivística a um documento, e estejam inseridos no conceito de variabilidade limitada, são considerados documentos arquivísticos digitais.

Porém, em se tratando do ambiente digital, questiona: seria esse documento tão somente aquele que aparece na tela, ou seja, o documento manifestado? E continua:

A questão assume maior complexidade quando se trata de documentos arquivísticos inseridos em sistemas interativos não dinâmicos. Isto porque em alguns casos os componentes

digitais integram um documento arquivístico armazenado, isto é, um documento digital com as seis características próprias de um documento arquivístico que é invisível ao usuário. Neste contexto tem-se duas entidades arquivísticas distintas mas relacionadas, quais sejam o documento arquivístico manifestado e o documento arquivístico armazenado (RONDINELLI, 2011, p. 243).

Rondinelli (2011, p. 244) explica que “para além do documento arquivístico manifestado há sempre um ou mais componentes digitais bem como, em alguns casos, um documento arquivístico armazenado”. Para exemplificar, ela cita o sistema de consulta *online* ao contracheque dos servidores do poder executivo federal. Há um banco de dados constituído de diferentes tabelas, cuja consulta utilizando diferentes possibilidades gera um contracheque, que é apresentado na tela. Na visão da autora, o contracheque se constitui num documento arquivístico manifestado e o banco de dados num documento arquivístico armazenado e, aos dois, correspondem vários componentes digitais, alguns comuns e outros específicos a cada um deles.

A autora estabelece, portanto, três conceitos, para a atribuição de credibilidade ao documento arquivístico no ambiente digital. São eles: acurácia, confiabilidade e autenticidade. A acurácia é definida como o “grau ao qual dados, informações, documentos e documentos arquivísticos são precisos, corretos, verdadeiros, livres de erros ou distorções, ou pertinentes ao assunto” (DURANTI; PRESTON, 2008, p. 796). Já a confiabilidade, se trata da

Credibilidade de um documento arquivístico enquanto uma afirmação do fato. Existe quando um documento arquivístico pode sustentar o fato ao qual se refere, e é estabelecida pelo exame da completeza da forma do documento e do grau de controle exercido no processo de sua criação (CÂMARA TÉCNICA DE DOCUMENTOS ELETRÔNICOS, 2011, p. 124).

A autenticidade, por sua vez, é entendida como a “Credibilidade de um documento arquivístico enquanto documento arquivístico, isto é, a qualidade de um documento arquivístico ser o que diz ser e que está livre de adulteração ou qualquer outro tipo de corrupção” (CÂMARA TÉCNICA DE DOCUMENTOS ELETRÔNICOS, 2011, p. 126).

Em síntese, a tese de Rondinelli (2011) conceitua registro, documento, dado, informação, documento arquivístico, arquivo administrativo, arquivo

histórico, ciclo vital, documentos contínuos, metadados, diplomática, documento arquivístico digital, além de estabelecer relações entre a Arquivologia e a Ciência da Informação.

De forma estruturada, Rondinelli (2011) apresenta ainda as características do documento arquivístico digital (forma fixa, conteúdo estável, relação orgânica, contexto identificável e pessoas envolvidas) e os elementos que o constituem (forma documental, anotações, contexto, suporte, atributos e componentes digitais).

Sobre os documentos digitais, a autora traz uma divisão entre estáticos e interativos, e destes, em dinâmicos e não dinâmicos, mencionando os documentos experienciais e os documentos arquivísticos em potencial. São apresentados, também, os conceitos de documento arquivístico digital manifestado e armazenado e propõem-se categorias funcionais para os documentos arquivísticos (digitais e não digitais), quais sejam, documentos prospectivos e retrospectivos. Por fim, a autora apresenta os elementos que atribuem credibilidade a um documento arquivístico digital (acurácia, confiabilidade e autenticidade).

Dada a multiplicidade de elementos considerados importantes para esta pesquisa encontrados no trabalho de Rondinelli (2011), dedicamos o item 3.2.1 para a apresentação dos conceitos por ela trabalhados.

Trataremos, no item 3.2.3, a discussão sobre as contribuições encontradas nos demais trabalhos analisados por esta pesquisa.

3.2.2 Modelos de categorização dos registros digitais

Os registros digitais arquivísticos, cada vez mais presentes no cotidiano das pessoas e das organizações, sejam como produtores ou recebedores de tais documentos, necessitam ser compreendidos enquanto entes dotados de características próprias.

Buscou-se conhecer as suas possibilidades de categorização, com vistas à ampliação do entendimento sobre seus elementos constitutivos, com a limitação das abordagens presentes nos autores selecionados por esta pesquisa. Dessa forma, foram localizados cinco modelos de classificação, que serão apresentados a seguir.

O estudo da pesquisadora brasileira Rondinelli (2011) divide os documentos digitais de quatro maneiras: (1) **documento digital estático**, aquele que não permite alteração na forma e no conteúdo além das determinadas pela tecnologia como abrir, fechar, diminuir, aumentar; (2) **documento digital interativo**, permite alteração de forma e/ou conteúdo por meio de regras fixas ou variáveis; podendo este ser um (3) **documento digital interativo não dinâmico**, em que as regras que gerenciam forma e conteúdo são fixas e o conteúdo é selecionado a partir de dados armazenados no sistema, como gráficos em *Excel* e catálogos de vendas *online*; ou ainda um (4) **documento digital interativo dinâmico**, em que as regras que gerenciam forma e conteúdo podem variar, como nos serviços de previsão do tempo e de cotação de moedas, cujos conteúdos estão sempre mudando.

Já o Conselho Internacional de Arquivos (2015), no documento destinado a orientar os arquivistas sobre as formas de gerenciamento dos registros digitais, embora não estabeleça exatamente categorias, indica a divisão de tais registros de acordo com dois critérios:

- 1 - **Pela função**, isto é, a relação entre dossiês e diferentes tipos de atividades e transações. Podemos apontar como exemplos os processos individuais, os processos de contencioso, os processos relativos a uma atividade específica, as coleções de correspondência, os sítios Web, etc.; e/ou
- 2 - **Pela forma e formato**. Exemplos incluem documentos resultantes de processamento de texto, bases de dados, documentos hipertextuais, imagens, folhas de cálculo, mensagens de correio electrónico, áudio, vídeo, etc.

Para o autor africano Katuu (2016), os registros digitais arquivísticos podem ser relacionados em 4 categorias: (1) **documentos criados usando aplicativos de escritório**, como documentos de texto, planilhas e apresentações; (2) **registros em ambientes *on-line*** ou baseados na *Web*, como intranet, extranets e registros de transações *online*; (3) **registros gerados por sistemas de informações de negócios**, como bancos de dados, sistemas de dados geoespaciais, sistemas de recursos humanos, sistemas financeiros, sistemas de fluxo de trabalho e sistemas de gerenciamento de conteúdo; e (4) **sistemas de comunicação digital**, tais como *e-mail*, serviços

de mensagens curtas, serviços de mensagens multimídia, intercâmbio eletrônico de dados e fax eletrônico.

Entretanto, para os autores chineses Zhao, Duan e Yang (2019), os registros digitais podem ser categorizados: (1) como fotos digitais, (2) **documentos digitais**, (3) **vídeos**, (4) **e-mails**, (5) **áudios** e (6) **outros recursos de rede**, categoria na qual se enquadrariam as mensagens instantâneas.

Por último, a pesquisadora britânica Garmendia (2020), chefe de catalogação do Arquivo Nacional do Reino Unido, em uma palestra ministrada durante o II Congresso Nacional de Arquivo e Documento Eletrônico, realizado em fevereiro de 2020, na cidade de Madrid, apresentou um modelo de categorização dos documentos digitais, ao qual ela se referiu como primeira geração de documentos digitais, dividido em três conceitos:

a) *born-digital* (natos digitais): criado em formato digital, indissociável de seus metadados, que se tornam uma parte fundamental do documento digital e lhe conferem confiabilidade;

b) *digital surrogates* (representantes digitais): cópias digitais de originais em papel, microformas ou outros documentos em formatos semelhantes. Substitutos digitais; e

c) *digitised records* (registros digitalizados): documentos originalmente criados em formato análogo ao digitalizado, com altos níveis de escaneamento (digitalização) e metadados de proveniência, incorporados ao objeto digital como elementos cruciais de confiança na sua autenticidade, de modo que possam ser convertidos em documento arquivístico oficial (GARMENDIA, 2020, grifo nosso).

Ao apresentar os arquivos digitais da segunda geração, Garmendia (2020) não apresenta uma proposta de divisão, mas destaca as seguintes características:

Múltiplas versões; Acumulações amorfas; Criadores ambíguos ou criação híbrida; Cronologia relaxada; Procedência não confiável; Conexão errática entre função e acumulação; Duplicação ou sobreposição; Autenticidade dúbia; Objetos intrincados e embutidos, conjuntos de dados; Documentos compostos de formatos múltiplos; Modelos de inteligência artificial, documentos algorítmicos; Acumulações digitais em que a necessidade de arquivamento ou migração é o único elemento comum (GARMENDIA, 2020).

Em síntese, Rondinelli (2011) utiliza conteúdo, mas compartilha a necessidade de considerar o formato com Zhao *et al* (2019) e com o CIA (2015), que acrescenta a identificação da função, enquanto Katuu (2016) considera o ambiente de produção.

Garmendia (2020) apresenta a proposta mais atual, inclui documentos digitalizados e seus metadados nos registros considerados de primeira geração; nos arquivos digitais de segunda geração, não estabelece categorias, mas atribui relevantes características à sua composição, dentre as quais destacamos formatos múltiplos e documentos algorítmicos.

Com o intuito de acrescentar mais uma possibilidade de reflexão acerca da natureza dos registros digitais e suas diferentes maneiras de categorizá-los, apresentaremos, a seguir, a proposta de categorização utilizada pela rede social acadêmica ORCID.

O ORCID (*Open Researcher & Contributor ID*), que pode ser traduzido como ID aberto de pesquisador e colaborador, é utilizado para compartilhamento de informações sobre pesquisadores e os resultados de suas pesquisas com os agentes interessados, e se configura como uma das alternativas disponíveis para a gestão dos objetos digitais com finalidade acadêmica.

De acordo com Haak *et al* (2012), o ORCID é uma iniciativa aberta interdisciplinar internacional sem fins lucrativos, iniciada em 2012, com o objetivo de resolver o problema da ambiguidade do nome do pesquisador em benefício de todas as partes interessadas, quais sejam, instituições de pesquisa, agências de fomento, editores e os próprios pesquisadores. Para os autores, o ORCID permite a troca de dados entre os diferentes sistemas de submissão de trabalhos, além de interagir com sistemas de identificação como *Scopus*, *RePEC*, *ResearcherID*, *VIVO* e, recentemente, com a plataforma *Lattes*.

Graças ao compartilhamento de informações entre os diferentes sistemas, o ORCID permite que trabalhos sejam importados automaticamente das bases *Airiti*, *Bienlefeld Academic Search Engine (BASE)*, *Crossref Metadata Search*, *Datacite*, *Deutsche Nationalbibliothek (DNB)*, *Europe PubMed Central*, *ISNI*, *JaLC*, *KoreaMed*, *MLA International Bibliography*, *Redalyc*, *Research Data Australia* e *Scopus – Elsevier*.

É possível, ainda, fazer a inclusão automática de trabalhos no ORCID utilizando os dados digitais de referência do ArXivID, DOI, PubMed ID ou BibTeX. Na última opção, para fazer a inclusão de um trabalho acadêmico de forma manual, são apresentadas 4 possibilidades iniciais de categorias, dentro das quais é possível realizar outras categorizações secundárias. O quadro 7 demonstra os dois níveis de categorização para trabalhos na plataforma:

Quadro 7 – Categorização de trabalhos no ORCID

Categoria de trabalho	Tipo de trabalho
Publicação	Capítulo de livro Revisão de livro Livro Entrada em dicionário Dissertação/Tese Livro editado Entrada de enciclopédia Artigo de revista (<i>Journal</i>) Edição de revista (<i>Journal</i>) Artigo de revista Manual Artigo em folheto informativo Artigo em jornal Recurso online Preprint Relatório Ferramenta de pesquisa Publicação estudantil supervisionada Teste Tradução Web site Artigo de trabalho
Conferência	Resumo de conferência Artigo de conferência Pôster de conferência
Propriedade intelectual	Divulgação Licença Patente Direito autoral registrado Marca comercial
Outro	Anotação Artístico/desempenho Conjunto de dados Invenção Palestra/discurso Objeto físico Técnica de pesquisa Software Empresa spin off Padrões e políticas Padrão técnico Outro

Fonte: Elaborado pelo autor (2020) com base em <http://www.orcid.org>.

De forma complementar, o repositório digital *Figshare* também se apresenta como uma das formas de guarda para objetos digitais, que pode receber diversos tipos de documentos. Para Thelwall e Kousha (2016), o *Figshare* é um repositório científico de uso geral, que começou em 2011, e

apoia pesquisadores de todas as disciplinas para depositar qualquer tipo de informação eletrônica *online* para uso por outras pessoas.

Thelwall e Kousha (2016, p. 334) destacam que “as coleções de arquivos depositados no *Figshare* recebem um identificador de objeto digital para incentivar os usuários a creditar aos autores os recursos utilizados citando-o em publicações formais”.

Nesse sentido, buscou-se verificar como *Figshare* divide os objetos digitais nele depositados. Inicialmente, é solicitada a escolha de uma categoria, nas quais estão as áreas do conhecimento e suas subdivisões. Em seguida, é solicitada a escolha do tipo de item. Têm-se as possibilidades do quadro 8:

Quadro 8 – Categorização de objetos no *Figshare*

Categoria	Tipo de item
Ciências Agrícolas e Veterinárias	Figura
Astronomia, Astrofísica, Ciência espacial	Meios de Comunicação
Ciências Biológicas	Conjunto de dados
Ambiente e Design	Pôster
Química	Contribuição de Revista
Comércio, Gestão, Turismo e Serviços	Apresentação
Ciências da Terra e do Meio Ambiente	Tese
Engenharia	Programas de computador
Ciências da Saúde	Recurso Online
Humanidades	Pré-Impressão
Ciências da Informação e da Computação	Livro
Língua, Comunicação e Cultura	Contribuição de Conferência
Matemática	
Meta Ciência	
Física	
Psicologia	
Ciências Sociais	
Estudos em Artes Criativas e Escrita	
Estudos em Sociedade Humana	
Tecnologia	
Sem categoria	

Fonte: elaborado pelo autor (2020) com base em <http://www.figshare.com>.

É possível perceber que ORCID e *Figshare* possuem aproximações no que se refere à categorização dos trabalhos, entretanto, não há uma uniformidade quanto à classificação nas duas plataformas.

Do mesmo modo, ao buscar estabelecer relações entre os modelos de categorização de registros digitais presentes na literatura da área de Arquivologia e Ciência da Informação, tem-se uma diversidade de

possibilidades que incluem formatos, funções, metadados e uma série de elementos.

Os modelos do ORCID e *Figshare* complementam a compreensão sobre a necessidade do estabelecimento de categorias para os registros digitais, contudo, este trabalho analisará os registros da amostra com base nos modelos presentes nas publicações dos cinco autores identificados.

A partir dos conceitos encontrados, e em cumprimento aos objetivos da pesquisa, buscar-se-á estabelecer relações entre as categorias existentes na literatura e os registros digitais produzidos pelo Centro de Ciências Jurídicas da Universidade Federal da Paraíba.

3.2.3 Perspectivas sobre os registros digitais arquivísticos

Iniciamos esta reflexão a partir do trabalho de Acker (2019, p. 308), que questiona “em que vez de perguntar ‘o que é um registro?’ quando nos deparamos com novas TIC, cabe aos arquivistas perguntar, ‘quando o registro está na infraestrutura?’”.

A reflexão acima está fortemente vinculada à necessidade de pesquisas arquivísticas que se debrucem também sobre o aspecto físico dos registros digitais, ou seja, a compreensão das características materiais que diferenciam os registros digitais, haja vista o fato de, na visão de Acker (2019), a maior parte dos trabalhos ter se voltado para os aspectos lógicos e conceituais, ou seja, de contexto computacional e de possibilidades de visualização dos registros.

Para Hofman *et al* (2017), “na teoria arquivística, os termos documento arquivístico – referindo-se a documentos produzidos no decorrer das atividades e mantidos para uso posterior ou referência – e registros são sinônimos”.

De acordo com os estudos de Cohen (2015, p. 28), “os documentos em papel são fixados a um suporte em uma forma documental, que é alterada ao longo do tempo por atos intencionais e efeitos naturais do meio ambiente”. Já os registros digitais, “conforme sejam criados, podem não ser fixados em uma forma documental”. Ao contrário disso, “são fixados em diferentes suportes, em diferentes formas documentais, que o autor (do registro/documento) nunca vê representado quando ‘salvo’”.

Em que pese o conceito de forma documental, e as definições de documento arquivístico manifestado e documento arquivístico armazenado de Rondinelli (2011), indagamos: é possível comparar completamente as características dos registros digitais arquivísticos com as características do documento arquivístico digital?

Frente a isso, Cohen (2015) destaca:

Muitos dos métodos que tornaram os registros analógicos confiáveis ao longo do tempo não foram traduzidos nas novas formas de gestão de registros digitais. Por exemplo, os processos cartoriais não têm analogia direta na arena digital. Não existe um órgão administrativo autorizado que ateste uma assinatura digital feita por um indivíduo identificado em sua presença na maioria dos processos normais usados no comércio digital (COHEN, 2015, p. 28).

Sabe-se que o Brasil possui uma norma que regula o uso de certificados digitais para a utilização, dentre outras aplicações, em assinaturas digitais. Trata-se da Medida Provisória convertida em lei, de nº 2.200, de 24 de agosto de 2001, que criou a infraestrutura de chaves públicas brasileira, conhecida como ICP-Brasil. Embora a iniciativa permita avanços quanto à questão mencionada por Cohen (2015), tal sistema está longe de cumprir, no ambiente digital, a função que os cartórios cumprem no ambiente dos documentos analógicos, sem mencionar as questões de verificação de autoria, autenticidade e autenticação envolvidas.

Não é a intenção direta desta pesquisa adentrar no universo de temas envolvidos com o papel do Estado na criação de políticas que viabilizem a preservação e utilização dos registros digitais de forma segura, embora esse seja um excelente tema para um próximo trabalho. Entretanto, na revisão de literatura realizada, encontramos os estudos de Guo (2016), que atribuem ao agente solucionador da necessidade de atestação de documentos e transações digitais, o nome de “**terceiro confiável**”, como sendo um intermediário entre produtores e usuários de registros digitais cuja autoria precisasse ser verificada, além da verificação quanto à existência ou não de adulterações dos dados ali contidos. No caso estudado por Guo (2016), o arquivo municipal funciona como repositório digital dos registros gerados por determinados tipos

de negócios, com transferências ou transmissões realizadas mensalmente pelos produtores.

O trabalho de Cohen (2015, p. 31) afirma ainda que “a confiabilidade é algo que a ciência forense digital atual não pode demonstrar realisticamente”, e conclui que “é necessário que os mecanismos digitais e suas propriedades exclusivas sejam levados em consideração para que os metadados associados à gestão dos registros digitais possam efetivamente geri-los” (COHEN, 2015, p. 42).

Nesse sentido, é preciso perceber que os arquivos, compreendidos como conjuntos documentais ou como instituições custodiadoras de documentos, encontram na tecnologia digital uma necessidade de reflexão importante, sobre sua constituição, seus métodos e suas práticas.

Nessa perspectiva, Klareld (2015, p. 176, grifo nosso) aponta que “o conceito de **arquivo é multifacetado**, [...] dá espaço para diferentes interpretações do que um arquivo é, ou poderia ser. Isso inclui o uso e definição do termo 'arquivo digital’”.

Para Anderson *et al* (2015), mesmo com novos desafios, o arquivo dos registros digitais ainda precisa realizar funções arquivísticas tradicionais – como é o caso da avaliação –, entretanto, sem excluir a necessidade de documentar a natureza e o contexto probatório **próprio dos registros digitais**. Os autores completam que “os contextos ricos e diversos nos quais os registros digitais são criados e usados devem ser o ponto de partida a partir do qual começamos a decidir o que será preservado a longo prazo”. Do mesmo modo, Conway (2015) afirma que,

Em particular, quando as informações nascem digitais, e não mais forem necessárias para seus propósitos originais, após serem avaliadas por valor contínuo e transferidas fisicamente para sua nova casa em servidores de arquivamento, o conceito tradicional de documento arquivístico servirá como um guia apropriado para o gerenciamento desta informação digital ao longo do tempo. As convenções da ciência arquivística e da gestão de documentos fornecem as orientações teóricas e práticas necessárias para a triagem de coleções digitais, para identificar aquelas com valor duradouro, uso a curto e longo prazo, e a viabilidade técnica e econômica (CONWAY, 2015, p. 15).

O trabalho de Zhang (2012) também aborda a necessidade de manutenção das funções clássicas, como a descrição arquivística, no ambiente digital, embora seja focado em documentos digitalizados. O artigo revela que os materiais de arquivo podem ser convenientemente representados no nível de coleções em formato digital, especialmente em ambiente *web*, e com a capacidade de fornecer informações contextuais de proveniência em tempo recorde. Mas ressalta que [...] “a ordem original dos registros de arquivo pode ser adotada para facilitar a pesquisa e navegação. No entanto, seu papel parece ser menos crucial e menos insubstituível”. Aponta ainda, na direção já mencionada por Klareld (2015), que

A representação de conteúdo, por outro lado, parece desempenhar um papel mais ativo na representação digital de arquivos. Esse papel ativo parece estar em processo de transformação do modelo tradicional de arquivamento hierárquico para um novo sistema de representação arquivística característico de múltiplos e facetados pontos de acesso e séries virtuais (ZHANG, 2012, p. 65).

Ocorre que, enquanto nos documentos analógicos (em papel) é possível fazer inferências quanto ao seu contexto quando eles já estão na idade permanente, com os registros digitais isso não será possível em todos os casos, seja porque eles não estarão acessíveis ou porque não existirão os elementos que tornem essa identificação possível.

Ou seja, a definição quanto ao contexto de ambos se dará por intermédio de condições necessariamente diferentes, de modo que os registros digitais precisam ser avaliados inicialmente na sua concepção. Sobre isso, o Conselho Internacional de Arquivos (2005), por meio de um manual específico sobre o tema, advertiu:

[...] para preservar os registros digitais, é importante que os requisitos de arquivamento sejam incorporados quando os sistemas são projetados e que os registros sejam cuidadosamente controlados durante todo o seu ciclo de vida para garantir sua qualidade e integridade contínuas (CONSELHO INTERNACIONAL DE ARQUIVOS, 2005, p. 5).

Nessa mesma linha, Moss *et al* (2018, p. 6) defendem que metadados sozinhos não constituem a solução. Os autores questionam: “além dos

metadados gerados no ambiente, como você os cria?” Segundo os autores mencionados, atribuir ao usuário a responsabilidade por criar e adicionar os metadados adicionais é irracional. Os estudos da revisão de literatura realizada por esta pesquisa indicam que robôs e inteligência artificial devem figurar no conjunto de soluções para a criação de metadados adicionais que contribuam para a gestão dos registros em ambientes digitais.

Conforme o trabalho de Allan (2014; 2015), a designação para a atribuição de metadados pelos produtores dos registros não está funcionando de forma eficiente nos países que já implantaram tal política. Para os autores das duas investigações produzidas pelo Arquivo Nacional do Reino Unido, junto aos órgãos públicos a ele vinculados, os resultados demonstraram que não se pode esperar que gestores ajam como arquivistas e saibam avaliar o que deve ser preservado e o que não deve no ambiente digital. Isto é, há que se elaborar um planejamento para a implementação de uma preservação digital sistêmica com a participação também de arquivistas, mas não somente. Índícios de como fazer isso estão presentes no trabalho de Conway (2015), quando afirma que

Esquemas de metadados, modelos de controle de processos e procedimentos para estabelecer e certificar a confiabilidade dos repositórios digitais, cada um a seu modo, contribuem para uma confiança emergente de que o gerenciamento de longo prazo do conteúdo digital em escala é desejável e tecnicamente viável (CONWAY, 2015, p. 15).

Também sobre metadados se concentra o trabalho de Rogers (2015), que conceitua registros digitais arquivísticos com base nos metadados, e atribui a eles a tarefa de subsidiar as ações de gestão e preservação dos registros digitais. Para a autora, utilizando os princípios diplomáticos básicos, os registros digitais podem ser identificados por meio da definição de quais metadados devem ser capturados, gerenciados e preservados.

Todos os esforços para a gestão dos registros arquivísticos em ambiente digital tem como objetivo a sua preservação e acesso posteriores por aqueles que necessitarem. Nesse sentido, Santos (2012) reforça:

A preservação de documentos arquivísticos digitais não é, *lato sensu*, diferente da preservação de quaisquer outros tipos de documentos digitais. É necessário confrontar a obsolescência tecnológica de *hardware*, *software* e formatos. Todavia é no conjunto de metadados, cujo registro é necessário na criação do documento e, também, na execução dos procedimentos de preservação, que se encontra o foco da área (SANTOS, 2012, p. 124).

Arquivistas americanos já trabalham em requisitos para gestão de registros digitais desde o final do século XX. De acordo com o trabalho de Bantin (1998), ao analisar o sistema de registros de uma universidade, foram identificados como sendo os principais problemas:

1) a falta de procedimentos, mecanismos e prazos para garantir que os metadados permaneçam unidos ou vinculados ao registro; 2) a falta de mecanismos definidos para migração (em oposição a backups simples) de registros; e 3) a falta de procedimentos para garantir que os registros permaneçam invioláveis ao longo do tempo (BANTIN, 1998, p. 362).

A literatura aponta para uma necessidade obrigatória de atenção aos metadados no trabalho de gestão e preservação dos registros digitais arquivísticos, embora ainda se tenha muito a fazer nesse sentido. Além disso, Moss *et al* (2018) ressaltam:

Enquanto o custo do armazenamento está caindo, o mesmo não pode ser dito para os custos do gerenciamento de registros. Consequentemente, o custo de “esquecer” os dados por exclusão seletiva exige mais esforço e é mais caro do que preservá-los. Isso enfaticamente inclina o padrão para a preservação ou talvez para “guardar as coisas” com mais precisão. Além disso, a exclusão de registros digitais é difícil. Cópias de *e-mails* são armazenadas no computador do remetente, no computador do destinatário ou em um servidor central e também podem ser armazenadas em um *buffer* ou armazenamento temporário (MOSS *et al*, 2018, p. 6).

Moss *et al* (2018) sugerem, ainda, uma ampliação da visão relacionada aos arquivos, em conformidade com o trabalho de Zhang (2012), em um trecho que consideramos relevante para esta análise.

Como todas as condições médicas que estamos familiarizados, tratamentos que lidam apenas com os sintomas nunca fornecerão uma cura. Podemos desenvolver a "pílula de *e-mail*" e a "pílula do *SharePoint*", mas o dano subjacente ao corpo do arquivo continuará a menos que alteremos nossa abordagem. Os arquivistas estão em uma posição única para observar o progresso da doença e desenvolver uma resposta sistêmica. Os arquivistas e seus usuários, principalmente a comunidade acadêmica, precisam mudar para uma posição em que veem os acervos arquivísticos como acervos on-line de dados que possuem características totalmente diferentes dos acervos analógicos tradicionais. Sua interpretação deve ser fluida e suscetível à análise por uma nova e crescente gama de ferramentas sofisticadas. As velhas fórmulas mágicas de metadados, ordem original e até registros originais perderam seu poder (MOSS *et al*, 2018, p. 7).

E os desafios não param com a proposta de Moss *et al* (2018) para inclusão dos dados no escopo da arquivística. De acordo com Hofman *et al* (2017), as pessoas e organizações provavelmente não vão parar de “confiar” seus registros a provedores em nuvem, em que a perda do contexto desses dados configura-se como uma ameaça. De acordo com os autores, os riscos envolvidos no armazenamento não controlado em nuvem precisam ser mitigados e uma solução possível seria o controle de privacidade no nível do registro.

No entanto, essa solução exigirá reformular o modelo de privacidade atual de um modelo focado em “informações” privadas – um termo amplo que engloba dados e registros de ambos – para um focado nos desafios coexistentes de dados privados (que sem dúvida existem fora dos registros), bem como de registros privados (HOFMAN *et al*, 2017, p. 10).

O trabalho de Hoffman *et al* (2017) acrescenta os algoritmos na discussão sobre confiabilidade dos registros digitais armazenados em nuvem ao dizer que

Nova pesquisa algorítmica poderia potencialmente desenvolver soluções para verificar e autenticar registros sem depender de terceiros, como autoridades certificadoras; esses algoritmos poderiam garantir aos usuários que os registros que eles confiaram aos seus provedores de serviços em nuvem são os registros que eles retornam, mesmo que os registros devam ser migrados (HOFMAN *et al*, 2017, p. 9-10).

Cabe destacar novamente o trabalho de Guo (2016), que tratou sobre o papel do arquivo, enquanto instituição de custódia, na questão da confiabilidade de registros digitais. O autor descreve a atuação do arquivo público como terceiro confiável neutro e imparcial para depósito de registros de instituições públicas e privadas. Guo (2016, p. 13) pontua três principais dificuldades: “a aplicabilidade dos métodos de encapsulamento, a exigência de gerenciar dados estruturados de forma fixa e o grande volume de dados envolvidos”. Entretanto, essa expansão no papel do arquivo trouxe, segundo o autor, “novas questões para os arquivistas quanto ao relacionamento entre arquivos públicos e registros privados, e o relacionamento entre arquivos públicos e serviços de armazenamento comercial”.

Sobre as ações que devem ser adotadas pelos governos para a preservação dos seus registros digitais, Ngoepe e Nkwe (2017) destacam a necessidade de consultas públicas antes da destruição de registros, e da importância do envolvimento dos gestores desses registros durante a fase corrente, bem como de autoridades arquivísticas públicas.

Recomenda-se que os registros digitais a serem preservados sejam identificados no momento de sua criação pelos gerentes de registros, com a assistência de arquivistas e o envolvimento do público através da apresentação de autoridades de descarte no parlamento e subsequente publicação na mídia nacional para convidar comentários públicos (NGOEPE; NKWE, 2017, p. 13).

No trabalho, os autores comentam a problemática envolvida na transferência de registros digitais para o Arquivo Nacional da África do Sul (NARSSA). Sobre a experiência do NARSSA, Ngoepe e Nkwe (2017, p. 14) ressaltam que “os registros de valor permanente são transferidos para um repositório de arquivos após um período de 20 anos. Vinte anos é muito tempo, pois registros de valor permanente podem ser perdidos, especialmente em ambientes digitais”.

Parte desse problema, conforme Netshakhuma (2018), que também estuda a gestão dos registros digitais em Botsuana, Moçambique, Namíbia e África do Sul, se dá pelas falhas da legislação de gerenciamento de arquivos e registros dos países africanos, principalmente no que se refere à atribuição de responsabilidade (ou falta dela) e definição de critérios de governança.

Em outro trabalho sobre os problemas da gestão de registros digitais no continente africano, Ngoepe e Katuu (2017) abordam o problema pelo viés educacional. De acordo com os autores, arquivistas não têm suas demandas devidamente contempladas na adoção de soluções tecnológicas em virtude de problemas relacionados à sua formação, pois “[...] os arquivistas e gerentes de registros são incapazes de se comunicar adequadamente com os profissionais de TI em termos e conceitos mutuamente compreensíveis” (NGOEPE; KATUU, 2017, p. 8). Isto é, a resolução do problema não pode ser vista apenas como a falta de requisitos nos sistemas, mas deve contemplar a criação de um ambiente favorável e capaz de promover um trabalho conjunto de todas as áreas envolvidas.

Em sintonia, Hofman *et al* (2017, p. 9) nomearam os arquivistas, profissionais de tecnologia da informação, advogados e profissionais da informação em geral como os responsáveis por esse trabalho conjunto.

A necessidade de adequação na prática e na formação do arquivista é corroborada por Garmendia (2020), quando afirma que “o documento digital transtorna e desestabiliza a prática arquivística”. A autora afirma ainda que “o documento digital exige mudanças na seleção e na formação de pessoas para trabalharem nos arquivos”.

Sobre a função arquivística da seleção no ambiente digital, a UNESCO, em parceria com instituições arquivísticas de diversos países, estabeleceu diretrizes com critérios para a preservação de longo prazo do patrimônio digital. Os critérios, segundo Crofts (2016, p. 13), são a redundância, definida como “a necessidade de existirem várias cópias que são armazenadas em pelo menos dois locais físicos diferentes”; o gerenciamento ativo, que consiste “no conjunto de procedimentos para garantir acessibilidade e integridade a longo prazo”; e o gerenciamento de metadados, segundo o qual os metadados devem ser considerados como “qualquer informação (em forma digital ou física), essencial para garantir que o material digital está sendo preservado, e permanece acessível, inteligível e utilizável ao longo do tempo”.

O documento da UNESCO lista os três tipos principais de metadados associados que precisam ser preservados:

Estruturais, necessários para a capacidade técnica de ler conteúdo digital); **Descritivos**, contendo informações contextuais bibliográficas, de arquivo ou de museu, que podem ser gerados pelo sistema ou criados por profissionais do patrimônio, conteúdo criadores e/ou usuários; e/ou **Administrativos**, documentando o gerenciamento de um objeto digital enquanto estiver em sua coleção (CROFTS, 2016, p. 14, grifo nosso).

São definidos também os cinco requisitos funcionais básicos para metadados digitais:

Identificação: os metadados devem permitir que cada objeto digital seja identificado exclusivamente e sem ambiguidade. Isso geralmente exigirá que um identificador globalmente exclusivo seja atribuído a cada item; **Localização:** os metadados devem permitir que cada objeto digital seja localizado e recuperado. A validade a longo prazo desses dados de localização precisa ser garantida para que os itens não sejam perdidos quando os sistemas são migrados ou atualizados; **Descrição:** é necessária uma descrição do objeto digital para facilitar a recuperação e interpretação. Os metadados descritivos se dividem em duas categorias: dados sobre conteúdo e sobre o contexto. Os dados sobre o conteúdo de um item geralmente podem ser recriados por exame e consulta. No entanto, é útil como uma ajuda para encontrar recursos de descoberta. Dados sobre o contexto, onde, quando e por quem um item foi criado, quais foi usado para, seu lugar em relação a um corpus geral de material, é muito mais difícil de recriar uma vez perdido; **Legibilidade:** os metadados sobre a estrutura, formato e codificação de objetos digitais são necessários para garantir que eles permaneçam legíveis ao longo do tempo. Este requisito funcional é particularmente importante para objetos digitais, pois eles não podem ser lidos sem mediação da tecnologia. Esses metadados devem identificar os padrões relevantes e fornecer referências à documentação técnica, arquivos de autoridade e outro material relacionado necessário para uma renderização completa do recurso digital. É necessário tomar cuidado para garantir que todas as múltiplas camadas de um objeto digital possam ser interpretadas: a partir do encapsular o formato do arquivo à representação e codificação dos próprios dados; **Gerenciamento de direitos:** direitos, condições de uso e restrições aplicáveis a cada item digital precisa ser registrado nos metadados. Esses metadados devem identificar convenções e leis aplicáveis e fornecer referências a normas legais relevantes quanto a documentação, contratos, etc., bem como os detentores de direitos (CROFTS, 2016, p. 15, grifo nosso).

Sobre o panorama internacional das normas e diretrizes de boas práticas para a gestão dos registros digitais, Katuu (2015) fez a seguinte analogia: os gestores de registros de cada país são como motoristas dirigindo em um nevoeiro. Enquanto existem poucos carros na rua, ainda será possível avançar, mesmo com dificuldades. Quando o volume de carros aumentar, postes de iluminação pública precisam ser instalados para diminuir a chance de acidentes. De acordo com o autor, novos conceitos como governança da informação e gestão do conhecimento aumentam a neblina. Ou seja, a comunidade internacional possui esforços para gerir os registros digitais, mas a falta de um consenso quanto aos termos já se configura como um dos primeiros problemas a serem enfrentados.

Estabelecendo ligações com o desenvolvimento sustentável que a correta gestão dos registros digitais pode oferecer a um país, o trabalho de Baron e Thurston (2016) também enfatizam a necessidade de reflexão quanto ao regramento dos registros digitais arquivísticos em nível internacional.

Dados os altos riscos e custos envolvidos, o investimento na construção de estruturas de controle eficazes, incluindo a institucionalização de padrões internacionais de gerenciamento de registros, fará uma grande diferença nas aspirações globais para uso das informações com o objetivo de alcançar o desenvolvimento sustentável (BARON; THURSTON, 2016, p. 12).

Políticas e práticas de gerenciamento de registros precisam ser atualizadas para a era digital, caso contrário, na visão de Baron e Thurston (2016), que tratam do assunto como uma questão legal e econômica, o aumento de informações poderá sobrecarregar os sistemas, levando a custos mais altos e perda de registros cada vez mais relevantes, diretamente relacionados ao desenvolvimento dos países.

Diante dos desafios que os registros digitais arquivísticos colocam para a sociedade, é incontestável a necessidade de se repensar as possibilidades de contribuição do arquivista nesse cenário.

[...] a mudança de suporte desloca a atividade do arquivista para uma outra fase do ciclo de vida do documento, ao da produção do documento. Se a explosão documental impeliu o arquivista a antecipar sua atividade da fase permanente para

tratar o documento na fase intermediária, a mudança de suporte promove nova mudança, desta vez deslocando a atenção do arquivista para um momento anterior à própria gênese do documento e criando uma nova área de atuação do arquivista, [...] que prefiro identificar como *arquivista digital*, que deve assegurar que o documento produzido em sistemas informatizados sejam classificados, descritos e produzidos de forma controlada em sua origem (SCHWAITZER, 2018, p. 15).

Outros autores reforçam a necessidade de gestão antes da produção dos registros digitais, como Santos (2012), que afirma que

É imprescindível que a preservação dos documentos arquivísticos digitais se inicie no ato de criação do documento, ou mesmo antes. Isso porque deve contemplar o planejamento quanto às potencialidades das tecnologias de informação e comunicação envolvidas e procedimentos necessários à conformidade legal e normativa institucional dos documentos produzidos, aí incluídos os sistemas informatizados que gerenciarão esses documentos (SANTOS, 2012, p. 118).

Para tanto, existem normas internacionais que já apontam para a utilização dos instrumentos de gestão de documentos tradicionais (plano de classificação e tabela de temporalidade) como possibilidades de solução para as novas demandas, feitas as devidas adaptações.

Dentre as utilidades do plano de classificação e da tabela de temporalidade podem-se destacar, além da organização intelectual das informações, a possibilidade de definir níveis de proteção e acesso adequados a um conjunto de documentos e a atribuição de períodos de guarda e ações de eliminação sobre os documentos (ISO 15.489-1, 2001, p.13).

Este capítulo apresentou a visão de autores e instituições a respeito dos registros digitais arquivísticos. A análise nos permitiu observar que os registros digitais arquivísticos demonstram-se como um desafio para a área e para seus profissionais, na medida que suas formas documentais não obedecem ao padrão dos documentos não digitais. Além disso, mesmo com a constatação da contribuição que a Arquivologia pode oferecer para a sua gestão e preservação, a falta de vinculação com o suporte em alguns casos, amplia o desafio e provoca os preceitos mais tradicionais da ciência arquivística.

A definição de arquivo é posta em discussão, a base conceitual da formação dos arquivistas também, visto que os registros digitais impõem novas necessidades e, conseqüentemente, novas habilidades para os profissionais responsáveis por sua gestão e preservação, dentre os quais o arquivista está inserido, embora não deva estar sozinho.

O diálogo que se estabelece sobre a função do arquivista não tem a intenção de diminuir a sua importância, mas, pelo contrário, de torná-lo capaz de lidar também com arquivos em que dados, algoritmos e diferentes tipos de registros devem ser considerados parte do acervo, com destaque especial para os metadados, matéria-prima da gestão e preservação dos registros digitais.

Seja na definição de regras institucionais, nacionais ou internacionais sobre os registros digitais arquivísticos, a literatura aponta para a um modelo de gestão e preservação digital em que a autenticidade dos documentos seja uma condição pensada antes da sua criação. Isto é, a definição dos seus critérios de armazenamento deve levar em conta uma avaliação que contemple o contexto próprio do ambiente digital, e use as ferramentas tecnológicas disponíveis para a implantação de requisitos prévios à criação dos registros digitais.

3.3 METADADOS DE PRESERVAÇÃO DIGITAL PARA REGISTROS DIGITAIS ARQUIVÍSTICOS

Com as revisões anteriores, verificou-se a necessidade de aprimorar determinado elemento acerca dos registros digitais, especificamente sobre os metadados relacionados à sua preservação digital, pois as atividades de gestão dos registros digitais arquivísticos estão fortemente relacionadas a este tópico. Realizou-se, portanto, uma revisão sistemática de literatura sobre o tema, cujos resultados serão descritos a seguir.

Os estudos da Ciência da Informação apontam que o acesso aos registros digitais no longo prazo só será possível mediante a adoção de políticas de preservação digital que incluam, dentre outros fatores, normas relacionadas aos metadados (SPENCE *et al* 2019; RIES, 2018; FORMENTON *et al*, 2017; RAMALHO *et al*, 2017; SVÄRD, 2017; LI; SUGIMOTO, 2014; SAYÃO, 2010; CANTARA, 2006; BRINDLEY *et al* 2004).

Sobre preservação digital, os fluxos de trabalho devem ser planejados em sintonia com o tipo de material a ser preservado. Por exemplo, nos estudos de Spence *et al* (2019), sugere-se que objetos digitais provenientes de digitalização devem ser tratados antes dos registros nato digitais. Considerando os objetivos deste trabalho, o estudo será destinado apenas aos registros nato digitais.

Similarmente a Rogers (2015), acerca da constituição dos registros digitais, Ries (2018) define os registros nato digitais como aqueles constituídos por

[...] objetos digitais, arquivos temporários, metadados e traços fragmentados do processo de criação, espalhado por vários locais do sistema, e que devem ser analisados em relação ao contexto histórico computacional específico, tendo a sua materialidade distribuída no conjunto de hardware, sistema operacional e aplicativo onde foi produzido (RIES, 2018, p. 1 tradução nossa).⁷

Para Svärd (2017), a instituição que deseja preservar o conteúdo dos registros digitais acessível ao longo do tempo, precisa ter um repositório, recursos humanos e financeiros compatíveis com a complexidade que a gestão eficaz do material digital requer.

Uma política de preservação digital não depende apenas da tecnologia, depende igualmente de recursos e de uma estratégia organizacional alinhada a esse propósito (SPENCE *et al*, 2017). Estratégia que inclua uma abordagem proativa da informação digital, em que os metadados sejam planejados antes mesmo da criação dos registros digitais (SVÄRD, 2017).

Metadados são considerados a chave para a preservação da informação no ambiente digital, pois facilitam a gestão de longo prazo, e documentam as condições de acesso necessárias para o material digital a ser preservado (BRINDLEY *et al*, 2004).

Nesse sentido, convém mencionar as principais definições para o termo “metadados” encontradas por esta pesquisa. Para Dappert *et al* (2013), metadados são informações sobre um objeto necessárias para gerenciá-lo. Na

⁷ No original: “[...] the born-digital record, consisting of digital objects, temporary files, metadata, and fragmented traces of the writing process scattered across multiple system locations, has to be analyzed with regard to the specific historical computing context, its distributed materiality ensemble of hardware, operating system, and application”.

Política de Preservação Digital, produzida pelo Grupo de Trabalho de Preservação Digital da Biblioteca da Universidade de Glasgow,

Metadados são informações adicionais que acompanham o recurso digital para garantir que ele possa ser acessado e compreendido ao longo do tempo, como palavras-chave usadas como tags para permitir uma busca e recuperação precisas. Os exemplos incluem o criador dos dados, a data e hora de criação, o tamanho do arquivo de dados, o tipo de arquivo (.docx, .pdf), etc. (DIGITAL PRESERVATION WORKING GROUP – UNIVERSITY LIBRARY – GLASGOW UNIVERSITY, 2020).

Saliente-se que, na literatura, metadados são divididos em categorias ou tipos específicos, a saber:

- Metadados descritivos ou de identificação – visam à pesquisa, recuperação e identificação. Podem conter elementos como, por exemplo, título, autor, assunto e palavras-chave.
- Metadados estruturais – vinculam de forma hierárquica distintos objetos digitais (textos, imagens, áudios etc.) integrantes de um mesmo documento ou recurso informacional.
- Metadados administrativos – dispõem informações que suportam a gerência dos recursos arquivísticos eletrônicos. Incluem de que forma e em que ocasião os recursos foram gerados, espécie de ficheiro ou de arquivo de computador e demais informações técnicas, além dos titulares com direitos ou permissões de acesso.
- Metadados técnicos – especificam os aspectos técnicos dos arquivos e dos seus formatos.
- Metadados de preservação – incluem informações requeridas ao arquivamento e salvaguarda dos objetos digitais ao longo do tempo (FORMENTON *et al*, 2017, p. 84-85).

Com base nos objetivos já apontados, o recorte utilizado por este trabalho são os metadados de preservação digital. Dessa forma, trataremos dos conceitos que se relacionam com os metadados de preservação, a começar pelo termo “longo prazo” que, para os fins desta pesquisa, será entendido de acordo com a definição da Administração Nacional Norte-Americana de Aeronáutica e Espaço (NASA), como “o tempo suficiente para que haja uma preocupação com a mudança de tecnologia”, conforme Sawyer e Reich (2004 *apud* CANTARA, 2006, p. 40).

Destaque-se que a referida agência americana, por meio do Comitê Consultivo para Sistemas de Informação Espacial (CCSDS), de acordo com

Brindley (2004), Cantara (2006) e Sayão (2010), é autora do mais importante modelo de requisitos de preservação de objetos digitais da atualidade, o *Reference Model for an Open Archival Information System*, mais conhecido como (OAIS), que pode ser definido como

uma infraestrutura conceitual que descreve o ambiente, as interfaces externas, os componentes funcionais e os objetos de informação, associados com um sistema responsável pela preservação de longo prazo de materiais digitais (SAYÃO, 2010, p. 13).

O autor acrescenta ainda uma relação entre o OAIS, os metadados e a preservação digital:

Quando oferece uma descrição de alto-nível dos tipos de informação que fluem no espaço onde se desenrolam os processos do que chamamos de preservação digital, o OAIS torna evidente o vínculo entre metadados e preservação digital e, dessa forma, reconstrói a ideia de metadados de preservação em bases mais sólidas. Como desdobramento, o modelo de informação OAIS vem constituindo o fundamento comum para a orientação e o desenvolvimento da maioria das iniciativas de metadados de preservação surgidas nos últimos anos (SAYÃO, 2010, p. 15).

Conforme um dos objetivos desta pesquisa, de identificar os metadados de preservação necessários para a preservação digital dos registros digitais arquivísticos, buscou-se na literatura os elementos constitutivos de tais metadados e conceitos que permitam a análise desse tema.

Primeiramente, o termo “preservação digital” pode ser entendido como um processo de acompanhamento do ciclo de vida dos objetos digitais, uma responsabilidade de gestão contínua, começando com a criação do recurso e continuando por um período indefinido (CANTARA, 2006).

Spence *et al* (2019) compactuam da necessidade de gestão dos objetos digitais antes mesmo da sua criação, como estratégia de preservação digital que garanta que os registros digitais sejam criados conjuntamente com os metadados necessários para a sua interpretação.

O trabalho de Dappert *et al* (2013) chama a atenção para a necessidade de preservar não apenas o conteúdo, mas a configuração e o ambiente computacional necessário para que os registros digitais possam ser acessados ao longo do tempo. Para esse objetivo, destaca a iniciativa PREMIS,

[...] uma iniciativa de metadados de preservação digital que integra metadados do ambiente computacional, com metadados de descrição dos objetos, eventos, agentes, e declarações de direitos em um nível direcionado a repositórios OAIS. Com um nível mais abrangente, incorporando uma abordagem de todo o ciclo de vida, o modelo conceitual do domínio da preservação digital (DAPPERT *et al*, 2013, p. 108).

Os metadados de preservação podem ser definidos como a informação (incluindo-se os diversos tipos de metadados) que apoia e documenta a preservação de longo prazo bem sucedida de materiais digitais (DAY, 2003; DAPPERT *et al*, 2013; SAYÃO, 2010), ao passo que as funções destes metadados de preservação podem ser descritas como:

- Registro do histórico, do hardware, software e outras dependências técnicas, das estruturas e das mudanças (de custódia legal e física ou de natureza tecnológica, social etc.) suportadas pelos recursos/objetos digitais, no decurso de todo o seu ciclo de vida (isto é, criação, seleção e descarte, identificação persistente, descrição e acesso, armazenamento e preservação).
- Documentação das tomadas de decisão e atos de preservação, dos métodos de preservação e seus resultados, das ações de gestão de coleções e direitos, além dos próprios metadados, de modo a apoiar os processos ou as iniciativas correntes e vindouras de preservação digital.
- Suporte à descoberta, localização, recuperação, referência única e persistente, preservação, acesso, utilização, comprovação de autenticidade e integridade, bem como gestão de direitos de propriedade intelectual de objetos arquivísticos e museológicos digitais ou eletrônicos.
- Assistência à uma gerência ideal dos objetos digitais, dos sistemas de informação (registros de correspondência e de documentos, registros de funcionamento dos sistemas e seus acessos por intermédio de práticas de auditorias mantidas pelos serviços de TIC, por exemplo), da comunidade usuária e das funcionalidades a serem oferecidas (FORMENTON *et al*, 2017, p. 85).

Em complemento, Sayão (2010) estabelece cinco categorias de informação baseadas em Lavoie e Gartner (2005), que devem ser traduzidas pelos metadados de preservação:

- 1) proveniência – os metadados de preservação devem registrar informações sobre a história do objeto desde sua origem, traçando a sua cadeia de custódia e de propriedade;
- 2) autenticidade – os metadados de preservação devem incluir informações suficientes para validar que o objeto é de fato o que diz ser e que não sofreu alterações – intencionais ou não – não documentadas;
- 3) atividades de preservação – os metadados de preservação devem documentar as ações tomadas ao longo do tempo para preservar o objeto digital e as consequências dessas ações sobre aparência, usabilidade e funcionalidades do objeto;
- 4) ambiente técnico – os metadados de preservação devem descrever as dependências técnicas necessárias para a apresentação e uso dos objetos digitais, tais como hardware, sistema operacional e software de aplicação;
- 5) gestão de direitos – os metadados de preservação devem registrar todos os itens relacionados às questões de propriedade intelectual que limitem as ações de preservação, de disseminação e uso por parte de usuários de hoje e do futuro (LAVOIE; GARTNER, 2005 *apud* SAYÃO, 2010, p. 11-12).

Os estudos de Sayão (2010), Day (2001), Brindley (2004) E Cantara (2006) informam que o pressuposto básico do OAIS é que um recurso de informação tenha dois componentes: o objeto que precisa ser preservado e as informações que tornem o objeto compreensível para os usuários do repositório. Ou seja, todo objeto de informação é composto por objeto de dados (físico ou digital) e por uma informação de representação que permite a completa interpretação dos dados em informações com significado.

É nesse sentido que o modelo OAIS estabelece três tipos de pacote de informação, definidos como: **pacote de informação de submissão**, constituído do conteúdo e metadados elaborados pelo produtor no momento do depósito no repositório; **pacote de informação de armazenamento**, constituído pelo conteúdo e metadados efetivamente armazenados e gerenciados pelo repositório por longo prazo; e **pacote de informação de disseminação**, formado pelo conteúdo e metadados entregues pelo repositório em resposta a uma requisição de acesso demandada pelo usuário (DAPPER *et al* 2013; SAYÃO, 2010; CANTARA, 2006).

Para o fim de preservação de longo prazo, conforme Sayão (2010, p. 17), o pacote objeto desta pesquisa é o pacote de informação de armazenamento, que se constitui de vários tipos de informação, quais sejam:

Informação de conteúdo – é a informação que o repositório tem obrigação de preservar, inclui a informação de representação, que são informações necessárias à apresentação e à interpretação da cadeia de bits que constituem o objeto armazenado como informação com significado para uma determinada comunidade alvo;

Informação de descrição de preservação – informação que apoia e documenta a preservação dos objetos arquivados no repositório;

Informação de empacotamento – informação que agrega todos os componentes de um pacote de informação – conteúdo e seus metadados – numa única unidade lógica;

Informação descritiva – informação que apoia o usuário na descoberta e na recuperação de objetos armazenados no repositório (CCSDS, 2002, p. 2-6).

Observa-se que, para além das categorias de informação apontadas anteriormente, o CCSDS estabeleceu um detalhamento da informação de descrição de preservação em quatro grupos distintos de dados, definidos em:

Informação de referência – tem origem na necessidade de identificar e de localizar um objeto ao longo do tempo para manter a sua integridade; a referência identifica ou, se necessário, descreve um ou mais mecanismos usados para assinalar identificadores aos objetos armazenados, de forma que eles possam ser identificados inequivocamente interna e externamente ao repositório. Por exemplo, um identificador local (um número de chamada) e um DOI ou um ISBN; pode incluir ainda informações que descrevem o objeto, por exemplo, um resumo.

Informações de contexto – está relacionado ao fato de que muitos objetos não podem ser adequadamente interpretados sem a compreensão do seu contexto; informação que documenta o relacionamento do objeto armazenado e seu ambiente; isso inclui a motivação da criação do objeto e como ela se relaciona com outros conteúdos; circunscreve as dependências técnicas – hardware, software, linkage, etc. – inclui ainda modo de distribuição, por exemplo, via rede.

Informação de proveniência – refere-se ao princípio de que parte da integridade de um objeto depende da sua história; informação que documenta a história do objeto armazenado; pode incluir informações sobre sua fonte ou origem, sua cadeia de custódia; registra também as ações de preservação sofridas pelo objeto e seus efeitos, por exemplo: as migrações efetuadas.

Informação de fixidez – refere-se a qualquer informação que documenta mecanismos particulares de autenticação usados para assegurar que o objeto armazenado não sofreu nenhuma alteração não documentada, e que sua integridade não foi comprometida, por exemplo, assinaturas digitais e *checksums* (CCSDS, 2002, p.4-28).

Infere-se, com base Sayão (2010, p. 18), que esses grupos de informação – que formam as bases das principais estruturas de metadados de preservação – “constituem a descrição mais geral de metadados necessários para instruir a preservação de longo prazo e o uso de materiais digitais”.

Observa-se, porém, que o trabalho de Sayão (2010) apontou preservação de tecnologia, emulação, migração e encapsulamento como estratégias de preservação digital, sem mencionar os metadados, enquanto que Adu e Ngulube (2016), além das opções já apresentadas, também consideram os metadados, juntamente com repositórios confiáveis e computação em nuvem, como parte das estratégias de preservação digital.

Diante do exposto, é possível afirmar que todas as iniciativas de preservação digital presentes neste estudo incluem metadados, sejam como estratégia ou parte da estratégia – necessários para garantir o acesso às informações criadas em ambientes digitais. Há, ainda, uma tendência pela adoção de esquemas padronizados de metadados, que diminuam as barreiras terminológicas e facilitem a interoperabilidade dos repositórios, com o objetivo comum de garantir o acesso pleno aos objetos digitais no futuro.

4. ANÁLISE DOS REGISTROS DIGITAIS ARQUIVÍSTICOS DO CCJ/UFPB

Com base no conceito de registro digital arquivístico de Rogers (2015), em consonância com o que é estabelecido por Ries (2018) e Acker (2019), esta análise se baseia nos registros e/ou documentos produzidos nos ambientes digitais, em decorrência da realização de atividades estritamente institucionais do Centro de Ciências Jurídicas da Universidade Federal da Paraíba.

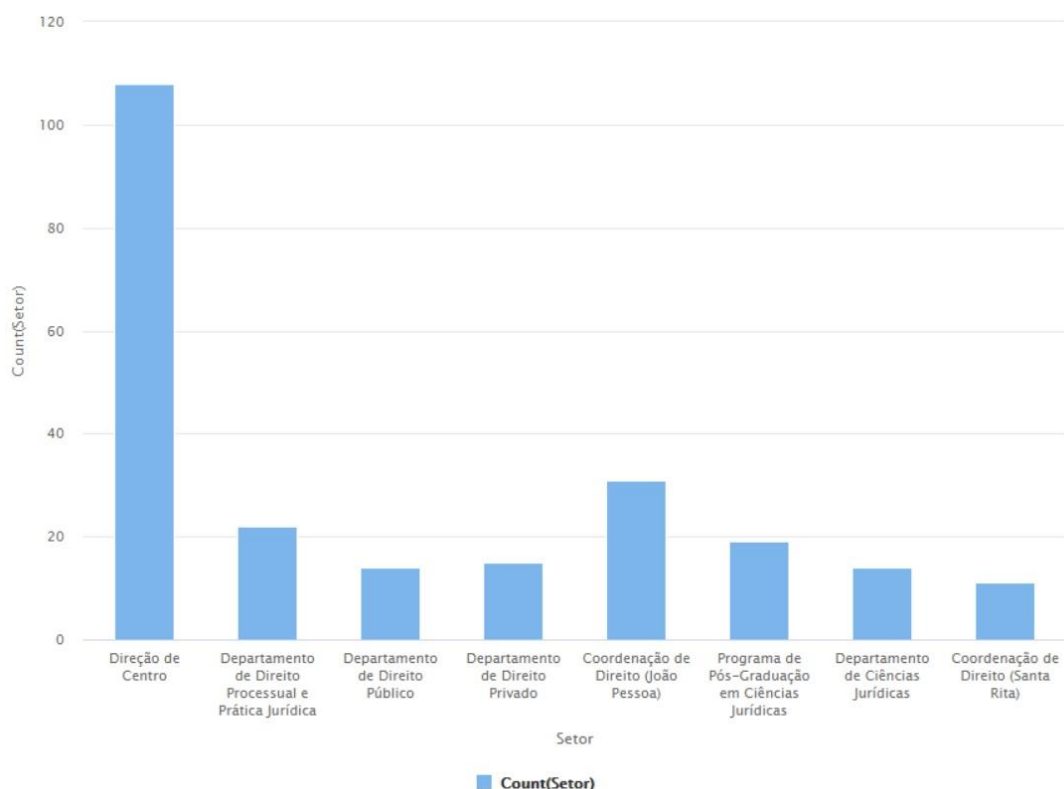
Tal conceito inclui, para além da definição de documento arquivístico digital, o conjunto de atributos e metadados necessários para a sua preservação e acesso nos diferentes ciclos de vida dos registros digitais arquivísticos.

Dessa forma, a análise será dividida em três seções. Na seção 4.1, serão analisados os conteúdos dos registros digitais arquivísticos, com o objetivo de descrevê-los e, na seção 4.2, de categorizá-los de acordo com os modelos encontrados na literatura. Já na seção 4.3 serão analisados os metadados dos registros digitais arquivísticos das instituições federais de ensino superior, com foco na preservação digital.

4.1 DESCRIÇÃO DOS REGISTROS DIGITAIS

Em atenção ao objetivo de descrever os registros digitais produzidos pelo Centro de Ciências Jurídicas da UFPB, esta pesquisa buscou fornecer as principais informações necessárias para a compreensão da natureza dos registros digitais, sua composição e os elementos necessários para a sua constituição como documento arquivístico digital.

O primeiro elemento de análise deste trabalho sobre os registros digitais arquivísticos produzidos pelo CCJ/UFPB foi o setor de origem. Tal elemento foi identificado de acordo com os critérios do diagnóstico documental e do levantamento de informações acerca da produção de tais registros (SILVA, 2018), conforme pode ser visto no gráfico 2.

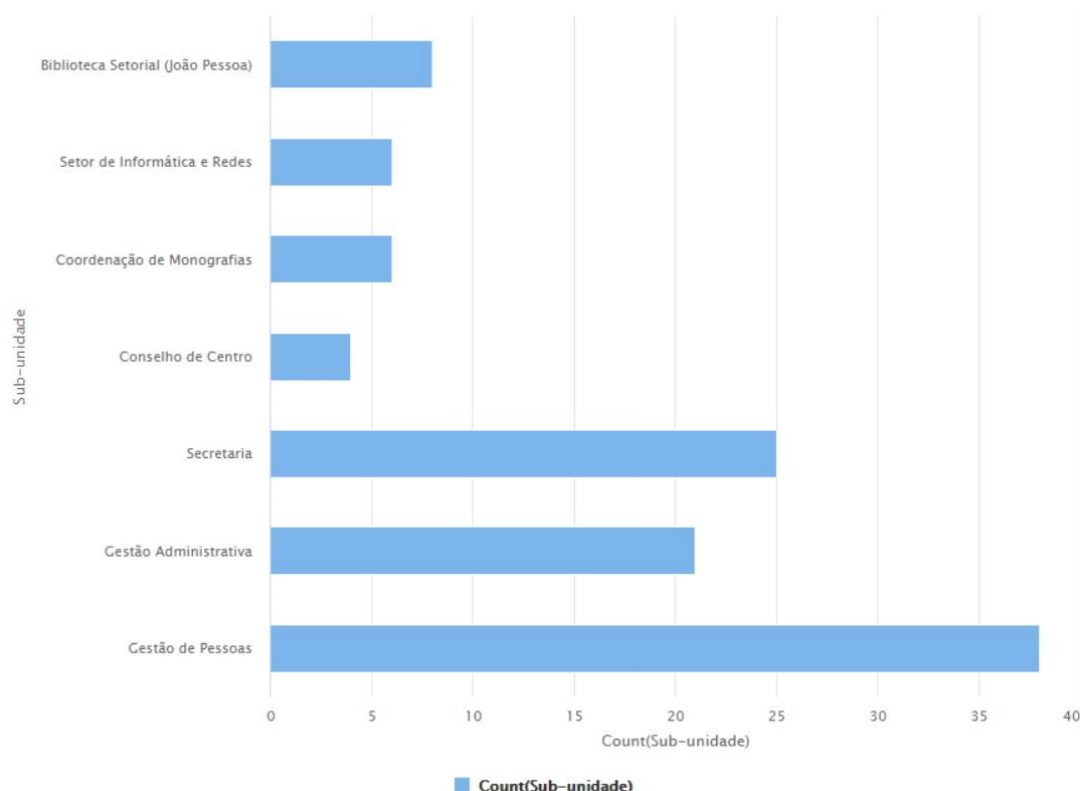
Gráfico 2 – Registros digitais arquivísticos do CCJ/UFPB por setor

Fonte: dados da pesquisa, 2020.

São oito os setores que compõem a base de dados utilizada para a análise. No gráfico é possível verificar que o setor Direção de Centro aparece como o maior produtor de registros, com um total de cento e oito registros, e a Coordenação de Direito de Santa Rita aparece como o menor produtor, com um total de onze registros.

O setor Direção de Centro se destaca em virtude de funcionar com diversos outros setores a ele subordinados, nesta análise denominados de subunidades, conforme expõe o gráfico 3.

Gráfico 3 – Registros digitais arquivísticos das subunidades da Direção do CCJ/UFPB



Fonte: dados da pesquisa, 2020.

São sete as subunidades vinculadas ao setor Direção de Centro. Dentre elas, a Gestão de Pessoas, com 38 possibilidades de criação de registros, que figura como a maior produtora, e o Conselho de Centro, com 4 registros, como o menor.

Ou seja, os oito setores, acrescidos às sete subunidades, totalizam quinze unidades produtoras de registros digitais arquivísticos do CCJ/UFPB, que, juntas, possuem duzentos e trinta e quatro possibilidades de registros digitais arquivísticos a serem produzidos de acordo com finalidades específicas.

É de extrema importância dizer que o número de registros por unidade ao qual esta análise se refere não está fundamentado na produção total de um registro, decorrente da finalidade que o gerou em determinado espaço de tempo ou ainda do número total de registros produzidos pelas unidades.

Longe disso, o presente estudo, de natureza teórica e qualitativa, define a quantidade de registros digitais **possíveis**, por unidade, derivado das

possibilidades de criação de tais registros em relação às competências institucionais definidas para cada unidade, quando do momento da realização desta pesquisa, de modo que uma análise quantitativa requer a multiplicação de cada uma dessas possibilidades de registro digital arquivístico pela incidência dos mesmos em determinado período.

A partir desse entendimento, a identificação de setores se justifica pela necessidade de relacionar as competências institucionais com a produção de registros e/ou documentos. Essa atividade permitiu o estudo das finalidades de cada unidade, com base no diagnóstico realizado por Silva (2018), e a sua relação com os registros digitais arquivísticos produzidos. Posteriormente, foi feito um recorte para os registros e/ou documentos produzidos pelas unidades, excluindo-se, portanto, aqueles recebidos por ela ou produzidos pelos servidores ali lotados no cumprimento de seus cargos de forma individual. A partir dessa etapa, o estudo se concentra nos registros criados no ambiente digital, no cumprimento das competências do Centro de Ciências Jurídicas da UFPB.

Para atender aos fins desta pesquisa, entende-se ser essencial o estudo da classificação dos documentos, considerando que a compreensão dos registros digitais arquivísticos relaciona-se fundamentalmente com a verificação da adequação de tais registros a conceitos clássicos da Arquivologia, como gênero e espécie documental, embora não se limite a estes.

O primeiro conceito, gênero documental, é definido pelo Arquivo Nacional do Brasil como

Reunião de espécies documentais que se assemelham por seus caracteres essenciais, particularmente o suporte e o formato, e que exigem processamento técnico específico e, por vezes, mediação técnica para acesso (1) como documentos audiovisuais, documentos bibliográficos, documentos cartográficos, documentos eletrônicos, documentos filmográficos, documentos iconográficos, documentos microográficos, documentos textuais (ARQUIVO NACIONAL – BRASIL, 2005, p. 99).

Nesse sentido, foram identificados no CCJ/UFPB registros digitais arquivísticos pertencentes a três gêneros documentais, conforme a seguir:

Gráfico 4 – Registros digitais do CCJ/UFPB por gênero e por setor

Fonte: dados da pesquisa, 2020.

Percebe-se que os registros digitais pertencentes ao gênero eletrônico são produzidos por todos os setores e representam a maioria dos registros produzidos pelo centro, num total de 231. Já os audiovisuais estão presentes em apenas dois setores, seguido pelo sonoro, produzido apenas pela Direção de Centro.

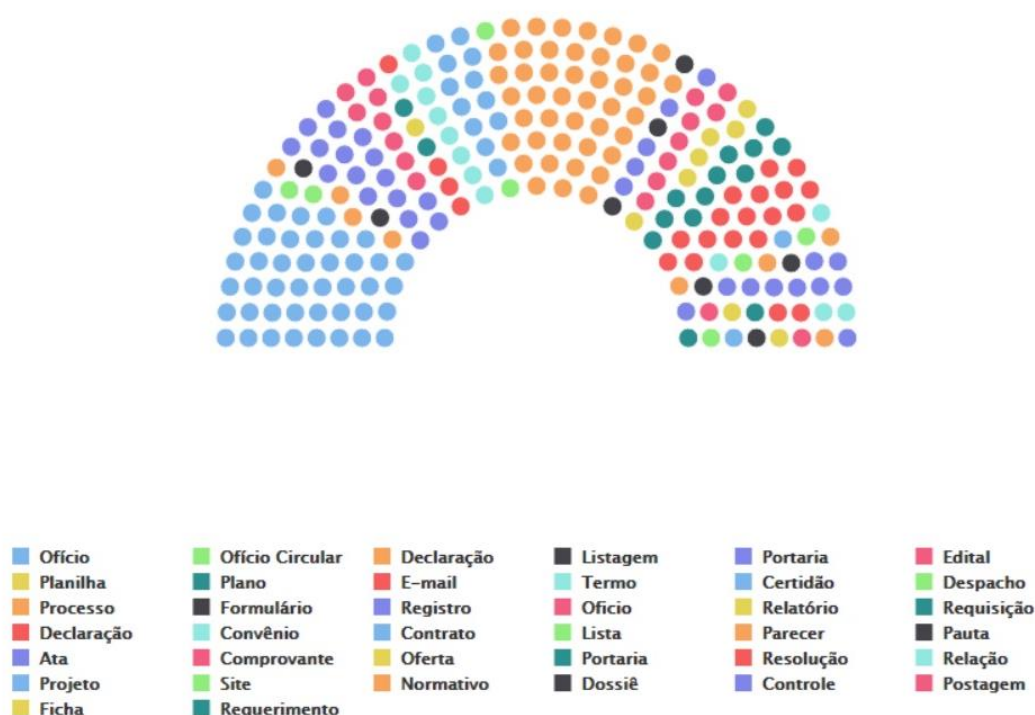
É necessário elucidar que, quando o diagnóstico foi realizado, no ano de 2018, as atividades da universidade aconteciam de forma presencial. Após o início da pandemia de COVID-19, desde março de 2020, diversas atividades passaram a ocorrer de forma remota. Como exemplo, as reuniões do conselho de centro, que eram gravadas em aplicativos de gravação de áudio, passaram a ser registradas em uma plataforma digital com múltiplas funções, que grava áudio, vídeo, mensagens trocadas pelos participantes, documentos apresentados e ainda registra como cada um dos conselheiros votou em cada uma das pautas deliberadas. Ou seja, aquele arquivo de áudio que era armazenado apenas nos computadores da direção do centro foi substituído por

um registro audiovisual complexo, que é gerado por um *software* externo, e inserido no SIPAC com metadados de descrição do contexto de produção e classificado como a espécie documental “registro”.

Espécie documental, de acordo com a literatura arquivística, é a “configuração que assume um documento de acordo com a disposição e a natureza das informações nele contidas” (CAMARGO; BELLOTTO, 2012, p. 46) ou, ainda, a “divisão de gênero documental que reúne tipos documentais por seu formato” (ARQUIVO NACIONAL – BRASIL, 2005, p. 85).

Sob esse aspecto, foram identificadas como espécie documental, trinta e oito possibilidades de registros digitais produzidos pelo CCJ/UFPB. A figura a seguir ilustra a representação de cada uma das espécies identificadas.

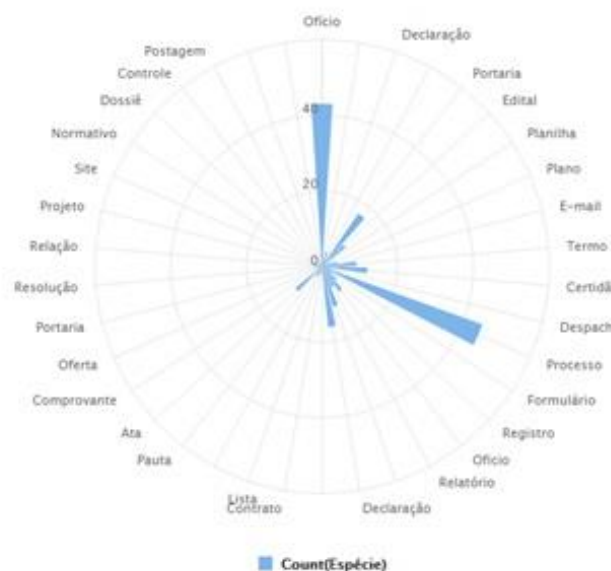
Gráfico 5 – Registros digitais do CCJ/UFPB divididos por espécie documental



Fonte: dados da pesquisa, 2020.

Das espécies identificadas, o processo é a que ocorre com maior frequência, sendo utilizado para quarenta e cinco finalidades. Já a menor incidência é a espécie documental site, que no âmbito do CCJ/UFPB é utilizado para uma única finalidade, sendo considerando por esta pesquisa como espécie.

Gráfico 6 – Principais espécies documentais dos registros digitais do CCJ/UFPB

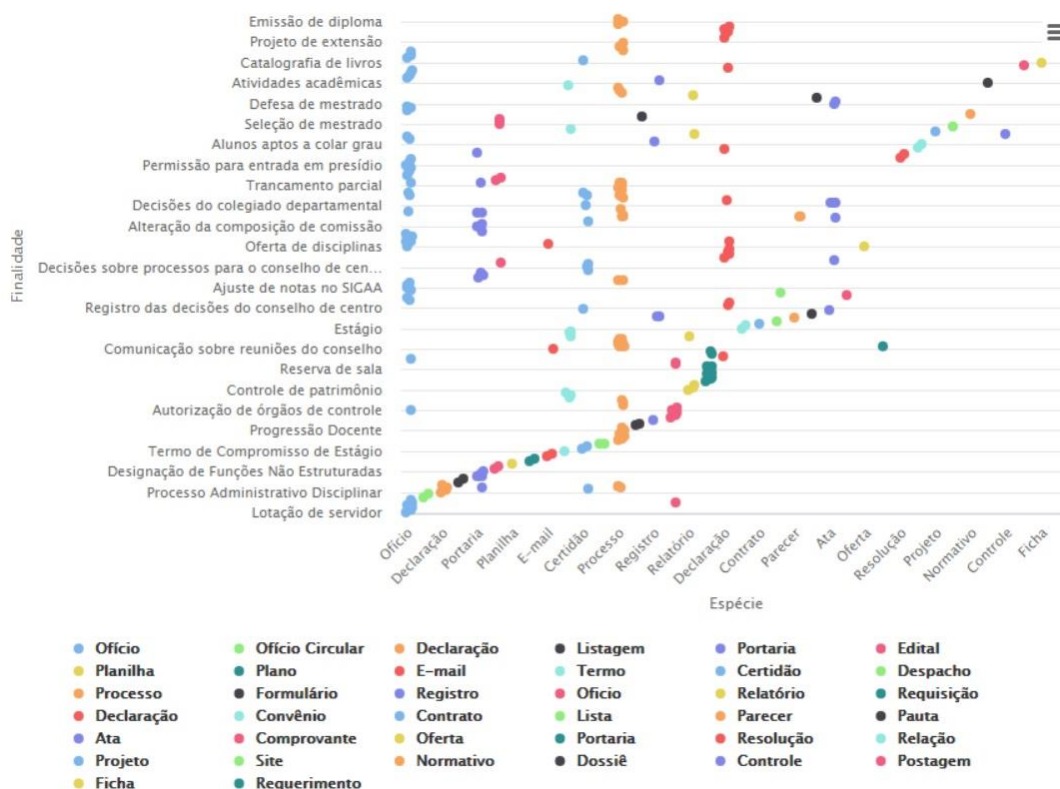


Fonte: dados da pesquisa, 2020.

O gráfico 6 ilustra as espécies documentais mais presentes dentre os registros digitais produzidos pelo CCJ/UFPB. Além do processo, destacam-se o ofício, com 43 finalidades; a portaria, com 17 finalidades; e a declaração, com 16 finalidades. Os demais têm incidência inferior a 10 finalidades.

A junção das finalidades com as espécies pode ser melhor visualizada por meio do gráfico 7.

Gráfico 7 – Espécies documentais x finalidades dos registros digitais do CCJ/UFPB



Fonte: dados da pesquisa, 2020.

Essa ilustração permite verificar que a predominância de algumas espécies documentais é justificada pela sua utilização para diferentes finalidades. Na Arquivologia, a junção da espécie documental com a sua finalidade dá origem ao conceito de tipo documental. Isto é, tipo documental é “a configuração que assume uma espécie documental de acordo com a atividade que a gerou” (CAMARGO; BELLOTTO, 2012, p. 80). Entretanto, em conformidade com os objetivos deste trabalho, embora tenham sido coletadas as finalidades dos registros, tal estudo buscou identificar se o registro é produzido para a realização das atividades fim ou das atividades meio da UFPB. Tal divisão tem como base o fato de a universidade utilizar um sistema para a produção de registros das atividades fim de forma preponderante, o SIGAA, e outro para as atividades meio, o SIPAC.

Nesse sentido, entende-se como atividades fim aquelas que constituem os objetivos constitucionais da universidade, a saber, o ensino, a pesquisa e a extensão; como atividades-meio, aquelas realizadas para a gestão da

instituição, tais como administração de pessoas, de materiais, de orçamento, entre outros. A divisão dos registros de acordo com as atividades pode ser visualizada no gráfico 8.

Gráfico 8 – Registros digitais do CCJ/UFPB por atividade



Fonte: dados da pesquisa, 2020.

É possível perceber que os registros digitais decorrentes da atividade meio prevalecem. Cabe aqui explicar que alguns registros, mesmo após identificada a espécie e a finalidade, ainda assim, poderão figurar em diferentes atividades. Por exemplo, a espécie despacho, vinculada à finalidade de instrução processual, tanto poderá ser referente a um processo de dispensa de disciplina (atividade fim), como a um processo de progressão funcional (atividade meio).

A predominância de registros das atividades meio nesta análise se deve, dentre outros fatores, ao recorte da pesquisa para os registros digitais criados pelas unidades no cumprimento das suas funções em detrimento dos registros produzidos pelos servidores em razão dos seus cargos. Por exemplo,

A plataforma “site” aparece como a segunda maior, com 5% dos registros. Destaca-se que os registros descritos como gerados por meio de “site” são todos produzidos nos *websites* da UFPB, tais como o *site* para criação de chamados da Superintendência de Tecnologia da Informação, o *site* para reserva de veículos e salas do CCJ/UFPB, dentre outros.

Salientamos, a existência das plataformas Gravador/SIPAC e Câmera/SIPAC, como a denominação utilizada para descrever o meio através do qual são criados os registros sonoros das reuniões do Conselho de Centro e os registros audiovisuais das entrevistas das seleções do Programa de Pós-graduação em Ciências Jurídicas. Tais documentos são criados com gravador e câmera, respectivamente, e posteriormente inseridos no sistema SIPAC.

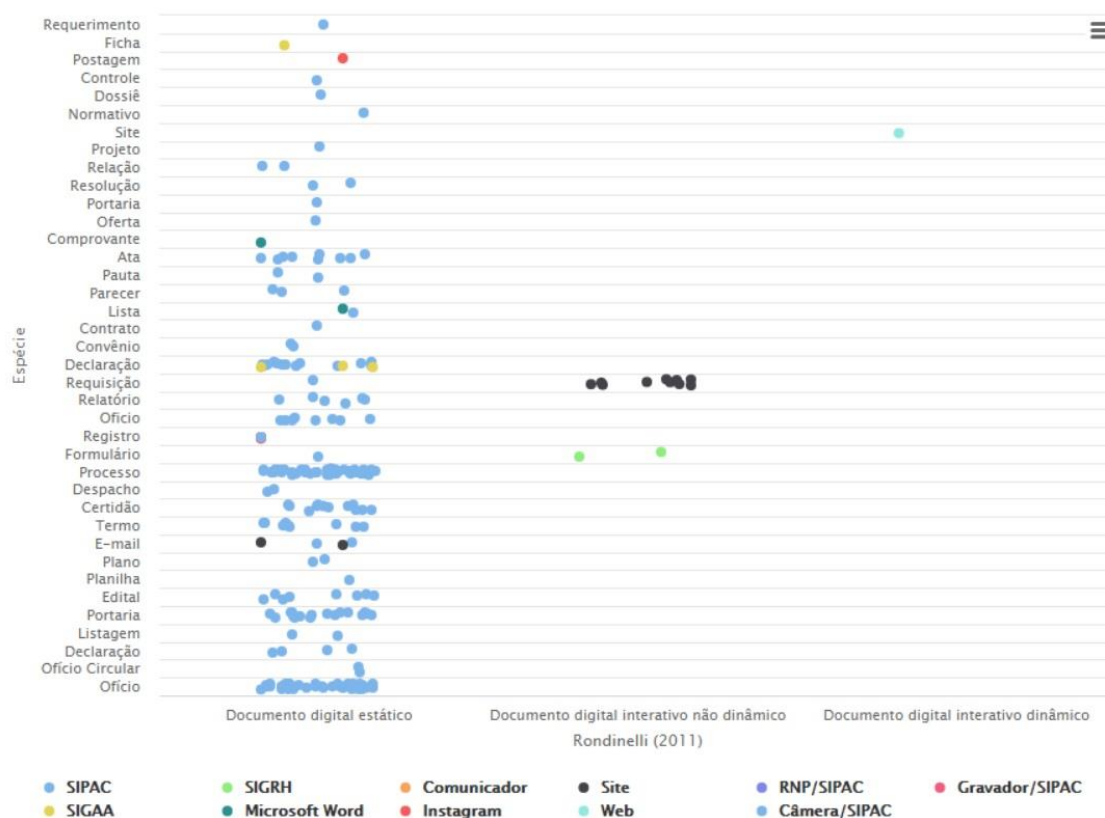
Destacamos, ainda, a plataforma *Instagram*, que já é utilizada pela Coordenação da Graduação em Direito de João Pessoa como ferramenta de comunicação com a comunidade, além de ser utilizada por projetos de extensão do Centro.

4.2 CATEGORIZAÇÃO DOS REGISTROS DIGITAIS

Visando o cumprimento do objetivo de categorizar os registros digitais arquivísticos produzidos pelo CCJ/UFPB, passaremos à análise da possibilidade de inserção dos registros digitais produzidos pelo Centro de Ciências Jurídicas da UFPB, nas divisões existentes na literatura para categorizá-los.

O primeiro modelo aplicado foi da autora Rondinelli (2011). De acordo com o estudo, os registros digitais produzidos pelo CCJ/UFPB dividem-se conforme o gráfico 10 abaixo:

Gráfico 10 – Registros digitais do CCJ/UFPB – espécie documental x modelo de Rondinelli (2011)



Fonte: dados da pesquisa (2020) com base em Rondinelli (2011).

O gráfico relaciona a categorização proposta por Rondinelli (2011) com as espécies documentais e destaca as cores referentes às plataformas utilizadas. Percebe-se que duzentos e vinte e um registros se enquadram como documento digital estático (o que equivale a noventa e cinco por cento da amostra), sendo a maioria processos, ofícios e declarações produzidos nos sistemas SIPAC e SIGAA, enquanto que doze registros (cinco por cento) são considerados documentos digitais interativos não dinâmicos, com predominância para requisições realizadas por meio de *site* ou do sistema SIGRH. Apenas um registro configura documento digital interativo dinâmico, que foi a categorização adotada para o *site* específico do CCJ/UFPB enquanto espécie documental.

Já de acordo com a proposta de Katuu (2016), os registros digitais arquivísticos do CCJ/UFPB são predominantemente registros gerados por sistemas de informações de negócios, conforme o gráfico 11.

Gráfico 11 – Registros digitais do CCJ/UFPB – gênero documental x modelo de categorias de Katuu (2016)



Fonte: dados da pesquisa (2020) com base em Katuu (2016).

Foi feita a análise das categorias propostas por Katuu (2016) em conjunto com as possibilidades de gênero documental existentes. Verifica-se que, de acordo com o modelo de Katuu (2016), noventa por cento dos registros digitais do CCJ/UFPB são classificados como registros gerados por sistemas de informações de negócios (com predominância do gênero eletrônico), seguidos de seis por cento constituídos de registros em ambiente *online* (no qual existem registros tanto do gênero eletrônico como audiovisual) e menos de um por cento constituem-se documentos criados usando aplicativos de escritório e sistemas de comunicação digital. Destaca-se, nessa categorização, que o registro sonoro das reuniões do conselho de centro até fevereiro de 2020, criado por meio do gravador de áudio do *Windows*, é considerado um documento criado usando aplicativos de escritório.

Chama a atenção o fato de um dos registros não ter sido contemplado pelo modelo do autor Katuu (2016). É o caso do registro audiovisual (vídeo),

produzido pelo Programa de Pós-Graduação em Ciências Jurídicas, com a finalidade de documentar a entrevista dos candidatos da seleção dos cursos de mestrado e doutorado da instituição. Esse mesmo registro também já é criado para outras finalidades institucionais com ocorrência frequente após a pandemia de COVID-19. De acordo com a análise, tal registro se configura como uma nova possibilidade de documentar atividades, que ainda requer mais investigações e pesquisas que contribuam na sua compreensão.

Há ainda a proposta de Zhao *et al* (2019), em que os registros do CCJ/UFPB se dividem de acordo com o gráfico 12.

Gráfico 12 – Registros digitais do CCJ/UFPB por atividade, com base em Zhao *et al* (2019)



Fonte: dados da pesquisa (2020) com base em Zhao *et al* (2019).

Prevalece o registro considerado como documento digital, constituído por noventa e sete por cento da amostra e ocorrendo como parte de atividades meio e fim, isoladamente ou de forma simultânea, seguido do *e-mail*, vídeo,

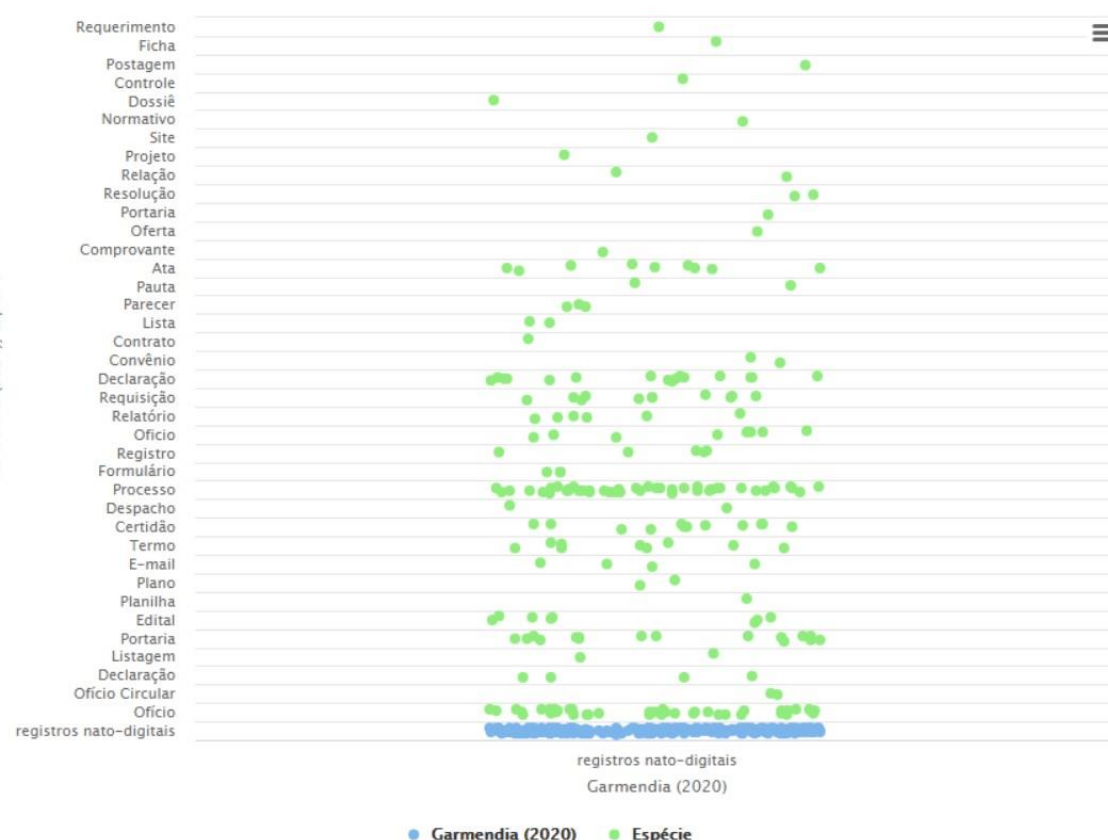
outros recursos de rede e áudio, todos com representatividade inferior a um por cento da amostra.

Cabe destacar, conforme gráfico 12, que, com exceção do registro em vídeo, presente tanto na atividade fim como na atividade meio, os registros de *e-mail*, áudio e recursos de rede estão presentes apenas na atividade meio.

Enfatizamos, ainda, que a pesquisa foi realizada antes da paralisação das atividades da UFPB, em decorrência da pandemia global de coronavírus. Considerando o início das atividades de trabalho remoto, é possível que haja um aumento na produção de registros em áudio e vídeo, e outros recursos de rede, em virtude da mudança de conjuntura nesse cenário.

Por fim, de acordo com os estudos de Garmendia (2020), os registros produzidos pelo CCJ/UFPB, conforme o recorte desta pesquisa, são todos considerados registros nato-digitais, conforme o gráfico 13.

Gráfico 13 – Registros digitais do CCJ/UFPB com base em Garmendia (2020)



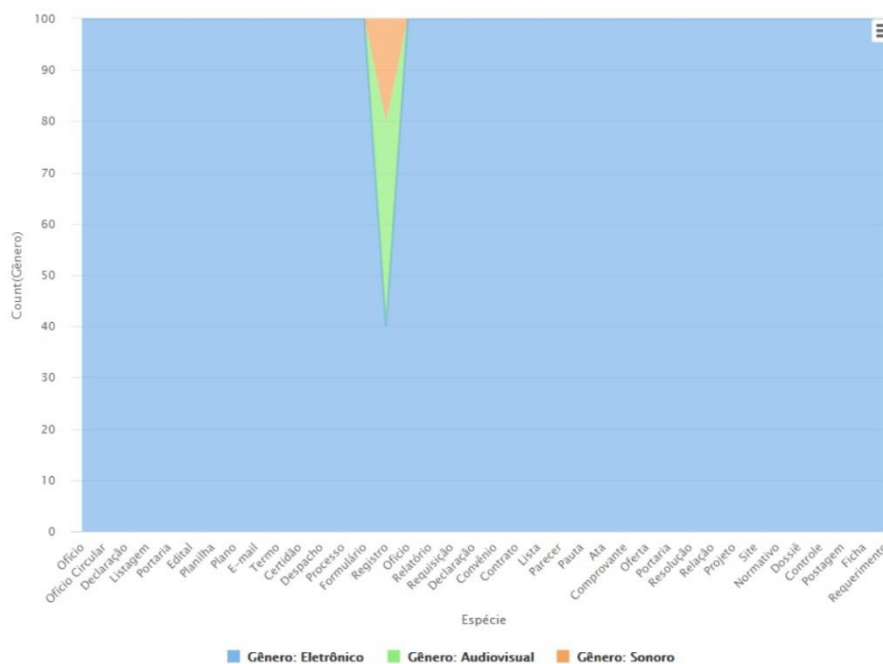
Fonte: dados da pesquisa (2020) com base em Garmendia (2020).

O cenário apresentado no gráfico que analisa a relação dos registros do CCJ/UFPB por espécie documental com as divisões proposta por Garmendia (2020), independente da espécie documental, não inclui nenhuma categoria além dos registros nato digitais, em virtude do recorte desta pesquisa ter se limitado aos documentos produzidos pelos setores utilizando os sistemas institucionais.

Entretanto, cabe ressaltar que, em alguns casos, documentos que deveriam ser gerados a partir do sistema oficial – como registros nato digitais, são produzidos fora do *software*, utilizando um aplicativo de escritório, por exemplo; em alguns casos são impressos, posteriormente digitalizados, e só então são inseridos nos sistemas de informação da universidade. Destaca-se que, embora essa prática exista, não é a recomendação da instituição para a produção dos seus registros, esta pesquisa se limitou a considerar a criação de registros nato-digitais, mas reconhece que parte dos registros digitais criados pelo CCJ/UFPB constitui-se de documentos provenientes de digitalização. Esse também se constitui como um dos desafios para a gestão arquivística dos registros digitais: a padronização da produção.

Em continuidade, foi feita a análise da relação dos registros digitais arquivísticos produzidos pelo CCJ/UFPB com os modelos disponíveis na literatura, destacando-se a constatação demonstrada no gráfico 14.

Gráfico 14 – A espécie documental “registro” e os gêneros documentais no caso do CCJ/UFPB



Fonte: dados da pesquisa, 2020.

Na visualização do gráfico de área que analisa as espécies e os gêneros documentais de forma associada, é possível verificar que a espécie documental “registro” se destaca como a única que possui participação em todos os gêneros documentais – eletrônico (informático), audiovisual e sonoro.

Ao considerar os conceitos da Arquivologia (CAMARGO; BELLOTTO, 1996; ARQUIVO NACIONAL – BRASIL, 2005) que definem espécie documental como “divisão de gênero de acordo com formato” e gênero documental como “reunião de espécies documentais que se assemelham por seus caracteres essenciais, particularmente o suporte e o formato”, deduz-se, com base em Rogers (2015), que os registros, sejam eles sonoros, filmográficos, audiovisuais ou de qualquer outro gênero, ao servirem de prova das atividades meio e fim do CCJ/UFPB em ambientes digitais, embora compartilhem o suporte eletrônico, por serem gerados, armazenados e exibidos em diferentes formatos e, portanto, necessitem de diferentes estratégias de gestão e preservação, carecem de estudos que ampliem o entendimento sobre a sua classificação.

Isto é, percebe-se que o **registro tanto assume a definição de entidade conceitual** composta de diversos elementos, dentre eles, os metadados, **como também pode assumir a função de espécie documental**. Tal constatação reforça o entendimento de que o termo registro ainda deve ser objeto de outras pesquisas que se dediquem ao aprofundamento teórico e epistemológico da sua natureza e amplitude.

De acordo com esta pesquisa, as classificações tradicionais da Arquivologia, para fins desta análise – gênero e espécie documentais –, **não reúnem os elementos suficientes para a compreensão da complexidade envolvida na criação dos registros, dadas as múltiplas possibilidades de geração, armazenamento e apresentação de tais documentos**.

Não é intenção deste trabalho estabelecer um conceito novo dentro das classificações existentes na Arquivologia ou definir o termo registro como uma nova categoria de classificação, porém, ao verificar a criação de registros em diferentes formatos e finalidades pelo CCJ/UFPB, considera-se pertinente a realização de pesquisas futuras que consigam avaliar a necessidade de ampliação do conceito de espécie e/ou de uma reflexão acerca das possibilidades teóricas atinentes ao termo registro, no âmbito da Arquivologia.

4.3 ANÁLISE DOS METADADOS DE PRESERVAÇÃO DIGITAL

A última análise realizada atende ao objetivo de identificação dos metadados necessários para a preservação dos registros digitais criados pelo CCJ, da Universidade Federal da Paraíba.

Com o intuito de ampliar a abrangência do estudo, e criar um ponto que possibilite uma análise comparativa, foram verificados os metadados de preservação digital de um conjunto de universidades federais, do qual a UFPB faz parte.

Para isso, foi elaborado um instrumento de coleta de dados, a partir de uma reflexão baseada no diálogo entre os trabalhos de Rogers (2015) e Sayão (2010), considerados relevantes para os fins desta pesquisa.

A combinação do conceito de Rogers (2015) sobre os registros digitais arquivísticos (que, além de relacioná-los aos dados de conteúdo gerados pelo usuário, os explicam como constituídos de quatro conjuntos diferentes de

metadados) à reflexão de Sayão (2010) acerca do OAIS, considerado o mais importante modelo de requisitos de preservação de objetos digitais da atualidade, nos levou à reflexão constante no quadro 9.

Quadro 9 – Metadados para preservação digital de registros digitais arquivísticos

(1) metadados de identificação de origem e localização;	infraestrutura conceitual que descreve o ambiente [...]
(2) metadados de aparência e desempenho;	[...] as interfaces externas [...]
(3) metadados de descrição automática pelo sistema operacional e/ou de sistema de arquivos e	[...] os componentes funcionais e os objetos de informação, associados com um sistema responsável pela preservação de longo prazo de materiais digitais
(4) metadados de descrição manual gerados pelo usuário.	
(ROGERS, 2015, p. 16,18).	(SAYÃO, 2010, p. 13).

Fonte: elaborado pelo autor (2020) a partir de Rogers (2015) e Sayão (2010).

O quadro acima reflete de forma conceitual os metadados necessários para a preservação digital dos registros digitais arquivísticos e, com base nele, foram elaboradas as questões enviadas às universidades federais que utilizam o SIPAC, conforme questionário abaixo:

Quadro 10 – Questionário sobre metadados de preservação digital para as universidades federais que utilizam o SIPAC

METADADOS DE PRESERVAÇÃO DIGITAL DE FORMA GERAL

- 1) Quais metadados são criados e armazenados pela sua instituição para preservação dos registros e documentos produzidos nos sistemas SIPAC e SIGAA (ou nos sistemas de produção de documentos acadêmicos e administrativos)?

METADADOS GERADOS DE FORMA AUTOMÁTICA

- 2) A sua instituição armazena metadados que identificam a origem e a localização dos usuários que produzem os registros e documentos? Caso sim, quais metadados são armazenados?
- 3) A sua instituição armazena metadados que diferenciem os registros e documentos gerados em ambiente digital (nato digitais) daqueles convertidos para o meio eletrônico (digitalizados)? Caso sim, quais metadados são armazenados em cada caso?
- 4) A sua instituição armazena metadados que descrevam o dispositivo e o sistema operacional no qual os registros e documentos foram produzidos? Caso sim, quais metadados são armazenados a respeito do dispositivo e do sistema operacional?
- 5) A sua instituição armazena metadados que descrevam os dados constantes nos registros e documentos criados/inseridos nos sistemas SIGAA/SIPAC (ou nos sistemas de produção de documentos acadêmicos e administrativos)? Caso sim, quais são os metadados armazenados para descrição dos dados?

METADADOS INSERIDOS MANUALMENTE PELO USUÁRIO

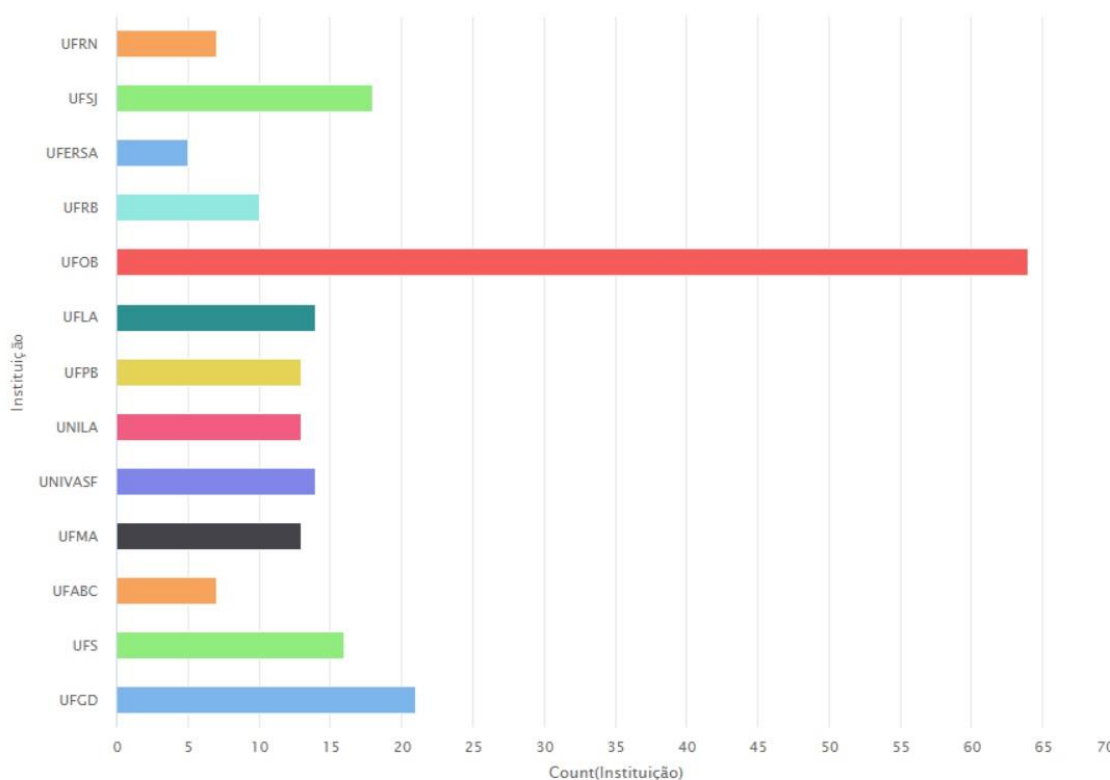
- 6) Existem metadados que descrevam os dados constantes nos registros e documentos criados/inseridos nos sistemas SIGAA/SIPAC (ou nos sistemas de produção de documentos acadêmicos e administrativos) que são incluídos pelo usuário de forma manual? Caso sim, quais são os metadados armazenados para descrição dos dados inseridos manualmente pelo usuário?

Fonte: elaborado pelo autor (2020).

O instrumento acima foi respondido de forma satisfatória para os objetivos deste trabalho por treze universidades federais que utilizam o SIPAC – Módulo Protocolo, por intermédio da Plataforma FalaB sistema oficial da Controladoria Geral da União para gestão do acesso à informação para órgãos e entidades da administração pública federal brasileira.

As respostas obtidas serão apresentadas a seguir, e comparadas aos requisitos preconizados pela literatura para a composição dos metadados necessários para a preservação digital dos registros digitais arquivísticos.

Passaremos agora à análise de cada um dos quesitos. Primeiramente, como resposta ao questionamento do nome do setor responsável pelas respostas, temos sua exibição no gráfico 15.

Gráfico 16 – Metadados de preservação digital nas IFES que utilizam o SIPAC

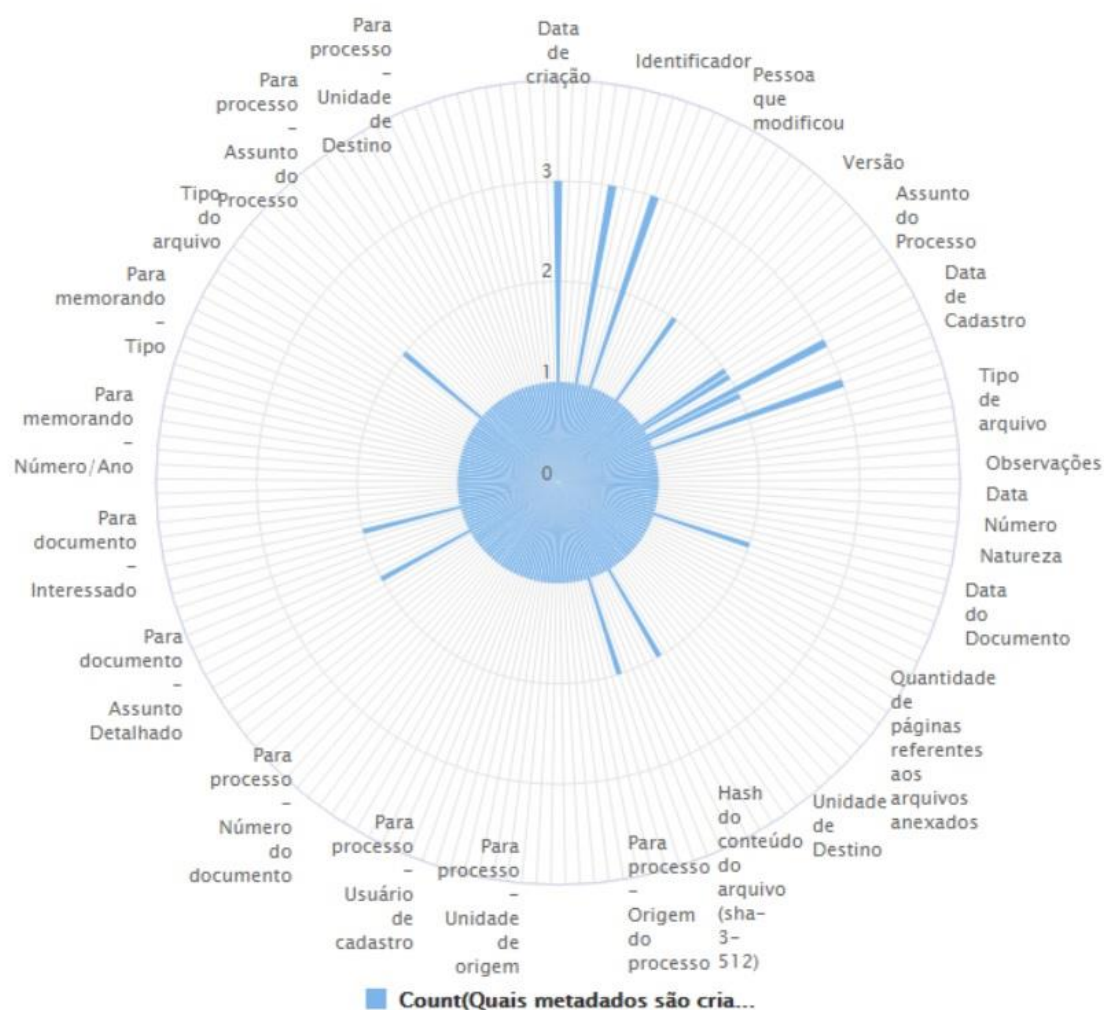
Fonte: dados da pesquisa, 2020.

O gráfico acima demonstra a amplitude da quantidade de metadados informados pelas universidades para a preservação dos seus registros digitais. Com destaques para o maior e menor resultado, respectivamente, a Universidade Federal do Oeste da Bahia (UFOB), com 64 (sessenta e quatro) tipos de metadados, e a Universidade Federal do Semiárido (UFERSA), com 5 (cinco) tipos de metadados. Cabe destacar que a UFPB informou que armazena 13 tipos de metadados, estando, portanto, abaixo da média das outras instituições em relação ao número de metadados de preservação digital armazenados.

Ressalta-se que a quantidade de campos de metadados indicados como resposta ao primeiro questionamento desta pesquisa não é um indicador da qualidade dos metadados de preservação digital das IFES participantes. A análise qualitativa a seguir oferecerá outros elementos para a compreensão da aderência dessas instituições aos preceitos de preservação digital por meio de metadados.

Dessa forma, para além da análise quantitativa, buscou-se compreender quais são os metadados informados pelas IFES para a preservação dos seus registros digitais. Sem realizar nenhuma alteração nos termos utilizados pelos respondentes, o resultado da questão 1 pode ser visualizado no gráfico 17.

Gráfico 17 – Metadados de preservação digital de registros digitais arquivísticos das IFES que utilizam o SIPAC/Módulo Protocolo



Fonte: dados da pesquisa, 2020.

Os tipos de metadados com incidência maior do que um foram destacados. Os destaques podem ser visualizados no quadro abaixo:

Quadro 11 – Metadados de preservação digital de registros digitais arquivísticos das IFES que utilizam o SIPAC/Módulo Protocolo

Grupo 1 - Três ocorrências	Grupo 2 - Duas ocorrências
Assunto Data de criação Identificador Número de páginas Unidade de origem	Assunto detalhado Data de cadastro Data do documento Data do <i>upload</i> Nome do arquivo Para documento – Assunto detalhado Para documento – Observação Tamanho Tamanho do arquivo Tipo de documento

Fonte: dados da pesquisa, 2020.

Inicialmente, analisaremos os campos de metadados do grupo um, que foram aqueles com três ocorrências, destacando os desafios relacionados à preservação digital para cada um deles. O primeiro termo, “**assunto**”, é utilizado por algumas instituições como o campo por meio do qual o usuário escolhe o código de classificação de documentos estabelecido pelo Conselho Nacional de Arquivos.

O poder de escolher o código do assunto exige do usuário um conhecimento abrangente dos instrumentos de gestão de documentos das atividades meio e fim da administração pública. Tal código é o fundamento legal para o estabelecimento da temporalidade do documento no sistema, ou seja, a definição do seu ciclo de vida. Do mesmo modo, essa escolha permite a definição da destinação do documento, isto é, a decisão a respeito da sua guarda temporária ou permanente. São vários os problemas decorrentes desse metadado, caso ele não seja adequadamente utilizado, seja pela falta de treinamento do usuário ou de adequação do sistema. Passemos a analisar cada um deles.

Primeiro, a escolha do código só irá garantir o cumprimento da temporalidade e destinação dos documentos se tal função estiver programada no sistema, de forma que ter a informação do código é importante, porém, não é suficiente para que a gestão do documento ocorra.

Segundo, permitir ao usuário escolher dentre um rol de centenas de códigos, em alguns casos com descrições com pouca ou nenhuma aproximação com a finalidade do documento, coloca em risco a finalidade de

gerir tais documentos, pois aumenta a chance de erros de classificação, podendo um documento de guarda permanente ser definido como de guarda temporária e vice-versa. Além disso, com ou sem um sistema que viabilize a transferência dos documentos do ambiente de gestão para o ambiente de preservação, tal cenário colabora para o retrabalho por parte dos profissionais de arquivo que posteriormente precisarão conferir tais classificações.

Ou seja, está posto como desafio para os profissionais de arquivologia e tecnologia da informação, o desenvolvimento de ferramenta capaz de realizar a classificação de forma automatizada, de maneira que o usuário não precise escolher o código de classificação, mas ao escolher o tipo de documento, com base em formulário pré-estabelecido, dicionário de termos e vocabulário controlado previamente estabelecido na programação do sistema, que cada tipo já esteja vinculado ao código de classificação do assunto correspondente, aumentando a eficiência já na fase da produção dos documentos.

Além das dificuldades já mencionadas, nem toda universidade usa esse campo “assunto” para o código de assunto do CONARQ. Existem ainda os casos em que esse campo é utilizado de forma indiscriminada, ampliando ainda mais os riscos à preservação desses registros.

Em se tratando do item “**data de criação**”, cuja importância é fundamental para a gestão do ciclo de vida de qualquer objeto digital, é possível perceber que, em algumas IFES, ele não é utilizado de modo a favorecer a preservação digital, haja vista que não é feita diferenciação entre o metadado gerado para os documentos natos digitais daqueles gerados para os documentos digitalizados. Isso quer dizer que, se um documento digitalizado for recebido pela instituição em uma determinada data, a ausência de campo específico para a separação da data de recebimento da data de criação do documento original pode facilitar a inserção de informação incorreta, ou seja, um documento criado em dezembro que foi recebido em janeiro pode ser descrito pelo usuário como criado em janeiro, em virtude de ter sido recebido naquela data, de modo que a gestão da sua temporalidade irá ocorrer de forma equivocada.

Além disso, a ausência de elementos que diferenciem os metadados de documentos natos digitais daqueles que são digitalizados fragiliza o ambiente de preservação digital, uma vez que trata todos os registros como se fossem

originais, quando os digitalizados deveriam ser tratados como cópias ou representantes digitais, com requisitos de autenticidade e validade jurídica absolutamente diferenciados.

Faz-se necessário explanar que, dentre as instituições pesquisadas, já existem aquelas que diferenciam os documentos entre nato digitais e digitalizados por meio de booleanos que marcam como “virtual” os documentos gerados em ambiente externo ao sistema, e como “escrito” aqueles produzidos diretamente no SIPAC. Essa análise será apresentada de forma detalhada na apresentação das respostas ao quesito três do instrumento de coleta de dados.

Passemos à análise do terceiro tipo de metadado indicado com maior frequência: o campo “**identificador**”. Verifica-se que nenhuma das instituições apresentou uma descrição para esse campo, embora outros campos tenham sido descritos. Isto é, as respostas apontam ser consenso entre os responsáveis pela criação dos metadados a existência de um campo com esta nomenclatura. Infere-se, portanto, que o identificador se caracteriza por uma numeração com a função de identificar cada registro criado de forma individual dentro da estrutura do sistema. Contudo, tem-se que fazer o destaque de que as instituições se utilizam de termos diferentes para representar a mesma informação no que se refere ao nome dos metadados. Por exemplo: enquanto a UFRN utiliza o termo “identificador persistente”, UFGD, UFMA e UNIVASF utilizam o termo “identificador”.

Essa situação se repete com inúmeros outros termos, e a falta de padronização na utilização de termos para finalidades semelhantes representa uma ameaça para a interoperabilidade dos sistemas e, conseqüentemente para a preservação digital dos registros. Constitui-se, portanto, como um desafio para a área de Arquivologia oferecer contribuições no sentido de diminuir a assimetria terminológica dos tipos de metadados de preservação digital do SIPAC.

O quarto metadado com a maior incidência, o campo “**número de páginas**”, apresenta desafios próprios. Primeiramente, o fato de os documentos escritos diretamente no sistema serem exibidos de forma contínua, isto é, sem a necessidade de paginação, torna tal campo desnecessário para os documentos nato digitais. Em segundo lugar, algumas instituições adotam a digitalização dos versos das páginas dos documentos

mesmo quando eles estão em branco, enquanto outras só digitalizam as páginas com conteúdo, de modo que, nos casos de processos administrativos digitalizados, dados de menor volume contidos em partes não significativas das páginas capturadas podem ser excluídos de forma errônea. Exatamente por isso, conforme ocorre com o campo “data de criação”, a utilização de termos comuns para registros nato digitais daqueles que são digitalizados pode acarretar prejuízos para a preservação dos registros em ambiente digital.

O quinto elemento de metadado com maior incidência entre as IFES é o campo “**unidade de origem**”, cuja função é indicar o setor de lotação do usuário responsável pela criação ou recebimento do documento que dá origem ao registro digital. Esse campo é passível de reflexões sobre a vinculação do setor de lotação do servidor no momento da criação do registro. Ou seja, se um servidor criou um registro num determinado ano em que estava lotado no setor X, mas após alguns anos já não se encontra mais lotado naquele setor, para fins de manutenção da organicidade e do contexto correto de produção daquele documento, é importante que esse campo de metadado fixe a unidade de lotação do servidor para aquela data específica, de maneira que a preservação dessa informação não seja alterada ou comprometida, seja pela movimentação do servidor ou pela alteração no nome do setor em possíveis reformas administrativas.

No segundo grupo de campos de metadados, no qual aparecem aqueles com duas ocorrências, iremos realizar a análise de forma coletiva para alguns elementos. O primeiro item, “**assunto detalhado**”, se configura como um campo de livre escrita para o usuário descrever o conteúdo ou a finalidade do documento. Entretanto, embora esse item possa ser considerado importante para facilitar a recuperação das informações, ele não é considerado um item obrigatório para a criação de um registro. Ou seja, sem uma descrição obrigatória e com a falta de classificação automática já mencionada na análise do campo “assunto”, as informações de contexto do documento podem não ser adequadamente preservadas.

São mencionados ainda os campos “**assunto detalhado – para documento**” e “**observação – para documento**”. O primeiro cumpre exatamente a mesma função do campo “assunto detalhado”, porém com a vinculação aos documentos avulsos. O segundo é destinado a apresentar

informações adicionais ao assunto detalhado, também no caso de documentos avulsos, sendo, na avaliação desta pesquisa, um item de metadado redundante em relação aos que já foram apresentados.

Percebe-se, de acordo com as respostas recebidas das IFES, que em algumas instituições os metadados associados a um processo administrativo são tratados de forma diferenciada dos metadados associados a documentos avulsos, por isso a expressão “para documento” após o termo que nomeia o campo de metadado. Nas instituições que adotam essa abordagem, também existe um campo “para processo” e, em algumas, outros campos “para memorando” e “para ofício”.

Os dados demonstram ser uma prática comum entre as IFES participantes desta pesquisa: separar os processos administrativos dos documentos avulsos, de modo que sejam criados e mantidos metadados diferentes para cada espécie. E ainda, a criação de um metadado específico para memorandos, que após alteração no manual de redação oficial da presidência da república, em alguns órgãos – inclusive a UFPB –, passou a ser substituído pelo ofício.

Os estudos realizados por esta pesquisa apontam para a necessidade de metadados que preservem o contexto e a descrição dos documentos. Entretanto, a criação de campos específicos para diferentes espécies não parece ser o melhor caminho para essa demanda.

Embora não seja intenção deste trabalho propor um modelo de metadados descritivos para a preservação digital, a literatura indica que um único campo intitulado “assunto” poderia ser utilizado para diferentes espécies documentais sem prejuízo de finalidade. Observa-se, no caso das IFES analisadas, que o estabelecimento de metadados com diferentes termos para processos administrativos, documentos avulsos e ofício, além de reunir numa mesma seara todos os documentos que não sejam processos administrativos ou ofícios, ocasionam duas dificuldades: limitam a três as possibilidades de descrição, e multiplicam por três a necessidade de metadados para a mesma finalidade, aumentando o retrabalho, diminuindo as possibilidades de interoperabilidade e, conseqüentemente, de eficiência.

Uma alternativa possível para a diminuição dos problemas de ordem terminológica resultantes da diferente aplicação de termos pelas instituições

usuárias do SIPAC seria a criação de um dicionário unificado no sistema, porém, tal ação requer um esforço interinstitucional e multiprofissional de padronização.

Destaque-se que a sugestão de unificação dos metadados relacionados à descrição do assunto para diferentes espécies não possui relação com a existência de interfaces diferentes para produção de processos e documentos avulsos. Tal sugestão se limita à gestão dos termos utilizados para o armazenamento dos metadados de preservação digital.

Dando continuidade, os itens “**data de cadastro**”, “**data do documento**” e “**data do upload**” reforçam a falta de padronização existente para a utilização de campos de metadados nas IFES. Insistimos que tal situação se amplia quando se percebe que algumas instituições não realizam a separação entre os metadados aplicados aos documentos digitalizados daqueles produzidos no ambiente digital, e outras fazem uma separação apenas por espécie documental como já mencionado.

Por exemplo, enquanto a UFGD possui um campo “data de criação” para documentos nato digitais e outro “data de inserção” para documentos digitalizados – cenário bastante coeso com a proposta da literatura –, a UFS possui um campo “data de autuação” para a espécie documental processo e outro “data de criação” para outras espécies documentais, porém, sem a distinção nato digital e digitalizado – cenário inadequado de acordo com a análise dos autores utilizados por esta pesquisa.

Colocar em um mesmo campo de metadado a data que um documento foi criado no sistema, independente da espécie, e a data em que ele foi recebido de forma digitalizada ou ainda a data em que ele foi inserido no sistema, configura-se como um equívoco para a preservação digital, pois torna ineficiente a gestão da temporalidade e a destinação desses documentos.

Situação semelhante ocorre com os campos “**tamanho**” e “**tamanho do arquivo**”. Embora esse campo tenha a finalidade de descrever o tamanho do registro em bits no sistema, a ausência de uniformidade pode prejudicar a necessidade de interoperabilidade futura entre os sistemas.

O campo “**nome do arquivo**” se apresenta como responsável pela preservação do nome original do arquivo antes da sua inserção, no caso de arquivos anexados. Destacamos a redação dada pela UFGD, que o configurou

como “nome do arquivo com extensão”, de maneira a preservar também o formato do documento digital inserido, ou seja, nessa instituição, se um documento produzido em PDF com nome “arquivo” for incluído, o metadado armazenado será “arquivo.pdf” e não apenas “arquivo”, de modo que, conforme preconiza a literatura, outro elemento de preservação do contexto técnico será mantido.

O último campo, “**tipo do documento**”, de acordo com as respostas, tem como função descrever qual a espécie do documento, se é ata, relatório, certidão, etc. Porém, de acordo com os conceitos arquivísticos, para de fato ser caracterizado como tipo documental, precisaria unir a espécie documental com a atividade, por exemplo, ata da reunião do conselho universitário, relatório de atividade de extensão, certidão de aprovação de afastamento docente para pós-doutorado, etc.

As IFES adotam diferentes possibilidades para esse campo, a começar pela forma como o nomeiam. Se consideradas todas as formas constantes nas respostas, o número de ocorrências deixaria de ser dois e passaria a sete, ocupando, portanto, o topo da lista, tendo sido informado por UFGD, UFS, UFMA, UNIVASF, UNILA, UFPB e UFRB, com destaque para a descrição informada pela UFGD, que escolheu o termo “tipo de conteúdo” para indicar a mesma finalidade.

Como se pode perceber, essa discussão é permeada de detalhes simples, mas importantes para definir se a preservação digital acontecerá de forma integral ou parcial. Nesse sentido, considera-se importante mencionar que, no caso dos marcadores booleanos utilizados para definir um registro como “virtual” ou “escrito”, é preciso lembrar que existem registros no SIPAC que não se configuram nem como documentos nato digitais, nem como documentos digitalizados. Tratam-se, portanto, dos processos e documentos que foram criados de forma analógica ou física, cuja tramitação foi registrada no SIPAC, isto é, o sistema contém apenas as informações de distribuição e movimentação dos documentos. A UFPB possui tais registros. Sendo assim, esse se torna outro elemento necessário para reflexão: podem os registros de tramitação ser tratados da mesma forma que os registros que contém conteúdo de documentos, sejam eles nato digitais ou digitalizados? Como deve ser a preservação dos processos que iniciaram físicos e tornaram-se digitais? Esses

documentos, conhecidos como híbridos, carecem de metadados específicos para a sua preservação, que contemplem, dentre outros fatores, a existência de normas institucionais que detalhem a necessidade ou não de digitalização do conteúdo em suporte analógico, bem como as informações quanto ao local de arquivamento dos documentos produzidos no momento anterior à migração para o suporte eletrônico.

Passaremos agora à análise das respostas ao quesito 2 – **A sua instituição armazena metadados que identificam a origem e a localização dos usuários que produzem os registros e documentos? Caso sim, quais metadados são armazenados?**

Das treze instituições participantes, apenas a UFERSA informou não armazenar metadados de identificação da origem e localização dos usuários que produzem registros e documentos em ambiente digital. As outras doze IFES informaram um total de 43 tipos de metadados utilizados para essa finalidade, cujos resultados podem ser visualizados no gráfico 18.

Gráfico 18 – Metadados armazenados pelas IFES para identificação de origem e localização dos usuários do SIPAC



Fonte: dados da pesquisa, 2020.

O campo de metadado que aparece em destaque, embora não esteja visível na ilustração gerada pelo *software Rapid Miner*, é o “**endereço IP**”, com quatro ocorrências (UFGD, UFABC, UNILA e UFRB), seguido do campo “**IP da máquina do usuário**”, com duas ocorrências (UFLA e UFRN); juntos, representam 14% das respostas sobre essa questão.

Considerando que a UFPB é o objeto deste estudo, destaca-se que a resposta fornecida, embora não contenha o campo IP, apresenta como opções de metadados os itens “login do usuário”, “nome do usuário”, “matrícula SIAPE do usuário”, “unidade de lotação do usuário” e “data de criação”.

Convém esclarecer que os metadados relacionados ao IP dizem respeito à identificação do computador utilizado pelo usuário quando do acesso ao sistema para o cadastro de um documento, sendo possível inferir que os dois termos utilizados pelas universidades possuem significação equivalente e a existência de dois termos se dá pela ausência de padronização terminológica da área.

A importância de se registrar o IP para a finalidade digital é inquestionável, entretanto, tal metadado também se configura como um importante elemento arquivístico de preservação dos registros digitais, haja vista que pode oferecer o contexto testemunhal que assegure ou não a autenticidade de um registro. A título de exemplo, a captura do IP de um documento que só possa ser produzido pelo reitor e tenha sido criado a partir de um dispositivo que ele não utiliza eventualmente pode contribuir na verificação da autenticidade desse documento.

Os únicos dois outros campos de metadados que apresentaram duas ocorrências, foram justamente “**login do usuário**” e “**unidade de origem – para processo**”, elementos presentes na resposta da UFPB.

Verifica-se que foram informados pelas IFES uma média 4 campos de metadados para identificação de origem e localização dos registros digitais. Embora a UFPB não tenha indicado o endereço IP como resposta a essa questão, é preciso destacar que a universidade informou que faz o armazenamento desse metadado para outra finalidade: a de descrição do dispositivo. Ou seja, considerando que a UFPB informou cinco itens de metadados para essa finalidade, a instituição armazena uma quantidade de

metadados de preservação digital sobre identificação e origem do usuário em quantidade superior às outras IFES pesquisadas.

Em sequência, depois desses quatro campos, todas as outras respostas obtiveram apenas uma ocorrência cada.

Destacamos que a UFOB respondeu que armazena o metadado “Assinante do documento”, entretanto, não faz diferença entre os metadados dos documentos nato digitais e dos digitalizados. No SIPAC, o assinante do documento nato digital é o servidor com a competência institucional para a assinatura daquele documento, portanto, autor ou participante da ação documentada. De forma oposta, no caso dos documentos digitalizados, o SIPAC identifica como assinante o servidor responsável pelo recebimento daquele documento para a inclusão no sistema, sendo, portanto, uma assinatura de autenticação, mas não de autoria.

A falta de metadados de identificação do assinante que diferenciem o autor do documento nato digital, daquele que é responsável pela autenticação de um representante digital, coloca em risco a preservação digital desses registros, uma vez que não se pode confundir o autor de um documento com alguém que autentica uma cópia de documento produzido por outrem, de tal maneira que a utilização de um único metadado para as duas situações configura-se como um elemento passível de interpretação equivocada quando da necessidade de acesso futuro a essa informação.

Ainda sobre esse quesito, destacamos a existência de metadados informados de exclusivamente por duas instituições. A UFRN foi a única universidade a indicar o armazenamento de metadado de geolocalização, o que a coloca em posição de destaque no que se refere aos metadados de identificação de localização em relação às outras IFES. Do mesmo modo, a UFGD foi a única a informar o armazenamento das telas e URL's acessadas pelo usuário, ou seja, a IFES realiza a captura e guarda de cada um dos endereços eletrônicos acessados pelos usuários do seu sistema.

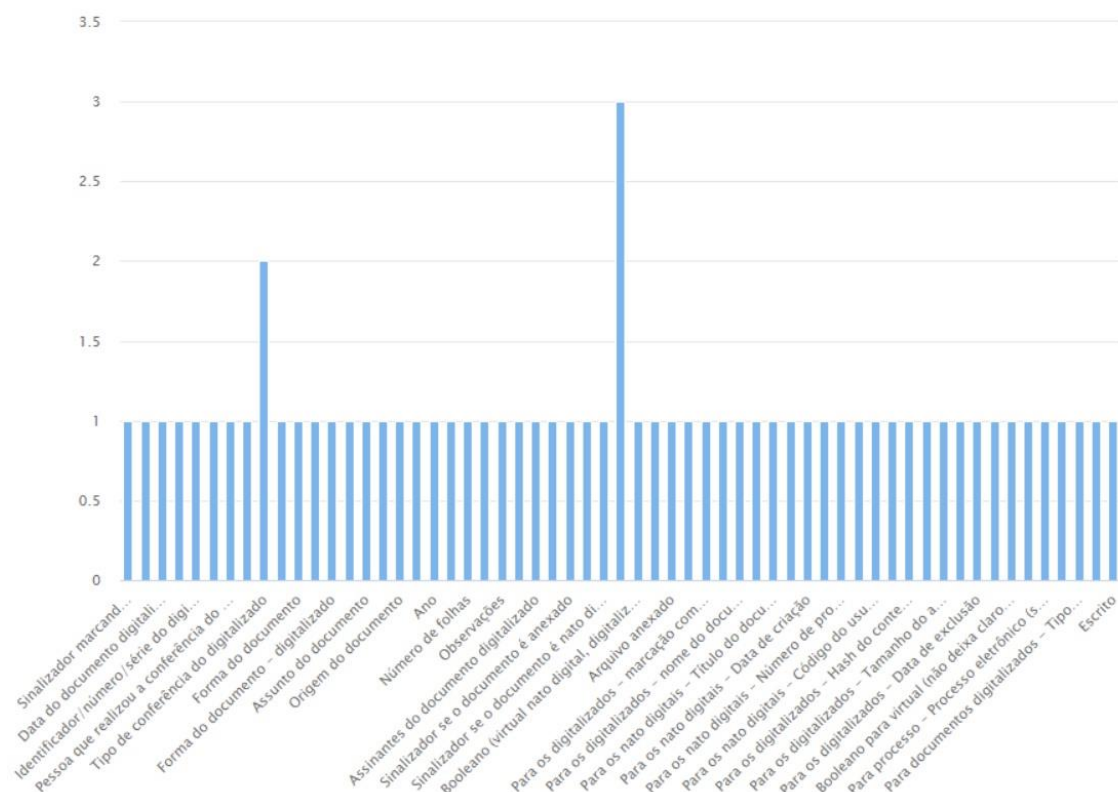
Partindo para a questão 3 - **A sua instituição armazena metadados que diferenciem os registros e documentos gerados em ambiente digital (nato digitais) daqueles convertidos para o meio eletrônico (digitalizados)?**

Três instituições disseram não realizar armazenamento desse tipo de metadados (UFERSA, UFOB e UNIVASF). A UFABC informou que, por não ser a desenvolvedora do *software*, preferia que as respostas fossem solicitadas à UFRN, responsável pela criação do sistema. Sendo assim, as respostas a essa questão foram fornecidas por UFGD, UFS, UFMA, UNILA, UFPB, UFLA, UFRB, UFSJ e UFRN.

Destaque-se que a UFOB, que contém o maior número de campos de metadados de preservação digital informado dentre as IFES respondentes, não diferencia metadados de digitalizados dos metadados dos registros nato digitais, o que corrobora o entendimento de que a quantidade de campos não reflete a qualidade dos metadados de acordo com os requisitos propostos pela literatura.

Das nove instituições respondentes a esse quesito, foram obtidos 62 (sessenta e dois) campos de metadados dedicados à diferenciação entre registros nato digitais e registros provenientes de digitalização, cuja ilustração pode ser visualizada no gráfico 19.

Gráfico 19 – Metadados armazenados pelas IFES para diferenciar registros nato digitais daqueles digitalizados



Fonte: dados da pesquisa, 2020.

Os dados obtidos revelam que o metadado mais utilizado pelas instituições para a distinção entre os registros gerados em ambiente digital daqueles digitalizados é justamente um campo que responde a essa questão, nomeado de diferentes formas, mas com a mesma finalidade. O item corresponde a 23% das respostas e suas variações podem ser identificadas no quadro 11.

Quadro 12 – Metadados de diferenciação entre nato digital e digitalizado nas IFES que utilizam o SIPAC

Nome do campo de metadado	IFES
1. Sinalizador marcando que se trata de um nato digital 2. Sinalizador marcando que se trata de um digitalizado	UFGD
1. Forma do documento – nato digital 2. Forma do documento – digitalizado	UFS
1. Sinalizador se o documento é anexado 2. Sinalizador se o documento é nato digital	UFMA
1. Booleano que indica que o documento é virtual e que foi escrito pelo próprio usuário dentro do sistema. Caso contrário, é um documento virtual cujo conteúdo está em um arquivo anexado ou é um documento não virtual (representação de um documento físico); 2. Booleano que indica se o documento é eletrônico ou físico;	UNILA
1. Para os nato digitais – marcação como escrito 2. Para os digitalizados – marcação como anexado	UFPB
1. Booleano para virtual 2. Booleano para escrito	UFRB
1. Para documentos – forma do documento (nato digital ou anexado) 2. Para processos – processo eletrônico (sim ou não)	UFSJ

Fonte: dados da pesquisa, 2020.

Convém elucidar que a UFLA, embora não tenha informado um campo específico para indicar se o registro é nato digital ou digitalizado, apresentou metadados diferentes para cada condição dos registros, indicando que a instituição faz a distinção entre os diferentes tipos de registros, embora não possua um campo específico para determinar essa finalidade.

Em sequência, o outro campo de metadado relacionado a esse quesito que se destaca é o item correspondente ao tipo de conferência realizada quando o registro se trata de um documento digitalizado. O item corresponde a 6% das respostas, conforme o quadro abaixo.

Quadro 13 – Metadados acerca do tipo de conferência para os documentos digitalizados nas IFES que utilizam o SIPAC

Nome do campo de metadado	IFES
Tipo de conferência do digitalizado	UFGD
Tipo de conferência do digitalizado	UFS
Tipo de conferência para digitalizados	UFMA
Para documentos digitalizados – tipo de conferência	UFSJ

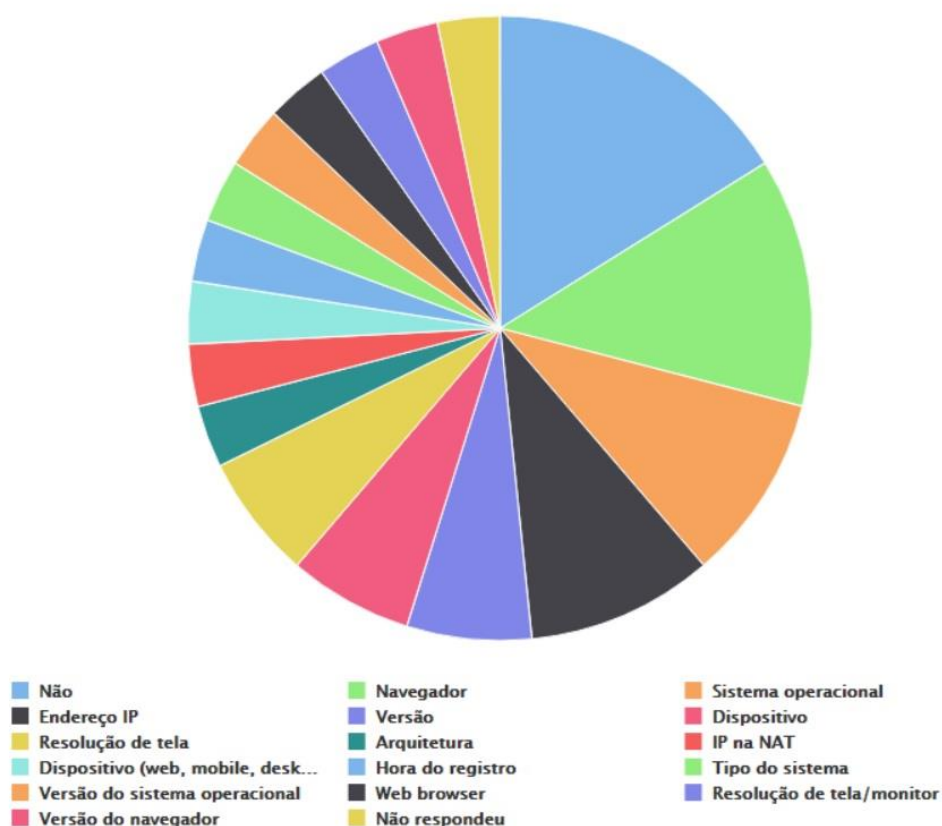
Fonte: dados da pesquisa, 2020.

A UFPB não informou a existência de metadado especificando o tipo de conferência para digitalizados na resposta a esse quesito, porém, indicou a existência do campo “Para digitalizados – Tipo de conferência (original, cópia autenticada em cartório, cópia autenticada administrativamente ou cópia simples)” na resposta ao quesito 6, tendo sido, inclusive, a única instituição a descrever qual tipo de conferência era realizada.

Verifica-se que, dentre as instituições que responderam a essa questão, a média de campos de metadados dedicados à finalidade de diferenciar nato digitais de digitalizados é de 20 itens. Na resposta desse item, a UFPB informou apenas 5, contudo, conforme apontado anteriormente, as respostas aos quesitos se complementam e o número de campos indicado não significa maior ou menor grau de aderência à categoria em questão. A análise dos aspectos positivos e negativos do conjunto de metadados utilizados pela UFPB será discutida até o final desta seção.

Em continuidade, passaremos a analisar as respostas ao quesito **4 - A sua instituição armazena metadados que descrevam o dispositivo e o sistema operacional no qual os registros e documentos foram produzidos? Caso sim, quais metadados são armazenados a respeito do dispositivo e do sistema operacional?**

Gráfico 20 – Metadados de descrição do dispositivo e sistema operacional armazenados pelas IFES que utilizam o SIPAC



Fonte: dados da pesquisa, 2020.

De acordo com o gráfico acima, é possível perceber que, das treze instituições respondentes, 5 cinco IFES responderam que não armazenam metadados de descrição do dispositivo e sistema operacional. São elas: UFRN, UFERSA, UFRB, UFOB e UFMA. Similarmente, 2 instituições – UFS e UFSJ – deixaram esse quesito sem resposta, diminuindo para 6 o número de universidades que informaram quais metadados são armazenados para essa finalidade: UFGD, UFABC, UNIVASF, UNILA, UFB e UFLA.

Dentre as instituições que indicaram uma resposta, foram apresentados 25 campos de metadados, ou seja, as IFES armazenam uma média de 4 campos de metadados para identificação do dispositivo ou sistema operacional. Dentre eles, aquele com maior incidência é o navegador ou *web browser*, com cinco ocorrências. Em seguida, está o endereço IP, com três ocorrências. Por fim, com duas ocorrências aparecem os campos “dispositivo”, “resolução de tela” e “versão”.

Destaque-se que o item “dispositivo” tem como finalidade a identificação de qual tipo de equipamento está sendo utilizado para a produção do registro, por exemplo, computador, tablet ou celular. Uma das instituições apresentou como possibilidades para esse campo as opções *web*, *mobile* e *desktop*. No campo “resolução de tela”, que também foi apresentado com a nomenclatura “resolução de monitor”, a intenção é obter a informação de qual era a disposição do “documento manifestado” para o usuário que o produziu ou inseriu. Já o item “versão”, não explicado de maneira clara pelos órgãos respondentes, indicaria a versão do sistema utilizado quando da criação do registro, de modo a preservar o ambiente técnico necessário para a preservação do documento ao longo do tempo de acordo com as características do *software*.

Chama a atenção o fato de a UFGD ser a única instituição a possuir um metadado destinado a preservar a arquitetura do registro. Entende-se que a adoção de um campo com a intenção de armazenar informações para visualização do documento, de acordo com as características computacionais necessárias para sua visualização da forma como foi produzido, atende de forma *sui generis* às recomendações da literatura para a preservação digital.

A UFPB apresentou 7 campos de metadados destinados à descrição do dispositivo e sistema operacional na resposta a esse questionamento.

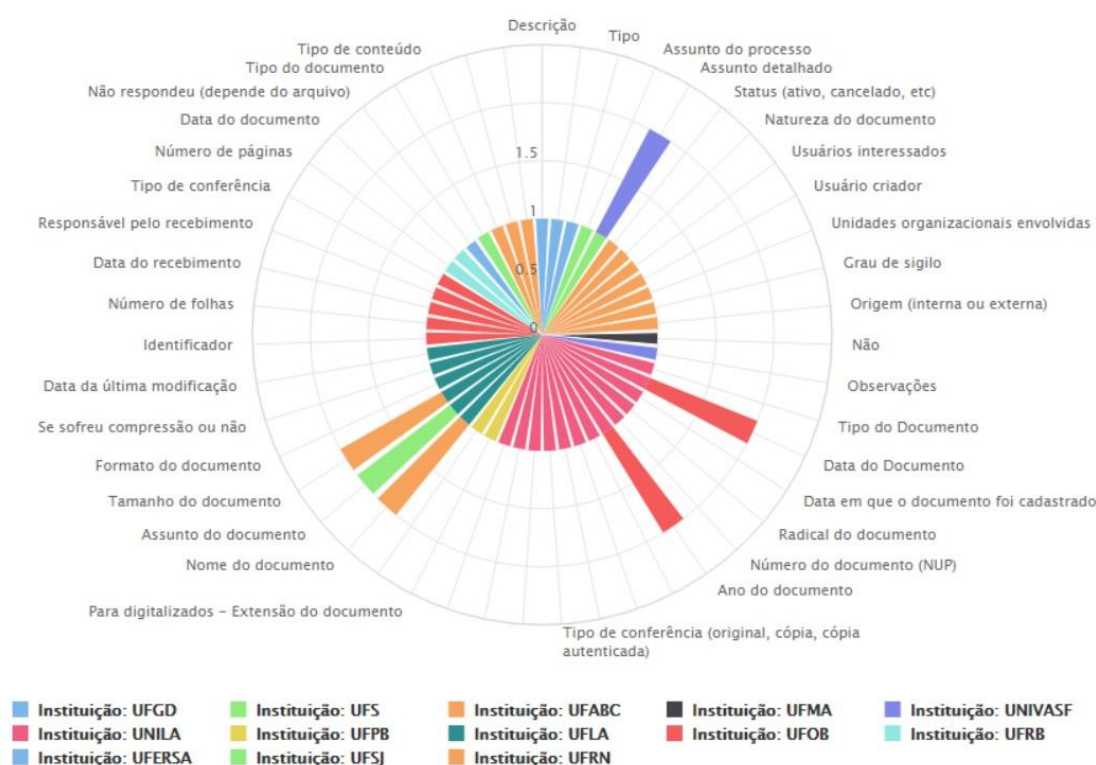
Outro fato que chama a atenção é o caso da UFRN, instituição proprietária do SIPAC, ter respondido que não armazena metadados destinados à descrição do dispositivo e sistema operacional nessa questão. Cabe destacar que, na resposta ao quesito 2, que trata dos metadados que identificam a origem e a localização dos usuários, a instituição já havia indicado o armazenamento dos seguintes campos: “informações do sistema operacional” e “informações do dispositivo”. Ou seja, a UFRN armazena metadados para a finalidade dessa questão, embora não tenha apresentado respostas no quesito 4.

Analisemos o quesito 5 – A sua instituição armazena metadados que descrevam os dados constantes nos registros e documentos criados/inseridos nos sistemas SIGAA/SIPAC (ou nos sistemas de produção de documentos acadêmicos e administrativos)? Caso sim, quais são os metadados armazenados para descrição dos dados?

Das treze IFES respondentes, apenas UNIVASF e UFERSA não apresentaram respostas a esse quesito. A primeira respondeu que não armazena esse tipo de metadado e a segunda, embora tenha dito que armazena, informou que “esses dados dependem do arquivo que está sendo inserido no sistema, ficando impossível de dizer exatamente quais são eles visto a variedade de tipos de documentos criados/inseridos nos SIGs”.

O gráfico 21 contém a ilustração das respostas obtidas pelas instituições participantes da pesquisa sobre os metadados que descrevem os dados constantes nos registros digitais.

Gráfico 21 – Metadados descritivos de dados armazenados pelas universidades federais que utilizam o SIPAC



Fonte: dados da pesquisa, 2020.

Identificam-se diversos campos de metadados nas respostas, entretanto, conforme o gráfico demonstra, apenas 6 campos obtiveram destaque quanto ao número de ocorrências, são eles: “ano do documento”, “assunto detalhado”, “assunto do documento”, “data do documento”, “nome do documento” e “tamanho do documento”. Todos esses campos já foram objeto

de análise desta pesquisa, com exceção do “ano do documento”, cuja função é registrar o ano em que o registro foi criado.

Considerando que as respostas dessa questão apresentaram campos de metadados bem diferentes quando comparados com as respostas das questões anteriores, e informados de maneira exclusiva por algumas instituições, destacamos os itens constantes no quadro abaixo:

Quadro 14 – Metadados de descrição de dados de universidades federais que utilizam o SIPAC

Descrição do campo de metadado	Universidade
Tempo de vida	UFRN
Se sofreu compressão ou não	UFLA
Código de verificação para averiguar a autenticidade do documento	UNILA
Numeração original do documento (no caso de documento externo)	
Unidades organizacionais envolvidas	UFABC

Fonte: dados da pesquisa, 2020.

Verifica-se, pela UFRN, a aplicação do conceito de tempo de vida. Sabe-se que “ciclo de vida” é uma abstração utilizada tanto pela Arquivologia para a definição das idades documentais, como por outras ciências, para a gestão de objetos digitais de diferentes naturezas. Embora não tenha ficado evidente se tal metadado está relacionado à temporalidade arquivística ou à definição do intervalo entre a data de criação e a data da consulta do registro, o armazenamento de tal informação – seja uma ou outra – constitui-se como importante elemento para a preservação digital.

A informação sobre “se sofreu compressão ou não”, constante na resposta da UFLA, demonstra uma importante preocupação com a preservação do contexto técnico do registro, fator preponderante na definição de uma política de preservação, uma vez que a não preservação desse tipo de dado pode comprometer, por exemplo, a visualização de imagens JPEG.

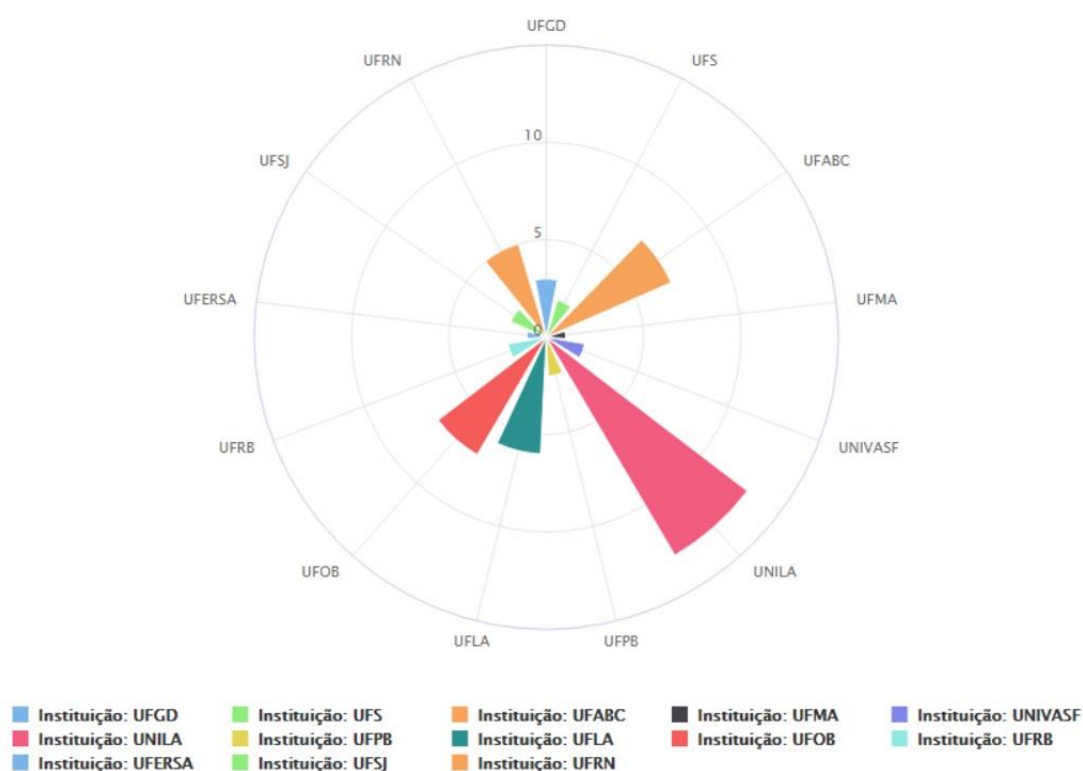
Do mesmo modo, os três outros campos de metadados informados pela UNILA e UFABC, “código de verificação de autenticidade”, “numeração original de documento externo” e “unidades organizacionais envolvidas”, fornecem elementos de contexto também considerados relevantes pela literatura de preservação digital de registros arquivísticos.

Dando sequência, considerando as respostas de todas as instituições, foram informados 51 campos de metadados destinados à preservação dos

dados contidos nos registros digitais pelas 11 universidades que responderam satisfatoriamente a essa questão, de modo que a média de itens armazenados para essa finalidade é de 5 campos.

O gráfico 22 ilustra a posição das IFES em relação à quantidade de metadados de descrição de dados informados.

Gráfico 22 – Universidades federais que utilizam o SIPAC em relação à quantidade de Metadados descritivos de dados



Fonte: dados da pesquisa, 2020.

Constata-se que a instituição com o maior número de metadados dedicados a essa finalidade é a UNILA, que informou armazenar 13 campos de metadados para descrição de dados dos seus registros. Em seguida, estão a UFABC e a UFOB, com 7 campos de metadados cada. Logo após a UFLA, com 6; seguida da UFRN, com 5 campos.

A verificação da aderência desses metadados ao modelo baseado em Rogers (2015) e Sayão (2010) aponta que nenhuma instituição apresenta um conjunto de metadados que possa ser considerado referência na descrição das informações de *hardware*, *software*, dependência entre os componentes

digitais e demais itens necessários à preservação digital do aspecto computacional. Ao contrário, as respostas obtidas demonstram um cenário bastante preocupante quanto ao armazenamento dessas informações.

A UFPB informou armazenar apenas dois campos de metadados para essa finalidade, a saber, “para digitalizados – extensão do documento” e “para nato digitais – conteúdo do documento em forma HTML”. Ou seja, a UFPB, embora quantitativamente abaixo da média das demais universidades em relação ao número de campos de metadados de descrição dos dados contidos em seus registros digitais, é a única a armazenar conteúdo em forma HTML de acordo com as respostas recebidas. É preciso destacar, ainda, que outros campos de metadados que também se configuram como descritores de dados já estão sendo armazenados pela UFPB, conforme respostas a questões anteriores.

Destaca-se que a simples diferenciação entre nato digitais e digitalizados ainda deixa de fora alguns elementos arquivísticos relevantes do contexto de produção dos registros digitais. Por exemplo, se uma certidão foi criada diretamente no sistema como um registro nato digital, logicamente ela é um documento produzido pela UFPB. Porém, uma certidão incluída como documento digitalizado tanto pode ter sido criada pela UFPB, como pode ter sido criada por um outro órgão, e não existe metadado sendo armazenado que estabeleça essa diferença. Isto é, esses registros estão sendo tratados de forma igual, quando são absolutamente diferentes do ponto de vista da responsabilidade institucional sobre eles.

Os digitalizados provenientes de documentos originais produzidos pela UFPB precisam ter informações quanto à localização física dos seus originais a eles associados, enquanto que os digitalizados provenientes de documentos recebidos pela UFPB precisam estar atrelados a metadados que traduzam de forma fidedigna o seu contexto de produção, e não apenas o tipo de conferência realizada para a digitalização.

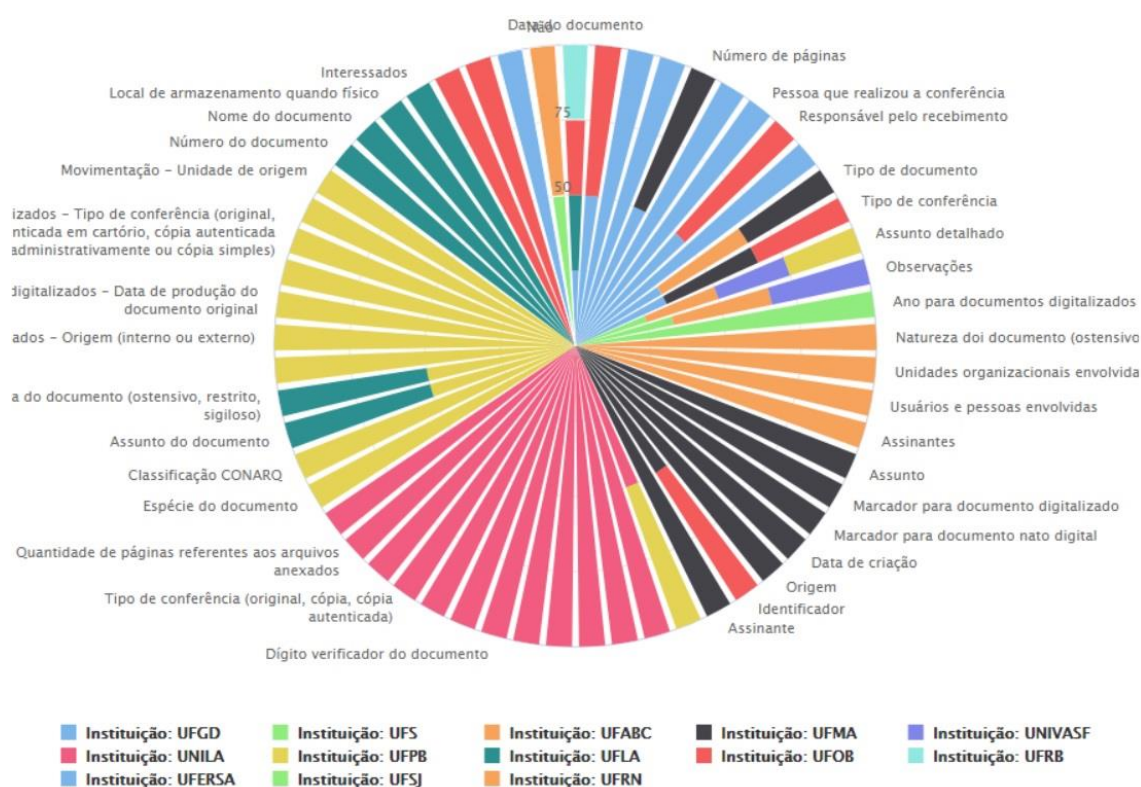
Passaremos agora à análise do quesito **6 - Existem metadados que descrevam os dados constantes nos registros e documentos criados/inseridos nos sistemas SIGAA/SIPAC (ou nos sistemas de produção de documentos acadêmicos e administrativos) que são incluídos pelo usuário de forma manual? Caso sim, quais são os**

metadados armazenados para descrição dos dados inseridos manualmente pelo usuário?

As instituições UFRN e UFSJ responderam que não armazenam esse tipo de metadado. Já a UFERSA, informou que armazena, contudo, que “esses dados dependem do arquivo que está sendo inserido no sistema, ficando impossível de dizer exatamente quais são eles visto a variedade de tipos de documentos criados/inseridos nos SIGs”.

As respostas das outras universidades federais que foram consideradas satisfatórias para as finalidades deste trabalho podem ser visualizadas no gráfico 23.

Gráfico 23 – Metadados descritivos incluídos de forma manual pelo usuário nas universidades federais que utilizam o SIPAC



Fonte: dados da pesquisa, 2020.

É possível perceber, com base no gráfico acima, que UNILA e UFPB se destacam como as IFES com o maior número de metadados descritivos de dados inseridos de forma manual pelo usuário. As duas instituições informaram armazenar 13 campos de metadados cada para esta finalidade.

Haja vista a possibilidade de dificuldade para leitura das legendas, consideramos conveniente descrever que os campos de metadados informados pela UFPB como resposta ao quesito 6 foram: “tipo do documento”, “espécie do documento”, “classificação CONARQ”, “assunto do documento”, “assunto detalhado”, “natureza do documento”, “hipótese legal da restrição de acesso – para documentos restritos”, “origem – para documentos digitalizados”, “data de produção do documento original – para documentos digitalizados”.

“data de recebimento do documento original – para documentos digitalizados”, “tipo de conferência – para documentos digitalizados”, “movimentação – unidade de destino” e “movimentação – unidade de origem”.

Salienta-se que, no campo “natureza do documento”, existem as opções de “ostensivo”, “restrito” ou “sigiloso”. No campo “origem – para documentos digitalizados”, as opções são “interno” ou “externo”. Já no campo “tipo de conferência – para documentos digitalizados”, conforme mencionado em análise anterior, as opções são “original”, “cópia autenticada em cartório”, “cópia autenticada administrativamente” ou “cópia simples”.

Ao considerar o conjunto das respostas de todas as IFES, os destaques que se apresentam foram os campos “assunto detalhado” e “data do documento”, com quatro incidências cada. Em seguida, com três ocorrências cada, “observações”, “tipo de conferência” e “tipo de documento”, todos já descritos em análises das questões anteriores.

Mediante o exposto, verifica-se que 100% das universidades federais brasileiras constantes nesta análise armazenam algum tipo de metadado para a preservação digital dos registros digitais arquivísticos criados no SIPAC. Contudo, **constata-se a falta de padronização na elaboração dos nomes utilizados para os campos, aspecto considerado um problema para a interoperabilidade interinstitucional de sistemas no futuro.**

Além disso, também é possível perceber que algumas IFES possuem problemas de ordem terminológica na composição dos seus metadados. Tal constatação se baseia na ausência de conexão entre o conceito de termos arquivísticos utilizados e sua aplicação prática no SIPAC.

No caso da UFPB, são armazenados metadados específicos para a preservação digital dos registros digitais criados no SIPAC, relacionados a todas as categorias analisadas por esta pesquisa. No entanto, as respostas da UFPB, assim como das outras IFES participantes, limitaram-se aos descritores automático e manuais do SIPAC Módulo Protocolo, sem mencionar os metadados existentes para os outros módulos, SIGAA e demais sistemas de criação de registros digitais, igualmente importantes para a preservação digital dos registros documentais da instituição em ambientes digitais.

A complexidade dos registros digitais arquivísticos envolve, embora não se limite a ela, a compreensão dos elementos arquivísticos necessários

para a sua preservação digital. Contudo, um dos principais desafios para essa demanda está em transpor requisitos pensados tradicionalmente sob a perspectiva dos documentos analógicos, para requisitos digitais com a finalidade de preservação do contexto e dos requisitos arquivísticos dos registros digitais.

De acordo com o modelo conceitual de requisitos para metadados de preservação digital, baseado no trabalho de Rogers (2015) e Sayão (2010), identificaram-se categorias sob as quais esses metadados precisam ser criados.

A primeira, relacionada à origem, localização e descrição da infraestrutura conceitual que descreve o ambiente, precisa incluir, para além das informações sobre setor de lotação do usuário que criou o registro, outros metadados que tragam luz ao caráter orgânico e contextual dos documentos arquivísticos, tais como informações sobre a vinculação daquele registro a alguma comissão, núcleo, projeto, grupo de trabalho ou qualquer outro tipo de unidade que possa não estar sendo identificada pelo sistema de produção do registro, como se percebe, no caso da UFPB, a falta de identificação do campus quando da produção de registros no SIPAC.

A segunda categoria, aparência e desempenho, nos metadados encontrados por esta pesquisa, faz menção ao formato do arquivo digital e outras informações de cunho tecnológico. Mas é preciso lembrar que, do ponto de vista arquivístico, existem elementos de aparência que traduzem aspectos diplomáticos e contextuais predominantemente arquivísticos e que não estão sendo capturados. Por exemplo, os registros audiovisuais criados pela plataforma de conferência *web* da Rede Nacional de Pesquisa possuem atributos visuais que os diferenciam daqueles criados pela plataforma *Google Meet* – ambos utilizados para gravação de atividades acadêmicas e administrativas da UFPB.

Em vista disso, preocupa-nos que, no futuro, sem essa informação, tais registros possam ter a sua integridade ou outros aspectos de sua autenticidade comprometidos. Isto é, não identificar a plataforma em que o registro foi criado originalmente antes de ser inserido no SIPAC é apenas um exemplo da necessidade de se criar um padrão para metadados de preservação arquivística de registros digitais.

Passando para a terceira categoria, que se refere à descrição do dispositivo e sistemas utilizados, entendemos que, embora eminentemente tecnológica, pode contribuir de forma efetiva com a manutenção dos requisitos arquivísticos, ao incluir informações sobre documentos inseridos provenientes de outros sistemas. Dessa forma, quando um documento que foi criado em outro *software* de governo precisar ser inserido no SIPAC, deveria ser possível identificar qual o sistema utilizado para a produção daquele registro, como por exemplo, o Sistema de Concessão de Diárias e Passagens (SCDP), Sistema Integrado de Administração Financeira (SIAFI), Sistema de Gestão de Pessoas (SIGEPE), entre outros. Embora já seja possível inserir essas informações em campo destinado a observações, a existência de um campo específico favoreceria a preservação e a recuperação de informações desses contextos.

Por fim, na categoria que trata dos metadados de descrição automática e manual dos dados, percebe-se que também existem possibilidades de ampliação que envolvam aspectos arquivísticos. Os eventos de gestão de processos administrativos, como desmembramentos, desentranhamentos, juntadas, são algumas das informações que não estão sendo objeto de metadados e que possuem relevante função arquivística. Do mesmo modo, embora a UFPB já possibilite a classificação dos registros de acordo com o Plano de Classificação de Documentos de Arquivo do CONARQ, existem alguns problemas do ponto de vista arquivístico que merecem uma reflexão.

Para cada código de classificação utilizado, existe uma previsão legal do tempo em que aquele registro digital precisa ser mantido no ambiente de gestão, se há a necessidade de guarda precaucional e, caso haja, por qual período, e qual deve ser a destinação final daquele documento, podendo ser a eliminação ou a guarda permanente. O que chama a atenção é o fato de, na UFPB, a informação quanto ao código de classificação ainda não estar sendo utilizada para a efetiva gestão do ciclo de vida dos documentos. Ou seja, embora existam registros no SIPAC classificados, não há ainda nenhuma ação efetiva de transferir documentos do ambiente de gestão para o ambiente de preservação, pois não estão sendo vinculados os metadados de temporalidade.

Tal situação nos leva para outra reflexão, que é a necessidade de interoperabilidade entre os sistemas para que a gestão dos documentos aconteça de forma plena. Sob esse aspecto, o CCJ possui um relevante

volume de documentos que já poderiam ser destinados, porém necessitam de informações constantes de outros sistemas da universidade para que a destinação aconteça e, como os sistemas não interoperam e a UFPB não armazena metadados que subsidiem essa gestão, a destinação fica condicionada à captura manual das informações de vinculação do registro ao evento que determina a sua temporalidade, e a um comando posterior que possibilite a sua eliminação, transferência ou recolhimento ao repositório arquivístico digital confiável, cuja existência e funcionamento ainda estão em processo de implantação.

Para ilustrar essa realidade, existem dois exemplos que consideramos significativos. Na atividade fim, os processos de dispensa de disciplina possuem como evento determinante da sua temporalidade a data de encerramento do vínculo do estudante com a instituição. Dessa forma, para agilizar a gestão dos documentos no ambiente digital, quando a Pró-reitoria de Graduação inserisse a data de colação de grau de um aluno no Sistema Integrado de Gestão das Atividades Acadêmicas (SIGAA), deveria ser criado um metadado com essa informação e este deveria ser associado a todos os registros criados para aquele aluno no SIPAC, de maneira que seria possível automatizar a destinação dos documentos de acordo com a determinação legal ou, ao menos, acelerar o processo de trabalho. Já na atividade meio, os processos de compra e diversos outros documentos financeiros e/ou que envolvem pagamento de despesas, possuem como evento determinante da sua temporalidade a data da aprovação das contas da instituição pelo TCU ou a data de apresentação do relatório de gestão.

Do mesmo modo, para tornar mais eficiente a gestão dos documentos no ambiente digital, quando a Reitoria publicasse o Relatório de Gestão, deveria ser criado um metadado com essa data, e este deveria ser associado a todos os registros cujos códigos de classificação fossem a ela condicionados, de maneira que seria possível automatizar a destinação dos documentos de acordo com as normas vigentes.

Destaca-se ainda a inadequação da UFPB quanto à classificação dos registros digitais das atividades meio, haja vista o código de classificação e tabela de temporalidade e destinação de documentos relativos às atividades-

meio do Poder Executivo Federal ter sido atualizado em fevereiro de 2020 e, até o mês de novembro de 2020, o SIPAC ainda não ter sido corrigido.

Verifica-se ainda, no caso específico da UFPB, que a instituição possui grupos de metadados para identificação de origem e localização, de distinção entre nato digitais e digitalizados, para descrição dos sistemas e dispositivos utilizados pelo usuário, para descrição automática de dados e para descrição manual de dados pelo usuário, mas não há a diferenciação precisa de metadados de preservação para os níveis lógico, físico e conceitual, três diferentes níveis necessários para a preservação no ambiente digital.

Deveria existir um conjunto de metadados específicos para a preservação do nível lógico (entendido como a cadeia de bits que precisa de informações contextuais em linguagem de máquina para ser compreendido), um conjunto de metadados relacionados à preservação do objeto digital no nível físico (entendido como o artefato ou o *hardware* utilizado e necessário para armazenamento e visualização do documento e as intervenções de migração e alterações de suporte realizadas ou planejadas), e um conjunto de metadados destinados à preservação do objeto digital no nível conceitual (ou seja, descritores do documento digital arquivístico manifestado que contribuam na organização e recuperação da informação pelo usuário).

Constata-se que os metadados podem e devem servir de parâmetro para a gestão dos registros digitais arquivísticos. Contudo, para isso, os sistemas precisam trabalhar de forma integrada, acompanhar as mudanças na legislação, mas também reunir informações contextuais de *hardware*, *software* e elementos descritivos que garantam a preservação dos requisitos arquivísticos dos registros digitais.

5. CONSIDERAÇÕES FINAIS

Ao longo da história o homem sempre buscou registrar a sua atividade, seja como indivíduo ou como grupo. No princípio, os registros eram feitos em pedras de argila, em papiros, em papel, de tal maneira que não se fazia necessária nenhuma mediação para que tais documentos fossem lidos, além do conhecimento sobre os símbolos que permitisse a sua decodificação.

No cenário atual, as tecnologias digitais de informação e comunicação, que não se limitam ao uso de computadores e *smartphones*, mas se consolidam na criação de vídeos, áudios, mensagens de texto e os mais diferentes tipos de registros digitais, alteraram de forma significativa o suporte no qual a informação que documenta as ações humanas e institucionais é registrada.

Nesse contexto, conhecer os símbolos utilizados nos registros não é mais suficiente para acessar a informação contida nos documentos. Um novo tipo de mediação é requerida, nesse caso, além da utilização de máquinas como mediadoras indispensáveis para a leitura e acesso aos documentos manifestados. Tanto essas máquinas, quanto o usuário que delas se utiliza para obter a informação que precisa terá que se valer de dados que descrevam os aspectos computacionais, lógicos, físicos e contextuais de maneira geral, dessas diferentes formas de registrar a informação.

Mais ainda: dado o caráter dinâmico dos recursos digitais e a forte obsolescência tecnológica decorrente desse universo, é preciso reunir informações que garantam ou ao menos busquem um nível de segurança razoável de que o ambiente computacional no qual a informação foi criada poderá ser reproduzido com o passar do tempo, de modo que os registros permaneçam acessíveis independentemente da existência do formato e das condições técnicas em que foram concebidos.

É justamente o conjunto desses atributos, aqui denominados de metadados, que permitem que os dados constantes de um documento em ambiente digital possam se tornar um documento manifestado e acessível ao usuário no decorrer do tempo, associado ao conteúdo que o fundamenta, e as transações e requisitos que o tornam válido, que constitui um registro digital,

que, quando criado para evidenciar e servir de prova das atividades de uma pessoa ou instituição, é um registro digital arquivístico.

Compreendemos que o conceito de registro digital arquivístico ainda precisa de outras reflexões que o esmiúcem e ampliem, porém, o estudo realizado até aqui nos leva à compreensão de que **o registro digital arquivístico se caracteriza como uma proposta conceitual mais ampla do que o documento arquivístico digital** (ACKER, 2019; RIES, 2018; ROGERS, 2015), pois ressignifica os limites do que é considerando relevante para a preservação dos requisitos arquivísticos dos documentos em ambiente digital, com destaque para os metadados.

No presente estudo, os registros digitais arquivísticos do CCJ da UFPB foram descritos de forma detalhada, como cumprimento do primeiro objetivo, com o entendimento de que os elementos que contribuem para a sua descrição, dentre outros, são: a plataforma utilizada, o setor de origem, a atividade vinculada, a espécie documental utilizada, o gênero documental ao qual o registro pertence, bem como do tipo documental resultante.

Ao perceber a existência de registros audiovisuais, mensagens de correio eletrônico, mensagens de texto em aplicativos, publicações em redes sociais, formulários em sítios institucionais na internet em conjunto com dezenas de espécies documentais tradicionais – agora criadas em ambiente eletrônico, coexistindo como meios de documentar as atividades meio e fim da UFPB que ocorrem no CCJ –, verifica-se a necessidade de ampliação deste estudo que, pelas questões já apresentadas, limitou-se aos documentos criados e inseridos no SIPAC.

Destacamos que um dos objetivos iniciais do trabalho era elaborar uma proposta de classificação dos registros digitais arquivísticos, com o intuito de colaborar com uma possível compreensão dos seus aspectos técnicos relacionados à gestão dos documentos arquivísticos digitais. Entretanto, considerando ser a classificação um conceito que agrega múltiplas interpretações, quais sejam, natureza do assunto, finalidade, função, grau de sigilo, entre outras, optou-se pela modificação do objetivo para a categorização, de modo a limitar o estudo a uma proposta de divisão não arquivística e mais ampla dos registros digitais, considerando não ser o escopo desta pesquisa

uma discussão teórica mais profunda dos aspectos arquivísticos da classificação dos registros digitais.

Contudo, outros pesquisadores poderão dedicar-se no futuro ao estudo das diferenças entre os registros digitais arquivísticos enquanto entidades conceituais que reúnem o conteúdo e os metadados dos documentos arquivísticos digitais, e os registros digitais arquivísticos enquanto possibilidade de configuração entre as espécies documentais e todos os aspectos teórico-metodológicos e semânticos derivados dessas possibilidades.

Sendo assim, os registros digitais arquivísticos do CCJ da UFPB foram categorizados de acordo com os modelos dos autores Garmendia (2020), Zhao *et al* (2019), Katuu (2016) e Rondinelli (2011) e, em cumprimento do segundo objetivo desta pesquisa, constituindo uma proposta que reforça o entendimento sobre a **necessidade de aprofundamento quanto à natureza dos registros digitais, suas características e a emergência da definição de ações de gestão arquivística próprias para tais recursos informacionais, com reflexões sobre a padronização da produção, da descrição e dos requisitos básicos para que o seu acesso seja possível ao longo do tempo, consideradas as variáveis físicas e lógicas envolvidas.**

Em cumprimento ao terceiro objetivo, foram identificados os metadados necessários para subsidiar a preservação digital dos registros digitais arquivísticos do CCJ da UFPB. A comparação das propostas constantes na literatura com as respostas das universidades federais pesquisadas, incluída a UFPB, sobre os campos de metadados existentes no SIPAC Módulo Protocolo, permitiu visualizar que as IFES brasileiras armazenam metadados para a preservação digital, porém as informações quanto ao aspecto material dos registros digitais são limitadas, não há padronização dos campos de metadados entre as instituições, existem incoerências terminológicas relacionadas aos conceitos arquivísticos, predomínio da atuação de profissionais de tecnologia em detrimento dos profissionais de arquivologia, lacunas de informação sobre as diferenças entre documentos nato digitais, digitalizados e híbridos, além de uma série de questões relacionadas com a ausência de uma política nacional ou no âmbito dos órgãos vinculados ao Ministério da Educação acerca dos critérios para a preservação digital dos registros digitais arquivísticos.

Destacamos que a UFPB reúne elementos de preservação digital para todas as áreas indicadas pela literatura, estando, portanto, alinhada ao modelo conceitual de preservação digital de registros digitais arquivísticos, no que se refere ao conjunto de metadados utilizado para o SIPAC Módulo Protocolo. Entretanto, verificou-se que em todas as categorias de análise existem oportunidades de criação de metadados que ampliem, sob o ponto de vista arquivístico, as possibilidades de preservação. Isto é, **não é suficiente ter metadados de preservação digital: é preciso que haja metadados de preservação digital específicos para registros digitais arquivísticos a fim de preservar o contexto arquivístico dos registros digitais.**

Embora a maioria dos metadados armazenados pela UFPB sejam relacionados ao contexto descritivo e conceitual dos documentos – aspecto relevante para a preservação digital, ainda é necessário o desenvolvimento de campos que consolidem esse nível da preservação e incluam metadados vinculados aos elementos de *hardware* e *software* necessários para a preservação de longo prazo.

Em suma, entendemos que o objetivo geral desta pesquisa, de estudar como se constituíam os registros digitais arquivísticos produzidos pelo CCJ/UFPB, foi alcançado, pois evidenciou-se um conjunto de elementos descritivos de tais registros, verificou-se a aplicação de modelos de categorização que contribuem para a sua compreensão e identificou-se metadados necessários para a sua preservação digital, com a análise daqueles que já existem e dos que precisam ser criados.

Propõe-se, como sugestão para investigação futura, a verificação da adequação da UFPB no que se refere à preservação digital das demais plataformas utilizadas para a criação de registros digitais arquivísticos (sítio, redes sociais, aplicativos institucionais e os outros módulos do SIG utilizados pelo órgão).

Constata-se a necessidade de reflexão sobre o profissional de arquivo frente a essas demandas. O Arquivista deve tornar-se um agente da curadoria digital, ou ainda, um Arquivista Digital, isto é, um planejador ativo das estratégias de preservação de longo prazo dos registros digitais arquivísticos sob sua responsabilidade. Para isso, o conhecimento sobre metadados (agora parte do acervo) deve fazer parte das suas competências profissionais para

que ele possa contribuir com a criação de informações: de conteúdo, que interpretem a cadeia de bits para as máquinas; e de descrição, que facilitem a recuperação para os usuários. Além disso, precisará estar preparado para criar e manter um repositório que contemple, dentre outros elementos, os desafios da privacidade e da propriedade intelectual, como também da gestão do uso e gestão do acesso.

A sociedade do futuro só conhecerá a sociedade do presente se os documentos arquivísticos digitais estiverem acessíveis, e isso só será possível a partir da compreensão dos registros digitais arquivísticos. Entendemos que esta pesquisa inicia a discussão, sem a intenção de esgotá-la ou limitá-la.

REFERÊNCIAS

ACKER, A. Quando é um registro? Estrutura de pesquisa para localização de registros eletrônicos na infraestrutura. GILLILAND, A. J.; MCKEMISH, S.; LAU, A. J. (org.). **Pesquisa no multiverso arquivístico**. Tradução de Ana C. Rodrigues. Salvador: 9Bravos, 2019.

ADU, K. K.; NGULUBE, P. Preserving the digital heritage of public institutions in Ghana in the wake of electronic government. **LIBRARY HI TECH**, ENGLAND, v. 34, n. 4, p. 748–763, 2016. Disponível em: <https://doi.org/10.1108/LHT-07-2016-0077>. Acesso em: 07 set. 2020.

ANDERSON, B. *et al.* Archival appraisal and the digital record: Applying past tradition for future practice. **New Review of Information Networking**, [s. l.], v. 20, n. 1, p. 3–15, 2015. ISSN: 13614576, DOI: 10.1080/13614576.2015.1114823.

ALLAN, A. **Records Review**. London: Cabinet Office and The National Archives, 2014. Disponível em: [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/370930/RECORDS_REVIEW - Sir Alex Allan.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/370930/RECORDS_REVIEW_-_Sir_Alex_Allan.pdf). Acesso em: 08 abr. 2020.

ALLAN, A. **Government Digital Records and Archives Review**. London: Cabinet Office and The National Archives, 2015. Disponível em: [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/486418/Report - Digital Records Review.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/486418/Report_-_Digital_Records_Review.pdf). Acesso em: 08 abr. 2020.

ARAÚJO, C. A. A. **Arquivologia, biblioteconomia, museologia e ciência da informação: o diálogo possível**. Brasília: Briquet de Lemos, 2014.

ARAÚJO, C. A. A. O que é ciência da informação. Belo Horizonte: KMA, 2018.

ARQUIVO NACIONAL (Brasil). **Dicionário Brasileiro de Terminologia Arquivística**. Rio de Janeiro: Arquivo Nacional, 2005.

BANTIN, P. Developing a Strategy for Managing Electronic Records – The Findings of the Indiana University Electronic Records Project. **The American Archivist**, [s. l.], v. 61, n. 2, p. 328–364, 1998. Disponível em: <https://www.americanarchivist.org/doi/abs/10.17723/aarc.61.2.f14rth861430lk74>. Acesso em: 06 mar. 2020.

BARON, J. R.; THURSTON, A. What lessons can be learned from the US archivist's digital mandate for 2019 and is there potential for applying them in lower resource countries? **Records Management Journal**, [s.l.], v. 26, n° 2, p. 206–217, 2016. ISSN: 0956-5698, DOI: 10.1108/RMJ-04-2015-0015.

BRASIL. **Decreto nº. 8.539**, de 8 de outubro de 2015. Dispõe sobre o uso do meio eletrônico para a realização do processo administrativo no âmbito dos

órgãos e das entidades da administração pública federal direta, autárquica e fundacional. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2015/decreto/d8539.htm. Acesso em: 26 mar. 2020.

BRASIL. **Decreto nº. 8.638**, de 15 de janeiro de 2016. Institui a Política de Governança Digital no âmbito dos órgãos e das entidades da administração pública federal direta, autárquica e fundacional. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2016/decreto/d8638.htm. Acesso em: 26 mar. 2020.

BRASIL. **Lei nº. 12.527**, de 18 de novembro de 2011. Regula o acesso a informações previsto no inciso XXXIII do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição Federal; altera a Lei nº 8.112, de 11 de dezembro de 1990; revoga a Lei nº 11.111, de 5 de maio de 2005, e dispositivos da Lei nº 8.159, de 8 de janeiro de 1991; e dá outras providências. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12527.htm. Acesso em: 27 mar. 2020.

BRASIL. **Estratégia Brasileira de Transformação Digital (e-Digital)**. Brasília. 2018. Disponível em: https://www.mctic.gov.br/mctic/export/sites/institucional/arquivos/ASCOM_PUB_LICACOES/estrategia_digital.pdf. Acesso em: 17 mar. 2020.

BRINDLEY, G.; MUIR, A.; PROBETS, S. Provision of digital preservation metadata: A role for ONIX? **Program**, [S. l.], v. 38, n. 4, p. 240–250, 2004. Disponível em: <https://doi.org/10.1108/00330330410699883>. Acesso em: 07 set. 2020.

BROOKS, J. Perspectives on the relationship between records management and information governance. **Records Management Journal**, School of History, University College Dublin, Dublin, Ireland, v. 29, nº 1/2, p. 5–17, 2019. ISSN: 09565698, DOI: <http://dx.doi.org/10.1108/RMJ-09-2018-0032>.

BUCKLAND, M. K. Information as thing. **Journal of the American Society of Information Science**, v. 42, n. 5, Jun. 1991, p. 351-360. Disponível em: <http://people.ischool.berkeley.edu/~buckland/thing.html>. Acesso em: 07 abr. 2020.

CÂMARA TÉCNICA DE DOCUMENTOS ELETRÔNICOS (CTDE). **Glossário**. Disponível em: <http://www.documentoseletronicos.arquivonacional.gov.br/media/2008ctdeglossario5.pdf>. Acesso em: 12 abr. 2020.

CAMARGO, A. M. de A.; BELLOTTO, H. L. (org.). **Dicionário de Terminologia Arquivística**. São Paulo: Núcleo Regional de São Paulo/Associação dos Arquivistas Brasileiros, Secretaria de Estado da Cultura, 1996. 142 p.
CAMARGO, A. M. de A.; BELLOTTO, H. L. **Dicionário de Terminologia Arquivística**. 3. ed. São Paulo: Associação de Arquivistas de São Paulo, 2012.

CANTARA, L. Long-term preservation of digital humanities scholarship. **OCLC Systems and Services**, [S. l.], v. 22, n. 1, p. 38–42, 2006. Disponível em: <https://doi.org/10.1108/10650750610640793>. Acesso em: 07 set. 2020.

CASTELLS, M. **A sociedade em rede: a era da informação: economia, sociedade e cultura**. Tradução Roneide Venâncio Majer. 6. ed. São Paulo: Paz e Terra, 2002.

CCSDS – Consultative Committee for Space Data System. **Reference Model for an Open Archival Information System (OAIS)**. Blue book (CCSDS 650.0-B-1). Washington, DC: CCSDS, 2002. Disponível em <http://public.ccsds.org/publications/archive/650x0b1.pdf>. Acesso em: 06 set. 2020.

COCCIOLO, A. Email as cultural heritage resource: appraisal solutions from an art museum context. **Records Management Journal**, [s.l.], v. 26, n° 1, p. 68–82, 2016. ISSN: 09565698, DOI: 10.1108/RMJ-04-2015-0014.

COHEN, F. B. Digital diplomatics and forensics: going forward on a global basis. **Records Management Journal**, [s.l.], v. 25, n° 1, p. 21–44, 2015. ISSN: 09565698, DOI: 10.1108/RMJ-03-2014-0016.

CONSELHO NACIONAL DE ARQUIVOS (BRASIL). Câmara Técnica de Documentos Eletrônicos. **e-ARQ Brasil: Modelo de Requisitos para Sistemas Informatizados de Gestão Arquivística de Documentos**. Rio de Janeiro: Arquivo Nacional, 2011. Disponível em: <http://www.siga.arquivonacional.gov.br/images/publicacoes/e-arq.pdf>. Acesso em: 26 mar. 2020.

CONSELHO NACIONAL DE ARQUIVOS (BRASIL). Câmara Técnica de Documentos Eletrônicos. **Resolução nº. 43**, de 4 de setembro de 2015. Estabelece diretrizes para a implementação de repositórios arquivísticos digitais confiáveis para o arquivamento e manutenção de documentos arquivísticos digitais em suas fases corrente, intermediária e permanente, dos órgãos e entidades integrantes do Sistema Nacional de Arquivos - SINAR. Disponível em: <http://conarq.gov.br/resolucoes-do-conarq/281-resolucao-n-39-de-29-de-abril-de-2014.html>. Acesso em: 26 mar. 2020.

CONWAY, P. Digital transformations and the archival nature of surrogates. **Archival Science**, [s.l.], v. 15, n° 1, p. 51–69, 2015. ISSN: 13890166, DOI: 10.1007/s10502-014-9219-z.

CROFTS, N. *et al.* Guidelines for the selection of digital heritage for long-term preservation. **Unesco**, [s. l.], p. 1–19, 2016. Disponível em: <https://unesdoc.unesco.org/ark:/48223/pf0000244280?posInSet=19&queryId=58499698-7748-4630-a78f-c502d470a74a>. Acesso em: 08 abr. 2020.

CUNNINGHAM, A. Digital curation/digital archiving: a view from the National Archives of Australia. **The American Archivist**, v. 71: p. 530–543, 2008. Disponível em: <https://meridian.allenpress.com/american->

[archivist/article/71/2/530/24117/Digital-Curation-Digital-Archiving-A-View-from-the](#). Acesso em: 08 abr. 2020.

DALLAS, C. Digital curation beyond the “wild frontier”: a pragmatic approach. **Archival Science**, v. 16, n^o 4, p. 421–457, 2015. ISSN: 1389-0166, DOI: 10.1007/s10502-015-9252-6.

DAPPERT, A. *et al.* Describing and preserving digital object environments. **New Review of Information Networking**, [s. l.], v. 18, n. 2, p. 106–173, 2013. Disponível em: <https://doi.org/10.1080/13614576.2013.842494>. Acesso em: 07 set. 2020.

DAY, M. Metadata for digital preservation: A review of recent developments. **Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)**, [s. l.], v. 2163, p. 161–172, 2001. Disponível em: https://doi.org/10.1007/3-540-44796-2_15. Acesso em: 07 set. 2020.

DAY, M. Preservation metadata. **UKOLN**, University of Bath, 2003. Disponível em: <http://www.ukoln.ac.uk/metadata/publications/ilyim-2003/>. Acesso em: 07 set. 2020.

DIGITAL PRESERVATION WORKING GROUP, UNIVERSITY LIBRARY, GLASGOW UNIVERSITY. **Glasgow University Digital Preservation Policy - Version 3.8**. Glasgow: [s. n.], 2020. Disponível em: https://www.gla.ac.uk/media/Media_598622_smxx.pdf. Acesso em: 07 set. 2020.

DURANTI, L. The Odyssey of Records Managers. **Records Management Quarterly**. v. 23, n. 3, p. 3–11, 1989.

DURANTI, L. **Diplomatics**: new uses for an old science. Society of American Archivists: Maryland, 1998. 186 p.

DURANTI, L. The Concept of Electronic Record. In: EASTWOOD, T.; MACNEIL, H. **Preservation of the integrity of electronic records**. Dordrecht: Kluwer Academic, 2002. p. 9–22.

DURANTI, L. (Ed.). **Creator guidelines**: making and maintaining digital materials: guidelines for individuals. Vancouver: InterPARES 2 Project, [2006]. 21 p.

DURANTI, L. **Diplomatics**. Encyclopedia of Library and Information Science. New York: Marcel Dekker, p. 1–9, 2009a.

DURANTI, L. From digital diplomatics to digital records forensics. **Archivaria**, n. 68, p. 39–66, 2009b.

DURANTI, L. Structural and formal analysis: the contribution of diplomatics to archival appraisal in the digital environment. In: DURANTI, L. **The Future of Archives and Recordkeeping**: A Reader. London: Facet, 2010. p. 65–68.

DURANTI, L.; THIBODEAU, K. The Concept of record in interactive, experiential and dynamic environments: the view of InterPARES. In: DURANTI, L.; PRESTON, R. **International research on permanent authentic records in electronic systems (InterPARES 2)**: experimental, interactive and dynamic records. Italia: ANAI, 2008. p. 402-459.

DURANTI, L.; PRESTON, R. **International research on permanent authentic records in electronic systems (InterPARES 2)**: experimental, interactive and dynamic records. Italia: ANAI, 2008. 844 p.

ENCONTRO NACIONAL DE ARQUIVISTAS DAS INSTITUIÇÕES FEDERAIS DE ENSINO SUPERIOR, 6., 2019, Goiânia. **Anais**. Sigla do evento: ENARQUIFES. (no prelo).

FENG, H. Identity and archives: return and expansion of the social value of archives. **Archival Science**, v. 17, n^o 2, p. 97–112, 2017. ISSN: 13890166, DOI: <http://dx.doi.org/10.1007/s10502-016-9271-y>.

FORMENTON, D. *et al.* Metadata standards as a technological resource to guarantee digital preservation. **Biblios – Revista de bibliotecologia y Ciencias de la Informacion**, USA, n. 68, p. 82–95, 2017. Disponível em: <https://doi.org/10.5195/biblios.2017.414>. Acesso em: 07 set. 2020.

FORCE, D. C.; ZHANG, J. Knowledge discovery from within: An examination of records management and electronic records management syllabi. **Records Management Journal**, [s.l.], v. 26, n^o 3, p. 259–278, 2016. ISSN: 09565698, DOI: 10.1108/RMJ-11-2015-0034.

GARMENDIA, J. **El archivo eletrônico en los sistemas de Reino Unido y Portugal**. (1 vídeo 46 min.) II Congreso Nacional de Archivo y Documento Eletrónico, Madri. 5 fev. 2020. Disponível: <https://youtu.be/CYjOdnUqusg>. Acesso em: 22 mar. 2020.

GIL, A. C. **Como Elaborar Projetos de Pesquisa**. 5^a ed. São Paulo: Atlas, 2010.

GUO, W. *et al.* Archives as a trusted third party in maintaining and preserving digital records in the cloud environment. **Records Management Journal**, [s.l.], v. 26, n^o 2, p. 170–184, 2016. ISSN: 09565698, DOI: 10.1108/RMJ-07-2015-0028.

HAAK, L. L. *et al.* ORCID: A system to uniquely identify researchers. **Learned Publishing**, [s. l.], v. 25, n. 4, p. 259–264, 2012. Disponível em: <https://onlinelibrary.wiley.com/doi/epdf/10.1087/20120404>. Acesso em: 09 mai. 2020.

HENTTONEN, P. Privacy as an archival problem and a solution. **Archival Science**, v. 17, n^o 3, p. 285–303, 2017. ISSN: 13890166, DOI: <http://dx.doi.org/10.1007/s10502-017-9277-0>.

HOSKINS, A. Digital network memory. In: ERLL, A; RIGNEY, A (eds) **Mediation, remediation, and the dynamics of cultural memory**. De Gruyter: Berlin, 2009. p. 91–106.

HOFMAN, D. *et al.* Trust in the balance: Data protection laws as tools for privacy and security in the cloud. **Algorithms**, [s. l.], v. 10, n. 2, 2017. Disponível em: <https://www.mendeley.com/catalogue/api/pdf/c7633a3d-0cd9-3c71-bac6-14245599d57d/?doi=10.3390/a10020047>. Acesso em: 08 abr. 2020.

HUVILA, I. The unbearable lightness of participating? Revisiting the discourses of “participation” in archival literature. **Journal of Documentation**, [s.l.], v. 71, n^o 2, p. 358–386, 2015. ISSN: 00220418, DOI: 10.1108/JD-01-2014-0012.

INTERNATIONAL COUNCIL ON ARCHIVES (ICA). **Electronic Records: A Workbook for Archivists**. ICA Study 16. International Council on Archives. 2005. Disponível em: https://www.ica.org/sites/default/files/ICA_Study-16-Electronic-records_EN.pdf. Acesso em 10 abr. 2020.

ISO. International Organization for Standardization. **ISO 15489-1**: Information and documentation - Records management – Part 1: General. Geneva/Switzerland: ISO, 2001.

KATUU, S. Managing digital records in a global environment. **The Electronic Library**, [s.l.], v. 34, n^o 5, p. 869–894, 2016. ISSN: 02640473, DOI: 10.1108/EL-04-2015-0064.

KLARELD, A. S. Isn't it information assets we're really talking about?: a discourse analysis of a panel discussion on digital archives. **Archives and Records**, [s.l.], v. 36, n^o 2, p. 167, 2015. ISSN: 23257962.

KLARELD, A. S. Recordkeeping in an outsourcing public agency. **Records Management Journal**, Bradford, v. 28, n^o 1, p. 99–114, 2018. ISSN: 09565698.

KRIESBERG, A. The future of access to public records? Public-private partnerships in US state and territorial archives. **Archival Science**, [s.l.], v. 17, n^o 1, p. 5–25, 2017. ISSN: 13890166, DOI: 10.1007/s10502-016-9268-6.

LAVOIE, B.; GARTNER, R. **Preservation Metadata**. OCLC, September 2005. Disponível em: <http://www.dpconline.org/docs/reports/dpctw05-01.pdf>. Acesso em: 05 set. 2020.

LE COADIC, Y. F. **A Ciência da Informação**. Tradução Maria Yeda S. F. de Filgueiras Gomes. Brasília: Briquet de Lemos, 1996.

LEVY, P. **Cibercultura**. São Paulo: Editora 34, 1999. Disponível em: <https://mundonativodigital.files.wordpress.com/2016/03/cibercultura-pierre-levy.pdf>. Acesso em: 12 nov. 2019.

LI, C.; SUGIMOTO, S. Provenance description of metadata using PROV with PREMIS for long-term use of metadata. **Proceedings of the International Conference on Dublin Core and Metadata Applications**. Dublin Core metadata initiative, 2014. p. 147–156. Disponível em: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85013660318&partnerID=40&md5=ca8571fafb61e76760ff385b7df0097f>. Acesso em: 07 set. 2020.

LUZ, C.; FLORES, D. Cadeia de custódia e de preservação: autenticidade nas plataformas de gestão e preservação de documentos arquivísticos. **Researchgate**. 2017. Disponível em: https://www.researchgate.net/profile/Charley_Luz/publication/325225229_Cadeia_de_custodia_e_de_preservacao_autenticidade_nas_plataformas_de_gestao_e_preservacao_de_documentos_arquivisticos/links/5afecd590f7e9b33d1284f84/Cadeia-de-custodia-e-de-preservacao-autenticidade-nas-plataformas-de-gestao-e-preservacao-de-documentos-arquivisticos.pdf. Acesso em: 21 mar. 2020.

MACNEIL, H. **Trusting records**: legal, historical and diplomatic perspectives. Dordrech: Kluwer Academic, 2000. 163 p.

MACNEIL, H. et al. Part One: establishing and maintaining trust in electronic records: authenticity task force report. In: DURANTI, L. (Ed.). **The InterPARES Project**: the long-term preservation of the authentic electronic records: findings of the InterPARES Project. Italia: Archilab, 2005. p. 20-63.

MARCONI, M. de A. LAKATOS, E. M. **Metodologia Científica**. 5ª ed. São Paulo: Atlas, 2011.

MATTERLART, A. **História da sociedade da informação**. Ed. Loyola: São Paulo, 2006.

MCKEMMISH, S. **Placing Records Continuum Theory and Practice**. Achival Science. Netherland: v. 1, p. 333 – 359, 2001.

MEADOW, C.; YUAN, W. Measuring the impact of information: defining the concepts. **Information processing & management**, v. 33, n. 6, p. 697-714, 1997.

MINISTÉRIO DA EDUCAÇÃO (BRASIL). **Portaria nº. 315**, de 4 de abril de 2018. Dispõe sobre os procedimentos de supervisão e monitoramento de instituições de educação superior integrantes do sistema federal de ensino e de cursos superiores de graduação e de pós-graduação lato sensu, nas modalidades presencial e a distância. Disponível em: <http://www.in.gov.br/web/guest/materia/->

[/asset_publisher/Kujrw0TZC2Mb/content/id/9177560/do1-2018-04-05-portaria-n-315-de-4-de-abril-de-2018-9177556](#). Acesso em: 27 mar. 2020.

MINISTÉRIO DA EDUCAÇÃO (BRASIL). **Portaria nº. 330**, de 5 de abril de 2018. Dispõe sobre a emissão de diplomas em formato digital nas instituições de ensino superior pertencentes ao sistema federal de ensino. Disponível em: http://www.in.gov.br/materia/-/asset_publisher/Kujrw0TZC2Mb/content/id/9365055/do1-2018-04-06-portaria-n-330-de-5-de-abril-de-2018-9365051. Acesso em: 27 mar. 2020.

MINISTÉRIO DA EDUCAÇÃO (BRASIL). **Portaria nº. 554**, de 11 de março de 2019. Dispõe sobre a emissão e o registro de diploma de graduação, por meio digital, pelas Instituições de Ensino Superior - IES pertencentes ao Sistema Federal de Ensino. Disponível em: http://www.in.gov.br/materia/-/asset_publisher/Kujrw0TZC2Mb/content/id/66544171/do1-2019-03-12-portaria-n-554-de-11-de-marco-de-2019-66543842. Acesso em: 27 mar. 2020.

MINISTÉRIO DA EDUCAÇÃO (BRASIL). **Portaria nº. 332**, de 13 de março de 2020. Dispõe sobre a alteração no prazo contido no caput do art. 45 da Portaria nº 315, de 4 de abril de 2018. Disponível em: <https://abmes.org.br/arquivos/legislacoes/Portaria-mec-332-2020-03-13.pdf>. Acesso em: 27 mar. 2020.

MINISTÉRIO DO PLANEJAMENTO, DESENVOLVIMENTO E GESTÃO (BRASIL). **Portaria nº. 9**, de 1 de agosto de 2018. Cria o Assentamento Funcional Digital - AFD no âmbito dos órgãos e entidades integrantes do Sistema de Pessoal Civil da Administração Federal. Disponível em: http://www.in.gov.br/materia/-/asset_publisher/Kujrw0TZC2Mb/content/id/35314259/do1-2018-08-03-portaria-n-9-de-1-de-agosto-de-2018-35314252. Acesso em: 27 mar. 2020.

MIRANDA, R. C. R. Conhecimento Estratégico: Caracterização e Identificação Baseada em Revisão Sistemática e Bibliométrica. **Informação & Sociedade**, João Pessoa, v. 28, n. 1, p. 23-34, 2018. Disponível em: <http://www.periodicos.ufpb.br/ojs2/index.php/ies/article/view/30805>. Acesso em: 22 jul. 2019.

MOHER, D. *et al.* Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. **PLOS Medicine**, v. 6, n. 4, p. 1-6, 2009. Disponível em: <http://journals.plos.org/plosmedicine/article/file?id=10.1371/journal.pmed.1000097&type=printable>. Acesso em: 18 jul. 2019.

MARTINS, J. M. **Gestão do conhecimento**: criação e transferência do conhecimento. Lisboa: Sílabo, 2010.

MOSS, M.; THOMAS, D.; GOLLINS, T. Artificial Fibers — The Implications of the Digital for Archival Access. **Frontiers in Digital Humanities**, [s. l.], v. 5, n. August, p. 1–9, 2018. Disponível em:

<https://www.mendeley.com/catalogue/api/pdf/b59f924d-18ec-3df4-b861-0a0fe73b92a9/?doi=10.3389/fdigh.2018.00020>. Acesso em: 08 abr. 2020.

NGOEPE, M.; KATUU, S. Provision of Records Created in Networked Environments in the Curricula of Institutions of Higher Learning in Africa}. **New Review of Information Networking**, [s.l.], v. 22, n° 1, p. 1–12, 2017. ISSN: 13614576, DOI: 10.1080/13614576.2017.1297731.

NGOEPE, M.; NKWE, M. Separating the wheat from the chaff with the winnowing fork: The eenymeenymynymo appraisal approach of digital records in South Africa. **Records Management Journal**, [s.l.], v. 28, n° 2, p. 130–142, 2018. ISSN: 09565698, DOI: 10.1108/RMJ-09-2017-0027.

NIU, J. Event-based archival information organization. **Archival Science**, [s.l.], v. 15, n° 3, p. 315–328, 2015. ISSN: 13890166, DOI: 10.1007/s10502-014-9222-4.

NIU, J. Original order in the digital world. **Archives and Manuscripts**, [s.l.], v. 43, n° 1, p. 61–72, 2015. ISSN: 01576895, DOI: 10.1080/01576895.2014.958863.

NKHOLEDZENI, S. N. The role of archives and records management legislation after colonialism in Africa. **Records Management Journal**, v. 29, n° 1/2, p. 210–223, 2019. ISSN: 09565698, DOI: <http://dx.doi.org/10.1108/RMJ-09-2018-0024>.

PAN, W. The implementation of electronic recordkeeping systems. **Records Management Journal**, [s.l.], v. 27, n° 1, p. 84–98, 2017. ISSN: 09565698, DOI: 10.1108/RMJ-04-2016-0014.

PEARCE-MOSES, R. **Glossary of Archival and Records Terminology** (Archival Fundamentals Series II). Society of American Archivists, Chicago, 2005. Disponível em: <http://files.archivists.org/pubs/free/SAA-Glossary-2005.pdf>. Acesso em: 22 mar. 2020.

PRODANOV, C. C.; FREITAS, E. C. **Metodologia do trabalho científico: métodos e técnicas da pesquisa e do trabalho acadêmico**. 2ª ed. Novo Hamburgo: Feevale, 2013. ISBN 978-85-7717-158-3. Disponível em: <http://www.feevale.br/Comum/midias/8807f05a-14d0-4d5b-b1ad-1538f3aef538/E-book%20Metodologia%20do%20Trabalho%20Cientifico.pdf>. Acesso em: 11 abr. 2020.

RAMALHO, J. C. *et al.* RODA-in A generic tool for the mass creation of Submission Information Packages. In: **12th Iberian Conference on Information Systems and Technologies (CISTI)**. Nova Iorque, Estados Unidos, 2017. Disponível em: <https://dx.doi.org/10.23919/CISTI.2017.7975761>. Acesso em: 07 set. 2020.

ROUSSEAU, J. Y.; COUTURE, C. **Os fundamentos da disciplina arquivística**. Lisboa: Dom Quixote, 1998. 356 p.

RIES, T. The rationale of the born-digital dossier genetique: Digital forensics and the writing process: With examples from the Thomas Kling Archive. **DIGITAL SCHOLARSHIP IN THE HUMANITIES**, ENGLAND, v. 33, n. 2, p. 391–424, 2018. Disponível em: <https://doi.org/10.1093/llc/fqx049>. Acesso em: 07 set. 2020.

RODRIGUES, A. de J. **Metodologia científica**. 1. Ed. São Paulo: Avercamp: 2006.

ROGERS, C. Diplomatics of born digital documents - considering documentary form in a digital environment. **Records Management Journal**, [s.l.], v. 25, nº 1, p. 6–20, 2015. ISSN: 09565698, DOI: 10.1108/RMJ-03-2014-0021.

RONDINELLI, R. C. **O Conceito de documento arquivístico frente à realidade digital: uma revisitação necessária**. 2011. Tese (Doutorado em Ciência da Informação) – Universidade Federal Fluminense, Niterói, 2011. 270 f.

RONDINELLI, R. C. **O conceito de documento arquivístico em ambiente digital**. São Paulo, 2013. Disponível em: <http://www.hu.usp.br/wp-content/uploads/sites/39/2015/01/rosely.pdf>. Acesso em: 11 fev. 2020.

SANTOS, V. B. dos. Preservação de documentos arquivísticos digitais. **Ciência da Informação**, v. 41, n. 1, 2012. DOI: 10.18225/ci.inf.v41i1.1357. Acesso em: 10 abr. 2020.

SAYÃO, L. F. Uma outra face dos metadados: informações para a gestão da preservação digital. **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, [S. l.], p. 1–31, 2010. Disponível em: <https://doi.org/10.5007/1518-2924.2010v15n30p1>. Acesso em: 07 set. 2020.

SAWYER, D.; REICH, L. **Reference model for an open archival information system (OAIS): partnerships in innovation: serving a networked nation**. 2004. Disponível em: www.archives.gov/electronic_records_archives/pdf/it_conference_sawyer.pdf. Acesso em: 15 mai. 2020.

SCHELLENBERG, T. R. **Modern archives: principles and techniques**. Chicago: University of Chicago Press, 1956. 247 p.

SCHWAITZER, L. de B. da S. O arquivista na aurora digital: diálogo com Bruno Delmas. **Revista do Arquivo**, n. 6, p. 9-15, Abr. 2018. ISSN: 2447-908X. Disponível em: http://www.arquivoestado.sp.gov.br/revista_do_arquivo/06/ensaio.php#inicio_artigo. Acesso em: 10 abr. 2020.

SLOYAN, V. Born-digital archives at the Wellcome Library: appraisal and sensitivity review of two hard drives. **Archives and Records**, [s.l.], v. 37, nº 1, p. 20–36, 2016. ISSN: 23257962, DOI: 10.1080/23257962.2016.1144504.

SILVA, P. F. C. da. **Diagnóstico Documental do Centro de Ciências Jurídicas da Universidade Federal da Paraíba**. João Pessoa, 2018. Disponível em:

<https://sipac.ufpb.br/public/verArquivoDocumento?idArquivo=1015768&key=ad6bc0472fdc44d167db3eedc39dbcfb&idDocumento=743140&downloadArquivo=true&publicPath=true>. Acesso em: 12 abr. 2020.

SPENCE, A. *et al.* Case study: the University of Glasgow's digital preservation journey 2017-2019. **Insights**, Research Information Administrator: digital preservation University, v. 32, p. 1–9, 2019. Disponível em: <https://search.proquest.com/docview/2203049840?accountid=149610>. Acesso em: 07 set. 2020.

SVÄRD, P. The woes of Swedish private archival institutions. **Records Management Journal**, [s.l.], v. 27, n^o 3, p. 275–285, 2017. ISSN: 09565698, DOI: 10.1108/RMJ-01-2016-0003. Acesso em: 13 mar. 2020

THELWALL, M.; KOUSHA, K. Figshare: A universal repository for academic resource sharing? **Online Information Review**, [s. l.], v. 40, n. 3, p. 333–346, 2016. ISSN: 1468-4527 Disponível em: <https://doi.org/10.1108/OIR-06-2015-0190>. Acesso em: 09 mai. 2020.

THIOLLENT, M. **Metodologia da pesquisa-ação**. 8. ed. São Paulo: Cortez, 1998.

THURSTON, A. What lessons can be learned from the US archivist's digital mandate for 2019 and is there potential for applying them in lower resource countries? **Records Management Journal**, [s.l.], v. 26, n^o 2, p. 206–217, 2016. ISSN: 0956-5698, DOI: 10.1108/RMJ-04-2015-0015.

TOGNOLI, N. B. *et al.* Definindo o conhecimento arquivístico: estruturas conceituais. **Informação & Informação**, v. 24, n. 2, p. 58-75, 2019. DOI: 10.5433/1981-8920.2019v24n2p58. Acesso em: 09 mai. 2020.

VASSILAKAKI, E.; MONIAROU-PAPACONSTANTINO, V. Beyond preservation: investigating the roles of archivist. **Library Review**, [s.l.], v. 66, n^o 3, p. 110–126, 2017. DOI: 10.1108/LR-09-2016-0077.

YEO, G. Concepts of record: evidence, information and persistent representation. **The American Archivists**, v. 70, p. 315-343, 2007.

YEO, G. Concepts of record: prototypes and boundary objects. **The American Archivists**, v. 71, p. 118-143, 2008.

ZHANG, J. Archival Representation in the Digital Age. **Journal of Archival Organization**, [s. l.], v. 10, n. 1, p. 45–68, 2012. Disponível em: <https://doi.org/10.1080/15332748.2012.677671>. Acesso em: 06 mar. 2020.

ZHAO, Y. *et al.* Postgraduates' personal digital archiving practices in China: problems and strategies. **The Journal of Academic Librarianship**, Chengdu, v. 45, n. 5, 2019. DOI: 10.1016/J.ACALIB.2019.06.002.

	Sector	Sub-área	Gênero	Espécie	Função	Atividade	Rondinelli (2011)	Katu (2016)	Zhao et al (2019)	Garmedia (2020)
1	Direção de Centro	Gestão de Pessoas	Elétrônico	Ofício	Unidade de servidor	Melo	SIPAC	registros gerados por sistemas de informações de negócios	documentos digitais	registros nato-digítals
2	Direção de Centro	Gestão de Pessoas	Elétrônico	Ofício	Informação da Direção	Melo	SIPAC	registros gerados por sistemas de informações de negócios	documentos digitais	registros nato-digítals
3	Direção de Centro	Gestão de Pessoas	Elétrônico	Ofício	Remoção de servidor	Melo	SIPAC	registros gerados por sistemas de informações de negócios	documentos digitais	registros nato-digítals
4	Direção de Centro	Gestão de Pessoas	Elétrônico	Ofício	Férs de servidor	Melo	SIPAC	registros gerados por sistemas de informações de negócios	documentos digitais	registros nato-digítals
5	Direção de Centro	Gestão de Pessoas	Elétrônico	Ofício	Conehos	Melo	SIPAC	registros gerados por sistemas de informações de negócios	documentos digitais	registros nato-digítals
6	Direção de Centro	Gestão de Pessoas	Elétrônico	Ofício	Viculos	Melo	SIPAC	registros gerados por sistemas de informações de negócios	documentos digitais	registros nato-digítals
7	Direção de Centro	Gestão de Pessoas	Elétrônico	Ofício	Pórtaro de funcionamento	Melo	SIPAC	registros gerados por sistemas de informações de negócios	documentos digitais	registros nato-digítals
8	Direção de Centro	Gestão de Pessoas	Elétrônico	Ofício Circular	Publicação de edital	Melo	SIPAC	registros gerados por sistemas de informações de negócios	documentos digitais	registros nato-digítals
9	Direção de Centro	Gestão de Pessoas	Elétrônico	Ofício Circular	Processo Administrativo Disciplinar	Melo	SIPAC	registros gerados por sistemas de informações de negócios	documentos digitais	registros nato-digítals
10	Direção de Centro	Gestão de Pessoas	Elétrônico	Declaração	Sindicância	Melo	SIPAC	registros gerados por sistemas de informações de negócios	documentos digitais	registros nato-digítals
11	Direção de Centro	Gestão de Pessoas	Elétrônico	Declaração	Anculo	Melo	SIPAC	registros gerados por sistemas de informações de negócios	documentos digitais	registros nato-digítals
12	Direção de Centro	Gestão de Pessoas	Elétrônico	Declaração	Controle de Docentes por Regime	Melo	SIPAC	registros gerados por sistemas de informações de negócios	documentos digitais	registros nato-digítals
13	Direção de Centro	Gestão de Pessoas	Elétrônico	Lisagem	Controle de Matrimonios	Melo	SIPAC	registros gerados por sistemas de informações de negócios	documentos digitais	registros nato-digítals
14	Direção de Centro	Gestão de Pessoas	Elétrônico	Portaria	Designação de Comissões	Melo	SIPAC	registros gerados por sistemas de informações de negócios	documentos digitais	registros nato-digítals
15	Direção de Centro	Gestão de Pessoas	Elétrônico	Portaria	Designação de Funções Gratificadas	Melo	SIPAC	registros gerados por sistemas de informações de negócios	documentos digitais	registros nato-digítals
16	Direção de Centro	Gestão de Pessoas	Elétrônico	Portaria	Designação de Funções Não Estruturadas	Melo	SIPAC	registros gerados por sistemas de informações de negócios	documentos digitais	registros nato-digítals
17	Direção de Centro	Gestão de Pessoas	Elétrônico	Portaria	Estágio	Melo	SIPAC	registros gerados por sistemas de informações de negócios	documentos digitais	registros nato-digítals
18	Direção de Centro	Gestão de Pessoas	Elétrônico	Edital	Eleições de representantes nos conselhos	Melo	SIPAC	registros gerados por sistemas de informações de negócios	documentos digitais	registros nato-digítals
19	Direção de Centro	Gestão de Pessoas	Elétrônico	Edital	Férias	Melo	SIPAC	registros gerados por sistemas de informações de negócios	documentos digitais	registros nato-digítals
20	Direção de Centro	Gestão de Pessoas	Elétrônico	Planilha	Plano Planos Setoriais de Trabalho	Melo	SIPAC	registros gerados por sistemas de informações de negócios	documentos digitais	registros nato-digítals
21	Direção de Centro	Gestão de Pessoas	Elétrônico	Plano	Planos Individuais de Trabalho	Melo	SIPAC	registros gerados por sistemas de informações de negócios	documentos digitais	registros nato-digítals
22	Direção de Centro	Gestão de Pessoas	Elétrônico	Plano	Comunicações Diversas	Melo	SIPAC	registros gerados por sistemas de informações de negócios	documentos digitais	registros nato-digítals
23	Direção de Centro	Gestão de Pessoas	Elétrônico	Email	Solicitações de documentos para processos	Melo	SIPAC	sistemas de comunicação digital	Email	registros nato-digítals
24	Direção de Centro	Gestão de Pessoas	Elétrônico	Email	Terno	Melo	SIPAC	registros gerados por sistemas de informações de negócios	documentos digitais	registros nato-digítals
25	Direção de Centro	Gestão de Pessoas	Elétrônico	Cartão	Homologação de processo eleitoral	Melo	SIPAC	registros gerados por sistemas de informações de negócios	documentos digitais	registros nato-digítals
26	Direção de Centro	Gestão de Pessoas	Elétrônico	Cartão	Negativa de processo administrativo	Melo	SIPAC	registros gerados por sistemas de informações de negócios	documentos digitais	registros nato-digítals
27	Direção de Centro	Gestão de Pessoas	Elétrônico	Cartão	Sindicância	Melo	SIPAC	registros gerados por sistemas de informações de negócios	documentos digitais	registros nato-digítals
28	Direção de Centro	Gestão de Pessoas	Elétrônico	Despacho	Instruções processuais	Melo	SIPAC	registros gerados por sistemas de informações de negócios	documentos digitais	registros nato-digítals
29	Direção de Centro	Gestão de Pessoas	Elétrônico	Processo	Seleção de estágio	Melo	SIPAC	registros gerados por sistemas de informações de negócios	documentos digitais	registros nato-digítals
30	Direção de Centro	Gestão de Pessoas	Elétrônico	Processo	Eleição de representante do conselho de centro	Melo	SIPAC	registros gerados por sistemas de informações de negócios	documentos digitais	registros nato-digítals
31	Direção de Centro	Gestão de Pessoas	Elétrônico	Processo	Eleição de representante do CONSEP	Melo	SIPAC	registros gerados por sistemas de informações de negócios	documentos digitais	registros nato-digítals
32	Direção de Centro	Gestão de Pessoas	Elétrônico	Processo	Eleição de representante do CONSUI	Melo	SIPAC	registros gerados por sistemas de informações de negócios	documentos digitais	registros nato-digítals
33	Direção de Centro	Gestão de Pessoas	Elétrônico	Processo	Progressão Docente	Melo	SIPAC	registros gerados por sistemas de informações de negócios	documentos digitais	registros nato-digítals
34	Direção de Centro	Gestão de Pessoas	Elétrônico	Processo	Progressão Técnico-administrativo	Melo	SIPAC	registros gerados por sistemas de informações de negócios	documentos digitais	registros nato-digítals
35	Direção de Centro	Gestão de Pessoas	Elétrônico	Formulário	Homologação de frequência do servidor	Melo	SIGRH	registros gerados por sistemas de informações de negócios	documentos digitais	registros nato-digítals
36	Direção de Centro	Gestão de Pessoas	Elétrônico	Formulário	Cadastro de jornada de trabalho	Melo	SIGRH	registros gerados por sistemas de informações de negócios	documentos digitais	registros nato-digítals
37	Direção de Centro	Gestão de Pessoas	Elétrônico	Formulário	Documentação digital interativo não dinâmico	Melo	SIGRH	registros gerados por sistemas de informações de negócios	documentos digitais	registros nato-digítals
38	Direção de Centro	Gestão de Pessoas	Elétrônico	Formulário	Documentação digital interativo não dinâmico	Melo	SIGRH	registros gerados por sistemas de informações de negócios	documentos digitais	registros nato-digítals

APÊNDICE B – QUESTIONÁRIO PARA AS UNIVERSIDADES FEDERAIS

Este pedido de informação faz parte de uma pesquisa de mestrado sobre a gestão dos registros e documentos digitais das universidades com foco na preservação digital, orientada pelo Professor Dr. Wagner Junqueira de Araújo, pertencente ao Programa de Pós-Graduação em Ciência da Informação da Universidade Federal da Paraíba.

Se você aceita participar desta pesquisa, solicito que responda as seis questões a seguir sobre os metadados dos registros e documentos digitais da sua instituição.

CONTEXTO

Considerando a necessidade de atendimento da Lei nº 8.159/1991, do decreto 4.073/2002, e das resoluções do Conselho Nacional de Arquivos nos. 20/2004, 24/2006, 25/2007, 32/2010, 36/2012, 37/2012, 40/2014 e 43/2015, em relação aos registros e documentos criados e/ou inseridos em ambiente digital nos sistemas SIPAC e SIGAA (ou nos sistemas de produção de documentos acadêmicos e administrativos), solicito as seguintes informações:

SOBRE A IDENTIFICAÇÃO DA IFES

NOME DA UNIVERSIDADE:

SETOR RESPONSÁVEL PELA RESPOSTA:

SOBRE OS METADADOS DE PRESERVAÇÃO DIGITAL DE FORMA GERAL

1) Quais metadados são criados e armazenados pela sua instituição para preservação dos registros e documentos produzidos nos sistemas SIPAC e SIGAA (ou nos sistemas de produção de documentos acadêmicos e administrativos)?

SOBRE OS METADADOS GERADOS PELO SISTEMA DE FORMA AUTOMÁTICA

2) A sua instituição armazena metadados que identifiquem a origem e a localização dos usuários que produzem os registros e documentos? Caso sim, quais metadados são armazenados?

3) A sua instituição armazena metadados que diferenciem os registros e documentos gerados em ambiente digital (nato digitais) daqueles convertidos para o meio eletrônico (digitalizados)? Caso sim, quais metadados são armazenados em cada caso?

4) A sua instituição armazena metadados que descrevam o dispositivo e o sistema operacional no qual os registros e documentos foram produzidos? Caso sim, quais metadados são armazenados a respeito do dispositivo e do sistema operacional?

5) A sua instituição armazena metadados que descrevam os dados constantes nos registros e documentos criados/inseridos nos sistemas SIGAA/SIPAC (ou nos sistemas de produção de documentos acadêmicos e administrativos)? Caso sim, quais são os metadados armazenados para descrição dos dados?

SOBRE OS METADADOS INSERIDOS MANUALMENTE PELO USUÁRIO

6) Existem metadados que descrevam os dados constantes nos registros e documentos criados/inseridos nos sistemas SIGAA/SIPAC (ou nos sistemas de produção de documentos acadêmicos e administrativos) que são incluídos pelo usuário de forma manual? Caso sim, quais são os metadados armazenados para descrição dos dados inseridos manualmente pelo usuário?

[illegible]

APÊNDICE D – QUESTIONÁRIO DO DIAGNÓSTICO DOCUMENTAL DO CCJ

Por quanto tempo a documentação é mantida no setor? Após o prazo de utilização mais frequente, qual é o procedimento realizado? Caso os documentos sejam transferidos, para qual unidade são enviados? Existem listas de transferência ou eliminação de documentos? Arquivista Quanto metros lineares existem? Qual a área total disponível para armazenamento? Da área disponível, quanto já foi utilizada? 4. ACESSO Entrevistado Quais documentos são mais buscados? Quais documentos são menos buscados? Quem é o principal utilizador dos documentos desta unidade? Existe algum recurso (informático ou não) para facilitar a recuperação de documentos? Existe alguma sugestão de instrumento de pesquisa que poderia ser utilizado? Que elementos descritores não poderiam faltar num sistema de recuperação de informações? Qual a natureza dos documentos de arquivo desta unidade? (ostensivo, sigilosos) Existe algum procedimento para o acesso aos documentos? (norma, manual, regra) 5. CONDIÇÕES DE ARMAZENAMENTO DO ACERVO Entrevistado A documentação fica próxima a instalações elétricas? A documentação fica próxima a instalações hidráulicas? A documentação fica próxima a instalações sanitárias? A documentação fica exposta a incidência direta de luz solar? Existem problemas como mofo ou fungos no ambiente? Existe ar condicionado no espaço destinado aos documentos? Existe proteção contra roubo de documentos? Existe proteção contra incêndio no ambiente? Qual a frequência de limpeza do ambiente? Qual a frequência de detetização do ambiente? 6. OUTROS Entrevistado Que tipo de ações já foram realizadas com os documentos desta unidade? Existe alguma ação já programada para os documentos desta unidade? Quais são os servidores responsáveis pelo arquivamento dos documentos? Quais são os servidores que possuem acesso liberado ao acervo? Centro de Ciências Jurídicas - CCJ - Universidade Federal da Paraíba Jardim Cidade Universitária, s/n - Castelo Branco - João Pessoa-PB CEP : 58051-900	UNIVERSIDADE FEDERAL DA PARAIBA CENTRO DE CIÊNCIAS JURÍDICAS DIREÇÃO DE CENTRO QUESTIONÁRIO PARA DIAGNÓSTICO DOCUMENTAL DOS ARQUIVOS SETORIAIS 1. IDENTIFICAÇÃO Entrevistado Unidade: Unidade a qual está subordinada: Responsável pelo setor: Entrevistado: Localização: Telefone: Data: 2. PRODUÇÃO DOCUMENTAL Entrevistado Quais são os principais documentos gerados? Quais são os documentos avulsos gerados? Quais são os processos gerados? Existe um modelo para os documentos? Existe um workflow para processos? Existem documentos produzidos em ambiente digital? Caso sim, quais são e quais sistemas são utilizados? Existem documentos recebidos em ambiente digital? Existem documentos pré-prints [formulários]? Arquivista Quais são os gêneros documentais produzidos? Quais são as espécies documentais produzidas? Quais são as tipologias documentais produzidas? 3. ARQUIVAMENTO Entrevistado Quais documentos avulsos são arquivados? Quais processos são arquivados? Que tipo de acondicionamento é utilizado? (envelope, pasta, capa, caixa) Que tipo de armazenamento é utilizado? (armário, gaveta, estante, prateleira) Como é feita a classificação dos documentos? Como é feito o arranjo documental? Qual método de arquivamento é utilizado? Centro de Ciências Jurídicas - CCJ - Universidade Federal da Paraíba Jardim Cidade Universitária, s/n - Castelo Branco - João Pessoa-PB CEP : 58051-900
---	--



Diagnóstico Documental do Centro de Ciências Jurídicas da Universidade Federal da Paraíba

Agosto 2018



Elaboração:
Pedro Felipy Cunha da Silva
Arquivista do Centro de Ciências Jurídicas da UFPB
pedrofelipy@ccj.ufpb.br

ANEXOS

ANEXO A – EXEMPLO DE RESPOSTA SOBRE METADADOS (UFPB)

02/10/2020

https://sipac.ufpb.br/publico/jsp/processos/documento_visualizacao.jsf?IdDoc=1514792

MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DA PARAÍBA

DESPACHO. Nº 46 / 2020 - STI - SE (11.00.51.07)

Nº do Protocolo: NÃO PROTOCOLADO

João Pessoa-PB, 02 de Outubro de 2020

DESPACHO

À PROPLAN - Serviço de Informação ao Cidadão

Saulo Eduardo de Carvalho Almeida

Prezado,

Ao tempo em que o cumprimentamos, segue adiante as respostas oferecidas pela Equipe de Sistema da Informação desta Superintendência relativamente aos questionamentos formulados:

1) Quais metadados são criados e armazenados pela UFPB para preservação dos registros e documentos produzidos nos sistemas SIPAC e SIGAA?

São armazenados dados sobre o usuário de criação do documento, tipo de documento, data de cadastro do documento, espécie do documento, classificação conarq, unidade em que o documento foi criado e unidade de destino, informações sobre graus de sigilo do documento, informações sobre o tipo de dispositivo utilizado.

2) A UFPB armazena metadados que identificam a origem e a localização dos usuários que produzem os registros e documentos? Caso sim, quais metadados são armazenados?

Sim. Para cada processo e documento criado são armazenadas informações sobre o usuário que produziu, contendo detalhes de login, nome, matrícula SIAPE, assim como a data de cadastro e a unidade em que o usuário estava vinculado no momento da criação do documento/processo.

3) A UFPB armazena metadados que diferenciam os registros e documentos gerados em ambiente digital (nato digitais) daqueles convertidos para o meio eletrônico (digitalizados)? Caso sim, quais metadados são armazenados em cada caso?

Sim. Para os documentos nato digitais é inserida a informação no banco de dados de que ele foi escrito, já para os documentos digitalizados é inserida a informação de que ele foi anexado, contendo ainda o caminho em que o documento se encontra, nome do documento e informações de criptografia.

4) A UFPB armazena metadados que descrevem o dispositivo e o sistema operacional no qual os registros e documentos foram produzidos? Caso sim, quais metadados são armazenados a respeito do dispositivo e do sistema operacional?

Sim. Na criação de processos e documentos são armazenadas informações sobre o dispositivo que realizou o cadastro, assim como o IP, hora do registro, tipo e versão do sistema operacional, web browser utilizado, resolução da tela/monitor.

5) A UFPB armazena metadados que descrevem os dados constantes nos registros e documentos criados/inseridos nos sistemas SIGAA/SIPAC? Caso sim, quais são os metadados armazenados para descrição dos dados?

Sim. São incluídos metadados como a extensão do documento, além disso também é possível encontrar o conteúdo do documento em formato HTML caso o documento tenha sido criado diretamente no sistema.

6) Existem metadados que descrevem os dados constantes nos registros e documentos criados/inseridos nos sistemas SIGAA/SIPAC que são incluídos pelo usuário de forma manual? Caso sim, quais são os metadados armazenados para descrição dos dados inseridos manualmente pelo usuário?

Sim, o usuário responsável pela criação do documento vai inserir informações sobre o tipo do documento, espécie do documento, classificação conarq, assunto do documento, assunto detalhado, natureza do documento (Documentos ostensivos, restritos, sigilosos), assim como a hipótese legal da restrição do documento, caso não seja ostensivo. Além disso, caso o documento não seja nato-digital e sim digitalizado, o usuário é responsável por incluir informações que indiquem se o documento é interno ou externo, a data que o documento original foi criado, a data de recebimento do documento original, o tipo de conferência (Documento original, cópia autenticada em cartório, cópia autenticada administrativamente ou cópia simples).

https://sipac.ufpb.br/publico/jsp/processos/documento_visualizacao.jsf?IdDoc=1514792

1/2