DISSERTAÇÃO DE MESTRADO PROFISSIONAL

Transformação Digital no Contexto da Inteligência de Estado: Análise e Mitigação das Vulnerabilidades do Documento Digital

Francisco Luziaro de Sousa

Orientador: Prof. Dr. Rafael Rabelo Nunes, ADM/FACE/UnB

Coorientadora: Profª. Drª. Lillian Maria Araujo de Rezende Alvares, FCI/UnB

Programa de Pós-Graduação Profissional em Engenharia Elétrica

DEPARTAMENTO DE ENGENHARIA ELÉTRICA
FACULDADE DE TECNOLOGIA
UNIVERSIDADE DE BRASÍLIA

UNIVERSIDADE DE BRASÍLIA
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA

Transformação Digital no Contexto da Inteligência de Estado: Análise e Mitigação das Vulnerabilidades do Documento Digital

Digital Transformation in the Context of State Intelligence: Analysis and Mitigation of Digital Document Vulnerabilities

Francisco Luziaro de Sousa

Orientador: Prof. Dr. Rafael Rabelo Nunes, ADM/FACE/UnB
Coorientadora: Profª. Drª. Lillian Maria Araujo de Rezende Alvares, FCI/UnB

PUBLICAÇÃO: **PPEE.MP.033**
BRASÍLIA-DF,

UNIVERSIDADE DE BRASÍLIA
Faculdade de Tecnologia

DISSERTAÇÃO DE MESTRADO PROFISSIONAL

Transformação Digital no Contexto da Inteligência de Estado: Análise e Mitigação das Vulnerabilidades do Documento Digital

Francisco Luziaro de Sousa

Orientador: Prof. Dr. Rafael Rabelo Nunes, ADM/FACE/UnB

Coorientadora: Profª. Drª. Lillian Maria Araujo de Rezende Alvares, FCI/UnB

Dissertação de Mestrado Profissional submetida ao Departamento de Engenharia Elétrica como requisito parcial para obtenção do grau de Mestre em Engenharia Elétrica

Banca Examinadora

Prof. Rafael Rabelo Nunes, Ph.D, FT/UnB _____
Presidente

Prof. Daniel Chaves Café, Dr., FT/UnB _____
Examinador interno

Prof. João José Costa Gondim, Dr., FT/UnB _____
Examinador interno

Prof. Washington Ribeiro, Dr., IBICT/MCTI _____
Examinador externo

FICHA CATALOGRÁFICA

DE SOUSA, FRANCISCO LUZIARO
Transformação Digital no Contexto da Inteligência de Estado: Análise e Mitigação das Vulnerabilidades do Documento Digital [Distrito Federal] 2022.
xvi, 51 p., 210 x 297 mm (ENE/FT/UnB, Mestre, Engenharia Elétrica, 2022).
Dissertação de Mestrado Profissional - Universidade de Brasília, Faculdade de Tecnologia.
Departamento de Engenharia Elétrica

1. Inteligência de Estado	2. Cadeia de Custódia Digital
3. Curadoria Digital	4. Repositório Digital Confiável
I. ENE/FT/UnB	II. Título (série)

REFERÊNCIA BIBLIOGRÁFICA

DE SOUSA, F.L. (2022). *Transformação Digital no Contexto da Inteligência de Estado: Análise e Mitigação das Vulnerabilidades do Documento Digital*. Dissertação de Mestrado Profissional, Departamento de Engenharia Elétrica, Universidade de Brasília, Brasília, DF, 51 p.

CESSÃO DE DIREITOS

AUTOR: Francisco Luziaro de Sousa
TÍTULO: Transformação Digital no Contexto da Inteligência de Estado: Análise e Mitigação das Vulnerabilidades do Documento Digital.
GRAU: Mestre em Engenharia Elétrica ANO: 2022

Francisco Luziaro de Sousa
Depto. de Engenharia Elétrica (ENE) - FT
Universidade de Brasília (UnB)
Campus Darcy Ribeiro
CEP 70919-970 - Brasília - DF - Brasil

DEDICATÓRIA

Dedico este trabalho a minha família, em especial a minha esposa *Viviane* e à minha mãe *Zulma*, mulheres fortes que me deram a cobertura e o suporte fundamentais para que eu pudesse me dedicar integralmente a esta pesquisa e chegar aos objetivos pretendidos.

Muito Obrigado Meus Amores!

AGRADECIMENTOS

Agradeço a ABIN por proporcionar esta oportunidade de crescimento acadêmico por meio da parceria com a Universidade de Brasília.

Agradeço aos professores e colaboradores do PPEE/UnB pela disponibilidade e assertividade com que conduziram mais esta jornada em prol do conhecimento.

Meus mais sinceros agradecimentos aos meus orientadores, *Prof. Rafael Rabelo* e *Profª Lillian Alvares*, que aceitaram o desafio de me orientar em uma temática transdisciplinar complexa que permeia a Engenharia Elétrica e a Ciência da Informação, e que me apoiaram integralmente durante a condução deste trabalho, principalmente nos momentos que mais demandaram resiliência.

RESUMO

Pesquisa de natureza aplicada, cujo objetivo é apontar medidas que assegurem as propriedades imprescindíveis ao documento de inteligência frente às vulnerabilidades advindas do documento digital. O trabalho emprega metodologia exploratória de abordagem qualitativa por meio de análise de conteúdo documental acadêmico selecionados a partir de Revisão Sistemática de Literatura combinada com método TEMAC. O estudo identifica relação de causa e efeito entre vulnerabilidades do documento digital e propriedades imprescindíveis ao documento de inteligência; por meio da descrição analítica aponta medidas eficazes na mitigação destas vulnerabilidades, entre elas a Cadeia de Custódia Digital e a Curadoria Digital previstas na hipótese de pesquisa; apresenta proposta de diretrizes aplicáveis à arquitetura de governança do órgão para a implantação das medidas de mitigação; aponta lacunas técnicas a superar para a efetiva integração entre Sistema Informatizado de Gestão Arquivística de Documento (SIGAD) e Repositório Digital Confiável (RDC-Arq); lança bases para estudos de caso voltados para implantação de repositórios digitais com certificação de confiabilidade; e sugere estudos aprofundados sobre *Long Term Digital Preservation as a Service* (LTDPaaS) como plataforma de armazenamento confiável para arquitetura de nuvem.

ABSTRACT

This research has an applied nature, its objective is to propose measures that ensure the essential properties of the intelligence document against the vulnerabilities arising from the digital document. The work employs an exploratory methodology with a qualitative approach through the analysis of documental content of academic papers selected from the Systematic Literature Review. The study identifies a relationship between vulnerabilities of the digital document and essential properties of the intelligence document; validates effective measures to mitigate these vulnerabilities, including the Digital Chain of Custody and Digital Curation foreseen in the research hypothesis; presents a proposal for guidelines applicable to the institution's governance architecture for the implementation of mitigation measures; points out technical gaps to be overcome for the effective integration between the Document Management Systems and the Trusted Digital Repositories; launches bases for case studies focused on the implantation and certification of trusted digital repositories; and suggests in-depth studies on *Long Term Digital Preservation as a Service* (LTDPaaS) as a storage platform in cloud architecture.

SUMÁRIO

LISTA DE FIGURAS

LISTA DE TABELAS

LISTA DE ABREVIATURAS E SIGLAS

ABIN	Agência Brasileira de Inteligência
AN	Arquivo Nacional
ACTDR	Audit and Certification of Trustworthy Digital Repositories
API	Application Programming Interface
BMP	Business Process Management
CAFe	Comunidade Acadêmica Federada
CCDA	Cadeia de Custódia Digital Arquivística
CCSDS	Consultative Committee for Space Data Systems
CD	Curadoria Digital
CONARQ	Conselho Nacional de Arquivos
CMIS	Content Management Interoperability Services
CMS	Content Manager System
CRL	Center for Research Libraries
CTDE	Câmara Técnica de Documentos Eletrônicos
DCC	Digital Curation Centre
ECM	Enterprise Content Manager
EDMS	Eletronic Document Management System
GED	Gerenciador Eletrônico de Documentos
GPL	General Public License
LOC	Library Of Congress
NASA	National Aeronautics and Space Administration
OAIS	Open Archival Information System
OASIS	Organization for the Advancement of Structured Information Standards
OCLC	Online Computer Library Center
POA	Planejamento Orçamentário Anual
PETIC	Plano Estratégico de Tecnologia da Informação e Comunicação
PDTIC	Plano Diretor de Tecnologia da Informação e Comunicação
RDC	Repositório Digital Confiável
RSL	Revisão Sistemática de Literatura
SEI	Sistema Eletrônico de Informações
SGD	Sistema de Gerenciamento de Documentos
SIGAD	Sistema Informatizado de Gestão Arquivística de Documento
SINAR	Sistema Nacional de Arquivos
SISBIN	Sistema Brasileiro de Inteligência
TEMAC	Teoria do Enfoque Meta Analítico Consolidado
TRAC	Trustworthy Repository Audit And Certification

1 INTRODUÇÃO

1.1 CONTEXTO DA PESQUISA

O movimento rumo ao digital, que se convencionou chamar de Transformação Digital (TD), foi uma inovação que se impôs naturalmente à realidade dos nossos dias, seguindo diferentes trajetórias com diferentes resultados, seja de forma disruptiva ou sustentada [1]. A TD ganhou impulso com a instituição do Sistema Nacional para a Transformação Digital (SinDigital) e com a formalização da Estratégia Brasileira para a Transformação Digital (E-Digital) – que difundiu a Estratégia de Governança Digital (EGD), um arcabouço normativo que estabelece diretrizes e estratégias para incremento da digitalização na relação estado-cidadão [2]. Esse arcabouço, contudo, não avança na especificação de objetivos ou metas de TD nas instituições, como a Agência Brasileira de Inteligência (ABIN).

Enquanto órgão central do Sistema Brasileiro de Inteligência (SISBIN) [3], a ABIN está à frente da Atividade de Inteligência do Estado Brasileiro, cuja missão é "produzir e difundir conhecimentos às autoridades competentes, relativos a fatos e situações que ocorram dentro e fora do território nacional, de imediata ou potencial influência sobre o processo decisório, a ação governamental e a salvaguarda da sociedade e do Estado", conforme Decreto Nº 8.793 [4].

Na ABIN, o principal processo finalístico pressupõe a coleta e análise de dados e informações para a produção de conhecimento, que é materializado em documentos de inteligência digitais. Este processo hoje conta soluções informatizadas para o armazenamento de artefatos digitais, para o compartilhamento e a elaboração colaborativa de documentos digitais nos mais variados formatos. Para que a TD seja uma inovação sustentada no órgão, deve entender suas necessidades, provocar o engajamento dos colaboradores, promover melhorias nos modelos de negócio e dos processos operacionais, conjugar tanto a digitalização como a digitização, elevando o patamar de desempenho institucional, sem, contudo, prescindir de preocupação com segurança das informações e comunicações desde o ambiente de gestão documental e arquivística ao de preservação e acesso [5].

Na visão da Ciência da Informação, onde se insere a gestão documental e arquivística, prevalece o paradigma da cadeia de custódia documental, que confere aos documentos analógicos (de suporte físico) os atributos de segurança e confiança indispensáveis à sua irrefutabilidade, enquanto fonte de prova ou de evidência documental, por meio da comprovação de sua autenticidade [6]. Segundo as definições do Conselho Nacional de Arquivos (CONARQ), propostas nas Diretrizes para a Presunção de Autenticidade de Documentos Arquivísticos Digitais [7], autenticidade é qualidade de um documento ser exatamente aquele que foi produzido, não tendo sofrido alteração, corrompimento e adulteração. Nesta visão, autenticidade deriva da combinação da identidade com a integridade, sendo identidade definida como o conjunto de atributos do documento arquivístico que o caracterizam como único, distinto dos demais; e integridade é conceituada com a qualidade do documento que mantém sua completude, livre de alteração ou adulteração indevidas.

Sob o paradigma da temporalidade documental – corrente, intermediária ou permanente, a gestão do-

cumental arquivística se volta para a guarda e a preservação do documento na fase permanente, estágio em que o documento físico era recolhido ao setor de arquivo do órgão, lá permanecendo armazenado apenas para consulta com vistas a garantir seu valor histórico, probatório e informativo [6].

Trazida para ambiente digital, a cadeia de custódia enfrenta desafios, tanto para a superação de vulnerailidades inerentes ao documento digital, quanto para o atendimento aos princípios fundamentais da segurança da informação, caros à ABIN, sintetizados pelo acrônimo CIA (*Confidentiality*, *Integrity* e *Availability*). Confidencialidade é o princípio afeto à capacidade de garantir que o nível necessário de sigilo seja aplicado em cada junção de dados em processamento, além da prevenção contra a divulgação não autorizada de parte ou do todo. O princípio da Integridade visa garantir o rigor e a confiança que não ocorrerão modificações não autorizadas em dados, informações ou documentos. Disponibilidade, por sua vez, é o princípio relacionado à capacidade dos recursos computacionais (sistemas, redes e dados) de se manterem disponíveis para acesso sempre que demandados [8]. Para o propósito desta pesquisa, segue-se a abordagem que considera os princípios fundamentais da segurança da informação como propriedades ou características essenciais aos documentos digitais, também aplicáveis ao documento de inteligência.

No contexto digital, a cadeia de custódia não rompe com a episteme advinda da Ciência da Informação, sobretudo da Arquivologia. Ao contrário, ratifica esse aprendizado e se apresenta como uma reação adaptativa da gestão documental e arquivística frente à inovação disruptiva advinda da TD, com o propósito não só de armazenamento, mas também de guarda e de preservação da autenticidade (identidade e integridade) do documento digital, assegurando que nenhum de seus bits seja alterado ou adulterado [6].

Outra vulnerabilidade que recai sobre os ativos digitais refere-se à dificuldade de preservação de acesso de logo prazo, sob os vários aspectos decorrentes da dependência tecnológica (deterioração dos mecanismos de armazenamento, descontinuidade de formatos e tecnologias digitais, entre outros) o cuidado com o material digital requer planejamento, estratégias, abordagens e práticas documentais e arquivísticas, que são coletivamente conhecidas como Curadoria Digital [9].

Neste sentido, a pesquisa produziu o artigo "Elementos-chave da Transformação Digital que influenciam na Curadoria Digital"[10], que traz um melhor entendimento acerca da relação da Transformação Digital (TD) com a Curadoria Digital (CD), onde se identifica a convergência de abordagens e tecnologias advindas da TD (*Blockchain, Cloud computing, Artificial intelligence, Data Digitization*, entre outras) para o propósito da CD.

Desta forma, as vulnerabilidades advindas da digitalização impõem que a ABIN adote, em seu processo de produção de conhecimento, medidas para a segurança, preservação e proteção dos documentos de inteligência nato-digitais[1] ou digitalizados, onde a cadeia de custódia e a curadoria digital se apresentam como soluções potenciais.

1.2 PROBLEMA OU HIPÓTESE DE PESQUISA

Principal processo finalístico da ABIN, a produção de documentos de inteligência faz uso de sistemas informatizados de gestão arquivística de documentos para o armazenamento, compartilhamento e elabo-

[1]Nato-digital: documento que já nasceu em formato digital, sem o suporte físico (papel)

ração colaborativa de documentos digitais. Neste ambiente, os artefatos de inteligência (documentos de variados formatos) são coletados e armazenados nativamente, onde são processados e resultam em documentos de inteligência, que são armazenados em repositório digital convencional. Este processo está sujeito às vulnerabilidades inerentes ao documento digital.

Em termos tecnológicos, o documento digital está sujeito aos riscos e vulnerabilidades inerentes às tecnologias digitais, tais como obsolescência, indisponibilidade e perda da interoperabilidade [11]. Sob o aspecto jurídico-administrativo, no que ser refere à autenticidade, o documento digital está sujeito à quebra da cadeia de custódia, e por conseguinte, ao risco de ter sua irrefutabilidade questionada, enquanto elemento probatório e evidência [6].

No que se refere aos aspectos legal e diplomático[2], o documento de inteligência deve ter presunção de autenticidade, em conformidade com as diretrizes definidas pela Câmara Técnica de Documentos Eletrônicos [7].

Sob o aspecto de sigilo documental, como ocorre na maioria das instituições que empregam o esquema convencional de armazenamento, em que a equipe de infraestrutura de TIC atua com permissão total sobre os repositórios dos ativos digitais, esta atuação termina por trazer para o time técnico a responsabilidade tácita de custodiante[3] de documentos digitais classificados. Esta condição implicará toda a equipe de implantação da solução de armazenamento em processo investigatório, em caso de quebra de sigilo ou de vazamento[4].

Em termos de preservação, o legado de documentos digitais do órgão não conta com mecanismos ou práticas que assegurem a longevidade dos documentos de inteligência para reuso ou pelo seu valor histórico, quando perderem sua classificação de sigilo [12].

Diante dessa problemática, a hipótese da pesquisa é que a adoção combinada de Cadeia de Custódia Digital Arquivística (CCDA) com práticas de Curadoria Digital (CD) promova a mitigação de vulnerabilidades inerentes ao documento digital e que contribua para a manutenção de propriedades que são imprescindíveis ao documento de inteligência digital.

1.3 OBJETIVOS

O objetivo geral do trabalho é propor diretrizes à Agência Brasileira de Inteligência (ABIN) para adoção de medidas que assegurem as propriedades imprescindíveis ao documento de inteligência frente às vulnerabilidades advindas do documento digital.

Os objetivos específicos são:

a) Identificar vulnerabilidades do documento digital que impactam as propriedades imprescindíveis ao documento de inteligência.

[2]Diplomático: referente à diplomacia, disciplina que tem como objeto o estudo da estrutura formal e da confiabilidade e autenticidade dos documentos

[3]Custodiante: Aquele que assume solidariamente a responsabilidade pela guarda e a proteção de algo ou alguém

[4]Vazamento: Divulgação indevida de informação ou documento com classificado de sigilo

b) Investigar e apontar medidas, dentre elas a Cadeia de Custódia Digital Arquivística (CCDA) e da Curadoria Digital (CD), para assegurar as propriedades imprescindíveis ao documento de inteligência frente às vulnerabilidades advindas do documento digital.

c) Propor diretrizes que viabilizem a adoção das medidas apontadas no contexto da arquitetura de governança do órgão.

1.4 JUSTIFICATIVA DA PESQUISA

O principal processo finalístico na cadeia de valor[5] da ABIN pressupõe a coleta e análise de artefatos informacionais (dados, informações e documentos) para a geração de conhecimento, que é materializado em documentos de inteligência digitais dos mais diversos formatos. O ambiente de produção de inteligência conta com soluções automatizadas para compartilhamento de artefatos e elaboração colaborativa de documentos, o que sujeita os documentos de inteligência às vulnerabilidades inerentes ao documento digital. A volatilidade digital também compromete o legado documental de inteligência do órgão (documentos nato-digitais e digitalizados), deixando sua preservação comprometida, seja por problemas de ordem tecnológica, seja por de quebra de sigilo ou por indisponibilidade.

Esta pesquisa pretende apontar caminhos para a mitigação das vulnerabilidades inerentes ao documento digital que afetam o processo de produção de documentos de inteligência de Estado, investigando a eficácia da CCDA na guarda e preservação da autenticidade do documento digital e da CD na longevidade dos documentos digitais.

A pesquisa se mostra viável e aplicável à necessidade do órgão, uma vez que está alinhada a seu objetivo estratégico de "Salvaguardar as comunicações governamentais, bens e conhecimentos sensíveis" [13], além de não demandar recursos adicionais de implementação, mas tão somente os recursos humanos e tecnológicos, já disponíveis.

1.5 RESULTADOS ESPERADOS

Espera-se que este trabalho aponte caminhos para a mitigação das vulnerabilidades inerentes ao documento digital que afetam as propriedades essenciais aos documentos de inteligência; valide a adoção da CCDA e CD, entre outras, como forma de assegurar essas propriedades; e reúna diretrizes para a adoção de medidas que aprimorem a segurança do processo de produção de conhecimento de inteligência.

Espera-se ainda propor medidas que aprimorem a segurança entorno dos repositórios digitais e retirem da equipe de infraestrutura de TIC a responsabilidade tácita de custodiante do material digital, que lhe recai por seu papel nas implantações de repositórios digitais tradicionais, reduzindo (ou eliminando) a necessidade de acesso privilegiado por parte desta equipe sobre os repositórios de documentos de inteligência.

[5] Cadeia de Valor: Método que possibilita que a instituição entenda como funciona a organização e a prática dos seus processos produtivos e estratégicos

1.6 ESTRUTURA DO TRABALHO

O trabalho está organizado em cinco partes, estruturadas conforme preconiza a Metodologia de Pesquisa para Ciência da Computação [14]. No Capítulo 1 está a Introdução, que contextualiza a pesquisa dentro da realidade do órgão, detalha o problema a ser investigado e a hipótese de pesquisa a ser testada, apresenta o objetivo geral e os objetivos específicos da pesquisa, explana sobre a justificativa e a viabilidade da execução da pesquisa, sua aplicabilidade e os resultados esperados em razão do trabalho.

O Capítulo 2 traz o referencial teórico que serve de base conceitual para o entendimento do trabalho, incluindo conceitos relacionados à transformação digital, como o de digitalização e de digitização, os conceitos e o relacionamento entre dado e inteligência, a evolução da noção de documento até chegar ao conceito de SIGAD, da origem da cadeia de custódia ao surgimento da Cadeia de Custódia Digital Arquivística (CCDA) e sua relação com o repositório digital confiável (RDC), o conceito de curadoria digital (CD) e o planejamento de ações de curadoria dentro do ciclo de vida documental e ainda a relação entre CCDA e CD.

O Capítulo 3 apresenta a classificação da pesquisa em termos metodológicos, a metodologia, as técnicas e procedimentos aplicados na pesquisa e como o encadeamento deles possibilitou o alcance dos resultados.

O Capítulo 4, que aborda resultados e discussão, conduz uma Revisão Sistemática de Literatura (RSL) que empregou o método TEMAC para identificação das vulnerabilidades do documento digital e das propriedades por elas afetadas, apresenta uma tabulação destes resultados e conduz uma discussão acerca de medidas eficazes de mitigação, onde se faz uma análise descritiva do impacto da CCDA e da implantação do RDC sobre as propriedades do documento digital, trazendo também um estudo comparativo entre implementações de RDC em software livre; do papel da CD para a longevidade documental digital; da importância do padrão CMIS para a interoperabilidade do documento digital; dos cenários de integração SIGAD e RDC-Arq; da importância da infraestrutura computacional para alta disponibilidade do documento digital. O capitulo culmina com a apresentação de diretrizes para a adoção das medidas apontadas em prol do documento de inteligência digital, considerando a arquitetura de governança do órgão.

A conclusão está no Capítulo 5, que apresenta a consolidação dos objetivos do trabalho, confirma a hipótese e apresenta os achados da pesquisa e traz também perspectivas de trabalhos futuros.

Ao final estão as referências, que servem de sustentação e de base bibliográfica à temática para a pesquisa.

2 REFERENCIAL TEÓRICO

2.1 TRANSFORMAÇÃO DIGITAL NA ABIN

O movimento rumo ao digital, batizado de Transformação Digital (TD), tem encontrado ambientação nos mais diversos segmentos da sociedade. Fortemente impulsionado pela Tecnologia da Informação e Comunicação (TIC), a mudança aponta para a convergência dos ativos informacionais das organizações cada vez mais para o formato digital [15].

O termo digitalização, recorrente no processo de TD das últimas décadas, por vezes se confundia com o próprio movimento de TD. Assim, digitalização consiste nas mudanças associadas à aplicação de tecnologias digitais em todos os aspectos da sociedade [16]. Acompanhando as evoluções tecnológicas, o conceito de digitalização ganhou mais elasticidade, referindo-se à adoção ou ao incremento de tecnologias digitais incorporadas aos processos de trabalho ou de negócio [5].

Considerando a linha do tempo da TD das últimas décadas, observa-se a digitalização em quatro fases, incluindo período, principal fenômeno, foco e atividades bem característicos. A primeira fase (pré-1990) é marcada pela experimentação das tecnologias para a geração de dados em formato digitais. Na segunda fase (de 1990 a 2000), o surgimento das plataformas digitais de comunicação reduz a intermediação nos processos de negócio e, por consequência, os custos operacionais. Na fase seguinte (de 2000 a 2010), o foco esteve na melhoria da eficiência pela exploração de soluções digitais voltadas para otimizar os fluxos de negócio, com destaque para os sistemas de gestão empresarial. Na fase contemporânea (após 2010), o fenômeno digital se consolida como perene e disruptivo, onde a conectividade e a integração sistêmica não são mais demandas pontuais, mas sim necessidades básicas e elementares dos processos de trabalho e de negócio. A Tabela 2.1 sintetiza essas transições.

Na fase contemporânea da TD, inserida no contexto da digitalização, surge a digitização como uma especialização da digitalização, buscando a conversão dos dados para o formato digital e a sua incorporação aos fluxos e processos de trabalho e de negócio, Figura 2.1. Digitização eventualmente é mencionada como capacidade de digitização[6], referindo-se à capacidade ou proficiência necessária para que as organizações gerem e operem dados digitais, criando valor efetivo aos processos de negócio pela transformação de entradas em saídas [5].

Na ABIN, o atual cenário de TD é compatível com quarta fase, onde o órgão busca incrementar sua capacidade de digitização institucional, trabalhando as dimensões da digitização para maior integração das soluções digitais, conforme representado na Figura 2.2.

A capacidade de digitização institucional é determinante para a qualidade dos dados e metadados referentes aos objetos digitais, são estas informações que irão assegurar a preservação da relação orgânica (onde, quando, porque, quem e como) dos documentos com a instituição [6].

[6] digitization capability

Tabela 2.1: Fases da transformação digital

	Fase 1	Fase 2	Fase 3	Fase 4
Período	Pré-1990	1990–2000	2000–2010	Pós-2010
Fenômeno	Dados digitais	Plataformas digitais de comunicação	Melhoria da eficiência digital	Fenômeno digital consolidado
Foco Principal	Experimentação	Desintermediação	Exploração	Integração
Atividades Dominantes	Tecnologias digitais são exploradas como uma nova forma de trabalhar e de desenvolver negócios	Tecnologias digitais são utilizadas para conectar com os clientes, seja diretamente de uma forma digital de baixo custo ou por meio de plataformas	Tecnologias digitais são usadas para otimizar os fluxos de negócios, principalmente para aumentar a eficiência dos processos de negócios conhecidos	Tecnologias digitais são amplamente difundidas e se tornam um fato aceito nos negócios, em vez de serem especiais ou extraordinárias

Fonte: Traduzido de *Digitization capability and the digitalization of business models in business-to-business firms: Past, present, and future* [5]

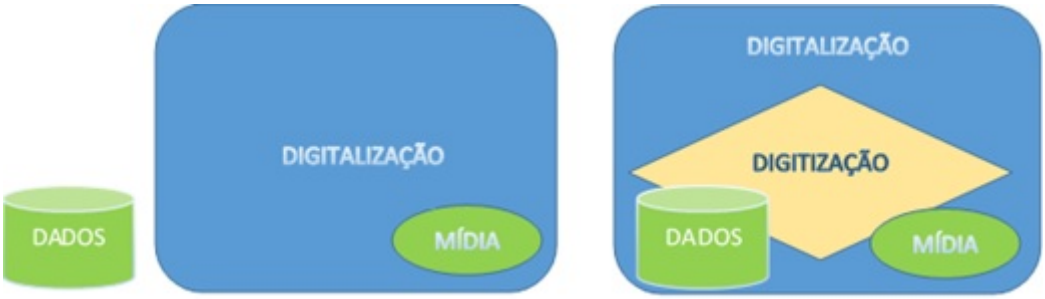

Figura 2.1: Digitalização x Digitalização com Digitização

Fonte: Elementos-chave da transformação digital que influenciam na curadoria digital [10]

2.2 DO DADO À INTELIGÊNCIA

Na visão clássica da Ciência da Informação, onde os conceitos dado, informação, conhecimento e inteligência têm sido amplamente discutidos, defende-se a inter-relação entre eles por meio de etapas de transformação e geração de valor [17]. Nesta abordagem, estes elementos figuram numa escala crescente de valor agregado que vai de dado à inteligência, onde dado é o objeto observado, capturado e armazenado, a matéria prima, o elemento bruto que não faz sentido isoladamente ou fora de um contexto que lhe traga significado. No nível acima da escala, informação é o dado que ganha significado útil, seja por interpretação, por processamento ou por contextualização. No terceiro nível, encontra-se o conhecimento, que é entendido como o conjunto de informações analisadas, organizadas e conectadas, que propiciam a compreensão de fato ou o entendimento de fenômeno, sendo a busca pelo conhecimento, a razão de ser da ciência [18]. No quarto nível, resultante da 3ª etapa, a inteligência figura como conhecimento de alto valor agregado [19], como representado na Figura 2.3.

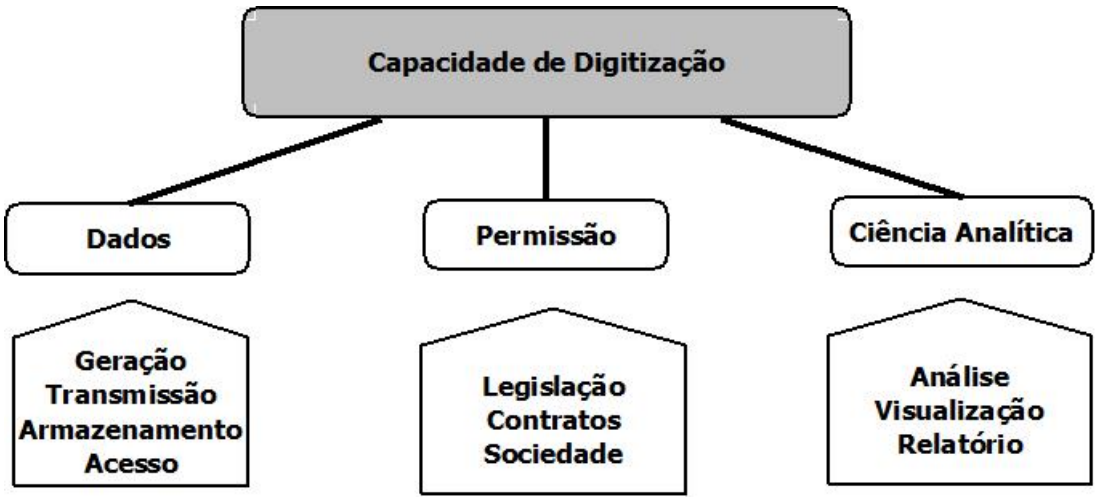

Figura 2.2: Dimensões da Digitização

Fonte: Digitization capability and the digitalization of business models in business-to-business firms [5]

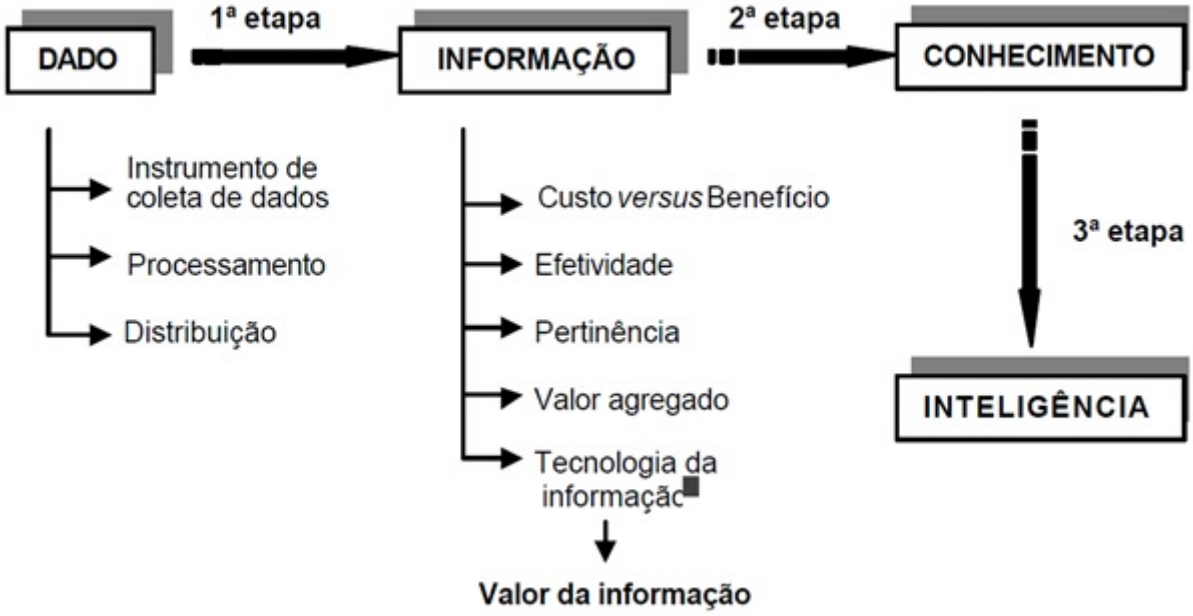

Figura 2.3: Etapas na Geração de Conhecimento e Inteligência

Fonte: Sociedade da informação e inteligência em unidades de informação [17]

Na perspectiva da Ciência da Informação destacam-se diferentes abordagens para o termo informação e inteligência. Para o termo informação, destacam-se as seis abordagens conceituais de Wersig & Neveling [20]:

i. Abordagem Estrutural: Informação é vista como coisa, contida numa evidência material autônoma podendo ser apreendida ou não (dados, textos, documentos e objetos);

ii. Abordagem do Conhecimento: Informação é um conhecimento em potencial, pode ser realizada, ou não, e se dá objetivamente, devendo ser adquirida por um sujeito e servindo a um fim específico.

iii. Abordagem da Mensagem: Informação é sinônimo de mensagem, associada ao próprio conteúdo comunicado.

iv. Abordagem do Significado: Uma visão orientada para a mensagem em que somente o significado da mensagem é tido como informação.

v. Abordagem do Efeito: Informação como um efeito desencadeado no receptor, como a consequência da transmissão ou abstração do conhecimento, a redução de incerteza, por exemplo.

vi. Abordagem do Processo: Informação como uma etapa do processo de conhecimento, onde Dado bruto é processado e transformado em Informação, que por sua vez é organizada e interprestada para resultar no Conhecimento.

Para inteligência, são consideradas as seguintes acepções [19]:

a. Inteligência como Aptidão - capacidade ou habilidade da qual se possa fazer uso para resolver problemas ou dificuldades de qualquer ordem.

b. Inteligência como Atividade de Informações – ações sistemáticas e integradas que objetivam a obtenção, análise e disseminação de conhecimentos sobre fatos e situações de imediata ou potencial influência sobre o processo decisório organizacional ou governamental.

c. Inteligência como Informação de Alto Valor – emprego ou uso efetivo e oportuno de informação ou conhecimento de alto valor agregado.

Assim, a partir das contribuições da Ciência da Informação, combinando as abordagens e acepções já apresentadas, pode-se oferecer um conceito mais completo de inteligência:

> atividades ou ações sistemáticas e integradas (b), que requerem capacidades e habilidades profissionais específicas (a), empenhadas na obtenção de dados e análise de informações (i) e disseminação (iii) de conhecimentos (ii) com valor agregado (c) sobre fatos e situações de imediata ou potencial influência (v) sobre o processo decisório (vi) organizacional ou governamental.

Além da Ciência da Informação, os diversos serviços de informações estatais se debruçaram por décadas na tentativa de formular uma definição para o termo inteligência. Sob esta ótica, a definição de inteligência é:

> a coleta e processamento de informações sobre países estrangeiros e seus agentes, necessárias a um governo para sua política externa e para a segurança nacional, a condução de atividades não atribuíveis no exterior para facilitar a implementação da política externa e a proteção de processo e produto, bem como pessoas e organizações preocupadas com eles, contra divulgação não autorizada [21].

No Brasil, a Atividade de Inteligência de Estado foi oficialmente estabelecida pela Lei nº 9.883/1999, que criou o Sisbin e instituiu a ABIN como órgão central, tendo a seu cargo

> planejar, executar, coordenar, supervisionar e controlar as atividades de inteligência do País. Atividade que tem o objetivo de produzir e difundir conhecimentos às autoridades competentes, relativos a fatos e situações que ocorram dentro e fora do território nacional, de imediata ou potencial influência sobre o processo decisório, a ação governamental e a salvaguarda da sociedade e do Estado, conforme Decreto Nº 8.793 - PR, 2016 [4].

Assim, no cumprimento de sua missão institucional, que é

> Antecipar fatos e situações que possam impactar a segurança da sociedade e do Estado brasileiros, de modo a assessorar o mais alto nível decisório do País, bem como salvaguardar conhecimentos sensíveis e aprimorar a Atividade de Inteligência de Estado [13],

a Inteligência de Estado entrega o conhecimento por meio de documentos digitais dos mais variados formados, conhecidos doutrinariamente como documentos de inteligência.

2.3 DO DOCUMENTO AO *CONTENT MANAGER*

Considerando que o conhecimento de inteligência materializa-se em documento de inteligência digital, é importante revistar o conceito de documento e conhecer o papel das ferramentas de automatização de gestão documental e arquivística. Na visão dos chamados patronos da documentação[7], documento seria qualquer coisa em que o conhecimento pudesse ser registrado e que se reconhece alguma propriedade informativa, onde o papel predominou como suporte físico informacional do documento. A partir da digitalização, o conceito de documento vem sofrendo mutações com o propósito de promover a transição do suporte informacional, fixo e rígido, para um suporte eletrônico, baseado no sistema binário, flexível e fluído.

O documento digital, inicialmente chamado de documento eletrônico, traz consigo a informação codificada em dígitos binários (bits), acessível e interpretável por meio de sistema computacional. De positivo, o documento digital oferece a facilidade de acesso, de armazenamento e de transferência; maior disponibilidade, legibilidade partilhada e flexibilidade de formato. Pelo lado negativo, a falta de estabilidade leva à volatilidade, seja pela fragilidade do armazenamento binário, baseado na combinação eletro-física de elementos com vida útil (disco rígido, disco ótico e chips, entre outros), seja pela obsolescência dos formatos digitais, que tendem a ser substituídos por formatos mais eficazes, muitas vezes sem a preocupação com a retro compatibilidade [22].

Apesar da digitalização, a noção de documento manteve um ponto comum entre o analógico e o digital, preservando sua característica de "unidade representativa da mensagem com potencial de utilização", Figura 2.4. Após a ruptura com o suporte informacional, saindo do suporte fixo e rígido para o flexível e fluído, a noção de documento naturalmente incorpora elementos inerentes ao material digital, de positivo a facilidade e a flexibilidade, de negativo a volatilidade, que implicará em muitos desafios a superar.

O documento de inteligência digital resultante da Atividade de Inteligência – coleta, análise e processamento de artefafos informacionais – é construído por meio de ferramenta de automação da gestão

[7]Patronos da documentação: Paul Otlet, Etienne-Gabriel Peignot, Pierre Larousse, Henri La Fontaine, entre outros

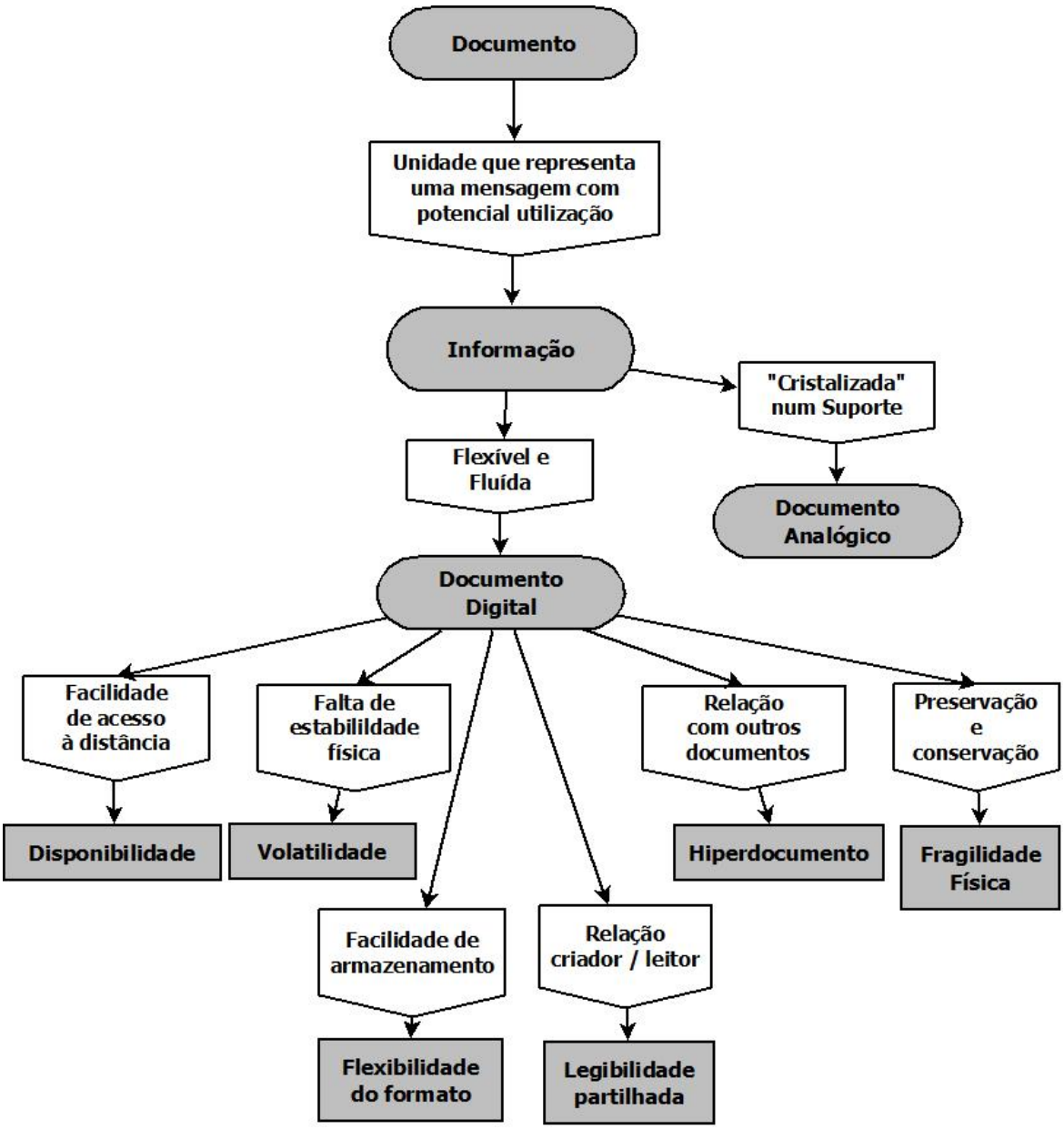

Figura 2.4: Mapa Conceitual de Documento Digital

Fonte: A noção de documento digital: uma abordagem terminológica [22]

documental, genericamente chamada de Sistema Informatizado de Gestão Arquivística de Documento (SIGAD). Advindo dos chamados GED (Gerenciador Eletrônico de Documentos) e *Eletronic Document Management System* (EDMS), a modalidade de sistemas *Content Management System* (CMS) possibilita a elaboração, o armazenamento e o compartilhamento de documentos digitais, em alinhamento com o modelo de requisitos para sistemas informatizados de gestão arquivística de documentos (e-Arq Brasil) [23].

As soluções SIGAD vêm dando lugar para as plataformas computacionais mais elaboradas, orientadas a pessoas, processos e conteúdos, permitindo o trabalho colaborativo por meio de sistemas distribuídos. Nestas plataformas, o componente *Business Process Management* (BPM) se ocupa da automatização dos processos de negócio, identificando, desenvolvendo, executando, monitorando e controlando-os para o

atingimento dos resultados; enquanto o gerenciamento de conteúdo (registro, documentação, classificação, armazenamento e indexação) fica a cargo do componente *Enterprise Content Management* (ECM) [24].

No contexto dos sistemas de gerenciamento de conteúdo (CMS), o modelo *Content Management Interoperability Services* (CMIS) tornou-se referência para a implementação de repositório de conteúdo interoperável. O modelo referencial CMIS oferece especificações detalhadas de modelo de dados, de objetos de dados (documento, diretório, item, lista, entre outros) e de serviços de repositório (gravação, recuperação, navegação, versionamento, controle de acesso e relacionamento entre objetos, entre outros) que permitem a interoperabilidade de conteúdo entre os repositórios que estiverem em conformidade com este padrão [25].

2.4 CADEIA DE CUSTÓDIA, CCDA E RDC

O conceito de custódia advém do campo jurídico, mais especificamente da Ciência Forense, onde cadeia de custódia é entendida como a adoção de um conjunto de tecnologias, procedimentos e práticas que assegurem as propriedades do objeto em trâmite (onde, quando, porque, quem e como), garantindo sua irrefutabilidade enquanto evidência material ou digital no processo investigativo [26].

Oriundo da gestão documental e arquivística tradicional, o conceito de cadeia de custódia documental tem por base o paradgima da temporalidade do documento em termos de arquivamento. Desta perspectiva, a idade do documento é definida em três fases conforme sua frequência de uso: corrente, intermediário e permanente. No ambiente de gestão de documentos, momento em que e suas referências ainda são suscetíveis a alterações, o documento permanece armazenado no chamado arquivo corrente ou intermediário. No ambiente de gestão arquivística, fase em que atinge uma condição estática e definitiva em termos de conteúdo, o documento chega a idade permanente, sendo necessário guardá-lo em área reservada, livre de adulteração, com vistas a preservar sua autenticidade (identidade + integridade), garantir seu valor histórico informativo e sua condição de elemento probatório confiável. No contexto inteiramente analógico, o arquivamento se dava pelo recolhimento do documento físico (suporte em papel) para o repositório arquivístico confiável, geralmente chamado de "arquivo" do órgão, especificamente concebido para atender a todos os requisitos de preservação arquivística [6].

No contexto de documentos nato-digitais ou digitalizados, a cadeia de custódia permanece fundamentada no paradigma da temporalidade ou idade documental (corrente, intermediária e permanente), mas conjuga os conceitos de Cadeia de Custódia (CoC)[8] com o de Cadeia de Preservação (CoP)[9] para formar a Cadeia de Custódia Digital Arquivística (CCDA), que surge com o propósito de garantir a transferência protocolar de pacotes de documentos do ambiente de gestão para a área de preservação de documentos sem que haja quebra na cadeia de custódia, e de superar os desafios inerentes ao documento digital, tais como o armazenamento, a preservação de autenticidade, a recuperação, o controle de acesso, entre outros. A adoção de uma CCDA na gestão documental e arquivística da instituição está necessariamente associada à implantação de Repositório Digital Confiável (RDC). O RDC então é uma especialização do repositório digital que tem o propósito da preservação e proteção das características do documento arquivístico,

[8] Chain of Custody
[9] Chain of Preservation

em especial a autenticidade (identidade e integridade) e a manutenção da relação orgânica (onde, quando, porque, quem e como) dos documentos com a instituição [6].

Desta forma, a CCDA pressupõe a implementação de RDC, que não é apenas um arcabouço tecnológico (hardware e software) empregado para o armazenamento e gestão de material digital, mas uma solução computacional que, além de prover armazenamento e acesso aos objetos digitais, tem a missão de prover confiança, preservação e acesso de longo prazo aos recursos digitais [27].

Os conceitos de RDC e RDC-Arq estão associados e podem ser mais bem entendidos por meio do mapa conceitual relacionado ao RDC (Figura 2.5), onde se apresenta, de forma ramificada, a relação de eventos, normas, modelos e recursos a ele relacionados. Por este mapa conceitual, o modelo conceitual *Open Archival Information System* (OAIS) resultou na norma ISO 14.721:2003, que define os padrões e as especificações de referência para implementação de repositório automatizado voltado para a gestão, preservação e acesso de documentos arquivísticos digitais. O RDC surgiu a partir dos requisitos de implantação definidos na ISO 16.363:2012, que junto com a ISO 16.919:2014, serviu de base para o estabelecimento da *Trustworthy Repository Audit and Certification* (TRAC), uma trilha para a certificação de "confiável"para repositórios digitais que contempla auditorias e a aplicação de critérios e de *checklist* definidos (OCLC and CRL, 2007). Um RDC que passa por auditoria e obtém a certificação TRAC passa a contar com a acreditação da comunidade internacional [28].

O RDC-Arq é a definição de RDC em nível nacional, que, por meio da Resolução nº 43 do Conselho Nacional de Arquivos – CONARQ, estabelece as Diretrizes para a Implementação de Repositórios Arquivísticos Digitais Confiáveis, uma recomendação que deve ser adotada para toda a Administração Pública Federal [29].

No contexto da gestão de documentos digitais, a CCDA pressupõe a custódia entre dois elementos computacionais que interagem no processo de transferência de dados e documentos. O chamado SIGAD (GED, ECM, ECMS, entre outros) atua na fase de gestão de documental – onde ficam armazenados os documentos corrente e intermediário, e o elemento RDC no papel de repositório arquivístico. O componente RDC fica a cargo da custódia na fase de Gestão de Preservação, tendo o papel de armazenar e preservar os documentos em idade permanente (definitiva). A transferência de documentos que se dá do ambiente de gestão para o ambiente de preservação configura uma alteração (não uma quebra) da cadeia de custódia, Figura 2.6.

As diretrizes do RDC-Arq estabelecem a conformidade com onze (11) padrões e especificações, o primeiro deles é o modelo conceitual *Open Archival Information System* (OAIS), que oferece uma abstração dos serviços de arquivamento, por meio do enpacotamento das informações documentais, em interface com os três atores da gestão arquivística: produtor, consumidor e administrador. Produtor é um papel desempenhado por pessoas, clientes ou sistemas que provê a informação a ser preservada. Consumidor é um papel desempenhado por pessoas, clientes ou sistemas que interage com os serviços do OAIS para encontrar e obter informação preservada de interesse. Um indivíduo ou sistema pode atuar com ambos os papéis, como produtor e consumidor dos serviços OAIS. O papel de administrador (ou gerente) é exercido por aqueles que definem a política geral de documentação da organização a ser aplicada ao repositório, Figura 2.7.

Nesse modelo, as interações entre as entidades se dão por meio de pacotes de protocolos específicos para cada fim. Na interação com o produtor (producer) é empregado o Pacote de Informações de Sub-

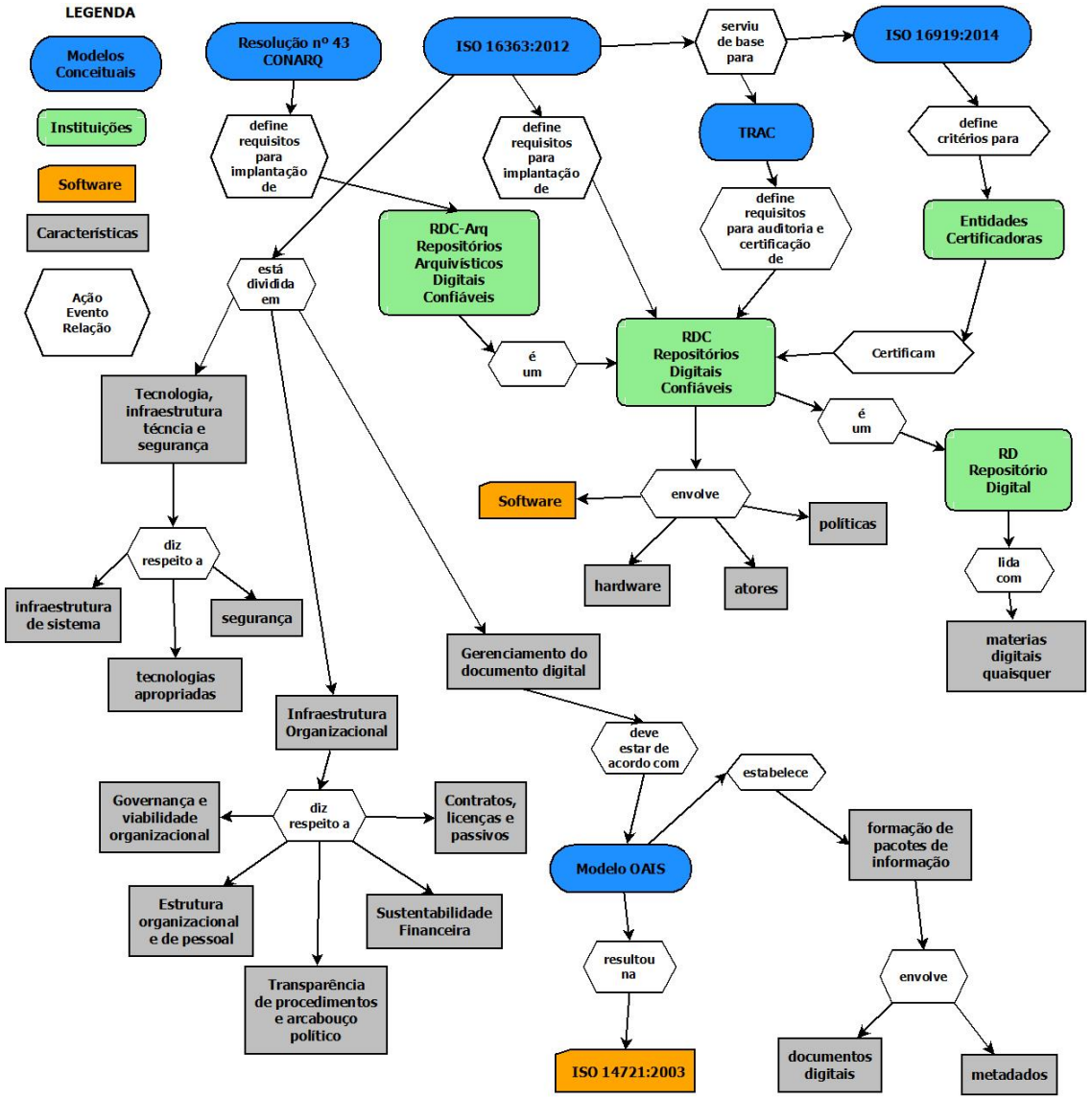

Figura 2.5: Mapa Conceitual Relacionado ao RDC-Arq

Fonte: Repositórios arquivísticos digitais confiáveis (rdc-arq)
como plataforma de preservação digital em um ambiente de gestão arquivística [30]

missão (SIP)[10]. Para as interações entre as entidades internas, o OAIS emprega o Pacote de Informações de Arquivamento (AIP)[11]. Na interface com o Consumidor (Customer), o modelo faz uso do Pacote de Informações de Disseminação (DIP)[12] [31].

As especificações dos pacotes SIP, AIP e DIP, conceitualmente criados pelo modelo OAIS, são abertamente divulgadas e atualizadas pela comunidade internacional [32]. A norma ISO 14721:2012, resultante desse modelo, deu as diretrizes para implementação de diversas soluções computacionais de gestão arquivística em repositório digital, muitas das quais desenvolvidas em código aberto com plena possibilidade de implantação.

[10] SIP: Submission Information Package
[11] AIP: Archival Information Package
[12] DIP: Dissemination Information Package

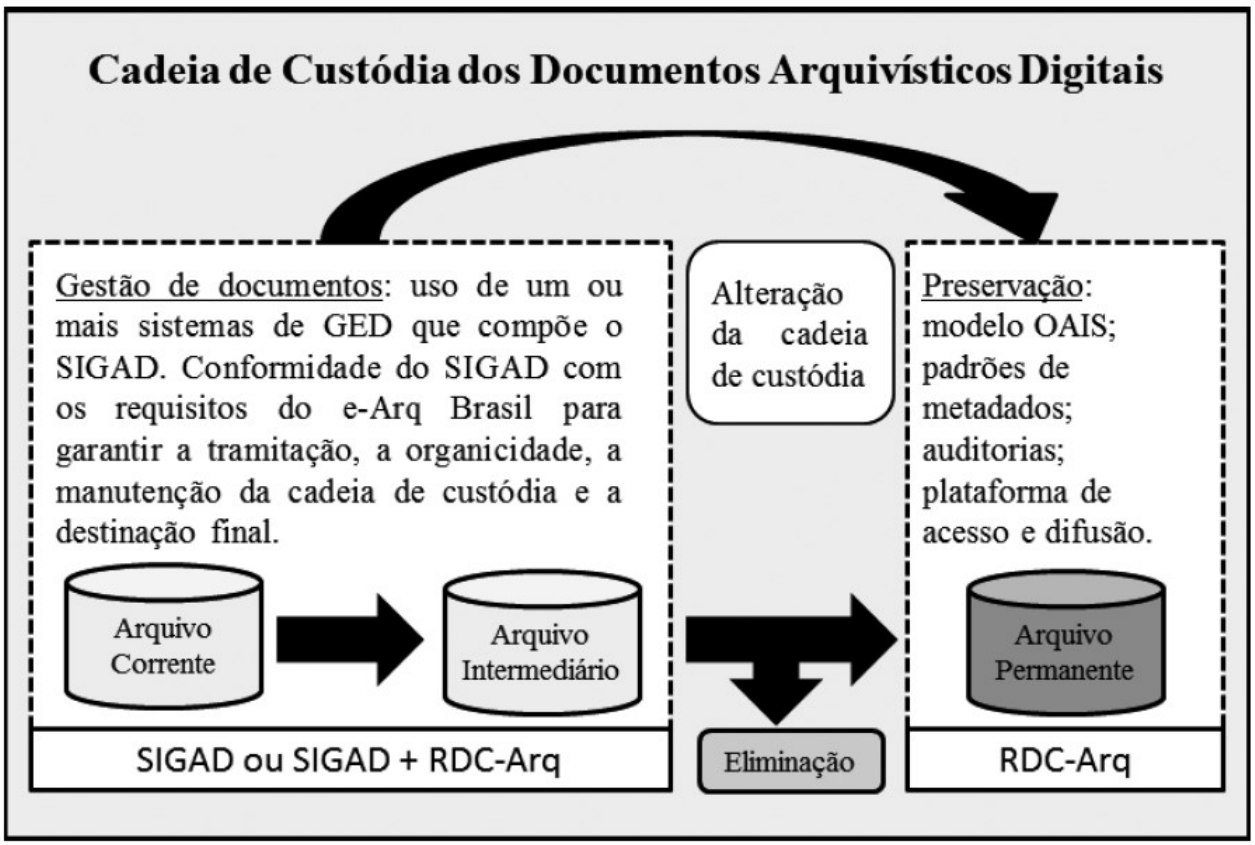

Figura 2.6: Cadeia de Custódia de Documentos Arquivísticos Digitais

Fonte: Cadeia de custódia para documentos arquivísticos digitais [6]

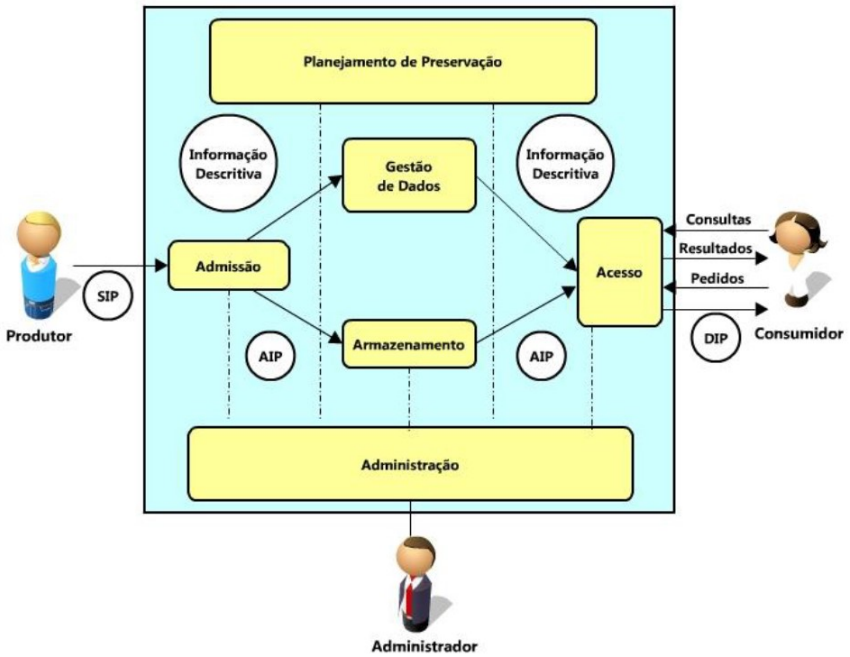

Figura 2.7: Entidades Funcionais do Modelo OAIS

Fonte: Reference Model for an Open Archival Information System (OAIS) [31]

Os demais padrões que fundamentam o RDC-Arq serão oportunamente abordados na sessão de resultados e discussão deste trabalho.

2.5 RELAÇÃO ENTRE CCDA E CD

Sob a perspectiva do material digital *per si*, com foco na qualidade da informação e dos dados coletados, como uma reação que visa dar sustentabilidade às inovações disruptivas advindas da TD, o Curadoria Digital evoluiu do conceito de curadoria de dados para um sentido mais amplo que engloba diversos conhecimentos, práticas e atividades voltados para o cuidado com o material digital [10].

Para alguns autores [33], Curadoria Digital é um termo hiperonímio, com conceito amplo que abrange atividades de diversas profissões, instituições, atores e setores, Figura 2.8.

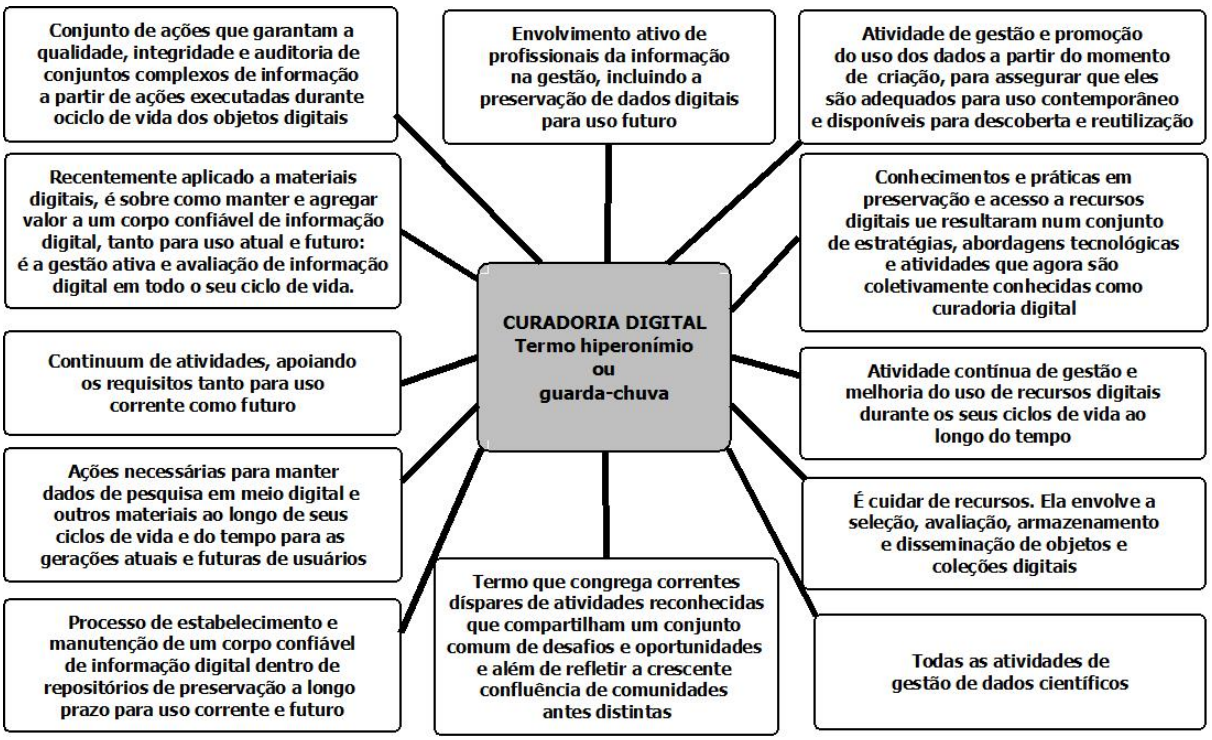

Figura 2.8: Alcance da Curadoria Digital

Fonte: Curadoria digital, custódia arquivística e preservação digital: relações possíveis [34]

Considerando o propósito de cada uma, CCDA e CD possuem uma relação de complementariedade, enquanto a CCDA permeia os processos de gestão de documentos, de gestão de arquivos e de gestão de conteúdos, a CD visa o planejamento, a qualidade, a manutenção, a longevidade e o reuso do material digital, ocupando-se da gestão ativa dos objetos digitais ao longo de todo o seu ciclo de vida [34], como representado na Figura 2.9.

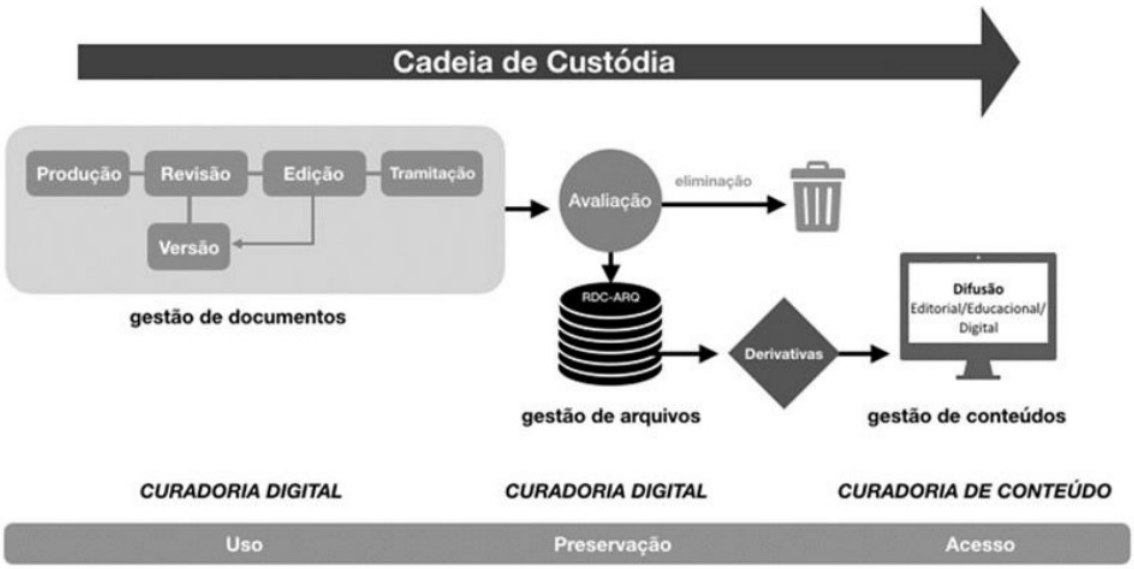

Figura 2.9: Relação: Gestão Documental e Arquivística, Curadoria Digital e Cadeia de Custódia

Fonte: Curadoria digital, custódia arquivística e preservação digital: relações possíveis [34]

A partir do propósito da CD, surgiram modelos capitaneados por diversos organismos independentes, como o *DCC Curation Lifecycle Model* [35], cujo núcleo do modelo compreende objetos digitais, em torno dos quais gravitam cuidados voltados para assegurar a qualidade, a estruturação, a representação, a recuperação e a preservação dos dados e objetos digitais, Figura 2.10.

Este modelo oferece a uma visão de alto nível para a gestão de materiais digitais, sua divisão em camadas concêntricas também auxilia no planejamento de atividades e orientam o uso de melhores práticas de curadoria em níveis mais granulares, possibilitando uma visão integral do fluxo e do gerenciamento de dados (núcleo do modelo) com os processos relacionados, em total aderência aos padrões relevantes e aos modelos de referência do RDC-Arq.

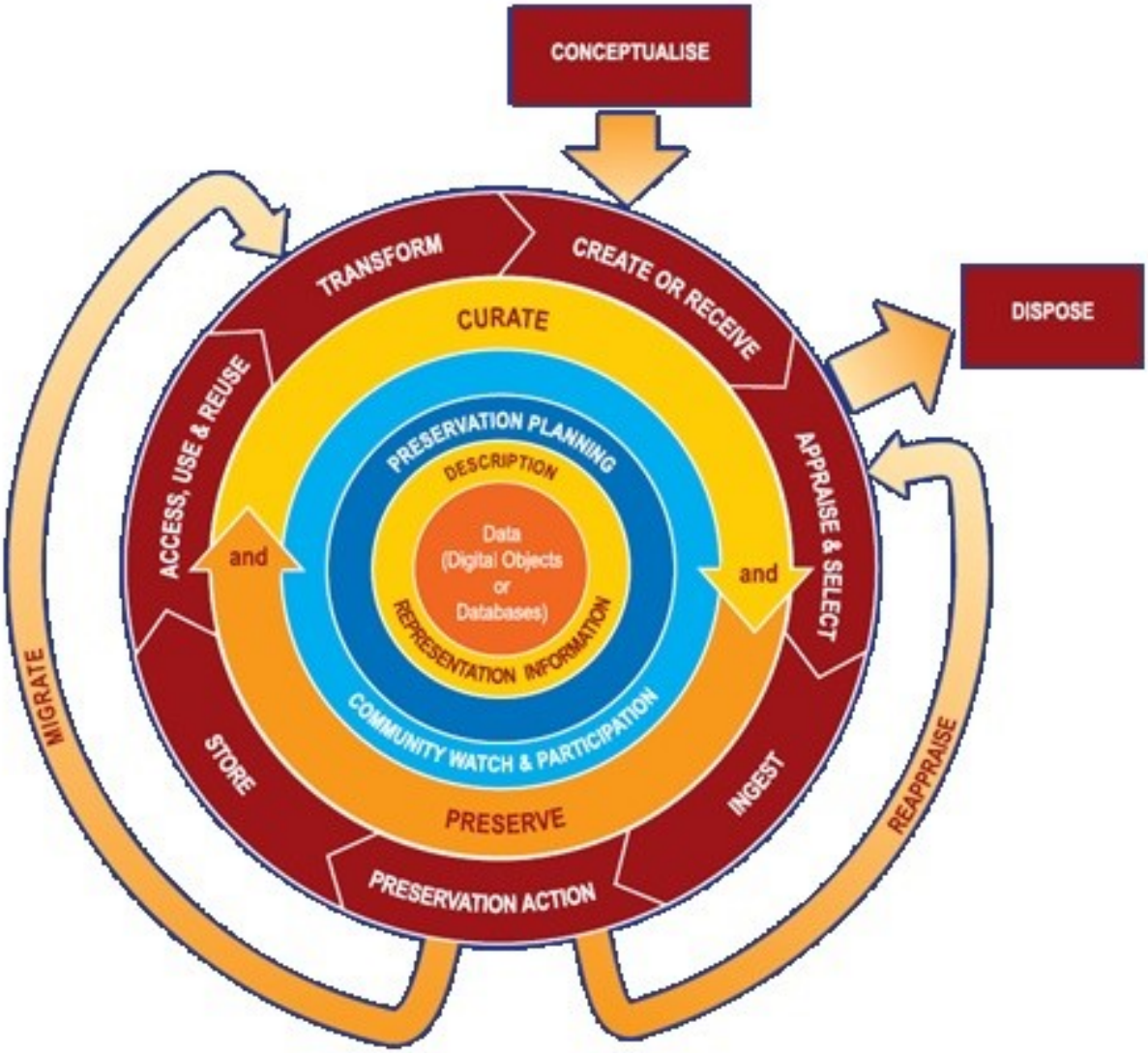

Figura 2.10: Modelo de Ciclo de Vida de Curadoria do DCC

Fonte: The DCC Curation Lifecycle Model [35]

3 METODOLOGIA

3.1 CLASSIFICAÇÃO DA PESQUISA

A pesquisa, quanto a sua classificação metodológica, é de natureza aplicada; é descritiva e exploratória para os objetivos; secundária quanto à origem dos dados coletados; é qualitativa quanto à abordagem; utiliza a pesquisa documental como procedimento técnico; e é transversal em termos de temporalidade, Figura 3.1.

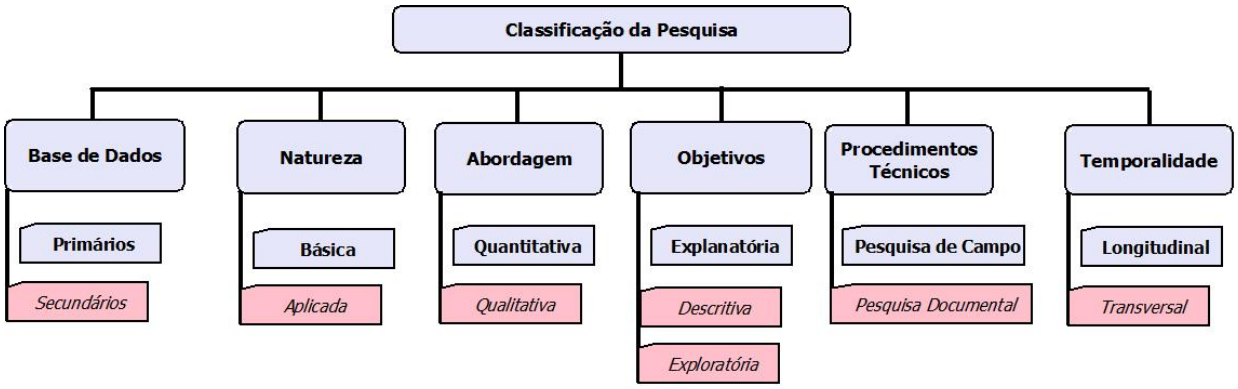

Figura 3.1: Classificação da Pesquisa

Fonte: Elaborado a partir de Metodologia da Pesquisa Científica: teoria e prática [36]

A natureza aplicada da pesquisa se mostra no objetivo geral do trabalho, que é voltado para a resolução de problema específico da instituição, ao fim do qual se propõe a adoção de medidas que assegurem as propriedades imprescindíveis ao documento de inteligência frente às vulnerabilidades advindas do documento digital.

Quanto aos objetivos, a pesquisa é classificada como exploratória e descritiva, na medida que emprega levantamentos e estudos bibliográficos existentes, e que investiga correlação entre elementos, conceitos e variáveis identificados nesta exploração.

A pesquisa é secundária quanto à coleta de dados, pois navega pelo conhecimento existentes acerca do tema de interesse em busca de descobrir solução ideal, ou pelo menos mais adequada ao contexto do problema. Para tanto, utiliza a pesquisa documental como procedimentos técnicos, que aliada à análise de conteúdo permite maior aprofundamento sobre os trabalhos científicos relacionados ao tema [14].

Em sua abordagem metodológica, a pesquisa se classificada como qualitativa, pois não requer o uso de métodos quantitativos e técnicas estatísticas, mas busca explicar o porquê dos fenômenos observados, bem como o que convém ser feito acerca do problema em questão, [36].

No que refere à classificação quanto à temporalidade, a pesquisa é transversal, uma vez que realiza coleta de dados única, considerando os trabalhos do período pós-2010, que é considerado como a 4ª das fases da transformação digital (Tabela 2.1).

Considerando o fluxo da pesquisa (Figura 3.2), o trabalho inicia com a revisão de literatura, que seleciona dados entre os trabalhos relevantes relacionados ao tema, resgata o referencial teórico necessário para dar o embasamento conceitual à pesquisa.

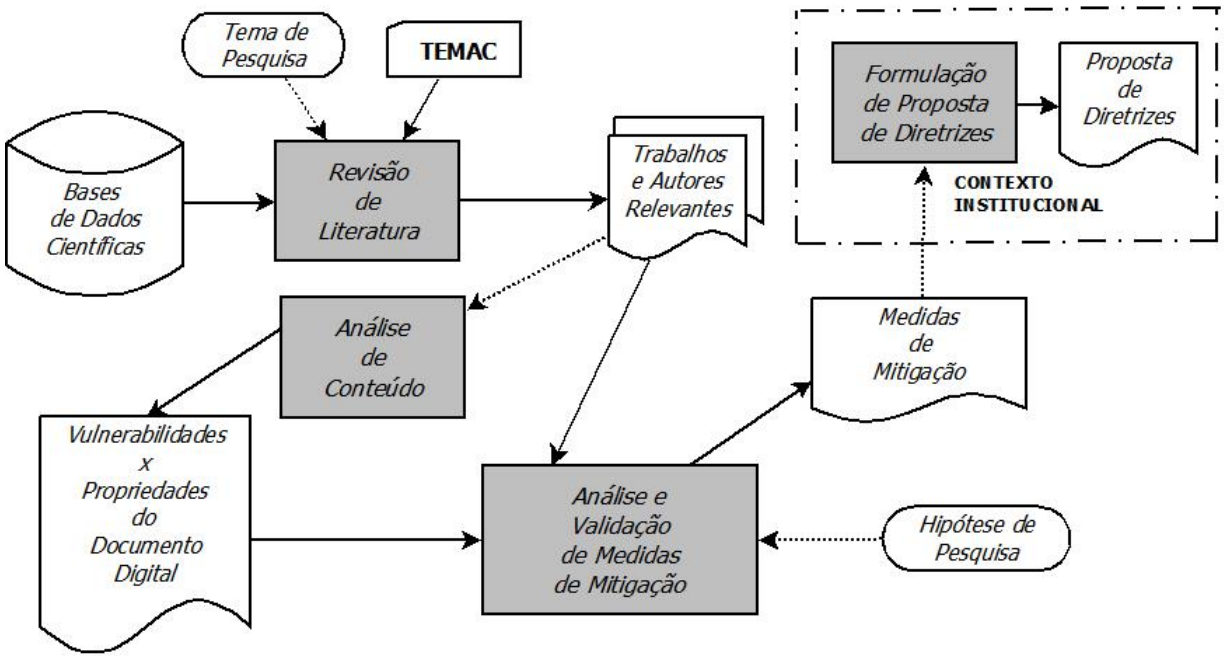

Figura 3.2: Fluxo da Pesquisa

Fonte: Elaborado pelo autor

3.2 SELEÇÃO DE DADOS DA PESQUISA

Considerando a problemática da pesquisa, que sugere a existência de vulnerabilidades inerentes ao documento digital que afetam o documento de inteligência, o trabalho emprega o método da Teoria do Enfoque Meta Analítico Consolidado (TEMAC) [37] para a seleção dos dados, por meio de Revisão Sistemática de Literatura (RSL). O roteiro TEMAC, pressupõe as 3 etapas: na primeira, chamada de "Preparação da pesquisa", serão definidos as palavras-chave, o espaço temporal da pesquisa, as bases de dados e as áreas de conhecimento da pesquisa; na segunda etapa, chamada de "Apresentação e inter-relação dos dados", os trabalhos são coletados, apresentados, classificados e selecionados sob diversos aspectos bibliométricos; na terceira e última etapa, denominada de "Detalhamento, modelo integrador e validação por evidências", são realizadas análises do rastreamento bibliográfico com o objeto de identificar os núcleos de cocitação e de acoplamento bibliográfico.

Desta forma, por meio de critérios objetivos combinados de RSL e TEMAC, os trabalhos e os autores mais relevantes são selecionados para exame e identificação das propriedades documentais afetadas pelas vulnerabilidades inerentes ao documento digital. Por propriedade documental, assume-se o conjunto de características do documento que lhe conferem qualidades e atributos desejáveis ou esperados.

3.3 ANÁLISE DOS DADOS: VULNERABILIDADES X PROPRIEDADES

A análise de dados da pesquisa recorreu às técnicas de análise documental, análise de conteúdo e descrição analítica propostas por Bardin [38]:

> *Análise Documental*: "...um conjunto de operações visando representar o conteúdo de um documento sobr uma forma diferente, a fim de facilitar sua consulta e rerenciação, ...tem por objetivo dar forma conveniente e representar de outro modo a informação."
>
> *Análise de Conteúdo*: "O cerne da análise de conteúdo está na interpretação da mensagem... em última instância, a interpretação do significado da mensagem de um emissor ao receptor pode ser decifrado pelas técnicas de análise de conteúdo, que permitam a inferência de conhecimentos relativos às condições de produção/recepção (variáveis inferidas) desta mensagem".
>
> *Descrição Analítica*: "A técnica de descrição analítica de conteúdo consiste da escrita descritiva da interpretação da mensagem extraída por medio da análise de conteúdo".

Por meio desta análise, os trabalhos (*papers*) selecionados foram examinados com vistas à identificação da relação entre vulnerabilidades e propriedades do documento digital, bem como para investigação e apontamento das medidas de mitigação aplicáveis às vulnerabilidades.

A identificação de relação entre vulnerabilidades e propriedades do documento digital, o exame textual dos trabalhos busca mapear qualquer ordem de vulnerabilidades inerentes ao documento digital, ao encontrar uma vulnerabilidade, busca-se em seu entorno, no contexto lógico e semântico do texto, inferir uma relação de *causa - efeito* entre uma e outra, além de quantificar o total de ocorrências desta no conjunto global dos trabalhos. O produto resultante desta análise irá conter uma tabulação da lista de vulnerabilidades, as propriedades por elas impactadas e a quantidade de incidências desta relação no conjunto dos documentos analisados.

3.4 ANÁLISE E VALIDAÇÃO DE MEDIDAS DE MITIGAÇÃO

As medidas de mitigação são ações ou iniciativas com o propósito de minimizar ou até de evitar danos ou efeitos negativos decorrentes de vulnerabilidades ou vulnerabilidades. Desta forma, a partir da tabulação de vulnerabilidades e de propriedades relacionadas ao documento digital, a pesquisa segue uma investigação exploratória sobre os trabalhos acadêmicos relacionados ao tema, para apontar, por meio da análise de conteúdo e da descrição analítica, quais medidas, dentre elas a hipótese de pesquisa, são eficazes de mitigar as vulnerabilidades identificadas. Como produto desta etapa, é apresentado um elenco de medidas com uma descrição analítica da eficácia na mitigação das vulnerabilidades, de forma a assegurar as propriedades do documento digital que afetam o documento de inteligência.

3.5 FORMULAÇÃO DE PROPOSTA DE DIRETRIZES

O elenco de medidas de mitigação apontadas na etapa anterior servirá de insumos para formulação de proposta de diretrizes aplicável ao contexto do órgão e do Sisbin. Para aumentar as chances de sucesso na implantação dessas medidas, formula-se proposta de diretrizes que nortearão o planejamento institucional desde o nível estratégico ao técnico-operacional, numa arquitetura de governança que respalde e assegure as condições de implantação das medidas apontadas. Como produto desta etapa, apresenta-se uma proposta com as diretrizes para a assegurar as condições de implantação do elenco de medidas de mitigação identificadas na etapa anterior.

4 RESULTADOS E DISCUSSÃO

4.1 REVISÃO SISTEMÁTICA DE LITERATURA

Uma vez que Revisão Sistemática de Literatura emprega o método TEMAC [37], a etapa de preparação da pesquisa requer a definição da string de consulta (*query string*), por meio da escolha das palavras-chaves que servirão de filtro de seleção dos trabalhos nas bases de dados de trabalhos acadêmicos, no caso as bases escolhidas foram SCOPUS e Web Of Science, que foram acessadas por intermédio da Comunidade Acadêmica Federada (CAFe) com as credencias da Universidade de Brasília (UnB).

Considerando o tema de pesquisa - vulnerabilidades do documento digital, foram definidas as palavras-chaves: *digital document*, *vulnerability*, *fragility* e *risk*. O filtro temporal da pesquisa definido abrange as publicações após o ano de 2010, em consonância com o enquadramento da instituição na 4ª fase da transformação digital. As strings de consulta formam elaboradas pelo arranjo lógico das palavras-chaves escolhidas visando um resultado abrangente de documentos afetos a esta temática, Tabela 4.1.

Tabela 4.1: Strings de Consulta às Bases de Dados

Base	String de Consulta	Resultado
Scopus	TITLE-ABS-KEY ("digital document") AND (PUBYEAR > 2009) AND (TITLE-ABS-KEY (vulnerability) OR TITLE-ABS-KEY (fragility) OR TITLE-ABS-KEY (risk))	46 documentos
Web Of Science	TS=("digital document") AND PY=(2010-2022) AND (TS=("vulnerability") OR TS=("fragility") OR TS=("risk")	19 documentos

Fonte: Elaborado pelo autor

A 2ª fase do roteiro TEMAC (Inter-relção de dados), onde se faz remoção das duplicidades, resultou em 46 documentos, dos quais a análise preliminar proposta pelo TEMAC mostra graficamente a evolução do tema ano-a-ano (Figura 4.1), os 10 autores que mais publicaram sobre o tema (Figura 4.2) e as 10 instituições que mais fomentaram esta linha de pesquisa (Figura 4.3).

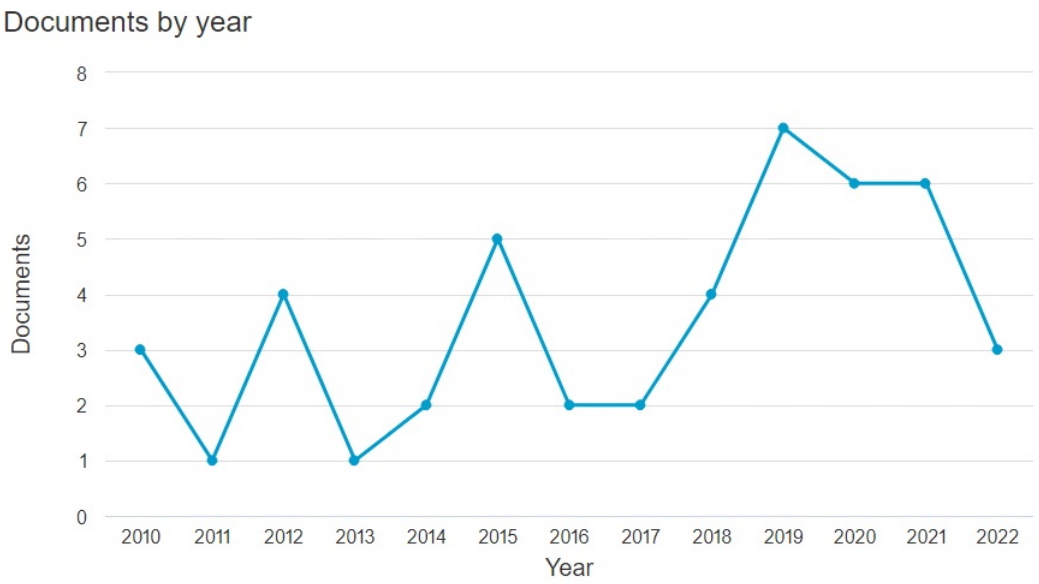

Figura 4.1: Documentos por Ano

Fonte: Gerado a partir do Scopus

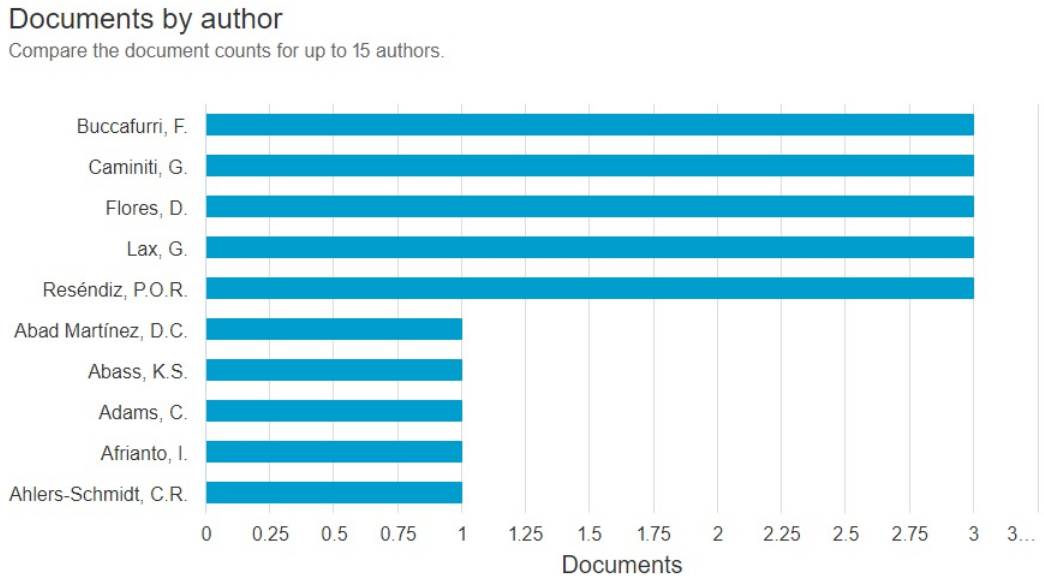

Figura 4.2: Documentos por Autor

Fonte: Gerado a partir do Scopus

Conforme previsto no TEMAC, a análise da pertinência dos conteúdos e a exclusão dos trabalhos inacessíveis reduziu para 9 (nove) o número de documentos relevantes ao contexto da pesquisa. Os trabalhos selecionados para a identificação das vulnerabilidades do documento digital estão elencados na bibliografia por meio das referências [11] [39], [40], [41], [42], [43], [44], [45] e [46].

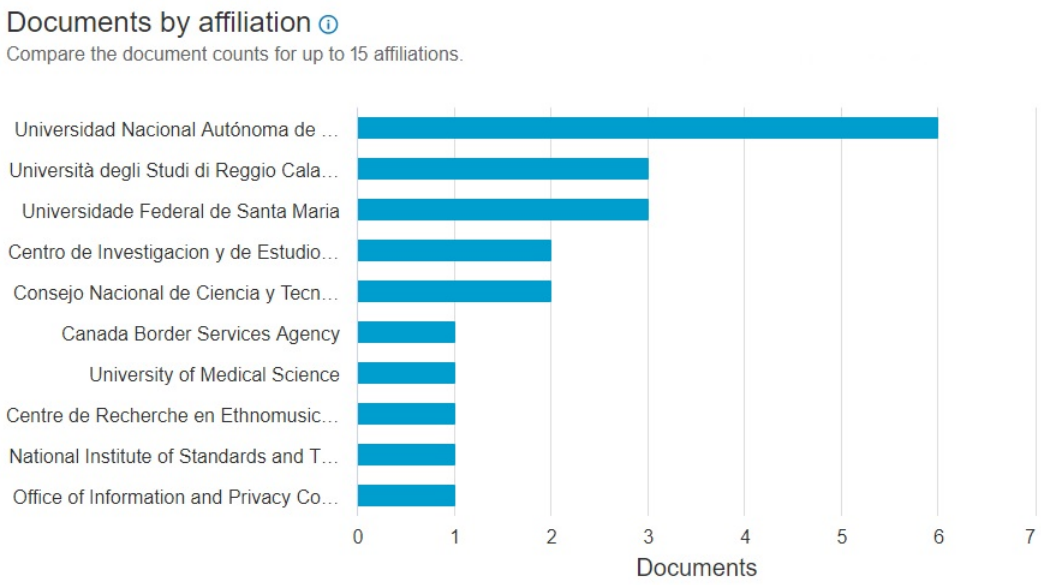

Figura 4.3: Documentos por Instituição

Fonte: Gerado a partir do Scopus

4.2 VULNERABILIDADES DO DOCUMENTO DIGITAL E PROPRIEDADES AFETADAS

Considerando o papel da ABIN, de subsidiar decisões de governo e políticas de Estado, é imprescindível que o conhecimento de inteligência produzido, materializado no documento de inteligência, seja dotado de propriedades essenciais que lhe assegurem a utilidade e oportunidade. A confiabilidade do documento de inteligência, no que se refere à veracidade da informação, à qualidade do conteúdo que ele apresenta, é um aspecto que está associado aos métodos e à doutrina empregados pela Atividade de Inteligência, e, portando, não é uma propriedade que esteja sujeita às vulnerabilidades advindas da transformação digital.

Neste ponto é importante relembrar a abordagem adotada para o propósito desta pesquisa, que considera os princípios fundamentais da segurança da informação (Confidencialidade, Integridade e Autenticidade) como propriedades ou características essenciais aos documentos digitais. Assim, a análise dos documentos selecionados apontou uma relação entre as vulnerabilidades e as propriedades do documento digital, onde o efeito dessas vulnerabilidades sobre as propriedades é verificada pelas repedidas menções, apuradas na íntegra dos documentos examinados, Tabela 4.2.

A digitalização do documento trouxe benefícios, mas também vulnerabilidades inerentes ao documento digital. A volatilidade e o risco de violabilidade podem comprometer a autenticidade, o sigilo, a integridade e a identidade do documento digital [42]. A dependência tecnológica tornou o documento digital sujeito aos riscos de obsolescência, estas vulnerabilidades afetam a interoperabilidade, a disponinibilidade e a longevidade [11].

Sob o aspecto jurídico-administrativo, a volatilidade do documento digital põe em suspeição sua acurária e sua autenticidade, enquanto elemento probatório e como evidência, o que levanta questionamentos também acerca da sua irrefutabilidade para sustentar fato ou evento a que se refere [6].

Tabela 4.2: Relação: Vulnerabilidades x Propriedades do Documento Digital

Categorias (Vulnerabilidades)	Subcategorias (Propridades)	Quantidade de Menções
Volatilidade Violabilidade	Autenticidade	7 vezes
	Confidencialidade	2 vezes
	Integridade	14 vezes
	Identidade	4 vezes
Dependência Tecnológica Obsolescência	Interoperabilidade	2 vezes
	Disponibilidade	10 vezes
	Longevidade	2 vezes

Fonte: Elaborado pelo autor

Por autenticidade, entende-se a "qualidade de um documento ser o que diz ser e de que está livre de adulteração ou de qualquer outro tipo de corrupção"[30]. A identidade do documento é o conjunto de atributos que o caracterizam como único e diferente dos demais. A integridade diz respeito à manutenção da forma e do conteúdo ao longo do tempo, que no contexto digital significa a preservação de cada bit que compõe o arquivo digital [7] [12]. Para o propósito desta pesquisa, integridade e identidade são consideradas componentes da autenticidade, uma vez que prevalece o entendimento que a autenticidade documental seja uma propriedade composta da combinação da identidade com a integridade [27] [47] [42].

A confidencialidade é a propriedade associada ao sigilo de certos dados, informações ou documentos que possuem classificação de sigilo e não podem ser disponibilizados ou divulgados sem autorização [29], a divulgação indevida pode acarretar danos intangíveis.

A longevidade é a propriedade voltada à preservação do documento digital, assegurando sua utilidade a despeito de sua idade, em contraponto ao risco de deterioração dos mecanismos de armazenamento e à descontinuidade de formatos e tecnologias digitais, o que termina por acarretar problemas de visualização, de embaralhamento (compressão, encriptação), de interação, de custódia e de tradução [48].

A interoperabilidade do documento digital refere-se à sua conformidade técnica no atendimento aos protocolos e/ou padrões que assegurem o diálogo sistêmico entre diferentes plataformas para a troca de dados, de metadados e do documento em si [49].

A disponibilidade do documento está intrinsicamente associada à infraestrutura computacional (computadores, dispositivos de rede e de armazenamento), sendo também vinculada aos níveis de serviços TIC de que o documento dependa.

Sob a perspectiva da segurança e da tecnologia, é imprescindível que o documento de inteligência produzido pelo órgão possua, entre suas propriedades, a autenticidade (do que é autêntico, legítimo), a confidencialidade (do que é sigiloso), a disponibilidade (do que é disponível, acessível), a interoperabilidade (do que dialoga) e a longevidade (do que tem vida longa).

4.3 ANÁLISE E VALIDAÇÃO DE MEDIDAS DE MITIGAÇÃO

Esta etapa tem o propósito de, por meio da descrição analítica, apontar e validar a eficácia de medidas de mitigação das vulnerabilidades do documento digital. O foco principal é assegurar, para o documento de inteligência produzido pelo órgão, a sustentação das propriedades afetadas pelos processos de digitalização e digitização, quais sejam: a autenticidade, a confidencialidade, a disponibilidade, a interoperabilidade e a longevidade.

A relação entre as propriedades e as medidas apontadas (Tabela 4.3) mostra, de forma sintética, onde uma medida promove (P) efetivamente ou contribui (C) para sustentar propriedade(s) do documento de inteligência digital. A CCDA e CD são medidas que têm uma inter-relação de complementariedade entre si em prol do documento de inteligência digital, uma vez que a CCDA, por meio do RDC, promove (P) a autenticidade e a confidencialidade, como também contribui (C) para a longevidade. A CD, que tem o *DCC Framework* como referência, promove (P) a longevidade e assim contribui (C) para a disponibilidade documental. A infraestrutura computacional na arquitetura *Private Cloud Computing*, na modalidade *On premises*, é uma medida capaz de promover (P) a alta disponibilidade requerida e também contribuir (C) para a confidencialidade do documento de inteligência. Por sua vez, a solução SIGAD, ao atender ao padrão *CMIS*, promove (P) a interoperabilidade documental, além de contribuir (C) com a autenticidade e confidencialidade, uma vez que também esteja em conformidade com o modelo *e-Arq Brasil*.

Tabela 4.3: Relação: Propriedades do Documento Digital x Medidas

	Medidas / Referências	CCDA	CD	Infraestrutura Computacional	SIGAD
		RDC-Arq	DCC Framework	Private Cloud Computing On Premises	e-Arq Brasil CMIS
Propriedades	Autenticidade	P			C
	Confidencialidade	P		C	C
	Longevidade	C	P		
	Disponibilidade		C	P	
	Interoperabilidade				P

Fonte: Elaborado pelo autor

4.3.1 Impacto da CCDA nas Propriedades do Documento Digital

No contexto analógico, a confiança depositada sobre o documento está intrinsicamente relacionada à presunção de autenticidade (composta pela identidade e integridade), que se baseia em dois pontos: (i) a evidenciação material de fatores inseparáveis, por meio de análise da forma, conteúdo e suporte; (ii) a existência de fluxo controlado, por meio de uma cadeia de custódia ininterrupta, que assegure a transferência inequívoca do documento de um custodiante a outro, sem levantar dúvida ou suspeição sobre sua autenticidade [50].

No contexto digital, a presunção de autenticidade enfrenta desafios para a evidenciação material, uma

vez que a análise de forma, conteúdo e suporte não assegura a identidade do arquivo digital (*checksum* ou *hash*[13]), tornando-o passível de duplicação e de adulteração. Assim a presunção de autenticidade do material digital toma por base a confiança depositada no ambiente de custódia e no repositório digital que o armazena.

Ainda que confiança seja uma grandeza intangível, de difícil mensuração no contexto digital, a presunção de confiança no repositório digital decorre da adoção de rígidos controles e padrões, evidenciada pela análise e verificação de fatores imprescindíveis, como mecanismo de armazenamento, permissão de acesso, registro de metadados, procedimentos de preservação, trilhas de auditoria, verificação de integridade, mecanismos de autenticação, entre outros, que, quando auditados e certificados, elevam o repositório à categoria de Repositório Digital Confiável (RDC).

No âmbito nacional, a confiança no repositório digital pressupõe o atendimento das Diretrizes para a Presunção de Autenticidade de Documentos Arquivísticos Digitais (RDC-Arq), estabelecidas pela Resolução Nº 43 do Conselho Nacional de Arquivos (CONARQ). Esta norma, que tem o objetivo de garantir a autenticidade (identidade e integridade), a confidencialidade, a disponibilidade, o acesso e a preservação de documentos arquivísticos digitais, está alicerçada na conformidade dos repositórios digitais com modelos e padrões consolidados, conferindo-lhes acreditação em nível internacional [29].

São estes os padrões e modelos que dão sustentação à confiabilidade do RDC-Arq:

1. Relatório RLG/OCLC-2002: Relatório resultante da força tarefa composta pelo Research Library Group - RLG e pelo Online Computer Library Center - OCLC estabelece os atributos e as responsabilidades indispensáveis aos repositórios digitais confiáveis (OCLC and CRL, 2007).

2. Open Archival Information System - OAIS (ISO 14721, 2003): Modelo que apresenta uma abstração funcional do repositório digital, dos metadados necessários à preservação e ao acesso dos materiais digitais, bem como dos serviços de arquivamento e de enpacotamento das informações documentais.

3. Trustworthy Repository Audit Certification – TRAC -2007: Estabelece uma trilha para certificação de repositório digital, que contempla auditorias e a aplicação de critérios e de checklist definidos. Um repositório digital com uma certificação TRAC conta com a acreditação de RDC (OCLC and CRL, 2007). A certificação TRAC evoluiu para Audit and Certification of Trustworthy Digital Repositories – ACTDR, padronizada por meio da ISO 16363:2012.

4. Requirements For Bodies Providing Audit and Certification Of Candidate Trustworthy Digital Repositories: Define uma prática sobre a qual devem se basear as operações de uma organização que realiza auditorias para avaliar a confiabilidade de repositórios digitais com vistas ao fornecimento da certificação apropriada [28].

5. Preservation Metadata Implementation Strategies – PREMIS: Apresenta dicionário de dados de referência para a construção e preservação de metadados [51]. O alinhamento com o padrão PREMIS confere ao RDC o emprego das melhores práticas no mapeamento e preservação dos metadados documentais, de forma a preservar a relação orgânica do documento no repositório.

[13]Checksum, hash: identificadores unívocos resultantes de algoritmo aplicado aos bits de objeto binário

6. General International Standard Archival Description - ISAD(G): Norma arquivística que tem o propósito de identificar e explicar o contexto e o conteúdo do documento arquivístico com vistas à promoção de seu acesso [52].

7. Norma Brasileira de Descrição Arquivística - NOBRADE: Uma adaptação da norma ISAD(G) à realidade brasileira, visa facilitar o acesso e o intercâmbio de informações, em âmbito nacional e internacional, por meio de descrições consistentes, apropriadas e autoexplicativas dos documentos digitais arquivísticos [53];

8. Modelo de Requisitos para Sistemas Informatizados de Gestão Arquivística de Documentos – e-ARQ Brasil: A parte II desse modelo apresenta um capítulo sobre metadados, onde se encontra um mapa de relacionamento entre o documento e as demais entidades a ele relacionadas, bem como um detalhamento dos metadados afetos a estes elementos [23].

9. Open Archives Initiative Protocol for Metadata Harvesting - OAI-PMH: Protocolo com especificação para coleta de metadados, baseado nos padrões abertos HTTP[14] e XML[15], que visa a facilitar a disseminação eficiente de conteúdo do repositório [54].

10. Metadata Encoding Transmission Standard - METS: Padrão baseado em XML para de codificação, o empacotamento e a transmissão de metadados descritivos, administrativos e estruturais relativos aos objetos digitais [55].

11. Encoded Archival Description - EAD: Padrão, também baseado em XML, para a descrição, estruturação e interoperabilidade dos metadados arquivísticos referenciais, permitindo a decodificação e apresentação das informações de forma estruturada [56].

No que concerne à gestão e preservação de documentos digitais, a confiança está alicerçada na adoção e manutenção da Cadeia de Custódia Digital Arquivística (CCDA), que pressupõe a transferência protocolar de documentos entre o ambiente de gestão documental e o ambiente de preservação e acesso, onde se faz uso de Repositório Digital Confiável (RDC) como plataforma de preservação arquivística [30].

Desse modo, é possível verificar que a CCDA, por meio o RDC, promove a autenticidade e a confiabilidade documental por meio da transferência protocolar do ambiente de gestão documental para o ambiente de preservação e acesso, além de contribuir para a longevidade documental fazendo uso de repositórios digitais ativos, que implementam diversas ações de preservação.

Existem várias implementações de repositórios digitais que atendem ao modelo OAIS (ISO 14721, 2003) [57], dentre elas algumas de código aberto, distribuídas sob licenciamento General Public License (GPL) [58], onde se destacam *Archivematica* [59] e *Repository of Authentic Digital Objects* (RODA) [60] por serem soluções maduras e sustentáveis no longo prazo.

Estudo comparativo entre RODA e Archivematica apesenta tabulações das diversas características de cada uma destas soluções [61]. O quadro comparativo que se segue apresenta as características básicas, um pouco da história, a evolução e os requisitos destas duas soluções, Tabela 4.4.

[14] Hyper Text Transfer Protocol
[15] eXtensible Markup Language

Tabela 4.4: Características Básicas do RODA e Archivematica

Nome da Solução	Tipo de Software	Fundador	País de Origem	Ano de Liberação	Licença GPL	Versão Corrente
RODA	Repositório Digital	Keep Solutions, Universidade de Minho	Portugal	2002	LGPLv3	v4.3.1 Junho/2022
Archivematica	Sistema de Preservação Digital	Artefactual.inc	Canada	2012	AGPLv3	v1.13.2 Junho/2022

Fonte: Open source software for digital archiving: A comparative study on roda and archivematica [61]

O quadro seguinte traz uma tabulação das características técnicas e dos pré-requisitos técnicos para a instalação das duas soluções, Tabela 4.5.

Tabela 4.5: Característica Técnicas do RODA e Archivematica

Sistemas Operacionais Suportados	Dependência Técnica	Estratégia de Preservação Digital	Plataforma Multilingual	Comunidade e Suporte	Integração Sistema de Terceiros	Padrões e Protocolos Suportados
RODA						
Windows, Linux, Mac OS	Java 8	Migração, Encapsulação, Emulação	Sim	Sim	Sim	OAI, OAI-PMH v2, Doublin, Core, PREMIS, METS, EAD 2002, EAD 3
Archivematica						
Linux	MySQL, Gearman, Elastic Search	Migração, Emulação	Não	Sim	Sim	OAI, Doublin, Core, PREMIS, METS

Fonte: Open source software for digital archiving: A comparative study on roda and archivematica [61]

As soluções Roda e Archivematica dão suporte a diferentes formados de documentos digitais, conforme Tabela 4.6.

A escolha da solução de repositório digital deve considerar diversos aspectos, as suas características técnicas, tais como sistema operacional, dependência, nível integração, comunidade de sustentação, protocolos e formatos suportados. A solução escolhida também pode contribuir com a longevidade documental por meio da automatização de ações da estratégia de preservação, entre outras.

Contudo, o reconhecimento da confiabilidade de repositório digital em nível nacional (RDC-Arq) não depende da solução de repositório *per si*, seja ela de código aberto ou proprietária, mas de processo de auditoria e certificação conforme a *Audit and Certification of Trustworthy Digital Repositories* (ACTDR -

Tabela 4.6: Formatos Digitais Suportados por Roda e Archivematica

Formato de Digital	RODA	Archivematica
Imagem	TIFF, JPG, JPEG, PNG, BMP, GIF, ICO, XMP, TGA	TIFF, JPG, JPEG, PNG, BMP, GIF, JP2, PCT, PSD, TGA, RAW
Audio	WAV, MP3, MP4, FLAC, AIFF, OGG, WMA	WAV, MP3, AC3, AIFF, WMA
Vídeo	AVI, MOV MPG, MPEG, VOB, MPV2, MP4, WMV, QT	AVI, MOV, MPG, MPEG, MP4, SWF, WMV, FLV
Texto	PDF, XML	TXT, PDF, RTF
Aplicação	DOC, DOCX, XLS, XLSX, PPT, PPTX, ODT, RTF, TXT, ODP, ODS	DOC, DOCX, XLS, XLSX, PPT, PPTX, WPD
Vetorial	AI, CDR, DWG	AI, EPS, SVG
Email	EML, MSG	PST, MailDir
Outros	BIN	X3F, 3FR, ARW, CR2, CRW, DCR, DNG, ERF, KDC,

Fonte: Open source software for digital archiving: A comparative study on roda and archivematica [61]

ISO 16363:2012), que permeia requisitos de diversos aspectos da Infraestrutura Organizacional, da Gestão dos Objetos Digitais e da Infraestrutura e Segurança da Gestão de Riscos, conforme os macro requisitos da ACTDR elencados abaixo, onde se manteve a numeração original para fins de comparação com o modelo de requisitos [62]:

3. INFRAESTRUTURA ORGANIZACIONAL

3.1. GOVERNANÇA E VIABILIDADE ORGANIZACIONAL - Definição da missão que o repositório deve assumir quanto ao compromisso com a preservação e o acesso em longo prazo, explicitada nos planos de sucessão e na política de recolhimento para custódia; onde devem estar declarados os compromissos nos casos de alteração na custódia.

3.2. ESTRUTURA ORGANIZACIONAL E DE PESSOAL - Gestão das competências de pessoal necessárias para o cumprimento das funções do repositório, com a definição do organograma eu explicita a divisão de funções e de responsabilidade; com o incentivo ao desenvolvimento continuado de habilidade do pessoal envolvido.

3.3. POLÍTICAS DE RESPONSABILIDADE E PRESERVAÇÃO - Políticas com definição de comunidade e de base de conhecimento, bem como de cumprimento de serviços, acompanhamento da evolução tecnológica, e de feedback de produtores e consumidores, além de ações de transparência, de preservação, registro de transformações e de avaliação de integridade.

3.4. SUSTENTABILIDADE FINANCEIRA - Definição de processos de planejamento para curto e longo prazo, com ajustes periódicos, bem como procedimentos financeiros transparentes com auditoria, além do monitoramento continuo dos riscos, benefícios, investimentos e despesas.

3.5. CONTRATOS, LICENÇAS E PASSIVOS - Manutenção de contratos e acordos de depósito apropriados aos materiais digitais de outra organização, capazes de especificar e transferir todos os direitos de preservação necessários, de forma que permita rastrear e gerenciar os direitos de propriedade in-

telectual e possíveis restrições sobre o uso do conteúdo, definindo as políticas com base na legislação para conteúdos digitais com propriedade ou direitos não especificados claramente.

4. GESTÃO DE OBJETOS DIGITAIS

4.1 ADMISSÃO: AQUISIÇÃO DE CONTEÚDO - Identificação das propriedades significativas dos objetos digitais que serão preservadas, que devem ser associadas às informações necessárias ao pacote SIP. Verificar se as fontes de proveniência dos objetos admitidos são autenticadas, e assegurar a execução de correções necessárias a cada pacote SIP submetido. Além manter o controle físico dos objetos digitais para preservá-los, e fornecer respostas adequadas ao produtor durante o processo de submissão.

4.2 ADMISSÃO: CRIAÇÃO DO AIP - Nomenclatura geral de pacotes AIP ou classe de informação, com a preservação das propriedades significativas, além da descrição da transformação do pacote SIP em AIP.

4.3 PLANEJAMENTO DA PRESERVAÇÃO - Identificação e documentação das estratégias de preservação, notificação quando a informação de representação entrar em risco de obsolescência, adequação dos planos de preservação conforme resultado dos monitoramentos e evidenciação do planejamento de preservação.

4.4 PRESERVAÇÃO DO AIP - Evidenciação da preservação das informações de conteúdo dos pacotes AIP por meio de estratégias de preservação documentadas de metadados e de ações aplicadas no seu tratamento; monitoramento continuado da integridade dos pacotes AIP, manutenção de registros de ações e processos administrativos pertinentes à preservação.

4.5 GESTÃO DA INFORMAÇÃO - Evidenciação de captura dos metadados de descrição necessários aos pacotes AIP para a identificação dos materiais, demonstrando que pacote AIP mantem a integridade referencial com as informações descritas.

4.6 GESTÃO DA ACESSO - Documentação e comunicação de opções de acesso e entrega que estão disponíveis à comunidade designada, registro de todas as solicitações de acesso, cumprimento dos acordos relacionados às condições de acesso; definição de política de acesso segura via sistema de gerenciamento, aos contratos de depósito; demonstração de que todas as solicitações de acesso resultam em uma resposta de aceitação ou rejeição, bem como o registro de todas as falhas de gerenciamento de acesso e analise dos casos de negação de acesso.

5. INFRAESTRUTURA E SEGURANÇA DA GESTÃO DE RISCOS

5.1 GESTÃO DE RISCOS DE INFRAESTRUTURA TÉCNICA - Garantia de que as funções de repositório sejam suportadas pelos principais sistemas operacionais, de modo que seja possível assegurar suporte de hardware e software adequado às funcionalidades de backup e suficientes aos conteúdos armazenados; detecção de corrupção ou perda de bits por meio de mecanismos de análise de erro, com notificação à administração de todos estes incidentes e as medidas adotadas para reparar ou substituir os dados afetados; definição dos processos de atualização das mídias de armazenamento,

do hardware e da segurança software, com registro de gestão de mudanças; conformidade das tecnologias de hardware e software com os serviços que presta à comunidade designada, bem como, a definição e aplicação de procedimentos para monitorar e avaliar a necessidade de mudanças nas tecnologias de hardware e/ou software utilizadas.

5.2 GESTÃO DO RISCO DE SEGURANÇA - Definição e execução de análise sistemática em relação a dados, sistemas, pessoal, planta física e segurança; determinação da funções, responsabilidades e autorizações relacionadas à implementação de mudanças no sistema, além da definição de plano de preparo e recuperação de desastres.

4.3.2 O Papel da Curadoria Digital na Longevidade do Documento Digital

A Curadoria Digital se ocupa do desenvolvimento de melhores práticas para assegurar a sustentabilidade e a longevidade do material digital – onde se insere o documento digital – seja por meio do planejamento ou da gestão ativa dos objetos digitais com foco na qualidade, na manutenção e no reuso ao longo de todo o seu ciclo de vida [34].

Nesta perspectiva, o *Digital Curation Centre* (DCC), comitê global para estudos e gerenciamento de material digital, publicou um framework baseado no *DCC Curation Lifecycle Model* [35]. Esse modelo permeia as abordagens e os elementos disruptivos advindos da transformação digital das últimas décadas [10].

O DCC framework oferece definições de objetos de dados e um plano de ações de curadoria dividido em etapas conforme a sazonalidade documental: Todo Ciclo de Vida, Ações Sequenciais e Ações Ocasionais. Esse plano contempla ações que vão desde o nascimento do documento (conceituar, criar e receber) até sua eventual morte (descartar) ou seu renascimento em outro formato (migrar), garantindo sua longevidade enquanto for útil.

Na etapa Ações de Todo Ciclo de Vida são previstas ações cotidianas dentro do ciclo de vida documental, voltadas para a representação, planejamento, compartilhamento tecnológico e preservação, conforme apresentado na Tabela 4.7.

A etapa de Ações Sequenciais descreve um plano de ações que vai desde a concepção do documento (conceituar, criar e receber), passando ações de cuidados com a sustentação, preservação e utilidade (avaliar e selecionar, ingerir, preservar, armazenar e acessar, usar e reutilizar e transformar), conforme a Tabela 4.8.

A etapa final do DCC framework estabelece ainda as Ações Ocasionais, que contempla a destruição segura do documento (descartar), sua revalidação (reavaliar) ou seu renascimento em formato atualizado (migrar), conforme a Tabela 4.9.

Tabela 4.7: DCC Framework, Ações de Todo Ciclo de Vida

Ações	Descrição
Atribuir Representação e Descrição de Informações	Atribuir metadados administrativos, descritivos, técnicos, estruturais e de preservação, usando padrões apropriados, para garantir a descrição e o controle adequados a longo prazo. Coletar e atribuir informações de representação necessárias para compreender e processar tanto o material digital quanto os metadados associados.
Preservar e Planejar	Planejar a preservação durante todo o ciclo de vida da curadoria de material digital. Isso incluiria planos para gerenciamento e administração de todas as ações do ciclo de vida da curadoria.
Participar e Monitorar a Comunidade	Manter-se atento às apropriadas atividades comunitárias e participar do desenvolvimento compartilhado de padrões, ferramentas e software adequados.
Curar e Preservar	Estar atendo e empreender ações de gerenciamento e administrativas planejadas para promover a curadoria e preservação durante todo o ciclo de vida da curadoria.

Fonte: The DCC Curation Lifecycle Model [35]

A relação entre CCDA e CD não é apenas possível, como sugere Luz [34], mas sobretudo complementar. A CCDA provê autenticidade e confidencialidade ao documento digital, enquanto CD se ocupa da sua preservação e longevidade. Esta parceria se reflete em nível de implementação de soluções automatizadas de repositórios digitais ativos, que pode ser observado quando os softwares, a exemplo de Archivematica e RODA, contemplam a execução de ações de curadoria digital, como as previstas no DCC framework, dentro do espaço de confinamento reservado ao repositório, sem intervenção de agente externo (usuário ou sistema).

4.3.3 Importância do Padrão CMIS para a Interoperabilidade do Documento Digital

O modelo referencial *Content Management Interoperability Services* (CMIS) oferece especificações detalhadas de modelo de dados, de objetos de dados e de serviços de repositório que permitem a interoperabilidade de conteúdo entre os repositórios que estejam em conformidade com este padrão [25]. O nível de detalhamento das especificações do padrão CMIS explica porque este modelo se tornou um referencial imprescindível para assegurar o intercâmbio de dados, metadados e documentos digitais entre repositórios.

O padrão CMIS inovou ao reunir especificações comuns e largamente usadas de objetos, propriedades e serviços de repositório, com objetivo de integrar um conjunto cada vez maior de sistemas de gerenciamento de conteúdo (CMS), independentemente e a despeito das tecnologias empregadas em cada um. O CMIS detalha o modelo de domínio, de dados e de repositório, contemplando objetos para documento, diretório, relacionamento, permissões, versão, consulta, entre outros; especificando os serviços essenciais ao repositório, implementações de serviços por meio de API[16] padronizadas, em geral Web Services[17].

[16] Applicaton Program Interface: interface funcional de plataformas de software
[17] Web Service: Interface funcional de plataforma implementada sob protocolo HTTP

Tabela 4.8: DCC Framework, Ações Sequenciais

Ações	Descrição
Conceituar	Conceber e planejar a criação de dados, incluindo método de captura e opções de armazenamento.
Criar e Receber	Criar dados incluindo metadados administrativos, descritivos, estruturais e técnicos. A preservação de metadados também podem ser adicionadas no momento da criação. Receber dados, de acordo com políticas de coleta documentadas, de criadores de dados, outros arquivos, repositórios ou centros de dados e, se necessário, atribuir metadados apropriados.
Avaliar e Selecionar	Avaliar os dados e selecionar para curadoria de longo prazo e preservação. Aderir a guias, políticas e requisitos legais documentados.
Ingerir	Transferir dados para um arquivo, repositório, centro de dados ou outro custodiante. Seguir as orientações, políticas ou requisitos legais documentados.
Ação de Preservação	Realizar ações para garantir a preservação e retenção a longo prazo da natureza autorizativa dos dados. As ações de preservação devem garantir que os dados permaneçam autênticos, confiáveis e utilizáveis, mantendo sua integridade. Ações incluem limpeza de dados, validação, atribuição de metadados de preservação, atribuição de informações de representação e garantindo estruturas de dados ou formatos de arquivo aceitáveis.
Armazenar	Armazenar os dados de maneira segura, respeitando os padrões relevantes.
Acessar, Usar e Reutilizar	Certificar-se de que os dados estejam acessíveis tanto para uso como para reuso, diariamente. Isso pode ser na forma de informações publicamente disponíveis. O robusto controle de acesso e procedimentos de autenticação podem ser aplicáveis.
Transformar	Criar novos dados a partir do original, por exemplo - Pela migração para um formato diferente. - Pela criação de um subconjunto, por seleção ou consulta, para criar resultados recém-derivados, talvez para publicação.

Fonte: The DCC Curation Lifecycle Model [35]

O padrão CMIS tornou-se a referência para a qual diferentes plataformas CMS/ECM têm de convergir para viabilizar a interoperabilidade entre si e o diálogo com o cliente produtor e consumidor da informação. Se por um lado, a convergência de CMS de mercado para o padrão CMIS, por meio da implementação de camada *CMIS Interface*, que assegura a interoperabilidade entre diferentes soluções de gerenciamento de conteúdo; por outro, estas plataformas implementam o armazenamento em repositórios digitais nativos e heterogêneos, fazendo uso de modelos proprietários de conteúdo e dados, ou seja, sem aderência e nem conformidade ao modelo OAIS, Figura 4.4.

Tabela 4.9: DCC Framework, Ações Ocasionais

Ações	Descartar
Conceituar	Descartar os dados que não foram selecionados para curadoria e preservação de longo prazo de acordo com as políticas documentadas, orientações ou requisitos legais. Normalmente, os dados podem ser transferidos para outro arquivo, repositório, centro de dados ou outro custodiante. Em alguns casos, os dados são destruídos. A natureza dos dados pode, por razões legais, exigir a destruição segura.
Reavaliar	Retornar dados que falham nos procedimentos de validação para avaliação posterior e nova seleção.
Migrar	Migrar os dados para um formato diferente. Isso pode ser feito de acordo com o ambiente de armazenamento ou para garantir a imunidade dos dados de obsolescência de hardware ou software.

Fonte: The DCC Curation Lifecycle Model [35]

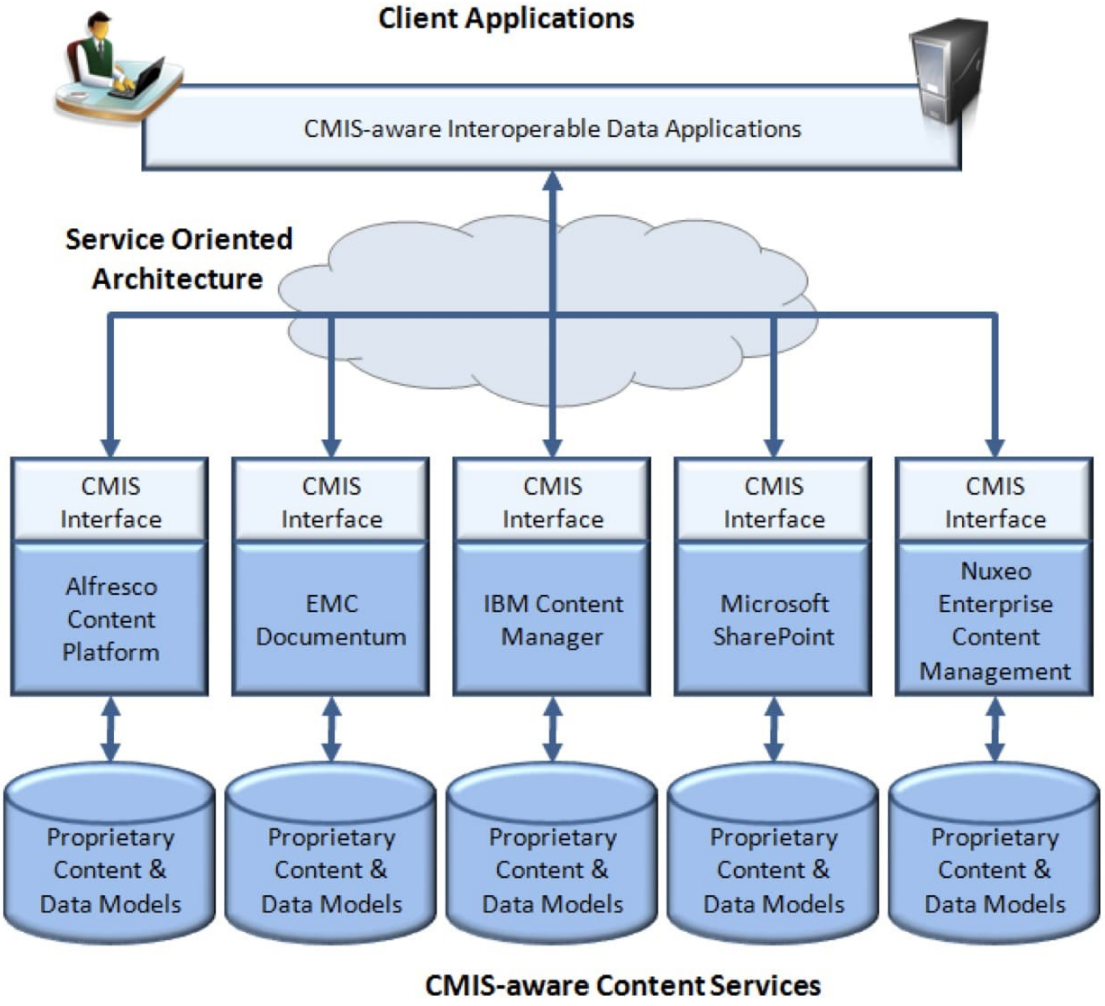

Figura 4.4: Interoperabilidade via CMIS

Fonte: Addressing Contemporary Requirements for Content Integration [63]

A conformidade com o padrão CMIS provê a interoperabilidade sistêmica entre os diferentes gerenciadores de conteúdo, a despeito da opção de CMS feita pelos membros do Sistema Brasileiro de Inteligência (SISBIN). Esta interoperabilidade é fundamental para a Abin, enquanto órgão central de inteligência, pois o processo de Produção de Conhecimento de Inteligência requer análise e cruzamento de dados e de documentos coletados das diversas e heterogêneas plataformas de conteúdo implementadas ao sabor de cada instituição-membro desse sistema.

4.3.4 Cenários de Integração entre SIGAD e RDC-Arq

A CCDA pressupõe uma linha contínua de custodiantes de documentos, sem que quebra na cadeia de custódia documental. No ambiente de gestão documental se dá a elaboração colaborativa de documentos por meio de Sistema Informatizado de Gestão Arquivística de Documento (SiGAD - ECM, CMS, SGD, EDMS, entre outras), já o ambiente de preservação e acesso visa a guarda e preservação do documento, onde conta com a implantação de Repositório Digital Arquivístico (RDC).

O Conselho Nacional de Arquivos preconiza distintos cenários de uso de RDC-Arq em interação com soluções SIGAD, conforme Orientação Técnica n.º 3 [64].

No cenário de integração parcial, o SIGAD armazena nativamente os documentos que estejam nas idades corrente e intermediária, ao passo que os documentos em idade permanente são encaminhados para armazenamento e consulta no RDC-Arq por meio do protocolos SIP e DIP, respectivamente, conforme a Figura 4.5.

Este cenário de uso parcial do RDC-Arq atende às necessidades da CCDA para fins de guarda e preservação documental, garantindo que o documento digital remetido pelo SIGAD mantenha sua autenticidade e confidencialidade dentro do RDC-Arq. Contudo, o SIGAD ainda armazena nativamente os documentos nas idades corrente e intermediária e pode manipulá-los, o que deixa estes documentos (de fases corrente e intermediária) sem a cobertura da confidencialidade e da autenticidade oferecida pelo RDC-Arq.

Em outro cenário de integração completa, a CCDA percorre todo ciclo de vida documental, uma vez que considera o uso de duas instâncias de RDC-Arq que atuam como repositórios digitais diretamente vinculadas ao SIGAD; um para atender ao ambiente de gestão de documental (documentos nas idades corrente e intermediária), e outro voltado para o ambiente de preservação e acesso (documentos na idade permanente). A transferência protocolar de documentos de uma instância de RDC para a outra não configura quebra na cadeia de custódia digital, mas tão somente uma alteração da CCDA [6], conforme a Figura 4.6.

Em qualquer destes cenários, pressupõe-se a integração entre SIGAD e RDC-Arq, contudo este diálogo não se dá de forma direta, mas demanda tradutores. A necessária integração entre SIGAD e RDC encontra alternativas por meio de componentes externos que servem de ponte entre um e outro. O componente Empacotador SIP faz a travessia no sentido SIGAD-RDC, enviando de dados e documentos por meio do protocolo Submission Information Package (SIP); a Plataforma de Acesso viabiliza a travessia no sentido RDC-SIGAD, recebendo dados e documentos do RDC através do protocolo Dissemination Information Package (DIP) e disponibilizando para o SIGAD, entre outras forma de consulta documental.

Se o padrão CMIS teve ampla aderência à camada superior das soluções SIGAD, o que viabilizou a in-

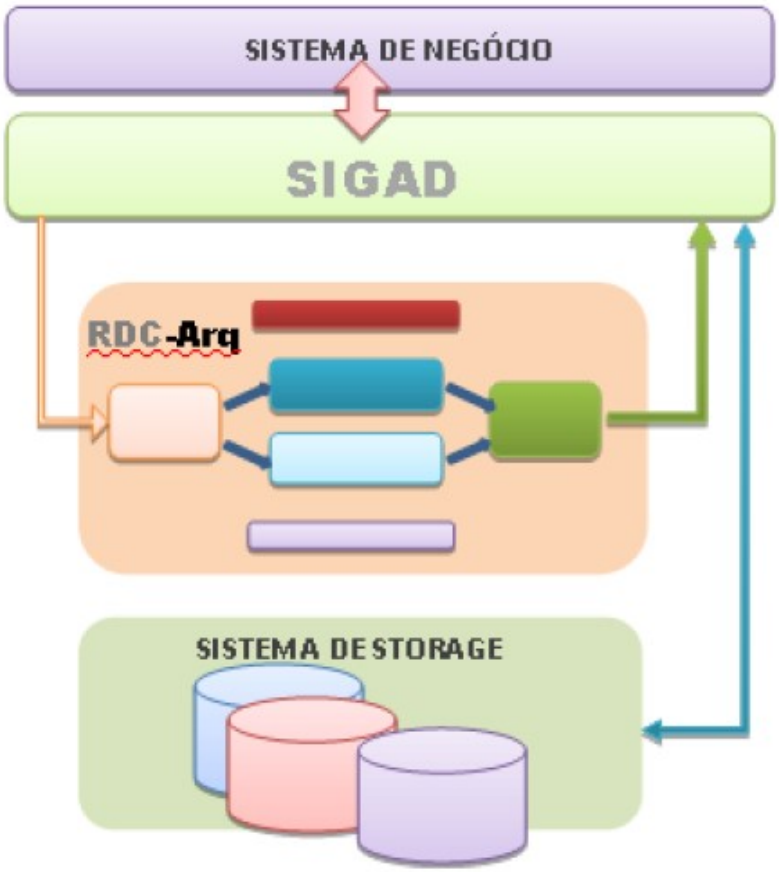

Figura 4.5: Cenário de Integração parcial entre SIGAD x RDC-Arq

Fonte: Cenários de uso de RDC-Arq em conjunto com SIGAD [64]

teroperabilidade entre elas, a despeito de adotarem modelos proprietários de armazenamento de conteúdo e de dados, conforme ilustrado pela *CMIS Interface* (Figura 4.4); o mesmo não ocorreu com o modelo OAIS (base do RDC-Arq). O alto acoplamento[18] entre as camadas mais baixas da arquitetura das soluções SIGAD (modelo de dados e de armazenamento) inviabiliza uma interface de integração direta entre SIGAD e RDC, ou seja, não existe uma suposta camada *OAIS Interface*. Para viabilizar esta integração, o componente de empacontamento SIP termina por recorrer à flexibilidade da *CMIS Interface* para obter os dados, metadados e documentos necessários para a geração do pacote SIP a ser remetido ao RDC-Arq, Figura 4.7.

Neste ponto é importante observar que, embora a Resolução nº 43/CONARQ tenha estabelecido desde Setembro/2015 o RDC-Arq como modelo nacional [29], e até proposto os cenários de uso de RDC-Arq em conjunto com SIGAD [64], a integração entre SIGAD e RDC ainda carece de implementações técnicas, uma vez que o alto acoplamento entre as camadas mais baixas da arquitetura das soluções SIGAD (modelo de dados e de armazenamento) inviabiliza uma interface de integração direta entre SIGAD e RDC. O desafio que permanece é a construção de componentes de empacotamento SIP específicos para cada gerenciador de conteúdo (SIGAD) que faça armazenamento no repositório digital padrão RDC-Arq (OAIS),

[18] Acoplamento: indica o nível de dependência entre as camadas de uma arquitetura tecnológica

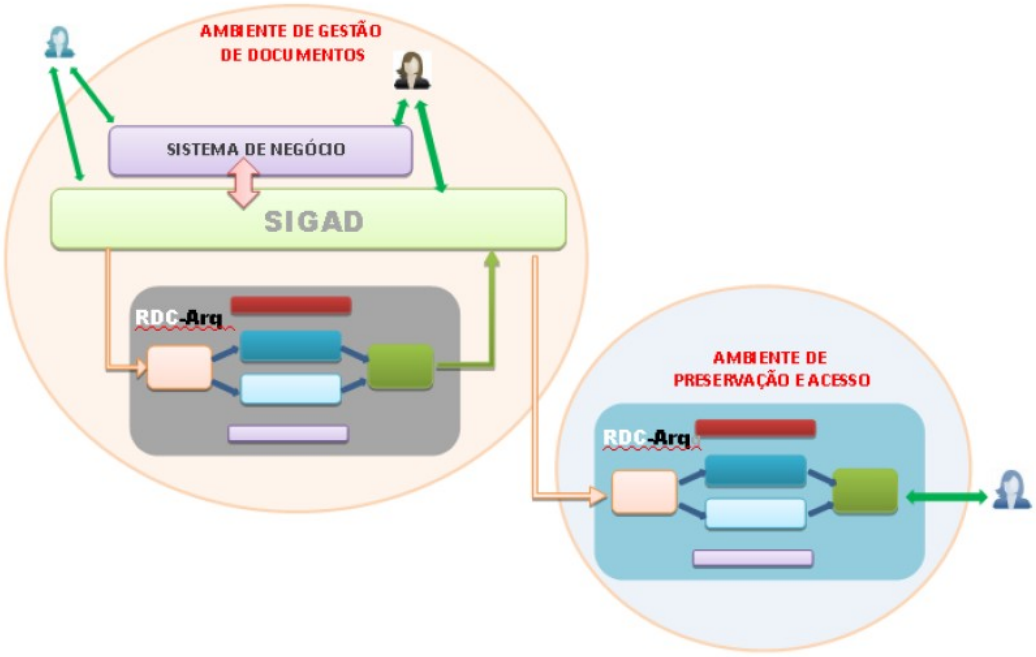

Figura 4.6: Cenário de Integração completa entre SIGAD e RDC-Arq

Fonte: Cenários de uso de RDC-Arq em conjunto com SIGAD [64]

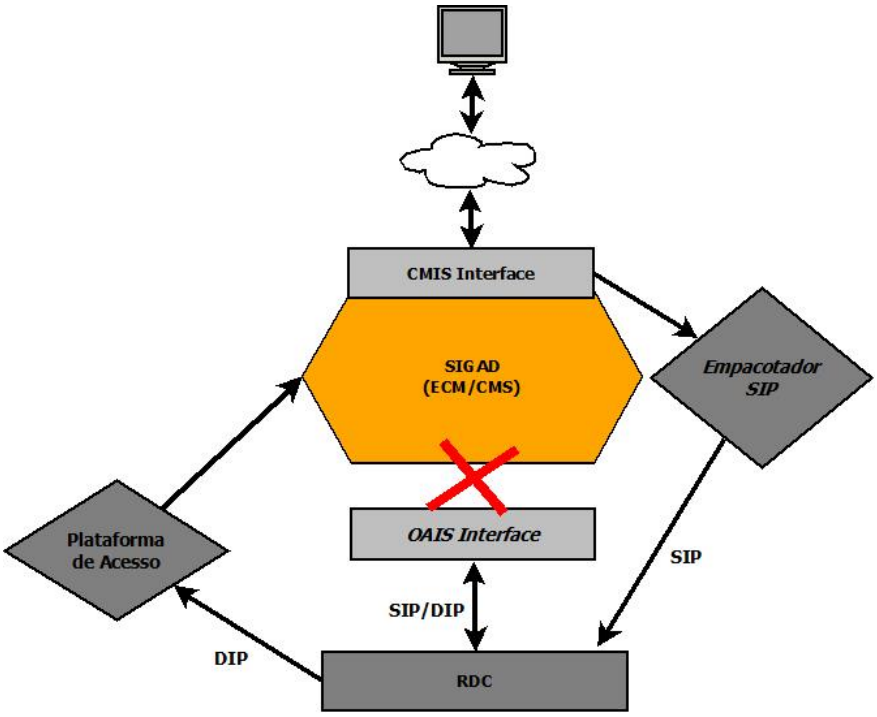

Figura 4.7: Abstração da Integração SIGAD x RDC-Arq

Fonte: Elaborado pelo autor

seja no cenário de integração parcial, apenas com o propósito de preservação digital, seja para o cenário de integração completa, com vistas à implantação da cadeia de custódia digital completa, que contempla todo

ciclo de vida documental.

Um exemplo de integração parcial, CCDA com vistas à preservação, pode ser observado em projeto de iniciativa pública. O Arquivo Nacional (AN), em parceria com o Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), patrocinou estudos e desenvolvimento de software para a integração entre SIGAD e RDC-Arq [65]. O produto resultante do projeto, batizado de HIPÁTIA, implementa um barramento de preservação arquivística para recepcionar pacotes arquivísticos (leia-se pacotes SIP) de diferentes sistemas para armazenamento em repositório digital confiável. Para isso a estrutura funcional do HIPÁTIA preconiza que cada sistema que se conecta ao barramento, no caso o Sistema Eletrônico de Informações (SEI), possua um componente *Crosswalk* próprio para cumprir o papel equivalente ao empacotador SIP, enquanto o Archivematica, na base para da estrutura, cumpre o papel de RDC-Arq, conforme ilustrado na Figura 4.8.

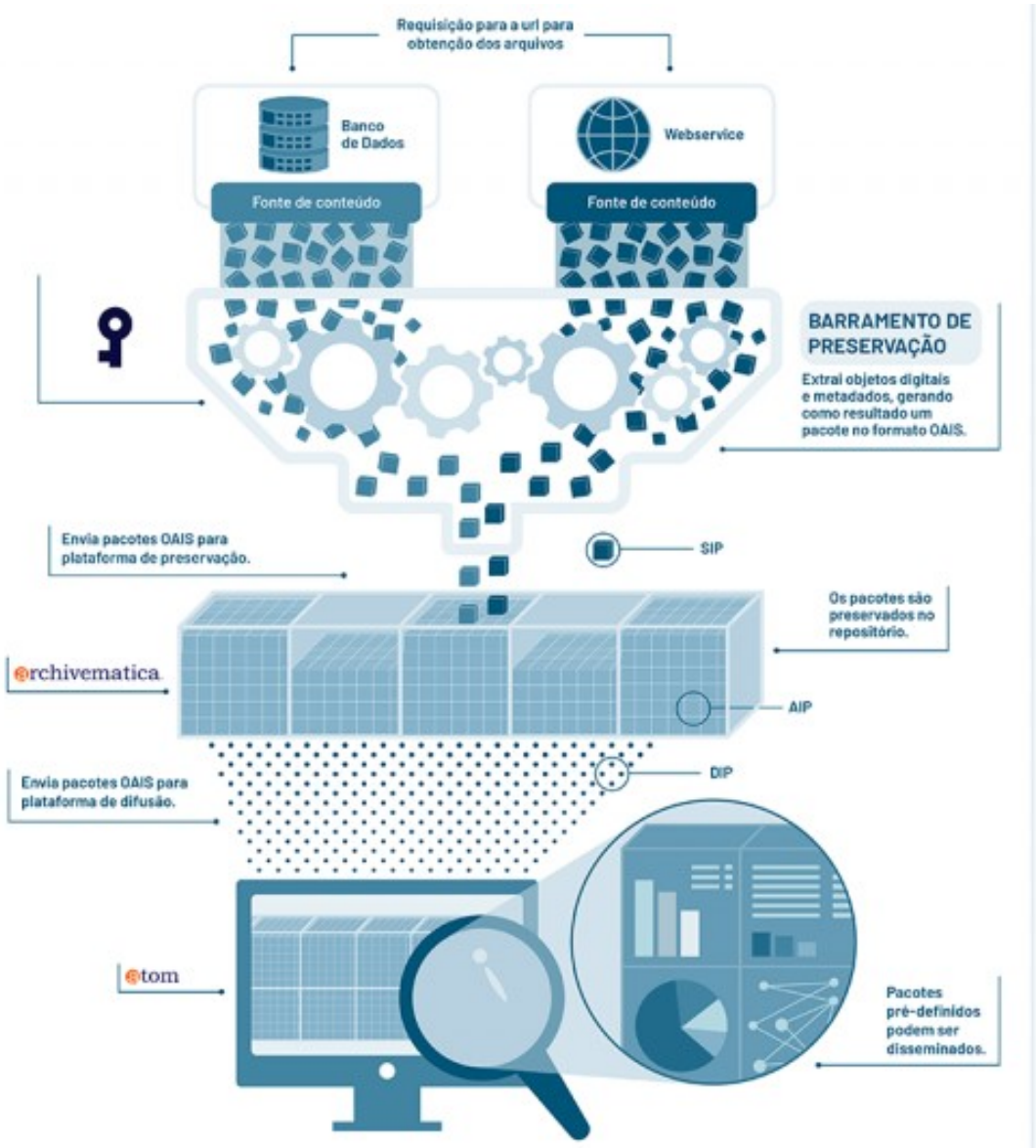

Figura 4.8: Estrutura Funcional do Hipátia

Fonte: Hipátia: Uma ferramenta livre no apoio à preservação digital [65]

4.3.5 Importância da Infraestrutura para a Disponibilidade do Documento Digital

A disponibilidade do documento digital está ligada à sua capacidade de estar acessível a qualquer tempo, o que está intrinsicamente associado à infraestrutura computacional e aos níveis de serviços TIC de que o documento dependa.

O provimento de infraestrutura computacional por meio da arquitetura de *cloud computing*[19] se impôs pelos vários benefícios que traz [66]. Dentre as vantagens desta arquitetura, tais como a simplicidade na administração, redução de custos, baixo impacto de intervenções, elasticidade, entre outras, a alta disponibilidade está entre os principais benefícios, uma vez que as soluções baseadas em nuvem contam com diversas estratégias de resiliência contra falhas e de melhoria de performance, a exemplo da replicação (ou redundância) de recursos e serviços e a da priorização do fluxo de dados lateral (entre serviços) em detrimento do vertical (entre serviço e usuário). Advinda desta arquitetura, surgiram novas abordagens para lidar com a infraestrutura computacional, tais como a Infraestrutura Como Código (IasC) e Infraestrutura Como Serviço (IasS), que permitiram a simplificação do provimento, do suporte e da administração dos recursos computacionais de forma sistemática, performática e escalável. A partir da combinação da arquitetura *cloud computing* com a abordagem IasS, a *Cloud Storage* oferece infraestrutura de armazenamento em repositórios virtuais como serviço [67].

A implementação de RDC encontra na arquitetura *cloud computing* a melhor disponibilidade que a infraestrutura computacional pode oferecer. Contudo, o nível de sensibilidade dos ativos digitais é determinante para a definição do modelo e da modalidade de nuvem computacional a ser empregada. Em termos de modelo a nuvem pode ser *Private Cloud*, *Community/Hybrid Cloud* ou *Public Cloud*, conforme seja de uso exclusivamente institucional, de uso de uma comunidade ou de uso público, respectivamente. A modalidade da *cloud* diz respeito à localização física de instalação da infraestrutura, ou seja, *cloud on premises* pressupõe que toda infraestrutura tecnológica (hardware e software) esteja instalada nas dependências da instituição, ao passo que a modalidade *off premises* faz uso de infraestrutura externa.

Conjugando a tendência de disponibilização de recursos computacionais como serviço em nuvem com a necessidade de preservação digital de longo prazo, a abordagem *Long Term Digital Preservation as a Service* (LTDPaaS) propõe a disponibilização de repositório digital confiável por meio de serviço, em conformidade com o padrão OAIS (ISO 14721, 2003) [68].

4.4 PROPOSTA DE DIRETRIZES PARA IMPLANTAÇÃO DE MEDIDAS

Esta proposta de diretrizes consiste de um conjunto de ações e iniciativas que visam respaldar a implantação das medidas apontadas na etapa antetior, de forma a mitigar as vulnerabilidades indentificadas e assegurar propriedades do documento digital imprescindíveis ao documento de inteligência, quais sejam: a autenticidade, a confidencialidade, a longevidade, a disponibilidade e a interoperabilidade.

Para aumentar a chance de sucesso, a implantação das medidas de mitigação de vulnerabilidades do documento digital deve permear o planejamento desde o nível estratégico ao técnico-operacional, numa

[19] Cloud Computing: modelo de arquitetura computacional, computação em nuvem

arquitetura de governança que respalde as iniciativas de mitigação, assegurando as condições apropriadas de sucesso em nível institucional.

Dentro da arquitetura de govenança, em termos estratégicos, o Plano Estratégico de TIC (PETI ou PETI) deve prever objetivo estratégico que vise definir a Política de Gestão Arquivística de Documentos e a Política de Segurança da Informação (POSIN). Como iniciativa estratégica, o PETIC deve considerar a implantação de Programa de Gestão Arquivística de Documentos, em que constem com as atividades [23]:

- Levantamento da estrutura organizacional e das atividades desempenhadas;
- Levantamento da estrutura organizacional e das atividades desempenhadas;
- Levantamento da produção documental, diferenciando os documentos arquivísticos dos não arquivísticos;
- Levantamento, caso existam, dos sistemas utilizados, internamente, para tratamento de documentos e informações;
- Definição, a partir do levantamento da produção documental, dos tipos de documentos que devem ser mantidos e produzidos, e das informações devem conter;
- Definição e/ou aperfeiçoamento da forma desses documentos;
- Análise e revisão do fluxo dos documentos;
- Elaboração e/ou revisão do plano de classificação e da tabela de temporalidade e destinação;
- Definição dos metadados a serem criados no momento da produção do documento e ao longo do seu ciclo de vida;
- Definição e/ou aperfeiçoamento dos procedimentos de protocolo e de arquivamento dos documentos;
- Definição e/ou aperfeiçoamento dos procedimentos para acesso, uso e transmissão dos documentos;
- Definição do ambiente tecnológico que compreende os sistemas (hardware e software), formatos, padrões e protocolos que darão sustentação aos procedimentos de gestão e preservação de documentos, integrando, quando possível, os sistemas legados;
- Definição da infraestrutura para armazenamento dos documentos não digitais, que compreende espaço físico, mobiliário e acessórios;
- Definição das equipes de trabalho de arquivo e de tecnologia de informação;
- Definição de programas de capacitação de pessoal;
- Elaboração e/ou revisão de manuais e instruções normativas.
- Definição dos meios de divulgação e de capacitação de pessoal; e
- Definição do plano de ação do programa de gestão, com seus objetivos, metas e estratégias de implantação, divulgação e acompanhamento, visando a melhoria contínua.

A POSIN, de formulação obrigatória para os entes da Administração Pública Federal, deve prever expressamente os papeis e as responsabilidades dos proprietários e dos custodiantes dos ativos digitais, bem como a necessidade do armazenamento dos ativos digitais em repositórios que garantam a preservação, a confidencialidade, a autenticidade e a acessibilidade durante todo ciclo de vida documental [69].

No campo tático, o Plano Diretor de Tecnologia da Informação e Comunicações – PDTIC (ou PDTI) deve prever, entre suas ações, a Adoção da Cadeia de Custódia Digital, a Salvaguarda dos Ativos Digitais e Promoção da Curadoria de Ativos Digitais.

No contexto operacional, o Planejamento Orçamentário Anual (POA) deve comtemplar recursos financeiros para a aquisição de soluções (software, hardware, entre outros), bem como para a capacitações que se fizerem necessárias ao cumprimento do PDTIC.

Sob a perspectiva técnica, no que se refere à escolha da solução com função de SIGAD (ECM, CMS, SGD entre outros), deve-se considerar os aspectos funcionais e de metadados previstos na Especificação para Sistemas Informatizados de Gestão Arquivística de Documentos [23], com destaque para a conformidade do SIGAD com o padrão CMIS (OASIS/CMIS) para assegurar a interoperabilidade de dados e sistêmica.

A escolha da solução de repositório digital deve considerar a conformidade com as Diretrizes para a Implementação de Repositórios Arquivísticos Digitais Confiáveis (RDC-Arq) [29], o alinhamento com a Estratégia de Preservação Digital a ser adotada (Classificação, Temporalidade, Padrões, Protocolos e Formatos suportados), o seu grau de maturidade e o nível de automatização oferecido para ações de curadoria digital (repositório digital ativo).

Em termos de interação entre SIGAD e RDC-Arq, deve-se buscar atingir o cenário de integração completa, uma vez que esta configuração implementa a CCDA que protege o documento de inteligência em todo seu ciclo de vida.

A infraestrutura computacional responsável pela disponibilidade do documento digital deve buscar atingir a alta disponibilidade, preferencialmente, pela adoção de arquitetura *private cloud computing* (nuvem privada), na modalidade *on premises* (instalada localmente). Esta arquitetura promove o agrupamento de recursos computacionais em blocos funcionais (*clusters* de aplicação, de banco de dados e de armazenamento), permitindo escalonamento e dimensionamento dos recursos ou serviços, com vistas à resiliência a falhas e ao melhor balanceamento de carga sobre os serviços oferecidos em cada bloco [66].

5 CONCLUSÃO, CONTRIBUIÇÕES E TRABALHOS FUTUROS

Perseguindo o objetivo proposto, que é propor diretrizes à Agência Brasileira de Inteligência (ABIN) para adoção de medidas que assegurem as propriedades imprescindíveis ao documento de inteligência frente às vulnerabilidades advindas do documento digital, esta pesquisa selecionou trabalhos por meio de Revisão Sistemática de Literatura combinada com método TEMAC, realizou a análise de dados, por meio de pesquisa documental e análise de conteúdo, para a identificação de relação de causa-efeito entre as vulnerabilidades e propriedades do documento digital que afetam o documento de inteligência, em especial a autenticidade, a confidencialidade, a interoperabilidade, a longevidade e a disponibilidade. A investigação percorreu os trabalhos acadêmicos e apontou, por meio de descrição analítica, medidas e iniciativas eficazes na mitigação das vulnerabilidades do documento digital e apresentou proposta de diretrizes que permeam a arquitetura de governança do órgão de forma que respalde e assegure as condições de implantação das medidas.

Este estudo comprovou a hipótese de pesquisa e apontou, dentre outras, a adoção da Cadeia de Custódia Digital Arquivística (CCDA) e da Curadoria Digital (CD) como medidas eficazes na mitigação das vulnerabilidades inerentes ao documento digital de maneira a assegurar propriedades imprescindíveis ao documento de inteligência no contexto da Produção de Conhecimento de Inteligência de Estado.

A CCDA, associada à implantação de Repositório Digital Confiável (RDC) - baseado no modelo de *Open Archival Information System (OAIS)*, mostrou-se capaz de promover a custódia documental entre os ambientes de gestão documental e de gestão de preservação, assegurando a autenticidade e a confidencialidade do documento de inteligência. Existem implementações de RDC em código aberto maduras, a exemplo das soluções *RODA* e *Archivematica*. O RDC deve ser um repostório digital ativo, em conformidade com as *Diretrizes para a Implementação de Repositórios Arquivísticos Digitais Confiáveis* (RDC-Arq), com nível de automatização para execução de ações de curadoria digital em espaço de confinamento.

A CD, por meio de planejamento, práticas e ações de curadoria, apresentou os instrumentos para promover a preservação do documento de inteligência e garantir sua longevidade enquanto este tiver valor informacional ou histórico, principalmente quando alicerçada em modelo de curadoria digital internacionalmente consolidado, como o *DCC Curation Lifecycle Model*. CCDA e CD apresentaram uma relação de complementariedade em prol do documento digital, esta relação é observada em repositórios digitais ativos, que implementam diversas ações de preservação previstas nos *frameworks* de curadoria digital.

A infraestrutura computacional mostrou-se um aspecto crítico para a disponibilidade do documento digital, a adoção da arquitetura de *cloud computing*, na modalidade *on premises* oferece robustez, escalabilidade e resiliência capazes de prover a alta disponibilidade requerida pelo documento de inteligência.

A escolha da solução SIGAD (ECM, CMS, SGD, entre outros) é fundamental para a interoperabilidade, a solução SIGAD deve estar em conformidade com o padrão *Content Management Interoperability Services (CMIS)*. A despeito de ser proprietária ou de código aberto, esta solução deve tomar, como re-

ferência, os aspectos funcionais e de metadados previstos na *Especificação para Sistemas Informatizados de Gestão Arquivística de Documentos (e-Arq Brasil)*, assim como a conformidade com o padrão *Content Management Interoperability Services - CMIS*, desenvolvido pela *Organization for the Advancement of Structured Information Standards – OASIS*, a fim de promover a interoperabilidade documental entre os diferentes membros do Sistema Brasileiro de Inteligência, o que é imprescindível para a ABIN, enquanto órgão central desse sistema.

As diretrizes para a implantação das medidas propostas deve permear o planejamento desde o nível estratégico ao técnico-operacional, cobrindo toda arquitetura de governança do órgão, de forma a dar respaldo para estas iniciativas e a assegurar as condições apropriadas para o sucesso em nível institucional. As diretrizes propostas indicam que, em termos estratégicos, o Plano Estratégico de TIC – PETIC (ou PETI) deve comtemplar a Política de Gestão Arquivística de Documentos, a Política de Segurança da Informação – POSIN e a implantação de Programa de Gestão Arquivística de Documentos. No campo tático, o Plano Diretor de Tecnologia da Informação e Comunicações – PDTIC (ou PDTI) deve considerar ações para Adoção da Cadeia de Custódia Digital, para Salvaguarda dos Ativos Digitais e para Promoção de Curadoria Digital de Ativos Digitais. No contexto operacional, o Planejamento Orçamentário Anual (POA) deve comtemplar recursos financeiros que viabilizem aquisições e capacitações que se fizerem necessárias ao cumprimento desse planejamento.

O reconhecimento de repositório digital como confiável (RDC-Arq) não depende da solução *per si*, seja ela de código aberto ou proprietária, mas da obtenção da certificação preconizada pela *Audit and Certification of Trustworthy Digital Repositories* (ACTDR), prevista no próprio modelo RDC-Arq, onde devem ser atendidos requisitos relacionados à Infraestrutura Organizacional, à Gestão dos Objetos Digitais e à Infraestrutura e Segurança da Gestão de Riscos.

A análise da integração de SIGAD com repositório digital mostrou que o padrão CMIS, voltado para a interoperabilidade, encontrou grande aderência por parte das soluções de gerenciamento de conteúdo abertas e de mercado; o que não ocorreu com o modelo de OAIS, cujo o foco está na preservação e na autenticidade documental, o que sugere maior preocupação com a interoperabilidade em detrimento da preservação.

Apesar da Resolução nº 43 do CONARQ ter estabelecido, desde 2015, o RDC-Arq como referência para a Presunção de Autenticidade de Documentos Arquivísticos Digitais, a necessária integração entre SIGAD e repositório digital ainda carece de soluções técnicas. O desafio reside na construção de componentes de empacotamento SIP específicos para cada gerenciador de conteúdo que armazene em repositório digital padrão RDC-Arq, seja apenas para fins de preservação, seja no propósito de cadeia de custódia completa. Pois mesmo o cenário de integração parcial, proposto pela Orientação Técnica nº 3 do CONARQ, Novembro de 2015, ainda encontra obstáculos técnicos a serem superados.

Para futuros trabalhos, esta pesquisa serve de base para estudo de caso de implementação da CCDA no âmbito da ABIN e do SISBIN; aponta para a necessidade de estudos aprofundados acerca da aplicação e da certificação de repositórios digitais em nível nacional; e sugere aprofundamento no processo de *Audit and Certification of Trustworthy Digital Repositories (ACTDR)*, com estudos sobre as ferramentas empregadas e o detalhamento acerca do processo de acreditação de entidade candidata à certificadora.

Por fim, este trabalho lança um olhar sobre a abordagem *Long Term Digital Preservation as a Service* (LTDPaaS), advinda da arquitetura *cloud computing* e voltada para o armazenamento e preservação de material digital de longo prazo, sobre a qual uma investigação pode trazer contribuições importantes em termos de preservação digital para ambientes de computação em núvem.

REFERÊNCIAS

1 HORN, M. B.; STAKER., H. Ensino híbrido: uma inovação disruptiva? uma introdução à teoria dos híbridos. Christensen Institute, 2013. Disponível em: <https://www.pucpr.br/wp-content/uploads/2017/10/ensino-hibrido_uma-inovacao-disruptiva.pdf>.

2 DIGITAL/ME, S. de G. *Estratégia de Governança Digital da Administração Pública Federal*. 2020.

3 CIVIL, P. de R. C. Lei no 9.883, de 7 de dezembro de 1999. 1999. Disponível em: <http://www.planalto.gov.br/ccivil_03/leis/l9883.htm>.

4 CIVIL/PR, P. da R. C. Política nacional de inteligência. 2016. Disponível em: <https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2016/decreto/D8793.htm>.

5 RITTER, T.; PEDERSEN, C. L. Digitization capability and the digitalization of business models in business-to-business firms: Past, present, and future. *Industrial Marketing Management*, Elsevier, v. 86, n. November 2019, p. 180–190, 2020. ISSN 00198501. Disponível em: <https://doi.org/10.1016/j.indmarman.2019.11.019>.

6 SANTOS, H. M. dos; FLORES, D. Cadeia de custódia para documentos arquivísticos digitais. p. 117–132, 2016.

7 Conselho Nacional de Arquivos - CONARQ. Diretrizes para a presunção de autenticidade de documentos arquivísticos digitais. p. 10, 2012. Disponível em: <http://www.conarq.gov.br/images/publicacoes_textos/conarq_presuncao_autenticidade_completa.pdf>.

8 MACHADO, F. *Segurança da informação: princípios e controle de ameaças*. 2014.

9 SAYÃO, L. F.; SALES, L. F. Curadoria digital: um novo patamar para preservação de dados digitais de pesquisa. *Informacao e Sociedade*, v. 22, n. 3, p. 179–191, 2012. ISSN 01040146.

10 SOUSA, F. L. D. Elementos-chave da transformação digital que influenciam na curadoria digital: Uma revisão sistemática de literatura sob o método temac. Revista Ibérica de Sistemas e Tecnologias de Informação - RISTI, p. 463–476, 2021.

11 SIEBRA, S. d. A. Curadoria Digital: uma área em expansão. *Archeion Online*, v. 6, n. 2, p. 1–6, 2019.

12 CIVIL, P. de R. C. Lei nº 12.527, de 18 de novembro de 2011. 2011. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12527.htm>.

13 GSI/PR, G. de S. I. Relatório de gestão 2018. p. 116, 2018. Disponível em: <https://www.gov.br/gsi/pt-br/acesso-a-informacao/auditorias/relatorio_gestao_gsi_2018.pdf>.

14 SIDNEI, W. R. *Metodologia de Pesquisa para Ciência da Computação*. [S.l.]: LTC, 2021.

15 AZEVEDO, M. T. d. Transformação digital na indústria: indústria 4.0 e a rede de água inteligente no Brasil. *Doctoral dissertation, Universidade de São Paulo*, p. 177, 2017. Disponível em: <http://www.teses.usp.br/teses/disponiveis/3/3142/tde-28062017-110639/>.

16 KAPLAN B., I. D. P.-W. D.-W.-H. A. T. Relevant theory and informed practice: Looking forward from a 20 year perspective on is research. *Proceedings of the IFIP WG 8.2. Conference*, 2004.

17 TARAPANOFF, K.; Araújo Júnior, R. H. de; CORMIER, P. M. J. Sociedade da informação e inteligência em unidades de informação. *Ciência da Informação*, v. 29, n. 3, p. 91–100, 2000.

18 Lemos Ribeiro, A. C. M.; Dos Santos, C. D. Isso não é uma pirâmide: Revisando o modelo clássico de dado, informação, conhecimento e sabedoria. *Ciencia da Informacao*, v. 49, n. 2, p. 1, 2020. ISSN 01001965.

19 SOUSA, F. L. D. Abordagens conceituais de informação aplicadas às acepções de inteligência. Revista Brasileira de Inteligência, p. 91–98, 2013.

20 WERSIG G., . N. U. The phenomena of interest to information science. p. 127–140, 1975.

21 BIMFORT, M. T. A definition of intelligence. CIA HISTORICAL REVIEW PROGRAM, p. 1–4, 1975.

22 SIQUEIRA, J. C. A noção de documento digital: uma abordagem terminológica. Em Questão, p. 125–140, 2012.

23 CONARQ; CTDE. Modelo de Requisitos para Sistemas Informatizados de Gestão Arquivística de Documentos e-Arq Brasil. p. 1–223, 2020. Disponível em: <https://www.gov.br/conarq/pt-br/assuntos/noticias/conarq-abre-consulta-publica-visando-a-atualizacao-do-e-arq-brasil/EARQ_v2_2020_final.pdf>.

24 MENDES, M. A. D. S.; BAX, M. P. *BPM e ECM: Similaridades, diferenças e limites conceituais e tecnológicos*. [S.l.]: Pontificia Universidade Catolica de Campinas, 2018. 95–105 p.

25 OASIS, O. for the Advancement of S. I. S. ontent management interoperability services (cmis) tc. Disponível em: <https://www.oasis-open.org/committees/tc_home.php?wg_abbrev=cmis>.

26 PRAYUDI, Y.; SN, A. Digital Chain of Custody: State of The Art. *International Journal of Computer Applications*, v. 114, n. 5, p. 1–9, 2015.

27 ROCHA, C. L. Repositórios para a preservação de documentos arquivísticos digitais. *Acervo*, v. 28, n. 2 jul-dez, p. 180–191, 2015. Disponível em: <http://revista.arquivonacional.gov.br/index.php/revistaacervo/article/view/608>.

28 CCSDS, C. C. for S. D. S. audit and certification of trustworthy digital repositories. 2012.

29 CONARQ. Diretrizes para a implementação de repositórios arquivísticos digitais confiáveis - rdc-arq. *DOU*, v. 148, p. 148–162, 2015.

30 GAVA T. B. S., . F. D. Repositórios arquivísticos digitais confiáveis (rdc-arq) como plataforma de preservação digital em um ambiente de gestão arquivística. Informação Informação, p. 74–99, 2020.

31 LEE, C. A. Reference Model for an Open Archival Information System (OAIS). *Recommendation for Space Data System Practices*, n. June, p. 135, 2012. Disponível em: <https://public.ccsds.org/pubs/650x0m2.pdf%0Ahttp://public.ccsds.org/publications/archive/650x0m2.pdf>.

32 DILCIS, D. I. L. I. S. B. 2021. Disponível em: <https://dilcis.eu/specifications>.

33 SIEBRA S.A., B. V. M. M. Curadoria digital: Um termo interdisciplinar. Brapci:Anais XVII Encontro Nacional de Pesquisa em Ciência da Informação, p. 1–17, 2016. Disponível em: <https://repositorio.unb.br/handle/10482/17324>.

34 LUZ, C. d. S. Curadoria digital, custódia arquivística e preservação digital: relações possíveis. *Páginas ab Arquivos Bibliotecas*, v. 10, n. 10, p. 92–103, 2018. ISSN 08735670.

35 HIGGINS, S. The DCC Curation Lifecycle Model. *International Journal of Digital Curation*, v. 3, n. 1, p. 134–140, 2008.

36 NASCIMENTO, F. P. d.; SOUSA, F. L. L. Classificação da pesquisa. natureza, método ou abordagem metodológica, objetivos e procedimentos. Thesaurus, 2016. Disponível em: <http://franciscopaulo.com.br/arquivos/ClassificaÃğÃčodaPesquisa.pdf>.

37 MARIANO A. M., . R. M. S. Revisão da literatura: Apresentação de uma abordagem integradora. XXVI Congreso Internacional de la Academia Europea de Dirección y Economía de la Empresa (AEDEM), v. 26, 2017.

38 BARDIN, L. *Análise de conteúdo*. São Paulo, 1977. 1-225 p. Disponível em: <https://ia802902.us.archive.org/8/items/bardin-laurence-analise-de-conteudo/bardin-laurence-analise-de-conteudo.pdf>.

39 KABIR, M. An efficient low bit rate image watermarking and tamper detection for image authentication. SN Applied Sciences, 2021. Disponível em: <https://doi.org/10.1007/s42452-021-04387-w>.

40 FORMENTON D., . d. S. G. L. Digital preservation challenges, requirements, strategies and scientific output. RDBCI-Revista Digital de Biblioteconomia e Ciência da Informação, 2020. Disponível em: <https://www.researchgate.net/profile/Danilo-Formenton/publication/342163605_Digital_Preservation_challenges_requirements_strategies_and_scientific_output/links/5ee6477b92851ce9e7e39d74/Digital-Preservation-challenges-requirements-strategies-and-scientific-output.pdf>.

41 GIUSTI M. R., . V. G. L. D. Revision of different implementations for digital preservation: towards a methodological proposal for preserving and auditing ir reliability. RDBCI-Revista Digital de Biblioteconomia e Ciência da Informação, 2018. Disponível em: <https://digital.cic.gba.gob.ar/items/008914bc-fb7e-40a1-b0ac-a0b2b9a791ae>.

42 KROTH M. L., . F. D. Authenticity of digital records: analysis of a leave of absence process. Journal of Librarianship and Information Science, p. 67–79, 2018. Disponível em: <https://brapci.inf.br/index.php/res/v/69588>.

43 SANTOS H. M., . F. D. D. The digital archival document as research source. Perspectivas em Ciência da Informação, 2016.

44 SANTOS H. M. D., . F. D. D. Digital preservation policies for archival documents. Perspectivas em Ciência da Informação, p. 197–217, 2015. Disponível em: <https://brapci.inf.br/index.php/res/v/69588>.

45 SANTOS H. M. D., . F. D. D. As vulnerabilidades dos documentos digitais: Obsolescência tecnológica e ausência de políticas e práticas de preservação digital. Journal of Librarianship and Information Science, p. 67–79, 2015. Disponível em: <https://brapci.inf.br/index.php/res/v/69588>.

46 GIRALT O., V.-P. C. . P.-S. C. Seguridad de los documentos de archivo: estudio de caso del archivo del ayuntamiento de barcelona. Revista Profesional de la información, p. 202–205, 2011. Disponível em: <https://doi.org/10.3145/epi.2011.mar.11>.

47 FLORES, D. Curadoria digital arquivística e software livre. Researchgate, 2015. Disponível em: <https://www.researchgate.net/publication/308605353_Curadoria_Digital_Arquivistica_e_Software_Livre>.

48 BESSER, H. Longevidade Digital. *Revista Acervo do Arquivo Nacional*, v. 23, n. 2, p. 57–70, 2010.

49 DIGITAL/ME, S. de G. Padrões de interoperabilidade de governo eletrônico - eping. 2018. Disponível em: <https://eping.governoeletronico.gov.br/>.

50 SANTOS, H. M. dos; FLORES, D. Preservação sistêmica para repositórios arquivísticos TT - Systemic preservation for archival repositories TT - Preservacion sistémica para repositorios de archivo. *RECIIS (Online)*, v. 14, n. 3, p. 764–781, 2020. ISSN 1981-6278. Disponível em: <https://www.reciis.icict.fiocruz.br/index.php/reciis/article/view/2089/2386>.

51 LOC, T. L. of C. Premis data dictionary for preservation metadata, version 3.0. 2021. Disponível em: <http://www.loc.gov/standards/premis/v3/>.

52 ICA, I. C. on A. Isad(g): General international standard archival description. 09 2021. Disponível em: <https://www.ica.org/en/isadg-general-international-standard-archival-description-second-edition>.

53 CONARQ, C. N. de A. Norma brasileira de descrição arquivística. 2021. Disponível em: <https://www.gov.br/arquivonacional/pt-br/canais_atendimento/imprensa/copy_of_noticias/serie-publicacoes-do-conarq-norma-brasileira-de-descricao-arquivistica>.

54 OAI, O. A. I. The open archives initiative protocol for metadata harvesting. 2015. Disponível em: <http://www.openarchives.org/OAI/openarchivesprotocol.html>.

55 LOC, T. L. of C. Metadata encoding transmission standard - mets. 2021. Disponível em: <http://www.loc.gov/standards/mets/>.

56 LOC, T. L. of C. Encoded archival description. 2021. Disponível em: <https://www.loc.gov/ead/>.

57 ISO, I. O. for S. Open archival information system - oais reference model. 2003. Disponível em: <https://www.iso.org/standard/24683.html>.

58 GNU. Gpl - general public license. 2022. Disponível em: <https://www.gnu.org/licenses/agpl-3.0.html>.

59 ARTEFACTUAL.INC. Archivematica. 2022. Disponível em: <https://www.archivematica.org>.

60 SOLUTIONS keep. Roda. 2022. Disponível em: <https://roda-community.org>.

61 SHUKLA AKHANDANAND, e. a. Open source software for digital archiving: A comparative study on roda and archivematica. NCDS - Digital Scholarship, 2020.

62 SANTOS, H. M. Manual para auditoria de repositórios arquivísticos digitais confiáveis. ResearchGate, Março 2018. Disponível em: <https://www.researchgate.net/profile/Henrique-Santos/publication/347438660_Manual_para_Auditoria_de_Repositorios_Arquivisticos_Digitais_Confiaveis/links/5fdb8fdc92851c13fe942dec/Manual-para-Auditoria-de-Repositorios-Arquivisticos-Digitais-Confiaveis.pdf>.

63 WALDT, D. Addressing contemporary requirements for content integration. The Gilbane Group, 2009. Disponível em: <https://gilbane.com/wp-content/uploads/2020/01/Gilbane-Beacon-CMIS.pdf>.

64 ARQUIVOS-CONARQ-CTDE, C. N. de. Cenários de uso de rdc-arq em conjunto com o sigad. DOU, 2022. Disponível em: <https://www.gov.br/conarq/pt-br/assuntos/camaras-tecnicas-setoriais-inativas/camara-tecnica-de-documentos-eletronicos-ctde/Orientacao_tecnica_3.pdf>.

65 SHINTAKU, M.; BRAGA, T. E. N.; OLIVEIRA, A. d. F. Hipátia - uma ferramenta livre no apoio à preservação digital. *Revista Brasileira de Preservação Digital*, v. 2, 2021.

66 MCSA. Comparison between cloud computing gridcomputing clustercomputing and virtualization. International Journal of Modern Computer Science and Applications - IMCSA, p. 42–47, 2015.

67 ABDALLA, P. A. Advantages to disadvantages of cloud computing for small-sized business. IEEE, 2019.

68 FRANKS, P. C. Government use of cloud-based long term digital preservation as a service: An exploratory study. *2015 Digital Heritage International Congress, Digital Heritage 2015*, IEEE, p. 371–374, 2015.

69 GSI/PR, G. de S. I. Instrução normativa nº 1, de 27 de maio de 2020. Imprensa Nacional, 2020. Disponível em: <https://www.in.gov.br/en/web/dou/-/instrucao-normativa-n-1-de-27-de-maio-de-2020-258915215>.