

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
FACULDADE DE BIBLIOTECONOMIA E COMUNICAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS DA INFORMAÇÃO

FELIPE PIRES DA SILVA

**GESTÃO DE DADOS EM ORGANIZAÇÕES:  
UMA ANÁLISE DA APLICAÇÃO DA PRESERVAÇÃO E CURADORIA DIGITAL  
NO CORPO DE CONHECIMENTO DO DMBOK**

Porto Alegre

2023

FELIPE PIRES DA SILVA

**GESTÃO DE DADOS EM ORGANIZAÇÕES:  
UMA ANÁLISE DA APLICAÇÃO DA PRESERVAÇÃO E CURADORIA DIGITAL  
NO CORPO DE CONHECIMENTO DO DMBOK**

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação da Universidade Federal do Rio Grande do Sul como requisito parcial para obtenção do título de Mestre em Ciência da Informação.

Orientador: Rafael Port da Rocha

Porto Alegre

2023

### CIP - Catalogação na Publicação

da Silva, Felipe Pires  
GESTÃO DE DADOS EM ORGANIZAÇÕES: UMA ANÁLISE DA  
APLICAÇÃO DA PRESERVAÇÃO E CURADORIA DIGITAL NO CORPO  
DE CONHECIMENTO DO DMBOK / Felipe Pires da Silva. --  
2023.  
127 f.  
Orientador: Rafael Port da Rocha.

Dissertação (Mestrado) -- Universidade Federal do  
Rio Grande do Sul, Faculdade de Biblioteconomia e  
Comunicação, Programa de Pós-Graduação em Ciência da  
Informação, Porto Alegre, BR-RS, 2023.

1. DMBOK. 2. gestão de dados. 3. preservação  
digital. 4. curadoria digital. I. da Rocha, Rafael  
Port, orient. II. Título.

## DEDICATÓRIA

Dedico esta pesquisa para meus pais, em especial minha mãe: guerreira, símbolo de resistência e resiliência frente aos desafios da vida. Minha fonte de inspiração.

## **AGRADECIMENTOS**

Agradeço primeiramente a Deus, meu maior companheiro desta jornada em dias e noites de muitas reflexões, porém, sempre na certeza de um desfecho vencedor.

Agradeço também pela parceria, generosidade e compreensão do meu orientador Professor Rafael Port da Rocha. Sou extremamente grato pela oportunidade de aprender com seu conhecimento, colaboração, paixão e dedicação diante de sua nobre missão de educar e contribuir para a ciência. Um especial agradecimento à banca de qualificação: Prof. Renê Faustino da UFRGS, Prof. Cláudio Ribeiro da UNIRIO e Profa. Sônia Caregnato da UFRGS, os quais tiveram a generosidade de contribuir para a evolução desta pesquisa. Aos professores e colegas do PPGCIN da UFRGS, a quem tive a honra de interagir durante o período de mestrado, sempre disponíveis e dispostos a compartilhar conhecimento. Aos meus familiares, em especial minha irmã Patrícia, outra guerreira, que me inspirou a não desistir em meio as adversidades desta caminhada. Ao meu pai, exemplo de sabedoria de vida, meu alicerce e meu guru nas tomadas de decisões da vida. No meio desta jornada encontrei pessoas humildes e cheias de bondade, que no final do dia, possuem a mais pura e verdadeira torcida para minha vitória. Pessoas que me ajudaram e empregaram esforços para que eu cumprisse a missão, à essas pessoas, meu profundo agradecimento.

## RESUMO

A gestão de dados tem se consolidado como um tema central nas organizações contemporâneas, especialmente diante da crescente produção de informações digitais. A integridade, disponibilidade e confidencialidade desses dados são imperativas, exigindo abordagens sistemáticas para sua gestão. O DMBOK (Data Management Body of Knowledge), ou Corpo de Conhecimento em Gestão de Dados do DAMA, surge como um modelo de referência para o gerenciamento de dados em uma organização, oferecendo diretrizes abrangentes com processos consistentes. Contudo, com a evolução tecnológica e a necessidade de preservação de ativos digitais a longo prazo, emerge a relevância da preservação digital e da curadoria digital. Estas práticas, voltadas para a manutenção, valorização e contextualização da informação, são fundamentadas e encontradas em modelos como OAIS, TRAC, Ciclo de Vida de Curadoria DCC, entre outros. Esta pesquisa analisa a presença de características de preservação digital e de curadoria digital nas áreas do Corpo de Conhecimento do DMBOK, utilizando os modelos OAIS e Ciclo de Vida de Curadoria DCC como referências. Relaciona os conceitos, objetivos e atividades trazidos pelos modelos DMBOK, OAIS e Ciclo de Vida de Curadoria DCC e analisa a existência de relação de cada área de conhecimento do DMBOK com os aspectos de preservação digital e curadoria digital. A conclusão aponta o DMBOK como um recurso valioso na gestão de dados organizacionais, cuja relevância é potencializada quando alinhada às práticas de preservação e curadoria digital. Propõe-se, ainda, a incorporação de características da preservação e curadoria digital ao corpo de conhecimento do DMBOK, fundamentado nos modelos OAIS e Ciclo de Vida de Curadoria DCC.

**Palavras-chave:** DMBOK, gestão de dados, preservação digital, curadoria digital

## ABSTRACT

Data management has become a central theme in contemporary organizations, especially in light of the increasing production of digital information. The integrity, availability, and confidentiality of this data are imperative, necessitating systematic approaches to their management. The DMBOK (Data Management Body of Knowledge) from DAMA emerges as a reference model for data management within an organization, offering comprehensive guidelines with consistent processes. However, with technological evolution and the need for long-term preservation of digital assets, the significance of digital preservation and digital curation becomes evident. These practices, aimed at maintaining, enhancing, and contextualizing information, are grounded in models such as OAIS, TRAC, DCC Curation Lifecycle, among others. This research examines the presence of digital preservation and curation characteristics in the areas of the DMBOK Body of Knowledge, using the OAIS and DCC Curation Lifecycle models as references. It relates the concepts, objectives, and activities brought by the DMBOK, OAIS, and DCC Curation Lifecycle models and analyzes the relationship of each area of knowledge in DMBOK with aspects of digital preservation and curation. The conclusion highlights DMBOK as a valuable resource in organizational data management, whose significance is amplified when aligned with digital preservation and curation practices. It is further proposed to incorporate features of digital preservation and curation into the DMBOK body of knowledge, grounded in the OAIS and DCC Curation Lifecycle models.

**Keywords:** DMBOK, data management, digital preservation, digital curation

## LISTA DE FIGURAS

Figura 1 - Modelo de ambiente de um OAIS .....	31
Figura 2 - Dados externos do Arquivo OAIS .....	34
Figura 3 - Entidades Funcionais do OAIS .....	36
Figura 4 - Modelo do Ciclo de Vida de Curadoria DCC .....	41
Figura 5 -Amsterdam Information Model .....	50
Figura 6 - Modelo de Alinhamento Estratégico.....	51
Figura 7 - Roda do DAMA Primeira Edição do DMBOK .....	56
Figura 8 - Roda do DAMA Segunda Edição do DMBOK .....	56
Figura 9 - Representação Hexágono dos Fatores Ambientais .....	59
Figura 10 - Diagrama de Contexto da Área de Conhecimento .....	60
Figura 11 - Diagrama de Contexto da Área de Governança de Dados.....	63
Figura 12 - Modelo gráfico de captação dos insumos da pesquisa .....	69



## LISTA DE QUADROS

Quadro 1 - Relação das atividades com Ciclo de Vida de DCC, Entidades e Processos de OAIS e Critérios de ACTDR.....	29
Quadro 1 - Relação das atividades com Ciclo de Vida de DCC, Entidades e Processos de OAIS e Critérios de ACTDR.....	30
Quadro 2 - Conceitos e objetivos das entidades funcionais do OAIS .....	39
Quadro 3 - Conceitos e objetivos das ações do Ciclo de Vida de Curadoria DCC ...	41
Quadro 4 - Componentes do Diagrama de Contexto .....	62
Quadro 5 - Coleta e questões de Objetivos de DMBOK, OAIS e CVCD/DCC .....	65
Quadro 6 - Coleta e questões de Conceitos de DMBOK, OAIS e CVCD/DCC .....	65
Quadro 7 - Coleta e questões de análise da área quanto aos Objetivos .....	66
Quadro 8 - Coleta e questões de análise da área quanto as Atividades .....	67
Quadro 9 - Objetivos de DMBOK, OAIS e CVCD/DCC .....	71
Quadro 10 - Conceitos de DMBOK, OAIS e CVCD/DCC .....	73
Quadro 11 - Entidades Funcionais de OAIS e Atividades de CVCD/DCC relacionadas a Governança de Dados pelos Objetivos .....	75
Quadro 12 - Entidades Funcionais de OAIS e Atividades de CVCD/DCC relacionadas a Governanca de Dados pelas Atividades.....	75
Quadro 13 - Entidades Funcionais de OAIS e Atividades de CVCD/DCC relacionadas a área de Armazenamento e Operação pelos Objetivos .....	78
Quadro 14 - Entidades Funcionais de OAIS e Atividades de CVCD/DCC relacionadas a área de Armazenamento e Operação por suas Atividades.....	79
Quadro 15 - Entidades Funcionais de OAIS e Atividades de CVCD/DCC relacionadas a área de Integração e Interoperabilidade pelos Objetivos.....	82
Quadro 16 - Entidades Funcionais de OAIS e Atividades de CVCD/DCC relacionadas a área de Integração e Interoperabilidade por suas Atividades .....	82
Quadro 17 - Entidades Funcionais de OAIS e Atividades de CVCD/DCC relacionadas a área de Segurança pelos Objetivos .....	85
Quadro 18 - Entidades Funcionais de OAIS e Atividades de CVCD/DCC relacionadas a área de Segurança por suas Atividades.....	86
Quadro 19 - Entidades Funcionais de OAIS e Atividades de CVCD/DCC relacionadas a área de Dados Mestre e de Referência pelos Objetivos.....	89
Quadro 20 - Entidades Funcionais de OAIS e Atividades de CVCD/DCC relacionadas a área de Dados Mestre e de Referência por suas Atividades .....	89
Quadro 21 - Entidades Funcionais de OAIS e Atividades de CVCD/DCC relacionadas a área de Arquitetura de Dados pelos Objetivos .....	91
Quadro 22 - Entidades Funcionais de OAIS e Atividades de CVCD/DCC relacionadas a área de Arquitetura de Dados por suas Atividades .....	92
Quadro 23 - Entidades Funcionais de OAIS e Atividades de CVCD/DCC relacionadas a área de Metadados pelos Objetivos .....	94
Quadro 24 - Entidades Funcionais de OAIS e Atividades de CVCD/DCC relacionadas a área de Metadados por suas Atividades .....	94
Quadro 25 - Entidades Funcionais de OAIS e Atividades de CVCD/DCC relacionadas a área de Modelagem e Design pelos Objetivos .....	97

Quadro 26 - Entidades Funcionais de OAIS e Atividades de CVCD/DCC relacionadas a área de Modelagem e Design por suas Atividades.....	97
Quadro 27 - Entidades Funcionais de OAIS e Atividades de CVCD/DCC relacionadas a área de Data Warehousing & Business Intelligence pelos Objetivos 99	
Quadro 28 - Entidades Funcionais de OAIS e Atividades de CVCD/DCC relacionadas a área de Data Warehousing & Business Intelligence por suas Atividades .....	100
Quadro 29 - Entidades Funcionais de OAIS e Atividades de CVCD/DCC relacionadas a área de Qualidade de Dados pelos Objetivos .....	102
Quadro 30 - Entidades Funcionais de OAIS e Atividades de CVCD/DCC relacionadas a área de Qualidade de Dados por suas Atividades.....	103
Quadro 31 - Entidades Funcionais de OAIS e Atividades de CVCD/DCC relacionadas a área de Gestão de Documentos e Conteúdo pelos Objetivos .....	105
Quadro 32 - Entidades Funcionais de OAIS e Atividades de CVCD/DCC relacionadas a área de Gestão de Documentos e Conteúdo por suas Atividades..	106

## LISTA DE ABREVIATURAS E SIGLAS

ACTDR	Audit and Certification of Trustworthy Digital Repositories
AIM	Amsterdam Information Management Model
AIP	Archival Information Packages
BABOK	Business Analysis Body of Knowledge
CCSDS	Consultative Committee for Space Data System
COBIT	Control Objectives for Information and Related Technologies
CRL	Center for Research Libraries
CVCD/DCC	Ciclo de Vida de Curadoria do DCC
DAMA	Data Management Association
DCC	Digital Curation Centre
DGF	Data Governance Framework
DIP	Dissemination Information Package
DMBOK	Data Management Body of Knowledge
DPE	Digital Preservation Europe
DRAMBORA	Digital Repository Audit Method Based on Risk Assessment
GDPR	General Data Protection Regulation
ISACA	Information Systems Audit and Control Association
ISKO	International Society for Knowledge Organization
ISO	International Organization for Standardization
ITBOK	Information Technology Body of Knowledge
ITIL	IT Infrastructure Library
NDSA	National Digital Stewardship Alliance
OAIS	Open Archival Information
OCLC	Online Computer Library Center
PMBOK	Project Management Body of Knowledge
PMI	Project Management Institute
SAM	Strategic Alignment Model
SIP	Submission Information Package
TRAC	Trustworthy Repositories Audit & Certification
XML	Extensible Markup Language

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO .....</b>	<b>10</b>
1.1	Questão de pesquisa .....	13
1.2	Objetivos .....	13
1.3	Justificativa .....	14
1.4	Estrutura da Dissertação .....	16
<b>2</b>	<b>REVISÃO DE LITERATURA.....</b>	<b>17</b>
2.1	Dado, Informação e Conhecimento.....	17
2.2	Dados como um ativo corporativo .....	20
2.3	Preservação Digital e Curadoria Digital.....	21
2.4	Modelo Open Archival Information System (OAIS).....	30
2.3.1	O ambiente de um OAIS.....	31
2.3.2	Os pacotes de Informação de um OAIS.....	32
2.3.3	Interações entre Produtor e Consumidor .....	34
2.3.4	Responsabilidades de um OAIS .....	35
2.3.5	Componentes Funcionais de um OAIS.....	36
2.5	Modelo do Ciclo de Vida de Curadoria do DCC .....	40
2.6	Gestão de Dados em Organizações .....	45
2.7	Referências Técnicas na Gestão de Dados .....	48
2.8	Corpo de Conhecimento em Gestão de Dados DMBOK .....	53
<b>3</b>	<b>METODOLOGIA .....</b>	<b>64</b>
3.1	Planejamento da Pesquisa.....	64
3.1.1	Percurso Metodológico do Objetivo A .....	65
3.1.2	Percurso Metodológico do Objetivo B .....	66
3.1.3	Percurso Metodológico do Objetivo C .....	67
3.2	Estratégia de Busca e Organização das Fontes Utilizadas .....	67
<b>4</b>	<b>ANÁLISE DOS DADOS .....</b>	<b>70</b>
4.1	DMBOK e os Modelos OAIS e CVCD/DCC.....	70
4.2	As Áreas de Conhecimento do DMBOK e Preservação e Curadoria Digital .....	74
4.2.1	Governança de Dados .....	74
4.2.2	Armazenamento e Operação .....	78

4.2.3	Integração e Interoperabilidade .....	81
4.2.4	Segurança .....	85
4.2.5	Dados Mestre e de Referência .....	88
4.2.6	Arquitetura de Dados .....	90
4.2.7	Metadados .....	93
4.2.8	Modelagem e Design .....	96
4.2.9	Data Warehousing & Business Intelligence.....	99
4.2.10	Qualidade de Dados .....	101
4.2.11	Gestão de Documentos e Conteúdo .....	105
4.3	Síntese da análise dos dados .....	107
4.3.1	Relação entre objetivos do OAIS e CVCD/DCC.....	107
4.3.2	Relação entre os principais conceitos abordados pelo DMBOK, OAIS e CVCD/DCC.....	108
4.3.3	Áreas de Conhecimento do DMBOK e sua relação com OAIS e CVCD/DCC .....	109
4.4	Proposições sobre a Preservação e Curadoria Digital em DMBOK.....	109
<b>5</b>	<b>CONSIDERAÇÕES FINAIS .....</b>	<b>113</b>
	<b>REFERÊNCIAS BIBLIOGRÁFICAS .....</b>	<b>116</b>
	<b>APÊNDICE A – MODELAGEM DO REPOSITÓRIO DE INFORMAÇÃO DAS BUSCAS REALIZADAS .....</b>	<b>122</b>
	<b>APÊNDICE B – EXEMPLO DO REPOSITÓRIO DE INFORMAÇÕES PREENCHIDO .....</b>	<b>123</b>

## 1 INTRODUÇÃO

A gestão de dados tem sido um tema amplamente debatido em organizações de todos os tamanhos e setores. Com a crescente quantidade de informações digitais geradas diariamente, nunca foi tão grande a importância de garantir a integridade, disponibilidade e confidencialidade sobre os dados.

Para desenvolver uma gestão de dados em organizações, torna-se necessário adotar uma abordagem sistemática. Isso envolve a implementação de políticas e práticas de gerenciamento de dados, bem como a utilização de tecnologias de armazenamento, proteção e recuperação. Devido ao avanço das tecnologias de análise de dados e inteligência artificial, a gestão de dados tornou-se ainda mais complexa e desafiadora. As organizações devem estar preparadas para lidar com grandes volumes de dados, seus vários tipos e garantir sua qualidade. Ademais, a transparência no uso de dados se tornou uma necessidade entre os consumidores, que, por sua vez, têm questionado cada vez mais as empresas a respeito da utilização de suas informações pessoais.

Ao longo dos últimos anos, acompanhou-se a mudança de compreensão sobre os dados. Um dos eventos que marcou essa mudança foi a crise financeira global de 2008. Segundo Peter Aiken e Juanita Walton (2013), a falta de capacidade de agregar, comparar e entender os dados antes dos acontecimentos impactou os tomadores de decisão, dificultando a reação e resposta aos acontecimentos daquela ocasião. Observar os dados sob outra perspectiva contribui para os resultados de negócio. Peter Aiken e Juanita Walton (2013, p.18) ressaltam que “os dados representam o único ativo estratégico não esgotável, não degradável e durável de uma organização”. Com base nessas características, consolida-se a visão do tratamento de dados como um ativo efetivo dentro de uma organização, tornando-se necessário seu devido tratamento para a manutenção a longo prazo desses ativos valiosos.

Em virtude da velocidade da evolução das tecnologias digitais, a gestão de dados tornou-se uma opção para as organizações, que precisam garantir a integridade, autenticidade e acessibilidade de seus ativos digitais. Uma gestão efetiva sobre os dados pode trazer benefícios significativos, como a melhoria na

tomada de decisão, aumento da eficiência operacional, criação de vantagens competitivas e a própria preservação da memória institucional. Métodos de gestão de dados têm sido debatidos e implementados por diversas companhias, e sua literatura divide-se majoritariamente entre áreas da tecnologia da informação e da administração, em função do contexto do uso massivo de ferramentas de tecnologias de informação e das formas de gerenciamento organizacionais.

O DMBOK (Data Management Body of Knowledge ou Corpo de Conhecimento em Gestão de Dados, tradução livre) é um modelo que fornece um conjunto de diretrizes e práticas recomendadas para o gerenciamento de dados em organizações. Ele é dividido em onze áreas de conhecimento que abrangem desde a estratégia de gerenciamento de dados até sua implantação técnica e operacional (DMBOK 2 ed., 2017). O DMBOK é um modelo amplo e abrangente, destacando-se de outros modelos mais específicos que tratam apenas de um tema ou pequenos grupos de disciplinas do contexto de dados. Por esse motivo, o DMBOK será o modelo de referência a ser utilizado nesta pesquisa.

As informações digitais transformaram-se em ativos valiosos e críticos para o sucesso dos negócios nas organizações, não somente a curto prazo, mas também para uso futuro. O imediatismo das grandes organizações em busca do pioneirismo gera um ponto de falha na preservação de dados. Em virtude da clara necessidade do uso de modelos de gestão de dados em organizações, torna-se cada vez mais relevante o tratamento da preservação de ativos digitais a longo prazo para benefício da própria organização. Preservar esses ativos requer a implementação de práticas de preservação digital que habilitam o controle e o monitoramento das informações focadas na prevenção e mitigação de riscos que possam comprometer a integridade destes ativos como, por exemplo, riscos inerentes à obsolescência da tecnologia de software e hardware, riscos decorrentes da carência de recursos tecnológicos, humanos, financeiros e de gestão.

A preservação digital envolve a adoção de padrões e normas que garantem o uso dos dados a longo prazo e preparam organizações para lidar com desafios como a obsolescência de tecnologias, formatos e tipos de documentos (OLIVER, HARVEY, 2016). Esse dinamismo e diversidade no qual os dados estão imersos cria uma abordagem mais ampla, na qual surge a necessidade das técnicas de

preservação digital, onde inúmeros estudos e normas já são conhecidos e aplicados em instituições de preservação juntamente com suas ferramentas de apoio.

A preservação digital caminha junto com processos que dão ênfase à manutenção da informação e atividades que criticam o conteúdo e a forma que serão disponibilizados para uma comunidade específica. Esses processos estão presentes na curadoria digital, que possui ações de coleta, organização, interpretação e compartilhamento de conteúdo digital de alta relevância para um público específico. Essas ações são voltadas para a valorização da informação, a fim de que o conteúdo seja facilmente acessível, compreensível e significativo para os usuários. A curadoria digital, portanto, se propõe a agregar valor e fornecer contexto ao conteúdo, agindo como um filtro que ajuda os consumidores de informações a encontrar e entender a mensagem de que necessitam em meio à avalanche de dados disponíveis na Internet. Existem vários modelos de práticas de preservação e curadoria digital que são amplamente reconhecidos e citados na literatura científica, como: OAI (Open Archival Information System), TRAC (Trustworthy Repositories Audit & Certification), DCC Curation Lifecycle Model, ISO 16363 e NDSA (National Digital Stewardship Alliance). Esses modelos são abordagens que contêm metodologias que se concentram na preservação e curadoria digital em perspectivas de repositórios digitais, ciclo de vida, dentre outras estruturas conceituais de preservação de arquivos digitais a longo prazo.

O escopo desta pesquisa concentra-se na análise do modelo de gestão de dados do DMBOK na busca por características e técnicas de preservação e curadoria digital que são encontradas em modelos amplamente citados e conhecidos na literatura científica, principalmente oriundas do campo da ciência da informação, para que possam incorporar as áreas de conhecimento de gestão de dados do DMBOK, potencializando assim a cultura da preservação de dados a longo prazo. A busca por relações da preservação digital e da curadoria digital frente ao modelo de gestão de dados do DMBOK expõe e cria oportunidades de geração de modelos avançados que aproximam a pauta da preservação de dados a longo prazo no âmbito da gestão de dados. A aproximação entre os modelos permite o diálogo entre a ciência da informação, tecnologia da informação e administração, unindo pontos fortes de cada campo científico com o propósito de potencializar o uso de



dados nas organizações e transformar dados em ativos de alto valor.

### **1.1 Questão de pesquisa**

As inúmeras possibilidades de investigação na temática de dados desperta ideias e reflexões de caráter atual, prático e teórico, estimulando propostas para este campo de atuação e convidando profissionais de outras áreas para colaboração e convergência de propósitos. Observou-se a necessidade de preservação de dados a longo prazo nas organizações, diante do uso de dados correntes para obtenção de informações imediatas e na ausência de práticas de curadoria após o uso desses dados. Durante as investigações preliminares em busca de literatura científica no campo da gestão de dados e na observação das práticas das organizações no tratamento dos dados e informações, formulou-se a seguinte questão de investigação:

- Diante da necessidade de preservação a longo prazo de dados em organizações, a gestão de dados apresentada pelo DMBOK aborda aspectos de preservação e curadoria digital para apoiar as organizações?

Com esta pesquisa, espera-se contribuir para aos estudos da gestão de dados através das sugestões feitas para a incorporação de práticas de preservação digital e curadoria digital no corpo de conhecimento DMBOK. Almeja-se que tais sugestões possam potencializar a utilização dos dados, transformando-os em ativos de valor, garantindo sua durabilidade a longo prazo, respeitando suas características intrínsecas e mantendo a necessária monitorização de seu contexto diante das alterações nos ambientes digitais.

### **1.2 Objetivos**

Frente à problemática apresentada, definiu-se um objetivo geral para guiar a

trajetória até a obtenção dos resultados esperados. Para percorrer tal trajetória, estabeleceram-se objetivos específicos, que consistem em etapas de construção que produzem insumos para atingir o objetivo geral da pesquisa.

O objetivo geral da pesquisa é:

- Analisar a presença de características de preservação digital e de curadoria digital nas áreas do corpo de conhecimento DMBOK.

Quanto aos objetivos específicos desta pesquisa, definiram-se os seguintes:

- a) Relacionar as áreas do DMBOK com os modelos de preservação e curadoria digital OAIS e CVCD/DCC;
- b) Analisar a relação de cada área de conhecimento do DMBOK com os modelos preservação digital e curadoria digital OAIS e CVCD/DCC;
- c) Ampliar a visão da gestão de dados através de sugestões de incorporação da preservação e curadoria digital no DMBOK.

### **1.3 Justificativa**

O avanço massivo da tecnologia possibilitou progressos em diversas áreas do conhecimento, alterou hábitos de consumo de produtos e serviços e propiciou a evolução das organizações frente às demandas de um mercado ávido por resultados. Por outro lado, a digitalização crescente de informações resultou em uma maior dependência de tecnologias, criando obstáculos para a gestão de dados em organizações de vários setores. Novos formatos de entrega de informações surgiram e seu valor foi reconhecido pelas organizações, o que fez com que os dados se transformassem numa espécie de moeda de troca através dos meios de comunicação digitais e da rede mundial de computadores. Decisões são tomadas com base em análises de dados, em modelos estatísticos aplicados em informações passadas e em modelos que podem prever eventos futuros. Para as organizações, a utilização de dados é reconhecida como ferramenta crucial para sua sobrevivência.

A preservação digital é frequentemente negligenciada nas organizações, o

que pode conduzir à perda de informações importantes e valiosas a longo prazo. Ainda existe uma falta de conhecimento sobre os benefícios da incorporação da preservação digital em todo o contexto organizacional e esta é frequentemente confundida com processos técnicos de tecnologia da informação, como as políticas de proteção de dados, comumente denominadas de backup. As ações para trabalhar com dados desde a sua criação até o seu descarte muitas vezes são deficientes, e a implementação de processos de curadoria digital, por exemplo, é desconhecida por uma grande parcela das equipes que trabalham com dados em organizações.

Nesse contexto, justifica-se esta pesquisa, que investigou a existência de relações dos aspectos da preservação e curadoria digital no contexto da gestão de dados, com o objetivo de contribuir para a evolução das práticas de gerenciamento de dados e promover a cultura da preservação a longo prazo nas organizações. É esperado que a pesquisa possa facilitar a criação de soluções inovadoras, potencializando tecnologias, conservando investimentos e otimizando processos de negócio. Ademais, pode-se prover através da pesquisa subsídios para o desenvolvimento de políticas e diretrizes para a gestão de dados digitais nas organizações, contribuindo para a sua sustentabilidade e sucesso na preservação a longo prazo.

Considerando-se a importância da preservação e curadoria digital na gestão de dados, o modelo DMBOK é apresentado como uma ferramenta valiosa para embasar a pesquisa. Sua estrutura bem definida e abrangente oferece possibilidades para a exploração e incorporação dos diversos aspectos da preservação e curadoria digital. Por exemplo, características de preservação digital podem ser encontradas em áreas de conhecimento do DMBOK como governança de dados, gestão de metadados, gerenciamento de informações, qualidade de dados, segurança de dados, entre outros. A curadoria digital, por sua vez, pode ser evidenciada em atividades e ações das áreas de conhecimento do DMBOK, destacando processos de sustentação dos objetivos de cada área do corpo de conhecimento.

Adicionalmente, é oferecida pelo DMBOK uma perspectiva holística da gestão de dados, onde as diferentes fases do ciclo de vida dos dados, desde a criação até a

destruição, são consideradas. Dado o interesse na relação com a preservação e curadoria digital, essa perspectiva se torna relevante, visto que a análise da literatura permite a identificação dos pontos críticos da gestão de dados e a definição das estratégias mais adequadas para a sua preservação.

#### **1.4 Estrutura da Dissertação**

A dissertação encontra-se estruturada em capítulos e seções nos quais são apresentados a introdução ao tema, o referencial teórico que serviu de âncora para os conceitos abordados na pesquisa, o desenvolvimento e o percurso metodológico da pesquisa e os resultados alcançados. A introdução realizada focou-se na exposição dos desafios da gestão de dados nos dias atuais e na necessidade de refletir sobre a preservação de informações de forma permanente, considerando a aplicação de um ciclo de vida de curadoria digital dentro dos conceitos da preservação digital. Nela, também foram apresentados o objetivo geral e específicos da pesquisa bem como sua justificativa de execução. O segundo capítulo apresenta o referencial teórico sobre a temática abordada, contextualizando e alicerçando os principais conceitos sobre as abordagens que serão exploradas no campo da gestão de dados, preservação digital e curadoria digital. Nas seções do segundo capítulo, são apresentados os modelos que serão objetos da pesquisa: o corpo de conhecimento do DMBOK, o modelo de referência OAIS e o modelo de ciclo de vida de curadoria digital do DCC.

O terceiro capítulo apresenta o percurso metodológico que foi realizado para alcançar os objetivos almejados na pesquisa e expõe a estrutura base da coleta e análise de dados. O quarto capítulo focou-se na exposição dos dados coletados e nas respectivas análises de acordo com seus objetivos específicos expostos no primeiro capítulo. O último capítulo apresenta as considerações finais da pesquisa, demonstrando sua aplicabilidade em pesquisas futuras e refletindo possíveis impactos positivos no desenvolvimento de modelos avançados de gestão de dados com abordagens de preservação e curadoria digital.

## 2 REVISÃO DE LITERATURA

Nas seções a seguir, a revisão da literatura é apresentada, abordando inicialmente os conceitos sobre dado, informação e conhecimento a fim de solidificar a compreensão dos conceitos de preservação digital, curadoria digital e gestão de dados. É explorado referencial teórico oriundos dos campos da Ciência da Informação, Tecnologia da Informação e Administração.

### 2.1 Dado, Informação e Conhecimento

Dados são frequentemente entendidos como representações simbólicas de fatos, observações ou eventos, os quais podem ser registrados, processados e analisados para formar informações e gerar conhecimentos (BORGMAN, 2015; FLORIDI, 2019). Esta perspectiva enfatiza a importância dos dados como matéria-prima para a criação de informações e conhecimento, bem como o papel dos profissionais da informação na organização, preservação e disseminação destes dados (BUCKLAND, 2017). Ou seja, a informação é vista como o resultado da interpretação e contextualização dos dados, transformando-os em algo útil e aplicável. No campo da administração, a definição de dados é frequentemente relacionada à sua utilização no contexto de tomada de decisões e gestão de negócios. Neste sentido, os dados são vistos como recursos valiosos que, quando devidamente analisados e interpretados, podem fornecer insights e orientações para a melhoria da eficiência operacional e a realização de objetivos estratégicos (KOKINA & DAVENPORT, 2017; KITCHIN, 2014). A busca por uma definição unívoca entre os autores sobre dados, informação e conhecimento é desafiadora. Os autores Peter Checkland e Sue Holwell (2011) posicionam os dados como fatos e apresentam uma camada chamada de Capta, que seria responsável por selecionar alguns destes dados com maior atenção ou até mesmo criar uma nova categoria para eles com o objetivo de colocá-los em contexto, gerando assim informação e, na sequência, conhecimento. Este é apenas um exemplo de perspectivas que são encontradas para explicar a formação do conhecimento baseada em dados.

Nas áreas da ciência, finanças e indústria, por exemplo, os dados são

considerados matéria-prima para o aprendizado estatístico e são usados para construir modelos que possam prever, estimar e classificar resultados. Esta perspectiva enfatiza os aspectos técnicos dos dados, incluindo sua representação, estrutura e formatos, bem como os algoritmos e técnicas usadas para sua manipulação e análise (HASTIE, TIBSHIRANI, FRIEDMAN, 2009). Embora as definições de dados possam variar ligeiramente entre diferentes campos, há um consenso de que os dados representam uma forma fundamental de registro e comunicação de informações. A crescente importância dos dados na sociedade atual reflete a necessidade de se compreender e gerenciar eficazmente esses recursos, a fim de extrair valor e conhecimento para a tomada de decisões informadas e a inovação (MAYER-SCHÖNBERGER & CUKIER, 2013).

Na era digital em que se vive, os dados são criados, modificados e trafegados por toda parte. São encontrados e coletados a partir de diversas fontes como em sensores de hardware, dispositivos móveis, redes sociais, dentre outros. No entanto, a simples coleta e armazenamento de dados não são suficientes para gerar valor para uma organização. É necessário que sejam transformados em informação útil e significativa para dar continuidade no processo da construção de conhecimento e sabedoria. Importante destacar que nem todos os dados são úteis para as organizações. Precisam ser selecionados e analisados os dados relevantes, de acordo com as necessidades e objetivos de cada organização. A informação só se torna conhecimento quando é interpretada e entendida em um contexto específico. Dessa forma, é fundamental que as empresas saibam como transformar esses dados em informações e conhecimentos úteis para o negócio. O conhecimento é um ativo intangível que pode ser criado a partir dos dados e informações, permitindo às empresas tomar decisões mais informadas e com maior probabilidade de sucesso. (CHEN E ZHANG, 2014).

É importante considerar que a informação e o conhecimento são constantemente evoluídos, refletindo as mudanças no ambiente social, político, econômico e tecnológico. Davenport (1998) direciona os dados como observações simples do estado das coisas, apresenta a informação como dados de maior relevância e propósito, e o conhecimento é apresentado como informação valiosa e mais sofisticada da mente humana repleta de reflexões, contextos e complexidade.

Nesse sentido, é fundamental que as organizações sejam capazes de gerenciar seus processos de informação e conhecimento de forma eficiente para adaptarem-se às mudanças e manterem sua vantagem competitiva. Destaca-se que a relação entre informação e conhecimento não é unidirecional, mas sim bidirecional. Ou seja, o conhecimento pode ser influenciado pela forma como as informações são buscadas, selecionadas e interpretadas, assim como a informação pode influenciar o conhecimento produzido. A Pirâmide do Conhecimento, apresentada pela International Society for Knowledge Organization (ISKO), estabelece sua hierarquia DIKW como: dados, informação, conhecimento e sabedoria. Este último elemento é complementar aos outros já vistos, sendo a sabedoria interpretada como diversos conhecimentos utilizados na prática para atingir um resultado desejado (FRICKÉ, 2019). O topo da pirâmide requer uma base sólida e consistente, e, neste caso, os dados são apresentados como a base dessa representação.

Para que o conhecimento surja, é necessária a presença de informação e, para que a informação possa existir, os dados precisam ser contextualizados. Esta hierarquia é importante para a composição do conhecimento, pois dados são fatos isolados que, por si só, não geram conhecimento (TUOMI, 1999). A relação entre informação e conhecimento, complexa e dinâmica, deve ser compreendida de maneira multidisciplinar, envolvendo aspectos cognitivos, sociais, culturais e tecnológicos. A ciência da informação desempenha um papel fundamental na compreensão e gestão desses processos, buscando contribuir para o desenvolvimento de organizações e sociedades mais informadas e conscientes.

Nos últimos anos, houve um aumento significativo no volume de dados gerados e coletados pelas organizações. Diante desse volume massivo de informações, torna-se fundamental que as empresas saibam transformar esses dados em informações e conhecimentos que possam ser usados para tomadas de decisão e melhorias nos negócios. A informação é vista como uma das principais vantagens competitivas das organizações atualmente. No entanto, a simples coleta de dados não é suficiente para garantir o sucesso empresarial. Faz-se necessário que as organizações saibam gerenciar e analisar esses dados para transformá-los em informações relevantes.

## **2.2 Dados como um ativo corporativo**

No contexto contemporâneo da economia digital, os dados emergem não apenas como um mero subproduto das transações, mas como um ativo central que impulsiona a inovação e a criação de valor. A centralidade dos dados na economia digital baseada em dados é multifacetada. Eles atuam como meio de transações digitais, representam bens intangíveis de capital e servem como infraestrutura essencial para uma economia cada vez mais digitalizada. Esta perspectiva transforma a compreensão tradicional de ativos, colocando os dados no epicentro da geração de riqueza e inovação (CIURIAK, 2018).

As grandes empresas de tecnologia, frequentemente referidas como Big Tech, têm redefinido a noção de ativos corporativos, colocando os dados pessoais como protagonistas de suas estratégias de valorização. Embora tradicionalmente os ativos tangíveis, como propriedade e equipamento, tenham sido vistos como os principais pilares do valor corporativo, a era digital testemunhou uma transição para ativos intangíveis, com dados pessoais emergindo como uma nova classe de ativo. Entretanto, não são os dados em si que são intrinsecamente valiosos, mas a capacidade de transformar esses dados em engajamento do usuário, que por sua vez pode ser monetizado. Esta reconfiguração do valor, onde o engajamento do usuário derivado de dados pessoais é central, desafia as noções tradicionais de ativos e valor corporativo e exige uma reavaliação das estratégias e práticas regulatórias no ambiente digital (BIRCH, COCHRANE, WARD, 2021).

No atual ambiente empresarial, os dados emergem como um ativo estratégico inigualável, distinguindo-se por sua durabilidade, não degradação e capacidade de não se esgotar. Ao contrário de outros ativos tradicionais, os dados, quando adequadamente gerenciados, não apenas mantêm sua integridade ao longo do tempo, mas também aumentam em valor à medida que avançam na cadeia de valor da informação. Esta progressão, desde a formulação de previsões empresariais até a concretização em transações e, posteriormente, servindo como base para futuras estratégias, destaca a natureza cíclica e auto-reforçadora dos dados (AIKEN, WALTON, 2013).



### 2.3 Preservação Digital e Curadoria Digital

A preservação digital e a curadoria digital são temas que têm se tornado cada vez mais relevantes em uma época na qual a produção e o consumo de informações ocorrem em ambientes digitais. Os dados assumem várias formas, tais como números, imagens, arquivos em formato XML, documentos, dentre muitos outros, no entanto, o que precisa ser preservado de fato são as informações codificadas por esses dados. A preservação de cada tipo de dado é conhecida por trazer uma série de benefícios para comunidades específicas e também por razões específicas. Informações são preservadas por muitos em função de questões legais, enquanto outros preservam por uma questão de memória, e alguns para exploração comercial, etc. (GIARETTA, 2022).

A preservação digital compreende métodos, habilidades e resultados bem diversos e podem ser vistos como complementares aos métodos tradicionais de preservação, contudo, novos ambientes surgem dentro do ecossistema de preservação digital, como repositórios que precisam ser confiáveis, o objetivo digital em si, assinaturas digitais, identificadores de persistência, dentre outros (JANTZ, GIARLO, 2005). Também são envolvidos diferentes processos, como a ingestão, armazenamento, recuperação, acesso e reuso dos recursos digitais. A preservação digital deve ser vista como um processo dinâmico e iterativo, que requer a constante revisão e atualização das estratégias de gestão dos recursos digitais. Oliver e Harvey (2016, p.90) definem a preservação digital como:

A preservação digital refere-se a atividades destinadas a garantir que possamos acessar dados, objetos digitais e bancos de dados no futuro — por mais tempo do que a vida útil do software e hardware usados em sua criação e manutenção inicial.

Além de atividades para garantir o acesso, a preservação digital envolve uma série de desafios técnicos, legais e éticos. Dentre os desafios técnicos, destaca-se, por exemplo: o monitoramento constante da obsolescência dos formatos de documentos digitais, a garantia da autenticidade e da integridade dos dados por meio de ações coordenadas de ingestão de dados, cópias de segurança e checagem periódica da integridade dos objetos preservados.

Já em relação aos desafios legais e éticos, torna-se importante destacar a

necessidade de garantir a privacidade e a segurança dos dados, uma vez que a preservação digital deve ser realizada de forma ética, dentro das políticas legais estabelecidas, garantindo a privacidade e os direitos dos usuários. Todas as ações devem ser realizadas de forma transparente e responsável, garantindo a proteção dos interesses dos usuários e da sociedade como um todo.

A preservação digital deve levar em consideração a interpretação e uso em diferentes contextos. Siebra e Borba (2021) ressaltam que a preservação a longo prazo não está apenas ligada ao suporte, como observado no ambiente analógico, mas passa a possuir três níveis: o físico, onde o objeto digital se encontra; o lógico, onde são trabalhados os formatos do objeto digital e seu respectivo hardware e software necessários; e o conceitual, que se concentra em processos para garantir a integridade e autenticidade do objeto digital. Para que isso ocorra, é necessário que as instituições responsáveis pela preservação digital desenvolvam uma visão estratégica que possa garantir a qualidade e a relevância dos dados ao longo do tempo, identificando e prevendo os possíveis riscos que rondam o ecossistema de preservação. “Desse modo, sem a visão estratégica, a preservação digital é apenas um amontoado de técnicas.” (GALINDO, 2021, p.24). Um dos desafios técnicos da preservação digital é a constante evolução das tecnologias de armazenamento e processamento de dados, o que exige a atualização constante das estratégias de preservação digital. A informação digital é frágil e se encontra em um ambiente repleto de riscos e impactos além da tecnologia. “Neste contexto, a preservação digital é vista como promotora do acesso a longo prazo, sendo uma necessidade para instituições de memória, de pesquisa e para organizações em geral” (TAVARES, FREIRE, 2019, p.4).

A curadoria digital, por sua vez, é entendida como a gestão ativa dos recursos digitais, com o objetivo de preservá-los e torná-los acessíveis a longo prazo. A curadoria digital é compreendida como um conjunto de estratégias com abordagens tecnológicas e atividades que envolvem uma gestão atuante na preservação de recursos digitais durante todo o ciclo de vida (SAYÃO & SALES, 2012). De acordo com Gracy e Kahn (2012), a curadoria digital possui um escopo maior que a preservação digital, pois envolve tanto atividades que podem ser caracterizadas como relacionadas à preservação, mas também abrange os aspectos do ciclo de

vida dos objetos digitais. Neste contexto, Gracy e Kahn (2012) destacam a definição do DCC para curadoria digital como sendo “manter, preservar e agregar valor aos dados de pesquisa digital em todo o seu ciclo de vida”.

Oliver e Harvey (2016, p.25) destacam que “a curadoria digital é um conceito mais inclusivo do que o arquivamento digital ou preservação digital” pelo fato de abordar uma variedade de processos direcionados ao objeto digital enquanto sua utilidade. Também é ressaltado que a curadoria digital inicia antes mesmo que o objeto digital seja criado, estabelecendo padrões que permitem o planejamento da adequada coleta de dados para que sejam atribuídos como homologados para a curadoria. Há uma ênfase sobre a agregação de valor nos conjuntos de dados e objetos digitais que vão além da preservação digital, incluindo, por exemplo, a participação de várias áreas demonstrando assim sua colaboração interdisciplinar. Siebra e Borba (2021) destacam que a preservação digital se coloca como uma atividade específica dentro do ciclo de vida da curadoria digital na qual volta sua atenção para as formas e ações necessárias para enfrentar a obsolescência tecnológica. Oliver e Harvey (2016) caracterizam a curadoria digital por alguns aspectos: pela variedade de processos aplicados no momento da criação de um objeto digital, pela sua capacidade de reproduzir esses objetos, por sua agregação de valor aos objetos digitais para seu reaproveitamento, pela sua diversidade de papéis oriundos de diversas áreas de atuação, por interesse em soluções de código aberto e pelo elo entre pesquisa e prática.

DeRidder (2018, p.3, tradução nossa) declara que: “A curadoria digital é um esforço multifacetado para garantir o acesso e uso atuais e de longo prazo do conteúdo digital”. Nesse contexto, é transmitido a ideia de que não se pode gerenciar efetivamente um conteúdo digital de forma isolada, necessita-se de um engajamento de várias perspectivas já que uma curadoria digital bem-sucedida está ligada à constante atualização do ecossistema que envolve o objeto digital em seu formato, seus sistemas de apoio e seu contexto em um determinado domínio. Como a curadoria digital é uma sequência de processos que devem ser aplicados ao objeto digital, torna-se necessário o uso de modelos de referência para que sejam colocados em prática. “No entanto, verifica-se que a diversidade de objetos digitais sendo criados é cada vez maior e a curadoria digital não possui uma “fórmula” única

que sirva para todo e qualquer caso.” (SIEBRA, 2019, p.3). As particularidades do objeto digital podem direcionar para o uso de um determinado modelo, porém o desafio de lidar com diferentes domínios pode ser útil para outros, entregando técnicas e lições aprendidas, compartilhando serviços e evitando retrabalho (PENNOCK, 2007).

A era digital e a crescente produção de conteúdo têm gerado desafios significativos no campo da Ciência da Informação, especialmente no que tange à preservação digital e curadoria digital. A salvaguarda de informações em um ambiente digital requer atenção especial, uma vez que a obsolescência tecnológica e a degradação dos suportes de armazenamento podem resultar na perda de dados valiosos e históricos (BORGMAN, 2015). E dentro desse contexto, é-se deparado com as abordagens de curadoria e preservação digital através de modelos de referência que são fundamentais para viabilizar a salvaguarda de informações a longo prazo, mantendo-se suas fiéis propriedades.

Modelos de referência em preservação e curadoria digital foram sendo construídos e testados ao longo dos anos. Alguns modelos possuem focos específicos, porém em todos eles existem processos bem definidos que permitem a implantação de uma gestão mais eficaz para garantir os objetivos da preservação e curadoria de ativos digitais a longo prazo. Dentre os modelos de referência mais conhecidos no campo da Ciência da Informação, encontra-se o modelo conceitual para repositórios de informações digitais OAIS - Open Archival Information System; modelos para certificação de repositórios digitais que atendem requisitos de preservação digital (TRAC - Trustworthy Repositories Audit & Certification, ISO 16363, Core Trust Seal; modelo para avaliação de risco no apoio ao planejamento da preservação digital (DRAMBORA); modelo que apoia a implementação progressiva da curadoria digital por meio de níveis de maturidade NDSA Levels of Preservation e o modelo de ciclo de vida da curadoria digital DCC Curation Lifecycle Model.

A importância da colaboração entre profissionais da Ciência da Informação e outras disciplinas é amplamente reconhecida na literatura. A integração desses campos de conhecimento contribui para o desenvolvimento de estratégias eficazes de curadoria e preservação digital, abordando desafios como a interoperabilidade, a

preservação de formatos e o acesso sustentável às informações digitais (TIBBO, 2015; OLIVER, 2016). Os modelos OAIS, TRAC, ISO 16363, NDSA Levels of Preservation e DCC desempenham um papel significativo na promoção das melhores práticas em curadoria e preservação digital, contribuindo para a construção de repositórios digitais confiáveis e sustentáveis. Ao serem adotados esses modelos, a integridade, a autenticidade e a acessibilidade das informações digitais podem ser garantidas pelas organizações ao longo do tempo, respondendo-se aos desafios emergentes e às demandas dos usuários no contexto da Ciência da Informação.

O modelo OAIS (Open Archival Information System) foi proposto pelo Consultative Committee for Space Data Systems (CCSDS, 2012), órgão americano que discute e desenvolve padrões para dados espaciais e sistemas de informação. Esse modelo aborda aspectos críticos da preservação de informações digitais, incluindo a ingestão, o armazenamento, a gestão, o acesso e a disseminação de informações, bem como a preservação a longo prazo e a manutenção da autenticidade e integridade dos dados (DERIDDER, 2018).

Identificam-se e especificam-se no modelo OAIS o ambiente, os componentes funcionais e o modelo de informação de um repositório de dados. O ambiente é definido como o contexto de pacotes de informações e suas interações com o produtor das informações a serem armazenadas, com o consumidor das informações e com o gestor da base de dados. A forma como a informação é armazenada é especificada pelo modelo de informação, identificando as estruturas e tipos de pacotes de informação (pacotes de submissão, armazenamento e disseminação da informação). O modelo funcional especifica seis componentes: Ingestão, Armazenamento, Gestão de Dados, Administração, Preservação e Acesso (GIARETTA, 2022; DERIDDER, 2018).

Conforme Ball (2010), dois objetivos principais são estabelecidos pelo modelo OAIS: estabelecer uma terminologia padrão para descrever os recursos dos repositórios e determinar um nível mínimo de funcionalidade para esses repositórios. Beneficiando-se da gama de pesquisas que envolvem o modelo OAIS, este modelo é selecionado para compor a análise desta pesquisa. Aspectos fundamentais da preservação digital são abordados pelo modelo OAIS, incluindo a ingestão,

armazenamento, gerenciamento, acesso e disseminação de informações, bem como a manutenção da autenticidade e integridade dos dados. A escolha desse modelo para análise possibilita o aprofundamento nos desafios e questões críticas relacionadas à preservação e acesso de informações digitais a longo prazo, proporcionando capacidade de explorar e adaptar as diretrizes e melhores práticas para um contexto específico de outro domínio, a fim de contribuir para o avanço do conhecimento e a aplicação eficaz da preservação digital (PATEL ET AL., 2009). Ademais, desafios emergentes e tendências no cenário da preservação digital, como a crescente complexidade e volume de dados, a necessidade de interoperabilidade entre sistemas e a importância da sustentabilidade na gestão de informações, são abordados pelo modelo OAIS.

TRAC, ISO 16363 e CoreTrustSeal são modelos de certificação de repositórios, que têm como base a especificação de repositório de OAIS. O Modelo TRAC (Trustworthy Repositories Audit & Certification) foi desenvolvido pelo Center for Research Libraries (CRL) e pela Online Computer Library Center (OCLC), com o objetivo de fornecer um conjunto de critérios para avaliar e certificar repositórios digitais confiáveis. Este modelo aborda aspectos como a governança e a estrutura organizacional, a preservação e a manutenção das informações digitais e a garantia de acesso aos usuários. O TRAC é frequentemente citado como um marco importante na promoção da confiabilidade e da transparência dos repositórios digitais (CONWAY, 2011).

A ISO 16363 (Space data and information transfer systems – Audit and certification of trustworthy digital repositories) é uma norma internacional que estende e formaliza os critérios do TRAC, fornecendo um padrão reconhecido globalmente para a auditoria e certificação de repositórios digitais confiáveis. A importância de práticas consistentes e eficazes de curadoria e preservação digital é enfatizada por esta norma, que auxilia na criação de repositórios sustentáveis e capazes de garantir o acesso a longo prazo às informações digitais (TRAC, 2007). A norma 16363 decorre da recomendação Audit and Certification Of Trustworthy Digital Repositories (ACTDR), desenvolvida pelo Consultative Committee for Space Data System (CCSDS).

O Core Trust Seal é uma certificação internacional desenvolvida pelo

CoreTrustSeal Board, com o objetivo de estabelecer padrões básicos para repositórios digitais confiáveis (CORETRUSTSEAL, 2022). Baseia-se em 16 requisitos que abordam áreas como governança, infraestrutura e gerenciamento de dados, incluindo aspectos como a missão do repositório, licenças e acesso, preservação e armazenamento de dados, entre outros. A confiabilidade e o compromisso com as melhores práticas de gestão e preservação de dados podem ser demonstrados pelos repositórios ao obterem a certificação Core TrustSeal. Com a exigência do modelo ISO16363 em vista, o Core TrustSeal é visto como um primeiro passo para uma certificação de um repositório, visando atingir futuramente uma certificação formal de acordo com ISO16363 (DILLO E LEEUW, 2018).

NDSA Levels of Preservation: trata-se de um modelo de referência com uma abordagem desenvolvida pela National Digital Stewardship Alliance (NDSA), dos Estados Unidos, que estabelece um conjunto de níveis progressivos de práticas de preservação digital (NDSA, 2022). Esses níveis, variando do mais básico ao mais complexo, auxiliam as instituições na identificação das etapas necessárias para melhorar a preservação de seus acervos digitais e na implementação de estratégias de acordo com suas necessidades e recursos (PHILIPS ET AL., 2013).

DRAMBORA (Digital Repository Audit Method Based on Risk Assessment) é um método de auditoria de repositórios desenvolvido pelo Digital Curation Centre (DCC) e pelo Digital Preservation Europe (DPE), baseado na avaliação de riscos que tem por objetivo identificar, avaliar e mitigar riscos associados à preservação digital, garantindo a acessibilidade e autenticidade das informações a longo prazo. A DRAMBORA serve como um facilitador para abrir canais de comunicação entre repositórios distintos, porém interligados, permitindo a identificação e a difusão de melhores práticas, tanto gerais quanto especializadas. Através da construção de categorias de repositórios, torna-se possível espelhar e comunicar realidades práticas. Em uma disciplina ainda em desenvolvimento, onde a falta de experiência e a incerteza dos envolvidos são obstáculos ao avanço, a disseminação de conhecimentos emergentes, adequados a contextos específicos pode proporcionar benefícios significativos (MCHUGH ET AL., 2008).

O DCC Curation Lifecycle Model, desenvolvido pelo Digital Curation Centre (DCC), propõe um ciclo de vida para a curadoria digital, abordando as atividades e

os processos que garantem a preservação e a acessibilidade das informações. Para fornecer uma visão abrangente das atividades e responsabilidades envolvidas na curadoria e gestão de dados digitais, o DCC desenvolveu o modelo de Ciclo de Vida de Curadoria (HIGGINS, 2008). “O Modelo de Ciclo de Vida de Curadoria do DCC é talvez o 'mapa' mais abrangente do que está envolvido na curadoria digital, pois cobre todo o ciclo de vida do conteúdo digital” (DERIDDER, p.21, 2018, tradução nossa). A cooperação entre bibliotecários, arquivistas e criadores de informações poderia assegurar que conteúdos fossem registrados em formatos arquivísticos, bem descritos, devidamente licenciados e armazenados de forma segura. No processo de criação de novos conteúdos digitais, materiais preexistentes poderiam ser incorporados ou modificados, gerando, assim, um "novo" objeto para a curadoria do ponto de vista da preservação digital. Isso está alinhado com um modelo que abrange todo o ciclo de vida do conteúdo, incluindo a eventual exclusão daqueles que não são selecionados para preservação de longo prazo ou removidos após reavaliação (DERIDDER, 2018).

O DCC Curation Lifecycle Model é uma ferramenta versátil que proporciona uma visão abrangente das atividades envolvidas na curadoria digital. Foi projetado para ser adaptado pelas organizações de modo a moldar suas práticas de curadoria digital, permitindo a identificação das ações, tecnologias, padrões e competências requeridas em cada etapa, com a flexibilidade de serem adicionadas ou removidas conforme necessário. O modelo não se destina a um único tipo de operação de curadoria digital ou a uma disciplina específica, podendo ser aplicado em uma variedade de contextos de curadoria digital (OLIVER, HARVEY, 2016).

Ademais, o Ciclo de Vida de Curadoria DCC (CVCD/DCC) mostra-se compatível com uma ampla variedade de disciplinas e domínios de pesquisa, tornando-se aplicável a diversos contextos e cenários. O modelo CVCD/DCC também tem sido o objeto de numerosos estudos e amplamente citado como referência de trabalhos em curadoria digital e, por esse motivo, é um dos objetos de análise desta pesquisa.

Oliver e Harvey (2016, p.61) reforçam os benefícios do uso do DCC/CVCD e solidificam o modelo como um dos modelos de análise desta pesquisa:



Os benefícios do Modelo são representados por possibilitar um melhor mapeamento das atividades em relação ao Ciclo de Vida; identificar pontos fracos na prática; auxiliar na identificação de colaboradores (por exemplo, criadores de dados) no processo de curadoria de dados; apoiar a documentação de políticas e processos; incentivar o desenvolvimento de padrões e tecnologias; e auxiliando na identificação de ferramentas e serviços para curadoria de dados. Foi desenvolvido um modelo de ciclo de vida de curadoria digital estendido baseado no Modelo DCC (Constantopoulos et al. 2009).

Como o modelo CVCD/DCC pode ser utilizado em conjunto com outros modelos de referência, trabalhos em domínios específicos exploram essa versatilidade e abrem possibilidades de extensão das práticas de preservação e curadoria digital. De acordo com Gracy e Kahn (2012), o CVCD/DCC pode ser empregado de maneira complementar ao modelo OAIS. Pires e Rocha (2020, p.18-19) correlacionam as atividades do CVCD/DCC com as atividades do modelo OAIS e ACTDR (ISO 16363) a favor de uma implementação em uma instituição de preservação, conforme apresentado no Quadro 1.

Quadro 1 - Relação das atividades com Ciclo de Vida de DCC, Entidades e Processos de OAIS e Critérios de ACTDR

Atividades	DCC	OAIS	ACTDR
Avaliar e Selecionar	Avaliar e selecionar, Criar e receber, Eliminar, Reavaliar	Administração: negociar acordo de submissão	Governança e viabilidade organizacional (3.1) Contratos, licenças e obrigações (3.5)
Ingerir	Ingerir	Ingestão: coordenar atualizações, gerar AIP, gerar informações descritivas, garantia de qualidade, receber submissão	Ingestão - aquisição de conteúdo (4.1) Ingestão - criação do AIP (4.2)
Armazenar a Longo Prazo	Armazenar	Armazenamento: fornecimento de dados, recebimento de dados	Preservação de AIP (4.4)
Prover Acesso	Acessar, usar e reusar	Acesso: coordenar atividades de acesso, fornecer resposta, gerar DIP	Gestão de Acesso (4.6)
Gerenciar Dados e Informação Descritiva	Descrição e representação da informação	Gestão de Dados: administrar base de dados, gerar relatórios, realizar consultas, receber atualização de base de dados	Gestão da Informação (4.5)

Continua

Fonte: Pires e Rocha (2020, p.18)

Quadro 2 - Relação das atividades com Ciclo de Vida de DCC, Entidades e Processos de OAIS e Critérios de ACTDR

Atividades	DCC	OAIS	ACTDR
Administrar o Ambiente Organizacional	Curar e preservar	Administração: Atendimento ao cliente, controle de acesso às instalações físicas	Governança e viabilidade organizacional (3.1) Estrutura organizacional e pessoal (3.2) Sustentabilidade financeira (3.4)
Gerenciar Documentação, Auditar e Promover a Transparência	Curar e preservar	Administração: estabelecer padrões e políticas, auditar submissão	Responsabilidade processual e estrutura de política de preservação (3.3) Critério de várias seções que determinam especificações e evidências de cumprimento
Planejar a Preservação	Planejamento da preservação	Planejamento da Preservação: desenvolver projetos de empacotamento e planos de migração, desenvolver estratégias e padrões de preservação	Planejamento da Preservação (4.3)
Desenvolver a Preservação Digital: Monitorar a Comunidade e Monitorar Gerenciar Riscos Associados à Infraestrutura Tecnológica	Participação e observação da comunidade, Ação de preservação Migrar, Transformar	Administração: ativar solicitações, atualizar informações do arquivo, gerenciar a configuração do sistema. Armazenamento: recuperação de desastres, verificação de erros, gerenciamento da hierarquia de armazenamento, substituição de mídia Planejamento da Preservação: desenvolver estratégias e padrões de preservação, monitorar comunidade designada, monitorar tecnologia	Gerenciamento de risco de infraestrutura técnica (5.1) Gerenciamento de risco de segurança (5.2)

Fonte: Pires e Rocha (2020, p.18-19)

## 2.4 Modelo Open Archival Information System (OAIS)

O OAIS é um modelo de referência que recomenda práticas para a formulação de uma estrutura de arquivamento com o objetivo de preservar informações e disponibilizá-las para uma comunidade específica. Apesar do modelo apoiar a preservação de informações físicas, o foco do OAIS é em informações digitais e, portanto, a modelagem e a preservação das informações físicas não são abordadas pelo modelo. (CCSDS, 2012). São atribuídas ao modelo uma ampla abordagem de funções de preservação de informações de arquivamento, que incluem ingestão, armazenamento, gerenciamento de dados, acesso e disseminação. A aplicação do modelo OAIS é estendida a todas as organizações,

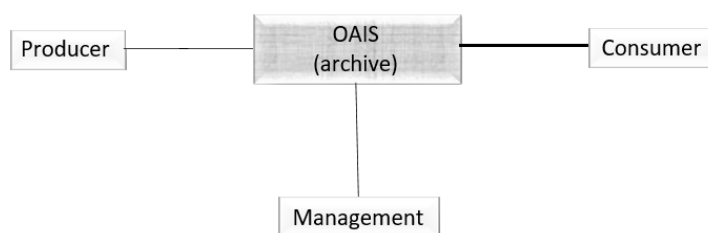
sendo especificamente aplicável àquelas comprometidas com a manutenção e disponibilização de informações a longo prazo. Variáveis importantes como mudanças tecnológicas ou possíveis mudanças na própria comunidade em si são discutidas no aspecto da preservação a longo prazo. A tecnologia talvez seja o fator mais complexo em curso, dado que em alguns casos existe a exigência de manter a aparência das informações preservadas, o que só é possível se toda a estrutura original que suporta a informação for preservada. O modelo OAIS não se detém nessas técnicas, mas provê um fundamento de arquitetura para protótipos e comparação de técnicas de simulação da estrutura original da informação.

### 2.3.1 O ambiente de um OAIS

O ambiente de um OAIS é composto pelo sistema de arquivamento, produtor, consumidor e pelo gerenciamento do próprio sistema de arquivamento (um OAIS), conforme é apresentado na Figura 1 .

O papel de Produtor (Producer) é desempenhado por pessoas ou sistemas clientes que fornecem as informações a serem preservadas. O papel de Gerenciamento (Management) é executado por aqueles que definem políticas de gestão do OAIS e o utilizam como componente de uma série de outras políticas mais abrangentes. A administração é responsável pela gestão do OAIS. O papel de Consumidor (Consumer) é desempenhado por pessoas ou sistemas clientes que buscam o OAIS para recuperar as informações designadas para preservação. Um OAIS pode atuar como produtor ou consumidor frente a outros OAIS, demonstrando interoperabilidade (CCSDS, 2012).

Figura 1 - Modelo de ambiente de um OAIS



Fonte: CCSDS (2012, p.2-2)

Conforme CCDS (2012, p.29, tradução nossa), "A informação é definida como qualquer tipo de conhecimento que pode ser trocado, e essa informação é sempre expressa por algum tipo de dado em uma troca". Na prática, a base de conhecimento necessária para interpretar a informação deve ser possuída por quem recebe os dados. Essa base de conhecimento é definida por uma comunidade, um grupo com conhecimento específico responsável por manter o conteúdo ao longo do tempo.

Para o sucesso da preservação da informação, é necessário que cada componente responsável por compor a informação seja identificado por um OAIS. No caso de informações digitais, é necessária a compreensão ao nível de bits. Outro aspecto relevante é a preservação do suporte físico no qual a informação é representada, pois, em alguns casos, é necessário manter a aparência original da informação, o que pode exigir a preservação de toda a cadeia de hardware e softwares.

### 2.3.2 Os pacotes de Informação de um OAIS

Um Pacote de Informações é conceituado como um recipiente que une dois tipos de informações, denominadas Informações de Conteúdo e Informações de Descrição de Preservação (conhecidas pela sigla em inglês PDI, de Preservation Description Information). Este recipiente é simbolicamente embalado e identificado por uma Informação de Embalagem (Packaging Information).

A Informação de Conteúdo é composta por dados e informações de representação associadas. Constitui-se no objeto físico ou digital (bits) e na descrição que possibilita a representação da informação para que uma comunidade ou grupo possa visualizar e compreender. A definição da Informação de Conteúdo deve ser realizada inicialmente para que se possam criar as Informações de Descrição de Preservação. As Informações de Descrição de Preservação são divididas em cinco tipos:

- Informação de Proveniência: descreve a fonte das Informações de Conteúdo, que tem a custódia dela desde a sua origem e sua história (incluindo histórico de processamento);
- Informação de Contexto: descreve o relacionamento entre as Informações de Conteúdo com outras informações fora do Pacote de Informações;
- Informação de Referência: fornece identificadores, ou sistemas de identificadores pelos quais as Informações de Conteúdo podem ser identificadas exclusivamente;
- Informação de Fixidez: é um esquema de proteção das Informações de Conteúdo que impedem alterações não documentadas;
- Informação de Direitos de Acesso: fornece os termos de acesso, incluindo preservação, distribuição e uso das Informações de Conteúdo.

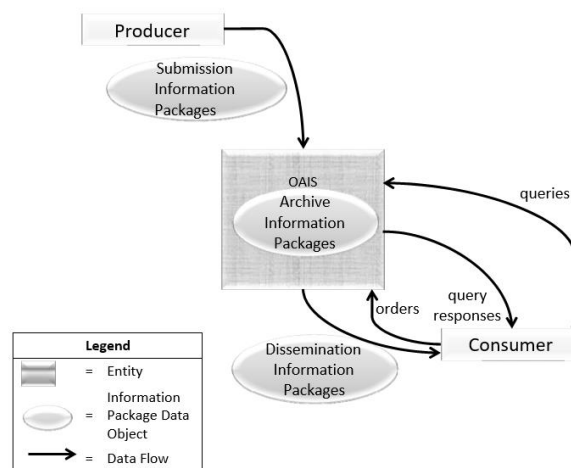
Dentro de um contexto onde um OAIS esteja presente, existem pacotes de informações padronizados que são criados conforme as regras do OAIS e pacotes que são enviados para um OAIS, mas que não cumprem completamente os padrões do OAIS. Assim, o CCSDS introduz as seguintes variantes: Submission Information Package (SIP), Archival Information Packages (AIP) e Dissemination Information Package (DIP).

Submission Information Package (SIP) são pacotes enviados para um OAIS por um Produtor e, na grande maioria desses pacotes, há informações de conteúdo e descrições. Este pacote resulta na produção de outro tipo que é o Archival Information Packages (AIP). Isto é, um Produtor pode submeter vários SIPs, gerando AIPs que podem ter uma relação de alta complexidade. Por sua vez, os Consumidores podem solicitar informações ao OAIS e, em resposta, o OAIS fornece um pacote chamado Dissemination Information Package (DIP). A qualidade do DIP depende do pedido do Consumidor, podendo possuir ou não um PDI completo ou incluir coleções de AIPs.

### 2.3.3 Interações entre Produtor e Consumidor

As interações entre as entidades em um ambiente OAIS possuem processos definidos entre si e com o ambiente externo. A Figura 2 apresenta, em alto nível, o fluxo dessas interações que são divididas em interação de gestão, interação com o produtor e interação com o consumidor.

Figura 2 - Dados externos do Arquivo OAIS



Fonte: CCSDS (2012, p.2-8)

A gestão é responsável por fornecer ao OAIS o escopo das atividades que serão conferidas ao Arquivo, bem como por determinar a amplitude dos grupos de produtores e dos consumidores do OAIS. Pode haver responsabilidades atribuídas à gestão frente ao OAIS, tais como:

- Financiar e prover diretrizes para uso de recursos como pessoas, equipamentos e instalações;
- Supervisionar e avaliar o desempenho de acordo com as metas e objetivos de longo prazo;
- Determinar ou endossar políticas de preços para os serviços OAIS;
- Resolver conflitos envolvendo a produtores, consumidores e administração interna do OAIS;

O trabalho do OAIS é iniciado a partir de uma solicitação feita pelo Produtor para a preservação de seus produtos de dados e, entre eles, um Acordo de Submissão é estabelecido. Os SIPs a serem apresentados são então identificados pelo OAIS, que analisa o acordo proposto que pode conter requisitos obrigatórios para fornecer informações ao OAIS ou outras ofertas de informações. Dentro de um Acordo de Submissão, existem Sessões de Envio de Dados que possuem um ou mais SIPs que devem cumprir os requisitos mínimos de um OAIS. Este processo nada mais é do que o reconhecimento de protocolos de entrada para um OAIS, ou seja, as regras do jogo.

A interação entre o Consumidor e o OAIS pode ser bastante ampla, pois podem existir solicitações de diversas naturezas. Um Contrato de Encomenda é estabelecido pelo Consumidor com o OAIS sobre informações e, a partir desse acordo, informações previamente tratadas são liberadas de acordo com sua natureza.

#### 2.3.4 Responsabilidades de um OAIS

Para que um OAIS cumpra os objetivos, deve-se obrigatoriamente observar os seguintes tópicos (CCSDS, 2012):

- Negociar e aceitar informações apropriadas de produtores de informações;
- Obter controle suficiente das informações fornecidas no nível necessário para garantir a Preservação a Longo Prazo;
- Determinar, por si só ou em conjunto com outras partes, quais comunidades devem se tornar a Comunidade Designada e, portanto, devem ser capazes de entender as informações fornecidas, definindo assim sua Base de Conhecimento;
- Certificar-se de que as informações a serem preservadas sejam independentemente compreensíveis para a comunidade designada. Em especial, a Comunidade Designada deverá poder compreender as

informações sem necessitar de recursos especiais, tais como a assistência dos peritos que as produziram;

- Seguir políticas e procedimentos documentados que garantam que as informações sejam preservadas contra todas as contingências razoáveis, incluindo o desaparecimento do Arquivo, garantindo que elas nunca sejam excluídas, a menos que permitido como parte de uma estratégia aprovada. Não deve haver supressões ad-hoc;
- Disponibilizar as informações preservadas à Comunidade Designada e permitir que as informações sejam divulgadas como cópias ou rastreáveis aos Objetos de Dados originais enviados, com evidências que apoiem sua Autenticidade.

### 2.3.5 Componentes Funcionais de um OAIS

Após o entendimento dos conceitos e dos componentes principais de um OAIS, o esquema é ampliado e separado em seis entidades funcionais e suas interfaces: Ingestão, Gestão de Dados, Acesso, Armazenamento, Administração e Planejamento da Preservação. A arquitetura de alto nível está representada na Figura 3.

Figura 3 - Entidades Funcionais do OAIS



Fonte: CCSDS (2012, p.4-1)



A entidade funcional de Ingestão tem a responsabilidade de fornecer serviços e funções que permitem a aceitação dos Pacotes de Informações de Submissão (SIPs) de Produtores ou de elementos internos, e preparar o conteúdo para armazenamento e gerenciamento dentro do OAIS. As funções desta entidade incluem receber os SIPs, garantir a qualidade dos SIPs, gerar AIP que obedecem aos padrões de formatação e documentação de dados do OAIS, extrair Informações Descritivas dos AIPs para inclusão no banco de dados do OAIS e coordenar atualizações para o Armazenamento Arquivístico e a Gestão de Dados.

A entidade funcional de Gestão de Dados é encarregada de fornecer serviços e funções para preencher, manter e acessar as Informações de Descrição que identificam e documentam os acervos do OAIS e seus respectivos dados administrativos. As funções desta entidade incluem a administração das funções do banco de dados do OAIS, a execução de atualizações na base de dados, a execução de consultas dos dados que estão sendo gerenciados para gerar respostas às consultas e a produção de relatórios a partir dessas respostas.

A entidade funcional de Acesso tem a responsabilidade de fornecer serviços que ajudam os Consumidores a identificar, descrever, localizar e obter informações armazenadas no OAIS. Essa entidade facilita a solicitação e o recebimento de informações sobre produtos pelos Consumidores. As funções de acesso incluem a interação com os Consumidores para atender solicitações, a aplicação de controles para restringir o acesso a informações protegidas, a coordenação das solicitações para sua conclusão efetiva, a geração de respostas (como Pacotes de Informações de Disseminação, respostas de consultas, relatórios) e a entrega dessas respostas aos Consumidores.

A entidade funcional de Armazenamento é encarregada de fornecer serviços para o armazenamento, manutenção e recuperação de AIPs. As funções desta entidade incluem receber AIPs da entidade funcional de Ingestão e adicioná-los ao armazenamento permanente, gerenciar a hierarquia de armazenamento, atualizar a mídia na qual os acervos de arquivamento estão armazenados, executar rotina de verificação de erros, fornecer recursos de recuperação de desastres e fornecer AIPs para a entidade funcional de Acesso.

A entidade funcional de Administração tem a responsabilidade de fornecer

serviços e funções para o funcionamento geral do sistema do OAIS, estabelecer e manter padrões e políticas do OAIS, fornecer suporte ao cliente e ativar solicitações armazenadas. As funções desta entidade incluem solicitar e negociar contratos de submissão com os Produtores, auditar os envios para garantir que eles atendam aos padrões do OAIS, manter a gestão de configuração de hardware e software do sistema, monitorar e melhorar as operações do OAIS, inventariar, relatar e migrar ou atualizar o conteúdo do OAIS.

A entidade funcional de Planejamento da Preservação é encarregada de fornecer serviços e funções para monitorar o ambiente do OAIS, fornecer recomendações e planos de preservação para garantir que as informações armazenadas no OAIS permaneçam acessíveis e compreensíveis para a Comunidade Designada a longo prazo.

No Quadro 3 são apresentados os conceitos e objetivos das entidades funcionais do OAIS que servem de referência para as análises desta pesquisa.

Quadro 3 - Conceitos e objetivos das entidades funcionais do OAIS

Entidade	Definição	Objetivos
Entidade Funcional de Ingestão	Trata-se do procedimento de assimilação de dados fornecidos pelos criadores de conteúdo. A Entidade Funcional de Ingestão tem a incumbência de acolher as informações provenientes dos produtores, organizando-as de modo apropriado para preservação e administração no interior do repositório	Fornecer serviços e funções para aceitar Pacotes de Informações de Submissão (SIPs) de Produtores  Preparar o conteúdo para o armazenamento e gerenciamento a partir de dentro do Arquivo.
Entidade Funcional de Armazenamento de Arquivamento	É o papel da preservação de arquivos no qual assegura que o ambiente do Arquivo continue protegido e devidamente acondicionado aonde gerencia a guarda, a conservação e o resgate dos AIPs (Pacotes de Informações de Arquivamento) sob a responsabilidade do arquivo	Fornecer serviços e funções para o armazenamento, manutenção e recuperação de Pacotes de Informações de Arquivo (AIP)
Entidade Funcional de Gerenciamento de Dados	Possui a responsabilidade em gerenciar e manter a consistência dos metadados e informações associadas aos objetos arquivados, garantindo a organização e a recuperação eficaz dos dados. Essa entidade desempenha um papel vital na preservação e acesso a informações, atuando como uma camada de gerenciamento entre as outras entidades funcionais do OAIS. Ao garantir a qualidade e a consistência dos metadados e informações associadas, a Entidade Funcional de Gerenciamento de Dados possibilita a recuperação eficiente de informações arquivadas e a manutenção da compreensibilidade e autenticidade dos dados ao longo do tempo.	Fornecer serviços e funções para preencher, manter e acessar Informações Descritivas
Entidade Funcional de Administração	É responsável por coordenar e gerenciar diversas atividades administrativas essenciais para o funcionamento adequado de um arquivo digital. Essa entidade abrange uma série de funções, incluindo a negociação de acordos de submissão, a gestão da configuração do sistema, o controle de acesso físico, o estabelecimento de padrões e políticas, a auditoria de submissões, a ativação de solicitações e o atendimento ao cliente. Ao desempenhar essas funções, a Entidade Funcional de Administração garante a preservação de longo prazo das informações, a integridade do arquivo e a capacidade de atender às necessidades dos usuários e das comunidades designadas.	Fornecer serviços e funções para a operação geral do sistema de Arquivamento
Entidade Funcional de Planejamento de Preservação	Possui um papel crucial para garantir a preservação de longo prazo e a acessibilidade das informações arquivadas em um ambiente digital. Essa entidade aborda os desafios relacionados à preservação das informações, considerando a evolução das tecnologias, formatos de dados e necessidades dos usuários.	Fornecer os serviços e funções para monitorar o ambiente do OAIS
Entidade Funcional de Acesso	Tem a incumbência de facilitar a interação dos usuários com o arquivo, permitindo a recuperação e disponibilização das informações arquivadas. Essa entidade garante que as solicitações dos usuários sejam atendidas de maneira eficiente e eficaz, convertendo as informações preservadas em formatos compreensíveis e acessíveis, conhecidos como Pacotes de formação de Disseminação (DIPs).	Fornecer os serviços e funções que apoiam os Consumidores na determinação da existência, descrição, localização e disponibilidade da informação armazenada no OAIS

Fonte: Elaborado pelo autor, 2023

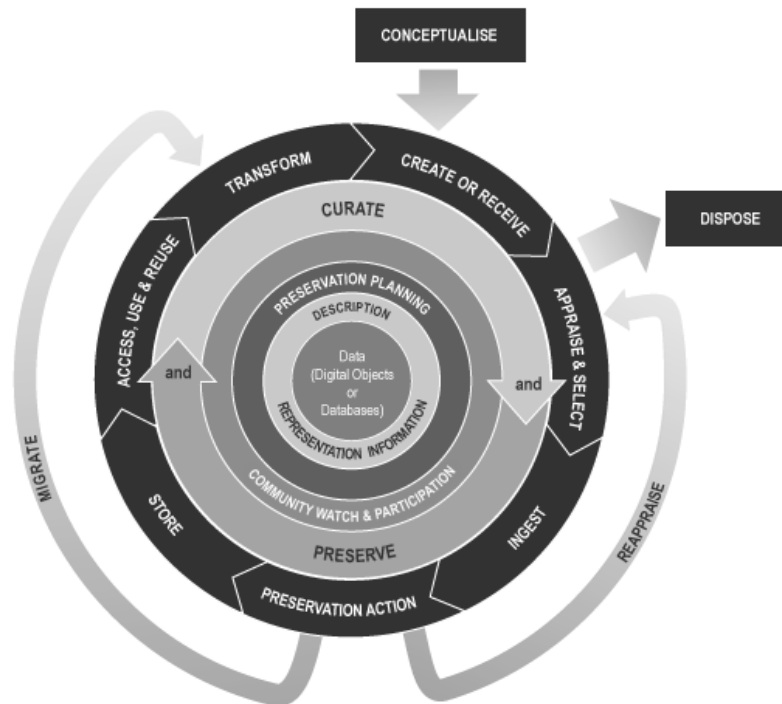
## 2.5 Modelo do Ciclo de Vida de Curadoria do DCC

Uma das contribuições mais significativas do DCC para o campo da curadoria digital é representada pelo Ciclo de Vida de Curadoria DCC, também conhecido como DCC Curation Lifecycle Model. Este modelo, proposto por Higgins (2008), fornece uma estrutura conceitual para a gestão de dados e informações digitais ao longo de todo o seu ciclo de vida, desde a criação até a preservação e reutilização. Baseia-se numa série de etapas e processos inter-relacionados, com o objetivo de garantir a autenticidade, integridade e acessibilidade dos dados ao longo do tempo.

O ciclo de vida apresenta como passos sequenciais: conceitualizar; criar ou receber; selecionar e avaliar; ingerir; realizar ações de preservação; armazenar; acessar, usar e reusar; e transformar. Apoiando todas as ações, o ciclo de vida especifica o planejamento da preservação; a descrição e a representação da informação; curar; e preservar, conforme pode ser observado na Figura 4. Esta abordagem holística permite a identificação de melhores práticas, políticas e ferramentas para cada etapa, assegurando a sustentabilidade e a longevidade dos dados digitais. A importância do Digital Curation Centre e do DCC Curation Lifecycle Model para a preservação e curadoria digital é incontestável. A adoção de boas práticas de curadoria digital é fundamental para garantir a integridade, autenticidade e acessibilidade de dados e informações digitais. Ao fornecer um modelo conceitual e uma base sólida de conhecimento, o DCC tem um papel crucial na promoção da sustentabilidade e da longevidade dos recursos digitais num mundo cada vez mais dependente da informação digital.

No Quadro 4 - Conceitos e objetivos das ações do Ciclo de Vida de Curadoria DCC, os conceitos e objetivos das ações do Ciclo de Vida de Curadoria DCC são apresentados, que servirão de referência para as análises desta pesquisa.

Figura 4 - Modelo do Ciclo de Vida de Curadoria DCC



Fonte: Higgins (2008, p. 136)

Quadro 4 - Conceitos e objetivos das ações do Ciclo de Vida de Curadoria DCC

Ações do Ciclo de Vida	Definição	Objetivos
Descrição e Representação da Informação	Esta ação envolve a criação e manutenção de metadados adequados e informações de representação associadas aos objetos digitais, permitindo que sejam descobertos, acessados e compreendidos por usuários atuais e futuros. Os metadados incluem informações descritivas, estruturais e administrativas, enquanto as informações de representação fornecem o contexto necessário para interpretar e utilizar os dados corretamente. Ao abordar a descrição e a representação da informação de maneira eficaz, os curadores digitais garantem que os recursos de dados sejam úteis e significativos para a comunidade designada, facilitando a descoberta e o acesso, bem como a preservação a longo prazo dos dados.	Garantir a descrição e o controle adequados a longo prazo  Entender e renderizar o material digital e os metadados associados

Ações do Ciclo de Vida	Definição	Objetivos
Planejamento de Preservação	Esta ação envolve a identificação e análise dos riscos e desafios associados à preservação digital, como a obsolescência tecnológica e a degradação de dados, e a definição de estratégias e planos de preservação adequados para enfrentar esses desafios. O planejamento de preservação inclui a seleção de formatos e padrões de dados sustentáveis, o desenvolvimento de políticas e procedimentos para garantir a autenticidade e a integridade dos dados e a avaliação e implementação de soluções de armazenamento e preservação. Ao abordar proativamente as questões de preservação e implementar estratégias eficazes, os curadores digitais garantem que os dados permaneçam acessíveis, compreensíveis e utilizáveis para a comunidade designada no futuro.	Planejar a gestão e administração de todas as ações do ciclo de vida da curadoria
Vigilância e Participação da Comunidade Designada	Esta ação envolve monitorar e se envolver ativamente com a comunidade designada, compreendendo suas necessidades e expectativas em relação aos recursos de dados e adaptando as práticas de curadoria para atender a essas demandas. A vigilância e a participação da comunidade designada também incluem a promoção do uso e reutilização dos dados, o incentivo à colaboração e a troca de conhecimentos entre os membros da comunidade e a identificação e incorporação de feedback para melhorar os processos de curadoria digital. Ao se concentrar no envolvimento e nas necessidades da comunidade designada, os curadores digitais garantem que os dados preservados permaneçam valiosos e aplicáveis, contribuindo para o avanço do conhecimento e a inovação.	Manter uma vigilância sobre as atividades apropriadas da comunidade  Participar do desenvolvimento de padrões compartilhados, ferramentas e software adequado
Curar e Preservar	Esta ação engloba uma série de atividades e processos que visam manter a integridade, a autenticidade e a usabilidade dos dados ao longo do tempo, enquanto enfrenta desafios, como obsolescência tecnológica, degradação de dados e mudanças nas necessidades dos usuários. Entre as principais atividades de cura e preservação estão a seleção e aquisição de dados, a descrição e representação de informações, a manutenção da qualidade dos dados, a migração de formatos e a garantia da segurança e privacidade dos dados. Ao realizar a curadoria e preservação eficazes dos dados, os curadores digitais asseguram que os recursos de dados sejam acessíveis, compreensíveis e reutilizáveis para a comunidade designada e outros usuários interessados no presente e no futuro.	Promover a curadoria e a preservação ao longo do ciclo de vida da curadoria

Ações do Ciclo de Vida	Definição	Objetivos
Conceituar	Desempenha um papel fundamental no início do processo de curadoria sendo uma ação que envolve a identificação, definição e planejamento de um projeto ou iniciativa que visa criar, preservar e disponibilizar recursos digitais de maneira eficiente e eficaz. Essa ação pode ser dividida em etapas distintas, que incluem a identificação das necessidades, definição de objetivos, planejamento estratégico, engajamento dos envolvidos, desenvolvimento de políticas e diretrizes.	Conceber e planejar a criação de dados, incluindo método de captura e opções de armazenamento
Criar e Receber	Esta atividade desempenha um papel fundamental na criação, aquisição e ingestão de recursos digitais de alta qualidade e na preparação para a sua gestão e preservação eficazes. Ao garantir a qualidade, autenticidade e integridade dos recursos digitais, esta ação contribui para a sustentabilidade e a reutilização dos dados e recursos digitais ao longo do tempo.	Criar dados, incluindo metadados administrativos, descritivos, estruturais e técnicos Receber dados de acordo com políticas de coleta documentadas a partir de criadores de dados, outros arquivos, repositórios ou data centers
Avaliar e Selecionar	É uma etapa fundamental para a tomada de decisões sobre quais dados e recursos digitais devem ser preservados e por quanto tempo. Essas ações ajudam a garantir a preservação eficiente e eficaz dos recursos digitais, otimizando o uso dos recursos e garantindo o acesso e a reutilização contínuos dos dados.	Avaliar os dados e selecionar para curadoria e preservação a longo prazo
Ingerir	É uma ação que desempenha um papel fundamental na garantia de que os recursos digitais sejam transferidos, armazenados e organizados de forma eficiente e segura em sistemas de gerenciamento de ativos digitais ou repositórios digitais. Ao abordar questões como a preparação de recursos, a transferência, a validação, a preservação e a documentação, essa ação contribui para a preservação e a reutilização eficientes e eficazes dos recursos digitais ao longo do tempo.	Transferir dados para um arquivo, repositório, data center ou outro custodiante
Ação de Preservação	É uma ação crítica que se concentra na proteção e manutenção dos recursos digitais ao longo do tempo, garantindo sua acessibilidade, integridade e autenticidade para as gerações futuras. Essa ação aborda os desafios e riscos associados à preservação digital, como a obsolescência tecnológica, a degradação dos dados e as mudanças nas necessidades e expectativas dos usuários.	Empreender ações para garantir a preservação e retenção a longo prazo da natureza autoritativa dos dados Garantir que os dados permaneçam autênticos, confiáveis e utilizáveis, mantendo sua integridade

Ações do Ciclo de Vida	Definição	Objetivos
Armazenamento	É uma etapa fundamental que se concentra na manutenção segura e eficiente dos recursos digitais. Essa ação garante que os dados e recursos digitais sejam armazenados de maneira adequada e acessível, preservando sua integridade e disponibilidade ao longo do tempo para uso e reutilização futuros.	Armazenar os dados de forma segura aderindo aos padrões relevantes
Acesso, Uso e Reutilização	É uma etapa essencial que se concentra em garantir que os recursos digitais sejam facilmente acessíveis, utilizáveis e reutilizáveis pelos usuários. Essa ação aborda a necessidade de disponibilizar os recursos digitais de maneira eficiente, segura e sustentável, promovendo a disseminação do conhecimento, a colaboração e a inovação.	Certificar-se de que os dados estejam acessíveis tanto aos usuários designados quanto aos reusuários no dia-a-dia
Transformação	É uma etapa importante que aborda a necessidade de adaptar e modificar os recursos digitais para atender às demandas em constante evolução dos usuários, às mudanças tecnológicas e às novas aplicações. Essa ação permite que os recursos digitais sejam atualizados, convertidos ou enriquecidos, de modo a manter sua relevância, acessibilidade e usabilidade ao longo do tempo.	Criar novos dados a partir do original
Descarte	É uma etapa crucial que lida com a eliminação segura e responsável dos recursos digitais que não são mais necessários ou úteis. Essa ação é importante para garantir a eficiência e a sustentabilidade dos sistemas de armazenamento e preservação digital, liberando espaço e recursos para a manutenção e a curadoria de materiais relevantes e de valor duradouro.	Descartar os dados que não foram selecionados para curadoria e preservação a longo prazo de acordo com políticas, orientações ou requisitos legais documentados
Reavaliação	É uma etapa importante que envolve a análise periódica dos recursos digitais, das políticas e dos processos de curadoria para garantir que continuem relevantes, eficazes e alinhados com as necessidades e expectativas dos usuários. Essa ação é crucial para manter a qualidade e a sustentabilidade dos recursos digitais e para identificar oportunidades de melhoria e aprimoramento.	Retornar dados que não cumpram os procedimentos de validação para posterior avaliação e reseleção
Migração	É uma etapa crítica que envolve a transferência de recursos digitais de um ambiente, sistema ou formato para outro. Essa ação é essencial para garantir a preservação, a acessibilidade e a usabilidade dos recursos digitais ao longo do tempo, especialmente em um cenário de constante evolução tecnológica.	Migrar dados para um formato diferente para garantir a imunidade dos dados contra a obsolescência de hardware ou software ou para adequar ao ambiente de armazenamento



## 2.6 Gestão de Dados em Organizações

A gestão de dados tem se tornado cada vez mais crítica para o sucesso das organizações na era digital. A quantidade crescente de dados gerados diariamente e o avanço das tecnologias de informação e comunicação têm provocado uma revolução na forma como as empresas operam, competem e inovam. (RODRIGUES, DUARTE, DIAS, 2018). A gestão de dados pode ser compreendida como o gerenciamento de um ciclo de vida. De acordo com o DMBOK, a gestão de dados é definida como o desenvolvimento, execução e supervisão de planos, políticas, programas e práticas que fornecem, controlam, protegem e aprimoram o valor dos ativos de dados e informações ao longo de seus ciclos de vida (DMBOK 2 ed., 2017).

Cada tipo de dado possui características distintas e, por sua vez, exigem diferentes ciclos de vida com requisitos de gerenciamento diferentes. (ENAP, 2021). Realizar uma gestão efetiva dos dados requer uma adaptação entre a teoria e prática e para que ambos coexistam em harmonia, alguns princípios devem ser seguidos. O ENAP (2021, p.19) aborda esses princípios como quatro asserções estratégicas:

- Os dados têm valor.
- Requisitos de gerenciamento de dados são requisitos de negócios.
- Gerenciamento de dados e do ciclo de vida.
- O gerenciamento de dados depende de diversas habilidades.

A eficácia da gestão de dados é considerada essencial para permitir que as organizações tomem decisões informadas e baseadas em dados, o que, por sua vez, contribui para a otimização dos processos de negócios, a melhoria da eficiência operacional e a inovação. Além disso, uma gestão de dados eficiente pode ajudar as organizações a lidar com os crescentes desafios de conformidade, segurança e privacidade que surgem à medida que a quantidade de dados aumenta e se torna mais complexa (HARRIS, DAVENPORT, 2017).

Em virtude da rápida evolução das tecnologias digitais e da crescente quantidade de dados gerados dentro e fora da organização, é necessário compreender a relação entre o tratamento de dados digitais e seus impactos na informação e na formação do conhecimento, bem como a importância da inserção

da preservação e curadoria digital a longo prazo no gerenciamento de dados. O tratamento adequado de dados digitais é visto como o primeiro passo para transformá-los em informações úteis e significativas.

O processo de organização, estruturação e análise de dados brutos é considerado essencial para extrair insights valiosos que permitam às organizações tomar decisões mais informadas e impulsionar a inovação. No entanto, a informação, por si só, não é suficiente. É imprescindível que as empresas transformem essas informações em conhecimento, que pode ser aplicado em processos de negócios, práticas e estratégias para criar valor sustentável (BORGMAN, 2015).

A crescente relevância da gestão de dados também está relacionada à necessidade de garantir a privacidade e a segurança dos dados. Bélanger e Crossler (2021) destacam a importância de adotar práticas de governança de dados robustas para proteger os direitos dos indivíduos e cumprir com as regulamentações de privacidade, como o Regulamento Geral de Proteção de Dados (GDPR) da União Europeia. A análise de dados e a aplicação de técnicas de aprendizado de máquina e inteligência artificial são consideradas essenciais para a transformação digital das organizações. Schwab et al. (2020) enfatizam a importância de integrar a ciência de dados e a análise avançada de dados nas estratégias de negócios das empresas, a fim de melhorar a tomada de decisões e impulsionar a inovação.

A ENAP (Escola Nacional de Administração Pública) tem enfatizado a importância da gestão de dados no contexto da administração pública, promovendo o uso de dados para melhorar a tomada de decisões, a eficiência operacional e a prestação de serviços públicos. A capacidade das organizações públicas de transformar dados em conhecimento é considerada essencial para melhorar a eficiência e eficácia das políticas públicas e programas governamentais. Nesse sentido, a gestão de dados é vista como um elemento-chave para promover a inovação e a aprendizagem organizacional no setor público (ENAP, 2022).

A gestão de dados também é considerada uma prática essencial no campo da pesquisa científica, desempenhando um papel significativo no sucesso dos projetos e na validade dos resultados. Uma gestão eficiente de dados de pesquisa é vista como fundamental para garantir a qualidade, a reprodutibilidade e a integridade das

pesquisas acadêmicas. Segundo Wilkinson et al. (2016), a gestão adequada de dados de pesquisa envolve a coleta, o armazenamento, a organização, a análise e a disponibilização dos resultados para a comunidade científica. Essa prática é considerada uma ajuda para assegurar que os dados sejam facilmente acessíveis, compreensíveis e reutilizáveis por outros pesquisadores. A qualidade e a integridade dos dados são aspectos cruciais na pesquisa científica.

A padronização e a documentação de dados são vistas como fundamentais para assegurar a confiabilidade e a comparabilidade dos resultados da pesquisa. A adoção de práticas de governança de dados é considerada essencial para garantir a conformidade com as regulamentações éticas e legais e preservar a integridade dos dados. A importância da gestão de dados de pesquisa também está relacionada à crescente ênfase na ciência aberta e na colaboração entre pesquisadores. A partilha de dados de pesquisa permite uma maior transparência e reprodutibilidade, além de promover a inovação e o avanço do conhecimento científico (MCKIERNAN et al., 2016). Para alcançar esse objetivo, é fundamental adotar práticas de gestão de dados que permitam a fácil localização, acesso e uso dos dados de pesquisa.

Um dos principais aspectos técnicos da gestão de dados é a organização dos dados. A organização de dados envolve a criação de esquemas de metadados, taxonomias e ontologias para descrever e categorizar dados de maneira consistente e padronizada. Isso facilita a recuperação de informações e a interoperabilidade entre sistemas. A curadoria de dados é outro aspecto técnico importante na gestão de dados. A curadoria de dados envolve atividades e práticas específicas para garantir que os dados sejam encontrados, compreendidos e utilizados no futuro. Atividades como: coleta, validação, preservação e manutenção de dados ao longo do tempo são cruciais para garantir a qualidade, a integridade e a reutilização dos dados em diversos contextos de pesquisa e aplicação.

A análise de dados é outro aspecto técnico fundamental da gestão de dados tanto para as organizações como na ciência da informação. A análise de dados inclui técnicas de mineração de dados, aprendizado de máquina e análise de texto para extrair padrões, tendências e conhecimento a partir de grandes volumes de dados. Essas técnicas permitem a geração de insights valiosos e a tomada de decisões informadas.

A segurança e a privacidade dos dados têm um papel extremamente relevante na condução de práticas de gestão de dados e são aspectos técnicos essenciais em todo o contexto. A importância de adotar práticas de governança de dados robustas para proteger os direitos dos indivíduos e cumprir as regulamentações de privacidade é enfatizada. Além disso, as medidas de segurança, como criptografia e controle de acesso, são cruciais para proteger os dados de ameaças e violações (BÉLANGER, CROSSLER, CORREIA, 2021).

A interoperabilidade é outro tema com aspecto técnico relevante na gestão de dados na ciência da informação. Segundo Borgman (2015), a interoperabilidade se refere à capacidade de sistemas e aplicativos diferentes trabalharem juntos para compartilhar e processar dados. A interoperabilidade pode ser alcançada por meio do uso de padrões abertos, protocolos e formatos de dados comuns.

Pode-se identificar a importância da incorporação da gestão de dados em várias áreas e tipos de organizações, e para tal, há uma necessidade de um guia que facilite o entendimento e a aplicação das mais diversas abordagens para um gerenciamento de dados efetivo. A salvaguarda das informações bem como a preocupação da manutenção da longevidade dos dados são aspectos que devem ser observados dentro do contexto da gestão de dados, oriundas da curadoria e preservação digital tratadas na ciência da informação.

## **2.7 Referências Técnicas na Gestão de Dados**

O desenvolvimento de novas tecnologias exigiu que muitos modelos de gestão fossem aprimorados e criados. Quando os dados são colocados no centro da discussão, como protagonistas da mudança organizacional, alguns processos surgem impactando diretamente os propósitos de uso desses tipos de registros informacionais digitais. Alguns modelos de gestão tornaram-se referência para organizações realizarem a gestão de seus dados, muitos deles baseados no contexto da própria administração de uma empresa.

A gestão de dados é um conjunto de funções independentes, cada uma com seus objetivos próprios, atividades e responsabilidades (LADLEY, 2020). Alguns modelos são referência para a prática da gestão em organizações, com seus

processos e atividades definidos e que podem ser empregados para a manipulação de dados, como os modelos de Governança de Dados em Organizações (Data Governance Framework - DGF), qualidade de dados em organizações (ISO 8000), governança de informação em organizações (Amsterdam Information Management Model - AIM), governança de tecnologia de informação (COBIT - Control Objectives for Information and Related Technologies) e alinhamento estratégico (Strategic Alignment Model - SAM).

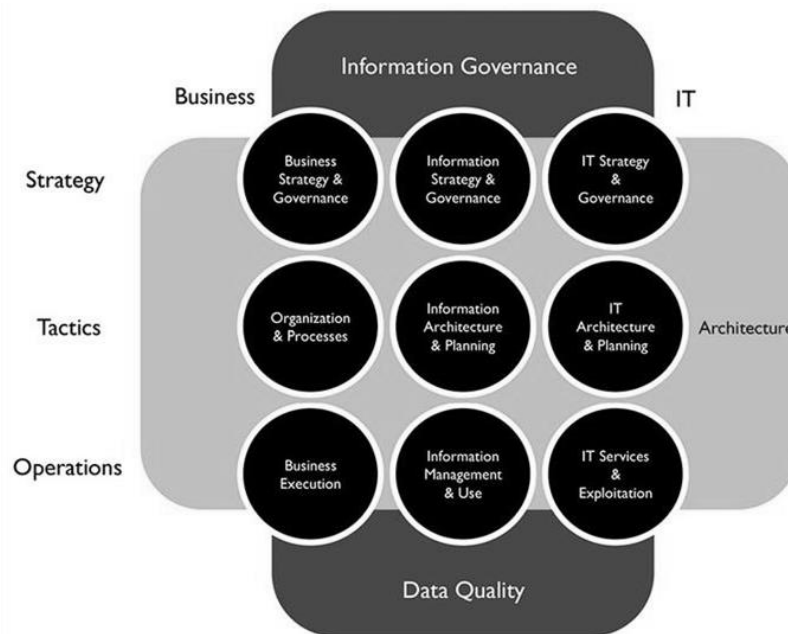
O arcabouço Data Governance Framework (DGF) provê uma abordagem estruturada para garantir que os dados sejam gerenciados de maneira eficiente, consistente e alinhada com os objetivos organizacionais (PANIAN, 2010). O DGF geralmente inclui diversos componentes, como políticas, processos, papéis e responsabilidades, indicadores de desempenho e ferramentas tecnológicas para apoiar a governança de dados (OTTO, 2011). Uma implementação bem-sucedida de um DGF pode resultar em melhor qualidade de dados, conformidade regulatória, tomada de decisões mais informada e maior eficiência operacional.

A norma ISO 8000 – Qualidade dos Dados é uma série de normas internacionais desenvolvidas pela Organização Internacional de Normalização (ISO) que aborda a qualidade e o gerenciamento de dados nas organizações. Esta norma estabelece princípios, requisitos e melhores práticas para garantir a qualidade dos dados, permitindo que as organizações tomem decisões informadas e sustentem a eficiência operacional (BATINI et al, 2016). A qualidade dos dados é um componente crítico para o sucesso das organizações em um ambiente de negócios cada vez mais orientado por dados. O gerenciamento de dados de baixa qualidade pode levar a decisões errôneas, perda de produtividade e insatisfação do cliente (WANG & STRONG, 2020). A ISO 8000 fornece diretrizes abrangentes para avaliar, melhorar e monitorar a qualidade dos dados em diversos domínios, incluindo a precisão, a integridade, a consistência, a completude e a atualidade dos dados.

O modelo Amsterdam Information Management Model (AIM), inspirado no artigo de Abcouwer, Maes e Truijens, traz as características estruturais deixando no centro a informação, a comunicação e o conhecimento. A influência das áreas de negócios com a área da tecnologia da informação está cada vez mais evidente, tornando assim a tecnologia transparente e fundindo as operações de negócio

(ABCOUWER, MAES, TRUIJENS,1997). Com essa integração acontecendo, a governança da informação concentra negócios, estratégia, táticas, operação e tecnologia, como apresentado na Figura 5.

Figura 5 -Amsterdam Information Model



Fonte: DAMA-DMBOK2 (2017, p.35)

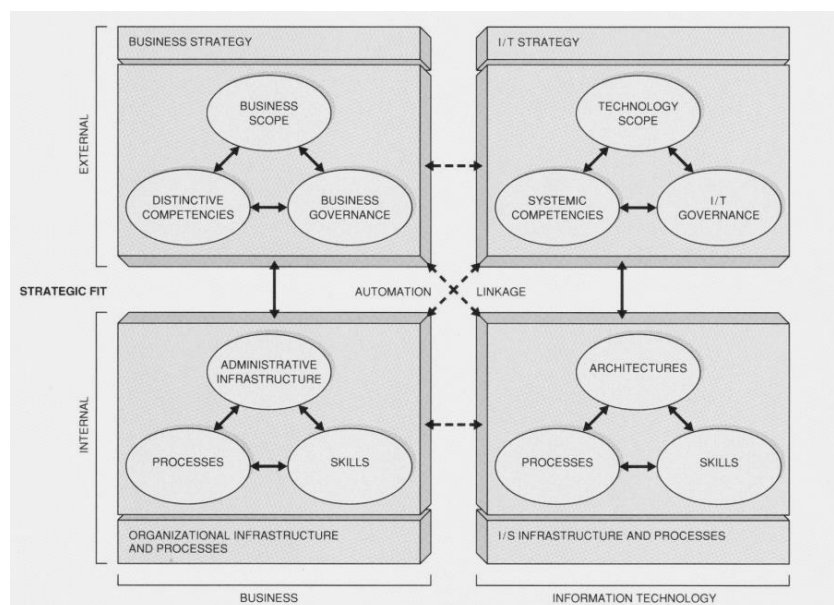
O arcabouço COBIT (Control Objectives for Information and Related Technologies) é um framework de governança e gestão de TI desenvolvido pela ISACA (Information Systems Audit and Control Association). Este framework fornece um conjunto abrangente de práticas recomendadas e objetivos de controle para ajudar as organizações a gerenciar seus recursos de TI de forma eficiente e alinhada com os objetivos do negócio (JUIZ, TOOMEY, 2015). Na perspectiva de gestão e governança de dados, um papel crucial é desempenhado pelo COBIT na promoção da qualidade, segurança e integridade dos dados em toda a organização.

Várias práticas de governança e gerenciamento de dados são estabelecidas pelo COBIT para ajudar a garantir a conformidade com políticas e regulamentações, melhorar a eficiência operacional e apoiar a tomada de decisões informadas. Uma das áreas-chave do COBIT relacionadas à governança de dados é a gestão de informações, que aborda a coleta, armazenamento, processamento e disseminação

de informações em toda a organização (TUREL, LIU & BART, 2017). A importância de garantir que as informações sejam precisas, completas, atualizadas e acessíveis para apoiar a tomada de decisões e os processos de negócios é enfatizada pelo COBIT.

Influenciado pela crescente posição estratégica que a Tecnologia da Informação vem conquistando, o Modelo de Alinhamento Estratégico (Strategic Alignment Model - SAM) foi desenvolvido justamente para aproximar as questões estratégicas de negócio de uma organização juntamente com a tecnologia da informação e processos. O modelo, conforme ilustrado na Figura 6, está pautado em quatro domínios fundamentais: estratégia de negócios, estratégia de tecnologia da informação, infraestrutura e processos organizacionais e infraestrutura e processos de tecnologia da informação (HENDERSON, VENKATRAMAN, 1999).

Figura 6 - Modelo de Alinhamento Estratégico



Fonte: Henderson, Venkatraman (1993, p. 476)

O Modelo de Alinhamento Estratégico proposto por Henderson e Venkatraman consegue ter seus domínios abstraídos para que se encaixem no gerenciamento de dados. A informação é associada à estratégia de negócios e ao uso operacional dos dados, enquanto os dados estão associados à tecnologia da informação e aos processos que suportam o gerenciamento físico de sistemas que

tornam os dados acessíveis para uso. Mesmo com a representação de alto nível do modelo, é possível entender os fatores organizacionais que influenciam as decisões sobre dados e gerenciamento de dados (DMBOK 2 ed., 2017).

Para a implantação de métodos de gestão, é de grande importância ter um caminho sólido para replicar melhores práticas à luz de conceitos bem definidos sobre o que se deseja gerenciar. Nesse sentido, guias de conhecimento foram sendo criados ao longo do tempo em seus determinados domínios para auxiliar na implantação de modelos de gestão mais eficientes e eficazes.

Os corpos de conhecimento, como o Project Management Body of Knowledge (PMBOK) e o Data Management Body of Knowledge (DMBOK), têm suas raízes no final do século XX e início do século XXI. Estes corpos de conhecimento foram criados em resposta à crescente necessidade de padronização e melhores práticas em diversas áreas, como gerenciamento de projetos e gestão de dados. O PMBOK, desenvolvido pelo Project Management Institute (PMI), é um conjunto de diretrizes e práticas amplamente reconhecido no campo do gerenciamento de projetos. Sua história começou em 1987, quando um relatório chamado "Ethics, Standards, and Accreditation Committee Final Report" foi publicado pelo PMI, o que se tornaria a base para o PMBOK. Em 1996, foi lançada pelo PMI a primeira edição do PMBOK Guide, estabelecendo um conjunto de padrões e melhores práticas para o gerenciamento de projetos. Desde então, o PMBOK Guide passou por várias atualizações e revisões, sendo a sexta edição publicada em 2017. O PMBOK Guide é considerado um recurso valioso para profissionais e estudantes de gerenciamento de projetos, oferecendo um quadro abrangente e flexível para planejar, executar e controlar projetos em diversos setores e contextos (GUIA PMBOK, 2017).

Outro corpo de conhecimento relevante, o Business Process Management Common Body of Knowledge, mais conhecido como BPM CBOK, é um conjunto consolidado de princípios e práticas gerenciais que destaca a gestão de processos de negócios. Este corpo de conhecimento é considerado um guia fundamental para profissionais e organizações que buscam compreender, implementar e gerir seus processos de negócios de maneira mais eficiente e efetiva. O BPM CBOK é organizado em nove áreas de conhecimento que abrangem desde a visão estratégica e o alinhamento dos processos com os objetivos do negócio, até a



modelagem, análise, desenho, medição, controle e transformação dos processos. Cada uma dessas áreas reflete aspectos críticos na gestão de processos, proporcionando um arcabouço holístico para compreender e implementar o BPM na prática (BPM CBOK, 2013).

O DMBOK, elemento principal desta pesquisa, foi desenvolvido pela Data Management Association (DAMA) para o campo da gestão de dados. O trabalho no DMBOK começou pela DAMA em 2005, com o objetivo de criar um conjunto de diretrizes e práticas que cobrissem todos os aspectos da gestão de dados. Em 2009, foi lançada pela DAMA a primeira edição do DMBOK, fornecendo uma estrutura abrangente e holística para a gestão de dados, incluindo áreas como governança de dados, arquitetura de dados, qualidade de dados e segurança de dados. A segunda edição do DMBOK foi publicada em 2017, refletindo as mudanças e desenvolvimentos no campo da gestão de dados (DMBOK 2 ed., 2017).

Para a união dos aspectos constituintes da gestão de dados, o modelo do DMBOK começa a fazer sentido como referência para a adoção e o estudo das práticas de gerenciamento de dados pelas organizações. Uma das principais razões para utilizar o DMBOK como modelo de estudo é a sua abordagem abrangente e holística para a gestão de dados.

Outros corpos de conhecimento semelhantes ao PMBOK e DMBOK incluem o IT Infrastructure Library (ITIL) para gerenciamento de serviços de TI, o Business Analysis Body of Knowledge (BABOK) para análise de negócios e o Information Technology Body of Knowledge (ITBoK) para gerenciamento de tecnologia da informação. Estes guias foram desenvolvidos para atender às necessidades específicas de cada campo e fornecer um conjunto de práticas e diretrizes padronizadas que podem ser adaptadas a diferentes contextos e organizações.

## **2.8 Corpo de Conhecimento em Gestão de Dados DMBOK**

A gestão de dados requer uma série de conhecimentos agregados oriundos de várias áreas, tais como administração, tecnologia da informação, projetos e até mesmo gestão de pessoas. Inspirado em modelos de gestão e práticas de mercado,

o corpo de conhecimento do DAMA (DMBOK) engloba conceitos, processos, atividades, responsabilidades e prepara uma organização para lidar com a gestão de dados de forma prática.

O DMBOK é uma compilação de processos e áreas de conhecimento geralmente aceitos como melhores práticas na disciplina de Gerenciamento de Dados. No contexto do DMBOK, a gestão de dados é um termo abrangente que descreve os processos usados para planejar, especificar, habilitar, criar, adquirir, manter, usar, arquivar, recuperar, controlar e purificar dados (CUPOLI, EARLEY, HENDERSON, 2014).

Glaucia et al. apontam o DMBOK para uma utilização corporativa: (2019, p.19, tradução nossa):

O contexto de dados corporativos tem uma abordagem econômica. O alto desempenho na manipulação de dados geralmente é importante, mesmo quando há alta taxa de simultaneidade de acesso, pois as empresas lidam com processos de negócios que, geralmente, visam o lucro. Assim, o perfil de princípios de gerenciamento de dados definido no DMBOK está relacionado a uma perspectiva empresarial, alinhada aos requisitos do negócio.

No DMBOK, todos os aspectos da gestão de dados são apresentados, incluindo planejamento de dados, governança de dados, arquitetura de dados, design de dados, desenvolvimento de dados, operações de dados, segurança de dados, qualidade de dados e suporte à tomada de decisão. Essa abordagem integrada permite que profissionais e pesquisadores identifiquem lacunas e áreas de melhoria em suas práticas de gestão de dados. Além disso, o DMBOK é valorizado pela sua flexibilidade e adaptabilidade, que permite a personalização do modelo às necessidades específicas de diferentes organizações e projetos. O DMBOK pode ser adaptado para atender aos requisitos de diferentes setores, tamanhos de organização e contextos regulatórios, sendo considerado um modelo valioso para uma ampla gama de situações e ambientes.

Ademais, o DMBOK é útil para estabelecer um vocabulário comum e padrões de melhores práticas no campo da gestão de dados. No DMBOK, uma linguagem unificada é fornecida, bem como um conjunto de diretrizes que facilitam a comunicação entre os profissionais de gestão de dados e auxiliam na promoção da

consistência nas práticas de gestão de dados em toda a organização. O DMBOK também é reconhecido por fornecer uma estrutura para a avaliação e melhoria contínua da gestão de dados. Através do DMBOK, as organizações podem avaliar suas práticas de gestão de dados em relação às melhores práticas do setor, identificar áreas de melhoria e desenvolver estratégias para aprimorar seus processos e resultados.

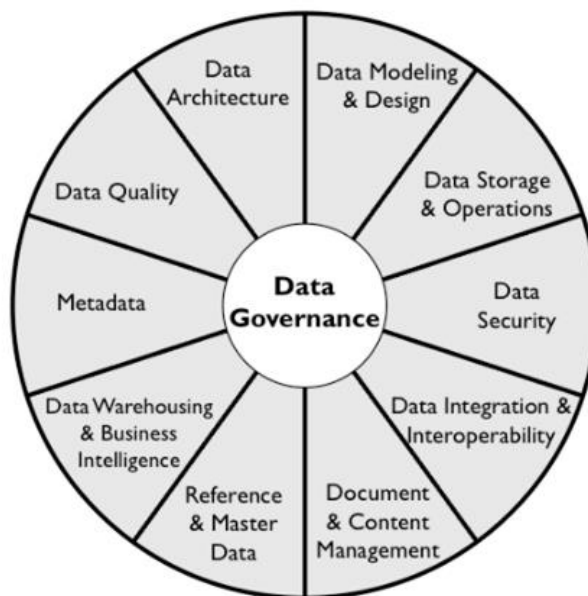
Na data da realização desta pesquisa, a segunda versão do DMBOK (DMBOK 2nd) está em vigência. As mudanças na Roda do DAMA entre a primeira versão lançada em 2009 (Figura 7) e a última versão lançada em 2017 (Figura 8) são notáveis. As disciplinas foram redefinidas e ajustadas ao contexto atual, sendo influenciadas por novas tecnologias e o uso de Big Data (dados semiestruturados e não estruturados). A importância da ética de dados foi enfatizada e a segurança em torno do uso dos dados foi reforçada. A Integração e Interoperabilidade de Dados foram adicionadas como uma nova área de conhecimento, a Governança agora é discutida tanto como uma função quanto uma área de conhecimento, e a Ética de Dados é trazida como um capítulo à parte, pois abrange a discussão em todo o processo de gerenciamento de dados. Capítulos separados são apresentados com os temas de Big Data, Ciência de Dados e Avaliação de Maturidade em Gerenciamento de Dados, expandindo assim as áreas de conhecimento fundamentais do arcabouço do DMBOK (DMBOK 2 ed., 2017).

Figura 7 - Roda do DAMA Primeira Edição do DMBOK



Fonte: DMBOK (2009, p.7)

Figura 8 - Roda do DAMA Segunda Edição do DMBOK



Fonte: DAMA 2 ed. (2017, p.67)

O corpo de conhecimento do DAMA foi constituído e estruturado de forma interdisciplinar para contemplar diversas áreas da gestão que envolvem dados. O DMBOK segue um arcabouço formado por três visões, que são:

- DAMA Wheel: “A Roda do DAMA”
- Environmental Factors Hexagon: o “Hexágono dos Fatores Ambientais
- Knowledge Area Context Diagram: o “Diagrama de Contexto da Área de Conhecimento”

The DAMA Wheel (Figura 8), traduzido livremente como "A Roda do DAMA", define as áreas do corpo de conhecimento do DMBOK, e proporciona uma visão clara das disciplinas que envolvem o gerenciamento de dados e que têm a capacidade de serem implementadas em momentos diferentes, respeitando os verdadeiros requisitos de uma organização (DMBOK 2 ed., 2017).

Conforme representado na Roda do DAMA, o corpo de conhecimento do DMBOK é estruturado em onze áreas, que são (DMBOK 2 ed., 2017):

1. Governança de dados: fornece direção e supervisão para gestão de dados, estabelecendo um sistema de direitos de decisão sobre os dados que respondem às necessidades da empresa.
2. Arquitetura de Dados: define o plano de gerenciamento de dados ativos, alinhado com a estratégia organizacional para estabelecer requisitos de dados estratégicos e projetos para atender a esses requisitos
3. Modelagem e Design de Dados: descobrir, analisar, representar e comunicar requisitos de dados em uma forma precisa, chamada modelo de dados.
4. Armazenamento e Operações de Dados: design, implementação e suporte de dados armazenados para maximizar seu valor. Fornecem suporte durante todo o ciclo de vida dos dados desde o planejamento até a eliminação de dados.
5. Segurança de Dados: garante que a privacidade e a confidencialidade dos dados sejam mantidas, que os dados não sejam violados e que os dados sejam acessados adequadamente.
6. Integração de dados e interoperabilidade: movimentação e

consolidação de dados entre armazenamentos de dados, aplicativos e organizações.

7. Gerenciamento de Documentos e Conteúdos: planejamento, implementação e atividades de controle usadas para gerenciar o ciclo de vida de dados e informações encontrados em uma série de mídia não estruturada, especialmente documentos necessários para requisitos de conformidade legal e regulamentar.
8. Dados mestre e de referência: reconciliação contínua e manutenção de dados compartilhados essenciais para permitir uso consistente em todos os sistemas da forma mais precisa, oportuna e versão relevante da verdade sobre entidades empresariais essenciais.
9. Data Warehousing e Business Intelligence: planejamento, implementação e controle para gerenciar dados de apoio à decisão e permitir que os trabalhadores do conhecimento obtenham valor dos dados por meio de análises e relatórios.
10. Metadados: planejamento, implementação e controle atividades para permitir o acesso a metadados integrados de alta qualidade, incluindo definições, modelos, fluxos de dados e outras informações crítico para entender os dados e os sistemas através dos quais eles é criado, mantido e acessado.
11. Qualidade de Dados: planejamento e implementação de técnicas de gestão da qualidade para medir, avaliar e melhorar a adequação dos dados para uso dentro de uma organização.

A Governança de Dados está no centro da roda, representando a gestão das atividades e constituindo um requisito para a orquestração e o diálogo entre as áreas de conhecimento para uma distribuição adequada de suas funções. As outras áreas de conhecimento estão distribuídas em torno da roda; todas as áreas são partes necessárias de um gerenciamento de dados maduro e consistente, mas podem ser implementadas em diferentes momentos, dependendo da maturidade e requisitos de uma organização.

The Environmental Factors Hexagon, traduzido livremente como "Hexágono dos Fatores Ambientais", apresenta as relações entre pessoas, processos e

tecnologia (Figura 9). Nesta visão, os objetivos e princípios estão no centro e orientam como as pessoas devem executar as atividades e utilizar as ferramentas para o gerenciamento de dados (DMBOK 2 ed., 2017). É apresentada uma relação entre pessoas, processos e tecnologia pelo Hexágono de Fatores Ambientais, para fornecer um entendimento prévio do Diagrama de Contexto da Área de Conhecimento. Nesta visão, os objetivos e princípios são colocados no centro do hexágono, fornecendo um guia de como as atividades devem ser executadas pelas pessoas e de como as ferramentas necessárias devem ser utilizadas para um gerenciamento de dados bem-sucedido.

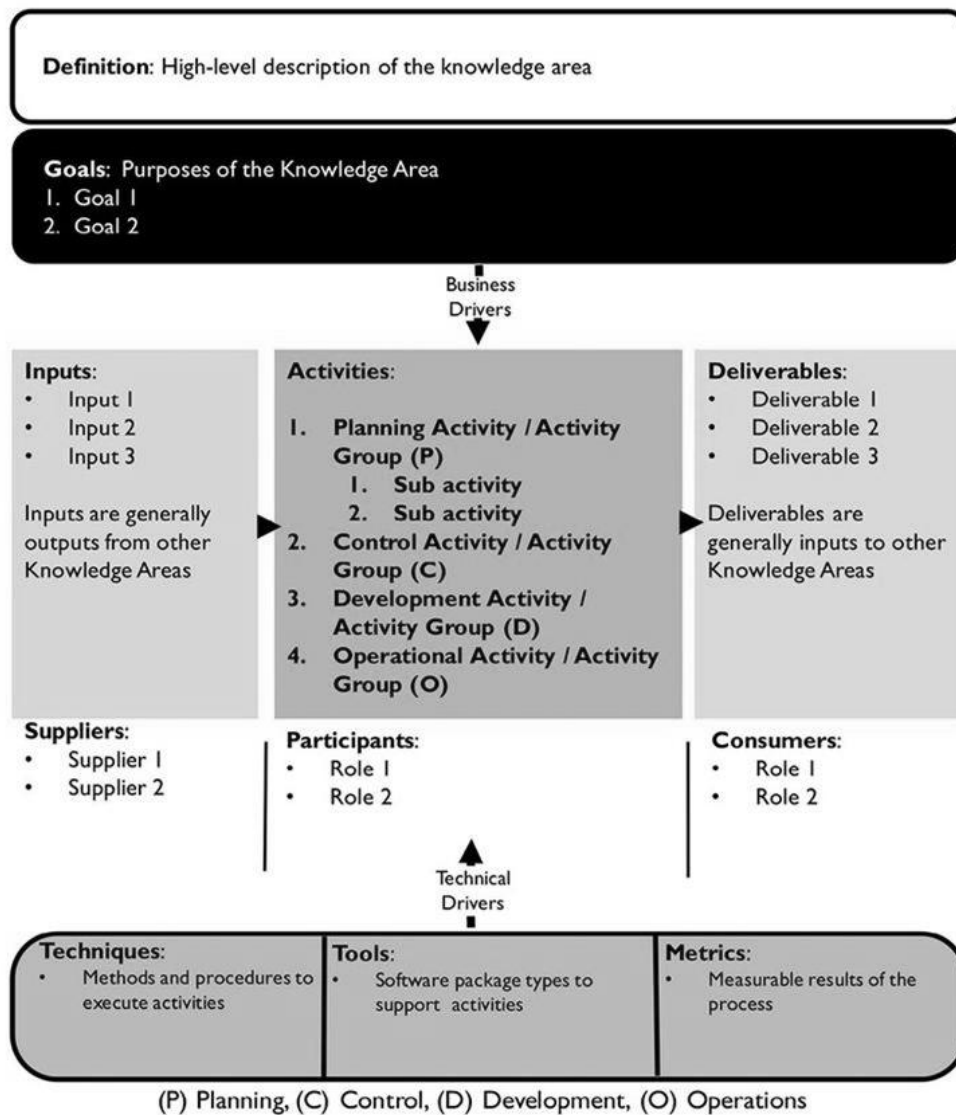
Figura 9 - Representação Hexágono dos Fatores Ambientais



Fonte: DAMA 2 ed. (2017, p.36)

"The Knowledge Area Context Diagram", traduzido livremente como "Diagrama de Contexto da Área de Conhecimento", descreve com riqueza de detalhes todas as atividades e processos que cada área do conhecimento possui (Figura 10).

Figura 10 - Diagrama de Contexto da Área de Conhecimento



Fonte: DAMA 2 ed. (2017, p.37)

O detalhamento das atividades de gestão de dados de cada área do DMBOK é explicitado no Diagrama de Contexto da Área de Conhecimento. As atividades se concentram no centro do diagrama, onde recebem ENTRADAS e geram SAÍDAS, possuindo também subatividades, tarefas e etapas. O diagrama descreve com detalhes as áreas de conhecimento, incluindo pessoas, processos e tecnologia com detalhes. Essa visão é baseada no SIPOC, sigla em inglês que significa: Suppliers, Inputs, Processes, Outputs, and Consumers (Fornecedor, Entradas, Processos, Saídas, e Consumidores).



Saídas e Consumidores). No Diagrama de Contexto, as atividades são colocadas no centro e, a partir delas, entregáveis são produzidos que atendem aos requisitos das partes interessadas. Cada área terá seu próprio diagrama de contexto, começando com suas definições e objetivos. O Quadro 5 apresenta os principais componentes do diagrama de contexto.

## Quadro 5 - Componentes do Diagrama de Contexto

1. Definição: Explicação de sua finalidade e do porquê de sua existência.		
2. Objetivos: Propósitos da Área de Conhecimento em questão e seus princípios fundamentais que guiam a performance das atividades em cada área.		
3. Entradas: As entradas são coisas tangíveis que cada Área de Conhecimento necessita para iniciar uma atividade. Muitas atividades possuem as mesmas entradas inclusive.	4. Atividades: Ações e tarefas que vão de encontro com os objetivos da Área de Conhecimento. Algumas atividades são descritas dentro de termos como subatividades, tarefas e passos. Atividades são classificadas em quatro categorias: Planejamento, Desenvolvimento, Operação e Controle.  a. Atividades de Planejamento (P) É um conjunto de estratégias e táticas para cumprir os objetivos da gestão de dados.  b. Atividades de Desenvolvimento (D) São atividades que são organizadas em torno do ciclo de vida do sistema cuja sigla em inglês é SDLC (analysis, design, build, preparation, and deployment).  c. Atividades de Controle (C) São atividades que garantem o andamento sobre a qualidade e integridade dos dados, confiabilidade e segurança de sistemas através dos quais dados são acessados e utilizados.  d. Atividades de Operação (O) São atividades que suportam o uso, manutenção e melhorias de sistemas e processos através dos quais os dados são acessados	5. Entregáveis Os entregáveis são saídas das atividades dentro de cada Área de Conhecimento. Os entregáveis devem ser um fim em si mesmo ou entradas dentro de outras atividades. Muitos entregáveis primários são criados por múltiplas funções.
6. Fornecedores: Pessoas que fornecem insumos para apoiar as atividades	7. Participantes: Pessoas que de alguma forma estão envolvidas nas atividades da área	8. Consumidores: Todo aquele que de alguma consome e se beneficia das atividades da gestão de dados
9. Técnicas	10. Ferramentas	11. Métricas
Métodos utilizados para a realização das atividades	Softwares e demais tecnologias que viabilizam os objetivos da área	Indicadores de desempenho, qualidade, eficiência ou outros determinados pela área em questão
12. Papéis e responsabilidades: descreve como indivíduos e o grupo contribuem para atividades dentro da Área de Conhecimento. Papéis são descritos conceitualmente com foco em grupos de responsabilidades que são requisitos em muitas organizações. Os papéis para cada indivíduo são definidos em termos de habilidades e qualificação requeridas. O Skills Framework for the Information Age (SFIA) foi utilizado para auxiliar na atribuição dos papéis e muitos desses papéis terão funções cruzadas.		

Fonte: Elaborado pelo autor, 2023

Para materializar o Diagrama de Contexto, a Figura 11 é apresentada, que exibe o diagrama na prática, utilizando como exemplo a área de Governança de Dados do DMBOK. Esse diagrama resume o que é apresentado na seção de DMBOK referente à área de Governança de Dados. Por meio desse diagrama, observa-se que a área é apresentada através de: sua definição, seus objetivos, atividades desenvolvidas, entradas e saídas das atividades, entregáveis, fornecedores, participantes, consumidores, técnicas, ferramentas e métricas. É seguida essa estrutura e um diagrama de contexto para todas as 11 áreas do DMBOK.

Figura 11 - Diagrama de Contexto da Área de Governança de Dados



### 3 METODOLOGIA

O objetivo desta pesquisa é analisar a presença de características da preservação digital e da curadoria digital no modelo de gestão de dados proposto pelo DAMA através do corpo de conhecimento do DMBOK.

Com base na literatura em preservação digital e curadoria digital, com foco nos modelos OAIS e CVCD/DCC, apresentados no Capítulo 2, o corpo de conhecimento do DMBOK é analisado. Diante da característica exposta do trabalho de pesquisa, uma abordagem qualitativa é aplicada e, do ponto de vista de seus objetivos, é definida como exploratória.

Quanto à natureza da pesquisa, pode ser classificada como pesquisa aplicada, pois "objetiva gerar conhecimentos para aplicação prática e dirigidos à solução de problemas específicos. Envolve verdades e interesses locais" (SILVA & MENEZES, 2005, p. 20). Apesar da dinâmica implementação de modelos de gestão de dados e de curadoria e preservação digital em organizações, esta pesquisa visa recuperar conceitos importantes através de referências sólidas para que, em futuras pesquisas, ocorra mais enriquecimento nas análises entre campos interdisciplinares envolvendo dados, informação e conhecimento.

A pesquisa então se beneficiará de materiais já publicados, envolvendo artigos científicos, livros, periódicos relevantes para as áreas de conhecimento envolvidas e de fontes descritas e acessíveis via internet.

#### 3.1 Planejamento da Pesquisa

Para alicerçar o bom andamento da pesquisa e obedecer aos critérios de coerência, consistência, originalidade e objetivação (SILVA & MENEZES, 2005), procedimentos técnicos e teóricos foram executados. De acordo com Mirian Goldemberg, a pesquisa é caracterizada por (2004, p.105):

Pesquisa é a construção de conhecimento original, de acordo com certas exigências científicas. É um trabalho de produção de conhecimento sistemático, não meramente repetitivo mas produtivo, que faz avançar a área de conhecimento a qual se dedica. Uma pesquisa necessita o cumprimento de três requisitos:

- a. a existência de uma pergunta que se deseja responder;
- b. a elaboração de um conjunto de passos que permitam chegar à resposta;
- c. a indicação do grau de confiabilidade na resposta obtida.

Seguindo os requisitos expostos por Goldemberg, as instruções norteadoras para a análise exploratória da pesquisa bibliográfica foram criadas, sendo assim divididas em atividades de exploração contendo o procedimento de pesquisa, os objetos a serem identificados e a descrição deles.

### 3.1.1 Percurso Metodológico do Objetivo A

O objetivo "Relacionar as áreas do DMBOK com os modelos de preservação e curadoria digital OAIS e CVCD/DCC" é desenvolvido por meio de pesquisa bibliográfica envolvendo DMBOK na sua segunda edição, OAIS e CVCD/DCC. O percurso metodológico desse objetivo é desenvolvido estabelecendo-se relações entre estes três modelos com base nos seus objetivos e conceitos. Para tal, a coleta e análise dos dados são determinadas pelos instrumentos dos Quadro 6 e Quadro 7.

Quadro 6 - Coleta e questões de Objetivos de DMBOK, OAIS e CVCD/DCC

Modelo	Objetivo
DMBOK	(Objetivos de DMBOK)
OAIS	(Objetivos de OAIS).
CVCD/DCC	(Objetivos de CVCD/DCC)
Análise conduzida por meio das seguintes questões norteadoras: <ul style="list-style-type: none"> <li>• Os objetivos de DMBOK envolvem direta ou indiretamente objetivos de OAIS?</li> <li>• Os objetivos de DMBOK envolvem direta ou indiretamente objetivos de CVCD/DCC?</li> </ul>	

Fonte: Elaborado pelo autor, 2023

Quadro 7 - Coleta e questões de Conceitos de DMBOK, OAIS e CVCD/DCC

Modelo	Conceitos
DMBOK	(Conceitos de DMBOK)
OAIS	(Conceitos de OAIS).
CVCD/DCC	(Conceitos de CVCD/DCC)
Análise conduzida por meio das seguintes questões norteadoras: <ul style="list-style-type: none"> <li>• DMBOK e CVCD/DCC compartilham conceitos? Quais?</li> <li>• DMBOK e OAIS compartilham conceitos? Quais?</li> </ul>	

Fonte: Elaborado pelo autor, 2023

### 3.1.2 Percurso Metodológico do Objetivo B

O objetivo é "Analisar a relação de cada área de conhecimento do DMBOK com os modelos de preservação digital e curadoria digital". Cada área de conhecimento do DMBOK é analisada em relação aos seus objetivos e atividades, com base nos modelos OAIS e CVCD/DCC. Os Quadro 8 e Quadro 9 apresentam as estratégias de coleta e análise de cada área em relação aos objetivos e atividades, respectivamente.

Quadro 8 - Coleta e questões de análise da área quanto aos Objetivos

Modelo	Entidade/Área	Objetivos
DMBOK	(Área em estudo)	(objetivos da área de DMBOK em estudo)
OAIS	(Entidade Funcional de OAIS relacionada com a área DMBOK em estudo pelos objetivos)	(objetivos da entidade funcional de OAIS)
	(Entidade Funcional de OAIS relacionada com a área DMBOK em estudo pelos objetivos)	(objetivos da entidade funcional de OAIS)
CVCD/DCC	(Atividade de CDCD/DCC relacionada com a área DMBOK em estudo pelos objetivos)	(objetivos da entidade funcional de CVCD/DCC)
	(Atividade de CVCD/DCC relacionada com a área DMBOK em estudo pelos objetivos)	(objetivos da entidade funcional de CVCD/DCC)
<p>Análise conduzida por meio das seguintes questões norteadoras:</p> <ul style="list-style-type: none"> <li>Os objetivos da área não possuem nem deveriam possuir o envolvimento de aspectos da curadoria e preservação digital?</li> <li><b>OU</b> área deve ou deveria contemplar nos seus objetivos aspectos da curadoria e a preservação digital? Quais aspectos principais?</li> </ul>		

Fonte: Elaborado pelo autor, 2023

Quadro 9 - Coleta e questões de análise da área quanto as Atividades

<b>Modelo</b>	<b>Entidade/Área</b>	<b>Atividades</b>
DMBOK	(Área em estudo)	(atividades da área de DMBOK em estudo)
OAIS	(Entidade Funcional de OAIS relacionada com a área DMBOK em estudo pelas atividades)	(atividades da entidade funcional de OAIS)
	(Entidade Funcional de OAIS relacionada com a área DMBOK em estudo pelas atividades)	(atividades da entidade funcional de OAIS)
CVCD/DCC	(Atividade de CDCD/DCC relacionada com a área DMBOK em estudo pelas atividades)	(atividades da entidade funcional de CVCD/DCC)
	(Atividade de CVCD/DCC relacionada com a área DMBOK em estudo pelas atividades)	(atividades da entidade funcional de CVCD/DCC)
<p>Análise conduzida por meio das seguintes questões norteadoras:</p> <ul style="list-style-type: none"> <li>• Atividades de DMBOK envolvem atividades de OAIS?</li> <li>• Atividades de DMBOK envolvem atividades de CVCD/DCC?</li> </ul>		

Fonte: Elaborado pelo autor, 2023

### 3.1.3 Percurso Metodológico do Objetivo C

Em relação ao percurso metodológico do objetivo "Propor sugestões para aproximar a preservação digital e curadoria digital no Corpo de Conhecimento DMBOK", a partir dos resultados obtidos na análise realizada pelo objetivo "B", sugestões são propostas no sentido de aproximar a preservação e curadoria digital ao DMBOK.

## 3.2 Estratégia de Busca e Organização das Fontes Utilizadas

Para o desenvolvimento da pesquisa, cinco etapas foram desenvolvidas para a estruturação dos procedimentos técnicos, visando selecionar e organizar as fontes utilizadas pela pesquisa:

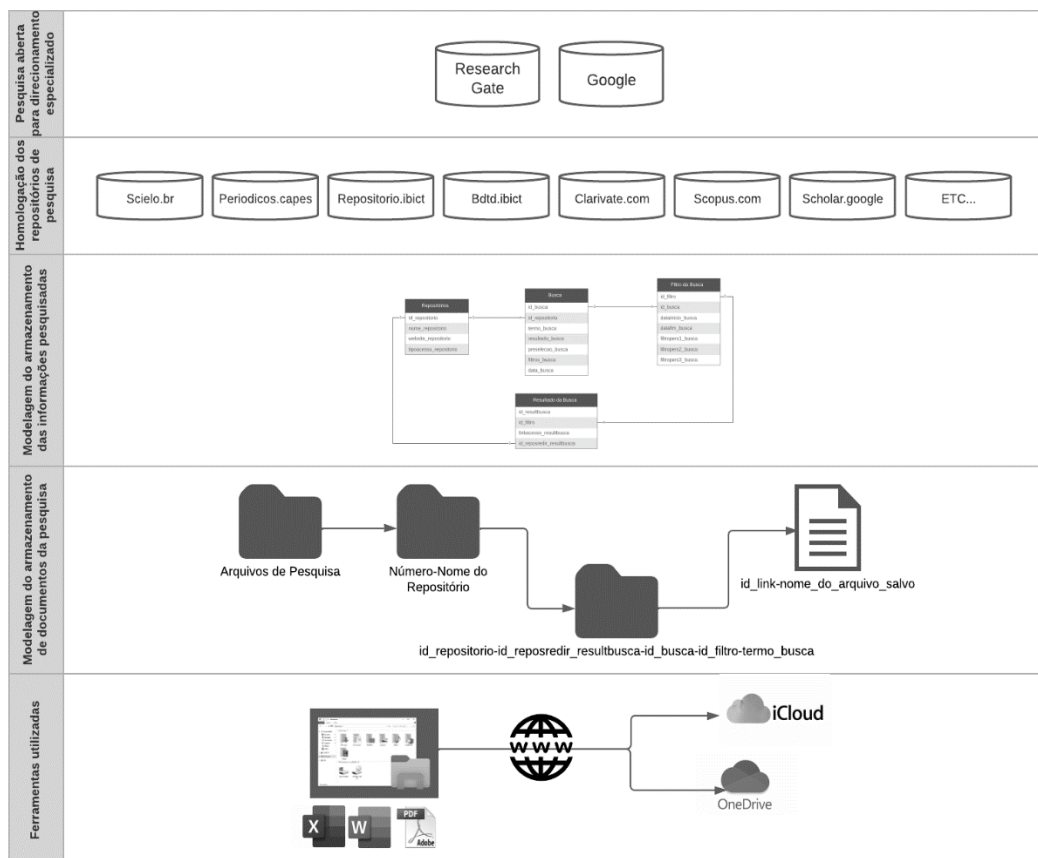
- Pesquisa aberta para direcionamento especializado: nesta etapa, fontes de pesquisa baseadas em redes sociais voltadas para a comunidade científica (como o Research Gate) e mecanismos de busca baseados em indexadores (como o Google) foram utilizadas. Essa primeira etapa foi inserida para direcionar e buscar por repositórios especializados, com o objetivo de explorar artigos, livros e documentos especializados e com notória relevância diante da comunidade científica e de domínios especializados;
- Homologação dos repositórios de pesquisa: nesta etapa, os repositórios que fizeram parte do processo de busca e filtragem por artigos científicos foram homologados. Ao longo da pesquisa, foram encontrados artigos citados por outros trabalhos, que também foram coletados e armazenados na sua íntegra para avaliação;
- Modelagem do armazenamento das informações pesquisadas: planejar a estrutura de armazenamento de informações se faz necessário diante da infinidade de ferramentas tecnológicas que facilmente mudam seu estado em função de atualizações diárias. Com isso, o intuito desta etapa foi estipular um modelo relacional de armazenamento de informações, servindo de fonte confiável de consulta pós-trabalho de pesquisa. A estrutura do modelo foi pensada para que todas as informações necessárias da pesquisa fossem armazenadas em um sistema gerenciador de banco de dados e para que seja passível de verificação do processo de pesquisa de forma pública;
- Modelagem do armazenamento de documentos da pesquisa: para garantir e reforçar o acesso posterior aos dados de pesquisa, foi criado um modelo estrutural e lógico para armazenar todos os documentos coletados oriundos dos processos de busca. Buscou-se seguir uma lógica hierárquica baseada na modelagem do armazenamento das informações pesquisadas;
- Ferramentas utilizadas: nesta etapa, as ferramentas utilizadas para os



procedimentos técnicos foram homologadas, sendo elas: Microsoft Word e Excel, uso de repositório de nuvem iCloud e Onedrive.

Como guia ilustrativo, foi criado um modelo baseado em aspectos da gestão de dados e assim, posto em prática um enxuto programa de gerenciamento de dados para os insumos da pesquisa, representado na Figura 12.

Figura 12 - Modelo gráfico de captação dos insumos da pesquisa



Fonte: Elaborado pelo autor, 2023

## **4 ANÁLISE DOS DADOS**

Esta seção apresenta as análises de dados entre os modelos DMBOK, OAIS e CVCD/DCC.

### **4.1 DMBOK e os Modelos OAIS e CVCD/DCC**

Esta seção apresenta a análise referente ao objetivo "Relacionar DMBOK com os modelos de preservação e curadoria digital OAIS e CVCD/DCC". Relações são estabelecidas entre o DMBOK e os modelos OAIS e CVCD/DCC com base nos objetivos e conceitos principais de cada modelo.

Para a análise da relação entre os objetivos dos três modelos, os objetivos de cada um dos modelos são reunidos no Quadro 10. A análise dos modelos é feita a seguir, com base nas seguintes questões norteadoras:

- Os objetivos de DMBOK envolvem direta ou indiretamente objetivos de OAIS?
- Os objetivos de DMBOK envolvem direta ou indiretamente objetivos de CVCD/DCC?

Quadro 10 - Objetivos de DMBOK, OAIS e CVCD/DCC

Modelo	Objetivos
DMBOK	<ul style="list-style-type: none"> <li>• Construir consenso para uma visão geral aplicável das áreas de conhecimento de gerenciamento de dados.</li> <li>• Fornecer definições padrão para áreas de conhecimento de gerenciamento de dados comumente usadas</li> <li>• Identificar princípios orientadores para o gerenciamento de dados.</li> <li>• Esclarecer o escopo e os limites das atividades de gerenciamento de dados.</li> <li>• Fornecer uma visão geral das boas práticas comumente aceitas, técnicas amplamente adotadas e abordagens alternativas significativas</li> <li>• Fornecer questões organizacionais e culturais comuns.</li> <li>• Identificar estratégias para análise de maturidade na gestão de dados.</li> <li>• Forneça recursos adicionais e material de referência para maior compreensão do gerenciamento de dados.</li> </ul>
OAIS	<ul style="list-style-type: none"> <li>• Fornece um quadro para a compreensão e uma maior sensibilização para os conceitos de repositório necessários para a preservação e o acesso a informações digitais a longo prazo;</li> <li>• fornece os conceitos necessários para que as organizações sejam participantes efetivos do processo de preservação;</li> <li>• fornece uma estrutura, incluindo terminologia e conceitos, para descrever e comparar arquiteturas e operações de repositórios existentes e futuros;</li> <li>• fornece uma estrutura para descrever e comparar diferentes estratégias e técnicas de preservação de longo prazo;</li> <li>• fornece uma base para comparar os modelos de dados de informações digitais preservadas pelos repositórios e para discutir como os modelos de dados e as informações subjacentes podem mudar ao longo do tempo;</li> <li>• fornece uma estrutura que pode ser expandida por outros esforços para abranger a preservação de longo prazo de informações que NÃO estão em formato digital (por exemplo, mídia física e amostras físicas);</li> <li>• expande o consenso sobre os elementos e processos para a informação digital de longo prazo</li> </ul>
CVCD/DCC	<p>Fornecer uma visão geral gráfica de alto nível dos estágios necessários para a curadoria bem-sucedida e a preservação dos dados da conceituação ou recebimento inicial.</p> <p>O modelo pode ser usado para planejar atividades dentro de uma organização ou consórcio para garantir que todas as etapas necessárias sejam realizadas, cada uma na sequência correta.</p> <p>O modelo permite que a funcionalidade granular seja mapeada em relação a ele; para definir papéis e responsabilidades e construir uma estrutura de padrões e tecnologias para implementar.</p> <p>Ele pode ajudar no processo de identificar etapas adicionais que podem ser necessárias, ou ações que não são exigidas por determinadas situações ou disciplinas, e garantir que os processos e políticas sejam adequadamente documentados</p>

Ao analisar os objetivos do DMBOK, observa-se que o corpo de conhecimento é direcionado para gestão de dados, trazendo consensos, definindo padrões, identificando princípios, estabelecendo escopo, fornecendo melhores práticas e identificando estratégias para análise de maturidade no contexto de uma organização. Também é abordado o contexto cultural e de uso de dados em geral para preparar uma organização para trabalhar com dados, explorando as áreas de conhecimento que envolvem o ecossistema do gerenciamento de dados.

Por outro lado, os objetivos do OAIS têm como foco a preservação digital em um repositório, sensibilizando, trazendo conceitos, estruturas, modelos de dados, processos, estratégias e técnicas de preservação. O modelo CVCD/DCC se concentra na curadoria digital, fornecendo uma visão gráfica de alto nível dos estágios necessários para uma curadoria bem-sucedida, visando o planejamento de atividades de curadoria dentro de uma organização em etapas, definição de responsabilidades, papéis e padrões.

Observa-se que CVCD/DCC e OAIS visam à preservação e à curadoria digital, sendo que o OAIS enfatiza repositórios e o DCC no ciclo de vida da curadoria em uma organização. Os objetivos do DMBOK são distintos dos do OAIS e do CVCD/DCC, pois o foco do DMBOK não é a preservação digital, mas sim a gestão de dados de seu corpo de conhecimento. Não se observa nos objetivos do DMBOK aspectos de preservação e curadoria digital, entretanto, os objetivos do DMBOK podem envolver indiretamente os objetivos do OAIS e do CVCD/DCC. A gestão de dados envolve a curadoria e a preservação digital e os modelos OAIS e CVCD/DCC podem complementar o DMBOK e suas áreas com aspectos de curadoria e preservação digital.

Para a análise da relação entre os conceitos dos três modelos, o Quadro 11 apresenta os conceitos de cada um dos modelos. Os conceitos dos modelos são analisados com base nas seguintes questões norteadoras:

- DMBOK e CVCD/DCC compartilham conceitos?
- DMBOK e OAIS compartilham conceitos?

Quadro 11 - Conceitos de DMBOK, OAIS e CVCD/DCC

Modelo	Conceitos
DMBOK	<ul style="list-style-type: none"> <li>• Governança de Dados</li> <li>• Arquitetura de Dados</li> <li>• Modelagem, de Dados Armazenamento de Dados</li> <li>• Segurança de Dados</li> <li>• Integração de Dados</li> <li>• Gerenciamento de documentos</li> <li>• Dados mestre/referência</li> <li>• Business Intelligence,</li> <li>• Metadados</li> <li>• Qualidade de Dados</li> </ul>
OAIS	<ul style="list-style-type: none"> <li>• Ambiente do Repositório, Produtor, Consumidor, Gestor</li> <li>• Pacotes de Informação de um Repositório</li> <li>• Funções de um Repositório</li> </ul>
CVCD/DCC	<ul style="list-style-type: none"> <li>• Curadoria</li> <li>• Preservação</li> <li>• Descrição e Representação da Informação</li> <li>• Ciclo de Vida da Curadoria Digital</li> </ul>

Fonte: Elaborado pelo autor, 2023

Os principais conceitos apresentados pelo DMBOK são aqueles que delimitam as suas áreas de conhecimento (Quadro 9). Identifica-se que esses conceitos abordam dados na perspectiva da sua Governança, Arquitetura, Modelagem, Dados mestre, Armazenamento, Segurança, Integração, Metadados, Qualidade, bem como a Gestão de Documentos e Business Intelligence e Data Warehousing. Os principais conceitos apresentados pelo OAIS são aqueles que envolvem suas especificações de ambiente, estrutura de informação e de componentes funcionais. Identifica-se que esses conceitos abordam um repositório digital na perspectiva de seus produtores, consumidores, gestores, dos pacotes de informação, e das funções de ingestão, acesso, armazenamento, planejamento da preservação digital. Por sua vez, os principais conceitos apresentados pelo CVCD/DCC são aqueles que envolvem a curadoria digital em uma organização na perspectiva de um ciclo de ações: representar informação, planejar a preservação, monitorar comunidade, preservar, curar, conceitualizar, criar, avaliar, selecionar, ingerir, armazenar, preservar, acessar, usar, reusar, transformar, eliminar.

Identifica-se que OAIS e DCC compartilham muitos conceitos, com enfoques

diferentes: OAIIS aplicado à repositório e DCC aplicado no ciclo de vida da curadoria digital em uma organização.

Os conceitos do DMBOK, voltados à gestão de dados, não enfatizam diretamente a preservação a longo prazo desses dados. Entretanto, é identificado que os conceitos do DMBOK podem envolver conceitos de preservação e curadoria digital na perspectiva de se ter uma governança de dados a longo prazo, desenvolver arquitetura, modelagem, dados mestres, segurança, armazenamento, metadados, integração de dados visando sua preservação a longo prazo.

## **4.2 As Áreas de Conhecimento do DMBOK e Preservação e Curadoria Digital**

Esta seção apresenta a análise referente ao objetivo "Analisar a relação de cada área de conhecimento do DMBOK com a preservação e curadoria digital" à luz dos conceitos e objetivos envolvidos. Nesta seção, cada área de conhecimento do DMBOK é analisada a partir de seus objetivos e atividades, com base nos modelos OAIIS e CVCD/DCC.

### **4.2.1 Governança de Dados**

Esta seção tem como foco a análise da relação da área de conhecimento Governança de Dados com a preservação e curadoria digital, levando em consideração seus objetivos e atividades. A base de análise são os objetivos e as atividades dos modelos OAIIS e CVCD/DCC em relação ao DMBOK. Os objetivos da área de Governança de Dados, juntamente com as entidades do OAIIS e as ações do CVCD/DCC que se relacionam com esses objetivos, são apresentados no Quadro 12. O Quadro 13 detalha as atividades da área de Governança de Dados, abrangendo as entidades do OAIIS e as ações do CVCD/DCC cujas atividades têm relação com a área do DMBOK em questão.

Quadro 12 - Entidades Funcionais de OAIS e Atividades de CVCD/DCC relacionadas a Governança de Dados pelos Objetivos

Modelo	Área/Entidade/Ações	Objetivos
DMBOK	Governança	Habilitar uma organização para gerenciar dados como ativos corporativos Definir, aprovar, comunicar e implementar princípios, políticas, procedimentos, métricas, ferramentas e responsabilidades para a gestão de dados Monitorar e direcionar políticas de conformidade, uso de dados e gerenciar as atividades
OAIS	Entidade Funcional Planejamento da Preservação	monitorar ambiente e prover planos e recomendações que garantam que a informação permaneça acessível e usável ao longo do tempo
CVCD/DCC	Curar e Preservar	Estar ciente e realizar ações de gestão e administrativas planejadas para promover a curadoria e a preservação ao longo do ciclo de vida da curadoria
	Planejamento de Preservação	Planejar a gestão e administração de todas as ações do ciclo de vida da curadoria
	Vigilância e Participação da Comunidade Designada	Manter uma vigilância sobre as atividades apropriadas da comunidade Participar do desenvolvimento de padrões compartilhados, ferramentas e software adequado

Fonte: Elaborado pelo autor, 2023

Quadro 13 - Entidades Funcionais de OAIS e Atividades de CVCD/DCC relacionadas a Governança de Dados pelas Atividades

Modelo	Área/Entidade/Ações	Atividades	Subatividades
DMBOK	Governança	Definir a Governança de Dados para a Organização	Desenvolver Estratégia de Governança de Dados
			Realizar avaliação de prontidão
			Realizar descoberta e alinhamento de negócios
			Desenvolver pontos de contato corporativos
		Definir a estratégia de Governança de Dados	Definir a Estrutura de Operação da Governança de Dados
			Desenvolver metas, princípios e políticas
			Subscrever projetos de gestão de dados
			Engajar Gestão de Mudança
			Engajar-se da Gestão de Problemas
			Avaliar requisitos regulatórios e de conformidade
		Implementar Governança de Dados	Patrocinar padrões e procedimentos de dados
			Desenvolver um glossário de negócios
			Coordenar em conjunto com grupo

Modelo	Área/Entidade/Ações	Atividades	Subatividades
			de arquitetura
			Avaliar os ativos de dados do patrocinador
		Incorporar a Governança de Dados	
OAIS	Entidade Funcional de Planejamento de Preservação		Monitorar Comunidade
			Monitorar Tecnologia
			Desenvolver estratégias e planos de preservação
CVCD/DCC	Curar e Preservar		Planejamento de Preservação
			Vigilância e Participação da Comunidade Designada

Fonte: Elaborado pelo autor, 2023

A função da Governança de Dados no corpo de conhecimento do DMBOK é assegurar que as demais áreas do modelo de gerenciamento de dados sejam aplicadas de acordo com as políticas, acordos e melhores práticas estabelecidas. Termos como estratégia, políticas, padrões e qualidade, supervisão, conformidade, gerenciamento de problemas, projetos de gerenciamento de dados e avaliação de ativos de dados compõem o escopo desta área de conhecimento. Nos objetivos apresentados no Quadro 12, nota-se a preocupação em habilitar, monitorar, direcionar, definir, aprovar, comunicar e implementar princípios, políticas, procedimentos, métricas, ferramentas e responsabilidades para a gestão de dados. A Governança de Dados aparece como uma entidade conhecedora dos direcionadores para qualquer questão que surja com relação a um Programa de Gerenciamento de Dados, funcionando como elo entre a estratégia de negócios e a condução de atividades voltadas ao cumprimento dos objetivos traçados em um plano de uso de dados de uma empresa.

Para que projetos de dados obtenham sucesso, é fundamental a existência de uma cultura de uso de dados entre os indivíduos. Assim, a governança está apta a direcionar responsabilidades e políticas para a mitigação de riscos, permitindo que a função de supervisão seja executada de forma fluida. Há também um papel relevante na melhoria de processos de todo o programa de gerenciamento de dados, uma vez que a dinâmica envolvendo novas regulamentações, novas tecnologias de armazenamento e coleta de dados, mudanças no consumo e na



produção de dados são temas em constante transformação e que requerem atenção para garantir a atualização do conhecimento das áreas e pessoas.

A governança não executa diretamente atividades de cada área ou disciplinas do ecossistema do gerenciamento de dados, mas supervisiona e garante que as atividades sejam realizadas. Verifica-se que os modelos OAIS e CVCD/DCC podem trazer à Gestão de Dados requisitos de preservação e curadoria digital para reforçar os principais pontos da preservação a longo prazo entre as áreas e ao longo da gestão de dados no processo como um todo.

O Modelo de Referência OAIS, que se concentra em um conjunto de melhores práticas e recomendações para a preservação a longo prazo de um repositório, engloba uma série de atividades e funções desde o conteúdo a ser preservado até a estrutura de objeto digital. O modelo enfatiza a necessidade de planejar a preservação digital por meio do desenvolvimento de estratégias e planos de preservação digital, bem como monitorar a tecnologia e a comunidade, observando mudanças que exigem a execução de ações planejadas de preservação digital (Quadro 12). Na perspectiva do DMBOK, a área de Governança não destaca objetivos de preservação digital voltados exclusivamente para um repositório, mas é possível observar uma relação dada a necessidade de supervisionar atividades que concretizarão seus objetivos de forma prática.

Ao analisar o modelo de Ciclo de Vida de Curadoria DCC, verifica-se o mesmo comportamento encontrado entre DMBOK e OAIS. O CVCD/DCC ressalta atividades de curadoria digital identificando ações, etapas, padrões e habilidades particulares observadas em cada fase do ciclo de vida da curadoria, destacando que curar e preservar envolvem planejar a preservação e monitorar a comunidade (Quadro 13). Do ponto de vista da área de Governança do DMBOK, a área é ampla e não possui em si objetivos e atividades voltadas para as ações de curadoria conforme apresentado pelo CVCD/DCC. Da mesma forma, o modelo do CVCD/DCC não foi concebido para uma operação de curadoria digital e, portanto, não possui características de como proceder na gestão organizacional (Digital Curation Second Edition, 2016).

#### 4.2.2 Armazenamento e Operação

Nesta seção, a relação da área de conhecimento Armazenamento e Operação com a preservação e curadoria digital é analisada na perspectiva dos seus objetivos e atividades. A análise é fundamentada nos objetivos e nas atividades dos modelos OAIS e CVCD/DCC em relação ao DMBOK. Os objetivos da área de Armazenamento e Operação são apresentados no Quadro 14, que também reúne as entidades do OAIS e as ações do CVCD/DCC cujos objetivos se relacionam com os objetivos da área em questão do DMBOK. As atividades da área de Armazenamento e Operação são detalhadas no Quadro 15, que igualmente agrupa as entidades do OAIS e as ações do CVCD/DCC cujas atividades têm relação com as atividades da área em questão do DMBOK.

Quadro 14 - Entidades Funcionais de OAIS e Atividades de CVCD/DCC relacionadas a área de Armazenamento e Operação pelos Objetivos

Modelo	Área/Entidade/Ações	Objetivos
DMBOK	Armazenamento e Operação	Gerenciar a disponibilidade dos dados durante o seu ciclo de vida Garantir a integridade dos dados Gerenciar o desempenho das transações dos dados
OAIS	Entidade Funcional de Armazenamento	Fornecer serviços e funções para o armazenamento, manutenção e recuperação de Pacotes de Informações de Arquivo (AIP)
	Entidade Funcional Planejamento da Preservação	monitorar ambiente e prover planos e recomendações que garantam que a informação permaneça acessível e usável ao longo do tempo
CVCD/DCC	Ação de Preservação	Empreender ações para garantir a preservação e retenção a longo prazo da natureza autoritativa dos dados Garantir que os dados permaneçam autênticos, confiáveis e utilizáveis, mantendo sua integridade
	Armazenamento	Armazenar os dados de forma segura aderindo aos padrões relevantes
	Migração	Migrar dados para um formato diferente para garantir a imunidade dos dados contra a obsolescência de hardware ou software ou para adequar ao ambiente de armazenamento

Fonte: Elaborado pelo autor, 2023

Quadro 15 - Entidades Funcionais de OAIS e Atividades de CVCD/DCC relacionadas a área de Armazenamento e Operação por suas Atividades

Modelo	Área/Entidade/Ações	Atividades	Subatividades
DMBOK	Armazenamento e Operação	Gerenciar tecnologia de banco de dados	Entender a tecnologia de banco de dados
			Avaliar a tecnologia de banco de dados
			Gerenciar e monitorar tecnologia de banco de dados
		Gerenciar operações de banco de dados	Entender os requisitos
			Planejar a continuidade dos negócios
			Desenvolver instâncias de bancos de dados
			Gerenciar o desempenho do banco de dados
Gerenciar conjuntos de dados de teste			
Gerenciar migração de dados			
OAIS	Entidade Funcional de Armazenamento de Dados		Receber dados para armazenar
			Gerenciar hierarquia de armazenamento
			Checar erros
			Substituir mídia
			Recuperação de desastres
			Fornecer dados armazenados
	Entidade Funcional de Planejamento de Preservação		Monitorar Comunidade
			Monitorar Tecnologia
			Desenvolver estratégias e planos de preservação
CVCD/DCC	Preservar		Armazenamento
			Migração
			Ação de Preservação

Fonte: Elaborado pelo autor, 2023

A área de conhecimento Armazenamento e Operação tem como objetivo sustentar o ciclo de vida dos dados, desde sua criação/aquisição até o descarte, garantindo o desempenho e a integridade dos dados (Quadro 14). As atividades da área de Armazenamento e Operação (Quadro 15) abordam aspectos relacionados ao banco de dados, isto é, a maneira (tecnologia e gestão) em que os dados serão armazenados e mantidos durante todo o seu ciclo de vida.

Com base nos tipos de dados que serão armazenados, pode-se escolher o melhor modelo de armazenamento para proporcionar um melhor desempenho de

transações, garantir uma melhor disponibilidade de acesso e garantir a integridade dos ativos de dados. É frequentemente encorajado aos responsáveis por atuar nesta área a automatizar o máximo de processos possível para reduzir erros e retrabalhos, minimizando assim os impactos no desenvolvimento das equipes que gerenciam esses tipos de repositórios. A concepção da melhor estrutura de banco de dados também é um fator relevante para esta área, que aborda o ciclo de vida dos dados e atribui aos responsáveis pelas atividades o papel de zelar por tarefas que visam um bom desempenho, integridade, qualidade e garantia da aplicação de políticas de retenção e eliminação de dados.

A disciplina de armazenamento de dados, abordada pelo DMBOK, envolve conceitos de tecnologia da informação, métodos e atividades operacionais práticas que se aproximam dos modelos de preservação e curadoria digital. Ao analisar os objetivos do modelo OAIS em relação à área de Armazenamento e Operação do DMBOK, é possível observar que, no OAIS, há uma grande preocupação com o armazenamento dos pacotes de informação, bem como sua disponibilidade, manutenção e acesso. O OAIS especifica atividades para receber e entregar dados armazenados na unidade de armazenamento, assim como para gerenciar as mídias e as hierarquias de armazenamento, verificar periodicamente a integridade dos pacotes armazenados e prevenir desastres (Quadro 14).

O DMBOK se concentra em gerenciar a tecnologia e as operações do banco de dados, mas não especifica o conteúdo (escolhas tecnológicas) e as ações necessárias para tal, como o OAIS faz quando se refere ao conteúdo que será armazenado como Pacote de Informações. Enquanto o modelo do DMBOK trata dados de maneira não especificada, o modelo OAIS trata os pacotes de informações com riqueza de detalhes na origem de seu armazenamento.

Na perspectiva das atividades, o OAIS apresenta ações que envolvem o armazenamento permanente das informações, gerenciando a hierarquia desse armazenamento e monitorando as mídias de armazenamento (Quadro 15). Ele também expõe ações para realizar rotinas a fim de mitigar erros especiais, fornecer recursos para recuperação em caso de desastres e administrar as funções do banco de dados do Arquivo. Em relação a essas atividades, é possível observar que há uma relação entre a área em questão do DMBOK e as entidades citadas do OAIS,

pois ambas estão intimamente ligadas pela preocupação com a manutenção a longo prazo do armazenamento de dados, fato que se observa principalmente nas subatividades abordadas pelo DMBOK na compreensão das tecnologias de bancos de dados, no gerenciamento e monitoramento de tecnologias, no gerenciamento de testes e no plano de continuidade de negócios.

Ao analisar o modelo CVCD/DCC, em relação aos seus objetivos, pode-se observar que as ações: Conceituar, Preservação, Armazenamento, Acesso, Uso e Reutilização e Migração, levantam questões relativas a todo o processo de manutenção de um repositório para armazenamento a longo prazo, a compreensão da criação dos dados para entender como capturá-los e escolher a melhor forma de armazená-los. Além disso, é garantido o armazenamento do dado original, autêntico e íntegro, bem como a observância da obsolescência tecnológica do ambiente de armazenamento para a adequada migração de dados. Diante dos objetivos da área de Armazenamento e Operação do DMBOK, é observado que o DMBOK aborda como objetivos a gestão de disponibilidade de dados em ciclo de vida, a integridade e o desempenho das transações de dados. Em relação às suas atividades, pode-se observar que o CVCD/DCC aborda de forma distinta o tratamento do conteúdo, se propondo a realizar a atribuição de informações de representação para garantir a estrutura de armazenamento de dados e a área em questão do DMBOK não aborda essa atividade específica da curadoria digital.

#### 4.2.3 Integração e Interoperabilidade

Esta seção é analisada em relação à área de conhecimento Integração e Interoperabilidade com a preservação e curadoria digital na perspectiva de seus objetivos e atividades, tendo como base de análise os objetivos e as atividades dos modelos OAIS e CVCD/DCC frente ao DMBOK. O Quadro 16 apresenta os objetivos da área de Integração e Interoperabilidade e reúne as entidades do OAIS e as ações do CVCD/DCC cujos objetivos se relacionam com os objetivos da área em questão do DMBOK. O Quadro 17 detalha as atividades da área de Integração e Interoperabilidade e reúne as entidades do OAIS e as ações do CVCD/DCC cujas atividades se relacionam com as atividades da área em questão do DMBOK.

Quadro 16 - Entidades Funcionais de OAIS e Atividades de CVCD/DCC relacionadas a área de Integração e Interoperabilidade pelos Objetivos

Modelo	Área/Entidade/Ações	Objetivos
DMBOK	Integração e Interoperabilidade	Fornecer dados com segurança, em conformidade com as regulamentações Reduzir o custo e a complexidade do gerenciamento de soluções Identificar eventos significativos e acionar alertas e ações automaticamente Apoiar inteligência de negócios, análise e gerenciamento de dados
OAIS	Entidade Funcional de Armazenamento	Fornecer serviços e funções para o armazenamento, manutenção e recuperação de Pacotes de Informações de Arquivo (AIP)
	Entidade Funcional de Ingestão	Fornecer serviços e funções para aceitar Pacotes de Informações de Submissão (SIPs) de Produtores
	Entidade Funcional de Acesso	Preparar o conteúdo para o armazenamento e gerenciamento a partir de dentro do Arquivo.
CVCD/DCC	Acesso, Uso e Reutilização	Certificar-se de que os dados estejam acessíveis tanto aos usuários designados quanto aos reusuários no dia a dia
	Migração	Migrar dados para um formato diferente para garantir a imunidade dos dados contra a obsolescência de hardware ou software ou para adequar ao ambiente de armazenamento

Fonte: Elaborado pelo autor, 2023

Quadro 17 - Entidades Funcionais de OAIS e Atividades de CVCD/DCC relacionadas a área de Integração e Interoperabilidade por suas Atividades

Modelo	Área/Entidade/Ações	Atividades	Subatividades
DMBOK	Integração e Interoperabilidade	Planejar e analisar	Definir integração de dados e ciclo de vida dos requisitos
			Executar a descoberta de dados
			Linhagem de dados do documento
			Perfil dos dados
			Examinar a conformidade com as regras de negócio
		Projetar soluções integradas	Componentes da solução de design
			Mapear origens para destinos
			Orquestrar design de dados
		Desenvolver soluções integradas	Desenvolver serviços de dados
Desenvolver orquestração			

Modelo	Área/Entidade/Ações	Atividades	Subatividades
			do fluxo de dados
			Desenvolver abordagem de migração de dados
			Desenvolver processamento de eventos complexos
			Manter metadados de integração
		Implementar e monitorar	
OAIS	Entidade Funcional de Ingestão		Receber informação do produtor
			Controle de qualidade
			Gerar AIP
	Entidade Funcional de Armazenamento de Dados		Receber dados para armazenar
			Fornecer dados armazenados
Entidade Funcional de Acesso a Dados		Gerar DIP	
CVCD/DCC	Descrição e Representação da Informação		Coletar e atribuir informações de representação necessárias
	Transformação		Migrar para um formato diferente
	Descarte		Transferir dados entre repositórios, datacenters ou outro custodiante

Fonte: Elaborado pelo autor, 2023

Conforme é apresentado no DMBOK, “[...] essa área descreve processos relacionados à movimentação e consolidação de dados dentro e entre armazenamentos de dados, aplicativos e organizações” (DMBOK, 2017, p.269). Essa área do Corpo de Conhecimento do DAMA é considerada dependente de outras áreas como Governança, Arquitetura, Segurança, Metadados, Armazenamento e Operação e Modelagem e Design (DMBOK, 2017). A abordagem desta área sobre o movimento de dados entre suportes demonstra afinidade entre outras áreas específicas para detalhamento de atividades em seu próprio domínio. A gama de bancos de dados distribuídos em uma organização através de vários sistemas distintos exige a integração de dados para a geração ou complemento de novas estruturas informacionais, tornando-se cada vez mais aplicada em projetos de análises em Big Data. O grande desafio é mantido nestes sistemas e integrações funcionais diante das mudanças de suporte e das próprias estruturas de dados no

que tange a extração, transformação e carregamento. Essa área do DMBOK apresenta conceitos considerando as melhores práticas para a entrega de uma melhor performance em um ambiente de dados, adicionando processos dentro do contexto da gestão de dados de uma organização.

Na perspectiva de um OAIS, a integração de fontes de dados é observada a todo momento, seja nos processos de Ingestão, Armazenamento e Acesso, quando ocorre uma interação com produtores e consumidores do OAIS (Quadro 16). As Entidades do OAIS de Ingestão, Armazenamento e Acesso (Quadro 17) participam de processos envolvendo, respectivamente, Pacotes de Submissão de Informação (SIPs), Pacotes de Arquivamento da Informação (AIPs) e Pacotes de Disseminação da Informação (DIPs), os quais são passíveis de integração de dados a todo momento, seja dentro do OAIS ou entre OAIS distintos. Portanto, pode-se observar que existem relações entre os objetivos do modelo OAIS e a área de Integração e Interoperabilidade do DMBOK no que tange à ingestão, armazenamento e acessos dos dados.

O modelo de referência do OAIS possui conceitos específicos relacionados à preservação a longo prazo de um Arquivo, enquanto o DMBOK aborda o tema de uma maneira mais ampla e genérica. Ao analisar os objetivos, pode-se observar com mais ênfase a questão do fornecimento de dados de acordo com políticas pré determinadas, o qual o OAIS expõe com bastante objetividade através de suas entidades funcionais de Ingestão, Armazenamento e Acesso.

Por sua vez, o CVCD/DCC complementa com sua ação de Migração, dando ênfase no objetivo do dado ser passível de movimentação entre ambientes. A área em questão do DMBOK na perspectiva de suas atividades é descrita com maior riqueza de detalhes desde o projeto até o desenvolvimento da integração de soluções e aborda componentes relativos ao fluxo de dados em soluções distintas.

No que tange à curadoria, através das ações do CVCD/DCC, a presença de atividades relacionadas à integração e interoperabilidade citadas no DMBOK é observada. Estão presentes nas Ações de Descrição e Representação da Informação, Transformação e Descarte (Quadro 17). Essas ações requerem que haja interações entre dados de forma a trabalhar na movimentação destes seja na ingestão, extração até o transporte para seu armazenamento. Essas ações não



fazem parte da área do DMBOK, contudo, existem relações entre as ações do CVCD/DCC e as atividades da área em questão do DMBOK, visto que dados trafegam desde sua origem até seu repouso em armazenamentos específicos.

As atividades da ação de Descrição e Representação da Informação consistem em coletar e atribuir informações aos dados e na área em questão do DMBOK há uma atividade de “manter metadados de integração” (Quadro 17). A atividade da ação de Transformação do CVCD/DCC é “migrar para um formato diferente” e no DMBOK existe uma atividade para desenvolver uma abordagem de migração de dados. As atividades da ação de Descarte do CVCD/DCC salientam a transferência de dados entre ambientes e o DMBOK aborda essa atividade como desenvolvimento da orquestração do fluxo de dados.

#### 4.2.4 Segurança

Esta seção é analisada em relação à área de conhecimento Segurança com a preservação e curadoria digital na perspectiva de seus objetivos e das atividades, tendo os objetivos e as atividades dos modelos OAIS e CVCD/DCC frente ao DMBOK como base de análise. O Quadro 18 traz os objetivos da área de Segurança e reúne as entidades do OAIS e as ações do CVCD/DCC cujos objetivos estão relacionados com os objetivos da área em questão do DMBOK. O Quadro 19 detalha as atividades da área de Segurança e reúne as entidades do OAIS e as ações do CVCD/DCC cujas atividades estão relacionadas com as atividades da área em questão do DMBOK.

Quadro 18 - Entidades Funcionais de OAIS e Atividades de CVCD/DCC relacionadas a área de Segurança pelos Objetivos

Modelo	Área/Entidade/Ações	Objetivos
DMBOK	<b>Segurança</b>	<b>Habilitar</b> o devido acesso e impedir acessos inapropriados aos dados corporativos <b>Entender</b> e cumprir todos os regulamentos e políticas relevantes de privacidade, proteção e confidencialidade <b>Garantir</b> que as necessidades de privacidade e confidencialidade de todas as partes interessadas sejam cumpridas e auditadas
	Entidade Funcional de	

Modelo	Área/Entidade/Ações	Objetivos
	Administração	
OAIS	Entidade Funcional de Acesso	Fornecer os serviços e funções que apoiam os Consumidores na determinação da existência, descrição, localização e disponibilidade da informação armazenada no OAIS
	Entidade Funcional de Ingestão	Fornecer serviços e funções para aceitar Pacotes de Informações de Submissão (SIPs) de Produtores
CVCD/DCC		

Fonte: Elaborado pelo autor, 2023

Quadro 19 - Entidades Funcionais de OAIS e Atividades de CVCD/DCC relacionadas a área de Segurança por suas Atividades

Modelo	Área/Entidade/Ações	Atividades	Subatividades
DMBOK	Segurança	Identificar requisitos relevantes de segurança de dados	Requisitos de Negócios
			Requisitos regulatórios
		Definir política de segurança de dados	
		Definir padrões de segurança de dados	Definir níveis de confidencialidade de dados
			Definir categorias de regulação dos dados
		Definir funções de segurança	
Avaliar os atuais riscos de segurança			
Implementar controles e procedimentos			
OAIS	Entidade Funcional de Administração		Negociar Acordo de Submissão
			Gerenciar configuração do sistema
			Atualizar conteúdos
			Controle de Acesso Físico
			Estabelecer padrões e políticas
			Envio de auditoria
			Auditar submissões
			Ativar requisições
			Atendimento ao Cliente
	Entidade Funcional de Acesso	Aplicar controles para limitar o acesso a informações	
	Entidade Funcional de Ingestão	Receber a informação	
CVCD/DCC	Acesso, Uso e Reutilização	Aplicar controles de acesso robustos e procedimentos de autenticação*	

Fonte: Elaborado pelo autor, 2023

Essa área do DMBOK envolve o planejamento, o desenvolvimento e a

execução de políticas e processos que são diretamente ligados à proteção dos dados contra acessos não autorizados, sendo passíveis de auditoria de todos os procedimentos executados no contexto do gerenciamento de dados. Regulamentações e questões legais são principais direcionadores que movem toda a área e a partir destes, atividades são planejadas e executadas que estejam alinhadas à governança da gestão de dados. A classificação qualitativa de dados em relação à sensibilidade e criticidade são algumas das questões a serem observadas quando se olha para a segurança das informações, e neste ponto, os conceitos de segurança de dados em sua infraestrutura de tecnologia, conteúdo, papéis e responsabilidades dos usuários de dados são expandidos.

A área de Segurança do modelo DMBOK e as entidades do modelo OAIS e as Ações de Curadoria do Ciclo de Vida do DCC compartilham alguns objetivos comuns relacionados à gestão de dados e informações, porém cada modelo ou ação tem foco em aspectos específicos e distintos.

A área de Segurança do DMBOK é responsável por garantir a proteção dos dados corporativos, cumprindo os regulamentos e políticas de privacidade, proteção e confidencialidade. Seus objetivos são voltados para o controle de acesso, privacidade, confidencialidade e auditoria das práticas de segurança de dados.

Por outro lado, as entidades do modelo OAIS e as ações do CVCD/DCC têm uma abordagem mais ampla, envolvendo a gestão, preservação e acessibilidade dos dados ao longo do tempo. Esses modelos e ações abrangem aspectos como armazenamento, manutenção, recuperação, descrição, localização e disponibilidade da informação, além de garantir a acessibilidade e reutilização dos dados para seus usuários. A segurança é um dos muitos aspectos abordados pelos modelos de preservação e ações de curadoria e constitui um elemento importante para a integridade das informações. No modelo OAIS e no modelo CVCD/DCC, etapas de verificação da integridade dos dados são existentes para analisar se não há modificações em seu estado. Essa característica de segurança é de suma importância para a garantia da integridade da informação e representa um elemento de segurança fundamental diante do modelo da área em questão do DMBOK.

Na perspectiva de suas atividades, a área em questão possui relações com atividades do modelo OAIS e do CVCD/DCC no que tange a aplicações de controle

de acesso. O modelo OAIS cita em sua Entidade Funcional de Acesso as questões de permissionamento de usuários, levando em consideração a disponibilização de informações devidamente autorizadas. Na construção de pacotes de conteúdo, informações de direitos de acesso também são carregadas, prática amplamente discutida no Corpo de Conhecimento do DAMA.

No Ciclo de Vida de Curadoria DCC, a ação de Acesso, Uso e Reutilização traz recomendações para a aplicação de controles robustos e procedimento de autenticação. As ações do modelo CVCD/DCC possuem objetivos definidos perante a manutenção ou sustentação de um processo de curadoria digital e algumas atividades dentro dessas ações podem possuir aspectos relacionados à segurança bem próximos com os aspectos do modelo do DMBOK.

#### 4.2.5 Dados Mestre e de Referência

Esta seção é analisada em relação à área de conhecimento Dados Mestre e de Referência com a preservação e curadoria digital na perspectiva de seus objetivos e das atividades, utilizando como base de análise os objetivos e as atividades dos modelos OAIS e CVCD/DCC em comparação ao DMBOK. O Quadro 20 apresenta os objetivos da área de Dados Mestre e de Referência e congrega as entidades do OAIS e as ações do CVCD/DCC cujos objetivos se relacionam com os objetivos da área em análise. O Quadro 21 detalha as atividades da área de Dados Mestre e de Referência e reúne as entidades do OAIS e as ações do CVCD/DCC cujas atividades se relacionam com as atividades da área em questão no DMBOK.

Quadro 20 - Entidades Funcionais de OAIS e Atividades de CVCD/DCC relacionadas a área de Dados Mestre e de Referência pelos Objetivos

Modelo	Área/Entidade/Ações	Objetivos
DMBOK	Dados Mestre e de Referência	Habilitar o compartilhamento de informações entre domínios de negócios e aplicativos dentro da organização Fornecer fonte confiável de dados mestre e de referência reconciliados e com qualidade certificada Reduzir custo e complexidade por meio de uso de padrões, modelos de dados comuns e padrões de integração
OAIS	Entidade Funcional Planejamento da Preservação	Monitorar ambiente e prover planos e recomendações que garantam que a informação permaneça acessível e usável ao longo do tempo
	Entidade Funcional Administração	Fornecer serviços e funções para a operação geral do sistema de Arquivamento
CVCD/DCC	Conceitualiza	Conceber e planejar a criação de dados, incluindo método de captura e opções de armazenamento.
	Planejamento da Preservação	Planejar a preservação durante todo o ciclo de vida de curadoria do material digital. Isso incluiria planos para gerenciamento e administração de todas as ações do ciclo de vida da curadoria.

Fonte: Elaborado pelo autor, 2023

Quadro 21 - Entidades Funcionais de OAIS e Atividades de CVCD/DCC relacionadas a área de Dados Mestre e de Referência por suas Atividades

Modelo	Área/Entidade/Ações	Atividades	Subatividades
DMBOK	Dados Mestre e de Referência	Identificar requisitos e direcionadores de negócio	Validar definições de dados
		Estimar e avaliar as fontes de dados	
		Definir abordagem de arquitetura	
		Modelar dados	
		Definir administração e manter processos	
		Estabelecer políticas de governança	
		Implementar serviços de compartilhamento/integração de dados	Adquirir fontes de dados para compartilhamento Publicar dados mestre e de referência
OAIS	Entidade Funcional Planejamento da Preservação		Desenvolver estratégias e planos de preservação
	Entidade Funcional Administração		Negociar Acordo de Submissão
CVCD/DCC	Conceitualiza		Estabelecer padrões e políticas
	Planejamento da Preservação		Conceber e planejar a criação de dados
	Planejamento da Preservação		Planejar a preservação durante todo o ciclo de vida de curadoria do material digital.

Fonte: Elaborado pelo autor, 2023

Em uma organização, geralmente há uma série de sistemas com sua respectiva fonte de dados, gerando informações constantemente. Muitas dessas informações têm origem em dados idênticos, ou seja, dados que são utilizados em uma área de vendas também podem ser utilizados em uma área de recursos humanos. Com o aumento crescente da distribuição de dados, surge também uma problemática em torno da gestão de dados de referência e dados mestre.

Os Dados Mestre são definidos como dados sobre as entidades de negócios, como por exemplo: funcionários, clientes, produtos, fornecedores, estruturas financeiras. Os Dados de Referência, por outro lado, são dados usados para classificar ou categorizar outros dados, como por exemplo: códigos de países. (DMBOK 2nd, 2017). Esta área do DMBOK tem como principal objetivo abordar tópicos como: requisitos de negócio, gestão da qualidade dos dados, gestão de custos e integração e a redução de riscos em um ecossistema com muitos dados iguais em bases ou sistemas distintos. A governança exige que seja institucionalizado um conjunto de dados de referência para que sejam constituídos dados mestre, que servirão como uma única fonte da verdade acordada na organização.

Na perspectiva do modelo OAIS, uma relação é estabelecida com a área do DMBOK, uma vez que o modelo de preservação carrega consigo a responsabilidade da descrição da informação (explicitamente no SIP e AIP). Muitos aspectos encontrados nos requisitos do modelo OAIS poderiam estar mais explicitamente presentes na área em análise do DMBOK. No contexto de dados, há uma ampla atuação na camada semântica dos dados em uma organização que deseja trabalhar com informações precisas. Uma observação similar é encontrada no modelo CVCD/DCC, onde se identificam as ações de contextualização dos dados e planejamento da preservação. Referenciar os dados e apontar para seu devido significado garante uma preservação de qualidade a longo prazo, pois a informação não perderá seu significado original ao longo do tempo.

#### 4.2.6 Arquitetura de Dados

Esta seção é dedicada à análise da relação entre a área de conhecimento

Arquitetura de Dados e a preservação e curadoria digital, na perspectiva dos seus objetivos e das atividades, com base nos objetivos e nas atividades dos modelos OAIS e CVCD/DCC em comparação com o DMBOK. O Quadro 22 apresenta os objetivos da área de Arquitetura de Dados e compila as entidades do OAIS e as ações do CVCD/DCC cujos objetivos se alinham aos objetivos da área em análise do DMBOK. O Quadro 23 detalha as atividades da área de Arquitetura de Dados e congrega as entidades do OAIS e as ações do CVCD/DCC cujas atividades se correlacionam com as atividades da área em análise no DMBOK.

Quadro 22 - Entidades Funcionais de OAIS e Atividades de CVCD/DCC relacionadas a área de Arquitetura de Dados pelos Objetivos

Modelo	Área/Entidade/Ações	Objetivos
DMBOK	Arquitetura de Dados	Identificar requisitos de armazenamento e processamento de dados Projetar estruturas e planos que atendam requisitos dados correntes a longo prazo da organização Preparar a organização para a evolução de produtos relacionados com dados
OAIS	Entidade Funcional de Ingestão	Fornecer serviços e funções para aceitar Pacotes de Informações de Submissão (SIPs) de Produtores Preparar o conteúdo para o armazenamento e gerenciamento a partir de dentro do Arquivo.
	Entidade Funcional de Armazenamento de Arquivamento	Fornecer serviços e funções para o armazenamento, manutenção e recuperação de Pacotes de Informações de Arquivo (AIP)
	Entidade Funcional Planejamento da Preservação	monitorar ambiente e prover planos e recomendações que garantam que a informação permaneça acessível e usável ao longo do tempo
CVCD/DCC	Armazenamento	Armazenar os dados de forma segura aderindo aos padrões relevantes
	Descrição e Representação da Informação	Coletar e atribuir informações de representação necessárias
	Planejamento da Preservação	Planejar a preservação durante todo o ciclo de vida de curadoria do material digital. Isso incluiria planos para gerenciamento e administração de todas as ações do ciclo de vida da curadoria.

Fonte: Elaborado pelo autor, 2023

Quadro 23 - Entidades Funcionais de OAIS e Atividades de CVCD/DCC relacionadas a área de Arquitetura de Dados por suas Atividades

Modelo	Área/Entidade/Ações	Atividades	Subatividades
DMBOK	Arquitetura de Dados	Estabelecer arquitetura corporativa de dados	Avaliar as especificações de arquitetura de dados existentes
			Desenvolver um roteiro
		Integrar arquitetura corporativa	Gerenciar requisitos corporativos em projetos
OAIS	Entidade Funcional de Ingestão		Receber a informação
			Gerar AIP
	Entidade Funcional de Armazenamento de Dados		Receber dados para armazenar (AIP)
			Fornecer dados armazenados (AIP)
Entidade Funcional Planejamento da Preservação	Desenvolver estratégias e planos de preservação		
CVCD/DCC	Planejamento da Preservação		
	Representação da Informação		
	Armazenamento		

Fonte: Elaborado pelo autor, 2023

A definição de arquitetura se estende a vários domínios, desde uma arquitetura de negócios até uma arquitetura técnica de tecnologia, por exemplo. De maneira simplificada, a arquitetura de dados tem como objetivo o planejamento para a construção de estruturas que envolvem dados em seu respectivo ecossistema. Esta área pode ser uma das mais complexas do corpo de conhecimento do DAMA, uma vez que abrange muitos domínios e incorpora diversos modelos de referência de arquiteturas em áreas distintas. Aqui, observa-se uma variedade de atividades que estabelecem conexões com outras áreas do DMBOK e que podem ser utilizadas e aplicadas em quase todas as esferas do arcabouço da gestão de dados.

Como a área em questão identifica, projeta e prepara, existem relações entre o DMBOK e os modelos de curadoria e preservação digital, principalmente onde se observa dentro das áreas em questão algum tipo de construção de planejamento para a posterior execução (Quadro 22 e Quadro 23), e é neste ponto que a análise



se estabelece.

Na perspectiva do modelo OAIS, o conceito se faz presente em algumas partes do modelo. O OAIS se concentra em fornecer insumos para que as arquiteturas possam ser adaptadas para interagir com os pacotes de informações, ditando pré-requisitos para a construção do ecossistema de um Repositório. Enquanto o DMBOK aborda a arquitetura de dados, o OAIS se concentra na criação de formatos para adequar seus pacotes de informações. Como os objetivos do OAIS são, em sua maioria, fornecer serviços e funções dos seus pacotes de informação, é possível observar que, indiretamente, este direciona a maneira como o ecossistema será projetado. O modelo OAIS possui em sua essência um viés de arquitetura, pois as entidades funcionais estabelecem uma arquitetura de dados que identifica suas camadas, se concentra no armazenamento - um dos atores da preservação digital - e interage com as outras camadas que compõem o ecossistema de informação.

Na análise do modelo CVCD/DCC, também se observa a ausência de uma ação específica envolvendo a arquitetura em si. As ações do ciclo de vida percorrem uma arquitetura específica, porém não direcionam diretamente para a construção de algum ambiente, deixando a cargo dos sistemas de preservação a adequação de acordo com suas ações de curadoria.

#### 4.2.7 Metadados

Esta seção é analisada em relação à área de conhecimento de Metadados com a preservação e curadoria digital sob a perspectiva de seus objetivos e atividades, tendo como base de análise os objetivos e as atividades dos modelos OAIS e CVCD/DCC em confronto com o DMBOK. O Quadro 24 apresenta os objetivos da área de Metadados e reúne as entidades do OAIS e as ações do CVCD/DCC cujos objetivos estão relacionados com os objetivos da área em questão do DMBOK. O Quadro 25 detalha as atividades da área de Metadados e reúne as entidades do OAIS e as atividades do CVCD/DCC cujas atividades estão relacionadas com as atividades da área em questão do DMBOK.

Quadro 24 - Entidades Funcionais de OAIS e Atividades de CVCD/DCC relacionadas a área de Metadados pelos Objetivos

Modelo	Área/Entidade/Ações	Objetivos
DMBOK	Metadados	Fornecer compreensão organizacional dos termos e uso de negócios Coletar e integrar metadados de diversas fontes Fornecer um padrão de acesso aos metadados Garantir a qualidade e a segurança dos metadados
OAIS	Entidade Funcional de Ingestão	Fornecer serviços e funções para aceitar os pacotes de informação e preparar o conteúdo para armazenamento e gerenciamento dentro do Arquivo
	Entidade Funcional de Gestão de Dados	fornece os serviços e funções para preencher, manter e acessar Informações Descritivas que identificam e documentam os acervos do repositório e dados administrativos usados para gerenciar o repositório
CVCD/DCC	Descrição e Representação da Informação	Garantir a descrição e o controle adequados a longo prazo
		Entender e renderizar o material digital e os metadados associados

Fonte: Elaborado pelo autor, 2023

Quadro 25 - Entidades Funcionais de OAIS e Atividades de CVCD/DCC relacionadas a área de Metadados por suas Atividades

Modelo	Área/Entidade/Ações	Atividades	Subatividades
DMBOK	Metadados	Definir estratégia de metadados	
		Entender os requisitos de metadados	Requisitos do usuário comercial
			Requisitos técnicos do usuário
		Definir arquitetura de metadados	Criar metamodelo
			Aplicar padrões de metadados
			Gerenciar armazenamentos de metadados
Criar e manter metadados	Integrar metadados		
	Distribuir e entregar metadados		
Consultar, relatar e analisar metadados			
OAIS	Entidade Funcional de Ingestão		Extrair informações descritivas dos AIPs para inclusão no banco de dados da Gestão de Dados
			Receber Pacotes de Informações de Arquivo (AIP) da Entidade de Ingestão
			Adicionar os AIPs ao armazenamento permanente
	Entidade Funcional de Gestão de Dados	Administrar base de dados (informação descritiva e informação do sistema)	

Modelo	Área/Entidade/Ações	Atividades	Subatividades
			Realizar consultas
			Gerar relatórios
			Realizar atualizações
CVCD/DCC	Descrição e Representação da Informação		Atribuir metadados administrativos, descritivos, técnicos, estruturais e de preservação usando padrões apropriados
			Coletar e atribuir informações de representação necessárias
	Criar e Receber		Adicionar metadados de preservação no momento da criação*
			Atribuir metadados apropriados*
	Ação de Preservação		Realizar a limpeza dos dados, validar e atribuir de metadados de preservação
			Atribuir informações de representação e garantir as estruturas de dados ou formatos de arquivo aceitáveis.

Fonte: Elaborado pelo autor, 2023

A descrição dos dados é considerada essencial para a formação de uma base de conhecimento sólida que direciona sistemas e a maneira como os dados são consumidos. Gerir dados sem metadados é vista como uma tarefa impossível e até mesmo os próprios metadados necessitam de gestão. O fornecimento de contexto para os dados permite que as informações sejam geradas com maior confiabilidade, qualidade e segurança. Os metadados auxiliam na representação de informações e impactam áreas como a segurança, ajudando a identificar dados sensíveis para, assim, criar e aplicar políticas de acesso. Metadados mal gerenciados podem causar sérias redundâncias de dados e, por consequência, uma falta de integridade informacional, com bases de conhecimento desatualizadas e instáveis.

Tanto o OAIS quanto o CVCD/DCC trabalham com metadados descritivos, estruturais, técnicos e administrativos de preservação digital. Além disso, especificam atividades que recebem, verificam e gerenciam os metadados. O DMBOK apresenta atividades para gerir metadados, como definir estratégias e requisitos de metadados, definir arquitetura de metadados, aplicar padrões, criar e manter metadados (Quadro 25). Para envolver a curadoria digital, os metadados do

OAIS e do CVCD/DCC precisam ser requisitos do DMBOK, observando-se seus padrões, sendo assim incorporados nas ações do DMBOK para gerenciar metadados.

No modelo OAIS, os metadados são recebidos e certificados na entidade funcional de Ingestão, são então inseridos no pacote AIP e são enviados para o componente de Gestão de Dados. Nessa perspectiva, entre seus objetivos está a preparação do conteúdo para armazenamento de dados, tornando o modelo OAIS específico, enquanto a área de Metadados no DMBOK se mostra mais ampla para atender a vários aspectos de dados em geral.

No modelo CVCD/DCC, observa-se que os objetivos também se relacionam com os objetivos da área em questão do DMBOK. A garantia da descrição e o controle, bem como a compreensão e a renderização do material digital, são objetivos dos Metadados do DMBOK. As relações são mais evidentes em relação às atividades, pois se encontram aspectos que relacionam as características das atividades de curadoria com as atividades do DMBOK. A atribuição de metadados, a coleta de informações de representação e a gestão de metadados constituem as atividades da área em questão do DMBOK, citadas no CVCD/DCC.

#### 4.2.8 Modelagem e Design

Nesta seção, a relação da área de conhecimento Modelagem e Design com a preservação e curadoria digital é analisada, tendo-se como base de análise os objetivos e as atividades dos modelos OAIS e CVCD/DCC em comparação ao DMBOK. O Quadro 26 apresenta os objetivos da área de Modelagem e Design e reúne as entidades do OAIS e as ações do CVCD/DCC cujos objetivos se relacionam com os objetivos da área em questão do DMBOK. O Quadro 27 detalha as atividades da área de Modelagem e Design e congrega as entidades do OAIS e as atividades do CVCD/DCC cujas atividades se correlacionam com as atividades de Integração e Interoperabilidade.

Quadro 26 - Entidades Funcionais de OAIS e Atividades de CVCD/DCC relacionadas a área de Modelagem e Design pelos Objetivos

Modelo	Área/Entidade/Ações	Objetivos
DMBOK	Modelagem e Design	Estimular e documentar a compreensão das perspectivas sobre dados Criar um alicerce para concluir iniciativas de escopo amplo
OAIS	Entidade Funcional de Ingestão	Fornecer serviços e funções para aceitar Pacotes de Informações de Submissão (SIPs) de Produtores Preparar o conteúdo para o armazenamento e gerenciamento a partir de dentro do Arquivo.
	Entidade Funcional de Armazenamento de Arquivamento	Fornecer serviços e funções para o armazenamento, manutenção e recuperação de Pacotes de Informações de Arquivo (AIP)
	Entidade Funcional Planejamento da Preservação	Monitorar ambiente e prover planos e recomendações que garantam que a informação permaneça acessível e usável ao longo do tempo
CVCD/DCC	Conceituar	Conceber e planejar a criação de dados, incluindo método de captura e opções de armazenamento
	Armazenamento	Armazenar os dados de forma segura aderindo aos padrões relevantes
	Planejamento da Preservação	Planejar a preservação durante todo o ciclo de vida de curadoria do material digital. Isso incluiria planos para gerenciamento e administração de todas as ações do ciclo de vida da curadoria.

Fonte: Elaborado pelo autor, 2023

Quadro 27 - Entidades Funcionais de OAIS e Atividades de CVCD/DCC relacionadas a área de Modelagem e Design por suas Atividades

Modelo	Área/Entidade/Ações	Atividades	Subatividades
DMBOK	Modelagem e Design	Planejar a modelagem de dados	
		Construir modelos de dados	Criar modelo conceitual de dados
			Criar modelo lógico de dados
			Criar modelo físico de dados
Revisar os modelos de dados			
		Gerenciar modelos de dados	
OAIS	Entidade Funcional de Planejamento de Preservação		Projeta modelos de Pacote de Informações Fornece assistência e revisão de design
	Entidade Funcional de Armazenamento de Dados		Fornecer dados armazenados (AIP)
	Entidade Funcional de Ingestão		Adicionar os AIPs ao armazenamento permanente
CVCD/DCC	Planejamento da Preservação		
	Conceituar		
	Armazenamento		

Fonte: Elaborado pelo autor, 2023

A área de modelagem de dados é aquela que descobre, analisa e define os requisitos para a compreensão eficaz dos ativos de dados. É crucial entender como os dados são percebidos em um determinado domínio para gerar um ecossistema em que uma organização produzirá e consumirá informações. Para tal, uma descrição minuciosa em torno de um modelo de dados se mostra eficaz, auxiliando a gestão de dados a orquestrar mudanças em todas as outras áreas do modelo. Uma documentação bem elaborada e alinhada com os objetivos de negócio pode apoiar nas decisões de investimentos em projetos e pode otimizar custos para uma organização que utiliza seus dados como fonte para tomada de decisões estratégicas. A abordagem do DMBOK para esta área está direcionada em levantar o máximo de questões que possam impactar um ambiente de dados, incluindo o próprio dado. A modelagem de dados passa pelo entendimento do formato do dado até sua melhor forma de armazenamento, considerando sua mudança dinâmica de estado e local. Dentro de um ecossistema heterogêneo, é considerado o nível mais baixo de um banco de dados para tipos estruturados e o melhor repositório para dados não estruturados. O DMBOK também incentiva o uso das melhores práticas de mercado para atender os requisitos de uma modelagem de dados bem aplicada, criando uma base sólida para outras áreas da gestão de dados.

Na perspectiva dos objetivos, o modelo OAIS expressa preocupação com a preparação dos pacotes de informações até seu devido armazenamento e manutenção. Analisando suas atividades, a entidade funcional de Planejamento de Preservação apresenta duas subatividades que projetam modelos de pacotes de informações e fornecem suporte para a revisão de modelos. O modelo OAIS aponta para uma modelagem específica de um pacote de informações, um conceito mais completo e de um domínio previamente conhecido, enquanto a área em questão do DMBOK trabalha com dados, uma abordagem mais ampla e aberta.

As ações de curadoria do CVCD/DCC, na perspectiva dos objetivos e atividades, não possuem especificamente uma ação de modelagem de todo o ecossistema como ocorre no DMBOK. Indiretamente, existem em suas ações de curadoria tarefas que planejam e monitoram, visto que os dados se encontram no centro do modelo CVCD/DCC. A área de Modelagem e Design do DMBOK não possui características do modelo em questão de curadoria frente à modelagem de

dados, porém o conhecimento dos dados que serão gerenciados pode ser relevante na construção de um bom modelo de dados, de forma a aplicar ações que garantam a longevidade dos dados.

#### 4.2.9 Data Warehousing & Business Intelligence

Esta seção tem como objeto de análise a relação da área de conhecimento Data Warehousing & Business Intelligence com a preservação e curadoria digital, na perspectiva de seus objetivos e das atividades. No Quadro 28 são apresentados os objetivos da área de Data Warehousing & Business Intelligence, e são reunidas as entidades do OAIS e as ações do CVCD/DCC cujos objetivos se relacionam com os objetivos da área em questão do DMBOK. O Quadro 29 detalha as atividades da área de Data Warehousing & Business Intelligence, agrupando as entidades do OAIS e as atividades do CVCD/DCC cujas atividades têm relação com as atividades da área em questão do DMBOK.

Quadro 28 - Entidades Funcionais de OAIS e Atividades de CVCD/DCC relacionadas a área de Data Warehousing & Business Intelligence pelos Objetivos

Modelo	Área/Entidade/Ações	Objetivos
DMBOK	Data Warehousing & Business Intelligence	Construir e manter o ambiente técnico e os processos técnicos e de negócio Apoiar e permitir análises de negócios eficazes e tomadas de decisões
OAIS	Entidade Funcional de Gerenciamento de Dados	Fornecer serviços e funções para preencher, manter e acessar Informações Descritivas
	Entidade Funcional de Acesso	Fornecer os serviços e funções que apoiam os Consumidores na determinação da existência, descrição, localização e disponibilidade da informação armazenada no OAIS
CVCD/DCC	Armazenamento	Armazenar os dados de forma segura aderindo aos padrões relevantes
	Acesso, Uso e Reutilização	Certificar-se de que os dados estejam acessíveis tanto aos usuários designados quanto aos reusuários no dia a dia

Fonte: Elaborado pelo autor, 2023

Quadro 29 - Entidades Funcionais de OAIS e Atividades de CVCD/DCC relacionadas a área de Data Warehousing & Business Intelligence por suas Atividades

Modelo	Área/Entidade/Ações	Atividades	Subatividades
DMBOK	Data Warehousing & Business Intelligence	Entender os requisitos	
		Definir e manter a arquitetura do DW e BI	
		Desenvolver o DW e os Data Marts	
		Carregar o DW	
		Implementar portfólio de inteligência de negócios	
		Manter produtos de dados	
OAIS	Entidade Funcional de Gerenciamento de Dados		Realizar consultas em dados de gerenciamento para produzir relatórios
	Entidade Funcional de Acesso		Gerar respostas (Pacotes de Informação de Divulgação, respostas a consultas, relatórios)
CVCD/DCC	Acesso, Uso e Reutilização		Publicar informações de acesso público*
	Transformação		Criar um subconjunto, por seleção ou consulta, para criar novos resultados derivados*

Fonte: Elaborado pelo autor, 2023

A estrutura de gestão de dados conhecida como Data Warehousing é facilitada para o armazenamento, organização e recuperação de grandes volumes de dados. É projetada para permitir um acesso eficiente a dados para fins analíticos, transformando dados brutos em informações significativas utilizáveis para tomada de decisões informadas. Por outro lado, Business Intelligence (BI) refere-se à prática que é facilitada para a coleta, integração, análise e apresentação de informações de negócios, facilitando assim a tomada de decisões. O BI utiliza o Data Warehousing como base para a implementação, fornecendo uma visão consolidada dos dados de negócios. A partir dessa perspectiva, o objetivo principal da área de Data Warehousing & Business Intelligence (DW&BI) do DMBOK é a construção e manutenção do ambiente de dados para a transformação dos mesmos em informações úteis para a tomada de decisões estratégicas. Tal transformação envolve uma série de atividades, desde a coleta e integração de dados até a análise



e apresentação dos resultados.

Os objetivos do modelo OAIS não possuem uma amplitude tão grande quanto a do DMBOK na área de DW&BI, mas possuem fortes elementos sobre o armazenamento de dados para o consumo. A proposta do DMBOK para a área em questão consiste na geração de uma camada analítica baseada nos dados armazenados. Todos os processos de garantia de integridade e acesso aos dados são empregados em um Data Warehousing, resultando em atividades específicas e direcionadas para preparação de dados para consumo. Com relação aos objetivos, o DMBOK não possui relação direta com o modelo OAIS nessa área, porém o OAIS apresenta serviços e funções que apoiam o consumo das informações.

Na perspectiva dos objetivos do CVCD/DCC, não existe uma relação direta com o DMBOK. As ações de curadoria não se concentram em análises de dados de forma a disponibilizar visões direcionadas ao negócio. Também não existe uma relação entre suas atividades e a área em questão do DMBOK, porém, o CVCD/DCC sugere publicar informações e criar subconjuntos para seleção ou consulta para a geração de novos resultados derivados.

#### 4.2.10 Qualidade de Dados

Esta seção tem como objeto de análise a relação da área de conhecimento Qualidade de Dados com a preservação e curadoria digital, considerando a perspectiva dos seus objetivos e atividades. A base de análise são os objetivos e as atividades dos modelos OAIS e CVCD/DCC em relação ao DMBOK. Os objetivos da área de Qualidade de Dados são apresentados no Quadro 30, que reúne as entidades do OAIS e as ações do CVCD/DCC cujos objetivos se relacionam com os objetivos da área do DMBOK em questão. O Quadro 31 detalha as atividades da área de Integração e Interoperabilidade, reunindo as entidades do OAIS e as ações do CVCD/DCC cujas atividades têm relação com as atividades da área do DMBOK em questão.

Quadro 30 - Entidades Funcionais de OAIS e Atividades de CVCD/DCC relacionadas a área de Qualidade de Dados pelos Objetivos

Modelo	Área/Entidade/Ações	Objetivos
DMBOK	Qualidade	Desenvolver abordagem para adequar os dados de acordo com os requisitos dos consumidores Definir padrões, requisitos e especificações para controles de qualidade de dados como parte do ciclo de vida dos dados Definir e implementar processos para medir, monitorar e relatar os níveis de qualidade dos dados Identificar e defender oportunidades para melhorar a qualidade dos dados, por meio de melhorias de processo e sistema
OAIS	Entidade Funcional de Ingestão	Fornecer serviços e funções para aceitar Pacotes de Informações de Submissão (SIPs) de Produtores Preparar o conteúdo para o armazenamento e gerenciamento a partir de dentro do Arquivo.
	Entidade Funcional de Acesso	Fornecer os serviços e funções que apoiam os Consumidores na determinação da existência, descrição, localização e disponibilidade da informação armazenada no OAIS
	Entidade Funcional de Administração	As funções de administração incluem solicitar e negociar contratos de envio com produtores, auditar envios para garantir que eles atendam aos padrões de arquivamento e manter o gerenciamento de configuração de hardware e software do sistema.
CVCD/DCC	Conceituar	Conceber e planejar a criação de dados, incluindo método de captura e opções de armazenamento
	Avaliar e Selecionar	Avaliar os dados e selecionar para curadoria e preservação a longo prazo
	Ação de Preservação	Empreender ações para garantir a preservação e retenção a longo prazo da natureza autoritativa dos dados
		Garantir que os dados permaneçam autênticos, confiáveis e utilizáveis, mantendo sua integridade
	Descarte	Descartar os dados que não foram selecionados para curadoria e preservação a longo prazo de acordo com políticas, orientações ou requisitos legais documentados
Reavaliação	Retornar dados que não cumpram os procedimentos de validação para posterior avaliação e reSeleção	

Fonte: Elaborado pelo autor, 2023

Quadro 31 - Entidades Funcionais de OAIS e Atividades de CVCD/DCC relacionadas a área de Qualidade de Dados por suas Atividades

Modelo	Área/Entidade/Ações	Atividades	Subatividades
DMBOK	Qualidade	Definir dados de alta qualidade	
		Definir uma estratégia de qualidade de dados	
		Definir escopo da avaliação inicial	Identificar dados críticos
			Identificar regras e padrões existentes
		Realizar a avaliação inicial da qualidade dos dados	Identificar e priorizar problemas
			Realizar análise de causa raiz dos problemas
		Identificar e priorizar melhorias	Priorizar ações com base no impacto nos negócios
			Desenvolver ações preventivas e corretivas
			Fortalecer as ações planejadas
		Desenvolver e implantar operações de qualidade de dados	Desenvolver procedimentos operacionais de qualidade de dados
Corrigir defeitos de qualidade de dados			
Medir e monitorar a qualidade dos dados			
Gerar relatórios sobre os níveis e descobertas de qualidade de dados			
OAIS	Entidade Funcional de Ingestão	Executar a garantia de qualidade em Pacotes de Informações de Submissão (SIPs)	
	Entidade Funcional de Administração	As funções de administração incluem solicitar e negociar contratos de envio com produtores, auditar envios para garantir que eles atendam aos padrões de arquivamento e manter o gerenciamento de configuração de hardware e software do sistema.	
CVCD/DCC			

Fonte: Elaborado pelo autor, 2023

A gestão eficaz da qualidade dos dados, conforme descrita no DMBOK, é apresentada como uma área crítica que demanda uma análise cuidadosa, considerando-se as suas implicações e o seu potencial para influenciar as operações e a estratégia de uma organização. A Qualidade de Dados é

compreendida como a condição dos dados em termos de sua adequação para atender aos seus propósitos pretendidos em operações, tomada de decisões e planejamento. Essa adequação é geralmente medida e descrita em termos de características como precisão, completude, consistência, confiabilidade, relevância e pontualidade dos dados. A gestão da qualidade dos dados envolve um conjunto de práticas que são implementadas para assegurar que os dados atendam a esses critérios e que eles permaneçam adequados para o seu uso pretendido ao longo do tempo. Os objetivos da Qualidade de Dados se centram em garantir que os dados sejam confiáveis, precisos, completos e relevantes para a tomada de decisões estratégicas e operacionais. Esses objetivos são alcançados por meio de uma variedade de atividades, que incluem a definição de padrões de qualidade de dados, a monitoração da conformidade com esses padrões, a identificação e correção de problemas de qualidade de dados e a implementação de melhorias contínuas na gestão de dados.

No modelo OAIS, a qualidade de dados é incorporada em suas entidades e funções. Desde o momento da ingestão dos dados, os pacotes de informações passam por uma espécie de garantia de qualidade antes de seu destino. Não há exatamente um objetivo tão claro como a área específica do DMBOK, mas sim existem atividades envolvendo o uso da qualidade de dados. A entidade funcional de Ingestão possui uma atividade chamada Garantia de Qualidade (Quality Assurance) na qual é validada a transferência do SIP (pacote de informações de envio) para a área de armazenamento temporário, e assim ocorre em várias outras funções das entidades. Pode-se observar que há uma aproximação entre as atividades, mas os conceitos de preservação são mais evidentes no OAIS em cada micro processo relacionado a seus pacotes de informações.

No modelo CVCD/DCC, há uma constância da aplicação de conceitos de qualidade de dados nas ações do ciclo de vida de curadoria para manter seu propósito. O DMBOK não apresenta os conceitos de curadoria de forma específica para as atividades da área de Qualidade, deixando os processos mais abertos e genéricos e colocando os esforços de entendimento nas etapas de planejamento de qualidade.

#### 4.2.11 Gestão de Documentos e Conteúdo

Esta seção é dedicada à análise da relação da área de conhecimento Gestão de Documentos e Conteúdo com a preservação e curadoria digital na perspectiva de seus objetivos e das atividades, tendo como base de análise os objetivos e as atividades dos modelos OAIS e CVCD/DCC em relação ao DMBOK. O Quadro 32 apresenta os objetivos da área de Gestão de Documentos e Conteúdo e agrega as entidades do OAIS e as ações do CVCD/DCC cujos objetivos se relacionam com os objetivos da área em questão do DMBOK. O Quadro 33 detalha as atividades da área de Gestão de Documentos e Conteúdo e agrupa as entidades do OAIS e as ações do CVCD/DCC cujas atividades têm relação com as atividades da área em questão do DMBOK.

Quadro 32 - Entidades Funcionais de OAIS e Atividades de CVCD/DCC relacionadas a área de Gestão de Documentos e Conteúdo pelos Objetivos

Modelo	Área/Entidade/Ações	Objetivos
DMBOK	Gestão de Documento e Conteúdo	Cumprir as obrigações legais e as expectativas dos clientes sobre a gestão dos registros Garantir armazenamento, recuperação e uso eficaz e eficientes de documentos e conteúdos Garantir capacidades de integração entre conteúdo estruturado e não estruturado
OAIS	Entidade Funcional de Ingestão	Fornecer serviços e funções para aceitar Pacotes de Informações de Submissão (SIPs) de Produtores Preparar o conteúdo para o armazenamento e gerenciamento a partir de dentro do Arquivo.
	Entidade Funcional de Armazenamento de Arquivamento	Fornecer serviços e funções para o armazenamento, manutenção e recuperação de Pacotes de Informações de Arquivo (AIP)
	Entidade Funcional de Gerenciamento de Dados	Fornecer serviços e funções para preencher, manter e acessar Informações Descritivas
	Entidade Funcional de Acesso	Fornecer os serviços e funções que apoiam os Consumidores na determinação da existência, descrição, localização e disponibilidade da informação armazenada no OAIS
CVCD/DCC	Conceituar	Conceber e planejar a criação de dados, incluindo método de captura e opções de armazenamento
	Armazenamento	Armazenar os dados de forma segura aderindo aos padrões relevantes

Fonte: Elaborado pelo autor, 2023

Quadro 33 - Entidades Funcionais de OAIS e Atividades de CVCD/DCC relacionadas a área de Gestão de Documentos e Conteúdo por suas Atividades

Modelo	Área/Entidade/Ações	Atividades	Subatividades
DMBOK	Gestão de Documento e Conteúdo	Planejar o gerenciamento do ciclo de vida	Desenvolver uma estratégia de conteúdo
		Criar políticas de manipulação de conteúdo, incluindo abordagem de descoberta eletrônica	
		Definir arquitetura da informação	
		Gerenciar o ciclo de vida	Capturar e gerenciar registros e conteúdo Reter, descartar e arquivar registros e conteúdo
		Publicar e entregar conteúdo	
OAIS	Entidade Funcional de Ingestão	Coordenar atualizações para o armazenamento do Arquivo e Gerenciamento de Dados	
	Entidade Funcional de Administração		Inventariar, relatar e migrar/atualizar o conteúdo do arquivamento
	Entidade Funcional de Acesso	Entrega de respostas aos consumidores	
CVCD/DCC	Descarte		Destruir dados de forma segura por razões legais*

Fonte: Elaborado pelo autor, 2023

A área de Gestão de Documento e Conteúdo do DMBOK é abordada, tratando de dados e informações que são armazenados fora de bancos de dados relacionais como documentos e outras informações não estruturadas ou semiestruturadas. Os objetivos dessa área são garantir, que incluem a integridade, a segurança, a acessibilidade e a usabilidade dos documentos e do conteúdo. Além disso, um esforço é feito nesta área para garantir a conformidade legal e regulamentar e para aumentar a eficiência e eficácia operacional. Para alcançar esses objetivos, uma série de atividades são realizadas, que incluem a definição de políticas e procedimentos para a gestão de documentos e conteúdo, a implementação de sistemas de gestão de conteúdo, a manutenção de um inventário de conteúdo e a garantia da conformidade regulamentar. Estas atividades envolvem

o gerenciamento de documentos e registros das organizações, como: e-mails, conversas de bate-papo, sites, documentos eletrônicos e dados brutos de aplicações e metadados. Esta área de conhecimento aborda este tipo de documento e a gestão do conteúdo desses registros.

Mecanismos de busca por documentos e conteúdos são apresentados como ferramentas auxiliares para o enriquecimento da gestão, como o uso de taxonomia, esquemas de classificação e etiquetagem, thesaurus, ontologias, etc. Nesta área, a gestão de dados possui afinidades com os modelos de preservação e curadoria digital. Na perspectiva dos objetivos, uma aproximação é observada entre os modelos OAIS e CVCD/DCC e a área em questão do DMBOK. Em virtude do modelo DMBOK dispor de objetivos específicos envolvendo a manipulação dos dados, a preocupação com a observância de padrões e políticas permeia ambos os modelos.

Há uma necessidade nas organizações de preservação a longo prazo de certos documentos e registros, seja por questões legais ou por preservação da memória. Nesse sentido, traços importantes são encontrados nos processos de gerenciamento de dados de preservação e nas ações de curadoria referentes ao conteúdo armazenado em um repositório ou arquivo. As ações que devem ser executadas para garantir a preservação das informações a longo prazo são ditadas pelo modelo CVCD/DCC, mas os caminhos técnicos de implementação não são especificados de forma técnica. O modelo OAIS também se comporta de maneira semelhante, desempenhando um papel importante para padrões específicos de preservação que o DMBOK não aborda de forma direta.

### **4.3 Síntese da análise dos dados**

Esta seção apresenta de forma objetiva uma síntese das análises realizadas entre a relação dos conceitos, objetivos e atividades do Corpo de Conhecimento DMBOK e os modelos de preservação digital e curadoria digital OAIS e CVCD/DCC.

#### **4.3.1 Relação entre objetivos do OAIS e CVCD/DCC**

- Os objetivos do DMBOK são distintos dos do OAIS e do CVCD/DCC,

pois o foco do DMBOK não é a preservação digital, mas sim a gestão de dados de seu corpo de conhecimento;

- Não se observa nos objetivos do DMBOK aspectos de preservação e curadoria digital, entretanto, os objetivos do DMBOK podem envolver indiretamente os objetivos do OAIS e do CVCD/DCC;
- A gestão de dados envolve a curadoria e a preservação digital e os modelos OAIS e CVCD/DCC podem complementar o DMBOK e suas áreas com aspectos de curadoria e preservação digital.

#### 4.3.2 Relação entre os principais conceitos abordados pelo DMBOK, OAIS e CVCD/DCC

- Os principais conceitos apresentados pelo DMBOK são aqueles que delimitam as suas áreas de conhecimento;
- Os principais conceitos apresentados pelo OAIS são aqueles que envolvem suas especificações de ambiente, estrutura de informação e de componentes funcionais;
- Os principais conceitos apresentados pelo CVCD/DCC são aqueles que envolvem a curadoria digital em uma organização na perspectiva de um ciclo de ações: representar informação, planejar a preservação, monitorar comunidade, preservar, curar, conceitualizar, criar, avaliar, selecionar, ingerir, armazenar, preservar, acessar, usar, reusar, transformar e eliminar;
- Identifica-se que OAIS e DCC compartilham muitos conceitos, com enfoques diferentes: OAIS aplicado à repositório e DCC aplicado no ciclo de vida da curadoria digital em uma organização;
- Os conceitos do DMBOK, voltados à gestão de dados, não enfatizam diretamente a preservação a longo prazo desses dados. Entretanto, é identificado que os conceitos do DMBOK podem envolver conceitos de preservação e curadoria digital na perspectiva de se ter uma governança de dados a longo prazo, desenvolver arquitetura, modelagem, dados mestres, segurança, armazenamento, metadados,



integração de dados visando sua preservação a longo prazo.

#### 4.3.3 Áreas de Conhecimento do DMBOK e sua relação com OAIS e CVCD/DCC

- As áreas do DMBOK relacionam-se com OAIS e CVCD/DCC na forma de tratamento do objeto digital;
- O Corpo de Conhecimento DMBOK aproxima-se de OAIS e CVCD/DCC em função dos micro processos de preservação e curadoria na garantia da informação a longo prazo;
- O DMBOK possui relação com OAIS em função da descrição das informações no que tange aos pacotes. No CVCD/DCC há aproximações pela mesma atividade de contextualização;
- Não há relação direta das áreas do DMBOK com os objetivos de OAIS e CVCD/DCC. Há aproximações que podem ser observadas na construção de pacotes de informações OAIS;
- As áreas do DMBOK aproximam-se com OAIS e CVCD/DCC no tratamentos dos repositórios de dados e sua manutenção;
- O DMBOK relaciona com OAIS pelo senso de administração e supervisão dos processos;
- As áreas do DMBOK aproximam-se com OAIS e CVCD/DCC em função das práticas de acesso aos dados e mitigação de riscos;
- As áreas do DMBOK possuem objetivos e atividades específicas porém aproxima com OAIS e CVCD/DCC pela aplicação de contexto em determinado domínio.

#### **4.4 Proposições sobre a Preservação e Curadoria Digital em DMBOK**

Nesta seção, o objetivo "Ampliar a visão da gestão de dados através de sugestões de incorporação da Preservação e Curadoria digital no DMBOK" é abordado. Com base nas análises dos objetivos específicos em relação aos objetivos e atividades do DMBOK e modelos de preservação e curadoria OAIS e CDCV/DCC, sugestões de incorporação de características da preservação e

curadoria digital no corpo de conhecimento DMBOK são apresentadas.

A eficiência e eficácia na gestão de dados em uma organização são necessárias para alimentar as tomadas de decisões estratégicas. Projetos relacionados a dados costumam possuir um investimento considerável, seja na contratação de serviços, pessoas ou tecnologias. A cultura do uso de dados é essencial para que um programa de gestão de dados seja institucionalizado e efetivamente utilizado pela organização, caso contrário, investimentos serão desperdiçados. Muitos conceitos da preservação e curadoria digital poderiam ser aplicados no gerenciamento de dados, mais especificamente no corpo de conhecimento do DAMA. Após as análises realizadas para cada área de conhecimento do DMBOK, esta pesquisa sugere o uso das características de preservação e curadoria digital no modelo do DMBOK.

De acordo com os conceitos da Governança no contexto do DMBOK, essa área engloba muitas responsabilidades envolvendo autoridade, controle, monitoramento e garantia da execução de atividades relacionadas à gestão dos ativos de dados em uma organização. Conforme observado nos Quadro 12 e Quadro 13, o DMBOK não aborda diretamente objetivos e atividades relacionados à preservação e curadoria digital. Diante de vários desafios expostos nesta pesquisa relacionados ao tratamento de dados a longo prazo, a sugestão para essa área de conhecimento do DMBOK é a incorporação da cultura da preservação a longo prazo, podendo inserir de forma clara e direta objetivos visando o tratamento de dados de forma permanente e inserindo atividades claras de curadoria digital no intuito de demonstrar os caminhos práticos para alcançar a preservação a longo prazo.

Estas proposições seriam parte da essência da área e deveriam ser incorporadas para garantir o sucesso da nova cultura no tratamento dos dados, podendo assim todas as áreas obter benefícios principalmente visando projetos que necessitam de massa de dados com histórico de qualidade e de acesso eficiente.

A área de Armazenamento e Operação do DMBOK destaca a importância de uma base consistente que garanta todas as demais atividades de gerenciamento de dados. É através de uma base sólida, confiável e disponível que ações envolvendo dados serão realizadas a longo prazo. Embora haja a aplicação do armazenamento para domínios diferentes, os conceitos e objetivos básicos são iguais para um

repositório de armazenamento de dados, entretanto, para o DMBOK, uma atenção é sugerida em torno do entendimento do tipo de dado que o armazenamento está recebendo. Onde se descreve "Entender os requisitos", poderiam ser aplicados os conceitos de preservação digital em prol de compreender a comunidade que irá consumir o determinado dado, aumentando as possibilidades da aplicação de modelos específicos de armazenamento e operação de dados para uso futuro em projetos de análises de dados e aprendizado de máquina concentrados em objetivos específicos de negócio para uma organização. As atividades de curadoria poderiam ser aplicadas para garantir ações de ciclo de vida observando o armazenamento e as características dos domínios específicos inseridos no contexto da preservação digital.

Compreender como trafegar dados entre ambientes distintos é o tema central da área do DMBOK. Há um ecossistema amplo para o cumprimento deste objetivo e envolve diretamente a estrutura de dados. Esta é uma área que conecta as outras áreas pelo seu critério de interoperabilidade e a partir disso, atividades importantes e sensíveis são desenvolvidas. Nos modelos de preservação e curadoria visitados na pesquisa, observou-se que não há um detalhamento específico do tratamento da interoperabilidade em si mas há uma ênfase maior em atividades que envolvem a descrição de dados para o entendimento de sua origem para que seja devidamente armazenado em seu destino. A sugestão para esta área do DMBOK é o reforço de políticas que atendam o suporte a longo prazo das integrações, haja visto que muitos suportes ficam obsoletos, os sistemas e processos de integração precisam estar aptos para acompanhar o ecossistema dos ambientes até sua atualização tecnológica.

A abordagem do DMBOK em relação à Segurança é apresentada em várias áreas e de forma completa através de sua área específica. Aspectos dos modelos de preservação e curadoria estão presentes no modelo, mas são muito mais abrangentes e discutidos no DMBOK do que nos modelos OAIS e CVCD/DCC propriamente ditos. Devido às constantes técnicas de invasão e roubo de dados, sugere-se para essa área o reforço de políticas que enfatizem a preservação de ativos digitais a longo prazo e a inclusão de aspectos de ações de curadoria para a constante sustentação do ciclo de vida da segurança dos dados.

A coleta de informações de descrição para o uso em dados é apresentada como um fator de impacto positivo para várias áreas da gestão de dados. Sua aplicação em projetos estruturais e analíticos pode aprimorar o entendimento dos dados e gerar classificações mais efetivas da construção das informações. Os modelos de preservação e curadoria expõem fatores de uso de metadados com foco na longevidade das informações, porém o DMBOK não enfatiza essa característica em seus objetivos e atividades. Sugere-se a incorporação do uso de metadados visando a permanência dos dados, com foco em uma comunidade específica, para acelerar a organização dos dados na construção de informações de maior qualidade.

O entendimento do tipo de dados que estão sendo gerenciados é apresentado como um ponto de partida importante para o gerenciamento adequado. Em uma organização, as políticas e normas distintas em várias áreas de negócio aumentam e criam a complexidade da supervisão da gestão de dados. Aspectos da curadoria e preservação digital podem ser incorporados em um programa de gestão de dados, com ênfase no momento do projeto de modelagem. Os chamados criadores de dados têm um papel importante nos modelos de preservação e curadoria apresentados, e é a partir desse momento que o restante do ecossistema será criado. A área de Modelagem e Design do DMBOK busca conhecer o dado para adequar e tornar funcional o ecossistema, porém não discute com os criadores de dados modelos que possam se adaptar a um ambiente existente. Considerando as ações de curadoria e a preservação de dados a longo prazo, o DMBOK pode absorver e dar maior ênfase na adaptação dos dados no momento de sua criação para serem aplicados em modelos existentes, a fim de auxiliar as organizações a maximizar seus investimentos já realizados.

A abordagem do DMBOK para a Qualidade de Dados tenta abranger a maioria dos casos, visando um consumo de informações precisas que entregam valor para as organizações. Apesar de possuir atividades bem definidas, a definição dos objetivos de negócio é deixada para as etapas de identificação e planejamento das atividades. Os modelos de preservação e curadoria apresentam objetivos definidos e, portanto, poderiam ser incorporados ao DMBOK sob o critério de um olhar de preservação a longo prazo.

## 5 CONSIDERAÇÕES FINAIS

Esta pesquisa foi realizada para compreender o corpo de conhecimento em gestão de dados DMBOK e analisar suas relações e aproximações em relação aos modelos de preservação e curadoria digital OAIS e CVCD/DCC. O DMBOK geralmente é direcionado para as áreas de tecnologia da informação e administração, concentrando processos, pessoas e tecnologia. Contudo, foram identificadas muitas oportunidades para reflexão sobre sua aproximação com outras áreas do conhecimento, como a ciência da informação, preservação e curadoria digital.

A temática da preservação e curadoria digital, abundantemente encontrada no campo da Ciência da Informação, carrega aspectos que expandem o contexto tecnológico e de processos, centralizando os debates na garantia dos dados a longo prazo e nas ações para alcançar esse objetivo. A garantia da contínua geração de conhecimento requer memória informacional permanente e muitos agentes dessa missão deixaram suas contribuições, das quais o mundo moderno ainda se beneficia, seja em uma descoberta arqueológica ou em algum insight de negócio em um grande volume de dados histórico de uma organização. Ambos os exemplos possuem uma característica comum de longevidade de seus registros, contudo, trabalhar com o mundo digital requer resiliência em uma sociedade mergulhada em tecnologias voláteis.

A primeira parte da pesquisa se concentrou no entendimento dos conceitos e objetivos do Corpo de Conhecimento em Gestão de Dados do DAMA (DMBOK), do modelo Open Archival Information System do CCSDS (OAIS) e do Ciclo de Vida de Curadoria do DCC (CVCD/DCC). O percurso realizado pôde apresentar um conhecimento com alta capacidade de aplicação para domínios distintos, como a Ciência da Informação, Administração e Tecnologia da Informação. Foi observado que o DMBOK oferece um detalhamento minucioso para cada área de conhecimento, permitindo a compreensão e implantação em uma organização que posiciona seus dados como ativos de valor. Devido ao ecossistema de produção, coleta, armazenamento e manipulação de dados estar envolvido em ambientes de tecnologia da informação, o corpo de conhecimento das áreas aborda

prioritariamente os processos, tecnologias e ferramentas de tratamento dos dados e não direciona diretamente as recomendações para um domínio específico.

Por sua vez, foi observado que os modelos OAIS e CVCD/DCC direcionam as recomendações e práticas para um domínio específico e se posicionam à luz da preservação digital e curadoria digital. O modelo OAIS foi observado tratando as ações de preservação centralizadas na informação e em seu devido armazenamento, para a manutenção de registros digitais a longo prazo. Suas entidades funcionais carregam seus papéis e responsabilidades com atividades definidas, as quais são integradas através de seus serviços e funções. O modelo OAIS não apresenta detalhes da materialização técnica das ações indicadas, as representações são conceituais e direcionadas a processos de preservação a longo prazo. Muitos "comos" permanecem em aberto, permitindo que haja uma expansão do OAIS para a formação de outros modelos ou até mesmo a incorporação de práticas envolvendo a preservação digital. O Ciclo de Vida de Curadoria do DCC concentra seu conhecimento nas atividades de curadoria digital, na identificação de ações e meios tecnológicos pelos quais o objeto digital pode passar. Foi observado que este modelo também não apresenta detalhes técnicos dos meios para os quais as ações de curadoria são realizadas, mas apresenta uma riqueza de conceitos sobre todo o processo para garantir a manutenção a longo prazo de ativos digitais.

A segunda parte da pesquisa foi direcionada para a realização de aproximações das áreas do DMBOK com os objetivos e atividades encontrados nos modelos de preservação e curadoria digital. Foi possível observar que em muitas áreas do DMBOK existem características dos modelos OAIS e CVCD/DCC e que em algumas delas há uma relação entre as entidades e as ações dos modelos analisados. Neste ponto da pesquisa, a contribuição que a preservação e curadoria digital podem trazer para o contexto da gestão de dados para uma organização foi facilmente observada. A compreensão do propósito adequado e dos 'porquês' as atividades são realizadas no cotidiano de uma organização revitaliza seus principais objetivos de negócio, e nesse quesito, os modelos OAIS e CVCD/DCC são empregados com muito pragmatismo. O uso para um domínio específico reforça e confere consistência aos modelos de preservação e curadoria digital, os quais podem viabilizar novos modelos aplicados para um contexto específico.

A terceira parte da pesquisa teve como objetivo a entrega de sugestões de incorporação de características dos modelos OAIS e CVDC/DCC em áreas do DMBOK a fim de enfatizar a cultura da preservação a longo prazo de dados. O uso de dados para análises avançadas, projetos de big data e, mais recentemente, de inteligência artificial, demanda dados históricos e de qualidade para que os modelos possam ser aplicados com alta precisão. A mudança de sistemas de informação devido a novas versões e novos produtos de mercado carrega consigo impactos em seu ambiente de tecnologia, tornando os dados vulneráveis a perdas, falta de integridade e contexto.

O DMBOK é uma fonte rica de conhecimento em gestão de dados para as organizações e sua importância pode ser maximizada através das práticas de preservação e curadoria digital. O diálogo com as abordagens de preservação e curadoria digital pode realizar uma mudança cultural entre pessoas, processos e tecnologia no tratamento de dados a longo prazo. Esta pesquisa pode apoiar na expansão do DMBOK, dando capacidade para organizações desenvolverem uma gestão de dados com aspectos de preservação e curadoria digital aproximando os modelos OAIS e CVCD/DCC para as áreas do Corpo de Conhecimento DMBOK.

## REFERÊNCIAS BIBLIOGRÁFICAS

ABCOUWER, A. W; MAES, R.; TRUIJENS, J. Contouren Van Een Generiek Model Voor Informatiemanagement. **PrimaVera Working Paper**, v. 97-07, 1997. Disponível em: <https://hdl.handle.net/11245/1.134544>. Acesso em: 20 ago. 2022.

ABPMP Brasil. **BPM CBOK**: Guia para o Gerenciamento de Processos de Negócio Corpo Comum de Conhecimento ABPMP BPM CBOK v3.0 1ª Edição. Brasil, ABPMP, 2013.

AIKEN, Peter; BILLINGS, W. Juanita. **Monetizing Data Management**: Unlocking the Value in Your Organization's Most Important Asset. 1 ed. New Jersey, 2013.

BALL, Alex. Preservation and Curation in Institutional Repositories. **DCC STATE OF THE ART REPORT**, versão 1.3, 2010. Disponível em: <https://www.dcc.ac.uk/sites/default/files/documents/reports/irpc-report-v1.3.pdf>. Acessado em: 20 ago. 2022.

BATINI, Carlo et al. **Data and Information Quality**: Dimensions, Principles and Techniques. Switzerland: Springer International Publishing, 2016.

BÉLANGER, France; CROSSLER, Robert E.; CORREIA, John. Privacy Maintenance in Self-Digitization: The Effect of Information Disclosure on Continuance Intentions. **ACM SIGMIS Database**: the DATABASE for Advances in Information Systems, v. 52, ed. 2, p. 7-24, 2021. Disponível em: <https://doi.org/10.1145/3462766.3462769>. Acessado em: 15 jan. 2023.

BIRCH, Kean., COCHRANE, DT., WARD, Callum. **Data as asset? The measurement, governance, and valuation of digital personal data by Big Tech**. Big Data & Society, v. 8 ed.1, 2021. Disponível em: <https://doi.org/10.1177/20539517211017308>. Acessado em: 12 jul 2023.

BORGMAN, Christine L. **Big Data, Little Data, No Data**: Scholarship in the Networked World. Londres: The MIT Press, 2015.

BUCKLAND, Michael. **Information and Society**. ed. The MIT Press Essential Knowledge Series, Londres: The MIT Press, 2017.

CHECKLAND, Peter, HOLWELL, Sue. Data, Capta, Information and Knowledge. *In*: HINTON, Matthew. **Introducing Information Management**. Routledge. Nova Iorque: The Open University, 2011.

CHEN, C.L. Philip; ZHANG, Chun-Yang. Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. **Information Sciences**, v. 275, p. 314-347, 2014. Disponível em: <https://doi.org/10.1016/j.ins.2014.01.015>. Acesso em: 10 out. 2022.



CIURIAK, Dan. **Rethinking Industrial Policy for the Data-Driven Economy**. SSRN 2018. Disponível em: <http://dx.doi.org/10.2139/ssrn.3223072>. Acessado em: 02 ago 2023.

CONWAY, P. Archival quality and long-term preservation: a research framework for validating the usefulness of digital surrogates. **Archival Science**, v. 11, p. 293–309, 2011. Disponível em: <https://doi.org/10.1007/s10502-011-9155-0>. Acessado em: 20 ago. 2022.

CORETRUSTSEAL. About. Coretrustseal, 2022. Disponível em: <<https://www.coretrustseal.org/about/> >. Acesso em: 15 de set. de 2022.

CRL, THE CENTER FOR RESEARCH LIBRARIES. **Trustworthy Repositories Audit & Certification: Criteria and Checklist**, versão 1, 2007. Disponível em: [https://www.crl.edu/sites/default/files/d6/attachments/pages/trac\\_0.pdf](https://www.crl.edu/sites/default/files/d6/attachments/pages/trac_0.pdf). Acessado em: 20 ago. 2022.

CUPOLI, Patricia; EARLEY, Susan; HENDERSON, Deborah. **DAMA-DMBOK2 Framework**. 2014. Disponível em: <https://www.itmatters.com.au/resources/DAMA-DMBOK2-Framework-V2-20140317-FINAL.pdf>. Acesso em: 15 set. 2022.

DAMA International. **DAMA-DMBOK: Data Management Body of Knowledge 2<sup>nd</sup> Edition**. Nova Jersey: Technics Publications, 2017.

DAVENPORT, Thomas H. **Ecologia da Informação: Por que só a tecnologia não basta para o sucesso na era da informação**. Tradução: Bernadette Siqueira Abrão. São Paulo: Futura, 1998. Título original: Information ecology.

DERIDDER, Jody L. **Digital Curation Fundamentals**. Londres: Rowman & Littlefield Publishers, 2018.

DILLO, Ingrid; LEEUW, Lisa de. CoreTrustSeal. **Mitteilungen der Vereinigung Österreichischer Bibliothekarinnen & Bibliothekare**, v. 71, n. 1, p.162-170, 2018. Disponível em: <http://eprints.rclis.org/34431>. Acessado em: 15 set. 2022.

ENAP. **Governança de Dados na Transformação Digital: Introdução à Gestão de Dados**. Brasília, DF: Fundação Escola Nacional de Administração Pública, 2022.

ENAP. **Governança de Dados: Princípios, Importância e Desafios do Gerenciamento de Dados, Módulo 2**. Brasília, DF: Fundação Escola Nacional de Administração Pública, 2021. Disponível em: <http://repositorio.enap.gov.br/handle/1/7092>. Acessado em: 13 out. 2022.

FIGUEIREDO, Glauca B. de *et al.* Aligning DMBOK and Open Government with the FAIR Data Principles. *In*: GUIZZARDI, Giancarlo; GAILLY, Frederik; MACIEL, Rita S. P. (orgs). **Advances in Conceptual Modeling: ER 2019 Workshops FAIR, MREBA, EmpER, MoBiD, OntoCom, and ER Doctoral Symposium Papers**, Salvador, Brasil,

2019. Disponível em: <https://doi.org/10.1007/978-3-030-34146-6>. Acessado em: 21 ago. 2022.

FLORIDI, Luciano. **The Logic of Information: A Theory of Philosophy as Conceptual Design**. Oxford: Oxford University Press, 2019.

FRICKÉ, Martin. The knowledge pyramid: the DIKW hierarchy. **ISKO Encyclopedia of Knowledge Organization**, 2018. Disponível em: <http://www.isko.org/cyclo/dikw>. Acesso em: 12 mar. 2022.

GALINDO, Marcos. Legado Digital: O futuro que nos é presente. *In*: SIEBRA, Sandra de Albuquerque, BORBA, Vildeane da Rocha (orgs.). **Preservação Digital e suas facetas**. São Carlos: Pedro & João Editores, 2021.

GIARETTA, David. **Thinking Digital Preservation: for the serious, intelligent, digital preserver**. Estados Unidos da América: Ollie Trotman Library of Congress, 2022.

GOLDENBERG, Mirian. **A arte de pesquisar: Como fazer pesquisa qualitativa em Ciências Sociais**, 8 ed. Rio de Janeiro: Record, 2004.

GRACY, Karen F.; KAHN, Miriam B. Preservation in the Digital Age: A Review of Preservation Literature. **Library Resources and Technical Services**, v. 56, n.1, 2012. Disponível em: <https://journals.ala.org/index.php/lrts/article/view/5495/6745>. Acessado em: 10 abr. 2022.

HARRIS, Jeanne G.; DAVENPORT, Thomas H. **Competing on Analytics: Updated, with a New Introduction (Updated, with a New Introduction: The New Science of Winning)**. Boston: Harvard Business Review, 2017.

HASTIE, Trevor; TIBSHIRANI, Robert; FRIEDMAN. **The Elements of Statistical Learning: Data Mining, Inference, and Prediction**. 2. Ed. Nova Iorque: Springer Series in Statistics. Springer, 2009.

HENDERSON, J. C.; VENKATRAMAN N. Strategic alignment: Leveraging information technology for transforming organizations. **IBM Systems Journal**, v. 32, ed. 1, 1993. Disponível em: <https://ieeexplore.ieee.org/document/5387398>. Acessado em: 21 ago. 2022.

HIGGINS, S. The DCC Curation Lifecycle Model. **The International Journal of Digital Curation**, v.3, ed. 1, p.134-140, 2008. Disponível em: <https://doi.org/10.2218/ijdc.v3i1.48>. Acessado em: 20 nov. 2021.

JANTZ, Ronald; GIARLO, Michael J. Digital Preservation: Architecture and Technology for Trusted Digital Repositories. **D-Lib Magazine**, v. 11, n. 6, 2005. Disponível em: <http://www.dlib.org/dlib/june05/jantz/06jantz.html>. Acessado em: 20 nov. 2021.

JUIZ, Carlos; TOOMEY, Mark. To govern IT, or not to govern IT? **Communications of the ACM**, v. 58, n. 2, p. 58-64, 2015. Disponível em: <https://dl.acm.org/doi/10.1145/2656385>. Acesso em: 18 abr. 2023.

KITCHIN, Rob. Big Data, new epistemologies and paradigm shift. **Big Data & Society**, v. 1, abr./jun.2014. Disponível em: <https://journals.sagepub.com/doi/10.1177/2053951714528481>. Acesso em: 16 jan. 2023.

KOKINA, Julia; DAVENPORT, Thomas H. The Emergence of Artificial Intelligence: How Automation Is Changing Auditing. **Journal of Emerging Technologies in Accounting**, v. 14, n. 1, p. 115-122, 2017. Disponível em: <https://doi.org/10.2308/jeta-51730>. Acesso em: 21 ago. 2022.

LADLEY, John. **DATA GOVERNANCE: How to Design, Deploy, and Sustain an Effective Data Governance Program, Second Edition**. United Kingdom: Elsevier, 2020.

MAYER-SCHÖNBERGER, Viktor; CUKIER, Kenneth. **Big Data: The Essential Guide to Work, Life and Learning in the Age of Insight**. Londres: John Murray Press, 2017.

MCHUGH, Andrew *et al.* Bringing Self-assessment Home: Repository Profiling and Key Lines of Enquiry within DRAMBORA. **The International Journal of Digital Curation**, v. 3, ed. 2, p.130-142, 2008. Disponível em: <http://www.ijdc.net/article/view/93/64>. Acessado em: 18 fev. 2023.

MCKIERNAN, Erin C. *et al.* Point of View: How open science helps researchers succeed. **eLife**, 2016. Disponível em: <https://doi.org/10.7554/eLife.16800>. Acessado em: 18 fev. 2023.

OLIVER, Gillian; HARVEY, Ross. **Digital Curation**. 2 ed. Chicago: ALA Neal-Schuman, 2016.

OTTO, Boris. A Morphology Of The Organisation Of Data Governance. **ECIS 2011 Proceedings**, 2011. Disponível em: <https://aisel.aisnet.org/ecis2011/272>. Acessado em: 13 abr. 2022.

PANIAN, Zeljko. Some Practical Experiences in Data Governance. **World Academy of Science, Engineering and Technology**, 2010. Disponível em: <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=173626431ed114b232c1dab8ae60a005a0d593dc>. Acessado em: 18 fev. 2023.

PATEL, M *et al.* The Role of OAIS Representation Information in the Digital Curation of Crystallography Data. **2009 Fifth IEEE International Conference on e-Science**, p. 132-139, 2009. Disponível em: <https://doi.org/10.1109/e-Science.2009.27>. Acessado em: 10 jul. 2022.

PHILLIPS, Megan *et al.* The NDSA Levels of Digital Preservation: Explanation and Uses. **OSFHOME**, p.216-222, 2013. Disponível em: <https://osf.io/dpnqs>. Acessado em: 11 out. 2022.

PENNOCK, M. Digital Curation: A Life-Cycle Approach to Managing and Preserving Usable Digital Information. **Library & Archives Journal**, ed. 1, 2007. Disponível em: [http://www.ukoln.ac.uk/ukoln/staff/m.pennock/publications/docs/lib-arch\\_curation.pdf](http://www.ukoln.ac.uk/ukoln/staff/m.pennock/publications/docs/lib-arch_curation.pdf). Acessado em: 12 out. 2022.

PROJECT Management Institute. **Um Guia do Conhecimento em Gerenciamento de Projetos**: Guia PMBOK Sexta Edição. Newton Square, PA. Project Management Institute, Inc., 2017.

ROCHA, Rafael P. da; PIRES, Cássio de O. Finalidade e Atividades da Curadoria Digital na Perspectiva de sua Implantação em uma Instituição. **Brazilian Journal of Information Science**: research trends, v. 14, n. 4, 2020. Disponível em: <https://doi.org/10.36311/1940-1640.2020.v14n4.10857>. Acessado em: 21 ago. 2022.

RODRIGUES, Adriana A.; DUARTE, Emeide N.; DIAS, Guilherme A. Desafios da gestão de dados na era do Big Data: perspectivas profissionais. **Informação & Tecnologia – Especial Enancib 2017**, v. 4, n. 2, p. 63–79, 2018. Disponível em: <https://periodicos.ufpb.br/index.php/itec/article/view/40538>. Acessado em: 12 nov. 2022.

SAYÃO, L. F.; SALES, L. F. Curadoria Digital: um novo patamar para preservação de dados digitais de pesquisa. **Informação & Sociedade**, v. 22, n.3, 2012. Disponível em: <https://www.brapci.inf.br/index.php/res/v/92680>. Acessado em: 21 jan. 2023.

SCHWAB-MCCOY, Aimee; BAKER, Catherine M. & GASPER Rebecca E. Data Science in 2020: Computing, Curricula, and Challenges for the Next 10 Years. **Journal of Statistics and Data Science Education**, v. 29, n. s1, 2021. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/10691898.2020.1851159>. Acessado em: 29 nov. 2022.

SIEBRA, Sandra de A. Curadoria Digital: uma área em expansão. **Archeion Online**, v. 6, n. 2, p. 1-6, 2019. Disponível em: <https://doi.org/10.22478/ufpb.2318-6186.2019v6n2.47089>. Acessado em: 21 ago. 2022.

SIEBRA, Sandra de Albuquerque; BORBA, Vildeane da Rocha (orgs.). **Preservação Digital e suas facetas**. São Carlos: Pedro & João Editores, 2021.

SILVA, Edna L.; MENEZES, Estera M. **Metodologia da Pesquisa e Elaboração de Dissertação**, 4. ed. rev. atual, Florianópolis: UFSC, 2005.

TAVARES, A. L. L.; FREIRE, I. M. Pesquisa sobre preservação digital no Brasil: reflexões preliminares. **Encontro Nacional de Pesquisa e Pós-graduação em**

**Ciência da Informação**, n. XX ENANCIB, 2019. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/123370>. Acessado em: 21 ago. 2022

THE CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEMS. **REFERENCE MODEL FOR AN OPEN ARCHIVAL INFORMATION SYSTEM (OAIS)**: Recommended Practice CCSDS 650.0-M-2. Washington, DC: CCSDS Secretariat, 2012.

TIBBO, Helen R. Digital Curation Education and Training: From Digitalization to Graduate Curricula to MOOCs. **Internal Journal of Digital Curation**, v. 10, n. 1, 2015. Disponível em: <https://doi.org/10.2218/ijdc.v10i1.352>. Acessado em: 21 ago. 2022.

TUOMI, Ilkka. Data is More Than Knowledge: Implications of the Reversed Knowledge Hierarchy for Knowledge Management and Organizational Memory. **Journal of Management Information Systems**, v. 16, n. 3, p. 103-117, inverno 1999/2000. Disponível em: <https://www.jstor.org/stable/40398446> Acesso em: 12 fev. 2023.

TUREL, Ofir; LIU, Peng & BART, Chris. Board-Level Information Technology Governance Effects on Organizational Performance: The Roles of Strategic Alignment and Authoritarian Governance Style. **Information Systems Management**, v. 34, p. 117-136, 2017. Disponível em: <https://doi.org/10.1080/10580530.2017.1288523>. Acesso em: 5 fev. 2023.

WANG, Richard Y. & STRONG, Diane M. Beyond Accuracy: What Data Quality Means to Data Consumers. **Journal of Management Information Systems**, v. 12, ed. 4, p. 5-33, 1996. Disponível em: <https://doi.org/10.1080/07421222.1996.11518099>. Acesso em: 26 abr. 2023.

## APÊNDICE A – MODELAGEM DO REPOSITÓRIO DE INFORMAÇÃO DAS BUSCAS REALIZADAS

Foi desenhado um repositório para a coleta de informações e organização dos dados pesquisados.



Fonte: Elaborado pelo autor, 2023

## APÊNDICE B – EXEMPLO DO REPOSITÓRIO DE INFORMAÇÕES PREENCHIDO

id_repositorio	nome_repositorio	website_repositorio	tipo_acesso
1	Scielo	<a href="https://search.scielo.org/">https://search.scielo.org/</a>	publico
2	Periódicos CAPES	<a href="https://www-periodicos-capes-gov-br.ez45.periodicos.capes.gov.br/">https://www-periodicos-capes-gov-br.ez45.periodicos.capes.gov.br/</a>	login ufrgs
3	Repositório ENAP	<a href="https://repositorio.enap.gov.br/">https://repositorio.enap.gov.br/</a>	publico
4	Google Scholar	<a href="https://scholar.google.com.br/">https://scholar.google.com.br/</a>	publico
5	International Journal of Digital Curation	<a href="http://www.ijdc.net/">http://www.ijdc.net/</a>	publico
6	ICIC Express Letters	<a href="http://www.icicel.org/">http://www.icicel.org/</a>	publico
7	ProQuest	<a href="https://www.proquest.com/">https://www.proquest.com/</a>	publico
8	Seres Unit	<a href="https://www.seres-unit.com/">https://www.seres-unit.com/</a>	publico
9	RU Radbound Universiteit	<a href="https://www.ru.nl">https://www.ru.nl</a>	publico
10	Revista Científica Hospital Santa Izabel	<a href="https://revistacientifica.hospitalsantaizabel.org.br/">https://revistacientifica.hospitalsantaizabel.org.br/</a>	publico
11	SBCOPENLIB	<a href="https://sol.sbc.org.br/">https://sol.sbc.org.br/</a>	publico
12	CIKI	<a href="https://proceeding.ciki.ufsc.br/">https://proceeding.ciki.ufsc.br/</a>	publico
13	UFMG	<a href="https://repositorio.ufmg.br/">https://repositorio.ufmg.br/</a>	publico
14	UFPB	<a href="https://repositorio.ufpb.br/">https://repositorio.ufpb.br/</a>	publico
15	Revista Gestão e Conhecimento	<a href="https://revistagc.com.br/ojs/index.php/rqc">https://revistagc.com.br/ojs/index.php/rqc</a>	publico
16	ATTENA	<a href="https://repositorio.ufpe.br/">https://repositorio.ufpe.br/</a>	publico
17	Research Gate	<a href="https://www.researchgate.net/">https://www.researchgate.net/</a>	publico
18	USP	<a href="https://www.leses.usp.br/">https://www.leses.usp.br/</a>	publico
19	Biblioteca Digital de Teses e Dissertações UCB	<a href="https://btd.ucb.br/8443/jspui/">https://btd.ucb.br/8443/jspui/</a>	publico
20	Repositório Digital de Monografias da FJP	<a href="http://monografias.fjp.mg.gov.br/">http://monografias.fjp.mg.gov.br/</a>	publico
21	Revista Águas Subterrâneas	<a href="https://aquassubterraneas.abas.org/asubterraneas">https://aquassubterraneas.abas.org/asubterraneas</a>	publico

Fonte: Elaborado pelo autor, 2023

id_busca	id_repositorio	termo_busca	resultado_busca	preselecao_busca	filtros_busca	data_busca
1	1	curadoria digital	3	2	não	null
2	1	preservação digital	16	6	não	null
3	1	gestão de dados	41	0	não	null
4	1	dmbok	0	0	não	null
5	2	curadoria digital	236	18	não	null
6	2	curadoria digital	102		sim	
7		preservação digital				null
8		gestão de dados				null
9		dmbok				null
10						null
11						null
12	5	data management	49	1	sim	01/04/2023
13	5	digital preservation	19	6	sim	01/04/2023
14	5	digital curation	24	5	sim	01/04/2023
15	5	lifecycle	39		não	01/04/2023
16						
17	4	dmbok	41	4	sim	02/04/2023
18	4	dmbok	128	16	sim	02/04/2023
19						
20						
21						

Fonte: Elaborado pelo autor, 2023

id_filtro	id_busca	datainicio_busca	datafim_busca	filtropers1_busca	filtropers2_busca	filtropers3_busca
1				relevância		
2				revisão por pares		
3				artigos		
4	12	01/01/2016	null			
5	13	01/01/2016	null			
6	14	01/01/2016	null			
7	15	null	null			
8	17	null	null	relevância	artigos de revisão	
9	18	null	null	relevância	qualquer tipo	páginas em português
10	19					
11	20					
12	21					
13	22					
14	23					
15	24					

Fonte: Elaborado pelo autor, 2023