



UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO DE TECNOLÓGICO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA E GESTÃO DO
CONHECIMENTO

Lidiane Visintin

**MODELO DE MATURIDADE DE DADOS ABERTOS: UMA MATRIZ DE
REFERÊNCIA PARA ORGANIZAÇÕES**

Florianópolis
2021

Lidiane Visintin

**MODELO DE MATURIDADE DE DADOS ABERTOS: UMA MATRIZ DE
REFERÊNCIA PARA ORGANIZAÇÕES**

Tese submetida ao Programa de Pós-graduação em Engenharia e Gestão do Conhecimento da Universidade Federal de Santa Catarina para a obtenção do título de doutor em Engenharia e Gestão do Conhecimento.

Orientador: Prof. José Leomar Todesco, Dr.

Coorientador: Prof. Fernando Álvaro Ostuni Gauthier, Dr.

Florianópolis

2021

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Visintin, Lidiane

MODELO DE MATURIDADE DE DADOS ABERTOS : UMA MATRIZ DE
REFERÊNCIA PARA ORGANIZAÇÕES / Lidiane Visintin ;
orientador, José Leomar Todesco, coorientador, Fernando
Álvaro Ostuni Gauthier, 2021.

211 p.

Tese (doutorado) - Universidade Federal de Santa
Catarina, Centro Tecnológico, Programa de Pós-Graduação em
Engenharia e Gestão do Conhecimento, Florianópolis, 2021.

Inclui referências.

1. Engenharia e Gestão do Conhecimento. 2. Dados
Abertos. 3. Maturidade. 4. Áreas Foco. 5. Governança de
Dados. I. Todesco, José Leomar . II. Gauthier, Fernando
Álvaro Ostuni . III. Universidade Federal de Santa
Catarina. Programa de Pós-Graduação em Engenharia e Gestão
do Conhecimento. IV. Título.

Lidiane Visintin

**MODELO DE MATURIDADE DE DADOS ABERTOS: UMA MATRIZ DE
REFERÊNCIA PARA ORGANIZAÇÕES**

O presente trabalho em nível de doutorado foi avaliado e aprovado por banca examinadora composta pelos seguintes membros:

Prof. Paulo Maurício Selig, Dr.
Universidade Federal de Santa Catarina

Prof. Denilson Sell, Dr.
Universidade Federal de Santa Catarina

Prof. José Maria Parente de Oliveira, Dr
Instituto Tecnológico de Aeronáutica

Prof. Bernadette Farias Lóscio, Dra
Universidade Federal de Pernambuco

Certificamos que esta é a **versão original e final** do trabalho de conclusão que foi julgado adequado para obtenção do título de doutor em Engenharia e Gestão do Conhecimento.

Coordenação do Programa de
Pós-Graduação

Prof. José Leomar Todesco, Dr.
Orientador

Florianópolis, 2021.

Este trabalho é dedicado aos meus pais, ao meu esposo
e aos meus sobrinhos.

AGRADECIMENTOS

Finalizar este trabalho não foi fácil, foi uma longa caminhada até este momento. Um caminho cheio de preocupações, angústias, muita ansiedade e incertezas, mas felizmente estou finalizando este ciclo. Assim o que me resta é agradecer, primeiramente a Deus pelas oportunidade e pela vida.

Agradeço a minha família por todo o apoio, mãe e pai vocês são demais. Agradeço a meu irmão ter me presenteado com dois sobrinhos, que são a minha distração e felicidade em diversos momentos. Agradeço ao meu esposo Murilo, por ter sobrevivido a convivência comigo, durante uma pandemia, ao trabalho remoto e a escrita da tese - (não posso esquecer de mencionar dos momentos de crises, mas esta parte foi mútua), obrigada por tudo!

Um agradecimento especial ao meu orientador José Leomar Todesco(Tite), agradeço pelas reuniões, por ter me escutado, pelas dicas, pelas correções e por todo esforço desempenhado na orientação deste trabalho. Agradeço ao meu coorientado professor Fernando Gauthier pela confiança.

Meu agradecimento aos amigos que fiz durante esta jornada, aos colegas dos laboratórios, entre estes: Marilei, Aline, Fernanda, Luciane, Jorge e Vinícius, dentre muitos outros deixo o meu muito obrigado, pelas nossa conversas, chimarrões e confraternizações. Também preciso agradecer aos meus colegas de trabalho do Instituto Federal Catarinense, por sempre me lembrarem que eu tinha uma tese para terminar, e por me auxiliarem em alguns momentos.

Agradeço aos especialistas entrevistados pela atenção, disponibilidade e sugestões que forneceram diferentes perspectivas sobre o trabalho desenvolvido. Agradeço a banca de qualificação que contribuiu com sugestões para a realização desta tese, bem como agradecer previamente a banca de defesa que aceitou o convite para contribuir e auxiliar no aperfeiçoamento do trabalho desenvolvido.

Por fim, agradeço ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo financiamento que possibilitou a realização de parte deste trabalho.

Não espere por uma vida sem problemas, isso não existe. Em vez disso, torça por uma vida cheia de problemas pequenos.(MARK MANSON)

RESUMO

Dados abertos podem ser vistos como um movimento para o compartilhamento do conhecimento e a ampliação da democracia, beneficiando tanto a sociedade quanto as organizações que disponibilizam dados. Este tema difundiu-se devido aos benefícios que podem ser observados e também devido a leis e regulamentações que alguns países impõem em busca de transparência, participação social, entre outros. No entanto, observa-se que há organizações que enfrentam desafios organizacionais - relacionados a processos internos, desafios técnicos - relacionados à ausência de metadados, a falta de interoperabilidade e a dificuldade em pesquisar e obter dados, entre outros. Deste modo percebe-se que essa perspectiva inibe o uso e também impacta na confiança dos dados que são disponibilizados nos portais. Tais inconsistências poderiam ser minimizadas com a definição, a implementação, o monitoramento e o aprimoramento das políticas e práticas para o tratamento de dados e com base neste contexto observa-se também que a governança de dados pode contribuir minimizar tais inconsistências. Neste sentido, o presente trabalho tem por objetivo propor um modelo de maturidade de dados abertos, considerando boas práticas aplicadas pelas organizações na abertura e na pós-publicação dos dados e princípios da governança de dados. A pesquisa foi conduzida através do *Design Science Research Methodology* (DSRM). Deste modo, o modelo desenvolvido permite avaliar políticas e processos, fatores organizacionais e padrões aplicados para dados abertos através de um questionário que viabiliza o preenchimento de uma matriz de referência para avaliação da maturidade. Para aplicar o modelo foi desenvolvida uma aplicação Web, esta que viabiliza a aplicação do questionário e a visualização do resultado obtido. Foram realizadas entrevistas com responsáveis por organizações ou departamentos que realizam a abertura de dados, totalizando três cenários de uso. Através destes cenários de uso foi possível observar o comportamento da estrutura elaborada, bem como coletar percepções e sugestões acerca do questionário aplicado, da estrutura do modelo e dos elementos que o compõem. Ainda para verificar, discutir, divulgar e obter *feedbacks* acerca do modelo foram realizadas entrevistas com 10 especialistas da área de dados abertos. Deste modo foi possível observar que as médias das percepções apontam que os especialistas concordam em relação à abrangência, relevância e a precisão da estrutura do modelo apresentado. Também foi possível verificar a importância dos elementos que compõem o modelo, dada pelas médias obtidas pelas percepções dos especialistas e ao final são elencadas as sugestões e comentários destes. Considerando a estrutura do modelo, os resultados obtidos através da obtenção dos cenários de uso e finalizadas as entrevistas com os especialistas concluiu-se que a pesquisa desenvolvida contribuiu fornecendo uma estrutura de ciclo de vida e de modelo de maturidade sendo que esta matriz e o ciclo de vida podem ser utilizados como modelos de referência, ou seja, podem ser estendidos, reduzidos ou até mesmo adaptados, permitindo assim a criação de estruturas mais específicas ou mais genéricas abrangendo as diferentes realidades.

Palavras-chave: Dados Abertos. Maturidade. Áreas Foco. Governança de Dados

ABSTRACT

Open data can be seen as a movement towards knowledge sharing and the expansion of democracy, benefiting both society and organizations that make data available. This theme has spread due to the benefits that can be observed and also due to laws and regulations that some countries impose in search of transparency, social participation, among others. However, it is observed that there are organizations that face organizational challenges - related to internal processes, technical challenges - related to the absence of metadata, the lack of interoperability and the difficulty in searching and obtaining data, among others. Thus, it is clear that this perspective inhibits the use and also impacts the reliability of the data that are made available on the portals. Such inconsistencies could be minimized with the definition, implementation, monitoring and improvement of policies and practices for the treatment of data and, based on this context, it is also observed that data governance can help to minimize such inconsistencies. In this sense, the present work aims to propose an open data maturity model, considering good practices applied by organizations in the opening and post-publication of data and data governance principles. The research was conducted through the Design Science Research Methodology (DSRM). In this way, the developed model allows the evaluation of policies and processes, organizational factors and standards applied to open data through a questionnaire that makes it possible to fill in a reference matrix for maturity assessment. To apply the model, a web application was developed, which enables the application of the questionnaire and the visualization of the obtained result. Interviews were conducted with those responsible for organizations or departments that perform data opening, where three scenarios of use were obtained. Through the usage scenarios, it was possible to observe the behavior of the elaborated structure, as well as collect perceptions and suggestions about the applied questionnaire, the model structure and the elements that compose it. Also to verify, discuss, disseminate and obtain feedback about the model, interviews were conducted with 10 experts in the area of open data. In this way, it was possible to observe that the averages of the perceptions show that experts agree in relation to the scope, relevance and precision of the structure of the model presented. It was also possible to verify the importance of the elements that make up the model, given by the averages obtained by the experts' perceptions and, at the end, their suggestions and comments are listed. Considering the structure of the model, the results obtained by obtaining the use scenarios and finalizing the interviews with experts, it was concluded that the research developed contributed by providing a lifecycle structure and maturity model, this matrix and the cycle of life can be used as reference models, that is, they can be extended, reduced or even adapted, thus allowing the creation of more specific or more generic structures covering different realities.

Keywords: Open Data. Maturity. Focus Areas. Data Governance

LISTA DE FIGURAS

Figura 1 – Estrutura da Tese.	31
Figura 2 – Nuvem LOD em 2009.	38
Figura 3 – Nuvem LOD em 2021.	39
Figura 4 – Dados na Web x Dados Abertos x Dados Conectados.	40
Figura 5 – Desafios de dados na Web.	50
Figura 6 – Inter-relação entre os domínios da governança.	52
Figura 7 – Relação entre governança e gestão.	54
Figura 8 – Governança de dados e gestão de dados.	55
Figura 9 – <i>Framework</i> de Governança de Dados.	58
Figura 10 – Resumo da evolução dos estudos acerca de modelos de maturidade.	61
Figura 11 – Cinco níveis do CMM.	65
Figura 12 – Três tipos de modelos de maturidade	67
Figura 13 – Modelo de maturidade da área de foco para o gerenciamento de arquitetura corporativa.	68
Figura 14 – Exemplificação das dependências	70
Figura 15 – Metamodelo do modelo de maturidade de áreas foco.	70
Figura 16 – Classificação baseada em estrelas	71
Figura 17 – Etapas da Metodologia <i>Design Science</i>	80
Figura 18 – Etapas da Pesquisa	81
Figura 19 – Resumo da análise dos trabalhos encontrados nas buscas realizadas.	90
Figura 20 – Etapas da construção de um modelo de maturidade	91
Figura 21 – Etapas da construção de um modelo de maturidade orientado a área foco	91
Figura 22 – Ciclo de vida de dados abertos.	97
Figura 23 – Estratégia e Apoio.	99
Figura 24 – Suporte a gestão e aos processos de publicação.	101
Figura 25 – Treinamento e apoio aos funcionários.	102
Figura 26 – Acessibilidade, interoperabilidade e padrões.	104
Figura 27 – Acessibilidade, interoperabilidade e padrões.	105
Figura 28 – Plataformas, ferramentas e serviços de dados abertos.	107
Figura 29 – Legislação, regulamentação e licença.	108
Figura 30 – Histórias de sucesso	109
Figura 31 – Incentivo ao uso	110
Figura 32 – Treinamento e suporte para usuários	111
Figura 33 – Feedback	112
Figura 34 – Pesquisa e Educação	113
Figura 35 – Colaboração	114

Figura 36 – Avaliação da iniciativa	115
Figura 37 – Ciclo de Vida de Dados Abertos	119
Figura 38 – Resumo das áreas foco e seus componentes.	124
Figura 39 – Modelo de Maturidade de Dados Abertos - Posicionamento de perguntas e dependências	155
Figura 40 – Representação do conhecimento abordado no modelo de maturidade.	158
Figura 41 – Cenário 1 - Área foco 1.	162
Figura 42 – Cenário 1 - Área foco 2.	162
Figura 43 – Cenário 1 - Área foco 3.	163
Figura 44 – Cenário 2 - Área foco 1.	164
Figura 45 – Cenário 2 - Área foco 2.	165
Figura 46 – Cenário 2 - Área foco 3.	166
Figura 47 – Cenário 3 - Área foco 1.	167
Figura 48 – Cenário 3 - Área foco 2.	167
Figura 49 – Cenário 3 - Área foco 3.	168

LISTA DE QUADROS

Quadro 1 – Desafios para a abertura de dados governamentais.	20
Quadro 2 – Definições e conceitos acerca de dados abertos.	24
Quadro 3 – Definições e conceitos acerca do modelo de maturidade.	25
Quadro 4 – Trabalhos que abordam discussões sobre modelos de maturidade	29
Quadro 5 – Trabalhos que abordam discussões sobre dados abertos e dados abertos conectados	30
Quadro 6 – Níveis de licença Creative Commons	35
Quadro 7 – Níveis de licença Open Data Commons	35
Quadro 8 – Trabalhos que abordam práticas e tratamento de dados para a pu- blicação	41
Quadro 9 – Boas Práticas para Dados na Web	46
Quadro 10 – Boas práticas para a publicação de dados abertos conectados . .	51
Quadro 11 – Métricas de Avaliação Norma UNE 178301:2015	73
Quadro 12 – Relação de domínios, subdomínios e variáveis do OD-MM	76
Quadro 13 – Trabalhos sobre boas práticas, ciclos de vida e fatores importantes para dados abertos	82
Quadro 14 – Trabalhos sobre governança de dados	83
Quadro 15 – Trabalhos sobre modelos de maturidade de áreas foco	84
Quadro 16 – Trabalhos resultantes após a análise do título e do resumo	86
Quadro 17 – Trabalhos selecionados para discussão	89
Quadro 18 – Comparativo entre fases e atividades-chave dos ciclos de vida . . .	117
Quadro 19 – <i>Stakeholders</i> do processo de abertura de dados	118
Quadro 20 – Definições acerca do escopo do modelo de maturidade	120
Quadro 21 – Definições acerca do design do modelo de maturidade	121
Quadro 22 – Definição dos Dados	126
Quadro 23 – Acesso e Entrega de Dados	127
Quadro 24 – Monitoramento e Medição	129
Quadro 25 – Gerenciamento de Mudança de Dados	130
Quadro 26 – 5 - Funções e Responsabilidades	132
Quadro 27 – Educação	133
Quadro 28 – Treinamento	134
Quadro 29 – Planejamento e Priorização da Abertura dos Dados	135
Quadro 30 – Planejamento da Publicação dos Dados	136
Quadro 31 – Fomento ao Uso	138
Quadro 32 – Licença	140
Quadro 33 – Proveniência e Qualidade	142
Quadro 34 – Versionamento	143

Quadro 35 – Identificação	144
Quadro 36 – Formato	145
Quadro 37 – Vocabulários	146
Quadro 38 – Metadados	147
Quadro 39 – Preservação	149
Quadro 40 – Enriquecimento	150
Quadro 41 – <i>Feedback</i>	151
Quadro 42 – Acesso	152
Quadro 43 – Matriz - Modelo de Maturidade de Dados Abertos	156
Quadro 44 – Descrição dos cenários de uso e dos entrevistados	160
Quadro 45 – Parte I - Percepções sobre aspectos do questionário e aspectos do modelo de maturidade	169
Quadro 46 – Parte II - Percepções sobre a importância dos elementos que compõem o modelo de maturidade	171
Quadro 47 – Perfil dos especialistas entrevistados	172
Quadro 48 – Estrutura do Modelo de Maturidade	174
Quadro 49 – Importância das áreas foco e dos componentes do modelo de maturidade	175
Quadro 50 – Trabalhos relacionados - comparativo	178
Quadro 51 – Seção I - Questionário e Estrutura do Modelo	204
Quadro 52 – Seção II - Importância das áreas foco e dos componentes do modelo de maturidade	205
Quadro 53 – Parte I - Perfil dos respondentes	207
Quadro 54 – Parte II - Importância das áreas foco e dos componentes do modelo de maturidade	208
Quadro 55 – Parte III - Estrutura do Modelo	210

LISTA DE TABELAS

Tabela 1 – Quantidade de arquivos resultantes para cada sintaxe de busca no ano de 2018	85
Tabela 2 – Quantidade de arquivos resultantes para cada sintaxe de busca no ano de 2021	86

LISTA DE ABREVIATURAS E SIGLAS

API	<i>Application Programming Interface</i>
CC	<i>Creative Commons</i>
CMM	<i>Capability Maturity Model</i>
CMMI	<i>Capability Maturity Model Integration</i>
CSV	<i>Comma Separated Values</i>
DSR	<i>Design Science Research</i>
EUA	Estados Unidos da América
FTP	<i>File Transfer Protocols</i>
HTML	<i>Hypertext Markup Language</i>
HTTP	<i>Hypertext Transfer Protocol</i>
IBM	<i>International Business Machines Corporation</i>
LGPD	Lei Geral de Proteção dos Dados
LOD	<i>Linked Open Data</i>
NSW	<i>New South Wales</i>
OCDE	Organização para a Cooperação e Desenvolvimento Econômico
OD-MM	<i>Open Data Maturity Model</i>
OGP	Open Government Partnership
PDF	<i>Portable Document Format</i>
PPEGC	Programa de Pós – Graduação em Engenharia e Gestão do Conhecimento
RDF	<i>Resource Description Framework</i>
SAMDA	Sistema de Apoio a Avaliação da Maturidade de Dados Abertos
SEI	<i>Software Enginneering Instituite</i>
SPARQL	<i>Protocol and RDF Query Language</i>
TI	Tecnologia da Informação
UFSC	Universidade Federal de Santa Catarina
UML	<i>Unified Modeling Language</i>
URI	<i>Uniform Resource Identifier</i>
W3C	<i>World Wide Web</i>
WoS	<i>Web of Science</i>
WTDO 2021	5º Workshop de Teses e Dissertações em Ontologias

SUMÁRIO

1	INTRODUÇÃO	17
1.1	CONTEXTUALIZAÇÃO	17
1.2	OBJETIVOS	23
1.2.1	Objetivo geral	23
1.2.2	Objetivos específicos	23
1.3	DELIMITAÇÃO DA PESQUISA	23
1.4	INEDITISMO, ORIGINALIDADE E CONTRIBUIÇÕES	26
1.5	ADERÊNCIA AO EGC	28
1.6	ORGANIZAÇÃO DO DOCUMENTO	30
2	FUNDAMENTAÇÃO TEÓRICA	32
2.1	DADOS ABERTOS	32
2.1.1	Licença Aberta	34
2.1.2	Dados Abertos x Dados Conectados	35
2.1.3	Dados Abertos Conectados	37
2.1.4	Os Ciclos de Vida de Dados Abertos	40
2.1.5	Boas Práticas para Dados na Web	45
2.2	GOVERNANÇA E GOVERNANÇA DE DADOS	52
2.2.1	<i>Framework</i> de Governança de Dados	56
2.3	MODELOS DE MATURIDADE	59
2.3.1	Definindo Modelos de Maturidade	61
2.3.1.1	Componentes Essenciais	62
2.3.1.2	Tipos de Modelos de Maturidade	64
2.3.1.3	Modelos de Maturidade de Áreas Foco	67
2.4	TRABALHOS RELACIONADOS	71
2.4.1	Classificação baseada em estrelas	71
2.4.2	Norma UNE 178301	72
2.4.3	Modelo de Maturidade de Dados Abertos	75
2.4.4	<i>Open Data Maturity Model</i>	76
3	PROCEDIMENTOS METODOLÓGICOS	79
3.1	ETAPAS DA PESQUISA	82
3.1.1	Etapa 1: Identificação do problema & Motivação	82
3.1.2	Etapa 2: Definir objetivos para uma solução	82
3.1.3	Etapa 3: Design & Desenvolvimento	90
3.1.4	Etapa 4: Demonstração	92
3.1.5	Etapa 5: Avaliação	92
3.1.6	Etapa 6: Comunicação dos resultados	92
4	CICLO DE VIDA DE DADOS ABERTOS	94

4.1	ATIVIDADES-CHAVE DO CICLO DE VIDA DE DADOS ABERTOS	98
4.1.1	Planejamento	98
4.1.2	Gerenciamento	100
4.1.3	Criação	103
4.1.4	Disponibilização/Manutenção	106
4.1.5	Uso	107
4.1.6	Aprimoramento	113
4.1.7	Descarte	116
4.2	COMPARATIVO ENTRE OS CICLOS DE VIDA	116
5	MODELO DE MATURIDADE DE DADOS ABERTOS	120
5.1	DEFINIÇÕES INICIAIS	120
5.2	ESTRUTURAÇÃO DO MODELO	122
5.2.1	Determinar as áreas foco	122
5.2.2	Determinar as capacidades e determinar as dependências	124
5.2.3	Posicionar as capacidades na matriz	154
6	APRESENTAÇÃO E AVALIAÇÃO DO MODELO DE MATURIDADE	159
6.1	APRESENTAÇÃO	159
6.1.1	Sistema de Apoio a Avaliação da Maturidade de Dados Abertos	
	SAMDA	159
6.1.2	Cenários de Uso	160
6.1.2.1	Percepções da pesquisadora sobre as entrevistas para obter os cenários de uso	168
6.1.2.2	Percepção dos Entrevistados nos Cenários de Uso	169
6.2	AVALIAÇÃO DO MODELO COM ESPECIALISTAS	172
7	CONSIDERAÇÕES FINAIS	180
7.1	CONCLUSÕES	180
7.2	TRABALHOS FUTUROS	182
	REFERÊNCIAS	184
	APÊNDICE A – SEÇÃO 1 - FORMULÁRIO APLICADO PARA OS	
	RESPONDENTES DE CADA CENÁRIO DE USO	204
	APÊNDICE B – SEÇÃO 2 - FORMULÁRIO APLICADO PARA OS	
	RESPONDENTES DE CADA CENÁRIO DE USO	205
	APÊNDICE C – SEÇÃO 1 - FORMULÁRIO APLICADO PARA OS	
	ESPECIALISTAS	207
	APÊNDICE D – SEÇÃO 2 - FORMULÁRIO APLICADO PARA OS	
	ESPECIALISTAS	208
	APÊNDICE E – SEÇÃO 3 - FORMULÁRIO APLICADO PARA OS	
	ESPECIALISTAS	210

1 INTRODUÇÃO

Nesta seção apresenta-se uma contextualização sobre a temática e a problemática da tese. Neste âmbito apresenta-se a pergunta de pesquisa, os objetivos (geral e específicos), o ineditismo, a originalidade e as contribuições, a delimitação ou escopo da pesquisa, a aderência ao Programa de Pós-graduação em Engenharia e Gestão do Conhecimento e por fim a estrutura deste documento.

1.1 CONTEXTUALIZAÇÃO

Os dados têm grande utilidade e alto valor porque fornecem os principais insumos para análises de indivíduos, instituições, empresas e da própria ciência, sendo que esta faz uso de dados para compreender e explicar o mundo. Ainda, por sua vez dados são utilizados para criar inovações, produtos, políticas e conhecimentos, sendo que estes moldam a forma de como as pessoas vivem.

Os dados são um recurso fundamental no mundo moderno (KITCHIN, 2014). No entanto, dada sua utilidade e valor, notam-se esforços e recursos dedicados a produzi-los e analisá-los. Visto que os dados formam a base da pirâmide do conhecimento, pois os dados precedem a informação, que precede o conhecimento, que precede o entendimento e a sabedoria (ADLER, 1986; WEINBERGER, 2014).

Com o uso de tecnologias e da Web, a criação, a análise e o consumo dos dados tornou-se uma realidade no cotidiano das organizações e das pessoas, sejam em forma de mensagem de texto, imagens, e-mails, até mesmo em rotas utilizadas por veículos que dispõem de GPS e entre outros. Neste sentido, as organizações fazem uso de dados para fornecerem produtos, serviços, reduzindo custos e controlando riscos. Agências governamentais, instituições de ensino e organizações sem fins lucrativos também precisam de dados para orientar suas atividades operacionais, táticas e estratégicas (MOSLEY *et al.*, 2009).

Considerando o contexto apresentado, diversas organizações, nos últimos anos, comprometidas com os anseios da sociedade passaram a promover a abertura de seus dados. O tema dados abertos surgiu em 1995 a partir de um documento da agência científica dos Estados Unidos da América (EUA) com o intuito de promover a troca aberta de dados geofísicos e ambientais entre agências governamentais para melhor compreensão de fenômenos associados (MACEDO; LEMOS, 2021). No entanto, observa-se que a abertura dos dados é comumente conhecida pela comunidade científica, visto que contribuições com estudos em andamento e até mesmo o compartilhamento de informações das mais diversas áreas viabilizam a ampliação do conhecimento, trazendo diversos benefícios para a sociedade. Desse modo, destaca-se que a comunidade acadêmica beneficia-se e explora a abertura e o compartilhamento de dados, permitindo que o conhecimento avance (CHIGNARD, 2013), antes mesmo de ter início as dis-

cussões sobre dados abertos. Contudo, o encontro entre teorias científicas, conceitos de software livre e código aberto- (*Openness*) viabilizaram o surgimento do conceito de dados abertos, que é amplamente discutido (CHIGNARD, 2013).

A abertura de dados pode ser vista como um movimento para compartilhamento de conhecimento e ampliação da democracia, beneficiando tanto a sociedade quanto as organizações que disponibilizam seus dados (BAUER; KALTENBÖCK, 2011). O que aumenta a responsabilidade, a transparência, a eficiência das organizações, além de incentivar o uso e a reutilização de dados e informações para ações dos mais diversos interesses (ATTARD *et al.*, 2015).

Destaca-se que atualmente confundem-se os conceitos de dados abertos e de *big data*. No entanto é válido considerar esses dois conceitos separadamente devido a suas características, por mais que compartilhem um objeto comum, os dados. O *big data* abrange as possibilidades oferecidas pela exploração de um volume de dados em crescimento exponencial (CHIGNARD, 2013). Dados abertos, por sua vez explora, discute e analisa a criação de valor independente da capacidade de compartilhar dados, e de disponibilizá-los a terceiros, e portanto não é primordial para este contexto o volume dos dados.

Neste sentido, tem-se que dados abertos necessitam atender dois critérios. O primeiro diz respeito a ser tecnicamente abertos, ou seja, estarem disponíveis em um formato padrão que possam ser recuperados, lidos e processados por máquinas, pois isso deve garantir que não sejam criadas barreiras sobre como os dados podem ser usados. O segundo critério diz respeito a serem legalmente abertos que significa que os dados disponíveis devem possuir uma licença, e esta licença deve permitir o seu uso gratuitamente, pois isto faz com que estes dados possam ser reutilizados sem restrições (ARMAN *et al.*, 2015; DIN *et al.*, 2015).

Dado o contexto, dados abertos podem fornecer benefícios à várias áreas. Pesquisas realizadas indicam impactos em setores como: educação, transporte, eletricidade, petróleo, saúde, e finanças (MANYIKA *et al.*, 2013; MACHADO *et al.*, 2019). O impacto positivo pode ser alcançado devido as pessoas obterem e processarem mais facilmente os dados. Neste contexto, o uso de dados abertos tem um grande valor econômico potencial, pois melhora o impacto do *big data* gerando transparência e expondo a variabilidade, permitindo a experimentação e possibilitando oportunidades de negócios (MANYIKA *et al.*, 2013; MACEDO; LEMOS, 2021).

Assim, pautas referentes a dados abertos passaram a compor a agenda de grandes organizações internacionais, como o G20 (Grupo das maiores economias do mundo), a Organização para a Cooperação e Desenvolvimento Econômico (OCDE) e do Banco Mundial (GRAY, 2014). Dados abertos tem um impacto positivo para à sociedade, pois pode melhorar a eficiência das administrações públicas, aumentar o crescimento econômico no setor privado, criar novos produtos baseados em dados e

contribuir para o espaço de formulação de políticas públicas (KAPOOR *et al.*, 2015). No entanto, dados abertos é um tema que necessita ser investigado, pois ainda há organizações em um estágio inicial acerca da abertura dos dados, sendo que há barreiras e desafios relacionados a dados abertos que precisam ser superados, para que as organizações obtenham sucesso através da abertura dos dados (KAASENBROOD *et al.*, 2015; WU, T.; MAHDZIR, 2018; HUSIN *et al.*, 2019).

A abertura de dados é amplamente difundida, compreendendo diferentes nações (HUIJBOOM; VAN DEN BROEK, 2011). Nesse cenário, as pessoas e organizações estão tendo acesso a dados de vários setores não só do contexto público. Observa-se este cenário através do número crescente de portais e de conjuntos de dados que estão acessíveis (CASTELLANI RIBEIRO *et al.*, 2015; TYGEL *et al.*, 2016; MCCRAE; ABELE, 2021), são exemplos: o DataPortals.org¹ que em maio de 2021 contabiliza cerca de 590 portais de dados abertos, considerando uma lista de portais acessíveis no mundo. Segundo Yi (2019) estima-se que os EUA em 2018 dispunham de 341.925 conjuntos de dados, disponíveis em seu portal de dados abertos. Ainda segundo Huang *et al.* (2019) o Reino Unido em 2017 dispunha de mais de 40.000 conjuntos de dados em seu portal.

Considerando os avanços na abertura de conjuntos de dados observam-se iniciativas como *Open Data Barometer*² foi possível evidenciar que países como Reino Unido e EUA estão indo na contramão de alguns compromissos de abertura de dados, ou seja, estão diminuindo sua pontuação (MACEDO; LEMOS, 2021). Também foi possível observar que o Brasil ampliou a abertura de dados, nos últimos anos, no entanto, ainda há espaço para aprimorar os dados que são disponibilizados, assim como as estratégias sobre dados abertos.

Deste modo outros estudos corroboram evidenciando possíveis aperfeiçoamentos que poderiam ser incluídos no âmbito de dados abertos, como é o caso do trabalho de Wang e Shepherd (2020) que abordam uma análise de 400 conjuntos de dados do Portal do Reino Unido, onde observa-se que vários recursos (cerca de 50%) estão em formato *Hypertext Markup Language* (HTML) ou *Portable Document Format* (PDF). Ainda neste mesmo estudo os autores destacam que há cerca de 2000 licenças diferentes em uso, que aproximadamente 2% dos conjuntos de dados não apresentam recursos vinculados e que 14% dos conjuntos de dados apresentam links quebrados (WANG; SHEPHERD, 2020). Neste sentido, Wang *et al.* (2019, p. 60) destacam: "Se o Reino Unido, como um pioneiro da iniciativa de dados abertos, está sobrecarregado com esses problemas, então os países menos avançados são propensos a experimentar obstáculos similares, ou até os mesmos".

¹ <http://datacatalogs.org/>

² mede as iniciativas de países que se comprometeram formalmente com a abertura de dados, seguindo os Princípios da Carta Internacional de Dados Abertos. Disponível em: <https://opendatacharter.net/principles/>

Na literatura são encontrados diversos estudos que abordam diferentes desafios, obstáculos e problemas relacionados a dados abertos, como: ausência de metadados, insuficiência de metadados, falta de interoperabilidade e dificuldade em pesquisar e navegar pelos dados (TYGEL *et al.*, 2016; ZUIDERWIJK *et al.*, 2012). Quando considerado os conjuntos de dados, também há indicativos que existem vários problemas nos dados disponibilizados, incluindo: conformidade de metadados, duplicação, inconsistência e integridade de valores (SADIQ; INDULSKA, 2017). Toda estas inconsistências sobre os dados podem ocasionar mal-entendidos, dificuldade de acesso e dificuldade de processamento dos dados disponíveis. Além de que a complexidade de obter os dados influenciará a vontade, a capacidade e a intenção do usuário em utilizar dados abertos (HUSIN *et al.*, 2019).

Há autores que destacam desafios não tão específicos, mas relacionados à abertura, o acesso e o uso dos dados. Segundo Kucera *et al.* (2015) estes desafios podem compor os seguintes grupos:

- Desafios organizacionais e desafios relacionados aos processos internos - problemas relacionados com as estruturas organizacionais e os processos internos através dos quais os dados abertos são entregues;
- Desafios jurídicos - problemas relacionados à abertura legal dos dados;
- Desafios técnicos - questões e desafios relacionados à tecnologia, dados e formatos ou infraestrutura necessária para a publicação.

Neste mesmo sentido, salienta-se o Quadro 1 que apresenta desafios para a abertura de dados governamentais.

Quadro 1 – Desafios para a abertura de dados governamentais.

Desafios	Descrição
Infraestrutura dos dados	A infraestrutura e o gerenciamento de dados governamentais se mostram deficientes e inconsistentes devido a alterações frequentes nos conjuntos de dados.
Acesso aos dados e usabilidade	A interface dos dados não possui um bom projeto de designer, afetando a capacidade de localização, consumo e integração. Os portais geralmente possuem uma abordagem técnica e orientada a desenvolvedores, não contemplando os usuários finais e cidadãos comuns.

Publicação	Os conjuntos de dados são geralmente incompletos, sem informações adicionais, com falta de padrão e muitas vezes sem metadados significativos, evidenciando problemas de interoperabilidade e em fontes de referência comuns. As publicações de forma geral não utilizam todo o potencial da Web de Dados preceitos de dados abertos conectados, essenciais para conectar conjunto de dados heterogêneos e descobrir recursos relacionados.
Políticos e sociais	Os dados de maior impacto social ainda permanecem amplamente fora do escopo de abertura. A abertura de dados capacita os empoderados, fortalecendo ainda mais aqueles que já têm poder e privando os demais.
Econômicos	Os dados abertos carecem de um modelo financeiro sustentável, além de sofrerem críticas por facilitar a neoliberalização através da comercialização dos serviços públicos.
Organização e processos internos	Se as autoridades públicas não perceberem o valor para abrir dados, não alocarão os recursos necessários para garantir a sua realização. Soma-se a isso a necessidade de proteger seus interesses pessoais como uma forte razão para que mantenham alguns dados secretos ou inacessíveis.
Legais e de privacidade	A natureza dos dados abertos impõe novas demandas ao governo para prover privacidade, segurança e impedir o uso indevido das informações. Os governos devem determinar a política sobre quais dados devem ser compartilhados e suas circunstâncias adequadas.
Capacitação e treinamento	Grande parte dos usuários não possui as habilidades necessárias para avaliar a qualidade e a adequação ao uso de dados, além de não possuir o alcance do que podem fazer com informações associadas. Soma-se a isso a necessidade de estudantes recém-formados e profissionais que lidam com tratamento e análise de dados serem devidamente treinados e capacitados tecnicamente em várias competências interdisciplinares.

Fonte: (MACEDO; LEMOS, 2021)

Dada a contextualização e a problematização, torna-se evidente de que a disponibilização dos dados não garante que os usuários farão uso. Sendo que, o sucesso de dados abertos não deve ser medido pela quantidade de dados, mas pela aceitação e uso dos dados (TALUKDER *et al.*, 2019). Além dos problemas e desafios já apresentados, os provedores de dados muitas vezes não fornecem suporte para que os usuários consigam obter benefícios através dos dados (KAASENBROOD *et al.*, 2015).

Considerando os problemas, desafios e obstáculos apresentados pela literatura, nota-se que vários destes estão relacionados aos padrões, políticas, processos e a fatores organizacionais. Nesta perspectiva, salienta-se que há vários estudos que fornecem orientações e conceitualizam as práticas para o tratamento de dados abertos (LÓSCIO *et al.*, 2016; HYLAND; WOOD, 2011; VILLAZÓN-TERRAZAS *et al.*, 2011). E, diversas inconsistências, problemas e desafios poderiam ser minimizados, com o uso adequado por parte dos provedores de dados das práticas para o tratamento de dados abertos.

Observando o contexto apresentado surgiu a hipótese de considerar os princípios da governança de dados, visto que a governança de dados é uma disciplina consolidada que considera avaliar, monitorar e direcionar a gestão de ativos de dados. Deste modo, ao considerar a governança de dados alinhado aos processos de abertura e pós-publicação espera-se que as organizações obtenham uma boa organização de seus processos contribuindo com o planejamento, com a execução de políticas e práticas viabilizando o aumento no valor dos ativos de dados e informações. Esta perspectiva será considerada ampliando e abrangendo as boas práticas já observadas para dados abertos.

Neste sentido, é possível observar diferentes iniciativas de organizações governamentais e não governamentais que compõem uma série de painéis e relatórios que se propõem a avaliar o engajamento das organizações na abertura de dados, dando destaque ao: Open Data Maturity³; Open Data Barometer⁴; Open Data Monitor⁵; e Global Open Data Index⁶. No entanto, nenhuma destas iniciativas tem como foco analisar e considerar as práticas e padrões adotados pelos provedores de dados na abertura, na pós-publicação dos dados e os princípios da governança de dados, e assim surgiu a seguinte pergunta de pesquisa: "*Como aplicar os princípios da governança de dados e as boas práticas para abertura e pós-publicação de dados abertos, considerando o contexto organizacional?*"

Portanto, apresentam-se a seguir os objetivos traçados para consolidar esta pesquisa.

³ Disponível em: <https://data.europa.eu/en/dashboard/2020>

⁴ Disponível em: https://opendatabarometer.org/?_year=2017&indicator=ODB

⁵ Disponível em: <https://www.opendatamonitor.eu/frontend/web/index.php?r=dashboard%2Findex>

⁶ Disponível em: <https://index.okfn.org/>

1.2 OBJETIVOS

1.2.1 Objetivo geral

Propor um modelo de maturidade de dados abertos, considerando princípios da governança de dados e as boas práticas aplicadas pelas organizações.

1.2.2 Objetivos específicos

Para o alcance do objetivo proposto foram elencados os seguintes objetivos específicos:

- Identificar boas práticas, etapas e atividades de abertura e pós-publicação de dados;
- Identificar componentes da governança de dados que podem corroborar com o modelo de maturidade;
- Compor um ciclo de vida para dados abertos, considerando as boas práticas aplicadas pelas organizações;
- Representar o conhecimento dos elementos que compõe o modelo de maturidade;
- Desenvolver um instrumento para a aplicação do modelo de maturidade;
- Avaliar o modelo de maturidade com especialistas da área.

1.3 DELIMITAÇÃO DA PESQUISA

Esta tese aborda diversos conceitos acerca de: dados abertos, governança de dados e modelos de maturidade. A exemplo de dados abertos são apresentadas as boas práticas de dados na Web e os ciclos de vida de dados abertos. Também busca-se apresentar aspectos que geram implicações sobre a abertura dos dados, sendo estes: **I)** legislação, regulamentações e licenças, **II)** colaboração e uso e **III)** sustentabilidade da iniciativa de dados abertos.

No âmbito de dados abertos, optou-se por não abordar questões sobre investimentos e questões financeiras, assim como a perspectiva legal não será aprofundada neste estudo, considerando de modo geral a conformidade e o uso de licença aberta. Sendo que para compreender e apresentar o contexto foram utilizados os conceitos e as respectivas definições presentes no Quadro 2. Deste modo, foi possível delimitar o domínio funcional do modelo de maturidade, sendo considerado para este o ciclo de vida de dados abertos composto através da realização desta pesquisa.

Quadro 2 – Definições e conceitos acerca de dados abertos.

Conceito	Definição
Dados	Os dados são representados por uma sequência de sinais e símbolos que não têm mais significado além da sua simples presença (NURNBERGER <i>et al.</i> , 2009). Também denominados com a menor unidade armazenada em um determinado recurso.
Recurso	Os "recursos", que mantêm os dados em si. Um recurso pode ser uma planilha CSV ou Excel, arquivo XML, documento PDF, arquivo de imagem, dados conectados no formato RDF, etc (CKAN, 2021). Os recursos são abordados pelo W3C como distribuição (LÓSCIO <i>et al.</i> , 2017).
Conjunto de Dados	Um conjunto de dados é uma parcela de dados - por exemplo, podem ser as estatísticas de crimes de uma região, os números de gastos de um departamento do governo ou as leituras de temperatura de várias estações meteorológicas (CKAN, 2021). E este pode conter diversos recursos.
Provedor de dados	Organização ou departamento responsável por fornecer os dados.
Organização responsável pelo portal	Organização que mantém o portal, esta pode ser diferente da organização responsável pela disponibilização dos dados.
Criador ou Autor	Quem cria, extrai, anonimiza, transforma ou mescla os dados.
Mantenedor	Quem é o responsável por manter os dados disponíveis.
Consumidor ou Usuário	Aquele que recupera e faz uso dos dados.

Fonte: Da Autora

Com a delimitação de escopo dessa pesquisa, foram abordadas as teorias e conceitos sobre modelos de maturidade, apresentando as principais definições que permeiam essas áreas de estudo. Também foram realizadas buscas em bases de dados para a obtenção de trabalhos relacionados ao que se propôs, isso se justifica pela observância de estudos similares e pelo fato de permitir, de forma propositiva o eventual e futuro desdobramento da pesquisa. Ainda para contextualizar as áreas foco do modelo de maturidade fez-se o uso dos princípios da governança de dados e assim optou-se pelo uso dos componentes do *framework* de Panian (2010).

Considerando o contexto apresentado, utilizou-se da *Design Science Research*, observando a proposta de Peffers *et al.* (2007), como metodologia de pesquisa. Com base nisso fez-se a composição das propostas de De Bruin *et al.* (2005) e Van Steenberg *et al.* (2010b), como o objetivo de compor e garantir a cientificidade do modelo de maturidade.

Neste sentido, para compreender o modelo de maturidade apresentado foi estruturado o Quadro 3, onde são abordados os principais conceitos e suas respectivas definições acerca de modelos de maturidade de áreas foco.

Quadro 3 – Definições e conceitos acerca do modelo de maturidade.

Conceito	Definição
Maturidade	Maturidade indica o grau de desenvolvimento.
Domínio Funcional	Um domínio funcional é o conjunto de atividades, meios, responsabilidades e atores envolvidos no cumprimento de uma função bem definida dentro de uma organização (VAN STEENBERGEN <i>et al.</i> , 2013).
Área Foco	Uma área foco é um subconjunto coerente bem definido de um domínio funcional. O conjunto total de áreas foco é uma partição do domínio funcional, ou seja, áreas de foco diferentes são disjuntas e a união de todas essas áreas de foco é o domínio funcional completo (VAN STEENBERGEN <i>et al.</i> , 2013).
Capacidade	Uma capacidade é a habilidade de atingir uma meta predefinida (VAN STEENBERGEN <i>et al.</i> , 2013).
Dependência	Uma capacidade depende de outra capacidade, se só puder ser obtida depois que outra capacidade for alcançada (VAN STEENBERGEN <i>et al.</i> , 2013).
Matriz de Maturidade	Uma matriz de maturidade fornece uma ordenação parcial de recursos dentro de um domínio funcional em áreas de foco em uma sequência de níveis de maturidade (VAN STEENBERGEN <i>et al.</i> , 2013).
Nível de Maturidade	Um nível de maturidade é um patamar evolucionário bem definido dentro de um domínio funcional (VAN STEENBERGEN <i>et al.</i> , 2013).
Instrumento de Avaliação	Um instrumento de avaliação é uma ferramenta para determinar a maturidade dentro de um domínio funcional (VAN STEENBERGEN <i>et al.</i> , 2013).
Questões de Avaliação	As perguntas de avaliação são usadas para determinar o nível de maturidade atual ou de destino de uma organização dentro de um domínio funcional (VAN STEENBERGEN <i>et al.</i> , 2013).
Avaliação Preenchida	Uma avaliação preenchida é o conjunto de questões de avaliação, juntamente com as respostas fornecidas durante uma avaliação específica (VAN STEENBERGEN <i>et al.</i> , 2013).
Perfil de Maturidade	Um perfil de maturidade é um conjunto específico de recursos dentro de um domínio funcional que foi alcançado por uma organização (VAN STEENBERGEN <i>et al.</i> , 2013).
Ação de Melhoria	Uma ação de melhoria é a descrição de uma atividade que deve resultar na obtenção de uma capacidade específica (VAN STEENBERGEN <i>et al.</i> , 2013).

Modelo de Maturidade	Um modelo de maturidade é um instrumento para avaliar e desenvolver a capacidade de uma organização para executar atividade dentro de um domínio funcional (VAN STEENBERGEN <i>et al.</i> , 2013).
----------------------	--

Fonte: Da Autora

Outra delimitação está relacionada à verificação do comportamento do artefato, onde fez-se o uso de cenários reais, com o intuito de observar as necessidades das organizações e com isso obter um modelo de maturidade adaptável e robusto para as diferentes organizações que realizam a abertura dos dados. Ainda foram entrevistados especialistas da área de dados abertos com o objetivo de identificar os componentes indispensáveis para o modelo de maturidade e com base nisso fornecer uma versão "simplificada" e uma versão completa do modelo proposto.

Todos os esforços da pesquisa realizada se deram com a intenção de se ter atividades no modelo que forneçam um roteiro para melhorias e não uma lista de tarefas para serem completadas antes dos dados serem liberados. De fato, é desejável que as organizações experimentem baixos níveis de maturidade para que desenvolvam processos adequados para si próprias. Visto que a tentativa de saltar para níveis mais elevados de maturidade podem resultar em processos impostos que não levam em conta a cultura, o ambiente ou as necessidades específicas das organizações. E neste sentido, a proposição do artefato objetiva-se em obter uma estrutura que não sirva como fator de desmotivação para as organizações, mas que proporcione uma estrutura adequada às diferentes realidades. Assim como o público alvo do modelo de maturidade são os *stakeholders* internos a organização, ou seja, a própria organização realizará a sua autoavaliação.

Nesta perspectiva, ressalta-se a importância da avaliação da maturidade no contexto de dados abertos, visto que os modelos de maturidade permitem não apenas garantir um *benchmark* de boas práticas e identificar áreas de melhoria, mas também possibilitam rastrear as metas de melhorias com base no status atual e a partir disso traçar estratégias na organização definindo a melhor maneira de alcançá-las.

1.4 INEDITISMO, ORIGINALIDADE E CONTRIBUIÇÕES

Os dados tem grande relevância para a sociedade do conhecimento, esta relevância é observada porque sem dados, não se pode, por exemplo, saber quantas pessoas nascem ou morrem ao ano, ou ainda, como o dinheiro público está sendo gasto e para qual finalidade. Assim observa-se que diversos estudos evidenciam a necessidade de se ter dados corretos (KLOSS, 2015), pois os dados e as informações

permitem que as pessoas e as organizações tomem decisões à partir do conhecimento obtido.

Com a crescente capacidade de coletar, transmitir e armazenar dados, juntamente com a crescente tendência à abertura (BARBOSA *et al.*, 2014; GOLDSTEIN *et al.*, 2013), criam-se oportunidades únicas que podem beneficiar o governo, a ciência, os cidadãos e a indústria.

Integrando e analisando vários conjuntos de dados, os governos podem compreender e possibilitar operações com melhor planejamento. Os cientistas podem se envolver em ciência orientada por dados, explorar processos e compreender o comportamento das pessoas, por exemplo (CASTELLANI RIBEIRO *et al.*, 2015). Contudo, disponibilizar os dados para os cidadãos tem o potencial de melhorar a governança e a participação, ainda para os empreendedores e corporações há a possibilidade de serem gerados novos produtos e serviços (CASTELLANI RIBEIRO *et al.*, 2015).

Os benefícios dos dados abertos são diversos e vão desde a melhoria da eficiência das administrações públicas, ao crescimento econômico no setor privado até mesmo o bem-estar social mais amplo (EUROPEAN DATA PORTAL, 2018). Deste modo, observa-se que a realização desta pesquisa contribui com a sociedade, com a academia e em especial com os provedores de dados, no sentido de se ter um artefato que viabilize avaliar a sua estratégia de abertura de dados e conseqüentemente refletir sobre o que está sendo feito e de que forma a organização pode aprimorar suas políticas, seus processos, suas atividades e seus padrões utilizados na disponibilização dos dados.

Destaca-se que a realização deste estudo é inédita e original, devido a considerar boas práticas de dados na Web, atividades e fatores importantes abordados pela literatura e também princípios da governança de dados, com o intuito de fornecer uma visão abrangente sobre ciclo de vida de dados abertos, servindo como modelo de referência para que este ciclo de vida possa ser estendido, reduzido e até mesmo adaptado, ou seja, o ciclo de vida possibilitará que modelos mais específicos ou mais genéricos sejam desenvolvidos, possibilitando o avanço das pesquisas da área.

Destaca-se que o ciclo de vida apresentado visa ser um modelo de referência para que este possa ser estendido, reduzido e até mesmo adaptado, ou seja, o ciclo de vida possibilitará que modelos mais específicos ou mais genéricos sejam desenvolvidos

Ainda sobre a originalidade e o ineditismo salienta-se que o modelo de maturidade organizado em áreas foco, oriundas da governança de dados e componentes viabilizam a avaliação da maturidade de modo a aprimorar cada componente de interesse da organização, tornando a estrutura do artefato adaptável aos diferentes cenários observados para dados abertos.

Neste sentido, também destaca-se que a abordagem utilizada para a composi-

ção do modelo de maturidade e a utilização da metodologia *design science research* para a concepção do modelo difere-se dos estudos evidenciados nas buscas realizadas nas bases de dados, sendo estes expostos na seção 2.4, desta forma contribui-se com o avanço das pesquisas destas áreas.

Por fim, destacam-se as contribuições da Engenharia do Conhecimento na realização deste estudo, pois ao final da concepção do modelo de maturidade foi possível se ter uma representação dos conhecimentos abordados, considerando como base alguns dos princípios de ontologias e desta maneira observa-se que podem ser derivados novos estudos no âmbito semântico e computacional.

1.5 ADERÊNCIA AO EGC

O Programa de Pós – Graduação em Engenharia e Gestão do Conhecimento (PPEGC) da Universidade Federal de Santa Catarina (UFSC) possui três áreas de concentração (PACHECO, 2016), a saber: a área de **Gestão do Conhecimento** que tem por objetivo estudar as bases conceituais e metodológicas para implantação da gestão organizacional baseada no conhecimento (EGC, 2021b), a área **Engenharia do Conhecimento** que tem por objetivo a pesquisa e o desenvolvimento de métodos, técnicas e ferramentas para a construção de modelos e sistemas de conhecimento em atividades intensivas em conhecimento (EGC, 2021b) e por fim a área de **Mídia e Conhecimento** que visa trabalhar o compartilhamento e disseminação do conhecimento, desenvolvimento e avaliação das mídias voltadas a catalisar a habilidade de grupos para pensar, comunicar, disseminar, preservar, apreender e criar conhecimento (EGC, 2021b).

Esta tese esta contextualizada na área de concentração de Engenharia do Conhecimento, sua linha de pesquisa é a Engenharia do Conhecimento Aplicada às Organizações. Esta linha estuda a concepção, desenvolvimento e implantação de soluções da Engenharia do Conhecimento nas organizações (EGC, 2021a). Ao ser considerado todo o conhecimento, as informações e os dados que são produzidas dentro de organizações, a Engenharia do Conhecimento torna-se uma importante área de pesquisa devido ao suporte tecnológico que esta viabiliza à resolução de problemas complexos. Assim, visa-se contribuir no avanço das pesquisas visto que dados abertos e dados abertos conectados são algumas das tecnologias que atualmente organizações públicas fazem uso, para a publicação e catalogação de dados em busca de transparência.

Neste sentido, foram realizadas buscas no banco de teses e dissertações do PPEGC / UFSC⁷, em junho de 2021, com o intuito de abordar estudos similares a esta tese, sobre dados abertos, dados abertos conectados e modelos de maturidade.

⁷ <http://btd.egc.ufsc.br/>

Após a leitura dos títulos e resumos destes trabalhos foram elencados os estudos apresentados no Quadro 4 que abordam discussões sobre modelos de maturidade.

Quadro 4 – Trabalhos que abordam discussões sobre modelos de maturidade

Fonte	Título	Área de Concentração	T/D
(HAUCK, 2011)	Um Método de Aquisição de Conhecimento para Customização de Modelos de Capacidade/Maturidade de Processos de Software.	Engenharia do Conhecimento	Tese
(HELOU, 2015)	Avaliação da Maturidade da Gestão do Conhecimento na Administração Pública	Gestão do Conhecimento	Tese
(PACHECO, 2015)	Maturidade em Gestão do Conhecimento da Prefeitura Municipal de Criciúma: O Caso dos Setores de Convênios.	Gestão do Conhecimento	Dissertação
(UENO, 2016)	Modelo de Avaliação da Maturidade do Processo de Inovação Como Estratégia Competitiva Empresarial	Gestão do Conhecimento	Tese
(PIMENTA, 2017)	Análise de maturidade da coprodução de conhecimento transdisciplinar: Um Estudo de Caso em uma Rede	Gestão do Conhecimento	Tese
(KRAEMER, 2018)	Modelo de Maturidade do Sistema de Educação Corporativa	Gestão do Conhecimento	Dissertação
(ZANUZZI, 2019)	Avaliação da Maturidade da Gestão do Conhecimento das Unidades Produtoras Integradas da Agroindústria Avícola	Gestão do Conhecimento	Dissertação
(FLORES, 2019)	Maturidade em gestão do conhecimento na administração pública: um estudo na Prefeitura Municipal de São José/SC	Gestão do Conhecimento	Dissertação
(ERPEN, 2019)	Pecuária Intensiva em Conhecimento: Modelo de Maturidade em Gestão do Conhecimento Aplicado a Bovinocultura de Corte	Gestão do Conhecimento	Tese
(FIGUEIRA, 2020)	Proposta de Um Modelo de Maturidade de Gestão de Projetos, que Incorpora os Processos Sistêmicos de Gestão do Conhecimento	Gestão do Conhecimento	Dissertação
(MAXIMO, 2021)	Maturidade da Gestão do Conhecimento: uma aplicação ao Modelo de Excelência da Gestão	Gestão do Conhecimento	Dissertação

Já o Quadro 5 apresenta os trabalhos que abordam o contexto de dados abertos e dados abertos conectados.

Quadro 5 – Trabalhos que abordam discussões sobre dados abertos e dados abertos conectados

Fonte	Título	Área de Concentração	T/D
(KLEIN, 2016)	Uma Proposta de Modelo Conceitual para Uso de Big Data E Open Data para Smart Cities.	Engenharia do Conhecimento	Dissertação
(SPERONI, 2016)	Modelo de Referência Para Indicadores de Inovação Regional Suportado por Dados Ligados	Engenharia do Conhecimento	Tese
(GOMES, 2017)	Proposta de Arquitetura para Ecossistema de Inovação em Dados Abertos.	Engenharia do Conhecimento	Dissertação
(PEREIRA, 2017)	OGDPub: Uma ontologia para Publicação de Dados Abertos Governamentais	Engenharia do Conhecimento	Dissertação
(CAMPOS, 2018)	Dados Abertos Governamentais: Desafios na Publicação	Engenharia do Conhecimento	Dissertação
(CASAES, 2019)	Governança de Dados Abertos Governamentais: Framework Conceitual para as Universidades Federais, Baseado em uma Visão Sistêmica	Engenharia do Conhecimento	Tese

Fonte: Da Autora

Deste modo, nota-se que há tese e dissertações que possuem temas similares ao desta tese e destaca-se que esta tese diferencia-se dos trabalhos previamente apresentados no programa considerando e convergindo o contexto de dados abertos e de modelos de maturidade, fomentando com isso, a importância da relação entre os temas escolhidos e contribuindo com os avanço das pesquisas no programa e nas respectivas áreas consideradas.

1.6 ORGANIZAÇÃO DO DOCUMENTO

O documento contempla seis capítulos, este capítulo que além de abordar a organização do documento também aborda o contexto, caracterizando o problema

de pesquisa, os objetivos do trabalho, a delimitação, o ineditismo e a aderência ao PPEGC.

Já o capítulo 2 aborda a fundamentação teórica acerca de dados abertos, dados abertos conectados incluindo o estudo dos ciclos de vida e das boas prática para dados na Web. Também é abordado o contexto de modelos de maturidade contemplando definições, componentes essenciais, tipos de modelos de maturidade e por fim neste capítulo apresentam-se os trabalhos relacionados.

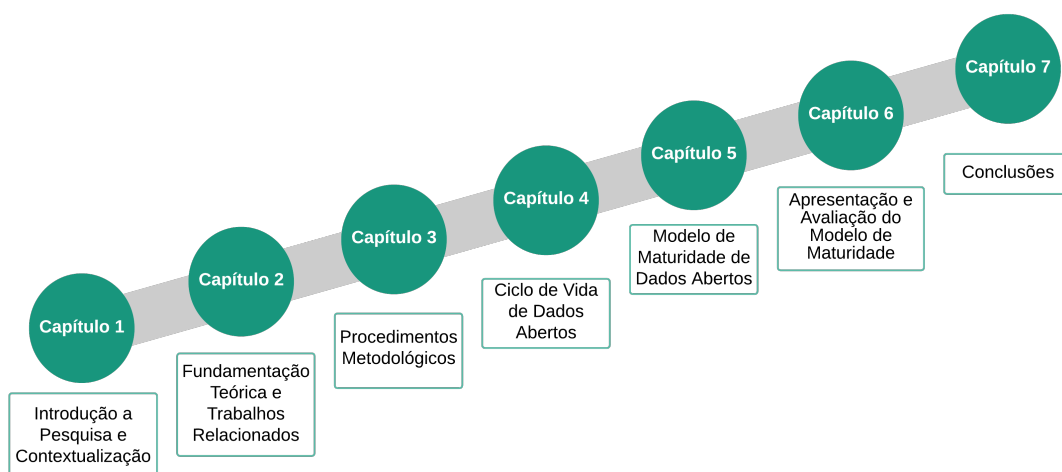
O capítulo 3 apresenta em detalhes os procedimentos metodológicos adotados na condução da etapas da *design science research* para a consolidação desta pesquisa.

O capítulo 4 aborda e discute o ciclo de vida de dados abertos, considerando os fatores importantes, boas práticas e atividades que corroboram com o contexto apresentado, sendo ao final traçado um comparativo com um ciclo de vida de estrutura similar, para evidenciar as características de ambos.

O capítulo 5 evidencia como se deu a composição do modelo de maturidade. O capítulo 6 possibilita vislumbrar a aplicação do modelo em cenários de uso reais, bem como evidenciar à percepção de especialistas de dados abertos sobre o modelo desenvolvido.

Por fim o capítulo 7 apresenta as conclusões e as possibilidades para pesquisas futuras. Uma visão geral da organização deste documento de tese é apresentado na Figura 1.

Figura 1 – Estrutura da Tese



Fonte: Da Autora

2 FUNDAMENTAÇÃO TEÓRICA

O objetivo deste capítulo é apresentar e discutir os resultados da revisão da literatura de dados abertos, dados abertos conectados, governança de dados e modelos de maturidade. Buscam-se também analisar as boas práticas e tratamentos de dados utilizadas na publicação e pós-publicação de dados abertos, além de investigar os componentes de um modelo de maturidade, bem como evidenciar *frameworks* de governança de dados com o intuito de identificar os itens que podem auxiliar na estruturação do modelo.

2.1 DADOS ABERTOS

O volume de dados produzidos nos últimos anos, no mundo teve um aumento significativo. Estima-se que cerca de 90% dos dados presentes na Web foram criados em anos recentes sendo que estes não obtiveram um aumento apenas no volume, mas também no nível de detalhamento, tudo isso devido às novas tecnologias (DATA REVOLUTION FOR SUSTAINABLE DEVELOPMENT, 2014).

As novas tecnologias são oriundas de movimentos como: IoT (BARNAGHI *et al.*, 2012), Big Data (HITZLER; JANOWICZ, 2013), Web Semântica (BERNERS-LEE *et al.*, 2001), Governo Aberto (MEIJER *et al.*, 2012), entre outros. O movimento de Governo Aberto, por exemplo, possibilitou que alguns países adotassem a abertura de dados principalmente no setor público, em busca de transparência, accountability, participação social e colaboração (MEIJER *et al.*, 2012). Através deste movimento e influenciado pelo conceito de *Openness*, surgiu o conceito de Dados Abertos (*Open Data*) que por sua definição são "dados que qualquer um pode livremente acessá-los, utilizá-los, modificá-los e compartilhá-los para qualquer finalidade, estando sujeito a, no máximo, exigências que visem preservar sua proveniência e sua abertura" (OPEN DEFINITION, 2018), como consequência houve um crescimento considerável na quantidade de dados que são disponibilizados na Web.

Desde o surgimento do conceito de dados abertos, algumas entidades e indivíduos manifestaram-se e publicaram recomendações, princípios, regras entre outras denominações, com o intuito de auxiliar aqueles que desejam abrir seus dados. Este é o caso do grupo de trabalho da Califórnia que em 2007 definiu oito princípios, para a abertura dos dados governamentais, sendo estes:

- Completos: Todos os dados públicos devem ser disponibilizados. Dado público é aquele que não está sujeito a restrições de privacidade, segurança ou outros privilégios.
- Primários: São apresentados tal como colhidos da fonte, com o maior nível possível de granularidade, sem agregação ou modificação.
- Atuais: Devem ser publicados o mais rápido possível para preservar seu valor. Em geral, têm periodicidade: quanto mais recentes e atuais, mais úteis.

- Acessíveis: São disponibilizados para a maior quantidade possível de pessoas, atendendo, assim, aos mais diferentes propósitos.
- Compreensíveis por máquina: Devem estar estruturados de modo razoável, possibilitando que sejam processados automaticamente.
- Não discriminatórios: Devem estar disponíveis para qualquer pessoa, sem necessidade de cadastro ou qualquer outro procedimento que impeça o acesso.
- Não proprietários: Nenhuma entidade ou organização deve ter controle exclusivo sobre os dados disponibilizados.
- Livres de licenças: Não devem estar submetidos a *copyrights*, patentes, marcas registradas ou regulações de segredo industrial. Restrições razoáveis quanto a privacidade, segurança e outros privilégios são aceitas, desde que transparentes e bem justificadas (PORTAL BRASILEIRO DE DADOS ABERTOS, 2018, p.1).

Neste mesmo sentido, a *Open Knowledge* define três regras fundamentais para potencializar a difusão de dados abertos (KNOWLEDGE, OPEN, 2018), sendo estas:

- Manter simples: Começar pequeno, simples e rápido, pois não há nenhuma exigência quanto ao tamanho do conjunto de dados (*datasets*) a ser aberto. Portanto pode-se começar com a abertura de um conjunto de dados ou até mesmo com apenas parte do conjunto. Possibilitando assim mover-se mais rapidamente e conseqüentemente aprender mais agilmente com base na experiência, assim como avaliar o fracasso ou o sucesso do conjunto de dados, afinal nem todos os conjuntos de dados serão úteis.
- Envolver-se cedo e envolver-se com frequência: Envolver-se com os usuários dos dados sejam eles cidadãos, empresas, desenvolvedores, entre outros permitirá entender as necessidades e dificuldades, possibilitando assim a abertura de dados mais relevantes em outras interações, quando possível.
- Abordar medos e mal-entendidos comuns: Para abrir os dados sempre surgirão dúvidas e medos, para isso é importante identifica-los e abordá-los o mais cedo possível.

Além das regras já apresentadas, são recomendadas quatro etapas chaves para a abertura dos dados, sendo que algumas podem ser realizadas simultaneamente (KNOWLEDGE, OPEN, 2018):

- Escolha de *datasets*: Escolha quais *datasets* ou quais partes pretende-se disponibilizar. Esta é uma etapa importante e provavelmente será retomada diversas vezes.
- Aplicar uma licença aberta: Determinar quais os direitos de propriedade intelectual existem sobre os dados. Aplicar a licença aberta que mais se adequar aos direitos e suporte à definição de abertura.
- Disponibilizar os dados: brutos e em formato útil (processáveis por máquina).

- Torna-los visíveis: postar na Web e possivelmente organizá-los em um catálogo para listar todos os *datasets* que foram disponibilizados.

No âmbito de auxiliar aqueles que desejam disponibilizar seus dados na Web emergiram também estudos sobre ciclos de vida, guias metodológicos, boas práticas, entre outros. Estes estudos serão detalhados no decorrer deste capítulo. No entanto, um dos contextos a ser destacado é a adequação da licença sobre os dados, pois a licença é uma declaração legal que permite fazer algo que não seria permitido de outra forma, como por exemplo, conceder a permissão de uso e distribuição dos dados (NEUMAIER, 2015).

2.1.1 Licença Aberta

Para a Open Definition (2018), uma licença é aberta se a mesma satisfaz um conjunto de permissões necessárias, sendo estas: uso, redistribuição, modificação, separação, compilação, propagação, aplicação para qualquer propósito e não houver cobrança. Além disso, eles incluem condições aceitáveis, que são exceções mínimas que não limitam e nem tornam incertas as permissões necessárias para dados abertos. A partir deste contexto tem-se três níveis de licença sobre dados ou *datasets* que podem ser aplicadas (NEUMAIER, 2015), sendo estas:

- Domínio Público: Não há nenhuma restrição sobre os dados ou conteúdo.
- Atribuição: deve-se fazer a atribuição do trabalho da forma especificada pelo autor ou licenciante.
- Atribuição-Compartilhada: deve-se fazer a atribuição do trabalho da forma especificada pelo autor ou licenciante e compartilhar qualquer conteúdo ou dado derivado sob a mesma licença.

Conforme Open Data Institute (2013) ainda há dois conjuntos de licenças abertas que podem ser aplicadas que são recomendados. Uma das licenças recomendadas é sobre os dados, ou seja, sobre conteúdo criativo (textos, imagens, entre outros) a licença atribuída é a *Creative Commons* (CC) Há três níveis dentro da licença CC que podem ser consideradas no uso de conteúdo aberto. Sendo estas especificadas no Quadro 6.

Quadro 6 – Níveis de licença Creative Commons

Nível da licença	Licença Creative Commons
Domínio publico	CC0
Atribuição	CC-by
Atribuição-Compartilhada	CC-by-sa

Fonte: <https://creativecommons.org/choose/>

De forma semelhante, a licença aplicadas sobre conteúdos também têm-se licenças para aplicar sobre um conjunto de dados especificamente para banco de dados que é a *Open Data Commons*, onde também tem-se três níveis conforme segue no Quadro 7.

Quadro 7 – Níveis de licença Open Data Commons

Nível da licença	Licença Open Data Commons
Domínio publico	PDDL
Atribuição	ODC-by
Atribuição-Compartilhada	ODbL

Fonte: <https://opendatacommons.org/licenses/>

Existem outros tipos de licenças, além destas abordadas, mas nem todas são licenças abertas, por exemplo, ao utilizar o seletor de licença *Creative Commons*, apenas as licenças que são descritas como licenças de “Cultura Livre” são licenças abertas (OPEN DATA INSTITUTE, 2013). E portanto observa-se a importância do uso de uma licença adequada para publicação de dados na Web, pois nem todos os dados presentes na Web estão disponíveis gratuitamente e estão abertos para qualquer um fazer uso e disseminá-los. Além disso, faz-se necessário esclarecer e compreender a diferença entre dados abertos e dados conectados.

2.1.2 Dados Abertos x Dados Conectados

Os dados conectados são um dos principais conceitos acerca da Web Semântica, também conhecida como a Web de Dados (HEATH; BIZER, Christian, 2011). A Web Semântica busca apresentar dados compreensíveis não somente aos seres humanos, mas também por máquinas. Desta forma, os dados conectados são definidos como: "um conjunto de boas práticas para conectar e publicar dados estruturados

na Web de diferentes fontes" (BERNERS-LEE, 2010), com a finalidade de torna-los legíveis por máquinas.

Há quatro recomendações básicas para a publicação e o consumo de dados conectados, (BERNERS-LEE, 2010), sendo estas:

1. Use URI como nomes para as coisas. O *Uniform Resource Identifier* (URI) é utilizado como identificador único de qualquer recurso, como: pessoas, lugares, documentos, conceitos, etc. Sendo que estes recursos representam entidades da Web.
2. Use *Hypertext Transfer Protocol* (HTTP) URIs para que as pessoas possam procurar esses nomes. Neste contexto restringe-se o uso de HTTP URIs, para evitar outros esquemas, como por exemplo, URNs. Além disso HTTP URIs geram nomes adequados para recursos por duas razões: fornece uma maneira simples de criar nomes exclusivos globalmente sem gerenciamento centralizado; e o URI trabalha não apenas como um nome, mas também como um meio de acesso a informações sobre um recurso na Web (BIZER, Chris *et al.*, 2007).
3. Quando alguém procura um URI, fornecer informações úteis, usando os padrões (*Resource Description Framework* (RDF)) ou SPARQL). O RDF foi desenvolvido pelo *World Wide Web* (W3C), e é um modelo de dados simples para descrever recursos (URIs) e seus relacionamentos. Neste modelo, os recursos e seus relacionamentos são representados como triplas (recurso-propriedade-valor), sendo que o valor pode ser outro recurso. Já o SPARQL é uma linguagem de consulta RDF, ou seja, uma linguagem de consulta semântica para bancos de dados, capaz de recuperar e manipular dados armazenados em RDF.
4. Incluir links para outros URIs para que eles possam descobrir mais coisas. As ligações se dão através das triplas do RDF. Sendo que as ligações RDF são a base da Web de Dados, pois elas permitem que através de um recurso referenciado por um URI seja possível alcançar diversos outros recursos ligados através de propriedades. Com isso, torna-se possível navegar na web de dados, descobrindo outras informações uteis referentes ao recurso (BIZER, Chris *et al.*, 2007).

Considerando o âmbito apresentado para dados conectados e a definição de dados abertos sendo: "dados que qualquer um pode livremente acessá-los, utilizá-los, modificá-los e compartilhá-los para qualquer finalidade, estando sujeito a, no máximo, exigências que visem preservar sua proveniência e sua abertura" (OPEN DEFINITION, 2018). Concluí-se que, dados abertos não são iguais aos dados conectados. Dados abertos podem ser disponibilizados à todos sem que estejam em um formato RDF. Ao

mesmo tempo, que os dados podem ser conectados sem estarem livremente disponíveis para serem utilizados ou distribuídos.

Portanto, existe uma relação entre dados abertos e a Web Semântica? Não há uma relação direta entre dados abertos e Web Semântica. Os esforços da comunidade e do W3C e de todos os defensores da abertura de dados são canalizados para enriquecer os dados abertos conectados, quando busca-se relacionar dados abertos a Web Semântica (VISINTIN *et al.*, 2017).

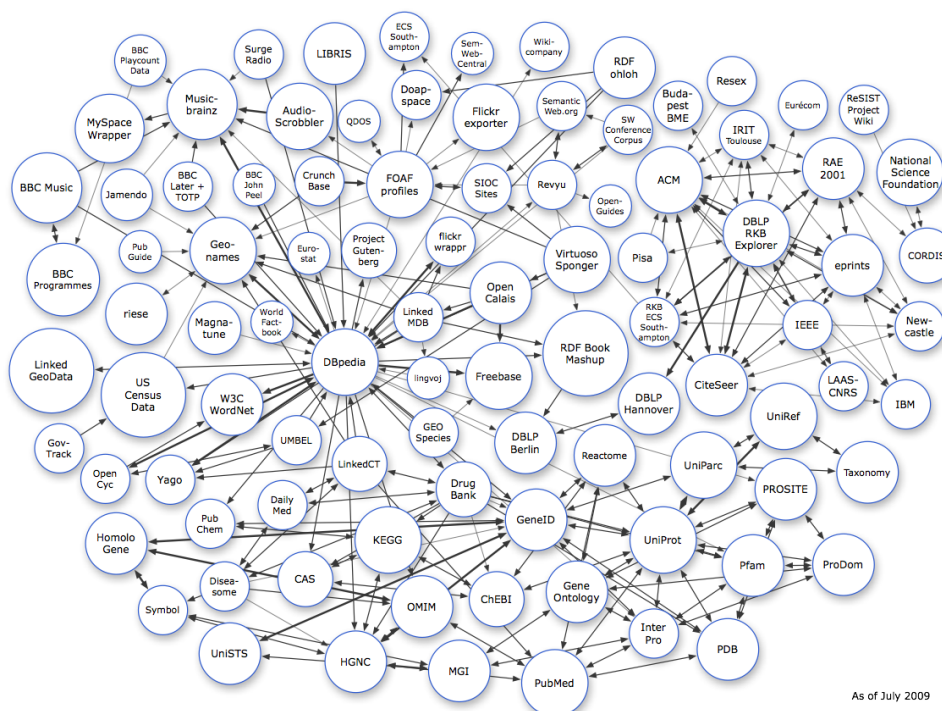
2.1.3 Dados Abertos Conectados

Em 2007 o projeto Dados Conectados (*Linked Data*) foi estendido, com o objetivo de construir a Web de Dados. A Web de dados é constituída de *datasets* com licença aberta, que fazem uso do formato RDF seguindo as recomendações de dados conectados, dando origem a Dados Abertos Conectados - *Linked Open Data* (LOD) (BIZER, Christian *et al.*, 2011).

Dados conectados e dados abertos conectados fazem uso de tecnologias presentes na arquitetura da Web Semântica (RDF e URIs), com o intuito de padronizar os dados, possibilitando assim o processamento, a conexão dos dados e dos *datasets*. Deste modo, os *datasets* em RDF podem ser associados a outros conjuntos de dados para formar uma grande base de dados conectados, ou seja, materializando assim a Web de Dados (ALSHEHHI *et al.*, 2013).

Os dados abertos conectados viabilizam o compartilhamento dos dados auxiliando as iniciativas de publicação de dados governamentais, de tal modo que todos os dados conectados ajudam a incrementar a nuvem de *Linked Open Data* (BIZER, Christian *et al.*, 2011). Esta nuvem é baseada em *datasets* conectados e publicados por indivíduos ou organizações que contribuem com o projeto. A seguir é apresentada na Figura 2 a nuvem LOD composta por 95 *datasets*.

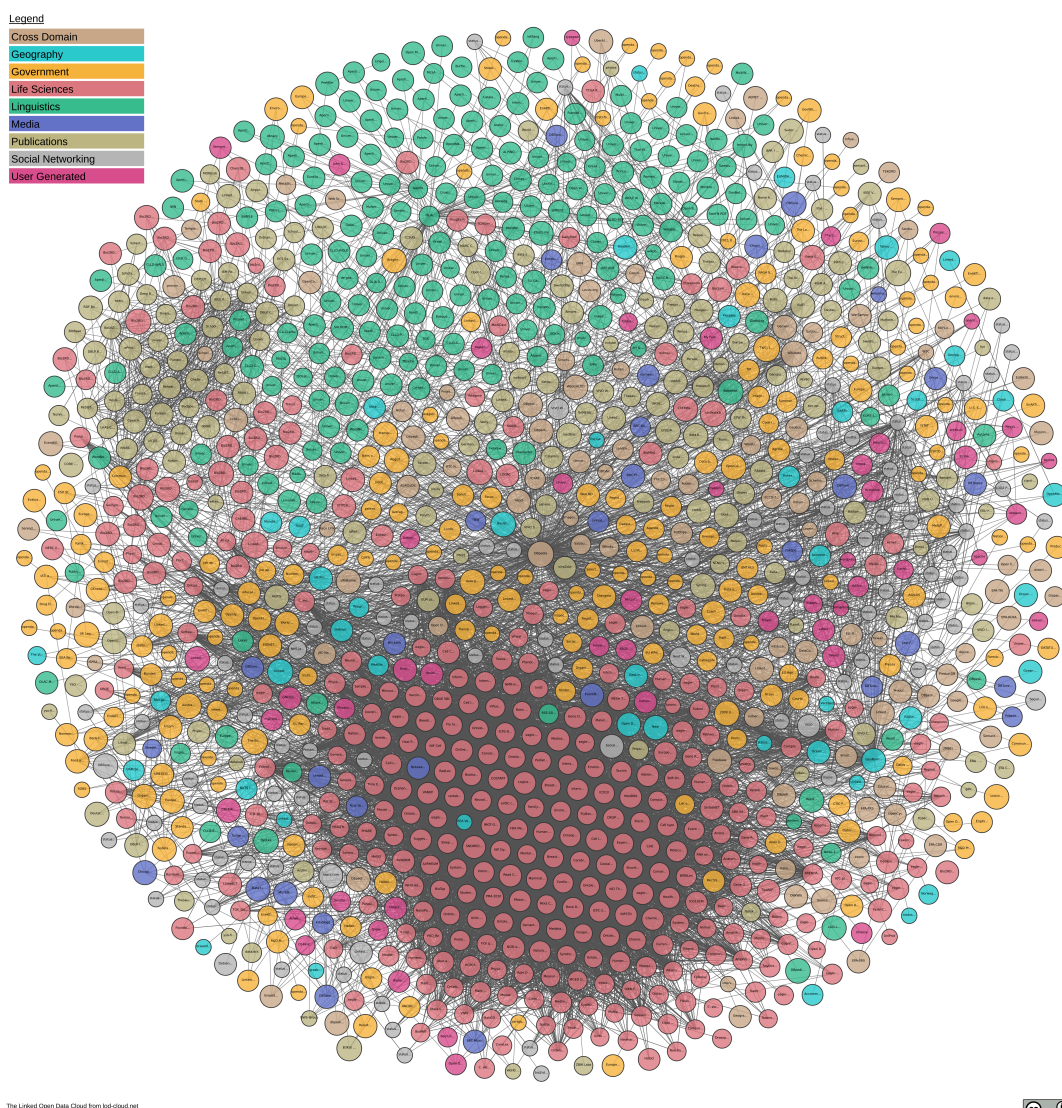
Figura 2 – Nuvem LOD em 2009



Fonte: (MCCRAE; ABELE, 2021)

Desde o início deste projeto nota-se um crescimento significativo no número de *dataset*, assim como o aumento de indivíduos que publicam seus dados. Ao fazê-lo, eles criaram esse espaço de dados interconectados globalmente - a Web de Dados. Estes *datasets* abrangem numerosos domínios e tópicos, tais como: pessoas, empresas, filmes, música, locais, livros e outras publicações, bem como dados governamentais. E esta evolução pode ser observada na Figura 3 com a nuvem LOD composta por 1301 *datasets* e 16283 links em maio de 2021 (MCCRAE; ABELE, 2021).

Figura 3 – Nuvem LOD em 2021

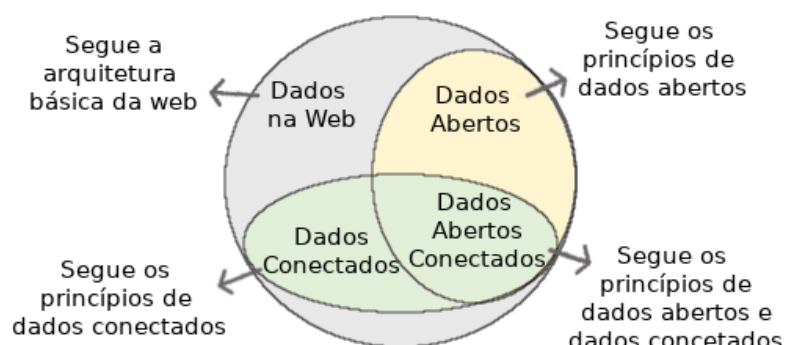


Fonte: (MCCRAE; ABELE, 2021)

A motivação por trás de dados abertos conectados é fornecer dados brutos, estruturados e ligados através da Web que possam ser acessados universalmente e ser facilmente compartilhados (BIZER, Christian *et al.*, 2011). A Web de Dados apresenta oportunidades para derivar *insight* e padrões, através dos dados que nela estão disponibilizados.

Em suma, Lóscio *et al.* (2016) resume em um diagrama esta distinção entre dados abertos, dados conectados e dados abertos conectados discutida até o momento e que pode ser observado na Figura 4. E para uma melhor compreensão quando relata-se neste trabalho dados abertos, também está sendo englobado neste estudo o contexto de dados abertos conectados.

Figura 4 – Dados na Web x Dados Abertos x Dados Conectados



Fonte: (LÓSCIO *et al.*, 2016)

Assim, surgiram diversas pesquisas com o intuito de guiar aqueles que desejam publicar dados abertos e dados abertos conectados na Web a fim de se ter dados padronizados, potencializando assim a utilização. Para isso, apresentam-se os trabalhos que abordam práticas e tratamento de dados com o intuito de disponibiliza-los na Web.


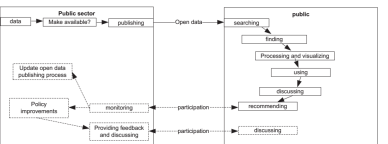
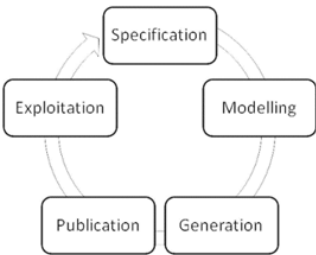
2.1.4 Os Ciclos de Vida de Dados Abertos

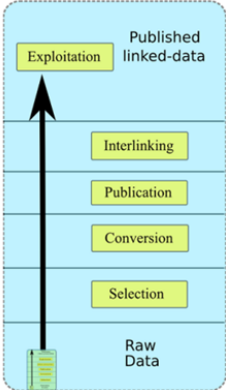
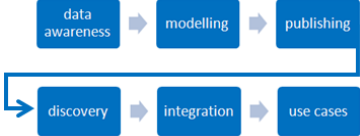
Na engenharia de software, é comum documentar um processo do mundo real por meio de um modelo de processo (LANE; RICHARDSON, 2011). Os modelos de processo de software vêm de várias formas, mas consistem em atividades referentes à organização, gerenciamento, medição e aprimoramento das atividades envolvidas no desenvolvimento de software. Todo projeto de desenvolvimento de software tem um processo, definido ou não de forma explícita ou implícita. Um dos sinônimos encontrados na literatura para processo de software é o termo ciclo de vida do software (DERNIAME *et al.*, 1999).



Considerando a contextualização acerca de ciclo de vida de software esta seção analisa as orientações que conceitualizam as práticas e tratamentos de dados, desde sua geração até a pós-publicação de dados na Web. Para isso, ao ser verificada a literatura, foram encontrados sete trabalhos que discutem as práticas para a publicação de dados, sendo estes dados abertos, dados conectados e dados abertos conectados, os estudos foram conduzidos sob diferentes terminologias. Eles foram chamados de ciclo de vida, processos ou ainda de orientações metodológicas. As diferentes terminologias ilustram diferentes propósitos. No entanto, todas as abordagens visam estruturar o tratamento dos dados em si, assim como guiar o processo de publicação, com intuito de obter dados padronizados e até mesmo conectados.

O Quadro 8 aborda os seguintes itens, para cada trabalho analisado: a terminologia, o objetivo que os autores propõem, quais são as etapas apresentadas, o foco de cada um e por fim o diagrama que o representa.

Quadro 8 – Trabalhos que abordam práticas e tratamento de dados para a publicação

Autor/Ano	Terminologia	Objetivo	Etapas	Foco	Diagrama
(HYLAND; WOOD, 2011)	Processo	Os autores abordam uma variedade de orientações sobre a criação, publicação e divulgação de dados conectados, com ênfase nas necessidades das organizações governamentais.	*Identificar, *Modelar, *Nomear, *Descrever, *Converter, *Publicar, *Manter	Nas orientações básicas para a publicação.	
(JANSSEN; ZUIDERWIJK, 2012)	Processo	Discutem a publicação de dados abertos como um processo.	*Pesquisa, *Encontra, *Processa e Visualiza, *Usa, *Discute.	Investigam a relação entre a publicação de dados abertos e os usuários dos dados.	
(VILLAZÓN-TERRAZAS <i>et al.</i> , 2011)	Orientações metodológicas	Englobam orientações metodológicas para gerar, publicar e explorar dados conectados do governo.	*Especificação, *Modelagem, *Geração, *Publicação, *Exploração	Nas orientações para gerar, publicar e explorar dados. O trabalho foi desenvolvido com base na experiência dos autores, sobre a produção de dados conectados governamentais	

(SCHARFFE <i>et al.</i> , 2012)	Ciclo de vida	Focado para o Datalift, sendo que é uma visão de três trabalhos apresentados na literatura, sendo estes: (VILLAZÓN-TERRAZAS <i>et al.</i> , 2011), (HYLAND; WOOD, 2011), (HAUSENBLAS, 2011).	*Seleção, *Conversão, *Publicação, *Interligação, *Exploração	Nos passos que são semelhantes à arquitetura implementada para a plataforma Datalift.	 <p>Um diagrama vertical que mostra o ciclo de vida de dados conectados. No topo, há um retângulo amarelo rotulado 'Exploitation' dentro de um contêiner maior rotulado 'Published linked-data'. Abaixo dele, há outros retângulos amarelos rotulados 'Interlinking', 'Publication', 'Conversion' e 'Selection' em ordem descendente. Na base, há um retângulo rotulado 'Raw Data'. Uma seta preta aponta para cima, conectando 'Raw Data' a 'Exploitation'.</p>
(HAUSENBLAS, 2011)	Ciclo de vida	Com base em vários anos de experiência em publicação e consumo de dados conectados ao longo dos últimos anos, eles identificam fases fundamentais, que englobam uma infinidade de ciclos de vida de dados conectados.	*Conhecimento dos dados, *Modelagem, *Publicação, *Descoberta, *Integração, *Casos de uso;	Tecnológico	 <p>Um diagrama de fluxo horizontal que mostra o ciclo de vida de dados conectados. No topo, há três retângulos azuis rotulados 'data awareness', 'modelling' e 'publishing' conectados por setas para a direita. Abaixo deles, há outros três retângulos azuis rotulados 'discovery', 'integration' e 'use cases' também conectados por setas para a direita. Uma seta curva no topo conecta 'publishing' de volta para 'discovery', completando o ciclo.</p>

<p>(AUER <i>et al.</i>, 2011)</p>	<p>Ciclo de vida</p>	<p>O objetivo é facilitar a distribuição e instalação de ferramentas e componentes de software que suportem o ciclo de vida de publicação de dados conectados.</p>	<p>*Armazenamento, *Autoria, *Fusão, *Enriquecimento, *Qualidade, *Evolução, *Exploração, *Extração;</p>	<p>Tecnológico</p>	
<p>(BROEK <i>et al.</i>, 2014)</p>	<p>Ciclo de vida</p>	<p>Com base na literatura apresentam as etapas de um ciclo de vida com foco na visão estratégica de publicação de dados.</p>	<p>*Identificação, *Preparação, *Publicação, *Reuso, *Avaliação;</p>	<p>Nas orientações básicas para a publicação de dados a nível estratégico.</p>	

Fonte: Da Autora

Em uma análise geral de todos estes trabalhos encontrados na literatura, percebe-se que em sua maioria a publicação de dados é descrita como um processo sequencial e unidimensional de fases, que um grupo muitas vezes não especificado de *stakeholders* realiza frequentemente para fornecer uma quantidade de dados à um público geral. Percebe-se também que estes trabalhos focam em sua maioria nos processos operacionais, ignorando muitas vezes a mensuração do consumo dos dados e o *feedback* dos usuários, sendo que isto viabiliza não somente melhorias quanto a qualidade dos dados disponibilizados, mas também potencializa novos negócios. Visto que, a qualidade, a disponibilidade e a utilização dos dados têm um papel central no sucesso daqueles que publicam seus dados (HAIDER, W.; HAIDER, A., 2013).

Para compreender e auxiliar a preparar as partes interessadas em publicar dados optou-se por discutir dois destes trabalhos. No entanto, é evidente que todos os trabalhos compartilham atividades comuns (e às vezes sobrepostas). Por exemplo, todos eles identificam a necessidade de especificar, modelar e publicar dados em um formato padrão. Em essência, eles capturam as mesmas tarefas necessárias, mas fornecem diferentes limites entre essas tarefas, devido ao escopo e a natureza do trabalho. Portanto, observa-se que um fluxo de trabalho não é melhor que o outro, eles são simplesmente maneiras diferentes de visualizar um processo e/ou abordar a disponibilização de dados na Web.

O primeiro ciclo analisado, foi proposto por Auer *et al.* (2011), o objetivo do projeto que construiu este ciclo é desenvolver e elencar ferramentas presentes na literatura, para apoiar a criação de dados conectados, sendo que o ciclo é composto por oito estágios, sendo estes: (1)Armazenamento, (2)Autoria, (3)Fusão, (4)Enriquecimento, (5)Qualidade, (6)Evolução, (7)Exploração e (8)Extração; Os autores deixam claro que estas etapas não devem ser tratadas separadamente, porém salientam que devem ser investigados métodos que facilitem benefícios mútuos para os estágios desenvolvidos, com o intuito de resolver os desafios que são apresentados.

Já o ciclo de vida proposto por Broek *et al.* (2014), o segundo ciclo analisado, foi desenvolvido baseado nos ciclos de vida encontrados pelos autores na literatura, sendo que este ciclo foi aperfeiçoado com base nas lições aprendidas em um estudo de caso desenvolvido em uma empresa semi-pública na Holanda. O ciclo de vida apresenta cinco fases, sendo estas: (1) identificação, (2) preparação, (3) publicação, (4) reuso, (5) avaliação.

Observa-se nos demais trabalhos presentes na literatura, assim como no ciclo de vida proposto por Auer *et al.* (2011), que estes levam em conta exclusivamente os processos operacionais de publicação de dados abertos conectados (como por exemplo, extração, limpeza, publicação e manutenção de dados), ignorando os processos estratégicos (como tomada de decisão, alinhamento com a alta gerência e fatores legais). Assim, as decisões sobre quais dados serão publicados, quem irá extrair os

dados, como os dados são editados, como os dados podem ser acessados, quais licenças estão disponíveis e questões como: privacidade e responsabilidade não são tratadas na maioria dos trabalhos encontrados na literatura. Estes processos estratégicos mais gerais, são abordados apenas pelo ciclo de vida proposto por Broek *et al.* (2014), sendo que os autores também definem *stakeholders* que auxiliam no processo de publicação de dados abertos conectados.

Através desta análise é possível perceber que os dois ciclos de vida, são complementares, ou seja, o ciclo proposto por Auer *et al.* (2011) fornece ferramentas, focado em como tratar os dados, com o intuito de minimizar o árduo trabalho de publicar dados abertos conectados e o ciclo proposto por Broek *et al.* (2014) apresenta as orientações a nível estratégico, assim como os *stakeholder* envolvidos em todo o processo. No entanto, um item que chama a atenção ao analisar estes estudos é que pouco se fala em licença aberta quando abordado o contexto de dados abertos e dados abertos conectados. Portanto, para complementar esta análise a próxima seção aborda as boas práticas sobre a publicação de dados na Web, com o intuito de encontrar práticas e padrões que possam complementar as estratégias abordadas nos ciclos de vida apresentados.

2.1.5 Boas Práticas para Dados na Web

A abertura e a flexibilidade da Web criou novos desafios para os editores de dados e para os consumidores de dados, por exemplo, como representar, descrever e disponibilizar dados de uma maneira que seja fácil de encontrar-los e compreendê-los. Ao contrário dos bancos de dados convencionais, onde há um único modelo de dados para representar os dados e um sistema de gerenciamento de banco de dados para controlar o acesso, os dados disponibilizados na Web estão disponíveis sobre diferentes representações e há diversas maneiras de ter acesso a estes dados (LÓSCIO *et al.*, 2017).

Contudo, ao publicar dados devem ser tomados alguns cuidados além do formato de disponibilização e o acesso, com objetivo que indivíduos possam fazer uso dos dados de forma confiável, assim como facilmente ligar a outros dados. Para isso se faz necessário conhecer os ciclos de vida que guiam a publicação de dados, bem como estar ciente das boas práticas propostas pelo W3C (LÓSCIO *et al.*, 2017, 2016), além do contexto já abordado no início da seção 2.1. Portanto, na sequência são abordadas e caracterizadas as boas práticas apresentadas pelo W3C:

Quadro 9 – Boas Práticas para Dados na Web

BP1: Fornecer metadados
Fornecer metadados tanto para usuários pessoas quanto para aplicações de computadores.
BP2: Fornecer metadados descritivos
Os seres humanos serão capazes de interpretar a natureza do conjunto de dados e suas distribuições, e os agentes de software serão capazes de descobrir automaticamente conjuntos de dados e distribuições.
BP3: Fornecer metadados estruturais
Os seres humanos serão capazes de interpretar o esquema de um conjunto de dados, e os agentes de software serão capazes de processar automaticamente os dados das distribuições.
BP4: Fornecer informações sobre a licença de dados
Os seres humanos serão capazes de compreender a licença de dados, descrevendo eventuais restrições impostas à utilização de certos dados, agentes de software serão capazes de detectar automaticamente a licença de dados de uma distribuição.
BP5: Fornecer informações de proveniência dos dados
Os seres humanos serão capazes de identificar a origem dos conjuntos de dados, e agentes de software serão capazes de processar automaticamente informações de proveniência
BP6: Fornecer informação de qualidade de dados
Os seres humanos e os agentes de software serão capazes de avaliar a qualidade e, portanto, a adequação de um conjunto de dados para a sua aplicação.
BP7: Fornecer indicador de versão
Os seres humanos e os agentes de software poderão facilmente determinar qual versão de um conjunto de dados.
BP8: Fornecer o histórico de versões
Os seres humanos e os agentes de software serão capazes de entender como o conjunto de dados muda de versão para versão e como duas versões específicas diferem.
BP9: Usar URIs persistentes como identificadores de conjuntos de dados
Os conjuntos de dados ou informações sobre conjuntos de dados poderão ser descobertos e citados ao longo do tempo, independentemente da sua disponibilidade ou do formato dos dados.
BP10: Usar URIs persistentes como identificadores dentro de conjuntos de dados
Os itens de dados serão relacionados em toda a Web criando um espaço global de informação acessível a humanos e máquinas.
BP11: Atribuir URIs para as versões dos conjuntos de dados e séries
Os seres humanos e os agentes de software serão capazes de referenciar versões específicas de um conjunto de dados, séries de conjunto de dados, bem como a versão mais recente de um conjunto de dados.

BP12: Usar formatos de dados padronizados legíveis por máquina
Máquinas serão capazes de ler e processar dados publicados na Web e os seres humanos serão capazes de usar ferramentas computacionais para manipular os dados.
BP13: Usar representações de dados que sejam independentes de localidade (locale neutral)
Os seres humanos e os agentes de software serão capazes de interpretar o significado do conjunto de caracteres (strings) que representam datas, horas, moedas, números, entre outros, com precisão.
BP14: Fornecer dados em vários formatos
Tantos usuários quanto possível serão capazes de utilizar os dados sem ter que transformá-los em seu formato preferido.
BP15: Reutilizar vocabulários, dando preferência aos padronizados
Interoperabilidade e consenso entre os publicadores e consumidores de dados serão reforçados.
BP16: Escolher o nível de formalização adequado
Os casos de aplicação mais prováveis serão apoiados com não mais complexidade do que o necessário.
BP17: Fornecer 'bulk download'
transferências de arquivos grandes, ou seja, que exigem mais tempo do que um usuário típico consideraria razoável, serão possíveis por meio de protocolos de transferência de arquivos dedicados.
BP18: Fornecer subconjuntos para conjuntos de dados grandes
Os seres humanos e as aplicações serão capazes de acessar subconjuntos de um conjunto de dados, em vez de todo o conjunto. Isso proporcionará aos consumidores o acesso aos dados com uma elevada proporção de dados que são realmente necessários em comparação aos dados desnecessários. APIs podem ser usadas para filtrar os dados disponíveis. A granularidade de acesso aos dados poderá ser definida de acordo com as necessidades do domínio e as demandas de desempenho das aplicações.
BP19: Usar <i>content negotiation</i> para servir os dados disponíveis em vários formatos
<i>Content negotiation</i> permitirá que diferentes recursos ou representações diferentes de um mesmo recurso possam ser servidas de acordo com a requisição feita pelo cliente.
BP20: Fornecer acesso em tempo real
Aplicações serão capazes de acessar os dados em tempo real ou quase em tempo real; em tempo real significa um intervalo de milissegundos até alguns segundos após a criação de dados.
BP21: Fornecer dados atualizados
Os dados na Web serão atualizados em tempo hábil para que os dados disponíveis on-line reflitam os dados mais recentes divulgados através de qualquer canal. Quando novos dados estiverem disponíveis, logo que possível, serão publicados na Web.

BP22: Fornecer uma explicação para os dados que não estão disponíveis
Os consumidores saberão que os dados que são referenciados a partir do conjunto de dados não estão disponíveis ou se estão disponíveis sob diferentes condições.
BP23: Tornar os dados disponíveis através de uma API
Os desenvolvedores terão acesso aos dados para uso em seus próprios aplicativos, com dados atualizados e sem a necessidade de esforço por parte dos consumidores. As aplicações serão capazes de obter dados específicos por meio de consultas à API.
BP24: Usar padrões Web como base para construção de APIs
Desenvolvedores que tenham alguma experiência com APIs baseadas em padrões Web, tais como o REST, já deverão ter um conhecimento inicial de como usar a API. Além disso, será mais fácil dar manutenção na API.
BP25: Fornecer documentação completa de sua API
Os desenvolvedores poderão obter informações detalhadas sobre cada chamada para a API, incluindo os parâmetros necessários e o retorno esperado, ou seja, todo o conjunto de informações relacionadas à API. O conjunto de valores - como usá-lo, avisos de alterações recentes, informações de contato e assim por diante - deve ser descrito e facilmente navegável na Web. Ele também permite que as máquinas acessem a documentação da API para ajudar os desenvolvedores a criar o software cliente da API.
BP26: Evitar alterações que afetem o funcionamento de sua API
O código do desenvolvedor deve continuar válido após alterações na API. Os desenvolvedores devem ser notificados das melhorias feitas na API e devem ser capazes de fazer uso delas. Quebrar alterações em sua API será raro e, se ocorrer, os desenvolvedores terão tempo e informações suficientes para adaptar o seu código, aumentando a confiança na API. Alterações na API deverão ser anunciadas no site da documentação da API.
BP27: Preservar identificadores
A URI de um conjunto de dados sempre levará para o conjunto propriamente dito ou então redirecionará para um recurso com informações sobre ele.
BP28: Avaliar a cobertura do conjunto de dados
Os usuários serão capazes de fazer uso de dados arquivados no futuro.
BP29: Coletar <i>feedback</i> dos consumidores de dados
Os consumidores de dados serão capazes de fornecer <i>feedback</i> e avaliações sobre conjuntos de dados e distribuições.
BP30: Compartilhar o <i>feedback</i> disponível
Os consumidores serão capazes de avaliar os tipos de erros que afetam o conjunto de dados, avaliar experiências de outros usuários e ter a certeza de que o publicador trata os problemas de forma adequada. Os consumidores também serão capazes de determinar se outros usuários já fizeram comentários semelhantes, poupando-lhes a submissão de relatórios desnecessários e poupando os publicadores de terem que lidar com duplicatas.
BP31: Enriquecer dados por meio da geração de novos dados

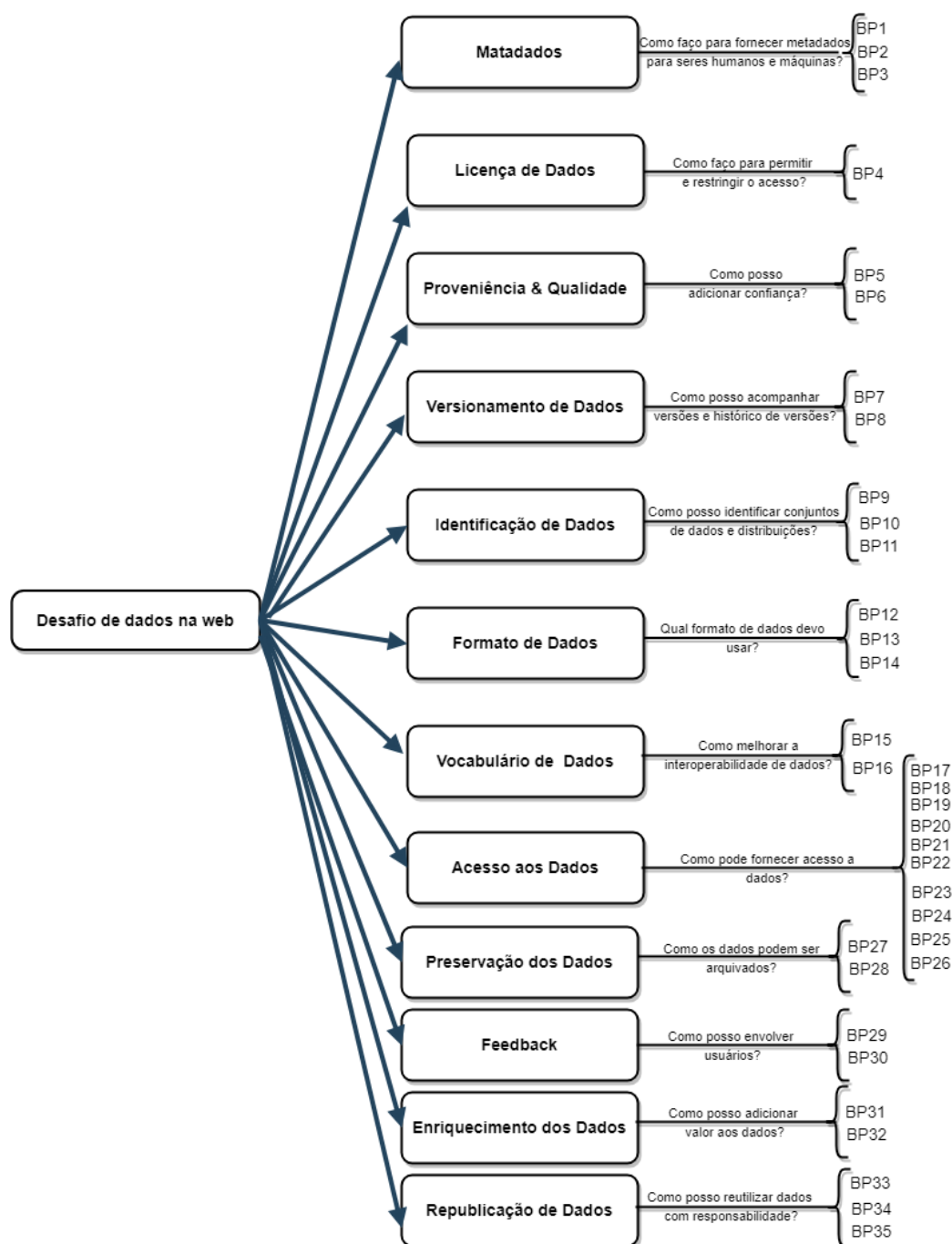
Os conjuntos de dados com valores nulos poderão ser "corrigidos" a partir do preenchimento de tais valores. Estrutura poderá ser conferida aos dados e sua utilidade poderá ser melhorada se forem adicionadas medidas ou atributos relevantes. Porém, tal adição só deverá ser feita se não alterar os resultados analíticos, o significado ou o poder estatístico dos dados.
BP32: Fornecer visualizações complementares
Complementar os conjuntos de dados com possíveis visualizações permitirá que os consumidores humanos tenham uma visão imediata sobre os dados, apresentando-os de forma que possam ser facilmente compreendidos.
BP33: Fornecer <i>feedback</i> para o publicador original
Uma melhor comunicação entre publicadores e consumidores fará com que seja mais fácil para os publicadores originais determinar como os dados que eles publicam estão sendo usados. Isso ajudará a justificar o investimento na publicação dos dados. Os publicadores também serão informados de medidas que podem ser tomadas para melhorar a qualidade dos seus dados.
BP34: Obedecer os termos de licença
Os publicadores serão capazes de confiar que seu trabalho está sendo reutilizado de acordo com os seus requisitos de licenciamento, tornando-os mais propensos a continuar com a publicação dos dados. Reutilizadores de dados vão ser capazes de licenciar adequadamente os trabalhos derivados a partir de dados previamente publicados.
BP35: Citar a publicação original do conjunto
Os consumidores finais serão capazes de avaliar a origem dos dados e os esforços dos publicadores originais serão reconhecidos. A cadeia de proveniência para os dados na Web será rastreável de volta ao seu publicador original.

Fonte: (LÓSCIO *et al.*, 2016)

Estas boas práticas descritas no Quadro 9 foram desenvolvidas para oferecer orientação técnica para a publicação de dados na Web, contribuindo para melhorar a relação entre publicadores e consumidores de dados (LÓSCIO *et al.*, 2016). Os autores deste trabalho abordam o contexto de dados na Web, no entanto percebe-se que estas boas práticas contemplam o âmbito de dados abertos e dados abertos conectados. Destaca-se que dados na Web possuem um universo mais amplo que dados abertos conforme pode ser observado na Figura 4, ou seja, nem todo dado na Web necessariamente é dado aberto.

Complementando as boas práticas identificadas no trabalho de Lóscio *et al.* (2017) estas foram relacionadas aos desafios de dados na Web, conforme pode ser observado na Figura 5.

Figura 5 – Desafios de dados na Web.



Fonte: (LÓSCIO *et al.*, 2017)

Dado o contexto, o W3C também descreve boas práticas para a publicação de dados abertos conectados (HYLAND *et al.*, 2014), sendo estas apresentadas no Quadro 10. Neste mesmo artigo do W3C discutem-se estas boas práticas e no primeiro passo "preparar os *stakeholders*" os autores discutem três ciclos de vida que ajudam a preparar os *stakeholders*, sendo estes os trabalhos de (HYLAND; WOOD, 2011; HAUSENBLAS, 2011; VILLAZÓN-TERRAZAS *et al.*, 2011) que já foram abordados e discutidos na seção 2.1.4.

Quadro 10 – Boas práticas para a publicação de dados abertos conectados

Passo 1: Preparar os <i>stakeholders</i>
Preparar as partes interessadas explicando o processo de criação e manutenção de dados abertos conectados.
Passo 2: Selecione um <i>dataset</i>
Selecione um conjunto de dados que ofereça benefícios a outras pessoas para reutilização.
Passo 3: Modelar os dados
Modelar dados conectados envolve representar objetos e como eles estão relacionados de maneira independente da aplicação.
Passo 4: Especificar uma licença apropriada
A reutilização de dados é mais provável de ocorrer quando há uma declaração clara sobre a origem, propriedade e termos relacionados ao uso dos dados publicados.
Passo 5: Bons URIs para dados conectados
A Web utiliza o URI como um sistema único de identificação global. O escopo global de URIs promove "efeitos de rede" em larga escala.
Passo 6: Use vocabulário padrão
Descreva objetos com vocabulários previamente definidos sempre que possível. Estenda os vocabulários padrão sempre que necessário e crie vocabulários (somente quando necessário) que sigam as boas práticas sempre que possível.
Passo 7: Converter dados
Converta dados em uma representação de dados conectados. Isso geralmente é feito por script ou outros processos automatizados.
Passo 8: Fornecer acesso aos dados por máquina
Forneça várias formas de mecanismos de pesquisa e outros processos automatizados acessarem seus dados usando mecanismos da Web.
Passo 9: Anunciar novos <i>datasets</i>
Lembre-se de anunciar novos <i>datasets</i> em um domínio autoritativo. Importante, lembre-se de que, como um editor de dados abertos conectados, um contrato social implícito está em vigor.
Passo 10: Reconhecer o contrato social
Reconheça sua responsabilidade em manter os dados depois de publicados. Assegure-se de que <i>datasets</i> permaneçam disponíveis onde sua organização diz que está e é que será mantido ao longo do tempo.

Fonte: (HYLAND *et al.*, 2014)

As boas práticas foram desenvolvidas com o intuito de auxiliar aqueles que desejam abrir seus dados. No entanto, percebe-se que mesmo com as boas práticas estipuladas ainda há problemas relacionados a dados abertos e que estes poderiam

ser minimizados se as boas práticas fossem consideradas tanto na publicação quanto na pós-publicação dos dados. Neste sentido, percebe-se a necessidade de verificar como estão sendo abertos os dados e a pós-publicação, quais são as estratégias para manter os dados publicados, assim como se há fomento para o uso dos dados disponibilizados.

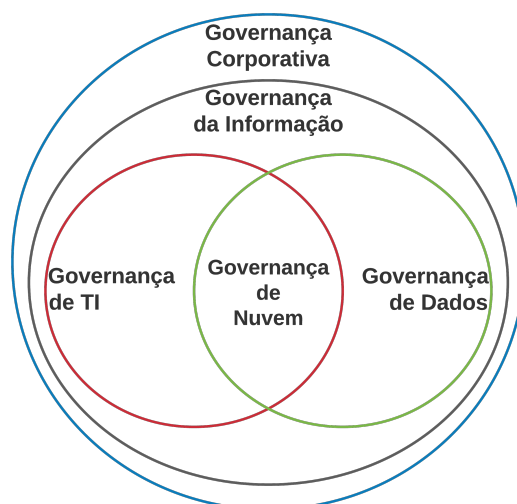
Nesta perspectiva, as práticas da governança de dados podem contribuir com a estratégia de abertura de dados das organizações, possibilitando evidenciar a necessidade de melhorias em diferentes aspectos; para isso, a próxima seção discute governança e governança de dados com o intuito de contribuir com este cenário.

2.2 GOVERNANÇA E GOVERNANÇA DE DADOS

O termo "governança", em geral, refere-se à forma como uma organização garante que as estratégias são definidas, monitoradas e alcançadas (WEBER *et al.*, 2009).

Com o surgimento de diversos domínios de governança como governança corporativa, governança de Tecnologia da Informação (TI), governança da informação e governança de dados, observa-se que alguns autores abordam esses domínios de governança como equivalentes e por este motivo diferenciar esses domínios principalmente com relação à governança de dados é essencial para a realização deste trabalho. Neste sentido, a Figura 6 ilustra uma representação gráfica da interrelação entre estes domínios.

Figura 6 – Inter-relação entre os domínios da governança



Fonte: (AL-RUITHE *et al.*, 2018), tradução nossa

A TI, recentemente tem sido abordada como a espinha dorsal das organizações e como consequência, o conceito de governança de TI se tornou importante para as organizações e constantemente é atrelado à governança corporativa. Assim,

a governança de TI, semelhante a governança corporativa, é o processo de estabelecer autoridade, responsabilidades e comunicação, junto com políticas, padrões, mecanismos de controle e medidas para permitir o cumprimento de funções definidas e responsabilidades (LEE, S. U. *et al.*, 2017).

No entanto a governança de TI é uma área madura em relação à governança de dados, isso é perceptível devido o alinhamento com as práticas da governança corporativa (AL-RUITHE *et al.*, 2018). Além da governança de TI, a governança de dados também tem um papel significativo no alinhamento dos negócios e da governança corporativa, visto que a governança de dados pode ser usada para resolver uma variedade de problemas relacionados aos dados e informação (KAMIOKA *et al.*, 2016).

Alguns autores diferenciam a governança de dados da governança de TI em princípios e práticas (WENDE, 2007; AL-RUITHE *et al.*, 2016). Em princípio, a governança de dados é projetada para a governança de ativos de dados, enquanto a governança de TI foca nas decisões sobre investimentos em TI ou seja, na prática, a governança de TI é projetada principalmente em torno do hardware e dos aplicativos de uma organização, e não de seus dados.

Em uma analogia com uma instalação hidráulica de uma casa a governança de TI preocupa-se com os encanamentos, os reservatórios e demais estruturas para garantir que a água flua conforme o esperado, já a governança de dados preocupa-se com os aspectos, como qualidade, turbidez, entre outros aspectos, acerca da água. Além disso a governança de TI é uma maneira de garantir que os investimentos no departamento de TI façam sentido nos negócios e criem valor para a organização (MERSINGER, 2016). Assim, percebe-se que a governança de TI se concentra em um âmbito mais amplo, por sua vez a governança de dados se concentra em dados brutos, sedo um nível mais rudimentar para implementar a governança (SMALLWOOD, 2014 apud MERSINGER, 2016).

Da mesma maneira como o conceito de governança de dados e de TI muitas vezes são abordados como equivalentes, os conceitos de gestão e governança muitas vezes são confundidos e, para isso é importante compreender e diferenciar as funções destes conceitos. Por exemplo, para o contexto da administração pública citam-se as seguintes funções para gestão:

- Implementar programas;
- Garantir a conformidade com as regulamentações;
- Revisar e reportar o progresso de ações;
- Garantir a eficiência administrativa;
- Manter a comunicação com as partes interessadas; e
- Avaliar o desempenho e aprender((TRIBUNAL DE CONTAS DA UNIÃO) p.31).

Já para a governança, são citadas as funções:

- Definir o direcionamento estratégico;

- Supervisionar a gestão;
- Envolver as partes interessadas;
- Gerenciar riscos estratégicos;
- Gerenciar conflitos internos;
- Auditar e avaliar o sistema de gestão e controle; e
- Promover a *accountability* (prestação de contas e responsabilidade) e a transparência((TRIBUNAL DE CONTAS DA UNIÃO) p.31).

Percebe-se que a gestão é inerente e integrada aos processos organizacionais, sendo responsável pelo planejamento, execução, controle, ação, enfim, pelo manejo dos recursos e poderes colocados à disposição de organizações para atingir seus objetivos, já a governança direciona, monitora, supervisiona e avalia a atuação da gestão, com vistas ao atendimento das necessidades e expectativas dos cidadãos e demais partes interessadas (TRIBUNAL DE CONTAS DA UNIÃO, 2014), como pode ser observado na Figura 7.

Figura 7 – Relação entre governança e gestão



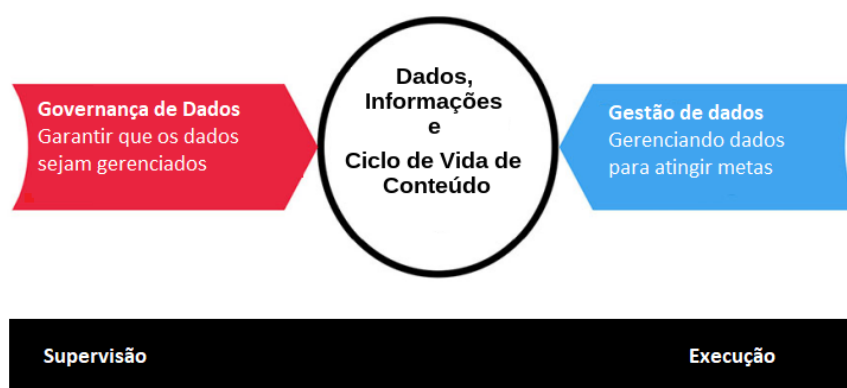
Fonte: (TRIBUNAL DE CONTAS DA UNIÃO, 2014)

A gestão de dados muitas vezes é vista apenas como atividades de TI, cujo objetivo é organizar e controlar seus recursos de dados para que eles sejam acessíveis, confiáveis e oportunos sempre que os usuários solicitarem. No entanto, vale ressaltar que a gestão de dados abrange todo o ciclo de vida de ativos de dados, desde a criação dos dados até a exclusão ou remoção destes (MOSLEY *et al.*, 2009). Portanto percebe-se que a gestão de dados é uma disciplina de negócios de alto nível abrangendo além das atividade de TI, o planejamento e a execução de políticas, práticas e projetos que adquirem, controlam, protegem e fornecem aumento no valor de ativos de dados e informações (MOSLEY *et al.*, 2009).

A governança de dados pode ser compreendida como: "o exercício de autoridade, controle e tomada de decisão compartilhada (planejamento, monitoramento e fiscalização) sobre a gestão de ativos de dados" (DAMA INTERNATIONAL, 2017). Wende (2007) também argumenta que a governança de dados é diferente da gestão de dados, sendo que a governança complementa a gestão, mas não a substitui. Neste

sentido, segundo (DAMA INTERNATIONAL, 2017) assim como um auditor controla os processos financeiros, mas na verdade não executa o gerenciamento financeiro, a governança de dados garante que os dados sejam gerenciados adequadamente sem a execução direta do gerenciamento de dados. Ou seja, a governança de dados representa uma separação inerente do dever entre supervisão e execução, veja a Figura 8.

Figura 8 – Governança de dados e gestão de dados



Fonte: (DAMA INTERNATIONAL, 2017)

Em resumo, observa-se que a governança de modo geral reúne um conjunto de práticas que visam aperfeiçoar à gestão, tendo como intuito preservar ao longo do tempo a relevância dos ativos que se aplicam às práticas. Na literatura três termos são emergentes e indicaram a execução da governança de dados, estes são: **1) definir, 2) implementar e 3) monitorar** (ALHASSAN *et al.*, 2016).

Portanto, nota-se que a governança de dados pode auxiliar as organizações a criarem uma missão, alcançar clareza, aumentar a confiança no uso de dados organizacionais, estabelecer responsabilidades, manter o escopo, focar e definir sucessos mensuráveis (CHEONG; CHANG, 2007). Além disso, discute-se que uma governança de dados eficaz poderá trazer diversos benefícios, sendo estes: treinamento da gerência e da equipe para adotar abordagens comuns para problemas de dados, construir padrões repetíveis de processos, reduzindo custos e aumentando a eficácia através da coordenação de esforços, e garantindo a transparência dos processos (KAMIOKA *et al.*, 2016; OTTO, 2011).

Neste sentido surgiram diversos estudos para contribuir com a governança de dados, são exemplos: Domínios de Decisões da Governança de Dados (KHATRI; BROWN, 2010), o *Framework* de Governança de Dados de NotreDame, (CHAPPLE, 2013), o *Framework* Dataflux (DATAFLUX CORPORATION, 2007), *Framework* de Governança de Dados para Gestão de Ativos (HAIDER, W.; HAIDER, A., 2013),

DAMA-DMBOK (DAMA INTERNATIONAL, 2017) e o *Framework* de Governança de Dados (PANIAN, 2010).

Nesta continuidade, para a realização deste trabalho, fez-se o estudo dos seguintes trabalhos:

- *Framework* DAMA-DMBOK, referência na área que visa promover o entendimento, desenvolvimento de prática de governança de dados e informações para apoiar as estratégias de negócios (MOSLEY *et al.*, 2009), visto que este é um dos trabalhos mais completos, citados da área e também por apresentar estudos recentes com atualizações em diversos aspectos (DAMA INTERNATIONAL, 2017).
- *Framework* de Governança de Dados de NotreDame (CHAPPLE, 2013), devido a algumas características que corroboram com o contexto de dados abertos, sendo estas: I) os autores abordam a tecnologia como a base da governança, mas destacam que esta não deve ser o todo e II) o acesso aos dados é colocado no topo do *framework* pois os autores destacam que todos os indivíduos envolvidos devem ter a capacidade de acessar os dados de que precisam de maneira oportuna e eficaz.
- E por fim, analisou-se o *Framework* de Governança de Dados de Panian (2010), que será utilizado no desenvolvimento deste trabalho, devido às particularidades que serão discutidas na sequência.

2.2.1 **Framework de Governança de Dados**

A governança de dados é uma disciplina que vem sendo aplicada pelas organizações com o intuito de aumentar o valor dos ativos de dados. Para isso as organizações precisam estabelecer padrões, políticas e processos para o uso o desenvolvimento e o gerenciamento dos dados, ainda neste sentido é necessário que se tenha uma infraestrutura tecnológica para suporte (PANIAN, 2010).

Conforme abordado no trabalho de Panian (2010) o *framework* viabiliza alcançar os objetivos da governança, sendo estes:

- Garantir que os dados atendam às necessidades do negócio.
- Proteger, gerenciar e desenvolver dados como um ativo valioso.
- Reduzir os custos de gerenciamento de dados.

A governança de dados é vista como uma disciplina importante e que visa garantir a conformidade dos dados (AL-RUITHE *et al.*, 2018). No trabalho de Panian (2010) a governança de dados estabelece padrões, políticas e processos exigidos

pelos regulamentos e políticas de governança corporativa e ajuda a automatizar tarefas relacionadas à conformidade, reduzindo custos, também tem por característica garantir a prestação de contas e auditoria dos dados.

Para isso, os autores discutem seis atributos-chave que devem ser considerados em uma iniciativa de governança de dados (PANIAN, 2010), sendo estes:

- **Acessibilidade:** garantir que todos os dados possam ser acessados, independentemente de sua fonte ou estrutura.
- **Disponibilidade:** garantir que os dados estejam disponíveis para os usuários e aplicações, quando, onde e como necessário.
- **Qualidade:** garantir a completude, acurácia e integridade dos dados.
- **Consistência:** garantir que o significado dos dados seja consistente e conciliados em todos os sistemas.
- **Auditabilidade:** garantir que haja controle e auditoria nos dados.
- **Segurança:** garantir o acesso seguro aos dados.

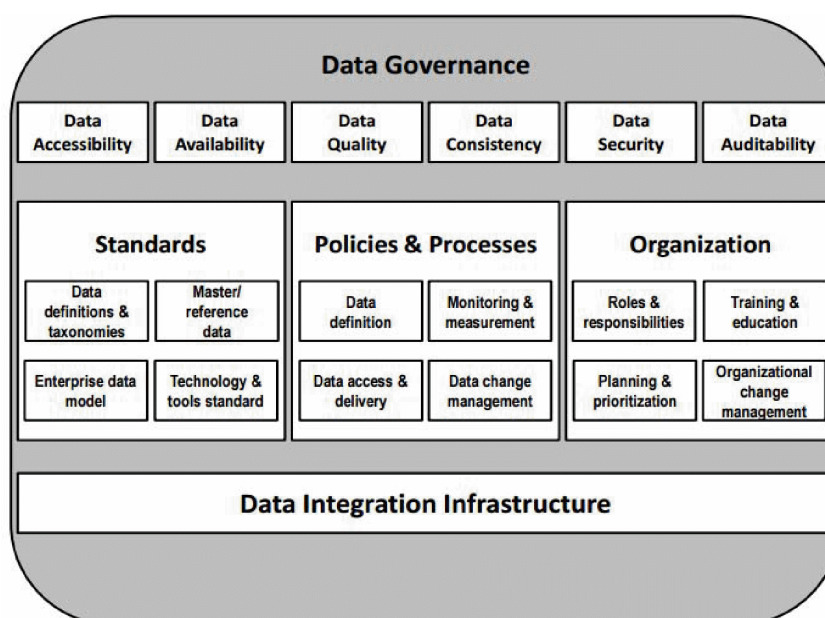
Além disso, Panian (2010) discute quatro componentes para garantir os seis atributos chave, sendo estes:

- **Padrões:** uma função fundamental da governança de dados é estabelecer os padrões para dados em uma empresa (FISHER, 2009). As empresas precisam estabelecer definições de dados e fazer cumprir os padrões técnicos e de desenvolvimento relacionados aos dados.
- **Políticas e Processos:** estabelecer e aplicar políticas e processos em torno da criação, desenvolvimento, controle, gestão e auditoria de dados é a base de uma prática eficaz de governança de dados. As empresas precisam definir dados e regras de negócios relacionadas a dados, controle de acesso e entrega de dados, estabelecer mecanismos de monitoramento e medição, e gerenciar alterações nos dados.
- **Organização:** indiscutivelmente a questão mais importante que as empresas devem abordar ao lançar uma iniciativa de governança de dados é como projetar a estrutura. As empresas precisam definir as funções e responsabilidades dentro da organização. A organização pode incluir vários e diferentes papéis em diferentes níveis, envolvendo pessoal de negócios e de TI - de conselhos executivos a implementadores do dia-a-dia, como administradores de dados e analistas de dados.

- **Tecnologia:** hipoteticamente, as empresas podem embarcar em uma iniciativa de governança de dados sem uma base de infraestrutura de tecnologia. Na verdade, muitas organizações lançam seus programas iniciais de governança de dados usando ferramentas manuais - planilhas, e documentos do *Word* - para capturar definições de dados e processos de documentos. No entanto, rapidamente percebe-se que este tipo manual de abordagem é severamente limitada. É difícil garantir alta qualidade e disponibilidade de dados, a segurança está em risco, dada a natureza *ad hoc* da abordagem e manter documentação detalhada é uma tarefa quase impraticável. De fato, é quase impossível atingir os objetivos finais de governança de dados usando uma abordagem manual. A tecnologia pode auxiliar a automatizar e dimensionar o desenvolvimento de padrões de governança de dados e de políticas e processos.

Considerando os componentes e os atributos chave, a Figura 9 apresenta os blocos de construção de uma estrutura de governança de dados eficaz. Segundo Panian (2010) essa estrutura leva em consideração os padrões, políticas e processos, a estrutura organizacional e a infraestrutura de tecnologia necessários para tornar os dados acessíveis, disponíveis, de alta qualidade, consistentes, auditáveis e seguros em toda a organização.

Figura 9 – Framework de Governança de Dados



Fonte: (PANIAN, 2010)

Não há uma abordagem única para implementar a governança de dados para todos os tipos de organizações (BEGG; CAIRA, 2012). Isso significa que a abordagem de cada organização para a governança de dados pode ser diferente. É, portanto, difícil capturar todas as diferentes visões sobre governança de dados. Neste sentido, na

revisão de literatura do trabalho de Al-Ruithe *et al.* (2018) aborda-se uma taxonomia da governança de dados, onde os autores apresentam três categorias principais que caracterizam o consenso dos autores da área, sendo: (I) pessoas e fatores organizacionais, (II) política e processo e (III) tecnologia.

Com isso, percebe-se que as categorias apresentadas encaixam-se nos componentes abordados no trabalho de Panian (2010). Sendo observado em **Organização** pessoas e fatores organizacionais, conforme previamente apresentado. **Tecnologia** é abordada nos dois estudos. **Políticas e processos** no trabalho de Al-Ruithe *et al.* (2018), além de políticas, processos e princípios também aborda **padrões** como elemento que compõe esta categoria, ou seja o *framework* apresentado fornece uma visão resumida sobre a governança de dados, onde podem ser observados componentes já discutidos por vários autores da área, sendo que estes viabilizam iniciativas de governança de dados.

Considerando o contexto apresentado, a fim de se obter uma visão detalhada acerca da iniciativa de governança pode-se avaliar a situação atual da organização. Para isso, pode ser feito o uso de modelos de maturidade, com o intuito de evidenciar a maturidade (capacidade) de cada processo, categoria, componente, área, ou dimensões em análise. Nesta perspectiva a próxima seção aborda o tema.

2.3 MODELOS DE MATURIDADE

Maturidade significa literalmente "maturação" ou "madureza"; isso implica em um estado ou qualidade de maduro, ou seja, um modelo de maturidade descreve a transição de um estado inicial para um estado mais avançado, possivelmente por meio de vários estados intermediários. Segundo P. Fraser *et al.* (2002, p.247, tradução nossa) a maturidade pode ser vista como “uma combinação da presença de um processo e da atitude da organização em relação a ele”.

Os modelos de maturidade refletem o grau em que os principais processos ou atividades são definidos, gerenciados e executados de forma eficaz. Eles tipicamente descrevem as características de uma atividade em vários níveis diferentes de desempenho (FRASER *et al.*, 2003). Deste modo, a mensuração da maturidade pode ser considerada útil e relevante para identificar o que medir e como medir determinado tópico em uma organização, com o intuito de verificar sua situação, e o que fazer para obter um nível de maturidade mais alto.

O contexto de modelo de maturidade de estágios foi aplicado pela primeira vez por Richard L. Nolan (NOLAN, 1973), da Universidade de Harvard, que em 1973 publicou os estágios de um modelo para organizações de TI. Conseqüentemente, as áreas de TI e engenharia de software têm estado por trás da criação de vários modelos de maturidade e são os fundamentos dos modelos de maturidade mais bem estabelecidos e abrangentes (CARALLI *et al.*, 2012).

Após o trabalho de Nolan, Watts Humphrey começou a desenvolver conceitos de maturidade de processo ainda enquanto trabalhava na *International Business Machines Corporation* (IBM) (HUMPHREY, Watts S, 1987). No entanto, o desenvolvimento efetivo de um modelo baseado nesses conceitos começou em 1986, quando Humphrey se juntou ao *Software Engineering Institute* (SEI) (HUMPHREY, Watts S; SWEET, 1987; HUMPHREY, W. S., 1988).

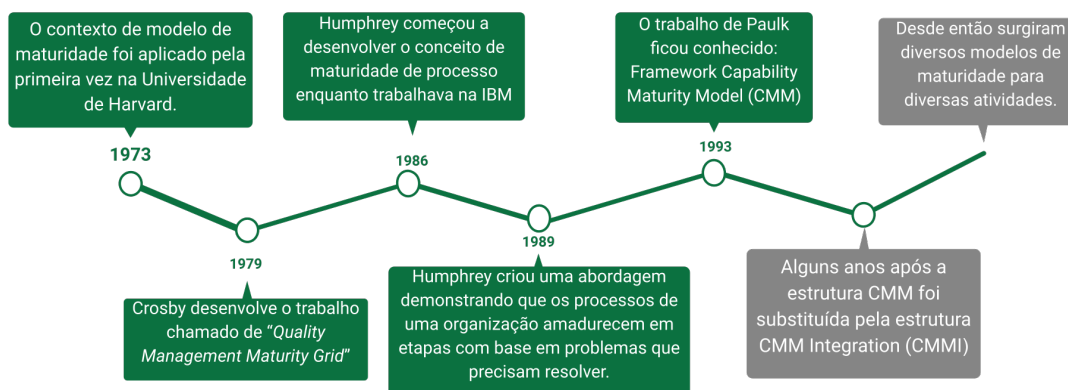
A Força Aérea dos EUA, solicitou à SEI que iniciasse a formalização da Estrutura de Maturidade do Processo, para que o Departamento de Defesa dos EUA pudesse avaliar os contratados de software, visto a necessidade de um modelo consistente para avaliar processos e práticas de negócios. Com o intuito de resolver os problemas relacionados ao desenvolvimento de software (projetos executados acima do orçamento e processos mais lentos do que o esperado) que ocorriam nas forças armadas. Para isso, Humphrey combinou seu trabalho com o trabalho desenvolvido por Crosby (1979), conhecido como "*Quality Management Maturity Grid*".

Assim, Humphrey criou uma abordagem demonstrando que os processos de uma organização amadurecem em etapas com base em problemas que precisam resolver (HUMPHREY, Watts S, 1989). Esse sistema de medição de maturidade de processo ficou conhecido mais tarde através do trabalho de Paulk *et al.* (1993) como: o *framework Capability Maturity Model* (CMM), que capturou as boas práticas organizacionais para o desenvolvimento de software.

Alguns anos depois, a estrutura do CMM foi substituída pela estrutura CMM Integration (*Capability Maturity Model Integration* (CMMI)) (SOFTWARE ENGINEERING INSTITUTE, 2018), que integrou modelos para engenharia de software, engenharia de sistemas, aquisição de software, sistemas e prestação de serviços em um único modelo com núcleo compartilhado (CARALLI *et al.*, 2012). Desde então, foram desenvolvidos diversos modelos de maturidade para diversas atividades, incluindo: desenvolvimento de software (PAULK *et al.*, 1993), inovação (CHIESA *et al.*, 1996), design de produto (FRASER *et al.*, 2001), gestão do conhecimento (HSIEH *et al.*, 2009), dentre outros.

Nesta sequência a Figura 10 aborda um resumo da linha temporal acerca dos da evolução dos estudos de modelos de maturidade.

Figura 10 – Resumo da evolução dos estudos acerca de modelos de maturidade.



Fonte: Da Autora

2.3.1 Definindo Modelos de Maturidade

Percebe-se que modelos de maturidade existem para diferentes áreas, pois viabilizam uma maneira das organizações abordarem problemas e desafios de maneira estruturada, fornecendo tanto um *benchmark* para avaliar recursos quanto um roteiro para aprimorá-los. De modo geral, um modelo de maturidade pode ser definido como um conjunto de características, atributos, indicadores ou padrões que representam a progressão em um domínio ou disciplina em particular (CARALLI *et al.*, 2012).

No entanto, "os modelos de maturidade permitem não apenas garantir um *benchmark* de boas práticas e identificar áreas de melhoria, mas também para rastrear as metas de melhoria com base no status atual e estratégia da organização, e definir a melhor maneira de alcançá-las" (ELMAALLAM; KRIOUILE, 2014, pg 1). Portanto, os modelos de maturidade podem ser vistos como artefatos para determinar o status de uma empresa e como adotar "medidas de melhoria" (BECKER *et al.*, 2009).

A avaliação da maturidade pode ser realizada por um auditor externo, ou por autoavaliação (FRASER, P. *et al.*, 2002). A autoavaliação pode ser realizada por um único indivíduo. No entanto, destaca-se o enriquecimento da avaliação se esta for exercida em equipe. Uma razão para isso é que ao envolver pessoas de diferentes grupos funcionais dentro de uma equipe, a avaliação tem uma perspectiva multifuncional e fornece uma oportunidade para o consenso (FRASER, P. *et al.*, 2002). Segundo Caralli *et al.* (2012) pode-se fazer uso de um modelo de avaliação de maturidade para melhorar processos, práticas e desempenho fornecendo às organizações e comunidades a capacidade de:

- **Desempenho interno de *benchmark*:** as organizações podem determinar onde estão em sua jornada de melhoria e definir metas para futuros investimentos.

- **Catalisar a melhoria de desempenho:** durante um período de tempo, as organizações podem usar o modelo como base para a melhoria contínua do desempenho.
- **Catalisar melhorias no desempenho da comunidade:** as organizações podem não apenas comparar seu desempenho com organizações semelhantes, mas também determinar um perfil de desempenho de "comunidade".
- **Criar e desenvolver uma linguagem comum:** cria-se uma maneira consistente de pensar e comunicar sobre um domínio que é incorporado na linguagem, modelo ou na taxonomia.

Segundo Paulk *et al.* (1993) a maturidade dos processos pode ser definida como a medida em que um processo específico é explicitamente definido, gerenciado, medido, controlado e eficaz. Ou seja, a maturidade implica que o processo seja bem compreendido, apoiado por documentação e treinamento, e aplicado de forma consistente em projetos em toda a organização e continuamente monitorado e aprimorado por seus usuários.

Neste sentido, observa-se que os modelos de maturidade são modelos conceituais baseados na ideia de que as capacidades organizacionais se desenvolvem através de uma série de estágios desejados ou lógicos, de um estado inicial até um estado maduro (GOTTSCHALK; SOLLI-SÆTHER, 2006). Maduro, neste contexto, significa melhor preparado para cumprir o seu propósito, ou seja, ter um maior nível de sofisticação, capacidade ou disponibilidade de características específicas (METTLER; ROHNER, 2009).

2.3.1.1 Componentes Essenciais

Diversas propostas de modelos de maturidade são apresentadas na literatura (PÖPPELBUSS; RÖGLINGER, 2011), embora tenham sido desenvolvidos sobre diferentes perspectivas eles compartilham algumas características como: ter um número de dimensões ou áreas de processo em vários estágios ou níveis de maturidade, com uma descrição de desempenho caracterizado em vários níveis de granularidade.

Alguns dos componentes que podem ou não estar presentes em cada modelo segundo P. Fraser *et al.* (2002), são:

- Um número de níveis (tipicamente de 3 a 6);
- Um descritor para cada nível (como inicial - repetível - definido - gerenciado - otimizado);
- Uma descrição genérica ou resumo das características de cada nível como um todo;

- Um número de dimensões ou áreas de processo;
- Um número de elementos ou atividades para cada área de processo;
- E uma descrição de cada atividade, como ela pode ser realizada em cada nível de maturidade;

Na literatura é possível encontrar diferentes abordagens sobre os componentes necessários para a estruturação de um modelo de maturidade. Segundo Salah *et al.* (2014) os modelos de maturidade devem abranger três componentes, sendo estes: (I) o modelo de referência, (II) a escala de desempenho e o (III) procedimento de avaliação.

Para Van Steenbergem *et al.* (2013) os componentes básicos de um modelo de maturidade são: (I) vários níveis gerais de maturidade, (II) vários aspectos ou áreas, que podem ser desenvolvidas ao longo de um caminho evolutivo para alcançar os níveis de maturidade definidos e (III) descrições de cada passo no caminho evolutivo. Sendo que um modelo de maturidade pode conter sugestões sobre como executar as várias etapas em termos de ações de melhoria.

Para Caralli *et al.* (2012, p. 8, tradução nossa) os componentes essenciais de um modelo de maturidade, são:

- **Níveis:** Os níveis representam os estados de transição em um modelo de maturidade. Dependendo da arquitetura do modelo, os níveis podem descrever uma etapa ou um patamar progressivo ou podem representar uma expressão de capacidade ou outro atributo que pode ser medido pelo modelo.
- **Domínios do modelo:** Os domínios são um meio de agrupar atributos semelhantes em uma área de importância para o assunto e a intenção do modelo. Nos modelos de maturidade de capacidade, os domínios são chamados de “áreas de processo” porque são uma coleção de processos que compõem um processo ou disciplina maior (como no caso da engenharia de software). Dependendo do modelo, os usuários podem se concentrar em melhorar um único domínio ou um grupo de domínios.
- **Atributos:** Atributos representam o conteúdo principal do modelo agrupado por domínio e nível. Eles são tipicamente baseados em práticas observadas, padrões ou outros conhecimentos especializados e podem ser expressos como características, indicadores, práticas ou processos. No caso de um modelo de maturidade de capacidade, os atributos também podem expressar qualidades de maturidade organizacional (como planejamento) que são importantes para apoiar a melhoria do processo (independentemente do processo que está sendo modelado).
- **Avaliação e Métodos de Pontuação:** Métodos de avaliação e pontuação são desenvolvidos para facilitar a avaliação usando o modelo como base. Podem ser formais ou informais, liderados por especialistas ou auto-aplicados. Os métodos de pontuação são algoritmos criados pela comunidade para garantir a consistência das avaliações e um padrão comum de medição. Os métodos de pontuação podem incluir a ponderação (de modo que os atributos importantes são valorizados em relação aos menos importantes) ou podem avaliar diferentes tipos de coleta de dados de maneiras diferentes (como fornecer notas mais altas para evidências documentadas em oposição a dados baseados em entrevistas).
- **Roteiro de Melhorias:** Além de serem usados para *benchmarking*, os modelos de maturidade podem ser usados para guiar os esforços de melhoria. Muitos desses modelos prescreveram métodos para identificar um escopo de melhoria, diagnosticar o estado atual e, em seguida, planejar e implementar a melhoria e verificar se ela ocorreu. Esses métodos

definem um ciclo clássico de planejar-fazer-chechar-agir (PDCA) no qual um modelo de maturidade se encaixa como base para a melhoria (SHEWHART; DEMING, William Edwards, 1986 apud CARALLI *et al.*, 2012), (DEMING, W Edwards, 1952 apud CARALLI *et al.*, 2012).

Na literatura também é possível observar uma série de características distintivas dos modelos de maturidade existentes, possibilitando evidenciar tipos distintos de modelos.

2.3.1.2 Tipos de Modelos de Maturidade

Em primeiro lugar é necessário compreender a natureza dos modelos de maturidade que esta pode ser descritiva, prescritiva ou comparativa (DE BRUIN *et al.*, 2005). Isso está diretamente relacionado ao propósito do modelo: se ele significa puramente avaliar o estado atual de uma organização (a natureza é descritiva), se pretende sugerir ações de melhoria (a natureza é prescritiva) ou se é direcionado à comparação com outras organizações (a natureza é comparativa). A maioria dos autores focam no propósito descritivo ou comparativo na criação de modelos de maturidade. Pouca atenção é dada à forma de selecionar ações de melhoria (natureza prescritiva) (PÖPPELBUSS; RÖGLINGER, 2011).

Ainda, de modo geral os modelos de maturidade podem ser categorizados conforme três tipos, sendo estes: modelos de progressão, modelos de capacidade e modelos híbridos.

- Modelos de Progressão

Os modelos de maturidade com foco em progressão representam escala simples de uma característica, indicador, atributo ou padrão, em que o aumento dos níveis de maturidade indica alguma progressão da maturidade do atributo.

Os modelos de progressão podem ser medidos independentemente e são tipicamente caracterizados por um foco nos atributos do modelo, em vez de atributos que definem especificamente a maturidade (CARALLI *et al.*, 2012). Observa-se portanto, que o principal objetivo de um modelo de progressão é fornecer um roteiro de progressão ou melhoria, conforme expresso por versões melhores de um atributo à medida que a escala progride. Caralli *et al.* (2012) abordam um exemplo de modelo de maturidade de progressão para a contagem dada por:

lápiz e papel → ábaco → calculadora → computador

Observa-se portanto que os níveis são rotulados em relação a uma "etapa" na progressão. No exemplo apresentado, os níveis poderiam ser caracterizados como: "primitivo" e "ferramentas ativas". Nota-se que modelos de progressão muitas

vezes são descritos como "modelos de maturidade de capacidade", quando na verdade eles não medem a capacidade ou a maturidade do processo (um conceito fundamental nos modelos de capacidade).

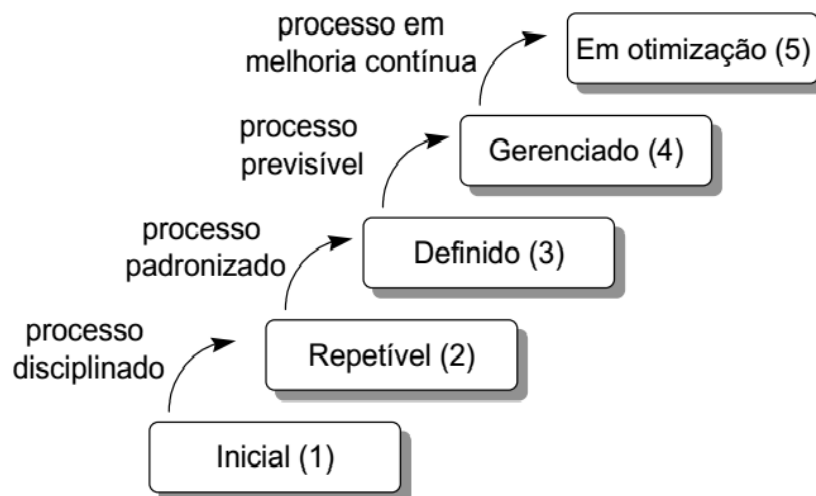
- Modelos de maturidade de capacidade

Nos modelos de maturidade de capacidade, a dimensão que está sendo medida é uma representação da capacidade organizacional em torno de um conjunto de características, indicadores, atributos ou padrões, muitas vezes expressos como "processos" (por isso, os modelos de maturidade de capacidade são frequentemente caracterizados como: "modelos de processo") (CARALLI *et al.*, 2012). Este tipo de modelo além de medir a capacidade de execução de uma determinada tarefa, sendo esta simples ou complexa permite analisar a capacidade organizacional de modo mais abrangente refletindo a maturidade da cultura e o grau em que as capacidades são incorporadas (ou "institucionalizadas") na cultura (CARALLI *et al.*, 2012). Portanto, observa-se que os modelos de maturidade de capacidade descrevem estados de maturidade organizacional relativos à maturidade do processo, como:

ad hoc → gerenciado → definido → gerenciado quantitativamente → otimizado

Devido à natureza genérica da escala de maturidade do processo, a estrutura de maturidade básica por exemplo do CMM, conforme pode ser observada na Figura 11, pode ser aplicada a outros domínios (CARALLI *et al.*, 2012).

Figura 11 – Cinco níveis do CMM



Fonte: (PAULK *et al.*, 1993)

- Modelos Híbridos

A ideia de se ter características do modelo progressivo e do modelo de maturidade de capacidade podem criar um modelo híbrido. Esse tipo de modelo reflete transições entre níveis semelhantes a um modelo de capacidade (ou seja, que descrevem a maturidade da capacidade), mas usam arquiteturalmente as características, indicadores, atributos ou padrões de um modelo de progressão (CARALLI *et al.*, 2012). Neste sentido, os modelos de maturidade híbridos fornecem o rigor de um modelo de maturidade de capacidade e incluem a facilidade de uso e compreensão dos modelos de progressão (CARALLI *et al.*, 2012).

Segundo P. Fraser *et al.* (2002) um questionário que faz uso de uma escala *Likert*, quando construído de uma maneira particular, pode ser considerado uma forma simples de modelo de maturidade. Neste caso, a questão é simplesmente uma declaração de 'boa prática' e solicita-se ao respondente que classifique o desempenho relativo da organização numa escala de 1 a n. Isso é equivalente a uma avaliação de maturidade na qual apenas as características de alto nível são descritas. Modelos que combinam uma abordagem de questionário com definições de maturidade são referidos como híbridos (FRASER, P. *et al.*, 2002).

Conforme já mencionado o pioneiro dos modelos é o CMM e a partir disso percebe-se que este e suas variações derivam a maioria dos demais modelos de maturidade desenvolvidos, ou seja os modelos desenvolvidos são baseados no modelo genérico de maturidade de 5 níveis definidos a partir do CMM, e dentre estes podem ser distinguidos dois tipos (VAN STEENBERGEN *et al.*, 2008):

1. **Os modelos de 5 níveis fixos** - Esses modelos distinguem cinco níveis de maturidade. Para cada nível, várias áreas são definidas especificamente para esse nível. Essas áreas de foco devem ser implementadas satisfatoriamente para que a organização atinja esse nível específico.
2. **Os modelos de 5 níveis contínuos** - Esses modelos também distinguem cinco níveis gerais de maturidade e várias áreas de foco. A diferença com o modelos de 5 níveis fixos é que as áreas de foco não são atribuídas a um nível, mas dentro de cada área de foco os níveis são distintos.

Ainda, segundo Van Steenbergen *et al.* (2008) observa-se um terceiro tipo de modelo oriundo da área de teste de software, sendo estes:

3. **Os modelos orientados à áreas de foco** - Esses modelos partem da ideia de que existem mais que cinco níveis genéricos de maturidade, ou seja, cada área de foco tem seu próprio número de níveis de maturidade específicos. A

maturidade geral de uma organização é expressa como uma combinação dos níveis de maturidade dessas áreas de foco.

A diferença entre estes tipos de modelos de maturidade podem ser percebidas na Figura 12. Nota-se em **a)** o modelos de 5 níveis fixos, com um número de áreas foco associadas com cada nível de maturidade; já em **b)** o modelo de 5 níveis contínuo em que para cada área foco há os mesmo níveis de maturidade genéricos; já em **c)** nos modelos orientados à áreas de foco cada área de foco tem um número de níveis de maturidade específicos.

Figura 12 – Três tipos de modelos de maturidade

	1	2	3	4	5
FA 1	X				
FA 2	X				
FA 3		X			
FA 4		X			
...					

a)

	1	2	3	4	5
FA 1	X	X	X	X	X
FA 2	X	X	X	X	X
FA 3	X	X	X	X	X
FA 4	X	X	X	X	X
...					

b)

	1	2	3	4	5	6	7	...
FA 1	X				X			
FA 2		X		X				
FA 3	X		X			X		
FA 4				X			X	
...								

c)

Fonte: (VAN STEENBERGEN *et al.*, 2008)

Apesar das diferenças nos tipos de modelos de maturidade, a maioria deles está de acordo com alguns princípios estruturais. Essa estrutura é importante pois fornece uma relação entre os objetivos, a avaliações e as boas práticas, além de facilitar os relacionamentos entre os recursos atuais e as possibilidades de melhoria (CARALLI *et al.*, 2012). Na sequência serão abordadas as características e os conceitos acerca de modelos de maturidade orientados à áreas de foco, também denominados por este trabalho como modelos de maturidade de áreas foco.

2.3.1.3 Modelos de Maturidade de Áreas Foco

Os modelos de maturidade orientados á áreas foco são oriundos da área de teste de software (KOOMEN; POL, 1999), sendo que posteriormente foram desenvolvidos modelos desse gênero, por exemplo, para os domínios de arquitetura corporativa (WAGTER *et al.*, 2005; BERG; STEENBERGEN, 2006; VAN STEENBERGEN *et al.*, 2010b) e gerenciamento de produtos (VAN DE WEERD *et al.*, 2006; METTLER; ROHNER, 2009; WEERD *et al.*, 2010) e a partir disso percebe-se que:

Um modelo de maturidade da áreas foco é usado para estabelecer a maturidade de uma organização em um domínio funcional específico. Um domínio funcional é descrito pelo conjunto de áreas foco que o constituem. Para cada área foco, um conjunto de capacidades é associada. As capacidades são justapostas em relação umas às outras em uma matriz de maturidade. Uma capacidade é posicionada à direita de outra capacidade para ser implementada após essa capacidade, pois depende de sua realização. Além de dependências difíceis, razões pragmáticas também podem informar esse posicionamento. Sediada sobre o posicionamento das capacidades na matriz de maturidade, vários níveis de maturidade podem ser distinguidos. Para estabelecer a maturidade real de uma organização no domínio funcional, é utilizado um instrumento de avaliação baseado na matriz de maturidade. O instrumento de avaliação contém questões de avaliação ligadas às capacidades. Respondendo as perguntas de avaliação para uma determinada organização produz uma avaliação detalhada. A avaliação preenchida pode ser traduzida em um perfil de maturidade para a organização, representado pelas células coloridas na matriz. Para orientar a organização no desenvolvimento incremental do domínio funcional, as ações de melhoria estão associadas às capacidades. Tudo isso junto constitui o modelo de maturidade da área de foco (VAN STEENBERGEN *et al.*, 2013, p.10, tradução nossa).

Para compreender os modelos de maturidade de áreas foco será feito o uso do trabalho de Van Steenbergen *et al.* (2010b). Este modelo de maturidade pode ser observado na Figura 13.

Figura 13 – Modelo de maturidade da área de foco para o gerenciamento de arquitetura corporativa.

Focus Area	Maturity Level	0	1	2	3	4	5	6	7	8	9	10	11	12	13
Development of architecture			A			B			C						
Use of architecture				A			B				C				
Alignment with business			A				B				C				
Alignment with the development process				A				B		C					
Alignment with operations						A			B			C			
Relationship to the as-is state						A				B					
Roles and responsibilities					A		B					C			
Coordination of developments								A			B				
Monitoring					A		B		C		D				
Quality management									A		B			C	
Maintenance of the architectural process								A		B		C			
Maintenance of architectural deliverables						A			B					C	
Commitment and motivation			A					B		C					
Architectural roles and training					A		B			C			D		
Use of an architectural method					A						B				C
Consultation				A		B				C					
Architectural tools								A				B			C
Budgeting and planning					A							B		C	

Fonte: (VAN STEENBERGEN *et al.*, 2010b)

A Figura 13 apresenta as áreas de foco na 1ª coluna da esquerda. As letras (A à D) à direita das áreas de foco indicam as capacidades de progressão relacionadas a cada uma das áreas de foco, conforme previamente explicado na seção 2.3.1.2. Por exemplo, a área foco Desenvolvimento de arquitetura (2ª linha da matriz) tem

as capacidades A: Arquitetura realizada como projeto, B: Arquitetura como um processo contínuo e C: Arquitetura como um processo de facilitação, representando uma progressão na maturidade.

O posicionamento das letras na matriz indicam a ordem na qual as capacidades das diferentes áreas de foco podem ser implementadas para desenvolver o domínio funcional da arquitetura corporativa de maneira equilibrada (VAN STEENBERGEN *et al.*, 2013). Com isso é possível perceber que os modelos de maturidade de áreas foco permitem distinguir mais de cinco níveis fixos de maturidade, indo um passo adiante e afirmando que cada área foco tem seu próprio número níveis de maturidade (VAN STEENBERGEN *et al.*, 2010a). Ao justapor todas as capacidades de todas as áreas foco em relação umas às outras, proporciona-se um caminho de desenvolvimento incremental e equilibrado. Essa justaposição de capacidade é feita posicionando as capacidades em uma matriz, conforme pode ser observado na Figura 13.

Na Figura 13 também é possível observar quatorze colunas que definem os níveis gerais de maturidade, sendo 0 o nível mais baixo e o nível 13 o mais alto. Esse nível geral de maturidade pode ser analisado pela coloração das células de cada linha até a primeira capacidade que ainda não foi implementada pela organização, criando um perfil de maturidade da organização.

A coluna mais à direita que contém apenas células coloridas indica o nível de maturidade da organização avaliada. Portanto, nota-se na Figura 13 que a organização avaliada possui o nível de maturidade 1, pois a **capacidade A do Uso de arquitetura** (3ª linha da matriz) na coluna que indica o nível de maturidade 2, não foi alcançado. Esta matriz também apresenta um desequilíbrio da área foco **Alinhamento com o processo de desenvolvimento**, sendo esta totalmente desenvolvida, enquanto outras áreas como, **Uso da arquitetura e Monitoramento**, ainda não são desenvolvidas (VAN STEENBERGEN *et al.*, 2010b). O posicionamento das capacidades em uma matriz tornam esse tipo de modelo adequado para expressar combinações, muitas vezes complexas, de diferentes fatores que determinam a eficácia de um domínio (VAN STEENBERGEN *et al.*, 2010a).

Portanto percebe-se que o desenvolvimento de um modelo de maturidade de áreas foco tem início com a definição de um domínio funcional e das áreas foco nas quais este domínio funcional específico consiste. Para cada uma dessas áreas foco, define-se uma série de capacidades que indicam a maturidade progressiva. As capacidades são específicas para cada uma das áreas foco.

O posicionamento das capacidades dentro da matriz devem representar as dependências existentes entre todas as capacidades de todas as áreas foco relativas umas às outras. Isso é feito colocando as capacidades que depende após a capacidade de dependência, mais à direita, na matriz, ver Figura14.

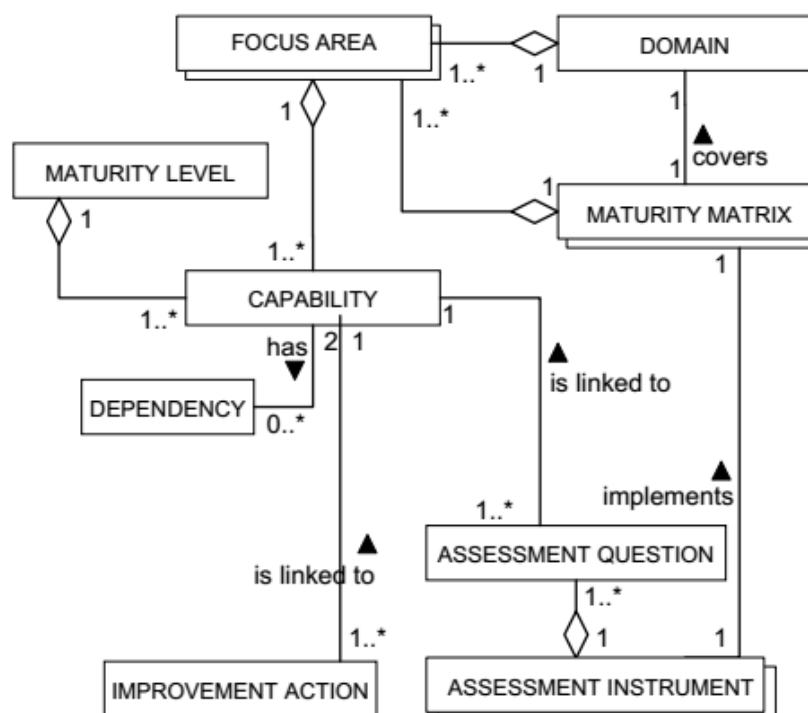
Figura 14 – Exemplificação das dependências

	Maturity Level	0	1	2	3
Focus Area					
Development of architecture			A		
Use of architecture				A	
Alignment with business			A		
Alignment with the development process				A	

Fonte: Da autora

A determinação da ordem das capacidades é feita com base em conhecimentos teóricos, conhecimentos empíricos e conhecimentos especializados do domínio funcional (VAN STEENBERGEN *et al.*, 2013). O metamodelo de concepção de modelos de maturidade de área foco é apresentado na Figura 15, onde aborda-se as relações entre os itens que o compõe. Este metamodelo faz uso de uma notação baseada em convenções *Unified Modeling Language* (UML), com alguns pequenos ajustes.

Figura 15 – Metamodelo do modelo de maturidade de áreas foco.



Fonte: (VAN STEENBERGEN *et al.*, 2010a)

Portanto, observa-se que ao considerar não somente cinco níveis fixos de maturidade, o modelo orientado à áreas foco torna-se mais flexível em relação de interdependências entre as áreas. Especialmente em domínios funcionais relativamente

novos, que exigem uma abordagem mais refinada para guiar a melhoria incremental (VAN STEENBERGEN *et al.*, 2010a).

Para dar continuidade na elaboração deste trabalho é necessário compreender os modelos de maturidade existentes para o contexto de dados abertos, o que possibilita evidenciar oportunidades de melhoria, assim como discutir e caracterizar o que estes trabalhos abordam. Nesta lógica a próxima seção apresenta esta discussão.

2.4 TRABALHOS RELACIONADOS

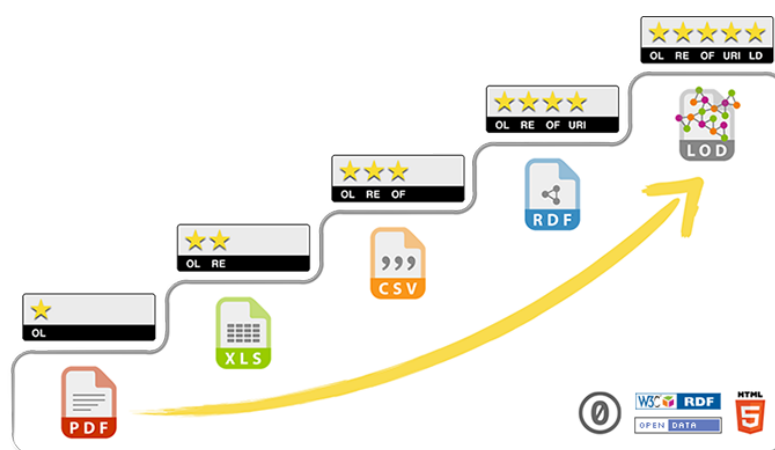
Para a elaboração desta seção foram realizadas buscas em base de dados, conforme será especificado na seção 3.1 item d, com o intuito de estudar as soluções existentes para modelos de maturidade de dados abertos, para com isso ser assertivo no desenvolvimento do artefato proposto. Caso encontre-se uma solução ideal para o problema que busca-se resolver, este poderia ser aperfeiçoado através desta pesquisa. Para isso, na sequência discute-se sobre os trabalho encontrados.

2.4.1 Classificação baseada em estrelas

Antes de abordar os modelos de maturidade encontrados na literatura, faz-se necessário discutir brevemente a classificação baseada em estrelas proposta por Berners Lee, visto que esta classificação é utilizada na avaliação da qualidade dos dados (NUSAPATI; SUNINDYO, s.d.; TORINO *et al.*, 2019; SISTO *et al.*, 2018). E pode ser equiparada a uma avaliação progressiva, conforme mencionada na seção 2.3.1.2.

A classificação baseada em estrelas proposta por Berners Lee foi criada para encorajar proprietários de dados, principalmente o governo à publicar seus dados; esta classificação pode ser observada na Figura 16.

Figura 16 – Classificação baseada em estrelas



Fonte: (BERNERS-LEE, 2011)

A proposta de Berners Lee faz uso de uma classificação de acordo com a quantidade de estrelas, sendo que a primeira estrela diz respeito a ser um formato legível por máquina e atribuída uma licença aberta, ou seja, há a necessidade de que os dados sejam abertos. O nível duas estrelas inclui a licença aberta mais um formato estruturado de dados, podendo fazer a utilização de um software proprietário, por exemplo, formato xls (Excel). Já para obter três estrelas é necessário se ter os anteriores mais um formato não proprietário, por exemplo, formato *Comma Separated Values* (CSV). Para obter quatro estrelas se faz necessário ter as estrelas anteriores e mais a utilização de padrões semânticos do W3C, como por exemplo, RDF e *Protocol and RDF Query Language* (SPARQL). No entanto, para se obter as cinco estrelas se faz necessário todas as estrelas anteriores acrescentando a ligação para dados de outras fontes com o intuito de obter um contexto mais amplo (BERNERS-LEE, 2011).

Esta classificação é uma das mais conhecidas e mais utilizadas, mas como pode se perceber leva em conta apenas critérios como: ser legível por máquina, a licença aberta e o formato para a classificação dos *datasets* disponibilizados. E por este motivo esta classificação é facilmente incorporada aos modelos de maturidade existentes.

2.4.2 Norma UNE 178301

Esta norma foi evidenciada através de um artigo sobre revisões de literatura para a avaliação da maturidade. No entanto, não foi possível recuperar o documento oficial na íntegra, referente a esta norma, e portanto para esta discussão fez-se o uso do trabalho de Silva e Pinheiro (2018) como base.

Esta norma aborda o contexto de cidades inteligentes, que visa ser uma referência para a administração e órgãos do setor público, no sentido de ajudar a definir, documentar e implantar projetos de dados abertos.

A norma UNE 178301 foi elaborada pelo comitê técnico de normalização da AENOR - AEN/CTN 178 - sobre Cidades inteligentes em janeiro de 2015. Esta norma estabelece um conjunto de métricas que levam em consideração aspectos que precisam ser contemplados nas iniciativas de dados abertos. As métricas e níveis propostos na norma, associados aos dados abertos, são considerados para o cálculo do indicador de dados abertos (UNE 178301, 2015 apud SILVA; PINHEIRO, 2018). Existem quatro níveis que caracterizam as posições alcançáveis (SILVA; PINHEIRO, 2018, p. 325), sendo estas:

- **Nível 0 - Resultados inexistentes:** considera que os valores medidos são esperados quando não existe uma iniciativa de abertura de dados e os resultados não são considerados relevantes.
- **Nível 1 - Resultados incipientes (informal):** considera que os valores medidos são esperados quando existe uma iniciativa informal de abertura de dados e seus resultados são considerados relevantes.

- **Nível 2 - Resultados existentes:** considera que os valores medidos são esperados quando existe uma iniciativa formal de abertura de dados e seus resultados são considerados relevantes.
- **Nível 3 - Resultados avançados:** considera que os valores medidos são esperados quando existe uma iniciativa formal de abertura dos dados que implementa as boas práticas.

A norma contempla cinco domínios: **i)** domínio estratégico, **ii)** domínio legal, **iii)** domínio organizacional, **iv)** domínio técnico e **v)** domínio econômico e social. Os domínios podem ser divididos em uma ou mais dimensões e cada dimensão tem as variáveis a serem avaliadas associadas a estes níveis. Para cada domínio analisado são atribuídos um peso e uma pontuação considerada para encontrar o nível de adequação à norma. O Quadro 11 apresenta o modelo de avaliação com os elementos e seu respectivos pesos.

Quadro 11 – Métricas de Avaliação Norma UNE 178301:2015

Domínios e Dimensões da Norma UNE 178301	Peso	Exemplo (pontuação obtida)
Domínio Estratégico		
Dimensão Estratégica		
Estratégia	25%	1
Liderança	50%	2
Compromisso de serviço	10%	0
Sustentabilidade econômica	15%	3
Domínio Legal		
Dimensão Legal		
Normas externas e internas	30%	1
Condições de uso e licenciamento	70%	2
Domínio Organizacional		
Dimensão Organizacional		
Unidade responsável	25%	1
Equipe de trabalho e capacitação	50%	2
Inventário	15%	0
Prioridade	10%	3
Domínio Medição		
Medição e cumprimento do processo	50%	2
Medição do uso e impacto	50%	0
Domínio Técnico		
Dimensão Disponibilidade		
Catálogo	30%	2
Presença no catálogo de informação pública	25%	0
Conjunto de dados documentados	15%	3

Categorização e busca	10%	1
Disponibilidade	10%	2
Referências persistentes e amigáveis	10%	0
Dimensão Acesso		
Acessibilidade/ Não discriminação	30%	2
Gratuidade	50%	0
Sistema de acesso	20%	3
Dimensão Qualidade dos Dados		
Dados primários	20%	2
Dados completos	20%	0
Dados documentados	20%	2
Dados tecnicamente corretos	10%	0
Dados georreferenciados	10%	3
Dados ligados	20%	2
Dimensão Atualização		
Processo de atualização	40%	2
Frequência de atualização	40%	0
Ampliação de conjunto de dados	20%	2
Domínio Econômico e Social		
Dimensão Reutilização		
Quantidade de dados	20%	1
Formato dos dados	40%	2
Vocabulário de dados padrão	40%	0
Dimensão Participação e Colaboração		
Transparência, participação e colaboração	30%	0
Resolução de reclamações e conflitos	20%	2
Fomento a reutilização	30%	0
Iniciativas de reutilizações desenvolvidas	20%	3

Fonte: (SILVA; PINHEIRO, 2018)

Para cada variável é calculado um valor por meio da seguinte fórmula: Valor = ((Pontuação ou nível * Peso) / 3 * 100). Ao final é somado o valor de todas as variáveis de todos os domínios e com isso é possível estabelecer um ranking. A pontuação máxima, nível 3 em todas as variáveis, é 1000. Os indicadores respectivos, em ordem crescente, são: de 0 a 200 = 1; de 201 a 400 = 2; de 401 a 600 = 3; de 601 a 800 = 4 e de 801 a 1000 = 5 (SILVA; PINHEIRO, 2018).

Os pesos atribuídos às variáveis não está detalhada na norma, também não há justificativas sobre a divisão dos domínios e dimensões. Esta norma é baseada

especificamente na legislação espanhola, portanto deve-se analisar cada contexto antes de aplicá-la (SILVA; PINHEIRO, 2018). Por este motivo e pela falta de acesso ao conteúdo completo da norma, optou-se por não considerar suas características na construção deste trabalho.

Na sequência serão discutidos dois trabalhos que serão utilizados como base para a estruturação do modelos de maturidade proposto.

2.4.3 Modelo de Maturidade de Dados Abertos

O Modelo de Maturidade de Dados Abertos apresentado por Dodds e Newman (2015) foi desenvolvido como um projeto conjunto entre o Instituto de Dados Abertos (ODI)¹ e o Departamento de Meio Ambiente, Alimentação e Assuntos Rurais (Defra)². O modelo oferece suporte à avaliação da eficiência da organização em suas atividades operacionais e estratégicas no que tange aos dados abertos. Também, fornece orientação às organizações sobre áreas potenciais para melhoria e possibilita comparar organizações para salientar seus respectivos pontos fortes e fracos, dando suporte a adoções mais amplas de boas práticas e ajudando a aperfeiçoar processos (DODDS; NEWMAN, 2015).

O Modelo foi desenvolvido com base em uma série de *workshops* com a participação de administradores de dados e especialistas em dados abertos do governo do Reino Unido e da comunidade de dados abertos em geral. Os *workshops* foram suplementados com pesquisa adicional sobre governança de dados e práticas de gestão. Algumas pesquisas comparativas sobre a concepção e o desenvolvimento de modelos de maturidade também foram esboçadas durante a análise. No entanto os autores pretendem revisar e ampliar o modelo no futuro, com base na experiência e com o uso do modelo (DODDS; NEWMAN, 2015). Uma ferramenta on-line também foi disponibilizada a *Open Data Pathway*³

O Modelo de Maturidade de Dados Abertos consiste de 15 atividades organizacionais: por exemplo, processos de liberação de dados ou adoção de normas comunitárias. Estas atividades estão agrupadas em cinco temas (**i**) processos de gestão de dados, **ii**) conhecimento e competências, **iii**) suporte ao cliente envolvido, **iv**) investimento e desempenho financeiro e **v**) supervisão estratégica que as categorizam. Isto reflete o objetivo de avaliar as diversas maneiras nas quais a prática de dados abertos pode impactar uma organização (DODDS; NEWMAN, 2015).

Neste modelo, as atividades também podem ser agrupadas de acordo com a sua relação com a publicação de dados, a reutilização de dados ou a ambas as áreas. Uma

¹ <http://theodi.org>

² <https://www.gov.uk/government/organisations/department-for-environment-food-rural-affairs>

³ <http://pathway.theodi.org/>

organização irá avaliar sua própria maturidade confrontando cada uma das atividades do modelo e produzindo uma pontuação de maturidade de 1-5 caracterizadas como: 1- Inicial, 2- Reprodutível, 3- Definido, 4- Gerenciado e 5- Otimizado.

Uma avaliação completa com base neste modelo dará à organização uma pontuação de maturidade em uma série de atividades, a saber, como um dado é liberado, como é governado e como conjuntos de dados são avaliados. A pontuação irá refletir a maturidade dos processos da organização em uma área específica e pode ser usada para identificar áreas para melhorias e definição de metas mensuráveis (DODDS; NEWMAN, 2015).

2.4.4 *Open Data Maturity Model*

O modelo de maturidade proposto por Solar *et al.* (2012), chamado *Open Data Maturity Model* (OD-MM) (Modelo de Maturidade de Dados Abertos), visa avaliar o comprometimento e as capacidades das agências públicas na busca dos princípios e práticas dos dados abertos, sendo que esta pesquisa abrange também os trabalhos de Solar *et al.* (2013), Solar *et al.* (2014a) e Solar *et al.* (2014b).

O modelo é baseado na definição de um conjunto de 3 elementos organizados em uma estrutura hierárquica. Seus principais elementos são os "Domínios", "Subdomínios" e "Variáveis Críticas". Os domínios são o coração do modelo, pois estão definidos em diferentes níveis de capacidade, sendo estes: estabelecimento e perspectiva legal, perspectiva tecnológica e reutilização de dados. Também foram definidos 9 subdomínios no total (3 por domínio). Sendo que existem 33 variáveis críticas distribuídas nos 9 Subdomínios (SOLAR *et al.*, 2012) e a definição dos pesos para as variáveis de cada subdomínio foi definida com no mínimo três especialistas em 3 rodadas, conforme pode ser observado no Quadro 12, para mais detalhes ver seção 4.1 do artigo de Solar *et al.* (2012).

Quadro 12 – Relação de domínios, subdomínios e variáveis do OD-MM

Domínio	Subdomínio	Variáveis	Peso	
Perspectiva Institucional e Jurídica	Estratégia	Estratégia	0.3	
		Liderança e Instituição	Liderança	0.4
			Instituição	0.3
	Leis e Regulamentos	Regulamentos externos	0.2	
		Regulamentos internos	0.4	
		Licenças	0.4	
	Gestão	Treinamento	0.3	
		Gestão de projetos	0.3	
		Avaliação de desempenho	0.4	

Perspectiva Tecnológica	Segurança e Disponibilidade	Segurança de sistemas	0.2
		Avaliação de dados	0.3
		Atualização de dados	0.3
		Ferramentas para mensurar nível de utilização	0.2
	Acesso	Leitura automatizada de dados	0.3
		Metadados	0.3
		Categorização e descoberta	0.2
		Uso de tecnologias semânticas	0.2
	Qualidade dos Dados	Formato dos dados	0.3
		Dados livres	0.25
		Dados primários	0.25
		Dados completos	0.2
Perspectiva do Cidadão e do Empreendedor	Reuso de Dados	Iniciativas de dados abertos desenvolvidas	0.3
		Número de dados abertos avaliados	0.3
		Ponto de acesso único	0.2
		Mensurar acesso aos dados	0.2
	Desenvolvedores	Dados gratuitos	0.2
		Reuso encorajado	0.4
		Queixas e conflitos	0.2
		Financiamento do projeto RISP	0.2
	Participação e Colaboração	Significado da participação e colaboração	0.3
		Transparência participativa	0.2
		Escuta ativa	0.3
		Mensurar o uso de dados (aplicações)	0.2

Fonte: (SOLAR *et al.*, 2014b)

Para o desenvolvimento deste modelo foram considerados trabalhos e artigos publicados, relatórios de comitês, bem como entrevistas telefônicas reuniões e teleconferências. Os autores contataram vários grupos interessados em dados abertos para entrevistá-los e usá-los como fontes primárias de informação. As entrevistas foram realizadas em duas rodadas. A primeira rodada foi com 9 especialistas que estiveram direta ou indiretamente envolvidos no desenvolvimento dos projetos de dados abertos, no governo do Chile, Colômbia e no estado de *New South Wales* (NSW), na Austrália.

As informações foram coletadas nesse processo e a partir da revisão da literatura, as questões foram estruturadas para a segunda rodada de entrevistas. A segunda rodada de entrevistas foi realizada com 4 especialistas com um domínio mais extenso e incluindo pessoas com experiência na região, governos latino-americanos e

representantes de fundações e defensores da sociedade civil do movimento de dados abertos (SOLAR *et al.*, 2012).

O nível de maturidade prescreve um caminho de melhoria para a organização (SOLAR *et al.*, 2012). Sendo que para cada subdomínio existe uma escala de medição incremental baseada em uma pontuação de 1 a 4, onde o nível 1 corresponde a capacidade inexistente, o nível 2 a capacidades emergentes (não formal), o nível 3 à capacidades existentes e o nível 4 corresponde a capacidades avançadas. Destaca-se que nível de maturidade de maneira geral é uma propriedade da organização como um todo.

Para validar o modelo OD-MM, este foi enviado a vários especialistas relacionados com o tema na América Latina. Representantes do Governo do Chile e da Colômbia, a sociedade civil e desenvolvedores de aplicativos de dados abertos, que fizeram seus comentários e depois de uma análise, levaram à modificações que permitiram melhorar o modelo para validá-lo (SOLAR *et al.*, 2012).

Com o modelo validado, em termos conceituais, foi desenvolvida uma ferramenta Web que permitiu a coleta de dados exigida pelo modelo para sua validação por meio de um piloto "in situ". Por essa razão, uma metodologia foi estabelecida para aplicar esse piloto MMOD⁴ (SOLAR *et al.*, 2012). Para a aplicação do projeto piloto, foram contatados especialistas do Chile, Colômbia, República Dominicana e El Salvador. Nesses países, uma amostra de 10 instituições públicas foram selecionadas, com base em critérios previamente definidos. Os resultados do piloto são abordados no trabalho de Solar *et al.* (2014b).

As iniciativas até agora descritas impulsionam o movimento de dados abertos e dados abertos conectados. Neste sentido, a próxima seção tem por objetivo abordar as etapas utilizadas para a realização desta pesquisa com o intuito de permitir a reprodutibilidade e com isso a assegurar a cientificidade.

⁴ Disponível em: <http://odmm.inf.santiago.usm.cl/> (interface em espanhol)

3 PROCEDIMENTOS METODOLÓGICOS

No desenvolvimento desta pesquisa optou-se pela pesquisa tecnológica. Observando que o conhecimento científico propõe teorias de aplicação ampla e diferencia-se do conhecimento tecnológico que é responsável pela geração de teorias restritas, cujo foco é a solução de problemas pontuais e específicos (JUNIOR *et al.*, 2017).

O conceito do *Design Science* foi introduzido por Herbert Simon, em 1969 e no Brasil, em 1981. Em seu trabalho o autor aborda a distinção entre a ciência natural e a *design science*, traduzida como ciência do projeto ou ciência do artificial (DRESCH *et al.*, 2015; SIMON, 1969). Neste contexto, observam-se diversas discussões acerca da necessidade, de se estudar o universo daquilo que é considerado artificial, observando que as "ciências do artificial se ocupam da concepção de artefatos que realizem objetivos" (SIMON, 1996).

A design science é a base epistemológica, quando trata-se do estudo do artificial. A Design Science Research (DSR), por sua vez, é o método que fundamenta e operacionaliza a condução da pesquisa quando o objetivo a ser alcançado é um artefato ou uma prescrição (DRESCH et al., 2015, p.67).

Cupani (2006) aborda uma visão tecnológica considerando que esta foca na construção de algo artificial, ou seja, um artefato. Onde um artefato não precisa ser necessariamente uma coisa, como (uma bicicleta ou remédio), mas pode ser a modificação do estado de um sistema natural como (o desviar de um rio), ou transformação de um sistema (ensinar alguém a ler). E neste sentido, os artefatos normalmente são projetados e construídos com o intuito de inserir alguma mudança em um sistema, resolver problemas e possibilitar o melhoramento de um desempenho.

Assim, alguns dos artefatos são caracterizados como (MARCH; SMITH, 1995):

- **Constructos:** também conhecidos como elementos conceituais que formam o vocabulário de um domínio, definindo os termos usados para descrever e especificar as possíveis soluções.
- **Modelos:** conjunto de proposições que expressam as relações entre os diversos conceitos de um domínio. Um modelo também pode ser considerado uma descrição ou representação de como as coisas são.
- **Métodos:** conjunto de passos usados para desempenhar determinada tarefa.
- **Instanciações:** é a execução de um artefato em seu ambiente, ou seja, são artefatos que operacionalizam outros artefatos(constructos, modelos e métodos).

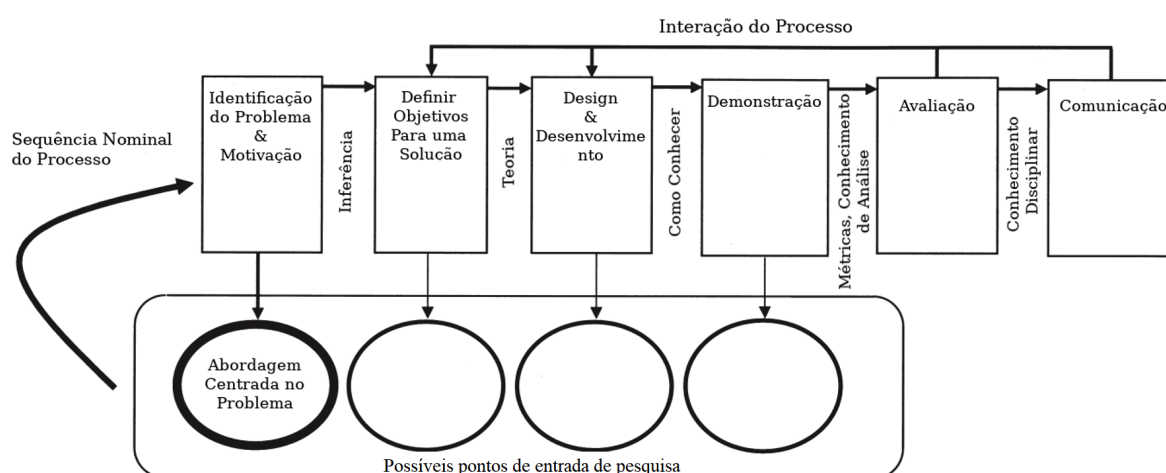
A adequada condução do método de pesquisa é um pré-requisito para a construção de um conhecimento científico confiável (DRESCH *et al.*, 2015). Sendo que,

observa-se a DSR aplicada como procedimento de criação de conhecimento para atingir dois objetivos diferentes em um projeto de pesquisa ao mesmo tempo: **I)** produzir conhecimento científico e **II)** resolver problemas organizacionais reais (DRESCH *et al.*, 2014).

Considerando o contexto apresentado fez-se um estudo sobre das diferentes formas para a condução das pesquisas fundamentadas pela *Design Science*. Dentre os trabalhos encontrados, destacam-se: (DRESCH *et al.*, 2015; PEFFERS *et al.*, 2007; WIERINGA, 2014; VAISHNAVI; KUECHLER, 2015) e (GLEASURE *et al.*, 2012).

Dentre os trabalhos encontrados na literatura observam-se similaridades entre estes, na sua maioria por utilizarem como base os princípios dos mesmos artigos, como é o caso de Peffers *et al.* (2007) e Vaishnavi e Kuechler (2015) e Gleasure *et al.* (2012). E neste âmbito, optou-se pelo método proposto por Peffers *et al.* (2007) para a condução desta pesquisa, sendo que este é um dos artigos mais citados sobre métodos para a condução da DSR e também por comparar sistematicamente as semelhanças entre os modelos de processo de DSR que o antecedem e a partir destes propor um processo geral (DENG; JI, 2018), este observado na Figura 17.

Figura 17 – Etapas da Metodologia *Design Science Research*

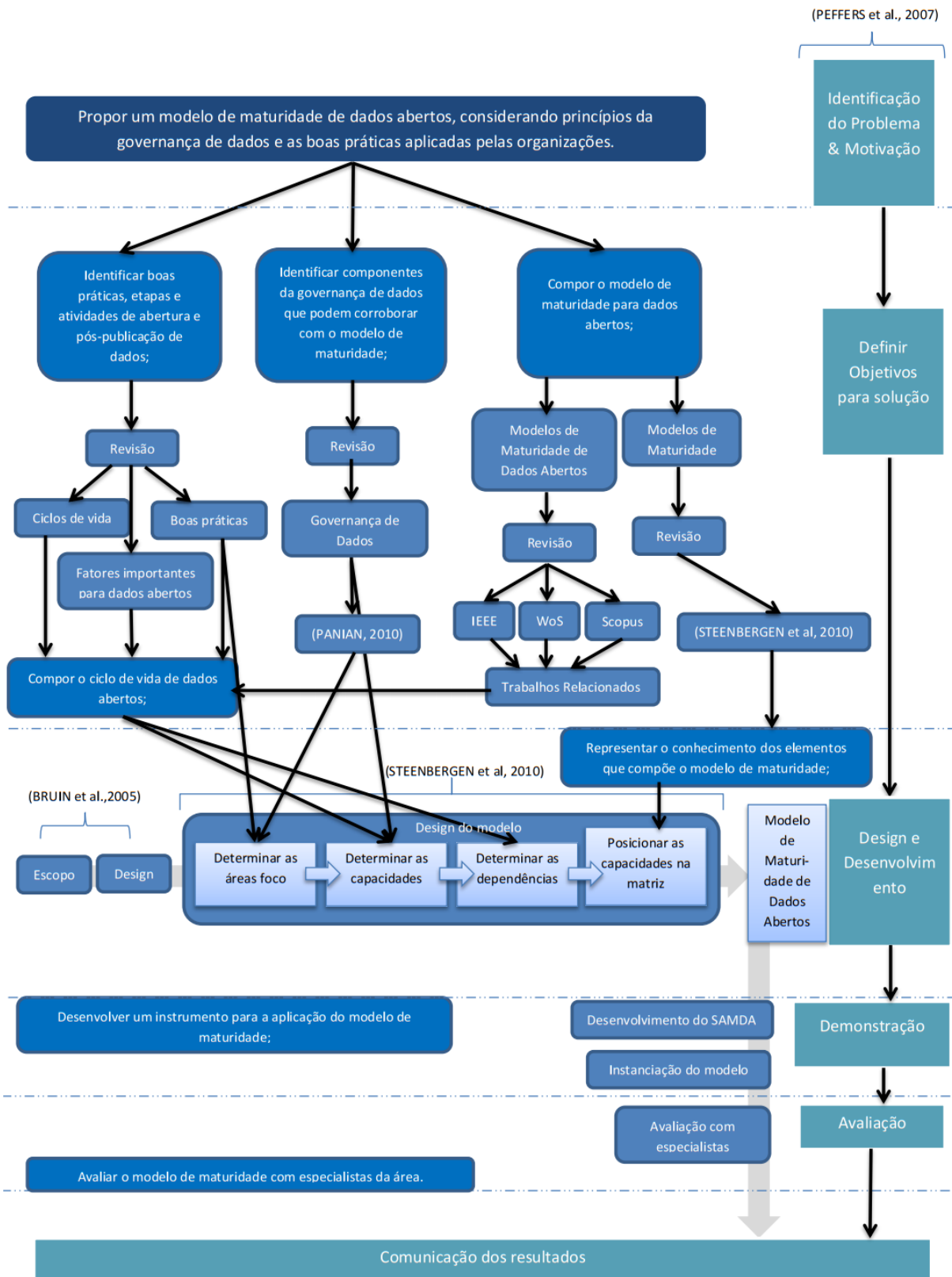


Fonte: (PEFFERS *et al.*, 2007)

Como o propósito deste trabalho é contribuir para a melhoria dos dados que são disponibilizados e que fazem uso de uma licença aberta, fez-se o uso de uma abordagem centrada no problema. Sendo esta definida por Peffers *et al.* (2007), como: "Os pesquisadores podem proceder nessa sequência se a ideia da pesquisa resultar da observação do problema ou de pesquisas futuras sugeridas em um trabalho de um projeto anterior".

Para representar cada etapa abordada na realização desta pesquisa a Figura 18 apresenta uma visão geral fazendo uso da metodologia Peffers *et al.* (2007). Na sequência estas etapas serão discutidas em detalhes.

Figura 18 – Etapas da Pesquisa



Fonte: Da Autora

3.1 ETAPAS DA PESQUISA

3.1.1 Etapa 1: Identificação do problema & Motivação

Para a elaboração deste trabalho a ideia emergiu de uma discussão no grupo de pesquisa sobre o porquê dados abertos não são amplamente consumidos. Consequentemente foram encontrados e sumarizados os problemas que foram abordados na seção 1.1, muitos destes problemas podem estar relacionados a inconsistências na abertura dos dados e até mesmo por falta de verificação se os dados disponibilizados estão de acordo com o que a organização almeja, e com este intuito observa-se a necessidade de realizar a avaliação de como as boas práticas acerca de dados abertos estão sendo empregadas na abertura do dados e até mesmo na pós-publicação. Para isso pensou-se em um modelo de maturidade, sendo que a motivação que guia a construção deste modelo é a possibilidade de se ter dados disponibilizados completos e acessíveis e possivelmente de melhor qualidade.

3.1.2 Etapa 2: Definir objetivos para uma solução

Antes de iniciar a elaboração da solução foi necessário compreender os conceitos acerca desta pesquisa, sendo estes: dados abertos, governança de dados, modelos de maturidade e modelos de maturidade de dados abertos.

- a) **Dados Abertos** No contexto de dados abertos foi realizada uma revisão de literatura através do uso do método de amostragem bola de neve (*snowball*), devido aos principais conceitos não estarem somente em artigos acadêmicos, mas também na literatura cinza (*Gray literature*) e a partir deste método chegou-se aos seguintes trabalhos:

Quadro 13 – Trabalhos sobre boas práticas, ciclos de vida e fatores importantes para dados abertos

Autor / Ano	Título
Hyland e Wood (2011)	<i>The joy of data-a cookbook for publishing linked government data on the Web.</i>
Villazón-Terrazas <i>et al.</i> (2011)	<i>Methodological guidelines for publishing government linked data.</i>
Hausenblas (2011)	<i>Linked data lifecycles.</i>
Auer <i>et al.</i> (2011)	<i>Introduction to linked data and its lifecycle on the Web.</i>
Janssen e Zuiderwijk (2012)	<i>Open data and transformational government.</i>
Scharffe <i>et al.</i> (2012)	<i>Enabling linked-data publication with the datalift platform.</i>

Broek <i>et al.</i> (2014)	<i>Walking the extra byte: A lifecycle model for linked open data.</i>
Lóscio <i>et al.</i> (2017)	<i>Data on the Web best practices.</i>
Hyland <i>et al.</i> (2014)	<i>Best practices for publishing linked data.</i>
Susha <i>et al.</i> (2015a)	<i>Critical factors for open data publication and use.</i>

Fonte: Da Autora

Outros trabalhos foram encontrados, mas destacam-se os trabalhos apresentados no Quadro 13 pois foram observadas características nestes estudos que contribuem significativamente com iniciativas de dados abertos.

Com base nos trabalhos apresentados no Quadro 13 foi estruturado o ciclo de vida que será apresentado na seção 4 com o intuito de delimitar o domínio funcional de maneira a embasar a estruturação do modelo de maturidade.

b) Governança de Dados

Para governança de dados também foi realizada uma revisão de literatura através do uso do método de amostragem bola de neve (*snowball*), devido aos mesmos fatores apresentados para dados abertos e a partir deste método chegou-se aos seguintes trabalhos:

Quadro 14 – Trabalhos sobre governança de dados

Autor / Ano	Título
Mosley <i>et al.</i> (2009)	<i>DAMA guide to the data management body of knowledge.</i>
DAMA International (2017)	<i>DAMA-DMBOK: Data Management Body of Knowledge (2nd Edition).</i>
Khatri e Brown (2010)	<i>Designing data governance.</i>
Chapple (2013)	<i>Speaking the same language: Building a data governance program for institutional impact.</i>
Dataflux Corporation (2007)	<i>The Data Governance Maturity Model Establishing the People, Policies and Technology That Manage Enterprise Data.</i>
Waqar Haider e Abrar Haider (2013)	<i>Governance structures for engineering and infrastructure asset management.</i>
Panian (2010)	<i>Some practical experiences in data governance.</i>

Fonte: Da Autora

Outros trabalhos foram encontrados, mas os trabalhos apresentados no Quadro 13 são destacados, pois estes foram selecionados para serem analisados com maior profundidade, para a escolha dos elementos que contribuiriam com a estruturação deste trabalho de tese, destes foi selecionado o trabalho de Panian (2010), conforme previamente já mencionado na seção 2.2.1.

c) Modelos de Maturidade

Para compreender o universo acerca de modelos de maturidade fez-se um estudo da origem até os modelos de maturidade aplicados à dados abertos, estes já abordados. Neste sentido, a seção 2.3 foi estruturada e apresenta as principais referências que embasaram a revisão de literatura, bem como a estruturação do modelo de maturidade.

Optou-se pelo uso de modelos de maturidade de áreas foco, devido a estes suportarem um número maior de níveis de maturidade, pois isto garante maior flexibilidade principalmente quando têm-se um domínio relativamente novo ou pouco explorado. Que é o caso dos modelos de maturidade de dados abertos. Sendo assim, os principais trabalhos que guiam a estruturação do modelo de maturidade são apresentados no Quadro 15.

Quadro 15 – Trabalhos sobre modelos de maturidade de áreas foco

Autor / Ano	Título
Van Steenbergem <i>et al.</i> (2008)	<i>A balanced approach to developing the enterprise architecture practice.</i>
Van Steenbergem <i>et al.</i> (2010a)	<i>The design of focus area maturity models.</i>
Van Steenbergem <i>et al.</i> (2013)	<i>Improving is functions step by step: the use of focus area maturity models.</i>
Van Steenbergem <i>et al.</i> (2010b)	<i>The dynamic architecture maturity matrix: Instrument analysis and refinement.</i>

Fonte: Da Autora

d) Modelos de Maturidade de Dados Abertos

Para analisar e discutir possíveis estudos que tenham sido realizados com o intuito de propor modelos de maturidade para dados abertos ou até mesmo realizar a avaliação da maturidade foram realizadas buscas nas bases de dados no mês de outubro de 2018 e mês de julho de 2021, onde foram utilizadas as seguintes bases: *Web of Science* (WoS), IEEE e Scopus, onde buscou-se as seguintes sintaxes "Open Data" AND "Maturity" e "Open Data" AND "benchmark".

Para as duas buscas realizadas foram considerados livros e capítulos de livros, assim como artigos de *journals* e de conferências.

As buscas foram realizadas com intuito de obter estudos similares ao que é proposto por esta pesquisa, no sentido de obter contribuições que auxiliem na estruturação do modelo de maturidade para dados abertos. Vale destacar que a busca realizada no ano de 2021 tem por objetivo abordar trabalhos incluídos na literatura após a banca de qualificação desta tese, ou seja, visa verificar e apresentar a continuidade das pesquisas da área. Os resultados quantitativos destas buscas podem ser observados nas Tabelas 1 e 2.

Considerando os critérios apresentados, em 2018 foram encontrados 223 arquivos, conforme pode ser confirmado através da soma de todos os documentos resultantes da duas buscas nas três bases de dados na Tabela 1. Com o auxílio do software EndNote foram removidos os arquivos duplicados restando um total de 142 arquivos, ainda através do EndNote fez-se a busca dos arquivos em formato PDF e com isso obteve-se um total de 78 arquivos a serem analisados.

Tabela 1 – Quantidade de arquivos resultantes para cada sintaxe de busca no ano de 2018

Sintaxe de busca	Scopus	WoS	IEEE	Duplicados	Disponíveis
"Open Data"AND "Maturity"	46	26	7	81	78
"Open Data"AND "Benchmark"	86	43	15		

Fonte: Da Autora

Em 2021, considerando os mesmos critérios, foram encontrados 470 arquivos, conforme pode ser confirmado através da soma de todos os documentos resultantes da duas buscas nas três bases de dados na Tabela 2. Com o auxílio do software EndNote foram removidos os arquivos duplicados restando um total de 320 arquivos, ainda através do EndNote fez-se a busca dos arquivos em formato PDF e com isso obteve-se um total de 105 arquivos para serem analisados.

Tabela 2 – Quantidade de arquivos resultantes para cada sintaxe de busca no ano de 2021

Sintaxe de busca	Scopus	WoS	IEEE	Duplicados	Disponíveis
"Open Data"AND "Maturity"	79	49	12	150	105
"Open Data"AND "Benchmark"	180	125	25		

Fonte: Da Autora

Em 2018 fez-se a análise do título e do resumo dos 78 artigos encontrados, e chegou-se a 15 artigos selecionados para a leitura na íntegra, conforme elencados no Quadro 16. Já em 2021, dos 105 artigos foram removidos os 78 que correspondiam aos artigos analisados em 2018 e assim foi realizada a análise do resumo e do título dos 27 artigos restantes. Com isso, foram adicionados 2 artigos ao Quadro de selecionados para a leitura na íntegra, ver Quadro 16.

Quadro 16 – Trabalhos resultantes após a análise do título e do resumo

Título	Referência
<i>A model to assess open government data in public agencies</i>	(SOLAR <i>et al.</i> , 2012)
<i>A Guide to Implement Open Data in Public Agencies</i>	(SOLAR <i>et al.</i> , 2013)
<i>A Framework for Benchmarking Open Government Data Efforts</i>	(SAYOGO <i>et al.</i> , 2014)
<i>Automatic Generation of Roadmaps for Open Data</i>	(SOLAR <i>et al.</i> , 2014a)
<i>A Model to Guide the Open Government Data Implementation in Public Agencies</i>	(SOLAR <i>et al.</i> , 2014b)
<i>Benchmarking open government: An open data perspective</i>	(VELJKOVIĆ <i>et al.</i> , 2014)
<i>Quality and maturity model for open data portals</i>	(OVIEDO <i>et al.</i> , 2015)
<i>Benchmarks for Evaluating the Progress of Open Data Adoption: Usage, Limitations, and Lessons Learned</i>	(SUSHA <i>et al.</i> , 2015b)
<i>A proposed checklist for the technical maturity of open government data: An application on GCC countries</i>	(ALROMAIH <i>et al.</i> , 2016)
<i>How to assess the success of the open data ecosystem?</i>	(WELLE DONKER; LOENEN, 2017)
<i>Standards and best practices related to the publication, exchange and usage of open data</i>	(ZENG <i>et al.</i> , 2017)
<i>Coretrustseal</i>	(DILLO; DE LEEUW, 2018)
<i>Development paths towards open government – an empirical analysis among heritage institutions</i>	(ESTERMANN, 2018)

<i>Culture of openness: The beginning of open government in municipalities in Germany</i>	(SIEBENLIST; MAINKA, 2018)
<i>Metrics for Open Government Data</i>	(SILVA; PINHEIRO, 2018)
<i>Maturity and usability of open data in North Rhine-Westphalia</i>	(AKYÜREK <i>et al.</i> , 2018)
<i>Assessing the Open Data Maturity of Government Agencies Having Different Years of Implementation Experience: A Study in Taiwan</i>	(YANG; WU, Y.-J., 2020)

Fonte: Da Autora

Considerando os trabalhos selecionados, o trabalho de Zeng *et al.* (2017) aborda a organização de um painel acerca de normas, padrões e boas práticas internacionais para lidar com dados abertos. Já o trabalho de Dillo e De Leeuw (2018) discute a criação de um repositório digital com uma certificação aplicada sobre os dados, visando dar maior confiabilidade para os dados que ali estão sendo disponibilizados, ou seja estes trabalhos possuem pouca, ou quase nenhuma relevância em relação a contribuições para o modelo de maturidade proposto.

Também foi possível verificar dentre estes artigos estruturas para avaliar, ou apenas focados na avaliação de portais de dados abertos (SAYOGO *et al.*, 2014; OVIEDO *et al.*, 2015; ALROMAIH *et al.*, 2016; AKYÜREK *et al.*, 2018), e neste sentido estes trabalhos abordam e discutem seus resultados e contribuições.

Destaca-se em meio a estes estudos um único trabalho que visa a avaliação de ecossistemas de dados abertos (WELLE DONKER; LOENEN, 2017), e para isso considera o fornecimento de dados, a governança de dados e as características dos usuários de dados de forma holística, com o intuito de caracterizar a maturidade do ecossistema de dados abertos.

Ainda sobre os trabalhos encontrados na literatura, Veljković *et al.* (2014), Estermann (2018) e Siebenlist e Mainka (2018) buscam discutir, elaborar e avaliar governo aberto ou iniciativas de governo aberto, sobre diferentes perspectivas, sendo estas através da maturidade ou de *benchmarks*. Assim como o trabalho de Yang e Yi-Jung Wu (2020) faz uso do trabalho de Solar *et al.* (2014a) para realizar a avaliação de iniciativas do governo.

Dentre os trabalhos encontrados, dois destes se destacam, pois abordam revisões de literatura, sendo estes Susha *et al.* (2015b) e Silva e Pinheiro (2018). O trabalho de Susha *et al.* (2015b) apresenta e discute uma meta-análise interpretativa acerca de cinco *benchmarks* de dados abertos e quatro modelos maturidade.

Os autores abordam e diferenciam *benchmark* de modelo de maturidade, e o interesse por este estudo se dá pelos modelos de maturidade abordados.

No entanto analisando os modelos de maturidade apresentados por Susha *et al.* (2015b) e verificando a pergunta a qual os autores se preocuparam em responder, sendo esta "Como os *benchmarks* de dados abertos existentes diferem na medição da maturidade de dados e o que pode ser aprendido com isso?", observa-se que apenas o trabalho de (SOLAR *et al.*, 2012), já recuperado através das buscas realizadas, pode contribuir com o modelo de maturidade proposto. Os demais modelos de maturidade, foram propostos por Kalampokis *et al.* (2011), Gwanhoo Lee e Kwak (2012) e Sayogo *et al.* (2014), sendo que além de avaliarem a sofisticação dos processos de dados abertos e não a maturidade (SUSHA *et al.*, 2015b), observa-se uma perspectiva diferenciada do que é proposto por esta tese, como por exemplo, abordam a maturidade de governo aberto.

O segundo trabalho que aborda revisão de literatura é o trabalho de Silva e Píneiro (2018) que tem por objetivo evidenciar métricas e indicadores acerca de dados abertos governamentais discutindo outros quatro trabalhos, sendo estes: *Metric for Releasing Open Data*-(MELODA), *Open Data Maturity Model*, *OpenGovB* e a Norma UNE 178301. Dois destes trabalhos, sendo estes *Open Data Maturity Model* que foi proposto por Solar *et al.* (2012) e o *OpenGovB*, oriundo do trabalho de Veljković *et al.* (2014) já estão presentes na busca realizada nas bases de dados. O MELODA não será discutido neste trabalho pois é uma proposta específica que propõe métricas para avaliar o reuso de dados abertos com base na legislação espanhola. A Norma UNE 178301 foi discutida na seção 2.4.

Com o intuito de complementar as buscas realizadas nas bases de dados foi realizada uma busca na Web, com as mesmas sintaxes de busca das bases de dados e foram encontrados dois trabalhos, sendo estes: (ATZ *et al.*, 2015) e (DODDS; NEWMAN, 2015). Sendo que, o trabalho de (ATZ *et al.*, 2015) aborda um *benchmark* que visa analisar o fornecimento, o uso e o impacto de dados abertos. Já o trabalho de Atz *et al.* (2015) também foi selecionado para ser discutido.

Com a realização desta busca alguns *benchmarks* também foram encontrados, por exemplo: **Open Data Barometer**¹, que foi produzido pela Open data Institute e pelo World Wide Web Foundation, **Global Open Data Index**², produzido pela Open Knowledge Foundation, **Open Data Readiness Assessment**³, produzido pela Open Government Working Group do Banco Mundial, entre outros. No entanto estes trabalhos não serão discutidos por esta tese.

¹ <https://opendatabarometer.org/>

² <https://index.okfn.org/>

³ <http://opendatatoolkit.worldbank.org/en/odra.html>

Com esta análise prévia, foi possível selecionar os trabalhos apresentados no Quadro 17, para serem discutidos de maneira mais aprofundada, com o intuito de agregar contribuições no desenvolvimento desta tese.

Destaca-se ainda que os trabalhos selecionados não englobam os *benchmarks* analisados devido a não possuírem o mesmo objetivo que os modelos de maturidade, sendo este o foco da tese, dado que "*benchmark* refere-se ao uso de métricas de desempenho para avaliar e comparar diferentes organizações ou países" (SUSHA *et al.*, 2015b, pag 3). Os trabalhos apresentados no Quadro 17 foram abordados na seção 2.4.

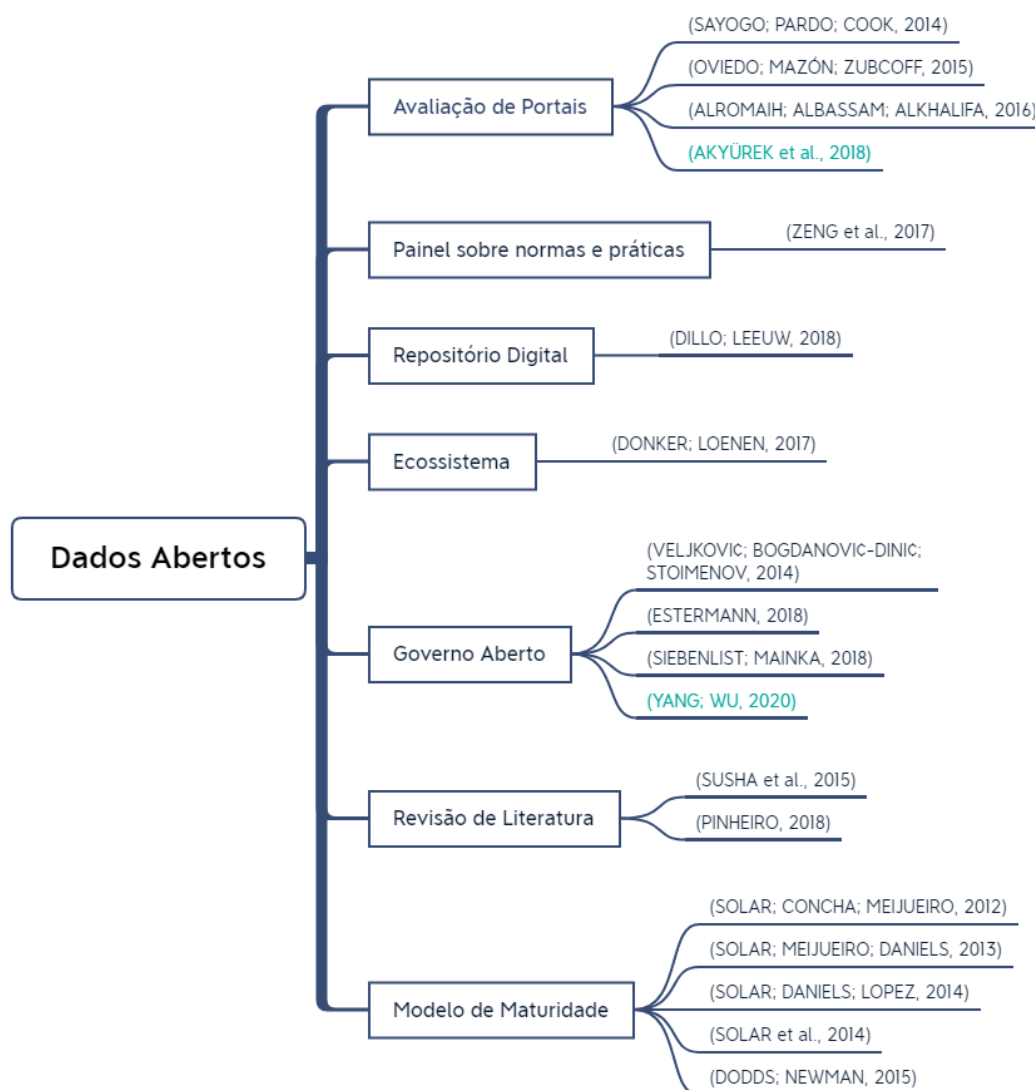
Quadro 17 – Trabalhos selecionados para discussão

Título	Referência
<i>A model to assess open government data in public agencies</i>	(SOLAR <i>et al.</i> , 2012)
<i>A Guide to Implement Open Data in Public Agencies</i>	(SOLAR <i>et al.</i> , 2013)
<i>Automatic Generation of Roadmaps for Open Data</i>	(SOLAR <i>et al.</i> , 2014a)
<i>A Model to Guide the Open Government Data Implementation in Public Agencies</i>	(SOLAR <i>et al.</i> , 2014b)
Guia do Modelo de Maturidade de Dados Abertos	(DODDS; NEWMAN, 2015)
Norma UNE 178301	Citado no trabalho de (SILVA; PINHEIRO, 2018)

Fonte: Da Autora

Com o objetivo de finalizar o raciocínio sobre os trabalhos encontrados através das buscas realizadas foi estruturado um mapa mental, que pode ser observado na Figura 19, que resume os trabalhos encontrados e abordados neste item. Destaca-se que os trabalhos na cor verde foram incluídos devido a busca realizada em julho de 2021.

Figura 19 – Resumo da análise dos trabalhos encontrados nas buscas realizadas



Fonte: Da Autora

3.1.3 Etapa 3: Design & Desenvolvimento

Os modelos de maturidade são numerosos e de aplicação diversificada. No entanto, a documentação sobre como desenvolver um modelo de maturidade que seja teoricamente sólido, rigorosamente testado e amplamente aceito é escassa (DE BRUIN *et al.*, 2005). Ainda, na literatura há revisões sobre a estruturação de modelos de maturidade e nota-se que estas observam diversos modelos criados sem um modelo reconhecível, além de outros sem método algum de construção (WENDLER, 2012).

Neste sentido, De Bruin *et al.* (2005) propõem um *framework* padrão para o desenvolvimento de modelos de maturidade, de qualquer domínio, podendo ser construído de modo descritivo, prescritivo ou comparativo. Para a estruturação do modelo de maturidade, optou-se pelo uso deste *framework*, em partes, sendo consideradas as

seguintes etapas: escopo e o design, conforme pode ser observado na Figura 20.

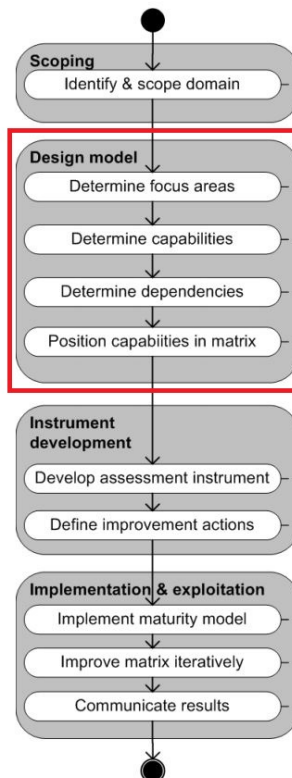
Figura 20 – Etapas da construção de um modelo de maturidade



Fonte: (DE BRUIN *et al.*, 2005)

As demais etapas do trabalho de De Bruin *et al.* (2005) não foram utilizadas na estruturação do modelo devido às particularidades a serem consideradas na elaboração de modelos de maturidade de áreas foco. Neste sentido, o desenvolvimento do trabalho se deu considerando a etapa "Design do modelo" e suas respectivas sub-etapas apresentadas no trabalho de Van Steenbergem *et al.* (2010a). Em destaque na Figura 21.

Figura 21 – Etapas da construção de um modelo de maturidade orientado a área foco



Fonte: (VAN STEENBERGEN *et al.*, 2010a)

Estas etapas são detalhadas e discutidas no capítulo 5. Optou-se pelo uso apenas da etapa "Design do Modelo", visto que as demais etapas propostas no trabalho

de Van Steenbergem *et al.* (2010a) também têm como base a DSR e por este motivo na continuidade consideram-se as etapas elencadas por Peffers *et al.* (2007).

3.1.4 Etapa 4: Demonstração

A etapa de demonstração do modelo se deu por meio do desenvolvimento de uma aplicação Web que viabiliza o acesso e a resolução do questionário para instanciar o modelo de maturidade e assim obter o perfil de maturidade correspondente. Através da aplicação Web foi possível coletar e analisar três cenários de uso reais, estes descritos na seção 6.

Na seção 6 é possível observar como se deu as entrevistas, o perfil de maturidade gerado para cada organização, algumas percepções observadas e os *feedbacks* oriundos dos entrevistados.

3.1.5 Etapa 5: Avaliação

Para realizar a avaliação do modelo de maturidade após a coleta dos três cenários de uso foi elaborado um formulário, para obter as percepções e contribuições dos entrevistados em cada cenário de uso referente a: **I)** estrutura do modelo de maturidade, considerando aspectos de compreensibilidade, facilidade de uso e utilidade e praticidade, ver Apêndice A; **II)** a importância dos elementos que compõem a estrutura; e sobre **III)** a relevância e abrangência do questionário aplicado na entrevista, ver Apêndice B. Os detalhes sobre os dados obtidos são discutidos na seção 6.1.2.2.

Ainda, para realizar a avaliação da estrutura do modelo de maturidade de dados abertos, foram realizadas entrevistas com especialistas da área. As entrevistas foram realizadas durante o mês de julho e agosto de 2021 de maneira *on-line*, sendo entrevistado um total de 10 profissionais da área de dados abertos. Para isso, foi apresentada e discutida a estrutura do modelo de maturidade e após foi solicitado aos entrevistados que respondessem um questionário, este nos apêndices C, D e E. Os detalhes obtidos através da entrevistas realizadas são abordados na seção 6.2.

3.1.6 Etapa 6: Comunicação dos resultados

A comunicação dos resultados é destinada à comunidade científica e sociedade civil com a apresentação dos resultados da pesquisa. Para isso, serão escritos e enviados artigos à periódicos e congressos, nacionais e internacionais relacionados a área de estudo. Neste sentido, já foi realizada a submissão de um artigo para o 5º Workshop de Teses e Dissertações em Ontologias (WTDO 2021).

Também foi realizado o registro da aplicação Web desenvolvida, para a coleta dos cenários de uso⁴. Sendo apresentado na página inicial da aplicação Web as

⁴ <https://lidianevisintin.studiokem.ufsc.br/>

informações relacionadas a pesquisa.

Destaca-se que a realização das entrevistas com os profissionais da área de dados abertos também possibilitou a divulgação e um momento de interação para coleta de *feedbacks* sobre a pesquisa e o artefato apresentado.

4 CICLO DE VIDA DE DADOS ABERTOS

O ciclo de vida visa organizar as orientações que conceitualizam as práticas e tratamentos de dados, desde sua geração até a pós-publicação. Neste sentido, apresenta-se a o ciclo de vida de dados abertos, sendo considerado os princípios da governança de dados e também as boas práticas para dados abertos. Este alinhamento se fez necessário para que fosse possível delimitar o domínio do modelo de maturidade a ser caracterizado.

Os dados assim como outros ativos organizacionais, têm um ciclo de vida (DAMA INTERNATIONAL, 2017). E conseqüentemente, dados abertos também, visto que:

- Eles serão gerados e disponibilizados a partir da percepção de uma necessidade;
- Transformados e entregues aos usuários;
- Poderão ser necessárias atividades de melhoria e manutenção;
- E removidos ou excluídos, ao final da sua vida útil.

Quando as organizações iniciam sua estratégia de abertura de dados, optam pela abertura em pequena escala, ou seja equipes pequenas e/ou até mesmo esforços individuais garantem a disponibilização dos dados de maneira eficiente. No entanto, quando se tem uma certa complexidade atrelada aos dados, na abertura (por exemplo, dados abertos conectados) ou até mesmo um aumento no volume de conjuntos de dados a serem abertos, faz-se necessário organizar as atividades e prever as próximas, para que estas sejam concluídas, medidas e até mesmo aprimoradas.

A literatura define que o objetivo do ciclo de vida é descrever o desenvolvimento de certos fenômenos e prever os próximos passos (BROEK *et al.*, 2014 apud LANE; RICHARDSON, 2011). Nessa perspectiva, introduz-se um ciclo de vida que descreve as atividades inerentes a dados abertos. No intuito de contribuir com as organizações para que estas conheçam, compreendam, planejem e melhorem seus processos, para que possam alcançar as estratégias traçadas acerca de dados abertos.

Uma estratégia definida sobre dados abertos não deverá apenas considerar os requisitos de conteúdo de dados, mas também seus requisitos de gerenciamento e planejamento de dados e organizacionais. Isso inclui políticas e estratégias com relação à criação, uso, qualidade, segurança, sustentabilidade, infraestrutura, treinamento entre outros aspectos. Vale ressaltar ainda que quando pensa-se em um ciclo de vida de dados abertos, considerando apenas a publicação ou a abertura dos dados é similar a considerar somente a programação no âmbito de um ciclo de vida de desenvolvimento de software, ou seja é eficiente para softwares de baixa complexidade. E neste sentido, observa-se a necessidade de se ter uma estrutura mais abrangente para o contexto de dados abertos.

Neste contexto, observam-se lacunas de pesquisa no âmbito de ciclos de vida de dados abertos, como já foi apresentado na seção 2.1.4, pois nota-se que a maioria dos estudos referentes a dados abertos focam em ciclos de vida de publicação de dados e não no ciclo de vida de dados abertos, ou seja, há a falta de uma estrutura holística. Neste cenário, nota-se ainda que a abertura de dados tem uma série de implicações importantes, e que algumas destas não tem grande ênfase nas discussões apresentadas junto a publicação de dados, devido a sua natureza, sendo exemplos destas:

- **Legislação, regulamentações e licenças**

Quando considerados **fatores externos** o primeiro item que se destaca são leis, regulamentações e licenças que regem ou que são especificadas por um país ou por uma determinada organização (KUCERA *et al.*, 2015). Dado o contexto, observa-se que dados abertos foram adotados como uma estratégia em busca de transparência por diversos países e que muitos destes propuseram diferentes leis, normativas e licenças para conjuntos de dados, para este último caso pode-se citar como exemplo o Reino Unido (NEUMAIER, 2015). E nesta perspectiva, se faz necessário considerar também que os países adotaram diferentes formas de auditar e aplicar punições, sejam estas brandas ou severas, a aqueles que não possuem conformidade em seus dados disponibilizados. Deste modo, este item é destacado como uma implicação importante, pois pode impactar diretamente e significativamente na abertura de dados, dada pela necessidade da conformidade dos processos de abertura e dos dados em si, seguindo legislações, regulamentações e licenças vigentes.

- **Colaboração e uso**

Quando considerada a ideia de uso, a literatura discute maneiras de interagir com a comunidade com o intuito de promover os dados que estão sendo disponibilizados, para que estes gerem valor, ou seja, para que com isso sejam gerados novos negócios, inovações, entre outros (EUROPEAN DATA PORTAL, 2018). No entanto, quando destacado o item uso, pouco se aborda sobre educação no contexto de dados abertos, bem como o fomento a pesquisas, sendo que esta abordagem poderia contribuir no sentido de diagnosticar e resolver questão sobre dados abertos. Neste sentido, observa-se que a colaboração pode trazer benefícios, principalmente para as organizações que estão iniciando na abertura de seus dados, para que estas consigam apreender com organizações que já estão à mais tempo disponibilizando seus dados, ou seja, compartilhando práticas, princípios e experiências com o intuito de tornar seus processos mais eficientes e conseqüentemente evoluir mais rapidamente sobre como abrir e manter seus dados, de maneira adequada. Assim, observa-se que a colaboração e o uso são

viabilizados por **fatores externos** a organização e deste modo, a organização necessitará relacionar-se com comunidades de dados abertos, com outras organizações e com a própria academia com o intuito de ampliar a colaboração para fomentar o uso dos dados disponibilizados, compreender e aprimorar suas atividades relacionadas a dados abertos.

- **Sustentabilidade da iniciativa de dados abertos**

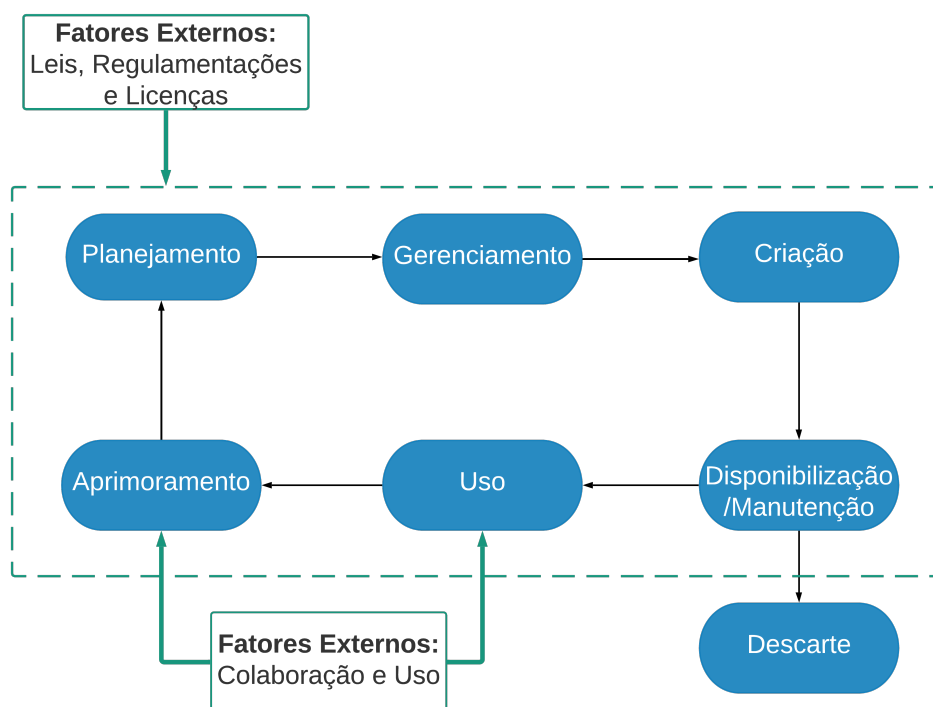
Outro item que vem ganhando destaque em pesquisas recentes é a sustentabilidade da iniciativa de dados abertos. Esta é outra implicação importante, visto que dados precisam ser relevantes, atuais e acessíveis para que possam ser utilizados, e sendo que há um custo para se produzir e disponibilizar estes dados (DAMA INTERNATIONAL, 2017). Contudo, a sustentabilidade da iniciativa não se dá somente pela quantidade e pela qualidade dos dados que são disponibilizados, mas pela capacidade da organização dar continuidade a disponibilização de dados atualizados regularmente com pelo menos o mesmo nível de qualidade e quantidade (GLOBAL PARTNERSHIP FOR SUSTAINABLE DEVELOPMENT DATA, 2016). Neste sentido, destaca-se a importância da sustentabilidade da iniciativa de dados abertos, sendo que este item pode contribuir com o consumo dos dados, visto que negócios criados a partir dos dados poderão depender do fornecimento destes, ou seja, consumir os dados disponibilizados ao longo do tempo.

Apresentadas as discussões iniciais, uma primeira proposta de ciclo de vida foi elaborada, conseqüentemente esta proposta foi adaptada para a natureza de dados abertos, no entanto a organização das atividades-chave que será discutida seguem os mesmos princípios das atividades-chave elencadas para compor o ciclo de vida de dados apresentado por DAMA International (2017).

Para compreender e construir o ciclo de vida de dados abertos, partiu-se da similaridade com o ciclo de vida de um produto, pois observa-se que se faz necessário ter uma etapa de criação, sendo que os dados podem ser transformados e armazenados, para que também possam ser mantidos, compartilhados, utilizados, bem como podem ser descartados, e por este motivo observa-se uma similaridade com o ciclo de vida de produto e não com o ciclo de vida de um software.

Ainda, ao longo do ciclo de vida, os dados podem ser limpos, transformados, mesclados, aprimorados ou agregados (DAMA INTERNATIONAL, 2017). Conforme os dados são utilizados ou aprimorados, novos dados poderão ser criados, similar a novos produtos. E neste sentido, poderá haver interações internas entre as atividades-chave definidas no ciclo de vida de dados abertos, no entanto estas interações não serão abordadas por este trabalho. Somente serão discutidas as atividades-chave apresentadas na Figura 22.

Figura 22 – Ciclo de vida de dados abertos



Fonte: Da Autora

Deste modo observa-se que o ciclo de vida proposto aborda uma estrutura abrangente de atividades-chave organizadas e relacionadas a fatores, atividades e boas práticas que contribuem para que a abertura de dados seja realizada com eficácia. Ou seja, o intuito deste ciclo é fornecer uma visão macro de dados abertos para que as organizações possam fazer uso deste estudo, considerando quais destes itens contribuem com a sua realidade e deste modo consigam organizar o planejamento, as atividades e até mesmo as equipes com o intuito de atingir as metas almejadas com dados abertos.

Outro ponto de destaque é que faz-se o uso do termo fatores importantes, pela natureza do problema, visto que, temos na literatura boas práticas, e não um conjunto delimitado de tarefas que precisam obrigatoriamente serem executadas para obter dados abertos e por este motivo optou-se pelo termo.

Assim sendo, para explicar sobre as atividades e ou boas práticas encontradas na literatura que cada uma das macroatividades está relacionada, optou-se pelo uso da categorização de fatores importantes apresentada pelo trabalho de (SUSHA *et al.*, 2015a). Portanto, o ciclo de vida está organizado da seguinte forma: 1º nível define as atividades-chave ou macroatividades, no 2º nível a(s) categoria(s), no 3º nível os fatores importantes que caracterizam cada categoria e no 4º nível as atividades e/ou boas práticas apresentadas pela literatura que corroboram com o que é proposto por cada um dos fatores. Conforme hierarquia exemplificada a seguir:

1° nível - **Planejamento**;

2° nível - **Estratégia e Apoio**;

3° nível - n fatores importantes relacionados;

4° nível - n atividades e/ou boas práticas;

1° nível - **Gerenciamento**;

2° nível - **Suporte a gestão e aos processos de publicação**;

3° nível - n fatores importantes relacionados;

4° nível - n atividades e/ou boas práticas;

2° nível - **Treinamento e apoio aos funcionários**;

3° nível - n fatores importantes relacionados;

4° nível - n atividades e/ou boas práticas;

1° nível - ...;

E neste sentido, se dará a contextualização das atividades-chave e suas respectivas subdivisões.

4.1 ATIVIDADES-CHAVE DO CICLO DE VIDA DE DADOS ABERTOS

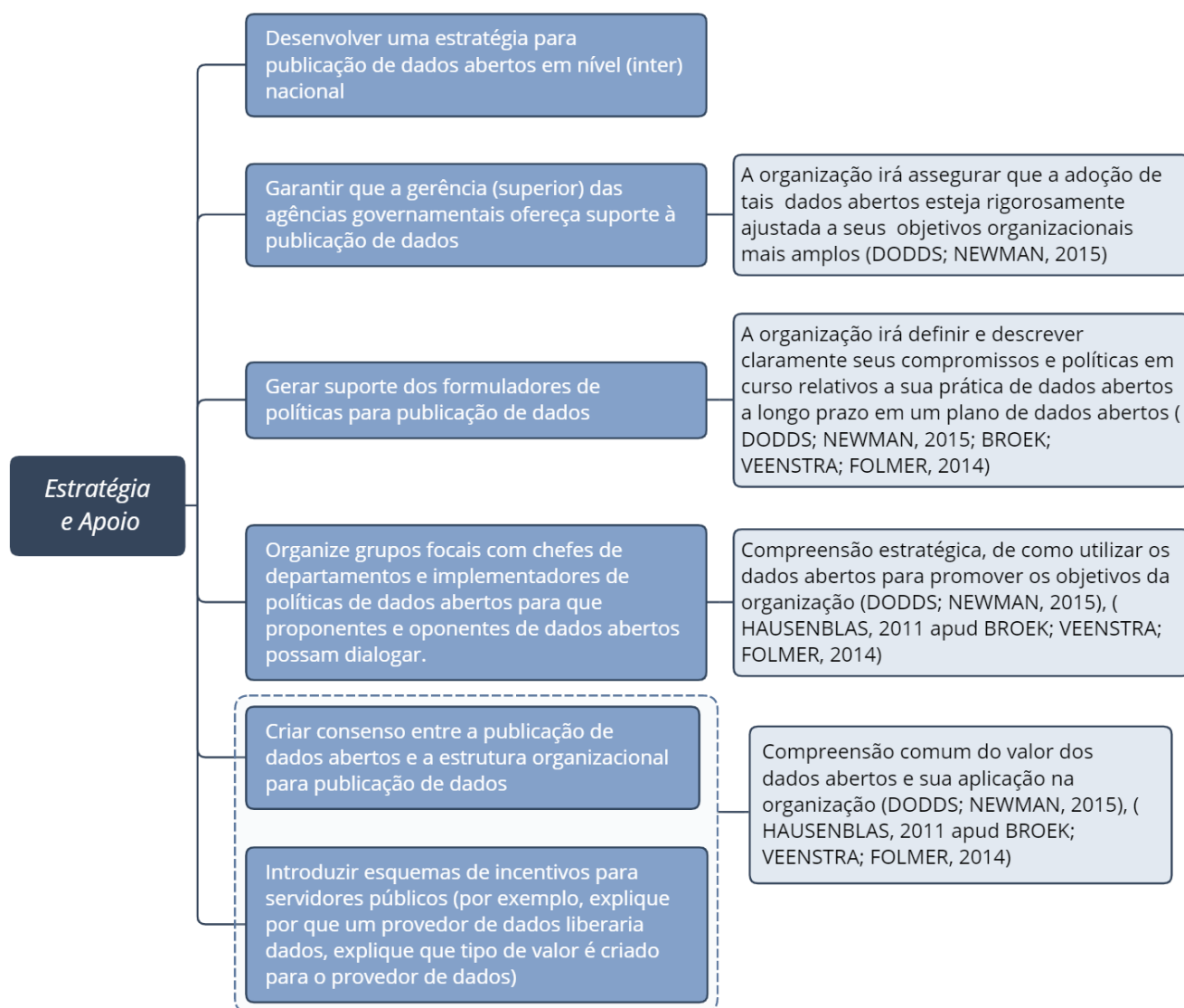
4.1.1 Planejamento

A primeira atividade-chave do ciclo de vida de dados abertos é definida como planejamento, ou seja, nesta macroatividade busca-se definir o plano de dados abertos, compreender e delimitar os objetivos da abertura dos dados, alinhar estes a estratégia organizacional e fornecer direcionamentos para que os envolvidos em todas as etapas do ciclo de vida consigam compreender o que se almeja com a abertura dos dados, e com isso consigam contribuir de maneira significativa com cada uma das atividades que compõem o ciclo de vida. A categoria que aborda os fatores importantes para dados abertos, desta atividade-chave é definida por:

(a) **Estratégia e Apoio**

Nesta categoria pode-se observar na Figura 23 em azul os fatores importantes destacados por (SUSHA *et al.*, 2015a), sendo que alguns destes foram relacionados com as boas práticas e/ou atividades, na Figura 23 destacadas em cinza, pois estas atividades corroboram com os respectivos fatores importantes.

Figura 23 – Estratégia e Apoio.



Fonte: Da Autora

De modo geral, podem ser observados que os fatores importantes, as boas práticas e/ou atividades apresentados na Figura 23 se remetem aos seguintes tópicos:

- Ter um plano de dados abertos;
- Alinhar os objetivos organizacionais com a estratégia de abertura de dados;
- Fornecer apoio àqueles que formulam os direcionamentos relacionados a abertura de dados;
- Ter a compreensão comum do valor de dados abertos para a organização.

Analisando a Figura 23, nota-se que o primeiro fator apresentado "Desenvolver uma estratégia para a publicação de dados abertos em nível (inter)nacional",

dependerá da estratégia adotada para a organização e também observa-se que não foi relacionado a nenhuma atividade e/ou melhor prática, devido a não ser encontrada nenhuma atividade e/ou melhor prática que corrobora com este fator importante, na literatura consultada. No entanto, este fator está sendo apresentado junto a esta atividade-chave, pois o intuito deste ciclo de vida é fornecer uma estrutura abrangente, para que a partir deste possam ser gerados *insights* e até mesmo outros ciclos de vida que contemplem a realidade de cada organização.

E neste sentido, as demais categorias que serão apresentadas poderão ter fatores importantes não relacionados a boas práticas e/ou atividades devido a não ter sido encontrado na literatura atividade e/ou melhor prática que corrobora com alguns destes fatores importantes abordados no ciclo de vida.

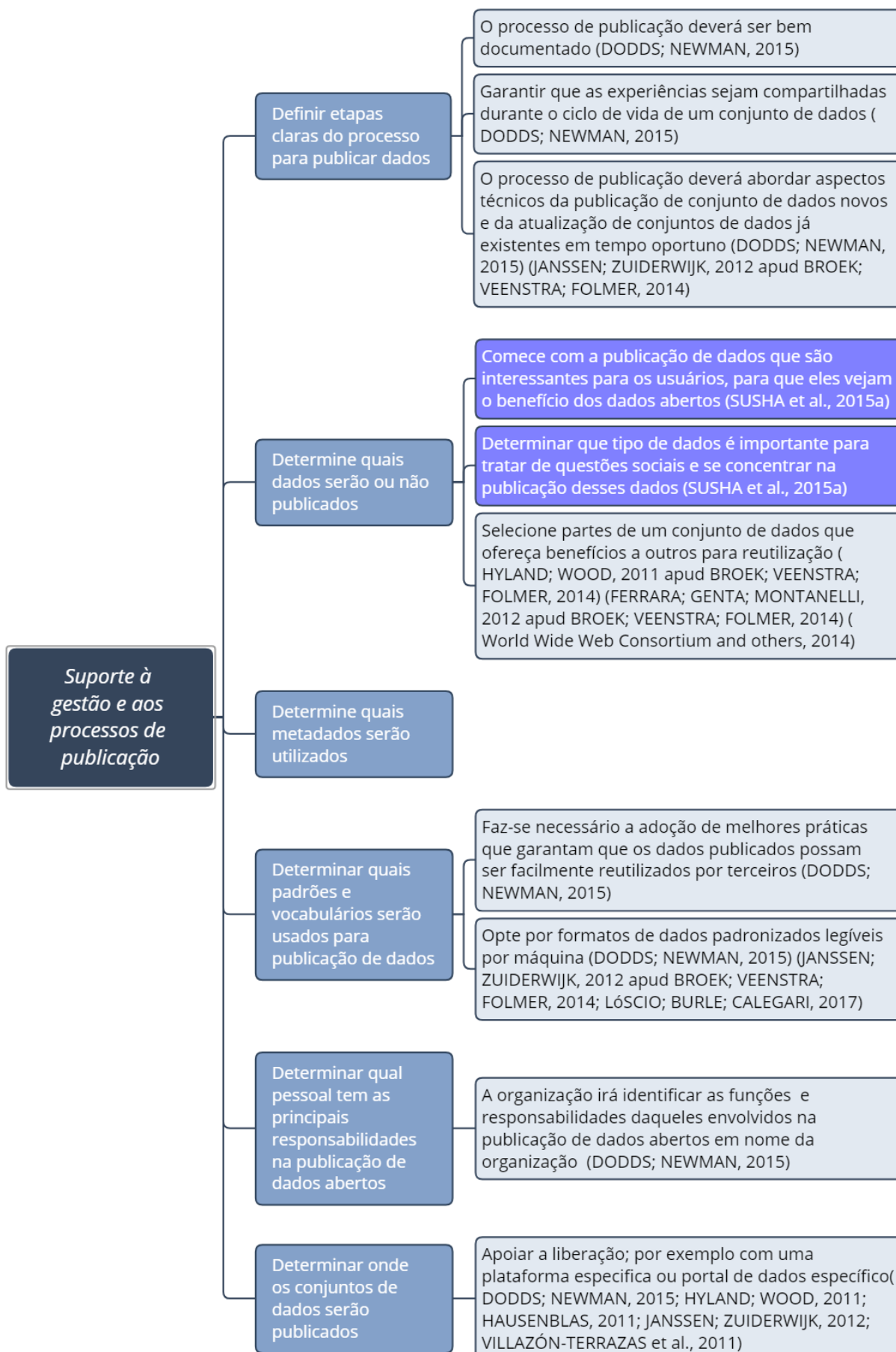
4.1.2 Gerenciamento

A segunda atividade-chave do ciclo de vida de dados abertos compreende atividades relacionadas a gerenciar, ou seja, tem por objetivo abordar atividades alinhadas com o planejamento já realizado, considerando os processos de abertura e de gestão de pessoas. Duas categorias abrangem os fatores importantes para dados abertos, sendo estas:

(a) ***Suporte à gestão e aos processos de publicação***

A primeira categoria utilizada que descreve os fatores importantes é descrita como "Suporte à gestão e aos processos de publicação". Na Figura 24 observam-se duas atividade destacadas em coloração diferenciada (na cor roxa), pois no trabalho de Susha *et al.* (2015a) são apresentadas como fatores importantes, no entanto percebe-se que estas duas atividade remetem a boas práticas para determinar quais dados serão ou não publicados, e por este motivo foram alinhadas ao nível das boas práticas e/ou atividades.

Figura 24 – Suporte a gestão e aos processos de publicação.



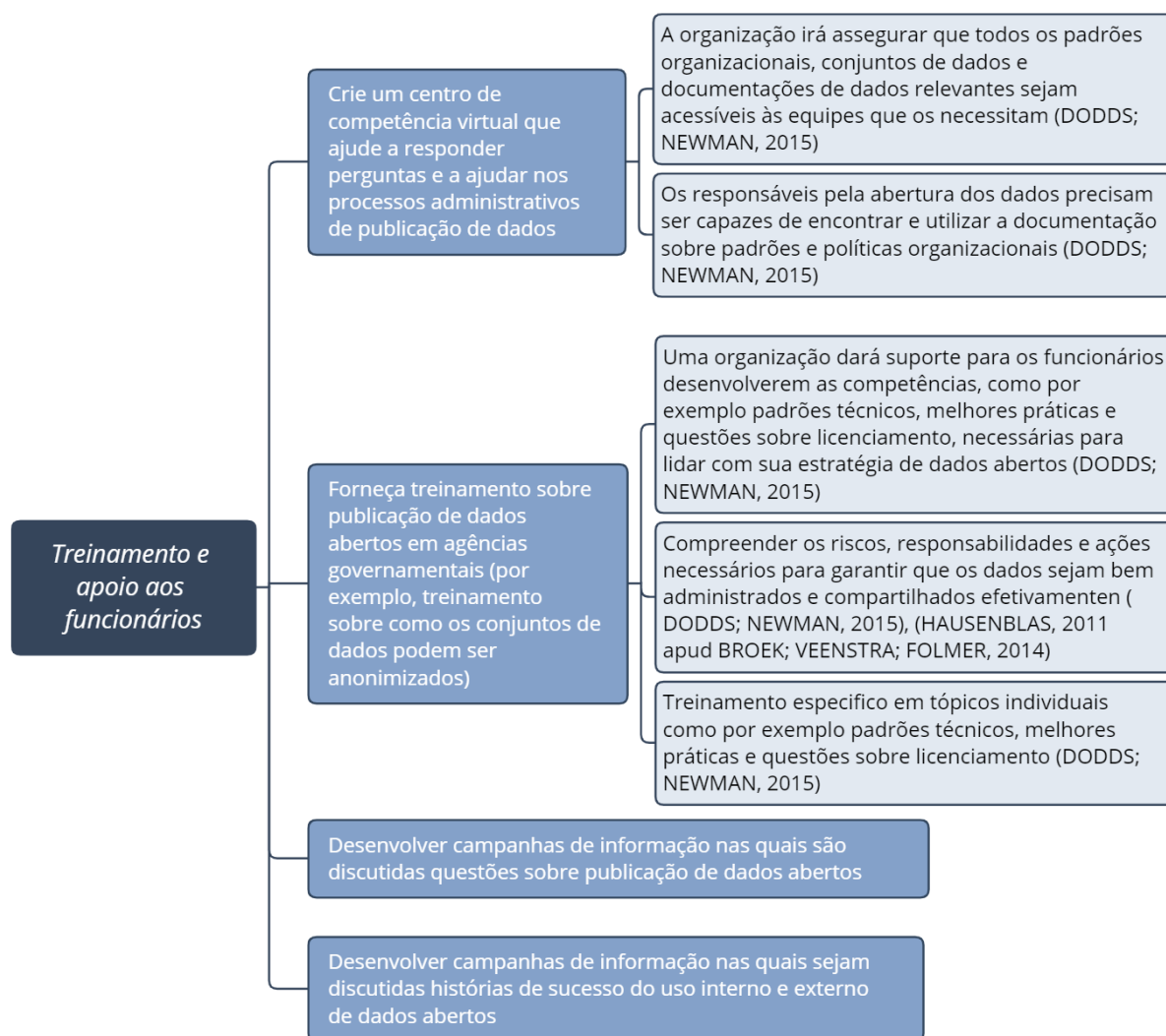
Nesta categoria observa-se ainda, que as boas práticas e os fatores importantes para a publicação de dados se remetem aos seguintes tópicos:

- Definir as funções e responsabilidades dos envolvidos na abertura de dados;
- Definir etapas claras do processo de abertura de dados;
- Definir quais conjuntos de dados ou partes de conjuntos de dados serão disponibilizados;
- Determinar quais padrões serão utilizados para dados e metadados;
- Determinar onde serão disponibilizados os dados;

Na sequência, apresenta-se a segunda categoria que esta inserida na atividade-chave gerenciamento, sendo esta: treinamento e apoio aos funcionários.

(b) **Treinamento e apoio aos funcionários**

Figura 25 – Treinamento e apoio aos funcionários.



Fonte: Da Autora

De modo geral, podem ser observadas que as boas práticas e os fatores importantes relacionados a esta categoria se remetem aos seguintes tópicos:

- Fornecer treinamento para os responsáveis pela abertura dos dados;
- Criar campanhas informativas.
- Criar uma base de conhecimento;

Observando o contexto apresentado, dado pelo alinhamento das boas práticas abordadas pela literatura com os fatores importantes, nota-se que esta atividade-chave pode contribuir com a capacitação dos envolvidos na criação, abertura e disponibilização dos dados, bem como viabilizar o gerenciamento dos processos de abertura de dados.

4.1.3 Criação

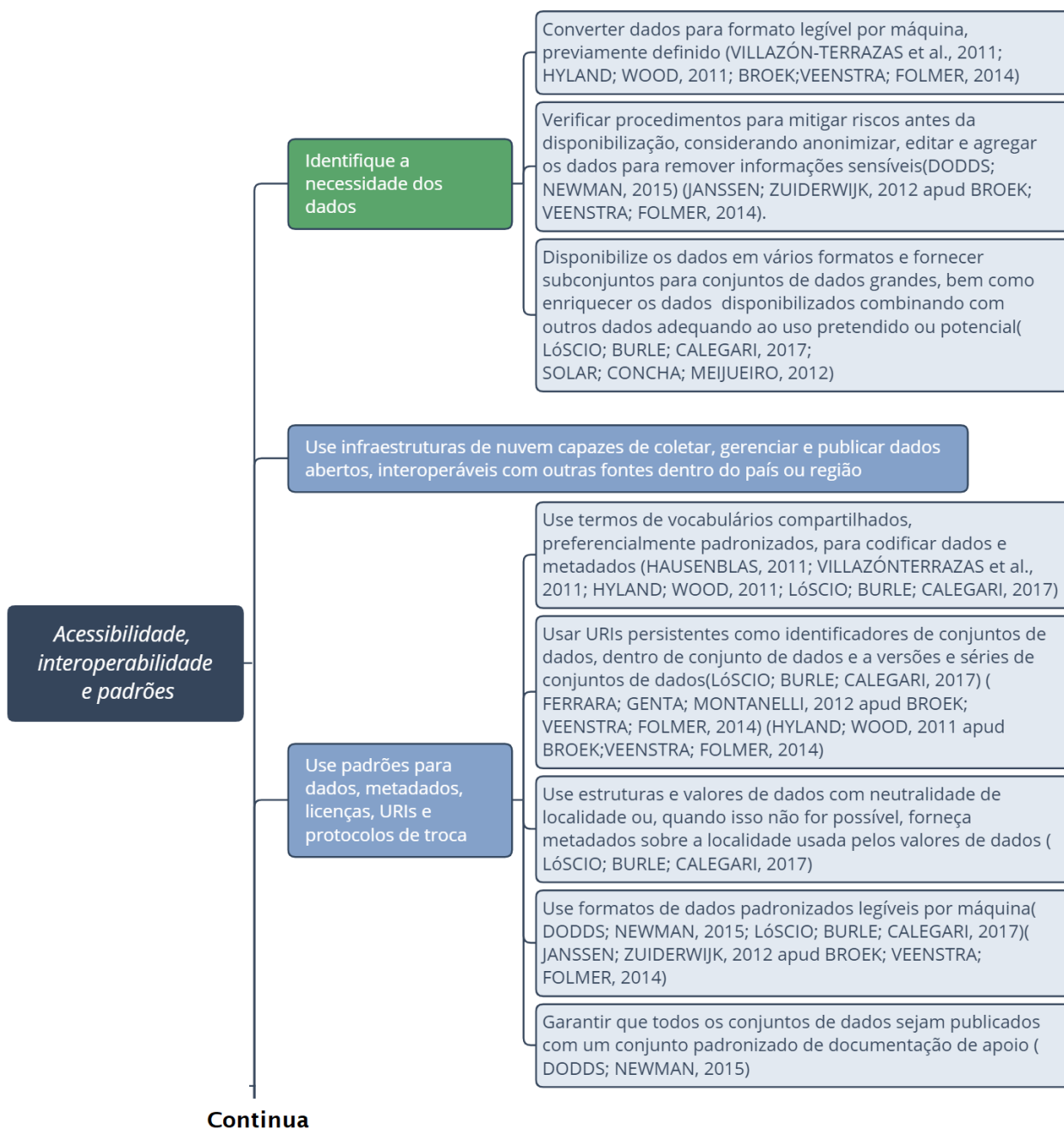
A terceira atividade-chave do ciclo de vida de dados abertos compreende atividades relacionadas à criação dos dados, ou seja, consideram-se aspectos técnicos e de padronização relacionados aos dados e às atividades de preparação e de disponibilização. Esta macroatividade contém uma única categoria de fatores importantes para dados abertos que é definida por:

(a) ***Acessibilidade, interoperabilidade e padrões***

Nesta categoria há uma grande quantidade de boas práticas relacionadas isto se deve à riqueza de estudos que discutem em nível detalhado, quais os padrões e como aplicá-los quando deseja-se abrir dados com o intuito de garantir interoperabilidade e acessibilidade, por exemplo, citam-se os seguintes trabalhos (LÓSCIO *et al.*, 2016; BROEK *et al.*, 2014; AUER *et al.*, 2011), entre outros já previamente citados.

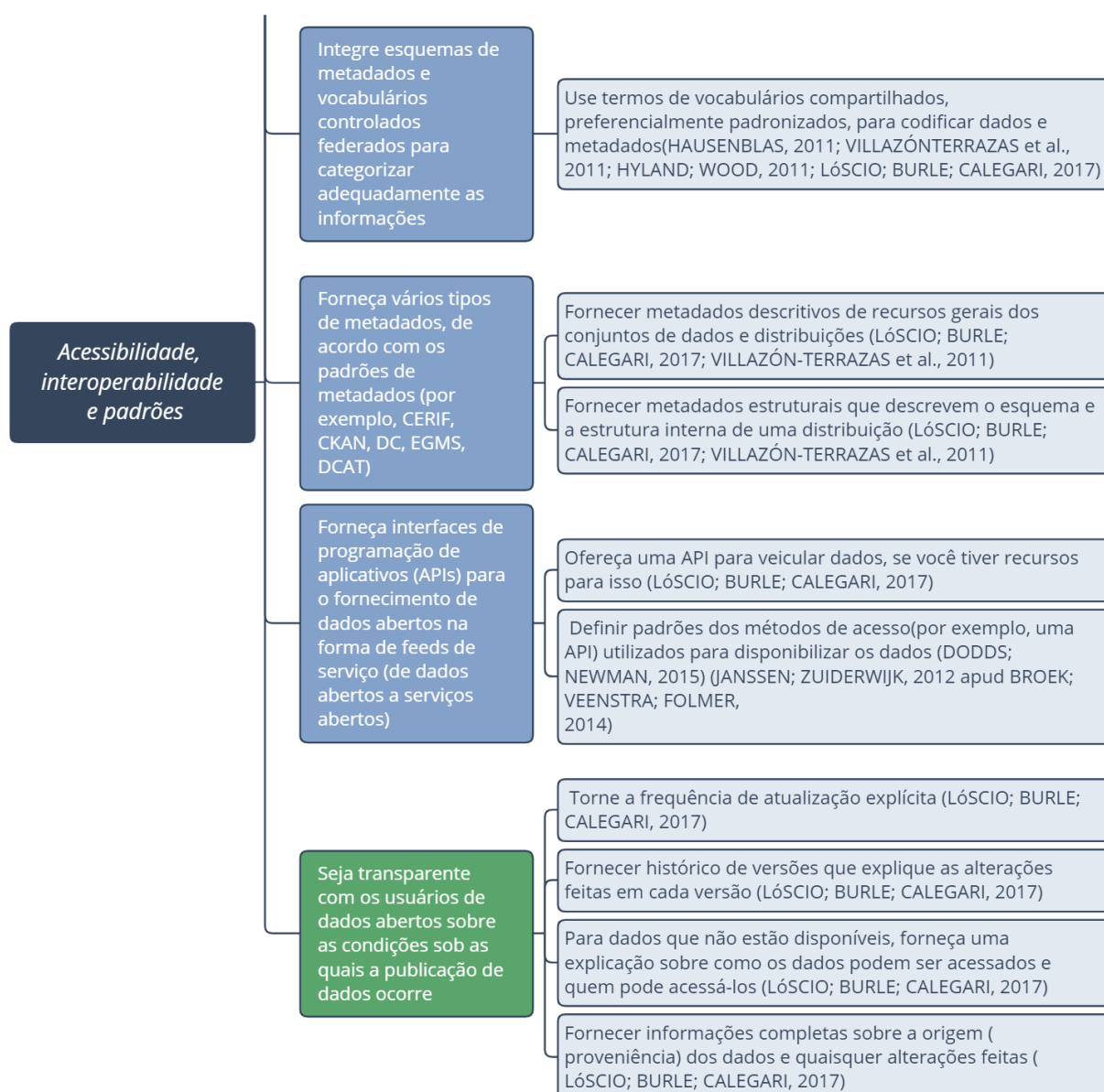
Outro ponto a ser destacado são os fatores importantes que estão na cor verde este fatores originalmente eram da categoria denominado de "Sustentabilidade da iniciativa de dados abertos" no trabalho de Sussha *et al.* (2015a). Contudo, conforme discutido inicialmente, a sustentabilidade é dada pela capacidade da organização dar continuidade à disponibilização de dados atualizados regularmente com pelo menos o mesmo nível de qualidade e quantidade (GLOBAL PARTNERSHIP FOR SUSTAINABLE DEVELOPMENT DATA, 2016). Optou-se por considerar o conceito de sustentabilidade como transversal ao ciclo de vida, com o intuito de garantir a continuidade do fornecimento dos dados. Sendo assim, observa-se que diversos outros fatores podem viabilizar a sustentabilidade da iniciativa e neste sentido foram distribuídos todos os fatores da categoria "Sustentabilidade da iniciativa de dados abertos" nas demais categorias que melhor os representam.

Figura 26 – Acessibilidade, interoperabilidade e padrões - A.



Fonte: Da Autora

Figura 27 – Acessibilidade, interoperabilidade e padrões - B.



Fonte: Da Autora

Em suma, podem ser observadas que as boas práticas alinhadas aos fatores importantes desta categoria, conforme apresentado nas Figuras 26 e 27, se remetem aos seguintes tópicos:

- Conhecer os dados a serem disponibilizados, no sentido de seguir adequadamente os padrões;
- Ter uma infraestrutura em nuvem capaz de coletar, gerenciar e publicar dados abertos;
- Uso de padrões para dados, metadados, licenças, URIs e protocolos de trocas;

- Fornecer *Application Programming Interface* (API)s;
- Ser transparente sobre a publicação dos dados e sobre os dados.

Apresentada na Figura 26, percebe-se que o trabalho de Susha *et al.* (2015a) alinhado às boas práticas abordadas pela literatura, nesta macroatividade contribuem com o que já é pesquisado amplamente e de certa forma orienta sobre onde encontrar boas práticas para viabilizar a acessibilidade, interoperabilidade, entre outras características necessárias para o uso de padrões no contexto de dados abertos.

4.1.4 Disponibilização/Manutenção

A quarta atividade-chave do ciclo de vida de dados abertos compreende as atividades relacionadas à disponibilização e a manutenção dos dados, também considera aspectos do contexto do meio utilizado para disponibilização. A categoria que descreve estes aspectos, é dada por:

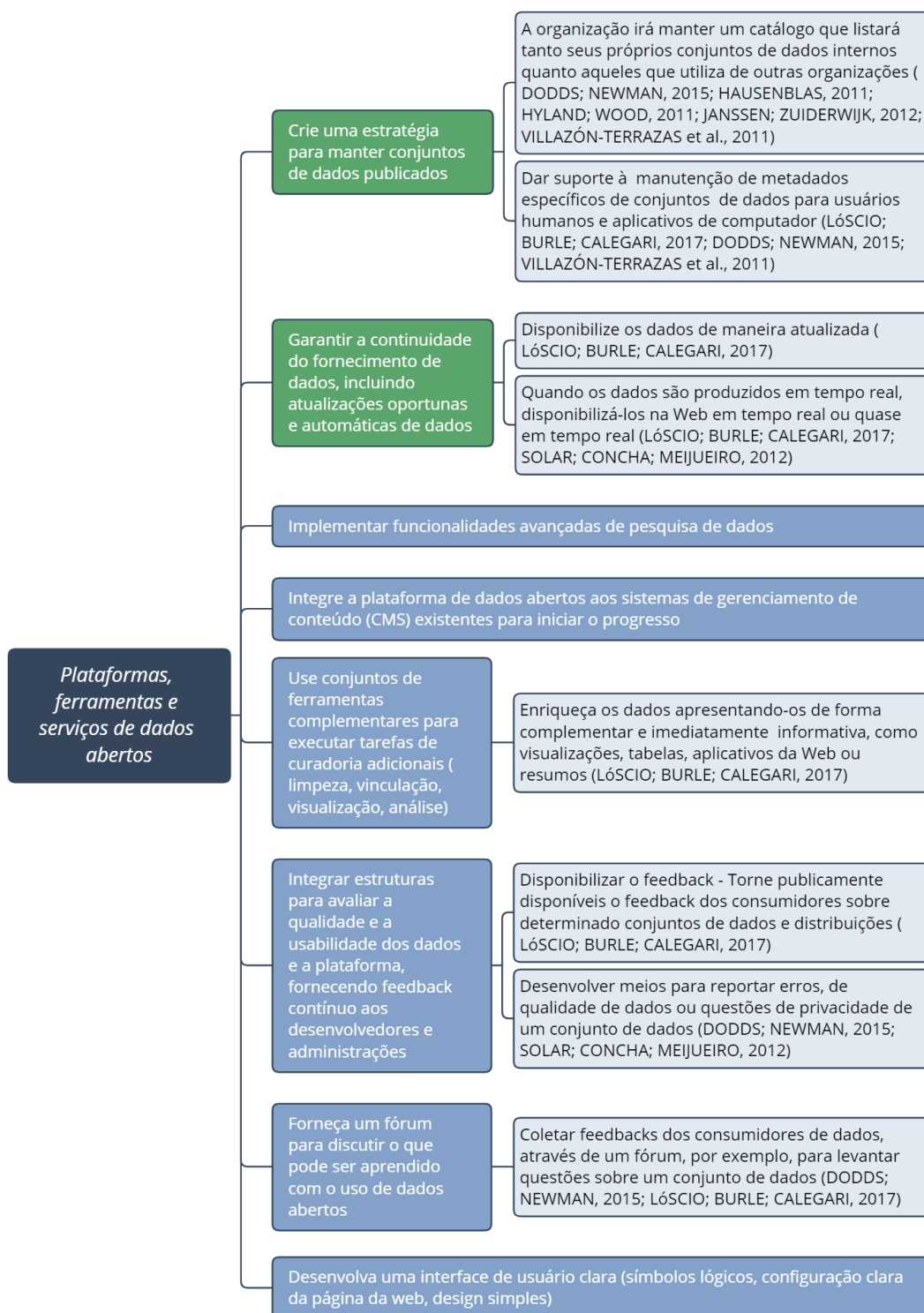
(a) **Plataformas, ferramentas e serviços de dados abertos**

Esta categoria foi ampliada e também aborda fatores originalmente da categoria "Sustentabilidade da iniciativa de dados abertos", conforme comentado previamente na seção 4.1.3 item(a), visto que para se ter uma estratégia, manter dados atualizados e uma constate disponibilização de dados faz-se necessário ter plataformas e serviços de dados abertos que contribuam com estes itens. Assim sendo, podem ser observadas na Figura 28 que as boas práticas relacionadas a esta categoria se remetem aos seguintes tópicos:

- Ter uma estratégia e implementar e garantir a continuidade do fornecimento de dados;
- Ter uma plataforma de dados abertos que permita combinar dados;
- Ter uma interface clara para a plataforma (boa usabilidade);
- Ter ferramentas de busca e enriquecimento dos dados;
- Ter meios de obter *feedback*.

Na Figura 28, percebe-se que os fatores e as boas práticas podem contribuir com a categoria anterior viabilizando que dados sejam constantemente disponibilizados conforme os padrões que foram especificados, bem como a atividade-chave posterior, ou seja, viabilizando o acesso, a visualização e o uso dos dados que serão disponibilizados.

Figura 28 – Plataformas, ferramentas e serviços de dados abertos.



Fonte: Da Autora

4.1.5 Uso

A quinta atividade-chave do ciclo de vida de dados abertos é definida por 6(seis) categorias, sendo que quando os dados são criados há um custo atrelado a este

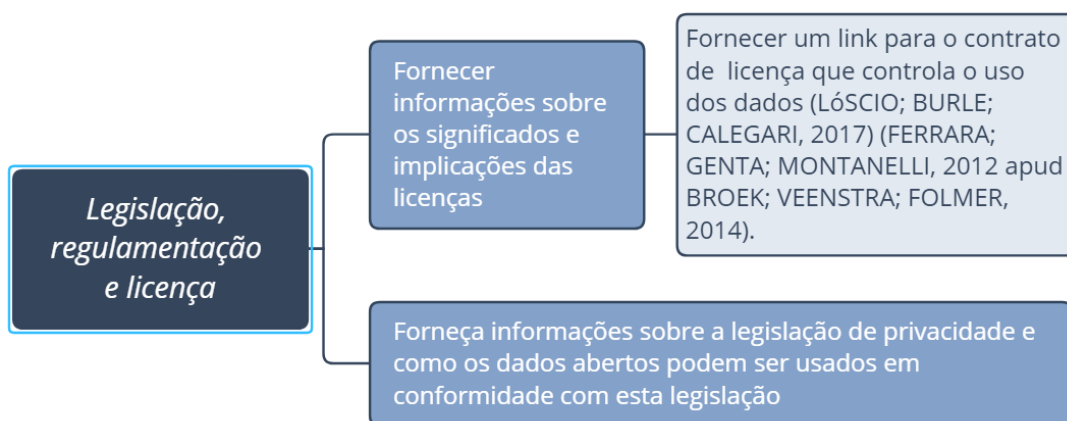
processo e somente haverá geração de valor se estes forem utilizados (DAMA INTERNATIONAL, 2017).

Diversos fatores podem influenciar na utilização dos dados e sendo que estes fatores podem ainda estar relacionados às atividades-chave iniciais, ou até mesmo a fatores externos, como por exemplo, a determinação de um objetivo que considera não somente a transparência através da divulgação dos dados, mas que busca a criação de novos negócios, assim como a facilidade de acesso e de compreensão dos dados. Portanto, optou-se por manter todas as categorias apresentados no trabalho de Sussha *et al.* (2015a) para descrever esta atividade-chave, sendo estas:

(a) **Legislação, regulamentação e licença**

A primeira categoria que descreve os fatores importantes que viabilizam o uso dos dados é descrita como: "Legislação, regulamentação e licença", esta é uma categoria que aparentemente coincide com o que já foi mencionado no início desta seção, no entanto optou-se por trazê-la nesta fase devido a importância de se ter informações sobre licenças e sobre a privacidade dos dados, visto que a licença fornece uma declaração legal que permite fazer algo, como por exemplo, conceder a permissão de uso e de distribuição dos dados.

Figura 29 – Legislação, regulamentação e licença.



Fonte: Da Autora

De modo geral, podem ser observadas que as boas práticas e fatores importantes da Figura 29 se remetem ao seguinte tópico:

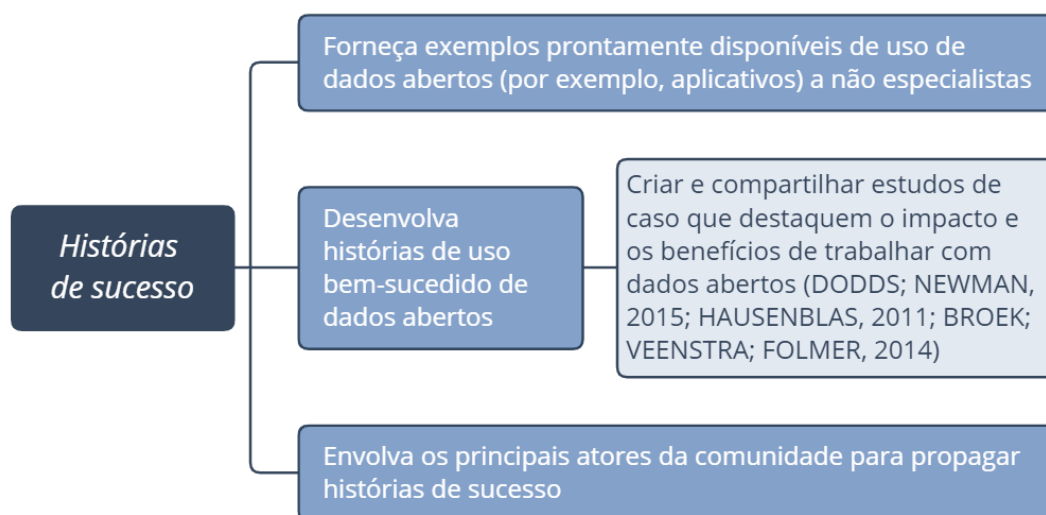
- Fornecer informações sobre as questões legais que regem os dados.

(b) **Histórias de sucesso**

A segunda categoria apresentada é descrita como "Histórias de sucesso", esta categoria aborda aspectos de divulgação, onde e como os dados estão sendo aplicados. Inclusive esta categoria foi considerada como fator importante na

atividade-chave de Gerenciamento, na categoria "Treinamento e apoio aos funcionários", dado pela importância de não apenas quem faz uso dos dados tomar ciência de onde e como os dados estão sendo utilizados, mas também quem os publica.

Figura 30 – Histórias de sucesso



Fonte: Da Autora

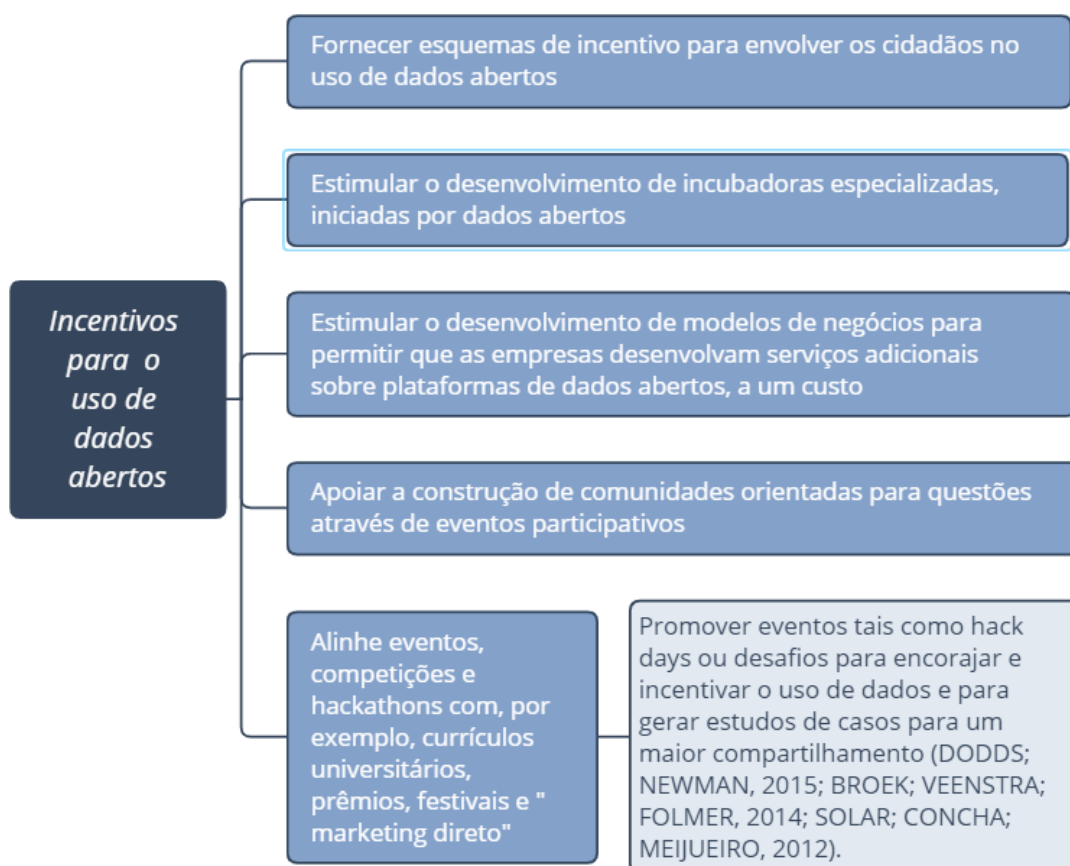
Em suma, podem ser observadas que as boas práticas e fatores importantes da Figura 30 se remetem ao seguinte tópico:

- Divulgar histórias de sucesso do uso de dados abertos.

(c) **Incentivo ao uso**

Assim como as duas primeiras categorias apresentadas, a terceira categoria "Incentivo para o uso de dados abertos" visa fomentar o uso dos dados de maneira mais ampla, ou seja, considerando articular diferentes formas para que os dados disponibilizados sejam utilizados, isto com o intuito de promover aquilo que já está disponível, bem como coletar informações sobre a necessidade de novos dados e quais padrões seriam mais adequados pra o uso.

Figura 31 – Incentivo ao uso



Fonte: Da Autora

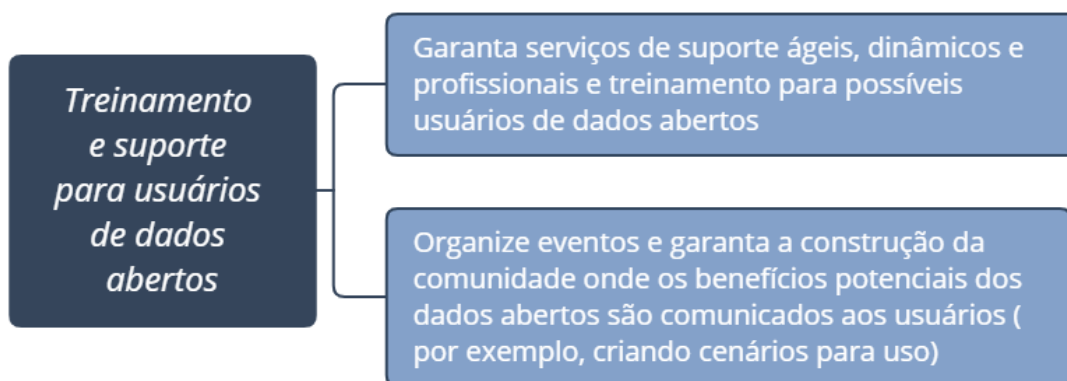
Desta forma, podem ser observadas que as boas práticas e fatores importantes da Figura 31 se remetem aos seguintes tópicos:

- Fornecer esquemas de incentivo para o uso dos dados;
- Estimular a criação de modelos de negócios e incubadoras;
- Promover eventos para o uso dos dados (*Hackdays*);
- Apoiar a construção de comunidades de dados abertos.

(d) **Treinamento e suporte para usuários**

A categoria "Treinamento e suporte a usuários", visa contribuir com a construção de uma comunidade de dados abertos, com o intuito de fomentar a utilização dos dados, bem como coletar informações sobre os dados que foram e que podem vir a ser disponibilizados.

Figura 32 – Treinamento e suporte para usuários



Fonte: Da Autora

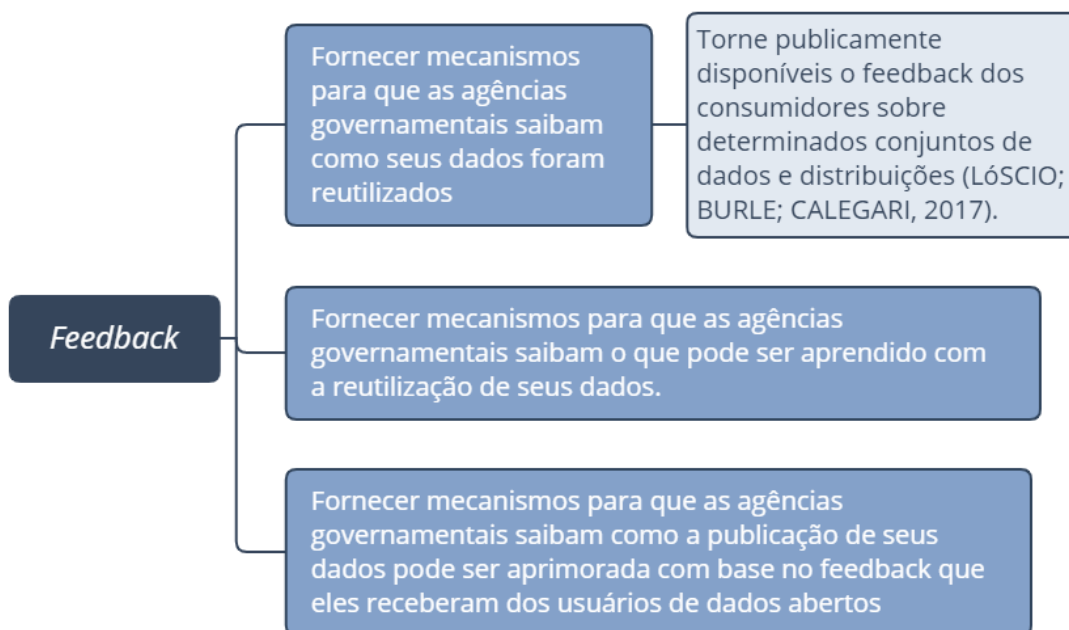
E neste sentido, podem ser observados que os fatores importantes da Figura 32 se remetem aos seguintes tópicos:

- Promover eventos para o uso dos dados;
- Garantir suporte(ágil) aos usuários dos dados.

(e) **Feedback**

A categoria "Feedback e Sustentabilidade" apresentada no trabalho de Sussha *et al.* (2015a) será abordada neste trabalho, apenas por "*Feedback*", devido ao conceito de sustentabilidade utilizado ser transversal ao ciclo de vida apresentado. Sendo assim, esta categoria tem o intuito de viabilizar a obtenção de *feedback* daqueles que fazem uso dos dados e de certa forma contribuir com a sustentabilidade, com a melhoria dos processos e dos dados que são disponibilizados.

Figura 33 – Feedback



Fonte: Da Autora

Portanto, na Figura 33 observa-se que as boas práticas e fatores importantes se remetem ao seguinte tópico:

- Compreender como seus dados estão sendo utilizados e o que se pode aprender com isso;
- Aprimorar seus dados com base nos *feedbacks*.

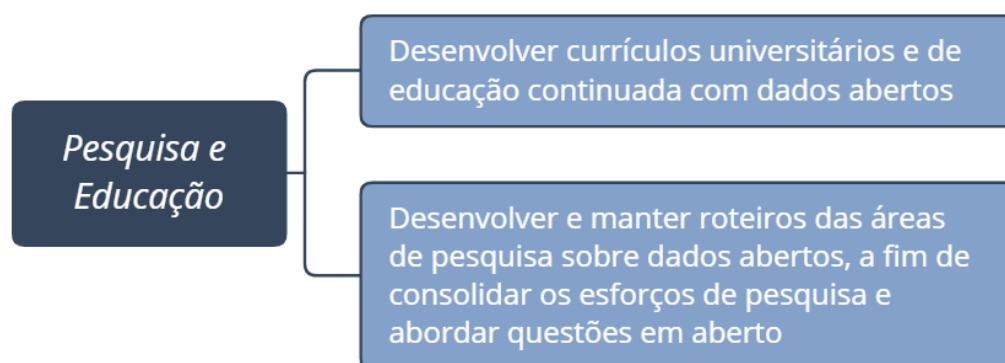
(f) **Pesquisa e Educação**

A sexta categoria e última, abordada por esta atividade-chave, é descrita por "Pesquisa e Educação", visto que pesquisas sobre dados abertos podem contribuir significativamente com a evolução dos processos e de boas práticas nesta área, entre outros aspectos. A inclusão de dados abertos na educação pode fomentar não somente o uso dos dados, mas criar cultura acerca de dados abertos.

E neste sentido, podem ser observados que os fatores importantes relacionados à categoria apresentada na Figura 34 se remetem de maneira geral ao seguinte tópico:

- Desenvolver pesquisa e educação focada em dados abertos.

Figura 34 – Pesquisa e Educação



Fonte: Da Autora

Destaca-se que as duas últimas categorias podem contribuir diretamente com a próxima atividade-chave a ser descrita, no sentido de melhorar não somente os dados que são disponibilizados, mas todo o contexto que engloba a abertura de dados.

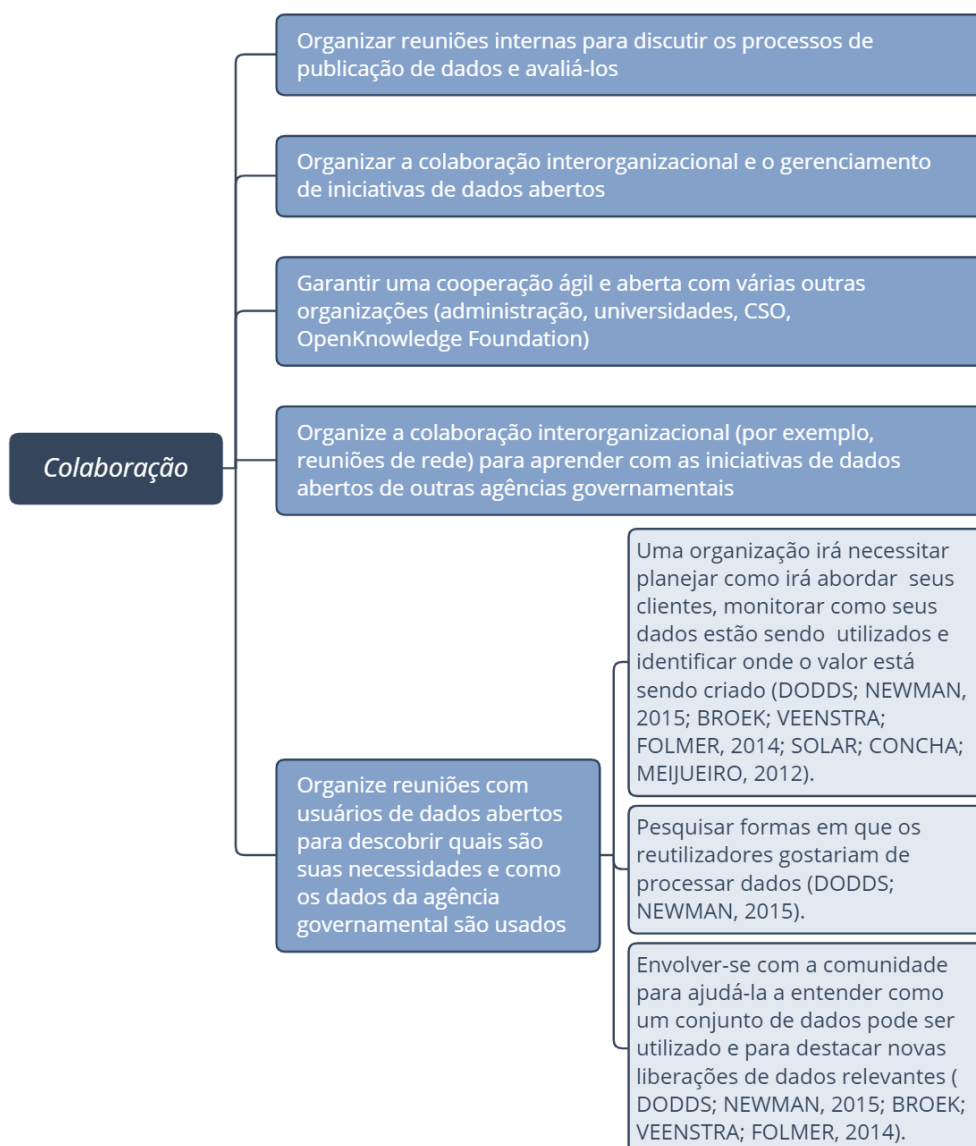
4.1.6 Aprimoramento

A sexta atividade-chave do ciclo de vida de dados abertos é apresentada como "Aprimoramento", sendo que a literatura aborda o contexto de melhoria dos dados. Todavia, o objetivo principal desta atividade-chave não é focar apenas na melhoria dos dados, mas também conseguir avaliar a estratégia de abertura de dados, assim como todo o contexto acerca de dados abertos pode ser aperfeiçoado. Nesta macroatividade optou-se por considerar não apenas a avaliação, mas também serão considerados aspectos de colaboração. Para descrever esta atividade-chave, tem-se as seguintes categorias:

(a) **Colaboração**

A primeira categoria que descreve esta atividade-chave denomina-se "Colaboração", sendo que esta visa à colaboração interna, inter-organizacional, de órgãos regulamentadores acerca de dados abertos e dos usuários de dados, ou seja, organiza diversas atividades para que estas partes possam colaborar com a qualidade dos dados e dos processos acerca de dados abertos, bem como com a estratégia e a abordagem adotada para a disponibilização destes dados.

Figura 35 – Colaboração



Fonte: Da Autora

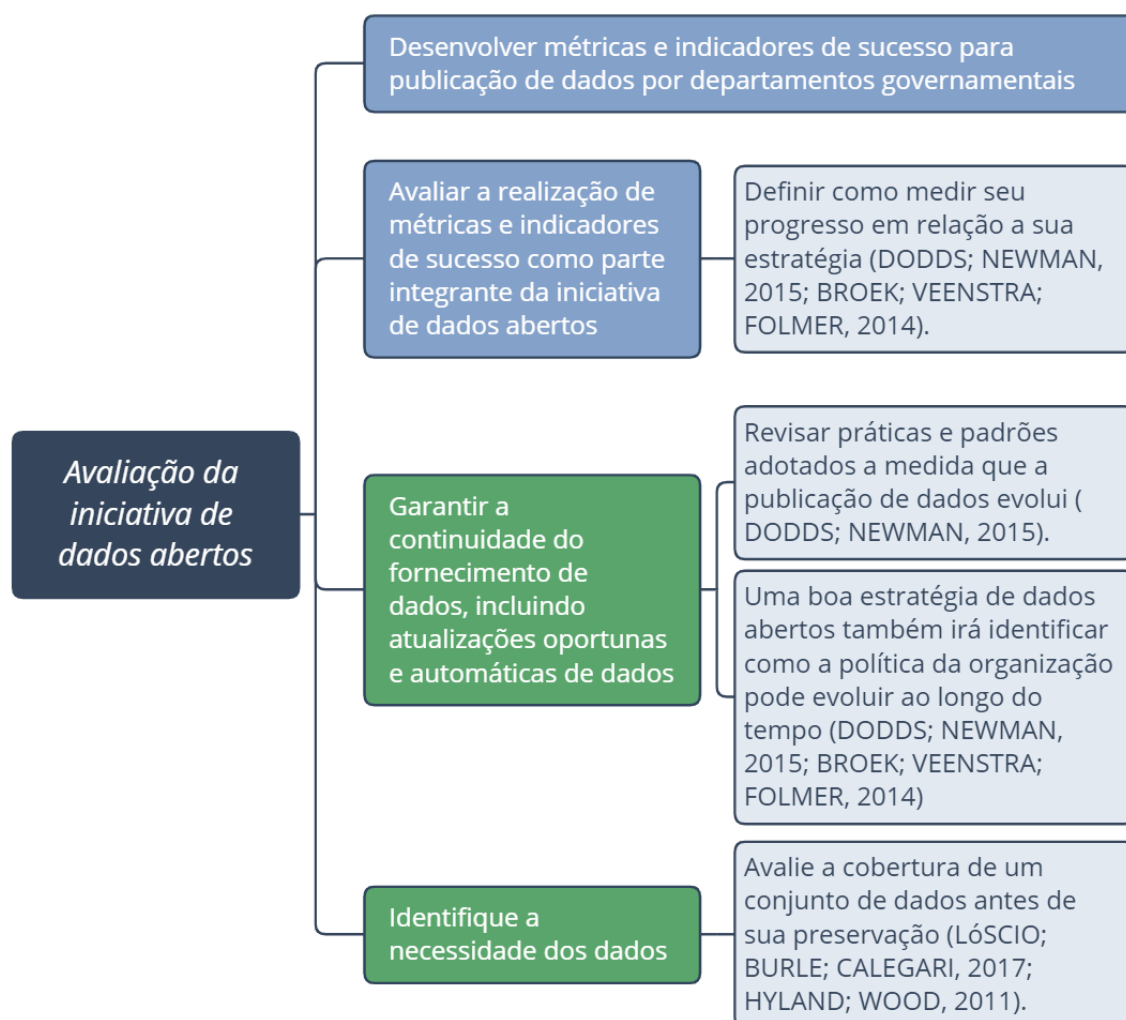
Neste sentido, pode-se resumir as boas práticas e fatores importantes apresentados na Figura 35 através do seguinte tópico:

- Garantir e viabilizar a colaboração interorganizacional, organizacional e dos usuários de dados.

(b) **Avaliação da iniciativa**

A segunda categoria para a atividade-chave aprimoramento é denominada "avaliação da iniciativa", considerando, que somente avaliando o que está sendo feito pode-se contribuir com a melhoria dos processos e até mesmo dos dados que são disponibilizados. Para isso a literatura aborda os fatores importantes e boas práticas que são apresentados na Figura 36.

Figura 36 – Avaliação da iniciativa



Fonte: Da Autora

Esta categoria oriunda do trabalho de Sussha *et al.* (2015a) foi ampliada com os fatores que estavam relacionados à sustentabilidade (na Figura 36 os que estão na cor verde). Portanto, nota-se a necessidade de se construir indicadores e métricas com o intuito de avaliar a estratégia utilizada para a abertura de dados e também considerando a utilização dos dados.

Neste sentido, são observadas que as boas práticas e fatores relacionados a esta categoria se remetem aos seguintes tópicos:

- Desenvolver métricas internas e externas para avaliar a iniciativa de dados abertos;
- Revisar as práticas e padrões utilizados e aprimorá-las a medida que a publicação de dados evolui;
- Avaliar a preservação de conjuntos de dados.

Este último tópico auxilia na tomada de decisão referente à preservação ou não de um conjunto de dados. No entanto, como esta é uma das últimas atividades-chave do ciclo de vida proposto, destaca-se que após ser aprimorada a iniciativa de dados abertos esta poderá ser iniciada novamente, ou seja, novos conjuntos de dados serão criados ou aprimorados gerando novas publicações e com isso viabilizando a sustentabilidade da iniciativa de dados abertos, cujo um dos objetivos é a continuidade do fornecimento dos dados.

4.1.7 Descarte

Esta última atividade-chave do ciclo é caracterizada pela remoção do conjunto de dados de uma determinada plataforma ou serviço de dados abertos e neste sentido a literatura apresenta apenas a recomendação de "Ao remover dados da Web, preserve o identificador e forneça informações sobre o recurso arquivado" (LÓSCIO *et al.*, 2017; HYLAND; WOOD, 2011), ou seja, esta última atividade-chave destaca o contexto da remoção dos dados, mas não da iniciativa como um todo, que esta última somente será encerrada com a finalização das atividades de abertura, remoção da plataforma ou de todos os conjuntos de dados que foram disponibilizados.

Evidencia-se que o ciclo de vida proposto poderia explicitar as atividades intermediárias ou sub-atividades que as categorias buscam evidenciar, agregando a elas um verbo de ação que contemple os fatores importantes que a descrevem, por exemplo, na atividade-chave **uso** poderia se ter a atividade intermediária descrita por: "Fornecer *feedback*", visto que os fatores importantes que descrevem a categoria também fazem uso do verbo fornecer, ver Figura 33.

Na sequência, a próxima seção visa abordar as principais diferenças entre o ciclo de vida abordado no trabalho de Broek *et al.* (2014) e o ciclo de vida apresentado.

4.2 COMPARATIVO ENTRE OS CICLOS DE VIDA

O ciclo de vida proposto por Broek *et al.* (2014) se assemelha ao ciclo de vida de dados abertos apresentado, pois os dois ciclos foram desenvolvidos com base na literatura, sendo que este trabalho além de considerar os trabalhos já abordados por Broek *et al.* (2014) também considera estudos mais recentes encontrados na literatura e a governança de dados. E neste sentido, esta seção visa comparar as duas estruturas com o intuito de enfatizar as principais contribuições.

As fases, passos por fase, atividades-chave e categorias de cada ciclo de vida são apresentadas no Quadro 18.

Quadro 18 – Comparativo entre fases e atividades-chave dos ciclos de vida

(BROEK <i>et al.</i> , 2014)		Ciclo de vida proposto	
Fases	Passos por fase	Atividades-chave	Categorias de fatores importantes
Identificação	Definir a estratégia	Planejamento	Estratégia e Apoio
	Selecionar os dados	Gerenciamento	Suporte a gestão e aos processos de publicação
Preparação	Definir requisitos		Criação
	Modelar e descrever os dados	Acessibilidade, interoperabilidade e padrões	
	Converter para um formato de dados legível por máquina	Disponibilização/Manutenção	Plataformas, ferramentas e serviços de dados
	Conectar os dados	Uso	Legislação, regulamentação e licenças
Publicação	Publicação de dados		Histórias de sucesso
	Publicação de metadados		Incentivo ao uso
Reuso	Exploração de dados publicados	Feedback	Treinamento e suporte para usuários
	Gestão de dados		Pesquisa e educação
Avaliação	Desenvolver propostas de negócio	Aprimoramento	Avaliação da Iniciativa
	Monitorar e melhorar os dados		Colaboração
		Descarte	

Fonte: Da Autora

Observa-se que existe um alinhamento entre alguns dos passos por fase e as categorias. Contudo, pode-se perceber que ainda há um grande enfoque em atividades relacionadas aos dados, na proposta de Broek *et al.* (2014), enquanto o ciclo de vida proposto engloba atividades relacionadas a dados, ao processo de abertura dos dados e a pós-publicação, considerando também aspectos estratégicos, de sustentabilidade da iniciativa, entre outros, ou seja, isso viabiliza uma visão abrangente sobre o ciclo de vida de dados abertos.

Vale ressaltar que o ciclo proposto por Broek *et al.* (2014) foi aperfeiçoado com base nas lições aprendidas em um estudo de caso desenvolvido em uma empresa semi-pública na Holanda e com isso chegou à representação gráfica apresentada no

Quadro 8, sendo que alguns dos passos por fase foram agrupados, com base no estudo de caso.

Ainda relacionado aos passos por fase são descritas 32 atividades que os autores encontraram na literatura referentes a 7 ciclos de vida que descrevem a abertura de dados abertos conectados. O ciclo de vida proposto considera estas atividades contemplando fatores importantes de publicação e uso de dados, oriundos do trabalho de Susha *et al.* (2015a) e boas práticas para a publicação de dados na Web, do trabalho de Hyland *et al.* (2014) e com isso buscou-se ampliar e atualizar as discussões realizadas, como por exemplo, incluindo questões relacionadas à sustentabilidade, sendo que os dados conectados são abordados como uma consequência da evolução ou aprimoramento da abertura dos dados.

Outro ponto observado no trabalho de Broek *et al.* (2014) é que não está claro como foram identificados os cinco *stakeholders* apresentados junto ao ciclo de vida, sendo este: a alta gerência, o gerente de informação, o consultor jurídico, o gerente de comunidade e o proprietário dos dados, observando ainda que ao longo do texto há um sexto (6º *stakeholder*), o gerente de projetos.

Neste contexto as contribuições e discussões conduzidas nesta tese são acerca de *stakeholders* em um nível mais abstrato, estes são descritos no Quadro 19. O intuito desta escolha é que a partir destes papéis possam ser derivados outros, conforme a necessidade da organização. Assim como quando se tem esforços iniciais de abertura de dados, o criador e o mantenedor podem ser uma única pessoa.

Quadro 19 – *Stakeholders* do processo de abertura de dados

Stakeholders	Responsabilidade que podem estar atreladas
Provedor de dados	Organização ou departamento responsável por fornecer dados
Criador ou Autor	Quem cria, extrai, anonimiza, transforma ou mescla os dados
Mantenedor	Quem é o responsável por manter disponíveis os dados
Consumidor ou Usuário	Aquele que recupera e faz uso dos dados

Fonte: Da Autora

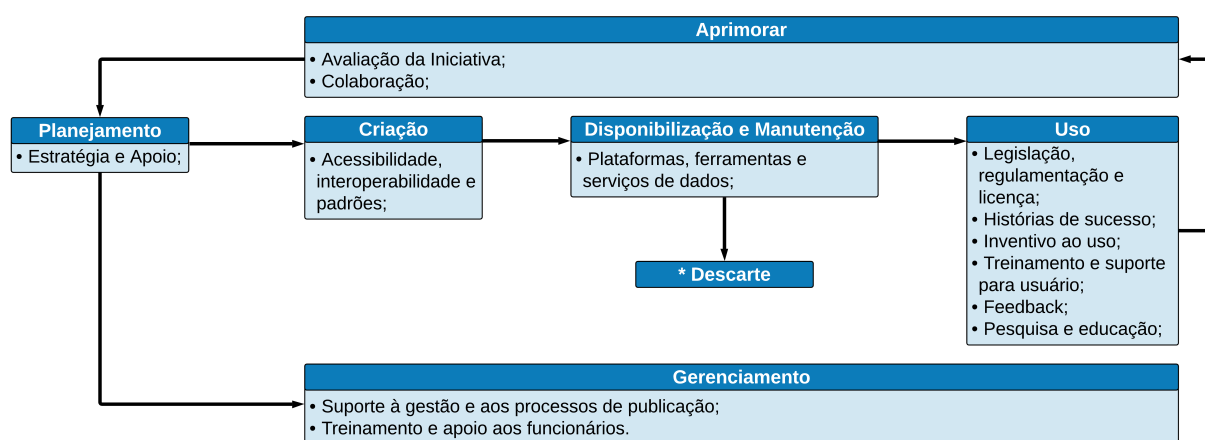
Através deste comparativo é possível perceber que o ciclo de vida proposto é abrangente e tem objetivo de servir como referência para a derivação de novos estudos, pois fornece em uma única estrutura as atividades relevantes apontadas pela literatura para dados abertos, ou seja, fornece um guia do "o que fazer?" quando a organização deseja abrir, ou que já abriu seus dados e necessita organizar e alinhar suas estratégias com a abertura de dados.

No entanto, destaca-se que os ciclos de vida de publicação, no caso do trabalho

de Auer *et al.* (2011), bem como as boas práticas de publicação de dados na Web, podem contribuir significativamente em especial com a atividade-chave de criação, viabilizando o "como fazer?", pois estes dois estudos apresentam esta perspectiva.

Em suma, apresenta-se a Figura 37, onde busca-se representar graficamente as atividades-chave e as categorias de atividades e/ou boas práticas que foram discutidos nesta seção; Destacam-se ainda as macroatividades de gerenciamento e aprimoramento, pois estas perpassam as demais macroatividades, conforme já discutido na seção.

Figura 37 – Ciclo de Vida de Dados Abertos



Fonte: Da Autora

Por fim, considerando o ciclo de vida apresentado, as discussões acerca da estruturação e o que foi discutido na seção 2 percebem-se as contribuições deste trabalho no sentido de organizar, alinhar e discutir o que há na literatura referente a dados abertos.

A estrutura apresentada contribui com uma visão abrangente e com isso podem ser derivados novos ciclos de vida, por exemplo, para contemplar a realidade de uma organização em específico. Assim como podem ser gerados *insights* sobre problemas, processos, tecnologias, entre outros fatores envolvidos no âmbito de dados abertos.

5 MODELO DE MATURIDADE DE DADOS ABERTOS

Apresenta-se a seguir como se deu a estruturação do modelo de maturidade de dados abertos. Primeiramente, são expostas as definições iniciais a partir dos quais o modelo foi idealizado. Em seguida, mostra-se a composição do modelo, com as áreas foco e seus respectivos componentes, as capacidades correspondentes a cada componente e as relações de dependência existentes. Por fim, é apresentado o instrumento de avaliação da maturidade, denominada de matriz de maturidade.

5.1 DEFINIÇÕES INICIAIS

As definições iniciais acerca do modelo se referem à **identificação e domínio do escopo**; para isso optou-se por fazer uso das fases genéricas **escopo** e **design** abordadas no trabalho de De Bruin *et al.* (2005). O **escopo** envolve questões mais gerais que guiam o desenvolvimento do modelo, se referindo ao foco e aos *stakeholders* que auxiliaram no desenvolvimento do modelo de maturidade.

O foco do modelo, ou seja o domínio do escopo é dados abertos como este domínio é específico e não foram encontrados muitos trabalhos que abordam modelos de maturidade, conforme observado na seção 2.4, foram utilizados artigos acadêmicos onde fossem evidenciadas as atividades e/ou boas práticas e aspectos importantes para a publicação e pós-publicação dos dados, conseqüentemente chegou-se ao ciclo de vida de dados abertos, apresentado na Figura 37.

Na Figura 37 observa-se uma atualização dos ciclos de vida encontrados na literatura, assim como uma complementação das atividades, devido a ser considerada uma perspectiva diferenciada do que é abordado na literatura. E neste sentido, observa-se que a academia contribui significativamente para mapear as atividades-chave do ciclo de vida de dados abertos, assim como viabilizou a estruturação do modelo de maturidade. Um resumo destas especificações são apresentadas no Quadro 20.

Quadro 20 – Definições acerca do escopo do modelo de maturidade

Critério	Característica	Justificativa
Foco do modelo	Ciclo de vida de dados abertos	Foco no contexto de dados abertos como um todo, não somente na publicação ou abertura.
Stakeholders do desenvolvimento	Academia	O ciclo de vida apresentado na seção 4 e o modelo de maturidade proposto foram desenvolvidos com base na literatura.

O **design** do modelo, conforme o trabalho de De Bruin *et al.* (2005) envolve as especificações referentes a público-alvo, ao método de aplicação, ao driver da aplicação, aos respondentes e à aplicação. Neste sentido, observa-se que público-alvo do modelo é o interno, pois a avaliação deve ser uma demanda da própria organização, ou de um departamento que está realizando a abertura de seus dados, visando a verificação das atividades e conseqüentemente observando-se como poderão ser melhoradas suas estratégias e atividades, quando consideradas as boas práticas ou atividades da literatura de dados abertos.

O método de aplicação é a autoavaliação, com execução por funcionários envolvidos com o processo da abertura dos dados. A autoavaliação poderá ser executada por um único indivíduo. No entanto, destaca-se o enriquecimento da avaliação se esta for exercida em equipe, pois possibilita uma perspectiva multifuncional e haverá uma oportunidade para o consenso (FRASER, P. *et al.*, 2002).

O driver da aplicação é a exigência interna, pois como mencionado previamente a ideia é que a própria organização ou departamento deverá observar a necessidade de avaliar. E por fim, a aplicação do modelo é realizada em uma entidade. Essas especificações são resumidas no Quadro 21.

Quadro 21 – Definições acerca do design do modelo de maturidade

Critério	Característica	Justificativa
Público	Interno (Executivos, Gestão)	O público-alvo a que serve o modelo de maturidade é interno às organizações, pois deve ser uma demanda da própria organização ou departamento, a verificação das suas atividades referente, às boas práticas acerca da abertura dos dados.
Método de Aplicação	Autoavaliação	O método de aplicação é autoavaliação para facilitar a disseminação do modelo e a aplicação por <i>stakeholders</i> envolvidos na abertura dos dados.
Driver de Aplicação	Exigência Interna	A aplicação pode partir da própria organização ou departamento.
Respondentes	Funcionários	Devido ao tipo de afirmações e conhecimento necessário para responder à avaliação, os respondentes deverão ser funcionários, ou seja, <i>stakeholders</i> envolvidos com as atividades de dados abertos.
Aplicação	1 entidade/ 1 região	A aplicação do modelo se dará em 1 organização.

Considerando as definições iniciais optou-se pela concepção de um modelo de maturidade descritivo, ou seja, com foco na verificação da situação atual de uma determinada organização ou de um determinado departamento. Esta escolha se deu devido à necessidade de se ter uma compreensão do domínio e como a maturidade será alcançada, visto que neste domínio poucas pesquisas exploram o contexto da maturidade.

Destaca-se ainda que um modelo de maturidade descritivo poderá evoluir para ser prescritivo, por meio da compreensão sólida da situação atual, pois assim podem ser incluídas melhorias substanciais e repetíveis (DE BRUIN *et al.*, 2005).

Para conduzir as demais etapas da construção do modelo de maturidade de dados abertos optou-se por fazer uso do trabalho de Van Steenbergem *et al.* (2010a), devido às particularidades do modelo, sendo estas discutidas na próxima seção.

5.2 ESTRUTURAÇÃO DO MODELO

A etapa de **Design do modelo**, conforme o trabalho de Van Steenbergem *et al.* (2010a) compreende a decisão sobre o conteúdo do modelo, necessitando identificar o que precisa ser medido na avaliação de maturidade e como isso pode ser medido. Para isso, têm-se os seguintes passos:

5.2.1 Determinar as áreas foco

As áreas foco foram definidas com base na governança de dados, com o objetivo de lidar com problemas de dados e garantir que os dados atendam às necessidades dos negócios, devido a governança de dados representar uma separação inerente do dever entre supervisão e execução (DAMA INTERNATIONAL, 2017).

Para isso, fez-se o estudo do *framework* DAMA-DMBOK, referência na área, e que visa promover o entendimento, desenvolvimento de prática de gerenciamento de dados e informações para apoiar as estratégias de negócios (MOSLEY *et al.*, 2009). No entanto, foi observado no Guia DAMA-DMBOK (DAMA INTERNATIONAL, 2017), um *framework* atualizado, completo e detalhado, que envolve a governança de dados e todas as áreas associadas. Nota-se que este estudo compreende detalhes avançados sobre a governança de dados, que não corresponde a realidade dos estudos acerca de dados abertos e por este motivo optou-se pelo uso de um *framework* sistematizado, com o intuito de obter um modelo de maturidade mais resumido em sua primeira versão.

Assim sendo, fez-se o uso do *framework* definido por Panian (2010), apresentado na Figura 9, para evidenciar as áreas foco que serão utilizadas na estruturação do modelo de maturidade. Foi utilizado o *framework*, devido a esta estrutura de governança de dados ser eficaz, sistematizada e por levar em consideração os padrões,

políticas e processos, estrutura organizacional e infraestrutura tecnológica necessária para tornar os dados acessíveis, disponíveis, de alta qualidade, consistentes, auditáveis e seguros (PANIAN, 2010). Sendo que este *framework* vem sendo utilizado em estudos recentes como base para estudos a respeito de Big Data (KIM; CHO, 2017), ou seja, é observado que pode ser aplicado quando há um grande volume de dados.

Observando os quatro componentes discutidos no trabalho de Panian (2010), sendo estes: **padrões, políticas e processos, organização, e tecnologia** e o contexto abordado por cada um destes, originaram as seguintes áreas de foco para o modelo de maturidade:

- **Gerenciamento de políticas e processos:** as organizações devem aplicar políticas e processos para a criação, desenvolvimento, controle, gerenciamento e auditoria dos dados (PANIAN, 2010). Para representar esta área foco serão utilizados os fatores elencados pelo mesmo autor, sendo estes: (1) definição de dados, (2) monitoramento e medição, (3) acesso e entrega de dados e (4) gerenciamento de mudança de dados.
- **Gerenciamento de fatores organizacionais:** uma questão importante que as organizações devem abordar ao lançar uma iniciativa de governança de dados é como projetar a estrutura organizacional. E neste sentido o componente "organização", abordado no trabalho de Panian (2010), será abordado nesta tese como fatores organizacionais, sendo representado pelos mesmos fatores apresentados pelo trabalho de origem, sendo estes: (1) funções e responsabilidade, (2) treinamento e educação¹, (3) planejamento e priorização² e (4) gestão de mudança organizacional³.
- **Gerenciamento de padrões:** Considerando aplicar padrões técnicos e de desenvolvimento relacionado aos dados, devido à necessidade da organização estabelecer e fazer uso de padrões para os dados disponibilizados. Como para este componente os fatores apresentados não representam adequadamente a realidade dos padrões para dados abertos, optou-se pelo uso dos "Desafios de dados na Web" apresentados no trabalho de Lóscio *et al.* (2017) para representar este componente, sendo estes: (1) licença, (2) proveniência e qualidade, (3) versionamento, (4) identificação, (5) formato, (6) vocabulários, (7) metadados, (8)

¹ **Treinamento e educação** - após a identificação das capacidades foram abordadas separadamente devido as características de cada um destes, considerando o que foi encontrado na literatura.

² **Planejamento e Priorização** - após a identificação das capacidades foi segmentado em dois, sendo estes "planejamento e priorização da abertura dos dados" e "planejamento da publicação dos dados", esta segmentação se deu devido às características das boas práticas encontradas na literatura que se relacionam com este fator.

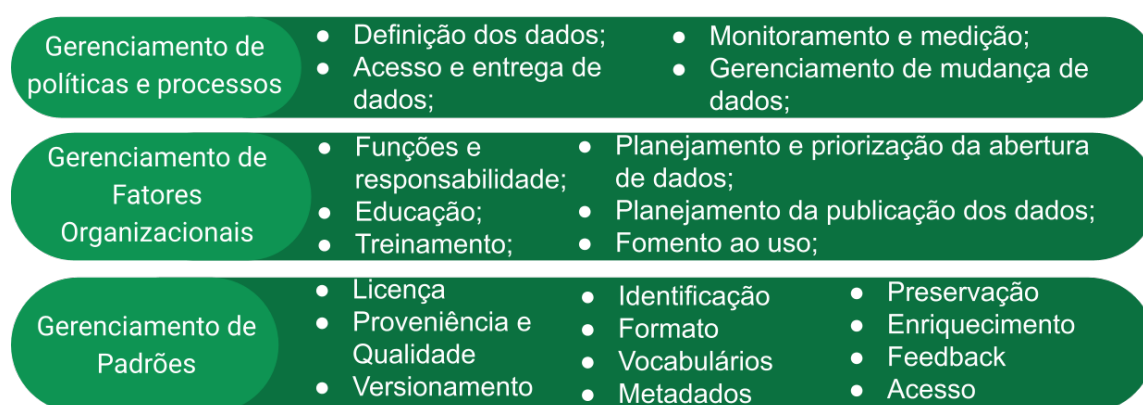
³ **Gestão da mudança organizacional** - após a identificação das capacidades, este fator foi removido devido à literatura não apresentar nenhum embasamento que possibilitasse a avaliação da maturidade

preservação, (9) enriquecimento, (10) *feedback*, (11) acesso e (12) reuso⁴, visto que o trabalho apresenta detalhadamente os padrões que podem ser aplicados sobre dados abertos.

- **Tecnologia:** A tecnologia não será abordada como área foco devido a ser necessária uma estrutura subjacente, ou seja, que dará suporte a diversas atividades compreendidas no contexto do modelo de maturidade.

Após uma primeira organização das capacidades, foram observadas algumas necessidade inerentes ao domínio funcional e com estas alterações chegou-se na estrutura apresentada na Figura 38.

Figura 38 – Resumo das áreas foco e seus componentes.



Fonte: Da Autora

Considerando a Figura 38 observam-se os elementos que compõem cada uma das três áreas foco. Neste sentido, estes elementos serão denominados de "componentes" e para cada um destes serão evidenciadas as capacidades para determinar a maturidade do componente e conseqüentemente da área foco. Na sequência, discute-se como foram evidências estas capacidades.

5.2.2 Determinar as capacidades e determinar as dependências

As capacidades foram elencadas com base em três ações que indicam a execução da governança de dados, sendo estas: definir, implementar e monitorar (ALHASSAN *et al.*, 2016). Ainda, conforme discutido junto ao ciclo de vida uma 4ª atividade denominada por "Aprimorar" será utilizada, com o intuito de garantir a continuidade no fornecimento de dados abertos, com no mínimo a mesma qualidade, bem como com a possibilidade de evoluir as políticas, processos e padrões utilizados pela organização.

Analisando o contexto apresentado, as atividades serão denominadas por:

⁴ **Reuso** - após a identificação das capacidades foi movido para a área foco gerenciamento de fatores organizacionais, devido às características das capacidades evidenciadas e assim sendo este denominado de **fomento ao uso**

- A** - que visa representar atividades que envolvem a **definição**.
- B** - que visa representar atividades que envolvem a **implementação**.
- C** - que visa representar atividades que envolvem o **monitoramento**.
- D** - que visa representar atividades que envolvem o **aprimoramento**.

Neste sentido, cada componente será representado pelas capacidades progressivas (A, B, C e D), sendo estas posteriormente posicionadas na matriz de maturidade, considerando as dependências existentes. Isso possibilitará determinar os níveis de maturidade de cada componente, bem como de cada área foco e por fim da organização como um todo.

A ideia da estruturação das capacidades partiu do princípio de descrever "o que" deve ser feito e não o "como" será feito. Seguindo esta lógica, na sequência serão apresentadas cada uma das áreas foco, com seus respectivos componentes, a descrição das capacidades evidenciadas e suas dependências. Destaca-se ainda, que as boas práticas observadas em cada quadro são oriundas da literatura e os fatores importantes estes foram observados no trabalho de (SUSHA *et al.*, 2015a), sendo que estes mesmos foram apresentados junto ao ciclo de vida e aqui reorganizados para abranger a realidade de cada área foco e de seus respectivos componentes.

- **Gerenciamento de políticas e processos**

Para esta área foco serão apresentados os componentes: I) definição dos dados, II) acesso e entrega dos dados, III) monitoramento e medição e IV) gerenciamento de mudança de dados. Em cada quadro será exibido o objetivo de cada componente, as perguntas elencadas para a avaliação das capacidades, as boas práticas e fatores importantes que auxiliaram na composição das perguntas, as capacidades e os pré-requisitos⁵, que consistem em todas as perguntas que terão que ter resposta afirmativa para que seja atribuída a capacidade, ou seja, as dependências existentes entre as questões. Em cada Quadro destaca-se em negrito o(s) termo(s) ou frase(s) que estão contemplados diretamente nas perguntas apresentadas.

Estabelecer políticas e/ ou processos em torno das práticas e tratamentos dos dados tornam a criação, o gerenciamento, o monitoramento e o aprimoramento dos dados eficaz (PANIAN, 2010). No Quadro 22 observa-se o contexto considerado para o componente definição dos dados.

⁵ **P1.1**, por exemplo, corresponde ao componente 1 e a pergunta de número 1, já **P1.2**, corresponde ao componente 1 e a pergunta de número 2 e assim sucessivamente.

Quadro 22 – Definição dos Dados

Componente: 1 - Definição dos dados		
Tem por objetivo verificar se as políticas e processos relacionados à mitigação de riscos, a dados a serem ou não disponibilizados, bem como sobre a conversão dos dados para um formato legível por máquina foram definidas, implementadas, monitoradas, mantidas e/ou aprimoradas.		
Fatores Importantes	Boas Práticas	Perguntas
Identificar a necessidade dos dados	<p>Mitigar riscos antes da disponibilização, considerando anonimizar, editar e agregar os dados para remover informações sensíveis (DODDS; NEWMAN, 2015), (JANSSEN; ZUIDERWIJK, 2012 apud BROEK <i>et al.</i>, 2014).</p>	<p>1.1-Existem políticas e/ou processos de mitigação de riscos antes da disponibilização dos dados?</p> <p>1.2-Existem políticas e/ou processos para definir quais dados serão ou não disponibilizados?</p> <p>1.3-Existem políticas e/ou processos de conversão de dados para um formato legível por máquina?</p> <p>1.4- As políticas e/ou os processos estão sendo executados para a mitigação de riscos?</p> <p>1.5- As políticas e/ou processos para definição da disponibilização de conjunto de dados estão sendo executadas?</p> <p>1.6- As políticas e/ou os processos estão sendo executados para a conversão de formato?</p>
	<p>Converter dados para formato legível por máquina, previamente definido (VILLAZÓN-TERRAZAS <i>et al.</i>, 2011; HYLAND; WOOD, 2011; BROEK <i>et al.</i>, 2014)</p>	
Definir quais dados serão ou não publicados	<p>Comece com a publicação de dados que são interessantes para os usuários, para que eles vejam o benefício (SUSHA <i>et al.</i>, 2015a)</p>	<p>1.7- Verifica-se regularmente se as políticas e/ou processos de mitigação de riscos estão ocorrendo conforme esperado?</p> <p>1.8- Verifica-se regularmente se as políticas e/ou processos de disponibilização de conjuntos de dados estão ocorrendo conforme esperado?</p>
	<p>Determinar que tipo de dados são importantes para tratar de questões sociais e se concentre na publicação destes (SUSHA <i>et al.</i>, 2015a)</p>	<p>1.9- Verifica-se regularmente se as políticas e/ou processos sobre a conversão de formato de dados estão ocorrendo conforme esperado?</p>
	<p>Selecione partes de um conjunto de dados que ofereça benefícios (HYLAND <i>et al.</i>, 2014), (HYLAND; WOOD, 2011 apud BROEK <i>et al.</i>, 2014), (FERRARA <i>et al.</i>, 2012 apud BROEK <i>et al.</i>, 2014).</p>	<p>1.10- Estas políticas e/ou processos de mitigação de riscos estão sendo mantidas e/ou aprimoradas?</p> <p>1.11- Estas políticas e/ou processos de disponibilização de conjuntos de dados estão sendo mantidas e/ou aprimoradas?</p> <p>1.12- Estas políticas e/ou processos de conversão de dados estão sendo mantidas e/ou aprimoradas?</p>

Pré-requisitos:	Capacidade
P1.1 e P1.2 e P1.3	A - Definição
P1.4 e P1.5 e P1.6	B - Implementação
P1.7 e P1.8 e P1.9	C - Monitoramento
P1.10 e P1.11 e P1.12	D - Aprimoramento

Em cada Quadro destaca-se em **negrito** o(s) termo(s) ou frase(s) que estão contemplados diretamente nas perguntas apresentadas.

Fonte: Da Autora

O Quadro 23 apresenta o contexto considerado para o componente acesso e entrega de dados. Destaca-se que as organizações utilizam de infraestruturas próprias, ou ainda fazem uso de infraestrutura de terceiros para disponibilizar seus dados⁶.

Quadro 23 – Acesso e Entrega de Dados

Componente: 2 - Acesso e entrega de dados		
Tem por objetivo verificar se as políticas e processos relacionados às etapas de publicação e de infraestrutura de disponibilização dos dados foram definidas, implementadas, monitoradas e mantidas e/ou aprimoradas.		
Fatores Importantes	Boas práticas	Perguntas
Definir etapas claras do processo para abertura de dados	Garantir que as experiências sejam compartilhadas durante o ciclo de vida de um conjunto de dados (DODDS; NEWMAN, 2015).	2.1- Há etapas definidas para a disponibilização dos dados?
	O processo de publicação deverá ser bem documentado (DODDS; NEWMAN, 2015)	2.2- As etapas de disponibilização de dados são bem documentadas? 2.3- Sabe-se onde os dados serão publicados?
Determinar onde os conjuntos de dados serão publicados	Apoiar a liberação; por exemplo com uma plataforma específica ou portal de dados específico (DODDS; NEWMAN, 2015; HYLAND; WOOD, 2011; HAUSENBLAS, 2011; JANSSEN; ZUIDERWIJK, 2012; VILLAZÓN-TERRAZAS <i>et al.</i> , 2011)	2.4- As etapas para disponibilização dos dados estão sendo executadas?

⁶ Optou-se por incluir apenas o contexto de infraestrutura visto que apenas organizações evoluídas na abertura de dados teriam condições de manter uma infraestrutura em nuvem.

Usar infraestruturas de nuvem capazes de coletar, gerenciar e publicar dados abertos, interoperáveis com outras fontes.		2.5 - Os dados estão sendo disponibilizados no local definido? 2.6- Verifica-se regularmente se as etapas de disponibilização dos dados estão ocorrendo conforme se espera? 2.7- Verifica-se regularmente se os dados estão sendo disponibilizados no local definido?
Garantir a continuidade do fornecimento de dados , incluindo atualizações oportunas de dados.	Disponibilize os dados de maneira atualizada (LÓSCIO <i>et al.</i> , 2017) Quando os dados são produzidos em tempo real, disponibilizá-los na Web preferencialmente em tempo real (LÓSCIO <i>et al.</i> , 2017; SOLAR <i>et al.</i> , 2012)	2.8- As etapas de disponibilização de dados estão sendo mantidas e/ou aprimoradas ? 2.9 - A infraestrutura permite a continuidade do fornecimento, de conjunto de dados e recursos?
Pré-requisitos:		Capacidade
P2.1 e P2.2 e P2.3		A - Definição
P2.4 e P2.5		B - Implementação
P2.6 e P2.7		C - Monitoramento
P2.8 e P2.9		D - Aprimoramento

Fonte: Da Autora

Considerando as políticas e/ou processos o Quadro 24 apresenta o contexto considerado para o componente monitoramento e medição.

Quadro 24 – Monitoramento e Medição

Componente: 3 - Monitoramento e Medição		
Tem por objetivo verificar se as políticas e processos acerca do monitoramento e da medição dos processos relacionados a dados abertos foram definidas, implementadas, monitoradas e mantidas e/ou aprimoradas.		
Fatores Importantes	Boas práticas	Perguntas
Organizar reuniões internas para discutir os processos de publicação de dados e avaliá-los		3.1-Existem reuniões internas para discutir como avaliar o processo de disponibilização de dados? 3.2- Foram desenvolvidas métricas ou indicadores de sucesso para disponibilização de dados?
Desenvolver métricas e indicadores de sucesso para publicação de dados por departamentos governamentais		3.3- Já foi realizada alguma avaliação com base nas métricas e indicadores evidenciados? 3.4- Verifica-se regularmente se a avaliação através da métricas e indicadores desenvolvidos é realizada com frequência?
Avaliar a realização de métricas e indicadores de sucesso como parte integrante da iniciativa de dados abertos.	Definir como medir seu progresso em relação a sua estratégia (DODDS; NEWMAN, 2015; BROEK <i>et al.</i> , 2014).	3.5- As métricas e indicadores desenvolvidos são mantidos e/ou aprimorados?
Pré-requisitos:		Capacidade
P3.1 e P3.2		A - Definição
P3.3		B - Implementação
P3.4		C - Monitoramento
P3.5		D - Aprimoramento

Fonte: Da Autora

Por fim, para políticas e/ou processos apresenta-se o Quadro 25 que aborda o contexto considerado para o componente gerenciamento de mudança de dados.

Quadro 25 – Gerenciamento de Mudança de Dados

Componente: 4 - Gerenciamento de mudança de dados		
Tem por objetivo verificar se as políticas e processos sobre alterações e atualizações de conjunto de dados foram definidas, implementadas, monitoradas e mantidas e/ou aprimoradas.		
Fatores Importantes	Boas práticas	Perguntas
Definir etapas claras do processo para publicar dados	O processo de publicação deverá abordar componentes técnicos da publicação de conjunto de dados novos e da atualização de conjuntos de dados já existentes em tempo oportuno (DODDS; NEWMAN, 2015), (JANSSEN; ZUIDERWIJK, 2012 apud BROEK <i>et al.</i> , 2014)	4.1 – Existem políticas ou processos que viabilizam o gerenciamento de mudanças, referentes aos conjuntos de dados e recursos? 4.2 – O processo de disponibilização de dados possibilita a alteração dos conjuntos de dados e recursos? 4.3 - Existem políticas ou processos que determinam a frequência de atualização dos conjuntos de dados?
Identificar a necessidade dos dados	Verificar procedimentos para mitigar riscos antes da disponibilização, considerando anonimizar, editar e agregar os dados para remover informações sensíveis (DODDS; NEWMAN, 2015), (JANSSEN; ZUIDERWIJK, 2012 apud BROEK <i>et al.</i> , 2014).	4.4- As políticas e/ou os processos estão sendo executados referentes a mudança de dados? 4.5- A infraestrutura possibilita a atualização oportuna e automática de conjunto de dados e de recursos? 4.6– Está sendo explicitada a frequência de atualização dos conjuntos de dados?
Ser transparente com os usuários de dados abertos sobre as condições sob as quais a publicação de dados ocorre	Fornecer histórico de versões que explique as alterações feitas em cada versão (LÓSCIO <i>et al.</i> , 2017)	4.7- Verifica-se regularmente se as políticas e/ou processos de mudança de dados estão ocorrendo conforme se espera?
	Torne a frequência de atualização explícita (LÓSCIO <i>et al.</i> , 2017)	4.8– Verifica-se regularmente se a atualização dos dados está ocorrendo conforme a frequência estipulada?
Garantir a continuidade do fornecimento de dados, incluindo atualizações oportunas e automáticas de dados.	Disponibilize os dados de maneira atualizada (LÓSCIO <i>et al.</i> , 2017).	4.9- Estas políticas e/ou processos sobre a mudança de conjunto de dados e recursos estão sendo mantidas e/ou aprimoradas?
	Quando os dados são produzidos em tempo real, disponibilizá-los na Web em tempo real ou quase em tempo real (LÓSCIO <i>et al.</i> , 2017; SOLAR <i>et al.</i> , 2012)	4.10- Estas políticas e/ou processos referentes à frequência de atualização estão sendo mantidas e/ou aprimoradas?

Pré-requisitos:	Capacidade
P4.1, P4.2 e P4.3	A - Definição
P4.4, P4.5 e P4.6	B - Implementação
P4.7 e P4.8	C - Monitoramento
P4.9 e P4.10	D - Aprimoramento

Fonte: Da Autora

• Gerenciamento de Fatores Organizacionais

Nesta área foco serão apresentados os componentes: I) funções e responsabilidades, II) educação, III) treinamento, IV) planejamento e priorização da abertura dos dados, V) planejamento da publicação dos dados e VI) reuso.

Em cada quadro será exibido o objetivo de cada componente, as perguntas elencadas para a avaliação das capacidades, as boas práticas e fatores importantes que auxiliaram na composição das perguntas, as capacidades e os pré - requisitos, para que determinada capacidade seja alcançada e serão destacados em negrito o(s) termo(s) ou frase(s) que estão contemplados diretamente nas perguntas apresentadas.

Ter funções ou responsabilidades definidas auxilia na atribuição de atividades específicas e responsabilidades individuais, possibilitando a conferência das atividades executadas (PANIAN, 2010). O Quadro 26 apresenta o contexto considerado para o componente funções e responsabilidades.

Quadro 26 – 5 - Funções e Responsabilidades

Componente: Funções e Responsabilidades		
Tem por objetivo verificar se as funções e responsabilidades foram definidas, implementadas, monitoradas e mantidas e/ou aprimoradas.		
Fatores Importantes	Boas práticas	Perguntas
Determinar quais pessoas tem as principais responsabilidades na publicação de dados abertos	A organização irá identificar as funções e responsabilidades daqueles envolvidos na publicação de dados abertos em nome da organização (DODDS; NEWMAN, 2015)	5.1– Existem papéis ou funções para os <i>stakeholders</i> envolvidos na abertura dos dados? 5.2– Está claro o papel ou função que cada <i>stakeholder</i> exerce no processo de abertura dos dados? 5.3 – Estão sendo executados os papéis ou funções definidos? 5.4 – Verifica-se regularmente se as atividades são desenvolvidas adequadamente por cada papel ou função? 5.5 – Os papéis ou funções estão sendo atualizados à medida que o volume de dados a serem disponibilizados aumentam ou quando observam-se novas necessidades?
Pré-requisitos:		Capacidade
P5.1 e P5.2		A - Definição
P5.3		B - Implementação
P5.4		C - Monitoramento
P5.5		D - Aprimoramento

Fonte: Da Autora

Possuir compreensão sobre os benefícios e as estratégias para dados abertos possibilita que os responsáveis pela abertura dos dados tenham maior engajamento (PANIAN, 2010). O Quadro 27 apresenta o contexto considerado para o componente educação.

Quadro 27 – Educação

Componente: 6 - Educação		
Tem por objetivo verificar se o incentivo à educação interna à organização foi definido, implementado, monitorado e mantido e/ou aprimorado.		
Fatores Importantes	Boas práticas	Perguntas
Organizar grupos focais com chefes de departamentos e implementadores de políticas de dados abertos para que proponentes e oponentes de dados abertos possam dialogar.	Compreensão estratégica, de como utilizar os dados abertos para promover os objetivos da organização (DODDS; NEWMAN, 2015), (HAUSENBLAS, 2011 apud BROEK <i>et al.</i> , 2014).	6.1– Existe o alinhamento de dados abertos à estratégia organizacional? 6.2- Fomenta-se a criação de um consenso sobre o valor que será criado para a organização através da disponibilização dos dados? 6.3- Verifica-se regularmente se há compreensão estratégica do uso de dados abertos? 6.4– Há campanhas de informação na organização, sobre histórias de sucesso no uso de dados abertos?
Criar consenso entre a publicação de dados abertos e a estrutura organizacional para publicação de dados		
Introduzir esquemas de incentivos para servidores públicos (por exemplo, explique por que um provedor de dados disponibilizaria dados, explique que tipo de valor é criado para o provedor de dados)		
Desenvolver campanhas de informação nas quais sejam discutidas histórias de sucesso do uso interno e externo de dados abertos		
Pré-requisitos:		Capacidade
P6.1		A - Definição
P6.2		B - Implementação
P6.3		C - Monitoramento
P6.4		D - Aprimoramento

Fonte: Da Autora

Fornecer treinamento para os envolvidos na abertura dos dados viabiliza a compreensão sobre os diferentes aspectos que devem ser observados, sobre boas práticas e sobre os padrões a serem utilizados. Neste sentido, o componente treinamento abordado no Quadro 28 mostra o contexto considerado para este.

Quadro 28 – Treinamento

Componente: 7 - Treinamento		
Tem por objetivo verificar se foram definidos, implementados, monitorados, mantidos e/ou aprimorados treinamentos acerca das práticas e padrões adotados para a publicação dos dados.		
Fatores Importantes	Boas práticas	Perguntas
Criar um centro de competência virtual que ajude a responder perguntas e a ajudar nos processos administrativos de publicação de dados	A organização irá assegurar que todos os padrões organizacionais, conjuntos de dados e documentações de dados relevantes sejam acessíveis às equipes que os necessitam (DODDS; NEWMAN, 2015).	7.1 – Existem treinamentos, por exemplo, sobre padrões técnicos, boas práticas e questões sobre licenciamento de dados abertos, entre outros tópicos? 7.2 – Existe um centro de competência que viabilize a recuperação de informações acerca dos processos de abertura de dados? 7.3 – Os treinamentos são acessíveis a todos os <i>stakeholders</i> envolvidos no processo de abertura de dados? 7.4- Verifica-se regularmente se os treinamentos, estão ocorrendo conforme se espera? 7.5 – Os treinamentos estão alinhados com as necessidades dos responsáveis pela abertura dos dados? 7.6 – Os treinamentos e o centro de competências estão sendo mantidos e atualizados?
	Os responsáveis pela abertura dos dados precisam ser capazes de encontrar e utilizar a documentação sobre padrões e políticas organizacionais (DODDS; NEWMAN, 2015).	
Fornecer treinamento sobre publicação de dados abertos em agências governamentais (por exemplo, treinamento sobre como os conjuntos de dados podem ser anonimizados)	A organização dará suporte para os funcionários desenvolverem as competências, como por exemplo, padrões técnicos, boas práticas e questões sobre licenciamento, necessárias para lidar com sua estratégia de dados abertos (DODDS; NEWMAN, 2015).	
	Compreender os riscos, responsabilidades e ações necessárias para garantir que os dados sejam bem administrados e compartilhados efetivamente (DODDS; NEWMAN, 2015), (HAUSENBLAS, 2011 apud BROEK <i>et al.</i> , 2014)	
Desenvolver campanhas de informação nas quais são discutidas questões sobre publicação de dados abertos	Treinamentos especificam em tópicos individuais como, por exemplo, padrões técnicos, boas práticas e questões sobre licenciamento (DODDS; NEWMAN, 2015).	
Pré- requisitos:		Capacidade
P7.1 e P7.2		A - Definição

P7.3	B - Implementação
P7.4	C - Monitoramento
P7.5 e P7.6	D - Aprimoramento

Fonte: Da Autora

Planejar e priorizar a abertura dos dados de modo geral auxilia a organizar, priorizar, definir e traçar metas a curto, médio e longo prazo. Assim é possível que a organização observe o quanto evolui com o passar do tempo em relação ao seu planejamento. Deste modo, o componente planejamento e priorização da abertura dos dados apresentado no Quadro 29 aborda a necessidade do plano de dados abertos em relação as capacidades.

Quadro 29 – Planejamento e Priorização da Abertura dos Dados

Componente: 8 - Planejamento e Priorização da Abertura dos Dados		
Tem por objetivo verificar se o planejamento e a priorização dos dados abertos foram definidos, implementados, monitorados, mantidos e/ou aprimorados.		
Fatores Importantes	Boas práticas	Perguntas
Desenvolver uma estratégia para publicação de dados abertos em nível (inter)nacional		8.1– Existe um plano de dados abertos? 8.2- O plano de dados abertos está sendo executado?
Garantir que a gerência (superior) das agências governamentais ofereça suporte à publicação de dados	A organização irá assegurar que a adoção de tais dados abertos esteja rigorosamente ajustada a seus objetivos organizacionais mais amplos (DODDS; NEWMAN, 2015)	8.3 - Verifica-se regularmente se o plano de dados abertos está sendo executado conforme o esperado? 8.4 - Existe a colaboração (inter) organizacional acerca de compartilhamento de experiências com o intuito de melhorar o planejamento e a priorização dos dados abertos?
Garantir uma cooperação ágil e aberta com várias outras organizações (administração, universidades, CSO, OpenKnowledge Foundation)		

Organize a colaboração interorganizacional (por exemplo, reuniões da rede) para aprender com as iniciativas de dados abertos de outras agências governamentais.		8.5- O plano de dados abertos está sendo atualizado ou aprimorado?
Determinar quais padrões e vocabulários será usado para publicação de dados	Faz-se necessário a adoção de boas práticas que garantam que os dados publicados possam ser facilmente reutilizados por terceiros (DODDS; NEWMAN, 2015).	
Pré- requisitos:		Capacidade
P2.3, P3.1, P5.1, P7.1 e P8.1		A - Definição
P2.5, P3.2, P5.3, P7.3 e P8.2		B - Implementação
P2.7, P3.4, P5.4, P7.4 e P8.3		C - Monitoramento
P8.4 e 8.5		D - Aprimoramento

Fonte: Da Autora

Ter um planejamento para publicar dados viabiliza que as práticas e tratamentos de dados sejam repetidos para cada novo recurso ou conjunto de dados a ser disponibilizado, além de priorizar conjuntos interessantes para uso possibilita que os usuários dos dados visualizem os benefícios. Deste modo, o componente planejamento da publicação de dados apresentado no Quadro 30 aborda a necessidade do planejamento e da priorização da publicação dos dados.

Quadro 30 – Planejamento da Publicação dos Dados

Componente: 9 - Planejamento da Publicação dos Dados		
Tem por objetivo verificar se o planejamento acerca dos dados a serem disponibilizados foi definido, implementado, monitorado, mantido e/ou aprimorado.		
Fatores Importantes	Boas práticas	Perguntas
Desenvolver uma estratégia para publicação de dados abertos em nível (inter)nacional		9.1– Existe um planejamento ou uma priorização acerca dos dados a serem disponibilizados?

Garantir que a gerência (superior) das agências governamentais ofereça suporte à publicação de dados	A organização irá assegurar que a adoção de tais dados abertos esteja rigorosamente ajustada a seus objetivos organizacionais mais amplos (DODDS; NEWMAN, 2015).	9.2 - O planejamento ou a priorização acerca dos dados a serem disponibilizados está sendo executada? 9.3 - Verifica-se regularmente se o planejamento ou a priorização definidos acerca dos dados a serem disponibilizados esta sendo executada conforme esperado?
Definir etapas de processo claras para a publicação de dados		9.4 - O planejamento ou a priorização dos dados a serem disponibilizados esta sendo atualizada ou aprimorada?
Começar com a publicação de dados que sejam interessantes para os usuários, para que eles vejam os benefícios dos dados abertos		
Determinar que tipo de dados são importantes para abordar questões sociais e se concentrar na publicação desses dados		
Determinar quais padrões e vocabulários serão usados para publicação de dados	Faz-se necessária a adoção de boas práticas que garantam que os dados publicados possam ser facilmente reutilizados por terceiros (DODDS; NEWMAN, 2015).	
Pré-requisitos:		Capacidade
P1.3, P2.3, P4.3, P8.1 e P9.1		A - Definição
P1.6, P2.5, P4.6, P8.2 e P9.2		B - Implementação
P1.9, P2.7, P4.8, P8.3 e P9.3		C - Monitoramento
P9.4		D - Aprimoramento

Fonte: Da Autora

Através do fomento ao uso dos dados as organizações podem se envolver com os usuários de dados abertos e compreender as suas necessidades, permitindo obterem *feedbacks* sobre a qualidade dos dados que estão sendo disponibilizados e estimularem o desenvolvimento de novos produtos e/ou serviços com base nos dados. Para isso, o Quadro 31 aborda o contexto considerado para o componente fomento ao uso.

Quadro 31 – Fomento ao Uso

Componente: 10 - Fomento ao Uso		
Tem por objetivo verificar se a organização tem definido, implementado, monitorado e mantido ou aprimorado o fomento ao uso dos dados.		
Fatores Importantes	Boas práticas	Perguntas
Fornecer exemplos prontamente disponíveis de uso de dados abertos (por exemplo, aplicativos) a não especialistas.		P10.1- A organização possui uma estratégia definida para fomentar o uso dos dados disponibilizados? P10.2- A organização promove a divulgação de histórias de sucesso sobre o uso dos dados ou fomenta a realização de eventos e competições que visem incentivar o uso de dados abertos?
Desenvolver histórias de uso bem-sucedido de dados abertos	Criar e compartilhar estudos de caso que destaquem o impacto e os benefícios de trabalhar com dados abertos (DODDS; NEWMAN, 2015; HAUSENBLAS, 2011; BROEK <i>et al.</i> , 2014).	P10.3- A organização envolve-se regularmente com os usuários de dados com o intuito de descobrir suas necessidades e verificar como seus dados estão sendo utilizados?
Envolver os principais atores da comunidade para propagar histórias de sucesso		
Fornecer esquemas de incentivo para envolver os cidadãos no uso de dados abertos		
Estimular o desenvolvimento de incubadoras especializadas, iniciadas por dados abertos.		P10.4- A organização está mantendo ou aprimorando a sua estratégia com relação ao fomento ao uso dos dados?
Apoiar a construção de comunidades orientadas para questões através de eventos participativos		
Alinhar eventos, competições e hackathons com, por exemplo, currículos universitários, prêmios, festivais e "marketing direto"	Promover eventos tais como <i>hackdays</i> ou desafios para encorajar e incentivar o uso de dados e para gerar estudos de casos para um maior compartilhamento (DODDS; NEWMAN, 2015; BROEK <i>et al.</i> , 2014; SOLAR <i>et al.</i> , 2012).	

<p>Organizar reuniões com usuários de dados abertos para descobrir quais são suas necessidades e como os dados da agência governamental são usados</p>	<p>Uma organização irá necessitar planejar como irá abordar seus clientes, monitorar como seus dados estão sendo utilizados e identificar onde o valor está sendo criado (DODDS; NEWMAN, 2015; BROEK <i>et al.</i>, 2014; SOLAR <i>et al.</i>, 2012).</p> <p>Pesquisar formas em que os reutilizadores gostariam de processar dados (DODDS; NEWMAN, 2015).</p> <p>Envolver-se com a comunidade para ajudá-la a entender como um conjunto de dados pode ser utilizado e para destacar novas liberações de dados relevantes (DODDS; NEWMAN, 2015; BROEK <i>et al.</i>, 2014).</p>	
<p>Garantir serviços de suporte ágeis, dinâmicos e profissionais e treinamento para possíveis usuários dos dados.</p>		
<p>Organizar eventos e garantir a construção da comunidade onde os benefícios potenciais dos dados abertos são comunicados aos usuários (por exemplo, criando cenários para uso).</p>		
<p>Desenvolver currículos universitários e de educação continuada com dados abertos</p>		
<p>Desenvolver e manter roteiros das áreas de pesquisa sobre dados abertos, a fim de consolidar os esforços de pesquisa.</p>		
<p>Pré-requisitos:</p>	<p>Capacidade</p>	
<p>P10.1</p>	<p>A - Definição</p>	

P10.2	B - Implementação
P10.3	C - Monitoramento
P10.4	D - Aprimoramento

Fonte: Da Autora

• Gerenciamento de Padrões

A área foco definida como gerenciamento de padrões aborda os componentes: I) Licença, II) Proveniência e Qualidade, III) Versionamento, IV) Identificação, V) Formato, VI) Vocabulários, VII) Metadados, VIII) Preservação, IX) Enriquecimento, X) *Feedback* e XI) Acesso, sendo estes apresentados abaixo, seguindo os mesmos padrões de apresentação já observados previamente. Observa-se nesta área foco em específico que a capacidade de "Aprimoramento" tem por objetivo elencar questões que colaborem com os padrões que visam à legibilidade por máquina.

Tem-se como primeiro componente a licença que informa o que pode e não pode ser feito com os dados, protege legalmente os dados e possibilita controlar o seu uso. Quanto menos restrições, a licença viabiliza maior a tendência de uso dos dados, assim também viabiliza-se a confiança (LÓSCIO *et al.*, 2017). Neste sentido definir quais licenças serão aplicadas sobre os dados e disponibilizar a informação tanto para humanos quanto para máquinas de qual licença está associada a um determinado conjunto de dados é essencial para o contexto de dados abertos. O Quadro 32 aborda o contexto considerado para o componente licença.

Quadro 32 – Licença

Componente: 11 - Licença		
Tem por objetivo verificar se os padrões sobre licença foram definidos, se estão sendo utilizados, monitorados e aprimorados.		
Fatores Importantes	Boas práticas	Perguntas
Fornecer informações sobre os significados e implicações das licenças	Fornecer um link para o contrato de licença que controla o uso dos dados (LÓSCIO <i>et al.</i> , 2017) (FERRARA <i>et al.</i> , 2012 apud BROEK <i>et al.</i> , 2014)).	P11.1- Há uma definição sobre qual licença será aplicada sobre os conjuntos de dados disponibilizados? P11.2- Faz-se o uso de uma licença aberta? P11.3- São disponibilizadas informações sobre as implicações da licença utilizada?

Fornecer informações sobre a legislação de privacidade e como os dados abertos podem ser usados em conformidade com esta legislação.		P11.4– Verifica-se regularmente se as informações sobre a licença estão sendo aplicadas sobre os conjuntos de dados, conforme o esperado? P11.5- Existem metadados sobre a licença utilizada nos conjuntos de dados disponibilizados? P11.6- Faz-se o uso de vocabulários para a licença?
Pré-requisitos:		Capacidade
P11.1		A - Definição
P11.2 e P11.3		B - Implementação
P11.4		C - Monitoramento
P11.5 e P11.6		D - Aprimoramento

Fonte: Da Autora

O segundo componente é denominado de proveniência e qualidade e visa informar sobre a origem, sobre quem são os responsáveis pela publicação, entre outras informações de rastreabilidade, já a qualidade informa diferentes dimensões de análise. Ter informações tanto para humanos, quanto para máquinas sobre a proveniência e sobre qualidade contribuem para que haja confiança e uso dados disponibilizados (LÓSCIO *et al.*, 2017), pois os usuários dos dados necessitam de meios para confiar na integridade dos dados que são disponibilizados. O Quadro 33 aborda o contexto considerado para o componente proveniência e qualidade.

Quadro 33 – Proveniência e Qualidade

Componente: 12 - Proveniência e Qualidade		
Tem por objetivo verificar se os padrões sobre proveniência e qualidade foram definidos, se estão sendo utilizados, monitorados e aprimorados.		
Fatores Importantes	Boas práticas	Perguntas
Ser transparente com os usuários de dados abertos sobre as condições que ocorrem à publicação de dados	Fornecer informações completas sobre a origem (proveniência) dos dados e quaisquer alterações feitas (LÓSCIO <i>et al.</i> , 2017)	P12.1- Há uma definição sobre quais informações de proveniência (responsáveis, origem, entre outras) serão disponibilizadas? P12.2 - As informações sobre a proveniência dos dados estão sendo disponibilizadas? P12.3 – Verifica-se regularmente se as informações sobre proveniência estão sendo disponibilizadas junto aos conjuntos de dados, conforme o esperado?
Integrar estruturas para avaliar a qualidade e a usabilidade dos dados e a plataforma, fornecendo <i>feedback</i> contínuo aos desenvolvedores e administrações.	Desenvolver meios para reportar erros, de qualidade de dados ou questões de privacidade de um conjunto de dados (DODDS; NEWMAN, 2015; SOLAR <i>et al.</i> , 2012)	P12.4- São disponibilizados metadados sobre a proveniência junto aos conjuntos de dados? P12.5 - Faz-se o uso de vocabulários para a proveniência? P12.6- Faz-se o uso de vocabulários para a qualidade?
Pré-requisitos:		Capacidade
P12.1		A - Definição
P12.2		B - Implementação
P12.3		C - Monitoramento
P12.4, P12.5, P12.6		D - Aprimoramento

Fonte: Da Autora

Os conjuntos de dados podem ter alteração e até mesmo sofrer mudanças ao longo do tempo. Para fins de histórico e com o intuito de permitir acompanhar a evolução de determinado conjunto de dados faz-se necessário ter informações legíveis por humanos e por máquinas sobre os versionamentos isso viabilizará confiança por parte dos usuários e conseqüentemente o seu uso (LÓSCIO *et al.*, 2017). Neste sentido, apresenta-se o terceiro componente denominado de versionamento, sendo que o Quadro 34 aborda o contexto considerado para este componente.

Quadro 34 – Versionamento

Componente: 13 - Versionamento		
Tem por objetivo verificar se os padrões sobre versionamento foram definidos, se estão sendo utilizados, monitorados e aprimorados.		
Fatores Importantes	Boas práticas	Perguntas
Seja transparente com os usuários de dados abertos sobre as condições sob as quais a publicação de dados ocorre	Fornecer histórico de versões que explique as alterações feitas em cada versão (LÓSCIO <i>et al.</i> , 2017)	<p>P13.1- Existe uma definição sobre as informações do histórico de versionamento a ser disponibilizado?</p> <p>P13.2- As informações acerca do histórico de versionamento estão sendo disponibilizadas?</p> <p>P13.3– Verifica-se regularmente se as informações sobre o histórico de versionamento está sendo disponibilizado, conforme o esperado?</p> <p>P13.4- Existem metadados referentes ao versionamento de um determinado conjunto de dados?</p> <p>P13.5- Faz-se o uso de vocabulários para o versionamento?</p>
Pré-requisitos:		Capacidade
P13.1		A - Definição
P13.2		B - Implementação
P13.3		C - Monitoramento
P13.4 e P13.5		D - Aprimoramento

Fonte: Da Autora

A utilização de URIs HTTP(ou HTTPS) únicas para a identificação de conjuntos de dados, versões, séries de conjuntos de dados e recursos permite que os usuários realizem o acesso aos dados disponibilizados na Web, conseqüentemente viabiliza-se o uso (LÓSCIO *et al.*, 2017). Deste modo o componente identificação apresentado no Quadro 35 aborda o contexto considerado para este componente.

Quadro 35 – Identificação

Componente: 14 - Identificação		
Tem por objetivo verificar se os padrões acerca da identificação dos conjuntos de dados foram definidos, se estão sendo utilizados, monitorados e aprimorados.		
Fatores Importantes	Boas práticas	Perguntas
Usar padrões para dados, metadados, licenças, URIs e protocolos de troca.	Usar URIs persistentes como identificadores de conjuntos de dados , dentro de conjunto de dados e a versões e séries de conjuntos de dados (LÓSCIO <i>et al.</i> , 2017), (FERRARA <i>et al.</i> , 2012 apud BROEK <i>et al.</i> , 2014), (HYLAND; WOOD, 2011 apud BROEK <i>et al.</i> , 2014)	P14.1- Existe uma definição sobre como serão identificados os conjuntos de dados? P14.2- As URIs utilizadas são padronizadas? P14.3- As URIs utilizadas são únicas? P14.4- Verifica-se regularmente se os padrões definidos para a identificação estão sendo aplicados conforme o esperado? P14.5- Os padrões definidos para URIs estão sendo aplicados nos conjuntos de dados e em versões e séries de conjuntos de dados?
Pré-requisitos:		Capacidade
P14.1		A - Definição
P14.2 e P14.3		B - Implementação
P14.4		C - Monitoramento
P14.5		D - Aprimoramento

Fonte: Da Autora

O formato em que os dados são disponibilizados é um fator importante para a reutilização, para isso há um consenso sobre a necessidade de disponibilizar dados em formatos legíveis por máquina. Sendo que a classificação de cinco estrelas proposta por (BERNERS-LEE, 2011), foi considerada para definir o grau de legibilidade possibilitado pelos formatos. Assim observa-se que a variedade de formatos para um mesmo conjunto de dados possibilita a facilidade de processamento por parte dos mais diversos usuários de dados. Para isso, o Quadro 36 aborda o contexto considerado para o componente formato.

Quadro 36 – Formato

Componente: 15 - Formato		
Tem por objetivo verificar se os padrões acerca do formato relacionados aos conjuntos de dados foram definidos, se estão sendo utilizados, monitorados e aprimorados.		
Fatores Importantes	Boas práticas	Perguntas
Determinar quais padrões e vocabulários que serão usados para publicação de dados	Opte por formatos de dados padronizados legíveis por máquina (DODDS; NEWMAN, 2015; LÓSCIO <i>et al.</i> , 2017), (JANSSEN; ZUIDERWIJK, 2012 apud BROEK <i>et al.</i> , 2014).	P15.1- Existe uma definição sobre quais formatos serão utilizados para a disponibilização de recursos atrelados a um conjunto de dados? P15.2- Faz-se o uso de um formato estruturado (por exemplo, csv, xlsx)?
Identificar a necessidade dos dados	Converter dados para formato legível por máquina , previamente definido (VILLAZÓN-TERRAZAS <i>et al.</i> , 2011; HYLAND; WOOD, 2011; BROEK <i>et al.</i> , 2014). Disponibilize os dados em vários formatos e fornecer subconjuntos para conjuntos de dados grandes, bem como enriquecer os dados disponibilizados combinando com outros dados adequando ao uso pretendido ou potencial (LÓSCIO <i>et al.</i> , 2017; SOLAR <i>et al.</i> , 2012)	P15.3- Verifica-se regularmente se o formato de dados utilizado é estruturado e legível por máquina? P15.4- A infraestrutura para disponibilização dos dados permite publicar dados abertos, interoperáveis com outras fontes dentro do país ou região? P15.5 - Faz-se o uso de padrões de formatos de dados semânticos, por exemplo, vocabulários, RDF? P15.6- O formato disponibilizado permite conectar dados a outras fontes (dados conectados)?
Usar padrões para dados, metadados, licenças, URIs e protocolos de troca	Use formatos de dados padronizados legíveis por máquina (DODDS; NEWMAN, 2015; LÓSCIO <i>et al.</i> , 2017), (JANSSEN; ZUIDERWIJK, 2012 apud BROEK <i>et al.</i> , 2014)	P15.7 - São disponibilizados recursos em mais de um formato, isso para um mesmo conjunto de dados?
Pré-requisitos:		Capacidade
P1.3 e P15.1		A - Definição
P15.2		B - Implementação
P15.3		C - Monitoramento
P15.4, P15.5, P15.6 e P15.7		D - Aprimoramento

Fonte: Da Autora

Os vocabulários definem conceitos e as relações existentes(ou não) entre estes,

normalmente utilizados para descreverem ou representarem uma área. Já os vocabulários controlados são criados por organizações de referência com objetivo de padronizar os vocabulários a fim de permitir o uso por qualquer indivíduo. Assim possibilita explicitar os conceitos e definições existentes para determinada área, contribuindo com a semântica dos dados. Deste modo, o Quadro 37 apresenta o contexto considerado para o componente vocabulários⁷.

Quadro 37 – Vocabulários

Componente:16 - Vocabulários		
Tem por objetivo verificar se os padrões acerca dos vocabulários foram definidos, se estão sendo utilizados, monitorados e aprimorados.		
Fatores Importantes	Boas práticas	Perguntas
Determinar quais padrões e vocabulários que serão usados para publicação de dados	Opte por formatos de dados padronizados legíveis por máquina (DODDS; NEWMAN, 2015; LÓSCIO <i>et al.</i> , 2017), (JANSSEN; ZUIDERWIJK, 2012 apud BROEK <i>et al.</i> , 2014).	P16.1- Existe uma definição sobre quais vocabulários serão utilizados? P16.2- Faz-se o uso de vocabulários controlados já existentes? P16.3- Verifica-se regularmente se os vocabulários definidos estão sendo utilizados, conforme o esperado?
Integrar esquemas de metadados e vocabulários controlados federados para categorizar adequadamente as informações	Use termos de vocabulários compartilhados, preferencialmente padronizados, para codificar dados e metadados (HAUSENBLAS, 2011; VILLAZÓN-TERRAZAS <i>et al.</i> , 2011; HYLAND; WOOD, 2011; LÓSCIO <i>et al.</i> , 2017).	P16.4- São utilizados vários tipos de vocabulários (por exemplo, DC, DCAT, DQV, PROV)?
Pré-requisitos:		Capacidade
P16.1		A - Definição
P16.2		B - Implementação
P16.3		C - Monitoramento
P16.4, P11.6*, P12.5*, P12.6* e P13.5*		D - Aprimoramento

Fonte: Da Autora

Os metadados viabilizam informações para que humanos e máquinas possam encontrar, compreender e fazer uso dos conjuntos de dados e recursos. Assim

⁷ As perguntas sinalizadas com * complementam a capacidade de aprimoramento deste componente.

há metadados que fornecem informações sobre versões, proveniência e licenças, por exemplo. Sendo que estes metadados contribuem com a legibilidade por máquina de componentes em específico, mas também complementam o componente metadados, visto que quanto maior a diversidade de metadados aumenta-se a compreensão, a legibilidade e conseqüentemente melhoram-se as chances de uso e reuso dos dados (LÓSCIO *et al.*, 2017). O Quadro 38 explicita o contexto considerado para o componente metadados⁸.

Quadro 38 – Metadados

Componente: 17 - Metadados		
Tem por objetivo verificar se os padrões para metadados foram definidos, se estão sendo utilizados, monitorados e aprimorados.		
Fatores Importantes	Boas práticas	Perguntas
Determine quais metadados serão utilizados		P17.1- Existe uma definição sobre quais metadados serão utilizados?
Usar padrões para dados, metadados , licenças, URIs e protocolos de troca	Use termos de vocabulários compartilhados, preferencialmente padronizados, para codificar dados e metadados (HAUSENBLAS, 2011; VILLAZÓN-TERRAZAS <i>et al.</i> , 2011; HYLAND; WOOD, 2011; LÓSCIO <i>et al.</i> , 2017)	P17.2- São disponibilizados metadados legíveis por humanos?
	Usar URIs persistentes como identificadores de conjuntos de dados, dentro de conjunto de dados e a versões e séries de conjuntos de dados (LÓSCIO <i>et al.</i> , 2017), (FERRARA <i>et al.</i> , 2012 apud BROEK <i>et al.</i> , 2014), (HYLAND; WOOD, 2011 apud BROEK <i>et al.</i> , 2014).	P17.3- São disponibilizados metadados legíveis por máquinas?
	Use estruturas e valores de dados com neutralidade de localidade ou, quando isso não for possível, forneça metadados sobre a localidade usada pelos valores de dados (LÓSCIO <i>et al.</i> , 2017).	P17.4- Verifica-se regularmente se os metadados estão sendo disponibilizados conforme o esperado?
	Garantir que todos os conjuntos de dados sejam publicados com um conjunto padronizado de documentação de apoio (DODDS; NEWMAN, 2015).	P17.5- São disponibilizados diferentes metadados para cada conjunto de dados?

⁸ As perguntas sinalizadas com * complementam a capacidade de aprimoramento deste componente.

Integrar esquemas de metadados e vocabulários controlados federados para categorizar adequadamente as informações	Use termos de vocabulários compartilhados, preferencialmente padronizados, para codificar dados e metadados (HAUSENBLAS, 2011; LÓSCIO <i>et al.</i> , 2017; VILLAZÓN-TERRAZAS <i>et al.</i> , 2011; HYLAND; WOOD, 2011).	
Fornecer vários tipos de metadados , de acordo com os padrões de metadados (por exemplo, CERIF, CKAN, DC, EGMS, DCAT)	Fornecer metadados descritivos de recursos gerais dos conjuntos de dados e distribuições (LÓSCIO <i>et al.</i> , 2017; VILLAZÓN-TERRAZAS <i>et al.</i> , 2011).	
	Fornecer metadados estruturais que descrevem o esquema e a estrutura interna de uma distribuição (LÓSCIO <i>et al.</i> , 2017; VILLAZÓN-TERRAZAS <i>et al.</i> , 2011).	
Pré-requisitos:		Capacidade
P17.1		A - Definição
P17.2 e P17.3		B - Implementação
P17.4		C - Monitoramento
P17.5, P11.5*, P12.4* e P13.4*		D - Aprimoramento

Fonte: Da Autora

A preservação de conjuntos de dados é essencial para manter aplicações e serviços que fazem uso dos dados disponibilizados. No entanto observa-se que há motivos e razões para que as organizações tornem um recurso indisponível. Neste caso o cenário ideal é fornecer informações para os recursos excluídos, se possível incluindo o motivo, pois isto aumenta a confiança por parte dos usuários, visto que remover um dado da Web é uma má prática (LÓSCIO *et al.*, 2017). Para isso, o Quadro 39 apresenta o contexto do componente preservação.

Quadro 39 – Preservação

Componente: 18 - Preservação		
Tem por objetivo verificar se os padrões para preservar os conjuntos de dados disponíveis foram definidos, se estão sendo utilizados, monitorados e aprimorados.		
Fatores Importantes	Boas práticas	Perguntas
Criar uma estratégia para manter conjuntos de dados publicados	A organização irá manter um catálogo que listará tanto seus próprios conjuntos de dados internos quanto aqueles que utilizam de outras organizações (DODDS; NEWMAN, 2015; HAUSENBLAS, 2011; HYLAND; WOOD, 2011; JANSSEN; ZUIDERWIJK, 2012; VILLAZÓN-TERRAZAS <i>et al.</i> , 2011).	P18.1- Há uma estratégia definida para manter os conjuntos de dados disponíveis para acesso? P18.2- Antes de remover um conjunto de dados faz-se uma análise das implicações que terão a remoção daquele conjunto?
	Dar suporte à manutenção de metadados específicos de conjuntos de dados para usuários humanos e aplicativos de computador (VILLAZÓN-TERRAZAS <i>et al.</i> , 2011; LÓSCIO <i>et al.</i> , 2017; DODDS; NEWMAN, 2015).	P18.3- Para dados que não estão mais disponíveis são fornecidas informações para humanos sobre como os dados podem ser acessados e quem pode acessá-los? P18.4- Verifica-se regularmente se há conjuntos de dados que não foram preservados e não há informações sobre como estes dados podem ser acessados e quem pode acessá-los?
Identificar a necessidade dos dados	Avalie a cobertura de um conjunto de dados antes de sua preservação (HYLAND; WOOD, 2011; LÓSCIO <i>et al.</i> , 2017).	P18.5 – Para dados que não estão mais disponíveis são fornecidas informações legíveis por máquinas?
Ser transparente com os usuários de dados abertos sobre as condições sob as quais a publicação de dados ocorre.	Para dados que não estão disponíveis, forneça uma explicação sobre como os dados podem ser acessados e quem pode acessá-los (LÓSCIO <i>et al.</i> , 2017)	
Pré-requisitos:		Capacidade
P2.5, P18.1 e P18.2		A - Definição
P18.3		B - Implementação
P18.4		C - Monitoramento
P18.5		D - Aprimoramento

Fonte: Da Autora

O princípio do componente enriquecimento é que se tenha um conjunto de ferra-

mentas que permitam aperfeiçoar os dados brutos disponibilizados, assim possibilitando a pesquisa, a análise e a visualização, conseqüentemente os usuários poderão ter uma melhor compreensão, assim como vislumbrar diferentes possibilidades de uso (LÓSCIO *et al.*, 2017). O Quadro 40 mostra o contexto considerado para o componente enriquecimento.

Quadro 40 – Enriquecimento

Componente: 19 - Enriquecimento		
Tem por objetivo verificar se os padrões para enriquecimento dos dados foram definidos, se estão sendo utilizados, monitorados e aprimorados.		
Fatores Importantes	Boas práticas	Perguntas
Implementar funcionalidades avançadas de pesquisa de dados		P19.1– São definidas ferramentas e/ou funcionalidades para pesquisa, visualização e análise dos dados disponibilizados? P19.2 – As funcionalidades e/ou ferramentas de pesquisa, análise e visualização de dados, foram disponibilizadas?
Fazer uso de conjuntos de ferramentas complementares para executar tarefas de curadoria adicionais (limpeza, vinculação, visualização, análise)	Enriqueça os dados apresentando-os de forma complementar e imediatamente informativa, como visualizações, tabelas, aplicativos da Web ou resumos (LÓSCIO <i>et al.</i> , 2017)	P19.3 – Verifica-se regularmente se as funcionalidades e/ou ferramentas de pesquisa, análise e visualização de dados, estão funcionando conforme esperado? P19.4 – As funcionalidades e/ou ferramentas de pesquisa, análise e visualização de dados, estão sendo mantidas ou aprimoradas?
Pré - requisitos:		Capacidade
P19.1		A - Definição
P19.2		B - Implementação
P19.3		C - Monitoramento
P19.4		D - Aprimoramento

Fonte: Da Autora

Os *feedbacks* são um meio importante para que os publicadores compreendam as necessidades dos usuários dos dados, sendo que os *feedbacks* também viabilizam melhorar a integridade dos dados disponibilizados, bem como permite e incentiva a disponibilização de novos conjuntos e recursos (LÓSCIO *et al.*, 2017). Neste sentido o Quadro 41 aborda o contexto considerado para o componente *feedback*.

Quadro 41 – *Feedback*

Componente: 20 - Feedback		
Tem por objetivo verificar se os padrões para coletar <i>feedback</i> foram definidos, se estão sendo utilizados, monitorados e aprimorados.		
Fatores Importantes	Boas práticas	Perguntas
Fornecer um fórum para discutir o que pode ser aprendido com o uso de dados abertos	Coletar <i>feedbacks</i> dos consumidores de dados, através de um fórum, por exemplo, para levantar questões sobre um conjunto de dados (LÓSCIO <i>et al.</i> , 2017; DODDS; NEWMAN, 2015).	P20.1 – Há uma definição sobre como ocorrerá o envio de <i>feedbacks</i> ? P20.2 – Há uma estrutura disponível para envio de <i>feedbacks</i> pelos usuários dos dados? P20.3 – Existe fórum para discutir o que pode ser aprendido com o uso dos dados abertos?
Integrar estruturas para avaliar a qualidade e a usabilidade dos dados e a plataforma, fornecendo <i>feedback</i> contínuo aos desenvolvedores e administrações	Torne publicamente disponíveis o <i>feedback</i> dos consumidores sobre determinados conjuntos de dados e distribuições (LÓSCIO <i>et al.</i> , 2017).	P20.4 – Verifica-se regularmente através dos <i>feedbacks</i> recebidos se os dados estão sendo utilizados? P20.5 – É possível aprimorar os dados disponibilizados com base nos <i>feedbacks</i> e também aprender com a reutilização dos dados?
Fornecer mecanismos para que as agências governamentais saibam como seus dados foram reutilizados		
Fornecer mecanismos para que as agências governamentais saibam o que pode ser aprendido com a reutilização de seus dados		

Fornecer mecanismos para que as agências governamentais saibam como a publicação de seus dados pode ser aprimorada com base no <i>feedback</i> que eles receberam dos usuários de dados abertos	
Pré-requisitos:	Capacidade
P20.1	A - Definição
P20.2 e P20.3	B - Implementação
P20.4	C - Monitoramento
P20.5	D - Aprimoramento

Fonte: Da Autora

Por fim, o componente acesso visa possibilitar o acesso aos dados, considerando o acesso de humanos e máquinas, para que assim possam se beneficiar dos dados que são disponibilizados na Web (LÓSCIO *et al.*, 2017). O Quadro 42 aborda o contexto considerado para o componente acesso.

Quadro 42 – Acesso

Componente: 21 - Acesso		
Tem por objetivo verificar se os padrões com relação ao acesso aos dados foram definidos, se estão sendo utilizados, monitorados e aprimorados.		
Fatores Importantes	Boas práticas	Perguntas
Determinar onde os conjuntos de dados serão publicados	Apoiar a liberação; por exemplo, com uma plataforma específica ou portal de dados específico (DODDS; NEWMAN, 2015; HYLAND; WOOD, 2011; HAUSENBLAS, 2011; JANSSEN; ZUIDERWIJK, 2012; VILLAZÓN-TERRAZAS <i>et al.</i> , 2011).	P21.1 – A interface de usuário é clara e intuitiva? P21.2 – O acesso aos dados se dá por meio de protocolos de troca, por exemplo, <i>File Transfer Protocols</i> (FTP)? P21.3 – É possibilitado obter acesso aos dados através de API?

Fornecer interfaces de programação de aplicativos (APIs) para o fornecimento de dados abertos na forma de <i>feeds</i> de serviço (de dados abertos a serviços abertos)	Ofereça uma API para veicular dados, se você tiver recursos para isso (LÓSCIO <i>et al.</i> , 2017) Definir padrões dos métodos de acesso (por exemplo, uma API) utilizados para disponibilizar os dados (DODDS; NEWMAN, 2015), (JANSSEN; ZUIDERWIJK, 2012 apud BROEK <i>et al.</i> , 2014).	
Desenvolver uma interface de usuário clara (símbolos lógicos, configuração clara da página da web, design simples).		
Ser transparente com os usuários de dados abertos sobre as condições sob as quais a publicação de dados ocorre.	Para dados que não estão disponíveis, forneça uma explicação sobre como os dados podem ser acessados e quem pode acessá-los (LÓSCIO <i>et al.</i> , 2017)	
Pré-requisitos:	Capacidade	
P2.3 e P18.1	A - Definição	
P2.5, P21.1 e P18.3	B - Implementação	
P20.4	C - Monitoramento	
P21.2 e P21.3	D - Aprimoramento	

Fonte: Da Autora

Observa-se que para cada componente foram descritas apenas dependências evidentes, pois a análise aprofundada de todas as dependências existentes que permeia a estrutura do modelo de maturidade demandará um grande esforço, sendo que este não está no escopo desta tese. Por este motivo a investigação sobre a análise em detalhes acerca das dependências será considerada na realização de trabalhos futuros.

Neste sentido nota-se que os componentes apresentados juntamente com as capacidades evidenciadas fornecem um ponto inicial para a avaliação de maturidade

nas organizações e conseqüentemente o modelo de maturidade como um todo poderá ser aperfeiçoado com base na experiência e nas práticas acerca de dados abertos.

5.2.3 Posicionar as capacidades na matriz

As capacidades foram posicionadas na matriz, de acordo com as dependências observadas para cada componente. E com isso chegou-se à matriz de maturidade apresentada na Figura 39, onde é possível observar em cada célula da matriz as perguntas correspondentes a cada nível de maturidade, assim como às dependências observadas⁹. Cada cor corresponde as capacidades apresentadas na legenda da Figura.

⁹ As perguntas de dependências são apresentadas entre parênteses (), em negrito e com tamanho menor de fonte.

Figura 39 – Modelo de Maturidade de Dados Abertos - Posicionamento de perguntas e dependências

Níveis de Maturidade		0	1	2	3	4	5	6	7	8	9	10	11	12	
Área foco	Gerenciamento de Políticas e Processos														
componentes	1-Definição dos dados		P1.1	P1.2	P1.3	P1.4	P1.5	P1.6	P1.7	P1.8	P1.9	P1.10	P1.11	P1.12	
	2-Acesso e entrega de dados		P2.1	P2.2	P2.3	P2.4	P2.5	P2.6	P2.7	P2.8	P2.9				
	3-Monitoramento e Medição		P3.1	P3.2	P3.3	P3.4	P3.5								
	4-Gerenciamento de mudança de dados		P4.1	P4.2	P4.3	P4.4	P4.5	P4.6	P4.7	P4.8	P4.9	P4.10			
Área foco	Gerenciamento de Fatores Organizacionais														
componentes	5-Funções e Responsabilidade		P5.1	P5.2	P5.3	P5.4	P5.5								
	6-Educação		P6.1	P6.2	P6.3	P6.4									
	7-Treinamento		P7.1	P7.2	P7.3	P7.4	P7.5	P7.6							
	8-Planejamento e priorização da abertura de dados					(P2.3, P3.1, P5.1, P7.1) P8.1		(P2.5, P3.2, P5.3, P7.3) P8.2		(P2.7, P3.4, P5.4, P7.4) P8.3	P8.4	P8.5			
	9-Planejamento da publicação dos dados						(P1.3, P2.3, P4.3, P8.1) P9.1		(P1.6, P2.5, P4.6, P8.2) P9.2			(P1.9, P2.7, P4.8, P8.3) P9.3	P9.4		
	10-Fomento ao uso		P10.1	P10.2	P10.3	P10.4									
Área foco	Gerenciamento de Padrões														
componentes	11-Licença		P11.1	P11.2	P11.3	P11.4	P11.5	P11.6							
	12-Proveniência e Qualidade		P12.1	P12.2	P12.3	P12.4	P12.5	P12.6							
	13-Versionamento		P13.1	P13.2	P13.3	P13.4	P13.5								
	14-Identificação		P14.1	P14.2	P14.3	P14.4	P14.5								
	15-Formato					(P1.3) P15.1	P15.2	P15.3	P15.4	P15.5	P15.6	P15.7			
	16-Vocabulários		P16.1	P16.2	P16.3	P16.4	P11.6*	P12.5*	P12.6*	P13.5*					
	17-Metadados		P17.1	P17.2	P17.3	P17.4	P17.5	P11.5*	P12.4*	P13.4*					
	18-Preservação							(P2.5) P18.1	P18.2	P18.3	P18.4	P18.5			
	19-Enriquecimento		P19.1	P19.2	P19.3	P19.4									
	20-Feedback		P20.1	P20.2	P20.3	P20.4	P20.5								
	21-Acesso								(P2.3, P18.1)		(P2.5, P18.3) P21.1	(P20.4)	P21.2	P21.3	
Legenda:	Não há práticas e tratamentos	A - Definição	B - Implementação	C - Monitoramento	D - Aprimoramento										

Fonte: Da Autora

As capacidades foram posicionadas em cada célula da matriz a qual corresponde a uma pergunta, sempre iniciado na coluna de maturidade 1, pois a maturidade 0 corresponde a inexistência de práticas relacionadas ao componente. Sendo que a capacidade A do componente "Definição dos dados" somente será atribuída, caso todas as perguntas correspondentes ao componente na cor azul tenham resposta afirmativa. Após foi realizado o posicionamento de acordo com as células que possuem dependência¹⁰, sempre posicionando mais à direita a pergunta da sua dependência.

As perguntas destacadas com * apresentada na Figura 39 agregam maior semântica aos seus respectivos componente, por este motivo foram consideradas no quesito aprimoramento, no sentido quanto mais se faz uso de vocabulários e metadados maior semântica será agregada a estes componentes. Desta forma, estas perguntas quando respondidas de forma afirmativa somam um nível de maturidade ao seu respectivo componente, ou seja, não há dependências existentes entre as perguntas destacadas com *.

Vale destacar que o modelo de maturidade apresentado na Figura 39 corresponde ao cenário ideal para que todos os componentes e áreas foco sejam desenvolvidos para dados abertos. No entanto, observa-se que a organização pode avaliar quais áreas foco e quais de seus componentes são interessantes para a sua estratégia de dados abertos e com isso viabilizem a melhoria de suas políticas e processos, de seus fatores organizacionais e seus padrões.

Também é possível representar a matriz de maturidade de dados abertos de forma resumida, esta apresentada no Quadro 43. Observa-se o posicionamento das capacidades, considerando as perguntas e dependências apresentadas na Figura 39.

Quadro 43 – Matriz - Modelo de Maturidade de Dados Abertos

Níveis de Maturidade	0	1	2	3	4	5	6	7	8	9	10	11	12
Gerenciamento de políticas e processos													
Definição dos dados				A			B			C			D
Acesso e entrega de dados				A		B		C		D			
Monitoramento e medição			A	B	C	D							
Gerenciamento de mudança de dados				A			B		C		D		
Gerenciamento de fatores organizacionais													
Funções e responsabilidade			A	B	C	D							
Educação		A	B	C	D								
Treinamento			A	B	C		D						
Planejamento e priorização da abertura dos dados					A		B		C		D		

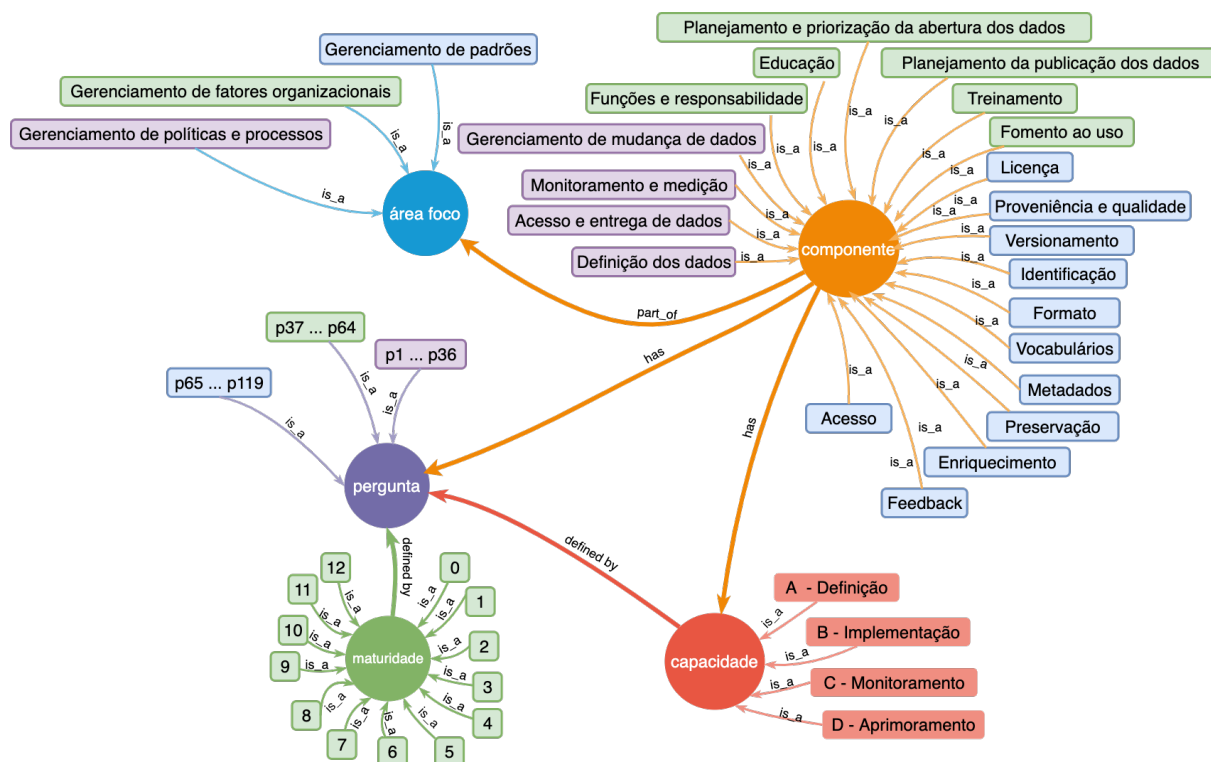
¹⁰ as dependências estão destacadas em negrito e entre parênteses ().

Planejamento da publicação dos dados						A		B			C	D	
Fomento ao uso		A		B		C		D					
Gerenciamento de padrões													
Licença		A		B	C		D						
Proveniência e qualidade		A	B	C			D						
Versionamento		A	B	C		D							
Identificação		A		B	C	D							
Formato					A	B	C				D		
Vocabulários		A	B	C					D				
Metadados		A		B	C				D				
Preservação								A	B	C	D		
Enriquecimento		A	B	C	D								
Feedback		A		B	C	D							
Acesso								A		B	C		D

Fonte: Da Autora

Na Figura 40 apresenta-se uma estrutura baseada na representação utilizada em ontologias, com o intuito de abordar e sumarizar os conhecimentos observados no decorrer desta seção e suas relações, esta representação considera também um esquema de cores para melhor representá-la. Sendo possível através desta figura observar e compreender as relações existentes entre as áreas foco, os componentes e as perguntas para com as capacidades e os níveis de maturidade.

Figura 40 – Representação do conhecimento abordado no modelo de maturidade



Fonte: Da Autora

6 APRESENTAÇÃO E AVALIAÇÃO DO MODELO DE MATURIDADE

Para viabilizar a coleta das respostas ao questionário e verificar o comportamento do modelo de maturidade, desenvolveu-se um sistema Web, com o intuito de instanciar e divulgar a pesquisa desenvolvida. Também foi realizada a avaliação do modelo de maturidade junto a especialistas da área com o objetivo de analisar e discutir a estrutura do modelo e obter possíveis comentários sobre melhorias acerca da abrangência, da relevância e da importância dos elementos que compõe o modelo de maturidade.

6.1 APRESENTAÇÃO

Para apresentar o modelo de maturidade foram coletadas respostas através de entrevistas, com o intuito de instanciar e divulgar a pesquisa desenvolvida. Para isso, fez-se necessário o desenvolvimento de um sistema Web, este descrito na sequência.

6.1.1 Sistema de Apoio a Avaliação da Maturidade de Dados Abertos SAMDA

No desenvolvimento do Sistema de Apoio a Avaliação da Maturidade de Dados Abertos (SAMDA), fez-se o uso das seguintes tecnologias: banco de dados Mysql, Flask, CSS, jQuery, Bootstrap, e Python. Sendo que o sistema está disponível no endereço: <https://lidianevisintin.studiokem.ufsc.br/aceso> para acesso.

O sistema desenvolvido possui três perfis de usuário, sendo estes:

- **Visitante** - somente possui acesso para visualizar perfis de maturidade já preenchidos no sistema sem acesso a informações de quem realizou o preenchimento;
- **Gestor** - possui acesso ao questionário para realizar a avaliação da maturidade e também é possível visualizar todos os perfis de maturidade da organização que o usuário está vinculado;
- **Administrador** - possui acesso ao questionário para realizar a avaliação da maturidade e também é possível visualizar todos os perfis de maturidade das organizações que realizaram a avaliação, sendo também possível alterar o perfil de outros usuário para fornecer ou remover privilégios.

Para ter acesso, o usuário precisará realizar o seu cadastro e por padrão terá acesso inicial como visitante, sendo possível solicitar ao administrador a mudança de perfil de acesso, caso o usuário queira fazer uso do questionário.

Este sistema foi desenvolvido com o intuito de auxiliar na avaliação da maturidade das organizações, visto que o controle das dependências demandaria um esforço extra caso fosse realizado através de planilhas, ou até mesmo de formulários do Google.

Destaca-se ainda que cada capacidade está relacionada a 1 até 4 questões do questionário, com possibilidade de resposta sim ou não. Todas as perguntas associadas a uma capacidade devem ser respondidas com sim para reivindicar a obtenção dessa capacidade em cada componente. Todo esse controle e a verificação das dependências foi realizado via software. Neste sentido, o sistema Web desenvolvido viabilizou a coleta dos cenários de uso que serão apresentados abaixo.

6.1.2 Cenários de Uso

Para obter os cenários de uso foram conduzidas reuniões online com os entrevistados; estas reuniões ocorreram nos meses de abril e maio de 2021. Nas entrevistas seguiu-se a seguinte sequência para todos os entrevistados:

- Identificação do perfil dos entrevistados;
- Identificação das características e dos objetivos da organização com a abertura dos dados;
- Apresentação da estrutura do modelo de maturidade;
- Aplicação do questionário;
- Apresentação e explicação do perfil de maturidade da organização;
- Identificação de possíveis sugestões por parte dos entrevistados;
- Agradecimento aos entrevistados e encerramento.

No Quadro 44 apresentam-se as principais características dos entrevistados e também o perfil das organizações e respondentes abordados em cada cenário de uso.

Quadro 44 – Descrição dos cenários de uso e dos entrevistados

	Cenários de Uso		
	Cenário 1	Cenário 2	Cenário 3
Perfil	1 respondente, sendo seu maior grau de escolaridade Mestrado em Tecnologia da Informação e Gestão do Conhecimento.	3 respondentes, sendo o maior grau de escolaridade destes: Mestrado em Administração (respondente 1), Doutorado em Engenharia e Gestão do Conhecimento (respondente 2) Especialização em Educação Escolar (respondente 3).	3 respondentes, sendo o maior grau de escolaridade destes: Graduado em Ciência da Computação (respondente 1), Graduado em Direito (respondente 2) e Especialização em Segurança da Informação (respondente 3).

<p>Descrição Geral</p>	<p>O cenário em análise corresponde a uma secretária em nível federal;</p> <ul style="list-style-type: none"> - Disponibiliza dados desde 2011; - Não há um número preciso de pessoas envolvidas na disponibilização dos dados, visto que há equipes descentralizadas; - A abertura dos dados é realizada com o intuito de obter transparência, gerar estudos acadêmicos, resolver problemas sociais, criar serviços e até mesmo otimizar serviços existentes; 	<ul style="list-style-type: none"> - O cenário em análise corresponde a uma instituição de ensino em nível federal; - A instituição disponibiliza dados desde 2019; - Há cerca de 5 ou 6 pessoas envolvidas na abertura dos dados; - A abertura é realizada com o intuito de obter transparência e por conta da lei de acesso à informação; 	<ul style="list-style-type: none"> - O cenário em análise corresponde a um centro de informática a nível estadual; - A instituição é responsável por preparar os dados para que outros departamentos do estado os disponibilizem; - Atualmente 5 pessoas estão envolvidas na preparação dos dados para a disponibilização; - A instituição realiza o preparo dos dados com objetivo de obter transparência e atender a legislação;
-------------------------------	---	---	--

Fonte: Da Autora

Na sequência serão apresentados os resultados obtidos através das entrevistas, para cada cenário.

• Cenário I

O entrevistado no início da reunião explicou que houve mudanças no setor onde ele trabalha, com o decorrer do tempo, sendo que houve a realocação de pessoal e também houve a atualização no plano de dados abertos e atualmente não há um planejamento único para todas as secretarias, observando que este cenário em análise disponibiliza dados desde 2011. Para o **cenário 1**, considerando a área foco **definição dos dados** obteve-se a avaliação da maturidade apresentada na Figura 41.

Figura 41 – Cenário 1 - Área foco 1

Níveis de Maturidade	0	1	2	3	4	5	6	7	8	9	10	11	12
Gerenciamento de Políticas e Processos													
Definição dos dados				A			B			C			D
Acesso e entrega de dados				A		B		C		D			
Monitoramento e Medição			A	B	C	D							
Gerenciamento de mudança de dados				A			B		C		D		

Fonte: Da Autora

Nesta área foco o entrevistado mencionou que não há uma definição específica de etapas para a disponibilização dos dados, pois quando há a necessidade de disponibilizar um novo conjunto de dados gera-se uma sprint do Scrum para atender determinada demanda. Referente ao monitoramento e medição e sobre o gerenciamento de mudança dos dados observou-se que não há ênfase para estes componentes.

Outro ponto observado na entrevista é referente à conversão dos dados para um formato legível por máquina, pois isso causou certa estranheza no entrevistado, visto que ele compreende que os dados apenas são coletados e disponibilizados. No entanto, o entrevistado destacou que priorizam formatos legíveis por máquina para a disponibilização dos dados.

Para o **cenário 1**, considerando a área foco **fatores organizacionais** obteve-se a avaliação da maturidade apresentada na Figura 42.

Figura 42 – Cenário 1 - Área foco 2

Níveis de Maturidade	0	1	2	3	4	5	6	7	8	9	10	11	12
Gerenciamento de Fatores Organizacionais													
Funções e Responsabilidade			A	B	C	D							
Educação		A	B	C	D								
Treinamento			A	B	C		D						
Planejamento e Priorização da Abertura de Dados					A		B		C		D		
Fomento ao uso		A	B	C	D								
Planejamento da publicação dos dados						A		B			C	D	

Fonte: Da Autora

Na Figura 42 observa-se que há ênfase em treinamento para os envolvidos na abertura dos dados, no entanto o fomento ao uso não é algo que o cenário observado possui como prioridade. Destaca-se ainda para **fatores organizacionais** que os componentes planejamento e priorização da abertura dos dados e planejamento da publicação dos dados não habilitaram perguntas, visto que as dependências para estes componentes não foram respondidas afirmativamente.

Para o **cenário 1**, considerando a área foco **gerenciamento de padrões** teve-se a avaliação da maturidade apresentada na Figura 43. Onde observa-se o componente versionamento em destaque; segundo o entrevistado este item recentemente foi aprimorado e aplicado ao último conjunto de dados disponibilizado.

Figura 43 – Cenário 1 - Área foco 3

Gerenciamento de Padrões												
Licença		A		B	C		D					
Proveniência e Qualidade		A	B	C			D					
Versionamento		A	B	C		D						
Identificação		A		B	C	D						
Formato					A	B	C				D	
Vocabulários		A	B	C					D			
Metadados		A		B	C				D			
Preservação								A	B	C	D	
Enriquecimento		A	B	C	D							
Feedback		A		B	C	D						
Acesso								A		B	C	D

Fonte: Da Autora

No **cenário 1**, considerando a área foco **gerenciamento de padrões** o componente identificação obteve maturidade 0, ou seja, o entrevistado citou que não há uma definição sobre como fazer uso de URIs, para a identificação de conjuntos de dados e recursos, de mesmo modo se deu para o componente enriquecimento.

Destaca-se que os componentes preservação e acesso não foram habilitados devido às dependências destes não terem sido respondidas.

Ao final da entrevista apresentou-se o resultado e discutiu-se sobre os componentes avaliados, com intuito de identificar se os componentes que obtiveram 0

como nível de maturidade, se estes realmente não foram desenvolvidos na organização. Neste sentido, o entrevistado destacou o monitoramento como apenas algumas ações pontuais sendo realizadas pelo cenário em análise.

Assim, o entrevistado citou como contribuições para o sistema a adição de uma breve descrição para cada componente, com o intuito de deixar claro os aspectos que são considerados para a avaliação. Também, o entrevistado destacou utilizar uma escala de "sempre" "as vezes" ou "nunca" para possíveis respostas, isso ao invés das opções de "sim" e "não".

No entanto, destaca-se que haveria uma dificuldade para atribuir um nível de maturidade e até mesmo determinar se uma dependência foi ou não realizada, uma vez que fosse possível admitir respostas do tipo "as vezes".

• Cenário II

Para este cenário foram entrevistadas 3(três) pessoas. No início da reunião um dos entrevistados declarou que a iniciativa de realizar a abertura dos dados é recente com objetivo de atender a necessidade de transparência e a lei de acesso à informação. Ainda, destacou que o plano de dados abertos existe desde 2017, no entanto a disponibilização dos dados deu-se somente em 2019, sendo que em 2020 não houve ênfase na abertura dos dados, devido à pandemia de Covid-19 que gerou diversas demandas. Neste sentido, destacaram que há a necessidade de atualizar o plano de dados abertos e conseqüentemente avançar na abertura dos dados.

Para este cenário (**cenário 2**), considerando a área foco **definição dos dados** obteve-se a avaliação da maturidade apresentada na Figura 44.

Figura 44 – Cenário 2 - Área foco 1

Níveis de Maturidade	0	1	2	3	4	5	6	7	8	9	10	11	12
Gerenciamento de Políticas e Processos													
Definição dos dados				A			B			C			D
Acesso e entrega de dados				A		B		C		D			
Monitoramento e Medição			A	B	C	D							
Gerenciamento de mudança de dados				A			B		C		D		

Fonte: Da Autora

Na Figura 44 é possível observar que os componentes monitoramento e medição e gerenciamento de mudança de dados não possuem ênfase.

Para a segunda área foco (**fatores organizacionais**) obteve-se a avaliação da maturidade apresentada na Figura 45, onde observam-se que os componentes planejamento e priorização da abertura dos dados estão zerados.

Figura 45 – Cenário 2 - Área foco 2

Níveis de Maturidade	0	1	2	3	4	5	6	7	8	9	10	11	12
Gerenciamento de Fatores Organizacionais													
Funções e Responsabilidade			A	B	C	D							
Educação		A	B	C	D								
Treinamento			A	B	C		D						
Planejamento e Priorização da Abertura de Dados					A		B		C		D		
Fomento ao uso		A	B	C	D								
Planejamento da publicação dos dados						A		B			C	D	

Fonte: Da Autora

Por fim, para este cenário (**cenário 2**) observa-se a área foco **gerenciamento de padrões** que resultou na avaliação da maturidade apresentada na Figura 46. Para esta área foco os entrevistados destacaram ter conhecimento sobre a necessidade de padrões para a identificação de conjuntos de dados e recursos, mas ainda não há uma percepção consolidada para a identificação. Os entrevistados informaram desconhecer a necessidade de vocabulários. Ainda com a avaliação foi possível observar que o componente *feedback* não possui ênfase.

Ao final da entrevista apresentou-se o resultado e discutiu-se sobre os componentes avaliados, com ênfase nos componentes que obteve-se 0 como nível de maturidade. Os entrevistados demonstraram surpresa ao obter o nível de maturidade final 0 (zero), devido a alguns componentes obterem este nível de maturidade.

Com relação aos componentes com nível de maturidade 0 (zero), os entrevistados citaram a necessidade de se ter métricas e indicadores de sucesso para que se tenha o planejamento e a priorização da abertura dos dados. Assim foi discutido que para obter a capacidade A no componente planejamento e priorização da abertura dos dados, bastava que a organização tivesse reuniões internas para discutir como avaliar o processo de disponibilização dos dados disponibilizados, destacou-se também que este seria o cenário ideal para dados abertos apresentado pela literatura.

Deste modo, os entrevistados conseguiram observar que os componentes que obtiveram nível de maturidade 0, não estão sendo desenvolvidos na organização e após as discussões sobre o que compete a cada componente destacaram que

Figura 46 – Cenário 2 - Área foco 3

Gerenciamento de Padrões												
Licença		A	B	C	D							
Proveniência e Qualidade		A	B	C	D							
Versionamento		A	B	C	D							
Identificação		A	B	C	D							
Formato				A	B	C				D		
Vocabulários		A	B	C					D			
Metadados		A	B	C					D			
Preservação								A	B	C	D	
Enriquecimento		A	B	C	D							
Feedback		A	B	C	D							
Acesso								A	B	C	D	

Fonte: Da Autora

a estrutura é robusta e abrangente o que tornou o questionário longo. Deste modo sugeriram a possibilidade da organização selecionar os componentes para cada área foco alinhados à estratégia organizacional, para melhor representar o que deseja-se avaliar.

Ainda durante a entrevista foi destacado que o fomento ao uso por mais que a organização não esteja objetivando agora é algo necessário e que estes deveriam pensar em como incluir este item na atualização do plano.

• Cenário III

Para este cenário foram entrevistadas 3 (três) pessoas. Este cenário difere-se dos demais analisados, visto que esta organização é responsável por receber e preparar os dados para serem disponibilizados por outros setores, secretarias e órgãos em nível estadual. Portanto, logo no início da entrevista ficou evidente de que certos componentes não seriam do interesse da organização devido as suas necessidades.

Para este cenário(**cenário 3**), considerando a área foco **definição dos dados** obteve-se a avaliação da maturidade apresentada na Figura 47.

Figura 47 – Cenário 3 - Área foco 1

Níveis de Maturidade	0	1	2	3	4	5	6	7	8	9	10	11	12
Gerenciamento de Políticas e Processos													
Definição dos dados				A			B			C			D
Acesso e entrega de dados				A		B		C		D			
Monitoramento e Medição			A	B	C	D							
Gerenciamento de mudança de dados				A			B		C		D		

Fonte: Da Autora

Para o **cenário 3**, considerando a área foco **fatores organizacionais** obteve-se a avaliação da maturidade apresentada na Figura 48.

Figura 48 – Cenário 3 - Área foco 2

Níveis de Maturidade	0	1	2	3	4	5	6	7	8	9	10	11	12
Gerenciamento de Fatores Organizacionais													
Funções e Responsabilidade			A	B	C	D							
Educação		A	B	C	D								
Treinamento			A	B	C		D						
Planejamento e Priorização da Abertura de Dados					A		B		C		D		
Fomento ao uso		A	B	C	D								
Planejamento da publicação dos dados						A		B			C	D	

Fonte: Da Autora

Por fim, no **cenário 3**, considerando a área foco **gerenciamento de padrões** obteve-se a avaliação da maturidade apresentada na Figura 49.

Ao final da entrevista, conforme realizado com os demais entrevistados apresentou-se o resultado e discutiu-se sobre os componentes avaliados, com ênfase nos componentes que obteve-se 0 (zero) como nível de maturidade. Os entrevistados observaram que os componentes em que obtiveram nota 0 (zero) de maturidade deveriam ser prioridade de quem disponibiliza os dados e no caso não compete a este cenário. No entanto, como esta organização oferece suporte a alguns quesitos relacionados a padrões, os entrevistados destacaram que foi interessante a discussão para compreender os benefícios de se fazer o uso de certos padrões, por exemplo, de vocabulários.

Figura 49 – Cenário 3 - Área foco 3

Gerenciamento de Padrões													
Licença		A		B	C		D						
Proveniência e Qualidade		A	B	C			D						
Versionamento		A	B	C			D						
Identificação		A		B	C	D							
Formato					A	B	C					D	
Vocabulários		A	B	C					D				
Metadados		A		B	C				D				
Preservação									A	B	C	D	
Enriquecimento		A	B	C	D								
Feedback		A		B	C	D							
Acesso									A		B	C	D

Fonte: Da Autora

No entanto destacaram que responder todo o questionário sem um conhecimento abrangente do assunto pode ser demorado e dificultar o entendimento do objetivo de certos componentes, mas ainda salientaram que a visualização através da representação de uma matriz facilitou o entendimento gráfico dos resultados.

Destacaram ainda que seria interessante aplicar o questionário com outros envolvidos no processo e comparar os resultados obtidos.

6.1.2.1 Percepções da pesquisadora sobre as entrevistas para obter os cenários de uso

Ao final da realização das entrevistas para obter os cenários de uso observou-se que:

- As características e necessidades das organizações provedoras dos dados devem ser consideradas dentro da avaliação da maturidade e por este motivo a sugestão de personalizar os componentes em cada área foco é interessante;
- A inclusão de uma resposta "as vezes" para as questões, inviabilizaria determinar se uma capacidade foi atingida ou se até mesmo uma dependência foi ou não realizada;

- A entrevista conduzida com apenas uma pessoa foi pouco produtiva, pois várias questões o entrevistado ficou em dúvida e acabou optando pelo "não". Sendo que as entrevistas com mais de um participante gerava discussões produtivas, visto que estes discutiam e chegavam a um consenso sobre o que estava sendo realizado na organização e até mesmo citavam a necessidade de realizar melhorias referentes aos componentes abordados no questionário aplicado;
- O nível de maturidade final apresentado deve ser removido do sistema, visto que alguns entrevistados demonstraram "certa insatisfação" ao observar o valor obtido ao final da avaliação, mesmo observando que os itens em que os cenários se destacaram em sua maioria eram aqueles priorizados pela organização analisada;
- A remoção no nível de maturidade ao final se faz necessário com o intuito em evitar desmotivações e garantir que seja possível as organizações realizarem uma reflexão sobre o que está sendo feito e como pode-se aprimorar todo o contexto acerca de dados aberto. O principal objetivo do artefato apresentado é possibilitar o vislumbre da maturidade que deve ser analisada componente a componente e não apenas como uma nota ao final.

6.1.2.2 Percepção dos Entrevistados nos Cenários de Uso

Após a coleta e a análise dos três cenários de uso foi elaborado um formulário sobre a percepção dos entrevistados referente a: **I)** estrutura do modelo de maturidade, considerando aspectos de compreensibilidade, facilidade de uso e utilidade e praticidade; **II)** a importância dos elementos que compõem a estrutura; e sobre **III)** a relevância e abrangência do questionário aplicado na entrevista.

Neste sentido, o Quadro 45 aborda as questões para a seção 1 do formulário (que dizem respeito ao questionário e a estrutura do modelo de maturidade), considerando as pontuações atribuídas por cada cenário de uso, onde faz-se o uso da seguinte escala: **0** - Discordo Fortemente, **1** - Discordo, **2** - Concordo, **3** - Concordo Plenamente, para mais detalhes ver (ApêndiceA).

Quadro 45 – Parte I - Percepções sobre aspectos do questionário e aspectos do modelo de maturidade

Cenários de Uso	1	2	3
As próximas afirmativas dizem respeito ao questionário que foi aplicado			
De modo geral, qual sua percepção sobre as seguintes afirmativas:			
As questões são relevantes para o domínio (Relevância)	3	3	2

As questões contribuem com aspectos importantes para cada componente (Relevância)	3	3	2
As questões são suficientes para representar cada componente (Abrangência)	2	3	2
De modo geral qual sua percepção sobre a estrutura do modelo de maturidade apresentados, considerando:			
Compreensibilidade			
Os níveis de maturidade são compreensíveis	2	3	2
A estrutura de avaliação é compreensível	2	3	2
Facilidade de uso			
O questionário é fácil de usar	3	2	1
As diretrizes para a avaliação são fáceis de usar	3	2	1
Utilidade e praticidade			
O modelo de maturidade é útil na realização de avaliações	3	3	2
O modelo de maturidade é prático para uso	3	3	2

Fonte: Da Autora

Referente às percepções sobre o **questionário**, nota-se que a média das respostas obtidas (observadas no Quadro 45) ficaram entre 2 (concordo) e 3 (concordo plenamente), demonstrando relevância e abrangência sobre o questionário.

Sobre a estrutura do modelo de maturidade para os quesitos **compreensibilidade** e **utilidade e praticidade**, nota-se que as médias das respostas obtidas (observadas no Quadro 45) ficaram entre 2 (concordo) e 3 (concordo plenamente), que denota compreensão, utilidade e praticidade observada pelos entrevistados. Referente ao quesito **facilidade de uso** as médias das respostas ficaram entre 1 (discordo) e 3 (concordo plenamente), sendo que o cenário que discordou sobre este aspecto havia sinalizado na entrevista que se faz necessário ter um conhecimento abrangente para compreender certos componentes, refletindo na facilidade de uso.

O Quadro 46 aborda as questões para a seção 2 do formulário (que dizem respeito aos elementos que compõe à estrutura do modelo de maturidade), considerando as pontuações atribuídas por cada cenário de uso, onde faz-se uso da seguinte escala: **0** - Irrelevante, **1**- Pouco Importante, **2** - Importante, **3** - Muito Importante, **4** - Indispensável, para mais detalhes ver (Apêndice B).

Quadro 46 – Parte II - Percepções sobre a importância dos elementos que compõem o modelo de maturidade

Cenários de Uso	1	2	3
Importância das áreas foco e dos componentes do modelo de maturidade			
Para você, qual a importância das seguintes áreas foco, quando considerado o domínio de dados abertos?			
Gerenciamento de Políticas e Processos	4	2	2
Gerenciamento de Fatores Organizacionais	4	2	3
Gerenciamento de Padrões	4	3	2
Para você, qual a importância dos componentes relacionados à área foco - Gerenciamento de políticas e Processos, considerando o âmbito de dados abertos?			
Definição dos dados	3	3	3
Acesso e entrega de dados	2	2	3
Monitoramento e Medição	2	3	4
Gerenciamento de mudança de dados	2	2	4
Para você, qual a importância dos componentes relacionados à área foco - Gerenciamento de Fatores Organizacionais, considerando o âmbito de dados abertos?			
Funções e Responsabilidade	4	3	2
Educação	4	2	2
Treinamento	4	3	4
Planejamento e Priorização da Abertura dos Dados	3	3	3
Planejamento da Publicação dos Dados	4	3	2
Fomento ao Uso	3	3	1
Para você, qual a importância dos componentes relacionados à área foco - Gerenciamento de Padrões, considerando o âmbito de dados abertos?			
Licença	4	3	2
Proveniência e Qualidade	2	3	2
Versionamento	2	4	1
Identificação	2	4	2
Formato	4	3	3
Vocabulários	3	3	1
Metadados	4	3	2
Preservação	2	3	2
Enriquecimento	2	3	1
<i>Feedback</i>	4	3	2
Acesso	3	3	3

Fonte: Da Autora

Em relação à percepção dos elementos que compõem o modelo de maturidade,

observados no Quadro 46, nota-se que as áreas foco possuem relevância no domínio de dados abertos, pois a média das respostas obtidas para estas, ficaram entre 2 (Importante) e 4 (Indispensável).

No entanto, a percepção da importância dos componentes é reflexo dos objetivos da organização acerca de dados abertos, por exemplo, para o **cenário 3** o componente **fomento ao uso** (Quadro 46), foi destacado na entrevista que este deveria ser prioridade da organização que disponibiliza os dados, visto que o cenário 3 apenas realiza o tratamento para a disponibilização dos dados.

Por fim, foram conduzidas entrevistas com especialistas da área de dados abertos para verificação da estrutura do modelo de maturidade. As impressões com base nas discussões são apresentadas na sequência.

6.2 AVALIAÇÃO DO MODELO COM ESPECIALISTAS

A verificação da estrutura do modelo de maturidade, principal finalidade da etapa de avaliação do artefato desenvolvido, teve como intuito discutir, divulgar e obter *feedbacks* de especialistas da área de dados abertos. Para isso foram realizadas entrevistas on-line onde realizou-se uma breve apresentação sobre a estrutura do modelo de maturidade e dos elementos que o compõem. Com isso foram discutidas e obtidas sugestões, observando e considerando os diferentes perfis dos entrevistados. Ao final de cada entrevista foi aplicado um formulário, que consiste de 3 partes, sendo estas: sobre **I)** perfil do entrevistado, sobre a **II)** importância das áreas foco e dos componentes e sobre a **III)** estrutura do modelo de maturidade de dados abertos. Os detalhes sobre os formulários estão nos Apêndices C, D e E.

Foram realizadas entrevistas com 10 especialistas em momentos diferentes, o perfil de cada entrevistado é apresentado no Quadro 47. Para a obtenção do perfil de cada especialista foram consideradas as respostas obtidas através do formulário do Apêndice C; optou-se por não apresentar o local de trabalho e o nome dos entrevistados com o intuito de manter a privacidade destes.

Quadro 47 – Perfil dos especialistas entrevistados

Especialista	Perfil
Especialista 1	Cinco anos, trabalhando em metodologias para dados abertos e publicando conjuntos de dados abertos científicos na Web de Dados. Consumo de dados abertos (3ª estrela).

Especialista 2	Desde 2012, liderou diversos projetos sobre dados abertos. Dentre alguns: primeira versão do Open Data for Development (OD4D), SPUK, Transparência Ativa com a CGU, edição do Data on the Web Best Practices (DWBP), TIC Web Dados Abertos, atualmente co-chair do Dataset Exchange Working Group (DXWG).
Especialista 3	Faz pesquisas com Web Semântica e Linked Data e já trabalhou com dados abertos em pesquisas.
Especialista 4	Pesquisa acadêmica e facilitação de debates - 11 anos de experiência.
Especialista 5	Trabalha com LOD há 7 anos, é professora e pesquisadora da área. Trabalha com pesquisas sobre metodologia para publicação e uso de LOD e desenvolvimento de estudo de casos de publicação de LOD.
Especialista 6	Desenvolveu pesquisa acadêmica sobre o assunto.
Especialista 7	Atuou como gerente de transparência e dados abertos na Controladoria-Geral do Estado de Santa Catarina por 2 anos.
Especialista 8	Trabalha há 32 meses na Controladoria-Geral do Estado de Santa Catarina com Dados Abertos.
Especialista 9	Trabalha há mais de dez anos com dados abertos. De 2009 – 2013, na sociedade civil, com dados abertos da educação; de 2013 a 2016 à frente da Coordenadoria de Promoção da Integridade da Controladoria do Município de São Paulo, tendo sido responsável pela política de abertura de dados da cidade; de 2016 a 2019, na Secretária de Educação da cidade de São Paulo, à frente da política de governo aberto da organização; e, desde então, na OKBR, coordenando projetos que usam dados abertos e estratégias de monitoramento e avaliação de dados abertos.
Especialista 10	Trabalhou durante a confecção das pesquisas e da Tese de Doutorado, quando desenvolveu um modelo de Dados Abertos Conectados, voltado para Indicadores de Inovação Regional.

Fonte: Da Autora

Observa-se através dos perfis dos entrevistados, que estes possuem diferentes conhecimentos e objetivos para dados abertos, mas todos possuem ampla experiência com o tema à mais de 2 anos. Isso também pode ser observado pelo tempo de trabalho que os entrevistados possuem e também porque alguns destes fazem uso de dados abertos, LOD e outros focam no quesito como transparência, por exemplo.

Com a realização das entrevistas são apresentadas as pontuações no Quadro 48, que correspondem à afirmativas sobre a estrutura do modelo. Para a obtenção destas pontuações fez-se uso do formulário do Apêndice E, sendo considerada a

seguinte escala: **0** - Discordo fortemente, **1** - Discordo, **2** - Concordo, **3** - Concordo plenamente.

Quadro 48 – Estrutura do Modelo de Maturidade

Especialista	1	2	3	4	5	6	7	8	9	10	Médias
Áreas Foco											
As áreas foco são suficientes para representar o domínio do modelo de maturidade (Abrangência)	1	3	3	2	3	1	3	3	2	3	2,40
As áreas foco são relevantes para o domínio (Relevância)	2	3	3	2	3	2	3	3	2	3	2,60
Componentes											
Os componentes são suficientes para representar cada área foco (Abrangência)	2	3	2	2	3	2	3	3	1	3	2,40
Os componentes são relevantes para cada área foco (Relevância)	2	3	2	3	3	2	3	3	2	3	2,60
Os componentes foram atribuídos corretamente às suas respectivas áreas foco (Precisão)	2	3	3	3	3	2	3	3	1	3	2,60
Capacidades											
As capacidades representam uma evolução progressiva para cada componente (Abrangência)	3	3	3	3	3	2	2	2	2	3	2,60
As capacidades são relevantes para o domínio (Relevância)	3	3	3	3	3	2	2	2	2	3	2,60

Fonte: Da Autora

Observa-se no Quadro 48 que as áreas foco, os componentes e as capacidades em sua maioria obtiveram uma percepção entre 2 (Concordo) e 3 (Concordo plenamente), sendo que os especialistas 1 e 6 discordaram referente a abrangência das áreas foco, mas não comentaram sobre motivo da pontuação e nem quais outras áreas que deveriam ser abordadas.

O especialista 9 destacou que discordou referente a abrangência e a precisão dos componentes, pois este necessitava compreender em detalhes as perguntas que foram englobadas em cada componente. Devido à dificuldade de agenda, não foi possível apresentar as perguntas que poderiam contribuir com este aspecto e por este motivo foi atribuída esta pontuação.

De modo geral as médias das percepções ficaram acima de 2, o que representa que os especialistas concordam com a estrutura do modelo de maturidade.

Com a realização das entrevistas obteve-se ainda as pontuações apresentadas no Quadro 49, referentes à importância das áreas foco e de seus respectivos componentes. Nesta etapa da pesquisa definiu-se o objetivo de priorizar os componentes de cada área foco, sendo que isso viabilizará a criação de uma versão "simplificada" do modelo de maturidade.

A versão "simplificada" do modelo de maturidade possibilitará um questionário menor e conseqüentemente as organizações terão acesso a um menor número de componentes para atingirem o aprimoramento. Deste modo, seria possível avaliar rapidamente e ter uma visão dos componentes indispensáveis para dados abertos.

Para a obtenção das respostas utilizou-se do formulário do Apêndice D, sendo considerada a seguinte escala: **0** - Irrelevante, **1**- Pouco Importante, **2**- Importante, **3** - Muito Importante e **4** - Indispensável.

Quadro 49 – Importância das áreas foco e dos componentes do modelo de maturidade

Especialista	1	2	3	4	5	6	7	8	9	10	Médias
Para você, qual a importância das seguintes áreas foco, quando considerado o domínio de dados abertos?											
Gerenciamento de Políticas e Processos	4	4	4	4	4	3	4	3	4	4	3,80
Gerenciamento de Fatores Organizacionais	4	4	4	4	4	3	3	3	4	4	3,70
Gerenciamento de Padrões	4	4	4	4	4	4	3	4	4	4	3,90
Para você, qual a importância dos componentes relacionados à área foco - Gerenciamento de políticas e Processos, considerando o âmbito de dados abertos?											
Definição dos dados	4	4	4	4	4	2	3	3	3	4	3,50
Acesso e entrega de dados	4	4	4	4	4	2	3	4	3	3	3,50
Monitoramento e Medição	3	4	3	2	3	4	3	3	2	2	2,90
Gerenciamento de mudança de dados	4	4	4	2	4	3	3	3	2	4	3,30
Para você, qual a importância dos componentes relacionados à área foco - Gerenciamento de Fatores Organizacionais, considerando o âmbito de dados abertos?											
Funções e Responsabilidade	3	4	3	3	3	3	4	3	3	3	3,20
Educação	4	3	3	2	3	2	4	4	3	4	3,20
Treinamento	3	4	3	2	4	3	3	4	2	2	3,00
Planejamento e Priorização da Abertura dos Dados	4	3	3	3	3	4	4	3	4	4	3,50
Planejamento da Publicação dos Dados	4	3	3	3	4	4	4	3	4	4	3,60
Fomento ao Uso	3	3	2	3	3	4	3	4	4	3	3,20

Para você, qual a importância dos componentes relacionados à área foco - Gerenciamento de Padrões, considerando o âmbito de dados abertos?											
Licença	4	4	3	2	4	2	2	3	2	3	2,90
Proveniência e Qualidade	4	4	4	3	3	3	2	4	4	4	3,50
Versionamento	4	4	4	4	4	4	2	3	2	3	3,40
Identificação	4	4	3	4	4	2	0	3	3	4	3,10
Formato	4	4	3	3	4	2	3	4	4	4	3,50
Vocabulários	4	4	3	3	4	3	3	4	2	4	3,40
Metadados	4	4	4	4	4	3	4	4	4	4	3,90
Preservação	4	4	4	3	3	3	3	4	2	4	3,40
Enriquecimento	4	4	3	3	4	3	2	3	2	4	3,20
<i>Feedback</i>	4	4	3	2	3	3	3	3	4	3	3,20
Acesso	4	4	4	3	4	4	3	4	3	4	3,70

Fonte: Da Autora

Considerando a média das pontuações obtidas, estas entre 2,9 (Importante) e 3,9 (Muito Importante), não foi possível compor uma versão simplificada do modelo, dada a importância dos componentes observada através das percepções dos especialistas.

Em relação à percepção dos elementos que compõem o modelo de maturidade, observados no Quadro 49, nota-se que as áreas foco possuem relevância no domínio de dados abertos, pois a média das respostas obtidas para estas, ficaram entre 3 (Muito Importante) e 4 (Indispensável).

Em relação à percepção dos componentes do modelo, observados no Quadro 49, nota-se que apenas **Monitoramento e Medição** e **Licença** obtiveram média geral inferior a 3. Os demais componentes ficaram acima de 3, mesmo obtendo pontuação 0, caso do componente **identificação**, pontuado pelo especialista ⁷¹.

Por fim, alguns especialistas discutiram durante a entrevista e até mesmo apontaram através do formulário sugestões para o modelo, sendo estas:

1. Pensar na inclusão do conceito de Governo Aberto, acrescentar na matriz, talvez em "Definição dos dados", a interação com a sociedade na priorização de quais dados serão abertos.
2. Incluir questões relacionadas à "Ligação com outras bases", ou algo neste sentido, na parte de gerenciamento de padrões. Nesse campo, poderia ser desde ligação com bases da própria organização, até bases externas.

¹ O especialista não especificou o motivo desta pontuação

3. Incluir algum item para avaliar os objetivos da organização;
4. Pensar na capacidade "Em construção";
5. Pensar sobre a quantidade de níveis (pois, pode desmotivar algumas organizações), visto que é um número considerável de questões a serem respondidas;
6. Incluir informações sobre os componentes e áreas foco sobre o que está sendo considerado para a avaliação;
7. Discutir os princípios da Lei Geral de Proteção dos Dados (LGPD) junto ao modelo de maturidade, no sentido da LGPD não ser utilizada para se sobressair em relação à abertura dos dados;
8. Considerar a introdução de um componente sobre financiamento e sustentabilidade das iniciativas;
9. A automação de processos, poderia ser considerada em níveis mais altos de maturidade;
10. Em organizações públicas, é importante considerar a dimensão de participação, especialmente na definição, no planejamento e priorização;

As sugestões apresentadas em sua maioria demandam análise e discussões para introduzir junto ao modelo e para isso estas serão apontadas nos trabalhos futuros. Com exceção das sugestões 2 e 6 que já encontram-se contempladas no trabalho. Os demais comentários acerca da pesquisa, estão na sequência:

- "Parabéns pela definição e esquematização da pesquisa. Com certeza a mesma poderá ajudar as organizações a analisarem sua maturidade relativa à publicação de dados abertos, assim como auxiliar no entendimento dos pontos que precisam ser melhorados";
- "Acredito que o modelo apresentado permitiu encapsular os elementos mais importantes em relação à publicação de dados abertos. Penso que pode ser aprimorado com uma explicação aos respondentes";
- "Excelente trabalho";
- "O modelo abrange muito bem os diversos aspectos a que uma organização deve atentar no processo de publicação de dados abertos. De um lado, demonstra uma preocupação com os critérios técnicos, mais relacionados à licenças, formatos, metadados e vocabulários, e que são objeto de foco específico dos profissionais responsáveis pela disponibilização dos dados, mas que atentam, também, para as necessidades daqueles que consomem tais dados. Sob outro aspecto, foram

relacionados os processos organizacionais e à Gestão do Conhecimento. Estes estão cobertos, portanto, critérios afeitos à educação e treinamento das equipes, observação de processos organizacionais, dentre outros que contribuem de forma geral para a organização. Parabéns pelo trabalho, que tem grande potencial de contribuição para o desenvolvimento da cultura da transparência e da máxima capacidade de uso de dados".

Após compor, apresentar e avaliar o modelo de maturidade de dados abertos o Quadro 50 visa abordar um comparativo entre os dois modelos de maturidade discutidos na seção 2.4 e o modelo de maturidade que foi elaborado através desta pesquisa, com base em como foram desenvolvidos, como abordam a avaliação da maturidade, quais os níveis de maturidade e como estes foram validados.

Quadro 50 – Trabalhos relacionados - comparativo

Modelo	Desenvolvido com base:	Avalia com base em:	Níveis de maturidade	Validação
(DODDS; NEW-MAN, 2015)	<i>Workshops</i> com a participação de administradores de dados e especialistas em dados abertos do Reino Unido e da comunidade de dados abertos em geral.	15 atividades organizacionais, sendo que estas atividades estão agrupadas em 5 temas.	1- Inicial, 2- Re-produtivo, 3- Definido, 4- Gerenciado, 5-Otimizado	Não há - No entanto, os autores pretendem avaliar e ampliar o modelo no futuro com base na experiência.
(SOLAR <i>et al.</i> , 2012)	Trabalhos e artigos publicados, relatórios de comitês, bem como entrevistas telefônicas, reuniões e teleconferências. A equipe de trabalho contactou vários grupos interessados em dados abertos para entrevistá-los e usá-los como fontes primárias de informação.	33 variáveis críticas, distribuídas em 9 subdomínios agrupadas em 3 domínios.	1- Capacidade Inexistente, 2- Capacidades emergentes, 3- Capacidades existentes, 4- Capacidades avançadas	O modelo foi validado através de entrevistas com especialistas da América Latina e através de uma ferramenta Web que permitiu a coleta de dados por meio de um piloto "in situ".

Modelo de maturidade de dados abertos	Boas práticas para dados na Web; Fatores importantes para dados abertos; Padrões para dados na Web; Atividades e princípios da governança de dados.	3 áreas foco e 21 componentes distribuídos nas 3 áreas foco	Considera 13 níveis de maturidade, sendo que 0 representa a inexistência até o 12 que representa a existência da (A) Definição, da (B) Implementação, do (C) Monitoramento e do (D) Aprimoramento para um determinado componente	Apresentação através de cenários de uso reais e avaliação através de entrevistas com especialistas da área.
---------------------------------------	---	---	--	---

Fonte: Da Autora

As iniciativas descritas no Quadro 50 contribuem fornecendo diferentes perspectivas de avaliação sobre dados abertos, no entanto percebe-se que em nenhum dos trabalhos encontrados abordam modelos de maturidade orientados a áreas foco e estes também não consideram as atividades e princípios da governança de dados.

7 CONSIDERAÇÕES FINAIS

As considerações finais destacam os resultados e avanços alcançados com a realização desta pesquisa, considerando os objetivos traçados e também abordam-se as possibilidades para trabalhos futuros.

7.1 CONCLUSÕES

A abertura dos dados por parte de organizações públicas ou privadas é necessária e importante para que se obtenha transparência, *accountability*, participação social e colaboração. Neste sentido, observa-se que diversas organizações realizam a abertura dos dados apenas com o intuito de cumprir leis e regulamentações e com isso acabam disponibilizando dados sem princípios mínimo para que estes possam ser utilizados.

Ao fazer uso de práticas e padrões para dados abertos, potencializa-se os benefícios acerca destes. Assim, o presente trabalho apresentou contribuições fornecendo uma percepção ideal de como a organização está em relação ao que propõe-se na literatura e neste sentido viabiliza-se para a organização uma reflexão sobre a necessidade de **I)** ter políticas e processos, mesmo que simples, mas aprimora-los com o passar do tempo, **II)** se envolver com os diversos *stakeholders* para compreender as necessidades, dificuldades e até mesmo compreender como estão sendo utilizados os dados e **III)** aprimorar as suas estratégias sobre dados abertos, considerando os recursos que a organização disponibiliza. Sendo estes itens contemplados no modelo de maturidade de dados abertos que foi apresentado.

Para realizar o objetivo de identificar boas práticas, etapas e atividades de abertura e pós-publicação de dados foi realizada uma revisão da literatura, sendo possível identificar as práticas, etapas, atividades e também fatores importantes para dados abertos. Com isso, obteve-se um ciclo de vida de dados abertos, este apresentado no capítulo 4, sendo este abordado e discutido no trabalho com o intuito de delimitar o domínio do modelo de maturidade e também para contemplar o objetivo do trabalho de: compor um ciclo de vida para dados abertos, considerando as boas práticas aplicadas pelas organizações.

Destaca-se que o ciclo de vida apresentado visa ser um modelo de referência para que este possa ser estendido, reduzido e até mesmo adaptado, ou seja, o ciclo de vida possibilitará que modelos mais específicos ou mais genéricos sejam desenvolvidos.

Deste modo, para alcançar o objetivo de identificar os componentes da governança de dados que corroborar com o modelo de maturidade fez-se novamente uma revisão de literatura e chegou-se ao trabalho de Panian (2010) e Alhassan *et al.* (2016),

estes utilizados na composição do modelo.

Assim, para alcançar o objetivo de representar o conhecimento dos elementos que compõem o modelo de maturidade de dados abertos fez-se o uso do domínio determinado pelo ciclo de vida e também utilizou-se das etapas de escopo e design apresentadas pelo trabalho de De Bruin *et al.* (2005). Ainda, na perspectiva de compor o modelo de maturidade fez-se o uso da etapa de design do modelo apresentada em Van Steenbergem *et al.* (2010a).

Assim, obteve-se uma estrutura de modelo de maturidade caracterizada por 3 áreas foco, estas: **I)** Gerenciamento de políticas e processos, **II)** Gerenciamento de fatores organizacionais e **III)** Gerenciamento de padrões. Cada área foco possui componentes e para cada um destes verifica-se as capacidades de A - Definir, B - Implementar, C - Monitorar e D - Aprimorar. Deste modo viabilizou-se a visualização da maturidade através de uma matriz, onde é possível observar os componentes que se destacam, assim como os componentes que necessitam de incentivos para que estes obtenham êxito e também foi apresentada uma representação dos conhecimentos abordados ao final do capítulo.

Neste sentido, para atingir o objetivo de desenvolver um instrumento para a aplicação do modelo de maturidade, foi formulado um questionário de avaliação da maturidade considerando as áreas foco, cada componente em específico e as capacidades definidas. E para auxiliar na resolução do questionário e na obtenção do resultado foi desenvolvida uma aplicação Web¹. Através desta aplicação Web, fez-se a coleta de três cenários de uso reais por meio de entrevistas, com isso foi possível compreender o comportamento do modelo e também obter algumas sugestões de possíveis melhorias. E deste modo também foi apresentado o modelo e seu comportamento, por meio de cenários de uso.

Por sua vez, após a realização das entrevistas para a coleta dos cenários de uso, surgiu a ideia de ter uma versão "simplificada" do modelo de maturidade em questão. Para isso, foi elaborado um novo questionário e foram realizadas entrevistas com especialistas da área de dados abertos. Com base nas respostas obtidas e nas discussões realizadas durante as entrevistas não foi possível chegar a uma estrutura "simplificada" para o modelo de maturidade, visto que a média das percepções elencadas para cada componente resultou em um grau de importância considerável. No entanto, destaca-se que a estrutura apresentada foi bem aceita o que demonstra a relevância e a importância na realização deste estudo. E assim, atingiu-se o objetivo de avaliar o modelo de maturidade com especialistas da área.

Com base nos resultados obtidos através dos cenários de uso analisados e em decorrência das entrevistas com os especialistas da área de dados abertos concluiu-se que o modelo de maturidade possibilita avaliar a maturidade de organização que

¹ Disponível em: <https://lidianewisintin.studiokem.ufsc.br/>

realizam a abertura de dados e esta pesquisa pode servir de referência para possíveis desdobramentos em novos modelos de maturidade, de modo a atender cenários em específico.

Deste modo, conclui-se que os objetivos específicos foram alcançados, conforme abordados previamente e o objetivo geral, "*propor um modelo de maturidade de dados abertos, considerando princípios da governança de dados e as boas práticas aplicadas pelas organizações.*" foi atingido e a pergunta de pesquisa, sendo esta: "Como aplicar os princípios da governança de dados e as boas práticas para abertura e pós-publicação de dados abertos, considerando o contexto organizacional?" foi respondida. Sendo que, o artefato apresentado abrange boas práticas, fatores importantes para dados abertos e alguns princípios da governança de dados, que são relevantes para as organizações.

7.2 TRABALHOS FUTUROS

Na realização deste trabalho, algumas possibilidades para trabalhos futuros foram vislumbradas, sendo estas apontadas nos itens abaixo:

- Realizar a adaptação do sistema para os usuários das organizações selecionam quais áreas foco e quais componentes são de seu interesse para que seja gerado um questionário adequado para cada cenário em específico. Com base nas necessidades de cada organização, o que viabilizaria flexibilidade da estrutura apresentada;
- Produzir estudos de caso considerando cenários similares para obter uma estrutura mínima necessária, por exemplo, para organização que tem por objetivo apenas obter a transparência através da abertura dos dados;
- Produzir estudos de caso, para evidenciar os componentes que devem ser primordiais para níveis de governo. Não considerando apenas o nível de governo, mas o objetivo apresentado no plano de dados abertos;
- Adaptar o sistema Web desenvolvido para torná-lo dinâmico, interativo e flexível para a resolução do questionário;
- Adaptar o sistema Web desenvolvido para explicitar o que visa-se com a avaliação de cada área foco e de cada componente, permitindo clarificar os termos utilizados e gerando autonomia para aqueles que desejam realizar a avaliação da maturidade em suas organizações;
- Analisar e discutir a possibilidade da inclusão dos princípios de Governo Aberto e de Open Government Partnership (OGP) no modelo de maturidade;

- Analisar e Discutir a LGPD, junto ao modelo de maturidade;
- Através da representação dos conhecimentos abordados, considerando como base alguns dos princípios de ontologias, desta maneira observa-se que podem ser derivados novos estudos no âmbito semântico e computacional.
- Apresentar uma ontologia com base no ciclo de vida de dados abertos;
- Apresentar uma ontologia com base nos conhecimentos abordados acerca do modelo de maturidade de dados abertos;
- Analisar e discutir a gestão do conhecimento junto ao ciclo de vida de dados abertos;
- Analisar e discutir a engenharia do conhecimento junto ao ciclo de vida de dados abertos;
- Contextualizar e discutir as contribuições da engenharia, da gestão e da mídia do conhecimento junto ao modelo de maturidade;
- Realizar estudos considerando a memória organizacional, com base no ciclo de vida de dados abertos e também considerando as avaliações realizadas para aferir a maturidade;
- Produzir novos estudo na perspectiva de acompanhar a evolução ao longo do tempo de uma determinada organização, percebendo se há evolução ou não no contexto de dados abertos;
- Realizar estudos comparando a maturidade identificada em relação ao que está sendo aplicado no portal de dados;
- Analisar e discutir a possibilidade da inclusão de aspectos de segurança dos dados acerca da disponibilização e das alterações realizadas em conjuntos e recursos, por exemplo uso de *blockchain*.

REFERÊNCIAS

- ADLER, M.J. **A Guidebook to Learning: For a Lifelong Pursuit of Wisdom.** [S.l.]: Macmillan, 1986. ISBN 9780025003408. Disponível em: <https://books.google.com.br/books?id=E0juAAAAMAAJ>.
- AKYÜREK, Haydar; SCHOLL, Cora; STODDEN, Regina; SIEBENLIST, Tobias; MAINKA, Agnes. **Maturity and usability of open data in North Rhine-Westphalia.** *In: PROCEEDINGS of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age.* [S.l.: s.n.], 2018. P. 1–10.
- ALHASSAN, Ibrahim; SAMMON, David; DALY, Mary. Data governance activities: an analysis of the literature. **Journal of Decision Systems**, Taylor & Francis, v. 25, sup1, p. 64–75, 2016.
- ALROMAIIH, Nouf; ALBASSAM, Hend; AL-KHALIFA, Hend. **A proposed checklist for the technical maturity of open government data: an application on GCC countries.** *In: ACM. PROCEEDINGS of the 18th International Conference on Information Integration and Web-based Applications and Services.* [S.l.: s.n.], 2016. P. 494–499.
- ALSHEHHI, Maryam; LEIDA, Marcello; HIRSCH, Benjamin; YOO, Paul D; TAHA, Kamal. **Visual analytics in the web of data.** *In: IEEE. ELECTRONICS, Circuits, and Systems (ICECS), 2013 IEEE 20th International Conference on.* [S.l.: s.n.], 2013. P. 102–103.
- ARMAN, Arry Akhmad; RAMADHAN, Gilang; FAJRIN, Muhammad. **Design of Data Management Guideline for Open Data Implementation: Case Study in Indonesia.** *In: PROCEEDINGS of the 2015 2Nd International Conference on Electronic Governance and Open Society: Challenges in Eurasia.* St. Petersburg, Russian Federation: ACM, 2015. (EGOSE '15), p. 17–23. DOI: 10.1145/2846012.2846024. Disponível em: <http://doi.acm.org/10.1145/2846012.2846024>.
- ATTARD, Judie; ORLANDI, Fabrizio; SCERRI, Simon; AUER, Sören. A systematic review of open government data initiatives. **Government Information Quarterly**, v. 32, n. 4, p. 399–418, 2015. ISSN 0740-624X. DOI: <https://doi.org/10.1016/j.giq.2015.07.006>.

ATZ, U; HEATH, T; FAWCET, J. Benchmarking open data automatically. **London: Open Data Institute**, 2015.

AUER, Sören; LEHMANN, Jens; NGOMO, Axel-Cyrille Ngonga. **Introduction to linked data and its lifecycle on the web**. In: REASONING Web. Semantic Technologies for the Web of Data. [S.l.]: Springer, 2011. P. 1–75.

BARBOSA, Luciano; PHAM, Kien; SILVA, Claudio; VIEIRA, Marcos R; FREIRE, Juliana. **Structured open urban data: understanding the landscape**. **Big data**, Mary Ann Liebert, Inc. 140 Huguenot Street, 3rd Floor New Rochelle, NY 10801 USA, v. 2, n. 3, p. 144–154, 2014.

BARNAGHI, Payam; WANG, Wei; HENSON, Cory; TAYLOR, Kerry. Semantics for the Internet of Things: early progress and back to the future. **International Journal on Semantic Web and Information Systems (IJSWIS)**, IGI Global, v. 8, n. 1, p. 1–21, 2012.

BAUER, Florian; KALTENBÖCK, Martin. Linked open data: The essentials. **Edition mono/monochrom, Vienna**, v. 710, 2011.

BECKER, Jörg; KNACKSTEDT, Ralf; PÖPPELBUSS, Jens. Developing maturity models for IT management. **Business & Information Systems Engineering**, Springer, v. 1, n. 3, p. 213–222, 2009.

BEGG, Carolyn; CAIRA, Tom. Exploring the SME quandary: Data governance in practise in the small to medium-sized enterprise sector. **Electronic Journal of Information Systems Evaluation**, Academic Conferences International Limited, v. 15, n. 1, 2012.

BERG, Martin van den; STEENBERGEN, Marlies van. **Building an Enterprise Architecture Practice: Tools, Tips, Best Practices, Ready-to-Use Insights (The Enterprise Series)**. [S.l.]: Springer-Verlag New York, Inc., 2006.

BERNERS-LEE, Tim. Linked Data Design Issues (July 2006). **URL= <http://www.w3.org/DesignIssues/LinkedData.html>**, última visita julho de 2018, 2010.

BERNERS-LEE, Tim. Linked Data Design Issues (July 2006). **URL= <http://www.w3.org/DesignIssues/LinkedData.html>**, última visita julho de 2018, 2011.

BERNERS-LEE, Tim; HENDLER, James; LASSILA, Ora. The Semantic Web: A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities [Electronic resource]. **Scientific American**, 2001.

BIZER, Chris; CYGANIAK, Richard; HEATH, Tom *et al.* How to publish linked data on the web. **URL=<http://wifo5-03.informatik.uni-mannheim.de/bizer/HowtoPublishLinkedData.html>**, última visita agosto de 2018, 2007.

BIZER, Christian; HEATH, Tom; BERNERS-LEE, Tim. **Linked data: The story so far**. *In*: SEMANTIC services, interoperability and web applications: emerging concepts. [S.l.]: IGI Global, 2011. P. 205–227.

BROEK, Tijds; VEENSTRA, Anne Fleur; FOLMER, Erwin. **Walking the extra byte: a lifecycle model for Linked Open Data**. Linked Open Data, 2014.

CAMPOS, Paula Assumpção. **Dados Abertos Governamentais: Desafios na Publicação**. 2018. Diss. (Mestrado) – Universidade Federal de Santa Catarina.

CARALLI, Richard; KNIGHT, Mark; MONTGOMERY, Austin. **Maturity models 101: A primer for applying maturity models to smart grid security, resilience, and interoperability**. [S.l.], 2012.

CASAES, Júlio César Costa. **Governança de Dados Abertos Governamentais: Framework Conceitual para as Universidades Federais, Baseado em uma Visão Sistêmica**. 2019. Tese (Doutorado) – Universidade Federal de Santa Catarina.

CASTELLANI RIBEIRO, Daniel; VO, Huy T.; FREIRE, Juliana; SILVA, Cláudio T. **An Urban Data Profiler**. *In*: PROCEEDINGS of the 24th International Conference on World Wide Web. Florence, Italy: ACM, 2015. (WWW '15 Companion), p. 1389–1394. DOI: 10.1145/2740908.2742135. Disponível em: <http://doi.acm.org/10.1145/2740908.2742135>.

CHAPPLE, Mike. Speaking the same language: Building a data governance program for institutional impact. **Educause Review**, v. 48, n. 6, p. 14–16, 2013.

CHEONG, Lai Kuan; CHANG, Vanessa. The need for data governance: a case study. **ACIS 2007 Proceedings**, p. 100, 2007.

CHIESA, Vittorio; COUGHLAN, Paul; VOSS, Chris A. Development of a technical innovation audit. **Journal of Product Innovation Management: AN INTERNATIONAL PUBLICATION OF THE PRODUCT DEVELOPMENT & MANAGEMENT ASSOCIATION**, Wiley Online Library, v. 13, n. 2, p. 105–136, 1996.

CHIGNARD, Simon. A brief history of Open Data. **Paris Tech Review**, v. 29, mar. 2013. <http://parisinnovationreview.com/articles-en/a-brief-history-of-open-data>.

CKAN, Comprehensive Knowledge Archive Network. **About Comprehensive Knowledge Archive Network**. Jun. 2021. Disponível em: <https://docs.ckan.org/en/2.9/contents.html>.

CROSBY, P.B. **Quality is Free: The Art of Making Quality Certain**. [S.l.]: McGraw-Hill, 1979. (Mentor book). ISBN 9780070145122. Disponível em: <https://books.google.com.br/books?id=n4IubCcpm0EC>.

CUPANI, Alberto. La peculiaridad del conocimiento tecnológico. **Scientiae Studia**, v. 4, n. 3, p. 353–371, 2006.

DAMA INTERNATIONAL. **DAMA-DMBOK: Data Management Body of Knowledge (2nd Edition)**. second. [S.l.]: Technics Publications, jul. 2017. v. 1.

DATA REVOLUTION FOR SUSTAINABLE DEVELOPMENT, Independent Expert Advisory Group on a. **A world that counts. Mobilising the data revolution for sustainable development**. [S.l.]: Independent Expert Advisory Group Secretariat New York, 2014.

DATAFLUX CORPORATION. **The Data Governance Maturity Model Establishing the People, Policies and Technology That Manage Enterprise Data**. [S.l.: s.n.], 2007. http://www.fstech.co.uk/fst/whitepapers/The_Data_Governance_Maturity_Model.pdf. Acesso em:25/10/2018.

DE BRUIN, Tonia; ROSEMANN, Michael; FREEZE, Ronald; KAULKARNI, Uday. **Understanding the main phases of developing a maturity assessment model**. In: AUSTRALASIAN CHAPTER OF THE ASSOCIATION FOR INFORMATION SYSTEMS. AUSTRALASIAN Conference on Information Systems (ACIS): [S.l.: s.n.], 2005. P. 8–19.

DEMING, W Edwards. **Elementary principles of the statistical control of quality: a series of lectures**. [S.l.]: Nippon Kagaku Gijutsu Remmei, 1952.

DENG, Qi; JI, Shaobo. A review of design science research in information systems: concept, process, outcome, and evaluation. **Pacific Asia journal of the association for information systems**, v. 10, n. 1, p. 2, 2018.

DERNIAME, Jean-Claude; KABA, Badara A; WASTELL, David. **Software process: principles, methodology, and technology**. [S.l.]: Springer Science & Business Media, 1999.

DILLO, Ingrid; DE LEEUW, Lisa. CoreTrustSeal. **Mitteilungen der Vereinigung Österreichischer Bibliothekarinnen und Bibliothekare**, v. 71, n. 1, p. 162–170, 2018.

DIN, Siti Haslinda Mohd; MANSOR, Nur Aziha; ARUS, Faiza Rusrianti Tajul. **OPEN DATA AND CHALLENGES FACED NY NATIONAL STATISTICS OFFICES**. *In*: MALAYSIA STATISTICS CONFERENCE (MyStats 2015). [S.l.: s.n.], nov. 2015. P. 77.

DODDS, Leigh; NEWMAN, Andrew. **Guia do Modelo de Maturidade de Dados Abertos**. [S.l.]: Open Data Institute, 2015. P. 1–53. Disponível em: http://ceweb.br/media/docs/publicacoes/13/Guia_Modelo_de_Maturidade_de_Dados_Abertos.pdf.

DRESCH, A.; LACERDA, D.P.; ANTUNES, J.A.V. **Design Science Research: A Method for Science and Technology Advancement**. [S.l.]: Springer, 2014. ISBN 9783319073750. Disponível em: <https://books.google.com.br/books?id=-SQAswEACAAJ>.

DRESCH, A.; LACERDA, D.P.; ANTUNES, J.A.V. **Design science research: método de pesquisa para avanço da ciência e tecnologia**. [S.l.]: Bookman Editora, 2015.

EGC. **Áreas de Concentração**. [S.l.: s.n.], jun. 2021a. <http://www.egc.ufsc.br/pos-graduacao/programa/areas-de-concentracao/>.

EGC. **Interação das Áreas na Busca do Objeto de Pesquisa do Programa**. [S.l.: s.n.], jun. 2021b. <http://www.egc.ufsc.br/index.php/pt/pesquisas/linhas-de-pesquisa>.

ELMAALLAM, Mina; KRIOUILE, Abdelaziz. A generic process for the development and the implementation of IS maturity models. **International Journal of Computer Science Issues (IJCSI)**, International Journal of Computer Science Issues (IJCSI), v. 11, n. 6, p. 34, 2014.

ERPEN, Julio Graeff. **Pecuária Intensiva em Conhecimento: Modelo de Maturidade em Gestão do Conhecimento Aplicado a Bovinocultura de Corte**. 2019. Tese (Doutorado) – Universidade Federal de Santa Catarina.

ESTERMANN, Beat. Development paths towards open government—an empirical analysis among heritage institutions. **Government Information Quarterly**, Elsevier, v. 35, n. 4, p. 599–612, 2018.

EUROPEAN DATA PORTAL. **Benefits of Open Data**. [S.l.: s.n.], nov. 2018. <https://www.europeandataportal.eu/en/using-data/benefits-of-open-data>.

FERRARA, Alfio; GENTA, Lorenzo; MONTANELLI, Stefano. **Tailoring linked data exploration through inCloud filtering**. In: ACM. PROCEEDINGS of the 2012 Joint EDBT/ICDT Workshops. [S.l.: s.n.], 2012. P. 140–143.

FIGUEIRA, Felipe Louro. **Proposta de Um Modelo de Maturidade de Gestão de Projetos, que Incorpora os Processos Sistêmicos de Gestão do Conhecimento**. 2020. Diss. (Mestrado) – Universidade Federal de Santa Catarina.

FISHER, Tony. **The data asset: How smart companies govern their data for business success**. [S.l.]: John Wiley & Sons, 2009. v. 24.

FLORES, Heriberto Alzerino. **Maturidade em gestão do conhecimento na administração pública: um estudo na Prefeitura Municipal de São José/SC**. 2019. Diss. (Mestrado) – Universidade Federal de Santa Catarina.

FRASER, P; FARRUKH, Cu; GREGORY, M. Managing product development collaborations—a process maturity approach. **Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture**, Sage Publications Sage UK: London, England, v. 217, n. 11, p. 1499–1519, 2003.

FRASER, P; MOULTRIE, J; HOLDWAY, R. **Exploratory studies of a proposed design maturity model**, 2001.

FRASER, P.; MOULTRIE, J.; GREGORY, M. **The use of maturity models/grids as a tool in assessing product development capability.** *In*: IEEE International Engineering Management Conference. [S.l.: s.n.], ago. 2002. 244–249 vol.1. DOI: 10.1109/IEMC.2002.1038431.

GLEASURE, Rob; FELLER, Joseph; O'FLAHERTY, Brian F. **Procedurally transparent design science research: A design process model**, 2012.

GLOBAL PARTNERSHIP FOR SUSTAINABLE DEVELOPMENT DATA. **SDG Data Revolution Roadmaps Toolbox: Open Data Module.** [S.l.: s.n.], jul. 2016. Disponível em: <http://www.data4sdgs.org/sites/default/files/2017-09/0pen%20Data%20for%20Sustainable%20Development%20-%20Data4SDGs%20Toolbox%20Module.pdf>.

GOLDSTEIN, B.; DYSON, L.; NEMANI, A. **Beyond Transparency: Open Data and the Future of Civic Innovation.** [S.l.]: Code for America Press, 2013. ISBN 9780615889085. Disponível em: <https://books.google.com.br/books?id=uzbxngEACAAJ>.

GOMES, Murilo Silveira. **Proposta de Arquitetura para Ecossistema de Inovação em Dados Abertos.** 2017. Diss. (Mestrado) – Universidade Federal de Santa Catarina.

GOTTSCHALK, Petter; SOLLI-SÆTHER, Hans. Maturity model for IT outsourcing relationships. **Industrial Management & Data Systems**, Emerald Group Publishing Limited, v. 106, n. 2, p. 200–212, 2006.

GRAY, Jonathan. **Towards a genealogy of open data.** *In*: THE paper was given at the General Conference of the European Consortium for Political Research in Glasgow. [S.l.: s.n.], 2014.

HAIDER, Waqar; HAIDER, Abrar. **Governance structures for engineering and infrastructure asset management.** *In*: 2013 Proceedings of PICMET '13: Technology Management in the IT-Driven Services (PICMET). [S.l.: s.n.], jul. 2013. P. 1229–1238.

HAUCK, Jean Carlo Rossa. **Um Método de Aquisição de Conhecimento para Customização de Modelos de Capacidade/Maturidade de Processos de Software.** 2011. Tese (Doutorado) – Universidade Federal de Santa Catarina.

HAUSENBLAS, Michael. **Linked Data life cycles**. [S.l.: s.n.], 2011.

<http://www.slideshare.net/mediasemanticweb/linked-data-life-cycles>.

HEATH, Tom; BIZER, Christian. Linked data: Evolving the web into a global data space.

Synthesis lectures on the semantic web: theory and technology, Morgan & Claypool Publishers, v. 1, n. 1, p. 1–136, 2011.

HELOU, Angela Regina Heinzen Amin. **Avaliação da Maturidade da Gestão do Conhecimento na Administração Pública**. 2015. Tese (Doutorado) – Universidade Federal de Santa Catarina.

HITZLER, Pascal; JANOWICZ, Krzysztof. Linked Data, Big Data, and the 4th Paradigm. **Semantic Web**, v. 4, n. 3, p. 233–235, 2013.

HSIEH, Ping Jung; LIN, Binshan; LIN, Chinho. The construction and application of knowledge navigator model (KNM™): An evaluation of knowledge management maturity. **Expert Systems with Applications**, Elsevier, v. 36, n. 2, p. 4087–4100, 2009.

HUANG, Ruhua; WANG, Chunying; ZHANG, Xiaoyu; WU, Dan; XIE, Qingwen. Design, develop and evaluate an open government data platform: a user-centred approach. **The Electronic Library**, Emerald Publishing Limited, 2019.

HUIJBOOM, Noor; VAN DEN BROEK, Tijs. Open data: an international comparison of strategies. **European journal of ePractice**, v. 12, n. 1, p. 4–16, 2011.

HUMPHREY, W. S. Characterizing the software process: a maturity framework. **IEEE Software**, v. 5, n. 2, p. 73–79, mar. 1988. ISSN 0740-7459. DOI: 10.1109/52.2014.

HUMPHREY, Watts S. **Characterizing the Software Process: A Maturity Framework**. [S.l.], 1987.

HUMPHREY, Watts S. **Managing the software process**. [S.l.]: Addison-Wesley Reading, MA, 1989. v. 1.

HUMPHREY, Watts S; SWEET, WL. **A Method for assessing the software engineering capability of contractors: preliminary version**. [S.l.], 1987.

HUSIN, Nik Nor Farah Annysa; ZAKARIA, Nor Hidayati; DAHLAN, Halina Mohamed. **Factors Influencing Open Data Adoption in Malaysia Based on Users Perspective.** *In*: 2019 6th International Conference on Research and Innovation in Information Systems (ICRIIS). [S.l.: s.n.], 2019. P. 1–5. DOI: 10.1109/ICRIIS48246.2019.9073396.

HYLAND, Bernadette; ATEMEZING, Ghislain; VILLAZÓN-TERRAZAS, Boris. **Best practices for publishing linked data.** World Wide Web Consortium, 2014. Disponível em: <https://www.w3.org/TR/ld-bp/>.

HYLAND, Bernadette; WOOD, David. The Joy of Data - A Cookbook for Publishing Linked Government Data on the Web. *In*: **Linking Government Data.** Edição: David Wood. New York, NY: Springer New York, 2011. P. 3–26. ISBN 978-1-4614-1767-5. DOI: 10.1007/978-1-4614-1767-5_1. Disponível em: http://dx.doi.org/10.1007/978-1-4614-1767-5_1.

JANSSEN, Marijn; ZUIDERWIJK, Anneke. **Open data and transformational government.** *In*: TRANSFORMING Government Workshop. Brunel University, United Kingdom. [S.l.: s.n.], 2012.

JUNIOR, Vanderlei FREITAS; FLAVIO, CECI; WOSZEZENKI, Cristiane Raquel; LEOPOLDO, Alexandre. Design Science Research Methodology Enquanto Estratégia Metodológica para a Pesquisa Tecnológica. **Revista Spacios**, v. 38, n. 6, 2017.

KAASENBROOD, Maaïke; ZUIDERWIJK, Anneke; JANSSEN, Marijn; JONG, Martin de; BHAROSA, Nitesh. Exploring the factors influencing the adoption of open government data by private organisations. **International Journal of Public Administration in the Digital Age (IJPADA)**, IGI Global, v. 2, n. 2, p. 75–92, 2015.

KALAMPOKIS, Evangelos; TAMBOURIS, Efthimios; TARABANIS, Konstantinos. **Open government data: A stage model.** *In*: ELECTRONIC government. [S.l.]: Springer, 2011. P. 235–246.

KAMIOKA, Taro; LUO, Xubin; TAPANAINEN, Tommi. **An empirical investigation of data governance: the role of accountabilities.** *In*: ASSOCIATION FOR INFORMATION SYSTEM. PACIFIC Asia Conference On Information Systems (PACIS). [S.l.: s.n.], 2016.

KAPOOR, Kawaljeet; WEERAKKODY, Vishanth; SIVARAJAH, Uthayasankar. **Open data platforms and their usability: Proposing a framework for evaluating citizen**

intentions. *In:* SPRINGER. CONFERENCE on e-Business, e-Services and e-Society. [S.l.: s.n.], 2015. P. 261–271.

KHATRI, Vijay; BROWN, Carol V. Designing Data Governance. **Commun. ACM**, ACM, New York, NY, USA, v. 53, n. 1, p. 148–152, jan. 2010. ISSN 0001-0782. DOI: 10.1145/1629175.1629210. Disponível em: <http://doi.acm.org/10.1145/1629175.1629210>.

KIM, Hee Young; CHO, June-Suh. **Data governance framework for big data implementation with a case of Korea.** *In:* IEEE. 2017 IEEE International Congress on Big Data (BigData Congress). [S.l.: s.n.], 2017. P. 384–391.

KITCHIN, R. **The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences.** [S.l.]: SAGE Publications, 2014. ISBN 9781473908260. Disponível em: <https://books.google.com.br/books?id=Gf0ICwAAQBAJ>.

KLEIN, Vinicius Barreto. **Uma Proposta de Modelo Conceitual para Uso de Big Data E Open Data para Smart Cities.** 2016. Diss. (Mestrado) – Universidade Federal de Santa Catarina.

KLOSS, Linda. **Implementing Health Information Governance: Lessons from the Field.** [S.l.]: American Health Information Management Association, 2015.

KNOWLEDGE, OPEN. **How to Open Data.** [S.l.: s.n.], 2018. <https://okfn.org/opendata/how-to-open-data>. 25/03/2018.

KOOMEN, Tim; POL, Martin. **Test Process Improvement: A step-by-step guide to structured testing.** [S.l.]: Addison-Wesley, 1999.

KRAEMER, Rodrigo. **Modelo de Maturidade do Sistema de Educação Corporativa.** 2018. Diss. (Mestrado) – Universidade Federal de Santa Catarina.

KUCERA, Jan; CHLAPEK, Dusan; KLIÉMEK, Jakub; NECASKÝ, Martin. **Methodologies and Best Practices for Open Data Publication.** *In:* DATESO. [S.l.: s.n.], 2015. P. 52–64.

LANE, Stephen; RICHARDSON, Ita. Process models for service-based applications: A systematic literature review. **Information and Software Technology**, v. 53, n. 5, p. 424–439, 2011. Special Section on Best Papers from XP2010. ISSN 0950-5849.

DOI: <https://doi.org/10.1016/j.infsof.2010.12.005>. Disponível em:
<http://www.sciencedirect.com/science/article/pii/S0950584910002211>.

LEE, Gwanhoo; KWAK, Young Hoon. An open government maturity model for social media-based public engagement. **Government information quarterly**, Elsevier, v. 29, n. 4, p. 492–503, 2012.

LEE, Sung Une; ZHU, Liming; JEFFERY, Ross. Data governance for platform ecosystems: Critical factors and the state of practice. **arXiv preprint arXiv:1705.03509**, 2017.

LÓSCIO, Bernadette; BURLE, Caroline; CALEGARI, Newton. Boas práticas para dados na WEB: desafios e benefícios. **Revista Principia**, Divulgação Científica e Tecnológica do IFPB, n. 32, p. 9–18, 2016.

LÓSCIO, Bernadette; BURLE, Caroline; CALEGARI, Newton. **Data on the Web Best Practices**. [S.l.: s.n.], 2017. <https://www.w3.org/TR/dwbp/>.

MACEDO, Dirceu; LEMOS, Daniela. Dados abertos governamentais: iniciativas e desafios na abertura de dados no Brasil e outras esferas internacionais. **AtoZ: novas práticas em informação e conhecimento**, v. 10, n. 2, p. 14–26, 2021. ISSN 2237-826X. DOI: 10.5380/atoz.v10i2.77737. Disponível em:
<https://revistas.ufpr.br/atoz/article/view/77737>.

MACHADO, Luís Miguel Oliveira; SOUZA, Renato Rocha; GRAÇA SIMÕES, Maria da. Semantic web or web of data? a diachronic study (1999 to 2017) of the publications of tim berners-lee and the world wide web consortium. **Journal of the Association for Information Science and Technology**, v. 70, n. 7, p. 701–714, 2019. DOI:
<https://doi.org/10.1002/asi.24111>. Disponível em:
<https://asistdl.onlinelibrary.wiley.com/doi/abs/10.1002/asi.24111>.

MANYIKA, James; CHUI, Michael; FARRELL, Diana; KUIKEN, Steve Van; GROVES, Peter; DOSHI, Elizabeth Almasi. **Open data: Unlocking innovation and performance with liquid information**. [S.l.: s.n.], out. 2013.
<https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/open-data-unlocking-innovation-and-performance-with-liquid-information>.

MARCH, Salvatore T; SMITH, Gerald F. Design and natural science research on information technology. **Decision support systems**, Elsevier, v. 15, n. 4, p. 251–266, 1995.

MAXIMO, Eduardo Zeferino. **Maturidade da Gestão do Conhecimento: uma aplicação ao Modelo de Excelência da Gestão**. 2021. Diss. (Mestrado) – Universidade Federal de Santa Catarina.

MCCRAE, John P.; ABELE, Andrejs. **The Linked Open Data Cloud**. [S.l.: s.n.], 2021. <https://lod-cloud.net/>. 21/04/2021.

MEIJER, Albert J; CURTIN, Deirdre; HILLEBRANDT, Maarten. Open government: connecting vision and voice. **International Review of Administrative Sciences**, SAGE Publications Sage UK: London, England, v. 78, n. 1, p. 10–29, 2012.

MERSINGER, Denise L. **The Role of Data Governance in Data Quality**. 2016. F. 106. Tese (Doutorado). ISBN 9781369401523. Disponível em: <https://search.proquest.com/docview/1859579973?accountid=26642>.

METTLER, Tobias; ROHNER, Peter. **Situational maturity models as instrumental artifacts for organizational design**. In: ACM. PROCEEDINGS of the 4th international conference on design science research in information systems and technology. [S.l.: s.n.], 2009. P. 22.

MOSLEY, Mark; BRACKETT, Michael H; EARLEY, Susan; HENDERSON, Deborah. **DAMA guide to the data management body of knowledge**. first. [S.l.]: Technics Publications, 2009.

NEUMAIER, Sebastian. **Open Data Quality Assessment and Evolution of (Meta-)Data Quality in the Open Data Landscape**. 2015. Tese (Doutorado) – Technische Universität Wien.

NOLAN, Richard L. Managing the Computer Resource: A Stage Hypothesis. **Commun. ACM**, ACM, New York, NY, USA, v. 16, n. 7, p. 399–405, jul. 1973. ISSN 0001-0782. DOI: 10.1145/362280.362284. Disponível em: <http://doi.acm.org/10.1145/362280.362284>.

NURNBERGER, Andreas; SEISING, Rudolf; WENZEL, Constanze. **On the fuzzy interrelationships of data, information, knowledge and wisdom**. In: IEEE. FUZZY

Information Processing Society, 2009. NAFIPS 2009. Annual Meeting of the North American. [S.l.: s.n.], 2009. P. 1–6.

NUSAPATI, Chairuni Aulia; SUNINDYO, Wikan Danar. **Ferramenta de publicação de dados semiautomática para o avanço do estudo de caso de nível de maturidade de dados de governo aberto da Indonésia: Badan pusat statistik Indonésia.** *In:* 2017 International Conference on Data and Software Engineering (ICoDSE). [S.l.: s.n.]. P. 1–6. DOI: 10.1109/ICODSE.2017.8285887.

OPEN DATA INSTITUTE. **Publisher’s Guide to Open Data Licensing.** [S.l.: s.n.], 2013. <https://theodi.org/article/publishers-guide-to-open-data-licensing/>. Acesso em : 30/08/2018.

OPEN DEFINITION. **Open Definition.** [S.l.: s.n.], 2018. <http://opendefinition.org/>. 10/03/2018.

OTTO, Boris. **A morphology of the organisation of data governance.** *In:* p. 1–6.

OVIEDO, Edgar; MAZÓN, Jose Norberto; ZUBCOFF, Jose Jacobo. **Quality and maturity model for open data portals.** *In:* IEEE. COMPUTING Conference (CLEI), 2015 Latin American. [S.l.: s.n.], 2015. P. 1–7.

PACHECO, Daniela Chagas. **Maturidade em Gestão do Conhecimento da prefeitura Municipal de Criciúma: o caso dos setores de convênios.** 2015. Diss. (Mestrado) – Universidade Federal de Santa Catarina.

PACHECO, R. C. S. **Em Interdisciplinaridade: Universidade e Inovação Social e Tecnológica.** *In:* [S.l.]: Organizado por Joana Maria Pedro e Patrícia de Sá Freire, Curitiba, CRV, 2016. Coprodução em Ciência, Tecnologia e Inovação: Fundamentos e Visões, p. 21–62.

PANIAN, Zeljko. Some practical experiences in data governance. **World Academy of Science, Engineering and Technology**, v. 62, n. 1, p. 939–946, 2010.

PAULK, MC; CURTIS, B; CHRISSIS, MB; WEBER, CV. **Capability maturity model sm for software, version 1.1. Relatório Técnico CMU.** [S.l.], 1993.

PEFFERS, Ken; TUUNANEN, Tuure; ROTHENBERGER, Marcus A; CHATTERJEE, Samir. A design science research methodology for information

systems research. **Journal of management information systems**, Taylor & Francis, v. 24, n. 3, p. 45–77, 2007.

PEREIRA, Larissa Mariany Freiberger. **OGDPub: Uma ontologia para Publicação de Dados Abertos Governamentais**. 2017. Diss. (Mestrado) – Universidade Federal de Santa Catarina.

PIMENTA, Rosângela Borges. **Análise de maturidade da coprodução de conhecimento transdisciplinar: Um Estudo de Caso em uma Rede**. 2017. Tese (Doutorado) – Universidade Federal de Santa Catarina.

PÖPPELBUSS, Jens; RÖGLINGER, Maximilian. **What makes a useful maturity model? a framework of general design principles for maturity models and its demonstration in business process management**. *In: ECIS*. [S.l.: s.n.], 2011. P. 28.

PORTAL BRASILEIRO DE DADOS ABERTOS. **O que são dados abertos**. [S.l.: s.n.], 2018. <http://dados.gov.br/dados-abertos/>. 25/07/2018.

AL-RUITHE, Majid; BENKHELIFA, Elhadj; HAMEED, Khawar. A conceptual framework for designing data governance for cloud computing. **Procedia Computer Science**, Elsevier, v. 94, p. 160–167, 2016.

AL-RUITHE, Majid; BENKHELIFA, Elhadj; HAMEED, Khawar. Data governance taxonomy: Cloud versus non-cloud. **Sustainability**, Multidisciplinary Digital Publishing Institute, v. 10, n. 1, p. 95, 2018.

SADIQ, Shazia; INDULSKA, Marta. Open data: Quality over quantity. **International Journal of Information Management**, Elsevier, v. 37, n. 3, p. 150–154, 2017.

SALAH, Dina; PAIGE, Richard; CAIRNS, Paul. **An evaluation template for expert review of maturity models**. *In: SPRINGER. INTERNATIONAL Conference on Product-Focused Software Process Improvement*. [S.l.: s.n.], 2014. P. 318–321.

SAYOGO, D. S.; PARDO, T. A.; COOK, M. **A Framework for Benchmarking Open Government Data Efforts**. *In: 2014 47th Hawaii International Conference on System Sciences*. [S.l.: s.n.], jan. 2014. P. 1896–1905. DOI: 10.1109/HICSS.2014.240.

SCHARFFE, François *et al.* **Enabling linked data publication with the Datalift platform**. *In: PROC. AAAI WORKSHOP ON SEMANTIC CITIES*. Toronto, Canada:

[s.n.], jul. 2012. no pagination. scharffe2012a. Disponível em:
<https://hal.inria.fr/hal-00768424>.

SHEWHART, Walter Andrew; DEMING, William Edwards. **Statistical method from the viewpoint of quality control**. [S.l.]: Courier Corporation, 1986.

SIEBENLIST, Tobias; MAINKA, Agnes. **Culture of openness: the beginning of open government in municipalities in Germany**. In: ACM. PROCEEDINGS of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age. [S.l.: s.n.], 2018. P. 59.

SILVA, Patrícia; PINHEIRO, Marta Kerr. Métricas para Dados Governamentais Abertos. **Revista Ibero-Americana de Ciência da Informação**, v. 11, n. 1, p. 314–332, 2018. ISSN 1983-5213. Disponível em:
<http://periodicos.unb.br/index.php/RICI/article/view/25726>.

SIMON, Herbert A. The sciences of the artificial. **Cambridge, MA**, 1969.

SIMON, Herbert A. **The sciences of the artificial**. [S.l.]: MIT press, 1996.

SISTO, Raffaele; GARCÍA LÓPEZ, Javier; PAÉZ, José Manuel; MÚGICA, Elena Maté. Open Data Assessment in Italian and Spanish Cities. In: SPRINGER. **SMART AND SUSTAINABLE PLANNING FOR CITIES AND REGIONS: RESULTS OF SSPCR 2017 2**. [S.l.: s.n.], 2018. P. 121–131.

SMALLWOOD, Robert F. **Information governance: Concepts, strategies, and best practices**. [S.l.]: John Wiley & Sons, 2014. v. 574.

SOFTWARE ENGINEERING INSTITUTE. **CMMI**. [S.l.: s.n.], nov. 2018.
<https://cmmiinstitute.com/cmmi>.

SOLAR, Mauricio; CONCHA, Gastón; MEIJUEIRO, Luis. **A model to assess open government data in public agencies**. In: SPRINGER. INTERNATIONAL Conference on Electronic Government. [S.l.: s.n.], 2012. P. 210–221.

SOLAR, Mauricio; DANIELS, Fernando; LOPEZ, Roberto. **Automatic Generation of Roadmaps for Open Data**. In: IOS PRESS. ELECTRONIC Government and Electronic Participation: Joint Proceedings of Ongoing Research, Posters, Workshop and Projects of IFIP EGOV 2014 and EPart 2014. [S.l.: s.n.], 2014a. P. 95.

SOLAR, Mauricio; DANIELS, Fernando; LÓPEZ, Roberto; MEIJUEIRO, Luis. A Model to Guide the Open Government Data Implementation in Public Agencies. **J. UCS**, v. 20, n. 11, p. 1564–1582, 2014b.

SOLAR, Mauricio; MEIJUEIRO, Luis; DANIELS, Fernando. **A guide to implement open data in public agencies**. *In*: SPRINGER. INTERNATIONAL conference on electronic government. [S.l.: s.n.], 2013. P. 75–86.

SPERONI, Rafael de Moura. **Modelo de Referência Para Indicadores de Inovação Regional Suportado por Dados Ligados**. 2016. Tese (Doutorado) – Universidade Federal de Santa Catarina.

SUSHA, Iryna; ZUIDERWIJK, Anneke; CHARALABIDIS, Yannis; PARYCEK, Peter; JANSSEN, Marijn. Critical factors for open data publication and use: A comparison of city-level, regional, and transnational cases. **JeDEM-eJournal of eDemocracy and Open Government**, v. 7, n. 2, p. 94–115, 2015a.

SUSHA, Iryna; ZUIDERWIJK, Anneke; JANSSEN, Marijn; GRÖNLUND, Åke. Benchmarks for evaluating the progress of open data adoption: usage, limitations, and lessons learned. **Social Science Computer Review**, SAGE Publications Sage CA: Los Angeles, CA, v. 33, n. 5, p. 613–630, 2015b.

TALUKDER, Md Shamim; SHEN, Liang; TALUKDER, Md Farid Hossain; BAO, Yukun. Determinants of user acceptance and use of open government data (OGD): An empirical investigation in Bangladesh. **Technology in Society**, Elsevier, v. 56, p. 147–156, 2019.

TORINO, Emanuelle; TREVISAN, Gustavo Lunardelli; VIDOTTI, Silvana A. B. Gregorio. Dados abertos capes: A look at the challenges of web data publishing. **Ciência da Informação**, p. 38–46, 2019.

TRIBUNAL DE CONTAS DA UNIÃO. **Referencial básico de governança aplicável a órgãos e entidades da administração pública**. [S.l.: s.n.], 2014. shorturl.at/iyPT5.

TYGEL, A.; AUER, S.; DEBATTISTA, J.; ORLANDI, F.; CAMPOS, M. L. M. **Towards Cleaning-Up Open Data Portals: A Metadata Reconciliation Approach**. *In*: 2016 IEEE Tenth International Conference on Semantic Computing (ICSC). [S.l.: s.n.], fev. 2016. P. 71–78. DOI: 10.1109/ICSC.2016.54.

UENO, Alexandre Takeshi. **Modelo de Avaliação da Maturidade do Processo de Inovação Como Estratégia Competitiva Empresarial**. 2016. Tese (Doutorado) – Universidade Federal de Santa Catarina.

UNE 178301. **Ciudades Inteligentes. Datos Abertos(Open Data)**. Ago. 2015.
Disponível em: <http://www.aenor.es/aenor/normas/normas/fichanorma.asp?tipo=N%5C&codigo=N0054318%5C&PDF=Si>.

VAISHNAVI, Vijay K; KUECHLER, William. **Design science research methods and patterns: innovating information and communication technology**. [S.l.]: Crc Press, 2015.

VAN DE WEERD, Inge; BRINKKEMPER, Sjaak; NIEUWENHUIS, Richard; VERSEDAAL, Johan; BIJLSMA, Lex. **Towards a reference framework for software product management**. *In*: IEEE. REQUIREMENTS Engineering, 14th IEEE International Conference. [S.l.: s.n.], 2006. P. 319–322.

VAN STEENBERGEN, Marlies; BERG, Martin van den; BRINKKEMPER, Sjaak. A Balanced Approach to Developing the Enterprise Architecture Practice. *In*: FILIPE, Joaquim; CORDEIRO, José; CARDOSO, Jorge (Ed.). **Enterprise Information Systems**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008. P. 240–253.

VAN STEENBERGEN, Marlies; BOS, Rik; BRINKKEMPER, Sjaak; VAN DE WEERD, Inge; BEKKERS, Willem. **The design of focus area maturity models**. *In*: SPRINGER. INTERNATIONAL Conference on Design Science Research in Information Systems. [S.l.: s.n.], 2010a. P. 317–332.

VAN STEENBERGEN, Marlies; BOS, Rik; BRINKKEMPER, Sjaak; WEERD, Inge van de; BEKKERS, Willem. Improving IS Functions Step by Step: the Use of Focus Area Maturity Models. **Scandinavian J. Inf. Systems**, v. 25, n. 2, p. 2, 2013.

VAN STEENBERGEN, Marlies; SCHIPPER, Jurjen; BOS, Rik; BRINKKEMPER, Sjaak. **The dynamic architecture maturity matrix: Instrument analysis and refinement**. *In*: SPRINGER. SERVICE-ORIENTED Computing. ICSOC/ServiceWave 2009 Workshops. [S.l.: s.n.], 2010b. P. 48–61.

VELJKOVIĆ, Nataša; BOGDANOVIĆ-DINIĆ, Sanja; STOIMENOV, Leonid. Benchmarking open government: An open data perspective. **Government**

Information Quarterly, v. 31, n. 2, p. 278–290, 2014. ISSN 0740-624X. DOI:
<https://doi.org/10.1016/j.giq.2013.10.011>. Disponível em:
<http://www.sciencedirect.com/science/article/pii/S0740624X14000434>.

VILLAZÓN-TERRAZAS, Boris; VILCHES-BLÁZQUEZ, Luis. M.; CORCHO, Oscar; GÓMEZ-PÉREZ, Asunción. Methodological Guidelines for Publishing Government Linked Data. *In: Linking Government Data*. Edição: David Wood. New York, NY: Springer New York, 2011. P. 27–49. ISBN 978-1-4614-1767-5. DOI:
10.1007/978-1-4614-1767-5_2. Disponível em:
http://dx.doi.org/10.1007/978-1-4614-1767-5_2.

VISINTIN, Lidiane; GOMES, Murilo Silveira; PEREIRA, Larissa F.; TODESCO, José. Dados Governamentais e Empresariais – Umq Discussão Acerca de Dados Abertos e Dados Conectados. **Congresso Internacional de Conhecimento e Inovação – Ciki**, v. 1, n. 1, 2017. ISSN 2318-5376. Disponível em:
<http://proceeding.ciki.ufsc.br/index.php/ciki/article/view/103>.

WAGTER, Roel; VAN DEN BERG, Martin; LUIJPERS, Joost; VAN STEENBERGEN, Marlies. **Dynamic enterprise architecture: how to make it work**. [S.l.]: John Wiley & Sons, 2005.

WANG, Victoria; SHEPHERD, David. Exploring the extent of openness of open government data—A critique of open government datasets in the UK. **Government Information Quarterly**, Elsevier, v. 37, n. 1, p. 101405, 2020.

WANG, Victoria; SHEPHERD, David; BUTTON, Mark. The barriers to the opening of government data in the UK: A view from the bottom. **Information Polity**, IOS Press, v. 24, n. 1, p. 59–74, 2019.

WEBER, Kristin; OTTO, Boris; ÖSTERLE, Hubert. **One Size Does Not Fit All—A Contingency Approach to Data Governance**. Association for Computing Machinery, New York, NY, USA, v. 1, n. 1, jun. 2009. ISSN 1936-1955. DOI:
10.1145/1515693.1515696. Disponível em:
<https://doi.org/10.1145/1515693.1515696>.

WEERD, Inge van de; BEKKERS, Willem; BRINKKEMPER, Sjaak. **Developing a maturity matrix for software product management**. *In: SPRINGER. INTERNATIONAL Conference of Software Business*. [S.l.: s.n.], 2010. P. 76–89.

WEINBERGER, D. **Too Big to Know: Rethinking Knowledge Now That the Facts Aren't the Facts, Experts Are Everywhere, and the Smartest Person in the Room Is the Room.** [S.l.]: Basic Books, 2014. ISBN 9780465085965. Disponível em: <https://books.google.com.br/books?id=1BiktAEACAAJ>.

WELLE DONKER, F.; LOENEN, B. van. **How to assess the success of the open data ecosystem?** v. 10, n. 3, p. 284–306, 2017. DOI: 10.1080/17538947.2016.1224938.

WENDE, Kristin. **A model for data governance—Organising accountabilities for data quality management**, 2007.

WENDLER, Roy. The maturity of maturity model research: A systematic mapping study. **Information and Software Technology**, v. 54, n. 12, p. 1317–1339, 2012. Special Section on Software Reliability and Security. ISSN 0950-5849. DOI: <https://doi.org/10.1016/j.infsof.2012.07.007>.

WIERINGA, R.J. **Design Science Methodology for Information Systems and Software Engineering.** [S.l.]: Springer Berlin Heidelberg, 2014. (Computer science). ISBN 9783662438398. Disponível em: <https://books.google.com.br/books?id=xLKLbQAAQBAJ>.

WU, Tong; MAHDZIR, Akbariah Mohd. An investigation of application developer's awareness and attitude towards open government data in Malaysia. **International Journal of Engineering & Technology**, v. 7, n. 4, p. 5127–5131, 2018. ISSN 2227-524X. DOI: 10.14419/ijet.v7i4.26558. Disponível em: <https://www.sciencepubco.com/index.php/ijet/article/view/26558>.

YANG, Tung-Mou; WU, Yi-Jung. **Assessing the Open Data Maturity of Government Agencies Having Different Years of Implementation Experience: A Study in Taiwan.** In: THE 21st Annual International Conference on Digital Government Research. [S.l.: s.n.], 2020. P. 348–350.

YI, Myongho. Exploring the quality of government open data. **The Electronic Library**, Emerald Publishing Limited, 2019.

ZANUZZI, Cinthya Mônica da Silva. **Avaliação da Maturidade da Gestão do Conhecimento das Unidades Produtoras Integradas da Agroindústria Avícola.** 2019. Diss. (Mestrado) – Universidade Federal de Santa Catarina.

ZENG, Marcia Lei; HLAVA, Marjorie; GREENBERG, Jane; NEEDLEMAN, Mark. Standards and best practices related to the publication, exchange and usage of open data. **Proceedings of the Association for Information Science and Technology**, Wiley Online Library, v. 54, n. 1, p. 610–611, 2017.

ZUIDERWIJK, Anneke; JANSSEN, Marijn; CHOENNI, Sunil; MEIJER, Ronald; ALIBAKS, Roexsana Sheikh. Socio-technical Impediments of Open Data. **Electronic Journal of e-Government**, v. 10, n. 2, 2012.

**APÊNDICE A – SEÇÃO 1 - FORMULÁRIO APLICADO PARA OS
RESPONDENTES DE CADA CENÁRIO DE USO**

Quadro 51 – Seção I - Questionário e Estrutura do Modelo

As próximas afirmativas dizem respeito ao questionário que foi aplicado	
De modo geral, qual sua percepção sobre as seguintes afirmativas:	
As questões são relevantes para o domínio (Relevância)	0 - Discordo Fortemente 1 - Discordo 2 - Concordo 3 - Concordo Plenamente
As questões contribuem com aspectos importantes para cada componente (Relevância)	0 - Discordo Fortemente 1 - Discordo 2 - Concordo 3 - Concordo Plenamente
As questões são suficientes para representar cada componente (Abrangência)	0 - Discordo Fortemente 1 - Discordo 2 - Concordo 3 - Concordo Plenamente
De modo geral, qual sua percepção sobre a estrutura do modelo de maturidade apresentados considerando:	
Compreensibilidade	
Os níveis de maturidade são compreensíveis	0 - Discordo Fortemente 1 - Discordo 2 - Concordo 3 - Concordo Plenamente
A estrutura de avaliação é compreensível	0 - Discordo Fortemente 1 - Discordo 2 - Concordo 3 - Concordo Plenamente
Facilidade de uso	
O questionário é fácil de usar	0 - Discordo Fortemente 1 - Discordo 2 - Concordo 3 - Concordo Plenamente
As diretrizes para a avaliação são fáceis de usar	0 - Discordo Fortemente 1 - Discordo 2 - Concordo 3 - Concordo Plenamente
Utilidade e praticidade	
O modelo de maturidade é útil na realização de avaliações	0 - Discordo Fortemente 1 - Discordo 2 - Concordo 3 - Concordo Plenamente
O modelo de maturidade é prático para uso	0 - Discordo Fortemente 1 - Discordo 2 - Concordo 3 - Concordo Plenamente

Fonte: da Autora

**APÊNDICE B – SEÇÃO 2 - FORMULÁRIO APLICADO PARA OS
RESPONDENTES DE CADA CENÁRIO DE USO**

Quadro 52 – Seção II - Importância das áreas foco e dos componentes do modelo de maturidade

Importância das áreas foco e dos componentes do modelo de maturidade	
Para você, qual a importância das seguintes áreas foco, quando considerado o domínio de dados abertos?	
Gerenciamento de Políticas e Processos	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
Gerenciamento de Fatores Organizacionais	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
Gerenciamento de Padrões	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
Para você, qual a importância dos componentes relacionados à área foco - Gerenciamento de políticas e Processos, considerando o âmbito de dados abertos?	
Definição dos dados	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
Acesso e entrega de dados	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
Monitoramento e Medição	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
Gerenciamento de mudança de dados	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
Para você, qual a importância dos componentes relacionados à área foco - Gerenciamento de Fatores Organizacionais, considerando o âmbito de dados abertos?	
Funções e Responsabilidade	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
Educação	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
Treinamento	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
Planejamento e Priorização da Abertura dos Dados	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
Planejamento da Publicação dos Dados	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
Fomento ao Uso	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
Para você, qual a importância dos componentes relacionados à área foco - Gerenciamento de Padrões, considerando o âmbito de dados abertos?	

Licença	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
Proveniência e Qualidade	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
Versionamento	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
Identificação	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
Formato	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
Vocabulários	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
Metadados	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
Preservação	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
Enriquecimento	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
<i>Feedback</i>	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
Acesso	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável

Fonte: da Autora

APÊNDICE C – SEÇÃO 1 - FORMULÁRIO APLICADO PARA OS ESPECIALISTAS

Quadro 53 – Parte I - Perfil dos respondentes

Perfil dos Entrevistados	
Nome:	
E-mail:	
Organização em que trabalha atualmente:	
Possui Familiaridade com o tema "Dados Abertos"?	Sim, Não, Talvez
Já trabalhou com dados abertos?	Sim, Não, Talvez
Em caso afirmativo na questão anterior, descreva brevemente a quanto tempo trabalha ou trabalhou e em quais atividades atuou?	

Fonte: Da Autora

APÊNDICE D – SEÇÃO 2 - FORMULÁRIO APLICADO PARA OS ESPECIALISTAS

Quadro 54 – Parte II - Importância das áreas foco e dos componentes do modelo de maturidade

Importância das áreas foco e dos componentes do modelo de maturidade	
Para você, qual a importância das seguintes áreas foco, quando considerado o domínio de dados abertos?	
Gerenciamento de Políticas e Processos	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
Gerenciamento de Fatores Organizacionais	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
Gerenciamento de Padrões	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
Para você, qual a importância dos componentes relacionados à área foco - Gerenciamento de políticas e Processos, considerando o âmbito de dados abertos?	
Definição dos dados	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
Acesso e entrega de dados	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
Monitoramento e Medição	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
Gerenciamento de mudança de dados	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
Para você, qual a importância dos componentes relacionados à área foco - Gerenciamento de Fatores Organizacionais, considerando o âmbito de dados abertos?	
Funções e Responsabilidade	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
Educação	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
Treinamento	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
Planejamento e Priorização da Abertura dos Dados	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
Planejamento da Publicação dos Dados	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
Fomento ao Uso	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
Para você, qual a importância dos componentes relacionados à área foco - Gerenciamento de Padrões, considerando o âmbito de dados abertos?	

Licença	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
Proveniência e Qualidade	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
Versionamento	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
Identificação	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
Formato	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
Vocabulários	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
Metadados	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
Preservação	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
Enriquecimento	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
<i>Feedback</i>	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável
Acesso	0 - Irrelevante 1- Pouco Importante 2- Importante 3 - Muito Importante 4 - Indispensável

Fonte: da Autora

APÊNDICE E – SEÇÃO 3 - FORMULÁRIO APLICADO PARA OS ESPECIALISTAS

Quadro 55 – Parte III - Estrutura do Modelo

Estrutura do Modelo de Maturidade	
De modo geral, qual sua percepção sobre a estrutura do modelo de maturidade apresentados, considerando as afirmativas sobre:	
Áreas Foco	
As áreas foco são suficientes para representar o domínio do modelo de maturidade (Abrangência)	0 - Discordo fortemente, 1 - Discordo, 2 - Concordo, 3 - Concordo plenamente
As áreas foco são relevantes para o domínio (Relevância)	0 - Discordo fortemente, 1 - Discordo, 2 - Concordo, 3 - Concordo plenamente
Componentes	
Os componentes são suficientes para representar cada área foco (Abrangência)	0 - Discordo fortemente, 1 - Discordo, 2 - Concordo, 3 - Concordo plenamente
Os componentes são relevantes para cada área foco (Relevância)	0 - Discordo fortemente, 1 - Discordo, 2 - Concordo, 3 - Concordo plenamente
Os componentes foram atribuídos corretamente às suas respectivas áreas foco (Precisão)	0 - Discordo fortemente, 1 - Discordo, 2 - Concordo, 3 - Concordo plenamente
Capacidades	
As capacidades representam uma evolução progressiva para cada componente (Abrangência)	0 - Discordo fortemente, 1 - Discordo, 2 - Concordo, 3 - Concordo plenamente
As capacidades são relevantes para o domínio (Relevância)	0 - Discordo fortemente, 1 - Discordo, 2 - Concordo, 3 - Concordo plenamente
Caso alguma das afirmativas tenha sido respondida com "discordo fortemente", quais pontos deveriam ser considerados para melhor representar o domínio?	
Caso queira deixar alguma sugestão ou comentário	

Fonte: da Autora