

**METADATA MANAGEMENT FOR OPEN GOVERNEMENT DATA: A PROPOSE FOR A
INDUSTRY OBSERVATORY****GESTÃO DE METADADOS PARA DADOS ABERTOS GOVERNAMENTAIS: UMA PROPOSTA
PARA UM OBSERVATÓRIO INDUSTRIAL****MARCELO COUTO**

<https://orcid.org/0009-0004-2627-7726/marcelodocouto1@gmail.com>
Universidade do Estado de Santa Catarina - UDESC, Florianópolis, Santa Catarina.

DORZELI SALETE TRZECIAK

<https://orcid.org/0000-0003-1967-1093/dorzeli@gmail.com>
Instituto Euvaldo Lodi de Santa Catarina – IEL, Florianópolis, Santa Catarina.

JÚLIO DA SILVA DIAS

julio.dias@udesc.br
Universidade do Estado de Santa Catarina - UDESC, Florianópolis, Santa Catarina.

Recebido em: 14/07/2024

Aprovado em: 17/10/2024

Publicado em: 09/12/2024

**RESUMO**

O objetivo deste artigo é apresentar uma proposta de *template* para catalogação de metadados das bases de dados de um Observatório de indústria. É uma pesquisa aplicada, com abordagem qualitativa e, ainda, exploratória e descritiva. Utiliza a pesquisa bibliográfica como procedimento técnico. Para a elaboração do *template* foram considerados os elementos relacionados a repositórios de metadados, tendo a iniciativa FAIR como principal referência e a tecnologia FAIR *Data Point* para auxiliar na recuperação de informação a partir dos repositórios de dados. O *template* desenvolvido descreve as informações básicas necessárias para entendimento das bases de dados deste Observatório de indústria, por meio de metadados que trazem suas características, sua localização e quais perguntas respondem. O *template* utilizou, majoritariamente, metadados propostos pela tecnologia FAIR *Data Point*, porém foi necessário acrescentar alguns metadados adaptados à realidade do Observatório para auxiliar no processo de recuperação de informações, a partir dos dados disponíveis. O estudo contribui para a redução do tempo de busca de dados sobre determinado tema, a gestão do conhecimento e o dimensionamento das potencialidades de utilização dos dados das bases para a evolução deste Observatório como um centro de inteligência, para o atendimento de demandas e a criação de produtos, projetos e serviços, atendendo aos seus objetivos de negócios.

Palavras-chave: dados abertos governamentais; gestão de conhecimento; gestão de metadados; governança de dados; observatórios.

ABSTRACT

The objective of this article is to present a template proposal for metadata cataloging of the databases of a industry Observatory. This is an applied research study with a qualitative, exploratory, and descriptive approach. It employs bibliographic research as a technical procedure. The development of the template considered elements related to metadata repositories, with the FAIR initiative serving as the primary reference, and the FAIR Data Point technology assisting in information retrieval from data repositories. The developed template outlines the basic information necessary for understanding the industry Observatory's databases through metadata that describe their characteristics, location, and the questions they address. The template primarily used metadata proposed by FAIR Data Point technology; however, it was necessary to add some metadata adapted to the Observatory's context to facilitate the information retrieval process from the available data. This study contributes to reducing the time required to search for data on specific topics, knowledge management, and assessing the potential uses of the database for the evolution of the industry Observatory as an intelligence center, meeting demands, and creating products, projects, and services in line with its business objectives.

Keywords: open data government; knowledge management; metadata management; data governance; observatories.

1 INTRODUÇÃO

Nos últimos anos, a percepção dos dados como ativos tem crescido exponencialmente. Isso ocorre porque a capacidade de armazenar e processar dados aumentou consideravelmente, além de serem gerados a uma velocidade e volume sem precedentes, ganhando cada vez mais valor e importância. Entretanto, nem todos os dados precisam ser mantidos indefinidamente em um data *lake*, um repositório centralizado de dados. Muitos dados podem ser utilizados para análises ou estudos específicos e, após esse uso, serem descartados sem nenhum prejuízo.

Por outro lado, a manutenção prolongada de dados subutilizados a longo prazo pode levar à criação de um grande espaço obsoleto, com informações que já não são mais utilizadas, impactando diretamente nos custos relacionados ao armazenamento, seja em nuvem ou servidor *on premise* da empresa. Por isso, é fundamental que haja uma governança de dados eficaz para gerir esses dados. A governança de dados tem um papel crucial em organizações que adotam a cultura de dados em sua estratégia, garantindo que as decisões sejam tomadas de forma mais assertiva, baseadas em informações relevantes e atualizadas (TALLON; RAMIREZ; SHORT, 2013).

Na oferta de dados e informações estratégicas para subsidiar a tomada de decisões, destacam-se os observatórios, como instrumentos que sistematizam várias fontes de informação em uma fonte de informação referencial. Observatórios disponibilizam dados acessíveis e confiáveis, facilitando o acesso à informação e ao conhecimento (ORTEGA; VALLE, 2010).

Nos últimos tempos, uma variedade de observatórios tem emergido com foco em diversas áreas, tais como ciência, tecnologia, inovação, saúde, social, software, prospecção, transporte e logística. Cada observatório tem um propósito específico, dependendo do ambiente em que foi criado, da sua natureza e do seu público-alvo. Essas características influenciam significativamente na estrutura, nos processos, nas metodologias e nos procedimentos adotados para o seu funcionamento (SOARES; FERNEDA; DO PRADO, 2018).

Um exemplo significativo é o Observatório da pesquisa, um centro de inteligência que abriga uma vasta gama de dados, tecnologias e soluções destinadas a auxiliar decisões que promovem a competitividade e o desenvolvimento sustentável da indústria e seus stakeholders. Este Observatório se apoia majoritariamente nos dados abertos governamentais para realizar suas análises e desenvolver essas soluções. Estas bases são originárias de diversas fontes do governo, e podem ser federais, estaduais e algumas até municipais, compreendendo diversas temáticas, como economia, saúde, educação, dentre outras.

O Observatório da indústria deste estudo será chamada de Observatório A. Ele possui uma infraestrutura robusta de big data e *analytics*, alimentada por cerca de 100 fontes de dados do Brasil e do mundo. Isso garante informações de qualidade para apoiar o processo decisório dos industriais catarinenses, exigindo a estruturação da governança dos dados de big data e a criação de orientações para a equipe.

Nesse contexto, esta pesquisa é realizada em uma organização que, conforme já mencionado, utiliza principalmente dados abertos governamentais. Isso implica um nível de complexidade maior do que se trabalhasse apenas com dados internos, já que eventuais mudanças que venham a ocorrer nos dados e na sua estrutura de disponibilização não dependem da organização, o que acaba trazendo uma vulnerabilidade para a organização, exigindo assim, uma gestão de dados eficaz para evitar que a qualidade dos dados fique comprometida, afetando a percepção do cliente sobre a sua credibilidade no fornecimento de informações. Além disso, essas

bases diferem não apenas na periodicidade de atualização, mas também na estrutura e na forma de organização dos dados.

Assim, é imprescindível a criação de um processo de catalogação e gestão dos metadados, visando aumentar a eficiência no processo inicial de busca dos dados e organizar todo o conhecimento que já foi gerado e que pode ser desenvolvido a partir das bases que o Observatório possui. Os metadados fornecem o contexto em que os dados estão inseridos e são elementos fundamentais na transformação do dado em informação e, posteriormente, em conhecimento (BARBIERI, 2019).

Para isso, este artigo tem como objetivo, propor um *template* para catalogação de metadados das bases de dados para o Observatório da indústria A, com uma estrutura que atenda aos objetivos de negócio, facilite o processo de *discovery* de dados e melhore a eficiência no atendimento de demandas, projetos, produtos e serviços que o Observatório oferece.

Espera-se que o *template* sugerido reduza significativamente o tempo necessário para localizar dados sobre um tema específico. Além disso, ele deverá auxiliar no processo de criação de produtos, desenvolvimento de projetos e serviços oferecidos pelo Observatório, que utiliza dados como sua principal “matéria-prima”.

No que se refere a gestão do conhecimento, o *template* contribuirá para evidenciar várias informações sobre as bases de dados, promovendo o entendimento da sua potencialidade para todos os colaboradores do Observatório, até mesmo para aqueles que não têm contato direto com elas.

Esse artigo está estruturado em: introdução, apresentando contexto, problemática e objetivo da pesquisa; procedimentos metodológicos, com a descrição das etapas adotadas para sua realização; referencial teórico com a abordagem da temática; construção da proposta de *template* para catalogação de metadados das bases de dados do Observatório A; considerações finais e; as referências utilizadas.

2 PROCEDIMENTOS METODOLÓGICOS

Esta pesquisa, quanto a sua natureza, caracteriza-se como aplicada. Segundo Marconi e Lakatos (2007) esse tipo de pesquisa objetiva produzir conhecimentos para aplicação de seus

resultados, contribuindo para melhorias nos resultados da instituição, buscando soluções para problemas encontrados na realidade. Já em relação à abordagem do problema, trata-se de uma pesquisa qualitativa, pois visa estudar alguns conceitos para a construção de uma proposta de *template* para um contexto organizacional.

No tocante aos seus objetivos, é uma pesquisa exploratória e descritiva. Esse tipo de pesquisa, segundo Silva (2018, p. 40) "visa explorar e descrever as características, contextos e significados atribuídos aos fenômenos, a partir de uma perspectiva interpretativa".

Quanto aos procedimentos técnicos, adotou-se uma abordagem de pesquisa bibliográfica, utilizando fontes primárias e secundárias para contextualizar as temáticas em discussão e identificar elementos para construir a proposta de *template* para catalogação de metadados. Conforme destacado por Minussi et al. (2018), a pesquisa bibliográfica tem como característica fundamental a reunião de diversos dados previamente dispersos no universo científico, ampliando assim, a área de abrangência da pesquisa.

O estudo tem como foco as bases de dados que são utilizadas e gerenciadas pelo Observatório FIESC, visto que servem de insumo para praticamente todas as suas atividades atualmente, necessitando de processos bem estabelecidos para facilitar a sua utilização e evitar retrabalhos e problemas relacionados à qualidade de dados. Para o desenvolvimento da proposta foram utilizados como referência os metadados presentes no *software* FAIR *Data Point* (2020), que utiliza os princípios FAIR como base para sua construção. Salienta-se que o objetivo deste estudo é desenvolver um catálogo de metadados para bases de dados armazenadas nos sistemas de repositórios de dados gerenciados pelo Observatório, não contemplando tabelas, indicadores e outros objetos de dados.

3 REFERENCIAL TEÓRICO

3.1 Observatórios

De acordo com Ortega e Valle (2010) não existe uma definição única sobre observatório, dificultando assim, a sua identificação e limitando a análise do fenômeno. Para Barbosa (2020), as definições, características e objetivos conceituais de observatório variam conforme o que se propõe a fazer e como pode se caracterizar uma ferramenta como observatório. Além disso, a diversidade

organizacional e operacional possibilitada pelo observatório aumenta a dificuldade de se ter uma definição única. Dessa forma, seu conceito acaba sendo definido, ao longo do tempo, de acordo com o viés do autor e do contexto em que o observatório está inserido (ASPINALL JACOBSON; CASTILLO-SALGADO, 2016; LÓPEZ et al., 2012).

Para Silva et al. (2013), os observatórios se caracterizam como um sistema de informação que objetiva a visualização de informações e dados de determinada área, além de fornecer análises e estudos, com base em indicadores estratégicos sobre a temática.

Trzeciak (2009. p. 31) ressalta que “o observatório é um mecanismo que fornece informações estratégicas para auxiliar na identificação de ameaças, oportunidades e tendências, além de oferecer subsídios à tomada de decisão de um determinado setor”.

Santoro e Xavier (2009) classificam os observatórios em territoriais, que estudam aspectos relativos a um espaço geográfico, seja cidade, estado ou região, e temáticos, focados em questões sociais, econômicas ou políticas. Os autores apontam que há também, aqueles que se caracterizam pelos dois tipos, pelo fato da questão geográfica fazer parte de um determinado interesse da sociedade.

O que se observa atualmente, é que há observatórios focados em diversas temáticas, tais como ciência, tecnologia e inovação, social, software, saúde, prospecção, transporte e logística, dentre outros. Cada observatório tem um propósito específico, dependendo da sua natureza, do ambiente em que foi criado e do público que pretende atender. Por consequência, essas características afetam e interferem na forma como são estruturados, nas suas metodologias, processos e procedimentos utilizados (SOARES; FERNEDA; DO PRADO, 2018).

Apoiar o processo decisório é um dos elementos essenciais do observatório, fundamentando a decisão de gestores com dados e informações tratados e qualificados. Aliás, coletar, compilar e tratar informações para subsidiar a tomada de decisão é uma das atividades dos observatórios (GUSMÃO, 2005; PHÉLAN, 2007; BATISTA et al., 2017). Reyes e Fernández (2019) reforçam que ser um centro de informações com coleta, tratamento e difusão de dados e informações é uma das atribuições do observatório.

Ressalta-se que, para executar as atividades mencionadas, o observatório deve possuir infraestrutura física e tecnológica, equipe qualificada, além de um robusto banco de dados estruturados e não estruturados, que forneçam informações para apoiar na tomada de decisão do seu público-alvo (SILVA et al., 2013).

Nesse cenário, a governança de dados desempenha um papel crucial na administração desses elementos. Seu principal objetivo é garantir a integração harmoniosa entre eles, de forma a assegurar que os objetivos estratégicos do Observatório sejam alcançados.

Por outro lado, dada a natureza dos Observatórios, que se baseia na extração de informações e conhecimento a partir de dados, é comum que muitos deles, por questões de limitação de recursos físicos e financeiros, recorram a dados públicos, mais especificamente, dados abertos governamentais. Esse fator necessita que a governança de dados se adapte às características desses dados, a fim de mitigar impactos negativos que possam ocorrer devido à falta de controle e de regras que normatizam o seu uso. Estes dados, por sua vez, são conhecidos por abrangerem uma grande variedade de temas e por retratar a realidade do país em relação a esses assuntos específicos. Com isso, se faz necessário um aprofundamento desse tema, o que será feito no próximo tópico.

3.2 Dados abertos governamentais

De acordo com a *Open Knowledge Foundation* (2022), um dado aberto é definido como um dado que pode ser acessado, utilizado, reutilizado, modificado e compartilhado por qualquer pessoa de forma gratuita. No contexto da gestão pública, o termo "dados abertos" foi introduzido pela primeira vez em um programa de cooperação internacional que visava operar satélites estadunidenses (POSSAMAI, 2016).

A *Open Knowledge Foundation* (2022) estabelece três normas fundamentais para considerar dados abertos: disponibilidade de acesso, reuso e distribuição e participação universal, que podem ser resumidas da seguinte forma (PORTAL BRASILEIRO DE DADOS ABERTOS, 2024):

- a) se o dado não pode ser encontrado ou indexado na web, ele não existe;

- b) se o dado não está disponível num formato aberto e legível por máquina, ele não pode ser reutilizado;
- c) se dispositivos legais não permitem que ele seja compartilhado, ele não é útil.

Dados abertos, são dados acessíveis ao público, representados em meio digital, estruturados em formato aberto, processáveis por máquina, referenciados na internet e disponibilizados sob licença aberta que permita sua livre utilização, consumo ou cruzamento, desde que se dê crédito à autoria ou à fonte (SANTOS, 2020).

No Brasil, a Administração Pública Federal é obrigada a publicar informações públicas, de acordo com o princípio da publicidade estipulado na Constituição de 1988, e os cidadãos têm o direito de receber essas informações. A Lei de Acesso à Informação, Lei 12.527 de 2011, entrou em vigor em 2019 e, com ela, foi lançado o Portal Brasileiro de Dados Abertos. Este portal tem como objetivo ser um ponto central para busca, acesso e uso dos dados abertos governamentais, além de promover o diálogo entre a sociedade e o governo para discutir o melhor uso dos dados em benefício de uma sociedade mais transparente e informada (PORTAL BRASILEIRO DE DADOS ABERTOS, 2024).

No entanto, apenas disponibilizar dados em portais eletrônicos não é suficiente para promover a transparência dos serviços públicos. É necessário garantir uma auditoria constante sobre o que é disponibilizado, e mais importante ainda, os usuários precisam ser capacitados para extrair e manipular esses dados para utilizá-los em benefício próprio e, por extensão, da sociedade.

Segundo Possamai (2016) todo dado público tem potencial para ser aberto. No entanto, o dado público não deve estar sujeito a limitações de privacidade, segurança, controle de acesso ou outros privilégios regulados por estatutos. Por isso, eles não incluem dados pessoais, sob a custódia estatal, ou dados sigilosos de organizações privadas. Além disso, alguns dados governamentais, devido a sua sensibilidade e por questões estratégicas de estado, demandam reserva e sigilo, desde que devidamente regulamentadas pelo estado.

Apesar de dados abertos governamentais possuírem características distintas em relação aos dados de organizações, quando utilizados internamente para subsidiar a tomada de decisão, naturalmente é necessário um programa de governança de dados, com regras, papéis e

responsabilidades para evitar custos desnecessários e garantir que o dado chegue ao tomador de decisão com qualidade. Esse tema é tratado no tópico a seguir.

3.3 Governança de dados

Governança de dados, segundo o DAMA (2017), refere-se ao exercício de autoridade e controle sobre o gerenciamento dos dados. Seu objetivo é maximizar a percepção dos gestores acerca da importância dos dados como ativos dentro da organização (CARRETERO et al., 2017). Todavia, poucas organizações atualmente consideram seus dados como ativos. Estas instituições geram, processam e utilizam dados de diversas maneiras, mas carecem de normas e políticas que regulamentem sua gestão.

A governança de dados é um termo relativamente novo, que surgiu no início deste século. Entretanto, a literatura apresenta diversos termos e abordagens para trabalhar com dados e informações, como a gestão total da qualidade de dados (WANG, 1998), a gestão da qualidade de dados (WANG; STRONG, 1996), entre outros (OTTO, 2012). É importante, no entanto, destacar a diferença entre governança e gestão de dados.

Governança de dados refere-se ao processo de tomada de decisão e à definição de quem tem autoridade para tomá-la, buscando garantir a eficiência da gestão e o uso adequado dos recursos. A gestão de dados, por sua vez, está relacionada à execução das decisões previamente tomadas.

Alguns autores descrevem a governança de dados como um *framework* que suporta a gestão de dados. E um dos *frameworks* mais amplamente utilizados é o DAMA-DMBOK, desenvolvido pela *Data Management Association International* (DAMA). Esse framework consiste numa coleção de processos considerados como melhores práticas dentro dos setores relacionados à gestão de dados (BENTO; NETO; CORTÊ-REAL, 2022).

A DAMA, fundada em 1980, visa auxiliar os profissionais de dados a aprimorarem seus conhecimentos e habilidades em gestão de dados. Além disso, busca disseminar boas práticas, educar e certificar os profissionais da área e conceitua a governança de dados como um framework transversal a vários setores, objetivando gerir os dados como um ativo estratégico para a organização (DAMA, 2017).

Para Abraham, Schneider e Brocke (2019) o objetivo da governança de dados é aprimorar a segurança da informação, reduzir custos, alinhar as atividades da empresa com seus dados e controlar a sua qualidade. Governança de dados, conforme argumentam, não se trata de uma ferramenta técnica, mas sim de um conjunto de normas, boas práticas e políticas para a utilização e distribuição de dados. Criar um ambiente de *big data* sem uma organização e regras bem definidas pode resultar em altos custos e perda de informações para a organização.

O objetivo principal de qualquer programa de governança de dados é organizar todas as atividades e serviços relacionados ao *big data* visando criar valor a partir dos dados e apoiar as análises gerenciais. Assim, a governança de dados em um contexto de big data deve implementar processos que evitem comportamentos inesperados, como vazamento de informações sigilosas, desconfiança nos dados apresentados e ineficiência na busca e recuperação das informações (KIM; CHO, 2017).

Esses sistemas de governança também devem permitir flexibilidade para acomodar mudanças nos requisitos ou padrões de governança à medida que a organização evolui. Isso inclui a revisão periódica das políticas e procedimentos e a realização de auditorias adicionais para garantir a conformidade contínua e a eficácia dos processos de governança de dados em um ambiente de big data em constante mudança.

Para estabelecer uma governança de dados eficaz, uma empresa deve implementar um modelo adaptado às suas necessidades específicas. Isso envolve mapear onde estão localizados os dados e desenvolver uma estratégia para coletar dados de diferentes sistemas que geram e armazenam informações dentro da organização (KIM, 2011). A tomada de decisão sobre o controle de acessos, responsabilidades e funções deve ser definida e especificada pelo modelo de governança da organização. Este modelo também é fundamental para estabelecer uma política de qualidade de dados, distribuir os dados dentro dos setores da empresa e fornecer normas e padrões para todos os que consomem os dados dentro da organização.

Dentre os elementos chave da governança de dados, a definição de padrões é essencial para garantir a qualidade dos dados dentro da organização. Isso é crucial para evitar problemas como duplicação, incompletude e desconfiança nas informações obtidas a partir dos dados. Além

disso, como argumentado por Kim e Cho (2017), com padrões claramente definidos, a seleção de tecnologias e ferramentas para o gerenciamento de dados torna-se mais precisa e eficaz.

Portanto, para garantir a qualidade de dados é essencial uma gestão de metadados que possibilite entender o contexto em que o dado está inserido, além de facilitar a criação de um repositório de dados sob responsabilidade da organização. Esse tema é apresentado a seguir.

3.4 Gestão de metadados

As informações em uma instituição, armazenadas em sistemas automatizados ou não, são elementos centrais para o seu funcionamento. Elas servem como insumos para os diversos processos de trabalho, permeando todas as áreas organizacionais. E para garantir a utilização eficiente e uniforme dessas informações, torna-se crucial mapeá-las, registrá-las e gerenciá-las de forma eficaz.

Nesse contexto, o gerenciamento de metadados se configura como uma ferramenta estratégica para otimizar a gestão da informação institucional. Por meio da criação e do gerenciamento de metadados, a instituição pode garantir a qualidade, a interoperabilidade e o uso eficiente da informação.

No modelo serial da *National Information Standards Organization* (NISO), citado por Sayão (2010), metadados é a informação estruturada que descreve, explica, localiza, ou possibilita que um recurso informacional seja fácil de recuperar, usar ou gerenciar.

A gestão de metadados é estratégica para a gestão de conhecimento organizacional (SHERMAN, 2006). Ela pressupõe a identificação de metadados que fazem sentido para os objetivos estratégicos da organização que, uma vez identificados nos processos e atividades da organização, devem ser gerenciados. Após constatar esses aspectos, devem ser definidos procedimentos práticos a serem implementados, visando uma gestão eficiente dos metadados.

Para entender o papel vital dos metadados na gestão de dados, informação e conhecimento pode-se imaginar uma grande biblioteca, com um grande volume de livros e revistas, mas sem catalogação. Sem um sistema de catalogação os leitores terão dificuldades para buscar um livro ou algum tópico específico. A catalogação não fornece apenas informações necessárias sobre onde o livro está, ela disponibiliza padrões para encontrá-lo, usando diferentes pontos de referência (área,

autor, título, etc.). Sem um catálogo, encontrar um livro específico é quase impossível. Da mesma forma, uma organização sem metadados é como uma biblioteca sem catalogação (DAMA, 2017).

Qin, Huang e Wang (2018), citam os seguintes benefícios da implementação de um sistema de metadados:

- descrever os atributos de conteúdo dos objetos informacionais;
- organizar os objetos de informação e estabelecer relações entre eles, proporcionando aos usuários uma experiência em vários níveis e canais, ajudando a localizar rápido e facilmente os recursos de informação que realmente necessitam;
- auxiliar os usuários a ter um conhecimento e entendimento básico dos objetos informacionais, sem a necessidade de navegar nos objetos de informação em si;
- salvar informações relacionadas ao processamento, arquivamento, uso e gerenciamento de recursos de informação, bem como gestão de direitos, direitos autorais, propriedade, direitos de uso, medidas anti-falsificação, marcas d'água eletrônicas e assinaturas eletrônicas;
- fornecer informações sobre a localização do próprio recurso de informação, como DOI (*Digital Object Identifier*), URL (*Uniform Resource Locator*), URN (*Uniform Resource Name*), etc., permitindo, assim, saber exatamente onde o objeto de informação está localizado, facilitando a obtenção de informações.

Na literatura de metadados existem alguns vocabulários que servem de modelos para auxiliar na implantação de catálogos em organizações. Por meio da sua utilização, é possível controlar o uso de termos em campos de metadados, garantindo a padronização e facilitando a recuperação dos dados. Rozsa, Dutra e Nhacuongue (2017) afirmam que os vocabulários permitem a definição de temas, propriedades e relações entre diferentes categorias e são usados na estruturação e contextualização dos dados na Web de Dados, para que possam ser compreendidos e processados automaticamente.

O vocabulário DCAT, ou *Data Catalog Vocabulary*, tem o propósito de orientar a publicação de catálogos de dados na web, e embora tenha sido originalmente criado para catálogos de dados

governamentais, pode ser utilizado em outros contextos (ALBERTONI et al., 2020). O DCAT é uma ferramenta essencial que possibilita aos editores descrever conjuntos de dados e serviços de dados em um catálogo de forma padronizada. Utilizando um modelo e vocabulário consagrados, o DCAT simplifica tanto o consumo quanto a agregação de metadados provenientes de diversos catálogos. Essa abordagem padronizada não apenas facilita a compreensão e o acesso aos dados, mas também promove a interoperabilidade e a integração eficaz de informações provenientes de fontes diversas (ALBERTONI et al., 2020). O DCAT também oferece a possibilidade de descrever relações entre conjuntos de dados e entre conjuntos de dados e outros recursos (ALBERTONI et al., 2020).

Outro vocabulário bastante utilizado na literatura, desenvolvido pelo *Dublin Core Metadata Initiative* (DCMI), é *Dublin Core*, que é um dos conjuntos de elementos de metadados mais amplamente utilizados no mundo. O *Dublin Core* inclui elementos básicos, como título, criador, data e formato, bem como elementos adicionais para descrever conteúdos mais complexos, como recursos educacionais, qualidades físicas de um recurso e direitos de uso. O DCMI tem um amplo alcance e é líder no desenvolvimento de padrões de metadados para a web (*DCMI USAGE BOARD*, 2020; WEIBEL et al., 1998). Também desenvolveu uma série de extensões ao *Dublin Core*, incluindo o *DCMI Type Vocabulary*, que fornece termos para descrever o tipo de recurso, e o *DCMI Terms*, que é um vocabulário semântico para descrever termos de assunto e conceitos.

Além desses dois vocabulários existe a iniciativa FAIR, que visa auxiliar na gestão de dados, com a premissa de tornar os dados mais fáceis de encontrar (*Findable*), acessíveis (*Acessible*), interoperáveis (*Interoperable*) e reutilizáveis (*Reusable*). Estes princípios não se configuram como regras e sim como um conjunto de orientações que podem ser seguidas ou não pelos repositórios (FELIPE; DOS SANTOS, 2022). A partir destes princípios foram criados os serviços FAIR, que buscam garantir que eles sejam aplicados na gestão de dados e possam contribuir para a interoperabilidade e reuso efetivo.

Esses princípios foram operacionalizados em uma ferramenta - o software FAIR Data Point, com a finalidade de recuperar informações mediante um software. Ele compreende os metadados que auxiliam na recuperação de informações por parte dos usuários, além de contribuir para que os dados se tornem FAIR (FELIPE; SANTOS, 2022)

O FAIR *Data Point* foi desenvolvido para lidar com diversos princípios FAIR relacionados a metadados, bem como para enfrentar desafios comuns encontrados por consumidores e produtores de objetos digitais. Uma das principais preocupações associadas aos princípios FAIR é garantir a clara distinção entre dados e metadados, juntamente com a implementação de mecanismos de acesso aos metadados que sejam consistentes e padronizados. Isso evita que os mesmos metadados sejam acessados de maneiras diferentes, minimizando assim o risco de inconsistências ou discrepâncias nas informações associadas a um mesmo conjunto de metadados (DA SILVA SANTOS, 2023).

Segundo Da Silva Santos (2023), os objetivos do software FAIR *Data Point* são:

- permitir que proprietários/criadores/editores exponham os metadados de seus objetos digitais de uma maneira que siga os princípios FAIR;
- permitir que consumidores/usuários descubram informações sobre objetos digitais de interesse;
- fornecer esses metadados de forma que sejam facilmente utilizados por máquinas.

O FAIR Data Point (2020) propõe algumas orientações de metadados para gerenciar as informações de bases de dados. Esses metadados são apresentados e descritos no Quadro 1, no tópico 4.2, e são a base para o desenvolvimento da proposta de *template* para catalogação de metadados das bases de dados do Observatório FIESC. Justifica-se esta escolha por acreditar que os princípios e objetivos do FAIR *Data Point*, descritos anteriormente, vão ao encontro das necessidades de gestão de metadados das bases de dados do Observatório FIESC, além de agregar vocabulários amplamente utilizados na literatura de metadados, como o DCAT e o vocabulário proposto pelo DCMI.

4 CONSTRUÇÃO DA PROPOSTA DE *TEMPLATE* PARA CATALOGAÇÃO DE METADADOS DAS BASES DE DADOS DO OBSERVATÓRIO A

Este item trata da construção da proposta de *template* para catalogação de metadados das bases de dados do Observatório A. Nesse sentido, inicialmente apresenta-se o Observatório A, foco do estudo e na sequência, a proposta de *template* com os elementos que o compõem.

4.1 O Observatório A

O Observatório A tem como foco a atuação em tecnologia e inovação, contribuindo para prover soluções para a inovação e o desenvolvimento tecnológico da indústria, e fortalecer a integração com os ecossistemas de inovação e tecnologia.

O Observatório A é um centro de inteligência com dados, tecnologias e soluções para suporte às decisões que impulsionam a competitividade e o desenvolvimento sustentável da indústria e seus *stakeholders*. Possui uma plataforma sistêmica e tecnologicamente customizada para suportar uma infraestrutura de *big data*, *data analytics* e sistemas de inteligência, abastecidos por cerca de 100 bases de dados públicos e privados do do Brasil e do mundo. Como suporte ao desenvolvimento de todas as suas atividades, o Observatório A conta com uma governança de dados, que contempla a adoção de um conjunto de políticas, estratégias, processos, práticas e ferramentas para garantir o gerenciamento e utilização eficaz dos dados, visando a segurança e o máximo de valor possível para o sistema indústria e seus *stakeholders*.

Atualmente, existe um *big data* hospedado em nuvem, que carece de regras, normas, padrões e definições de papéis e responsabilidades para garantir uma boa utilização e organização do dado dentro desse ambiente. Portanto, faz-se necessário um processo de estruturação da governança dos dados do *big data*, e também a criação de um conjunto de elementos que irão guiar a equipe em sua estruturação e no gerenciamento dos dados. Um dos elementos fundamentais de um programa de governança de dados, são os metadados que precisam ter um processo de gerenciamento e manutenção constante. Devido às características das suas bases de dados, o Observatório necessita de um controle dos metadados para auxiliar no entendimento da equipe em relação às bases de dados disponíveis, bem como as informações que podem ser geradas a partir delas. Com isso, este estudo propõe um *template* para catalogação dos metadados das bases utilizadas pelo Observatório. É um processo inicial, porém fundamental para começar uma gestão de metadados e contribuir com a governança de dados do Observatório.

4.2 Proposta de *template* para catalogação de metadados das bases de dados do Observatório A

Para elaborar um modelo de catalogação de metadados das bases de dados para o Observatório A, como já mencionado na revisão de literatura, utilizou-se a iniciativa FAIR Data Point (2020), a fim de estruturar o seu formato e atender tanto as características dos dados trabalhados

pele Observatório, como os objetivos para o seu uso. Esta iniciativa tem foco em metadados descritivos, ou seja, metadados que buscam detalhar o significado daquele dado. Os metadados indicados por essa alternativa, de acordo com FAIR Data Point (2020) são apresentados no Quadro 1.

Quadro 1 - Metadados do modelo FAIR *Data Point* (2020)

Vocabulário	Caráter	Descrição
<i>Type</i>	Obrigatório	Necessário ser do tipo DCAT: <i>dataset</i>
<i>Title</i>	Obrigatório	Nome do conjunto de dados
<i>Publisher</i>	Obrigatório	Organização ou pessoa responsável pelo conjunto de dados
<i>HasVersion</i>	Obrigatório	Versão do conjunto de dados
<i>Description</i>	Opcional	Descrição do conjunto de dados com a <i>tag</i> de idioma
<i>ConformTo</i>	Opcional	Especificação do esquema de metadados, do conjunto de dados (ShEx)
<i>MetadataIssued</i>	Opcional	Data da criação da entrada do conjunto de dados
<i>MetadataModified</i>	Opcional	Data da última modificação do conjunto de dados
<i>Language</i>	Opcional	Idioma do conjunto de dados
<i>License</i>	Opcional	Licença associada ao conjunto de dados
<i>Rights</i>	Opcional	Direitos sobre o conjunto de dados
<i>AccessRights</i>	Opcional	Descrição dos direitos do conjunto de dados
<i>IsPartOf</i>	Obrigatório	Relação com os metadados pai

<i>Keyword</i>	Opcional	Palavras-chave relacionadas ao conjunto de dados com a <i>tag</i> de idioma
<i>MetadataIdentifier</i>	Obrigatório	Identificador de metadado
<i>Distribution</i>	Obrigatório	Lista de URLs de identificação
<i>Theme</i>	Obrigatório	Lista de termos que descrevem o assunto do conjunto de dados
<i>ContactPoint</i>	Opcional	Ponto de contato responsável pelo conjunto de dados
<i>LandingPage</i>	Opcional	Página inicial do conjunto de dados

Fonte: FAIR *Data Point* (2020)

A partir do Quadro 1, buscou-se fazer uma adaptação à realidade do Observatório FIESC, levando em conta os seus conjuntos de dados e suas relações com seus produtos, projetos, demandas e serviços. Como já apresentado, segundo o FAIR Data Point (2020), os metadados são de caráter obrigatório ou opcional.

Os itens de caráter obrigatório para o *template* para a catalogação de metadados das bases de dados do Observatório são:

- *Type*: se refere a como o tema em relação ao que aquele conjunto de dados se relaciona, por exemplo: economia, saúde, educação etc.;
- *Title*: se refere ao nome original da fonte em que este conjunto de dados foi extraído, por exemplo CAGED;
- *Publisher*: se refere ao responsável por atualizar e garantir a qualidade daquele *dataset*;
- *HasVersion*: é um importante metadado que ajudará no versionamento das bases de dados do Observatório;

- *IsPartOf*: não se aplica ao presente trabalho, que tem como limitação apenas trabalhar com as bases de dados completas e não os *datamarts* gerados a partir delas;
- *MetadataIdentifier*: é uma aplicação importante para codificar e facilitar a identificação das bases de dados;
- *Distribution*: pode servir para localizar onde é possível encontrar a base de dados completa, por meio do endereço em que ela se encontra nos sistemas do Observatório;
- *Theme*: pode ser utilizado para facilitar a compreensão do que a base de dados se refere, por meio de um conjunto de termos relacionados à base.

Além dos elementos mencionados anteriormente, é possível incluir alguns itens que, embora sejam considerados opcionais pelo FAIR *Data Point* (2020), são tratados como obrigatórios no *template* para o Observatório FIESC. São eles:

- *MetadataModified*: indica o momento em que a base de dados foi modificada ou atualizada;
- *Keyword*: objetiva trazer palavras chaves que ajudam a identificar quais respostas aquele conjunto de dados pode dar.

Já os metadados de caráter opcional, tanto para o FAIR *Data Point*, como para o modelo proposto do artigo, são:

- *Description*: para apresentar uma explicação mais detalhada sobre a base de dados;
- *MetadataIssued*: para registrar a data que aquela base foi inserida nos sistemas do Observatório a partir da fonte original;
- *License*, *rights* e *accessRights*: serão utilizadas somente quando a base de dados for privada, ou seja, foi adquirida pelo Observatório e possui termos de uso;

- *ContactPoint*: terá um caráter opcional, visto que a equipe não é grande e já existe a identificação do responsável no metadado *Publisher*.

Os outros metadados propostos por FAIR Data Point (2020) e não descritos, não são aplicáveis aos conjuntos de dados do Observatório. Porém, devido a experiência do autor do presente artigo em seu contexto, o *template* terá outros metadados, de caráter obrigatório, que são cruciais para atender os princípios FAIR. São eles:

- *Location*: utilizado para saber onde a base de dados se encontra armazenada nos sistemas do Observatório;
- *Frequency*: utilizado para indicar a periodicidade de atualização;
- *Operator*: diferente do *Publisher* que é o responsável pela base de dados se manter atualizada e com dados fidedignos, o *Operator* vai realizar a atividade de atualizar a base de dados, por meio de scripts de ETL (*extract, treatment, load*), ou de outra forma que foi definida pelo *Publisher*.

Com base no exposto, no Quadro 2, apresenta-se uma proposta de *template* para catalogação de metadados das bases de dados do Observatório FIESC.

Quadro 2 - Proposta de *template* para catalogação de metadados das bases de dados do Observatório FIESC

Vocabulário	Caráter	Descrição
<i>Type</i>	Obrigatório	Necessário ser do tipo DCAT: <i>dataset</i>
<i>Title</i>	Obrigatório	Nome do conjunto de dados
<i>Publisher</i>	Obrigatório	Organização ou pessoa responsável pelo conjunto de dados
<i>HasVersion</i>	Obrigatório	Versão do conjunto de dados
<i>MetadataIdentifier</i>	Obrigatório	Identificador de metadado

<i>Distribution</i>	Obrigatório	Lista de URLs de identificação
<i>Theme</i>	Obrigatório	Lista de termos que descrevem o assunto do conjunto de dados
<i>Description</i>	Obrigatório	Descrição do conjunto de dados com a <i>tag</i> de idioma
<i>MetadataModified</i>	Obrigatório	Data da última modificação do conjunto de dados
<i>License</i>	Opcional	Licença associada ao conjunto de dados
<i>Rights</i>	Opcional	Direitos sobre o conjunto de dados
<i>AccessRights</i>	Opcional	Descrição dos direitos do conjunto de dados
<i>ContactPoint</i>	Opcional	Ponto de contato responsável pelo conjunto de dados
<i>Location</i>	Obrigatório	Local onde a base foi armazenada após ser extraída
<i>Frequency</i>	Obrigatório	Frequência que ela é atualizada
<i>Operator</i>	Obrigatório	Quem atualiza a base de dados

Fonte: Elaborado pelo autor (2024)

Para exemplificar a aplicação do template gerado, foi selecionada a base do CAGED, que é uma das bases mais utilizadas pelo Observatório A, tanto para análises econômicas quanto para o desenvolvimento de projetos e produtos (Quadro 3).

Quadro 3 - Aplicação do *template* de metadados - CAGED

Vocabulário	Caráter	Descrição
<i>Type</i>	Obrigatório	Economia
<i>Title</i>	Obrigatório	Novo CAGED

<i>Publisher</i>	Obrigatório	Analista A
<i>HasVersion</i>	Obrigatório	Décima segunda atualização
<i>MetadataIdentifier</i>	Obrigatório	BD01
<i>Distribution</i>	Obrigatório	http://pdet.mte.gov.br/microdados-rais-e-caged
<i>Theme</i>	Obrigatório	Movimentação de empregos, admitidos, desligados
<i>Description</i>	Obrigatório	Conjunto de dados que trazem a movimentação de empregos de todo país
<i>MetadataModified</i>	Obrigatório	16/01/2024
<i>License</i>	Opcional	Não se aplica
<i>Rights</i>	Opcional	Não se aplica
<i>AccessRights</i>	Opcional	Não se aplica
<i>ContactPoint</i>	Opcional	analistaa@gmail.com
<i>Location</i>	Obrigatório	G:/drivecompartilhado/basesdedados/caged
<i>Frequency</i>	Obrigatório	Mensal
<i>Operator</i>	Obrigatório	Analista B

Fonte: Do autor (2024)

Outros dois exemplos de aplicação do template são apresentados nos Quadros 4 e 5, com bases de outras temáticas, também utilizadas para o desenvolvimento de projetos e produtos do Observatório que são o INEP Educação Básica e o DATASUS leitos.

Quadro 4 - Aplicação do *template* de metadados - INEP Educação básica

Vocabulário	Caráter	Descrição
-------------	---------	-----------

<i>Type</i>	Obrigatório	Educação
<i>Title</i>	Obrigatório	INEP Censo básico
<i>Publisher</i>	Obrigatório	Analista D
<i>HasVersion</i>	Obrigatório	Vigésima primeira atualização
<i>MetadataIdentifier</i>	Obrigatório	BD03
<i>Distribution</i>	Obrigatório	https://dados.gov.br/dados/conjuntos-dados/inep-microdados-do-censo-escolar-da-educacao-basica#:~:text=Os%20Microdados%20do%20Censo%20Escolar,para%20a%20menor%20unidade%20de
<i>Theme</i>	Obrigatório	Ensino básico, censo
<i>Description</i>	Obrigatório	Conjunto de dados que trazem as características da educação brasileira nível básica, com granularidade de município
<i>MetadataModified</i>	Obrigatório	23/10/2023
<i>License</i>	Opcional	Não se aplica
<i>Rights</i>	Opcional	Não se aplica
<i>AccessRights</i>	Opcional	Não se aplica
<i>ContactPoint</i>	Opcional	analistad@gmail.com
<i>Location</i>	Obrigatório	G:/drivecompartilhado/basesdedados/inep_ensino_basico
<i>Frequency</i>	Obrigatório	Anual
<i>Operator</i>	Obrigatório	Analista F

Fonte: Do autor (2024)

Quadro 5 - Aplicação do *template* de metadados - DATASUS leitos

Vocabulário	Caráter	Descrição
<i>Type</i>	Obrigatório	Saúde
<i>Title</i>	Obrigatório	DATASUS Leitos
<i>Publisher</i>	Obrigatório	Analista G
<i>HasVersion</i>	Opcional	Trigésima segunda atualização
<i>MetadataIdentifier</i>	Obrigatório	BD05
<i>Distribution</i>	Obrigatório	http://tabnet.datasus.gov.br/cgi/tabcgi.exe?cnes/cnv/leiintsc.def
<i>Theme</i>	Obrigatório	Hospitais, leitos, SUS
<i>Description</i>	Obrigatório	Conjunto de dados que trazem a quantidade de leitos públicos e privados de todos os municípios do país
<i>MetadataModified</i>	Obrigatório	02/02/2024
<i>License</i>	Opcional	Não se aplica
<i>Rights</i>	Opcional	Não se aplica
<i>AccessRights</i>	Opcional	Não se aplica
<i>ContactPoint</i>	Opcional	analistaG@gmail.com
<i>Location</i>	Obrigatório	G:/drivecompartilhado/basesdedados/datasus_leitos
<i>Frequency</i>	Obrigatório	Mensal

<i>Operator</i>	Obrigatório	Analista H
-----------------	-------------	------------

Fonte: Do autor (2024)

Com esses três exemplos ilustrados nos quadros anteriores é possível visualizar as informações que são necessárias sobre cada base, desde quem é responsável por manter a base atualizada e com dados de qualidade, passando pelo tema, até onde está armazenada. Essa proposta é uma primeira etapa para implementar uma gestão de metadados no Observatório A. Esse catálogo disponível para toda a equipe, ajudará a entender o que existe de recurso para responder às demandas, desenvolver produtos, apoiar em projetos e realizar análises, otimizando o processo de entendimento em relação a base, auxiliando na busca desses dados nos sistemas internos e apoiando na gestão de qualidade de dados do Observatório.

É importante ressaltar que esse catálogo de metadados foi proposto apenas para base de dados, mas pode servir como suporte para desenvolver outros tipos de catálogos, seja de indicadores, projetos e/ou documentos. Com isso, a gestão de conhecimento do Observatório vai ganhando maturidade e os riscos inerentes a retenção e falta de disseminação do conhecimento vão sendo mitigados.

Para iniciar a aplicação da proposta gerada para a catalogação de metadados, sugere-se a utilização de planilhas de *excel* e, conforme o catálogo vai ganhando aderência, recomenda-se a avaliação de compra de ferramentas mais robustas e apropriadas, já que utilizar apenas planilha de *excel* pode tornar o trabalho de atualização do catálogo muito manual e com o tempo poderá ficar em desuso. Como estes metadados foram propostos com base na orientação de FAIR *Data Point* (2020), que tem como um dos objetivos tornar os metadados facilmente utilizados por máquinas, esse processo de transição ficará facilitado.

5 CONCLUSÃO

Instituições de grande porte geram e utilizam uma grande quantidade de dados. Nessas instituições, diferentes indivíduos possuem diferentes níveis de conhecimento em relação aos dados, mas dificilmente um deles terá domínio completo sobre todos os dados. Assim, um programa de governança de dados é fundamental, não apenas por estabelecer boas práticas de gerenciamento de dados, mas também para disseminar as informações relacionadas aos dados.

Uma das atividades essenciais de governança de dados é a documentação desses dados de modo adequado. Caso contrário, a instituição corre o risco de perder valiosos conhecimentos sobre seu passado, sua existência e sobre a possibilidade de aplicação desses conhecimentos para o desenvolvimento de soluções que atendam aos objetivos de negócio.

O Observatório A tem como objetivo fornecer informações para a tomada de decisões, principalmente de gestores industriais. Para isso, conta com mais de 100 bases de dados, principalmente de dados abertos governamentais. Devido a quantidade e a natureza dessas bases, gerir as informações e o conhecimento que são extraídos a partir delas se torna uma tarefa complexa. Foi nesse contexto que esse artigo apresentou uma proposta de *template* para catalogação de metadados das bases de dados do Observatório.

O *template* foi desenvolvido tendo como base o FAIR *Data Point* (2020), uma das iniciativas encontradas na literatura, que visa auxiliar na gestão de dados, com a premissa de tornar os dados mais fáceis de encontrar, acessíveis, interoperáveis e reutilizáveis.

Os elementos do *template* desenvolvido compreendem metadados essenciais, tais como título, temática, publicação e responsável, entre outros, além da distinção entre metadados obrigatórios e opcionais. Salieta-se que foi necessário adicionar alguns metadados adaptados à realidade do Observatório A, para auxiliar no processo de recuperação de informações a partir dos dados disponíveis. As bases CAGED, INEP Educação básica e DATASUS leitos foram utilizadas para exemplificar a aplicação do *template* gerado.

A relevância do *template* para o Observatório A se apoia nas etapas do processo de busca de dados para responder às demandas, desenvolver produtos e serviços, já que os dados, majoritariamente dados abertos governamentais são os principais insumos para estas atividades. Tendo um catálogo de dados atualizado e bem estruturado, acredita-se que haverá maior otimização desse processo, minimizando os riscos associados à qualidade de dados.

Ademais, o *template* proposto pode contribuir para reduzir o tempo de busca de dados relativos a um determinado tema, para a gestão de conhecimento e ainda, para mensurar as potencialidades de utilização dos dados das bases, visando o fortalecimento do Observatório A como um centro de inteligência de referência.

Cabe destacar que a manutenção do *template* gerado inicialmente será realizada de forma manual. Contudo, frente ao volume considerável de bases de dados do Observatório, é imperativo avaliar a possibilidade de adquirir uma ferramenta que automatize esse processo. Enquanto a manutenção permanecer manual, é recomendável estabelecer um conjunto de procedimentos padronizados para que todos os intervenientes possam contribuir de forma consistente para o catálogo e utilizá-lo para aprimorar suas atividades.

Por fim, é fundamental validar a eficácia do *template* gerado, o que requer a colaboração de especialistas da área. Recomenda-se, portanto, para futuras pesquisas, a realização de testes de conteúdo, tanto quantitativos (medição do Índice de Validade de Conteúdo - IVC) quanto qualitativos (envolvendo um comitê de especialistas), utilizando formulários estruturados. Além disso, é possível considerar a expansão do catálogo de metadados para incluir indicadores, produtos e documentos relacionados às atividades do Observatório A.

REFERÊNCIAS

ABRAHAM, Rene; SCHNEIDER, Johannes; VOM BROCKE, Jan. Data governance: a conceptual framework, structured review, and research agenda. *International journal of information management*, v. 49, p. 424-438, 2019. **International Journal of Managing Projects in Business**, 2016.

ALBERTONI, Riccardo et al. Data catalog vocabulary (DCAT)-version 2. **World Wide Web Consortium**, 2020.

ASPINALL, Peter J.; JACOBSON, Bobbie; CASTILLO-SALGADO, Carlos. Establishing and sustaining health observatories serving urbanized populations around the world: scoping study and survey. **The European Journal of Public Health**, v. 26, n. 4, p. 681-686, 2016.

BARBIERI, Carlos. **Governança de dados: práticas, conceitos e novos caminhos**. Alta Books, 2019.

BARBOSA, Jessyca Lorena Pereira. **Um estudo sobre observatórios através de um mapeamento sistemático da literatura**. 2020. Dissertação (Mestrado em Ciência da Computação) – Universidade Federal de Pernambuco, Recife, 2020.

BENTO, Patrícia; NETO, Miguel; CÔRTE-REAL, Nadine. How data governance Frameworks can leverage data-driven decision making: A sustainable approach for data governance in

organizations. In: **2022 17th Iberian Conference on Information Systems and Technologies (CISTI)**. IEEE, 2022.

CARRETERO, Ana et al. MAMD 2.0: Environment for data quality processes implantation based on ISO 8000-6X and ISO/IEC 33000. **Computer Standards & Interfaces**, v. 54, p. 139-151, 2017.

DA SILVA SANTOS, Luiz Olavo Bonino et al. FAIR data point: a FAIR-oriented approach for metadata publication. **Data Intelligence**, v. 5, n. 1, p. 163-183, 2023.

DAMA, I. **DAMA-DMBOK**: data management body of knowledge. 2017.

BOARD, DCMI Usage. **Dublin core metadata initiative (DCMI) metadata terms**. 2020.

FAIR DATA POINT. 2020. Disponível em: <https://www.fairdatapoint.org/>. Acesso em: 10 mar. 2024.

FELIPE, Carla Marques; SANTOS, Raimunda Fernanda. Avaliação de metadados em repositórios de dados de pesquisa sobre biodiversidade. **Em Questão**, 2022.

GUSMÃO, Regina. Estruturas e dispositivos nacionais de produção e difusão de indicadores de CT&I: deficiências e possíveis avanços. **Parcerias Estratégicas**, Brasília, v. 20, p. 1075–1098, 2005.

KIM, Seok-soo. A Case study of implementation data governance for enterprise architecture. **Journal of Information Technology and Architecture**, v. 8, n. 3, p. 255-265, 2011.

KIM, Hee Young; CHO, June-Suh. Data governance framework for big data implementation with a case of Korea. In: **2017 IEEE International Congress on Big data (Big Data Congress)**. IEEE, 2017. p. 384-391.

LÓPEZ, Juan Ignacio Torregrosa et al. **Proyecto de creación de un observatorio ambiental de la industria de bolivia**. 2012.

MARCONI, Marina. A.; LAKATOS, Eva. M. **Técnicas de pesquisa**: planejamento e execução de pesquisas, amostragens e técnicas de pesquisa, elaboração, análise e interpretação de dados. São Paulo: Atlas, 2007.

MINUSSI, Sandro Gindri et al. Considerações sobre estado da arte, levantamento bibliográfico e pesquisa bibliográfica: relações e limites. **Gestão Universitária**, v. 9, n. 2, 2018.

OPEN KNOWLEDGE FOUNDATION. **For a fair, sustainable and open future**. 2022. Disponível em: <https://okfn.org>. Acesso em: 10 mar. 2024.

ORTEGA, Cristina; VALLE, Roberto San Salvador. Nuevos retos de los observatorios culturales. **Boletín Gestión Cultural Nº19: Observatorios culturales en el mundo**, enero, 2010.

OTTO, Boris. Managing the business benefits of product data management: the case of Festo. **Journal of Enterprise Information Management**, v. 25, n. 3, p. 272-297, 2012.

PHÉLAN, Mauricio. La red observatorios locales de Barcelona, España. Un estudio de casos para diseñar una propuesta nacional. Fermentum. **Revista Venezolana de Sociología y Antropología**, v. 17, n. 48, p. 96-122, 2007.

PORTAL BRASILEIRO DE DADOS ABERTOS. Disponível em: <https://dados.gov.br/home>. Acesso em: 12 jan. de 2019.

POSSAMAI, Ana Júlia. **Dados abertos no governo federal brasileiro: desafios de transparência e interoperabilidade**. 2016. Tese (Doutorado em Ciências Políticas) – Universidade Federal do Rio Grande do Sul, Porto Alegre, 2016.

QIN, Duo; HUANG, Lei; WANG, Ying. Construction of Railway Metadata Management System Based on Metadata Content Model and CWM Exchange Mechanism. In: **2018 8th International Conference on Logistics, Informatics and Service Sciences (LISS)**. IEEE, 2018.

REYES, Yoel Ramón Sarmiento; FERNÁNDEZ, Mercedes Delgado; ABREU, Marta Beatriz Infante. Observatorios: clasificación y concepción en el contexto iberoamericano. **Revista Cubana de Información en Ciencias de la Salud (ACIMED)**, v. 30, n. 2, 2019.

ROZSA, Vitor; DUTRA, Moisés-Lima; NHACUONGUE, Januário Albino. Linked open data no contexto acadêmico: identificação e análise de vocabulários utilizados na academia e na pesquisa científica. **Brazilian Journal of Information Science**, v. 11, n. 3, p. 34-52, 2017.

SANTORO, Paula; XAVIER, Iara. **Observatórios e sistemas de informação em São Paulo, Brasil: resultados**. São Paulo: Pólis, 2009.

SANTOS, Ana Cristina Gomes. Dados abertos e ciência aberta: como as Universidades Federais brasileiras se apresentam nesse horizonte. **Biblios. Revista de Bibliotecología y Ciencias de la Información**, Pittsburgh, n. 78, p. 01-16. 2020.

SAYÃO, Luís Fernando. Uma outra face dos metadados: informações para a gestão da preservação digital. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, v. 15, n. 30, p. 1-31, 2010.

SHERMAN, Rick. Align metadata and business initiatives. **Information Management**, v. 16, n. 1, 2006.

SILVA, Antonio Waldimir Leopoldino da et al. Observatórios brasileiros de meio ambiente e sustentabilidade: diagnóstico e análise. In: **XV Congresso Latino-Iberoamericana de Gestão de Tecnologia**. 2013.

SILVA, Edna L. Pesquisa qualitativa: exploratória, descritiva e explicativa. **Revista de Pesquisa Qualitativa**, v. 5, n. 2, 2018.

SOARES, Lilian Campos; FERNEDA, Edilson; DO PRADO, Hércules Antonio. Observatórios: um levantamento do estado do conhecimento. **Brazilian Journal of Information Science**, v. 12, n. 3, p. 86-110, 2018.

TALLON, Paul P.; RAMIREZ, Ronald V.; SHORT, James E. The information artifact in IT governance: toward a theory of information governance. **Journal of Management Information Systems**, v. 30, n. 3, p. 141-178, 2013.

TRZECIAK, Dorzeli. S. **Modelo de observatório para arranjos produtivos locais**. 2009. Tese (Doutorado em Engenharia de Produção) – Universidade Federal de Santa Catarina, Florianópolis, 2009.

WANG, Richard Y. A product perspective on total data quality management. **Communications of the ACM**, v. 41, n. 2, p. 58-65, 1998.

WANG, Richard Y.; STRONG, Diane M. Beyond accuracy: what data quality means to data consumers. **Journal of Management Information Systems**, v. 12, n. 4, p. 5-33, 1996.

WEILL, Peter; ROSS, Jeanne W. IT governance: how top performers manage IT decision rights for superior results. **Harvard Business Press**, 2004.