



UNIVERSIDADE FEDERAL DO RIO DE JANEIRO – UFRJ
ESCOLA DE COMUNICAÇÃO – ECO
INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA – IBICT
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO – PPGCI

NÚBIA HELENA MOREIRA SILVA CERVEIRA

**NANOPUBLICAÇÃO COMO ABORDAGEM PARA A RECUPERAÇÃO DA
INFORMAÇÃO EM DOCUMENTOS TÉCNICOS DE ENGENHARIA**

Dissertação de Mestrado

Fevereiro de 2025



NÚBIA HELENA MOREIRA SILVA CERVEIRA

**NANOPUBLICAÇÃO COMO ABORDAGEM PARA A RECUPERAÇÃO DA
INFORMAÇÃO EM DOCUMENTOS TÉCNICOS DE ENGENHARIA**

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação do convênio entre o Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) e a Escola de Comunicação da Universidade Federal do Rio de Janeiro (Eco UFRJ), como requisito parcial à obtenção do título de mestre em Ciência da Informação.

Orientadora Prof.^a Dr.^a Luana Farias Sales

Rio de Janeiro

2025

CIP - Catalogação na Publicação

C419n Cerveira, Núbia Helena Moreira Silva
Nanopublicação como abordagem para a recuperação da informação em documentos técnicos de engenharia / Núbia Helena Moreira Silva Cerveira. -- Rio de Janeiro, 2025.
122 f.

Orientadora: Luana Farias Sales.
Dissertação (mestrado) - Universidade Federal do Rio de Janeiro, Escola da Comunicação, Instituto Brasileiro de Informação em Ciência e Tecnologia, Programa de Pós-Graduação em Ciência da Informação, 2025.

1. Nanopublicação. 2. Documentação de engenharia. 3. Curadoria digital. 4. Recuperação da informação.
I. Sales, Luana Farias, orient. II. Título.

NÚBIA HELENA MOREIRA SILVA CERVEIRA

**NANOPUBLICAÇÃO COMO ABORDAGEM PARA A RECUPERAÇÃO DA
INFORMAÇÃO EM DOCUMENTOS TÉCNICOS DE ENGENHARIA**

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação do convênio entre o Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) e a Escola de Comunicação da Universidade Federal do Rio de Janeiro (Eco UFRJ), como requisito parcial à obtenção do título de mestre em Ciência da Informação.

Rio de Janeiro, 26 de fevereiro de 2024.

Prof.^a Dr.^a Luana Farias Sales – (Orientadora)

PPGCI IBICT/Eco-UFRJ

Prof.^a Dra.^a Lillian Maria Araujo de Rezende Alvares

PPGCI/IBICT – ECO/UFRJ (Membro Efetivo)

Prof.^a Dra.^a Silvana Aparecida Borsetti Gregorio Vidotti

UNESP (Membro Efetivo)

Prof.^a Dr.^a Naira Christofolletti Silveira

PPGCI/IBICT (Membro suplente)

A Deus, pela vida e pelas oportunidades encontradas ao longo da caminhada.

A meus pais, por me proporcionarem amor e educação.

Aos garotos de minha vida, Ivaldo, Gabriel e Guilherme.

AGRADECIMENTOS

A Deus, meu motivo de força, fé e coragem, sem Ele nada teria sido possível.

A todas as oportunidades de aprendizado, de novos conhecimentos, de desbravar, sempre com frio na barriga, mas nunca sem fé e coragem.

À Luana Sales, minha orientadora, pela generosidade em ter aceitado esse desafio e por todo carinho durante essa caminhada, foi longa, mas acabou, professora! Gratidão!

Aos professores do PPGCI/IBICT, que contribuíram para a construção desta dissertação, nas aulas preciosas.

À Professora Lena Vânia (In Memoriam), de quem tive a honra de ser aluna, e pude observar a sua paixão pela Ciência da Informação e pela docência.

Aos membros da Banca por aceitarem o convite para compartilhar este momento de conhecimento e aprendizado.

À Gabriel e Guilherme, que me ensinam e me inspiram todos os dias.

À Lene, minha amiga, minha irmã, meu suporte. Ela não imagina a contribuição que me deu para conseguir elaborar este trabalho.

À minha querida amiga Magna que, mesmo no turbilhão de emoções e surpresas na sua vida, comemorou comigo o primeiro parágrafo que consegui escrever desta preciosidade.

À minha amiga e “chefa” Márcia Maria, pela oportunidade e liberação para assistir as aulas do mestrado.

A todos os colegas do mestrado, que durante as aulas, em suas discussões maravilhosas, contribuíam para o meu conhecimento.

A Ivaldo, companheiro, amigo, que suportou minha ausência e minha fala sobre “nanopublicação”. Deu certo marido!!! Obrigada!

Ao meu amado Pai (In Memoriam), que durante esse processo nos deixou. Ele estaria feliz por minha conquista! Sei que está comemorando.

À minha mãezinha, que já não lembra meu nome, mas sei que se lembrasse, estaria feliz e comemorando comigo.

“The goal is to have a world in which all of the science literature is online, all of the science data is online, and they interoperate with each other.”

Jim Gray

RESUMO

CERVEIRA, Núbia Helena Moreira Silva. **Nanopublicação como abordagem para a recuperação da informação em documentos técnicos de engenharia**. 2025. Dissertação (Ciência da Informação) - Instituto Brasileiro de Informação em Ciência e Tecnologia e a Escola de Comunicação da Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2025.

Investiga a aplicabilidade das nanopublicações para aprimorar a recuperação, o uso e o reuso da informação em documentos técnicos de engenharia, visando aumentar a qualidade da gestão da informação técnica dentro do ciclo de vida de projetos de engenharia. Para verificar sua aplicabilidade, foi feito um levantamento bibliográfico sobre os campos de atuação da nanopublicação, para que fosse possível identificar em quais domínios do conhecimento este conceito já está sendo utilizado. Esta pesquisa justifica-se pela necessidade de tornar a recuperação da documentação de engenharia mais eficaz, bem como tornar seus dados interoperáveis. São explorados modelos e práticas já consolidados, como os princípios FAIR (Findable, Accessible, Interoperable, Reusable), que promovem o uso estruturado de dados e a curadoria digital, que se aplicados aos documentos técnicos de engenharia, podem garantir a preservação e o reuso eficiente dos dados ao longo do tempo, facilitando a rastreabilidade de documentos em projetos de grande escala. A pesquisa é de natureza aplicada e exploratória, com um levantamento bibliográfico sobre nanopublicação e um estudo de caso sobre documentos de engenharia em uma subestação de obras. Com o estudo de caso e a extração dos conceitos dos documentos de engenharia, foi possível evidenciar que pequenos fragmentos de informações de um documento de engenharia podem se tornar informações granulares e com contexto para que sejam interoperáveis, com possibilidade de reuso e preservação. Assim, foi possível verificar a possibilidade de aplicação da nanopublicação na engenharia, para mitigar problemas recorrentes na gestão documental, promovendo uma maior integração entre dados técnicos, melhorando a rastreabilidade e possibilitando a reutilização de conhecimento técnico, de maneira mais eficiente e confiável. O estudo contribui para a área ao sugerir uma inovação na organização da documentação técnica e abre caminho para futuras pesquisas sobre interoperabilidade e automação, no setor de engenharia.

Palavras-chave: nanopublicação; documentação de engenharia; objetos digitais; curadoria digital; recuperação da informação.

ABSTRACT

CERVEIRA, Núbia Helena Moreira Silva. **Nanopublicação como abordagem para a recuperação da informação em documentos técnicos de engenharia**. 2025. Dissertação (Ciência da Informação) - Instituto Brasileiro de Informação em Ciência e Tecnologia e a Escola de Comunicação da Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2025.

This study investigate the applicability of nanopublications to improve the retrieval, use, and reuse of information in technical engineering documents, with the goal of improving the quality of technical information management within the lifecycle of engineering projects. To assess its applicability, a bibliographic survey was conducted on the fields of application of nanopublications to identify in which domains of knowledge this concept is already being utilized. This research is justified by the need to make the retrieval of engineering documentation more effective, as well as to make its data interoperable. Established models and practices are explored, such as the FAIR principles (Findable, Accessible, Interoperable, Reusable), which promote the structured use of data, and digital curation, which, when applied to engineering, can ensure the preservation and efficient reuse of data over time, facilitating document traceability in large-scale projects. The research is applied and exploratory in nature, comprising a bibliographic review on nanopublications and a case study on engineering documents in a substation construction project. Through the case study and the extraction of concepts from engineering documents, it was possible to demonstrate that small fragments of information from an engineering document can become granular and contextualized data, making them interoperable, reusable, and preservable. Thus, the study verified the potential application of nanopublications in engineering to mitigate recurring issues in document management, promote greater integration of technical data, improve traceability, and enable the efficient and reliable reuse of technical knowledge. The study contributes to the field by suggesting an innovation in the organization of technical documentation and paves the way for future research on interoperability and automation in the engineering sector.

Palavras-chave: nanopublication; engineering documentation; digital object; digital curation; information retrieval.

LISTA DE FIGURAS

Figura 1 - Estrutura Base da dissertação	16
Figura 2 - Mapa Conceitual do Referencial Teórico	25
Figura 3 - Representação genérica de um ciclo de vida do Projeto.....	27
Figura 4 - Metodologia de gerenciamento FEL	29
Figura 5 - Emissão de documentos.....	30
Figura 6 - Tipos de documentos emitidos por fase	33
Figura 7 - Ciclo de vida de um documento digital de engenharia.....	40
Figura 8 - Estrutura de uma Nanopublication	45
Figura 9 - Exemplo de uma Nanopublicação	46
Figura 10 - The DCC Curation Lifecycle Model	53
Figura 11 - Fluxograma PRISMA	62
Figura 12 - Frequência de palavras	64
Figura 13 - Composição do código do documento.....	79
Figura 14 - Tela do SIGEM apresentando a estrutura do projeto.....	80
Figura 15 - Tela do SIGEM com visão dos documentos	81
Figura 16 - Tela dos metadados do documento.....	81
Figura 17 - Planta com detalhes gerais da Casa de Painéis.....	87
Figura 18 - Classes do Estudo de caso Subestação	89
Figura 19 - Relacionamentos na modelagem	90
Figura 20 - Exemplo de arquitetura a ser adotada para documentação técnica.....	94
Figura 21 - Nanopublicações a partir de um documento técnico	95
Figura 22 - Nanopublicação - Caixa de passagem	96
Figura 23 - Fluxograma do processo de emissão de documentos	99
Figura 24 - Partes interessadas	100
Figura 25 - Gerenciamento da documentação técnica.....	106
Gráfico 1 - Áreas do conhecimento identificadas	67

LISTA DE QUADROS

Quadro 1 - Percurso Metodológico.....	61
Quadro 2 - Artigos mais citados.....	64
Quadro 3 - Conceitos de Nanopublicação.....	66
Quadro 4 - Ciência da Computação.....	68
Quadro 5 - Bioinformática.....	70
Quadro 6 - Ciências da Saúde.....	72
Quadro 7 - Ciência da Vida.....	73
Quadro 8 - Ciência da Informação.....	73
Quadro 9 - Humanidades Digitais e Ciências Sociais.....	74
Quadro 10 - Documentos para análise.....	83
Quadro 11 - Categoria de Documentos.....	83
Quadro 12 - Temática Engenharia Civil.....	85
Quadro 13 - Temática Engenharia Elétrica.....	86
Quadro 14 - Temática Arquitetura e Urbanismo.....	87
Quadro 15 - Planta Geral de Elétrica.....	88
Quadro 16 - Documentos selecionados para extração da informação.....	93
Quadro 17 - Extração de informação dos documentos técnicos.....	94
Quadro 18 - Dublin Core – Para descrição dos documentos de engenharia.....	95

LISTA DE ABREVIATURAS E SIGLAS

ABNT	Associação Brasileira de Normas Técnicas
AIDA	Atomic, Independent, Declarative, Absolut
BIM	Building Information Modeling
CAD	Computer Aided Design
CAPES	Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
CONFEA	Conselho Federal de Engenharia e Agronomia
DCC	Digital Curation Centre
DCC SCARP	Sharing, Curation, Reuse and Preservation
DCU	Digital Curation Unit
DWG	Drawing
EVTE	Estudo de Viabilidade Técnica e Econômica
FAIR	Findable, Accessible, Interoperable, Reusable
FEL	Front-End Loading
GED	Gerenciamento Eletrônico de Documentos
GML	Generalized Markup Language
IITKM	Immortal Information and Through-Life Knowledge Management
IPA	Independent Project Analysis
ISO	International Organization for Standardization
KWIC	Key Words in Context
NP	Nanopublicação
OAIS	Open Archival Information Systems
PDCA	Plan, Do, Check, Action
PDF	Portable Document Format
Petrobras	Petróleo Brasileiro S.A.
PLM	Product Life Management
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta- Analyses
RDF	Resource Description Framework,
RI	Recuperação da Informação
SIGEM	Sistema de Gerenciamento de Empreendimentos
URI	Universal Resource Identifier
WOL	Web Ontology Language
XML	Extensible Markup Language

Sumário

1 INTRODUÇÃO	15
2 APRESENTAÇÃO DO TEMA	17
2.1 PROBLEMA DE PESQUISA	20
2.2 JUSTIFICATIVA	22
2.3 OBJETIVOS	23
2.3.1 Objetivo Geral	23
2.3.2 Objetivos Específicos (OE):	24
3 REFERENCIAL TEÓRICO	25
3.1 PROJETOS DE ENGENHARIA	26
3.1.1 Gestão da informação em projetos de engenharia	30
3.2 RECUPERAÇÃO DA INFORMAÇÃO EM PROJETOS DE ENGENHARIA	34
3.3 REPRESENTAÇÃO DO OBJETO DIGITAL	38
3.4 METADADOS E DADOS FAIR: UMA BASE DE SUSTENTAÇÃO	40
3.5 NANOPUBLICAÇÃO	43
3.6 CURADORIA DIGITAL	48
3.6.1 Conceitos	48
3.6.2 Modelos	51
3.6.2.1 DCC Curation Lifecycle Model	53
3.6.3 Curadoria Digital para projetos de Engenharia	54
4 METODOLOGIA	58
5 ESTUDO BIBLIOGRÁFICO SOBRE NANOPUBLICAÇÃO E SUAS APLICAÇÕES	61
5.1 OBJETIVO DA REVISÃO BIBLIOGRÁFICA	61
5.2 CRITÉRIOS DE ELEGIBILIDADE	61
5.3 ANÁLISE DOS RESULTADOS	62
5.3.1 Análise bibliométrica dos artigos selecionados	63
5.3.2 Conceitos	64
5.3.3 Análise crítica dos resultados	65
5.3.3.1 Ciência da Computação	67
5.3.3.2 Bioinformática	69
5.3.3.3 Ciências da Saúde	71
5.3.3.4 Ciências da Vida	72
5.3.3.5 Ciência da Informação	72
5.3.3.6 Humanidades Digitais e Ciências Sociais	73
5.3.4 Resultados	74

6 APLICAÇÃO DO CONCEITO DE NANOPUBLICAÇÃO PARA DOCUMENTOS DE ENGENHARIA	77
6.1 DOCUMENTOS TÉCNICOS DE ENGENHARIA: CARACTERÍSTICAS	78
6.2 PROJETOS NA PETROBRAS E GESTÃO DE DOCUMENTOS TÉCNICOS	80
6.3 ESTUDO DE CASO	81
6.4 MODELO DE NANOPUBLICAÇÃO PARA DOCUMENTO DE ENGENHARIA	86
6.4.1 Ontologia dos Documentos de Engenharia Subestação	88
6.4.2 Especificação da Nanopublicação para Documento de Engenharia.....	90
7 MODELO DE CURADORIA DIGITAL APLICADO AOS DOCUMENTOS DE ENGENHARIA	97
7.1 FLUXO DE GERENCIAMENTO DA DOCUMENTAÇÃO TÉCNICA	98
7.2 CARACTERÍSTICAS DOS MODELOS EXISTENTES	102
7.3 CURADORIA DIGITAL PARA DOCUMENTOS TÉCNICOS DE ENGENHARIA	105
7.3.1 Conceitualização	106
7.3.2 Criação e Recebimento.....	107
7.3.3 Descrição e informação da representação	108
7.3.4 Plano de Preservação.....	109
8 CONCLUSÃO	110
REFERÊNCIAS	112

1 INTRODUÇÃO

A Engenharia é, no seu sentido mais geral, a transformação de uma ideia em realidade, criação e utilização de ferramentas para realizar uma tarefa ou cumprir um propósito (Blockley, 2012).

A documentação técnica de engenharia desempenha um papel fundamental no ciclo de vida de projetos e empreendimentos, fornecendo registros essenciais para planejamento, execução, construção e montagem. No entanto, a complexidade dos documentos e o volume de informações, apresentam desafios para a recuperação da informação, especialmente em momentos que demandam soluções rápidas e precisas.

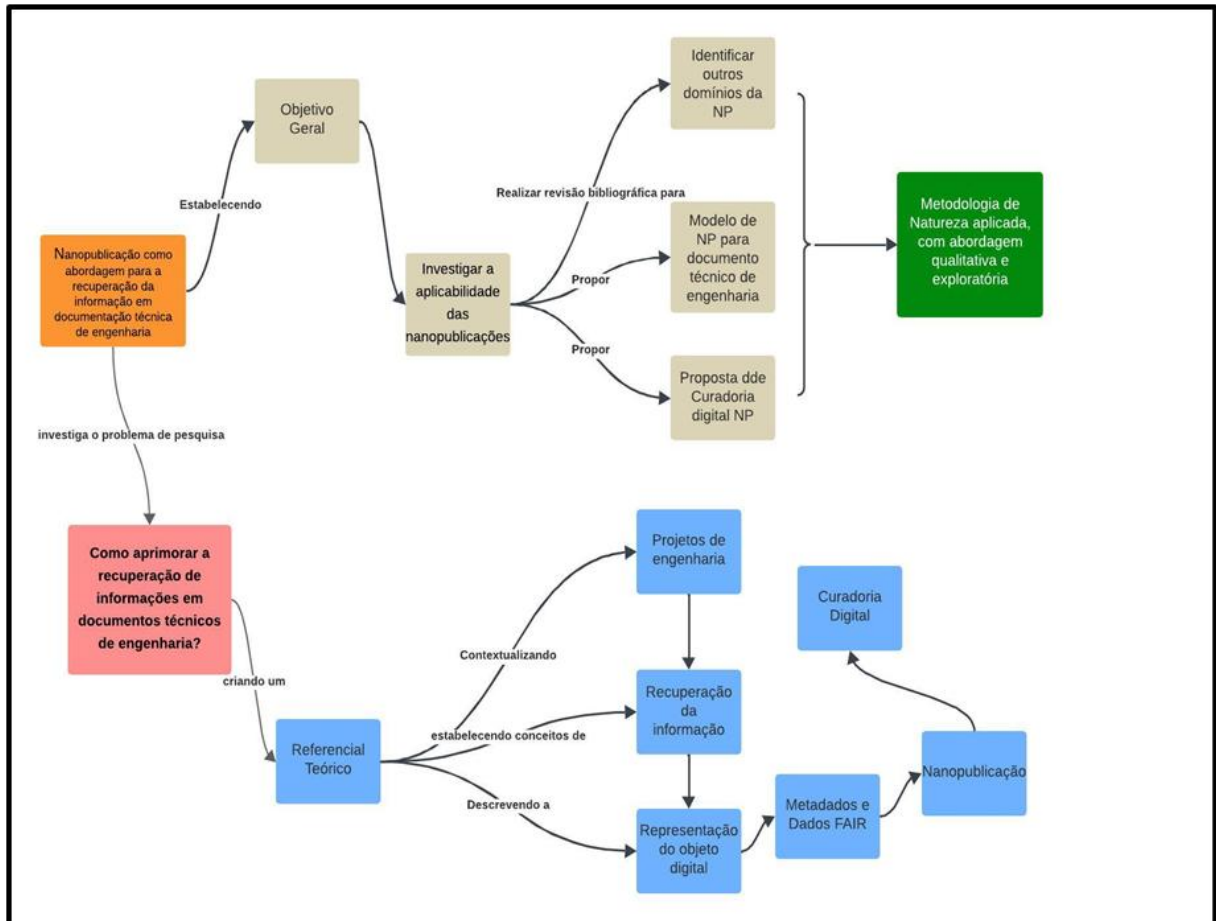
No contexto da engenharia, onde os documentos técnicos desempenham um papel crucial na concepção, execução e manutenção de projetos, a gestão eficiente dessas informações é vital para garantir a precisão, a rastreabilidade e a reutilização ao longo do ciclo de vida de um projeto. Apesar dos avanços na chamada “Engenharia digital”, a recuperação e o reuso de informações técnicas ainda enfrentam desafios significativos, como a dispersão de dados, a ausência de padronização e a dificuldade em assegurar a qualidade e a confiabilidade das informações disponíveis. Nesse cenário, a adoção de modelos estruturados de representação da informação, como a *Nanopublicação*, surge como uma possibilidade promissora para propor um novo método de organização e gestão documental, com vistas à recuperação das informações contidas nos documentos.

Os desafios enfrentados na gestão da documentação técnica, não apenas impactam diretamente na execução dos projetos de engenharia, mas também têm implicações mais amplas para a sociedade, como atrasos em obras de infraestrutura, custos adicionais e a má utilização de recursos humanos e operacionais. Assim, explorar soluções que melhorem a gestão de documentos no setor de engenharia, não é somente uma necessidade técnica, mas também uma necessidade estratégica e econômica.

A proposta desta pesquisa investiga a aplicabilidade das nanopublicações como uma abordagem para aprimorar a recuperação de informações de documentos técnicos de engenharia. As nanopublicações, que são unidades mínimas de conhecimento com metadados ricos e interligados, podem melhorar a representação contextualizada da informação, o uso e o reuso de informações técnicas. Ao pesquisar essa abordagem, buscamos aumentar a qualidade da gestão da documentação técnica, contribuindo para uma melhor integração de dados ao longo do ciclo de vida de projetos de engenharia.

A figura 1 a seguir apresenta a estrutura da dissertação, em forma de resumo, apresentando a pergunta de pesquisa, objetivos geral e específico, bem como o referencial teórico utilizado para subsidiar a pesquisa.

Figura 1 - Estrutura Base da dissertação



Fonte: Autoria própria (2024)¹

¹ Mapa conceitual desenvolvido com o aplicativo Lucidchart.

2 APRESENTAÇÃO DO TEMA

O interesse pelo tema da pesquisa surgiu a partir da experiência profissional desta autora, na empresa Petróleo Brasileiro S.A. (Petrobras), onde atuou na gestão da documentação técnica de engenharia e em práticas de gestão de projetos. Nessa atividade, interagiu com diversas disciplinas de Engenharia acompanhando a emissão de diferentes tipos de documentos durante o ciclo de vida de um projeto, onde percebe-se que, embora o compartilhamento de informações ocorra, ele se limita ao grupo que está desenvolvendo a atividade, restringindo assim, o fluxo de informações e conhecimento entre as equipes.

Na Petrobras, a documentação técnica produzida ou recebida é cadastrada em sistemas de gerenciamento eletrônico de documentos (GED), mas antes de ser cadastrada, ela já percorreu um caminho longo que quase sempre é impossível refazer. Esse caminho passou pela automação de projetos, engenharia digital, modelagem de projetos, elaboração de tabelas, cálculos, ou seja, várias informações e representação de dados que serviram de base para a emissão final de um documento estático, com informações e especificações de um sistema ou equipamento.

Corroborando com essa percepção, Vanja (2005) aponta que dentro da engenharia são utilizadas diferentes tecnologias, e geralmente é difícil ou quase impossível, encontrar os documentos, dados ou informações corretas na hora certa (Vanja, 2005, p.367).

Durante a busca de um documento técnico cadastrado no GED, o usuário tem grande dificuldade e é necessário interpretar inclusive a sua necessidade de informação, para que a busca seja efetiva.

Em estudo realizado por Wild e colaboradores (2010), eles assinalam que os métodos de busca e recuperação de informações fornecem meios para retornar documentos candidatos como fontes potenciais de informações que estão sendo procuradas, no entanto, eles geralmente deixam o usuário com a tarefa de explorar o conteúdo do documento na busca pelas informações em si (Wild et al., 2010, p.47).

A indústria também percebe essa dificuldade, como apontado no relatório “Construction Disconnected”, onde os autores relatam que em entrevista com quase 600 profissionais da construção foi identificado que eles gastam 5,5 horas do seu tempo procurando por dados de projeto, 4,7 horas na resolução de conflitos e 3,9 horas lidando com erros e retrabalho de projetos (Thomas et al., 2018).

O aumento exponencial de dados e informações para a elaboração de projetos de Engenharia e a dificuldade de recuperação de informação de forma rápida e precisa, despertou

o interesse desta autora em pesquisar sobre como eles poderiam ser organizados, tratados e preservados para uma melhor recuperação e reutilização.

Considerando que no desenvolvimento de um projeto de engenharia as disciplinas produzem informações e dados heterogêneos, surgiu o interesse de pesquisar sobre Nanopublicação (NP)² e propor um modelo para a documentação técnica de engenharia, com vistas a sua rápida recuperação e preservação ao longo do ciclo de vida do projeto ou produto.

As Nanopublicações surgem no âmbito das Ciências da Vida, para representar conhecimento científico, e são uma forma, legível por máquina, de publicar unidades mínimas de informação, contendo uma afirmação (conteúdo principal da NP), metadados de proveniência (como a afirmação surgiu) e metadados da NP em si. A sua estrutura semântica poderá desempenhar um papel significativo na curadoria digital, pois facilitará a organização e a rastreabilidade das informações, atendendo as necessidades de recuperação e rastreabilidade da informação contida nos documentos técnicos.

Considerando que no campo da e-Science o “dilúvio” de dados de pesquisa aumenta exponencialmente e que existe uma concordância entre os pesquisadores para que os dados sejam abertos e compartilhados (Faniel; Zimmerman, 2011), seguimos nessa tendência com foco no dilúvio de dados gerados por projetos de engenharia, mas que não devem ser compartilhados livremente, devido às informações que podem ser sensíveis ao negócio que está sendo projetado.

No entanto, a falta de tratamento e organização dessa documentação ou informação técnica com políticas e pesquisas voltadas para sua curadoria, é uma crise digital. Thomas Kuhn (2000, p.158) entende que “o significado das crises consiste exatamente no fato de que indicam que é chegada a ocasião para renovar os instrumentos”.

Sendo assim, o interesse no tema de pesquisa surge da necessidade percebida no dia a dia e no desejo de ampliar esse campo do conhecimento. A pesquisa se propõe investigar a Nanopublicação, com foco no domínio específico da documentação técnica de engenharia, pensando em sua recuperação e preservação a longo prazo.

A trilha acadêmica do Programa de Pós-Graduação em Ciência da Informação do IBICT contribuiu para a consolidação da necessidade da pesquisa, levando em consideração a curiosidade e a abrangência do tema de pesquisa sobre Nanopublicação, em um domínio específico e tão pouco discutido no âmbito da academia.

² Este assunto será mais bem detalhado na Seção 2.

A documentação técnica é um produto emitido nas diversas fases de um projeto de engenharia. Cada projeto tem um objetivo, um propósito com itens finais, entregas e resultados bem definidos, para alcançar um benefício específico. É uma atividade única e temporária (Nicholas; Steyn, 2021, p.3).

Segundo o guia internacional *Project Management Body of Knowledge (PMBOK®)* (Project Management Institute, 2021), um projeto é um esforço temporário com o objetivo de criar um produto, serviço ou resultado específico. É elaborado e gerenciado por fases, chamadas de ciclo de vida, fornecendo a estrutura básica para o gerenciamento do projeto.

O documento técnico de engenharia pode descrever as características que um equipamento deve ter: dimensões, medidas, conceitos e aplicações, e essas informações e dados estão referenciados em diversas outras bases, mas quando a versão final é emitida, ele é um arquivo em, por exemplo, “DWG,” “pdf”, ou seja, um documento digital com representação das informações de projetos necessárias para a execução do empreendimento.

No âmbito da Petrobras, toda essa documentação técnica de Engenharia produzida é cadastrada e suportada por um GED/EDMS, o SIGEM – Sistema de Gerenciamento de Empreendimentos, e por outros sistemas de planejamento e apoio à gestão. Para que as informações sejam identificadas e facilmente recuperadas, se faz necessário uma boa descrição dos metadados nos sistemas.

O projeto de engenharia é composto por vários tipos documentais emitidos a partir de disciplinas técnicas, como Arquitetura e urbanismo, Engenharia Civil, Engenharia Mecânica, Engenharia Elétrica, entre outras, e conforme o projeto, existe uma série de revisões desses documentos até que ele esteja liberado para ser implementado.

No SIGEM, o acesso às informações técnicas ocorre por meio da recuperação dos documentos, com as buscas geralmente realizadas por código do documento (ID), título, palavra-chave ou, ainda, navegando pela estrutura de pastas do projeto ou busca por conteúdo.

Essa busca resulta em uma grande quantidade de documentos, dificultando a rápida identificação da informação necessária. Isso ocorre porque os dados e informações estão dispersos, exigindo uma análise às vezes manual. Esse problema poderia ser mitigado com o uso de Nanopublicações, pois uma unidade mínima de informação necessária e publicável facilitaria o acesso e a rastreabilidade durante a execução de um projeto que precisa de respostas rápidas para colaborar na tomada de decisões.

Em projetos de engenharia, existe uma sequência definida de documentos interdependentes, onde cada um fundamenta a emissão do próximo. Atualmente, os registros

desses documentos são realizados com base em atributos, como tipo de documentos (desenho técnico, folha de dados etc.), título, revisão e outros metadados. Entretanto, essas informações não estão vinculadas diretamente ao processo que pertencem, resultando em dados fragmentados e desassociados, cuja principal referência é um código numérico, título do documento ou assunto.

Com vistas à recuperação e reuso da informação produzida durante o ciclo de vida do projeto de Engenharia, identificou-se a necessidade de um estudo voltado para a Nanopublicação de documentos de Engenharia, para realizar a recuperação e a rastreabilidade da informação desses documentos na forma de curadoria digital.

O pressuposto desta pesquisa é que a arquitetura semântica de Nanopublicação / Nanopublication (NP) – modelo utilizado na comunicação científica, principalmente no domínio das Ciências da Vida, pode ser útil ao contexto da Engenharia, atendendo as necessidades de recuperação e rastreabilidade da informação contida nos documentos técnicos.

A recuperação da informação técnica poderá proporcionar a reutilização de informações em diferentes fases do projeto e ser aplicada a outros, além de minimizar o tempo dedicado a essas buscas pelos profissionais de Engenharia, contribuindo para uma tomada de decisão mais assertiva durante a execução dos projetos.

2.1 PROBLEMA DE PESQUISA

Segundo Lowe *et al.* (1999), as informações do projeto de Engenharia, quando organizadas e gerenciadas, são reconhecidas como vitais para obter vantagem competitiva. A ausência de gerenciamento dessas informações traz um risco ao desenvolvimento do projeto e, conseqüentemente, ao negócio.

Os riscos são inerentes aos projetos de engenharia devido sua natureza, mas são mapeados e administrados durante o planejamento. Em relação à falta de gerenciamento de informações, podem ocorrer riscos técnicos quando um projeto não é executado com o documento correto, riscos de projetos que podem ser relacionados à possibilidade de falha em um projeto ou na aquisição de um suprimento no projeto por uma informação errada ou não revisada em um documento (McMahon; Busby, 2005, p.288).

A falta de gerenciamento das informações, além dos riscos citados, envolve também a perda de tempo dos profissionais na busca de uma informação, como já citado sobre o relatório “Construction disconnected”, que em pesquisa em nível mundial identificou que basicamente os profissionais ficam dois dias na busca de informações.

O processo de desenvolvimento de um projeto é uma sequência de atividades com esforço intelectual e organizacional, pois durante esse desenvolvimento, grande parte do conhecimento é representado na forma de um documento técnico, desenhos, relatórios, entre outros documentos que suportam o projeto.

A engenharia trabalha com diversos tipos de informação, que são dinâmicas e têm relações entre si (Shah; Rogers, 1993). Documentos técnicos de Engenharia são caracterizados por terem informações multivariadas, dinâmicas e complexas, além de serem interdisciplinares em quase todo o ciclo de vida do projeto.

A documentação técnica é a compilação de documentos que especificam uma variedade de informações e instruções sobre um ativo operacional, ela é fonte essencial de informação que descreve o produto e é a base do conhecimento de como um ativo operacional deve ser gerenciado (Koukias; Kiritsis, 2015).

O sistema de Gerenciamento Eletrônico de Documentos/Eletronic Document Management Systems (GED/EDMS) deve ser flexível para reconhecer as especificidades de cada projeto de construção e apoiar a experiência e as preferências individuais e a mudança de requisitos de um determinado projeto (Gabrielaitis; Baušys, 2009).

Um dos fatores que causa a perda de tempo na busca de informações pelos Engenheiros, é não saber em qual sistema determinado documento está mesmo sabendo que ele existe, ou não saber como recuperar esse documento, caso não saiba o ID (código) ou não tenha sido cadastrada a palavra-chave correta.

A maioria dos sistemas de recuperação da informação é desenvolvida para recuperar o documento e com arquitetura desenhada para que a recuperação seja pelo código ou por assunto ao qual o documento corresponde ou ainda por estruturas de pastas, com classificações por assunto.

Para McMahon et al. (2004), cada vez mais, os sistemas computacionais são utilizados para a gestão e disseminação de informações. No entanto, para que sejam efetivos, devem ser construídos considerando a percepção dos engenheiros e como eles usam a informação. Essas informações, para serem recuperadas, precisam ser cadastradas pensando no usuário e contar com dados confiáveis

Percebe-se a necessidade de se estabelecer metadados ricos para que seja possível ter informações acessíveis ao longo do ciclo de vida de um projeto, mantendo a rastreabilidade e possibilitando sua recuperação e reuso. Entretanto, mesmo que a falta de rastreabilidade, ausência de metadados confiáveis e as dificuldades de reuso dessas informações sejam um

problema, é importante considerar a privacidade e a segurança desses dados e informações de documentos técnicos, pois muitos são sensíveis ao negócio, bem como, considerar a dinâmica de um desenvolvimento de projeto, onde nem sempre é possível atualizar as informações no tempo necessário.

Essa necessidade de recuperar informações com maior precisão e rapidez, garantindo a segurança do projeto e a confiabilidade dos dados, leva à seguinte questão de pesquisa: **Como aprimorar a recuperação de informações em documentos técnicos de engenharia?**

2.2 JUSTIFICATIVA

Com o avanço da tecnologia, a engenharia que já gerava uma quantidade considerável de documentos, vem enfrentando desafios frente à recuperação de informações, tendo em vista a natureza do seu conteúdo, sua diversidade de formatos e sua terminologia técnica.

O volume crescente dessa massa de documentos torna cada vez mais a documentação descentralizada ou controlada em pequenos “silos”, isso por conta da falta de confiabilidade que os Engenheiros acabam tendo nos sistemas que armazenam os documentos.

O desenvolvimento de um projeto de engenharia depende em grande parte do conhecimento adquirido entre os engenheiros e das informações disponíveis. Para avançar no projeto e ajudar na tomada de decisão, são necessárias várias buscas de informações e documentos e, em algumas situações, essa busca por informação pode levar a uma perda de 30% do tempo de um profissional, dedicado somente a realizar uma busca para dar continuidade ao projeto, de acordo com Lowe *et al.* (1999).

Alguns estudos identificados, relacionados à Recuperação da Informação em documentos de Engenharia, têm o escopo voltado para a modelagem de sistemas de gerenciamento de documentos e informações (McMahon *et al.*, 2004), melhoria da experiência do usuário (Lowe, 2002), e para a recuperação de informação com modelagem de imagens, classificação facetada de informações (Porto, 2007).

Percebe-se, então, a falta de uma abordagem focada na gestão dos documentos de engenharia que permita a recuperação e a rastreabilidade de informações técnicas específicas contidas nos documentos, visando o reuso dessas informações e o compartilhamento do conhecimento em diferentes fases do projeto, além de sua aplicação em projetos futuros.

Segundo Pahl *et al.* (2007, p.1, tradução própria), “uma das principais tarefas dos Engenheiros é aplicar seu conhecimento à solução de problemas técnicos e otimizar essas soluções”.

Com a complexidade técnica envolvida em um desenvolvimento de projeto, em muitas situações, pequenos fragmentos de informações para dúvidas objetivas podem economizar horas de trabalho e de buscas de informações e documentos.

A proposta deste trabalho é estudar a aplicabilidade do uso de nanopublicações, no âmbito da documentação técnica, para tornar a recuperação da informação mais rápida e assertiva, evitando retrabalho e proporcionando reuso de informações de engenharia durante a execução de novos projetos.

A pesquisa no uso de nanopublicações, para informação técnica, irá ampliar o uso desse conceito, hoje voltado mais para a área das Ciências Biomédicas.

Considerando que os projetos de Engenharia são emitidos com informações técnicas consolidadas, a proposta do uso de nanopublicações, com pequenas “afirmações técnicas” e sua devida proveniência, poderiam ser compartilhados e reutilizados permitindo a recuperação de informações específicas de forma mais eficiente.

Nanopublicações são constituídas por dados estruturados que se interligam com outros dados (Mons; Velterop, 2009). Devido a sua estrutura de Linked data, podem melhorar o compartilhamento e acesso às informações pelos Engenheiros, fazendo com que eles tenham informações mais precisas e atualizadas, elevando assim o nível de qualidade dos projetos de Engenharia, permitindo que a tomada de decisão seja mais assertiva e reduzindo os riscos.

2.3 OBJETIVOS

Para o cumprimento dos objetivos desta pesquisa, pretendeu-se investigar como as Nanopublicações podem ser aplicadas ao domínio da engenharia, mais especificamente aos documentos técnicos de engenharia.

Alinhada com a necessidade de recuperação e reutilização de informações, foi feita uma proposição de um modelo conceitual de nanopublicação e de curadoria digital para os documentos técnicos de engenharia, com vistas a criar um ciclo onde possa ser possível implementar e acompanhar a preservação de dados.

2.3.1 Objetivo geral

Investigar a aplicabilidade das nanopublicações para aprimorar a recuperação, o uso e o reuso da informação em documentos técnicos de engenharia, visando aumentar a qualidade da gestão da informação técnica dentro do ciclo de vida de projetos de engenharia.

2.3.2 Objetivos específicos (OE):

OE1 - Realizar revisão bibliográfica sobre o conceito de nanopublicações e suas aplicações atuais em diferentes domínios;

OE2 - Propor, por meio de estudo de caso, um modelo de nanopublicação para documentos técnicos de engenharia;

OE3 - Propor uma curadoria digital para documentação técnica de Engenharia, que atenda às necessidades de representação, recuperação, acesso, interoperabilidade, reuso e preservação de dados e informações.

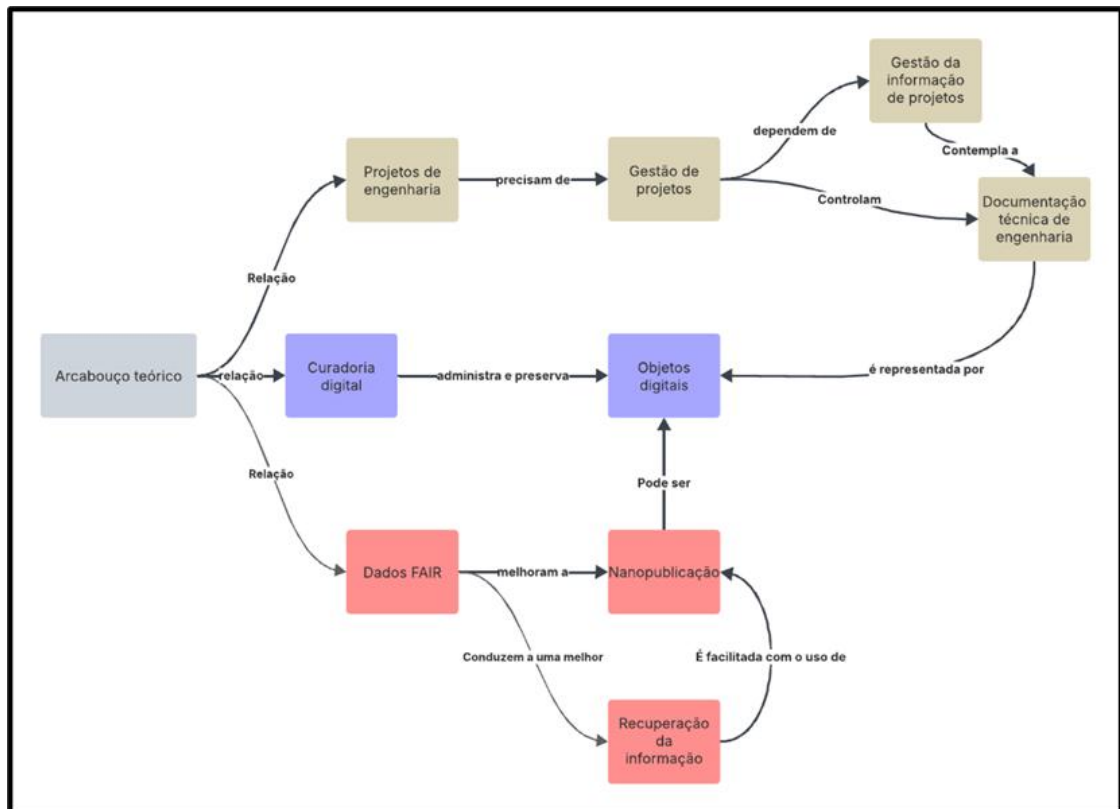
3 REFERENCIAL TEÓRICO

Esta pesquisa investigou a nanopublicação aplicada à documentação técnica de engenharia, analisando seu fluxo, processo de criação e implementação, com o objetivo de aprimorar a recuperação e rastreabilidade das informações técnicas.

O foco da pesquisa nos remete ao clássico conceito de Ciência da Informação (CI), de Borko (1968, p.3), ao dizer que a CI “investiga as forças que governam os fluxos de informação, os significados do processamento da informação, visando acessibilidade e usabilidade”. Pretendeu-se contextualizar projetos de engenharia e as informações geradas por eles, identificando o processamento da informação e sua representação com vistas a sua recuperação e reutilização.

É um estudo transdisciplinar que aborda, a curadoria digital, focando em como a informação contida em projetos de engenharia deve ser gerenciada e preservada. A figura 2 contém o mapa conceitual do referencial teórico.

Figura 2- Mapa Conceitual do Referencial Teórico



Fonte: Autoria própria (2024).

Para fundamentar a pesquisa e estabelecer relacionamentos entre conceitos e teorias, estudamos as características e produtos dos projetos de engenharia, a recuperação da

informação em projetos de engenharia, a Nanopublicação como uma proposta para o uso das informações técnicas, os princípios FAIR (Findable, Accessible, Interoperable, Reusable) e os objetos digitais, e descrevemos a curadoria digital, que não apenas preserva, mas também potencializa a utilização de documentos de engenharia por meio da nanopublicação.

3.1 PROJETOS DE ENGENHARIA

A Engenharia transforma ideias em produtos tangíveis e, para melhor compreender sua complexidade, se faz necessário diferenciar tecnologia e engenharia, pois, embora usados de maneira equivalente, possuem definições e implicações diferentes.

Arthur (2009) define tecnologia em três conceitos distintos: o primeiro é que a tecnologia é um meio para cumprir um propósito humano; o segundo é que é um conjunto de práticas; e o terceiro é um conjunto de práticas e dispositivos de engenharia disponíveis em uma cultura, sendo um significado coletivo, é um conceito usado para discutir como a tecnologia afeta nossas vidas e a sociedade (Arthur, 2009, p.11).

A tecnologia pode ser considerada como produto que a engenharia muitas vezes contribui para criar.

Blockley (2012) descreve a engenharia como a aplicação do conhecimento científico e técnico para projetar, criar, operar e desmontar dispositivos, máquinas, estruturas, sistemas e processos complexos que dão suporte ao esforço humano.

Com esse conhecimento aplicado são otimizadas soluções dentro dos requisitos e restrições estabelecidos e então, engenharia e tecnologia se integram, mesmo sendo distintas, por meio de métodos e técnicas, buscando soluções viáveis e sustentáveis para melhorar o ambiente onde estão inseridas, geralmente com o desenvolvimento de um novo projeto ou produto.

Projetos de engenharia são a materialização dessa aplicação em ações coordenadas, envolvendo planejamento, monitoramento e execução, para atender a um escopo específico.

Para Maximiano (2014, p.5), os projetos são temporários, mas seus resultados são duradouros, “[...] estamos cercados por todos os lados de resultados de projetos, que vêm sendo realizados desde o tempo dos faraós”.

Kerzner (2002) define projeto como uma atividade única, com um objetivo bem definido e com restrições, e que geralmente fornece um produto, serviço ou entrega exclusivo.

Pahl et al. (2007) abordam projeto como um processo sistemático para o desenvolvimento de soluções para problemas técnicos, envolvendo diversas fases, desde a identificação de necessidades e requisitos até a elaboração detalhada da solução.

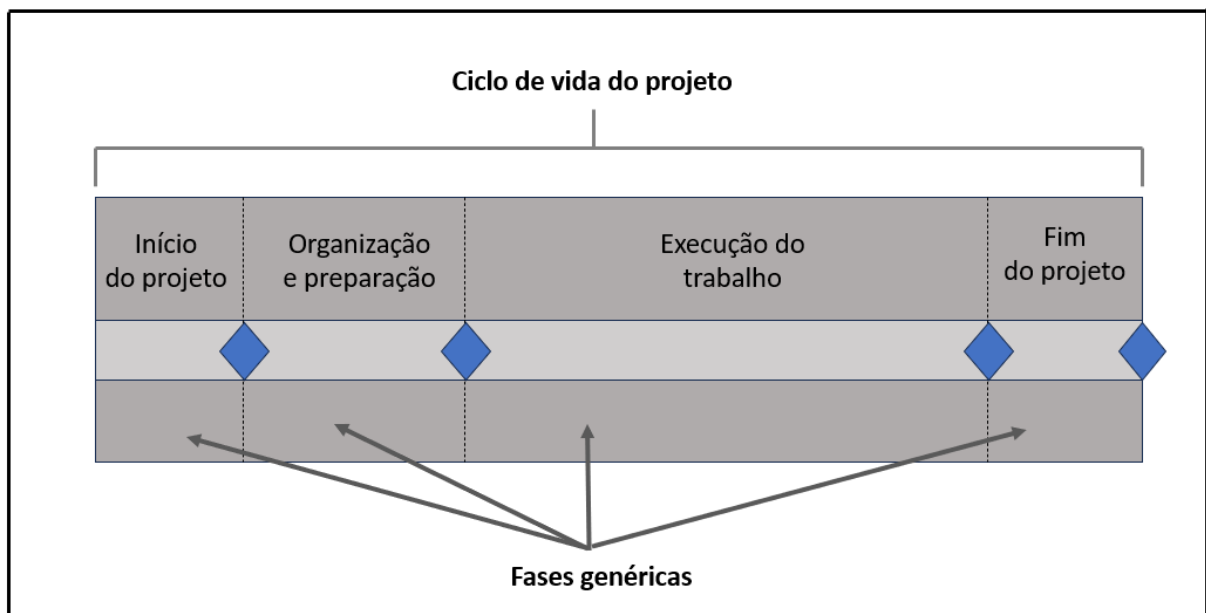
De acordo com o guia internacional *Project Management Body of Knowledge* (PMBOK®) (Project Management Institute, 2021), projeto é um esforço temporário empreendido para criar um produto, serviço ou resultado único.

Para que seja possível realizar controle de todas as fases de um projeto, se faz necessário ter uma gestão com ações coordenadas e acompanhamento por meio de gerenciamento de projetos, cujo mais importante guia é o Project Management Body of Knowledge (PMBOK®) (Project Management Institute, 2021), pois é um guia com terminologia e padronização de escopo mundial com os princípios necessários para gerenciar um projeto.

Projetos são elaborados em fases, que contemplam o chamado “ciclo de vida do projeto”, fornecendo uma estrutura básica para o gerenciamento do projeto, e segundo o PMBOK®, é a série de fases pelas quais um projeto passa, do início ao término. A fase é um conjunto de atividades relacionadas de maneira lógica que culmina na conclusão de uma ou mais entregas. Elas podem ser sequenciais, iterativas ou sobrepostas (Project Management Institute, 2021, p.547).

Na figura 3, temos a representação de um ciclo de vida do projeto com fases genéricas, pois embora todos os projetos tenham início, meio e fim, as entregas variam de acordo com o projeto e o meio onde ele está sendo desenvolvido.

Figura 3- Representação genérica de um ciclo de vida do Projeto



Fonte: Project Management Institute (2021).

Além do PMBOK® utilizado para controlar todo o ciclo de vida do projeto, documentando as entradas e saídas e todo o seu andamento, existem outras metodologias, como

por exemplo a metodologia FEL – *Front End Loading*, implementada pela IPA – Independent Project Analysis. Essa metodologia é utilizada para avaliar projetos, principalmente na indústria de petróleo e gás, e verificar se apresenta viabilidade técnica e econômica para seguir com as demais fases do projeto.

Conceituando FEL, segundo IPA³:

[...] é um processo pelo qual uma empresa traduz suas oportunidades de negócios e tecnologia em projetos de capital. O objetivo do FEL é obter uma compreensão detalhada do projeto para minimizar o número de alterações durante as fases posteriores de execução do projeto. O FEL prossegue até que o projeto “certo” seja selecionado e não termina até que um pacote completo de base de projeto seja concluído. A qualidade do FEL é um indicador chave dos resultados do projeto (Front-End Loading, 2024).

As fases utilizadas na metodologia FEL são:

- a) Fase 1 - Oportunidade do Negócio (FEL 1), onde identifica-se o objetivo do negócio e seu valor em alinhamento com a estratégia corporativa, implica na definição de escopo.
- b) Fase 2 - Engenharia Conceitual (FEL 2), define-se o objetivo do projeto, escolhendo a melhor alternativa para executar essa oportunidade de negócio, são as análises das soluções tecnológicas.
- c) Fase 3 - Engenharia Básica (FEL 3), detalhamento de escopo, custo e prazo, resultando na aprovação da execução do projeto ou não, ocorre a solução refinada de engenharia selecionada na fase 2.

Existe ainda a chamada fase 4, que será a fase de execução e construção e montagem, e a fase 5, que é a operação.

Segundo Romero e Andery, a utilização da metodologia FEL inicia pela aplicação de conceitos de gestão de portfólios e gerenciamento de projetos, e que por si só já é considerada um projeto (Romero; Andery, 2016, p.24).

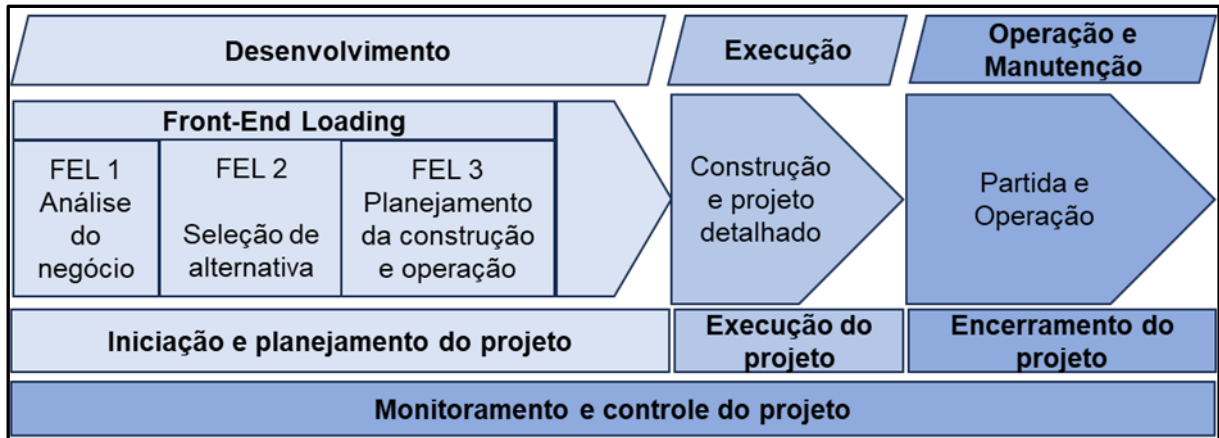
Uma das principais vantagens da FEL é que com as etapas de passagem de fase, as equipes podem verificar riscos e desafios no início do projeto, ao realizar os estudos de viabilidades e projetos preliminares (Serrão, 2023, p.3).

A metodologia é fortemente marcada pela emissão de documentos nas passagens dos portões, fortalecendo assim o empreendimento que será resultado da aprovação das fases.

³ [Glossário IPA | Análise de Projeto Independente \(IPA\)](#)

A figura 4 mostra essas fases e o relacionamento com a aprovação do Estudo de Viabilidade Técnica e Econômica – EVTE, com etapas de projetos gerados a partir de cada fase, representando o uso da metodologia FEL.

Figura 4- Metodologia de gerenciamento FEL



Fonte: Adaptado de Romero e Andery (2016).

Os projetos são conduzidos nas organizações para o desenvolvimento de produtos, implantação de empreendimentos, mudanças em um processo, modernização de plantas industriais, adequação e atendimentos à legislação e a órgãos ambientais, entre outros exemplos que poderiam ser citados.

São complexos e multidisciplinares, sendo assim, a gestão da informação está no cerne do gerenciamento de projetos, pois informações e conhecimentos precisam ser compartilhados e disponibilizados para todos os envolvidos nos projetos, evitando retrabalho e otimizando recursos e soluções.

A documentação dos Projetos de Engenharia é desenvolvida com informações formais estruturadas para atingir seus objetivos, que é a execução do projeto no prazo e com custo estimado.

Para que a documentação se consolide em uma estrutura formal, o fluxo e as informações produzidas ao longo do desenvolvimento do projeto, em diversos sistemas de gestão ou de controle de projeto, precisam ser mapeadas para que se tenha rastreabilidade ao longo do ciclo de vida.

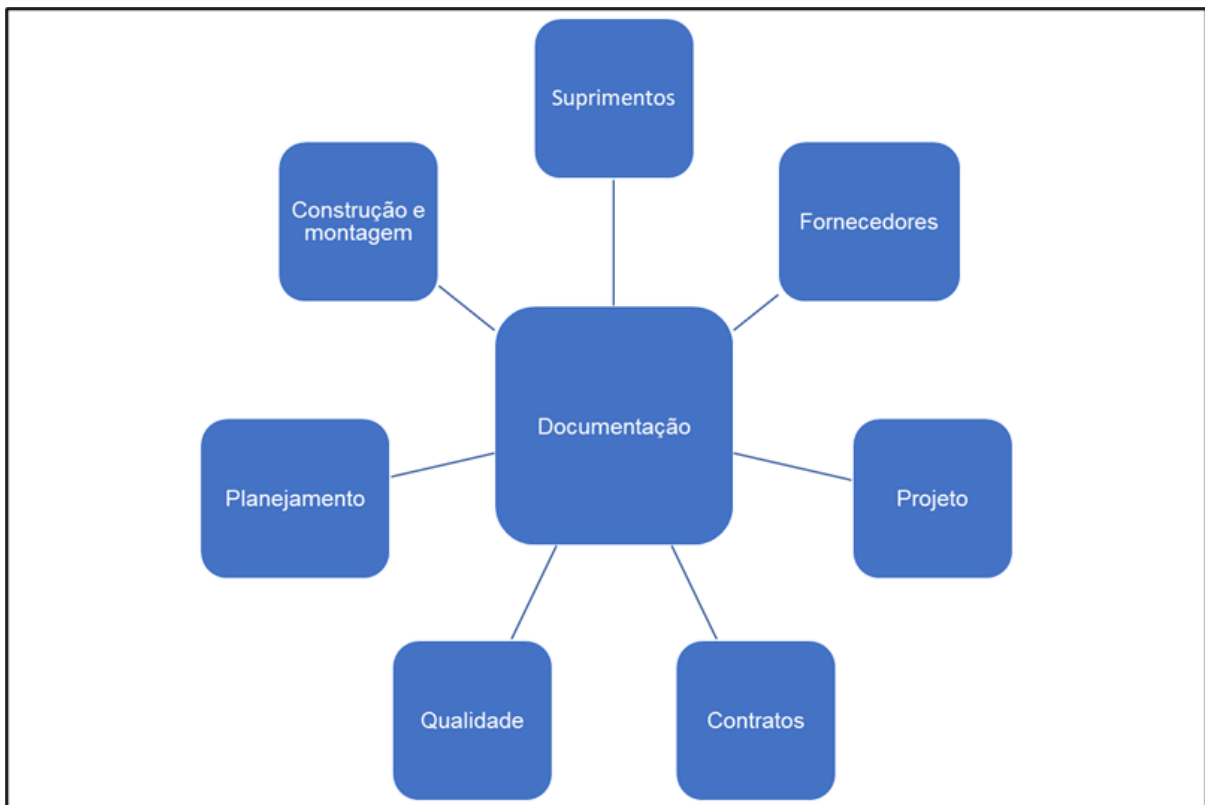
O presente estudo foca nos dados, informações e documentos oriundos das fases dos projetos de engenharia, para contextualizar a necessidade de aprimorar a recuperação da informação.

Em relação a projetos, os documentos emitidos por etapas ou fases são referentes à execução em si ou aplicação da tecnologia no empreendimento. Podem ser gerados por detalhamento de projetos, detalhamento de fabricantes de equipamentos ou oriundos da construção e montagem.

3.1.1 Gestão da informação em projetos de engenharia

Durante a execução de um empreendimento existe uma série de documentos que são produzidos e recebidos, iniciando na documentação para validação da implantação do projeto, passando por documentos de contratação, até a documentação técnica finalizada e pronta para ser entregue ao cliente final, que chamamos de operação da unidade.

Figura 5 – Tipos de documentos emitidos



Fonte: Autoria própria (2024).

Para a execução de um projeto, contemplando todo o seu ciclo, os tipos de documentos abaixo são os mais comuns:

a) Documentação de gestão

A documentação de gestão visa atender o gerenciamento do projeto em si, inclui os planos de gerenciamento, documentos de planejamento, documentos organizacionais,

atas de reunião corporativas, enfim, é uma documentação geralmente produzida pela gestão dos processos que ocorrem nos projetos.

b) Documentação administrativa e contratual

Documentos referentes à execução dos contratos assinados para o empreendimento, são documentos como contratos, anexos contratuais, correspondências contratuais, atas de reunião, medições e notas de pagamentos contratuais, documentos de planejamento e controle físico, administrativo e financeiro. Essa documentação é restrita e sua gestão deverá ter procedimento específico.

c) Documentação da qualidade

Documentos do sistema de gestão da qualidade adotado no empreendimento, visando atender aos requisitos da NBR ISO 90014. Os documentos gerados são: política da qualidade, procedimentos, manuais de qualidade, documentos de planejamento e controle dos registros, modelos, regras e instruções para apoiar a gestão do empreendimento e seus processos.

d) Documentação de engenharia

É um conjunto de documentos originados nos projetos conceitual, básico e executivo, emitidos por projetistas e por fabricantes de equipamentos, cujas informações estão contidas em critérios, requisitos, dados de instalação, plantas de arranjo, fluxogramas, desenhos técnicos em geral e são essenciais para possibilitar a construção, a montagem e a operação de uma planta industrial.

Essa documentação, em conjunto com a documentação de compras, de fornecedores e de construção e montagem, irá formar a documentação técnica do empreendimento.

A documentação técnica é um dos principais ativos das plantas industriais, sendo insumo para os processos de manutenção, operação e ampliação dessas instalações. Durante a execução desses empreendimentos, até o seu funcionamento total, são definidos e registrados todos os dados que pertencem a essa instalação.

Essa documentação é formada por diversos tipos documentais como, memoriais descritivos, especificações técnicas, desenhos, folhas de dados, listas de equipamentos, dentre outros, que registram a informação técnica para construir, montar e operar uma unidade operacional, ou seja, registram uma informação técnica durante o ciclo de vida do projeto e, depois, durante o ciclo de vida do produto.

⁴ É uma norma que estabelece requisitos para a gestão da qualidade de organizações. Ela faz parte da série de normas ISO 9000.

É um período dinâmico onde a geração, atualização e tramitação de documentos é constante. Essa geração e atualização de documentos acontece para cada fase do projeto, cada uma com suas características e funcionalidades.

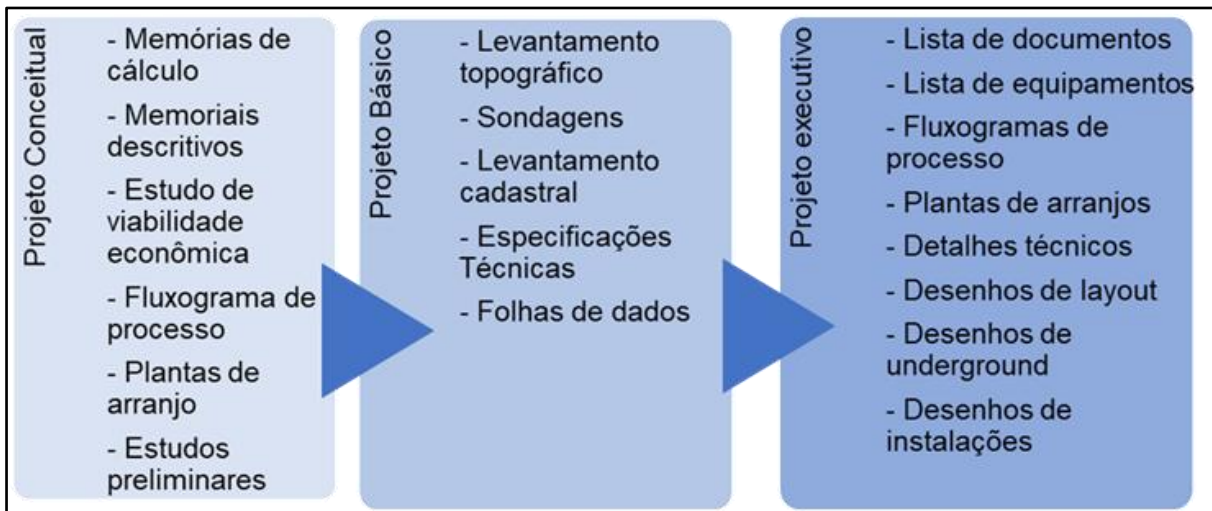
Devido ao grande volume de documentos gerados em todas as fases, o gerenciamento dessa documentação é de fundamental importância para garantir sua rastreabilidade e confiança.

Historicamente, a 1ª e 2ª fases do projeto geram relativamente pouca documentação, na 3ª fase a documentação cresce devido à aquisição de equipamentos e, conseqüentemente, da documentação de fornecedor recebida.

O desenvolvimento que ocorre nessas fases gera a documentação necessária para a execução do empreendimento ou produto, conforme descrição abaixo (Figura 6):

- a) **Projeto conceitual:** É um esboço do projeto, desenvolvido a partir de estudos técnicos preliminares e com o objetivo de determinar a melhor solução técnica, definindo diretrizes e características a serem adotadas na etapa do projeto básico.
- b) **Projeto básico:** A Lei nº 14.133/2021 define o projeto básico como “o conjunto de elementos necessários e suficientes, com nível de precisão adequado para definir e dimensionar a obra ou serviço, ou complexo de obras ou de serviços objeto da licitação, elaborado com base nas indicações dos estudos técnicos preliminares, que assegure a viabilidade técnica e o adequado tratamento do impacto ambiental do empreendimento” (Brasil, 2021, art. 6, XXV);
- c) **Projeto executivo:** Utilizando como referência a Decisão Normativa nº 106 de 2015, do CONFEA, e a Lei nº 8.666 de 1993, podemos definir Projeto Executivo como o conjunto de elementos necessários e suficientes à execução completa da obra ou serviço, de acordo com as normas pertinentes (Brasil, 1993, art. 6, X; Conselho Federal de Engenharia e Agronomia 2015).

Figura 6 - Tipos de documentos emitidos por fase



Fonte: Autoria própria (2024).

Na fase de construção e montagem, 4ª fase, esse crescimento da documentação se deve aos registros de qualidade emitidos durante a execução do empreendimento e a emissão dos documentos finalizados, com status “Conforme construído”.

Cada fase do projeto utiliza a documentação gerada anteriormente para detalhar melhor e produzir novos documentos, é um sistema em cadeia e todos os envolvidos visam o mesmo objetivo, que é entregar o produto para o cliente final.

A necessidade de um fluxo de informação com controle e retroalimentação eficiente se torna cada vez mais imprescindível para o sucesso da gestão de projetos. Aldo Barreto (1994) contextualiza a produção da informação da seguinte maneira:

A produção da informação, definida por nós como estruturas significantes, operacionaliza-se através de práticas bem definidas e se apoia em um processo de transformação orientado por uma racionalidade técnica que lhe é específica; representa atividades relacionadas à reunião, seleção, codificação, redução, classificação e armazenamento de informação. Todas essas atividades orientam-se para a organização e controle de estoques de informação, para uso imediato ou futuro. Este repositório de informação representa um estoque potencial de conhecimento e é imprescindível para que este se realize no âmbito da transferência de informação (Barreto, 1994, p. 26).

Toda informação produzida, transformada em conhecimento e agrupada em um documento precisa ter uma gestão. De acordo com Davenport e Prusak (1997), a gestão da informação é um conjunto estruturado de atividades que espelha a forma pela qual uma organização captura, distribui e usa informação e conhecimento.

Na documentação produzida em ambientes de projeto, especificamente a documentação técnica gerada em empreendimentos como obras de engenharia, essa sinergia entre as partes

deve ser algo constante para que se possa garantir a entrega do produto ao cliente, atendendo os requisitos acordados previamente e a segurança da documentação a ser entregue.

Todo esse processo de emissão e gestão de informações e documentos possui suportes diferentes para registros, no entanto, eles não possuem uma conexão que facilite a recuperação e o reuso, visando conseguir estabelecer uma maior rastreabilidade, facilitando assim a reutilização de informações.

Esta pesquisa apresenta um modelo baseado no conceito de nanopublicações, que possui rastreabilidade e é legível por máquina, facilitando assim a interoperabilidade, além de possuir proveniência dos dados e metadados.

No entanto, para que seja possível emitir parte do documento técnico como nanopublicação, é preciso antes trabalhar os dados. E para fortalecer essa teoria, seguiremos abordando sobre a recuperação da informação e como estabelecer dados FAIR, para ter uma melhor assertividade na recuperação da informação e uma melhor condição de manter a preservação através da curadoria digital de documentos técnicos de engenharia.

3.2 RECUPERAÇÃO DA INFORMAÇÃO EM PROJETOS DE ENGENHARIA

Em 1945, o engenheiro Vannevar Bush, no seu artigo “*As We May Think*”, questionou a explosão informacional gerada pela Segunda Guerra, argumentando que a ciência proporcionou uma comunicação mais rápida entre os homens, fornecendo o registro de suas ideias, permitindo assim que o conhecimento evoluísse e fosse compartilhado, mas, para que fosse útil à ciência, ele deveria ser armazenado e, acima de tudo, consultado. Assim, em sua publicação, ele faz uma descrição simples de sua máquina de recuperação da informação, o MEMEX, “[...] um dispositivo no qual um indivíduo armazena todos os seus livros, registros e comunicações, e que é mecanizado para que possa ser consultado com extrema velocidade e flexibilidade” (Bush, 1945, p.106, tradução própria)⁵.

Cinco anos depois, em 1950, no documento “*The theory of digital handling of non-numerical information and its implication to machine economics*”, Calvin Mooers declara que “Os requisitos da Recuperação da informação, de encontrar informação cuja localização ou a própria existência é desconhecida, exigem agora que seja possível, através de alguma técnica eficiente, especificar uma seleção por meio de algum vocabulário...” (Mooers, 1950, p.5,

⁵ “A memex is a device in which an individual stores all his books, records, and communications, and which is mechanized so that it may be consulted with exceeding speed and flexibility. It is an enlarged intimate supplement to his memory” (Bush, 1945, p.106).

tradução própria)⁶. Na sequência, em 1951, ele definiu a recuperação da informação como sendo o “processo ou método pelo qual um potencial usuário da informação é capaz de converter sua necessidade de informação em uma lista de citações de documentos armazenados contendo informações úteis para ele” (Mooers, 1951, p. 25, tradução própria)⁷.

Hans Peter Luhn, em 1958, citado por Swanson (1988, p.93), desenvolveu um programa automático para a indexação de documentos, e nesse mesmo ano publica o artigo “*Bibliography and index: Literature on information retrieval and machine translation*”, que trazia títulos indexados pelo sistema *Key Words in Context*, ou KWIC. O trabalho dele estimulou o interesse nos problemas e no processo de indexação, no final da década de 1950.

A sequência dos fatos citados acima, como exemplos de preocupações e estudos voltados para a informação e seu resgate, refletem como a informação sempre foi tema para que seu registro, sua indexação e sua recuperação fossem realizadas, garantindo assim o compartilhamento e reuso.

Para a proposta desta pesquisa, foi realizado o levantamento de literatura com foco na recuperação da informação em documentos de Engenharia, buscando registros de estudos aplicados ao domínio de engenharia, projetos e documento técnicos, desenvolvimento de sistemas de recuperação da informação e estudos com arquitetura de websemântica, considerando a ambiguidade das informações de projetos de engenharia.

No âmbito da Engenharia, existem estudos de recuperação da informação que trabalham desde a necessidade de pesquisa do usuário até a modelagem de sistemas para recuperar essa informação (Ahmed-Kristensen; Wallace, 2004; McMahon *et al.*, 2004; Wild *et al.*, 2010).

A Recuperação da Informação (RI) surge da necessidade de atender a explosão informacional do pós-guerra. Manning, Prabhakar e Schutze (2008) definem a RI como a maneira de localizar documentos não estruturados, que satisfaça a necessidade de informação em grandes coleções.

Segundo Baeza-Yates e Ribeiro-Neto (2011), a RI é uma área da Ciência da Computação cujo objeto de estudo é prover fácil acesso às informações para os usuários. Entende-se que no âmbito da Ciência da Computação a RI tem como objeto de estudo facilitar o acesso às

⁶ “The requirements of information retrieval, of finding information whose location or very existence is a priori unknown, now requires that it be possible by some eficiente technique to specify a selection of complex C, by means of any set or combinations of descriptors chosen in any way from the vocabular. There must be complete Independence in the choice and use of descriptors” (Mooers, 1950, p.5).

⁷ “Information retrieval is the name for the process or method whereby a prospective user is able to convert his need for Information into an actual list of citations to documents in storage containing Information useful to him” (Mooers, 1951, p.25).

informações de determinado domínio, proporcionando suporte tecnológico e desenvolvendo sistemas computacionais que possam recuperar informações de maneira eficiente, traduzindo a necessidade do usuário que precisa recuperar a informação.

Na Ciência da Informação, o objeto de estudo é a organização e a representação da informação, e o gerenciamento de informações e documentos, com base nas necessidades do usuário, facilitando a busca e recuperação dessa informação, porém, ambas têm como foco principal a interação com o usuário e a interação com o sistema de busca.

As técnicas de recuperação da informação atuais ou recuperam muitos resultados ou esses resultados são irrelevantes para a busca feita pelos engenheiros, podem inclusive estar em uma forma que os usuários não possam navegar e explorar para localizar, enfim, o que precisam, isso significa que os resultados da pesquisa devem satisfazer os usuários que estão em busca de alguma informação que corresponda à sua compreensão, e uma compreensão inclui, entre outras coisas, as relações entre os termos e a capacidade de desambiguação e inferência, nesse ponto, as técnicas estatísticas de recuperação da informação baseadas em palavra-chave falham com os usuários e frustram seus objetivos (Li; Raskin; Ramani, 2008).

Hahm *et al.* (2014) identificaram alguns estudos voltados para a recuperação da informação no domínio da engenharia e diferenciaram esses estudos em dois tipos: navegação e pesquisa. Na navegação, o usuário olha para a estrutura de informações e se move pelas visualizações das informações disponíveis. Na pesquisa, o usuário deve preencher uma consulta com base nas suas necessidades de informação. De acordo com os tipos de documentos de recuperação, o método de busca pode ser dividido em recuperação de documentos estruturados e recuperação de documentos não-estruturados.

O levantamento realizado por eles identificou que os estudos anteriores focaram na recuperação de documentos estruturados ou abordagens de navegação, e para documentos não estruturados, os autores fizeram uma proposta de pesquisa semântica geral, que chamaram de “uma expansão de consulta”, que seria necessária para superar a ambiguidade da linguagem natural.

Para a recuperação da informação em documentos de engenharia, existe uma limitação referente à similaridade, pois alguns sistemas de busca se baseiam na correspondência exata das palavras, e palavras sozinhas não conseguem capturar a semântica ou significado do documento e da consulta, devido a sua ineficácia na compreensão, tendo em vista que documentos técnicos descrevem vários processos, com terminologia técnica e abreviações e, como já citado, são documentos com grande nível de ambiguidade

A recuperação da Informação em documentos de engenharia enfrenta desafios devido às variações na sintaxe e na semântica.

A documentação técnica de engenharia utiliza terminologia e jargões técnicos, abreviaturas, siglas e sinônimos, o que dificulta a recuperação da informação devido à ambiguidade, e para que esta seja tratada nas informações técnicas, alguns estudos sobre recuperação da informação em projetos de engenharia são dedicados ao uso de tecnologias semânticas.

A seguir um breve resumo de alguns desses estudos:

- 1) Com o propósito de promover a reutilização de fontes existentes e hierarquicamente estruturadas, o estudo apresenta um esboço sobre o desenvolvimento de taxonomias baseado em diretrizes de desenvolvimento de ontologias (Lowe, 2002);
- 2) Recuperação de fragmentos de documentos de engenharia baseada na decomposição de documentos, com marcação desses fragmentos utilizando a linguagem semântica XML⁸. A marcação de documentos de engenharia com XML proporciona aos usuários uma maior facilidade de transferir, armazenar e publicar dados dos documentos (Liu et al., 2006);
- 3) Recuperação de documentos de engenharia não estruturados com descrições textuais ou textos associados ao CAD⁹, usando uma abordagem baseada em ontologia. Os pesquisadores desenvolveram uma ontologia em PROTÉGÉ¹⁰ categorizando os produtos e componentes dos dispositivos projetados, com isso foi possível construir uma taxonomia com os temas categorizados e criar uma lista de termos léxicos usados para combinar palavras em documentos ou consultas (Li; Raskin; Ramani, 2008).

A documentação de engenharia é composta, em sua maioria, por descrições textuais e desenhos, mas com a ambiguidade inerente da disciplina Engenharia, os profissionais encontram dificuldades de localizar as informações que necessitam.

A recuperação da informação para documentos de engenharia, em grande parte dos sistemas, é realizada por palavra-chave, isso porque os documentos técnicos são cadastrados

⁸ Extensible Markup Language (XML) é um formato de texto simples e flexível. É um padrão do W3C Consortium. Disponível em: <https://www.w3.org/XML/>. Acesso em: 6 jun. 2024.

⁹ Abreviação de Computer Aided Design, que em português significa "Projeto e Desenho Assistidos por Computador".

¹⁰ Plataforma gratuita de código aberto que fornece ferramentas para a construção de modelos de domínio e aplicações em ontologia. Disponível em: <https://protege.stanford.edu/>. Acesso em: 6 jun. 2024.

nos sistemas segundo seus metadados e de acordo com o descritivo do documento. Sendo assim, este estudo também pretendeu apresentar uma relação entre metadados e a Recuperação da informação, ou seja, a importância de ter metadados confiáveis para que a recuperação da informação ocorra satisfatoriamente.

3.3 REPRESENTAÇÃO DO OBJETO DIGITAL

Para atender o objetivo desta pesquisa, que é apresentar um modelo conceitual de curadoria digital para documentos de engenharia, é necessário contextualizar o objeto digital enquanto representação desse documento e conciliar a sua representação digital com os dados e informações de um projeto.

Thibodeau (2002) abordou a necessidade de compreender o que são objetos digitais, considerando a definição genérica de que “um objeto digital é um objeto de informação, de qualquer tipo de informação ou formato, expresso em formato digital”. Complementa que se examinarmos o que significa a informação expressa em formato digital, iremos reconhecer a característica básica dos objetos digitais, que é a herança de três classes distintas, esclarecendo que:

Objetos digitais são entidades com múltiplas inerências; ou seja, as propriedades de qualquer objeto digital são herdadas de três classes. Todo objeto digital é um objeto físico, um objeto lógico e um objeto conceitual, e suas propriedades em cada um desses níveis podem ser significativamente diferentes. Um objeto físico é simplesmente uma inscrição de sinais em algum meio físico. Um objeto lógico é um objeto que é reconhecido e processado pelo software. O objeto conceitual é o objeto tal como é reconhecido e compreendido por uma pessoa ou, em alguns casos, reconhecido e processado por uma aplicação informática capaz de executar transações (Thibodeau, 2002, p.6, tradução própria),

Hui (2014) argumenta que um objeto digital é uma entidade complexa, com várias camadas de significado e que exige uma análise além de sua estrutura técnica: “Os objetos digitais aparecem em três fases, que são interdependentes entre si, mas não podem ser reduzidas ou generalizadas em unidade: objetos, dados e redes” (Hui, 2014, p.63).

Essa complexidade citada pelo autor deve ser considerada na curadoria digital e, se aplicada essa teoria aos documentos de engenharia, entendemos que eles são documentos estáticos quando emitidos em formatos de “.pdf”, “.doc” por exemplo, mas também são entidades dinâmicas interligadas a outros objetos e sistemas.

Documentos de engenharia como “objetos” possuem propriedades que podem ser manipuladas, mas o valor desse documento está na relação que ele possui com os demais documentos. Para estabelecer uma curadoria digital para essa documentação, esta não deve

focar somente na preservação do arquivo, mas nas relações que cada documento possui com os outros.

Esses documentos possuem “dados e metadados”, sendo que os dados representam as informações de engenharia, enquanto os metadados nos apresentam o contexto que precisamos para entender essas informações, como autor, data de criação, versão etc. Essa atualização e integridade dos metadados é fundamental para o projeto durante todo o seu ciclo de vida.

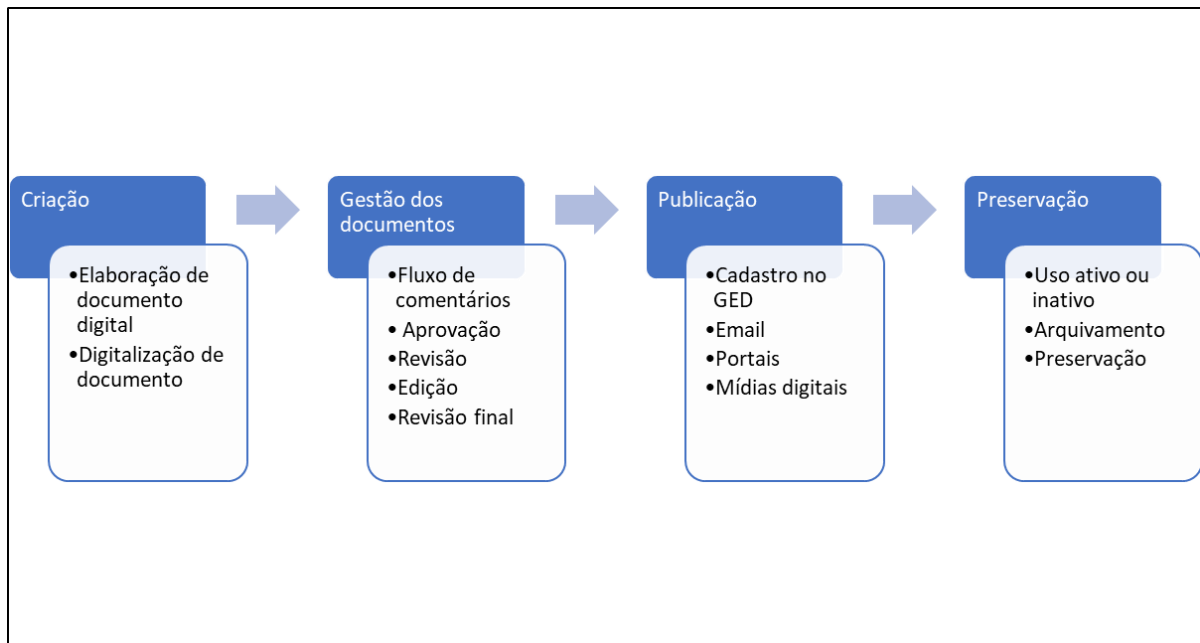
Ainda sobre a teoria de Hui, as “redes” se aplicam ao ambiente colaborativo necessário aos documentos de engenharia, elas fazem parte de uma estrutura onde estão conectados pessoas, projetos, sistemas.

Esse documento digital de engenharia e sua representação através de dados e metadados tem uma trajetória até a sua disponibilidade e uso, enquanto objeto digital emitido e cadastrado em repositórios digitais ou em gerenciador eletrônico de documentos – GED. É composto por um conjunto de tecnologias e é utilizado como repositório centralizado para organizar e gerenciar a documentação digital, permitindo, no caso de documentos de engenharia, que se tenha controle de revisões e versões, e uma melhor rastreabilidade, além da segurança dos registros e da garantia da integridade da documentação.

Relaciona-se diretamente com os objetos digitais, uma vez que os documentos geridos são, na verdade, objetos digitais compostos por dados e metadados, que devem ser bem descritos, administrados e preservados para manter sua integridade e rastreabilidade.

Na figura 7 está representada, de forma simples, uma trajetória desse objeto digital e seu ciclo de vida no projeto, independentemente de sua fase, considerando o gerenciamento no GED.

Figura 7- Ciclo de vida de um documento digital de engenharia



Fonte: Adaptado de Ged System (UNISOCIESC, 2015).

A proposição de representar esse objeto digital como uma nanopublicação surge da necessidade de preservar as conexões e interdependências entre os objetos promovendo um fluxo de trabalho integrado por meio da curadoria digital e, assim, facilitar a recuperação e o reuso dessas informações de projetos.

3.4 METADADOS E DADOS FAIR: UMA BASE DE SUSTENTAÇÃO

Sayão (2010, p. 1), em seu artigo “Uma outra face dos metadados”, argumenta que o “conceito tradicional de metadado pode ser ampliado para abrigar um conjunto de informações que apoiem as atividades de gestão da preservação de materiais digitais”. De acordo com Riley (2017, p. 1, tradução própria), “Metadados, as informações que criamos, armazenamos e compartilhamos para descrever as coisas, nos permitem interagir com essas coisas para obter o conhecimento de que precisamos”¹¹.

Richard Gartner (2008) contextualiza no relatório “*Metadata for digital libraries: state of the art and future directions*”:

Os metadados são o núcleo de qualquer sistema de recuperação de informações e, portanto, suas implicações para qualquer biblioteca digital são profundas: a escolha de um esquema de metadados sustenta a capacidade de qualquer biblioteca de entregar

¹¹ Metadata, the information we create, store, and share to describe things, allows us to interact with these things to obtain the knowledge we need. (Riley, 2017, p.1)

objetos de maneira significativa e afeta muito sua capacidade de longo prazo de manter e preservar seus ativos digitais (Gartner, 2008, p. 5, tradução própria)¹².

Richard Gartner (2008) citou o “núcleo de qualquer sistema de recuperação da informação”, que podemos ampliar para os sistemas de Gerenciamento Eletrônico de Documentos – (GED) ou os sistemas de gestão de documentos de engenharia – EDMS, que comportam e armazenam os documentos de engenharia cujos metadados que fazem a representação formal do documento podem ser por exemplo: título, código, disciplina (especialidade da engenharia), dependendo da situação do documento, a sua ação, ou seja, para qual finalidade aquele documento foi emitido.

A diversidade dos locais de armazenamento durante a execução e elaboração de documentos de engenharia torna a recuperação das informações pelos engenheiros, exaustiva, envolvendo buscas em sistemas e interações informais com os demais engenheiros devido à complexidade das pesquisas (Auriscchio; Bracewell; Wallace, 2010; Hertzum; Pejtersen, 2000).

Propor uma forma de trabalhar a recuperação estruturada de documentos e informações é uma solução para as limitações dos sistemas tradicionais de recuperação da informação, na gestão de documentos de engenharia, e a utilização de metadados ricos na recuperação de informação melhora significativamente a eficiência e a precisão dos resultados de busca.

David Haynes (2018) descreve o seguinte sobre metadados:

Os metadados são usados em catálogos de recursos de informação digitais e físicos. Os requisitos para livros em um catálogo de biblioteca podem ser muito diferentes dos metadados incorporados em uma página da Web, mas os conceitos gerais de metadados se aplicam a ambos. Seu uso para recursos digitais e impressos fornece alguns exemplos úteis. Linguagens de marcação de documentos, como SGML e XML, são amplamente utilizadas para expressar padrões de metadados (Haynes, 2018, p.19, tradução própria)¹³.

Os modelos de metadados são uma tecnologia semântica de representação do conhecimento, enquanto os metadados descrevem os dados, por isso é necessário que exista uma infraestrutura para armazenar e preservar os dados tornando-os semanticamente interoperáveis, ou seja, um dado que atenda aos princípios FAIR (Horsch *et al.*, 2021).

¹² Metadata is the core of any information retrieval system and so its implications for any digital library are profound: the choice of a metadata scheme underpins any such library's ability to deliver objects in a meaningful way and greatly affects its long-term ability to maintain and preserve its digital assets. (Gartner, 2008, p.5)

¹³ Metadata is used in catalogues of digital and physical information resources. The requirements for books in a library catalogue might be very different from the metadata embedded in a web page, but the general concepts of metadata apply to both. Its use for digital and printed resources provides some helpful examples. Document mark-up languages such as SGML and XML are widely used to express metadata standards (Haynes, 2018, p.19).

Nessa abordagem, os metadados têm grande importância na organização e recuperação da informação. Percebemos assim que os elementos que descrevem os dados e metadados de uma informação ou documento técnico são essenciais na construção de um ambiente de pesquisa de documentos que tenha confiabilidade.

Os metadados fornecem informações essenciais sobre os dados, e no caso da documentação técnica de engenharia, os metadados descritivos detalham sua origem, conteúdo, aplicação, em alguns casos proveniência, facilitando assim a identificação e localização das informações relevantes para o negócio.

Em paralelo às práticas de padronização, visando uma recuperação da informação consistente e segura e a interoperabilidade dos dados, é possível seguir em direção aos princípios FAIR dos dados de projetos.

Wilkinson *et al.* (2016) apresentam esses princípios como diretrizes para melhorar a infraestrutura dos dados científicos, mas, atualmente, esses princípios são fortemente aplicados em diversos tipos de dados.

Os princípios FAIR, em resumo, são:

Findability (*Recuperabilidade*): os dados devem ser fáceis de encontrar tanto por humanos como por máquinas. Isso é alcançado usando identificadores únicos e persistentes e metadados ricos.

Accessibility (*Acessibilidade*): os dados devem ser acessíveis através de protocolos padronizados. Os metadados devem permanecer acessíveis mesmo se os dados não estiverem mais disponíveis.

Interoperability (*Interoperabilidade*): os dados devem ser compatíveis e utilizáveis em diferentes plataformas utilizando uma linguagem formal e compartilhada para representação do conhecimento.

Reusability (*Reusabilidade*): os dados devem ser reutilizáveis e manter sua relevância ao longo do tempo. Isso é possível com a descrição detalhada dos metadados e aderência aos padrões.

Os princípios FAIR têm como objetivo não só tornar os dados acessíveis e reutilizáveis, mas também permitir e garantir que eles sejam encontráveis e interoperáveis por meio de métodos ou sistemas de recuperação da informação eficientes e eficazes.

Documentos técnicos de Engenharia são considerados, em sua maioria, como documentos não estruturados, devido ao uso de abreviaturas, termos técnicos, simbologias, gráficos e tabelas, sendo assim, entende-se que a descrição dos metadados dessa documentação

é essencial para a sua recuperação a longo prazo, portanto, ao reconhecer a importância e a necessidade de metadados na organização e descrição dos dados para posterior recuperação e reuso, estamos colaborando para um ambiente de compartilhamento e pesquisa de documentos de engenharia transparente, colaborativo e eficaz.

A emissão do documento técnico é realizada para atender uma necessidade de conformidade do projeto, a sua ação caracteriza a fase do projeto até a sua conclusão final, podemos dizer que um documento técnico é vivo enquanto está na fase de execução de um projeto, e mesmo quando concluído ele continua com sua “ação” para uso na operação de uma unidade ou para atestar a finalização de um projeto.

A elaboração de projetos com padrões de metadados seguindo os princípios FAIR, nos possibilitará manter uma rastreabilidade e reuso para outros projetos diminuindo o tempo de criação de novos objetos digitais. Uma possibilidade de se manter esses dados interoperáveis é através da criação de pequenas notas técnicas ou “nano informações técnicas” com o conceito de Nanopublicações para as informações dos documentos técnicos de engenharia.

3.5 NANOPUBLICAÇÃO

No mundo da publicação científica, o conceito de *nanopublicação* ampliou a importância de sua aplicação e uso nos últimos anos, devido ao crescimento da geração dos dados de pesquisas, no chamado “dilúvio de dados”, onde se tornou cada vez mais difícil administrar e tratar esses dados.

As nanopublicações, doravante serão chamadas de NP, são unidades mínimas de uma informação científica, citável, rastreável, compartilhável e reutilizável, sendo altamente eficiente para a troca do conhecimento científico. Poderíamos dizer, ainda, que uma NP é uma unidade mínima de informação científica com contexto associado, baseado em *linked data* para publicar os resultados de pesquisas e seus metadados em pequenas unidades científicas (Groth; Gibson; Velterop, 2010; Mons; Velterop, 2009).

NP foi desenvolvida para facilitar a localização, conexão e curadoria de declarações científicas e para determinar sua atribuição, qualidade e procedência, elas são pequenos trechos com arquitetura em RDF – Resource Description Framework, com grafos nomeados (Kuhn; Krauthammer, 2012).

Em termos conceituais, observa-se que a NP é fundamentada em tecnologias semânticas apresentando alto nível de granularidade de trechos de informações atômicas e que fornece uma declaração formal (asserção), a procedência da declaração, ou seja, com informações sobre de

onde vem essa declaração (proveniência) e os metadados sobre a NP (procedência), é legível por máquina e é representada por triplas RDF (Sujeito-predicado-objeto) (Groth; Gibson; Velterop, 2010; Kuhn *et al.*, 2018; Mons; Velterop, 2009).

O conceito do modelo de NP surgiu com Barend Mons e Jan Velterop, em 2009, em uma apresentação no “*Workshop on Semantic Web Applications in Scientific Discourse*”¹⁴, sendo publicado nos Anais do evento com o título “*Nano-publication in the E-science era*” (Mons; Velterop, 2009).

Com a produção crescente dos dados, tornava-se urgente trabalhar com dados e informações estruturadas, sendo assim, publicações em formatos clássicos já não eram suficientes para disseminar conhecimento, pois na maioria das vezes o que o pesquisador mais necessita durante uma busca são pequenas unidades de informação.

Barend Mons e Jan Velterop (2009) citam cinco passos para implementar o conceito de NP, são eles: **1º passo**: Dos termos a conceitos; **2º passo**: Dos conceitos para declarações; **3º passo**: Anotações de declarações com contexto e procedência; **4º passo**: Tratando declarações ricamente anotadas como NP; e **5º passo**: Removendo redundâncias, meta-analisando declarações da web (triplas brutas para triplas refinadas).

Os passos apresentados por Mons e Velterop (2009) constituem o conceito para uma NP, considerando-se imprescindível atender a identificação dos conceitos e declarações de forma única, bem como a referência aos conceitos e declarações que foram identificados, permitindo, assim, a interoperabilidade e o compartilhamento.

A partir da publicação do artigo intitulado “The anatomy of a Nano-publication”, de Paul Groth, Andrew Gibson e Johannes Velterop, em 2010, surge uma compreensão mais consolidada sobre a estruturação e definição técnica do modelo de uma NP. Ela deverá ter a capacidade de apresentar uma declaração científica com contexto, para que ela seja por si só considerada uma publicação.

O modelo principal apresentado no artigo atende requisitos de práticas de publicações existentes e a necessidade de agregar informações de fontes distribuídas, sendo assim,

Semelhantes às publicações científicas-padrão, as nanopublicações precisam ser confiáveis, atribuíveis e revisáveis. Além disso, elas precisam ser facilmente curadas. As nanopublicações devem ser facilmente agregadas e identificadas em toda *web*. Finalmente, precisam ser extensíveis para atender a novas formas de metadados e descrição (Groth; Gibson; Velterop, 2010, p. 1, tradução própria)¹⁵.

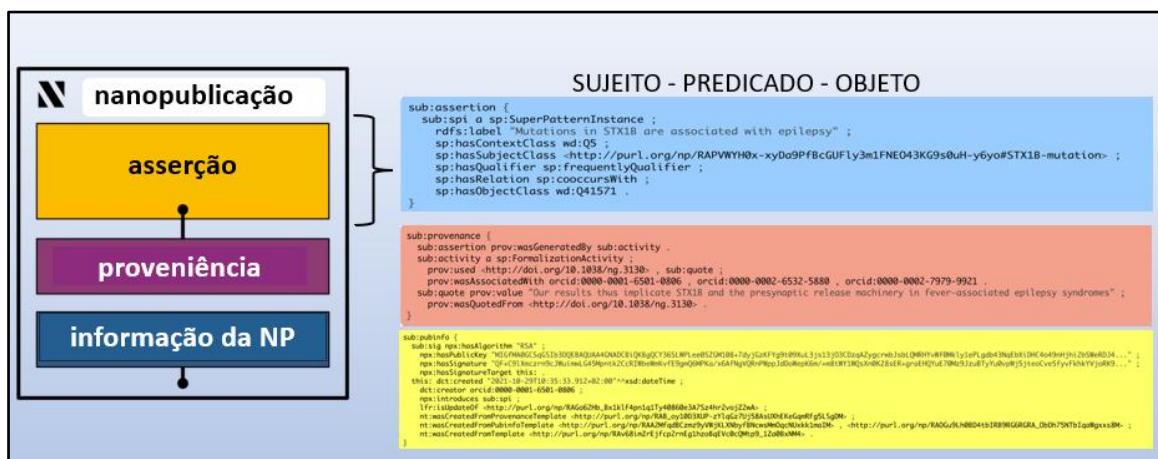
¹⁴ Disponível em: <https://www.w3.org/wiki/HCLS/ISWC2009/Workshop>. Acesso em: 3 jun. 2024.

¹⁵ “Similar to standard scientific publications, nano-publications need to be citable, attributable, and reviewable. Furthermore, they need to be easily curated. Nano-publications must be easily aggregated and identified across the Web. Finally, they need to be extensible to cater for new forms of both metadata and description”.

Em termos conceituais, observa-se que a NP é um conceito e uma tecnologia baseada em *Linked Data* para publicar resultados científicos e seus metadados em pequenas unidades de publicação, representadas em RDF, portanto, são interpretáveis por máquina (Mons; Velterop, 2009), com três elementos básicos (nanopub.net, 2022, p. 1), conforme figura 8.

1. **Asserção** (Assertion): é o conteúdo principal da Nanopublicação, como por exemplo, o dado científico pesquisado. É a afirmação que está sendo feita na Nanopublicação ou a declaração simples do resultado científico sob a forma de uma pequena unidade atômica de informação. O grafo da asserção contém as afirmações que constituem a reivindicação principal das nanopublicações;
2. **Proveniência** (Provenance): descreve como se chegou à afirmação referida, é a procedência dessa declaração, como por exemplo, qual estudo foi conduzido para derivar nessa asserção ou de quais documentos ela foi extraída ou quais métodos científicos foram utilizados para gerar a afirmação, asserção. O grafo da proveniência contém uma ou mais triplas RDF que contêm as informações sobre como surgiu essa asserção.
3. **Informação da publicação** (Publication info): contém metadados sobre a Nanopublicação como um todo, tais como quando e por quem foi criada e os termos da licença para a sua reutilização. Este grafo contém também uma ou mais triplas RDF com essas informações sobre a própria Nanopublicação.

Figura 8 - Estrutura de uma Nanopublication



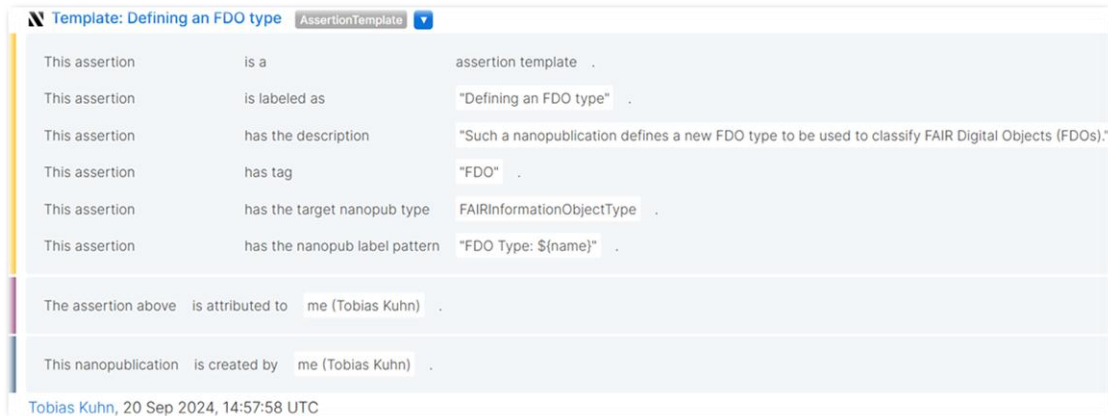
Fonte: Adaptado de Nanopub.net (2022).

A NP é implementada e alinhada com tecnologias da *web* semântica, como RDF, OWL e SPARQL. Elas surgiram como uma abordagem inovadora para lidar com o excesso de informações, inicialmente com foco nas ciências da vida, sendo mais adequadas para a apresentação das relações entre dados de pesquisa e troca de conhecimento do que os artigos tradicionais.

A figura 9, abaixo, é um exemplo de uma Nanopublicação extraída da interface *Nanodash*, que é uma plataforma para publicação de nanopublicações com diversos templates

disponíveis, criada para usar as nanopublicações e uma rede descentralizada de serviços, permitindo, assim, que os usuários não especialistas publiquem diretamente pequenas instruções de dados vinculados por meio da interface antes chamada *nanobench* (Kuhn *et al.*, 2021).

Figura 9 - Exemplo de uma Nanopublicação



Fonte: Nanodash [2024]¹⁶

No exemplo da figura 9, a NP está representada da seguinte maneira:

- Asserção: “Defining na FDO type”
- Proveniência: Tobias Kuhn
- Info da NP: autor: Tobias Kuhn

Segundo Kuhn *et al.* (2013), as NPs podem ser usadas também para outros tipos de afirmações, e poderiam se tornar a base para toda *web* semântica, compartilhando qualquer informação que se deseje na forma de uma ou mais NPs, podendo incluir afirmações científicas e dados experimentais, opiniões, relações sociais, eventos e metadecarações sobre outras NPs.

A proposta dos autores foi expandir o conceito de NP utilizando sentenças em inglês para representar as afirmações científicas informais e subespecificadas, que fornecem uma representação uniforme e sucinta das afirmações científicas.

Tobias Kunh e Dumontier (2015), a fim de tornar as NPs verificáveis e imutáveis, sugerem que URI – Universal Resource Identifier confiáveis podem ser usados como identificadores que incluem *hash* criptografados calculados sobre o conteúdo da NP, tornando assim não apenas os recursos específicos, mas toda a árvore de referência verificável, deste

¹⁶ <https://nanodash.knowledgepixels.com/>

modo, as NPs podem citar outras através de suas URIs, criando redes complexas de citação (Kuhn; Dumontier, 2014; 2015).

Em termos práticos, podemos argumentar que uma NP carrega consigo os princípios da atomicidade, interoperabilidade, procedência, granularidade e *Linked data* (Kuhn *et al.*, 2018), e que se assemelham aos princípios de dados FAIR, que são uma diretriz geral para melhorar as práticas de gerenciamento e curadoria de dados e metadados com foco em atividades orientadas por homens e máquinas (Wilkinson *et al.*, 2016).

As NPs possuem as características necessárias para o compartilhamento do conhecimento técnico, seguindo a ampliação do uso além das Ciências da Vida, ampliando o conceito para uso em informações técnicas, especificamente em informações técnicas de documentos de engenharia. No entanto, teríamos dificuldades quanto ao conceito de acesso livre, considerando que uma informação técnica gerada em/ou para um ambiente industrial tem um certo grau de sigilo, que, dependendo da proporção de sua divulgação, pode causar impactos ao negócio.

Por outro lado, considerando que no ambiente de Engenharia as informações circulam em pequenos grupos de conhecimento, onde essas redes informais entre os engenheiros pode ser a força que apoia muitas medidas tomadas pelas empresas de engenharia, devido às especialidades, a oportunidade de compartilhar e ampliar uma informação técnica em uma “árvore de referência verificável” de um determinado projeto pode oferecer soluções para o projeto, além de contribuir para a celeridade do prazo (McMahon; Lowe; Culley, 2004).

Em um projeto de Engenharia, ao vincular asserções e procedência usando representações semânticas, não apenas os dados se tornam interoperáveis, mas seu valor pode ser estimado de forma independente.

As NPs “técnicas” podem fornecer uma informação comum para a troca de dados e, portanto, permitir a atribuição de valor a esses dados individuais, facilitando o que hoje temos de maior dificuldade em projetos de Engenharia: a interoperabilidade e o compartilhamento de dados para a elaboração de projetos e emissão de documentos técnicos de Engenharia.

Entendemos que para que seja possível tornar esse dado de Engenharia interoperável, uma das ações que devemos tomar é tornar esse dado FAIR, ou seja, com princípios que promovam a acessibilidade e a usabilidade desses dados técnicos de Engenharia.

3.6 CURADORIA DIGITAL

Com o aumento exponencial na escala e diversidade dos dados, a necessidade de curadoria digital torna-se cada vez mais premente em todas as áreas do conhecimento. No âmbito da Engenharia, especificamente dos documentos técnicos de engenharia, escopo desta pesquisa, a curadoria surge para atender a gestão da ambiguidade dos dados utilizados nos projetos e da necessidade de tornar esses dados interoperáveis, considerando a alta obsolescência dos programas utilizados para a elaboração de cálculos, desenhos, ou outros tipos documentais necessários à incorporação do projeto.

Os progressos da tecnologia da informação e os avanços significativos na produção de dados de pesquisa, com simulações, gerações de dados por meio de instrumentos de medição, experiências, enfim, pesquisas que produzem cada vez mais dados nos mais variados formatos, dados estruturados em bancos de dados, dados semiestruturados ou não estruturados, nos levam ao desafio da gestão de dados, que atinge diversas áreas, os dados se multiplicam em arquivos, planilhas ou bancos de dados armazenados em discos rígidos, cadernos digitais, sites, *blogs* e *wikis* (Bell; Hey; Szalay, 2009).

3.6.1 Conceitos

A curadoria digital surge como resposta à explosão de dados. Sua criação, distribuição, uso e reuso, fazem parte de uma discussão geral que aos poucos foi envolvendo várias áreas do conhecimento como, Ciência da informação, Ciência da Computação, Engenharias, Biblioteconomia e Arquivologia. Assim sendo, se fez necessário uma área que contemplasse “a gestão e preservação de material digital para garantir a acessibilidade a longo prazo” (Abbott, 2008, p. 1, tradução própria). O Digital Curation Centre - DCC¹⁷ define curadoria como “manter e agregar valor a um corpo confiável de dados de pesquisa digital para uso atual e futuro, engloba o gerenciamento ativo durante todo o ciclo de vida da pesquisa” (Digital Curation Centre, 2004).

Sayão e Sales (2012, p. 180) entendem que “O desafio não é somente preservar o conjunto de dados, mas sobretudo a capacidade de transmitir conhecimento para uso futuro das comunidades interessadas”. Portanto, o esforço é para que os registros digitais ou dados sejam salvaguardados com vistas à preservação desses recursos digitais durante todo o ciclo de vida.

O termo Curadoria Digital é citado pela primeira vez em 2001, no seminário “*Digital curation: digital archives, libraries and e-Science seminar*”¹⁸, realizado em Londres, com

¹⁷ Disponível em: <http://www.dcc.ac.uk/>. Acesso em: 3 jun. 2024.

¹⁸ Disponível em: <https://www.ariadne.ac.uk/issue/30/digital-curation/>. Acesso em: 7 jun. 2024.

relatório de Neil Beagrie e Philip Pothen (Dallas, 2016). Esse foi um evento sobre preservação digital, e um dos pontos mais importantes foi a aprovação do modelo de referência *Open Archival Information Systems* – OAI, sobre o qual falaremos mais à frente.

Em 2003, Lord e McDonald no relatório “*Data curation for e-Science in the UK: an audit to establish requirements for future curation and provision*”, estabeleceram algumas definições para as três atividades principais na curadoria: “curadoria”, “arquivamento” e “preservação” (Lord; Macdonald, 2003).

Essas definições são a base para o ciclo de vida da curadoria digital, considerando a ênfase na gestão de dados, de sua criação até a preservação, sendo a “curadoria” a atividade de gerenciar e promover o uso de dados, desde sua criação, garantindo sua adequação ao propósito do dado e disponível para a recuperação e reutilização.

Para a atividade de “arquivamento”, os autores citam que é uma atividade de curadoria que garante que os dados sejam adequadamente selecionados, armazenados e que possam ser acessados, mantendo sua integridade lógica e física ao longo do tempo, incluídas a segurança e a autenticidade.

A “preservação” completa a cadeia da curadoria digital, sendo uma atividade dentro do arquivamento, onde os itens específicos de dados são mantidos ao longo do tempo para que possam ser acessados e compreendidos por meio da mudança tecnológica e, assim, com a preservação, falamos do item de interoperabilidade dos dados, importante nas ciências em geral e não menos importante para os dados de Engenharia (Lord; MacDonald, 2003, p.12).

O Digital Curation Centre - DCC surge em 2004, e a curadoria digital, segundo a definição do DCC, implica o gerenciamento, a preservação e o aperfeiçoamento de dados de pesquisa digital ao longo de seu processo. Trata-se de manutenção, preservação e agregação de valores aos dados garantindo uso e reuso dos dados.

Beagrie (2004) define curadoria digital como uma base ampla de curadoria de dados, abordando o uso científico e a preservação do patrimônio histórico e cultural em arquivos. Ele cita o seguinte:

O termo "curadoria digital" está sendo cada vez mais usado para as ações necessárias para manter dados de pesquisa digital e outros materiais digitais ao longo de todo o seu ciclo de vida e ao longo do tempo para as gerações atuais e futuras de usuários. Implícitos nessa definição estão os processos de arquivamento digital e preservação digital, mas também inclui todos os processos necessários para uma boa criação e gestão de dados, e a capacidade de agregar valor aos dados para gerar novas fontes de informação e conhecimento (Beagrie, 2004, p.7).

O processo de curadoria digital engloba todas as ações desde a criação, passando pelo acesso e uso, até a transformação. A preservação é uma das ações realizadas ao longo do

processo de curadoria, para assegurar a manutenção e retenção prolongada dos objetos digitais (Higgins, 2008).

Harvey (2010) complementa essa definição ao enfatizar que a curadoria digital é um conceito mais amplo do que arquivamento digital ou preservação digital. Aborda todo o leque de processos aplicados a objetos digitais ao longo de seu ciclo de vida. A curadoria digital começa antes da criação dos objetos digitais, estabelecendo padrões para a coleta de dados, descrição desses dados, o que resulta no planejamento desses objetos digitais prontos para terem uma curadoria e em condições de garantir que possam ser mantidos e usados no futuro.

Grande parte da literatura, ao definir curadoria digital, enfatiza as atividades de preservação digital, curadoria de dados, gerenciamento de dados, gestão de registros eletrônicos e outros ativos digitais ao longo de seus ciclos de vida completos que abrangem todas as etapas necessárias para a criação e gestão de dados com qualidade (Lee; Tibbo, 2007; Yakel, 2007).

A fragilidade dos materiais digitais e a suscetibilidade das mudanças tecnológicas requerem ações em cada estágio do ciclo de vida dos objetos digitais, para que seja possível manter a autenticidade e a integridade com vistas à reutilização (Pennock, 2007).

Na maioria das definições, percebe-se abordagens de ciclo de vida e preservação. Sendo assim, para efeito desta pesquisa, a curadoria digital pode ser definida como práticas que envolvem avaliar, gerenciar, preservar, agregar valor, arquivar e reutilizar os dados ao longo de seu ciclo de vida.

Na curadoria digital, a abordagem de ciclo de vida é utilizada para "garantir a manutenção da autenticidade, confiabilidade, integridade e usabilidade do material digital, o que, por sua vez, garante a maximização do investimento em sua criação" (Higgins, 2008, p. 135).

Sendo assim, podemos interpretar que as ações para o ciclo de vida da curadoria digital envolvem questões técnicas, gerenciais e tecnológicas, que devem ser mantidas com práticas de gestão de dados, estabelecendo utilização de padrões de metadados e critérios FAIR, contribuindo assim para a sua interoperabilidade e reuso.

Para tratar e gerenciar todos os registros digitais ou objetos digitais, dados oriundos de pesquisas científicas ou dados produzidos na indústria, para a elaboração de diversas frentes, a interdisciplinaridade da curadoria digital pode ser representada conceitualmente. Segundo Abbot (2008, p. 1, tradução própria),

A curadoria digital é a gestão e preservação de dados digitais a longo prazo. Todas as atividades envolvidas no gerenciamento de dados, desde o planejamento de sua criação, melhores práticas em digitalização e documentação, e garantir sua

disponibilidade e adequação para descoberta e reutilização no futuro, fazem parte da curadoria digital.

Uma visão predominante na literatura é a de que o conceito de curadoria digital é relativamente novo e que surge com a necessidade de preservação dos dados digitais, conhecimentos e práticas acumuladas na última década, que resultaram em um conjunto de estratégias e abordagens tecnológicas que atualmente são conhecidas como “curadoria digital” (Sayão; Sales, 2012).

3.6.2 Modelos

Existem vários modelos de ciclo de vida desenvolvidos na curadoria digital, isso reforça a necessidade de acompanhar a gestão dos dados antes mesmo de sua criação, definindo critérios para emissão, tornando-os acessíveis por máquinas para que em toda a cadeia eles estejam preservados, para que venham a ser utilizados e reutilizados quando necessário.

“Cada vez mais os dados digitais são um meio para novas descobertas. Podemos reutilizá-los para extrair valor adicional ou simplesmente evitar a duplicação do trabalho existente” (Lord; McDonald, 2003, p.9).

Os modelos descrevem como os *stakeholders* devem preservar e agregar valor aos dados, desenvolvendo padrões, identificando boas práticas de gestão, técnicas e tecnológicas, permitindo assim que o modelo em uso evolua e siga o que existe de mais básico em um ciclo de PDCA – *Plan, Do, Check, Action*, cuja função básica é auxiliar no diagnóstico, análise e problemas de uma organização, mas nesse caso seria um ciclo voltado para a melhoria contínua de um modelo de ciclo de vida de curadoria digital.

Do ponto de vista do ciclo de vida da informação, o modelo de curadoria digital abrange processos para alcançar a confiabilidade dos recursos digitais, a organização, arquivamento e preservação a longo prazo, agregando valor e novos usos para os recursos informacionais (Constantopoulos; Dallas, 2008).

Na nossa revisão de literatura encontramos alguns modelos de ciclo de vida da curadoria digital ou de preservação, mas para esta pesquisa decidimos trabalhar com os estudos de Alex Poole (2016) e Oliver & Harvey (2016).

Alex H. Poole (2016), em seu artigo “*The Conceptual Landscape of Digital Curation*”, cita oito modelos de ciclo de vida considerados principais:

- Digital Curation Center Lifecycle Model
- I2S2 Idealized Scientific Research Activity Lifecycle Model
- Data Documentation Initiative (DDI) Combined Life Cycle Model

- ANDS Data Sharing Verbs
- DataONE Data Lifecycle
- Research360 Institutional Research Lifecycle
- Capability Maturity Model for Scientific Data Management
- UK Data Archive Data Lifecycle

Quanto ao modelo Open Archival Information System – OAIS, o autor comenta que o modelo inovou e tornou-se um padrão, mas não constitui um modelo de ciclo de vida completo, pois desconsidera a especificação de diretrizes para a criação ou reutilização de dados (Poole, 2016, p.4).

Oliver e Harvey (2016) descrevem três modelos conceituais para curadoria digital, são eles:

- Digital Curation Center Lifecycle Model
- Data Curation Continuum
- Open Archival Information System Reference Model

Os autores consideram que o modelo de referência OAIS não oferece orientações sobre criação de dados ou sobre utilização e reuso de dados, corroborando com o mesmo posicionamento de Alex Poole.

Além dos modelos citados acima, outro que é importante ser citado, devido a sua contribuição para o modelo original do DCC, é o modelo apresentado por Constantopoulos *et al.* (2009), o DCC&U, um modelo de ciclo de vida digital estendido, pois é uma fusão do modelo apresentado pelo Digital Curation Centre - DCC e o modelo apresentado pela *Digital Curation Unit* – DCU.

Considerado como uma abordagem alternativa à visualização dos processos englobados na curadoria digital, o modelo ainda conta com as ações do ciclo de vida do *DCC Curation Lifecycle* e enfatiza a necessidade de registrar e manter como as informações armazenadas, curadas e preservadas são utilizadas e acessadas pelos usuários, considerando aqui a experiência como diferenciador, pois suas interações com o mundo digital criam e modificam ativos de informação (Constantopoulos *et al.*, 2009, p.41).

Faremos uma breve descrição sobre o *DCC Curation Lifecycle Model*, considerado flexível e genérico, portanto, passível de adaptação para as necessidades de curadoria em diversos campos do conhecimento. Para o propósito desta pesquisa, poderá ser utilizado como insumo para um modelo de curadoria digital de documentos técnicos de engenharia, tendo em vista que não foi identificado na literatura estudo voltado para esse domínio com escopo

completo, voltado a todo o ciclo de um dado de Engenharia digital ou documento técnico digital.

3.6.2.1 DCC Curation Lifecycle Model

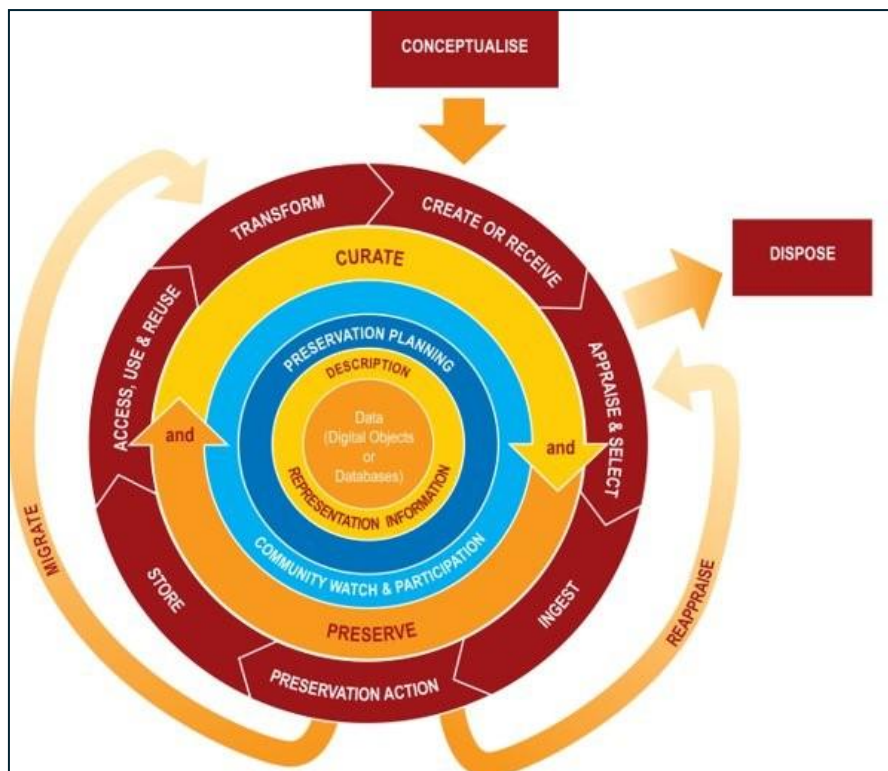
O modelo desenvolvido pelo *Digital Curation Centre* – DCC apresenta uma visão integrada ao desenvolvimento e otimização da infraestrutura de gestão de dados de pesquisa, fornecendo uma visão dos estágios necessários para uma curadoria e preservação (Digital Curation Centre, 2024).

O DCC Curation Lifecycle Model é o mais flexível e mais utilizado para planejar as atividades de curadoria, como descreve Higgins (2008, p. 135, tradução própria),

De natureza genérica, o modelo é indicativo e não exaustivo. Quando usado como uma ferramenta de planejamento organizacional, é adaptável a diferentes domínios e extensível para permitir que as atividades de curadoria e preservação sejam planejadas em diferentes níveis de granularidade.

Como consta na figura 10, o modelo é apresentado na forma de círculos concêntricos que representam o elemento-chave do modelo: dados, ações de ciclo de todo ciclo de vida, ações sequenciais e ações ocasionais. Os dados abrangem mais do que conjuntos de dados em um banco de dados; incluem objetos digitais de natureza textual e não textual (Sabharwal, 2015).

Figura 10 - The DCC Curation Lifecycle Model



Fonte: Higgins (2008).

Descrever o campo da curadoria digital é importante, mas se as definições são muito amplas, elas podem perder o foco e podem ser interpretadas de maneiras distintas, sendo assim, é importante enfatizar o conjunto de ações que agregam valor aos dados originais, que permitem a preservação, acessibilidade e reuso.

No âmbito dos projetos de Engenharia, percebemos um ecossistema digital disperso e com pouca discussão sobre como se beneficiar e compartilhar esses dados, são poucos os estudos sobre curadoria digital e, mesmo assim, com especificidades.

Para o desenvolvimento do modelo conceitual de curadoria digital para documentos técnicos de engenharia será utilizado o modelo já consolidado do DCC.

3.6.3 Curadoria Digital para projetos de Engenharia

Existe na literatura muita preocupação com os dados e o seu crescimento exponencial, principalmente em relação a dados científicos, considerando que um estudo voltado para as Ciências Biológicas, por exemplo, pode gerar milhões de dados que precisam ser compartilhados e preservados.

No domínio do conhecimento da Engenharia, encontramos alguns estudos voltados para o discurso sobre a necessidade de preservação e algumas tentativas de criação de sistemas visando recuperar informações e estabelecer camadas de dados mais leves, pensando na sua recuperação.

Traçamos, a seguir, uma linha cronológica ou *cronográfica* dos estudos identificados sobre curadoria digital na Engenharia, cujo conceito pode ser aplicado ao ciclo de vida de um projeto.

Projetos de engenharia ou criação, elaboração e manutenção de produtos, onde podemos citar o ciclo de vida do produto, geram inquietação em relação à curadoria e preservação de dados e documentos.

Em 2005, McMahon, Giess e Cullet realizaram um estudo de quatro aspectos de gestão da informação para apoiar o processo de suporte ao longo da vida do produto, *Product Life Management* – PLM, são eles: representação do modelo de engenharia, os sistemas utilizados, a organização dos dados e a gestão, já com foco no desafio de como projetar e preservar os produtos ou projetos nesse ambiente dinâmico, fazendo com que as informações relacionadas aos produtos ou projetos sejam reutilizadas pelas gerações futuras.

Nesse trabalho foram identificadas as seguintes lacunas:

- a) Representação do conhecimento – como o produto e o projeto de engenharia devem ser modelados e armazenados para facilitar o acesso a longo prazo;
- b) Questões relacionadas a sistemas – como os dados digitais devem ser arquivados e preservados;
- c) Organização dos dados – como as grandes coleções de informações podem ser organizadas e recuperadas por *stakeholders* ao longo do tempo com segurança e confiabilidade adequada;
- d) Gestão de comportamento – como as organizações irão tratar a gestão da informação ao longo do ciclo de vida do produto e da cadeia produtiva.

Considerando que o estudo é de 2005, 18 anos depois ainda não temos um estudo de curadoria digital específica para os dados de Engenharia. A reflexão de McMahon, Giess e Cullet (2005) continua atual:

O conteúdo semântico dos modelos de engenharia é baixo. Os modelos geométricos são essencialmente coleções de geometria com pouco ou nenhum conteúdo semântico formal, e embora tenha havido progresso no aprimoramento da semântica de modelos por meio de modelagem baseada em recursos, ou na identificação de conteúdo semântico por meio do reconhecimento de recursos – por exemplo, no planejamento de processos – o progresso tem sido limitado devido à dificuldade de reconhecimento e à questão da dependência do ponto de vista – cada especialidade de engenharia tem sua própria semântica e precisa de suas próprias representações de modelo (McMahon; Giess; Cullet, 2005, p.29, tradução própria).

Immortal Information and Through-Life Knowledge Management – IITKM ou KIM Project, foi um projeto multidisciplinar que iniciou em 2005 e foi realizado pelo consórcio de oito centros de pesquisa e onze Universidades do Reino Unido, além de participação da indústria, empresas do setor da construção, empresas aeroespaciais e fornecedores e consultores de *software*. O objetivo do projeto era estabelecer boas práticas quanto à concepção e uso de sistemas de informação e suporte ao conhecimento através de melhorias na gestão de informações, facilitando a aprendizagem contínua e gerenciando efetivamente o conhecimento ao longo do ciclo de vida do produto ou serviço (Ball *et al.*, 2006).

O projeto KIM enfatizava a necessidade de estratégias robustas de gerenciamento do conhecimento e curadoria digital para garantir que os dados técnicos, modelos e informações de engenharia permanecessem acessíveis e relevantes ao longo do tempo. Com base nessa diretriz, o projeto tinha três áreas principais para desenvolver o trabalho, que consistia em padrões e ferramentas, são elas: Representação e gerenciamento de informações; Aprendizado ao longo do ciclo de vida do produto; e Gerenciamento do ciclo de vida do conhecimento.

O projeto durou três anos e gerou alguns produtos com vistas a resolver os problemas de acessibilidade da informação, preservação e interoperabilidade, principalmente de projetos gerados em CAD, que possui grandes problemas de emulação e migração, pois dificilmente oferece soluções satisfatórias (Ball *et al.*, 2006; Ball; Neilson, 2010; Ball; Patel; Ding, 2008; McMahon *et al.*, 2009).

Percebe-se um esforço em estabelecer padrões e meios de assegurar a gestão da informação, autenticidade e acessibilidade dos documentos de Engenharia, estabelecendo o que alguns autores chamaram de gestão “semântica”, citando sobre a importância do uso de metadados e habilitando, assim, a cadeia de preservação e, conseqüentemente, a de curadoria.

Um desses estudos é baseado na utilização do modelo OAIS para o desenvolvimento de ferramentas de planejamento e preservação da documentação de engenharia (Ball; Patel; Ding, 2008; Bahloul; Buzon; Bouras, 2008).

Em 2010, Alex Ball e Colin Neilson publicaram o estudo “*Curation of research data in the disciplines of Engineering*” (Ball; Neilson, 2010). O relatório faz parte do projeto conduzido pela Digital Curation Centre – DCC, “DCC SCARP”, acrônimo de Sharing, Curation, Reuse and Preservation.

O estudo ressalta o contexto comercial e industrial para trabalhar com a pesquisa na Engenharia devido à confidencialidade de alguns dados que requer um nível de controle na divulgação das informações. Talvez esse seja um dos motivos de não existir repositórios públicos com dados de engenharia. Percebe-se que para trabalhar a curadoria, pensando no compartilhamento desses dados, deve existir um nível de acesso para esses dados, evitando quebra de confidencialidade e problemas para as empresas envolvidas.

Essas pesquisas e discussões destacam a necessidade urgente de desenvolver e implementar estratégias eficazes de curadoria digital na Engenharia, reconhecendo a importância crítica de preservar e gerenciar informações técnicas para o avanço contínuo da disciplina e para garantir a sustentabilidade e a confiabilidade de produtos e sistemas de engenharia ao longo do tempo.

Na condução desta pesquisa, não foi possível identificar literatura mais recente com foco na curadoria digital para os documentos de Engenharia ou estudos na área de preservação digital voltados para todo o ciclo de vida de um projeto de Engenharia.

Existem alguns estudos específicos para algumas disciplinas, como Engenharia Civil ou Arquitetura, preocupados com o conjunto de dados que estão dispersos e com propostas de uso dos modelos de preservação e dos critérios FAIR, para esses dados de engenharia gerados

(Patlakas; Musso; Larkham, 2021), com isso percebemos uma lacuna na literatura voltada para a criação de um modelo de curadoria digital específico para a Engenharia, que como vimos, é complexa devido a sua interdisciplinaridade e múltiplas formas de geração de dados.

Esta base teórica descreve e apresenta a nanopublicação como uma ferramenta para garantir rastreabilidade, interoperabilidade e usabilidade. Estudamos projetos de engenharia e as metodologias de gerenciamento de projetos, como o PMBOK® (Project Management Institute, 2021), que descrevem o ciclo de vida e as práticas de gerenciamento para projetos de engenharia.

Os princípios do gerenciamento de dados FAIR (Wilkinson *et al.*, 2016) e a integração de tecnologias semânticas, conforme discutido por Kuhn *et al.* (2013) e outros, apoiam o foco do estudo na estruturação de informações para recuperação e reutilização da informação.

Contemplamos os desafios específicos da Engenharia, como a aplicação de modelos de curadoria digital (Poole, 2016) e as necessidades de preservação articuladas por McMahon, Giess e Culley (2005), construindo uma base teórica para explorar a abordagem da nanopublicação na documentação de Engenharia.

Assim, encerramos esta seção sobre Referencial Teórico, tendo sido estabelecido um arcabouço para embasar a teoria que trabalhamos ao longo da pesquisa.

4 METODOLOGIA

A pesquisa aqui projetada é de natureza aplicada, com abordagem qualitativa, busca a compreensão dos conceitos e aplicações das teorias existentes sobre Nanopublicação, históricos, aplicações e projetos existentes para que seja possível construir um modelo voltado para a documentação técnica de engenharia.

A pesquisa qualitativa, segundo Yin (2016, p.6), “contribui com revelações sobre conceitos existentes ou emergentes e esforça-se por usar múltiplas fontes de evidência em vez de se basear em uma única fonte”. Além disso, é explicitada por Creswell e Creswell (2021) como sendo “uma abordagem voltada para a exploração e para o entendimento do significado que indivíduos ou grupos atribuem a um problema social ou humano” (Creswell; Creswell, 2021, p.1).

Pela necessidade de compreensão de fatos teóricos, a pesquisa qualitativa seguiu com o objetivo de ser exploratória, pois busca proporcionar maior familiaridade com o problema e tem, conforme citado por Gil (2022, p.27),

[...] como principal finalidade desenvolver, esclarecer e modificar conceitos e ideias, tendo em vista a formulação de problemas mais precisos ou hipóteses pesquisáveis para estudos posteriores. Dessa maneira visa proporcionar uma perspectiva do tema investigado, por meio de um processo que engloba pesquisa bibliográfica, estudo de caso e análise documental.

Para estabelecer a base teórica dessa pesquisa, foi realizado levantamento, entre abril e julho de 2023, em bases de dados interdisciplinares, com busca delimitada a artigos de revisão por pares, artigos de conferências ou congressos essencialmente ligados à área de Engenharia ou Curadoria Digital.

As bases consultadas foram: *Scopus* e *Web of Science*, por se tratar de bases mais voltadas para estudos das Engenharias e tecnologias.

Para realizar as buscas foram utilizados os seguintes termos: (“Information retrieval” AND "engineering document") OR ("Project document") OR ("Technical information"); “Engineering project”; (“digital curation” AND "engineering document"); “nanopublication or nanopublicação” e “nanopub*”.

As buscas não foram limitadas aos termos citados, e a pesquisa contemplou também artigos referenciados nas publicações selecionadas.

Apesar das buscas iniciais terem sido realizadas no período citado acima, foram realizadas também algumas atualizações com corte temporal até dezembro de 2023 e com atualizações em julho de 2024, para incluir buscas na base *IEEE Explorer*.

Considerando que o tema é relativamente novo, foi necessário estabelecer também um estudo bibliográfico para o tema “Nanopublicação”, com vistas a conhecer as áreas que já usam o conceito e ampliar a possibilidade de uso na Engenharia.

A pesquisa bibliográfica, que iremos considerar a partir de agora como uma Revisão Sistemática de Literatura, foi realizada a partir de buscas em bases de dados de publicações científicas, com o intuito de verificar os estudos científicos existentes sobre curadoria digital e preservação digital voltados para projetos de Engenharia.

Sendo o estudo multidisciplinar, pois envolve as áreas de Ciência da informação, Engenharia e Ciência da Computação, ampliamos os resultados das buscas verificando as referências dos autores citados nos textos selecionados para leitura e interpretação do estudo.

As buscas foram feitas nas bases de dados *SCOPUS* e *Web Of Science*, além do *Google Scholar*, utilizado para verificar as citações dos estudos escolhidos. As bases foram selecionadas a partir do Portal de Periódico CAPES.

A pesquisa bibliográfica traz a possibilidade de entender sobre estudos anteriores e posteriores de alguns autores escolhidos. Neste sentido, elaboramos uma pesquisa usando o aplicativo de Inteligência artificial “Conect Papers”, cujos resultados foram transferidos para o gerenciador eletrônico Zotero e analisado conforme o grau de citação dos estudos.

Creswell e Creswell (2021, p. 11), em relação ao estudo de caso, entendem que ele permite que o pesquisador desenvolva uma análise profunda de um caso, geralmente um projeto, um evento, uma atividade, um processo. Para a nossa pesquisa, o estudo de caso é voltado para o processo de gestão de documentos técnicos de engenharia, analisando seus fluxos, procedimentos utilizados, políticas, diretrizes e sistemas adotados, além de um experimento em um *corpus* documental com vistas ao desenvolvimento do modelo.

O *corpus* documental adotado para usar na pesquisa é referente ao projeto executivo de uma subestação de entrada desenvolvido para um empreendimento na Petrobras.

A presente pesquisa possui autorização para uso da análise documental do projeto citado.

O percurso metodológico foi seguido nas etapas descritas abaixo e consolidadas no Quadro 1.

Quadro 1 – Percurso Metodológico

Objetivo	Etapa Metodológica	Resultado
OE1 - Realizar revisão bibliográfica sobre o conceito de nanopublicações e suas aplicações atuais em diferentes domínios	Revisão Bibliográfica Classificação Material bibliográfico Fichamento dos materiais selecionados	Estudo Bibliográfico sobre Nanopublicação e seu uso em outros domínios
OE2 - Propor, por meio de estudo de caso, um modelo de nanopublicação para documentos técnicos de engenharia	<ul style="list-style-type: none"> • Estudo de caso em um corpus documental – Projeto Subestação de obras • Identificação dos documentos • Modelagem dos conceitos dos documentos 	Descrição de um modelo conceitual de Nanopublicação para documentos de engenharia
OE3 - Propor uma curadoria digital para documentação técnica de Engenharia, que atenda às necessidades de representação, recuperação, acesso, interoperabilidade, reuso e preservação de dados e informações.	<ul style="list-style-type: none"> • Descrever o fluxo de informações durante a execução de um projeto de engenharia buscando a proposição do modelo. 	Especificação de curadoria digital para documentos técnicos de engenharia

Fonte: Autoria própria (2024).

5 ESTUDO BIBLIOGRÁFICO SOBRE NANOPUBLICAÇÃO E SUAS APLICAÇÕES

O conceito de Nanopublicação (Nanopublication) é utilizado com maior intensidade nas Ciências da Vida, sendo um conceito relativamente novo, com o seu surgimento datado de 2009 (Mons; Velterop, 2009).

A NP, citada inicialmente por Mons e Velterop (2009), tem por objetivo transformar dados de pesquisa em informações legíveis por máquinas, facilitando assim a sua interoperabilidade e o seu reuso.

As pesquisas, publicadas como NP, trazem na sua estrutura pequenas informações científicas, chamadas "afirmações" ou "declarações", que podem ser compostas em outros artigos e documentos.

As afirmações ou declarações são acompanhadas por metadados que indicam a sua proveniência, como a fonte original e o contexto em que foram coletadas, além de metadados com as informações da NP em si.

O objetivo desta seção é identificar e analisar os campos de aplicação do conceito de NP, bem como compreender a sua contribuição para a representação, interpretação e compartilhamento de dados.

O método aqui utilizado foi apresentado por Levy e Ellis (2006), por se ter uma metodologia de revisão de literatura voltada para Sistemas de Informação. No conceito destes autores, a revisão segue etapas para coletar, conhecer, compreender, aplicar, analisar, sintetizar e avaliar a literatura, a fim de fornecer uma base sólida para um tópico, além de demonstrar que a pesquisa proposta contribui com algo novo para o corpo geral de conhecimento (Levy; Ellis, 2006, p.189).

Esta seção foi escrita seguindo as diretrizes do protocolo PRISMA 2020 (Figura 11).

5.1 OBJETIVO DA REVISÃO BIBLIOGRÁFICA

O objetivo a ser atendido com a revisão é identificar a aplicação de nanopublicações em diferentes domínios.

5.2 CRITÉRIOS DE ELEGIBILIDADE

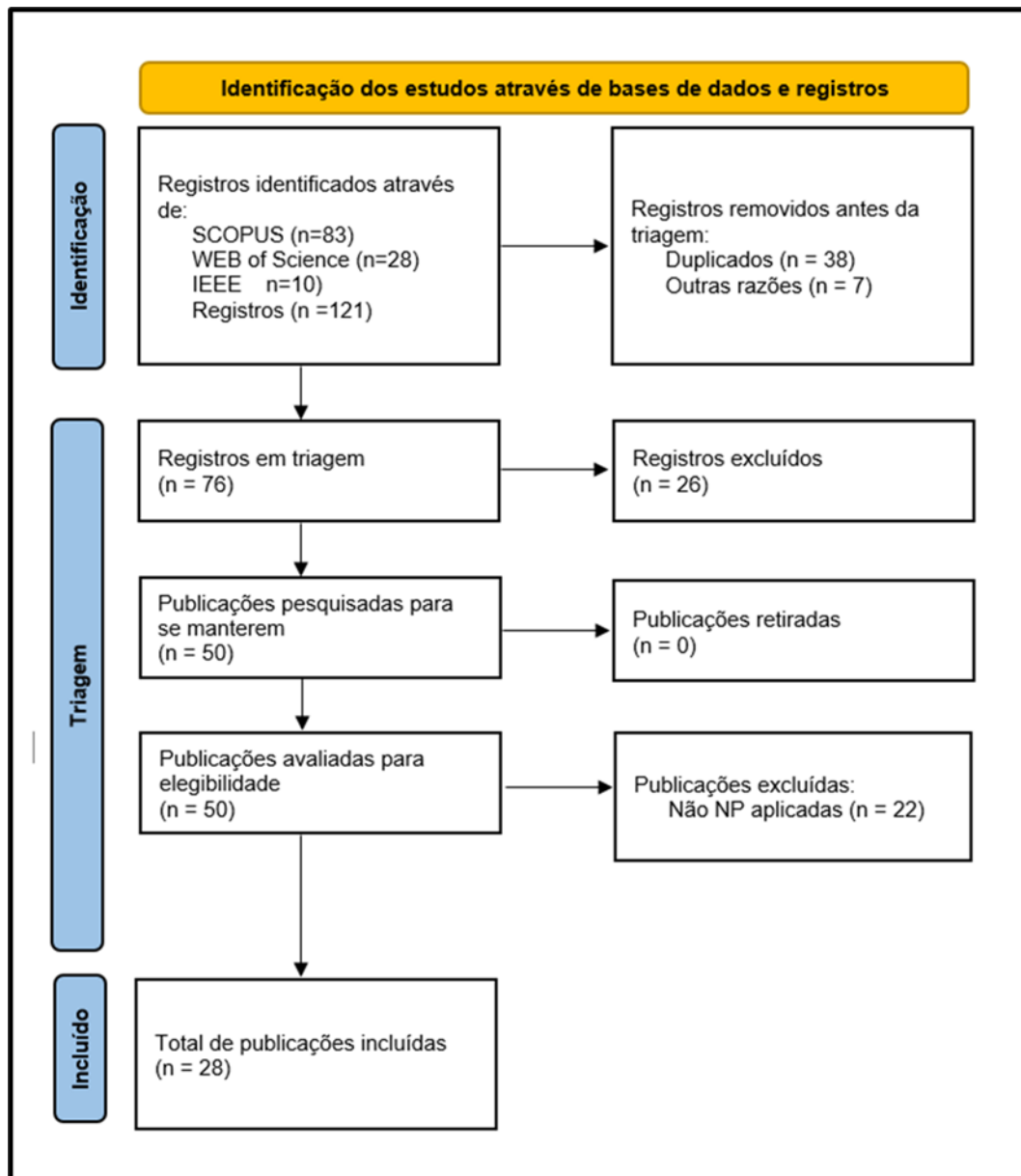
a) Critério de inclusão:

- Estudos cujo foco é o desenvolvimento, aplicação ou evolução de Nanopublicação em diferentes domínios.

b) Critério de exclusão:

- Estudos não relacionados com a aplicabilidade de Nanopublicação;

Figura 11 - Fluxograma PRISMA



Fonte: Page *et al.* (2021).

5.3 ANÁLISE DOS RESULTADOS

A análise dos resultados foi construída considerando o objetivo principal da revisão de literatura. Para uma maior completude no escopo, realizamos uma análise bibliométrica, identificando os artigos mais citados, frequência de palavras e os conceitos mais utilizados na literatura científica.

5.3.1 Análise bibliométrica dos artigos selecionados

Com base nos estudos selecionados, identificamos que os anos com maior número de publicações foram: 2015, com um total de dez publicações, seguido por 2012 e 2018, com oito publicações e, por último, 2023, com seis publicações.

O Quadro 2 e a Figura 12 abaixo apresentam uma análise bibliométrica dos estudos selecionados. Análise dos artigos mais citados e análise por nuvens de palavras.

a) Análise dos artigos mais citados

Na fase de seleção dos artigos avaliamos os artigos mais citados, por entender que são estudos de grande impacto sobre o assunto “Nanopublicação”, onde percebemos também os autores com maior quantidade de publicações: Barend Moons, Jan Valterop e Tobias Kuhn.

Essa verificação das citações foi feita com base na ferramenta *Publish or Perish – PoP*¹⁹, um programa que recupera e analisa citações acadêmicas utilizando várias fontes de dados. Com o resultado do PoP foi realizada uma comparação com o *i10-index* do Google Acadêmico e optou-se por apresentar esse resultado na pesquisa, onde selecionamos os 5 artigos mais citados, conforme apresentado a seguir (Quadro 2).

Quadro 2 – Artigos mais citados

Autor	Título	Ano	Google Scholar citações
<i>P. Groth, A. Gibson, J. Velterop</i>	The anatomy of a nanopublication	2010	459
<i>B. Mons, J. Velterop</i>	Nano-publication in the e-science era	2009	143
<i>T. Kuhn, P.E. Barbano, M.L. Nagy, M. Krauthammer</i>	Broadening the scope of nanopublications	2013	92
<i>G.P. Patrinos, D.N. Cooper, E. van Mulligen, V. Gkantouna, G. Tzimas, Z. Tatum, E. Schultes, M. Roos, B. Mons</i>	Microattribution and nanopublication as means to incentivize the placement of human genome variation data into the public domain	2012	78
<i>T. Kuhn, A. Merono-Penuela, A. Malic, J.H. Poelen, A.H. Hurlbert, E.C. Ortiz, L.I. Furlong, N. Queralt-Rosinach, C. Chichester, J.M. Banda, E. Willighagen, F. Ehrhart, C.Evelo, T.B. Malas, M. Dumontier</i>	Nanopublications: A growing resource of provenance-centric scientific linked data	2018	42

Fonte: Autoria própria (2024).

b) Análise por Nuvens de palavras

¹⁹ <https://harzing.com/home>

A escolha pelos autores deu-se devido a sua ocorrência nas publicações, os conceitos são fontes dos estudos seminais (Quadro 3).

Quadro 3 – Conceitos de Nanopublicações

Autor	Nanopublicações - Conceito
Mons, Barend; Velterop, Jan. Nanopublication in the e-Science era , 2009.	São definidas como “uma unidade mínima de informação publicada que descreve uma única evidência empírica”. Consistem em três partes: um sujeito, um predicado e um objeto, expressos em RDF. As nanopublicações são projetadas para serem compartilhadas, combinadas e analisadas por máquinas, permitindo uma extração de conhecimento mais eficiente e abrangente a partir de dados científicos
Groth, Paul; Gibson, Andrew; Velterop, Jan. The anatomy of a nanopublication , 2010	São declarações científicas centrais com contexto associado que visam enfrentar os desafios da comunicação científica, como atribuição, qualidade e procedência. É um modelo de comunicação científica que consiste numa unidade mínima de informação publicável, que pode ser validada cientificamente somente se o seu contexto for levado em consideração.
Kuhn, T. <i>et al.</i> A growing resource of provenance-centric linked data , 2018	A nanopublicação consiste num fragmento atômico, de uma declaração formal (“asserção/afirmação”), que vem com informações da origem do conhecimento (proveniência da afirmação) e com metadados sobre a Nanopublicação em si (proveniência da nanopublicação, como foi criada, essa é a parte chamada “informações da publicação”. Todas as três partes representadas em RDF

Fonte: Dados da pesquisa (2024).

5.3.3 Análise crítica dos resultados

A seguir, iremos apresentar uma análise das contribuições identificadas nos artigos durante a revisão de literatura.

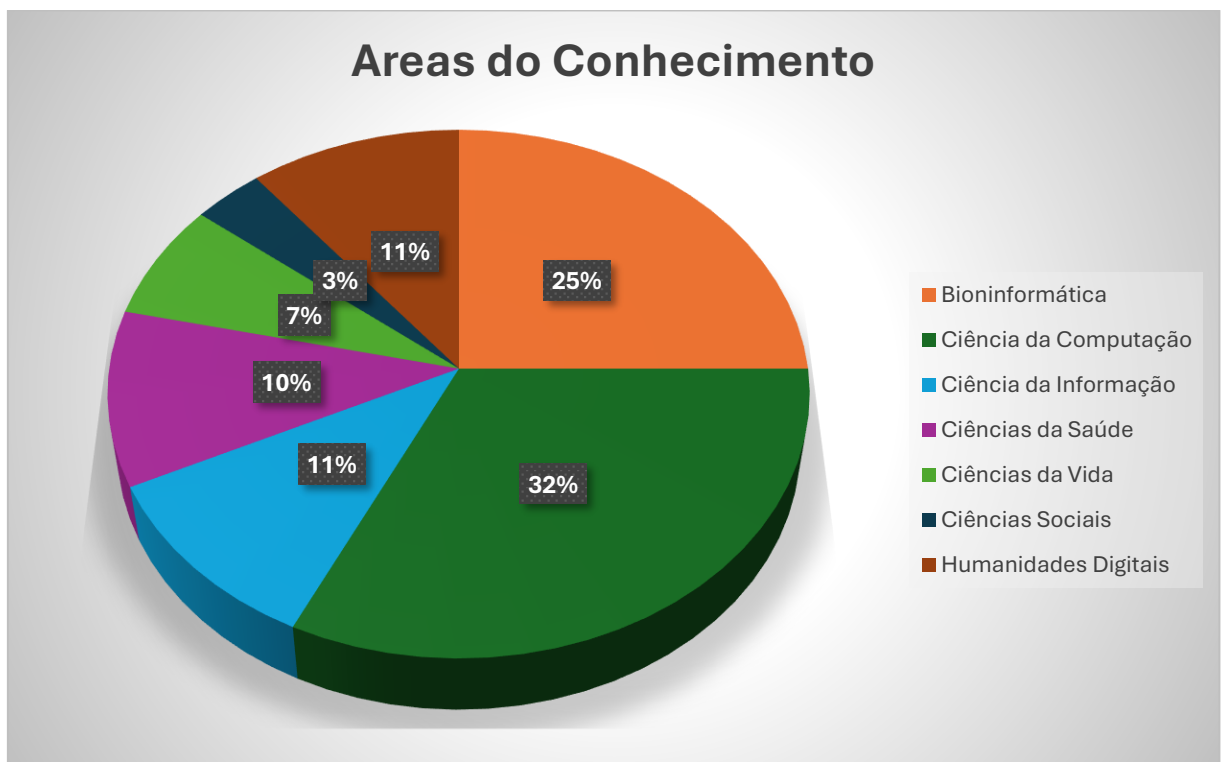
Esta análise visa explorar os estudos, para identificar as principais colaborações proporcionadas por esse conceito, NP, nas diversas áreas do conhecimento, exemplificando que as NPs se estendem para além da comunicação científica, como citado por Kuhn et al. (2013)

Nanopublicações podem se tornar a base para toda a Web Semântica. Qualquer informação que se queira compartilhar, pode ser publicada na forma de uma ou mais nanopublicações. Podem incluir afirmações científicas e dados experimentais, mas também opiniões, relações sociais, eventos, propriedades de outras nanopublicações ("meta-nanopublicações"), e muito mais (Kuhn *et al.*, 2013, p.491, tradução própria).

Corroborando com a visão citada por Kuhn, Golden e Shaw (2016), entendem que a Nanopublicação é uma maneira útil de localizar a produção de dados em um contexto mais amplo, sendo uma abordagem viável até mesmo para áreas fora das ciências experimentais e observáveis.

O gráfico 1 representa as áreas de conhecimento mapeadas durante a revisão de literatura, e pretende atender ao objetivo específico **OE1**): “Realizar revisão bibliográfica sobre o conceito de nanopublicações e suas aplicações atuais em diferentes domínios”.

Gráfico 1- Áreas do conhecimento identificadas



Fonte: Elaboração própria (2024)

O foco da revisão foi identificar em quais áreas do conhecimento foram utilizados e aplicados o conceito de Nanopublicação. Sendo assim, foi identificado que a área com maior número de estudos foi Ciência da Computação, representando 32% dos artigos. Na sequência, temos Bioninformática com 25% dos estudos, Ciência da Informação e Humanidades digitais (11%) cada, Ciências da Saúde (10%) e Ciências da Vida e Ciências Sociais, (7%) e (3%), respectivamente.

Para identificação das áreas do conhecimento nos artigos, foi utilizada a tabela de áreas do conhecimento da CAPES22, considerando sempre a área de maior abrangência do artigo ou o seu domínio de aplicação, por exemplo, o artigo claramente tratava sobre Websemântica ou Tecnologia da informação, o enquadramento ficou em “Ciência da Computação”.

5.3.3.1 Ciência da Computação

Os artigos classificados como Ciência da Computação tratam de estudos voltados ao desenvolvimento de aplicações, visando melhorar a qualidade, acessibilidade e interoperabilidade dos dados de pesquisa, publicações científicas e dados biomédicos.

Nesses domínios percebe-se que a grande motivação para utilizar a NP ocorre devido à interoperabilidade, necessidade de integração de grande volume de dados e a rastreabilidade através da proveniência.

Ao discorrerem sobre essa dificuldade de gestão de grandes volumes de dados, Sernadela e colaboradores (2014) entendem que a NP é uma resposta natural à complexidade por trás da comunicação científica, superando a inconsistência, ambiguidade e redundância das publicações tradicionais.

McCusker et al. (2018) ao desenvolverem um framework para criar grafos do conhecimento, baseiam-se na NP, pois o seu uso simplifica e padroniza a produção do conhecimento estruturado, desse modo, com uma estrutura flexível, baseada em NP, esses grafos podem integrar, estender e publicar conhecimento de fontes heterogêneas.

O estudo produzido por Lek et al. (2020) nos sugere que a NP pode ser aplicada a diversas áreas do conhecimento, devido a sua capacidade de integração de diferentes tipos de dados. Eles apresentam uma representação de corpora linguísticos baseado em Linked data, na forma de NPs e concluem que esse conjunto de dados pode se beneficiar da estrutura refinada e consistente da proveniência das NPs.

A seguir apresentamos os artigos classificados como Ciência da Computação (Quadro 4), com o contexto e a motivação de uso.

Quadro 4 – Ciência da Computação

Autores	Contexto	Motivação do uso de NP
Bucur, Cristina-Iulia, Kuhn, Tobias, Ceolin, Davide, 2020	Publicações Científicas	Tornar o processo de publicação científica representada em pequenos trechos dos artigos, usando a natureza semântica e interoperável da NP. Esses

²² <https://www.gov.br/capes/pt-br/aceso-a-informacao/acoes-e-programas/avaliacao/instrumentos/documentos-de-apoio/tabela-de-areas-de-conhecimento-avaliacao>

Autores	Contexto	Motivação do uso de NP
		pequenos trechos podem se conectar a uma entidade maior e formar uma rede complexa de links.
Sernadela, Pedro; Oliveira, José, 2014	Dados de pesquisa Biomédica	Tornar os dados não estruturados resultantes da pesquisa biomédica, legível por máquina, superando assim a inconsistência, ambiguidade e a redundância das publicações tradicionais
Kuhn, Tobias et al, 2013	Publicações Científicas	A ampliação do conceito de NP com AIDA facilita a busca e comunicação de hipóteses de pesquisa, reivindicações e opiniões em uma estrutura de uma NP existente, pois essa representação no modelo AIDA torna o conceito original mais flexível e amplia a sua aplicação, na prática.
McCusker, James et al, 2018	Dados Biomédicos	Melhorar a curadoria e a gestão de dados científicos, facilitando a construção de grafos do conhecimento com foco em proveniência, além de aumentar a interoperabilidade e integração dos dados com o uso de NPs.
Lopes, Pedro; Sernadela, Pedro; Oliveira, José, 2013	Dados Biomédicos	O crescente volume de dados científicos, particularmente em biomedicina, e a necessidade de métodos mais eficazes para integrar, compartilhar e reutilizar esse conhecimento.
Clare, A. et al, 2011	Notas e manuscritos científicos	Disseminar o conhecimento acumulado nas notas eletrônicas e manuscritos e que não são compartilhados, bem como centralizar as publicações e anotações de blogs e wikis para consumo e retroalimentação das NPs com novas hipóteses.
Fabris, Erika, 2019	Nanopublicações	Promover a interoperabilidade, integração e compartilhamento de dados científicos, bem como promover o uso das NPs por não especialistas, ampliando assim a aplicação do conceito.

Autores	Contexto	Motivação do uso de NP
Feijoó, Matheus P.P. et al, 2023	Dados de pesquisa	Melhorar a qualidade, acessibilidade e interoperabilidade dos dados de pesquisa, seguindo os princípios FAIR.
Lek, Timo et al, 2020	Corpora Linguístico	Os desafios de interoperabilidade e proveniência em corpora linguísticos. Com as NPs pode haver a integração de dados de diferentes corpora de maneira uniforme, facilitando a fusão automática de anotações e a recuperação de informações de forma mais transparente e confiável.

Fonte: Autoria própria (2024).

5.3.3.2 Bioinformática

Seguindo com a análise dos artigos selecionados, vamos apresentar no Quadro 5 os estudos identificados para a área de Bioinformática, que revelam aplicações para tratar os dados biológicos e biomédicos, cujos problemas são parecidos com os citados para os estudos da área de Ciência da computação: tratar e integrar grandes volumes de dados diversificados.

Quadro 5 – Bioinformática

Autores	Contexto	Motivação do uso de NP
McCusker, J. et al, 2014	Dados Biológicos	Facilitar a integração de dados de diversas fontes, promovendo a interoperabilidade e reutilização dos dados, com o uso da NP, o framework possibilita uma visualização e exploração de redes biológicas complexas, com foco específico na descoberta de novas associações entre drogas e doenças por meio da integração desses dados.
Sernadela, Pedro et al, 2014	Dados de pesquisa Biomédica	Sobrecarga de informações no domínio biomédico. As nanopublicações representam uma solução para

Autores	Contexto	Motivação do uso de NP
		apresentar de forma granular as associações entre dados e autores.
Sernadela, Pedro et al, 2018	Dados Biomédicos	Tornar os dados não estruturados resultantes da pesquisa biomédica, legível por máquina, superando assim a inconsistência, ambiguidade e a redundância das publicações tradicionais.
Giraldo, Olga L. et al, 2022	Dados abertos de Biologia Molecular	As nanopublicações são usadas para representar de forma semântica e granular cada etapa dos protocolos, tornando-os FAIR. A intenção é melhorar a pesquisa, organização e manutenção dos protocolos, além de facilitar sua integração em fluxos de trabalho automatizados.
Queral-Rosinach, Núria et al, 2016	Dados genéticos do banco de dados DisGeNET	São os desafios de rastreabilidade, integração e verificação de dados nas associações entre genes e doenças, pois há uma necessidade crescente de métodos mais eficientes para representar e compartilhar informações científicas de forma automatizada e confiável.
Chichester, Christine et al, 2014	Base de conhecimento de proteínas humanas neXtProt	Os desafios de integração de dados biológicos, como variações de nomenclatura, ontologias incompatíveis e dificuldades de rastreamento de proveniência podem ser resolvidos usando as nanopublicações devido sua estrutura semântica padronizada que permite representar dados de forma granular, com atribuição clara e rastreabilidade, facilitando assim as consultas complexas na área da Biologia.
Gibson, Andrew et al, 2012	Dados Biológicos	Melhorar a avaliação da qualidade e da confiabilidade das evidências científicas nas ciências da vida. As nanopublicações permitem que as afirmações científicas sejam publicadas com metadados de proveniência detalhados, incluindo fonte, contexto e atribuição, o que facilita o rastreamento e a verificação das evidências.

Fonte: A autoria própria (2024).

5.3.3.3 Ciências da Saúde

Na área das Ciências da Saúde identificamos três estudos onde o foco eram as aplicações do conceito de NP para realizar integração de dados, pensando na confiabilidade dos dados e da sua proveniência. No Quadro 6 apresentamos esses estudos.

Quadro 6 – Ciências da Saúde

Autores	Contexto	Motivação do uso de NP
Sernadela, Pedro; Lopes, Pedro; Oliveira, José Luis, 2013	Dados de farmacovigilância (registros eletrônicos de pacientes de vários países europeus, informações sobre prescrições de medicamentos, vacinas, resultados de testes laboratório)	Como a farmacovigilância envolve dados de diversas fontes, as NPs permitem integrar essas fontes de dados em um formato unificado e semântico, melhorando a interoperabilidade entre os diferentes sistemas de dados, além de que as NPs permitem representar essas associações de maneira granular e rastreável, incluindo dados de proveniência garantindo a transparência e a confiabilidade das informações, portanto uso de NPs na farmacovigilância busca melhorar a integração, análise e transparência dos dados, abordando desafios comuns de heterogeneidade, volume e rastreabilidade no campo da segurança de medicamentos.
Gonzalez, Alejandro Rodriguez et al, 2014	Dados clínicos de pacientes	As NPs permitem armazenar informações de proveniência detalhadas, garantindo a reprodutibilidade e auditoria das decisões diagnósticas. Essa abordagem facilita o desenvolvimento de sistemas de suporte à decisão mais precisos e personalizados, ao registrar como e por que cada diagnóstico foi feito, bem como a origem das regras diagnósticas.
Banda, Ruan M. et al, 2015	Dados de interações de medicamentos	A necessidade de consolidar múltiplas fontes de interações medicamentosas em um único formato interoperável, permitindo fácil acesso e análise conjunta de diferentes fontes.

Fonte: Autoria própria (2024).

5.3.3.4 Ciências da Vida

Os estudos relacionados no Quadro 7 focam na interoperabilidade e integração dos dados com o intuito de descobrir novas associações genômicas e de incentivar os pesquisadores a compartilharem os seus dados científicos de forma semântica.

Quadro 7 – Ciências da Vida

Autores	Contexto	Motivação do uso de NP
Van der Horst et al, 2015	Dados genômicos e dados de pesquisas científicas	A integração de dados usando NPs geram novas descoberta de associações genéticas com avanços na reprodutibilidade e nas citações de dados genômicos.
Thompson, Mark; Schultes, Erik, 2012	Dados clínicos de pacientes	A motivação surge da interoperabilidade e promoção de reastreabilidade, permitindo a integração semântica de dados das ciências da vida proporcionado pela capacidade granular das NPs.

Fonte: Autoria própria (2024).

5.3.3.5 Ciência da Informação

Estudos realizados para a área de Ciência da informação versando sobre Recuperação da informação e indexação. Dois estudos são dos mesmos pesquisadores, mas se complementam e por esse motivo ficaram na lista dos selecionados (Quadro 8).

Quadro 8 – Ciência da Informação

Autores	Contexto	Motivação do uso de NP
Lipani, Aldo et al, 2014	Desenvolvimento de ontologia para Recuperação da informação, formalizando conceitos, componentes e relações fundamentais no campo	A dificuldade de reprodução e comparação de resultados de pesquisas na área de RI, os dados gerados não são totalmente incluídos nos artigos publicados ou não estão em formato legível por máquina, o que limita a capacidade de avaliação e verificação dos resultados.

Lipani, Aldo et al, 2014a	Recuperação da informação	Promover a reprodutibilidade na pesquisa de recuperação de informação (IR), pois na maioria das vezes os artigos de RI não fornecem detalhes suficientes sobre os experimentos, dificultando a replicação dos resultados por outros pesquisadores capacidade granular das NPs.
Paula, Lorena; Moura, Maria Aparecida, 2015	Ciência da informação com foco em representação e organização da informação em ambientes digitais	Necessidade de otimizar a representação, disseminação e recuperação de informações em ambientes digitais.

Fonte: Autoria própria (2024).

5.3.3.6 Humanidades Digitais e Ciências Sociais

Nos estudos referentes a área das Humanidades digitais podemos perceber que a aplicação do uso de NPs vem da necessidade de manter dados de proveniência de objetos históricos e culturais. Quanto ao estudo selecionado sobre Ciências Sociais, o seu foco é nos estudos científicos e aborda a necessidade de comparar os resultados de pesquisa (Quadro 9).

Quadro 9 – Humanidades Digitais e Ciências Sociais

Autores	Contexto	Motivação do uso de NP
Mariani, Fabio, 2022	Proveniência de objetos artísticos e culturais das instituições culturais GLAM (Galerys, Libraries, Archive and Museuns)	Melhorar a transparência na documentação da proveniência de objetos artísticos, propondo o uso de NP para representar a subjetividade e a incerteza e permitindo o rastreamento da proveniência dos dados.
Golden, Patrick; Shaw, Ryan, 2016	A aplicação de tecnologias de dados nas humanidades	Explorar como as NPs podem ser aplicadas às humanidades para realizar a representação formal de

Autores	Contexto	Motivação do uso de NP
		definições de períodos históricos mantendo a integridade semântica e o contexto discursivo das ciências humanas.
Maiatsky, Michail et al, 2018	Humanidades digitais abordando a complexidade da cultura filológica	A possibilidade de representação e integração do conhecimento acadêmico na área de humanidades digitais, de maneira estruturada e padronizada do conhecimento.
Asif, Imran et al, 2021	Publicações científicas nas ciências sociais, onde métodos como revisões sistemáticas e meta-análises são comumente utilizados para sintetizar conhecimentos.	A Nanopublicação pode melhorar a capacidade de detectar, comparar e explicar afirmações contraditórias em estudos científicos de forma automatizada, facilitando a comparação desses resultados de pesquisa, as vezes oriundos de meta-análise ou revisões sistemáticas.

Fonte: Autoria própria (2024).

5.3.4 Resultados

A revisão de literatura foi realizada, com estudos compreendidos entre 2011 e 2023. Podemos perceber que a aplicação da NP em outras áreas do conhecimento ocorre a partir de 2014, anteriormente o foco era voltado para as áreas das Ciências da vida, berço do conceito de Nanopublicação.

Com base nos estudos analisados, percebeu-se que o crescente volume de dados, em todas as áreas do conhecimento, é um dos motivadores para manter-se os dados controlados, interoperáveis e integrados, por meio de critérios FAIR e com a Nanopublicação como representação do objeto digital.

As NPs surgiram como uma solução inovadora no campo das Ciências da Vida, fornecendo uma maneira de encapsular declarações científicas de forma granular e estruturada, em proposta apresentada por Groth, Gibson e Velterop (2010). Os autores propõem um modelo inicial de NP com uma instância de formato baseado em ontologias.

Percebendo a necessidade de expandir o conceito para ser amplamente utilizado, Tobias Kuhn e colaboradores (2013) descrevem a AIDA (Atomic, Independent, Declarative, Absolut), uma Nanopublicação ampliada, que oferece maior flexibilidade e aplicabilidade a uma variedade de resultados científicos.

Posteriormente, pensando em tornar o processo de publicação de dados vinculados mais acessível, Kuhn et al. (2021) apresentam uma abordagem para usar NPs em uma rede descentralizada de serviços, “Nanobech”²³, onde os usuários podem publicar diretamente pequenas instruções de linked data, isso amplia o uso de NPs para outras áreas do conhecimento e para não especialistas.

Segundo Bucur e colaboradores (2022), NPs foram propostas para expressar conhecimento científico em linked data, como pequenos pacotes de publicações independentes, permitindo proveniência e metadados ricos (Bucur et al., 2022, p.4),

Expressar conhecimento científico ou não em pequenos pacotes de publicações independentes, motivou as áreas do conhecimento tanto das ciências da vida, com foco nos seus dados científicos que poderiam ser combinados e reutilizados, como mais recentemente nas ciências sociais e humanidades, que buscaram o uso de nanopublicações pensando na proveniência dos dados, conforme citam Golden e Shaw (2016, p. 1):

Nanopublicações incluem representações úteis e utilizáveis da proveniência de afirmações estruturadas. Essas representações de proveniência são úteis, pois permitem que os consumidores dos dados publicados façam conexões com outras fontes de informação sobre o contexto desses dados.

Podemos perceber nos estudos que a aplicação da NP em qualquer domínio do conhecimento é possível, pois a própria estrutura da NP permite essa adaptação.

No estudo de Van der Horst *et al.* (2015), a metodologia aplicada pode servir de base conceitual para a integração de dados de áreas como Ciências Sociais, Ciências Ambientais, para a integração de dados climáticos, ou seja, aplicada a campos onde exista a necessidade de análise de dados complexos e variados.

²³ “Nanobech” teve o nome revisado para “Nanodash” e pode ser consultado no endereço: <https://nanodash.petapico.org/>

A revisão de literatura realizada destaca que a NP tem se firmado como uma abordagem robusta para a representação de dados, promovendo uma estrutura mais formalizada, interoperável e reutilizável dos dados científicos visando superar as dificuldades em relação à quantidade de resultados desses dados (Fabris, 2019; Groth; Gibson; Velterop, 2010; Kuhn *et al.*, 2013).

O conceito de NP tem sido adotado por vários campos científicos, levando à criação de mais de dez milhões de nanopublicações disponíveis (Kuhn *et al.*, 2018), no entanto, ainda observamos uma lacuna na aplicação de NPs em áreas das Ciências Exatas e da Terra, como a Engenharia, onde a complexidade e o volume de dados sinalizam carência na gestão dos dados.

A complexidade da documentação de engenharia e a sua real disponibilização é um problema causado por grandes conjuntos de dados heterogêneos que são frequentemente produzidos para simulações ou experimentos (Samad; Osman; Nayan, 2023).

Visando estudar uma alternativa para a recuperação da informação de documentos de engenharia, com mais segurança e condições de reusabilidade, propomos a seguir um modelo conceitual de Nanopublicação, adaptado para documentos de engenharia, explorando a viabilidade e benefícios na estruturação, compartilhamento e reuso dessas informações técnicas.

6 APLICAÇÃO DO CONCEITO DE NANOPUBLICAÇÃO PARA DOCUMENTOS DE ENGENHARIA

A proposta desta seção é avaliar se o conceito de Nanopublicação pode ser aplicado de maneira prática para documentos de Engenharia, atendendo assim a um dos objetivos específicos dessa pesquisa.

A Nanopublicação é uma publicação semântica de nível granular que contempla três partes principais: uma declaração, uma proveniência associada e metadados de descrição. Essas partes são representadas em triplas RDF (Groth; Gibson; Velterop, 2010; Kuhn et al., 2013; Kuhn; Dumontier, 2015).

Asif, Tiddi e Gray (2021) no artigo “Using Nanopublications to detect and explain contradictory research claims” discorrem sobre o problema de detectar informações conflitantes nos resultados de pesquisa e propõem usar as NPs como estrutura padrão para representar essas alegações nas pesquisas e usar os dados de proveniência como um indicador da fonte de contradições entre as diferentes afirmações.

Apesar do estudo dos autores ter foco em Ciências Sociais, podemos ampliar e utilizar como exemplo o uso nos documentos de Engenharia e considerar uma base para estabelecer os conflitos ou diferenças entre as diversas revisões de um documento técnico, por exemplo, facilitando ao especialista a comparação de uma alteração entre desenhos técnicos e estabelecendo um indicador de qualidade do projeto com base nesses dados.

Essa unidade mínima de informação ou conhecimento estruturado, com contexto, pode encapsular diversos tipos de dados ou informações. Como mapeado na seção anterior, ela pode ser aplicada a diversas áreas do conhecimento, para melhorar a sua rastreabilidade, interoperabilidade e possibilidade de reuso, como ao universo da documentação de engenharia, com sua complexidade e dados de diferentes fontes, que pode ser beneficiado com esse conceito de representação de objetos digitais.

A documentação técnica de engenharia é um registro detalhado e sistemático de todas as fases de um projeto, da sua concepção até a conclusão. Quintana *et al.* (2010, p. 497) citam o seguinte: O objetivo fundamental de um documento de engenharia é transportar, controlar e manter a definição de um produto de forma precisa e clara, sem risco de má interpretação. Os desenhos técnicos fornecem um meio de comunicar a complexidade de modo compreensível e eficiente.

Madsen e Madsen (2017, p. 46) descrevem que os desenhos de engenharia comunicam uma variedade de conceitos, como requisitos de engenharia e instruções, e que um desenho de

engenharia ou um conjunto completo de desenhos fornece todos os dados necessários para fabricar ou construir um equipamento ou produto.

O documento técnico de engenharia é caracterizado por ter informações referentes ao escopo no qual o documento será emitido, ou seja, são informações representativas, e emitidas conforme normas e padrões.

Visando estabelecer uma relação entre o documento técnico de engenharia e o conceito de NP, avaliando assim a possibilidade de uso do conceito nas informações técnicas de Engenharia, foi realizado um estudo de caso em documentos emitidos na fase do projeto executivo, para uma Subestação de obras na Petrobras.

Para realizar o estudo de caso, inicialmente iremos caracterizar documentos técnicos de engenharia da Petrobras, sendo o contexto organizacional do campo de estudo.

6.1 DOCUMENTOS TÉCNICOS DE ENGENHARIA: CARACTERÍSTICAS

Conforme perspectiva arquivística, documentos são vistos como provas das atividades e situações que ajudaram a criar, modificar, eliminar ou manter. Trazem nas características dos registros a imparcialidade, autenticidade, naturalidade, inter-relacionamento e unicidade, além de atuarem como provas, pois capturam fatos, suas causas e consequências, preservando a memória (Duranti, 1994).

Documento técnico é qualquer documento que defina um projeto, suporte ou processo de fabricação de um produto. São caracterizados como documentos de projeto, fabricação e montagem e abarcam diversas tipologias como desenhos, especificações, listas e manuais (Watts, 2008).

A documentação técnica de engenharia é amparada por normas ABNT e normas ISO, que definem responsabilidades, formatos, estruturas e conteúdo, caracterizando assim um padrão que serve como guia para a emissão dos documentos técnicos.

No âmbito da Petrobras, a documentação técnica de engenharia é emitida nas fases dos projetos conceituais, básicos e executivos, que além das normas já citadas acima, atende também as Normas Técnicas Petrobras, coordenadas pelo Comitê de Normas da Petrobras – CONTEC/NORTEC.

Para a elaboração de projetos, as normas são emitidas pela Subcomissão 12 – Normas Gerais de Projeto, conforme listadas abaixo:

- N-0058 Símbolos gráficos para fluxogramas de processo e de engenharia
- N-0075 Abreviaturas para projetos industriais
- N-0381 Formulários para emissão de documento técnico de engenharia

- N-1521 Identificação de equipamentos industriais
- N-1674 Projetos de arranjo de instalações industriais terrestres de petróleo
- N-1710 Codificação de documentos técnicos de engenharia
- N-1913 Preparação de requisição de material
- N-2064 Emissão e revisão de documentos de projeto
- N-2065 Elaboração de informações básicas de empreendimentos

Para o nosso estudo, a norma basilar será a N-1710, por ser a norma que codifica os documentos, e é por meio dessa codificação que recuperamos os documentos emitidos, na maioria das nossas buscas.

A norma N-1710 “visa uniformizar e sistematizar a codificação de documentos técnicos de engenharia emitidos em papel ou meio eletrônico relativos às instalações da PETROBRAS para permitir o seu arquivamento ordenado e facilitar a recuperação de informações” (Petrobras, 2014).

O código é formado por um “conjunto alfanumérico que identifica o documento, sendo obtido pela associação ordenada dos códigos representativos dos diversos grupos básicos que compõem o conjunto” (Petrobras, 2014).

Esta norma cita 5 grupos para a composição do código do documento, são eles (Figura 13):

- Identificação do Idioma (Grupo 0);
- Anexo A - Categoria (Grupo 1);
- Anexo B - Identificação da Instalação (Grupo 2);
- Anexos C ou E - Área de Atividade (Grupo 3);
- Anexos D ou F - Classe de Serviço, Materiais e Equipamentos (Grupo 4);
- Código de Origem do Documento (Grupo 5);
- Cronológico (Grupo 6).

Figura 13 - Composição do código do documento

	Anexo A	Anexo B	Anexo C (ou E)	Anexo D (ou F)		
Grupo 0	Grupo 1	Grupo 2	Grupo 3	Grupo 4	Grupo 5	Grupo 6
Identificação do idioma	Categoria do documento	Identificação da instalação	Áreas de atividades	Classe de serviços equipamentos e materiais	Origem do documento	Cronológico
L	- LL	- ABBB.CC	- ABBBB	- AAA	- LLL	- AAA

Fonte: Petrobras (2014).

Para a análise dos documentos que iremos trabalhar, teremos as seguintes composições para a formação do código:

Grupo 2: Identificação da instalação: 5305.00 – Refinaria Premium I

Grupo 5: Código de origem: CB3 – Consórcio SIEMAR

Exemplo de código, consoante os 5 grupos citados:

DE-5305.00-8000-700-CB3-001

6.2 PROJETOS NA PETROBRAS E GESTÃO DE DOCUMENTOS TÉCNICOS

A Petrobras é uma empresa de economia mista especializada na indústria de óleo, gás natural e energia. Atua nas áreas de exploração e produção, refino, geração e comercialização de energia (Petrobras, 2025).

No seu organograma contempla uma gerência responsável pelo macroprocesso “implementar sistemas de superfície, gás e energia” para garantir a implantação de empreendimentos do refino, gás e energia, através da elaboração do projeto de engenharia, construção, montagem e comissionamento.

A gestão da documentação técnica integra o processo de gerenciamento de projetos, sendo uma das suas atividades a administração do SIGEM – Sistema Integrado de gerenciamento de empreendimentos, o GED/EDMS da Petrobras, para armazenar, controlar, definir fluxos, gerenciar revisões e garantir a segurança dos dados.

A documentação é cadastrada pelo código N-1710 e a sua ordenação no GED/EDMS é feita por estruturas de pastas configuradas, conforme necessidade do projeto (Figura 14).

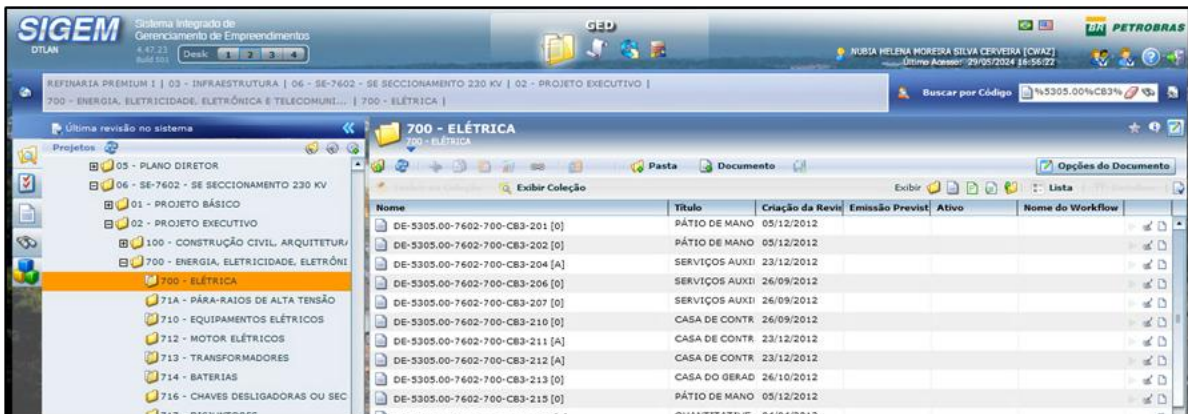
Figura 14 - Tela do SIGEM apresentando a estrutura do projeto



Fonte: Tela capturada pela autora (2024).

A busca dos documentos no SIGEM pode ser feita pelo código ou navegando na estrutura de pastas, conforme figura 15.

Figura 15 - Tela do SIGEM com visão dos documentos



Fonte: Tela capturada pela autora (2024).

Os documentos cadastrados têm os seguintes campos de metadados preenchidos para manter o histórico do documento e a sua recuperação (Figura 16):

Figura 16 - Tela dos metadados do documento

Propriedades
DE-5305.00-7602-946-CB3-201

Informações | Atributos de Tipo | GA_190 - ARQUITETURA | Outros

Dados Básicos | Tags | Linhas | Palavras-Chave

Código do Documento: DE-5305.00-7602-946-CB3-201

Revisão *: 0

Título: SETOR 230kV - DIAGRAMA UNIFILAR GERAL SIMPLIFICADO

Status de Documento *: Não Liberado

Classificação do Documento *: Projeto

Substituído por: []

Substitui: []

Situação de Documento: []

Fonte: Tela capturada pela autora (2024).

6.3 ESTUDO DE CASO

Será analisado um *corpus* de trinta e dois (32) documentos para caracterizar partes de um documento técnico e verificar a aplicabilidade do conceito de NP. Essa documentação é da fase do projeto executivo para a construção e montagem de uma subestação de obras.

No quadro 10 estão os tipos de documentos, identificados por tipo e disciplina, incluídos no estudo de caso.

Quadro 10 – Documentos para análise

Categoria dos documentos/Disciplinas	Quantidade
DE	23
100 - CONSTRUÇÃO CIVIL, ARQUITETURA, URBANISMO	10
700 - ENERGIA, ELETRICIDADE, ELETRÔNICA E TELECOMUNICAÇÕES	8
900 - SERVIÇOS E CUSTOS DIVERSOS	5
ET	2
700 - ENERGIA, ELETRICIDADE, ELETRÔNICA E TELECOMUNICAÇÕES	2
LI	3
700 - ENERGIA, ELETRICIDADE, ELETRÔNICA E TELECOMUNICAÇÕES	3
MC	3
100 - CONSTRUÇÃO CIVIL, ARQUITETURA, URBANISMO	2
700 - ENERGIA, ELETRICIDADE, ELETRÔNICA E TELECOMUNICAÇÕES	1
RL	1
700 - ENERGIA, ELETRICIDADE, ELETRÔNICA E TELECOMUNICAÇÕES	1
Total Geral	32

Fonte: Autoria própria (2024).

Seguindo com o detalhamento, os documentos selecionados para exemplificar esse conceito estão classificados segundo no Grupo 1, categoria do documento, conforme a Norma Petrobras N-1710 e apresentados no Quadro 11.

Quadro 11 – Categoria de documentos

Categoria do documento	Descrição	Definição do tipo documental e exemplos de uso da categoria do documento
DE	Desenho	Documento que possui informações técnicas representadas graficamente conforme normas e/ou padrões aplicáveis. Exemplos: Plantas de tubulação, de situação, de arranjo geral, diagrama, fluxogramas, anteprojeto, plantas de arranjo, diagrama unifilar.
ET	Especificação Técnica	Documento com os requisitos de qualidade, critérios de projeto, especificação de processos, códigos e rotinas quanto à mão-de-obra, materiais, sistemas, equipamentos, restrições da instalação, e serviços a serem realizados no projeto de construção, na operação, manutenção, inspeção ou descomissionamento de uma instalação. Exemplos: Informações Básicas de Empreendimentos; Critérios de Projeto,

Categoria do documento	Descrição	Definição do tipo documental e exemplos de uso da categoria do documento
		Fabricação e Montagem; Especificação de Materiais; Especificação de Sistemas; Especificação de equipamentos não padronizados.
LI	Lisa	Documento que apresenta informações em colunas e linhas e contém a relação de equipamentos, de instrumentos ou de materiais, necessários à aquisição, fabricação ou montagem de instalações em um determinado projeto. Exemplos: Lista de cabos, de instrumentos, de material, de equipamentos etc.
MC	Memória de Cálculo	Documento que contém instruções, procedimentos, rotinas e demais informações necessárias à operação e manutenção de uma instalação, inclusive em situações de resposta à emergência. Exemplos: Manual de operação, de manutenção, de equipamento, de instrumentação, de embarcação, de garantia da qualidade.
RL	Relatório	Documento que apresenta um conjunto de informações utilizadas para reportar resultados parciais ou totais de uma determinada atividade, experimento, projeto, ação, pesquisa, ou outro evento concluído ou ainda em andamento. Exemplos: Relatório de teste hidrostático, de avaliação de integridade, de acompanhamento de atividades, de levantamentos de campo e de avaliações técnicas.

Fonte: Petrobras (2024).

A análise detalhada dos documentos irá permitir avaliar, teoricamente, como a NP pode ser aplicada para representar documentos digitais de engenharia de maneira granular e verificável.

O escopo dos documentos selecionados para o estudo de caso se refere à elaboração de uma subestação de obras, que segue normas ABNT/NBR específicas da especialidade de elétrica.

No tema geral “Subestação de obras”, iremos trabalhar com os conceitos que descrevem os documentos técnicos. Esse conceito irá ajudar a definir sobre o que a NP irá representar.

Não foi utilizado nenhum *software* de análise de texto, pois a maioria dos documentos é composta por desenhos técnicos e não havia disponibilidade de *software* que realizasse a captura de dados em documento com múltiplas informações.

Realizamos a descrição dos documentos, identificando as informações relevantes para a recuperação e representação da informação do documento de engenharia.

A seguir, apresentaremos a extração dos dados separados por grupos temáticos de disciplinas nos quadros 12, 13 e 14, respectivamente.

Quadro 12 – Temática Engenharia Civil

Tipo Documental	Conceitos	Contexto
Fundações	Sapata, Forma, armação, subestação	Detalha as especificações e requisitos técnicos para a fundação do tipo "sapata" em uma subestação, incluindo dimensões, materiais e normas
Estruturas	Subestação, Cobertura, formas, Sala de painéis, Tag PN-8921310	Detalha as dimensões e detalhes da estrutura de cobertura da subestação, incluindo sala de painéis
Estruturas	Subestação, estrutura , Tag BCP-8921; Tag TF-8921610	Documento detalha as dimensões da subestação contemplando as estruturas de vigas
Estruturas	Subestação, Laje, Piso, Armação	Documento detalha a armação inferior e superior do radier, com detalhes da armadura da malha de base e relação da ação
Estruturas	Subestação, Fundação, cintamento, armação	Documento detalha e especifica a relação do aço das vigas com dimensões e peso total aplicado
Estruturas	Subestação, Fundação, cintamento, armação	Documento detalha e especifica a relação do aço das vigas com dimensões e peso total aplicado
Plot Plan - Arranjo Geral	Casa de painéis, equipamentos	Apresenta a distribuição dos equipamentos na locação da subestação, com tags e referências
Memória de Cálculo de estrutura	Casa de painéis, fundações, memória de cálculo	Descreve os critérios técnicos e memorial justificativo de cálculo das fundações do tipo sapata projetadas de concreto armado da Casa de Painéis de 13,8 kV

Memória de Cálculo de estrutura	Casa de painéis, concreto, estrutura	Descreve critérios técnicos e memorial justificativo de cálculo das estruturas de concreto armado da Casa de Painéis de 13,8 kV,
---------------------------------	--------------------------------------	--

Fonte: Autoria própria (2024).

Quadro 13 - Temática Engenharia Elétrica

Tipo Documental	Conceitos	Contexto
Planta Geral de Elétrica	Setor 13,8Kv, Planta geral, Planta Se de obras, Grupo gerador de diesel, Tag 7602701A/B, PN-8921401,TF-8921610	Detalha a planta geral do setor 13,8Kv
Planta de Gerador	Planta SE 13,8Kv , GRUPO GERADOR DE DIESEL	Detalha a planta de distribuição da SE 13,8Kv, Serviços gerais para a subestação
Diagrama Elétrico	Subestação, Diagrama, Serviços auxiliares	Diagrama esquemático da subestação de obras com detalhes , referências e cargas
Diagrama Elétrico	Casa de Painéis, Subestação, quadro de iluminação, Tag PLE-8921801A/B, Tag PN-8921701	Diagrama trifilar com disposição de tomadas, quadro de iluminação e tomadas
Diagrama Elétrico	Casa de Painéis, iluminação, tomadas	Detalhes para iluminação e tomadas da casa de painéis para a subestação de obras
Planta de Aterramento	Pátio, subestação, aterramento	Documento apresenta a malha de aterramento do pátio da subestação de obras, com dimensões e documentos de referência
Diagrama Elétrico	Subestação, Diagrama	Detalha o diagrama unifilar simplificado com referências e orientações sobre atendimento a NOS e legendas
Diagrama Elétrico	Subestação, diagrama, 220/127v	Apresenta o diagrama unifilar da subestação de obras para 220/127 volts
Equipamentos	Subestação, especificação técnica, transformador trifásico, 13800-380/220 V, NBR 5034,5051,5280,5356	Documento especifica as características e requisitos básicos de fornecimento e instalação de transformadores auxiliares
Equipamentos	Subestação, especificação técnica, carregador retificador 380Vca-125 Vcc, 13800-380/220 V, NBR 5180,5356,5361	Documento especifica os requisitos para fornecimento e instalação de carregadores retificadores

Lista de Malhas	Subestação, Malha de aterramento	Detalha a lista de materiais da malha de aterramento da subestação de obra
Lista de Equipamentos	Subestação, Equipamentos elétricos, SE-8921001	Detalha a lista dos equipamentos elétricos principais da subestação de obras 13,8 kV, SE-8921001
Relatório de Cabo condutor	Cabo condutor	Cálculos do esticamento do cabo condutor representado em tabela

Fonte: Autoria própria (2024).

Quadro 14 – Temática Arquitetura e Urbanismo

Tipo Documental	Conceitos	Contexto
Planta baixa	Casa de painéis da Subestação, Arquitetura, Tag PN-8921310	Detalha seções da planta de arquitetura da casa de painéis, descrevendo materiais e especificações
Planta Baixa	Casa de painéis da Subestação, Arquitetura, Tag PN-8921311	Detalha seções da planta de arquitetura da casa de painéis, descrevendo materiais, dimensões e especificações
Planta Baixa	Casa de painéis da Subestação, Arquitetura, viga	Detalha seções típicas de vigas e canaletas, contém lista de materiais e descrição

Fonte: Autoria própria (2024).

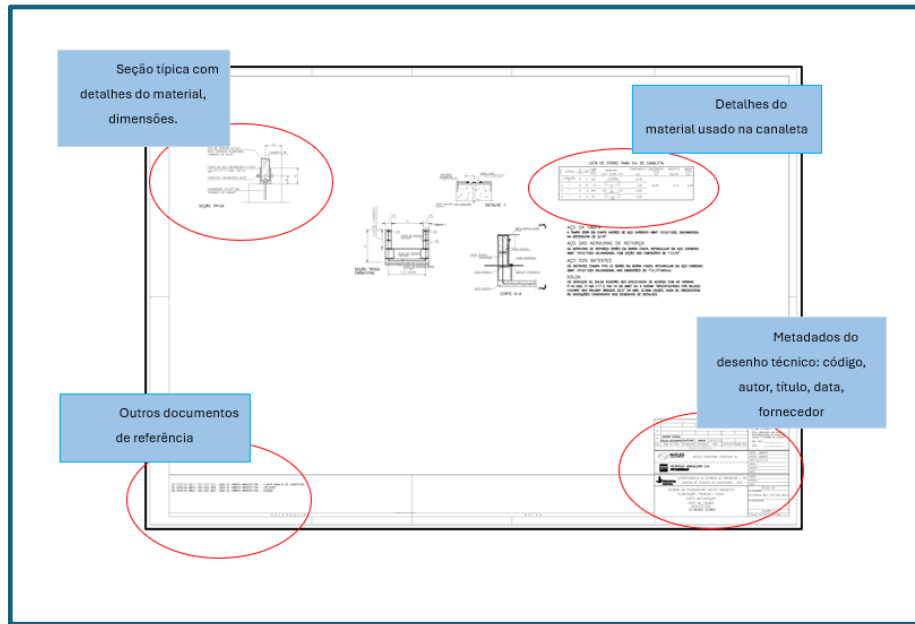
6.4 MODELO DE NANOPUBLICAÇÃO PARA DOCUMENTO DE ENGENHARIA

O *corpus* analisado corresponde a um projeto que resultou em trinta e dois (32) documentos únicos, cujas informações se complementam. Portanto, um único documento não transmite a mensagem completa, pois ele faz referências a outros documentos, citações de normas e de documentos complementares, como representado na figura 17.

A necessidade de se obter informações dos diversos documentos, em um projeto de engenharia, se depara com o desafio de representação dessas informações. Como discutido e apresentado durante o desenvolvimento desta pesquisa, os documentos de um projeto de engenharia são interdependentes e são gerados por diversas fontes.

A emissão dessas informações, como NP, facilitaria a interoperabilidade e a rastreabilidade das versões. Considerando que documentos de projetos possuem diversas revisões, enquanto estão sendo elaborados, seria facilitado o próprio acompanhamento e monitoramento de revisões, estabelecendo uma base de conhecimento que poderia ser reutilizada em outros projetos.

Figura 17 - Planta com detalhes gerais da Casa de Painéis



Fonte: Adaptado de Petrobras (2024).

O documento da figura 17 é uma Planta (Desenho técnico) com detalhes da casa de Painéis, ela está classificada como “Planta Baixa” no conjunto de documentos e possui as seguintes informações:

- ✓ Seções típicas com detalhes de materiais
- ✓ Detalhes do material da canaleta
- ✓ Documentos de referência
- ✓ Título
- ✓ Código
- ✓ Nome do fornecedor
- ✓ Nome do autor
- ✓ Data

Das informações listadas acima, poderíamos usar como exemplo de emissão de uma NP a descrição conceitual utilizada para descrever o documento (Quadro 15):

Quadro 15 – Planta Geral de Elétrica

Tipo Documental	Conceitos	Contexto
Planta baixa	Casa de painéis da Subestação, Arquitetura, Tag PN-8921310	Detalha seções da planta de arquitetura da casa de painéis, descrevendo materiais e especificações

Fonte: Autoria própria (2024).

A informação do documento técnico poderia ser representada da seguinte maneira como NP:

- Asserção: “O TAG PN-8921401 está representado na Planta Baixa de Arquitetura”
- Proveniência: DE-5305.00-8921-190-CB3-206
- Info NP: Criado por Núbia Cerveira, 30/10/2024

O *Tag* citado acima é uma identificação do equipamento. É um *código* que o equipamento recebe durante a elaboração do projeto, para que ele tenha uma identificação na unidade operacional. Na Petrobras, há uma norma específica que controla esse processo, a N-1521 – Identificação de equipamentos.

Ao emitir uma NP com essa asserção/declaração, construiremos uma rede dos documentos de fornecimento do painel elétrico, dos documentos de projeto e dos documentos de construção e montagem, todos poderão ser acessados a partir de um único ponto e poderão ser reutilizados e readequados nos projetos, conforme necessidade da revisão.

Essas “reutilização” e “centralização” das informações técnicas, tornam o conceito de Nanopublicação eficiente e possível de ser aplicado aos documentos técnicos de engenharia.

A proposta de um modelo de Nanopublicação passa pela construção de uma ontologia, com base no *corpus* de documentos do estudo de caso e apresenta uma especificação, descrevendo um formato conceitual de *Nanopublication* para documento de engenharia.

6.4.1 Ontologia dos Documentos de Engenharia Subestação

A criação de uma ontologia para a elaboração da NP é importante, pois irá fornecer um modelo formal, com a descrição dos conceitos para descrever as informações, facilitando a integração dos dados heterogêneos.

Groth, Gibson e Velterop (2010) enfatizam essa importância de utilizar ontologias, já existentes ou não, e adaptando-se quando necessário, para facilitar a interoperabilidade, pensando em como as nanopublicações podem ser estruturadas e enriquecidas com metadados. Kuhn (2014) e Clark, (2014) corroboram com esse pensamento, com foco na integração de dados e na rastreabilidade das afirmações científicas, pois uma ontologia bem definida permite essa rastreabilidade, aumentando assim a confiabilidade das informações.

Para nosso trabalho, iremos criar uma ontologia para representar os documentos técnicos de uma subestação, relativos ao estudo de caso, com a finalidade de relacionar os conceitos e estruturar as informações dos documentos. Optou-se por desenvolver a ontologia

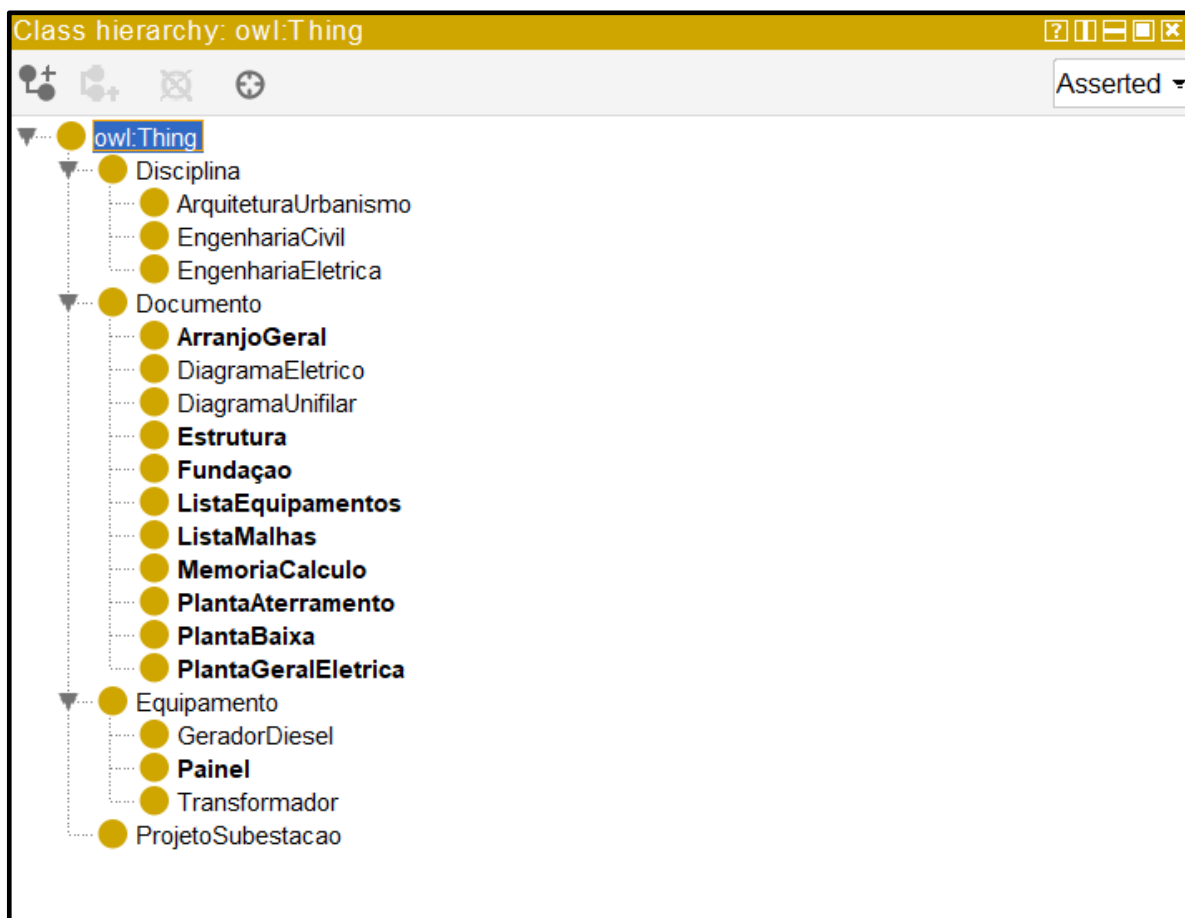
sabendo-se que a proposta aqui apresentada é conceitual, pois não foi aplicada em campo, para testes.

Ontologias são especificações formais explícitas de conceitualizações compartilhadas de um domínio de interesse (Gruber, 1995). Ela é formal, pois é expressa em uma linguagem de representação de conhecimento legível por máquinas; é explícita, pois o conhecimento está descrito em conceitos, propriedades e relações; e compartilhada, por refletir uma concordância entre pessoas a respeito do domínio especificado.

Para a representação dos documentos produzidos no estudo de caso, foi utilizada a ferramenta *Protégé*²⁴, que é um editor de ontologias gratuito, de código aberto e uma estrutura para a construção de sistemas inteligentes.

A definição das classes da ontologia foi feita com base nos quadros 5, 6 e 7 dos grupos temáticos de disciplina, contextualizados para explicitar os documentos do estudo de caso. Os conceitos identificados foram organizados em classes e subclasses conforme a figura 18 abaixo:

Figura 18 - Classes do Estudo de caso Subestação

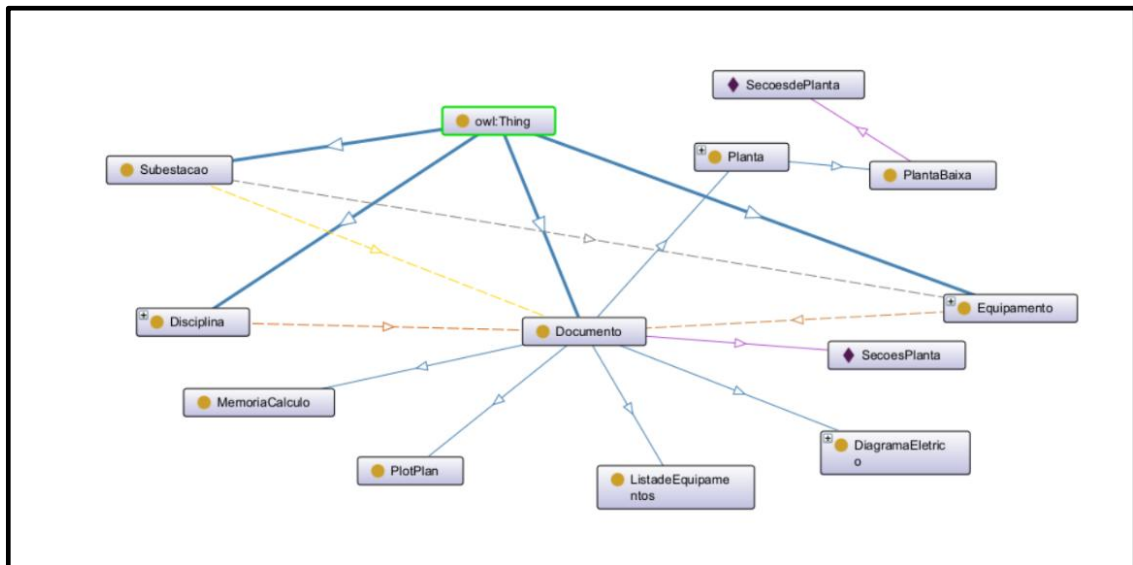


Fonte: Autoria própria (2024).

²⁴ <https://protege.stanford.edu/>

Na figura 19, temos a representação desses conceitos com seus relacionamentos.

Figura 19 - Relacionamentos na modelagem



Fonte: Autoria própria (2024).

6.4.2 Especificação da Nanopublicação para Documento de Engenharia

No contexto da Engenharia, a disseminação e o compartilhamento de conhecimento e informações são fundamentais para o avanço dos projetos.

A documentação técnica de engenharia é extensa e diversificada, dificultando sua busca, compartilhamento e acesso. Nanopublicações podem fragmentar informações complexas em unidades menores e mais gerenciáveis.

Uma Nanopublicação é composta por *Asserção*, *Proveniências* e *Metadados* da própria NP. Com base nessa estrutura, especificaremos uma NP para documentos técnicos de engenharia. A seguir descreveremos conceitualmente um modelo de Nanopublicação para tais documentos, ressaltando que ele não foi testado.

1º) Identificação do escopo a ser emitido como NP

Seleção das informações a serem publicadas como Nanopublicações, identificando toda a sequência de documentos técnicos a serem emitidos em um determinado projeto, por disciplina. O conteúdo selecionado deve conter informações sobre dimensões, materiais, processos de fabricação, normas a serem utilizadas, locação de algum equipamento e demais informações técnicas necessárias para a execução do projeto.

Considerando o estudo citado na subseção 6.3, no quadro 16, abaixo, apresentamos os documentos emitidos para o projeto executivo, para a temática de Engenharia Elétrica, considerando a emissão por códigos dos documentos.

Quadro 16 – Documentos selecionados para extração de informações

Documento	Título
DE-5305.00-7602-700-CB3-201	PÁTIO DE MANOBRA - DUTOS E CANALETAS - PLANTA GERAL
DE-5305.00-7602-700-CB3-202	PÁTIO DE MANOBRA - DUTOS E CANALETAS - DETALHES
DE-5305.00-7602-700-CB3-204	SERVIÇOS AUXILIARES - PLANTA GERAL
DE-5305.00-7602-700-CB3-212	CASA DE CONTROLE - ILUMINAÇÃO E TOMADAS - PLANTA E DETALHES
DE-5305.00-7602-700-CB3-213	CASA DO GERADOR DIESEL - ILUMINAÇÃO E TOMADA - PLANTA
DE-5305.00-7602-781-CB3-201	CASA DE CONTROLE - SPDA - PLANTA, CORTES, VISTA E DETALHES
DE-5305.00-7602-946-CB3-203	SETOR 230kV - SERVIÇOS AUXILIARES - DIAGRAMA UNIFILAR GERAL
DE-5305.00-7602-946-CB3-204	SETOR 230kV - SERVIÇOS AUXILIARES - DIAGRAMA UNIFILAR - 125Vcc
DE-5305.00-7602-946-CB3-205	SETOR 230kV - SERVIÇOS AUXILIARES - DIAGRAMA UNIFILAR - 0,40kV
DE-5305.00-7602-946-CB3-206	SETOR 230kV - SERVIÇOS AUXILIARES - DIAGRAMA UNIFILAR - 220/127V
ET-5305.00-7602-714-CB3-201	ESPECIFICAÇÃO TÉCNICA - CARREGADOR RETIFICADOR 380 Vca - 125 Vcc
ET-5305.00-7602-714-CB3-202	ESPECIFICAÇÃO TÉCNICA - BATERIA DE ACUMULADORES - 125 Vcc
LI-5305.00-7602-700-CB3-202	PÁTIO DE MANOBRA - DUTOS E CANALETAS - LISTA DE MATERIAIS
LI-5305.00-7602-700-CB3-208	SETOR 230kV - CASA DE CONTROLE - ILUMINAÇÃO E TOMADAS - LISTA DE MATERIAL
LI-5305.00-7602-700-CB3-211	LISTA DE EQUIPAMENTOS ELÉTRICOS
LI-5305.00-7602-781-CB3-201	CASA DE CONTROLE - SPDA - LISTA DE MATERIAL
MC-5305.00-7602-700-CB3-211	MC - CASA DE CONTROLE - ILUMINAÇÃO E TOMADAS

Fonte: Autoria própria (2024).

2ª) Estruturação da informação que será transformada em NP

Os dados e informações selecionados ou identificados devem ser estruturados em pequenos blocos, cada um com uma afirmação específica e verificável. No Quadro 17, abaixo, apresentamos uma amostragem desses pequenos blocos de informações para alguns documentos do quadro 16.

Quadro 17 – Extração de informações dos documentos técnicos

Documento selecionado	Conteúdo
DE-5305.00-7602-700-CB3-201	O croqui representa a planta baixa das fundações da subestação, conforme norma NBR 6492
	A caixa de passagem PES-000-45415 é especificada para passagem de cabos elétricos de média tensão e atende à norma IEC 61439
	A caixa de passagem tipo CAC-0809-CP1 a CP26, com altura H=1000 mm, é projetada para instalações elétricas subterrâneas conforme NBR 15751
	A canaleta reforçada PES-000-45402 possui resistência mecânica conforme norma IEC 61537 e está especificada para suportar cargas de até 200 N/ PES-000-45402
DE-5305.00-7602-700-CB3-202	O detalhamento dos dutos e canaletas inclui o Transformador de Potencial (TP) de 230KV, conforme a norma IEC 61869-5
	O detalhamento dos dutos e canaletas inclui o Transformador de Corrente (TC) de 230KV, projetado para medição e proteção conforme IEC 61869-2
	A Seção DAL S/LT faz parte do detalhamento dos dutos e canaletas, sendo projetada para abrigar cabos de distribuição elétrica em média tensão.
	A Seção DALC/LT do detalhamento dos dutos e canaletas é projetada para organizar e proteger os cabos de controle da subestação
	O detalhamento dos dutos e canaletas inclui o Disjuntor Tripolar de 230KV, especificado conforme IEC 62271-100 para interrupção de corrente em alta tensão
DE-5305.00-7602-700-CB3-204	Planta Geral dos Serviços Auxiliares
	Planta da subestação 13,8 Kv
	Apresenta a localização do PN 13,8KV
	Apresenta a localização do Transformador TF-7602701A/B com dimensões em milímetro
	Apresenta locação do Grupo Gerador de Diesel com dimensões em milímetro
	Sala de painéis com os PN-892110, CB-8921801A, CB-8921801B
DE-5305.00-7602-700-CB3-212	O detalhamento do sistema de iluminação e tomadas da casa de controle foi especificado conforme a norma NBR 5410, garantindo segurança e eficiência energética na distribuição elétrica
	O quadro de iluminação e tomadas foi especificado conforme NBR 5410
	Detalhamento da casa de controle com iluminação e tomadas conforme norma
	A planta apresenta diagramas e tabelas para a segurança da subestação
	A planta apresenta um sistema de aterramento para segurança da subestação
	O detalhamento do diagrama trifilar do PL-7602701 conforme norma para a casa de controle
DE-5305.00-7602-700-CB3-213	Planta de cortes e detalhes da casa do grupo gerador
	A planta apresenta um quadro principal com cargas e barramentos para a casa do grupo gerador
	A planta apresenta quadro com o número de circuitos da casa do grupo gerador
DE-5305.00-7602-946-CB3-203	A potência do transformador TF-7602701A/B são 13,8/0,4 Kv
	A potência do gerador GE-7602701 é 400v
	O banco de baterias da BT-7602801A está alimentando o CB-7602801A
	O banco de baterias da BT-7602801B está alimentando o CB-7602801B
	O painel de corrente contínua está interligado com os painéis de proteção e controle
Os painéis de proteção e controle possuem cargas de 125 Vcc	

Fonte: Autoria própria (2024).

3ª) Identificação de Metadados

Cada Nanopublicação deve ser acompanhada de metadados descrevendo sua origem, autoria, o contexto e a validade, para facilitar a localização, interoperabilidade e reutilização da NP por

outros profissionais. Para a identificação dos metadados dos documentos de engenharia iremos utilizar o Dublin Core²⁵, devido a sua flexibilidade (Quadro 18).

Quadro 18 - Dublin Core – Para a descrição do documento de engenharia

Dublin Core	Descrição	Exemplo
identifier	Código do documento. Identificador único do documento	"DE-5305.00-7602-700-CB3-213=0"
title	Título do documento	"Planta baixa da subestação"
creator	Engenheiro responsável	"Eng ^a Helena Cerveira"
date	Data de criação	"2025-02-01"
format	Formato do arquivo do documento	"dwg"
language	Idioma do documento	"pt-BR"
subject	Especialidade	"Engenharia Elétrica - Projetos de Subestação"
coverage	Local espacial	"Subestação Obra, 230 kV"
relation	Relação com outros documentos	"Relacionado ao projeto civil SE-8901"
source	Fonte do documento	"SIGEM"
publisher	Empresa responsável pelo projeto	"XPTO Engenharia"
rights	Níveis de acesso ao documento	"Interno"

Fonte: Autoria própria (2024).

4ª Arquitetura da Nanopublicação para documentos de engenharia

Os documentos de engenharia podem ser modelados como nanopublicações contendo metadados sobre seu conteúdo, autoria, versão e contexto de aplicação.

Com base no estudo de caso da subseção 6.3, os principais tipos e formatos de documentos incluem:

- a) **Engenharia elétrica:** Diagramas unifilares, trifilares, distribuição e de aterramento, Especificações técnicas, listas de materiais e diagramas de proteção.
- b) **Engenharia civil:** Plantas, cortes, fundação e memórias de cálculos estruturais.

Para transformar documentos de engenharia em nanopublicações, deverá ser utilizado o modelo de nanopublicação, que é composto por grafos RDF (Groth; Gibson; Velterop, 2009; Gray *et al.*, 2024; Kuhn *et al.*, 2013).

- **Afirmação/Asserção (Assertion):** Representa a unidade mínima de informação técnica de um determinado documento de engenharia, normas técnicas

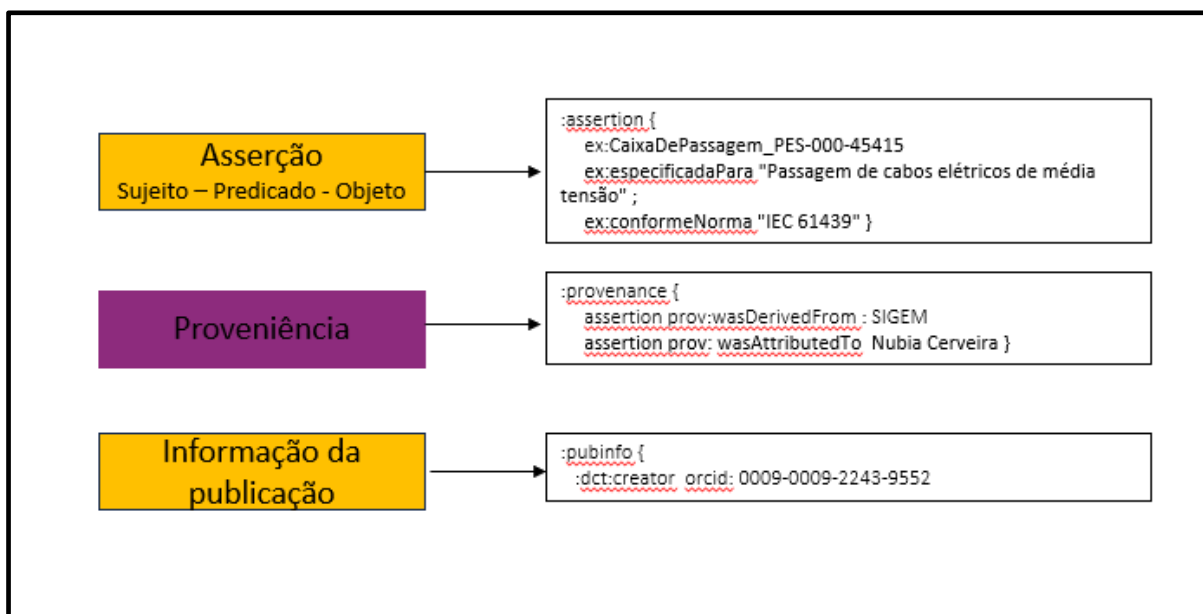
²⁵ www.dublincore.org/specifications/dublin-core/dcmi-terms/#http://purl.org/dc/elements/1.1/publisher

associadas e relações com outros documentos. Contém a informação principal que precisa ser declarada.

- Proveniência (Provenance): Fornece detalhes sobre a origem dos dados usados para derivar a afirmação, como qual repositório de origem da afirmação, link para a norma usada para fundamentar a afirmação etc. Aqui teremos a origem e a evidência da informação técnica.
- Metadados da Publicação (Publication Info): Contém informações sobre a própria nanopublicação, como autor, data e versão.

A figura 20 apresenta essa arquitetura com exemplo das asserções retiradas de um documento técnico:

Figura 20 - Exemplo de arquitetura a ser adotada para documentação técnica



Fonte: Autoria própria (2024).

5ª Ontologias sugeridas

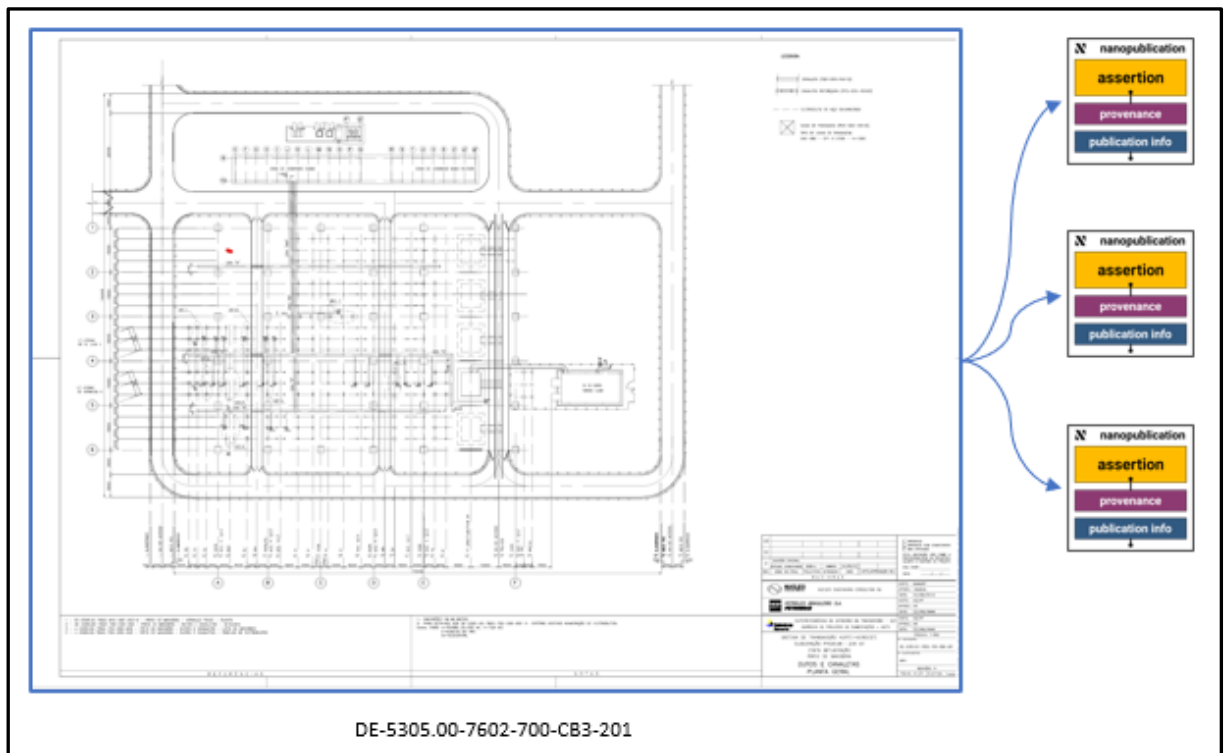
- ISO 15926: utilizada para integrar dados de ciclo de vida em plantas de processo, incluindo as de produção de petróleo e gás, pode ser usada para configurar um vocabulário próprio para projetos ou operações. Utiliza tecnologias como: Web Ontology Language (WOL), Resource Description Framework (RDF), Generalized Markup Language (GML) (Klüwer; Skjæveland; Valen-Sendstad, 2008).
- PROV-O – Provenance Ontology: fornece um conjunto de classes, propriedades e restrições que são utilizadas para representar informações de proveniência geradas em diferentes sistemas e contextos. Utiliza a linguagem OWL2 (Groth; Moreau, 2013).

6ª Modelagem das nanopublicações:

Na figura abaixo temos um documento que deu origem a três (3) declarações/afirmações sobre uma planta da subestação de obras.

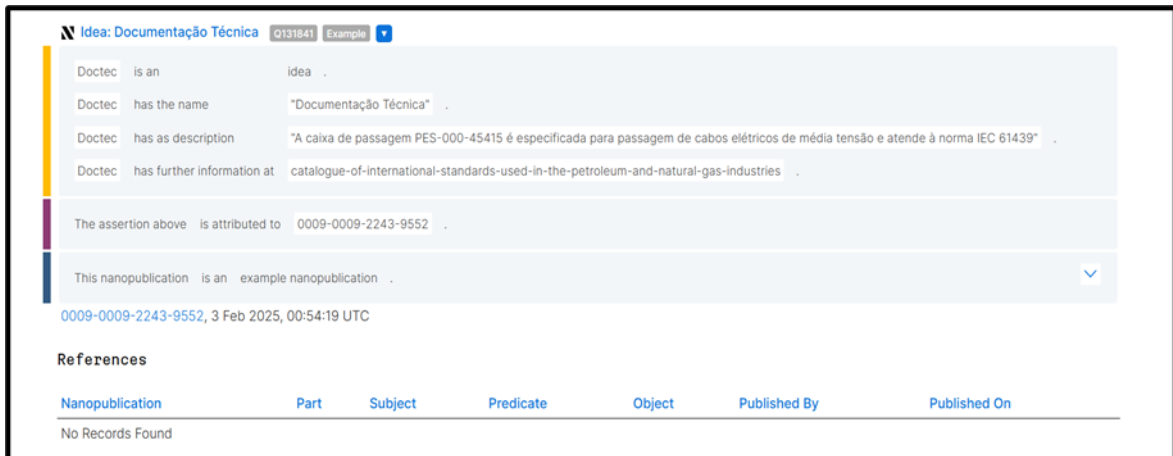
A modelagem foi realizada com base nas declarações derivadas desse documento técnico, conforme figura 21 abaixo.

Figura 21 - Nanopublicações a partir de um documento técnico



Fonte: Autoria própria (2024).

Na sequência apresentamos uma modelagem feita no Nanodash, considerando o template existente na plataforma (Figura 22).

Figura 22 - Nanopublicação - Caixa de passagem

Fonte: Nanodash²⁶

O objetivo da pesquisa é responder, por meio do levantamento bibliográfico e pelo estudo de caso, que é possível utilizar documentação técnica e emitir pequenas fragmentos dessas informações como Nanopublicação, esta publicação traria maior rastreabilidade, interoperabilidade e confiabilidade, permitindo integração em diversos sistemas e facilitando a gestão da documentação técnica.

O uso da Nanopublicação para documentos técnicos de engenharia trará uma redução na ambiguidade que é própria dos documentos de engenharia, devido a sua heterogeneidade de informações e dados.

No entanto, se faz necessário pontuar algumas desvantagens, como por exemplo a necessidade de validação e certificação, que precisa ter em alguns documentos de engenharia. O uso de nanopublicações, nesse caso, poderia invalidar ou dificultar, devido a exigência de validação em cada nanopublicação.

Apesar da desvantagem pontuada, o uso de NP para documentos de engenharia, ainda segue sendo uma alternativa inovadora para a gestão documental, com vista à recuperação da informação de maneira rápida e segura.

Em complemento a uma Nanopublicação de documentos de engenharia, sugere-se a implantação de uma política de curadoria digital para documentos técnicos de engenharia, tornando assim a documentação confiável, como tem que ser quando é destinada a operação de unidades operacionais, como na Petrobras.

²⁶ <https://nanodash.knowledgepixels.com/>

7 MODELO DE CURADORIA DIGITAL APLICADO AOS DOCUMENTOS DE ENGENHARIA

Esta seção apresenta uma proposta de Curadoria Digital, com foco na documentação técnica digital de engenharia, apresentando uma descrição conceitual para atender ao objetivo específico (OE3) desta pesquisa.

A Curadoria digital busca gerenciar os dados digitais agregando valor e garantindo sua preservação, integridade, interoperabilidade e acessibilidade a longo prazo. O foco vai além da preservação, contemplando também a criação, descrição, acesso e reuso dos objetos digitais (Constantopoulos *et al.*, 2009; Silva; Siebra, 2017; Yakel, 2007).

A documentação produzida e tratada na engenharia abrange diversos aspectos, tais como: linhas geométricas por meio de um desenho técnico; dimensões em plantas de arranjo ou desenhos dimensionais; modelos para análise de elementos; processos de cálculos e modelos conceituais; além de modelos relacionados aos processos de fabricação, processos de comissionamento e construção e montagem.

Durante a execução de um projeto de engenharia percebe-se o grande desafio de realizar a tramitação de informações e documentos de forma eficiente, tanto internamente quanto externamente, com fornecedores, empresas contratadas, clientes e órgãos reguladores durante o ciclo de vida do projeto e durante a vida útil do ativo. Nesse sentido, é fundamental que tais informações sejam preservadas e mantidas compreensíveis ao longo do tempo, atendendo a exigências legais e garantindo a capacidade de atualização dos projetos.

No centro do que os engenheiros fazem estão os dados e informações de documentos técnicos, enquanto isso a maior parte desses dados está disponível em formatos não padronizados e desorganizados.

Somando-se a toda complexidade do projeto, todas as partes interessadas precisam de formas eficazes de trabalhar com dados e informações que estão disponíveis em milhares de tipos de arquivos e formatos.

Pschybilla e Homann (2020) corroboram com esse ponto de vista quando tratam sobre o ciclo de vida de um processo na engenharia mecânica. Os autores citam o desperdício de informações e a baixa qualidade dos dados, o que acarreta um atraso no projeto desde sua criação. A informação organizada otimiza os processos de ponta a ponta, focando em fluxos de informações enxutas e assertivas.

Nesse contexto, a curadoria abrange todas as atividades necessárias para assegurar que os dados permaneçam adequados para uso durante todo o ciclo. Isso inclui a manutenção,

preservação e organização de informações e documentos técnicos, de modo a possibilitar sua reutilização e interpretação contínuas.

Para exemplificar, iremos descrever o fluxo de gerenciamento da documentação técnica de engenharia, e na sequência, iremos descrever a proposição de curadoria.

O fluxo detalhado é fruto de experiência profissional da autora, portanto seu contexto de aplicação é a atuação como profissional de informação e documentação na Petrobras.

7.1 FLUXO DE GERENCIAMENTO DA DOCUMENTAÇÃO TÉCNICA

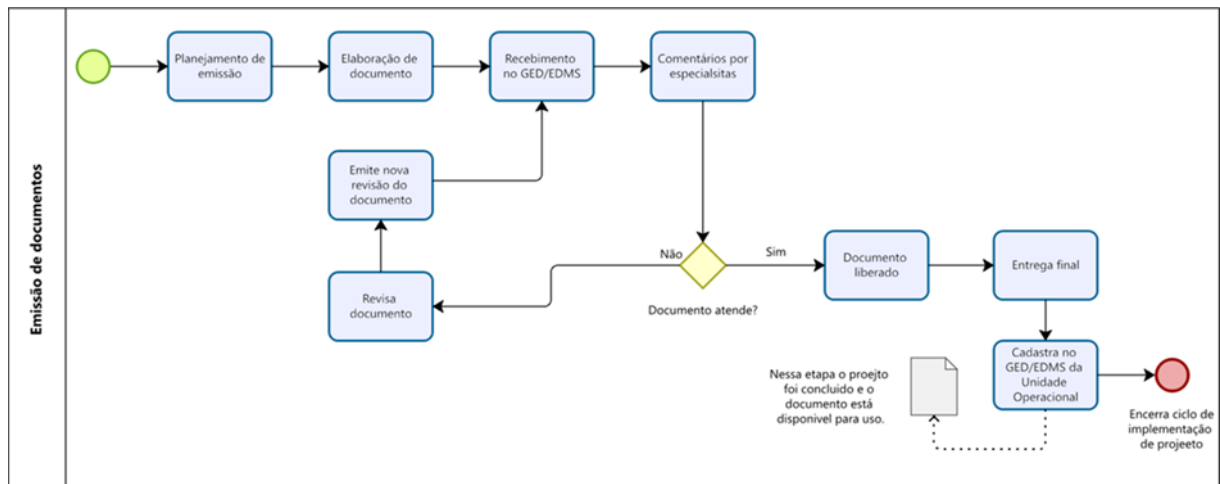
O processo de gerenciamento da documentação técnica de engenharia tem como base a emissão dos documentos, durante a execução de um projeto.

O fluxo de emissão de documentos utiliza práticas de gerenciamentos de projetos como o PMBOK, por exemplo, inicia na fase de planejamento onde ficam distribuídos as especialidades e os tipos documentais produzidos por cada uma, nessa etapa cria-se o diagrama de procedência de emissão de documentos, definido geralmente com base no escopo do projeto contratado.

A documentação emitida pela projetista, passa pelo processo de qualidade e possui os responsáveis pela execução, verificação e aprovação, na sequência é enviada para o Sistema de gerenciamento eletrônico de documentos, que no caso da Petrobras é o SIGEM, nessa etapa do GED/EDMS, conforme fluxo definido internamente, os documentos são comentados e analisados pelo corpo de especialistas das disciplinas técnicas, quando o documento é liberado para execução e após sua emissão final, conforme construído, esse mesmo documento é transferido para outro sistema de gerenciamento de documentos, a depender da área de negócio da Petrobras.

Nessa etapa de ser gerenciado pelo sistema da área de negócio, o documento é de caráter permanente e utilizado para a operação da unidade. As demais revisões de projeto ficam no SIGEM como memória técnica.

O fluxograma abaixo apresenta as etapas de emissão de documentos de engenharia desde a concepção na projetista, seguindo do envio para análise dos especialistas e retornando para implementação de comentários ou não, seguindo seu fluxo até a emissão final e entrega para a unidade operacional.

Figura 23 - Fluxograma do processo de emissão de documentos

Fonte: Autoria própria (2024).

O papel do gerenciamento da documentação técnica em projetos de engenharia é fundamental, pois desde o início participa do fluxo criando estruturas, definindo papéis, informando sobre normas e qualidade do que deve ser emitido, enfim, é essencial para que a documentação emitida ou recebida esteja o mais completa possível, garantindo assim a sua rastreabilidade e disponibilidade.

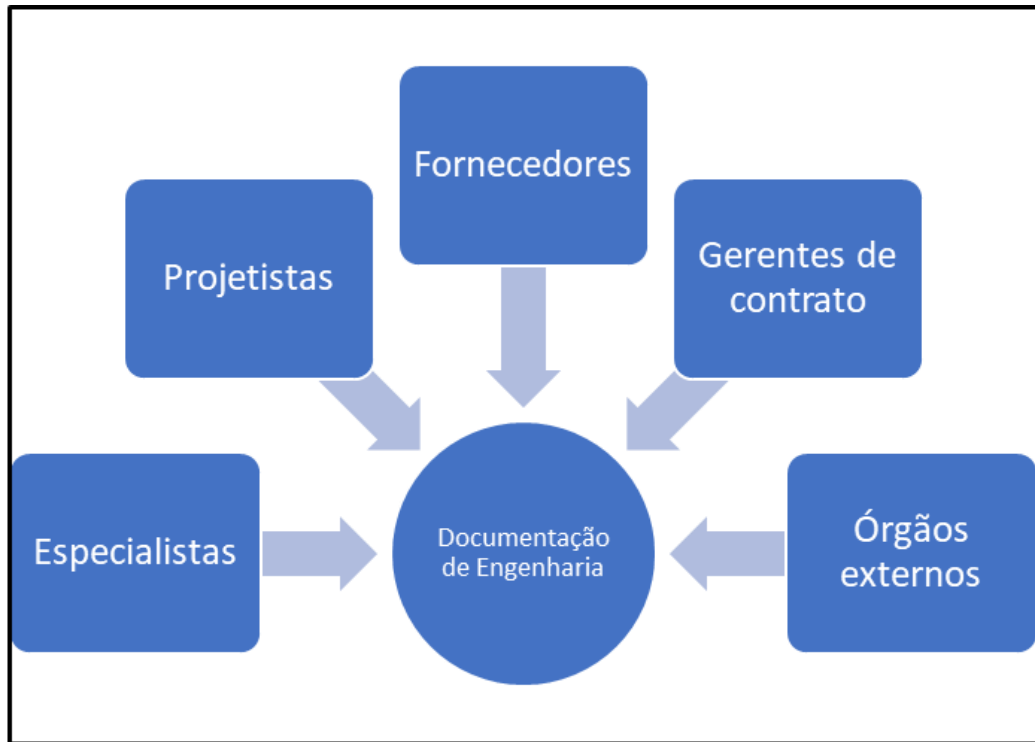
Alguns estudos tratam sobre essa preocupação com as práticas de gerenciamento de projetos que não envolvem a gestão documental, como citado abaixo:

Argumentamos que as práticas existentes de gerenciamento de projetos subestimam as inter-relações entre as tarefas individuais e outros componentes do projeto. Isso deixa as interdependências mal gerenciadas, promovendo um único evento que prejudica a busca por melhorias contínuas. Percebemos a necessidade de existir uma abordagem unificada para o gerenciamento de projetos. Nessa abordagem, é dada grande ênfase à forma como os gestores organizam e estruturam as informações do projeto e suas interdependências (Froese, 2010, p.534).

A figura 24 representa as partes envolvidas na emissão de documentos técnicos de engenharia, durante a execução de um projeto, todos são produtores de informações que precisam estar disponíveis e precisam de acompanhamento para garantir a sua confiabilidade e integridade, bem como sua disponibilidade e acessibilidade nos diversos sistemas de gerenciamento de obras, como sistemas de planejamento – MsProject²⁷, Primavera²⁸ ou sistemas de GED/EDMS, como o SIGEM, no caso da Petrobras.

²⁷ Software de gerenciamento de projetos da Microsoft para planejamento, execução e acompanhamento de projetos

²⁸ Software de gerenciamento de projetos da Oracle para planejamento, avaliação e acompanhamento de projetos

Figura 24 - Partes interessadas

Fonte: Autoria própria (2024).

O gerenciamento da documentação técnica inicia muito antes da assinatura de um contrato para a execução do projeto ou pedido de compra para aquisição de um equipamento. A gestão inicia estabelecendo critérios a partir dos anexos contratuais e está intrinsicamente ligada ao gerenciamento do projeto.

A preparação para acompanhar a execução de um projeto na gestão da documentação técnica, envolve alguns domínios de gerenciamento, como abaixo:

a) Planejamento

Segundo o *PMBOK*® “o objetivo do planejamento é desenvolver proativamente uma abordagem para criar entregas do projeto” (Project Management Institute, 2021, p.129).

O processo de planejar o gerenciamento da documentação inicia desde o levantamento dos requisitos necessários, identificação e partes interessadas externas e internas, que vão influenciar na entrega de um produto.

Esse planejamento no gerenciamento da documentação técnica deve estabelecer critérios, rotinas e procedimentos que definam como será a estruturação do gerenciamento da documentação, desde o início do projeto até a entrega final para o cliente.

Portanto, todos os itens citados devem estar descritos em um plano de gestão da documentação que deverá contemplar os seguintes tópicos:

- Classificação de documentos

Definir que tipos de documentos serão gerenciados no projeto, como documentos de contrato, documentos de gestão do projeto, planejamento, engenharia, comissionamento, suprimentos. Essa definição é necessária para que sejam mapeados os sistemas que serão utilizados e outros critérios.

- Codificação da documentação

Definir critérios de codificação de documentos, técnicos, administrativos ou de gerenciamento, visando estabelecer uma uniformidade nas informações recebidas para que elas sejam de fácil recuperação.

- Emissão de documentos

Definir como será a emissão dessa documentação, em qual sistema será gerenciada.

- Controle da documentação

Estabelecer listas de documentos recebidos e enviados, para que seja possível acompanhar o planejamento da emissão dos documentos.

- Distribuição da documentação

Estabelecer fluxo de comunicação e prazo de tramitação.

- Arquivamento da documentação

Definir como será o sistema de armazenamento dessa documentação.

- Matriz de responsabilidades

Estabelecer matriz com os envolvidos no processo definindo atividades e responsabilidades a serem executadas durante a execução do projeto.

b) Monitoramento e controle

O monitoramento e controle da documentação técnica emitida e recebida se dá através de:

- Reuniões de acompanhamento

Participar das reuniões para acompanhamento do projeto juntamente com os responsáveis pelo mesmo, gerentes de projeto, fiscais de contrato, projetistas, fornecedores, e estabelecer periodicamente reuniões específicas de documentação para tratativas e melhorias do processo.

- Monitoramento diário

Controle diário dos documentos emitidos, através da verificação dos itens de qualidade que devem ser atendidos, como padrão de numeração, formatação, emissão de acordo com a LD, fluxo de informação definido anteriormente etc. As inconsistências

encontradas são monitoradas através de relatório e devem ser enviadas periodicamente para o responsável pela emissão do projeto.

- **Relatórios**

Verificação da quantidade e qualidade de documentos emitidos por *status* mensalmente pelas projetistas e fornecedores, analisando as disciplinas com maior índice de reprovação e falta de consolidação. Essa análise serve de insumo para o relatório mensal do empreendimento e para averiguar o período de maior emissão de documentos.

- **Auditorias de qualidade**

Auditoria na qualidade do projeto desenvolvido através da aplicação de Listas de Verificação dos requisitos do Projeto, com tópicos específicos de Documentação Técnica.

c) Encerramento

No processo de gestão da documentação técnica, o encerramento está diretamente ligado à transferência da documentação para o cliente final, para que seja possível fazer a operação da planta industrial ou uso do equipamento ou produto.

Para que essa documentação técnica de engenharia esteja finalizada e concluída para entrega, ela deverá estar com o status de “Conforme construído” na documentação de engenharia, projeto e status de “Certificado” na documentação de fornecedor.

Essa documentação encerrada é um dos produtos do projeto e será transferida para o cliente como obrigatoriedade para aceite das instalações, conforme fluxograma citado na figura 23.

7.2 CARACTERÍSTICAS DOS MODELOS EXISTENTES

Existem na literatura alguns modelos que orientam o gerenciamento, a preservação e a acessibilidade dos objetos digitais. Cada modelo tem suas características, que podem influenciar sua aplicabilidade em diferentes contextos.

A seguir, relatamos alguns desses modelos de curadoria digital, destacando suas características e implicações para a prática a ser sugerida para a documentação técnica de engenharia. São eles: a) DCC – Digital Curation Lifecycle Model; b) o DCC&U – Extended digital curation lifecycle model; c) UK Data Archive Data Lifecycle; d) DataOne Data Lifecycle.

a) DCC - Digital Curation Lifecycle Model

O modelo realiza uma abordagem do ciclo de vida onde garante que todos os estágios necessários sejam identificados e planejados, e as ações necessárias implementadas, na sequência correta. Isso pode garantir a manutenção da autenticidade, confiabilidade, integridade e usabilidade do material digital, o que, por sua vez, garante a maximização do investimento em sua criação (Higgins, 2008).

A variedade de sistemas que modelam ou elaboram projetos é um desafio a ser enfrentado na engenharia, e modelos como o DCC, com foco na preservação e que possui adaptabilidade e interoperabilidade, possibilita o uso na documentação técnica de engenharia, pois tem aplicação em diversos domínios.

b) DCC&U - Extended Digital Curation Lifecycle Model

O modelo de ciclo de vida proposto por Constantopoulos *et al.* (2009) destina-se a apoiar o planejamento e a organização do gerenciamento de conteúdo digital, identificando atividades importantes que não foram incluídas no modelo DCC.

Como uma extensão ao modelo do DCC, sua característica principal é a consideração dos recursos de informação contextual como objeto de curadoria (Constantopoulos *et al.*, 2009).

A ênfase na usabilidade e na adição de valor através do contexto torna esse modelo particularmente interessante para a documentação técnica, considerando o objeto digital como nanopublicação, que traz consigo uma proveniência que precisa ser preservada, além das informações da própria NP.

c) UK Data Archive Data Lifecycle:

Este modelo foca na gestão eficiente de dados de pesquisa para reutilização por outros pesquisadores. Sua estrutura simples e adaptável pode ser útil para a curadoria da documentação técnica, especialmente em ambientes colaborativos e de revisões entre disciplinas.

A documentação técnica é frequentemente reutilizada em novos projetos, o que exige organização e sistemas de busca eficientes. O estágio de reutilização no modelo da UK Data Archive é importante para otimizar o trabalho dos engenheiros.

d) DataONE - Dataone Data Lifecycle

O modelo DataONE visa garantir a qualidade, descrição, preservação, publicação, descoberta e reutilização de dados digitais.

O modelo possui oito (8) etapas para gerenciamento de dados: planejar, coletar, garantir/assegurar a qualidade dos dados, descrever os dados, preservar, descobrir dados úteis, integrar esses dados e analisar

Anjos e Dias (2019) contextualizam que

É válido ressaltar que as etapas componentes do Ciclo de Vida dos Dados - DataONE não são dependentes entre si, e não são completamente obrigatórias para sua otimização. As etapas podem e devem ser realizadas de acordo com as necessidades de cada pesquisador, comunidade científica, entre outros (Anjos; Dias, 2019, p.88).

Os modelos são voltados para curadoria digital de dados de pesquisa, e para a proposição de um modelo de curadoria para documentos de engenharia, mesmo que conceitual, será necessário ajustar às necessidades específicas dessa documentação, a saber:

✓ **Proveniência**

É preciso registrar a origem dos documentos, o contexto de sua criação, os autores e as equipes envolvidas. Isso garante a rastreabilidade e a compreensão da evolução do projeto.

✓ **Preservação**

Pelo viés arquivístico, a documentação técnica é classificada como permanente, ou enquanto o ativo estiver em operação, portanto, a documentação técnica precisa ser preservada ultrapassando o ciclo de vida do projeto, considerando inclusive a sua característica de memória técnica de uma organização.

✓ **Reutilização**

Existe a necessidade de reutilização da documentação em novos projetos, o que ocorre de modo não sistematizado devido às várias fontes de origem de emissão de documentos e sistemas.

✓ **Padronização**

A necessidade de padronizar metadados e processos é crucial para a interoperabilidade e compartilhamento da documentação técnica com foco nas normas específicas adotadas na Engenharia.

✓ **Acessibilidade**

Em um contexto industrial, o compartilhamento de informações é definido dentro de um ambiente de trabalho baseado em projetos e contratos, que é gerenciado e projetado para proteger os interesses comerciais, por esse motivo a informação gerada para projetos de engenharia deve ter níveis de acesso, que devem ser mapeados e refletidos no plano de curadoria.

Para a proposta nesta pesquisa, iremos tratar o objeto digital e seu ciclo, políticas de gestão da informação e promoção de incentivos para a colaboração, fomentando assim uma cultura de compartilhamento e uso dos dados de engenharia existentes, mantendo os níveis de acessibilidade.

De acordo com Sayão (2001, p.83), “[,] um modelo é uma criação cultural, um “mentefato”, destinado a representar uma realidade, ou alguns dos seus aspectos, a fim de torná-los descritíveis qualitativa e quantitativamente e, algumas vezes, observáveis”.

Com base na descrição citada por Sayão, este modelo pretende ser uma representação do processo de curadoria, com vistas a torná-lo aplicável na engenharia.

O modelo que está sendo proposto foi desenvolvido utilizando como base o *DCC Curation Lifecycle Model* devido à sua adaptabilidade a diferentes domínios, mas não representará a totalidade do domínio da Engenharia e não contemplará todas as ações do DCC.

7.3 CURADORIA DIGITAL PARA DOCUMENTOS TÉCNICOS DE ENGENHARIA

Em projetos de engenharia, assim como na indústria de construção, a representação de informações de um ambiente construído por meio de desenhos bidimensionais, perspectivas, cálculos de engenharia ou especificações técnicas nos leva ao desafio de interoperabilidade, que pode ser mais bem gerenciado se existir uma política de curadoria (Neilson, 2009, p. 11).

A proposta é uma visão geral das ações necessárias para uma curadoria digital que seja possível aplicar aos documentos de engenharia, contemplando as ações para todo o ciclo de vida da curadoria, ações sequenciais e ações ocasionais no contexto da Engenharia e da representação da informação digital de seus documentos.

Pensando em um cenário de ciclo de vida de projetos, é vital a curadoria desses dados e documentos para promover a reutilização.

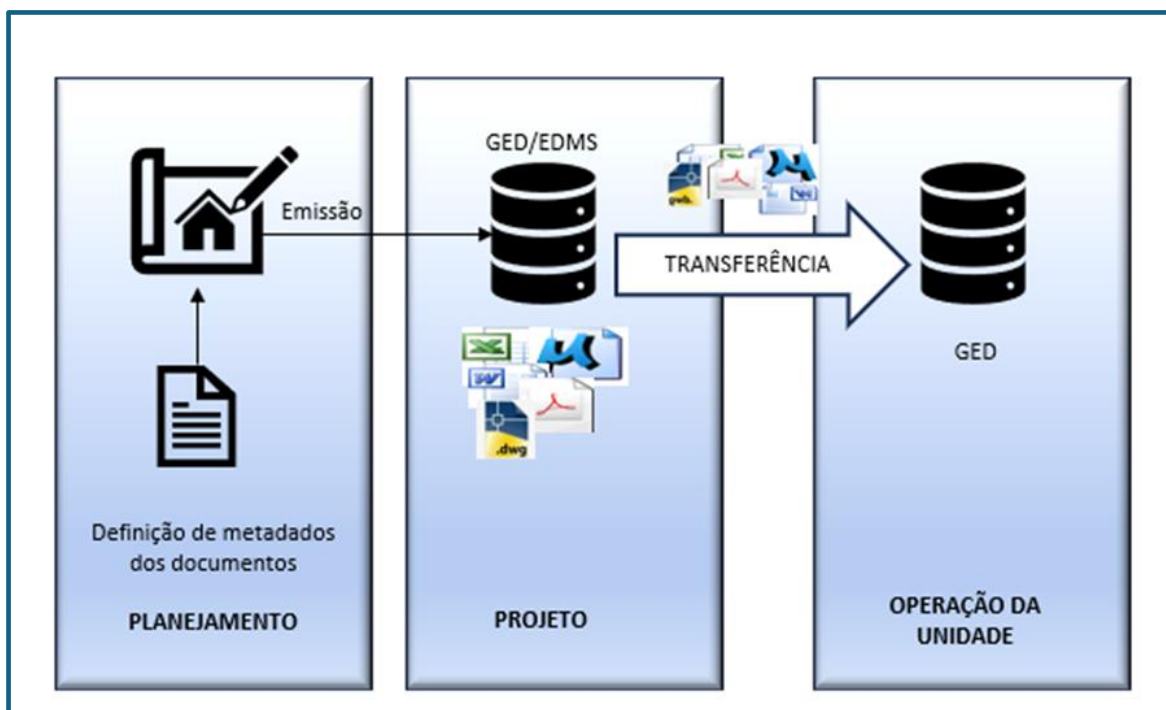
Documentos de engenharia, como modelos CAD 3D, dados de simulação, e resultados de experimentos, são essenciais para o progresso da disciplina (Smith, 2009; Esteva, Janse e Balandrano, 2018). No entanto, esses dados frequentemente se perdem devido à obsolescência tecnológica, fragilidade das mídias e falta de políticas de preservação adequadas.

A implementação de um projeto de curadoria digital é essencial para:

- Reduzir a duplicação de esforços na criação de dados.
- Garantir a acessibilidade e reutilização de dados de pesquisa.
- Preservar o contexto e a integridade dos dados ao longo do tempo.

A figura abaixo representa o processo de gestão da documentação técnica realizado durante o desenvolvimento de um projeto, sugerimos que seja desenvolvido uma política de curadoria digital que contemple desde o planejamento inicial de emissão dos documentos por empresas projetistas externas ou internas.

Figura 25 - Gerenciamento da documentação técnica



Fonte: Adaptado do Procedimento da Petrobras (2024).

7.3.1 Conceitualização

É a etapa onde os fundamentos e diretrizes gerais para a criação, organização e preservação dos documentos devem ser definidos. Essa etapa estabelece o propósito, o escopo e os padrões que guiarão todas as fases subsequentes.

De acordo com o DCC, a conceitualização concebe e planeja a criação de objetos digitais, incluindo métodos de captura de dados e opções de armazenamento.²⁹

McMahon, Giess e Culley (2005) discutem a importância do gerenciamento de informações para suporte ao produto durante toda a vida útil, enfatizando a necessidade de sistemas para dar suporte à introdução do produto e ao suporte durante toda a vida útil. Isso inclui armazenar dados de gerenciamento do ciclo de vida do produto (PLM) para acesso de

²⁹ <https://www.dcc.ac.uk/>

longo prazo e capturar a lógica do projeto, resultados de decisão e informações do processo de projeto.

Da mesma forma, Fan e Zhao (2022) destacam a importância dos modelos de gerenciamento digital para desenhos de projetos de energia, com foco no uso da tecnologia *blockchain* para abordar questões como versões inconsistentes e a necessidade de documentos impressas.

Nessa etapa, se faz necessário identificar os requisitos e práticas de governança para a preservação dos documentos de engenharia.

Alguns itens são importantes para desenvolver no planejamento:

- Definir os requisitos dos documentos de engenharia que devem ser preservados
- Identificar padrões e normas relevantes para a gestão documental
- Escolher os formatos que devem ser emitidos ou recebidos os documentos
- Estabelecer os metadados para garantir a rastreabilidade e recuperação futura
- No caso de nanopublicações, identificar o repositório onde serão cadastradas

Aqui deverá ser feita a definição dos objetivos, os metadados necessários e os critérios de avaliação. No contexto de engenharia, isso pode envolver a criação de um plano de gestão de informações que especifique os formatos de arquivo preferenciais, os padrões de nomenclatura, os vocabulários controlados e os sistemas de indexação a serem utilizados.

7.3.2 Criação e Recebimento

“Crie dados incluindo metadados administrativos, descritivos, estruturais e técnicos.”³⁰

As ações criar e receber correspondem à criação de objetos digitais e inclusão de metadados administrativos, descritivos, estruturais, técnicos e de preservação. Também consiste na recepção de dados e metadados coletados em outras ações do modelo (HIGGINS, 2008).

No domínio da documentação de engenharia, a etapa criar entende-se como emissão de documentos, e é a fase de criação do documento de engenharia conforme normas e padrões. Dentro de um fluxo de gestão documental essa etapa pode ser criar ou receber documentos de parceiros externos ou internos.

Essa documentação recebida ou criada deve contemplar metadados mínimos para que seja possível a recuperação e preservação desses objetos digitais, no caso de documento técnicos de engenharia, esses metadados são:

³⁰ <https://www.dcc.ac.uk/>

- Nome do emissor (autor, engenheiro responsável, equipe técnica, empresa de projeto);
- Código do documento (ID)
- Disciplina técnica;
- Data de criação e versão/revisão do documento;
- Formato do arquivo e software utilizado;
- Suporte físico e digital de armazenamento.

No caso de o objeto digital ser uma Nanopublicação, esses metadados já estarão identificados pela própria criação da NP e a NP já contém a proveniência da declaração ou informação técnica que foi modelada.

Os documentos técnicos de engenharia ou documentos de projeto de engenharia são emitidos em uma sequência lógica para a execução do projeto, para sua emissão são utilizados diversos sistemas, de cálculos e de modelagens que depois se consolidam em um tipo documental.

Para essa produção de documentos de engenharia se faz necessário mapear todos os tipos que serão emitidos durante o desenvolvimento do projeto visando viabilizar a interoperabilidade e a preservação desses arquivos.

7.3.3 Descrição e informação da representação

A descrição e informação da representação é uma das ações para todo ciclo de vida do *DCC Model Lifecycle* e aplicando aos documentos técnicos, o que será curado durante todo o ciclo de vida será o documento digital de engenharia.

Para os documentos de engenharia os metadados precisam ser atribuídos durante a emissão do documento, se faz necessário que os padrões de metadados estejam descritos no planejamento e nas políticas de curadoria, para garantir a descrição e o controle adequados a longo prazo.

No contexto de documentos digitais de engenharia, metadados devem incluir detalhes como a data de criação, o autor, a versão do documento, e as modificações feitas.

De acordo com Oliver e Harvey (2016), a chave para identificar, localizar, acessar, entender e usar dados e objetos digitais no futuro é adicionar informações sobre eles que forneçam um contexto e, se necessário, forneçam as ferramentas para usar esses dados.

A ação de descrever a informação da representação deve ocorrer em todos os pontos do ciclo de vida e na documentação de engenharia é fundamental pois a versão de uma ferramenta

de desenvolvimento de um documento pode em uma revisão pode ser uma e na revisão seguinte pode ser outra, e ter esse metadado descrito e controlado ajudará na preservação.

7.3.4 Plano de Preservação

O plano de preservação tem como objetivo garantir que os documentos digitais permaneçam acessíveis e íntegros ao longo do tempo. Para documentos técnicos de engenharia, é essencial considerar fatores como formatos de arquivo, segurança, integridade dos dados, acessibilidade e requisitos regulatórios.

No plano devem ser identificados os tipos de documentos técnicos a serem preservados, com consulta aos especialistas para identificar os documentos mais críticos e que poderão ser reutilizados em projetos futuros.

Devem ser descritos nessa fase os formatos de Arquivo com adoção de formatos que garantam longevidade e acessibilidade, como PDF/A, XML, DWG e formatos abertos recomendados por normativas técnicas e sempre com atenção aos níveis de acessibilidade dos documentos técnicos.

Os itens referentes a obsolescência com implementação de estratégias como migração de formatos e emulação para manter a acessibilidade de documentos elaborados com versões antigas de softwares devem constar no plano com esses tipos de arquivos.

Como plano a ser implementado na Curadoria digital, deve constar o monitoramento com revisões regulares para verificar a integridade dos documentos e assegurar que as estratégias de preservação estejam sendo seguidas.

A curadoria digital de documentos de engenharia deve envolver ações que garantam a preservação, acesso e reutilização das informações técnicas essenciais. A combinação de preservação digital, metadados padronizados e monitoramento constante fortalece a confiabilidade dos registros e promove a sustentabilidade do conhecimento técnico na engenharia.

A transformação dos documentos de engenharia em nanopublicações podem promover uma curadoria de fato, bem como preservação das informações técnicas essenciais, considerando que a modelagem utilizada para uma NP é baseada em RDF projetada para garantir interoperabilidade semântica e integração de dados entre diferentes sistemas e domínios do conhecimento.

8 CONCLUSÃO

A gestão de dados é um dos problemas da era da informação, sua diversidade de formatos e crescente volume nos encaminha cada vez mais a necessidade de padronizar os metadados e descrever os objetos digitais fortalecendo o seu domínio. Na engenharia esses dados estão distribuídos em vários sistemas, modelados em diversos softwares e organizados sob os mais variados padrões

A presente pesquisa explorou a aplicabilidade das nanopublicações na recuperação de informações contidas na documentação técnica de engenharia, propondo um modelo que visa melhorar a acessibilidade, rastreabilidade, interoperabilidade e reutilização das informações ao longo do ciclo de vida de projetos de engenharia.

Apesar dos avanços na engenharia digital, a gestão eficiente da documentação técnica ainda enfrenta desafios significativos, como a dispersão de dados e a ausência de padronização.

Nesse contexto, a nanopublicação, já utilizada em outros domínios do conhecimento além das ciências da vida, pode ser utilizada também na engenharia, com foco no documento técnico, permitindo a fragmentação e organização de dados em unidades mínimas de informação, facilitando sua recuperação e reuso.

Ao adotar as nanopublicações, torna-se possível atender aos princípios FAIR, promovendo um ambiente mais colaborativo e eficiente para a gestão da informação técnica.

A pesquisa evidenciou que a utilização de metadados ricos e a estruturação semântica das informações técnicas são fundamentais para superar as limitações dos sistemas tradicionais de recuperação da informação, especialmente em setores complexos e multidisciplinares, como a engenharia.

Portanto, conclui-se que a nanopublicação, ao integrar tecnologias da web semântica e conceitos de curadoria digital, pode transformar significativamente a forma como a informação técnica é gerida e recuperada, contribuindo para a eficiência dos projetos de engenharia e a qualidade das decisões tomadas com base nessas informações.

Sugere-se, como trabalhos futuros, a implementação prática deste modelo em sistemas de gerenciamento de documentos e a análise de seus impactos na produtividade e na mitigação de riscos em projetos de engenharia.

A Engenharia já trabalha somente com dados para algumas especialidades, a necessidade de emissão de documentos como representação da informação será superada por tecnologias que já existem como BIM - Building Information Modeling, Digital Twin entre

outras, mas é urgente que se discuta padrões de dados e metadados para essas tecnologias, para que essa explosão de dados no domínio da engenharia não fique dispersa.

REFERÊNCIAS

- ABBOTT, Daisy. **DCC Briefing Paper: What is digital curation?** 2008. Disponível em: <https://era.ed.ac.uk/handle/1842/3362>. Acesso em: 3 jun. 2024.
- AHMED-KRISTENSEN, Saeema; WALLACE, K. M. Identifying and supporting the knowledge needs of novice designers within the aerospace industry. **Journal of Engineering Design**, v. 15, n. 5, p. 475–492, out. 2004.
- ANJOS, Renata Lemos dos; DIAS, Guilherme Ataíde. Atuação dos profissionais da informação no Ciclo de Vida dos Dados–DATAONE: um estudo comparado. **Informação & informação**, v. 24, n. 1, p. 80-101, 2019.
- ARTHUR, W. Brian. **The nature of technology: what it is and how it evolves**. New York: Simon & Schuster/Free Press, 2009.
- ASIF, Imran; TIDDI, Ilaria; GRAY, Alasdair J. G. Using nanopublications to detect and explain contradictory research claims. *In: IEEE INTERNATIONAL CONFERENCE ON eSCIENCE, 17th, 2021. Proceedings [...]*. Innsbruck, Austria: IEEE, 2021. p. 1-10.
- AURISICCHIO, Marco; ACEWELL, Rob; WALLACE, Ken. Understanding how the information requests of aerospace engineering designers influence information-seeking behaviour. **Journal of Engineering Design**, v. 21, n. 6, p. 707–730, 2010.
- BAEZA-YATES, Ricardo; RIBEIRO-NETO, Berthier. **Modern information retrieval: the concepts and technology behind search**. New Jersey, USA: Addi-son-Wesley Professional, 2011.
- BAHLOUL, Khaled; BUZON, Laurent; BOURAS, Abdelaziz. Archival Initiatives in the Engineering Context. *In: YAN, Xiu-Tian; ION, William J.; EYNARD, Benoit (orgs.). Global Design to Gain a Competitive Edge*. London: Springer, 2008. p. 313–321.
- BALL, Alexander James *et al.* A grand challenge: immortal information and through-life knowledge management (KIM). **International Journal of Digital Curation**, v. 1, n.1, p. 53–59, 2006.
- BALL, Alexander James; NEILSON, Colin. **Curation of research data in the disciplines of Engineering**. 2010. (SCARP Case Study n. 7). Disponível em: https://www.dcc.ac.uk/sites/default/files/documents/publications/case-studies/SCARP_B4812_EngCase_v1_2.pdf. Acesso em: 7 jun. 2024.
- BALL, Alexander James; PATEL, Manjula; DING, Lian. Towards curation and preservation architecture for CAD engineering models. *In: INTERNATIONAL CONFERENCE ON PRESERVATION OF DIGITAL OBJECTS, 5. (iPRES 2008). Proceedings [...]*. London: 2008. p.107-114. Disponível em: <https://researchportal.bath.ac.uk/en/publications/towards-a-curation-and-preservation-architecture-for-cad-engineer>. Acesso em: 3 jun.2024.
- BARRETO, Aldo de Albuquerque. A questão da informação. **São Paulo em perspectiva**, v. 8, n. 4, p. 3–8, 1994.
- BEAGRIE, Neil. The Digital Curation Centre. **Learned Publishing**, v. 17, n. 1, p. 7–9, 2004. Disponível em: <https://onlinelibrary.wiley.com/doi/epdf/10.1087/095315104322710197>. Acesso em: 8 jun. 2024.

BELL, Gordon; HEY, Tony; SZALAY, Alex. Beyond the Data Deluge. *Science*, v. 323, n. 5919, p. 1297–1298, 6 Mar. 2009.

BLOCKLEY, David. **Engineering: A Very Short Introduction**. Oxford: Oxford University Press, 2012

BORKO, Harold. Ciência da Informação: o que é isto. *American Documentation*, v. 19, n. 1, p. 3-5, 1968

BRASIL. **Lei nº 8.666, de 21 de junho de 1993**. Regulamenta o art. 37, inciso XXI, da Constituição Federal, institui normas para licitações e contratos da Administração Pública e dá outras providências. Brasília, DF: Casa Civil, 1993. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/18666cons.htm. Acesso em: 4 fev. 2025.

BRASIL. **Lei nº 14.133, de 1º de abril de 2021**. Lei de Licitações e Contratos Administrativos. Brasília, DF: Presidência da República, 2021. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2019-2022/2021/lei/114133.htm. Acesso em: 4 fev. 2025.

BUCUR, Cristina-Iulia *et al.* **Nanopublication-Based Semantic Publishing and Reviewing: A Field Study with Formalization Papers**. 2022. Disponível em: <https://arxiv.org/pdf/2203.01608>. Acesso em: 5 fev. 2025.

BUSH, Vannevar. As we may think. *The Atlantic Monthly*, July 1945. Disponível em: <https://csis.pace.edu/~marchese/CS835/Lec3/bush.pdf>. Acesso em: 6 jun. 2024.

CLARK, Tim. Next generation scientific publishing and the web of data. *Semantic Web*, v. 5, n. 4, p. 257-259, 2014.

CONSELHO FEDERAL DE ENGENHARIA E AGRONOMIA. **Decisão Normativa nº 106, de 17 de abril de 2015**. Conceitua o termo “Projeto” e define suas tipificações. Brasília, DF: CONFEA, 2015. Disponível em: https://www.febrageo.org.br/downloads/Decis%C3%A3o%20Normativa%20106-15_Defini%C3%A7%C3%A3o%20de%20projeto.pdf. Acesso em: 4 fev. 2025.

CONSTANTOPOULOS, Panos; DALLAS, Costis. **Aspects of a digital curation agenda for cultural heritage**. 2008. 6 p. Disponível em: <http://www.dcu.gr/wp-content/uploads/2016/10/Aspects-of-a-digital-curation-agenda-for-cultural-heritage.pdf>. Acesso em: 3 jun. 2024.

CONSTANTOPOULOS, Panos *et al.* DCC&U: An extended digital curation lifecycle model. *The International Journal of Digital Curation*, v. 4, n. 1, 2009. Disponível em: <https://ijdc.net/index.php/ijdc/article/view/100/75>. Acesso em: 3 jun. 2024.

CRESWELL, John W.; CRESWELL, J. David. **Projeto de pesquisa: Métodos qualitativo, quantitativo e misto**. 5. ed. Porto Alegre: Penso Editora, 2021.

DALLAS, Costis. Digital curation beyond the “wild frontier”: A pragmatic approach. *Archival Science*, v. 16, n. 4, p. 421-457, 2016.

DAVENPORT, Thomas H.; PRUSAK, Laurence. **Information ecology: Mastering the information and knowledge environment**. New York: Oxford University Press, 1997.

DIGITAL CURATION CENTRE. **Curation Lifecycle Model**. 2024. Disponível em: <https://www.dcc.ac.uk/guidance/curation-lifecycle-model>. Acesso em: 7 jun. 2024.

DIGITAL CURATION CENTRE. **What is digital curation?** 2004. Disponível em: <https://www.dcc.ac.uk/about/digital-curation>. Acesso em: 5 fev. 2025.

DURANTI, Luciana. Registros documentais contemporâneos como provas de ação. **Revista Estudos Históricos**, v. 7, n. 13, p. 49–64, 1994.

ESTEVA, Maria; JANSEN, Craig; CORONEL, Josue Balandrano. Designing and Building Interactive Curation Pipelines for Natural Hazards in Engineering Data. *International Journal of Digital Curation*, v.13, n.1, p.306-317, 2018. Disponível em: <https://ijdc.net/index.php/ijdc/article/view/661/534>

FABRIS, Erika. Extending the Use of Nanopublications: Retrieval, Citation and Statement Verification. *In: FDIA@ ESSIR, 9th, 2019. Proceedings [...]*. Milan: 2018. p. 48-53. Disponível em: <https://ceur-ws.org/Vol-2537/paper-09.pdf>. Acesso em: 5 fev. 2025.

FAN, Jiayuan; ZHAO, Jun. A digital management model of power design drawings based on blockchain. *In: ASIA CONFERENCE ON POWER AND ELECTRICAL ENGINEERING (ACPEE), 7th, 2022. Proceedings [...]*. Hangzhou, China: IEEE, 2022. p.565-569.

FANIEL, Ixchel M.; ZIMMERMAN, Ann. Beyond the data deluge: a research agenda for large-scale data sharing and reuse. **International Journal of Digital Curation**, v. 6, n. 1, p. 58-69, 2011.

FROESE, Thomas M. The impact of emerging information technology on project management for construction. **Automation in construction**, v. 19, n. 5, p. 531-538, 2010.

FRONT-END LOADING. *In: INDEPENDENT PROJECT ANALYSIS. IPA Glossary 2024*. Disponível em: <https://www.ipaglobal.com/about/ipa-glossary>. Acesso em: 6 jun. 2024.

GABRIELAITIS, Linas; BAUŠYS, Romualdas. **Management of project properties in “Virtual Archive” for building design industry**. 2009. Disponível em: <https://etalpykla.vilniustech.lt/handle/123456789/124685>. Acesso em: 3 jun. 2024.

GARTNER, Richard. **Metadata for digital libraries: state of the art and future directions**. Bristol: JISC Technology and Standards Watch, 2008. Disponível em: https://kclpure.kcl.ac.uk/ws/portalfiles/portal/5949619/tsw_0801pdf.pdf. Acesso em: 3 jun. 2024.

GIL, Antonio Carlos. **Como Elaborar Projetos de Pesquisa**. 7. ed. São Paulo: Editora Atlas, 2022.

GOLDEN, Patrick; SHAW, Ryan. Nanopublication beyond the sciences: the PeriodO period gazetteer. **Peer Journal Computer Science**, v. 2, p. e44, 2016. Disponível em: <https://peerj.com/articles/cs-44.pdf>. Acesso em: 5 fev. 2025.

GRAY, Alasdair J. G. et al. Nanopublication Guidelines: Nanopublications Working Draft 04 February 2025. Disponível em: https://nanopub.net/guidelines/working_draft/. Acesso em: 7 fev. 2025.

GROTH, Paul; GIBSON, Andrew; VELTEROP, Jan. The anatomy of a nanopublication. **Information Services and Use**, v. 30, n. 1-2, p. 51-56, 2010.

GROTH, Paul; MOREAU, Luc (ed.). **PROV-overview**; An Overview of the PROV Family of Documents. 2013, p. 881-906. (W3C Working Group Note).

GRUBER, Thomas R. Toward principles for the design of ontologies used for knowledge sharing? **International Journal of Human-Computer Studies**, v. 43, n. 5-6, p. 907-928, 1995.

HAHM, Gyeong June *et al.* A personalized query expansion approach for engineering document retrieval. **Advanced Engineering Informatics**, v. 28, n. 4, p. 344–359, 2014. Disponível em: http://kirc.kaist.ac.kr/papers/journal/2014_personalized.pdf. Acesso em: 3 jun. 2024.

HARVEY, Ross. **Digital curation: a how-to-do-it manual**. London: Facet Publishing, 2010.

HAYNES, David. **Metadata for Information Management and Retrieval: Understanding metadata and its use**. London: Facet Publishing, 2018.

HERTZUM, Morten; PEJTERSEN, Annelise Mark. The information-seeking practices of engineers: searching for documents as well as for people. **Information Processing & Management**, v. 36, n. 5, p. 761–778, 2000.

HIGGINS, Sarah. Digital curation: the emergence of a new discipline. **International Journal of Digital Curation**, v. 6, n. 2, p. 78-88, 2011.

HIGGINS, Sarah. The DCC curation lifecycle model. **International Journal of Digital Curation**, v. 3, n. 1, p. 134–140, 2008.

HORSCH, Martin Thomas *et al.* Research Data Infrastructures and Engineering Metadata. *In*: HORSCH, Martin Thomas *et al.* **Data Technology in Materials Modelling**. Cham: Springer, 2021. p. 13–30.

HUI, Yuk. What is a digital object? *In*: HALPIN, Harry; MONNIN, Alexandre. (ed.). **Philosophical engineering: Toward a philosophy of the web**. New Jersey: Wiley, c2014. p. 52-67.

KERZNER, Harold. **Gestão de projetos: uma abordagem global**. Porto Alegre Editora Bookman, 2002.

KLÜWER, Johan W.; SKJÆVELAND, Martin G.; VALEN-SENDSTAD, Magne. ISO 15926 templates and the Semantic Web. *In*: W3C WORKSHOP ON SEMANTIC WEB IN ENERGY INDUSTRIES. **Proceedings [...]**. 2008. Disponível em: <https://www.w3.org/2008/11/ogws-papers/kluwer.pdf>. Acesso em: 5 fev. 2025.

KOUKIAS, Andreas; KIRITSIS, Dimitris. A Step-by-Step Approach to Ensure Asset Performance Based on Technical Documentation. **IFAC-PapersOnLine**, v. 48, n. 21, p. 820–825, 2015.

KUHN, Tobias. A survey and classification of controlled natural languages. **Computational linguistics**, v. 40, n. 1, p. 121-170, 2014.

KUHN, Thomas Samuel. **A estrutura das revoluções científicas**. 5. ed. São Paulo: Perspectiva, 2000. (Debates Ciência, 115).

KUHN, Tobias *et al.* Broadening the scope of nanopublications. **Lecture Notes in Computer Science**, v. 7882. p. 487–501, 2013.

KUHN, Tobias *et al.* **Nanopublications: A growing resource of provenance-centric scientific linked data**. 2018. 10 p. Disponível em: <https://arxiv.org/pdf/1809.06532>. Acesso em: 8 jun. 2024.

KUHN, Tobias *et al.* Semantic micro-contributions with decentralized nanopublication services. **Peer Journal Computer Science**, v. 7, p. e387, 2021.

KUHN, Tobias; DUMONTIER, Michel. Making digital artifacts on the web verifiable and reliable. **IEEE Transactions on Knowledge and Data Engineering**, v. 27, n. 9, p. 2390–2400, 2015.

KUHN, Tobias; DUMONTIER, Michel. Trusty URIs: Verifiable, immutable, and permanent digital artifacts for linked data. **Lecture Notes in Computer Science**, v. 8465, p. 395–41, 2014.

KUHN, Tobias; KRAUTHAMMER, Michael. Underspecified scientific claims in nanopublications. **CEUR Workshop Proceedings**, v. 906, p. 29–32, 2012.

LEE, Christopher A.; TIBBO, Helen R. **Digital curation and trusted repositories: steps toward success.** 2007. Disponível em: <https://jodi-ojs-tdl.tdl.org/jodi/index.php/jodi/article/download/229/183>. Acesso em: 3 jun. 2024.

LEK, Timo *et al.* **Provenance for Linguistic Corpora Through Nanopublications.** 2020. Disponível em: <https://arxiv.org/abs/2006.06341>. Acesso em: 5 fev. 2025.

LEVY, Yair; ELLIS, Timothy J. A systems approach to conduct an effective literature review in support of information systems research. **Informing Science Journal**, v. 9, p. 181-212, 2006. Disponível em: <http://www.scs.ryerson.ca/aferworn/courses/CP8101/CLASSES/ConductingLiteratureReview.pdf>. Acesso em: 3 jun. 2024.

LI, Zhanjun; RASKIN, Victor; RAMANI, Karthik. Developing engineering ontology for information retrieval. **Journal of Computing and Information Science in Engineering**, v. 8, n. 1, 011003, 2008.

LIU, Shaofeng *et al.* An Approach for Document Fragment Retrieval and its Formatting Issue in Engineering Information Management. **Lecture Notes in Computer Science**, v. 3981, p. 279–287, 2006.

LORD, Philip; MACDONALD, Alison. **Data curation for e-Science in the UK: an audit to establish requirements for future curation and provision.** 2003. Disponível em: <https://digitalpreservation.gov/news/2004/e-ScienceReportFinal.pdf>. Acesso em: 3 jun. 2024.

LOWE, Alistair *et al.* A method for the study of information use profiles for design engineers. *In: ASME 1999 DESIGN ENGINEERING TECHNICAL CONFERENCES. Proceedings [...].* Las Vegas: ASME, 1999. p. 109–119.

LOWE, Alistair. **Studies of information use by engineering designers and the development of strategies to aid in its classification and retrieval.** Bristol: University of Bristol, 2002.

MADSEN, David A.; MADSEN, David P. **Engineering drawing & design.** 6th ed. Boston: Cengage Learning, 2017.

MANNING, Christopher D.; PRABHAKAR, Raghavan; SCHÜTZE, Hinrich. **Introduction to information retrieval.** Cambridge: Cambridge University Press, 2008.

MAXIMIANO, Antonio Cesar Amaru. **Administração de projetos: como transformar ideias em resultados.** São Paulo: Atlas, 2014.

MCCUSKER, Jamie P. *et al.* Developing Scientific Knowledge Graphs Using Whyis. *In: SEMSCI@ ISWC, 2018. Proceedings [...].* Monterey, 2018. p. 52-58.

MCMAHON, Chris *et al.* **The development of a set of principles for the through-life management of engineering information.** 2009. Disponível em:

<https://purehost.bath.ac.uk/ws/portalfiles/portal/230070112/kim40rep007mjd10.pdf>. Acesso em: 3 jun. 2024.

MCMAHON, Chris *et al.* Waypoint: an integrated search and retrieval system for engineering documents. **Journal of Computing and Information Science in Engineering**, v. 4, n. 4, p. 329–338, 2004.

MCMAHON, Chris; BUSBY, Jerry. Risk in the design process. In: Design process improvement: a review of current practice, 2005.

MCMAHON, Chris; GIESS, Matt; CULLEY, Steve. Information management for through life product support: the curation of digital engineering data. **International Journal of Product Lifecycle Management**, v. 1, n. 1, p. 26-42, 2005.

MCMAHON, Chris; LOWE, Alistair; CULLEY, Steve. Knowledge management in engineering design: personalization and codification. **Journal of Engineering Design**, v. 15, n. 4, p. 307-325, 2004.

MONS, Barend; VELTEROP, Jan. **Nano-publication in the e-science era**. 2009. Disponível em: [https://www.w3.org/wiki/images/4/4a/HCLSS\\$ISWC2009\\$\\$Workshop\\$Mons.pdf](https://www.w3.org/wiki/images/4/4a/HCLSS$ISWC2009$$Workshop$Mons.pdf). Acesso em: 8 jun. 2024.

MOOERS, Calvin N. **The Theory of Digital Handling of Non-Numerical Information and its Implications to Machine Economics**. Boston: Zator Company, 1950. 34 p.

MOOERS, Calvin N. Zatorcoding applied to mechanical organization of knowledge. **American Documentation**, v. 2, n.1, p. 20-32, 1951.

NANOPUB.NET. **Nanopublications**: What is a Nanopublication? 2022. Disponível em: <https://nanopub.net/>. Acesso em: 7 jun. 2024.

NEILSON, Colin. **Digital Curation approaches for Architecture**. (SCARP Case Study N. 6). Edinburgh: Digital Curation Centre, 2009. Disponível em: https://www.dcc.ac.uk/sites/default/files/documents/scarp/SCARP_Architecture.pdf. Acesso em: 5 fev. 2025.

NICHOLAS, John M; STEYN, Herman. **Project management for engineering, business and technology**. 2021.

OLIVER, Gillian; HARVEY, Ross. **Digital curation**. Chicago: American Library Association, 2016.

PAGE, Matthew J. *et al.* The PRISMA 2020 statement: an updated guideline for reporting systematic reviews, **BMJ**, v. 372, n.71, 2021. Disponível em: doi: 10.1136/bmj.n71. Acesso em: 5 fev. 2025

PAHL, Gerhard *et al.* **Engineering design**: A systematic approach. London: Springer, 2007.

PATLAKAS, Panagiotis; MUSSO, Marta; LARKHAM, Peter. A digital curation model for post-occupancy evaluation data. **Architectural Engineering and Design Management**, v. 18, n. 2, p. 128–148, 2021.

PENNOCK, Maureen. Digital curation: A life-cycle approach to managing and preserving usable digital information. **Library & Archives**, v. 1, n. 1, p. 1–3, 2007.

PETROBRAS. **N-1710 - Codificação de Documentos Técnicos de Engenharia**. Rio de Janeiro: CONTEC, 2014.

PETROBRAS. **Perfil**: conheça mais sobre a nossa empresa. Rio de Janeiro, [2025]. Disponível em: <https://petrobras.com.br/quem-somos/perfil>. Acesso em: 5 fev, 2025.

POOLE, Alex H. The conceptual landscape of digital curation. **Journal of Documentation**, v. 72, n. 5, p. 961–986, 2016.

PORTO, Renata Maria Abrantes Baracho. **Sistema de recuperação de informação visual em desenhos técnicos de engenharia e arquitetura**: modelo conceitual, esquema de classificação e protótipo. Orientadora: Beatriz Valadares Cendón. 2007. 273 f. Tese (Doutorado em Ciência da Informação) - Escola de Ciência da Informação – Universidade Federal de Minas Gerais, 2007.

PROJECT MANAGEMENT INSTITUTE. **The standard for project management and a guide to the project management body of knowledge (PMBOK guide)**. 7th ed. Newtown Square: Project Management Institute, 2021.

PSCHYBILLA, Thomas; HOMANN, Alex. Evaluation of end-to-end process and information flow analyses through digital transformation in mechanical engineering. *Procedia cirp*, v. 93, p. 298-303, 2020.

QUINTANA, Virgilio *et al.* Will Model-based Definition replace engineering drawings throughout the product lifecycle? A global perspective from aerospace industry. **Computers in industry**, v. 61, n. 5, p. 497-508, 2010.

RILEY, Jenn. **Understanding Metadata**: What is Metadata, and What is it for? Baltimore: NISO, 2017. Disponível em: <https://groups.niso.org/higherlogic/ws/public/download/17446/Understanding%20Metadata.pdf>. Acesso em: 7 jun. 2024.

ROMERO, Fernando; ANDERY, Paulo. **Gestão de megaprojetos**: Uma abordagem lean. Rio de Janeiro: Brasport, 2016.

SABHARWAL, Arjun. **Digital curation in the digital humanities**: Preserving and promoting archival and special collections. Waltham: Chandos Publishing, 2015.

SAMAD, Norhaslinda Abdul; OSMAN, Kamisah; NAYAN, Nazrul Anuar. Computational Thinking through the Engineering Design Process in Chemistry Education. **International Journal of Educational Methodology**, v. 9, n. 4, p. 771-785, 2023.

SAYÃO, Luís Fernando. Modelos teóricos em Ciência da Informação – abstração e método científico. **Ciência da Informação**, Brasília, v. 30, n. 1, p. 82-91, jan./abr. 2001.

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.

SAYÃO, Luis Fernando; SALES, Luana Farias. Curadoria digital: um novo patamar para preservação de dados digitais de pesquisa. **Informação & Sociedade: Estudos**, v. 22, n. 3, p. 179–191, 2012.

SERNADELA, Pedro *et al.* A nanopublishing architecture for biomedical data. *In: INTERNATIONAL CONFERENCE ON PRACTICAL APPLICATIONS OF COMPUTATIONAL BIOLOGY & BIOINFORMATICS (PACBB 2014)*, 8th. **Proceedings [...]**. London: Springer International Publishing, 2014. p. 277-284.

SERRÃO, Samara Moraes. Metodologia FEL (Front-End Loading) como recurso para o gerenciamento de projetos. **Revista Técnico-Científica do CREA-PR**, n.32, 2023. Disponível em: <https://revistatecie.crea-pr.org.br/index.php/revista/article/view/799/664>. Acesso em: 5 fev. 2025.

SHAH, Jami J.; ROGERS, Mary T. Engineering data management: achieving integration through database technology. **Computing & Control Engineering Journal**, v. 4, n. 3, p. 119–126, 1993.

SILVA, Faysa Maria de Oliveira; SIEBRA, Sandra de Albuquerque. Análise de modelos de ciclos de vida para curadoria de objetos digitais. *In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO (ENANCIB)*, 18, 2017. **Anais [...]**. Marília: Unesp, 2017. Disponível em: http://enancib.marilia.unesp.br/index.php/XVIII_ENANCIB/ENANCIB/paper/viewFile/615/875. Acesso em: 5 fev. 2025.

SMITH, Mackenzie. Curating architectural 3D CAD Models. **International Journal of Digital Curation**, v.4, n.1, 2009. Disponível em: <https://ijdc.net/index.php/ijdc/article/view/81/80>. Acesso em: 3 fev. 2025

SWANSON, Don R. Historical note: Information retrieval and the future of an illusion. **Journal of the American Society for Information Science**, v. 39, n. 2, p. 92–98, 1988.

THIBODEAU, Kenneth. Overview of Technological Approaches to Digital Preservation and Challenges. *In: THE STATE OF DIGITAL PRESERVATION: AN INTERNATIONAL PERSPECTIVE. Proceedings [...]*. Washington: CLIR, 2002.

THOMAS, Eric *et al.* **Construction Disconnected**: Rethinking the management of project data and mobile collaboration to reduce cost and improve schedules. 2018. Disponível em: https://pg.plangrid.com/rs/572-JSV-775/images/Construction_Disconnected.pdf. Acesso em: 3 fev. 2025.

UNISOCIESC. **O que é GED?** 2015. Disponível em: <https://gedsystem.blogspot.com/p/o-que-e-ged.html>. Acesso em: 4 fev. 2025.

VAN DER HORST, Eelke *et al.* Finding Novel Associations Across Domains Using Linked Data: a case study on genetic variants disrupting transcription start sites. *In: SWAT4LS INTERNATIONAL CONFERENCE*, 2015. **Proceedings [...]**. Cambridge, 2015, p. 2-12.

VANJA, Sándor. Workflow design. *In: CLARKSON, John; ECKERT, Claudia (ed.). Design process improvement: a review of current practice*. London: Springer, 2005. p. 366-385.

WATTS, Frank B. **Engineering documentation control handbook**: configuration management in industry. 3rd ed. Norwich: William Andrew, 2008.

WILD, Peter J. *et al.* A diary study of information needs and document usage in the engineering domain. **Design Studies**, v. 31, n. 1, p. 46-73, 2010.

WILKINSON, Mark D. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. **Scientific data**, v. 3, n. 1, p. 1–9, 2016.

YAKEL, Elizabeth. Digital curation. **OCLC Systems & Services: International digital library perspectives**, v. 23, n. 4, p. 335–340, 2007.

YIN, Robert K. **Pesquisa qualitativa do início ao fim**. Porto Alegre: Penso Editora, 2016.