

SUBMETIDO 03/12/2021  
APROVADO 02/06/2022  
PUBLICADO ON-LINE 15/06/2022  
PUBLICADO 10/10/2023  
EDITORA ASSOCIADA  
Luzia Góes Camboim

DOI: <http://dx.doi.org/10.18265/1517-0306a2021id6540>

ARTIGO ORIGINAL

## Extração de conhecimento a partir da modelagem de processos e grafo semântico: uma abordagem no processo de manutenção *offshore*

 Yago Pacheco Teixeira <sup>[1]</sup>  
 Benedito Adilson Domiciano Neto <sup>[2]</sup>  
 Mark Douglas de Azevedo Jacyntho <sup>[3]</sup>  
 Simone Vasconcelos Silva <sup>[4]\*</sup>

[1] [yago.btkn@gmail.com](mailto:yago.btkn@gmail.com)

[2] [netodomiciano@hotmail.com](mailto:netodomiciano@hotmail.com)

[3] [mjacyntho@iff.edu.br](mailto:mjacyntho@iff.edu.br)

[4] [simonevs@iff.edu.br](mailto:simonevs@iff.edu.br)

Instituto Federal de Educação, Ciência e  
Tecnologia Fluminense (IFF), Brasil

**RESUMO:** A indústria petrolífera possui grandes desafios e incertezas relativos à complexidade do negócio, os quais foram agravados devido à pandemia de covid-19. Uma das formas de enfrentar esses desafios é aumentar a eficiência por meio de tomadas de decisões mais assertivas. O objetivo deste trabalho é propor um método para extração de conhecimento a partir de dados oriundos do processo de manutenção de uma organização do setor petrolífero. Esse método aborda a modelagem de processos e a geração de um grafo semântico capaz de proporcionar indicadores para auxílio à tomada de decisões. Primeiramente, foram realizadas uma revisão bibliográfica e uma seleção de trabalhos relacionados ao tema tratado. Em seguida, elaborou-se a modelagem do processo de manutenção, possibilitando a identificação dos dados gerados. Foi desenvolvido um grafo semântico a partir do mapeamento dos dados de acordo com as ontologias adequadas, viabilizando consultas capazes de extrair o conhecimento necessário para geração de indicadores. Como resultado, o método proposto (consulta ao grafo semântico), comparado ao método utilizado pela organização (consulta a uma planilha), obteve retorno (consultas aos dados) de forma mais rápida, mais organizada e com menor número de interações (redução de 50%), proporcionando um melhor resultado na criação de indicadores.

**Palavras-chave:** conhecimento; dados; grafo semântico; manutenção; modelagem de processos.

### *Knowledge extraction from process modeling and semantic graph: an approach to the offshore maintenance process*

**ABSTRACT:** The oil industry has great challenges and uncertainties related to the complexity of the business, which were aggravated due to the covid-19 pandemic. One of the ways to face these challenges is to increase efficiency

\*Autor para correspondência.

through more assertive decision-making. The objective of this work is to propose a method for extracting knowledge from data from the maintenance process of an organization in the oil sector. This method addresses the modeling of processes and the generation of a semantic graph capable of providing indicators to aid decision making. First, a bibliographic review and a selection of works related to the topic were carried out. Then, the maintenance process was modeled, enabling the identification of the generated data. A semantic graph was developed from the mapping of data according to the appropriate ontologies, enabling queries capable of extracting the necessary knowledge to generate indicators. As a result, the proposed method (query to the semantic graph) compared to the method used by the organization (query to a spreadsheet) obtained a return (queries to the data) in a faster, more organized way and with a lower number of interactions (50% reduction), providing a better result in the creation of indicators.

**Keywords:** data; knowledge; maintenance; process modeling; semantic graph.

## 1 Introdução

Os gestores de empresas da indústria petrolífera estão carregados de desafios relativos às incertezas relacionadas à complexidade do negócio, ao tempo de maturação de projetos, à dificuldade de projeções de longo prazo e à volatilidade do mercado de acordo com a política mundial. Essas incertezas foram agravadas pela pandemia de covid-19, que paralisou as atividades em todo o mundo, gerando retração na economia global e causando variações abruptas nos preços do petróleo (ABRÃO, 2020).

Algumas organizações estão mudando seus modelos de negócio para se adaptarem aos novos desafios, mas essa não é uma opção para aquelas cujo negócio é inteiramente dedicado à extração de petróleo e gás. Para essas organizações, uma das maneiras de enfrentar esses desafios é melhorar a eficiência e reduzir custos (FRASER; ANASTASELOS; RAVIKUMAR, 2018).

As operações de manutenção são significativas para a segurança e consistência de sistemas offshore (ISLAM *et al.*, 2020). A manutenção como função estratégica das organizações é responsável direta pela disponibilidade dos ativos, além de possuir grande importância nos resultados obtidos. Estes serão melhores à medida que a gestão da manutenção for mais eficaz. Com o passar dos anos e com o avanço tecnológico, foi necessário implementar o planejamento e controle de manutenção no setor de petróleo e gás. A gestão da manutenção auxilia no planejamento e controle de todo o ciclo desse processo, cujas vantagens são inúmeras, visando sempre à redução de custo, do desperdício e do retrabalho (VITÓRIO; NOGUEIRA, 2016).

Para auxiliar o processo de tomada de decisão, são utilizadas várias técnicas, ferramentas, métodos, modelos, *frameworks* e abordagens que estão disponíveis na gestão estratégica. Os métodos para extração de dados também podem auxiliar no processo de tomada de decisão (POKRAEV, 2009). Uma das formas de tratar dados é por meio da Web Semântica, pois esta possui o objetivo de apresentar, automatizar, integrar e reutilizar dados entre aplicações (W3C, 2013b). A Web Semântica trabalha com formatos comuns de integração e combinação de dados oriundos de diversas fontes (W3C, 2015).

Nesse contexto, o objetivo deste artigo é propor um método para extração de conhecimento a partir de dados oriundos dos processos de manutenção de uma organização do setor petrolífero. Esse método aborda a modelagem dos processos e a geração de um grafo semântico capaz de proporcionar indicadores para auxílio à tomada de decisões.

Por meio do método proposto, é possível verificar as vantagens do uso de grafo semântico em relação ao sistema utilizado atualmente pela organização para a consulta aos dados do processo de manutenção e a geração de indicadores.

A partir da introdução, este artigo encontra-se organizado nas seguintes seções: a Seção 2 aborda uma revisão bibliográfica; a Seção 3 descreve a metodologia utilizada; a Seção 4 apresenta os resultados obtidos; e a Seção 5 descreve as conclusões deste artigo.

## 2 Revisão bibliográfica

Esta seção apresenta uma breve revisão da literatura referente à modelagem de processos e à Web Semântica. Além disso, também aborda os trabalhos relacionados.

### 2.1 Modelagem de processos

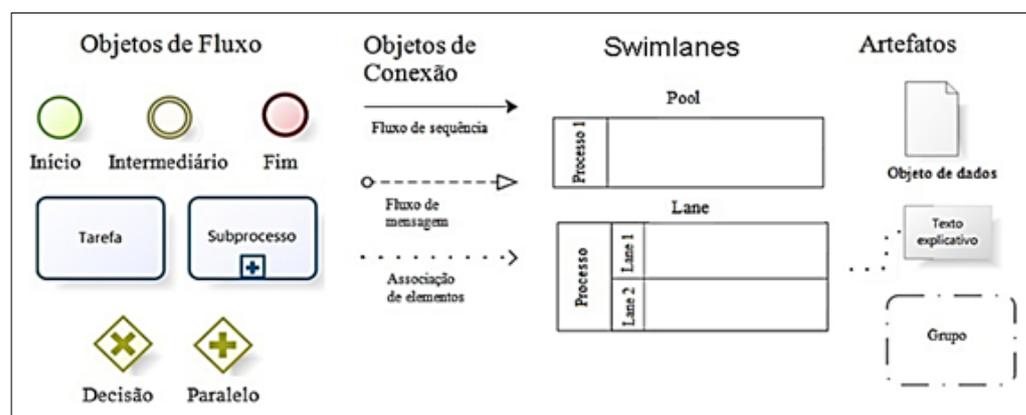
O *Business Process Management* (BPM) pode ser entendido como um conjunto de técnicas associado à gestão de processos de negócio com o auxílio de ferramentas tecnológicas. Visa à melhoria dos processos de negócio das organizações por meio do uso de métodos e ferramentas para modelar, analisar e controlar processos de negócio envolvendo os aspectos estratégicos, organizacionais, humanos e os sistemas aplicativos (ABPMP, 2020).

A modelagem de processos tem como finalidade representar graficamente o fluxo das atividades que formam um processo, apresentando as dimensões de interesse conforme os objetivos que cada situação sugere, assim como entradas, saídas, interfaces com outros processos, pessoas responsáveis por cada atividade, informações necessárias, entre outros elementos (CAVALCANTI, 2017).

Uma das notações mais utilizadas atualmente para modelagem de processos é a BPMN (*Business Process Model and Notation*), cujo objetivo principal é fornecer uma notação que seja facilmente compreensível por todos os usuários, desde os analistas de negócios até os desenvolvedores técnicos e, finalmente, os gestores que irão gerenciar e monitorar os processos. Assim, a BPMN fecha a lacuna entre o design do processo e a sua implementação. A apresentação gráfica dos processos facilita sua análise e modificação, ao mesmo tempo em que oferece uma ampla gama de aplicações da notação como método básico para o gerenciamento de processos de negócio (JASIULEWICZ-KACZMAREK *et al.*, 2018).

A Figura 1 (próxima página) apresenta os elementos principais utilizados na notação BPMN, disponíveis para representar os mais variados tipos de processos de negócio.

**Figura 1** ▶  
Principais elementos da  
notação BPMN.  
Fonte: Amarilla e  
Iarozinski Neto (2018)



## 2.2 Web Semântica

A Web Semântica permite a criação de ambientes que tenham características semânticas úteis para integrações e buscas; nela, tecnologias e padrões possuem a intenção de apresentar informações em uma forma estruturada, permitindo o processamento automático e o reuso de informações por agentes computacionais (KISELEV; YAKUTENKO, 2020).

Um conceito importante da Web Semântica é o de ontologia, que, de acordo com Carchedi *et al.* (2018), é um artefato de representação da informação muito útil para integração de dados e para garantir interoperabilidade semântica das informações que ela representa. Para Torino *et al.* (2020), quanto melhor definida uma ontologia, a qual identifica e conecta os dados, mais bases ela poderá alcançar, com mais dados podendo ser interpretados e modelados; com isso, possibilita-se a aplicação de uma ontologia para um determinado escopo com bases de dados de diversos locais que tenham os mesmos tipos de dados que possam ser referenciados.

De acordo com Santarém Segundo, Coneglian e Lucas (2017), são várias as linguagens da Web Semântica que se destacam. Neste trabalho, são utilizadas duas dessas linguagens: *Resource Description Framework* (RDF) e *SPARQL Protocol and RDF Query Language* (SPARQL).

Por meio da linguagem RDF, é possível criar uma rede de informações a partir de dados distribuídos, gerando os grafos semânticos, também conhecidos como grafos de conhecimento. É possível aplicar resoluções de um conjunto de problemas utilizando esses grafos criados (SANTARÉM SEGUNDO, 2014).

Os elementos básicos para construção de um modelo RDF são as triplas, compostas de sujeito ou recurso, predicado ou propriedade, e objeto ou valor da propriedade. Assim, triplas RDF são uma forma de representação para escrever um código RDF (ALVES, 2005). Os grafos semânticos consistem em vértices (ou nós) e arestas que são formadas por triplas RDF (LOPES, 2020).

A linguagem de consulta SPARQL é baseada em triplas RDF (W3C, 2013a). A estrutura da SPARQL baseia-se em consultas simples na forma de padrões gráficos simples, gerados para recuperar e manipular dados no formato RDF, em que os dados, que são armazenados em triplas, podem ser consultados e tratados (CARVALHO, 2020).

Para verificar a corretude das consultas, são utilizadas perguntas de competência que podem ser empregadas para verificar se uma determinada ontologia satisfaz os seus requisitos (SILVA, 2019).

## 2.3 Trabalhos relacionados

Esta seção apresenta quatro trabalhos relacionados ao tema tratado neste artigo.

### 2.3.1 Jasiulewicz-Kaczmarek *et al.* (2018)

Nesse trabalho, os autores apresentaram uma aplicação da gestão de processos por meio da modelagem de manutenção de uma companhia manufatora de dispositivos médicos, que possui diversos departamentos, com variadas máquinas e linhas de produção compostas de diferentes componentes como mecânicos, elétricos e hidráulicos.

A modelagem fez-se necessária devido à implementação de um sistema de supervisão e coordenação. Escolheu-se um processo chamado “Processo de Recuperação” para elaborar a modelagem. Esta foi baseada em entrevistas com colaboradores do processo de recuperação, tais como operadores de máquinas, coordenador de produção, técnicos de manutenção, gerente de compras, e outros.

Por meio das entrevistas, foi elaborado o modelo atual do processo (AS-IS) e, por fim, foi elaborado o processo futuro (TO-BE), visando à redução de burocracia, remoção de tarefas duplicadas, redução de tempo no processo, implementação de linguagem simples, padronização de atividades, definições de pessoas e de regras de negócios e utilização de tecnologia da informação.

Obteve-se a redução do tempo de reparo em 19% em média, a redução no atraso de aquisições em 46%, o aumento da precisão do *lead time* em aproximadamente 43% e a redução do tempo de espera por peças de reposição.

### 2.3.2 Barz *et al.* (2017)

A notação BPMN foi utilizada para modelar e controlar uma análise de dados semânticos em tempo real de várias fontes, captados de uma arquitetura de sistemas físicos cibernéticos para fábricas inteligentes propostas no trabalho. Para a modelagem do processo, foi utilizado um caso de uso da manutenção de uma turbina a gás Siemens.

Foram relatadas as vantagens dessa abordagem para fábricas inteligentes, como, por exemplo, na detecção e em tratamentos automatizados de anomalias, inclusive utilizando técnicas como o uso de canetas digitais (pelo qual os dados inseridos são digitalizados e armazenados em metadados e vinculados a fontes de conhecimentos semânticos) e empregando os modelos de processos, os modelos de estruturas e os modelos de usuários.

### 2.3.3 Araujo e Rodrigues (2014)

Esses autores realizaram a modelagem de um processo chamado “Gerenciamento de Problemas de Serviço de Manutenção” da companhia Eletrobras Eletronorte para o melhor entendimento do processo e para sua otimização. Após a modelagem do processo, foram elaboradas sugestões qualitativas de melhorias para as raias do modelo cujas responsabilidades são da companhia.

Junto à indicação de melhorias, foi sugerida a criação de indicadores KPI (*Key Performance Indicators*). Foi recomendada também a utilização de um sistema computadorizado integrado a todas as plataformas de gerenciamento da rede da companhia, para tornar a análise de eventos mais eficiente.

#### 2.3.4 Joschko *et al.* (2013)

Esses autores apresentaram uma modelagem e simulação de parques eólicos *offshore*, incluindo sua operação e manutenção. Para a modelagem, contatou-se *stakeholders* do ramo, desde operadores até donos de companhias do setor, e, por meio de um questionário, foram coletadas as percepções destes quanto aos processos de negócios, recursos e dados. Também foram realizadas visitas *in loco*.

Os modelos de processos de negócio elaborados foram adequados para dar suporte técnico nas operações e em treinamentos.

Foi indicado o uso desses modelos como uma interface de comunicação cooperativa entre parceiros, sendo fundamental para otimizar os processos e interações. A ferramenta de simulação também permitiu comparar diferentes estratégias de manutenção por todo o ciclo de vida do parque eólico.

### 3 Metodologia

O método de pesquisa tem como objetivo conduzir o pesquisador à busca das respostas necessárias ao problema de pesquisa que se propõe (SAUNDERS; LEWIS; THORNHILL, 2019). A metodologia aqui empregada incorpora princípios, práticas e procedimentos necessários para a realização da pesquisa, visando cumprir os três objetivos propostos por Dresch, Lacerda e Miguel (2015): ser consistente em relação à literatura prévia do tema pesquisado; estabelecer um modelo para a condução e realização da pesquisa; e proporcionar ao pesquisador um modelo para a apresentação e avaliação da pesquisa. Esse tipo de metodologia é conhecido como *Design Science Research* (DSR) e possibilita a aproximação entre a teoria e a prática.

O DSR consiste no método que operacionaliza a pesquisa e tem como objetivo a construção de um artefato. Assim, o DSR parte da compreensão do problema para a construção e avaliação de artefatos que possibilitem a transformação de uma determinada situação para estados melhores ou desejáveis (SANTOS; KOERICH; ALPERSTEDT, 2018). Dessa forma, a metodologia proposta encontra-se dividida nas seguintes etapas:

#### 3.1 Etapa I – Definição do problema

“Como extrair conhecimento a partir de dados gerados por processos de manutenção?”

Para responder a esta pergunta, primeiramente, foi realizada uma revisão bibliográfica sobre modelagem de processos e Web Semântica, visando a uma melhor compreensão dos conceitos. E para conhecer outros estudos semelhantes, foram selecionados trabalhos relacionados ao tema da modelagem de processos de manutenção.

Para a seleção dos trabalhos relacionados, utilizou-se o método denominado PRISMA (LIBERATI *et al.*, 2009), o qual tem por objetivo ajudar a melhorar o relato de revisões sistemáticas e meta-análises (OLIVEIRA; GOMES; PORTO, 2012). Para tal, foram utilizadas três bases de conhecimento: SCOPUS, Web of Science e Google Scholar. E a estratégia de busca utilizada foi composta pelo seguinte tesauro: (“*process management*” OR BPM\*) AND (“*equipment maintenance*”).

A busca do tesauro na base SCOPUS retornou 89 resultados e na base Web of Science retornou 53 resultados. A busca foi realizada em relação ao título, resumo e palavras-chave. Já na base Google Scholar, devido a uma quantidade muito grande de resultados, a busca do tesauro foi limitada apenas ao título, retornando então 8 resultados. Dessa forma, foram totalizados 150 resultados e, ao remover os duplicados entre as bases, restaram 105 resultados.

Foi realizada uma análise dos títulos e resumos para a triagem dos estudos relacionados ao tema deste trabalho. Assim, dos 105 resultados, 12 foram selecionados para análise dos critérios de elegibilidade e 93 foram descartados por falta de aderência ao tema tratado. Em seguida foi feita uma análise dos trabalhos de acordo com os critérios de elegibilidade (C1 – somente artigo, C2 – em idioma inglês, português ou espanhol, e C3 – acesso ao texto completo).

Dos 12 estudos selecionados, 1 foi descartado por não atender ao critério C2 e 7 foram descartados por não atenderem ao critério C3. Dessa forma, 4 estudos compõem os trabalhos relacionados; todos estes utilizaram a notação BPMN para realizar a modelagem de seus processos de manutenção, conforme apresentados na subseção 2.3.

Após a revisão da literatura, foi realizada a seleção da organização objeto deste estudo. Para tal, era preciso que esta atendesse a dois requisitos: pertencer ao setor *offshore* e possuir processos de manutenção.

### 3.2 Etapa II – Desenvolvimento do artefato

Essa etapa representa a elaboração do método proposto neste trabalho.

Dessa forma, primeiramente foi realizada uma análise dos processos de manutenção da organização selecionada, com o objetivo de compreender os tipos de manutenção existentes e como eles ocorrem em relação à geração dos dados. Com isso, foram identificados dois tipos de manutenção: a preventiva e a detectiva – a primeira gera dados não estruturados (em formato de textos) e a segunda gera dados estruturados.

Para o desenvolvimento de grafo semântico, é necessário que todos os dados estejam no formato estruturado, portanto este trabalho abordará apenas o processo de manutenção detectiva.

O processo de manutenção detectiva foi analisado e descrito de forma detalhada em relação ao seu fluxo de atividades. A partir dessa descrição, foi elaborada a modelagem desse processo, utilizando a notação BPMN, sendo esta validada em seguida. A função dessa modelagem é prover, de forma visual, uma correta compreensão do fluxo das atividades organizacionais que gera o dado estruturado.

A partir da versão definitiva da modelagem do processo de manutenção, foi identificada a atividade responsável pelo armazenamento dos dados. Com isso, foram estabelecidas as variáveis que compõem esses dados e, em seguida, estas foram convertidas para as variáveis correspondentes nas ontologias adequadas, gerando um

código RDF e, a partir deste, um grafo semântico em relação ao modelo do processo de manutenção detectiva.

### 3.3 Etapa III – Avaliação

Essa etapa representa a avaliação do método proposto neste trabalho, sendo dividida em duas validações: da modelagem do processo e do grafo semântico.

O objetivo da validação da modelagem do processo é demonstrar que o modelo é fidedigno à realidade do processo na organização. E para a validação, foram definidas seis questões: Q1 – Corretude das raias; Q2 – Adequação dos *stakeholders* que participam dos processos; Q3 – Corretude das tarefas; Q4 – Corretude dos desvios (*gateways*); Q5 – Corretude dos artefatos; Q6 – Corretude dos fluxos das informações.

As respostas para as questões acima foram obtidas por meio de uma análise da forma como a organização executa o processo de manutenção detectiva.

Para validação do grafo e demonstração de como este pode auxiliar na tomada de decisões, foi proposta uma pergunta de competência, a qual foi respondida por meio de consulta SPARQL, proporcionando os indicadores necessários para auxiliar a gestão.

Dessa forma, segue a pergunta de competência (P1) proposta: “Quantas e quais tarefas (ID, nome, descrição, área e período) indicaram problemas na execução em um determinado dia da semana considerado crítico pela organização?”

## 4 Resultados obtidos

Esta seção aborda os resultados obtidos na execução das três primeiras etapas da metodologia proposta.

### 4.1 A organização foco do estudo e seus processos de manutenção

A organização foco deste estudo atua no ramo petrolífero e encontra-se sediada no Rio de Janeiro. Opera em mais de 14 países, no segmento de energia, prioritariamente nas áreas de exploração, produção, refino, comercialização e transporte de petróleo, gás natural e seus derivados, e conta com mais de 40 mil colaboradores em seu quadro de funcionários. Este estudo está focado especificamente na área *offshore* de exploração e produção de petróleo e gás.

A organização utiliza um sistema integrado de gestão empresarial (*Enterprise Resource Planning* – ERP) conhecido como sistema SAP (Sistemas, Aplicativos e Produtos para Processamento de Dados), incluindo o módulo de manutenção. Pela complexidade das operações dessa organização, existem diversos processos para manutenções específicas, destacando-se a manutenção preventiva e a detectiva.

Segundo Slack, Brandon-Jones e Johnston (2018), a manutenção preventiva visa eliminar ou reduzir as probabilidades de falhas por manutenção, por meio de limpeza, lubrificação, substituição e verificação das instalações em intervalos de tempo pré-planejados. Já a manutenção detectiva tem como objetivo desenvolver ações que permitam o aumento da confiabilidade dos equipamentos por meio de sistemas

de proteção capazes de detectar falhas ocultas e não perceptíveis aos profissionais envolvidos nesse processo.

Este estudo encontra-se relacionado ao processo de manutenção detectiva, que gera dados estruturados.

## 4.2 Descrição do processo de manutenção detectiva

O processo de manutenção detectiva é proveniente de uma iniciativa de modernização do processo de manutenção offshore da organização estudada, utilizando dispositivos móveis específicos para auxiliar na execução das tarefas.

O planejamento e controle do processo é feito por uma célula de programação intitulada “Célula de Planejamento e Programação”, que é composta por diversos colaboradores, os quais são responsáveis por:

- filtrar possíveis notas e ordens de manutenção, as quais são liberadas automaticamente no SAP de acordo com a sua última execução e periodicidade do plano de manutenção;
- verificar a existência de impedimento para execução das tarefas, como, por exemplo, equipamento inoperante, necessidade de parada de algum equipamento em específico ou condição de mão de obra especializada;
- registrar as tarefas com possíveis informações, materiais e alertas;
- elaborar um nivelamento semanal por mês, em relação à quantidade de horas necessária para a execução das tarefas. A criação das tarefas, da periodicidade e dos planos é feita por uma equipe de Engenharia.

Esse processo de manutenção é controlado por um aplicativo Android e tem a periodicidade de cada agrupamento de tarefas previamente definida, sendo as tarefas liberadas nas datas definidas para cada executante pré-determinado. Assim, as tarefas aparecem para o executante nos dias previstos. O executante verifica se as tarefas podem ser executadas naquele momento, verificando condições como equipamentos em funcionamento, condições climáticas e outras situações que possam vir a impedir ou afetar a execução.

Caso não haja impedimento de execução, as tarefas são realizadas, preenchidas e finalizadas no aparelho celular. Se houver algum impedimento de execução, o usuário insere a justificativa de impedimento de execução da tarefa e então a finaliza.

Após a finalização das atividades, elas ficam armazenadas no aparelho celular até este se conectar à rede Wi-Fi. Ao fazer isso, o aparelho exporta automaticamente as tarefas preenchidas e finalizadas para um arquivo no formato JSON (*JavaScript Object Notation*), para uma pasta na rede corporativa. O JSON é um formato compacto, de padrão aberto e que possibilita a troca de dados entre sistemas.

Semanalmente um colaborador do setor de suporte de manutenção realiza backup desses arquivos para uma pasta específica na rede corporativa e executa a concatenação de todas as tarefas exportadas em único arquivo JSON por meio de script, ou seja, um programa computacional desenvolvido na linguagem VBA (*Visual Basic for Applications*) do editor de textos Microsoft Word.

Em seguida, é aberto um arquivo de planilha eletrônica no Microsoft Excel denominado “Rotas”, no qual se executa um segundo script desenvolvido em VBA, o qual importa o

arquivo JSON contendo todas as rotas concatenadas, cadastrando as informações no banco de dados do sistema SAP.

Ao ser finalizado o processo de cadastro no banco de dados, o colaborador do suporte de manutenção salva o arquivo e realiza um backup das informações para uma planilha eletrônica auxiliar, contendo as datas atualizadas das execuções de cada tarefa ou as justificativas de sua não execução. Essa planilha auxiliar é utilizada como apoio para elaboração do texto de um e-mail relatando o acompanhamento semanal, em que são indicadas as tarefas e seus status (executadas e finalizadas, não executadas e justificadas, e não executadas nem justificadas).

O e-mail semanal é enviado para o coordenador de manutenção, que verifica as justificativas de não execução, solicitações de abertura de notas de tratamento e demais informações, direcionando para os responsáveis as ações determinadas.

### 4.3 Modelagem do processo e validação

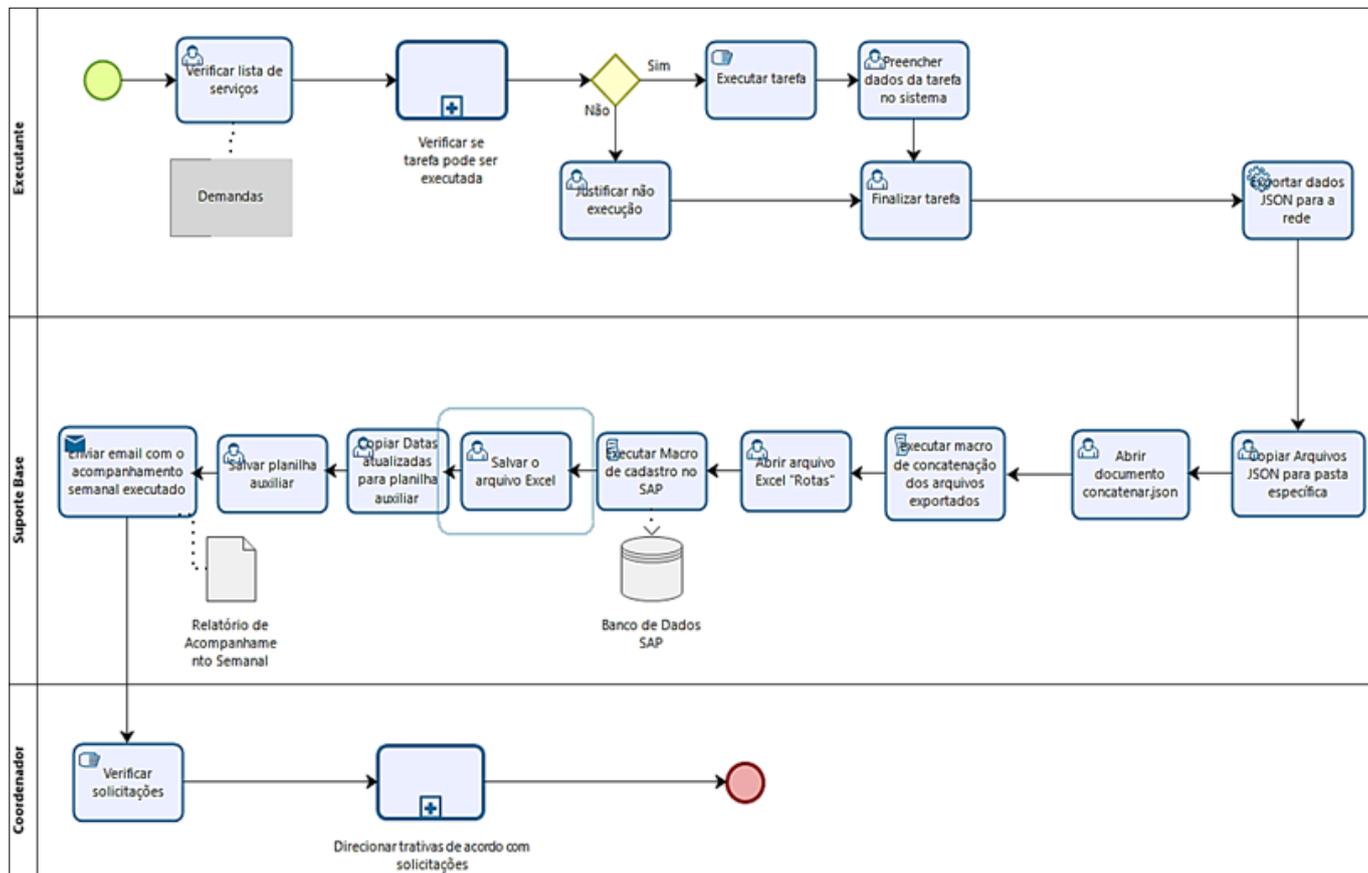
Após a análise dos documentos referentes ao processo de manutenção detectiva, foi elaborada a sua modelagem utilizando a notação BPMN. A versão inicial dessa modelagem foi validada por meio de uma verificação da forma como a organização executa esse processo de manutenção.

Após a análise da execução do processo, ocorreu a validação deste de acordo com as seis questões propostas (Q1 a Q6). Como resultado, não foi identificada a necessidade de melhorias na modelagem do processo da manutenção detectiva. Essa modelagem encontra-se representada pela Figura 2.

**Figura 2 ▼**

Modelagem do processo de manutenção detectiva.

Fonte: elaborada pelos autores (2021)



## 4.4 Dados gerados pelos processos

O armazenamento dos dados ocorre por meio do sistema SAP; o acesso ao banco de dados é realizado por meio da tarefa “Executar Macro de Cadastro no SAP” (Figura 2). A maioria das tarefas geram variáveis relacionadas aos dados do processo, conforme mostra o Quadro 1.

**Quadro 1** ►  
Variáveis para os dados do processo.  
Fonte: elaborada pelos autores (2021)

Nome da variável	Descrição do conteúdo
GrandeArea	Área de trabalho responsável pela tarefa
DescriçãoPto	Tarefa a ser executada
Id	Número que identifica a tarefa executada no momento
LocalInstalação	Código do local de execução da tarefa
Nome1	Nome do local de execução da tarefa
Descrição	Descrição do local de execução da tarefa
Nome2	Nome do equipamento
Período	Periodicidade da tarefa
DiaBase	Dia da semana para execução da tarefa
DataAbertura	Data em que a tarefa se encontra disponível para o executante
DataFinalização	Data de finalização da tarefa pelo executante
Hora	Hora de finalização da execução da tarefa
Valid	Validação se a tarefa apresentou problemas ou não

## 4.5 Geração do RDF

Os dados resultantes do processo de manutenção detectiva foram exportados para um arquivo no formato .xls ou .xlsx, ou seja, arquivo de planilha eletrônica criado por meio do software Microsoft Excel. Com isso, obteve-se um arquivo contendo 3.011 tarefas, e cada uma delas pode conter dados distribuídos em até 23 colunas.

A seleção das ontologias foi feita de acordo com uma análise de suas classes e características, de forma que estas representem adequadamente os dados do processo de manutenção, possibilitando o mapeamento das variáveis.

As ontologias selecionadas foram: i) FOAF – descreve as pessoas, suas atividades e suas relações com outras pessoas e objetos (BRICKLEY; MILLER, 2014); ii) DCMES – descreve os objetos digitais tais como vídeos, sons, imagens e textos (DUBLIN CORE, 2012); iii) SCHEMA – vocabulário de dados estruturados que define as entidades, as ações e os relacionamentos (SCHEMA.ORG, 2021); iv) SAREF – vocabulário com extensão para indústria e manufatura (ROODE *et al.*, 2020); e v) Time Ontology – conceitos temporais para descrever as propriedades temporais de recursos OWL (W3C; OGC, 2020).

Foram utilizadas onze variáveis das ontologias citadas acima para o mapeamento das treze variáveis oriundas dos dados gerados pelo processo, conforme mostra o Quadro 2. Essa quantidade menor de variáveis das ontologias se justifica pelo fato de uma destas representar mais de uma variável dos dados.

[1] GraphDB. Disponível em: <https://www.ontotext.com/products/graphdb/>. Acesso em: 10 mar. 2022

Para o mapeamento e a criação do grafo semântico foi utilizada a plataforma GraphDB<sup>1</sup>, que consiste em um banco de dados semântico em conformidade com os padrões W3C (*World Wide Web Consortium*).

Foi realizado o tratamento inicial dos dados, transformando o tipo das colunas para “número”, “data” ou “texto”. Cada coluna da planilha eletrônica (arquivo .xls ou .xlsx) foi mapeada em propriedades e classes dos vocabulários das ontologias. O Quadro 2 aponta esse mapeamento entre as variáveis dos dados dos processos e das ontologias.

**Quadro 2** ▶  
Mapeamento das variáveis versus ontologia.  
Fonte: elaborada pelos autores (2021)

Variáveis dos dados/processo	Ontologia	Variáveis na ontologia
Nome2	FOAF	<i>Name</i>
DataAbertura e DataFinalização	DCMES	<i>Date</i>
DescriçãoPto Descrição Id Valid Nome1 e LocalInstalação	Schema	<i>Name</i> <i>Description</i> <i>Identifier</i> <i>Value</i> <i>Location</i>
GrandeArea	Saref	<i>Area</i>
Período DiaBase Hora	Time	<i>Weeks</i> <i>DayofWeek</i> <i>inDateTime</i>

O mapeamento das variáveis ocorreu por meio da plataforma GraphDB, utilizando sua ferramenta OntoRefine, que possui a funcionalidade de limpeza e transformação de dados em outros formatos.

Com o mapeamento, obteve-se as triplas-RDF (sujeito, predicado e objeto), contendo as variáveis ontológicas da tarefa (*identifier, name, area, weeks, dayofweek, location e description*), como pode-se observar no Quadro 3. Essas triplas foram exportadas em arquivo com formato RDF e armazenadas num repositório criado no GraphDB.

Por meio da plataforma GraphDB, também foi possível observar a geração do grafo semântico. A Figura 3 apresenta um exemplo do grafo semântico gerado por meio das triplas RDF após o mapeamento. Além disso, na mesma figura, observa-se o nó vermelho, que indica o agrupamento das variáveis que contêm os dados da tarefa executada, e o nó azul, que representa o identificador da tarefa que funciona como um link entre os dados. Ainda no Quadro 3, é possível observar uma parte (*date, description, location e name*) das onze variáveis ontológicas que compõem o nó vermelho.

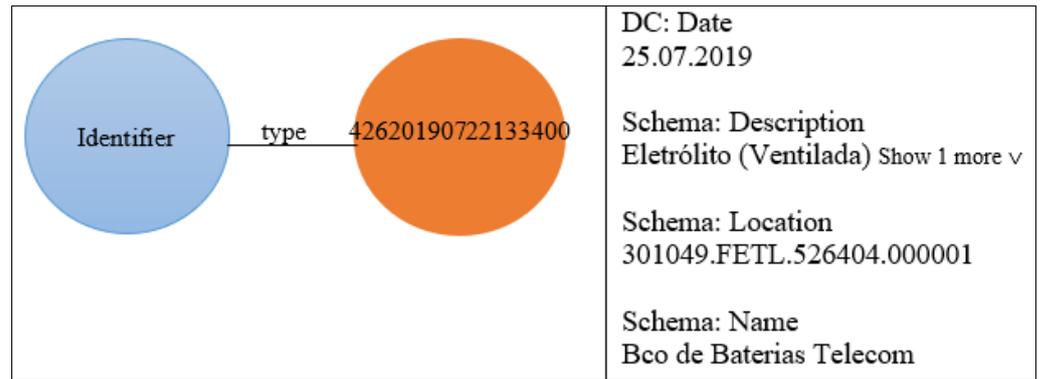
**Quadro 3** ▶  
Exemplo das triplas geradas.  
Fonte: elaborado pelos autores (2021)

a	⊕ 🗑	Schema: Identifier
Schema: Name	⊕ 🗑	“Bco de Baterias Telecom”
Saref: Area	✎ ⊕ 🗑	“FACILIDADE 02”
Time: Weeks	✎ ⊕ 🗑	“1”
Time: DayOfWeek	✎ ⊕ 🗑	“QUA”
Schema: Location	✎ ⊕ 🗑	“301049.FETL.526404.000001”
Schema: Description	✎ ⊕ 🗑	“P43 0100PF 0006”

**Figura 3** ▶

Exemplo de parte do grafo semântico criado.

Fonte: elaborado pelos autores (2021)



#### 4.6 Validação do grafo

Para validar o grafo semântico criado, são realizadas consultas utilizando a linguagem SPARQL visando responder à pergunta de competência (P1) proposta: “Quantas e quais tarefas (ID, nome, descrição, área e período) indicaram problemas na execução em um determinado dia da semana considerado crítico pela organização?”. Essa pergunta deve ser respondida por meio de consultas ao grafo semântico a partir dos dados estruturados.

As variáveis das ontologias utilizadas para responder à P1 são: *Value* (“False” se ocorrer problema, caso contrário “True”) e *DayofWeek*. Os valores para essas variáveis foram configurados da seguinte forma: *Value* = “False” e *DayofWeek* = “QUA”. Dessa forma, deseja-se saber quais tarefas apresentaram problemas em sua execução em uma quarta-feira. O recorte dessa consulta elaborada por meio da linguagem SPARQL pode ser observado no Quadro 4.

**Quadro 4** ▶

Consulta SPARQL da pergunta de competência no grafo de conhecimento.

Fonte: elaborado pelos autores (2021)

```

PREFIX
time: < http://www.w3.org/2006/time>

Select  ?o_description ?o_area ?o_ID
        ?o_weeks ?o_name

Where {
    ?s a schema:identifier ?o_ID.
    ?s schema:description
    ?o_description.
    ?s saref:area ?o_area.
    ?s time:weeks ?o_weeks.
    ?s schema:name ?o_name.
    ?s time:DayOfWeek
    ?o_DayOfWeek.
    ?s schema:value ?o_value FILTER
    (?o_value="false" &&
    ?o_DayOfWeek="QUA"). }
    
```

Essa consulta resultou nas informações de 146 tarefas (*ID, name, description, area e weeks*) que apresentaram algum problema de execução em uma quarta-feira. Assim, o gestor pode analisar qual dia da semana é o mais crítico em relação aos problemas de execução e identificar possíveis causas e soluções.

As consultas por meio do grafo semântico foram feitas de forma rápida, precisa e com apenas uma interação com o sistema, ou seja, com apenas uma ação do usuário.

Com o objetivo de comparar o número necessário de interações entre o usuário e o sistema para efetuar a consulta referente à pergunta de competência, pode-se perceber que a consulta realizada por meio do grafo semântico gerou apenas uma interação, enquanto na consulta realizada por meio do sistema atual da organização (uma planilha eletrônica com os dados estruturados) foram necessárias duas interações, ou seja, o dobro do número de interações.

Dessa forma, observa-se que o número de interações necessárias no sistema atual para uma consulta de dados é maior do que no grafo semântico proposto. Além disso, a consulta no sistema atual da organização retorna todas as informações, inclusive as não solicitadas, poluindo a visualização das informações desejadas.

Um número maior de interações é indicativo de maior tempo e custo de processamento exigido, principalmente em grandes bancos de dados. E um retorno com informações não solicitadas pode atrapalhar a visualização e, por consequência, a tomada de decisão.

#### 4.7 Comparação do método proposto com os trabalhos relacionados

O Quadro 5 mostra uma comparação dos principais temas abordados nos trabalhos relacionados e no presente trabalho, seguindo a legenda: Modelagem (BPMN), Dados Semânticos (DdsS) e Grafo Semântico (GrafS).

**Quadro 5** ►  
Comparação dos trabalhos.  
Fonte: elaborado pelos  
autores (2021)

Artigos	BPMN	DdsS	GrafS
Jasiulewicz-Kaczmarek <i>et al.</i> (2018)	✓		
Barz <i>et al.</i> (2017)	✓	✓	
Araujo e Rodrigues (2014)	✓		
Joschko <i>et al.</i> (2013)	✓		
Este trabalho	✓	✓	✓

É possível observar no Quadro 5 que, como mencionado anteriormente, todos os trabalhos utilizaram a modelagem de processos por meio da notação BPMN. Os dados semânticos foram utilizados por dois trabalhos, mas apenas o presente trabalho abordou os grafos semânticos.

A principal contribuição do presente trabalho está no fato de este propor um método capaz de gerar indicadores extraídos automaticamente por meio de consultas ao grafo semântico, apoiando a tomada de decisões por parte dos gestores organizacionais.

## 5 Conclusão

Neste trabalho, foi elaborada a modelagem do processo de manutenção detectiva que gera os dados estruturados. O modelo foi validado por meio de uma análise da forma como a organização executa o processo.

Os dados gerados foram mapeados para as ontologias adequadas e, com isso, foi elaborado um grafo semântico capaz de possibilitar consultas que respondam à pergunta de competência proposta.

Dessa forma, este estudo alcançou o objetivo proposto, contribuindo para responder à questão de pesquisa que definiu o problema tratado: “Como extrair conhecimento a partir de dados gerados por processos de manutenção?”

A consulta por meio do grafo semântico ocorreu de forma rápida, com menos interações (redução de 50%) e de forma visual mais atrativa e objetiva do que a consulta realizada pelo atual sistema da organização foco deste estudo. Após a geração de um grafo semântico, as consultas são realizadas de forma rápida, mesmo em cenários nos quais as organizações possuem uma quantidade expressiva de dados.

Para trabalhos futuros, sugere-se a inserção do processo da manutenção preventiva por meio da aplicação de mineração de texto nos dados não estruturados gerados por esse processo, visando torná-los estruturados. Assim, esses dados poderão ser integrados aos dados estruturados da manutenção detectiva abordada neste estudo.

Com isso, será possível gerar um novo grafo semântico, permitindo uma consulta em uma base de dados única, otimizando as informações disponíveis para uma tomada de decisão mais abrangente, ou seja, que represente os dois principais tipos de manutenção da organização.

## Financiamento

Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq.

## Conflito de interesses

Os autores declaram não haver conflito de interesses.

## Referências

ABPMP – ASSOCIATION OF BUSINESS PROCESS MANAGEMENT PROFESSIONALS INTERNATIONAL. **CBOK V4**: guia para o conjunto de conhecimento em gestão de processos de negócios (BPM CBOK®). Brasília, DF: ABPMP Brasil, 2020.

ABRÃO, R. A. F. Covid-19 e os impactos na indústria petrolífera: considerações iniciais. **Boletim de Economia e Política Internacional**, n. 27, p. 1-9, maio-ago. 2020. Brasília, DF: IPEA, 2020. DOI: <http://dx.doi.org/10.38116/bepi27>.

ALVES, R. C. V. **Web semântica**: uma análise focada no uso de metadados. 2005. Dissertação (Mestrado em Ciência da Informação) – Faculdade de Filosofia e Ciências,

Universidade Estadual Paulista, Marília, 2005. Disponível em: <http://hdl.handle.net/11449/93690>. Acesso em: 8 jun. 2022.

AMARILLA, R. S. D.; IAROZINSKI NETO, A. Análise comparativa dos principais processos de negócio de empresas do subsetor de edificações da construção civil. **Gestão & Produção**, v. 25, n. 2, p. 269-283, 2018. DOI: <https://doi.org/10.1590/0104-530X2406-16>.

ARAUJO, M. C.; RODRIGUES, M. A. The telecommunications business expansion at Eletrobras Eletronorte and the needs for changes in operation and maintenance. **Electra – CIGRE’s Digital Magazine**, n. D2-303, 2014.

BARZ, M.; POLLER, P.; SCHNEIDER, M.; ZILLNER, S.; SONNTAG, D. Human-in-the-loop control processes in gas turbine maintenance. In: MARIK, V.; WAHLSTER, W.; STRASSER, T.; KADERA, P. (ed.). **Industrial applications of holonic and multi-agent systems (HoloMAS 2017)**. Cham: Springer, 2017. p. 255-268. (Lecture Notes in Computer Science, v. 10444.) DOI: [https://doi.org/10.1007/978-3-319-64635-0\\_19](https://doi.org/10.1007/978-3-319-64635-0_19).

BRICKLEY, D.; MILLER, L. **FOAF Vocabulary Specification 0.99**. Namespace document 14 January 2014 – Paddington Edition. 2014. Disponível em: <http://xmlns.com/foaf/spec>. Acesso em: 8 jun. 2022.

CARCHEDI, L. C.; SOUZA, J.; BARRÉRE, E.; MENDONÇA, F. Onto4LA: uma ontologia para integração de dados educacionais. In: CONGRESSO BRASILEIRO DE INFORMÁTICA NA EDUCAÇÃO (CBIE), 7., 2018, Fortaleza. **Anais [...]**. Fortaleza: Sociedade Brasileira de Computação, 2018. DOI: <http://doi.org/10.5753/cbie.wcbie.2018.439>.

CARVALHO, G. M. **Desenvolvimento de uma camada semântica para um protótipo de repositório de dados de pesquisas de análise por ativação neutrônica**. 2020. Dissertação (Mestrado em Ciências) – Instituto de Pesquisas Energéticas e Nucleares, Universidade de São Paulo, São Paulo, 2020. DOI: <https://doi.org/10.11606/D.85.2020.tde-17122020-112928>.

CAVALCANTI, R. **Modelagem de processos de negócios: roteiro para realização de projetos de modelagem de processos de negócios**. Rio de Janeiro: Brasport, 2017.

DRESCH, A.; LACERDA, D. P.; MIGUEL, P. A. C. A distinctive analysis of case study, action research and design science research. **Review of Business Management**, v. 17, n. 56, p. 1116-1133, 2015. DOI: <http://doi.org/10.7819/rbgn.v17i56.2069>.

DUBLIN CORE. **Dublin Core™ Metadata Element Set, Version 1.1**: reference description. 2012. Disponível em: <http://www.dublincore.org/specifications/dublin-core/dces>. Acesso em: 8 jun. 2022.

FRASER, M. S.; ANASTASELOS, T.; RAVIKUMAR, G. V. V. **The disruption in oil and gas upstream business by Industry 4.0**. Bengaluru: Infosys, 2018. Disponível em: <https://www.infosys.com/engineering-services/white-papers/documents/disruption-oil-gas-upstream.pdf>. Acesso em: 8 jun. 2022.

ISLAM, R.; ANANTHARANAMAN, M.; KHAN, F.; ABBBASSI, R.; GARANIYA, V. A hybrid human reliability assessment technique for the maintenance operations of

marine and offshore systems. **Process Safety Progress**, v. 39, n. S1, e12118, 2020. DOI: <https://doi.org/10.1002/prs.12118>.

JASIULEWICZ-KACZMAREK, M.; WASZKOWSKI, R.; PIECHOWSKI, M.; WYCZÓŁOKOWSKI, R. Implementing BPMN in maintenance process modeling. *In*: ŚWIĄTEK, J.; BORZEMSKI, L.; WILIMOWSKA, Z. (ed.). **Information systems architecture and technology**: Proceedings of 38th International Conference on Information Systems Architecture and Technology (ISAT 2017) – Part II. Cham: Springer, 2018. p. 300-309. (Advances in Intelligent Systems and Computing, v. 656.) DOI: [https://doi.org/10.1007/978-3-319-67229-8\\_27](https://doi.org/10.1007/978-3-319-67229-8_27).

JOSCHKO, P.; WIDOK, A. H.; PAGE, B.; APPEL, S.; GREINER, S.; ALBERS, H. Modeling and simulation of offshore wind farms including the mapping and analysis of relevant O&M processes. *In*: INTERNATIONAL CONFERENCE ON INFORMATICS FOR ENVIRONMENTAL PROTECTION, 27., 2013, Hamburg. **Enviroinfo 2013**: Environmental Informatics and Renewable Energies. Aachen: Shaker Verlag, 2013. Disponível em: <http://enviroinfo.eu/sites/default/files/pdfs/vol7995/0001.pdf>. Acesso em: 8 jun. 2022.

KISELEV, B.; YAKUTENKO, V. An overview of massive open online course platforms: personalization and semantic web technologies and standards. **Procedia Computer Science**, v. 169, p. 373-379, 2020. DOI: <https://doi.org/10.1016/j.procs.2020.02.232>.

LIBERATI, A.; ALTMAN, D. G.; TETZLAFF, J.; MULROW, C.; GOTZSCHE, P. C.; IOANNIDIS, J. P. A.; CLARKE, M.; DEVEREAUX, P. J.; KLEIJNEN, J.; MOHER, D. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. **PLoS Medicine**, v. 6, n. 7, e1000100, 2009. DOI: <https://doi.org/10.1371/journal.pmed.1000100>.

LOPES, D. C. F. **Grafos de conhecimento**: perspectivas e desafios para a organização e representação do conhecimento. 2020. Dissertação (Mestrado em Ciência da Informação) – Universidade Federal de São Carlos, São Carlos, 2020. Disponível em: <https://repositorio.ufscar.br/handle/ufscar/13055>. Acesso em: 8 jun. 2022.

OLIVEIRA, D. S.; GOMES, G. F. F.; PORTO, L. C. M. A importância da auditoria interna no processo de gestão das organizações em um ambiente globalizado e cada vez mais competitivo. **Revista de Ciências Gerenciais**, v. 1, n. 1, p. 1-19, 2012. Disponível em: <http://repositorio.enap.gov.br/handle/1/4739>. Acesso em: 8 jun. 2022.

POKRAEV, S. V. **Model-driven semantic integration of service-oriented applications**. Thesis (PhD on Telematics and Information Technology) – University of Twente, Enschede, 2009. Disponível em: <https://research.utwente.nl/en/publications/model-driven-semantic-integration-of-service-oriented-application-2>. Acesso em: 8 jun. 2022.

ROODE, M.; FERNÁNDEZ-IZQUIERDO, A.; DANIELE, L.; POVEDA-VILLALÓN, M.; GARCÍA-CASTRO, R. SAREF4INMA: A SAREF extension for the industry and manufacturing domain. **Semantic Web**, v. 11, n. 6, p. 911-926, 2020. DOI: <http://doi.org/10.3233/SW-200402>.

SANTARÉM SEGUNDO, J. E. Web Semântica: introdução a recuperação de dados usando SPARQL. *In*: ENCONTRO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIAS DA INFORMAÇÃO (ENANCIB), 15., 2014, Belo Horizonte.

**Anais** [...]. Belo Horizonte: BRAPCI, 2014. Disponível em: <https://brapci.inf.br/index.php/res/v/189508>. Acesso em: 8 jun. 2022.

SANTARÉM SEGUNDO, J. E.; CONEGLIAN, C. S.; LUCAS, E. R. O. Conceitos e tecnologias da Web semântica no contexto da colaboração acadêmico-científica: um estudo da plataforma Vivo. **Transinformação**, v. 29, n. 3, p. 297-309, 2017. DOI: <https://doi.org/10.1590/2318-08892017000300007>.

SANTOS, G. F. Z; KOERICH, G. V.; ALPERSTEDT, G. D. A contribuição do design research para a resolução de problemas complexos na administração pública. **Revista de Administração Pública**, v. 52, n. 5, p. 956-970, 2018. DOI: <https://doi.org/10.1590/0034-761220170014>.

SAUNDERS, M. N. K.; LEWIS, P.; THORNHILL, A. **Research methods for business students**. 8. ed. Amsterdam: Pearson, 2019.

SCHEMA.ORG. **Organization of schemas**. 2021. Disponível em: <https://schema.org/docs/schemas.html>. Acesso em: 17 jul. 2021.

SILVA, C. B. **Uma abordagem de modularização de ontologias baseada na satisfação lógica de questões de competência**. 2019. Tese (Doutorado em Ciência da Computação) – Universidade Federal de Pernambuco, Recife, 2019. Disponível em: <https://repositorio.ufpe.br/handle/123456789/35377>. Acesso em: 8 jun. 2022.

SLACK, N.; BRANDON-JONES, A.; JOHNSTON, R. **Administração da produção**. 8. ed. São Paulo: Atlas, 2018.

TORINO, E.; TREVISAN, G. L.; CONEGLIAN, C. S.; BOTEGA, L. C.; SANTAREM SEGUNDO, J. E.; VIDOTTI, S. A. B. G. Enriquecimento semântico para a disponibilização de dados abertos: teoria e prática. **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, v. 25, n. 1, p. 1-19, 2020. DOI: <https://doi.org/10.5007/1518-2924.2020.e67991>.

VITÓRIO, L. S. M.; NOGUEIRA, D. A. C. Gestão da manutenção e planejamento de parada programada de um navio plataforma, estocagem e produção de petróleo: Floating Production Storage Offloading (FPSO). **Leopoldianum**, ano 42, n. 116-118, p. 135-147, 2016. DOI: <https://doi.org/10.58422/releo2016.e692>.

W3C. **Semantic Web**. 2015. Disponível em: <https://www.w3.org/standards/semanticweb>. Acesso em: 20 ago. 2021.

W3C. **SPARQL 1.1 Query Language**. W3C Recommendation 21 march 2013. 2013a. Disponível em: <https://www.w3.org/TR/2013/REC-sparql11-query-20130321>. Acesso em: 8 jun. 2022.

W3C. **W3C Semantic Web Activity**. 2013b. Disponível em: <https://www.w3.org/2001/sw>. Acesso em: 8 jun. 2022.

W3C; OGC. **Time Ontology in OWL**. W3C candidate recommendation 26 march 2020. 2020. Disponível em: <https://www.w3.org/TR/owl-time>. Acesso em: 8 jun. 2022.