

Uma abordagem ontológica para dados de sensores e informacionais em plantas de produção

Lívia Cristina Silva do Nascimento¹, Mara Abel¹

¹Instituto de Informática– Universidade Federal do Rio Grande do Sul (UFRGS)– 91501-970– Porto Alegre– RS– Brazil

Abstract

This article presents a research proposal for the unification of the Informational Artifacts Ontology (IAO) and the Semantic Sensor Networks Ontology (SSN) to broaden the semantic coverage of Informational Content Entities (ICEs) in industrial contexts, especially in production plants. The proposed unification aims to enable a coherent and integrated representation of informational and sensory data, offering a comprehensive ontological structure encompassing formal records and dynamic data from sensors and monitoring systems. By associating sensory data with relevant informational artifacts, this approach seeks to improve interoperability, data analysis capabilities, and the traceability of industrial processes. The model will be evaluated using real industrial data, focusing on the oil and gas domain.

Keywords

information content entity, informational artifact ontology, ontology, ssn, sensor data

1. Introdução

Na era da digitalização e da Internet das Coisas (IoT), a quantidade de dados gerados por sensores e dispositivos conectados tem crescido exponencialmente. Estes dados são cruciais para a indústria, pois fornecem informações valiosas sobre operações, condições de máquinas, ambientes e processos em tempo real. No entanto, a simples coleta de dados não é suficiente; é necessário transformá-los em informações úteis e acionáveis.

Em processos industriais, dados são gerados por sensores a todo momento, e uma das maiores dificuldades no tratamento desses dados é garantir que os sistemas que os produzem e os que fazem uso deles sejam interoperáveis entre si [1]. No contexto das plantas de produção, a gestão eficiente dos artefatos de informação é fundamental para assegurar a fluidez dos processos e a integridade dos dados gerados ao longo das operações.

A anotação de dados gerados por sensores com significado semântico é essencial para as tarefas de automatização e interoperabilidade requeridas pela Indústria 4.0 [2]. Diante desse desafio, a aplicação de ontologias fornece um grande arcabouço para a representação e padronização do conhecimento. Ontologias oferecem esquemas de organização e recuperação de informação, pois permitem descrever as relações e propriedades das entidades.

Para a representação de informações de forma ontológica, a Informational Artifact Ontology (IAO) é uma abordagem promissora para a estruturação e padronização de artefatos e entidades que armazenam informações, como documentos, bancos de dados e imagens. A IAO foi criada com o objetivo de ser "um recurso de modelagem de domínio neutro para a representação de entidades de conteúdo informacional (ICE)" [3], explicando a relação existente entre as entidades informacionais e a realidade à qual estão se referindo, descrevendo ou falando sobre.

Manter registros é um conceito fundamental para salvar e recuperar informações. No contexto do gerenciamento de conhecimento, é importante entender como essas informações e conceitos se relacionam, bem como são especificados por meio de ontologias [4]. De forma ampla, a IAO pode ser

Proceedings of the 17th Seminar on Ontology Research in Brazil (ONTOBRAS 2024) and 8th Doctoral and Masters Consortium on Ontologies (WTDO 2024), Vitória, Brazil, October 07-10, 2024

✉ livia.nascimento@inf.ufrgs.br (L. C. S. d. Nascimento); marabel@inf.ufrgs.br (M. Abel)

🌐 <https://www.petwin.org/ontology/> (L. C. S. d. Nascimento); <https://www.petwin.org/> (M. Abel)

🆔 0000-0002-9589-2616 (M. Abel)



© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

utilizada na tarefa de modelar dados, descrevendo a relação entre entidades físicas e as informações sobre elas que são materializadas em meios físicos ou digitais. No entanto, essa abordagem apresenta limitações ao tratar de dados de sensores e da representação de medidas, uma vez que não inclui conceitos detalhados para modelar observações, amostras ou o processo de sensoriamento, que são centrais para a representação desses dados.

No que se refere à modelagem de dados de sensoriamento, o padrão de projeto da Semantic Sensor Network Ontology (SSN) é especializado nessa tarefa, sendo amplamente consolidado e utilizado para a representação de sensores, observações e medidas [5]. A SSN fornece uma especificação formal de uso geral para modelar a interação entre as entidades envolvidas nos atos de sensoriamento e amostragem [6], oferecendo uma estrutura detalhada que abrange desde a definição dos sensores até as características das observações e os resultados obtidos.

Este artigo apresenta uma proposta de investigação para a integração das ontologias IAO e SSN, com o objetivo de expandir a cobertura semântica das Entidades de Conteúdo Informacional (ICEs) em contextos industriais, especificamente em plantas de produção, onde o modelo construído será aplicado a dados do domínio de petróleo e gás. Com essa integração, visa-se possibilitar uma representação coerente e integrada de dados informacionais e sensoriais, proporcionando uma estrutura ontológica abrangente que abarca tanto registros formais quanto dados dinâmicos provenientes de sensores e sistemas de monitoramento. Adicionalmente, a proposta inclui a análise das relações estabelecidas entre as entidades informacionais e a realidade que elas representam, sejam estas entidades materiais ou imateriais, e como essas relações podem ser descritas.

2. Embasamento conceitual e teórico

A ciência da computação e a ciência da informação utilizam ontologias como modelos para a organização, representação e recuperação de informações digitais. Ontologias são estruturas formais que representam conhecimentos em domínios específicos, definindo conceitos, entidades e as relações entre eles de maneira explícita e compartilhável [7].

A Information Artifact Ontology ¹ (IAO) é uma ontologia desenvolvida para representar artefatos de informação, como documentos, bancos de dados e registros eletrônicos. Ela surgiu como parte dos esforços da Open Biomedical Ontologies (OBO) Foundry para criar um conjunto de ontologias interoperáveis [8]. A IAO se originou da necessidade de padronizar a representação de informações em diversas disciplinas, facilitando a troca e a integração de dados entre diferentes sistemas e organizações. Dessa forma, suas principais aplicações incluem a anotação e o gerenciamento de dados, especificando o modelo dos dados e suas relações com os objetos físicos aos quais expressam.

Em [3], são apresentados conceitos fundamentais sobre a representação de informação no âmbito das ontologias. As Entidades de Conteúdo Informacional (ICEs) são tipos mais gerais descritos na IAO que carregam ou expressam informação. As ICEs representam todo tipo de conteúdo informacional, como documentos, gráficos, mensagens, dados, especificações técnicas ou qualquer forma de informação que possa ser armazenada, transmitida ou processada. Neste trabalho, são apresentados os fundamentos da representação de informação, destacando-se a relação de *aboutness*, que liga a entidade informacional à entidade física à qual ela se refere.

Em [9], os autores apresentam uma discussão sobre questões que ainda estão em aberto no campo de estudo das entidades informacionais. Entre essas questões, destaca-se o desafio de explorar as bases ontológicas das entidades de informação para uma melhor compreensão e modelagem. Na IAO, quando as ICEs são definidas, determina-se que essas entidades podem existir em múltiplos suportes físicos, que são seus portadores, e modelar essa relação de forma consistente é complexo. Ao observarmos cenários reais na indústria, podemos identificar entidades que não possuem uma relação direta com uma entidade material, mas que se referem, na realidade, a outras entidades informacionais.

A Semantic Sensor Network Ontology ² (SSN) foi criada para fornecer uma estrutura semântica

¹Disponível em: <https://obofoundry.org/ontology/iao.html>

²Disponível em: <https://www.w3.org/TR/vocab-ssn/>

para a descrição de sensores, suas observações e os sistemas de sensores em rede [10]. Desenvolvida inicialmente pelo W3C Semantic Sensor Network Incubator Group, a SSN evoluiu para se tornar um padrão amplamente aceito para representar dados de sensores na web semântica. A SSN permite a integração e interoperabilidade de dados de sensores provenientes de diferentes fontes, facilitando a criação de aplicações que requerem monitoramento ambiental, controle industrial e gestão de recursos naturais [5]. Suas principais aplicações incluem redes de sensores para monitoramento ambiental, onde a SSN é usada para descrever e compartilhar dados sobre condições ambientais, como temperatura e qualidade do ar; a Internet das Coisas (IoT), onde ajuda a integrar dados de sensores de diversos dispositivos conectados; e sistemas de cidades inteligentes, onde a SSN contribui para a coleta e análise de dados de infraestrutura urbana para melhorar a gestão de serviços públicos [6].

3. Domínio de Aplicação

A Indústria 4.0 é caracterizada pela integração de tecnologia nos processos de manufatura e produção. Sensores desempenham um papel fundamental nessa transformação, pois são os dispositivos que permitem a coleta de dados em tempo real a partir do ambiente de produção, máquinas e sistemas. Esses dados são essenciais para a automação, monitoramento e otimização dos processos industriais.

Entidades de conteúdo informacional são geradas a todo momento por sensores, como dados de medidas de temperatura, vazão e pressão. Essas medições são conjuntos de dados marcados pelo tempo e podem ser recuperadas por meio de tags de historiadores industriais, que são projetados para lidar com a alta velocidade e o grande volume de dados gerados por equipamentos e sensores industriais, garantindo a integridade, segurança e acessibilidade dos dados.

Um tag é um identificador único atribuído a uma fonte de dados ou ponto de medição específico dentro do ambiente dos historiadores industriais. Cada tag representa uma variável de processo contínua ou discreta, como temperatura, pressão, velocidade, etc., e está associado a um fluxo de dados que pode ser monitorado e registrado continuamente ao longo do tempo. O tag permite o armazenamento, a recuperação e a análise de dados associados a esse ponto de medição específico, facilitando a gestão e a análise de grandes volumes de dados industriais.

Os registros feitos por sensores necessitam de um processo de coleta, que consiste na observação das características das entidades observadas. Nesse contexto, o padrão desenvolvido pela SSN define uma estrutura formal para descrever sensores, seus processos, observações e o contexto em que operam. Esta ontologia oferece uma base de modelagem que considera todo o processo de amostragem, modelando conceitos de sensores, redes de sensores e amostragem. Isso permite representar de forma consistente os processos e condições sob os quais os registros ocorrem no mundo real [11].

4. Trabalhos relacionados

A aplicação de uma camada semântica em dados é fundamental porque adiciona significado e contexto às informações, permitindo uma interpretação mais precisa e eficaz dos dados. Esta seção é dedicada a apresentar diferentes abordagens que tentam tratar o problema da semântica em dados de sensores.

Em [12] é apresentada uma abordagem para associar dados de sensores a ontologias de domínio. O processo envolve a classificação dos dados dos sensores, tornando-os instâncias da ontologia SSN (Semantic Sensor Network), com o objetivo de facilitar a integração, interoperabilidade e reutilização dos dados de sensores em diferentes domínios. Isso é feito estruturando formalmente um modelo semântico comum e, posteriormente, mapeando essas instâncias para conceitos na ontologia de domínio. Esta solução é proposta para tratar a heterogeneidade dos dados coletados por sensores, que pode variar em termos de formato, nomenclatura e estrutura dependendo do domínio de aplicação.

Em [13] é apresentada uma ontologia que é proposta por meio da extensão das ontologias SSN e SOSA, utilizando duas ontologias auxiliares para localização e tempo, com o objetivo de interagir e

gerenciar uma plataforma baseada no ThingsBoard³. A ontologia visa cobrir semanticamente elementos de rede, nós e gateways, assim como os dados produzidos pelos mesmos e sua criação é justificada através da necessidade de desenvolver uma plataforma capaz de controlar e atuar sobre diferentes dispositivos, além de considerar diferentes tipos de dados.

Em [14] é proposta uma ontologia específica para artefatos textuais, que leva em consideração a distinção entre a forma material e o conteúdo informacional das entidades de informação. Em [15], os autores concluem que, embora a IAO forneça uma estrutura útil para entender a informação em termos de representação e conteúdo, a ontologia de artefatos textuais oferece uma perspectiva adicional importante ao considerar o design e o uso dos artefatos de informação. Isso permite uma visão mais completa e prática de como as entidades de informação são criadas e utilizadas no mundo real.

5. Proposta de pesquisa

A proposta de pesquisa consiste em estender a IAO, que modela artefatos de informação como documentos, relatórios e bancos de dados, para tratar dados de sensoriamento através da integração com a SSN, que é especializada em descrever sensores, observações, amostras e os processos de sensoriamento em ambientes industriais. Essa integração permitirá uma representação semântica mais rica e detalhada, onde os dados sensoriais podem ser diretamente associados a documentos e registros de produção, manutenção e controle, aumentando a interoperabilidade e a capacidade de análise dos dados.

Objetivos:

- **Cobertura Semântica Ampliada:** A integração das ontologias poderá permitir uma modelagem mais completa dos ICEs, incluindo propriedades e relacionamentos das diferentes ICEs identificadas na planta de produção.
- **Interoperabilidade:** Facilitar a integração de sistemas de informação e sensoriamento em ambientes industriais, permitindo que dados de diferentes formatos sejam padronizados e tratados por meio da ontologia.
- **Rastreabilidade e Contextualização:** Melhorar a rastreabilidade dos processos industriais ao associar dados sensoriais diretamente com os artefatos informacionais relevantes, como relatórios de desempenho ou registros de manutenção.
- **Avaliação do Modelo:** Avaliar o modelo utilizando dados reais coletados da indústria.

Metodologia:

- **Revisão de Literatura:** Revisar a literatura para identificar abordagens existentes e lacunas na integração das ontologias IAO e SSN.
- **Definição do Modelo Ontológico:** Desenvolver um modelo preliminar integrando os conceitos e relações das ontologias IAO e SSN.
- **Implementação:** Implementar o modelo ontológico em OWL⁴, assegurando a representação coerente de dados informacionais e sensoriais.
- **Avaliação:** Testar a ontologia com dados reais do setor de petróleo e gás, avaliando a cobertura semântica, interoperabilidade e utilidade prática.

Resultados esperados:

- **Melhoria na Interoperabilidade:** Espera-se que a integração das ontologias melhore a interoperabilidade entre sistemas de informação e sensoriamento.
- **Rastreabilidade Aumentada:** A associação de dados sensoriais a artefatos informacionais deverá aumentar a rastreabilidade dos processos industriais.
- **Modelo Avaliado com Sucesso:** A ontologia deve ser validada com dados reais, comprovando sua eficácia na representação de dados informacionais e sensoriais, através de casos de uso.

³O ThingsBoard é uma plataforma de gerenciamento de dispositivos IoT de código aberto que permite que as empresas coletem, processem e visualizem dados em tempo real.

⁴A Web Ontology Language (OWL) é uma linguagem de representação de conhecimento utilizada para criar e compartilhar ontologias na web. Ler mais em: <https://www.w3.org/2001/sw/wiki/OWL>

6. Próximos passos

Para o desenvolvimento desta proposta, será necessário reforçar a revisão de literatura para uma maior fundamentação da proposta e compreensão dos conceitos centrais, bem como das limitações das ontologias IAO e SSN e das necessidades específicas do contexto industrial. Será preciso desenvolver um framework ontológico que combine os conceitos principais das ontologias IAO e SSN, seguindo uma metodologia de construção de ontologias e discutindo os termos-chave que a compõem. Por fim, será necessário avaliar o modelo criado com dados de sensoriamento coletados da indústria e analisar sua capacidade de representação de entidades informacionais.

7. Conclusão

Neste artigo, foi apresentada uma proposta de integração das ontologias IAO e SSN com o objetivo de ampliar a cobertura semântica das Entidades de Conteúdo Informacional (ICEs) em contextos industriais. A proposta visa investigar a representação consistente tanto dos dados informacionais quanto dos dados dinâmicos capturados por sensores e sistemas de monitoramento. Ao estabelecer uma ligação semântica clara entre os registros formais e as informações em tempo real, a integração das ontologias IAO e SSN contribui para uma representação mais rica e contextualizada das atividades e processos industriais.

Adicionalmente, a proposta também envolve a investigação das relações semânticas entre as entidades informacionais e as realidades que elas expressam, sejam materiais ou imateriais. Esse aprofundamento permitirá uma melhor compreensão das interações e dependências entre os elementos documentados e as condições operacionais monitoradas.

Nos próximos passos, pretende-se aplicar a ontologia unificada em testes baseados em dados no domínio industrial, especificamente no setor de petróleo, para validar o modelo. Dessa forma, espera-se que esta abordagem contribua para o avanço das metodologias de representação e gestão de dados nestes ambientes.

Agradecimentos

Os autores agradecem a CAPES-Brazil Finance Code 001, A agencia brasileira CNPq e ao projeto Petwin financiado pela FINEP e Libra Consortium (Petrobras, Shell Brasil, Total Energies, CNOOC, CNPC).

Referências

- [1] D. Šormaz, B. Kulvatunyou, M. Drobñjaković, S. Seeharit, Comparative study of approaches for an ontology of digital artifacts, in: International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, volume 87295, American Society of Mechanical Engineers, 2023, p. V002T02A057.
- [2] A. P. Vedurmudi, J. Neumann, M. Gruber, S. Eichstädt, Semantic description of quality of data in sensor networks, *Sensors* 21 (2021) 6462.
- [3] W. Ceusters, B. Smith, Aboutness: Towards foundations for the information artifact ontology (2015).
- [4] M. Löw, R. Rocha, Ontologias de documentos: Usos para a representação da informação, 2019. URL: <https://conferencias.ufsc.br/index.php/enancib/2019/paper/view/1229/817>.
- [5] A. Haller, K. Janowicz, S. J. Cox, M. Lefrançois, K. Taylor, D. Le Phuoc, J. Lieberman, R. García-Castro, R. Atkinson, C. Stadler, The modular ssn ontology: A joint w3c and ogc standard specifying the semantics of sensors, observations, sampling, and actuation, *Semantic Web* 10 (2019) 9–32.
- [6] K. Janowicz, A. Haller, S. J. Cox, D. Le Phuoc, M. Lefrançois, Sosa: A lightweight ontology for sensors, observations, samples, and actuators, *Journal of Web Semantics* 56 (2019) 1–10.
- [7] T. Gruber, What is an ontology, 1993.

- [8] E. C. Merrell, R. M. Kelly, Improving biomedical data by improving the information artifact ontology., in: ICBO, 2021, pp. 77–83.
- [9] E. M. Sanfilippo, Ontologies for information entities: State of the art and open challenges, *Applied ontology* 16 (2021) 111–135.
- [10] M. Compton, P. Barnaghi, L. Bermudez, R. Garcia-Castro, O. Corcho, S. Cox, J. Graybeal, M. Hauswirth, C. Henson, A. Herzog, et al., The ssn ontology of the w3c semantic sensor network incubator group, *Journal of Web Semantics* 17 (2012) 25–32.
- [11] K. Taylor, A. Haller, M. Lefrançois, S. J. Cox, K. Janowicz, R. Garcia-Castro, D. Le Phuoc, J. Lieberman, R. Atkinson, C. Stadler, The semantic sensor network ontology, revamped., in: JT@ ISWC, 2019.
- [12] J. Liu, Y. Li, X. Tian, A. K. Sangaiah, J. Wang, Towards semantic sensor data: an ontology approach, *Sensors* 19 (2019) 1193.
- [13] M. Vila, M.-R. Sancho, E. Teniente, X. Vilajosana, Semantics for connectivity management in iot sensing, in: *International Conference on Conceptual Modeling*, Springer, 2021, pp. 297–311.
- [14] H. Weigand, P. Johannesson, B. Andersson, An artifact ontology for design science research, *Data & Knowledge Engineering* 133 (2021) 101878.
- [15] H. Weigand, P. Johannesson, *Information entities and artifact ontology* (2020).