

ONTOLOGIA COMPUTACIONAL – UM MODELO DE DADOS PARA REPRESENTAÇÃO DO CONHECIMENTO

MACEDO, Rita de Cássia Catini de
Universidade de Mogi das Cruzes(UMC)
rita.catini@fatec.sp.gov.br

RODRIGUES, Silvia Cristina Martini
Universidade de Mogi das Cruzes(UMC)
silviac@umc.br

MACEDO, Paulo Cesar de
Faculdades Santa Lúcia
paulo.macedo@fatec.sp.gov.br

RESUMO

A representação e armazenamento do conhecimento computacional é comumente feita por meio dos bancos de dados tradicionais. Quando uma empresa inicia suas operações, normalmente são criadas bases de conhecimento para começar o negócio seja qual for a área. Com o advento da Internet e a grande demanda de usuários e informação, novos mecanismos precisaram ser criados, tanto para o compartilhamento quanto para agilizar o acesso evitando o retrabalho. Esse artigo apresenta uma linguagem da Web Semântica projetada para representar conhecimento rico e complexo sobre coisas, grupos de coisas e relações entre elas. Apresenta ainda a criação de uma ontologia computacional para melhor compreensão e entendimento do impacto desse assunto na maneira em que utilizamos a Internet. Espera-se que o leitor vislumbre como as chamadas Big techs trabalham com seus motores de busca do conhecimento e ainda como reaproveitar uma base de dados que já foi concebida nesse formato além de compartilhar o conhecimento armazenado nela.

PALAVRAS-CHAVE: *conhecimento; inteligência artificial; web semântica; engenharia de software; arquitetura da informação.*

INTRODUÇÃO

É notório que atualmente a tecnologia da informação está cada vez mais presente em nossa vida, seja por meio do computador, *smartphones*, eletrodomésticos inteligentes ou ainda invisíveis como sensores nos relógios e nos ambientes. É amplamente conhecido também que existe uma preocupação muito grande com relação de como as tecnologias podem nos ajudar na execução das mais diferentes tarefas cotidianas. As Redes Sociais têm uma grande parcela da difusão do conhecimento e uso dessas tecnologias, onde algumas empresas na área de entretenimento como *Netflix*, *Spotify*, entre outras, buscam facilitar a vida das pessoas oferecendo um certo nível de inteligência a fim de, em seus algoritmos, indicar baseado em conhecimento prévio, qual será o seu próximo filme, ou qual o seu gênero preferido de músicas a serem ouvidas ou ainda sua próxima compra. Na indústria não é diferente. A automação avança, e cada vez mais os dispositivos que outrora realizavam tarefas simples passam a necessitar de algoritmos inteligentes, evoluindo tecnologicamente. Na área da saúde existem inúmeros dispositivos médicos, aplicativos ou programas capazes de auxiliar o corpo médico ou o indivíduo a buscar uma melhor qualidade de vida. Percebe-se então que todas as áreas do conhecimento humano tendem a aderir ou ser influenciada de alguma forma pelo advento tecnológico atual.

O objetivo desse artigo é apresentar o conceito de ontologias e como elas podem nos ajudar na difusão do conhecimento de um domínio. Este artigo demonstra também a criação de uma ontologia parcial em formato próprio para uma experimentação desse contexto de forma a perceber a transição e influência que esse conceito citado pode causar.

Além dessa introdução este artigo está dividido em uma “Fundamentação teórica”, na qual se destaca de forma empírica, baseados em outros artigos, o estado da arte e apresenta as bases conceituais importantes para a compreensão do assunto, técnicas e soluções presentes, “Metodologia”, onde se relata a forma como a pesquisa será conduzida, em e em seguida um capítulo de “Experimento”, um cenário para um caso, onde utiliza-se como exemplo uma ontologia. Finalmente os resultados obtidos bem como as possibilidades de avanços para novos trabalhos nessa linha de pesquisa. (Grifos nossos).

2. ONTOLOGIAS

A *Internet* é uma ampla e crescente fonte de informações e serviços que precisam ser compartilhados por pessoas e aplicativos, a fim de disseminar conhecimento, notícias, informações etc. Nesse contexto, ou seja, na área da computação, as ontologias desempenham um papel importante no apoio à troca e ao compartilhamento de informações. Em linhas gerais, as ontologias *web* tem a capacidade de descrever formalmente objetos e suas características.

Para Roche (2003), uma ontologia é a conceituação de um domínio ao qual um ou vários vocabulários podem ser associados e que participam do significado dos termos. Definida para um determinado objetivo, uma ontologia expressa um ponto de vista compartilhado por uma comunidade. Uma ontologia é representada em uma linguagem (ontologia explícita) cuja teoria (semântica) garante as propriedades da ontologia em termos de consenso, coerência, compartilhamento e reutilização.

Ontologias podem ser formais (*hard ontologies*) ou suportar uma abordagem de interpretação aberta (*soft ontologies*). De acordo com Kaipainen *et al.* (2008), é contra a natureza dinamicamente evolutiva de muitas aplicações contemporâneas serem analisadas apenas em termos de ontologias formais convencionais. As chamadas *hard ontologies*, dependem de categorias fixas baseadas em expertise e estrutura hierárquica. Porém, existem contextos em que é difícil fazer uma estrutura fixa e permanente do conteúdo. O desenvolvimento de *soft ontologies* nos permite trabalhar com áreas onde é difícil estabelecer relações claras e organizar o conhecimento para criar ferramentas da *Web Semântica* (VERA *et al.*, 2013).

De acordo com Shih, Yang e Tseng (2011), tradicionalmente, a construção de ontologias é um processo demorado e árduo pois depende principalmente da contribuição de especialistas de domínio. Na tentativa de melhorar esse processo, pesquisas recentes voltaram-se para a construção e manutenção de ontologias automáticas e semiautomáticas.

Isotani e Bittencourt (2015) ainda destacam que existem duas maneiras de representar ontologias: a representação formal, usada para que as ontologias possam ser interpretadas por computadores. Para tanto, linguagens de descrição de ontologias têm sido propostas, tais como RDF e OWL mais bem explicadas a seguir. Outra maneira, é utilizar representação gráfica, usada para a compreensão humana, como por exemplo grafos, UML, estrutura de árvore, entre outras. Ambas são importantes, uma vez que a falta de uma dessas representações afeta a qualidade/uso da ontologia.

2.1 Resource Description Framework (RDF)

Na *web* semântica o objetivo é permitir que os computadores interpretem e explorem facilmente a *web*.

O *World Wide Web Consortium* (W3C) (2014) define que o *Resource Description Framework* (RDF) é uma estrutura para representar informações na *web*. A estrutura central do modelo de dados é um conjunto de triplas, cada uma consistindo de um sujeito, um predicado e um objeto (*s,p,o*).

Catarino e Souza (2012) afirmam que o modelo apresentado para a descrição de recursos RDF possibilita a representação dos metadados dos recursos *Web*, tais como título, autor, data de modificação de uma página da *Web*, *copyright* e informações sobre licenças de documentos *web* ou informações sobre o compartilhamento.

Pode-se concluir então que o RDF é uma estrutura para descrever recursos na *web* e foi projetado para ser lido e compreendido por computadores. Dessa forma, RDF não foi projetado para ser exibido para pessoas. Sabe-se também que essa estrutura é escrita em XML onde a sigla significa *eXtensible Markup Language*, que é basicamente um formato de arquivo universal usado para criar documentos com dados organizados.

O RDF faz parte da Atividade Semântica na *Web* do W3C e é recomendado pelo W3C desde 10 de fevereiro de 2004 (W3C, 2014).

2.2 Ontology Web Language (OWL)

Na ciência da computação, uma ontologia define e/ou especifica os conceitos, relacionamentos e outras distinções que são relevantes para a modelagem de um domínio. Essa especificação assume a forma de definições de vocabulário representacional (classes, relações e assim por diante), que fornecem significados para o vocabulário e restrições formais em seu uso coerente. A *Web* Semântica por sua vez é conduzida pelo *World Wide Web Consortium* e é projetada geralmente por meio de sintaxes que usam *Uniform Resource Identifiers* (URIs) para representar os dados. A *Web* Semântica é uma *Web* de Dados - de datas, títulos, números de peças, propriedades químicas e quaisquer outros dados que alguém possa conceber. A coleção de tecnologias da *Web* Semântica envolve o conceito de *Ontology Web Language* (OWL) conhecida como Linguagem de Ontologia da *Web* e fornece um ambiente onde qualquer aplicativo, desde que aponte (direcione) a ela, possa consultar esses dados e fazer inferências usando vocabulários, por exemplo (W3C, 2014).

Para Liu (2011), OWL é uma linguagem de marcação semântica para publicação e compartilhamento de ontologias, projetada para descrever classes e relacionamentos entre elas, que é atualmente recomendada pelo W3C.

Já para Isotani e Bittencourt (2015), em uma definição mais sucinta, cita a OWL como uma linguagem para definição e instanciação de ontologias *Web*.

Percebe-se que uma ontologia OWL pode formalizar um domínio, definindo classes e propriedades destas classes, definir indivíduos e afirmações sobre eles e, usando-se a semântica formal OWL, especificando como derivar consequências lógicas, isto é, fatos que não estão presentes na ontologia, mas são vinculados pela semântica. Foi projetada para ser usada por aplicações que necessitem processar o conteúdo de informações, ao invés de somente apresentar a visualização destas informações, como os tradicionais bancos de dados.

3. METODOLOGIA

Como elencado anteriormente, percebe-se a existência de inúmeras regras para criação de ontologias que devem ser consideradas. Sua representação está diretamente ligada a orientação a objetos no intuito a facilitar a implementação pelos desenvolvedores de *software*. Uma ontologia deve considerar um domínio e uma hierarquia de classe e subclasses desse domínio, além de suas relações, ou seja, associações que podem apresentar um conjunto de instâncias derivadas do conhecimento.

Importante notar que a base de dados representada em uma ontologia deve ser conectada, altamente compartilhável, que permita a interoperabilidade e a possibilidade de lidar com o acúmulo de conhecimento.

Em muitas áreas do conhecimento, existem inúmeras aplicações que precisam da criação de bancos de dados e seu preenchimento por meio de digitação manual e esse artigo não propõe uma forma de substituir essa criação e sim aproveitar as bases já existentes, como as bases de conhecimento ontológicas, conceito estudado na área de dados abertos conectados.

O uso de uma ontologia pode ser feito de duas maneiras: *i- online*: indicando o endereço do repositório onde ela se encontra. Permite a consulta dos dados da base de conhecimento sempre em sua versão mais atual, porém é totalmente dependente de acesso à *Internet*; ou *ii- offline*: a partir do *download* da base de conhecimento para uso local e independente de uma conexão com a *Internet* para consulta dos dados. Porém, neste caso, eventuais atualizações dos dados da base de conhecimento local devem

ser feitas manualmente ou com um novo *download* do arquivo disponível na *web*. Internamente as informações, análogas às tabelas de um banco de dados tradicional, são ligadas por meio de *hyperlinks* URI, já detalhados anteriormente.

Para exemplificar essa utilização, o presente trabalho propõe a criação de uma base de apoio a uma loja de bolos que poderia ser aproveitada simplesmente apontando qualquer aplicação a uma ontologia previamente criada, sem a necessidade da digitação e evitando o retrabalho. Para tanto, os seguintes passos e métodos foram empregados na construção da solução proposta neste artigo:

- Estudo preliminar sobre o tema para criação do conhecimento da base, como: tipo de bolos, ingredientes, recheios entre outros.
- Modelagem das classes de objetos, instâncias e suas relações;
- Escolha da ferramenta adequada para construção da ontologia;
- Construção da Ontologia e exportação em formato compatível.

3.1 Experimento

Como parte deste trabalho, propõe-se a criação de um protótipo que permita entender como seria a aplicação real de uma ontologia, bem como validar as configurações adequadas, seguida das interfaces necessárias para o entendimento do funcionamento do algoritmo.

a. Criando a ontologia

Para a criação da ontologia foi utilizado o *software* Protégé¹, uma distribuição gratuita disponível na *Web*. Na **Figura 1** pode-se notar que foram criadas entidades para suportar o entendimento básico em que um *Confeiteiro* necessita para indicação da receita do bolo, mostra ainda as principais classes e subclasses criadas na ontologia Ontobolo, e a relação hierárquica entre elas.

b. Apresentação da Ontologia

O bolo é um tipo de alimento feito à base de farinha, geralmente com açúcar e ovos, além de outros ingredientes, geralmente de forma arredondada ou retangular, cozida ou assada (MICAHELIS, 2022).

¹O Protégé é um editor de ontologias gratuito e de código aberto e um sistema de gerenciamento de conhecimento, desenvolvido pela Universidade de Stardford.

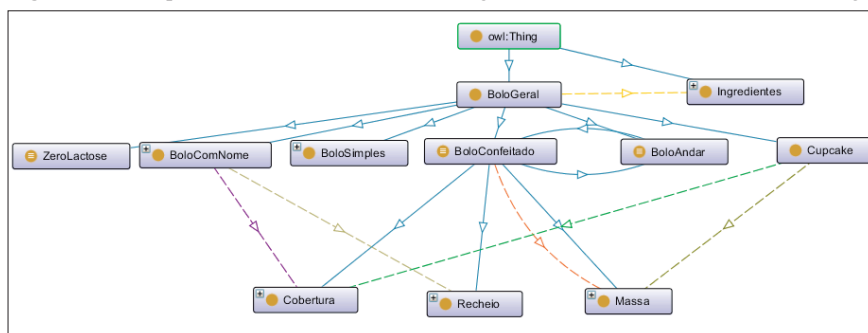
A ontologia de bolos (Ontobolo) foi criada para representar o conhecimento relacionado ao domínio de bolos. O objetivo da ontologia é servir de base de conhecimento ontológico e auxiliar aplicações de padarias, confeitarias e supermercados ou qualquer outra aplicação ligada à sua área de conhecimento. Na ontologia é possível a identificação e classificação de diversos tipos e bolo, desde os simples à bolos confeitados, seus ingredientes e combinações de coberturas, massas e recheios para fazê-los.

c. Descrição da Ontologia:

Para a construção da ontologia Ontobolo, inicialmente foi definido que o domínio que seria coberto por ela seria o de bolos doces em geral e foi feito o desenho conceitual da ontologia. A efetiva criação da ontologia foi feita na ferramenta Protégé (MUSEN, 2015) e seguiu os seguintes passos:

- I. Definição das classes;
- II. Especificação das propriedades;
- III. Criação de indivíduos;
- IV. Especificação de restrições.

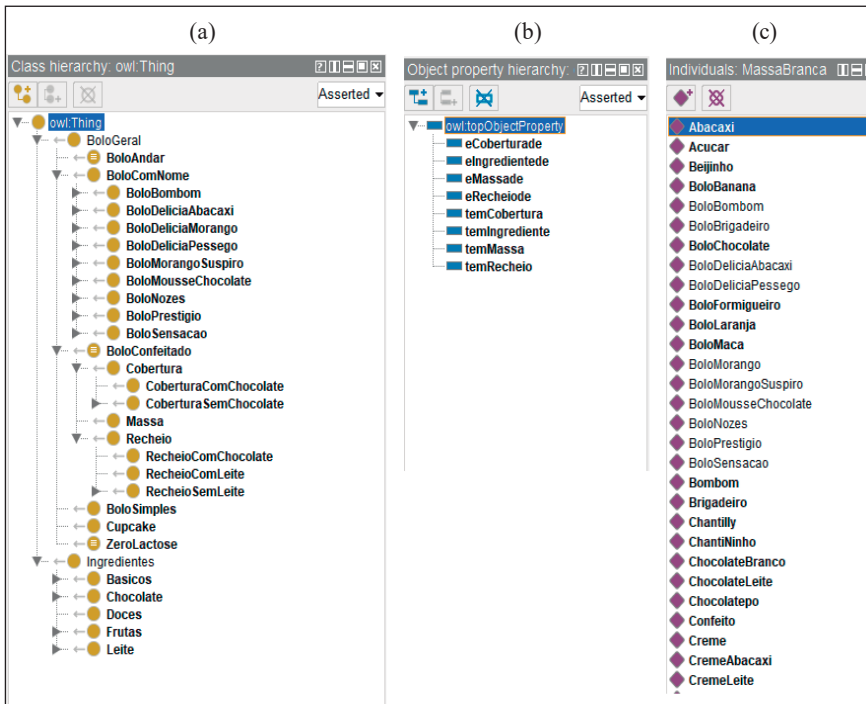
Figura 1 – Principais classes e subclasses da ontologia Ontobolo, criada na ferramenta Protégé.



Na ontologia, foram também criadas diversas propriedades de objetos, para definir as relações e restrições entre as classes e instâncias das classes, para a representação dos objetos individuais do domínio.

Com o objetivo de facilitar a compreensão dos exemplos, detalhes dos elementos da ontologia, como a) hierarquia de classes, b) propriedades de objetos e c) instâncias das classes, podem ser visualizados na **Figura 2**.

Figura 2 – a) Hierarquia de classes, b) propriedades de objetos e c) instâncias das classes da ontologia Ontobolo.



A seguir serão apresentados alguns exemplos de propriedades, restrições e relações entre classes que foram definidas na ontologia Ontobolo.

d. Exemplos de Propriedades Atribuídas:

- Inversas:
 - “RecheioComChocolate” \leftrightarrow Chocolate “temIngrediente” e “IngredienteDe”
- Assimétrica:
 - “temLactose” é uma propriedade assimétrica do domínio Ingrediente para os *ranges* Leite e Chocolate.

e. Exemplo de Restrição:

- Os bolos confeitados devem ter 1 tipo de cobertura, 1 tipo de massa e até 2 recheios (de qualquer tipo);

f. Exemplo de Relação de Classes:

- “BoloAndar” é equivalente à “BoloConfeitado” e ambos são disjuntivos de “BoloSimples”;
- “Bolo Mousse de Chocolate” é disjuntivo dos demais bolos com nomes e é subclasse de “BoloComNome”, “MassaChocolate”, “CoberturaComChocolate” e “RecheioComChocolate”.
- A classe “ZeroLactose” corresponde a qualquer bolo cuja propriedade “temIngrediente” não tenha qualquer derivado de Leite ou Chocolate.

Após os testes pode-se constatar as seguintes vantagens com relação ao desenvolvimento tradicional:

- Desenvolvimento apenas das interfaces e regras de negócio: o desenvolvedor pode se concentrar no desenvolvimento da ferramenta com foco em suas funcionalidades e sem se preocupar com a parte relacionada aos dados do produto “bolo”.
- Aproveitamento de dados e menos retrabalho: esse cenário permitiria que qualquer empresa relacionada a esse ramo de atuação pudesse aproveitar os dados e ligações, fornecidos pela base de conhecimento ontológica, sem ter que redigitar tudo.
- Evitar a cópia de conhecimento: não se trata de copiar a base de uma loja para outra, significa que todas as lojas de bolo que quiserem podem apontar para mesma base de conhecimento.
- Economia no trabalho profissional: as empresas que optarem por soluções desse tipo, estarão economizando em horas de planejamento e modelagem dos dados pelo profissional da área de banco de dados.

CONSIDERAÇÕES FINAIS

A ontologia criada nesse artigo permitiu demonstrar como os objetos e instâncias podem ser criados e acessados sem necessidade de retrabalho.

Percebe-se, nesse sentido, no ponto de vista técnico, que algumas vantagens podem ser elencadas, tais como o compartilhamento dos dados que podem transitar de um negócio para outro, a possibilidade de ter uma base de conhecimento sempre atualizada, considerando que, se a ontologia for acessada de forma *online*, o criador pode acrescentar novas informações para que todos aproveitem, além da independência da Internet (consumo local), caso a empresa opte por fazer o *download* da ontologia pronta.

Na *Internet*, existem ontologias já criadas que podem ser consumidas e adaptadas para diversas áreas do conhecimento. Empresas como *Google*,

Microsoft e Amazon, já utilizam esse conceito em seus motores de busca e aplicações de vendas, além de outras empresas menores que utilizam para os mais diversos setores.

O fato de a W3C considerar importante e disponibilizar o aprendizado desse assunto em seus conteúdos, indica que, em um futuro próximo, mais desenvolvedores estarão aptos a utilizar e oferecer soluções mais inteligentes a seus consumidores.

Pode-se concluir que o conceito de ontologias é promissor, já impacta a área de desenvolvimento positivamente. Espera-se que, em pesquisas futuras, possa ainda juntar conhecimento por meio de ontologias flexíveis (*soft ontologies*), nas quais os conceitos podem tomar formas diferentes dependendo da situação do ambiente, ou do sentimento das pessoas, influenciado por sensores.

REFERÊNCIAS

CATARINO, M. E.; DE SOUZA, T. B. A representação descritiva no contexto da web semântica. *Transinformação*, v. 24, n° 2, p. 77-90, 2012.

ISOTANI, S.; BITTENCOURT, I.I. **Dados Abertos Conectados**: Em busca da Web do Conhecimento. Novatec Editora, 2015.

KAIPAINEN, M. *et al.* Soft ontologies, spatial representations and multi-perspective exploitability. *Expert Systems*, v. 25, n° 5, p. 474-483, 2008.

LIU, C. *et al.* Fuzzy reasoning over RDF data using OWL vocabulary. In: **Proceedings of the 2011 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology**-Volume 01. IEEE Computer Society, 2011. p. 162-169.

MICHAELIS. **Dicionário**. Disponível em <https://michaelis.uol.com.br/modernoportugues/busca/portugues-brasileiro/bolo/>. Acesso em abril de 2022.

MUSEN, M. A. **The Protégé project**: A look back and a look forward. *AI Matters*. Association of Computing Machinery Specific Interest Group in Artificial Intelligence, 1(4), June 2015. DOI: 10.1145/2557001.25757003.

PROTÉGÉ. **Protégé Ontology Editor**. Disponível em: <https://protege.stanford.edu/>. Acesso em fevereiro de 2022.

ROCHE, C. Ontology: a survey. In: **8th Symposium on Automated Systems Based on Human Skill and Knowledge IFAC**. 2003.

SHIH, W.; YANG, C.; TSENG, S. Fuzzy folksonomy-based index creation for e-learning content retrieval on cloud computing environments. In: **Fuzzy Systems (FUZZ)**, 2011 IEEE International Conference on. IEEE, 2011. p. 965-970.

VERA, M. S. *et al.* Practical experiences for the development of educational systems in the semantic web. **Journal of New Approaches in Educational Research (NAER Journal)**, v. 2, n° 1, 2013.

W3C. World Wide Web Consortium *et al.* **RDF 1.1 concepts and abstract syntax**. 2014. Disponível em <https://www.w3.org/RDF/>. Acesso em junho de 2022.

