



PROPOSIÇÃO DE UM MODELO DE ÍNDICE DE COMPARAÇÃO SEMÂNTICA BASEADO EM ONTOLOGIAS PARA APOIO AO DOCENTE NA AVALIAÇÃO DO ESTUDANTE

Israel Alves Rocha¹, José Leonardo Oliveira Lima¹
israel_iarb@hotmail.com, jjleo@ueg.br (<https://orcid.org/0000-0001-8869-3056>)

¹ Universidade Estadual de Goiás, Sistemas de Informação, Anápolis, Goiás

RESUMO: A Educação a Distância evidencia um desafio para o docente no que se refere à avaliação da aprendizagem dos discentes por meio dos trabalhos feitos baseados nos materiais disponibilizados para estudo. Essa pesquisa tem como objetivo propor um modelo de um índice de similaridade semântica, baseado em ontologias, de forma a auxiliar o docente a verificar o uso do material indicado para estudo pelo docente nas produções escritas feitas pelos discentes. A pesquisa é exploratória e metodológica, com caráter bibliográfico e de pesquisa de laboratório. O estudo está em desenvolvimento e já foram levantadas metodologias baseadas no uso de ontologias, que se aproximam a proposta da pesquisa. Os estudos preliminares possibilitam inferir que o modelo que está sendo idealizado se mostra viável, mas são necessárias análises detalhadas para compreender a forma como o índice deve ser construído e como deve ser a comparação semântica.

Palavras-Chave: Educação a distância, Ontologia, Similaridade semântica, Comparação semântica, Domínio de conhecimento.

PROPOSAL OF A SEMANTIC COMPARISON INDEX MODEL BASED ON ONTOLOGIES TO SUPPORT TEACHERS IN STUDENT ASSESSMENT

ABSTRACT: Distance Education shows a challenge for the teacher while evaluating the learning of students through their writing work done based on the materials made available for studying. This research aims to propose the model of a semantic similarity index based on domain ontologies, in order to help the teacher to verify the presence of the material provided for studying in the written productions made by the students. The research is exploratory and methodological, with a bibliographic and laboratory research feature. The study is under development and methodologies based on the use of ontologies have been raised, which are closer to the research proposal. Preliminary studies make it possible to infer that the model being idealized proves to be viable, but detailed analyzes are needed to understand how the index should be built and how the semantic comparison should be.

Keywords: Distance education, Ontology, Semantic similarity, Semantic comparison, Knowledge domain.



1. INTRODUÇÃO

A rápida expansão da Educação a Distância (EaD) no mundo, verificada de forma expressiva no Brasil somente na última década, é a realidade atual de grande parcela dos alunos de ensino superior. Ações como a adoção do Ensino Remoto Emergencial (ERE) foram utilizadas para minimizar o impacto causado no ensino durante a pandemia de Covid19, provocando o aumento no uso de Tecnologias de Informação e Comunicação (TIC) como mediadores das dinâmicas educacionais, que ocorrem por meio das videoaulas e dos Ambientes Virtuais de Aprendizagem (AVA) (BROILO; NETO, 2021, LIMA, 2016, SANTOS, 2009).

A modalidade de ensino híbrida identificada no ERE e a modalidade EaD, exigem do aluno o estudo dos materiais, além do acompanhamento de aulas síncronas e assíncronas, disponibilizadas pelo professor, que servirão como base para a construção do conhecimento relativo ao objeto de estudo da disciplina. Segundo Lima (2016), o processo de avaliação e acompanhamento por parte do docente em ambientes *online* tem se tornado cada vez mais desafiador pois, devido à grande carga de trabalho em virtude do grande número de alunos, torna-se complexo identificar por exemplo, se o estudante está construindo conhecimento por meio dos materiais disponibilizados para estudo.

A Educação Híbrida e a Educação a Distância são duas modalidades de ensino distintas, mas que apresentam similaridades na utilização das Tecnologias de Informação e Comunicação (TICs) e dos Ambientes Virtuais de Aprendizagem.

Considerando a dificuldade exposta na avaliação da construção do conhecimento por parte do aluno, seja na educação híbrida ou na EaD, urge a necessidade da criação de mecanismos de auxílio ao docente na verificação da construção de conhecimento do aluno com o material didático desenvolvido e/ou indicado pelo professor para estudos de um determinado tema. Portanto, a pesquisa tem como objetivo o desenvolvimento de um modelo de índice de comparação semântica baseado em ontologias para apoio à avaliação docente.

2. REVISÃO BIBLIOGRÁFICA

Campos (2004) aponta que o processo de conhecimento é equivalente à construção de modelos do mundo/domínio, que propiciam descrever e explicar os fenômenos que são observados, como uma forma de conhecer e classificar o mundo.

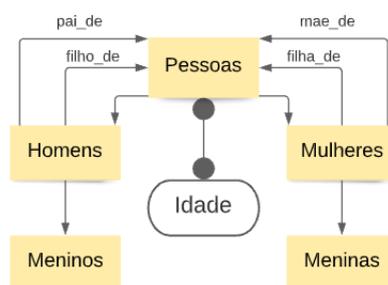
A representação do conhecimento busca compreender, representar e descrever o ser. As diversas formas de representações foram fundamentais para perpetuar o conhecimento (LIMA; ALVARES, 2012).

Conforme Campos e Marcondes (2008), os sistemas de inteligência artificial consideram que, tudo aquilo que pode ser representado, existe; dessa forma, grupos de pesquisadores trabalham para sistematizar e estruturar em domínios específicos chamados de Ontologias, o conhecimento científico.

2.1 Ontologia

Pickler (2007) discorre que o termo ontologia foi introduzido na filosofia como forma de distinguir o estudo do ser do estudo dos vários tipos de seres vivos existentes, fornecendo categoricamente um sistema para organização da realidade. Pickler (2007) escreve também que o termo foi adaptado para o contexto da Web e Inteligência Artificial (IA) como uma forma de representação que favorece a proposição de relações entre termos e conceitos.

Figura 1 – Exemplo estrutura de Ontologia sobre o domínio Pessoa



Fonte: Elaborado pelo autor adaptado de Lopes e Vieira (2010).

Pickler (2007) destaca que a natureza da Ontologia para a Ciência da Computação, é de instrumento de representação do conhecimento, tratando a informação e tornando inteligível ao computador, visando a inferência automatizada.

2.2 Redes Semânticas

Ao estudar um texto, o leitor busca a compreensão do conteúdo por meio do uso do conhecimento prévio existente. Djik e Kintsch (1983 apud GIRALDELLO, 2017, p. 25) afirmam que a compreensão se vale não somente da representação de uma base textual, mas, também, do chamado modelo situacional. A base textual remonta a ideias que se relacionam com situações do texto, permitindo que seja construído um modelo situacional; são esses dois modelos que permitem ao leitor compreender a mensagem transmitida.

Giraldello (2017), explica que durante a leitura de um texto, a compreensão do texto significa representar de forma mental e progressiva esse texto em redes semânticas, utilizando-se de forma constante o uso de conhecimento e processos inferenciais que levam o leitor a compreender o que está sendo estudado.

Para Cunha (2013), uma rede semântica pode ser modelada por grafos, onde cada vértice dessa rede representam as palavras, e as arestas representam as ligações entre essas palavras. Cunha explica também que se analisadas palavras que ocorrem juntas numa mesma sentença, poderá se constatar que são evocadas de forma a construir uma ideia por meio da associação.

2.3 Similaridade Semântica

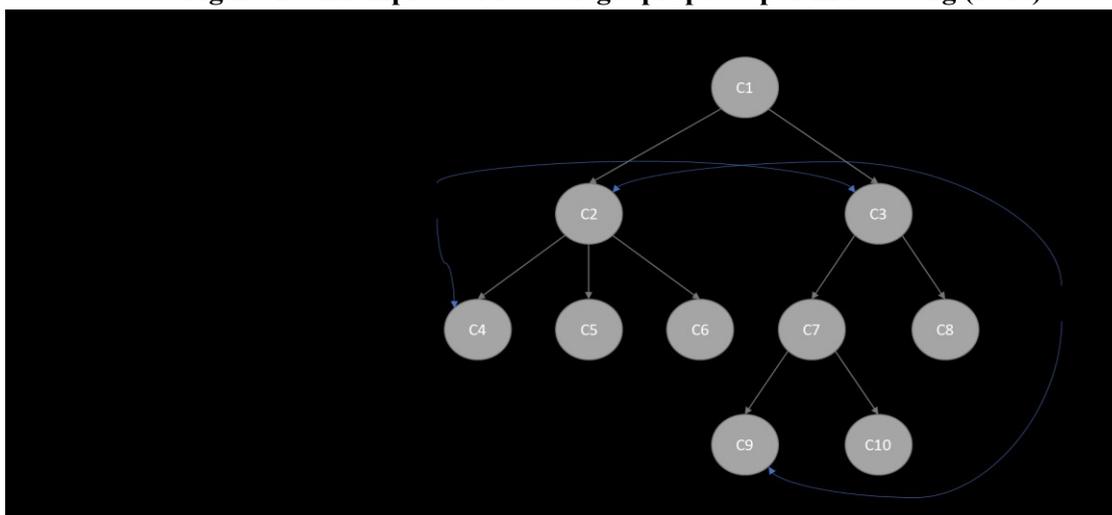
À medida que os estudos com uso das redes semânticas são aplicados nos processos de

tratamento e recuperação de informação, emerge a possibilidade de identificar a similaridade entre o conteúdo desses textos, verificando a estrutura textual e a construção semântica.

Conforme Arruda (2017), a similaridade é a medida utilizada na verificação da correspondência entre os pares de palavras, sentenças ou documentos. A verificação da similaridade semântica advém do uso de uma base de dados semântica, podendo ser uma ontologia, ou outro tipo base de dados de mesmo gênero.

A proposta de Liu e Wang (2014) trabalha com quatro passos definidos para o cálculo da similaridade semântica entre sentenças de textos, o primeiro passo identifica os termos chaves da sentença (T1, T2 para sentença 1 e Ta, Tb para a sentença 2), o segundo passo mapeia os nós da rede semântica verificando os nós próximos aos termos chaves, o terceiro passo forma o vetor considerando o peso (p) dos termos ($0 < p < 1$), por fim, o quarto passo calcula a similaridade dos vetores (V1 e V2) pela similaridade de cossenos.

Figura 2 – Exemplo da Metodologia proposta por Liu e Wang (2014)



Fonte: Elaborado pelo autor adaptado de Liu e Wang. (2014).

3. METODOLOGIA

Seguindo classificação proposta por Vergara (2016) e Marconi e Lakatos (2007), o tipo de pesquisa quanto aos fins apresenta características de uma pesquisa exploratória e metodológica. Quanto aos meios, a pesquisa apresenta caráter bibliográfico e laboratorial.

Para o levantamento de dados, foram utilizadas as plataformas de busca SciELO, Google Acadêmico, Periódicos da CAPES, além de plataformas como a Minha Biblioteca, entre outras. Os materiais levantados estão sendo tratados de forma qualitativa, em conformidade com a análise e síntese dos materiais, observando a relevância para a proposta dessa pesquisa.

A pesquisa apresenta caráter laboratorial, pois, contempla a execução de testes em situações

controladas, observando e analisando os resultados obtidos, esperados e inesperados.

4. RESULTADOS PRELIMINARES

Com base no trabalho já realizado, a proposição do modelo de índice de verificação da similaridade semântica com o uso de ontologias é possível observando as metodologias propostas por Liu e Wang (2014), e Arruda (2017), baseadas em ontologias, que têm mostrado maior aderência quanto aos requisitos da pesquisa.

O modelo de índice, em processo de estudo e desenvolvimento tem o intuito de identificar a similaridade semântica entre pares de textos. Na atual fase de estudo, já emerge pontos importantes para o modelo. O modelo requer a obtenção da estrutura semântica dos pares de documentos a serem verificados. Outro ponto de relevância observado é o cálculo do índice de similaridade semântica, que permite verificar a proximidade dos conteúdos.

5. CONSIDERAÇÕES FINAIS

Levando em consideração o caminho percorrido até então, é possível vislumbrar o modelo do índice de similaridade semântica, sendo que as metodologias estudadas apresentam grande aderência à proposta, de forma a tornar viável o modelo a ser proposto, objeto do presente estudo.

REFERÊNCIAS

ALVARENGA, Lídia. Representação do conhecimento na perspectiva da ciência da informação em tempo e espaço digitais. **Encontros Bibli: Revista eletrônica de biblioteconomia e ciência da informação**, 2003, v. 8, n. 15, p.18-40. <https://doi.org/10.5007/1518-2924.2003v8n15p18>.

ARRUDA, Claudineia Gonçalves de. **SOM4SimD**: um método semântico baseado em ontologia para detectar similaridade entre documentos. 2017. Dissertação (Mestrado em Ciência da Computação) – Universidade Federal de São Carlos, São Carlos, 2017. Disponível em: <https://repositorio.ufscar.br/handle/ufscar/8961>. Acesso em: 21 jun. 2021.

BROILO, Liane; NETO, Gilberto Broilo. Pandemia 2020 e a EaD: O impacto do Covid-19 no ensino brasileiro. **Revista Ecom**, [s. l.], 1 jan. 2021. Disponível em: <http://www.publicacoes.fatea.br/index.php/ECCOM/article/view/1238>. Acesso em: 21 jun. 2021.

CAMPOS, M. L. de A.; Modelização de domínios de conhecimento: uma investigação de princípios fundamentais. **Ciência da Informação**, [S. l.], abr. 2004. Disponível em: <https://www.scielo.br/j/ci/a/CyYd3Km3xzTdmf5DzxxQd3h/abstract/?lang=pt>. Acesso em: 05 ago. 2021.

COSTA, Douglas de Jesus. **Análise de algoritmo de comparação semântica de textos**: Uma abordagem sob a ótica cognitivista de Piaget. 2017. 120 p. Monografia (Curso de Sistemas de Informação) - Universidade Estadual de Goiás, Anápolis, 2017.



CUNHA, M. DO V. Redes semânticas baseadas em títulos de artigos científicos. **Repositório da Produção Científica e Intelectual**, [s.l.], nov, 2013. Disponível em: <http://repositoriosenaiba.fieb.org.br/handle/fieb/747>. Acesso em: 21 jun. 2021.

GIRALDELLO, Ademir Paulo. **Compreensão leitora: Como se (re)constrói a significação textual?** Disponível em: https://core.ac.uk/display/235124863?utm_source=pdf&utm_medium=banner&utm_campaign=pdf-decoration-v1. Acesso em: 6 set. 2021.

LIMA, J. L. O. **Avaliação discente em cursos de graduação a distância mediados por ambientes virtuais de aprendizagem:** Necessidade de informações dos docentes na visão de especialistas europeus e brasileiros. 2016. 298f. Tese (Doutorado em Ciência da Informação) – Universidade de Brasília, Brasília, 2016.

LIMA, J. L. O.; ALVARES, L. Organização e representação da informação e do conhecimento. *In:* ALVARES, L. (org.). **Organização da informação e do conhecimento:** conceitos, subsídios interdisciplinares e aplicações. São Paulo: B4 Editores, 2012. P. 21-48.

LIU, H.; WANG, P. Assessing Text Semantic Similarity Using Ontology. **Journal of Software**, v. 9, n. 2, p. 490–497, 2014. Disponível em: <http://www.jsoftware.us/vol9/jsw0902-32.pdf>. Acesso em: 2 ago. 2021.

LOPES, L.; VIEIRA, R. Processamento de linguagem natural e o tratamento computacional de linguagens científicas. *In:* PERNA, Cristina Lopes; DELGADO, Heloísa Koch; FINATTO, Maria José (Orgs.). **Linguagens especializadas em corpora:** modos de dizer e interfaces de pesquisa, p. 183–201., EDIPUCRS, 2010.

MARCONDES, Carlos Henrique; CAMPOS, Maria Luiza de Almeida. Ontologia e web semântica: o espaço da pesquisa em ciência da informação. **PontodeAcesso**, Salvador, p. 107-136, jun. 2008. Disponível em: <https://periodicos.ufba.br/index.php/revistaici/article/view/2669>. Acesso em: 21 jun. 2021.

MARCONI, Marina de Andrade; LAKATOS, Eva Maria. **Metodologia científica:** ciência e conhecimento científico, métodos científicos, teoria, hipóteses e variáveis. 5. ed. São Paulo: Atlas, 2007.

PICKLER, Maria Elisa Valentim. Web Semântica: ontologias como ferramentas de representação do conhecimento. **Perspectivas em ciências da informação**, [s. l.], p. 65-83, abr. 2007. Disponível em: <https://www.scielo.br/j/pci/a/HHdw6KMMPG45HxwShcwTmFSs/?lang=pt>. Acesso em: 21 jun. 2021.

SANTOS, Edméa. Educação online para além da EaD: um fenômeno da cibercultura. **X Congresso Internacional Galego-Português de Psicopedagogia**, [s. l.], set. 2009. Disponível em: <https://www.educacion.udc.es/grupos/gipdae/documentos/congreso/Xcongreso/pdfs/t12/t12c427.pdf>. Acesso em: 21 jun. 2021.

VERGARA, S. C. **Projetos e relatórios de pesquisa em Administração**, 16ª edição. Grupo GEN, 2016. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788597007480/>. Acesso em: 02 Jun 2021.