UMA LEITURA SOBRE AS NOVAS DIREÇÕES PARA A CIÊNCIA DA INFORMAÇÃO

Áquila Sartori Mesquita Rocha
Mestranda | UFCG
aquila.sartori@estudante.ufcg.edu.br

MANDL, Thomas. **Novas direções para a Ciência da Informação**: inteligência artificial e processamento de textos. In: Projeto SESA ON - LINE. 2020. Disponível em https://www.youtube.com/watch?v=HlUjCkc5TRo&list=PLxN9uB8ODjKmszvX0wANWVEQ0Zx6X8HYw&index=18&t=3487s. Acesso em 23 de dezembro de 2021.

O projeto de pesquisa e extensão SESA (Seminário dos Saberes Arquivísticos) tem o objetivo de desenvolver atividades que possam qualificar o estudante, dando-lhe oportunidade de atuar em práticas sociais em Instituições de Ensino Superior - IES. Partindo desse pressuposto, surge o SESA *On-line*, que estabelece uma cooperação acadêmica com relações internacionais, via plataformas digitais e ambientes virtuais de aprendizagem.

O SESA *On-line* nasceu durante a época de pandemia da Covid-19, em que a interação estava restrita ao ensino e práticas emergenciais. Nesse sentido, a ideia do projeto *on-line*, aberto e gratuito, consiste na realização de palestras, mesas-redondas ou entrevistas relacionadas às temáticas de Literacia, Tecnologia e Fazeres e Saberes em Arquivologia e Ciência da Informação a partir da propositura de pesquisadores parceiros da rede. A proposta inclui pesquisadores da Ciência da Informação, estudantes e associações, oportunizando intercambiar conhecimentos e experiências por meio do canal no Youtube.

A palestra “Novas direções para a Ciência da Informação: inteligência artificial e processamento de textos” pronunciada por Mandl (2021) pelo projeto SESA *On-line* abordou os avanços da inteligência artificial como influência significativa na área de processamento e análise de textos, avanços esses que irão acarretar importantes mudanças nas práticas profissionais em bibliotecas e arquivos. O palestrante

CC BY NC ND

apresenta as técnicas inovadoras e discute as consequências nas práticas documentais.

No momento inicial da palestra, Mandl apresentou brevemente o local onde reside e trabalha na Alemanha por meio de fotos e mapas, assim, os alunos e todos os internautas puderam conhecer melhor o contexto ao qual o doutor e professor em Ciência da Informação é inserido.

Para uma melhor compreensão, na primeira parte, o palestrante elucidou de forma introdutória e resumida sobre as ciências envolvidas na discussão, que são Arquivologia, Ciência da Informação e Inteligência Artificial. A respeito de Arquivologia, Mandl (2021) registra que atualmente o interesse da área por Inteligência Artificial vem crescendo consideravelmente e é necessário acompanhar esses progressos.

Comenta ainda sobre alguns livros que apresentam os primórdios da Aquivologia: *Von Registratur*, de Jacob von Rammingen (1571), autor alemão que possivelmente foi o primeiro a falar sobre arquivística e a influenciar, a partir de seus escritos, os pensamentos da importância de considerar o arquivo como uma ferramenta auxiliar para historiadores; *De Archivis*, de Baldassare Castiglione (1632); *Modern Archives: Principles and Techniques,* de Schellenberg (2003). É interessante ressaltar a atualização na maneira de pensar o arquivo, avaliação e relação entre itens, e que hoje a Inteligência Artificial tem uma participação fundamental na ciência que trabalha com arquivos.

Nesse sentido, o acesso aos documentos exige do arquivista, em qualquer tempo, organização e gestão documental dos objetos para recuperá-los. Para exemplificar melhor as maneiras como essa gestão acontece, Mandl exibe imagens que ilustram como em épocas diferentes foram e são contemplados acesso à informação nas bibliotecas físicas e virtuais.

Segundo o palestrante, a área da Ciência da Informação se preocupa com o gerenciamento de informações criadas a partir de dados simples. Em sua maioria, essas informações, após serem constituídas por meio de análises ou transformações, permitem a produção de conhecimento. A indexação de informações pode ser automática ou semi-automática utilizando uma interface ou software, ou seja, por meio da computação ou atividade humana, assim, utiliza-se o conhecimento como

CC BY NC ND

interação no processo de armazenamento, recuperação ou disseminação da informação.

Sobre Inteligência Artificial - área da tecnologia que utiliza sistemas ou máquinas que imitam a inteligência e comportamento humano com finalidades diversas - Mandl (2021) destaca que existem dois conceitos, uma IA fraca, que acontece quando há uma reprodução ou imitação de comando, sem capacidade de raciocínio próprio; e uma IA forte, que reproduz processos cognitivos ao analisar dados padronizados, como por exemplo os comandos de uma máquina em um jogo de xadrez. O palestrante também destaca que a IA por substituir determinadas ações ameaça funções humanas, como um funcionário de fábricas, por outro lado a IA possibilita geração e aperfeiçoamento em áreas como: arte, medicina ou até seguros de carro.

Ao entrar em um subcampo da IA, Mandl apresenta a Aprendizagem de Máquina, um aprendizado automático em que o sistema tem habilidades de aprender a partir de dados amostrais, regras e padrões. Nesse sentido, são fornecidas pelo "ator" exemplos de entradas e saídas desejadas com o objetivo de mapear essas ações. Essa tarefa é esclarecida pelo palestrante e chamada de aprendizado supervisionado.

Como uma breve análise de Mandl, a Inteligência Artificial atingirá diretamente a área da Ciência da Informação e Arquivologia ligada à representação de textos. E como forma de elucidar o internauta a respeito da representação de texto, foi importante ressaltar métodos utilizados para que aconteça o *deep learning* (aprendizado profundo). Um dos principais e mais comuns métodos é por frequência de dados, que acontece a partir da significação representativa de palavras, em que é feita uma análise da frequência do uso de palavras e a quantidade de vezes que foi utilizada em determinado texto. Com base nesses dados é possível perceber a relevância de dado assunto, assim fica mais fácil captar um tipo de informação, porém, com esse método há ainda uma falha no que se refere a dificuldade na identificação semântica e sintática das palavras nos textos.

Em virtude dos avanços na área de representação de textos, surge então o método *Embedding* que supera as dificuldades de pouca informação semântica das palavras no texto, pois nesse método o sistema analisa o texto e encontra conexão e relação entre as palavras. Segundo o palestrante, esse sistema também tenta

CC BY NC ND

adivinhar as palavras que podem ser usadas entre as frases através de um percurso que inicia com o *input* (palavras de entrada), passam por uma *black box* ou caixa preta (programação pelo sistema de *Embedding*) e finaliza na *output* (palavra sugerida), todo o processo é gerado por representações numéricas e programados com um objetivo particular. Portanto, quanto mais informações forem chegando para o sistema, mais específico será o resultado do método.

A partir desse mesmo método, o professor Mandl apresenta outra função que pode ser utilizada nas áreas que trabalham com representação de textos. Nesse sentido, o *Embedding* também poderá indicar próximas palavras e representar frases completas, colecionando informações sobre um contexto específico. Dessa forma, o palestrante conclui que esse método após tanto ser usado e aperfeiçoado consegue hoje gerar textos com níveis tão bons quanto os elaborados por humanos.

Nessa perspectiva, Mandl avalia que a área de Arquivologia precisa estar a par de toda atualização que envolve a Inteligência Artificial, pois os processos manuais vêm sendo substituídos por ferramentas que utilizam a automatização, não só como pesquisa de arquivos, como também na manipulação de geração de textos e arquivos. Portanto, essa palestra, ofertada pelo projeto SESA, é de grande importância para o conhecimento sobre os avanços da IA na área de processamento e análise de textos. O professor Thomas Mandl apresentou métodos relevantes como *Embedding*, aprendizagem de máquina e modelos gerativos de maneira direta e acessível e evidencia a todo momento a importância da continuação dos estudos por parte dos atuantes e estudantes da área.

REFERÊNCIAS

PROJETO SESA. Disponível em: https://sites.google.com/view/projetosesa/home?authuser=1. Acesso em: 23 de dezembro de 2021.

CC BY NC ND