



## GPT-3, PROCESSAMENTO DE LINGUAGEM NATURAL E APLICAÇÃO DO DIREITO

### Resumo

Sérgio Rodrigo de Pádua

Desde o lançamento do *GPT-3* houve impacto na comunidade científica. O *GPT-3* (*Generative Pre-training Transformer 3*) é uma IA de processamento de linguagem natural (*NLP*). A *OpenAI* fornece os dados em *paper* publicado e em sua página no GitHub (<https://github.com/openai/gpt-3>). O *GPT-3* é uma IA *task agnostic* via rede neural profunda (*DNN*) com 175 bilhões de parâmetros. O treinamento do sistema *deep learning* usou o seguinte *dataset* em número de *tokens*: *Common Crawl* (410 bilhões – peso 60%), *WebText2* (19 bilhões), *Books1* (12 bilhões), *Books2* (55 bilhões) e *Wikipedia* (3 bilhões). A maior base de dados é o *Common Crawl*, que é um *crawler* de leitura de conteúdos existentes em toda a internet. Em relação à língua dos textos utilizados no treinamento tem a seguinte distribuição por número de palavras: a) inglês – 92,64708% – 181.014.683.608; b) francês – 1,81853% – 3.553.061.536; c) alemão – 1,46937% – 2.870.869.396; d) espanhol – 0,77289% – 1.510.070.974; e) italiano – 0,60793% – 1.187.784.217; f) português – 0,52483% – 1.025.413.869; e outros 112 idiomas em menores percentuais. O *GPT-3* está em testes (via *API* do sistema). No Direito a *API* do *GPT-3* pode ser aplicada em conjunto com outros sistemas de IA, como por exemplo, a *CARA A.I.* da empresa *Casetext* (<https://casetext.com>), cujo sistema é utilizado para fins de pesquisa de jurisprudência, leis e doutrina, gerando texto com o emprego de conceitos jurídicos. Assim, a aplicabilidade do *GPT-3* ao Direito tem se mostrado possível, sendo um auxiliar aos profissionais do Direito. Não há ainda uma visão de substituição dos operadores do Direito pela IA referida, pois esta tem menor acurácia que os humanos na atribuição de significado (76% do *GPT-3 Small* e 52% do *GPT-3* – o modelo com 175 bilhões de parâmetros tem baixa acurácia em interpretar o real significado das frases geradas). O *GPT-3* será base para outros sistemas de IA, mas não é IA do tipo geral (com compreensão de mundo). A IA aplica conceitos jurídicos, mas não foi treinada especificamente em Direito, o que pode ser auxiliado por IA jurídica. O *GPT-3* pode ser aplicado para geração de documentos em inglês (contratos e petições, por exemplo), já sua aplicação em português (0,52% da base de dados) pode ser dificultosa (ainda que o sistema consiga realizar tradução). Logo, o *GPT-3* indica o caminho de desenvolvimento para sistema *NLP task agnostic* com grande potencial de aplicação no Direito, mas que deve ser melhor desenvolvido. O resultado da pesquisa é parcial no sentido de possibilidade de utilização do *GPT-3* como sistema de processamento de linguagem natural auxiliar aos advogados especificamente. Conclui-se que são necessários maiores estudos sobre o sistema e suas potencialidades de aplicação ao direito brasileiro.

**Palavras-chave:** *GPT-3*; Processamento de Linguagem Natural; Inteligência Artificial Jurídica; Direito.