

INTELIGÊNCIA ARTIFICIAL CONSTITUCIONAL: TEORIA E PRÁTICA

Victoria Midori Deguchi Teramoto, Aluna do 9º período do curso de Direito da FAE Centro Universitário. Bolsista do Programa de Apoio à Iniciação Científica (PAIC 2024/2025). Ana Luiza Camargo dos Santos, Aluna do 7º período do curso de Direito da FAE Centro Universitário. Voluntária do Programa de Apoio à Iniciação Científica (PAIC 2024/2025). Aryadne Izabelle dos Santos Santana, Aluna do 9º período do curso de Direito da FAE Centro Universitário. Voluntária do Programa de Apoio à Iniciação Científica (PAIC 2024/2025). Bruna Cheiko Wuicik, Aluna do 7º período do curso de Direito da FAE Centro Universitário. Voluntária do Programa de Apoio à Iniciação Científica (PAIC 2024/2025). Giovanni Cantarelli Arten, Aluno do 7º período do curso de Direito da FAE Centro Universitário. Voluntário do Programa de Apoio à Iniciação Científica (PAIC 2024/2025). João Victor Archegas, Orientador da Pesquisa. Mestre em Direito pela Universidade Federal do Paraná. Professor de Direito na FAE.

Contatos: victoria.teramoto@mail.fae.edu.

santos.luiza@mail.fae.edu

aryadne.santana@mail.fae.edu

bruna.wuicik@mail.fae.edu

giovanni.arten@mail.fae.edu

joao.archegas@fae.edu

RESUMO

O presente artigo investiga a possibilidade de desenvolvimento de uma Inteligência Artificial Constitucional, isto é, um modelo de linguagem treinado a partir de princípios constitucionais explícitos e verificáveis, capazes de orientar sua atuação em contextos normativos. Parte-se da análise crítica do conceito de constitucionalismo digital, destacando a ascensão de plataformas tecnológicas como novos centros de poder transnacional e os riscos que esse fenômeno representa à proteção de direitos fundamentais. Em seguida, examina-se o funcionamento técnico dos *Large Language Models* (LLMs), suas limitações semânticas, ausência de agência moral e riscos de vieses, opacidade e inconsistência factual. A partir dessa fundamentação, explora-se o modelo de IA Constitucional proposto pela Anthropic com o Claude, discutindo os desafios éticos e políticos relacionados à definição da “constituição” algorítmica. A pesquisa adota abordagem empírica e qualitativa, realizando experimentos com diferentes modelos (ChatGPT, Claude, DeepSeek, CoPilot), os quais foram instruídos a operar sob referenciais normativos diversos, como a Constituição Federal de 1988, a Teoria do Sopesamento de Robert Alexy, a Teoria da Justiça de Rawls, a Bíblia Sagrada e o Manifesto Comunista. Os resultados indicam que os modelos conseguem simular argumentação jurídica coerente com os referenciais fornecidos, embora apresentem diferentes graus de rigidez ética e suscetibilidade a instruções subversivas. O estudo contribui para o debate sobre regulação da IA com base em direitos fundamentais, destacando os limites e possibilidades da IA Constitucional.

Palavras-chave: Teoria Constitucional. Constitucionalismo Digital. *Large Language Models*. Inteligência Artificial Constitucional. Governança Digital. *Accountability*.