



EXPLICABILIDADE E INTELIGÊNCIA ARTIFICIAL JUDICIAL

Sérgio Rodrigo de Pádua

Resumo

A hipótese estudada no presente resumo se refere a definição do direito à explicabilidade das decisões judiciais auxiliadas por IA. A presente pesquisa é de cunho qualitativo e mediante método exploratório-descritivo, mediante a análise interdisciplinar entre Direito e Ciência da Computação, de maneira a delinear os aspectos relevantes do direito à explicabilidade das decisões tomadas com auxílio de sistemas de inteligência artificial jurídica. Especificamente em relação ao Poder Judiciário, com a recente inclusão dos sistemas de IA judicial como novos atores no auxílio à tomada de decisão, a necessidade de explicabilidade toma corpo. A explicabilidade é trazida como um elemento da transparência das decisões tomadas por IA através de diversas iniciativas, como na Recomendação 449 da OCDE, na *First Version of a Draft Text of a Recommendation on the Ethics of Artificial Intelligence* da UNESCO, nos *Ethical Guidelines for Trustworthy AI* da União Europeia, no art. 20 da Lei Geral de Proteção de Dados (Lei 13.709/2018) e na Resolução 332/2020 do Conselho Nacional de Justiça. O art. 19 da Resolução 332/2020 do Conselho Nacional de Justiça prevê que os sistemas de inteligência artificial utilizados em auxílio à elaboração de decisões judiciais “*observarão, como critério preponderante para definir a técnica utilizada, a explicação dos passos que conduziram ao resultado*”. A transparência no uso da IA judicial, segundo o art. 8º, VI, da Resolução 332/2020 do CNJ, tem como um de seus elementos “*o fornecimento de explicação satisfatória e passível de auditoria por autoridade humana quanto a qualquer proposta de decisão apresentada pelo modelo de Inteligência Artificial*”. Sendo assim, a explicabilidade dos resultados do processamento advindo de algoritmos de IA passa pela ideia de *accountability*, sendo a explicabilidade um aprimoramento desta abordagem. Dessa maneira, a explicabilidade no desenvolvimento de modelos de IA judicial inclui a transparência no processo construção do modelo algorítmico, o que consiste na definição do objetivo (primeira fase), na escrita do código algorítmico (*coding* – segunda fase) e na implementação do modelo (terceira fase). Nessa linha, a explicabilidade do processo de construção de modelos de IA tem pressupostos característicos que são distintos da explicabilidade das saídas (*outputs*) e resultados (*outcomes*) propiciados por um modelo de IA. Portanto, o resultado parcial da pesquisa demonstra: a necessária delimitação do alcance do direito em relação às técnicas de inteligência artificial explicável (*xAI*), a fim de se garantir à legitimidade jurídica dos sistemas de IA judicial; e o necessário reconhecimento do direito à explicabilidade da decisão judicial tomada com auxílio de IA.

Palavras-chave: Inteligência artificial judicial; *xAI*; Direito à Explicabilidade.