

AGÊNCIA ARTIFICIAL E RESPONSABILIDADE MORAL. Lucas Benevides Ghiotto, Marcos Antonio Alves. UNESP, lucas.ghiotto@unesp.br, marcos.a.alves@unesp.br.

Resumo: O recente desenvolvimento de sistemas de Inteligência Artificial dotados de alto grau de autonomia impulsionou as discussões em torno do conceito de agência. Segundo tradição clássica na filosofia, a agência é definida, dentre outras coisas, pela capacidade de agir intencionalmente. Essa concepção é compartilhada por filósofos da corrente intencionalista, tal como Fred Dretske. Para ele, a agência exige que a cognição, através de uma relação funcional entre os estados internos de um sistema cognitivo, não somente cause, mas também explique sua ação. Sistemas de baixa complexidade informacional, como termostatos ou certos seres vivos muito simples, exibem intencionalidade de ordem mínima, insuficiente para a existência de agência genuína. O progresso na inteligência artificial (IA) propiciou o surgimento de máquinas autônomas capazes de processar uma enorme quantidade de dados e operar em contextos complexos, sem a supervisão humana constante. Em virtude da sua capacidade de processamento informacional com eficácia causal, não haveria, em princípio, impedimentos de atribuição de agência à IA. Entendemos que a definição dretskeana da agência e de seus possíveis desdobramentos éticos possui algumas complicações. Ao atribuímos agência genuína a sistemas de IA, enfrentamos um paradoxo: a equivalência entre a agência computacional e a humana ameaçaria deslocar a responsabilidade moral dos programadores ou usuários para os próprios algoritmos, gerando um vácuo de *accountability*. A fim de sanar este problema, Luciano Floridi propõe uma distinção entre diversos tipos de agência, distinguindo a agência natural, presente em organismos biológicos, da Agência Artificial (AA). Segundo sua abordagem, embora as AAs exibam interação, autonomia e adaptabilidade, esses sistemas operam sob uma orientação a objetivos ditados por objetivos humanos, desprovidos da complexidade mental e biológica que caracteriza a agência natural genuína. Argumentamos, neste trabalho, que a IA deve ser compreendida como um instrumento cujos danos e efeitos no mundo, embora vastos, não a tornam um agente moral responsável de modo que o poder de agir desses sistemas não lhes confere responsabilidade ética. Defendemos que a obrigação moral deve recair, necessariamente, sobre os criadores e usuários destes artefatos ou de seus responsáveis. Ao reconhecermos a IA como um agente técnico, preservamos a integridade dos mecanismos de justiça, impedindo que o avanço tecnológico sirva de pretexto para a irresponsabilidade e impunidade humanas.

Palavras-chave: Agência artificial, intencionalidade, responsabilidade moral, Floridi.