Responsabilidade criminal pelo produto "inteligente": reflexões e desafios

Susana Aires de Sousa— *Investigadora Integrada do Instituto Jurídico*

I. Contextualização: *setting the stage*

Imagine-se um mundo em que os veículos prescindem de condução humana (são, por isso, autónomos), utilizam energia limpa, movendo-se em comunicação com as estradas e outros veículos, interagindo com todo o sistema rodoviário e de transporte. Um mundo onde "as coisas" comunicam entre si, detetando pedestres, prevendo percursos seguros, mais eficientes, e prevenindo (evitando) quaisquer acidentes. Um mundo em que o erro humano é eliminado (estima-se que 94% dos acidentes de trânsito graves têm como causa decisões humanas erradas). Neste cenário, a lesão de bem jurídicos em contexto rodoviário seria reduzida à condição de mero acontecimento fortuito.

Este mundo ainda está por vir, fazendo parte de um futuro provável. O presente, porém, oferece um outro cenário: aquele em que o número de acidentes ligados a carros "inteligentes", autónomos ou automatizados, que circulam sistemas rodoviários desadequados, é crescente, desafiando modelos de responsabilidade e categorias jurídicas clássicas. Um exemplo recente liga-se à manipulação visual de sinais de trânsito, insignificante ao olho humano, mas suficiente para que o algoritmo interprete errada e perigosamente o limite de velocidade permitido. Um outro exemplo, que merece ser referido pela sua visibilidade, mas também por ilustrar os desafios lançados à responsabilidade jurídica, diz respeito ao primeiro atropelamento mortal causado por um carro autónomo da Uber, em fases de testes, ocorrido em março de 2018, em Tempe, no Arizona. A vítima, Elaine Herzberg, foi mortalmente atropelada, quando atravessava uma estrada empurrando uma bicicleta, por um carro, de marca Volvo, modificado pela Uber e autorizado a circular na via pública. Vários fatores contribuíram para este desenlace fatal, desde a dificuldade sentida pelo algoritmo em identificar aquele obstáculo como uma pessoa, reagindo tardiamente, até ao alheamento da *designated driver* – a pessoa humana que no interior do veículo devia monitorizar o seu desempenho –, a uma "cultura empresarial de segurança inadequada", ou ainda à conduta da vítima que atravessava a rodovia, à noite, num local sem sinalização. Em março de 2019, o Ministério Público deduziu acusação por homicídio negligente contra a pessoa humana. O julgamento terá, em breve, lugar.

II. Desafios: o *responsability gap*

As decisões tomadas por algoritmos ocorrem em muitos outros domínios económicos e sociais. A utilização destes sistemas integra o quotidiano por diversas formas, mais ou menos visíveis: meios de informação, comunicação ou aconselhamento (técnico ou económico); *internet*, computadores ou *smartphones*; utilização de sistemas de diagnóstico ou de robôs cirúrgicos; *trading* algorítmico; transportes através de veículos autónomos (carros, *shuttles*, barcos, *drones*), etc. Contudo, decisões tomadas por sistemas computacionais complexos dinâmicos, imprevisíveis à pessoa humana, desafiam modelos e categorias clássicas em que assenta a atribuição de responsabilidade. É justamente nesta autonomia de aprendizagem (e de decisão) que reside o *responsability gap* ou *AI criminal gap*.

Em causa estão as categorias que suportam o juízo de imputação do evento desvalioso a uma conduta, como a causalidade e a culpa. Alguns algoritmos funcionam como autênticas caixas negras na forma como processam os dados (*input*) e alcançam um determinado resultado (*output*). Isto é, o tratamento algorítmico dos dados, segundo uma estrutura complexa, torna opaco o processo que conduz a determinado resultado, não obstante a sua capacidade de grande precisão na determinação de nova informação. A opacidade será tanto maior quanto mais complexos (e precisos) sejam os modelos de *machine learning* utilizados, sendo que, em alguns casos, o estado atual de desenvolvimento tecnológico não permite determinar, atendendo ao grau de complexidade do sistema, como se chegou àquele resultado, seja ele um juízo de previsibilidade, um aconselhamento, ou uma decisão.

Por sua vez, a imprevisibilidade e a natureza dinâmica dos sistemas computacionais complexos fundamentam dúvidas sobre a imputação subjetiva do dano exigida pelos tipos legais de crime. Da perspetiva da pessoa humana ligada ao fabrico, à programação ou à utilização do sistema, a intervenção da máquina

torna imprevisível o evento desvalioso. A opacidade do sistema e a imprevisibilidade do resultado danoso dificultam quer a possibilidade de representação humana daquele resultado, quer uma prova, suficientemente sustentada, da sua existência.

III. Revisitar possíveis soluções

Em alguns casos, as dificuldades podem ser superadas por regras clássicas, já instituídas em contexto de responsabilidade (civil e criminal) pelo produto. Contudo, casos haverá reveladores de especiais particularidades associadas à complexidade computacional do produto dito "inteligente". Este produto distingue-se, no risco que lhe é inerente, pela sua imprevisibilidade e incontrolabilidade. O *risco inerente* é um conceito importante em matéria de responsabilidade pelo produto porque constitui um parâmetro para a intervenção do direito enquanto instrumento de controlo de riscos. A grande autonomia de alguns sistemas inteligentes, associada ao contexto em que são aplicados (por exemplo, tráfego rodoviário) ou às circunstâncias em que são utilizados (*v. g.*, domínio militar), impõe um dever acrescido de cuidado que deve concretizar-se em medidas jurídicas que diminuam esse risco, procurando-se, dessa forma, aumentar a *confiança* numa utilização segura, capaz de modificar o grau de risco para um nível aceitável ou permitido.

Deste modo, em casos de incerteza sobre a amplitude dos riscos associados ao produto "inteligente", o princípio da precaução constitui fundamento para a imposição de medidas e deveres especiais, como um dever de vigilância e monitorização do produto ou a implementação de regulação dinâmica (*v. g.*, a *sandbox aproach)*. Estas medidas contribuem para um conhecimento gradual do produto em contexto real, diminuindo assim a sua opacidade (ou *black box*).

Estes deveres, assentes numa ideia de plausibilidade do risco, são imputáveis a pessoas jurídicas (humanas e empresas), cujo cumprimento pode ser reforçado por normas sancionatórias, essencialmente de natureza não penal. Todavia, esta regulação deve estar sujeita a um *princípio de revisibilidade* que acompanha o grau de conhecimento do produto. Ou seja, na medida em que o contexto de plausibilidade evolua para um estado de previsibilidade, deve rever-se a necessidade de intervenção penal. Assim, se o risco de produção do dano se torna previsível, deve a autoridade pública averiguar da necessidade de criminalização da conduta, designadamente através da construção de incriminações especificamente voltadas para a chamada *IA forte*.