

A guerra no Médio Oriente e o papel da inteligência artificial na desinformação

 *Cristiana Pereira*

cristianapereira136@gmail.com

<https://orcid.org/0009-0003-5206-6730>

ISCAP, Instituto Politécnico do Porto

Resumo

A intensificação dos conflitos contemporâneos tem sido acompanhada por transformações significativas no ecossistema informacional, particularmente com a crescente utilização de Inteligência Artificial (IA). No contexto da guerra no Médio Oriente, a desinformação emerge como um fenómeno central, influenciado pela capacidade tecnológica de produzir e disseminar conteúdos manipulados. Este artigo desenvolve uma reflexão crítica sobre este fenómeno, analisando três dimensões fundamentais: a produção automatizada de conteúdos falsos, a amplificação algorítmica nas plataformas digitais e os efeitos cognitivos associados à exposição a desinformação. A análise demonstra que a IA não apenas aumenta o volume de informação falsa, como também altera a forma como os conflitos são percecionados. Na discussão, são exploradas implicações e possíveis respostas, evidenciando a necessidade de abordagens integradas que articulem regulação, tecnologia e literacia mediática.

Palavras-chave: Inteligência Artificial, Desinformação, Médio Oriente, Deepfakes, Algoritmos, Media Digitais

Abstract

The intensification of contemporary conflicts has been accompanied by significant changes in the information ecosystem, particularly with the growing use of artificial intelligence. In the context of the war in the Middle East, disinformation has emerged as a central phenomenon, driven by the technological capacity to produce and disseminate manipulated content. This article offers a critical reflection on this phenomenon, analysing three fundamental dimensions: the automated production of false content, algorithmic amplification on digital platforms, and the cognitive effects associated with exposure to disinformation. The analysis demonstrates that AI not only increases the volume of misleading information but also alters the way in which conflicts are perceived. The discussion explores implications and possible responses, highlighting the need for integrated approaches that combine regulation, technology and media literacy.

Keywords: Artificial Intelligence, Disinformation, Middle East, Deepfakes, Algorithms, Digital Media

Introdução

Os conflitos armados contemporâneos decorrem num contexto profundamente marcado pela digitalização da informação e pela ambiguidade associada ao funcionamento das plataformas digitais, que desempenham atualmente um papel central na forma como os acontecimentos são comunicados e interpretados. No caso da guerra no Médio Oriente, utilizando para fins analíticos, a guerra em Gaza desencadeada desde outubro de 2023, a circulação de conteúdos ocorre num ambiente caracterizado não apenas pela rapidez com que a informação se propaga, mas também pela grande diversidade de fontes disponíveis e pela crescente dificuldade de verificação da veracidade dos conteúdos. Esta combinação de fatores contribui para um cenário em que a desinformação assume um papel cada vez mais significativo e preocupante, tornando-se um elemento estruturante do próprio conflito informacional.

Esta preocupação torna-se particularmente visível no contexto da guerra no Médio Oriente, onde conteúdos visuais gerados, ou manipulados por Inteligência Artificial (IA), passaram a circular amplamente nas redes sociais. O *Digital News Report 2024* do *Reuters Institute* assinala que imagens artificiais relacionadas com a guerra foram partilhadas online, contribuindo para aumentar a incerteza sobre a autenticidade dos conteúdos visuais e para dificultar a distinção entre informação factual, propaganda e manipulação digital. Este cenário reforça a ideia de que os conflitos armados contemporâneos são também disputas informacionais, na quais a credibilidade das imagens e das fontes se torna um elemento central da perceção pública do conflito (Newman, 2024).

A desinformação, também entendida como a propagação intencional de informação falsa ou falaciosa, assume neste contexto um papel estratégico, influenciando de forma direta e indireta as perceções do público relativamente aos acontecimentos. Embora este fenómeno não seja propriamente novo, a sua escala, intensidade e complexidade aumentaram de forma significativa com o desenvolvimento das tecnologias digitais e com a generalização do acesso à internet e às redes sociais.

Os autores Wardle e Derakhshan (2017), subdividem a desordem informacional em três categorias interdependentes: a *misinformation* (informação falsa partilhada de forma inadvertida, sem intenção maliciosa de causar prejuízo); a *disinformation* (conteúdo deliberadamente fabricado ou manipulado com o propósito explícito de enganar, orientar a opinião pública ou causar danos geopolíticos); e a *malinformation* (informação baseada em factos reais, mas recontextualizada e divulgada de forma maliciosa para infligir danos a indivíduos, grupos ou nações). No âmbito deste trabalho, a atenção centra-se nas manifestações de *disinformation* e na sua simbiose com os sistemas automatizados. Atualmente, os utilizadores não são apenas consumidores passivos de informação, mas também produtores e disseminadores ativos, o que contribui para a multiplicação de conteúdos e para a dificuldade em distinguir entre fontes credíveis e fontes duvidosas.

O surgimento da IA, em particular dos sistemas generativos, veio acrescentar uma nova dimensão a este fenómeno. Estas tecnologias permitem a criação de textos, imagens e vídeos com um grau de realismo extremamente elevado, possibilitando a produção de conteúdos que simulam acontecimentos ou distorcem factos de forma convincente. Como consequência, torna-se mais difícil distinguir entre informação autêntica e conteúdos manipulados. Para além disso, importa considerar os efeitos desta desinformação ao nível cognitivo, nomeadamente na forma como os indivíduos interpretam, processam e avaliam a informação que consomem no seu quotidiano.

Neste contexto, torna-se essencial compreender de que forma a IA contribui para a intensificação da desinformação e quais são as implicações desse fenómeno na perceção pública dos conflitos armados.

Este artigo sustenta que a IA está a intensificar a desordem informacional (Wardle & Derakhshan, 2017) acerca da guerra no Médio Oriente, e por conseguinte, a forma como esta é percebida, porque facilita a produção de conteúdos manipulados que parecem verdadeiros, amplifica a sua propagação através de sistemas algorítmicos e intensifica os efeitos cognitivos que dificultam a avaliação crítica da informação, e por isso, implica a necessidade de respostas integradas que combinem a regulação tecnológica bem como a responsabilização das plataformas, de forma a que as pessoas fiquem mais informadas e possam fazer uma distinção mais precisa do que é real e do que é produzido por IA.

Revisão de Literatura

Produção de conteúdos manipulados através da Inteligência Artificial

A capacidade da IA para produzir conteúdos manipulados com elevado grau de realismo constitui um dos fatores mais relevantes para compreender a intensificação da desinformação nos dias de hoje. O desenvolvimento constante de tecnologias como os modelos generativos e os *deepfakes* permite criar imagens, vídeos e textos que simulam acontecimentos de forma extremamente convincente, dificultando cada vez mais a distinção entre conteúdos autênticos e conteúdos artificiais. Esta evolução tecnológica tem vindo a transformar profundamente a forma como a informação é produzida e consumida.

A este propósito, Paris e Donovan (2019) destacam o impacto e a tipologia da manipulação de evidências visuais e sonoras, propondo uma distinção fundamental entre *deepfakes* e *cheapfakes*. Enquanto os *deepfakes* recorrem a sistemas de IA e aprendizagem profunda para gerar simulações sintéticas altamente sofisticadas, os *cheapfakes* baseiam-se em métodos de edição mais simples, acessíveis e convencionais, como a alteração da velocidade do áudio, cortes tendenciosos, rotulagem enganosa ou a mera descontextualização de registos reais. Os autores sublinham que, apesar de os *deepfakes* atraírem maior alarme social devido ao seu carácter tecnológico inovador, os *cheapfakes* continuam a ser massivamente mais comuns e igualmente perigosos. Ambas as abordagens são altamente eficazes em enganar o público, demonstrando que a desordem informativa não depende unicamente de tecnologias de ponta, mas sim da vulnerabilidade histórica da evidência audiovisual face à manipulação intencional. No ambiente hiperconectado atual, o público tende a aplicar a heurística de que "ver é acreditar", o que transforma tanto as manipulações rudimentares como as geradas por IA em ferramentas políticas poderosas.

De acordo com Khanjani et al. (2023), estas tecnologias utilizam algoritmos de aprendizagem profunda para gerar conteúdos sintéticos, como vídeos e áudios que imitam, com grande precisão, características humanas, incluindo expressões faciais, entoação de voz e padrões de comportamento. O estudo destes autores evidencia que, no contexto da desinformação, os riscos associados a estas tecnologias são cada vez mais elevados, destacando-se a manipulação da opinião pública, a difamação e a perda de confiança nas fontes de informação. À medida que a tecnologia evolui, os conteúdos gerados tornam-se progressivamente mais sofisticados e difíceis de detetar, o que aumenta a sua eficácia enquanto instrumentos de manipulação. Este fenómeno pode levar a uma situação particularmente preocupante, em que os indivíduos começam não só a acreditar em conteúdos falsos, mas também a duvidar da veracidade de conteúdos reais, assumindo que estes possam ser produtos de IA. Esta inversão de confiança constitui um dos efeitos mais complexos da desinformação contemporânea, uma vez que fragiliza a credibilidade da informação no seu todo e dificulta a distinção entre o que é verdadeiro e o que é manipulado.

A literatura recente confirma que a Inteligência Artificial Generativa (IAG) tem um papel ambivalente no ecossistema da desinformação. Por um lado, pode apoiar processos de verificação e deteção, por outro, reduz os custos técnicos, temporais e criativos da produção

de conteúdos enganosos. Park e Nan (2026), numa revisão de âmbito sobre IAG e *misinformation*, mostram que os modelos generativos podem contribuir para a criação de textos, imagens e outros conteúdos plausíveis, aumentando a escala e a sofisticação das mensagens falsas ou manipuladas. No contexto de conflitos armados, esta capacidade torna-se especialmente problemática, porque permite produzir rapidamente conteúdos emocionalmente apelativos, ajustados a narrativas políticas específicas e difíceis de verificar em tempo real.

No contexto da guerra no Médio Oriente, esta capacidade traduz-se na circulação de conteúdos que podem distorcer a interpretação dos acontecimentos, contribuindo para a construção de narrativas alternativas e, por vezes, contraditórias. Estas narrativas podem ser utilizadas para influenciar a opinião pública, tanto a nível nacional como internacional, moldando perceções e posicionamentos relativamente ao conflito. A facilidade de acesso às ferramentas de IA reduz significativamente as barreiras à produção de desinformação, permitindo que um número cada vez maior de utilizadores participe neste processo, mesmo sem conhecimentos técnicos avançados.

A aplicação prática destas dinâmicas na guerra em Gaza revela exemplos empíricos severos de desordem visual. Um dos fenómenos amplamente documentados envolve a proliferação de imagens hiper-realistas geradas por IAG, exibindo cenários dramáticos de falsas vítimas infantis ou acampamentos de refugiados fictícios, massivamente partilhadas para polarizar a opinião pública mundial e angariar suporte moral. Noutro extremo, observa-se o recurso estratégico a *cheapfakes* e descontextualizações gráficas brutas, como a disseminação de gravações digitais extraídas do videojogo de simulação militar *Arma 3*, apresentadas em plataformas como o X (antigo Twitter) e o TikTok como se fossem filmagens fidedignas de bombardeamentos e de interceções de defesas antiaéreas ocorridas em tempo real. Estes casos demonstram que a maleabilidade estrutural dos formatos digitais distorce a capacidade do espectador comum de validar autonomamente as evidências da guerra.

Para além da circulação de conteúdos efetivamente manipulados, importa considerar outro efeito relevante: a suspeição generalizada criada pela própria possibilidade de uso de IA. Uma análise do *Institute for Strategic Dialogue* sobre Gaza mostra que, em vários casos, imagens e vídeos autênticos de vítimas e destruição foram acusados de serem gerados por IA, contribuindo para desacreditar provas visuais e narrativas de sofrimento civil. Este fenómeno aproxima-se do chamado *liar's dividend*, conceito analisado pela TrueScreen (2026), num ambiente em que a manipulação tecnológica é plausível, até conteúdos verdadeiros podem ser rejeitados como falsos. Assim, a IA não intervém apenas através da produção de inverdades, mas também através da erosão da confiança na evidência visual autêntica.

Segundo Shoaib et al. (2023), a IA não apenas aumenta o volume de conteúdos manipulados, como também melhora a sua qualidade, tornando-os mais difíceis de distinguir de conteúdos autênticos. Esta evolução representa uma mudança significativa, na medida em que a desinformação se torna mais sofisticada, mais credível e, consequentemente, mais eficaz na sua capacidade de influenciar os indivíduos.

Outro aspeto importante prende-se com a velocidade de produção, a IA permite gerar grandes quantidades de conteúdo num curto espaço de tempo, contribuindo para a saturação do espaço informacional, este excesso de informação dificulta a verificação dos conteúdos e aumenta a probabilidade de disseminação de informação falsa. Neste contexto, os mecanismos tradicionais de verificação tornam-se menos eficazes, uma vez que não conseguem acompanhar o ritmo acelerado de produção e circulação de conteúdos.

Para além disso, importa considerar que muitos destes conteúdos são produzidos com objetivos específicos, nomeadamente influenciar narrativas políticas ou sociais, o que reforça a sua relevância em contextos de conflito. Assim, a produção de conteúdos manipulados

assume um papel central na intensificação da desinformação, contribuindo para a sua propagação em larga escala e para a crescente dificuldade em distinguir entre verdade e falsidade.

Amplificação algorítmica nas plataformas digitais

Para além da produção de conteúdos, a disseminação da desinformação está profundamente ligada ao funcionamento das plataformas digitais e aos algoritmos que estas utilizam para organizar e distribuir a informação. As plataformas recorrem a sistemas de recomendação baseados em IA, que selecionam, filtram e organizam conteúdos com base no comportamento dos utilizadores, nomeadamente no histórico de interações, preferências individuais e padrões de consumo de informação ao longo do tempo.

De acordo com Gonçalves et al. (2024), estes sistemas são concebidos com o objetivo de maximizar o envolvimento dos utilizadores, privilegiando conteúdos que apresentam maior probabilidade de gerar interação, como partilhas, comentários ou reações. Este tipo de conteúdos tende a ser emocionalmente apelativo, o que significa que muitas vezes inclui informação sensacionalista, polémica ou controversa, capaz de captar rapidamente a atenção dos utilizadores. Ao priorizarem o envolvimento em detrimento da veracidade da informação, os algoritmos contribuem de forma significativa para a amplificação da desinformação, estes conteúdos manipulados ou enganadores podem alcançar rapidamente um grande número de pessoas, especialmente quando conseguem gerar reações emocionais intensas, como a indignação, o medo ou a empatia. Esta lógica de funcionamento favorece a circulação de conteúdos que captam a atenção, mesmo que estes não sejam rigorosos ou verdadeiros.

A amplificação da desinformação não depende apenas da existência de conteúdos falsos, mas também da forma como as plataformas digitais organizam a visibilidade desses conteúdos. Milli et al. (2025), através de uma auditoria algorítmica ao Twitter/X, mostram que sistemas de ordenação baseados em *engagement* podem amplificar conteúdos divisivos e emocionalmente carregados, quando comparados com uma ordenação cronológica. Esta evidência é particularmente relevante para contextos de guerra, nos quais imagens de sofrimento, acusações morais e narrativas de medo ou indignação tendem a gerar fortes reações. Assim, os algoritmos não criam necessariamente a desinformação, mas podem aumentar o seu alcance quando esta é capaz de mobilizar emoção, conflito e interação.

No contexto da guerra no Médio Oriente, esta dinâmica torna-se particularmente evidente. Conteúdos que apelam à emoção, como imagens de violência, sofrimento ou injustiça, tendem a ser amplamente partilhados nas redes sociais, muitas vezes sem qualquer verificação prévia. Quando estes conteúdos são manipulados ou descontextualizados, podem contribuir para a construção de perceções distorcidas do conflito, influenciando significativamente a forma como este é interpretado pelo público. A personalização algorítmica constitui outro fator relevante, ao apresentar conteúdos alinhados com as crenças, interesses e opiniões dos utilizadores, os algoritmos contribuem para a formação de bolhas informativas, dentro destas bolhas, os indivíduos são expostos predominantemente a informação que confirma as suas próprias perspetivas, reduzindo o contacto com pontos de vista alternativos e dificultando o desenvolvimento de uma visão crítica e equilibrada.

A discussão sobre “bolhas informativas” deve também ser tratada com maior precisão conceptual. Nguyen (2020) distingue *epistemic bubbles* e *echo chambers*: nas primeiras, determinadas vozes ou perspetivas ficam ausentes do ambiente informacional; nas segundas, as fontes externas são ativamente desacreditadas. Esta distinção é importante para compreender a desinformação em plataformas digitais, porque a personalização algorítmica pode limitar a exposição a perspetivas alternativas, mas a radicalização da desconfiança depende também de dinâmicas sociais e identitárias que tornam certas fontes

sistematicamente rejeitadas. No contexto de conflitos armados, isto pode contribuir para leituras polarizadas, nas quais os utilizadores recebem sobretudo conteúdos alinhados com a sua visão prévia do conflito e desconfiam de informação proveniente de outros grupos ou fontes.

Esta distinção é atualizada por Turner (2025), que analisa bolhas epistémicas e câmaras de eco em ambientes online, mostrando que a limitação da abertura a perspetivas alternativas não resulta apenas da tecnologia, mas também das práticas de seleção, confiança e rejeição de fontes pelos próprios utilizadores. Assim, a personalização algorítmica deve ser entendida como parte de um processo mais amplo, os sistemas digitais podem estreitar o campo de informação disponível, mas as comunidades online também reforçam normas internas de credibilidade e descredibilização. Esta abordagem permite evitar uma explicação demasiado simplista, segundo a qual os algoritmos seriam os únicos responsáveis pela formação de perceções enviesadas sobre o conflito.

Estas dinâmicas estruturais encontram uma forte fundamentação teórica na tese de Zuboff (2019) sobre o conceito de *capitalismo de vigilância*. A autora argumenta que o modelo de negócio das plataformas digitais contemporâneas assenta na extração, modificação e mercantilização da experiência humana através da recolha massiva de dados comportamentais. Sob esta lógica económica, a captação e a retenção da atenção dos utilizadores tornam-se os objetivos primordiais dos sistemas algorítmicos, uma vez que o tempo de permanência nas redes dita diretamente o potencial de geração de receita publicitária. Consequentemente, os algoritmos são intencionalmente desenhados para maximizar o envolvimento, criando uma arquitetura digital onde conteúdos sensacionalistas, divisivos ou falsificados são artificialmente privilegiados devido à sua capacidade intrínseca de desencadear reações emocionais imediatas e duradouras. Assim, na perspetiva de Zuboff (2019), a proliferação da desinformação em momentos de crise geopolítica não constitui um mero erro técnico do sistema, mas sim o resultado direto e funcional de um modelo de negócio que lucra com a monetização do comportamento e com a desordem informativa por ele gerada.

Importa ainda referir que muitos utilizadores não têm plena consciência do funcionamento destes algoritmos, o que dificulta a sua capacidade de reconhecer os mecanismos que influenciam a informação que consomem. Esta falta de transparência contribui para uma aceitação mais passiva dos conteúdos apresentados, aumentando o impacto da desinformação.

Deste modo, os algoritmos das plataformas digitais desempenham um papel fundamental na disseminação da desinformação, contribuindo para a sua rápida circulação, para o aumento do seu alcance e para o reforço de perceções enviesadas, especialmente em contextos de conflito como o do Médio Oriente.

Efeitos cognitivos da desinformação e perceção do conflito

A exposição contínua a conteúdos manipulados tem implicações significativas ao nível cognitivo, influenciando de forma profunda a maneira como as pessoas percecionam os acontecimentos e avaliam a informação que consomem no seu dia a dia. Num contexto em que a quantidade de informação disponível é extremamente elevada, a dificuldade em distinguir entre conteúdos verdadeiros e falsos torna-se cada vez mais evidente, comprometendo a capacidade de análise crítica dos indivíduos e afetando os seus processos de tomada de decisão.

Segundo Danry et al. (2024), conteúdos gerados por IA influenciam diretamente a crença das pessoas, aumentando de forma significativa a probabilidade de aceitação de informação falsa, este fenómeno não depende apenas do conteúdo em si, mas também da forma como

este é apresentado, nomeadamente do seu grau de coerência, da sua estrutura e da forma como se enquadra nas expectativas dos utilizadores. De facto, conteúdos que apresentam uma aparência credível, organizada e coerente tendem a ser mais facilmente aceites, especialmente em contextos marcados por uma elevada carga emocional, como é o caso dos conflitos armados. Quando os indivíduos são confrontados com imagens ou narrativas que evocam emoções fortes, como medo, indignação ou empatia, a sua capacidade de avaliação crítica pode ser reduzida, tornando-os mais vulneráveis à influência da desinformação, para além disso, a existência de múltiplas versões sobre os mesmos acontecimentos contribui para o fenómeno da fadiga informacional, a constante exposição a informações contraditórias dificulta a capacidade dos indivíduos de analisar criticamente os conteúdos e de formar uma opinião fundamentada. Com o tempo, esta sobrecarga informativa pode levar a uma diminuição do interesse em verificar a veracidade das informações, favorecendo a aceitação de conteúdos sem validação, esta situação pode ainda conduzir ao desenvolvimento de uma atitude de desconfiança generalizada, em que os indivíduos deixam de confiar não apenas em conteúdos duvidosos, mas também em fontes tradicionalmente consideradas credíveis. Como consequência, verifica-se uma erosão da confiança no espaço público, o que dificulta a construção de consensos baseados em factos e enfraquece o debate informado.

A dimensão visual da desinformação é particularmente relevante em conflitos armados, porque imagens e vídeos tendem a funcionar como provas imediatas do acontecimento. Hameleers (2025) sublinha que a guerra informacional contemporânea tem uma forte natureza visual, na medida em que conteúdos visuais podem mobilizar emoções, construir interpretações morais do conflito e influenciar a confiança nas fontes. Esta dinâmica é agravada pela circulação de imagens manipuladas ou suspeitas de manipulação, pois a dúvida sobre a autenticidade dos conteúdos pode gerar uma desconfiança mais ampla em relação ao ecossistema informativo. Deste modo, a desinformação visual não afeta apenas aquilo em que as pessoas acreditam, mas também a sua disposição para confiar em qualquer evidência apresentada como factual.

Assim, os efeitos cognitivos da desinformação assumem um papel central na forma como os conflitos são percecionados, contribuindo para a distorção da realidade e reforçando o impacto da IA no contexto da desinformação contemporânea.

Discussão

A análise dos tópicos apresentados ao longo deste artigo permite compreender que a intensificação da desinformação no contexto da guerra no Médio Oriente não resulta apenas da evolução tecnológica, mas sobretudo da forma como essa tecnologia se articula com fatores sociais, económicos e comportamentais que moldam o ecossistema digital contemporâneo. A IA, neste sentido, não atua de forma isolada, mas integrada num conjunto mais amplo de dinâmicas que envolvem as plataformas digitais, os modelos de negócio associados e o comportamento dos utilizadores.

De facto, a IA acaba por assumir um papel duplo neste processo, por um lado, facilita significativamente a criação de conteúdos manipulados, reduzindo as barreiras técnicas e tornando acessível a produção de desinformação a um número cada vez maior de pessoas. Por outro lado, contribui para que esses mesmos conteúdos se disseminem de forma mais rápida e eficaz, através dos sistemas algorítmicos que estruturam o funcionamento das plataformas digitais e que privilegiam conteúdos com maior potencial de envolvimento, esta combinação de fatores faz com que a desinformação não só aumente em quantidade, mas também em intensidade e impacto social.

Perante esta realidade, torna-se evidente que não existe uma solução única ou imediata para o problema. A regulação das tecnologias de IA surge como uma necessidade importante, sobretudo no sentido de limitar a produção e circulação de conteúdos manipulados, no

entanto, a aplicação destas medidas enfrenta desafios significativos, nomeadamente devido à dimensão global do fenómeno, à diversidade de enquadramentos legais e à rapidez com que a tecnologia evolui, a criação de normas eficazes exige, assim, um esforço coordenado entre diferentes países e instituições. Nesse sentido, o relatório do *European Digital Media Observatory* (Bontcheva, 2024) sobre IAG e desinformação mapeia detalhadamente os riscos associados e sublinha a urgência de respostas regulatórias robustas e de um envolvimento ativo das plataformas digitais para conter essas ameaças sistémicas.

Para além disso, as plataformas digitais desempenham um papel central neste processo e, como tal, devem assumir uma maior responsabilidade, tornar os algoritmos mais transparentes, melhorar os critérios de recomendação e reforçar os mecanismos de moderação de conteúdos pode contribuir para reduzir a disseminação de desinformação. Ainda assim, estas medidas levantam questões importantes, sobretudo no que diz respeito ao equilíbrio entre a necessidade de limitar conteúdos prejudiciais e a preservação da liberdade de expressão, que constitui um princípio fundamental nas sociedades democráticas.

Por outro lado, importa reconhecer que a própria tecnologia pode também ser utilizada como parte da solução. O desenvolvimento de ferramentas de deteção de conteúdos falsos, baseadas em IA, representa uma estratégia relevante para combater a desinformação, no entanto, estas soluções enfrentam uma dificuldade constante, à medida que os métodos de manipulação evoluem e se tornam mais sofisticados, os sistemas de deteção têm de acompanhar esse ritmo, o que exige um processo contínuo de atualização e adaptação.

Neste contexto, a literacia mediática assume um papel particularmente importante. Quanto mais preparados estiverem os cidadãos para analisar criticamente a informação, menor será a probabilidade de serem influenciados por conteúdos manipulados ou enganosos, o desenvolvimento destas competências permite não só identificar desinformação, mas também compreender os mecanismos através dos quais esta se propaga, embora esta não seja uma solução imediata, constitui uma estratégia fundamental a médio e longo prazo.

Em síntese, a resposta à desinformação deve assentar numa abordagem integrada, que combine diferentes dimensões e reconheça a complexidade do fenómeno.

Conclusão

A guerra no Médio Oriente evidencia de forma particularmente clara o impacto da IA na transformação do ecossistema informacional em contextos de conflito, a facilidade com que atualmente é possível criar conteúdos manipulados, aliada à rapidez com que estes se disseminam através das plataformas digitais, contribui para a formação de um ambiente informacional cada vez mais complexo, dinâmico e difícil de interpretar por parte dos cidadãos.

Ao longo deste artigo, procurou-se demonstrar que a desinformação não deve ser entendida como um fenómeno isolado, mas sim como o resultado da interação entre diferentes fatores, nomeadamente tecnológicos, sociais e cognitivos. A IA desempenha um papel central neste processo, não apenas ao facilitar a produção de conteúdos manipulados, mas também ao potenciar a sua disseminação e ao influenciar a forma como estes são percecionados pelos indivíduos em diferentes contextos.

A reflexão crítica desenvolvida permite concluir que as respostas a este fenómeno devem ser necessariamente abrangentes, articuladas e adaptadas à complexidade do problema, a regulação das tecnologias de IA, a responsabilização das plataformas digitais e o reforço da literacia mediática surgem como elementos fundamentais para mitigar os efeitos da desinformação. No entanto, nenhuma destas medidas, considerada de forma isolada, será suficiente para resolver o problema de forma eficaz.

Importa ainda reconhecer que a eliminação total da desinformação não constitui um objetivo realista, tendo em conta a natureza dinâmica e em constante evolução deste fenómeno. Assim, o foco deve centrar-se na redução do seu impacto, na melhoria dos mecanismos de controlo e, sobretudo, no fortalecimento da capacidade dos indivíduos para lidar com a informação de forma crítica, consciente e informada.

Por fim, como perspetiva futura, torna-se essencial continuar a aprofundar o estudo deste fenómeno, não apenas no contexto do Médio Oriente, mas também em outros cenários de conflito e noutras realidades informacionais. A compreensão mais detalhada das dinâmicas associadas à desinformação permitirá desenvolver estratégias mais eficazes, sustentáveis e adaptadas às transformações constantes do ambiente digital contemporâneo.

Declaração de uso de IA

Durante a elaboração deste artigo, a autora utilizou a ferramenta *ChatGPT* para revisão de literatura e aperfeiçoamento gramatical. Todo o conteúdo gerado foi submetido a uma análise crítica, sendo revisto, validado e complementado pela autora garantindo o seu rigor conceptual e a sua conformidade com princípios éticos e científicos.

Referências

- Bontcheva, K. (Ed.). (2024). *Generative AI and disinformation: Recent advances, challenges, and opportunities*. vera.ai project. <https://lirias.kuleuven.be/retrieve/758830>
- Danry, V., Pataranutaporn, P., Groh, M., Epstein, Z., & Maes, P. (2024). *Deceptive AI systems that give explanations are more convincing than honest AI systems and can amplify belief in misinformation* [Preprint]. arXiv. <https://doi.org/10.48550/arXiv.2408.00024>
- Gonçalves, A., Torre, L., & Melo, P. V. (Eds.). (2024). *Inteligência artificial e algoritmos: desafios e oportunidades para os media*. LabCom, Universidade da Beira Interior. <https://labcom.ubi.pt/inteligencia-artificial-e-algoritmos-desafios-e-oportunidades-para-os-media/>
- Hameleers, M. (2025). The visual nature of information warfare: The construction of partisan claims on truth and evidence in the context of wars in Ukraine and Israel/Palestine. *Journal of Communication*, 75(2), 90-100. <https://doi.org/10.1093/joc/jqae045>
- Khanjani, Z., Watson, G., & Janeja, V. P. (2023). Audio deepfakes: A survey. *Frontiers in Big Data*, 5, Article 1001063. <https://doi.org/10.3389/fdata.2022.1001063>
- Milli, S., Carroll, M., Wang, Y., Pandey, S., Zhao, S., & Dragan, A. D. (2025). Engagement, user satisfaction, and the amplification of divisive content on social media. *PNAS Nexus*, 4(3), pgaf062. <https://doi.org/10.1093/pnasnexus/pgaf062>
- Newman, N. (2024). *Overview and key findings of the 2024 Digital News Report*. Reuters Institute for the Study of Journalism, University of Oxford.
- Nguyen, C. T. (2020). Echo chambers and epistemic bubbles. *Episteme*, 17(2), 141-161. <https://doi.org/10.1017/epi.2018.32>
- Paris, B., & Donovan, J. (2019). *Deepfakes and cheapfakes: The manipulation of audio and visual evidence*. Data & Society Research Institute.. https://datasociety.net/wp-content/uploads/2019/09/DS_Deepfakes_Cheap_FakesFinal-1-1.pdf
- Park, S., & Nan, X. (2026). Generative AI and misinformation: A scoping review of the role of generative AI in the generation, detection, mitigation, and impact of misinformation. *AI & Society*, 41, 1501–1515. <https://doi.org/10.1007/s00146-025-02620-3>

- Shoaib, M. R., Wang, Z., Taleby Ahvanooley, M., & Zhao, J. (2023). Deepfakes, misinformation, and disinformation in the Era of Frontier AI, generative AI, and large AI Models. In *2023 International Conference on Computer and Applications (ICCA)* (pp. 1-7). IEEE. <https://doi.org/10.48550/arXiv.2311.17394>
- TrueScreen. (2026). *The liar's dividend: when authentic content gets dismissed as fake*. <https://truescreen.io/articles/liars-dividend-digital-trust-crisis/>
- Turner, C. (2025). Online Echo Chambers, Online Epistemic Bubbles, and Open-Mindedness. *Episteme*, 22(1), 232-257. <https://doi.org/10.1017/epi.2023.52>
- Wardle, C., & Derakhshan, H. (2017). *Information disorder: Toward an interdisciplinary framework for research and policymaking*. Council of Europe. <https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-for-research/168076277c>
- Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs.