O QUE PRECISAS DE SABER SOBRE INTELIGÊNCIA ARTIFICIAL



Powered by

Centro Internet Segura

CNCS Centro Nacional de Cibersegurança PORTUGAL

EduQA Instituto de Educação, Qualidade e Avaliação

IPDJ Instituto Português do Desporto e Juventude

APAV Apoio à Vítima

Microsoft

Cofinanciado pela União Europeia

Apoio

Forum Estudante

#1
Introdução

#2
Cuidados a ter com a IA

#3
Desinformação. Como saber se algo é fake?

#4
Como manter a tua privacidade com a IA

#5
Melhora a tua utilização de IA

#6
IA e Saúde Mental. Como te podes proteger?

#7
Sabe mais: Recursos e ferramentas

#1 Introdução



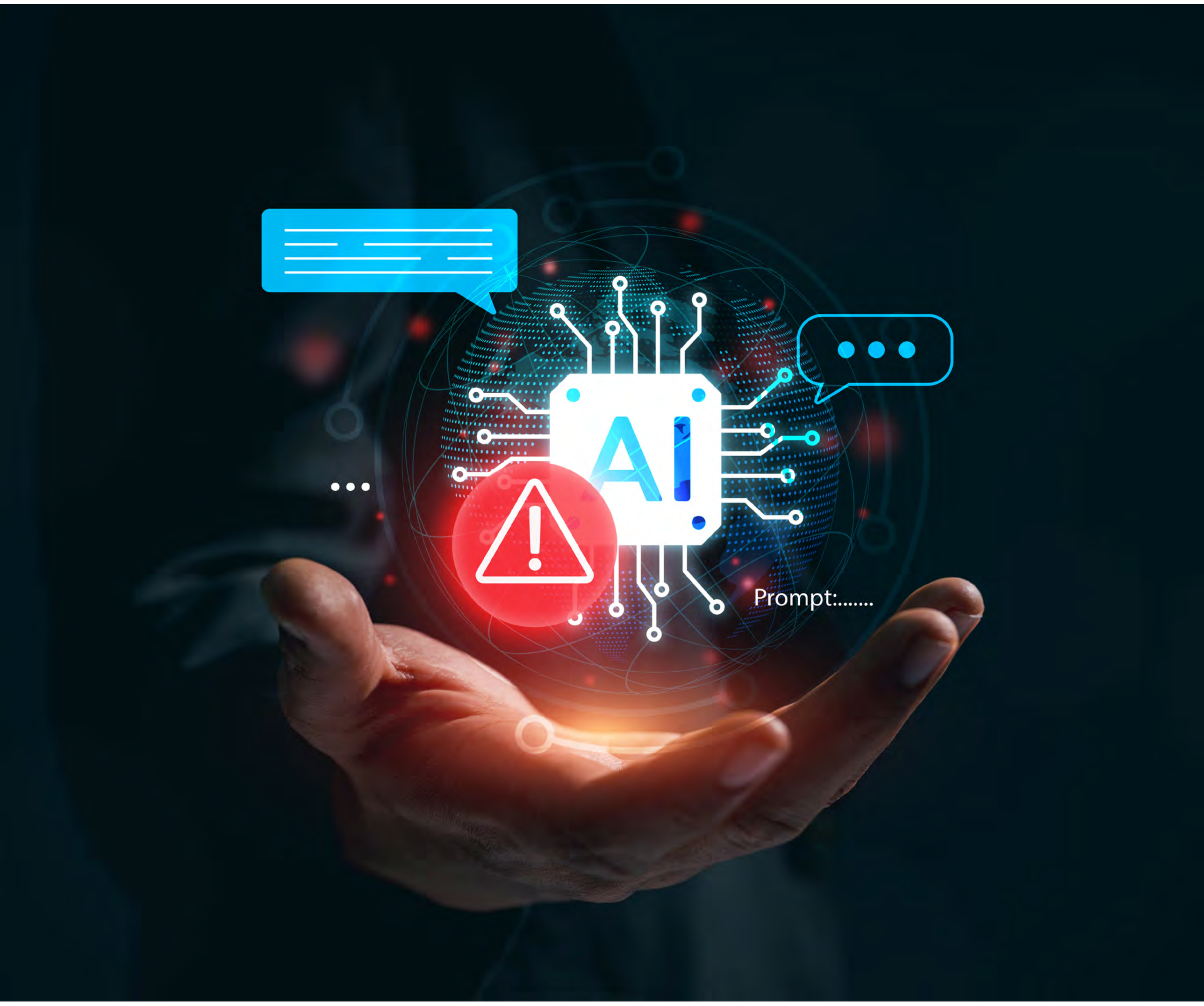
Ao longo dos últimos anos, as tecnologias de Inteligência Artificial (IA) tornaram-se parte do nosso dia a dia. Podemos encontrá-las em aplicações *chatbot*, mas também em contextos mais específicos como algoritmos de redes sociais ou filtros para edição de imagem. Para além das muitas potencialidades que estas tecnologias

nos trazem, existem também riscos, desafios e responsabilidades que importa compreender.

Que cuidados devemos ter com a IA? Podes confiar na veracidade da informação que ela te oferece? Quais os passos para melhorar a tua cibersegurança? E quais as práticas que te garantem uma utilização mais produtiva e saudável? Estas são algumas das questões a que vamos responder ao longo desta publicação, para que possas utilizar a IA de forma consciente, mais segura e desenvolvas boas-práticas de utilização.

Esta tecnologia é uma ferramenta poderosa, que pode fazer a diferença na tua vida. Descobre, ao longo das próximas páginas, como podes tirar partido das suas potencialidades mantendo-te seguro/a.

#2 Cuidados a ter com a IA



Descobre, neste capítulo, alguns cuidados que deves ter em conta durante a utilização de ferramentas de IA e os "vícios" que estas tecnologias evidenciam quando produzem respostas.

Tendo em conta a utilização crescente da IA no nosso dia a dia, há que ter em conta os riscos associados à sua utilização. Para começar, a partilha de dados sensíveis com modelos, ferramentas e aplicações de IA pode levar a que, mais tarde, estes dados sejam expostos. As LLM's (ver caixa) memorizam essas informações e podem, mais tarde, reproduzir esses dados em vez de gerarem algo novo.

O que são Large Language Models (LLM's)?

Os LLM's são sistemas de IA avançados compreendem e geram linguagem natural. São sistemas que integram informação através de dados que serviram de base à sua preparação através de aprendizagem automática. Podem ser utilizados em contextos pessoais ou profissionais.

Em agosto de 2025, o Business Digital Index da Cybernews revelou que 5 de 10 distribuidores de LLM's já sofreram fugas de dados. Outro risco é o de informação falsa ou inventada. Um fenómeno apelidado pela revista *Nature* de "alucinação", estas ferramentas podem dar respostas sem fundamento ou então espalhar informação falsa.

Através de modelos de IA também é possível criar imagens, vídeos ou áudios – conhecidos como *deepfakes* que podem ser utilizados para manipular opiniões, cometer fraudes, espalhar desinformação ou conteúdos nefastos ou ilegais.

Outros impactos que podem ser sentidos ou notados com a utilização da IA: extração de dados de treino a partir de alguns modelos; erros de utilização em contextos diferentes, como a saúde; infringir propriedade intelectual; e impacto no mercado de trabalho, gerando desigualdades socioeconómicas.



Que cuidados devo ter

Aqui ficam algumas recomendações sobre cuidados que deves ter ao utilizar estas ferramentas:

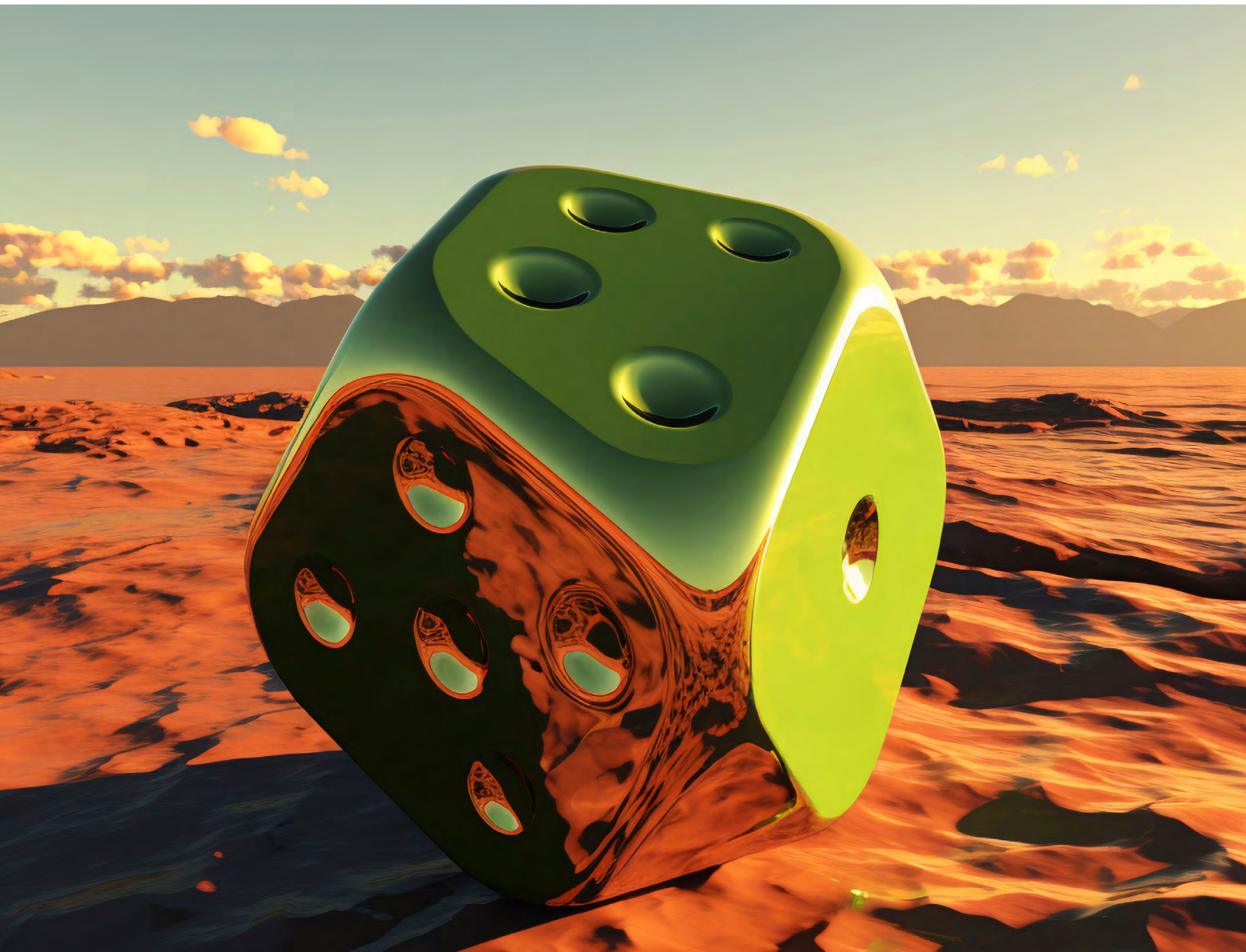
1. Não partilhes informação pessoal (nomes, números, *palavras-passe*);
2. Ao pesquisar informação **verifica** sempre os factos;
3. **Pede sempre fontes** para as informações que recebes;

4. **Não delegues** as decisões finais num sistema automático;
5. **Limita o teu uso de tempo e faz intervalos regulares**;
6. **Evita utilizar IA ou *chatbots* para suporte emocional**, pois estes podem responder de forma inadequada. Tem cuidado, em particular, para **não personalizares a IA e desenvolveres uma falsa relação afetiva**. Se te sentires sozinho ou precisares de falar com alguém sobre o que estás a sentir, recorre a amigos, familiares ou profissionais especializados;
7. Garante **equilíbrio entre o mundo real e a vida digital**;
8. **Consulta adultos ou profissionais** quando sentires algum desconforto ao utilizar estas ferramentas;
9. Confirma que a IA que vais utilizar **tem políticas claras** quanto à privacidade, conteúdo, moderação e monitorização.

Os “vícios” da IA

Estas tecnologias continuam a ser aperfeiçoadas, por isso, podem ocorrer imprecisões ou falhas que se manifestam do seguinte modo:

- desinformação, gerando informação falsa ou fabricada. A plataforma *Newsguard* detetou mais de mil notícias com falsas narrativas, em diferentes idiomas (incluindo português), produzidas com recurso a inteligência artificial;



- tendência para reproduzir ou amplificar estereótipos criando um enviesamento social/demográfico. Num artigo, a *Bloomberg* indica que, numa análise feita a mais de 5 mil imagens, o modelo *Stable Diffusion* estereotipa profissões de acordo com a raça e género das pessoas;
- em vez de apresentar uma perspetiva neutra pode mostrar enviesamentos quando apresenta resultados. A University of Southern California (USC) aponta que 38,6% dos factos apresentados pela IA apresentam um enviesamento;
- desigualdade ou conteúdo enviesado nos dados de treino, a IA pode amplificar este viés, podendo reproduzir padrões de discriminação ou estereótipos;
- também podem censurar resultados, evitando abordar certos assuntos tendo em conta aspetos políticos sensíveis. A *NBC News* reportou, em janeiro de 2025, que a IA chinesa *Deepseek*, quando questionada

acerca de temas politicamente sensíveis, respondia sempre que: “o tema estava fora do seu campo de conhecimento”;

- os LLM’s podem tornar-se demasiado lisonjeiros ou adaptar as suas respostas de maneira a agradar o utilizador, mesmo indo contra fatos ou ética.

Ao longo das próximas páginas vamos aprofundar várias dimensões ligadas a estes riscos: a desinformação, a cibersegurança e a saúde mental.

#3 Desinformação. Como saber se algo é fake?



Sabes quando uma publicação na internet não é confiável? Sabes como são produzidas notícias falsas e *deepfakes*? E sabes identificar estes conteúdos? Nós explicamos-te como.

Num inquérito realizado pela UNESCO, 85% dos participantes revelou estar preocupado com o efeito da desinformação na população do seu país. O sentimento está associado ao impacto real deste fenómeno. No total, 78% das pessoas leem desinformação nas redes sociais de forma intencional. Quem se cruza com notícias falsas acredita nelas 75% das vezes.

Os jovens com idades entre os 11 e os 14 anos acreditam não estar expostos a desinformação *online*. No entanto, são aqueles com idades compreendidas entre os 12 e os 22 anos que estão mais vulneráveis à desinformação através das redes sociais. Em Portugal são os jovens com idades entre os 25 e os 34 anos quem menos confia nas notícias. A grande maioria dos jovens portugueses (81%) considera já ter estado exposta a desinformação e notícias falsas.

Existem 7 tipos de informação falsa| desinformação

Menos prejudicial

Sátira ou paródia

Não tem intenção mas pode enganar

Conteúdo enganador

Uso enganador de informação para enquadrar um tema ou indivíduo

Conteúdo Impostor

Quando se fazem passar por fontes genuínas

Conteúdo Fabricado

Informação 100% falsa, criada com o intuito de enganar e causar dano.

Ligações falsas

Quando a imagem e o conteúdo não estão relacionados

Contexto falso

Quando conteúdo genuíno é partilhado com uma falsa contextualização

Conteúdo Manipulado

Quando imagens ou informações verdadeiras são editadas de maneira enganosa

Mais prejudicial

Fonte: Rapid Analysis: Digital misinformation / disinformation and Children, UNICEF Office of Global Insight and Policy (adaptado).

Porque existe desinformação?

A desinformação circula facilmente nas redes sociais. Através dos seus algoritmos as publicações mais partilhadas e visualizadas ganham destaque, promovendo assim a viralização e a criação de boatos ou teorias.

Estes conteúdos, como são chamativos, geram cliques e partilhas, tornam-se facilmente monetizáveis, fazendo com que os criadores de conteúdo ganhem dinheiro com estes *posts*.

As pessoas que criam estes conteúdos podem fazê-lo por convicção política ou porque são pagas para o fazer. Podem distorcer notícias, que vão de temas desde a política à saúde, e usar diferentes perfis para difundir a informação falsa, tendo como propósito causar uma reação. O Conselho da UE distingue de forma

clara as informações erradas, falsas e a desinformação. “As informações erradas são informações falsas que foram criadas ou divulgadas de forma não intencional. A desinformação é também uma informação falsa, mas criada deliberadamente e com o objetivo de causar danos.”



Como funciona a desinformação?

- Uma sociedade aberta e tolerante é muito mais resistente à desinformação. A desinformação é utilizada para dividir as pessoas e exacerbar divergências internas e conflitos ou controvérsias existentes.
- O objetivo dos divulgadores de desinformação é inundar-nos de informação para nos confundir e assoberbar de tal forma que já não sabemos em quem confiar.
- A desinformação pode propagar-se muito facilmente nas redes sociais, mediante caça-cliques (*clickbait*), robôs digitais (*bots*), histórias falsas, etc.
- *Deepfake* — a IA pode ser utilizada para criar vídeos ou imagens falsas convincentes.
- Os meios de comunicação social tradicionais também são usados para fins de desinformação.

Adaptado do documento "Conjunto de Ferramentas para Professores – Como Detetar e Combater a Desinformação", produzido pela Comissão Europeia e disponível **aqui**.

Características da desinformação

A desinformação pode surgir em diversos contextos e em numerosas plataformas diferentes. Alguma da desinformação visa especificamente os jovens e é concebida para os influenciar. Algumas das características comuns da desinformação assentam no facto de esta:

1. Falar às emoções da pessoa visada, levando a que esta tenha dificuldade em pensar de forma lógica e crítica.
2. Atacar o opositor, promovendo visões da realidade baseadas na noção de «Nós contra Eles».
3. Simplificar os factos e excluir o contexto.
4. Repetir uma ideia vezes sem conta.
5. Ignorar os contornos dos factos, apresentando apenas um lado da questão.
6. Manipular imagens de várias formas, como retoques e recortes.
7. Extrair imagens dos respetivos contextos originais, combinando-as com outras imagens, música/sons e textos para criar novos significados.

8. Recorrer a pessoas famosas e celebridades que o grupo-alvo admira.
9. Recorrer cada vez mais a falsificações baratas e falsificações profundas.
10. Apelam às emoções e são resistentes a qualquer tipo de prova.

Adaptado do documento "Orientações para professores e educadores sobre o combate à desinformação e a promoção da literacia digital", produzido pela Comissão Europeia e disponível **aqui**.



Sabes identificar notícias falsas?

Antes de clicares ou partilhares, verifica a fonte e cruza a informação:

1. Consulta fontes oficiais e confiáveis;
2. Lê as notícias até ao fim antes de partilhares;
3. Desconfia de títulos sensacionalistas;
4. Quando alguém partilhar uma notícia, verifica se o link é autêntico e seguro: examina se o domínio tem erros ortográficos, verifica o destino real dos links encurtados e verifica se o conteúdo corresponde ao título da notícia;
5. O uso da Inteligência Artificial traz maior sofisticação à desinformação pela capacidade de simular imagens, vozes e textos, que parecem verdadeiros. Compara com outras fontes.

Fonte: **Centro Nacional de Cibersegurança**

Como identificar *deepfakes*?

Em 2024 foi verificado a nível mundial um aumento de 245% em conteúdos *deepfake* em relação a 2023. Portugal não foi exceção e contou com um crescimento na ordem dos 1700%. Em 2025 foram partilhados mais de 8 milhões de conteúdos *deepfake*.

A Europol estima que 90% do conteúdo *online* poderá ser gerado de forma artificial já em 2026, com os *deepfakes* a serem rapidamente difundidos pelas redes sociais, apps de mensagens e de vídeo, tornando-os numa ferramenta poderosa para disseminar desinformação e informação falsa.

Desde 2 de agosto de 2026, os deepfakes (imagens, vídeos ou áudio que tenham sido editados ou gerados com recurso a IA) terão de ser rotulados. Os conteúdos gerados ou alterados por IA

terão também de incluir marcas legíveis por máquina, para que possam ser detetados com mais facilidade.

A plataforma **Better Internet for Kids (BIK)** ajuda-te a identificar um *deepfake*, *presta atenção a:*

- Falhas: como esta tecnologia ainda está a ser aperfeiçoada é dada a falhas, procura imperfeições no discurso, nos lábios, na pele, nos movimentos do locutor e até nas mãos (número de dedos, por exemplo);
- Áudio: Nos discursos, fica atento à voz, ao movimento dos lábios e a alterações de tom;
- Situações improváveis: às vezes surgem vídeos ou imagens que representam as pessoas em situações que não as imaginamos. O mesmo para publicações nas redes sociais;
- Faz *factcheck*: confirma autenticidade do vídeo, por exemplo, se for de uma pessoa com um cargo político, confirma em plataformas noticiosas de confiança;

- Usa Ferramentas de deteção – existem ferramentas em desenvolvimento e disponíveis ao público que permitem analisar conteúdos de vídeo à procura de indícios de deepfake. Como se trata de uma área em constante evolução e desenvolvimento, não há garantia de que estas ferramentas sejam (ou continuem a ser) rigorosas, mas são úteis para verificar se há pistas em conteúdos de vídeo.

Pensa como um *fact checker*!



1. Verifica

Começa pelo meio onde a notícia é publicada. Se for um site, pesquisa mais informação sobre o mesmo. Se o conteúdo vier das redes sociais, procura saber que fontes partilharam a informação.

2. Manipulação

Não te deixes enganar pelo título. Este pode ser utilizado para manipular as tuas emoções. Da mesma forma, um clip de vídeo pode isolar um excerto de uma entrevista ou discurso para que fique fora do contexto e contribuir para influenciar a tua opinião acerca de um tema ou indivíduo.

3. Desconfia

Desconfia do conteúdo sempre que não mencione fontes. Tenta descobrir a origem das imagens ou então vê se foram utilizadas fontes credíveis. Podes fazer pesquisas com recurso a imagem, seja num motor de busca ou num LLM.

4. Verifica a data

O *repost* também é uma técnica utilizada no mundo da desinformação, através da publicação de notícias ou vídeos antigos, alterando o seu contexto. Verifica sempre a data de publicação.

5. Demasiado insólito?

Há artigos falsos que são produzidos a partir de textos humorísticos. Se a situação te parecer exagerada, pesquisa e verifica a autenticidade.

6. Confirma

Se continuares com dúvidas, confirma com alguém da tua confiança (pai/mãe, professor/a, amigo/a).

7. Tem em conta as tuas emoções

Os conteúdos falsos têm como um dos objetivos influenciar a forma como te sentes em relação a determinado assunto. Se algum conteúdo te provocar uma reação emocional forte ou inesperada, pergunta sempre: quem é que vai beneficiar com isto?

8. Questiona

Confere e cruza informação, respondendo às perguntas: Quem? O quê? Quando? Como? Porquê? Onde?

#4 Como manter a tua privacidade com IA



Para usar a internet em segurança, é importante saber como proteger os nossos dados pessoais. Descobre como os podes proteger.

Nos primeiros oito meses da implementação do Regulamento Geral de Proteção de Dados (maio de 2018 – janeiro de 2019), de acordo com a **União Europeia** (UE), foram registadas 59 mil violações de dados pessoais no Espaço Económico Europeu.

A plataforma **IT Governance** indica que, só em dezembro de 2023, 204 violações de dados foram reportadas na Europa, com o número de dados pessoais afetados a situar-se acima dos 100 milhões. Já na União Europeia, de 131 incidentes, mais de 41 milhões de dados foram impactados. Globalmente, registaram-se 1.351 violações de dados, com um número superior a 2 mil milhões de dados comprometidos.

Há que ter em conta também o uso da IA como ferramenta utilizada pelos cibercriminosos para clonagem de vozes e geração de identidades falsas de modo a burlar outras pessoas.

O que são o RGPD, o RSD e o AI Act?

O Regulamento Geral de Proteção de Dados (RGPD) “estabelece as regras relativas ao tratamento, por uma pessoa, uma empresa, ou uma organização, de dados pessoais relativos a pessoas na UE”. Já o RSD é o Regulamento dos Serviços Digitais que inclui medidas para proteger os jovens e as crianças online. Podes saber mais **aqui**.

Desde 2 de agosto de 2026, o artigo 50.º do Regulamento Europeu da Inteligência Artificial, conhecido como AI Act, obriga a identificar claramente os deepfakes e outros conteúdos (imagem, vídeo, áudio ou texto) gerados ou alterados por IA. Estes conteúdos devem ainda incluir marcas legíveis por máquina, que facilitam a sua deteção. Existem exceções, como o uso puramente pessoal e as obras criativas, artísticas, satíricas ou ficcionais. Podes saber mais **aqui**.

Quem é mais impactado por estas violações de dados?

Um estudo divulgado no início de 2025, focado na realidade dos Estados Unidos da América, indica que mais de mil milhões de pessoas reportaram ter visto os seus dados ser comprometidos. Regista-se um aumento de 312%, quando comparado com os 419 milhões de 2023. São os dados de clientes e consumidores de empresas que costumam ser comprometidos. Os funcionários das empresas também costumam ver os seus dados expostos.

Que riscos estão associados à fraca proteção de dados pessoais?

A **SecurityBrief** identifica alguns dos riscos associados à exposição dos dados pessoais *online*, como o roubo de identidade, que em 2024 aumentou 150% apenas na Europa. Um relatório da *Verizon* indica métodos como o *phising* ou o roubo de credenciais como as principais formas de aceder a contas pessoais– e-mail, redes sociais ou serviços como *banking*.

De acordo com um **relatório do CNCS**, a perceção dos profissionais de Cibersegurança é que, em Portugal, as ameaças mais prevalentes são casos de pishing/smishing, ransomware, engenharia social, exploração de vulnerabilidade, software malicioso e comprometimento de conta.

Por que motivo devo proteger os meus dados?

Ao assegurares a proteção dos teus dados (nome, identificação, morada, dados de pagamento), impedes que criminosos os utilizem para abrir contas ou fazer compras em teu nome. A fuga destas informações tem igualmente custos financeiros associados.



Evitas também que se comprometam as tuas contas pessoais, como e-mail, redes sociais e outros serviços associados. A proteção dos teus dados assegura a proteção da tua privacidade.

Dicas para protegeres os teus dados

Usar uma rede segura

Uma rede segura é importante para manter os teus dados protegidos. Se a tua rede não estiver protegida fica mais fácil para os cibercriminosos conseguirem acesso às tuas informações.

Mantém as tuas redes sociais privadas

Através de métodos de engenharia social, podem conseguir extrair informações pessoais conhecendo a tua pegada digital. Fica atento a quem adicionas nas tuas plataformas sociais e assegura-te que a tua palavra-passe cumpre todos os requisitos de segurança.

Não guardes informação

Os *websites* e os *browsers* permitem que guardes informação sensível, como palavras-passe ou o número do teu cartão bancário. Guarda as tuas palavras-passe num gestor de *passwords* e salva qualquer informação sensível em ficheiros encriptados.

Confirma antes de clicar

Os *links*, *pop-ups* e *downloads* podem conter vírus e *malware*. Confirma sempre a fonte antes de clicar.

Desativa o *Bluetooth* e *Wi-fi*

As definições de *wi-fi* e *bluetooth* podem ser exploradas caso estejam ativas enquanto não estão a ser usadas. Não faças tarefas que envolvam a utilização de dados sensíveis quando estás numa rede pública.

Cria uma palavra-passe forte

O CNCS recorda que é importante usar palavras-passe fortes (pelo menos 12 caracteres de tipo diversificado, sem palavras associadas ao utilizador), ter uma diferente por cada plataforma que se usa, alterar as que são atribuídas pelo fabricante ou instalador (no Wi-Fi doméstico, por exemplo) e não as partilhar com ninguém.

Evita *sites* desconhecidos

Verifica os protocolos e certificações de segurança de um site, bem como os métodos de pagamento. Os cibercriminosos podem criar sites falsos, à semelhança dos verdadeiros, para espalhar *malware* e fazer ataques como *phishing*.

Faz *back-ups* com segurança

Guarda regularmente o teu trabalho ou ficheiros num dispositivo externo ou então num local onde consigas controlar quem tem acesso.

Usa autenticação multifactor (MFA)

O múltiplo fator de autenticação é uma forma de garantir que a segurança das contas online não fica exclusivamente dependente da palavra-passe. De acordo com a brochura "Ler Antes de Clicar", do CNCS, através deste mecanismo, sempre que alguém procura aceder a uma conta, tem de acionar pelo menos um segundo fator de autenticação para entrar na conta, para além da palavra-passe – por exemplo, a introdução de um código enviado por SMS para o telemóvel do utilizador. Por esta razão, sempre que possível, deve ativar-se esta funcionalidade.

#5 IA e Saúde Mental. Como te podes proteger?



Quais os riscos e desafios associados à utilização de Inteligência Artificial, do ponto de vista da saúde mental? Neste artigo, contamos-te tudo.

“Deslizar para dentro do abismo”- foi desta forma que um conjunto de especialistas em saúde mental **descreveu ao *The Guardian*** os riscos da utilização de *chatbots* de IA por parte de pessoas vulneráveis. O conjunto de psicoterapeutas alertou para o facto desta utilização poder aumentar os sintomas de ansiedade, alimentar a dependência emocional, aumentar o recurso ao autodiagnóstico ou amplificar pensamentos sombrios ou delirantes: “Embora algumas pessoas recebam conselhos úteis, estamos preocupados com outras que recebem informação incorreta e enganadora”.

«Os sistemas baseados em LLM’s estão a ser usados como companheiros, confidentes e terapeutas e algumas pessoas obtêm daí benefícios reais. Mas encontrámos riscos significativos e é importante falar sobre algumas das diferenças fundamentais»

Investigadores da Universidade de Stanford

Esta preocupação é corroborada pelas conclusões de um **estudo de 2025 da Universidade de Stanford** – os *chatbots* de IA não só não têm a mesma eficácia quando comparados com terapeutas humanos, como também contribuem para a manutenção de estigmas e respostas perigosas. Um dos investigadores, deixa um aviso: "Encontrámos riscos significativos e é importante falar sobre algumas das diferenças fundamentais".

Os investigadores concluíram que os vários serviços de *chatbot* especializados em saúde mental contribuíram para estigmatizar certas condições ou transtornos como a dependência de álcool e a esquizofrenia. De igual forma, estas aplicações de IA não evidenciaram capacidade para "ajudar o paciente a repensar" a situação em que se encontrava, chegando mesmo a "potenciar comportamentos perigosos".

Se te sentires vulnerável emocionalmente, deves sempre ter em consideração a importância de falar com amigos, familiares, professores ou profissionais de saúde mental. A Linha Internet Segura está disponível das 08h00 às 23h00, de forma gratuita, através do número **800 21 90 90** ou do linhainternetsegura@apav.pt e formulário de apoio.

Um estudo da *Harvard Business Review* de 2025 concluiu que, dentro dos vários usos de inteligência artificial generativa, a terapia e os "*chatbots* de companhia" estão no topo da lista. A revista *Nature* acrescenta: "Existe um número crescente de casos em que utilizadores vulneráveis ficam enredados em dependência emocional e, por vezes, em interações prejudiciais com *chatbots*". Duas das consequências negativas para a saúde mental são a "dependência emocional disfuncional" e "o sentimento de ausência psicológica".

A questão torna-se mais problemática pelo facto de, ao longo dos últimos tempos, existirem vários casos noticiados de pessoas que alegam ter recebido instruções sobre como incorrer em práticas de automutilação ou suicídio. O **relatório "Fake Friend**", produzido pelo Centro de Combate ao Ódio Digital (Center for Countering Digital Hate) concluiu que "ao final de alguns minutos de interações simples com o ChatGPT, o sistema produziu instruções relacionadas com automutilação, planeamento de suicídio, distúrbios alimentares e abuso de substâncias". De 1200 respostas recebidas, 53% continham conteúdo prejudicial, conclui a investigação.

Nem todos os *Chatbots* são iguais. Se precisares tirar dúvidas sobre Segurança na Internet, o ***Chatbot* da Linha Internet Segura** é anónimo, confidencial, só tem informação fidedigna e baseada na lei Portuguesa e podes pedir a qualquer momento para seres contactado por um técnico especializado da **Linha Internet Segura**.

Os vários modelos de IA generativa implementam salvaguardas para estes casos, contudo, em julho deste ano, **dois investigadores concluíram** que "ao colocar as questões num quadro de investigação académica, conseguiam frequentemente contornar [essas] salvaguardas". Um dado que nos deve levar a pensar sobre a forma como interagimos com IA. A importância fica clara nas palavras de um dos investigadores: "Não há forma de garantir que um sistema de IA seja 100% seguro".

É muito importante que evites humanizar as plataformas de IA como conselheiros afetivos. Tem em conta que estas soluções tecnológicas são desenhadas tendo como objetivo de simular empatia, manter o contexto e personalizar o tom da conversa. Do ponto de vista científico, não há efetivamente uma relação de amizade recíproca. Embora o resultado pareça humano, mas não existe consciência ou emoção nestas plataformas.

Contactos e serviços de apoio emocional



Sabe quem podes e deves contactar, em caso de emergência ou simples necessidade de falar, desabafar ou pedir ajuda.

Linha Internet Segura (LIS)

08h00-23h00 (dias úteis) → ✆800 21 90 90
linhainternetsegura@apav.pt

Chatbot LIS

https://cislab.cncs.gov.pt/chatbot-linha-internet-segura/

SOS Voz Amiga

15h30 – 00h30 → ✆213544545 | ✆912802669 | ✆963 524660
21h00 – 24h00 → ✆800 209 899 (gratuita)

Conversa Amiga*

15h00 – 22h00 → ✆808 237 327 | ✆210 027 159

**0,07€ no 1º minuto e, a partir do 2º minuto, 0,026€ (horário normal) e 0,008€ (horário económico*

Vozes Amigas de Esperança de Portugal

16h00 – 22h00 → ✆222 080 707

Voz de Apoio

21h00 – 24h00 → ✆225 506 070

SNS24 – Serviço Nacional de Saúde

✆808 24 24 24

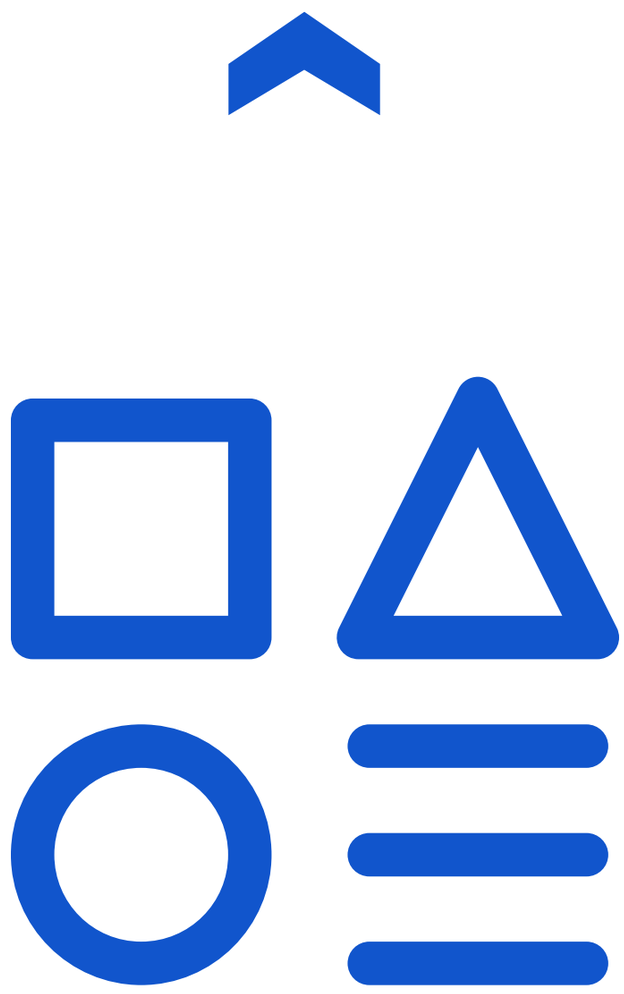
Número Nacional de Emergência Médica

✆112

#6 Melhora a tua utilização da IA



Queres tornar-te um *power user* de Inteligência Artificial? Segue este conjunto de dicas que preparámos para ti para garantires que fazes a melhor utilização destas tecnologias.



#1 Conhece os elementos básicos da tecnologia

Começa por dominar os conceitos tecnológicos básicos envolvidos. Os LLM's, por exemplo, consistem em ferramentas que, a partir da análise de padrões de linguagem, preveem qual a resposta mais provável. Isto significa que o modelo não tem consciência nem opiniões próprias, apenas gera conteúdo tendo por base o que já analisou. Esta é informação muito relevante para que interpretes as respostas que te são fornecidas por um LLM, conhecendo as suas potencialidades, mas também as suas limitações.



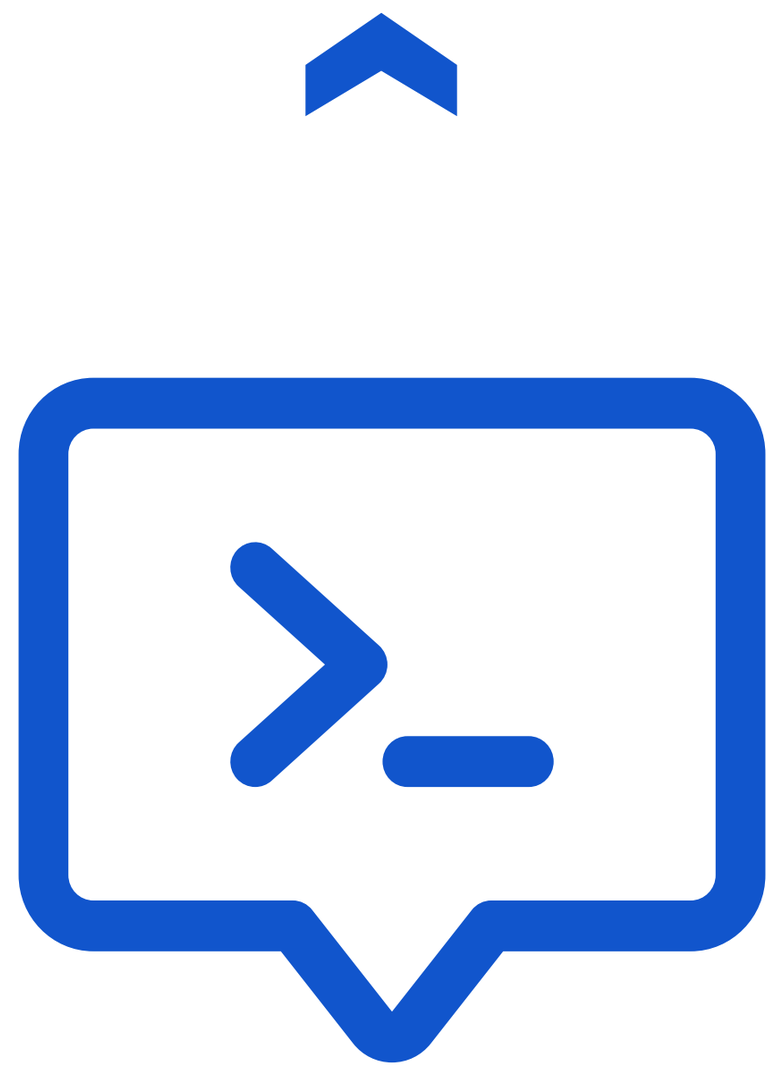
#2 Mantém sempre o pensamento crítico

É muito importante que mantenhas sempre o pensamento crítico quando estás a utilizar ferramentas de Inteligência Artificial. Pelas razões explicadas no ponto anterior, nem tudo o que uma IA gera está correto. Por isso, deves sempre verificar as informações com outras fontes, sobretudo se estiveres prestes a tomar decisões importantes. Na dúvida, a melhor via de ação será sempre conversar com alguém sobre o tema ou informação que estás a investigar, para evitar que estejas a dar passos errados. Outra estratégia que promove pensamento critíco é pensares primeiro no tema e só depois pedires ajuda à IA. Além de teres uma base tua, podes discutir o tema com a IA, tornando a discussão mais rica e não te deixando levar pela opinião geral antes de formares a tua.



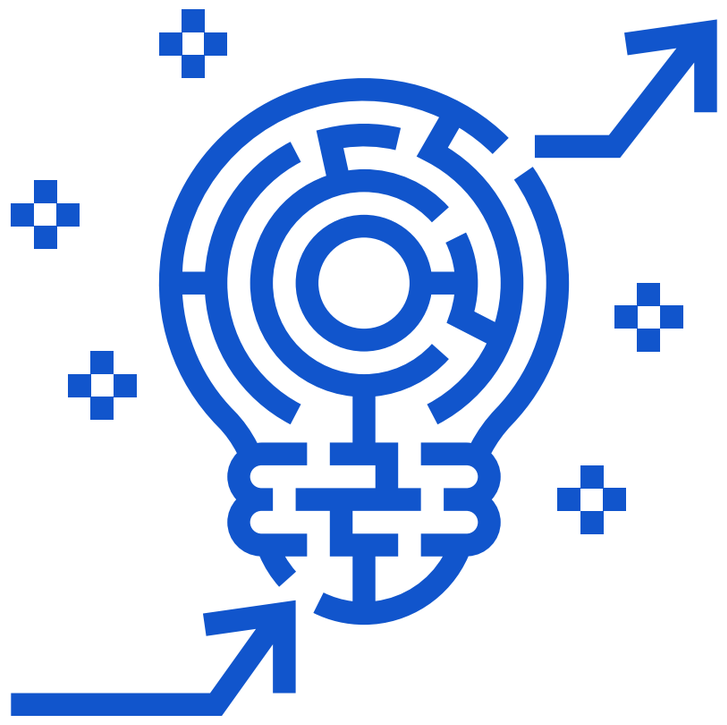
#3 Faz uma utilização ética da IA

Fazer uma utilização ética da IA envolve várias questões. Desde logo, implica não recorrer a estas tecnologias para criar conteúdos falsos ou prejudiciais, mas também respeitar os direitos autorais e a propriedade intelectual. Por outro lado, a própria utilização de IA tem impacto ambiental, pelo que deves escolher os momentos em que o fazes. É muito importante que evites depender da IA para todas as tarefas, garantindo tempo para pensar e criar sem tecnologia. Por fim, se utilizaste IA de forma estrutural num determinado conteúdo, deves ser transparente e identificá-lo.



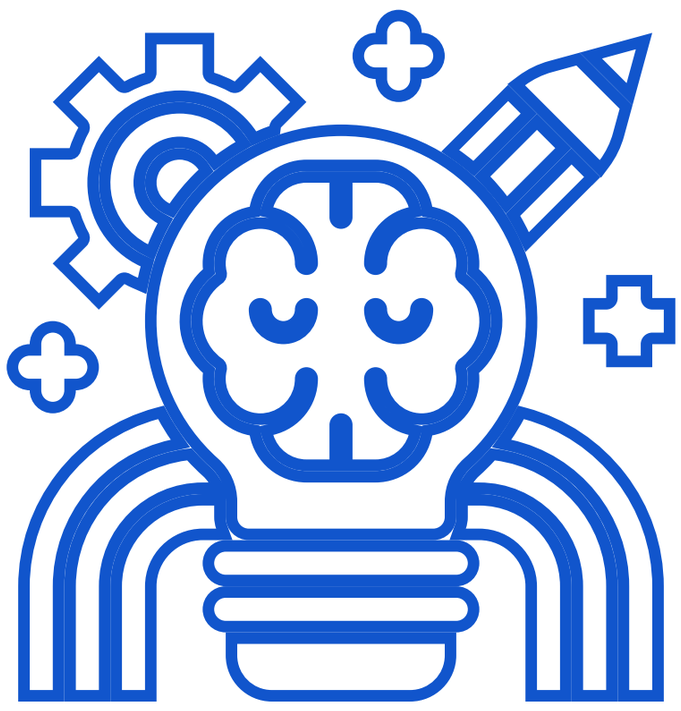
#4 Aprende a criar os melhores *prompts*

A forma como pedes algo a uma IA (escrever um *prompt*) vai influenciar muito o resultado que obténs. Começa por ser o mais claro e específico relativamente ao que pretendes (qual a dimensão do texto, o contexto, tom, o objetivo, etc...). Recorrer a comparações também te vai ajudar a obter respostas mais realistas, sendo que também podes limitar ou orientar a "criatividade" de um LLM, pedindo factos concretos. Por fim, não te fiques pela primeira resposta: procura aperfeiçoar o resultado com novos *prompts* ou pedidos.



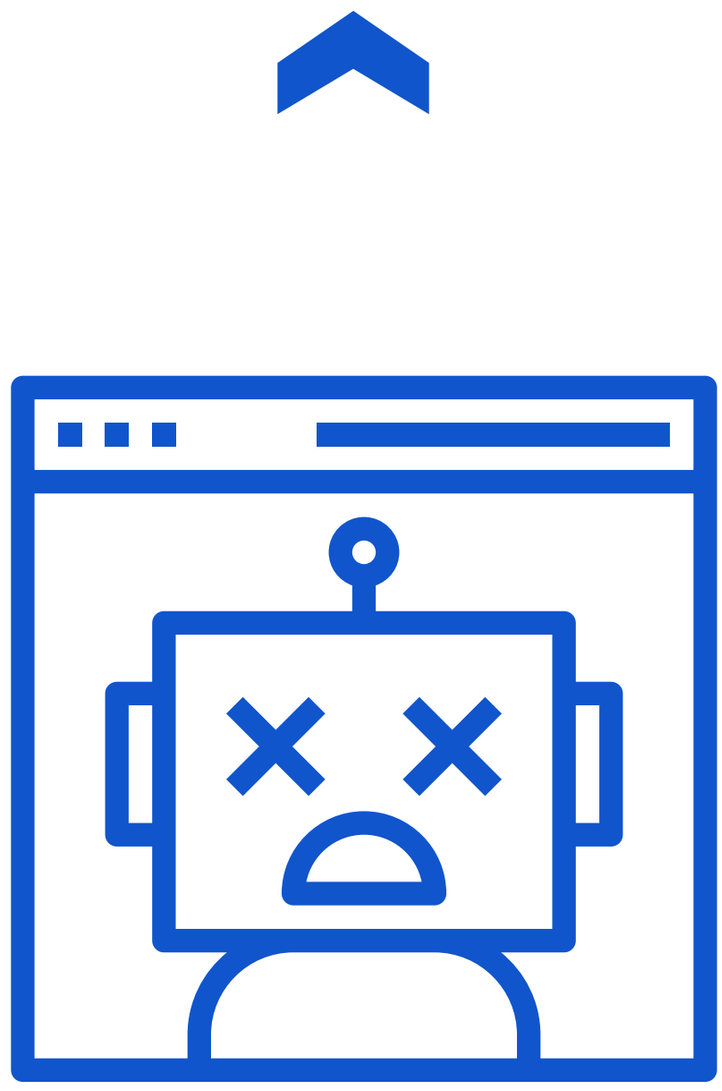
#5 Utiliza todas as potencialidades

Existem muitas ferramentas de IA como *chatbots*, editores de imagem e de música, análise de dados, entre outros. Conheceres as várias ferramentas especializadas vai permitir-te utilizar soluções para necessidades específicas. Por outro lado, existem potencialidades que podes explorar, como configurar as funcionalidades de uma ferramenta para as tuas necessidades (quanto ao estilo, linguagem ou privacidade, por exemplo), de forma a conseguires resultados mais produtivos e seguros.



#6 Usa a IA como um apoio para a criatividade

Um dos princípios fundamentais para uma utilização produtiva de IA é ver estas tecnologias como um apoio ou um parceiro de criação de conteúdo, evitando ver estas ferramentas como meras “máquinas” que fazem o trabalho por ti. A melhor utilização da IA é aquela que a vê como um apoio para a tua criatividade e talento, potenciando o resultado final. Nesse sentido, podes recorrer à IA para gerar ideias ou explorar novos temas, mas tendo sempre em conta que a tua ação deve ser central nessa utilização.



#7 Analisa os erros da IA

Já aqui referimos os erros que a Inteligência Artificial pode cometer. Mais do que apenas identificar estes lapsos, procura entender qual a razão ou razões que levaram a que esse erro fosse cometido. Para tal, podes contar com a ajuda da ferramenta ou aplicação na interpretação do lapso que foi cometido (informação falsa, alucinação, entre outros). Desta forma, vais poder identificar padrões e estar de sobreaviso para uma situação semelhante, no futuro.

#7 Sabe mais: recursos e ferramentas



Conhece aqui algumas das formas como podes aumentar os teus conhecimentos sobre Inteligência Artificial, rumo a uma utilização segura e produtiva.

Cursos gratuitos

Plataforma NAU

“Cidadão Ciberinformado”, Centro Nacional de Cibersegurança (CNCS)

Compreende e aprende a identificar a veracidade de uma notícia ou de outro tipo de informação publicada online.

“Cidadão Ciberseguro”, Centro Nacional de Cibersegurança (CNCS)

Garante um conjunto de competências que te permitem utilizar o ciberespaço e navegar de forma segura.

“Inteligência Artificial Generativa”, Academia Portugal Digital

Sabe como a IA Generativa pode transformar a criação de conteúdos, explorando o futuro desta revolução digital.

“Introdução à Inteligência Artificial”, Academia Portugal Digital

Entende os princípios básicos da Inteligência Artificial Generativa, bem como algumas das suas aplicações práticas.

“Aplicações de Inteligência Artificial no dia a dia”, Academia Portugal Digital

Descobre como a IA está presente no teu dia a dia, aprendendo a tornar a vida mais prática, segura e eficiente.

Coursera

IA Generativa para crianças, pais e professores

Aprende a explorar com IA generativa, nomeadamente sabendo como criar os melhores *prompts* com este curso gratuito da Universidade de Vanderbilt.

Sites e aplicações

CIS-LAB – Laboratório de Cibersegurança

Este é um espaço para todos aprenderem sobre o mundo digital e cibersegurança. Descobre oportunidades, treina competências e partilha experiências!

Experience AI

Uma iniciativa conjunta da Google e da Fundação Raspberry Pi, esta plataforma permite aceder a uma introdução às tecnologias de IA e de Machine Learning.

Classificador de Imagens GenAI

Esta iniciativa, de professores finlandeses, consiste numa plataforma que permite entender conceitos básicos de IA generativa, treinando um modelo personalizado.

IBM Skills Build

Domina os conceitos básicos de IA Generativa com esta plataforma que reúne vários cursos criados pela IBM.

Minecraft AI Detective

Este DLC gratuito do jogo Minecraft promete oportunidades únicas de aprendizagem e criatividade sobre IA.

CRAFT | Recursos de Literacia de IA

Este portal da Universidade de Stanford agrega conteúdos educativos sobre Inteligência Artificial para estudantes do ensino secundário.

Relatórios e Estudos

Relatório **“Me, Myself and AI research”, Internet Matters**

Compreende os riscos e benefícios da interação de jovens com chatbots de Inteligência Artificial.

Relatório **“A child rights audit of GenAI in EdTech”, Digital Futures For Children**

Quais os direitos das crianças e jovens no que diz respeito à sua privacidade, segurança e educação?

Relatório **“Talk, Trust, and Trade-Offs: How and Why Teens Use AI”**

Relatório **“Guidance on AI and children Version 3.0 | Recommendations for AI policies and systems that uphold child rights”**

Referências

→ https://www.edpb.europa.eu/news/news/2024/edpb-opinion-ai-models-gdpr-principles-support-responsible-ai_en

→ https://www.nature.com/articles/s41586-024-07421-0

→ https://www.usenix.org/conference/usenixsecurity21/presentation/carlini-extracting

→ https://www.who.int/publications/i/item/9789240029200

→ https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

→ https://www.ilo.org/publications/generative-ai-and-jobs-refined-global-index-occupational-exposure

→ https://arxiv.org/pdf/2311.05232

→ https://www.who.int/europe/news/item/25-09-2024-teens--screens-and-mental-health

→ https://news.stanford.edu/stories/2025/08/ai-companions-chatbots-teens-young-people-risks-dangers-study

→ https://www.ed.ac.uk/news/2024/teens-who-up-screen-time-raise-mental-health-risks

→ https://jopm.jmir.org/2025/1/e69449

→ https://onlinelibrary.wiley.com/doi/full/10.1002/aaai.12188

→ https://www.newsguardtech.com/special-reports/ai-tracking-center/

→ https://www.mdpi.com/2673-2688/6/5/106

→ https://www.bloomberg.com/graphics/2023-generative-ai-bias/

→ https://arxiv.org/abs/2508.15798

→ https://viterbischool.usc.edu/news/2022/05/thats-just-common-sense-usc-researchers-find-bias-in-up-to-38-6-of-facts-used-by-ai/

→ https://arxiv.org/abs/2401.16558

→ https://www.sciencedirect.com/science/article/pii/S0736585324001151

→ https://arxiv.org/abs/2508.15798

→ https://www.unesco.org/sites/default/files/medias/fichiers/2023/11/unesco_ipsos_survey.pdf

→ https://www.unesco.org/sites/default/files/medias/fichiers/2023/11/unesco_ipsos_survey.pdf

→ https://www.projectreal.co.uk/sessions.html

→ https://expresso.pt/sociedade/tecnologia/2025-08-04-jovens-acreditam-que-nao-estao-expostos-a-desinformacao-na-internet-76167772

→ https://expresso.pt/geracao-e/2025-06-12-nao-ha-consciencia-do-perigo-das-redes-sociais-jovens-so-querem-saber-o-que-se-passa-e-aceitam-toda-a-informacao-que-chega-9c25955b

→ https://www.publico.pt/2024/05/15/p3/noticia/jovens-portugueses-25-34-anos-sao-menos-confiam-noticias-2090486

→ https://sol.sapo.pt/2025/02/17/maioria-dos-jovens-portugueses-ja-foi-exposto-a-desinformacao/

→ https://www.internetsegura.pt/FakeNews

→ https://sumsub.com/newsroom/deepfake-cases-surge-in-countries-holding-2024-elections-sumsub-research-shows/

→ https://sumsub.com/newsroom/deepfake-cases-surge-in-countries-holding-2024-elections-sumsub-research-shows/

→ https://www.europarl.europa.eu/RegData/etudes/BRIE/2025/775855/EPRS_BRI%282025%29775855_EN.pdf

→ https://better-internet-for-kids.europa.eu/en/learning-corners/teachers-and-educators/deepfakes

→ https://gdpr.eu/gdpr-requirements-data-breach-reporting/

- https://www.itgovernance.eu/blog/en/data-breaches-and-cyber-attacks-in-europe-in-december-2023-100884532-records-breached

- https://securitybrief.co.uk/story/identity-fraud-in-europe-surges-by-150-deepfakes-on-the-rise

- https://www.enisa.europa.eu/topics/cyber-threats/threat-landscape

- https://newsroom.ibm.com/2024-07-30-ibm-report-escalating-data-breach-disruption-pushes-costs-to-new-highs

- https://www.theguardian.com/society/2025/aug/30/therapists-warn-ai-chatbots-mental-health-support

- https://hai.stanford.edu/news/exploring-the-dangers-of-ai-in-mental-health-care

- https://www.webpronews.com/mit-study-ai-companions-ease-loneliness-but-risk-dependency-and-abuse

- https://counterhate.com/wp-content/uploads/2025/08/Fake-Friend_CCDH_FINAL-public.pdf

- https://www.latimes.com/science/story/2025-07-31/chatgpt-perplexity-and-other-ai-self-harm-suicide-information

- https://www.nau.edu.pt/pt/curso/inteligencia-artificial-generativa

→ https://www.nau.edu.pt/pt/curso/introducao-a-inteligencia-artificial

→ https://www.nau.edu.pt/pt/curso/aplicacoes-de-inteligencia-artificial-no-dia-a-dia

→ https://www.coursera.org/learn/generative-ai-for-kids

→ https://experience-ai.org/en

→ https://tm.gen-ai.fi/image/general

→ https://skillsbuild.org/students/course-catalog/artificial-intelligence

→ https://www.minecraft.net/en-us/article/reed-smart-ai-detective

→ https://craft.stanford.edu

→ https://www.internetmatters.org/hub/research/me-myself-and-ai-chatbot-research

Powered by

Centro Internet Segura

CNCS Centro Nacional de Cibersegurança PORTUGAL

EduQA Instituto de Educação, Qualidade e Avaliação

IPDJ Instituto Português do Desporto e Juventude

APAV Apoio à Vítima

Microsoft

Cofinanciado pela União Europeia

Apoio

Forum Estudante