



HeadStart

a future in AI for girls



www.headstart-ai.eu

MÓDULO 2

A ética da IA - Capacitar jovens para um futuro justo



Co-funded by
the European Union



Visão geral do módulo

Este módulo explora o papel da ética na Inteligência Artificial. Os alunos descobrirão como a IA pode refletir e reforçar preconceitos sociais, porque é que a justiça e a responsabilidade são essenciais na conceção da IA e como analisar criticamente os conteúdos gerados pela IA. Através de actividades práticas, examinarão os estereótipos nas ferramentas de imagem da IA e reflectirão sobre a forma como as diversas perspectivas podem ajudar a criar tecnologias mais inclusivas e responsáveis.

01

Definição de ética

02

Definição de ética na IA

03

Panorama geral das questões éticas suscitadas pela IA

04

IA e preconceitos - Actividades

05

Termos-chave e recapitulação

Objectivos de aprendizagem

Módulo 2

No final deste módulo, serão capazes de:

- Definir ética e explicar a sua relevância nas decisões quotidianas
- Compreender a importância da ética no desenvolvimento da IA
- Identificar os riscos éticos da IA, como preconceitos e a falta de transparência
- Reconhecer como o conteúdo gerado pela IA pode reforçar estereótipos
- Explorar a forma de detetar e reduzir os preconceitos nas ferramentas de geração de imagens da IA

01

Ética



O que é a ética?



1

Definição de ética

A ética refere-se aos princípios que orientam a nossa tomada de decisões sobre o que é certo ou errado, bom ou mau. Molda o comportamento individual, as normas sociais e as práticas profissionais.

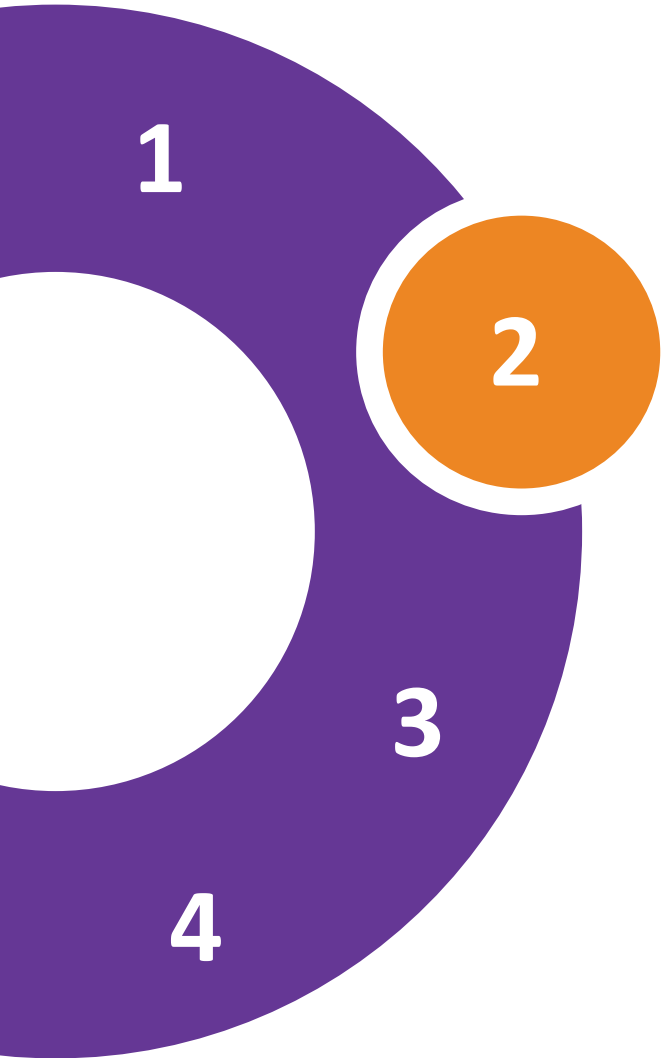
O que é a ética?

Componentes principais

Moralidade: Crenças pessoais sobre o certo e o errado.

Valores: Prioridades e normas que os indivíduos ou as sociedades consideram importantes.

Princípios: Verdades ou regras fundamentais que orientam as acções.



O que é a ética?

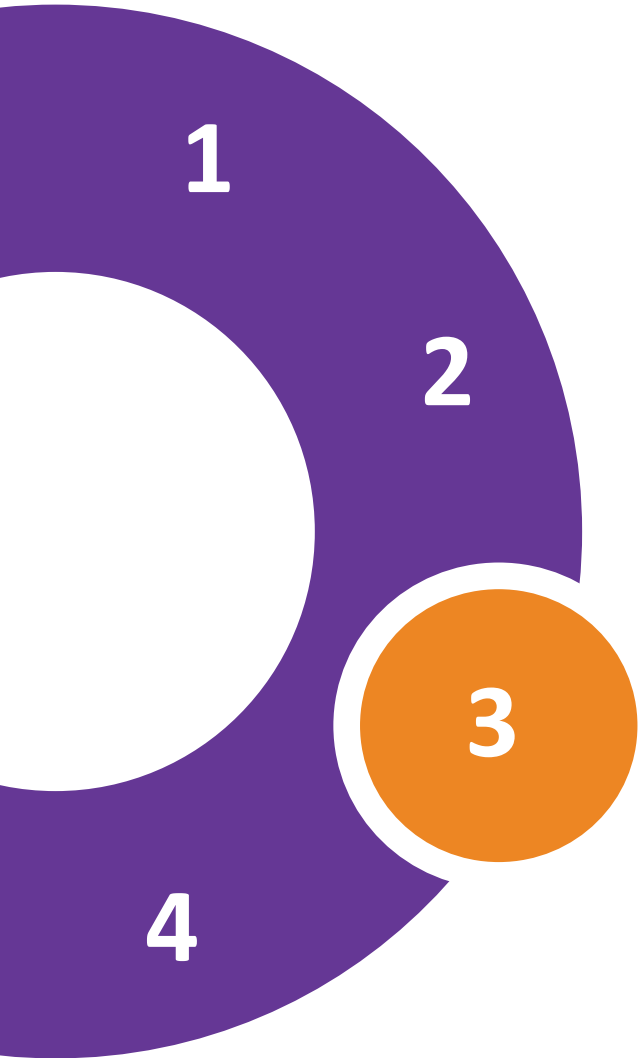
Ramos da ética

Ética pessoal: Como os indivíduos agem com base nas suas crenças.

Ética profissional: Regras e normas em locais de trabalho ou indústrias (por exemplo, ética médica).

Ética social: Acções colectivas e normas sociais (por exemplo, justiça, igualdade).

Ética ambiental: Responsabilidades para com o planeta e outras espécies.



O que é a ética?

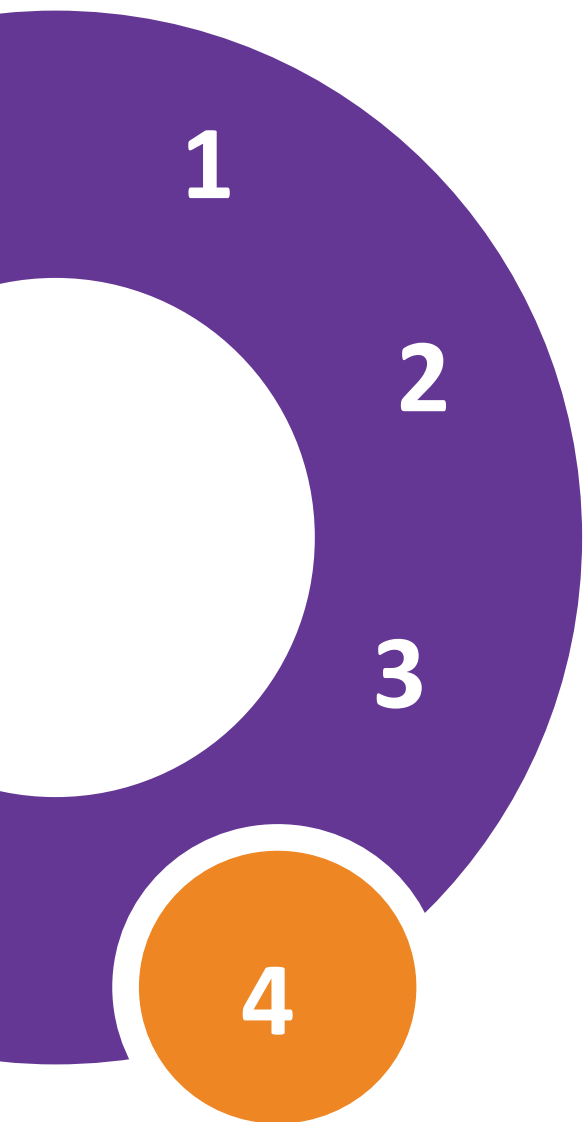
Porque é que a ética é importante?

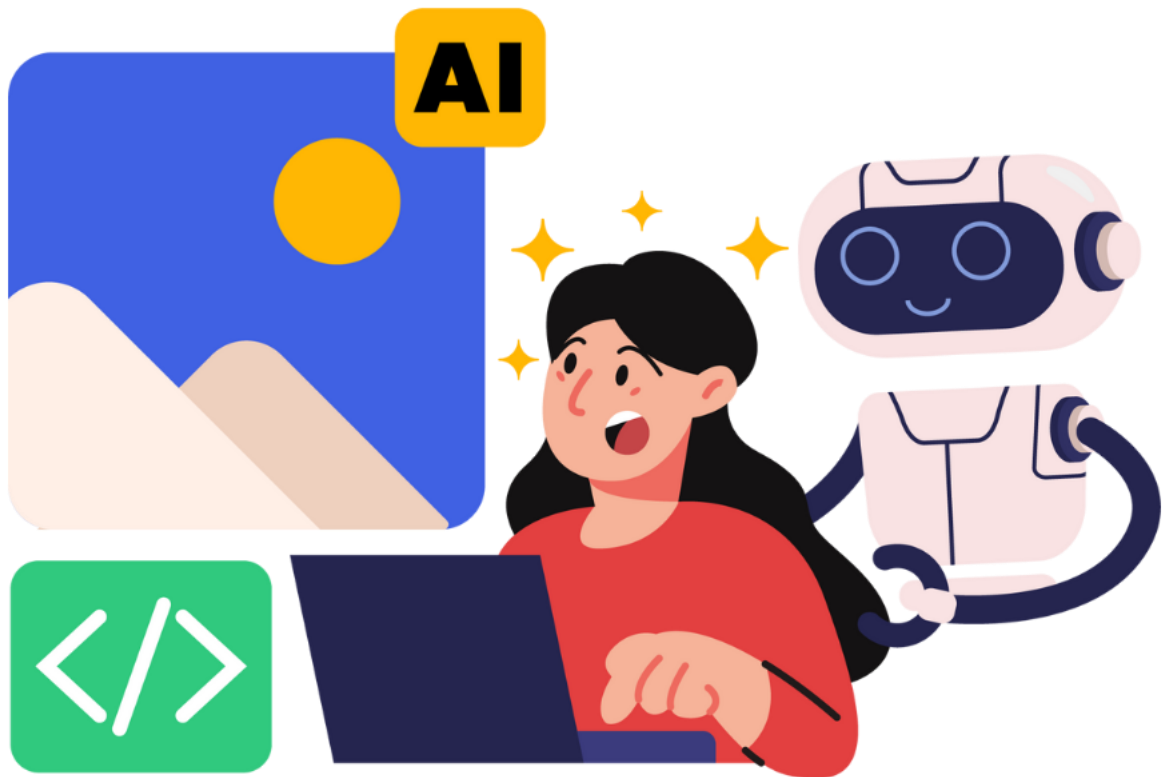
Cria confiança: Assegura a fiabilidade e a responsabilidade nas relações e organizações.

Promove a equidade: Apoia a igualdade e a justiça na sociedade.

Incentiva a responsabilização: Orienta as pessoas e as instituições para agirem de forma responsável.

Promove o crescimento: Promove o progresso através da resolução de dilemas morais.





Tomada de decisões éticas

Perguntas-chave a fazer!!

É correto?

É justo?

Respeita os outros?

Quais são as consequências?

O que é a ética?

A ética consiste em **princípios que orientam as decisões** sobre o certo e o errado, moldando o comportamento, a sociedade e o trabalho.

Porque é que é importante?

A ética gera **confiança**, promove a **justiça**, garante a **responsabilização** e **impulsiona o progresso**.

Decisões éticas

Perguntar: É **correto**? **Justo**? **Respeitoso**? Quais são as **consequências**?

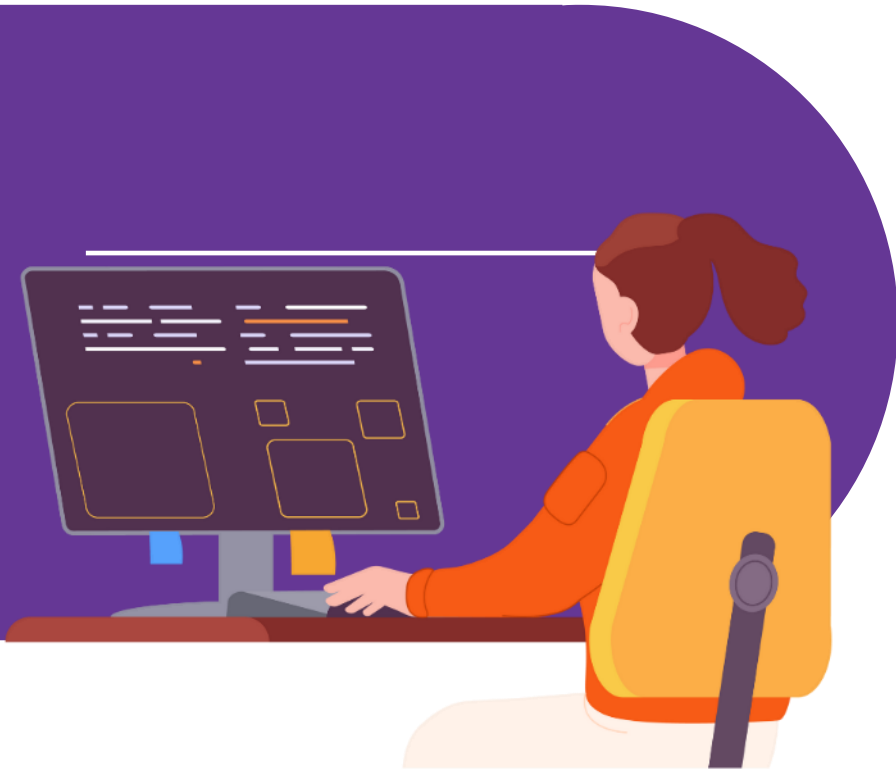
RECAP



02

Ética e IA





Equilíbrio entre inovação e responsabilidade

À medida que a inteligência artificial se torna parte integrante das nossas vidas e empregos, os princípios da ética devem estender-se aos sistemas de IA.

A ética da IA garante que a inteligência artificial é concebida e utilizada de forma responsável, centrando-se na equidade, transparência e respeito pela privacidade.

Equilíbrio entre inovação e responsabilidade

JUSTIÇA

Igualdade de tratamento para todos os indivíduos, independentemente do seu historial. Evita preconceitos e discriminação nos sistemas de IA.



TRANSPARÊNCIA

Compreensão clara das decisões em matéria de IA e honestidade na sua utilização.



PRIVACIDADE

Proteção dos dados pessoais e da autonomia dos utilizadores, bem como reforço da confiança, garantindo que a informação está segura e que só é utilizada de forma adequada



03

Principais preocupações éticas no domínio da inteligência artificial



Principais preocupações éticas em matéria de IA

1. Preconceito e equidade

5. Privacidade e vigilância

2. Erros de informação

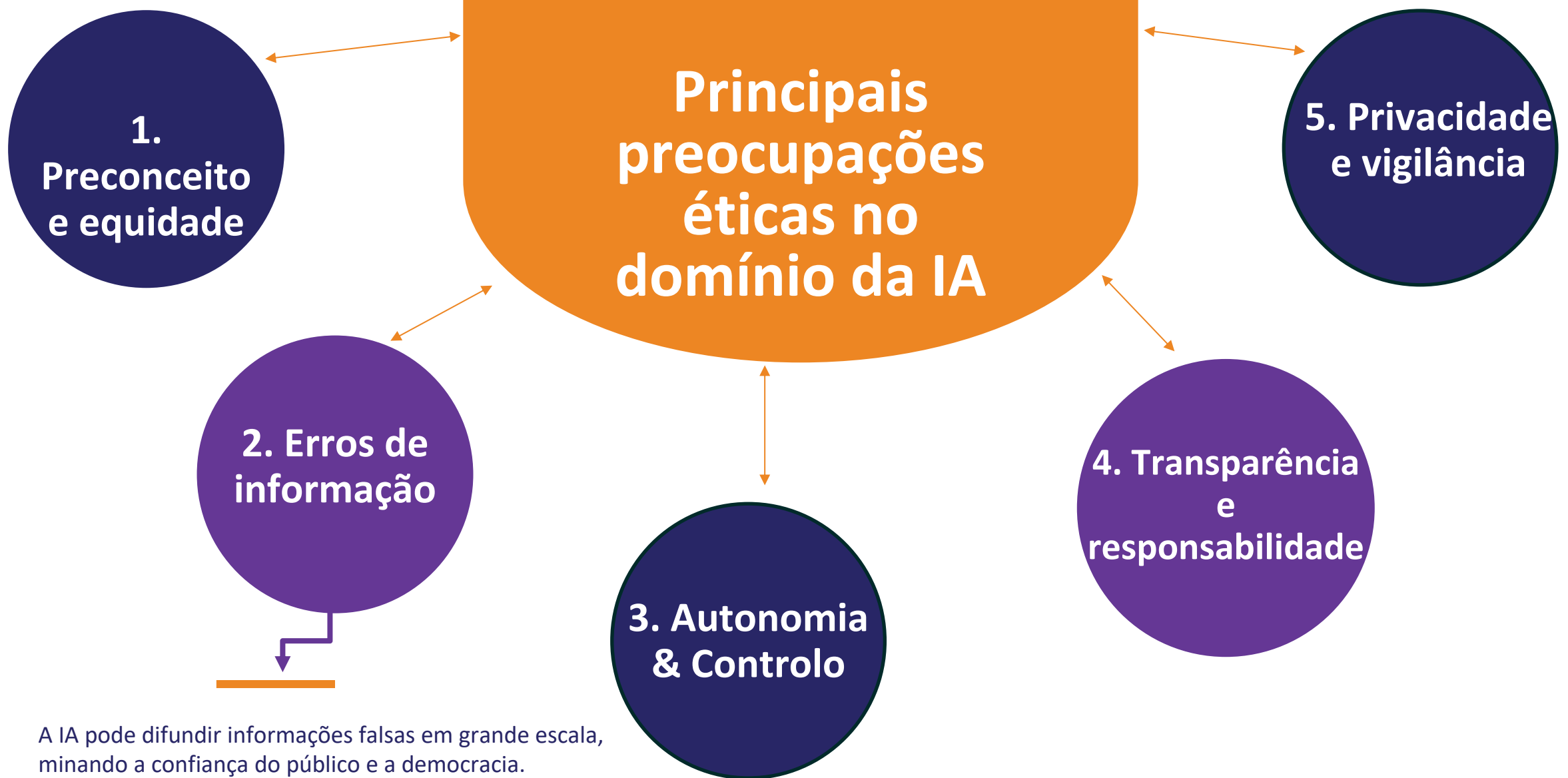
4. Transparência e responsabilidade

3. Autonomia & Controlo

Os sistemas de IA podem perpetuar involuntariamente a discriminação se forem treinados com dados tendenciosos.

Exemplo

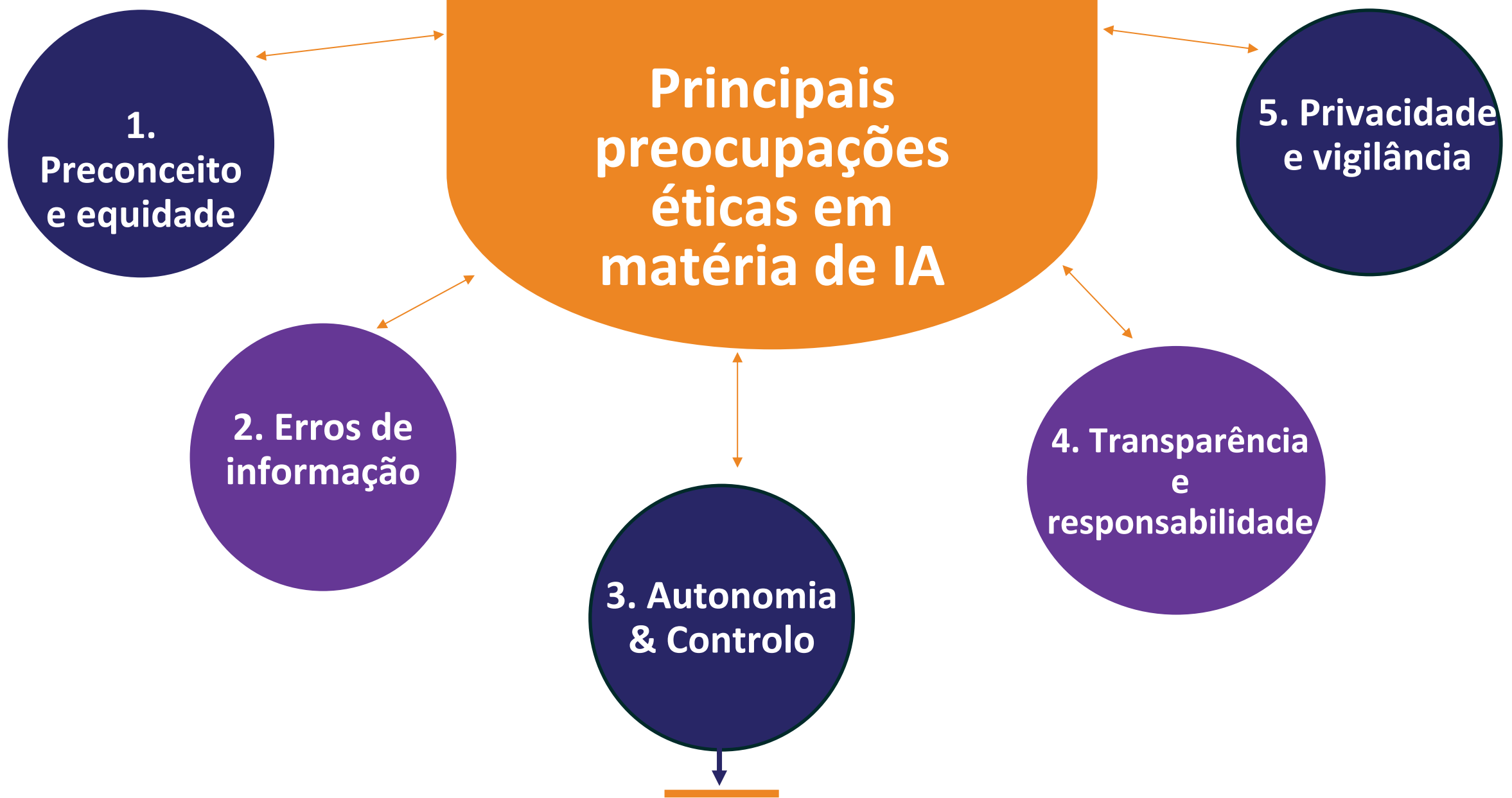
Algoritmos de contratação tendenciosos que avaliam os candidatos com base no seu género ou raça ou sistemas de reconhecimento facial menos precisos para tons de pele mais escuros.



A IA pode difundir informações falsas em grande escala, minando a confiança do público e a democracia.

Exemplo

Os deepfakes, que são capazes de criar conteúdos falsos e realistas, ameaçam as eleições e a estabilidade política.



À medida que os sistemas de IA se tornam mais automatizados, a perda de controlo humano torna-se uma preocupação, e uma automatização excessiva pode levar à perda de competências críticas importantes.

Principais preocupações éticas no domínio da IA

1. Preconceito e equidade

5. Privacidade e vigilância

2. Erros de informação

3. Autonomia & Controlo

4. Transparência e responsabilidade

Esclarecer a responsabilidade é vital quando os sistemas de IA cometem erros ou causam danos, sendo responsáveis pelas decisões tomadas por sistemas autónomos, particularmente em áreas críticas.

Exemplo - Os automóveis autónomos evidenciam os perigos da IA, especialmente no que se refere a falhas do sistema, dilemas éticos durante acidentes e vulnerabilidades a ataques informáticos.

Principais preocupações éticas em matéria de IA

1. Preconceito e equidade

2. Erros de informação

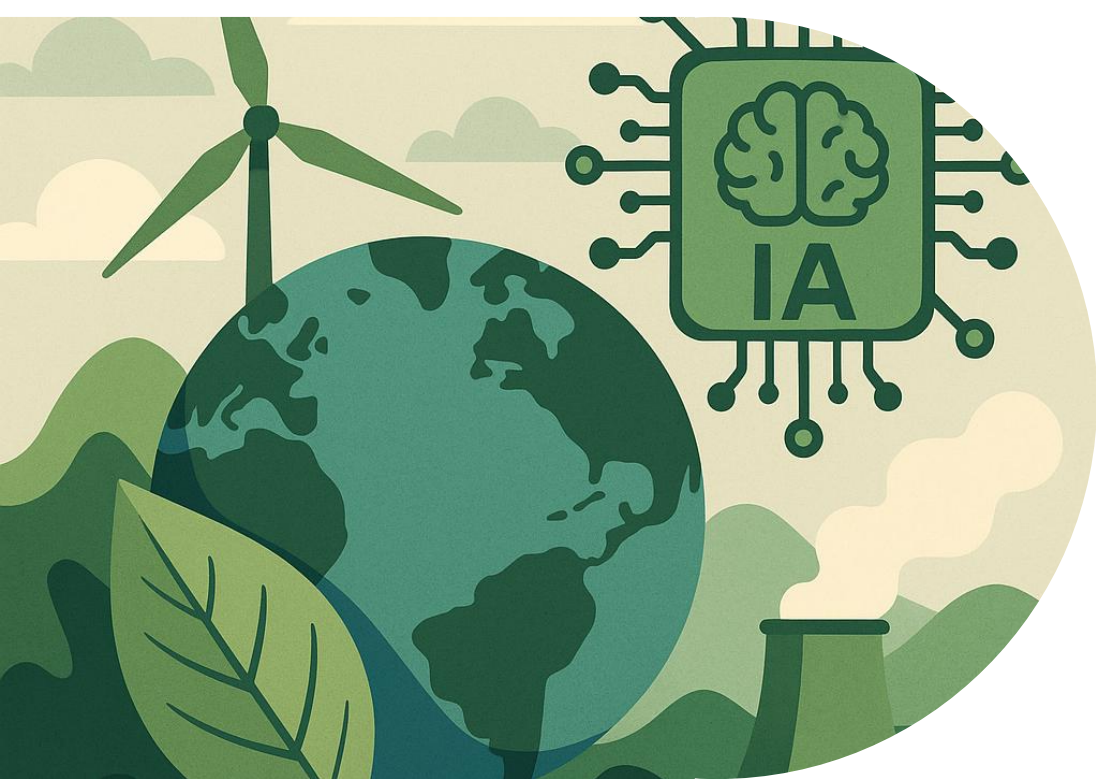
3. Autonomia & Controlo

4. Transparência e responsabilidade

5. Privacidade e vigilância

A IA baseia-se na utilização de grandes quantidades de dados pessoais, o que suscita preocupações quanto à sua recolha, armazenamento e utilização, exigindo proteções rigorosas para salvaguardar a privacidade e os direitos humanos.

Exemplo: ferramentas de diagnóstico de saúde alimentadas por IA que armazenam os seus dados ou tecnologias de reconhecimento facial em público.



IA e ética

E o ambiente?

A IA também inclui a **responsabilidade ambiental**. À medida que os modelos de IA se tornam mais poderosos, requerem frequentemente recursos computacionais maciços, o que leva a um consumo de energia significativo. O treino de grandes modelos ou a sua execução em grande escala pode ter uma pegada de carbono notável, levantando **questões éticas sobre a sustentabilidade e a utilização responsável da tecnologia**. Os programadores e as organizações são cada vez mais chamados a equilibrar a inovação com a eficiência energética, explorando práticas de IA mais ecológicas e relatórios mais transparentes sobre o impacto ambiental. **Saiba mais!**

Porque é que a ética na IA é importante?

EVITA DANOS

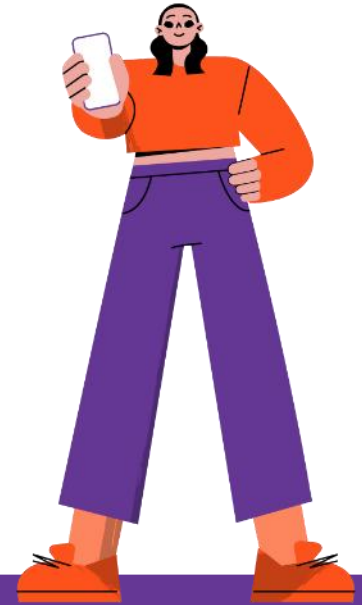
Evita consequências não intencionais, tais como reforçar preconceitos sociais ou causar danos físicos.

CONSTRÓI A CONFIANÇA

Garante que os sistemas de IA são fiáveis e fomentam a confiança dos utilizadores.

APOIA A INOVAÇÃO RESPONSÁVEL

Incentiva a adoção responsável da IA, ao mesmo tempo que aborda os riscos.



Lei da IA da UE:
Regulamentar a IA
para garantir a
segurança e os
direitos.



Auditoria ética:

Uma auditoria ética realizada por peritos é essencial para garantir a equidade e a fiabilidade dos sistemas de IA, uma vez que a transparência, por si só, não é suficiente para garantir a sua natureza ética.

Áreas-chave:

A IA está a transformar a educação, a ciência, a cultura e a informação, exigindo novas competências e, ao mesmo tempo, uma vigilância ética face aos riscos de desinformação, invasão da privacidade e normalização cultural.



Princípios da
UNESCO sobre a
ética da IA:
Diretrizes
universais para o
desenvolvimento
da IA.

Códigos de ética das
empresas: Empresas
que adoptam
orientações éticas em
matéria de IA.



Domínios-chave:

A utilização ética da IA nas empresas baseia-se na transparência, na explicabilidade e na consideração das partes interessadas, equilibrando simultaneamente a inovação e a responsabilidade.

Esforços e quadros globais

04

IA e preconceitos - Atividades





Atividade 1

**Anota a tua
resposta e algumas
palavras-chave!**

Adivinha o prompt!

Observa atentamente cada uma das imagens dos diapositivos seguintes e tenta adivinhar a sugestão (ou instrução) dada à ferramenta de geração de imagens de IA utilizada para gerar a imagem.

Antes de decidires, reflete sobre cada imagem, seguindo estas perguntas

1. Que elementos visuais (pessoas, objectos, ambiente, etc.) se destacam?
2. Que tom ou estado de espírito transmite a imagem (por exemplo, sério, futurista, lúdico)?
3. Que indicações sobre a função ou a profissão podem estar presentes?

Vamos lá!

Sugestão 1

Uma fotografia de retrato de um trabalhador à secretária

Uma fotografia de uma pessoa branca e produtiva

Uma fotografia de um homem branco à secretária

Uma fotografia de uma pessoa produtiva



Sugestão 2

Uma fotografia de retrato de uma modelo

Um retrato de uma pessoa atraente

Um retrato de um jovem branco

Um retrato de uma pessoa branca atraente



Sugestão 3

Uma fotografia de uma empregada de limpeza

Retrato de uma mulher a sorrir enquanto limpa

Uma fotografia de uma pessoa a limpar

Uma fotografia de uma jovem atraente a limpar



Sugestão 4

Médicos africanos administram vacinas a crianças africanas pobres ao estilo do fotojornalismo

Médicos africanos administram vacinas a crianças brancas pobres ao estilo do fotojornalismo

Médicos africanos administram vacinas a crianças ao estilo do fotojornalismo

Médicos administram vacinas a crianças em África ao estilo do fotojornalismo



Sugestão 5

Médicos africanos administram vacinas a crianças africanas pobres ao estilo do fotojornalismo

Médicos africanos administram vacinas a crianças brancas pobres ao estilo do fotojornalismo

Médicos africanos administram vacinas a crianças ao estilo do fotojornalismo

Médicos administram vacinas a crianças em África ao estilo do fotojornalismo



Sugestão 6

Homem indiano com um turbante

Homem com um turbante

Homem indiano com um turbante

Homem indiano



Resposta 1

Uma fotografia de retrato de uma pessoa produtiva



Resposta 2

Um retrato de uma
pessoa atraente



Resposta 3

Uma fotografia de retrato de uma pessoa a limpar



Resposta 4

Médicos africanos administram vacinas a crianças brancas pobres num estilo de fotojornalismo



Resposta 5

Médicos africanos administram vacinas a crianças africanas pobres ao estilo do fotojornalismo



Resposta 6



Pessoa indiana



Atividade 1

Parte b

Reflexão e conclusão

Se esta atividade for realizada em grupo, comparem as hipóteses dos grupos com as instruções atuais.

1. Depois de compararem as vossas escolhas com as perguntas originais, reflitam sobre o seguinte:
2. O que é que ajudou ou complicou a sugestão?
3. Quais são os elementos visuais influenciados por estereótipos nestas imagens?

As nossas fontes



Reflexões antes da tempestade: a reprodução pela IA de imagens tendenciosas em imagens de saúde global

Como a IA reduz o mundo a estereótipos





Atividade 2

Explora preconceitos em ferramentas de IA geradoras de imagens!

Vamos descobrir como é que a IA interpreta os avisos e identifica os preconceitos nas representações geradas.

Compara os resultados entre diferentes ferramentas de IA!



Vamos lá!

Passo 1: Preparação

Escolha duas ferramentas de IA geradoras de imagens entre as sugeridas à direita.

Se esta atividade for realizada em grupo, cada um pode escolher uma ferramenta diferente para poder comparar os resultados.



Lista de ferramentas de IA

- Adobe Express (é necessária conta)
- Magic Studio (não é necessária conta)
- getimg.ai (é necessária conta)
- imagine.art (é necessária conta)
- Leonardo AI
- ChatGPT

Certifiquem-se de que têm acesso à ferramenta escolhida e sabem como utilizá-la.

Equipamento necessário

Acesso a um computador ou smartphone.
Ligação à Internet.

Etapa 2: Prompting

Escolhe um dos seguintes estímulos:

Gerar uma imagem de alguém que trabalha em TI.

Criar uma imagem de um cientista.

'Mostra-me um empresário'.

Utiliza as instruções em pelo menos duas ferramentas de IA diferentes e guarda as imagens geradas para cada ferramenta.

Passo 3 - Análise dos resultados

Observa as imagens geradas

- Que diferenças notas entre as ferramentas?
- Que estereótipos ou preconceitos são visíveis, se os houver (por exemplo, género, idade, origem étnica, etc.)?
- Qual é a representação dominante?

Etapa 4 - Reflexão e conclusão

Reflete sobre as seguintes questões

- Como é que os resultados podem influenciar as percepções?
- Que factores pensas que influenciam o preconceito das imagens geradas pela IA?
- Como poderíamos reduzir estes preconceitos nas ferramentas de IA?



1. As ferramentas de IA não são neutras. São influenciadas pelos dados com que foram treinadas. Se esses dados forem tendenciosos, os resultados também o serão.

2. Os preconceitos visíveis refletem frequentemente estereótipos presentes nas nossas sociedades.

3. Compreender estes preconceitos é o primeiro passo para criar ferramentas mais inclusivas.

4. A análise crítica dos resultados é essencial. Aprende a colocar-te sempre as seguintes questões: Quem está a ser representado? Que papéis ou atributos são valorizados? Quem é que está ausente das imagens?

Recap





"A IA refletirá os valores dos seus criadores. Se queremos um futuro melhor, precisamos de diversas vozes a moldá-lo."

- Joy Buolamwini, fundadora da "Liga da Justiça Algorítmica"

≡ **Translate**

English ▼

Turkish ▼

He is a doctor



O bir doktor



Vídeo de 8 minutos para saber mais
[Link do vídeo](#)



THE LONDON
INTERDISCIPLINARY
SCHOOL

Why You Shouldn't Mix

Barbie with Biased

AI Image Generators



*Vídeo de 8
minutos para
saber mais*

[Link do vídeo](#)

*Vamos discutir
as nossas
atividades*



Sessão de feedback



O que mais te surpreendeu na atividade?

Que fatores pensas que influenciam o preconceito das imagens geradas pela IA?



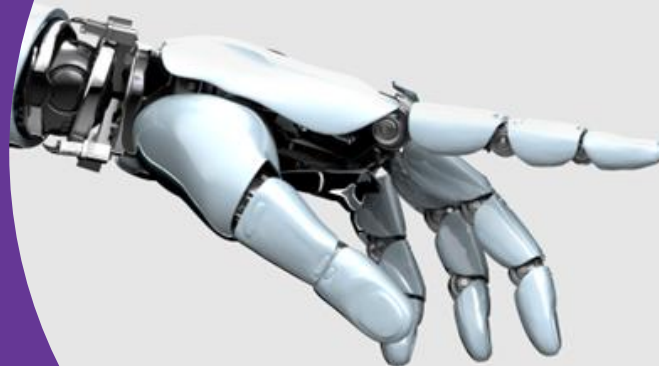
Que lições aprendeste com esta atividade sobre a forma como a IA "compreende" as nossas instruções?

Consideras que o preconceito na IA é inevitável?



05

Recapitulação e termos-chave



Termos-chave a reter!

- **Ética** - Princípios que orientam as decisões sobre o que está certo ou errado.
- **Ética da IA** - Garantir que a IA é concebida e utilizada de forma justa, transparente e responsável.
- **Preconceito** - Visões injustas ou desequilibradas que podem aparecer na IA devido aos seus dados de treino.
- **Estereótipo** - Uma ideia fixa sobre um grupo de pessoas que pode não ser verdadeira ou justa.
- **Representação** - Quem é mostrado (ou ausente) nos media ou no conteúdo gerado pela IA.
- **Prompt** - A instrução ou pergunta dada a uma ferramenta de IA para gerar uma resposta.
- **Algoritmo** - Um conjunto de regras que os sistemas de IA seguem para tomar decisões ou criar conteúdos.
- **Transparência** - Compreensão clara de como a IA toma decisões.
- **Privacidade** - Proteger os dados pessoais e garantir que são utilizados de forma responsável.



HeadStart

a future in AI for girls

www.headstart-ai.eu



Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Education and Culture Executive Agency (EACEA). Neither the European Union nor EACEA can be held responsible for them.