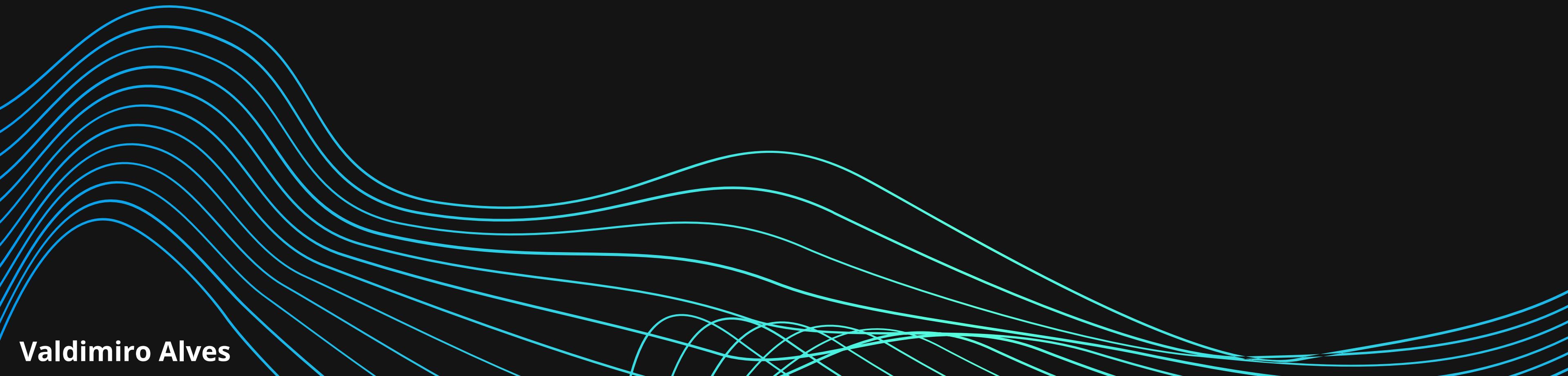


Deepfakes: A Realidade por Trás da Ilusão Digital

- Entendendo, Identificando e Navegando na Nova Era da Mídia



UNIVERSIDADE
FEDERAL DO CARIRI



Sobre o que vamos falar?

- Inteligência Artificial (IA)
- LLMs
- Geração de Vídeo por IA
- Deepfakes
- Ferramentas
- Usos Positivos e Criativos
- Riscos e Desafios
- Casos Reais e Consequências
- A Solução de Longo Prazo

Inteligência Artificial (IA) em poucas palavras

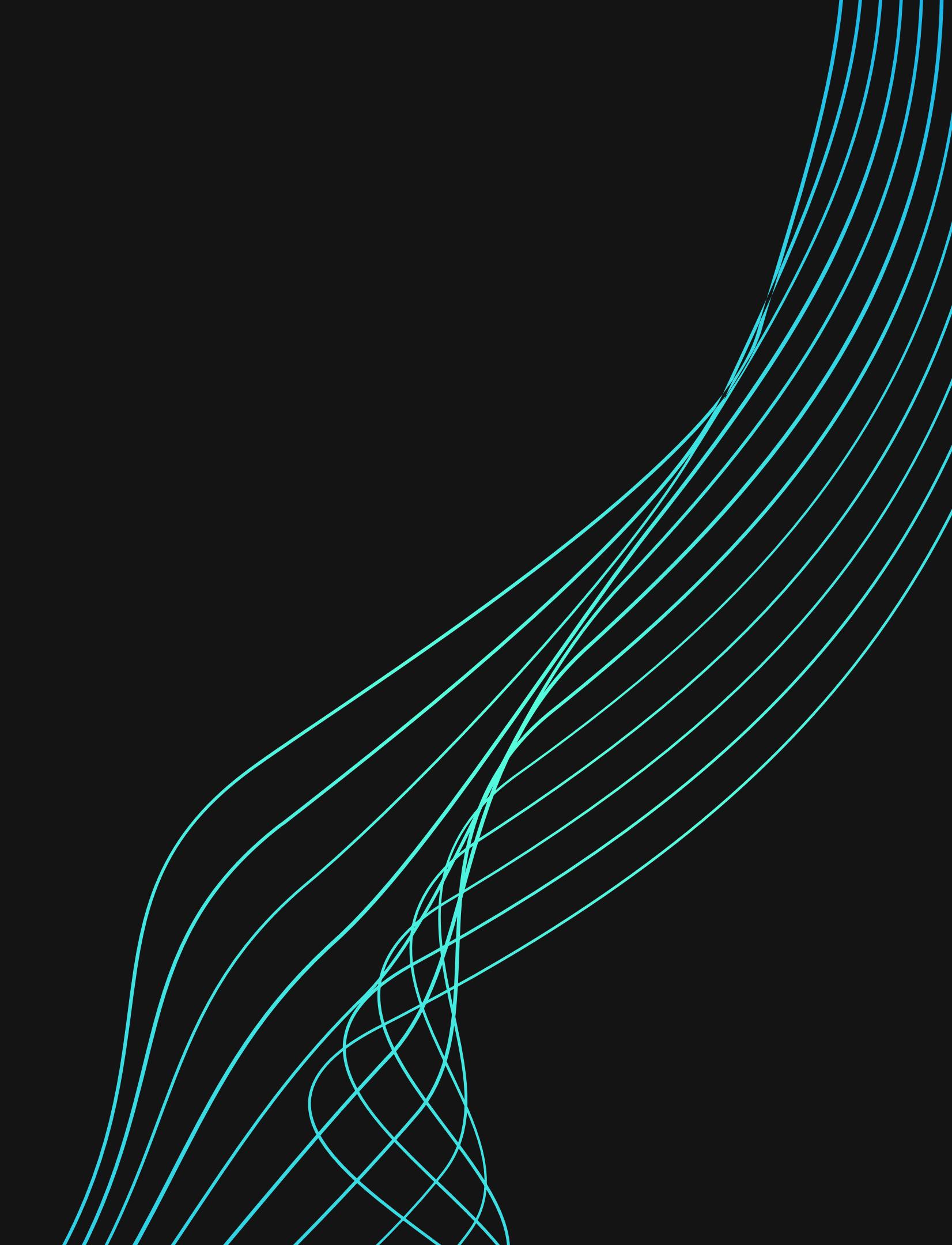


Inteligência Artificial (IA) em poucas palavras

- **Conceito Central:**

"Pense na IA como o campo de estudo que busca ensinar as máquinas a **aprender, raciocinar e tomar decisões** de forma semelhante aos humanos."

LLMs (Large Language Models)



LLMs (Large Language Models)

- **O que são?**
- São um tipo específico de IA treinada com uma quantidade massiva de texto da internet.

LLMs (Large Language Models)

- **O que são?**
- São um tipo específico de IA treinada com uma quantidade massiva de texto da internet.
- São capazes de entender o contexto, a gramática, o estilo e as nuances da linguagem humana.

LLMs (Large Language Models)

- **O que são?**
- São um tipo específico de IA treinada com uma quantidade massiva de texto da internet.
- São capazes de entender o contexto, a gramática, o estilo e as nuances da linguagem humana.
- **Exemplo Prático:** o ChatGPT e o Gemini são atualmente dois dos LLMs mais usados.

Se a IA pode escrever, ela pode "desenhar"?



Se a IA pode escrever, ela pode "desenhar"?

- **Conceito da Transição:**
 - Pegaram a mesma ideia dos LLMs e aplicaram a imagens.

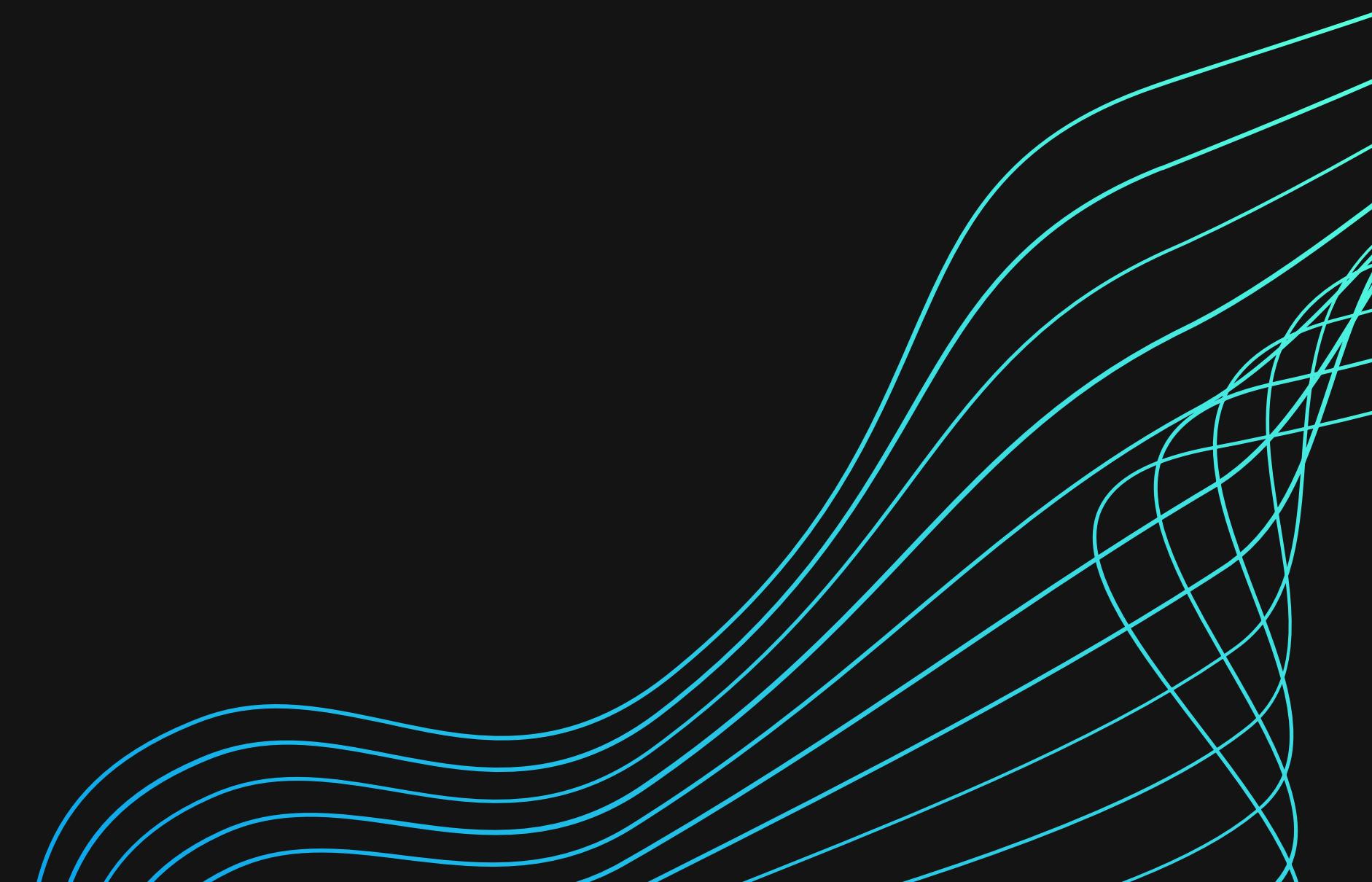
Se a IA pode escrever, ela pode "desenhar"?

- **Conceito da Transição:**
 - Pegaram a mesma ideia dos LLMs e aplicaram a imagens.
 - Se um LLM aprende a 'linguagem das palavras', outros modelos podem aprender a 'linguagem das imagens' e dos vídeos.

Se a IA pode escrever, ela pode "desenhar"?

- **Conceito da Transição:**
 - Pegaram a mesma ideia dos LLMs e aplicaram a imagens.
 - Se um LLM aprende a 'linguagem das palavras', outros modelos podem aprender a 'linguagem das imagens' e dos vídeos.
 - **IA Generativa** - Uma categoria de IA que não apenas analisa, mas CRIA conteúdo novo: texto, imagens, música e vídeos.

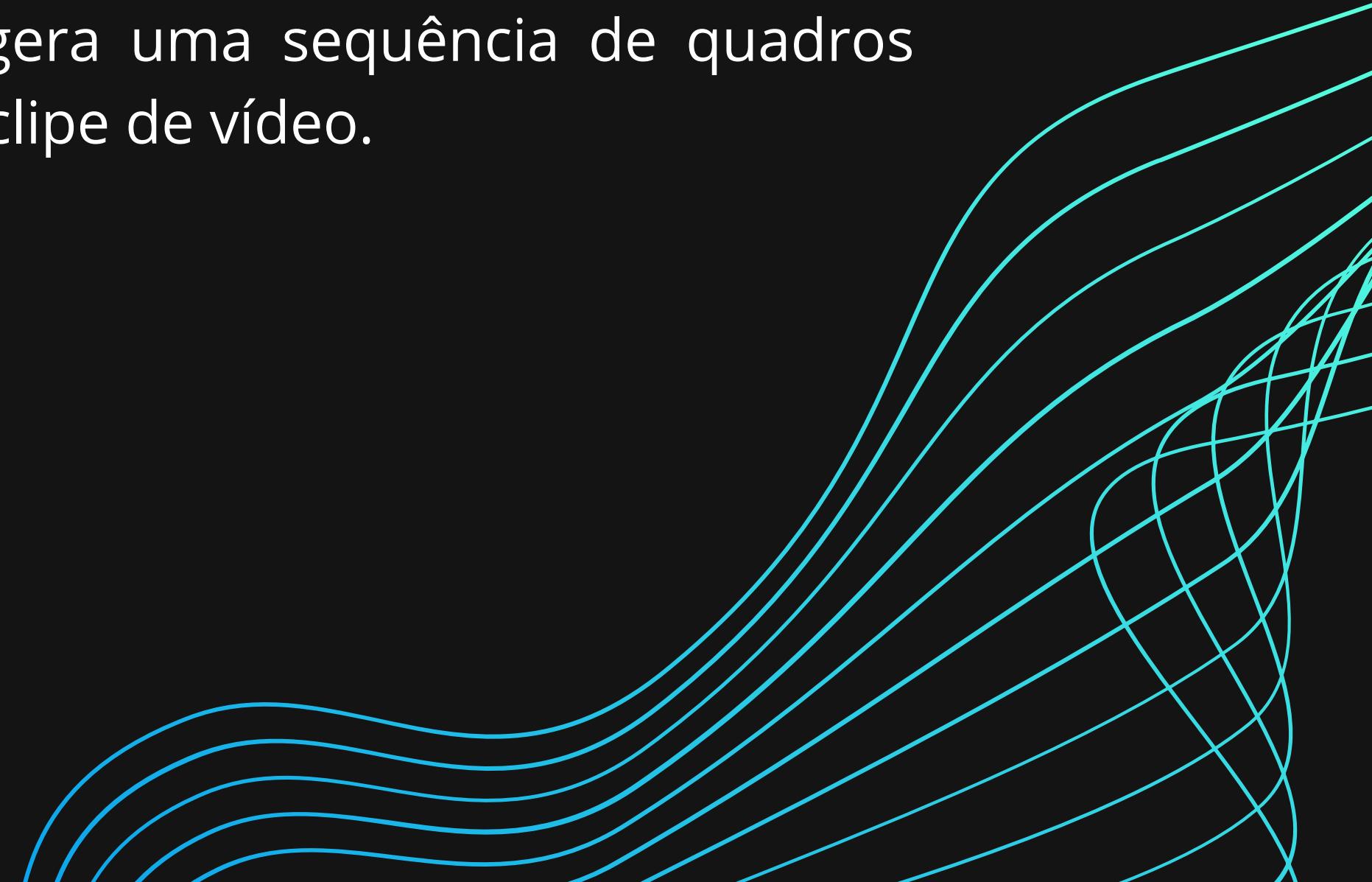
Geração de Vídeo por IA



Geração de Vídeo por IA

- **Como Funciona?**

O modelo de IA analisa uma descrição de texto (o prompt) para entender objetos, ações e o estilo visual. Depois, ele gera uma sequência de quadros (frames) consistentes que, juntos, formam um clipe de vídeo.



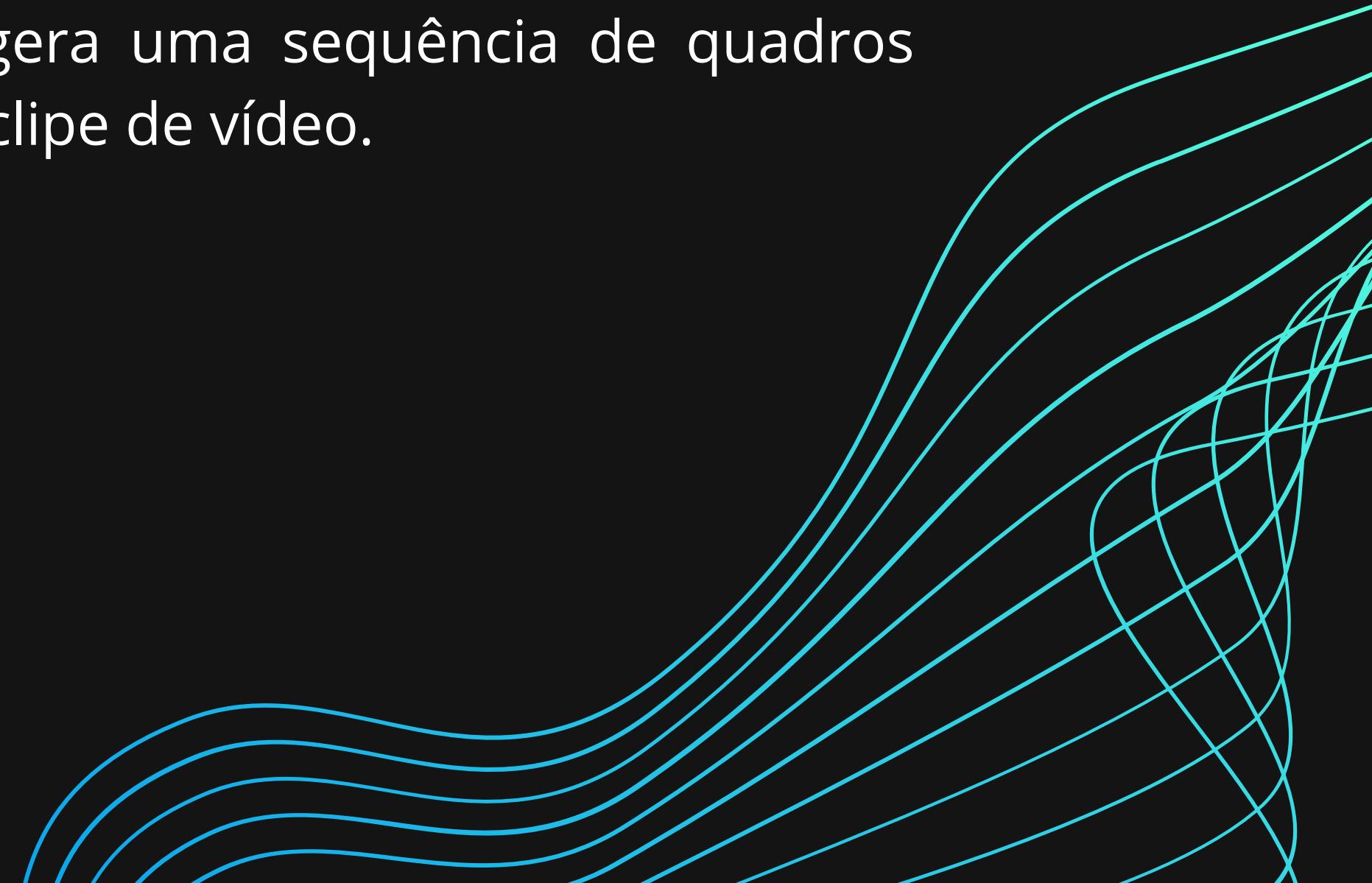
Geração de Vídeo por IA

- **Como Funciona?**

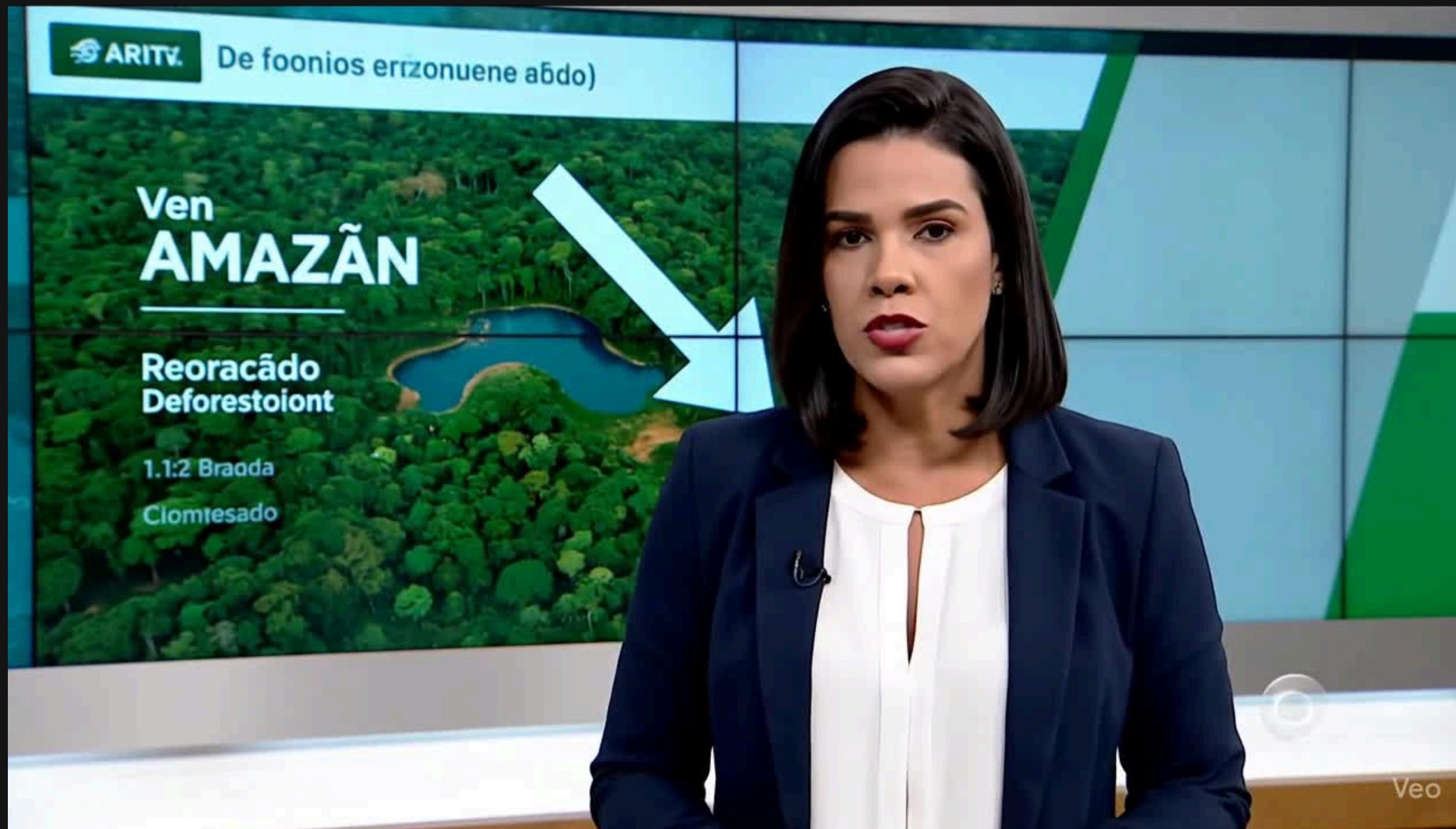
O modelo de IA analisa uma descrição de texto (o prompt) para entender objetos, ações e o estilo visual. Depois, ele gera uma sequência de quadros (frames) consistentes que, juntos, formam um clipe de vídeo.

- **O que ela pode criar?**

- Cenas Fantásticas;
- Simulações Realistas;
- Protótipos Rápidos;
- Arte Generativa.



prompt usado: “Gere um vídeo de uma reportagem falando sobre a diminuição do desmatamento.”



VOCÊ CONFIARIA NO QUE ACABOU DE VER?

Outro exemplo:



VOCÊ CONFIARIA NO QUE ACABOU DE VER?



O que são Deepfakes?



O que são Deepfakes?

DEEP LEARNING (APRENDIZAGEM PROFUNDA)

Um tipo avançado de Inteligência Artificial que imita o cérebro humano para aprender com grandes volumes de dados.



O que são Deepfakes?

DEEP LEARNING (APRENDIZAGEM PROFUNDA)

Um tipo avançado de Inteligência Artificial que imita o cérebro humano para aprender com grandes volumes de dados.

FAKE (FALSO):

Conteúdo sintético, manipulado.



O que são Deepfakes?

DEEP LEARNING (APRENDIZAGEM PROFUNDA)

Um tipo avançado de Inteligência Artificial que imita o cérebro humano para aprender com grandes volumes de dados.

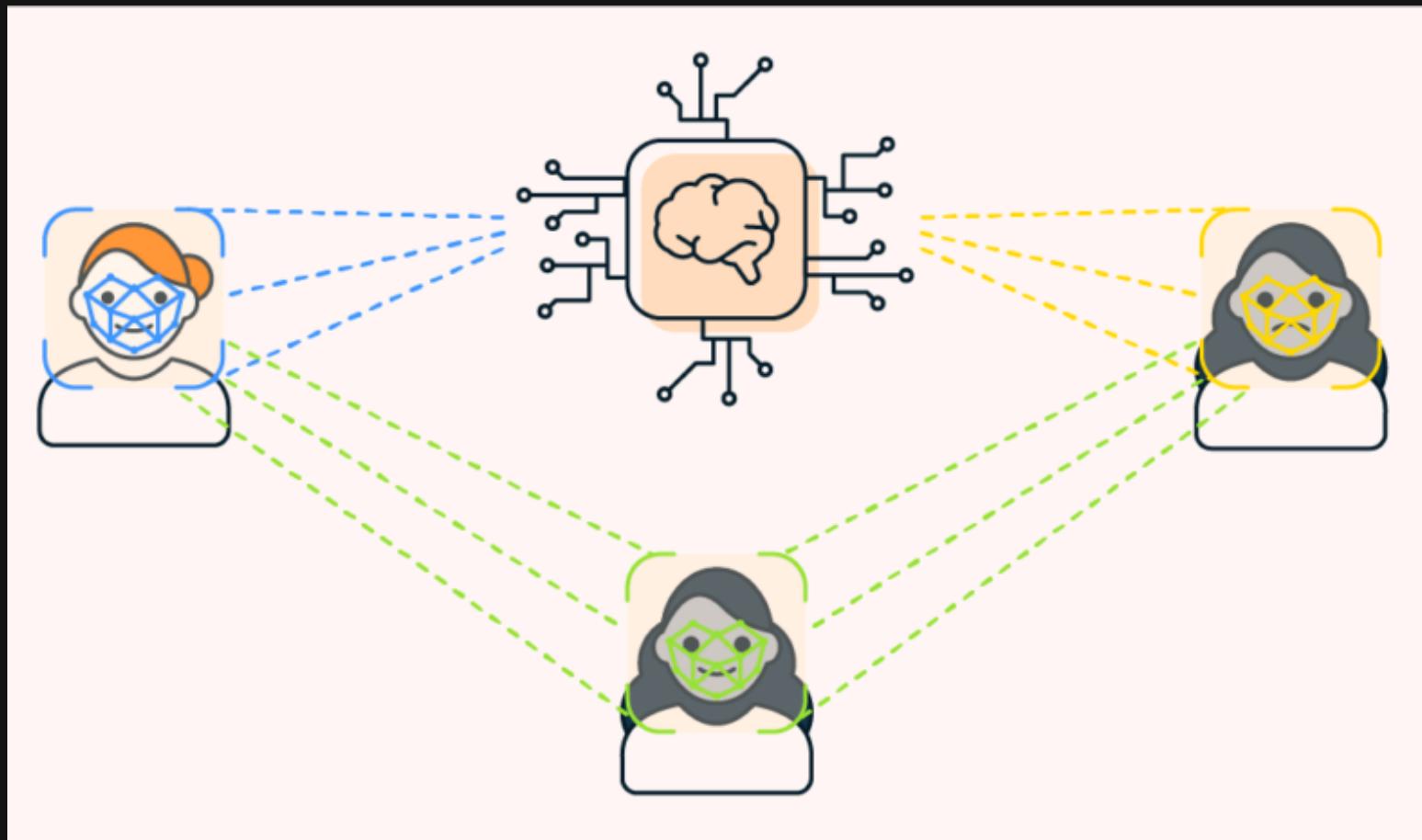
FAKE (FALSO):

Conteúdo sintético, manipulado.

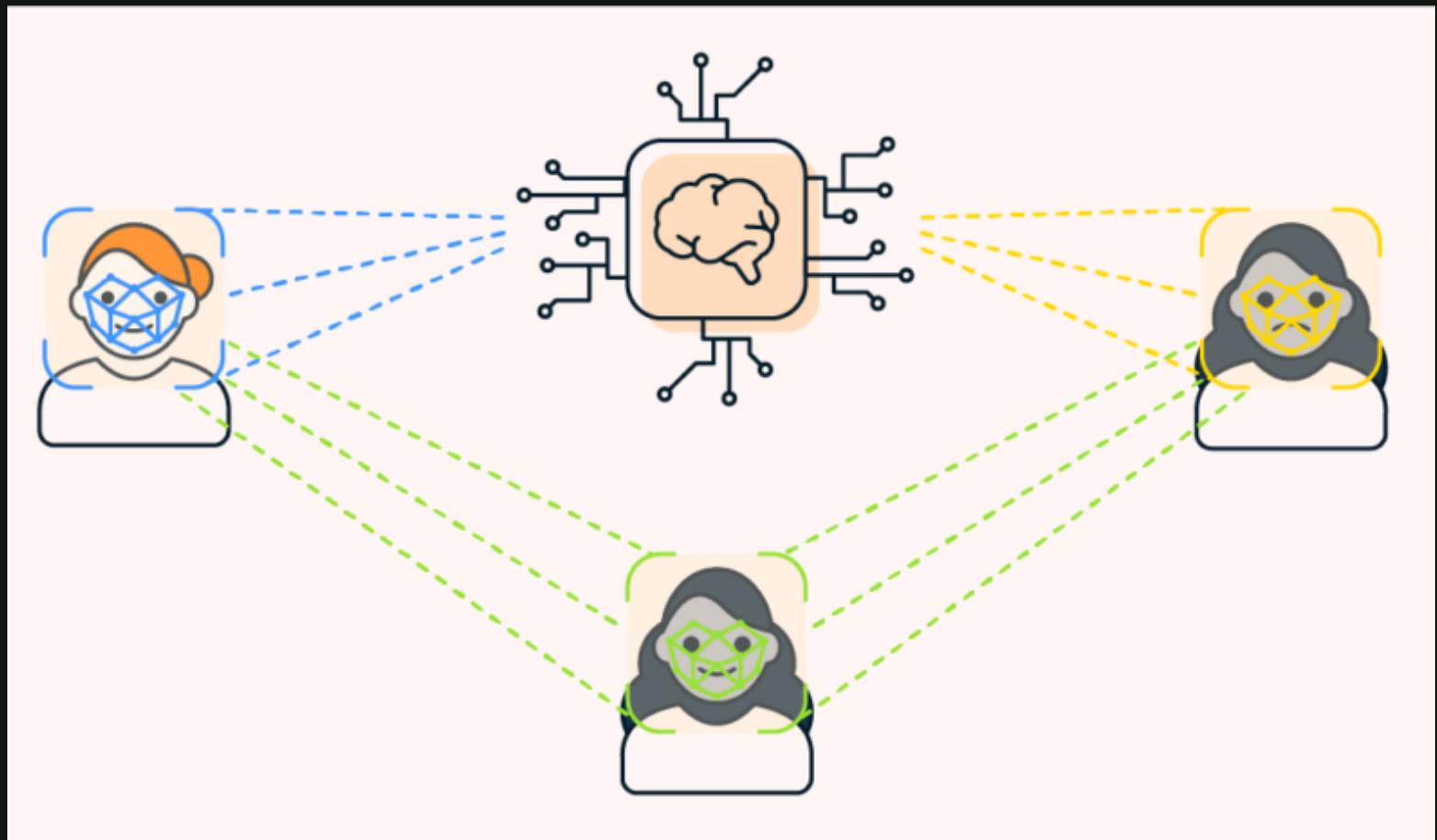
DEFINIÇÃO:

Mídia sintética (vídeo ou áudio) onde a imagem ou voz de uma pessoa é substituída pela de outra, de forma hiper-realista.

Como Funciona?



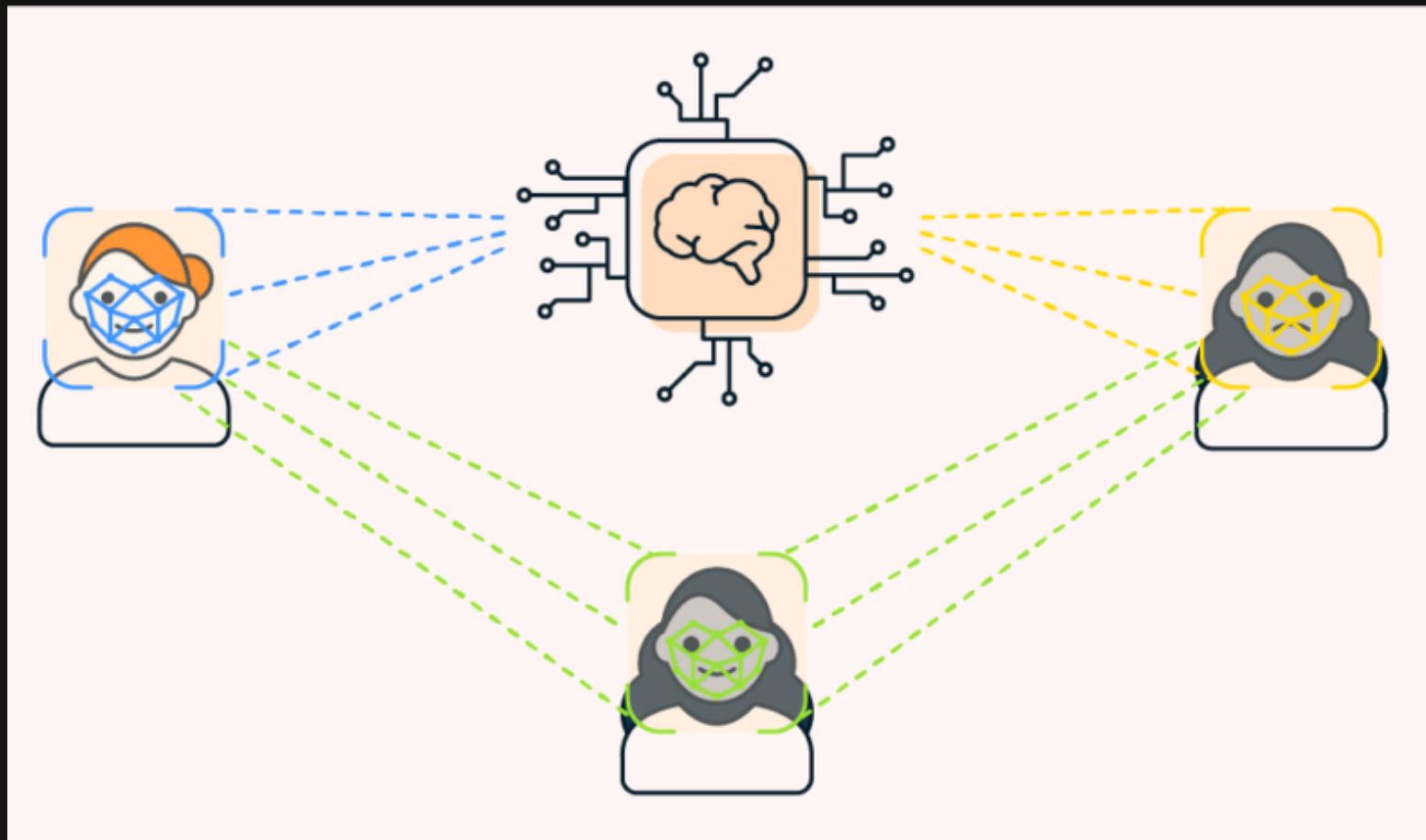
Como Funciona?



COLETA DE DADOS:

O algoritmo precisa de centenas ou milhares de imagens e vídeos da pessoa-alvo (quem será "imitada") e da pessoa-fonte (cujo rosto será usado no vídeo original).

Como Funciona?



COLETA DE DADOS:

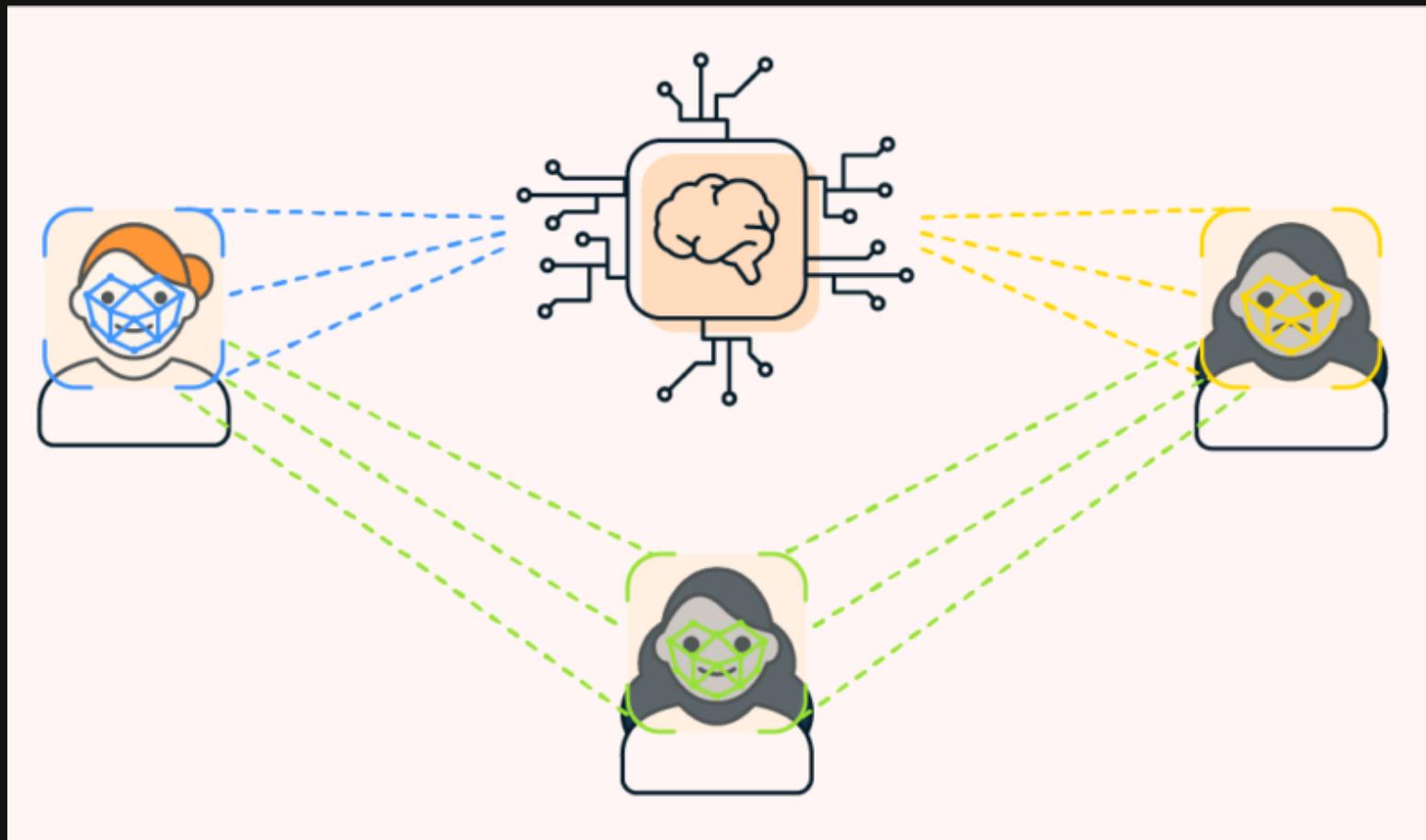
O algoritmo precisa de centenas ou milhares de imagens e vídeos da pessoa-alvo (quem será "imitada") e da pessoa-fonte (cujo rosto será usado no vídeo original).

TREINAMENTO DA IA:

Utilizando Redes Adversariais Generativas (GANs), dois algoritmos competem entre si:

- O Gerador cria as imagens falsas.
- O Discriminador tenta identificar se a imagem é real ou falsa.
- Eles "treinam" um ao outro até que o Gerador consiga enganar o Discriminador.

Como Funciona?



COLETA DE DADOS:

O algoritmo precisa de centenas ou milhares de imagens e vídeos da pessoa-alvo (quem será "imitada") e da pessoa-fonte (cujo rosto será usado no vídeo original).

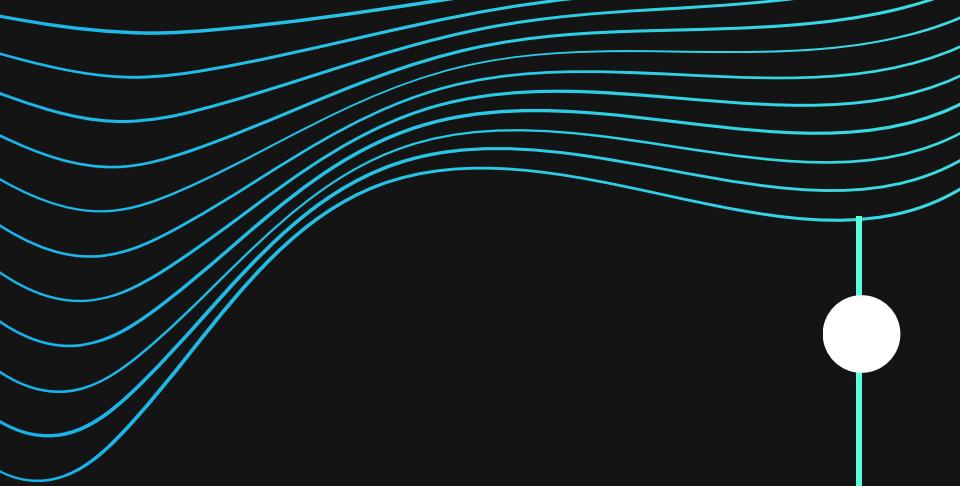
TREINAMENTO DA IA:

Utilizando Redes Adversariais Generativas (GANs), dois algoritmos competem entre si:

- O Gerador cria as imagens falsas.
- O Discriminador tenta identificar se a imagem é real ou falsa.
- Eles "treinam" um ao outro até que o Gerador consiga enganar o Discriminador.

GERAÇÃO E REFINAMENTO:

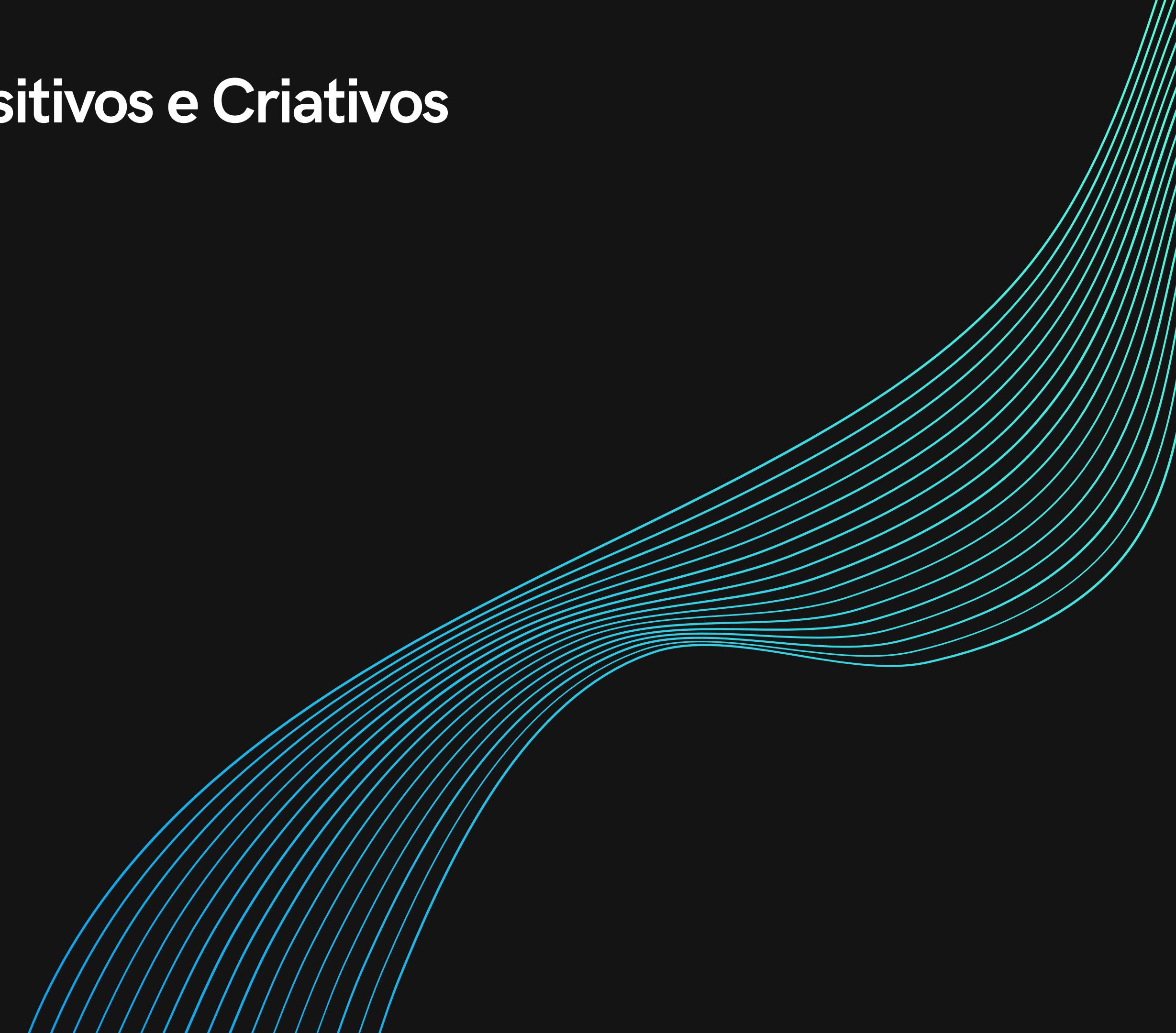
O modelo treinado é aplicado ao vídeo original, substituindo o rosto quadro a quadro.



Ferramentas

- [FACESHOW](#)
- [DEEPFACELAB](#)
- [DEEPFACELIVE](#)

Usos Positivos e Criativos



Usos Positivos e Criativos

Entretenimento:

- Rejuvenescimento de atores em filmes.
- Dublagem de filmes em outras línguas com sincronia labial perfeita.

Usos Positivos e Criativos

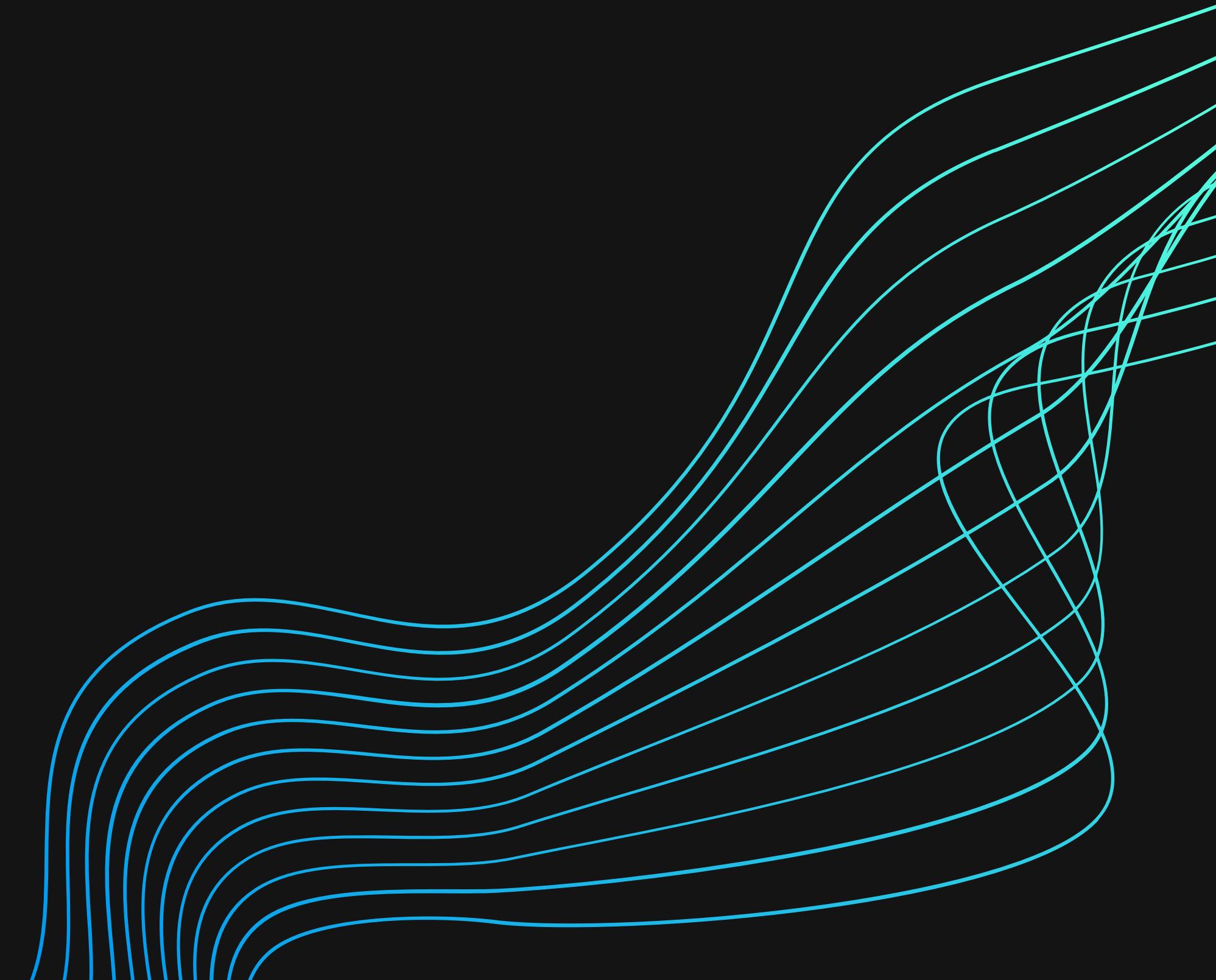
Entretenimento:

- Rejuvenescimento de atores em filmes.
- Dublagem de filmes em outras línguas com sincronia labial perfeita.

Educação e Cultura:

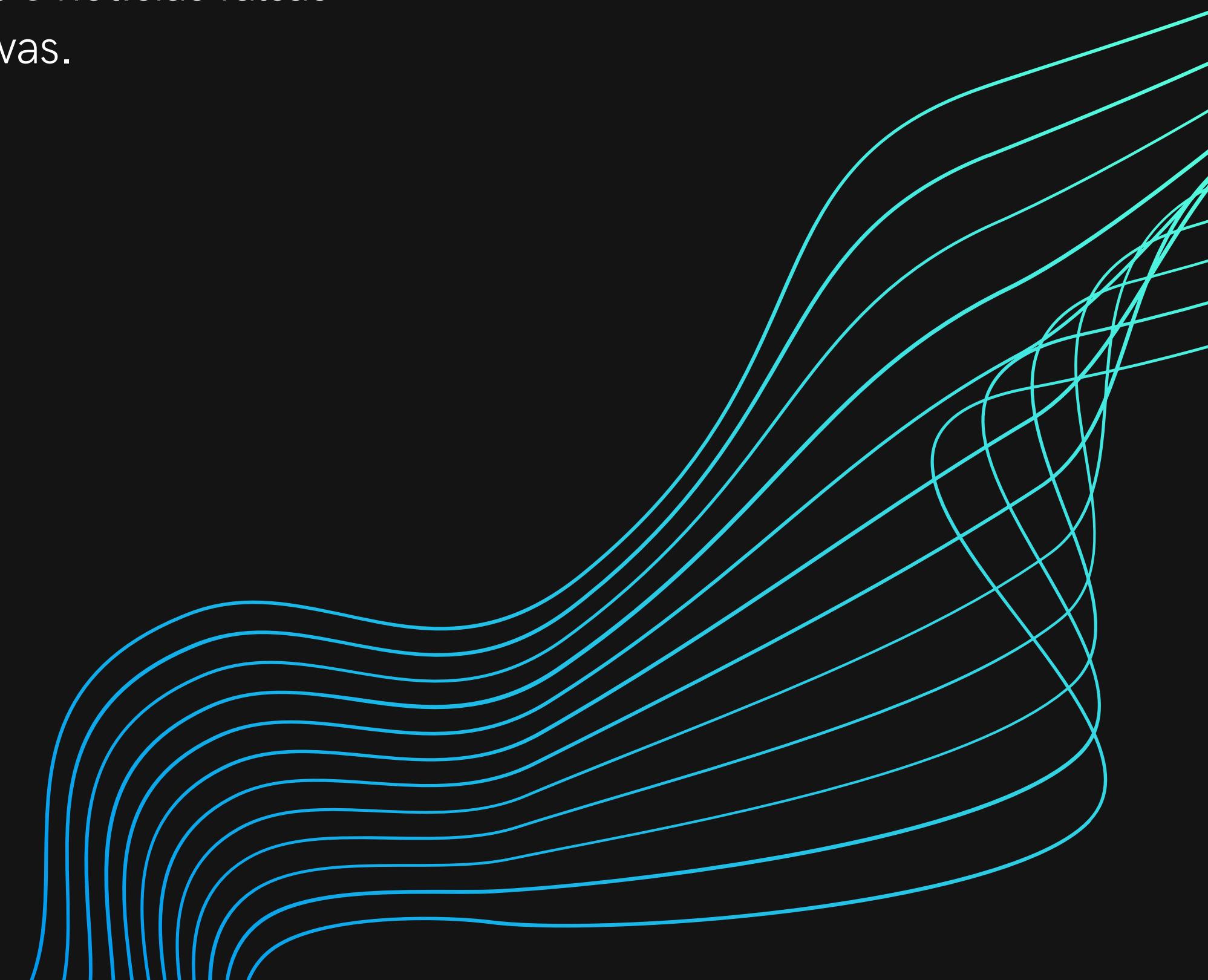
- "Trazer à vida" figuras históricas em museus e aulas.

Riscos e Desafios



Riscos e Desafios

- Desinformação e Fake News: Espalhar boatos e notícias falsas com vídeos manipulados para parecerem provas.

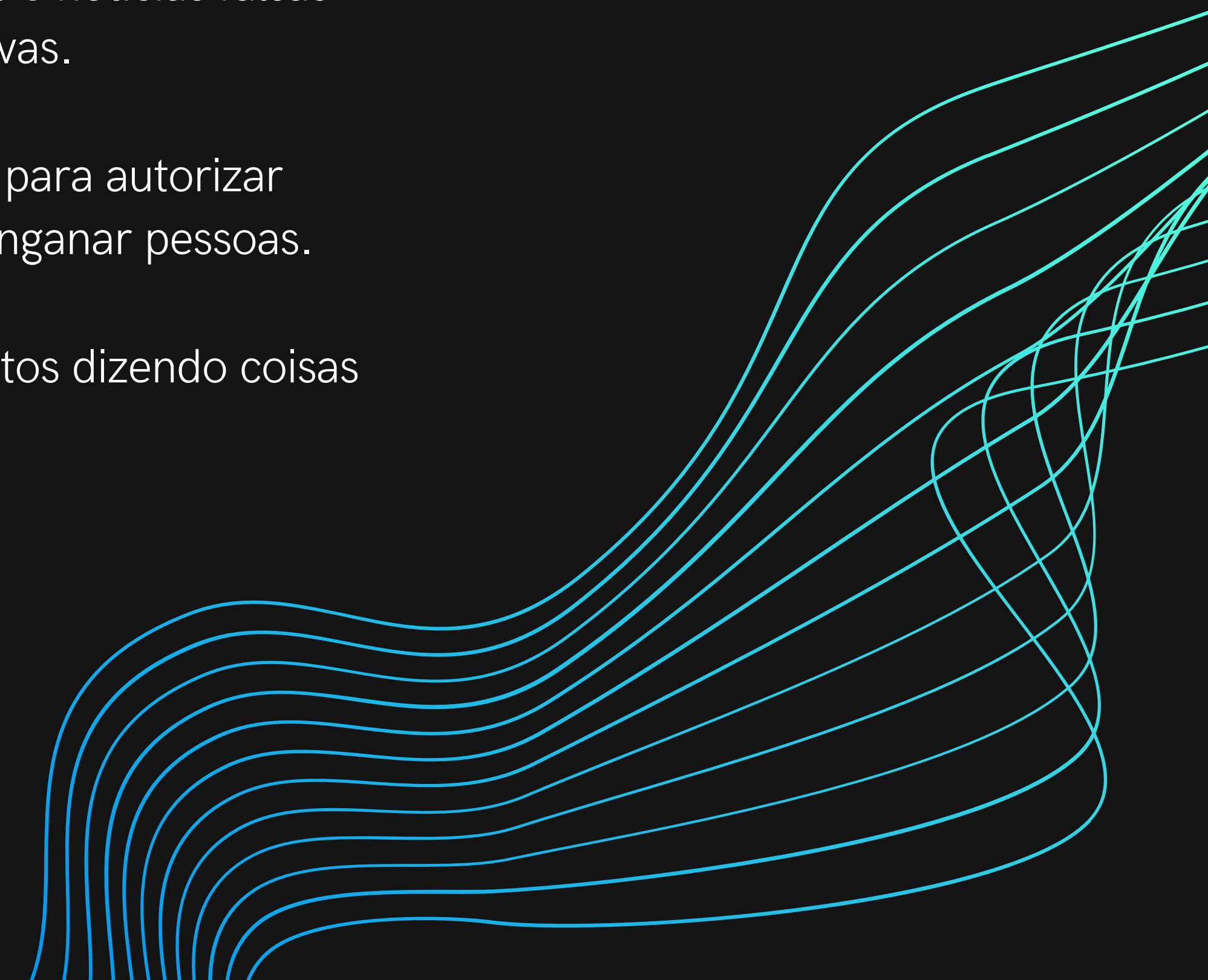


Riscos e Desafios

- Desinformação e Fake News: Espalhar boatos e notícias falsas com vídeos manipulados para parecerem provas.
- Fraudes e Golpes: "Sequestro" de identidade para autorizar transações financeiras (deepfake de voz) ou enganar pessoas.

Riscos e Desafios

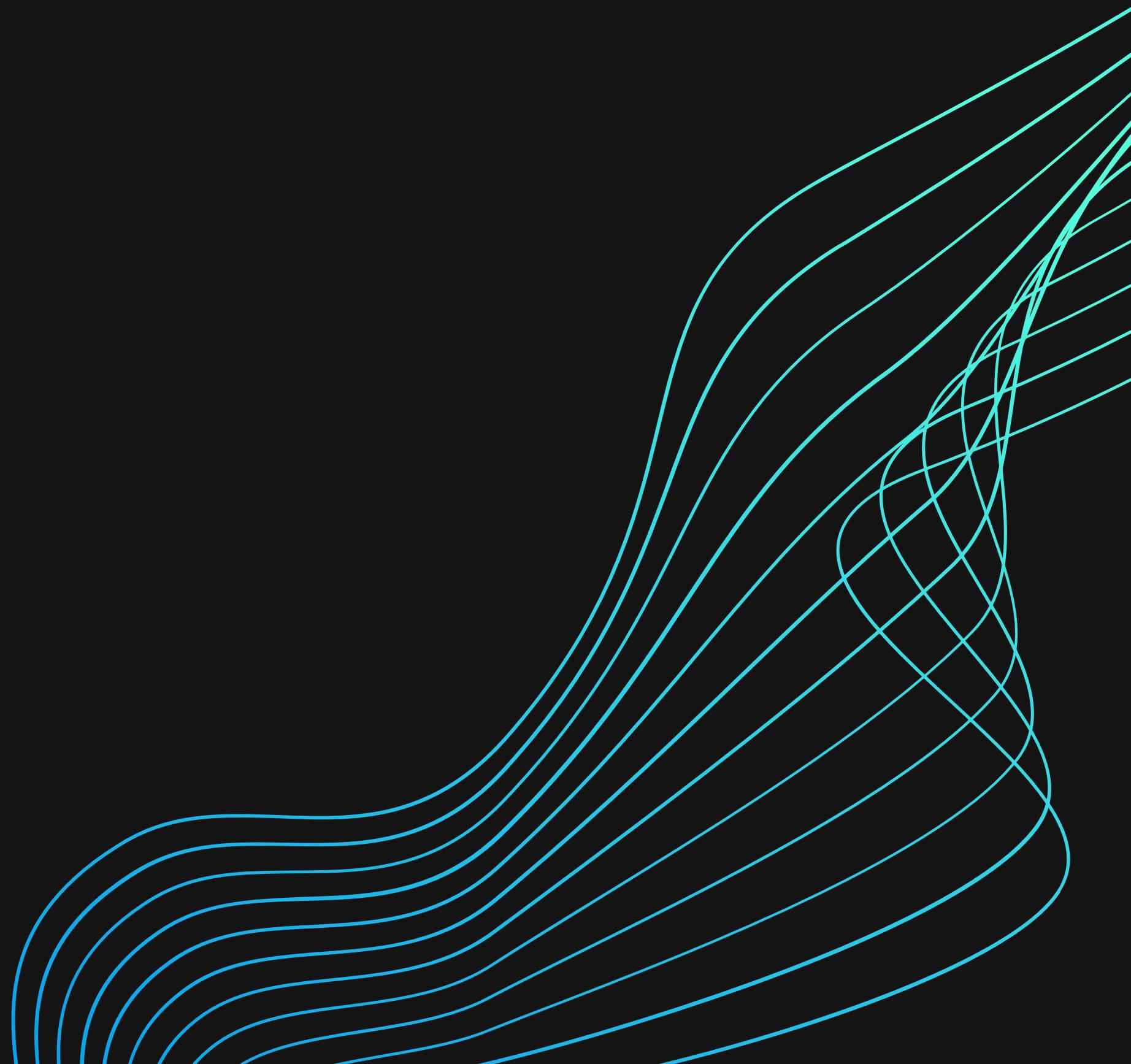
- Desinformação e Fake News: Espalhar boatos e notícias falsas com vídeos manipulados para parecerem provas.
- Fraudes e Golpes: "Sequestro" de identidade para autorizar transações financeiras (deepfake de voz) ou enganar pessoas.
- Manipulação Política: Criar vídeos de candidatos dizendo coisas que nunca disseram para influenciar eleições.



Riscos e Desafios

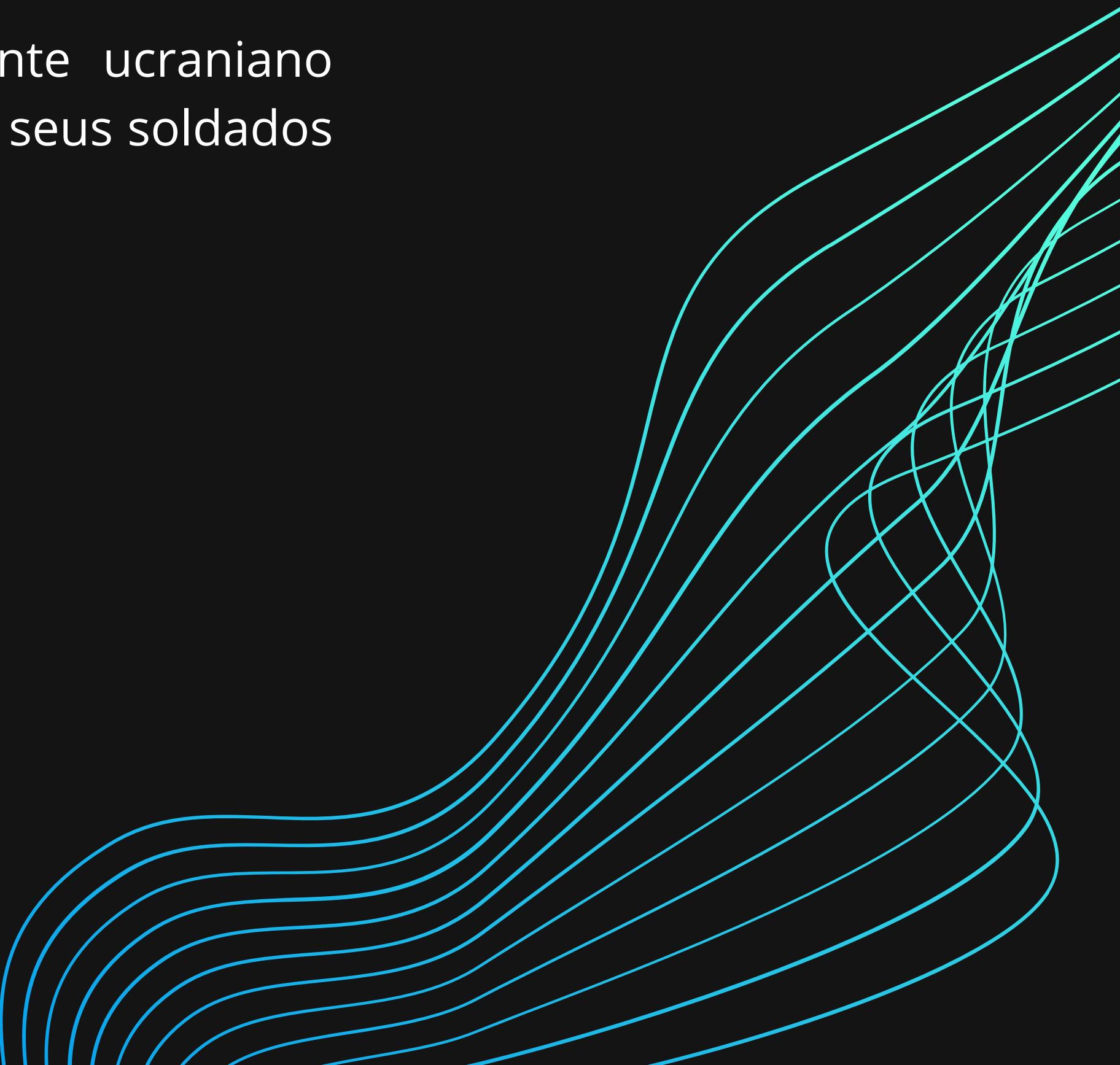
- Desinformação e Fake News: Espalhar boatos e notícias falsas com vídeos manipulados para parecerem provas.
- Fraudes e Golpes: "Sequestro" de identidade para autorizar transações financeiras (deepfake de voz) ou enganar pessoas.
- Manipulação Política: Criar vídeos de candidatos dizendo coisas que nunca disseram para influenciar eleições.
- Cyberbullying e Abuso: Criação de conteúdo pornográfico não consensual ou humilhante para atacar indivíduos.

Casos Reais e Consequências



Casos Reais e Consequências

- **Exemplo Político:** O deepfake do presidente ucraniano Volodymyr Zelenskyy pedindo a rendição de seus soldados no início da guerra.



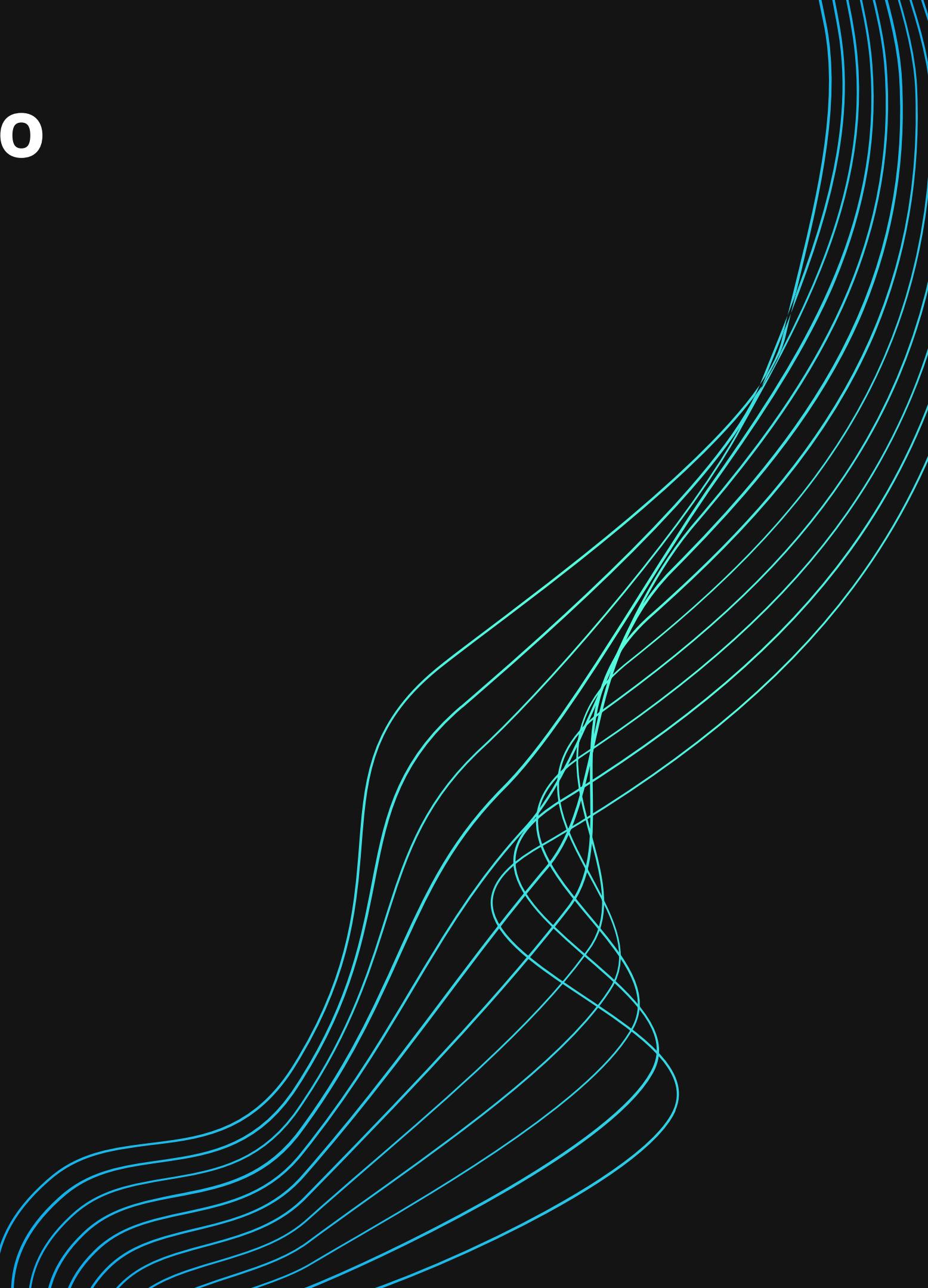
Casos Reais e Consequências

- **Exemplo Político:** O deepfake do presidente ucraniano Volodymyr Zelenskyy pedindo a rendição de seus soldados no início da guerra.
- **Exemplo de Fraude:** CEO de uma empresa do Reino Unido transferiu \$243.000 após receber uma chamada de um deepfake de voz do seu chefe.

Casos Reais e Consequências

- **Exemplo Político:** O deepfake do presidente ucraniano Volodymyr Zelenskyy pedindo a rendição de seus soldados no início da guerra.
- **Exemplo de Fraude:** CEO de uma empresa do Reino Unido transferiu \$243.000 após receber uma chamada de um deepfake de voz do seu chefe.
- **Impacto Social:** O "Dividendo do Mentiroso" - a existência de deepfakes faz com que pessoas desacreditem de vídeos reais, alegando que são falsos para escapar de responsabilidades.

A Solução de Longo Prazo



A Solução de Longo Prazo

- **Alfabetização Midiática:** Ensinar, desde a escola, como consumir, analisar e criar mídia de forma responsável.

A Solução de Longo Prazo

- **Alfabetização Midiática:** Ensinar, desde a escola, como consumir, analisar e criar mídia de forma responsável.
- **Pensamento Crítico:** A habilidade de questionar informações, identificar vieses e buscar evidências.

A Solução de Longo Prazo

- **Alfabetização Midiática:** Ensinar, desde a escola, como consumir, analisar e criar mídia de forma responsável.
- **Pensamento Crítico:** A habilidade de questionar informações, identificar vieses e buscar evidências.
- **Responsabilidade das Plataformas:** A necessidade de redes sociais e empresas de tecnologia desenvolverem ferramentas de detecção e políticas claras. ex: Marco civil da internet

A Solução de Longo Prazo

- **Alfabetização Midiática:** Ensinar, desde a escola, como consumir, analisar e criar mídia de forma responsável.
- **Pensamento Crítico:** A habilidade de questionar informações, identificar vieses e buscar evidências.
- **Responsabilidade das Plataformas:** A necessidade de redes sociais e empresas de tecnologia desenvolverem ferramentas de detecção e políticas claras. ex: Marco civil da internet
- A tecnologia não é a bala de prata. A defesa mais forte é uma sociedade educada e cética.



Recapitulando:

Recapitulando:

- Deepfakes são uma tecnologia poderosa com um enorme potencial para o bem e para o mal.

Recapitulando:

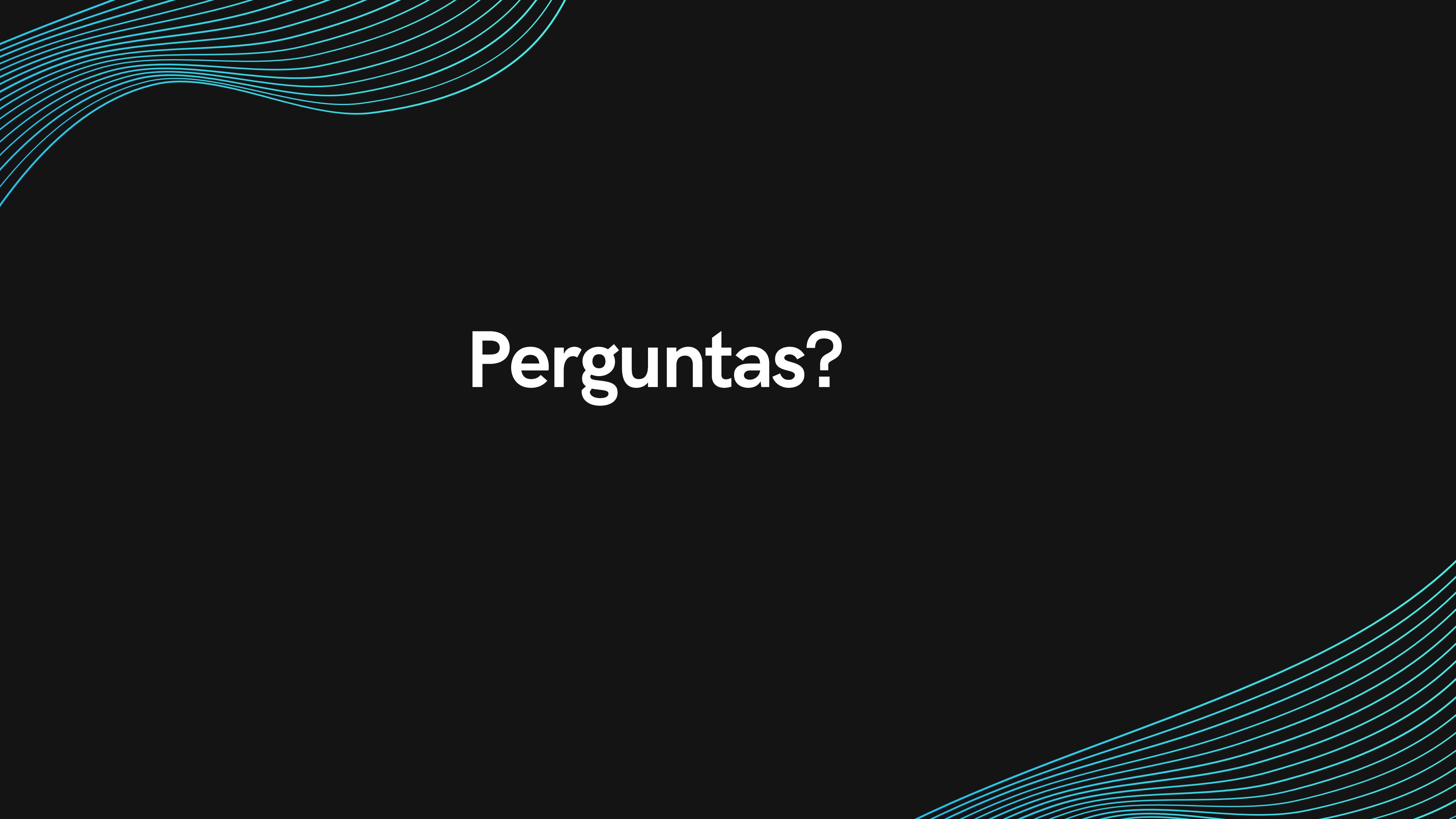
- Deepfakes são uma tecnologia poderosa com um enorme potencial para o bem e para o mal.
- Ignorar não é uma opção. Entender é o primeiro passo para a proteção.

Recapitulando:

- Deepfakes são uma tecnologia poderosa com um enorme potencial para o bem e para o mal.
- Ignorar não é uma opção. Entender é o primeiro passo para a proteção.
- O futuro da verdade depende da nossa capacidade de nos adaptarmos a essa nova realidade.

Recapitulando:

- Deepfakes são uma tecnologia poderosa com um enorme potencial para o bem e para o mal.
- Ignorar não é uma opção. Entender é o primeiro passo para a proteção.
- O futuro da verdade depende da nossa capacidade de nos adaptarmos a essa nova realidade.
- Agora que entendemos a teoria, vamos para a próxima etapa: uma análise prática .



Perguntas?