

CODA Amazônia 2024

**Mobilizando bancos de  
casos para enfrentar a  
discriminação algorítmica**

Me. Fernanda Rodrigues



# Olá! Meu nome é Fernanda.

Sou doutoranda em Direito, Tecnociências e Interdisciplinaridade pela UFMG. Mestre em Direitos na Sociedade em Rede e graduada em Direito pela UFSM. Coordenadora de pesquisa e pesquisadora no Instituto de Referência em Internet e Sociedade (IRIS). Professora. Consultora em pesquisas e projetos na área de raça, direito e novas tecnologias.

# Agora é com você ;)

Qual o seu nome?

De onde você vem?

No que você atua?

Por que se interessou pela oficina?



# Tópicos da oficina

- 01** Inteligência artificial
- 02** Regulação de IA no Brasil
- 03** Sistemas de IA de alto risco e risco excessivo
- 04** Importância do registro dos casos de danos algorítmicos
- 05** Bancos de casos de danos algorítmicos

# Inteligência artificial

**Definição pela OCDE (Organização para a Cooperação e Desenvolvimento Econômico):**

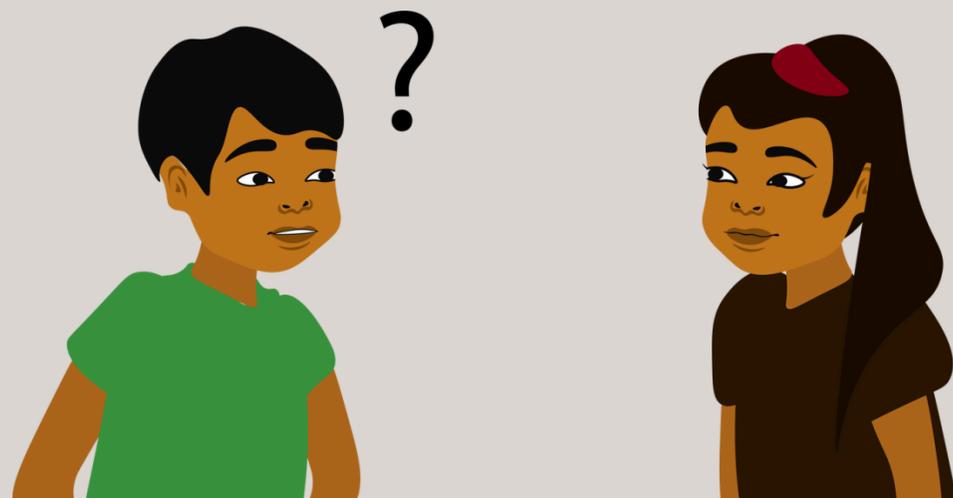
“Um sistema de IA é um sistema baseado em máquina que, para um determinado conjunto de objetivos explícitos ou implícitos, infere, a partir das informações que recebe, como gerar resultados, como previsões, conteúdo, recomendações ou decisões que podem influenciar ambientes físicos ou virtuais. Diferentes sistemas de IA variam nos seus níveis de autonomia e adaptabilidade após a implantação”.



# Regulação da IA no Brasil

- Disputa narrativa em torno dos possíveis danos x possíveis benefícios da tecnologia
- Regulação x inovação
- Regulação principiológica x regulação baseada em direitos e riscos

**Mas o que é  
gradação de riscos  
de tecnologias de IA?**



# Gradação de riscos da IA

Risco excessivo

Risco alto

Risco moderado

Risco baixo



# Risco excessivo

Alguns exemplos do que hoje (27/06) está previsto como **risco excessivo** na proposta de regulação para IA no Brasil:

- sistemas capazes de explorar vulnerabilidades para induzir comportamentos que podem causar danos à saúde e outros direitos
- sistemas de escore de crédito social
- sistemas que permitem produção, disseminação ou facilitação de criação de material de abuso ou exploração sexual de crianças e adolescentes
- **armas autônomas**
- sistemas de reconhecimento facial, **salvo uma série de exceções**



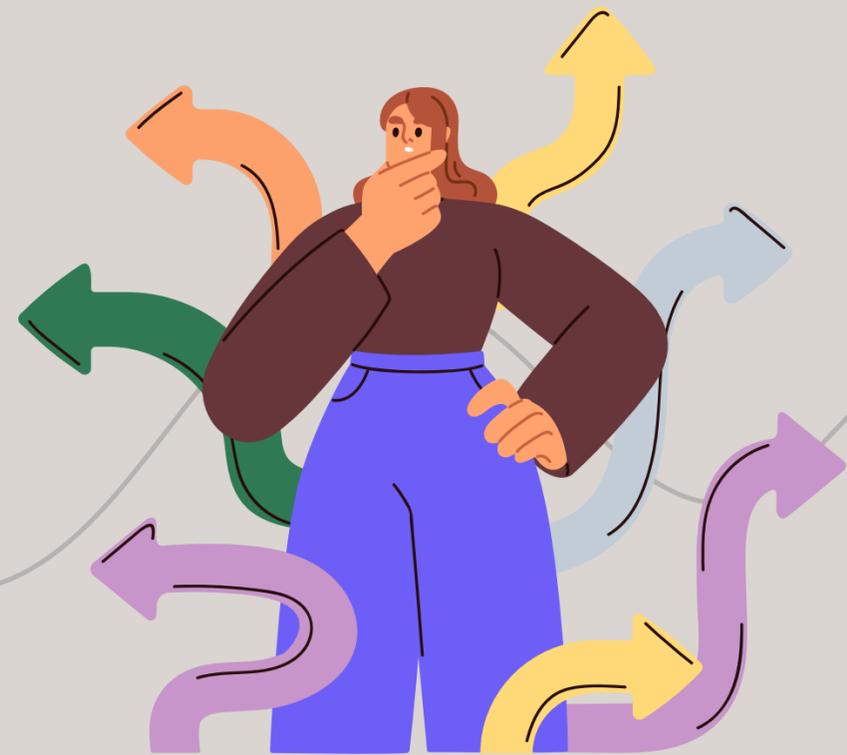
# Risco alto



Alguns exemplos do que hoje (27/06) está previsto como **risco alto** na proposta de regulação para IA no Brasil:

- recrutamento e avaliação de candidatos em processos seletivos de emprego
- veículos autônomos em espaços públicos, quando o uso puder gerar risco relevante à integridade física de pessoas
- sistemas que podem auxiliar diagnósticos e procedimentos médicos, havendo risco relevante à integridade física e mental
- identificação biométrica para reconhecimento de emoções
- estudo de crimes para identificar padrões e perfis comportamentais

**E por que é importante  
olhar para os casos de  
discriminação  
algorítmica?**



**A lista de tecnologias de IA de alto risco  
poderá ser atualizada pelo Sistema Nacional  
de Regulação e Governança de Inteligência  
Artificial.**

## Atualização da lista de IAs de risco alto

Deverá levar em consideração a probabilidade e a gravidade dos impactos adversos sobre pessoa ou grupos afetados e pelo menos um critério específico, dentre os quais se destaca:

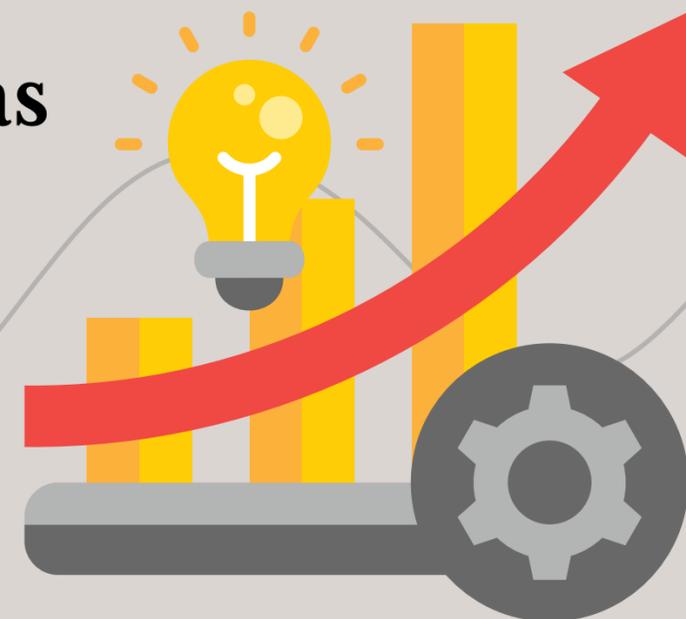
- produção ilícita e/ou abusiva de efeitos jurídicos relevantes e impacto negativo sobre acesso a serviços públicos ou essenciais
- alto potencial danoso de ordem material ou moral e viés discriminatório ilegal ou abusivo
- afetar negativamente pessoas de grupo vulnerável
- **histórico danoso, de ordem material ou moral**
- alto potencial danoso sistêmico, tais como à segurança cibernética, higidez do processo eleitoral e violência contra grupos vulneráveis



**Bancos de casos de  
discriminação algorítmica  
podem contribuir para  
demonstrar um ou mais critérios  
de atualização.**

# E além disso...

- Contribuem para traçar **estratégias de mitigação adequadas**
- Auxiliam a compreender **quais as comunidades mais impactadas** por determinada tecnologia
- Podem indicar **quais tecnologias são inadequadas** para um contexto em específico
- **Dentre outras**



# Mapeamento de danos algorítmicos (Desvelar)

## Desvelar

Justiça Racial na Inteligência Artificial e TICs

Desvelar /

<https://desvelar.org/casos-de-discriminacao-algoritmica/>



# Bases de dados de interseções sobre IA (Fundação Mozilla)

m AI Intersections Database Database About Protect Your Data Get Involved Contribute to the database

## AI Intersections Database

Search the database:  Search

Search by keyword, actor, or issue

Browse database by justice areas:

- Community Health and Collective Security
- Disability Justice
- Economic Justice
- Environmental Justice
- Gender Justice
- Human Rights
- Racial Justice

<https://aidatabase.mozilla.org/>

# Bases de datos de incidentes de IA (Responsible AI Collaborative)

The screenshot shows the AI Incident Database (AIID) website. The header features the AIID logo and the text "AI INCIDENT DATABASE". Below the header, there are "Discover" and "Submit" buttons. A search bar contains the text "Search over 3000 reports of AI harms" and has "Search" and "Discover" buttons. A sidebar on the left lists various navigation options: "Welcome to the AIID", "Discover Incidents", "Spatial View", "Table View", "Entities", "Taxonomies", "Word Counts", "Submit Incident Reports", "Submission Leaderboard", "Blog", "AI News Digest", "Risk Checklists", "Random Incident", and "Subscribe". The main content area displays a list of incidents. The first incident is titled "Incident 721: Fake AI-Generated Students Are Reportedly Enrolling in Online College Classes". Below the title is a quote: "AI spambots invade community college classroom, professor puzzled" with a "Latest Incident Report" tag. The source is cited as "cybernews.com 2024-06-18". A "Read More" button is located at the bottom right of the incident card.

<https://incidentdatabase.ai/>

# Monitor de incidentes de IA (OCDE)

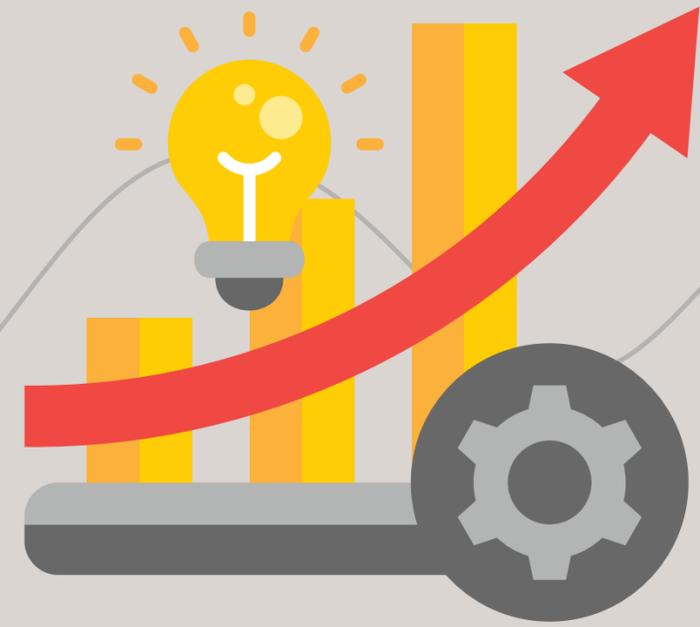
The screenshot shows the OECD AI Incidents Monitor (AIM) website. The header includes the OECD.AI Policy Observatory logo, navigation links (Blog, Live data, Countries, Priority issues, Tools, Resources, About), and a search icon. The main content area features the title 'OECD AI Incidents Monitor (AIM)' and a subtitle 'A repository of AI incidents and hazards'. A paragraph describes the AIM's purpose: 'AI-related legislation is gaining traction, and effective policymaking needs evidence, foresight and international cooperation. The OECD AI Incidents Monitor (AIM) documents AI incidents and hazards to help policymakers, AI practitioners, and all stakeholders worldwide gain valuable insights into the risks and harms of AI systems. Over time, AIM will help to show risk patterns and establish a collective understanding of AI incidents and hazards and their multifaceted nature, serving as an important tool for trustworthy AI.' Below this is a link to 'Please see methodology and disclosures here.' A search bar with a 'Reset' button is present. The main content is divided into two sections: 'Summary visualisations of AI incidents & hazards' and 'Summary statistics of AI incidents & hazards'. The first section shows a line chart titled 'Evolution of incidents and hazards' with a y-axis labeled 'Incidents and hazards' ranging from 600 to 800. The second section is a table with the following data:

	Incidents & hazards	Articles
All time total	11419	63495
Current month's total	594	3053
Last month's total	759	4182
Peak month	2024-02	2024-02

<https://oecd.ai/en/incidents>

# Formas de fazer a análise dos dados

- Teoria fundamentada nos dados
- Análise de conteúdo
- Análise de discurso





# Obrigada!

fernandarodrigues.blog  
dsrs.fer@gmail.com