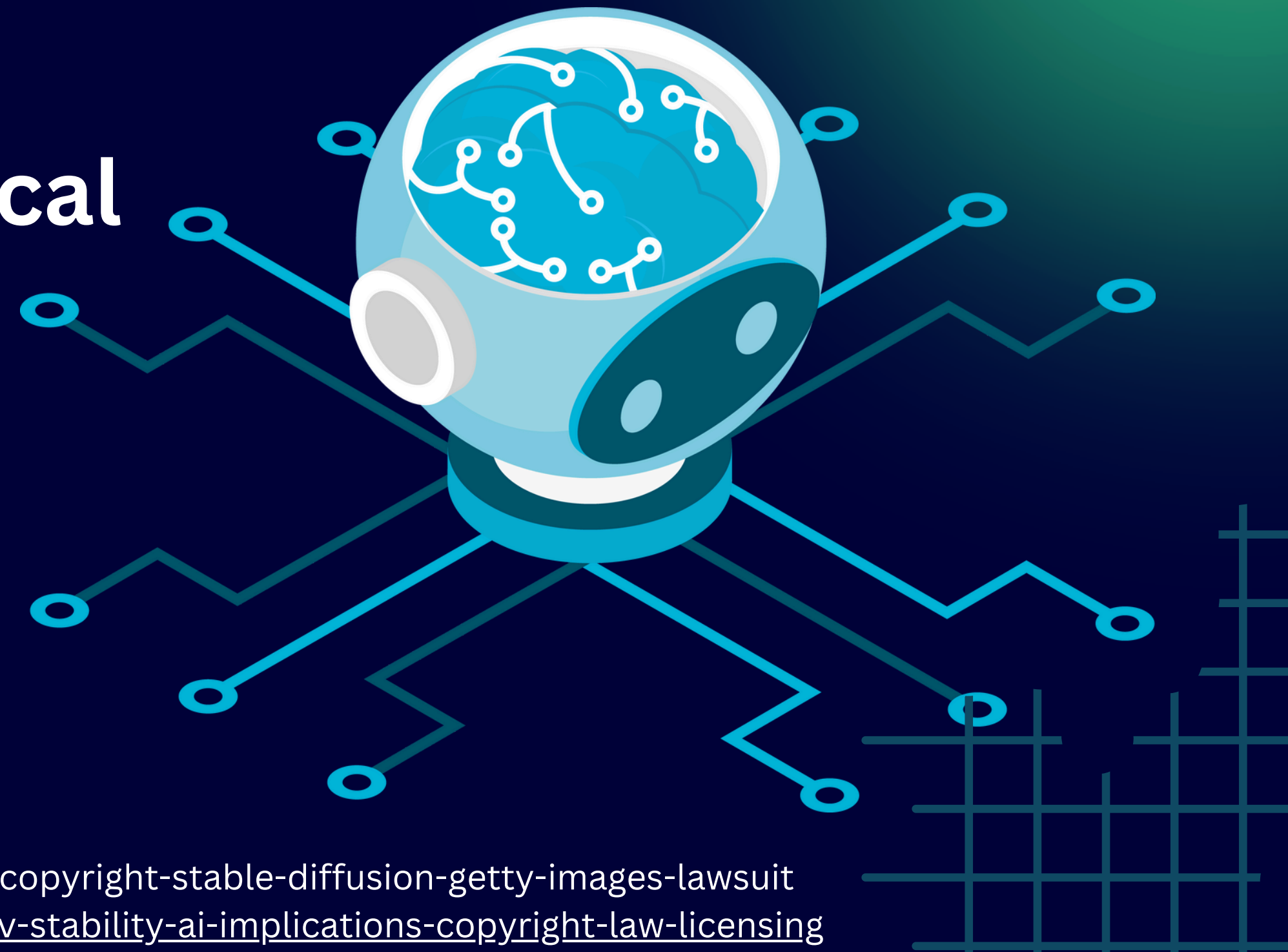


# Getty Images vs. Stability AI

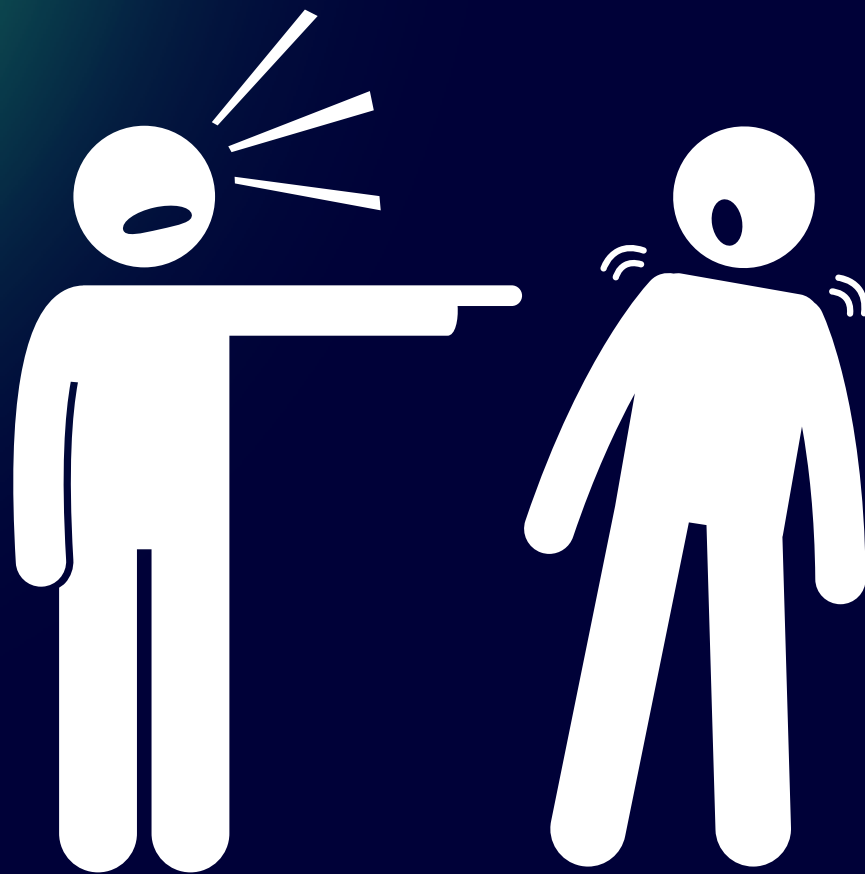
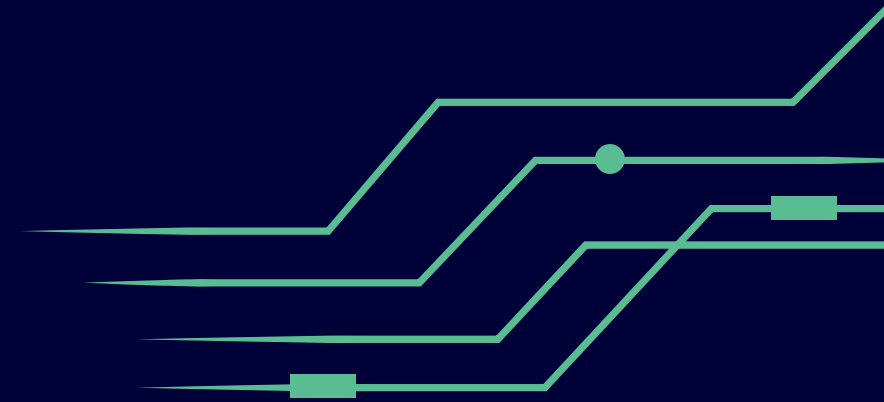
## A Brief Overview of Ethical Issues in AI Data Collection

Ali Muhammad Asad



sources: <https://www.theverge.com/2023/1/17/23558516/ai-art-copyright-stable-diffusion-getty-images-lawsuit>  
<https://www.pinsentmasons.com/out-law/analysis/getty-images-v-stability-ai-implications-copyright-law-licensing>

# Getty Images sues Stability AI for scraping content



## 1 What Happened?

Stability AI allegedly used millions of Getty's copyrighted images without permission to train its AI model (Stable Diffusion).

## 2 Harm Caused

- Copyright violation
- Financial Loss
- Devaluation of professional content
- Loss of trust in AI companies

# Analysis of Causes

## Data and Model Design

- Indiscriminate usage of web scraping without permission
- Lack of copyright attribution
- Lack of filtering / auditing mechanisms
- No watermark detection or image tracking

## Business Decisions

- Preference of rapid development over proper licensing and practices
- Most likely due to assumption - fair use would compensate usage of copyright images
- Major oversight indicating lack of legal expertise

## Ethical Oversight


- Failure to consider the impact on professional content creators
- Failure to consider potential market devaluation of artists





# Preventive Measures

Plenty of measures could be taken to prevent such incidents, including technical and non-technical measures



## ETHICAL DATA COLLECTION

Development of comprehensive data collection frameworks ensuring licensing, attribution, and filtering

## LEGAL COMPLIANCE

Ensurance of AI compliance with fair use guidelines, engaging legal experts in the process.

## STAKEHOLDER ENGAGEMENT

Collaborate with creative industries to mitigate negative impacts.

## ONGOING MONITORING

On going monitoring and data auditing to ensure data is compliant with the fair guidelines.





# Thank You

