

# Investigating Harmful AI News

Ali Muhammad Asad  
aa07190

## 1. Headline and Source

**Headline:** Getty Images is suing the creators of AI art tool Stable Diffusion for scraping its content **Link:** Verge Article by James Vincent (January 27, 2023)

## 2. Summary of Incident

- **What Happened?**

Stability AI, the creator of Stable Diffusion (a popular AI image generation tool), was sued by Getty Images (a supplier of stock images, editorial photography, video and music for business and consumers). The AI company was allegedly used millions of Getty's copyrighted images to train their model without permission. The case highlighted crucial issues about AI training data rights and creative industry impact (something which artists of all kinds have been concerned about for a long time).

- **Harm:**

Getty images suffered a major loss due to direct copyright violations as Stable AI unauthorizedly used millions of their professional photographs without any compensation to the original creators, which is a major ethical violation, and a financial loss. This also causes a potential market value destruction for photographers and artists who create such content. In addition, this hurts the economy as well due to the devaluation of professional photography, and the disruption of stock photo business models. Thus, it causes reduced income for professional content creators and photographers who have invested time and effort. Moreover, it also causes a loss of trust in AI companies and uncertainty in AI training practices, raising questions about fair use in AI development, threatening traditional creative industry's sustainability.

## 3. Analysis of Causes

1. **Data Collection Practices:** This highlights the indiscriminate usage of web scraping without permission of the original owners, lack of any filtering mechanism for copyrighted content, and a severe lack of attribution system in the model.
2. **Model Design:** This also highlights the lack of built-in watermark detection, inability to track source images, and no mechanism to respect copyright status.
3. **Business Decisions:** The business definitely preferred rapid development over proper licensing and data collection practices, which is a major ethical violation, and highlights the lack of proper legal and ethical guidelines in AI development. This most likely occurred due to assumption that fair use would cover the training data containing copyrighted

images, which is a major oversight, and also shows the lack of legal expertise in the AI company.

4. **Ethical Oversight:** As mentioned above, the company failed to consider the ethical implications of using copyrighted images without permission, and failed to consider the impact on the creative industry, and the potential market value destruction for photographers and artists. This raises a major concern regarding ethical considerations for data collection practices, and inadequate stakeholder engagement in AI development.

#### 4. Preventive Measures

1. **Ethical Data Collection Framework:** AI companies should develop a comprehensive ethical data collection framework that ensures proper licensing, attribution, and filtering mechanisms for copyrighted content. This should also include a robust watermark detection system, and a mechanism to track source images.
2. **Legal Compliance:** AI companies should ensure legal compliance with fair use guidelines, and should engage legal experts to ensure that their data collection practices are in line with legal requirements.
3. **Stakeholder Engagement:** AI companies should engage with stakeholders in the creative industry to understand the impact of their data collection practices, and should take steps to mitigate any negative consequences on the creative industry.
4. **Ongoing Monitoring:** AI companies should implement ongoing monitoring and auditing of their data collection practices to ensure compliance with ethical and legal guidelines, and should take corrective action if any violations are detected.
5. **Policy Changes:** For AI companies can develop clear data collection guidelines, create ethical review boards, implement fair compensation models, and develop stakeholder consultation processes. Industries can create legal frameworks to protect content creators, and ensure fair use of their content.

#### 5. Conclusion

The case of Getty Images suing Stability AI highlights the importance of ethical data collection practices in AI development, and the need for legal compliance and stakeholder engagement. AI companies should develop comprehensive ethical data collection frameworks, engage with legal experts and stakeholders, and implement ongoing monitoring and auditing of their data collection practices to ensure compliance with ethical and legal guidelines. This case also raises important questions about fair use in AI development, and the impact of AI on the creative industry. It is essential for AI companies to consider the ethical implications of their data collection practices, and take steps to mitigate any negative consequences on the creative industry.