



hiQ Labs v. LinkedIn



&

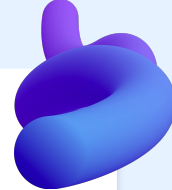
A.T. v. OpenAI



OpenAI

Akshat Lal & Brandon Le

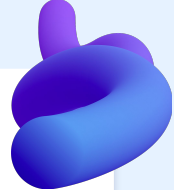
key terms



- **Data scraping:** The automated collection of data from a website. In this case, LinkedIn argued that such activity, even for **public data**, could be restricted.
- **Computer Fraud and Abuse Act (CFAA):** A U.S. law that criminalizes unauthorized access to computers and networks. The debate in this case focused on whether scraping public data constituted "unauthorized access."
- **Public Data:** Information that is accessible without login restrictions.

A.T. v. OpenAI LP

key terms



- **Data Scraping:** OpenAI/Microsoft have been hit with lawsuits over scraping of copyrighted materials and personal data from across the internet to train their generative AI systems.
- **Accountability:** The firms bringing the lawsuits hope to “hold BigAI accountable for their mass theft of personal information and violations of privacy, property and consumer rights”.
- **Professional obsolescence:** The two SWEs that brought the lawsuit are concerned that companies have incorporated their “skills and expertise” into products that could take their jobs one day.

case 1:

hiQ Labs Inc v. LinkedIn Corp

Landmark
Case:

LinkedIn vs
hiQ Labs



hiQ Labs Inc v. LinkedIn Corp

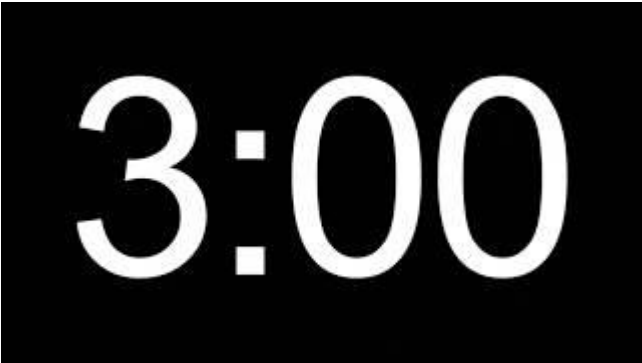
overview

- hiQ Labs was scraping LinkedIn's public profiles for data analytics. LinkedIn attempted to block hiQ's access, citing user privacy concerns and violation of its terms of service.
- The case revolved around data scraping and whether accessing public data on a platform like LinkedIn without permission constitutes a violation of the Computer Fraud and Abuse Act (CFAA) and LinkedIn's terms of service.

hiQ Labs Inc v. LinkedIn Corp

Question

- Take 3 minutes. Try to find LinkedIn Corp terms of service and where it states that you can't scrape information from the website



3:00

hiQ Labs Inc v. LinkedIn Corp

Verdict

- Settlement in December 2022, favoring LinkedIn.
- hiQ agreed to a permanent injunction
- Court ruled that hiQ's scraping likely did *not* violate the CFAA
- Terms of Service was viewed as a contract

case 2:

A.T. v. OpenAI LP



A.T. v. OpenAI LP

overview

- Alleges that **OpenAI** and **Microsoft** used web scraping to collect private data
- Argument is that they violate us consumer privacy laws
- Accusing the companies of using stolen private information to improve models
- This is the second class-action lawsuit filed against OpenAI and Microsoft after one in June 2024.

A.T. v. OpenAI LP

Verdict

- Still ongoing



hiQ Labs Inc v. LinkedIn Corp

Question

- Do you agree or disagree with the statement that it was better for OpenAI to ask for forgiveness than to ask for permission? That is, should OpenAI have reached out to various companies to create partnerships before gathering data and creating ChatGPT?

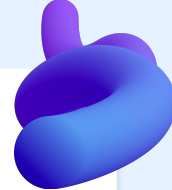
3:00



Can enforcing terms of service over public data create monopolies on information? What are the implications?

Examples:

- Amazon & E-commerce data
- Google & Search behavior data
- Facebook & Social interaction data
- LinkedIn & Professional networking data
- Bloomberg & Financial market data
- Netflix & Viewing preferences data

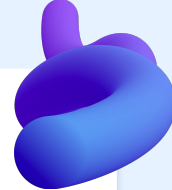


How can regulatory bodies balance the need for innovation in AI with the protection of consumer privacy? What risks arise from too much or too little regulation in this space?

Does more regulation hamper innovation?

Should companies be allowed to block third parties from scraping public data on their platforms?

What is the difference between a company web scraping and an individual scraping data privately?



Should terms of service have the power to override a user's right to access and collect public data, or does this set a problematic precedent for information control on the internet, potentially conflicting with **First Amendment rights**?

First Amendment

Congress shall make no law respecting an establishment of religion, or prohibiting the free exercise thereof; or abridging the freedom of speech, or of the press; or the right of the people peaceably to assemble, and to petition the Government for a redress of grievances.

“right to receive information and ideas”

Should terms of service have the power to override a user's right to access and collect public data (first amendment) ?

Does restricting data scraping hinder innovation and competition, or does it protect user privacy and company interests?

In your opinion which is more important?



Should AI companies have explicit consent from users when scraping publicly available data from the internet? Why or why not?



Is it fair for companies to monetize AI products trained on personal data without compensating the individuals whose data was used? Why or why not?



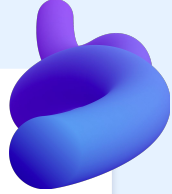
In your opinion, is scraping publicly available data, a violation of the Computer Fraud and Abuse Act (CFAA) ?

CFAA: A U.S. law that criminalizes **unauthorized** access to computers and networks.

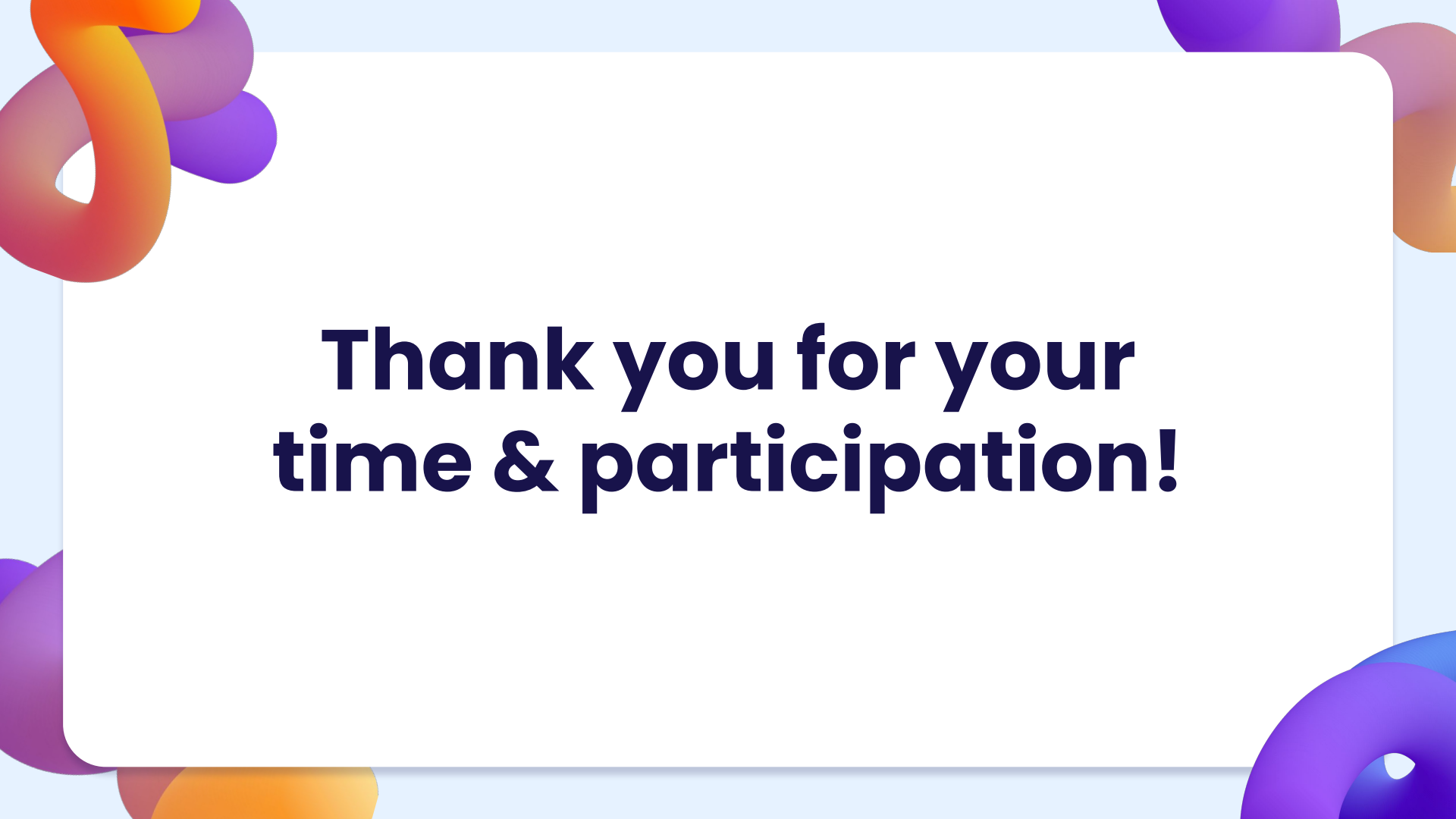
Who truly owns the data that is publicly available on the internet, and how should ownership rights be determined in the context of AI model training?

Is web scraping legal?





How might this case influence future regulatory frameworks for AI, especially regarding consumer privacy, privacy policies and data usage?

The background is a light blue gradient. It is decorated with several large, colorful, abstract shapes that resemble thick, twisted ribbons or loops. These shapes are in shades of orange, yellow, and purple, and are positioned around the edges of the central white box. One large orange and yellow loop is in the top-left corner, and another is in the top-right corner. There are also purple and orange shapes in the bottom-left and bottom-right corners.

**Thank you for your
time & participation!**