imperva

What is web scraping?

Web scraping is the process of using bots to extract information from a website. In recent years, the debate over web scraping is growing more complex as business intelligence and data privacy issues arise.

Potential legitimate use cases for web scraping

- Search engine bots that crawl a website to analyze and rank its content
- Price comparison sites deploying bots to auto-fetch prices and product descriptions for allied seller websites
- Market research companies using scrapers to pull data from forums and social media

How bad actors use web scraping

- Undercutting prices for an unfair competitive advantage
- Stealing copyrighted content and intellectual property

Is web scraping legal? An ongoing debate

The legality of web scraping is unsettled, and legal claims have been pursued by website owners to prevent scraping of their sites.

For example, LinkedIn has been engaged in a multi-year legal battle with hiQ Labs, asserting, among other claims, that hiQ Labs' scraping of LinkedIn's site violated the Computer Fraud and Abuse Act (CFAA). In April 2022, the U.S. Court of Appeals for the Ninth Circuit Court rejected this argument, reconfirming its earlier ruling that, in this case, scraping publicly available information on a website did not violate the CFAA. The Court's ruling did not foreclose the possibility of other remedies that may be available to a website owner (e.g., contract or common law trespass claims).

Given the prevalence of web scraping, organizations understanding the need to protect their proprietary data while maintaining a legitimate flow of traffic are adopting preventative measures to mitigate automated threats.

Contact us now to schedule a demo or begin a free trial

imperva.com