

# Data scraping policy briefing

As a valued university partner, we want to keep you informed of an important policy matter that requires your awareness and, in some cases, direct action. We have observed an increase in automated data scraping activity originating from student accounts at several of our university partner institutions. This briefing explains what data scraping is, why it violates our Terms of Service, how we detect it, and what consequences apply when it is identified.

## What is data scraping?

Data scraping—also referred to as web scraping or automated data extraction—refers to the use of software tools, scripts, bots, or browser extensions to systematically and automatically collect large volumes of data from the PitchBook Platform at a speed or scale that exceeds normal human usage.

This is distinct from a user manually browsing and reviewing information. Scraping involves:

- Automated scripts or programs that submit rapid, repeated requests to PitchBook servers
- Browser extensions or plugins designed to extract and export bulk data
- Bots or headless browser tools that simulate user activity to harvest data
- Third-party applications that integrate with PitchBook to pull data en masse without authorization
- Copy-paste automation tools used to systematically extract large datasets
- Translation or data transformation services that access PitchBook content programmatically to process or reformat data at scale

## Why data scraping is prohibited

PitchBook's Terms of Service explicitly prohibit the automated extraction of data from the platform. This policy exists for several important reasons:

**Intellectual property:** PitchBook's data is proprietary. Unauthorized extraction undermines the significant investment made to source, validate, and maintain this data.

**Fair access:** Automated scraping consumes server resources at a disproportionate rate, degrading performance for all legitimate users.

**License compliance:** University licenses grant access for individual research and educational use—not bulk data extraction for redistribution or commercial purposes.

**Data integrity:** Scraped data is often incomplete, outdated, or misused, leading to inaccurate analysis.

## How we detect data scraping

**Fully automated detection** PitchBook employs automated, real-time monitoring systems designed to identify scraping behavior as it occurs. When scraping activity is detected, our systems flag the account immediately without manual review delay. There is no grace period—accounts are acted upon as soon as the threshold is crossed.

Our detection methodology includes, but is not limited to:

- Monitoring request frequency and volume per user session
- Identifying non-human browsing patterns (e.g., sequential data pulls, absence of normal page navigation)
- Flagging known scraping tool signatures and browser automation fingerprints
- Detecting API misuse or unauthorized integration attempts

## Consequences of scraping activity



**Immediate and permanent account termination** Any account found to be engaged in data scraping will be permanently terminated immediately and automatically upon detection. The student will lose access to PitchBook for the remainder of their enrollment at the institution, without prior notice.

Violations may also affect the institution's overall license status. In addition to permanent account termination, the following may apply:

- Escalation to the university's IT and compliance teams
- Review of the institution's broader access privileges
- Potential legal action in cases of significant or willful violations

## Your role as point of contact

We ask that you take the following actions to help prevent scraping incidents within your institution:

- Share the enclosed student-facing notice with your student user population
- Include scraping policy language in any onboarding materials or orientations related to PitchBook access
- Direct students with legitimate large-dataset needs to contact their university library for guidance
- Notify us if you become aware of organized or coordinated scraping activity



**Legitimate data needs?** If students or faculty have academic research requirements that involve larger data sets, we encourage you to reach out to us directly. We have data export options and research programs that may be able to accommodate legitimate needs within the bounds of our licensing terms.

## Questions & contact

For questions about this policy, a specific account suspension, or support for legitimate data needs, please reach out to your account manager.