Security update: Incident involving unauthorized admin access

Diego Comas, Head of Security August 30, 2023

TL;DR: Sourcegraph experienced a security incident that allowed a single attacker to access some data on Sourcegraph.com. This was limited to:

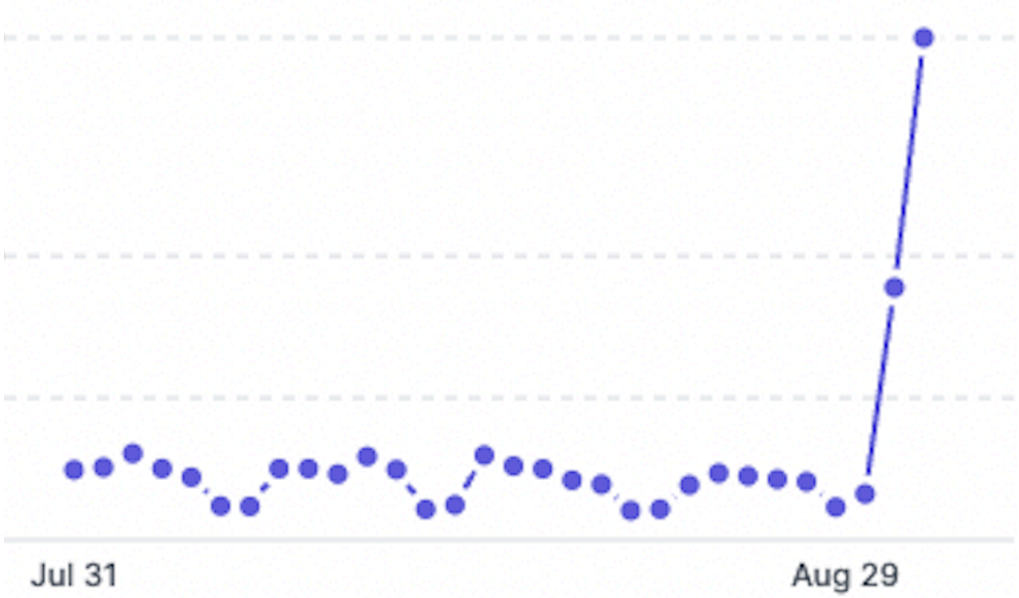
- Paid customers:
  - The license key recipient's name and email address.
  - A small subset of customers' Sourcegraph license keys may have been accessed (note that license keys *do not* enable access to Sourcegraph instances). We are reaching out directly to those who may have been impacted to rotate license keys.
- Community users:
  - Sourcegraph account email addresses. No action is required.

No other customer info, including private code, emails, passwords, usernames, or other PII, was accessible.

Background

Sourcegraph experienced a security incident on August 30, 2023 where a malicious actor used a leaked admin access token in our public Sourcegraph instance at Sourcegraph.com. The malicious external user used their privileges to increase API rate limits for a small number of users.

On August 30, 2023 our team noticed a significant increase in API usage and began investigating the cause.



The spike in usage was ruled as isolated and inorganic and our security, engineering, and support teams quickly assembled to understand what was going on.

Our security team identified a code commit from July 14 where a site-admin access token was accidentally leaked in a pull request and was leveraged to impersonate a user to gain access to the administrative console of our system.

In the spirit of transparency, we want to share the full timeline of the incident and what we have done to resolve this incident, as well as additional steps we're taking to prevent this kind of leak in the future.

Timeline

On July 14, 2023 *(2023-07-14 22:01:00 UTC)* a Sourcegraph engineer accidentally committed a code change that contained an active site-admin access token. The site-admin access token had broad privileges to view and modify account information on Sourcegraph.com.

Sourcegraph.com is an instance of Sourcegraph that contains public code only. It's also used for authentication for free-tier Cody users. It is separate from all paid customer instances (both on-premises and cloud). The instance also hosts our license management for all customers.

Our internal control systems, including automated code analysis, failed to catch the access token being committed to the repository.

On August 28, 2023 *(2023-08-28 13:18:36 UTC)*, a user created a brand new Sourcegraph account.

On August 30, 2023 *(2023-08-30 06:47:59 UTC)*, using the leaked site-admin access token, this user elevated their account privileges to a site-admin and gained unauthorized access to the admin dashboard.

The malicious user continued to probe the system by changing their access from a site-admin to regular user multiple times.

The malicious user, or someone connected to them, created a proxy app allowing users to directly call Sourcegraph's APIs and leverage the underlying LLM. Users were instructed to create free Sourcegraph.com accounts, generate access tokens, and then request the malicious user to greatly increase their rate limit.

On August 30 *(2023-08-30 13:25:54 UTC)*, the Sourcegraph security team identified the malicious site-admin user, revoked their access, and kicked off an internal investigation for both mitigation and next steps.

Impact

The promise of free access to Sourcegraph API prompted many to create accounts and start using the proxy app. The app and instructions on how to use it quickly made its way across the web, generating close to 2 million views.

As more users discovered the proxy app, they created free Sourcegraph.com accounts, adding their access tokens, and accessing Sourcegraph APIs illegitimately.

The impact of the malicious user having admin access was limited to a subset of:

- **Paid Customers**

  - The license key recipient’s name and email address
  - Sourcegraph license key
- **Free-Tier Community Users**
  - Email addresses

We have no indication that any of this data was viewed, modified, or copied, but the malicious user could have viewed license key recipients' emails and community user email addresses as they navigated the admin dashboard.

Regarding paid customer license key exposure, we saw that the user accessed a page in the admin dashboard where they would have only seen the first 20 license key items. We were able to determine which items those were at the time of viewing because of stable sorting.

Important Note: Customers’ private data or code was not viewed during this incident. Customer private data and code resides in isolated environments and were therefore not impacted by this event.

How we're mitigating

As soon as we understood the scope of the incident we took the following steps:

- Identified the malicious account and fully revoked its access
- Proactively rotated a subset of Sourcegraph customer license keys that may have been viewed
- Temporarily reduced the rate limits for all free community users
- Created new processes and tests and will continue to monitor for malicious activity and abuse

Expanding our secret scanning through additional static analysis tests will ensure we can better detect and prevent this kind of leak in the future.

If you’re a Community user, we know these rate limit reductions aren’t ideal for devs who are using Cody to help them write and understand code. This reduction will be

short-term while we investigate the issue further.

Next steps

Our teams are actively working to create a long-term solution for our community and customers to prevent future incidents like this. While we are not ready to publicly share our additional mitigation options at this time as our internal investigation is still ongoing, know that we are working around the clock to implement a solution that is least disruptive to the Sourcegraph community at large.

Stay tuned for more updates and be sure to join our Discord community for the latest.

FAQ

Is my code host data compromised?

No private customer data or code was accessible.

Is there any action that I need to take?

If you are part of the subset of customers whose Sourcegraph license keys may have been accessed, your account team will reach out with steps and a new license key as soon as possible. *Note: The Sourcegraph license key does not enable any access to a customer instance.*

For free-tier users with a Sourcegraph.com account: No action is needed.

What email addresses were viewable?

- Paid customers: When a Sourcegraph customer receives their license, they provide one email address to associate with their license key. Only this recipient email is stored in Sourcegraph.com and no other customer emails were accessible.
- Community users: Sourcegraph.com account email addresses.

I have more questions, who can I contact?

Reach out to your Account team (Technical Advisor or Account Executive) or our Support team at support@sourcegraph.com.

Updated August 31, 2023: Added a detail to the ***`Impact`*** *section clarifying how we determined which license keys could have been viewed.*