One-page design challenge: Reinforce use of machine user for organization code host connections

Date: Nov 9, 2021
Author: Quinn Keast
Contributors: Milan Freml

Problem Statement

When the users who create organizations on Cloud set up a code host connection, they provide a personal access token that Sourcegraph later uses to fetch a list of repositories that can be added to Sourcegraph. This personal access token is never subsequently revealed to users.

However, for early access, all members of an organization are “admins” for that organization, and can view and make changes to which repositories that organization has added to Sourcegraph, based on the originally-provided personal access token—even if that personal access token belongs to someone else in their organization.

As a result, if the creator uses their own personal access token to create the code host connection, other members of the organization can view a list of all of that user’s public and private repositories in the “Manage repositories” view.

In the future, we will migrate to a GitHub app, which will resolve this problem. However, for early access, we will instead strongly recommend use of “machine users” for code host connections. Machine users are “fake” user accounts to which the organization can grant granular repository access on the code host, and then use the machine user’s personal access token to create the code host connection on Sourcegraph.

We’ve observed in user testing that the current design implementation for the personal access token modal is **very familiar and easy** for developers **(PROBLEM)**, which is great on the surface, but risks creators may use their own personal access token without recognizing or acknowledging the risk of doing so **(USER PROBLEM)**. If we improve this, we’ll create a better shared responsibility model for early access, where users either stop and set up a machine user before creating the code host connection, or acknowledge and opt-in to the exposure created by using their own personal access token **(POSITIVE OUTCOME)**.

Hypotheses

We believe we will prevent unintentional data exposure and loss of trust in Sourcegraph Cloud **(POSITIVE OUTCOME)** by...

1. Introducing friction into the process of adding a personal access token while creating code host connections for organization
2. Clearly documenting the value and process of setting up a machine user

We'll know this is true if:

- In internal hallway testing, creators set up a machine user, or directly acknowledge they are comfortable with the data exposure to other members of their organization
- During early access, our participating organizations confirm with us that they have set up a machine user for their code host connection
- During early access, our participating organizations confirm with us that they used their own personal access token, but that they were aware of the risk and comfortable with the data exposure to other members of their organization

Design exploration

- Low-fidelity collaborative exploration in Excalidraw
- ⭐High-fidelity in Figma