

# SOURCEBOT PRIVACY POLICY

Version: May 29<sup>th</sup>, 2026 WEB

Taqla Inc. (“**Sourcebot**”, “**we**”, “**us**”) builds Sourcebot, a code understanding tool available as a self-hosted (on-premises) deployment and as a cloud-hosted deployment at [app.sourcebot.dev](https://app.sourcebot.dev). This Privacy Policy explains what personal information we collect, why, and what we do with it.

This policy covers our **website** ([sourcebot.dev](https://sourcebot.dev)), our **cloud deployment** ([app.sourcebot.dev](https://app.sourcebot.dev)), and the limited data our **self-hosted deployments** send back to us. It does **not** cover data that stays inside your self-hosted deployment. That data never reaches us. For an in-depth description of the policies that govern your use of the software, please refer to our [Software License Agreement](#).

If you have questions, email us at [legal@sourcebot.dev](mailto:legal@sourcebot.dev).

## 1. QUICK SUMMARY

- **Self-hosted Sourcebot:** Your code, queries, AI inputs/outputs, and user accounts stay inside your deployment. We never see them. The only things we receive are (a) product usage telemetry, (b) a daily Service Ping with aggregate instance metrics, and (c) information necessary to perform checkout when you start a trial or purchase a subscription. **Sensitive data such as your code, search queries, AI inputs or outputs, repository names, branch names, file contents, usernames, or email addresses is never transmitted.**
- **Cloud Sourcebot ([app.sourcebot.dev](https://app.sourcebot.dev)):** We process your account information, search queries, AI inputs and outputs, and usage telemetry in order to provide the service.
- **We don’t sell your personal information.**
- **You can disable product telemetry** in self-hosted deployments by setting the environment variable `SOURCEBOT_TELEMETRY_DISABLED=true`. Service Ping cannot be disabled via an environment variable, but you can block outbound network requests to our servers (with the caveats described in Section 2b).

## 2. WEBSITE (SOURCEBOT.DEV)

When you visit our marketing site or contact us, we collect:

- Information you submit in forms (name, email, company, message)
- Standard web analytics data (pages visited, referrer, approximate location from IP, browser type)
- Session recordings via PostHog, which capture how you interact with the site (clicks, navigation, scrolls). We use these to understand and improve the user experience. Sensitive form inputs are masked.
- Cookies used to remember your preferences and measure site usage

We use this to respond to your inquiries and improve the site.

## 3. SELF-HOSTED SOURCEBOT

Self-hosted Sourcebot is designed to keep your data inside your infrastructure. Your repositories, code, search queries, AI inputs and outputs, credentials, access tokens, and user accounts are never transmitted to us.

There are only three situations where data leaves your deployment and reaches us. Two of them, product telemetry and Service Ping, include the hostname of your deployment so we can distinguish one deployment from another, but they do not include identifiable information about individual users (no usernames, emails, IP addresses, queries, code, or AI content).

## 2a. Product telemetry

By default, self-hosted Sourcebot sends product usage events through PostHog to help us improve the product. These events are tied to an installation identifier (`install_id`, a UUID generated on first run) and include the hostname of your deployment and the version of Sourcebot you are running.

The events cover things like:

- Lifecycle (install, upgrade) and backend operations (repository indexing, connection sync results)
- Search activity (timing, file/match counts, languages searched, but **never the search query itself or its results**)
- Chat and “Ask Sourcebot” interactions (chat IDs, message counts, tool usage, but **never the message content, AI input, or AI output**)
- Code navigation actions (go-to-definition pressed, find references pressed)
- Ask Connector actions (MCP server URL, tool name)
- Authentication method used (e.g. GitHub, Google, magic link), without identifying the user

Telemetry events **do not include**: your code, search queries, AI inputs or outputs, IP addresses, repository names, branch names, file contents, usernames, or email addresses. PostHog’s standard URL, referrer, and IP properties (`$current_url`, `$pathname`, `$referrer`, `$ip`, and related fields) are explicitly blocked.

We do **not** collect session recordings from self-hosted deployments, because session recordings could capture personal or sensitive information from your environment.

You can disable product telemetry entirely by setting `SOURCEBOT_TELEMETRY_DISABLED=true`. See our [telemetry documentation](#) for more information.

## 2b. Service Ping

Once per day, self-hosted Sourcebot sends a Service Ping to our servers. This helps us understand how Sourcebot is used across deployments, supports our product, support, and sales teams, and, for Pro tier deployments with an online license, verifies your license.

The Service Ping payload includes:

- **installId**: a UUID identifying your installation
- **hostname**: the hostname of your deployment
- **version**: the version of Sourcebot you are running
- **deploymentType**: the deployment type (e.g. `docker-compose`, `kubernetes`, `other`)
- **userCount**: the number of users provisioned in the deployment
- **repoCount**: the number of indexed repositories (count only, not names or URLs)
- **dauCount**, **wauCount**, **mauCount**: daily, weekly, and monthly active user counts as aggregate numbers
- **isTelemetryEnabled**: whether product telemetry (Section 2a) is enabled on this deployment
- **activationCode**: for licensed deployments only, the activation code associated with your license

For deployments running under a paid license, the activation code links the Service Ping to the customer record created at checkout (see Section 2c). For free deployments, no activation code is present.

The Service Ping **does not** include your code, search queries, AI inputs or outputs, repository names, branch names, file contents, usernames, email addresses, or IP addresses.

There is no environment variable to disable the Service Ping. Because you control the network on which your deployment runs, you can block outbound requests to our Service Ping endpoint at your firewall. However:

- **Free tier:** Blocking Service Ping does not affect your ability to use Sourcebot.
- **Pro tier with an online license:** The Service Ping endpoint is the same endpoint that verifies your license. Blocking it will prevent license verification and, after an extended period, cause the software to suspend operation as described in our [Software License Agreement](#). If you need to operate in a fully air-gapped environment on the Pro tier, contact us about an offline license.

## 2c. Trial and subscription checkout

When you check out from inside your self-hosted deployment to start a trial or purchase a subscription, we receive the information necessary to perform the checkout. This consists of:

- The full name, email address, and organization name that you provided to Sourcebot during onboarding.
- For paid subscriptions, your payment information, which is collected and processed directly by Stripe, our payment processor. We do not see or store your payment card details.

This is the only situation in which identifiable information about your users is transmitted to us from a self-hosted deployment, and it only happens when you initiate a checkout.

We use this information to:

- Mint and validate your license activation code
- Process your payment (via Stripe) and manage your subscription
- Contact you about your trial or subscription
- Send you product updates (which you can unsubscribe from at any time)

We do not share this information with third parties except the subprocessors listed below.

## 4. CLOUD SOURCEBOT (APP.SOURCEBOT.DEV)

If you use our cloud deployment, we process substantially more data than we do for self-hosted users, because we operate the service on your behalf. This includes:

- **Account information:** Email address, name, and any other profile information you provide.
- **Authentication data:** Login records, IP address, session information.
- **Usage content:** Search queries you run, inputs you provide to AI features, and outputs the AI returns to you. We process this to deliver the service and may retain it for operational, debugging, and abuse-prevention purposes.
- **Telemetry:** The same product usage events as self-hosted deployments, but associated with your account and IP address rather than anonymized. We also use PostHog to capture session recordings (clicks, navigation, scrolls) to help us understand and improve the product experience. Sensitive fields are masked.

We do not use your queries, AI inputs, or AI outputs to train machine learning models. We do not sell this information.

## 5. HOW WE USE INFORMATION

Across both deployment modes, we use the information we collect to:

- Provide, operate, and improve Sourcebot
- Communicate with you about your account, trial, or subscription
- Send product updates and marketing communications you can opt out of
- Detect and prevent abuse, fraud, and security incidents
- Comply with our legal obligations

## 6. SHARING AND SUBPROCESSORS

We share personal information only with service providers who help us run our business, under contractual confidentiality and data protection obligations. These currently include:

- **Stripe:** payment processing for subscriptions, including checkouts initiated from inside self-hosted deployments
- **PostHog:** product analytics, telemetry, and session recordings (website and cloud only; session recordings are not collected from self-hosted deployments)
- **Cloud infrastructure providers:** for hosting the cloud deployment and supporting services
- **Email and communications providers:** for transactional and product update emails

We may also disclose information when required by law, to protect our rights or those of our users, or in connection with a merger, acquisition, or sale of assets.

## 7. DATA RETENTION

- **Self-hosted product telemetry:** Retained in aggregate; identifiable session-level events deleted within 12 months.
- **Service Ping data:** Retained for the lifetime of the deployment plus 24 months, to support historical trend analysis and customer success.
- **Trial and subscription records:** Retained for the duration of your relationship with us and for a reasonable period afterward to comply with legal and tax obligations.
- **Cloud accounts:** Retained while your account is active. You can delete your account at any time; we will delete your personal data within 30 days, except where retention is required by law.
- **Website analytics:** Retained for up to 24 months.

## 8. YOUR RIGHTS

Depending on where you live, you may have rights to access, correct, delete, export, or restrict our processing of your personal data, and to object to processing or withdraw consent. To exercise any of these rights, email [legal@sourcebot.dev](mailto:legal@sourcebot.dev).

## **9. INTERNATIONAL TRANSFERS**

We are based in the United States, and the information we collect is processed there. If you access Sourcebot from outside the US, you consent to your information being transferred to and processed in the US, subject to appropriate safeguards where required.

## **10. CHILDREN**

Sourcebot is not intended for anyone under 16. We do not knowingly collect personal information from children.

## **11. CHANGES**

We may update this policy from time to time. When we do, we'll update the "Version" date above, and for material changes we'll notify you by email or in-product notice.

## **12. CONTACT**

Taqla Inc.

Email: [legal@sourcebot.dev](mailto:legal@sourcebot.dev)