

Limits

From the SourceLace manual: How far SourceLace goes for your organization: how much work the AI may do for one question, how long agents may run, how long people stay signed in, how long queries may take and how large a file SourceLace reads. Admins see every limit on the Limits page and can make so

SourceLace version 0.1.0 · Made October 7, 2026

The latest version is at <https://sourcelace.onrender.com/docs>

Limits

How far SourceLace goes for your organization: how much work the AI may do for one question, how long agents may run, how long people stay signed in, how long queries may take and how large a file SourceLace reads. Admins see every limit on the **Limits** page and can make some of them stricter.

Who sets a limit

Your plan sets each limit, and your SourceLace contact can set a different value for your organization (for example for a pilot). That value is what **SourceLace allows** on the **Limits** page.

For the limits with a box in the **Your value** column, an admin can choose a stricter (lower) value, for example shorter sign-in sessions for your security policy, or fewer agent runs at once. You can never choose more than SourceLace allows; ask your SourceLace contact for more. Leave a box empty to use what SourceLace allows. Changes apply within a minute, and a shorter sign-in session also ends sessions that are already older than it.

Some limits are set by SourceLace only, such as the longest answer the AI may write. The page shows them so you know them.

Two more limits come with your plan or trial but are not on the **Limits** page:

- **Changes to records.** A plan or trial without them refuses every change (SourceLace only reads), whatever an admin has allowed on **Manage sources**. Trials do not include them by default.
- **Sources in all.** A plan or trial can cap the total number of sources, whatever their connector class (a trial allows 2 by default). At that number, adding a source is refused until one is removed.

Every limit

AI in the app

Model calls per question

How many times the AI may call the model to answer one chat question (looking up schemas, running queries, joining). More answers harder questions and costs more. An admin can choose a lower value.

Output tokens per model call

The longest answer one model call may write.

Lookups at the same time per question

When the AI needs several things at once to answer a question, such as the fields of three objects or a query in two systems, it looks them up at the same time instead of one after another, so answers come sooner. This is how many lookups run at once. Changes and documents always wait until the lookups are done and then happen one at a time, in order. If one lookup fails, the others still finish. An admin can choose a lower value; 1 does every lookup in turn.

AI agents

Steps per agent run

The most tool calls one agent run may make. Each agent's own limit must be at most this. An admin can choose a lower value.

Credits per agent run

The most credits (AI use plus steps) one agent run may use. Each agent's own limit must be at most this. An admin can choose a lower value.

Time per agent run

How long one agent run may take before it is stopped. An admin can choose a lower value.

Chat messages per agent run

How many Slack or Teams messages one agent run may post or propose. An admin can choose a lower value.

Agent runs at the same time

How many of the organization's agent runs may run at once, across all servers, so one organization cannot take every place. Others wait for the next turn. An admin can choose a lower value.

Days a change waits for approval

How long a change an agent proposed waits for someone to approve it before it expires. An admin can choose a lower value.

Reviewed shadow runs before going live

How many runs of an agent in shadow mode people must review before it may go live. Here more is stricter: an admin can choose a higher value, never a lower one.

Days in shadow before going live

How many days an agent must stay in shadow mode before it may go live, on top of its reviewed runs. 0 means no minimum. Here more is stricter: an admin can choose a higher value, never a lower one. Admins can also set this on **Agent settings**; the higher of the two counts. See [Days in shadow and approved changes](#).

Agent lookups at the same time

Like **Lookups at the same time per question**, for agents: how many of one run's lookups run at once, so runs finish sooner. Record changes, email drafts and chat messages always wait until the lookups are done and then happen one at a time, in order. Each lookup still counts as one step, and a run never makes more steps than its limit. An admin can choose a lower value.

Credits

Credits per agent step

Credits each tool call of an agent run costs, on top of its AI use.

Credits per question

How many credits make one "question" when credits are shown to people.

Warn at this share of included AI use

Admins see a notice when the month's included AI use reaches this share (0.8 is 80%). An admin can choose a lower value.

Warn at this share of the extra maximum

Admins see a notice when extra AI use reaches this share of their monthly maximum. An admin can choose a lower value.

Warn at this share of seats

Admins see a notice when this share of the plan's maximum seats is used. An admin can choose a lower value.

Sign-in

Web app session

How long someone stays signed in to the SourceLace web app before signing in again. An admin can choose a lower value.

AI app sign-in (refresh token)

How long an AI app (MCP) stays signed in to SourceLace without the person signing in again. An admin can choose a lower value.

AI app access token

How long one access token for an AI app lasts before the app renews it. An admin can choose a lower value.

Queries and files

Query time

How long one query to a database or data warehouse (SQL databases, MongoDB, Snowflake, BigQuery) may run before SourceLace stops it. An admin can choose a lower value.

Largest file read

The largest file SourceLace downloads to read (Google Drive, SharePoint, OneDrive, S3, Slack, Teams). An admin can choose a lower value.

Most bytes one lake query reads

How much one query on a [lake source](#) may read from storage, data and metadata together. A query that would read more is stopped. An admin can choose a lower value.

Most files one lake query reads

How many files (data and metadata) one query on a lake source may open. An admin can choose a lower value.

Health record results kept

How long SourceLace holds a result read from a health record (FHIR) source before deleting it: 5 minutes unless set otherwise, at most 30. These results hold protected health information, so they are kept for less time than other results; your [data retention](#) settings can shorten it further, never lengthen it. An admin can choose a lower value. See [Health records](#).

Asking across systems

How far one question over several systems may go ([Asking across systems](#), Preview). An admin can choose lower values for each of these except the monthly rows.

Tables in one question

How many tables (objects) one question across systems may read. 8 by default.

Systems asked at once

How many of one question's tables SourceLace reads at the same time (6 by default), and how many reads run at once against one system for your whole organization (2 by default), so one busy question cannot flood a system.

Rows per table and per question

The most rows one table may return for a question (2 million by default) and the most from all tables together (5 million). A question that needs more is stopped with the count from each system, never cut short without saying so.

Rows fetched without asking first

Above this many rows (250,000 by default), SourceLace shows the plan and its cost and fetches only once you confirm.

Time per question

How long one question across systems may take, fetching and computing together (5 minutes by default). If it runs out, SourceLace says which system was still answering and how far each had got.

Rows per month

How many rows your organization may fetch for questions across systems each month, set by your plan. Credits for them are shown before a large question runs.

Masked values per conversation

How many different masked values (the placeholders data protection puts in place of real values) one conversation may hold, 100,000 by default. A question that would need more is stopped and says so: select fewer masked fields, narrow the question, count instead of listing rows, or start a new conversation.

Organization

Access groups

How many access groups the organization may have.

Skills

How many skills the organization may have, and each person, and each project.

On-premise agents

How many SourceLace Agents (for on-premise systems such as SAP ECC) the organization may register.