Measuring your agents

From the SourceLace manual: Agents → Performance shows what your agents finished and how long the whole journey and each stage took, so you can see where to improve. This chapter defines every figure in plain words: what it counts, over what window, how to read it, what good looks like and what to d

SourceLace version 0.1.0 · Made November 22, 2026
The latest version is at https://sourcelace.ai/docs

Measuring your agents

Agents → Performance shows what your agents finished and how long the whole journey and each stage took, so you can see where to improve. This chapter defines every figure in plain words: what it counts, over what window, how to read it, what good looks like and what to do about it. SourceLace's own people read the same definitions, so a number means the same thing to everyone.

What is measured

Everything on the page is worked out **once a day, shortly after midnight (UTC), from the runs that finished the day before**, the approvals decided that day and the deliveries that were sent or failed that day. It never reads your data: SourceLace keeps counts, kinds (which tool, which connection, which error type) and durations, and nothing else. A run's own page shows its timeline as soon as it ends; the Performance page catches up with it the next day.

- **Period**: the last 7, 30 or 90 days up to the latest day worked out. A run counts on the day it **finished** (its whole journey is only known then), an approval on the day it was **decided**.
- **Who sees what**: admins see the whole organization and every agent; an agent's owner (and the co-owners of a shared agent) see that agent's own view from its page. Nothing leaves your organization.
- **Dry runs and test runs are left out**: they change nothing, so they would make the figures look better than they are.
- **Times read as "up to"**: durations are counted into fixed time bands (5 s, 15 s, 30 s, 1 min, 2 min, 5 min, 10 min, 15 min, 30 min, 1 hour, 4 hours, 1 day, 3 days, 7 days, over 7 days), so a median shows as "up to 5 min": half the runs took that long or less. The **90th percentile** is the band nine in ten runs fit in. This keeps the page fast and keeps no per-run detail beyond the run itself.

Meaningful work

Figure	What it counts	How to read it
Runs finished	Runs that ended in the period: completed, finished without a change, failed, stopped or skipped	The volume of work attempted
Runs completed	Runs that ended normally and whose proposals, if any, were not all turned down (outcomes)	The work that landed
Success rate	Completed out of finished. Shown from 3 finished runs (an admin can change this on the Limits page), so one run never reads as 0% or 100%	Above 90% is healthy for a mature agent; a new agent often starts lower while its instructions settle
Changes applied	Proposals a person approved, plus changes, drafts and messages the agent was allowed to make itself	The changes that actually happened in your systems

Figure	What it counts	How to read it
Changes proposed, rejected, expired, stale	The runs' proposals by what became of them. **Rejected**: a person said no. **Expired**: nobody decided before the approval window ended. **Stale**: approved, but the record had changed meanwhile, so nothing was written	Many rejected or stale proposals mean the agent proposes the wrong thing, or too late
Records changed	Records the runs proposed or changed, as each run counts them (a proposal counts even if it was not approved)	Compare with changes applied to see how much proposed work landed
Outputs delivered	Results chats, files, email drafts and chat messages the runs made, plus deliveries that were sent (outputs)	What people received
Credits	The AI credits the finished runs used, and per completed run and per change (credits)	The cost of the work

Outcomes of a run

Every finished run has exactly one outcome:

Outcome	Meaning	What to do
Completed	The run ended normally. If it proposed changes, at least one was approved, or it made changes it was allowed to make itself	Nothing: this is the work you asked for
Finished, change not made	The run ended normally and proposed changes, but every one was rejected, expired or stale	Look at what it proposes: the instructions, the fields, or the approval window
Failed	Something went wrong: a system refused, a sign-in expired, the AI was unavailable, or SourceLace restarted twice at the same step	Open the run: its error says why. Most failures are a connection to fix
Stopped	A stop switch, or the run's step, time or credit limit	Raise the agent's limit, or narrow its task
Skipped	A shared agent's run for one person could not start because of that person's own setup (for example a system they have not connected)	That person connects the system

Runs still queued, running or waiting for approval have no outcome yet and are not counted until they end. The **Agents → Overview** and **Approvals** pages show what is pending now.

The journey

A run's journey runs from the moment it was queued to the moment it finished, in three stages:

Stage	From	To	What makes it long
Queue wait	The run was queued	Its work began	Your organization's runs at the same time were all taken, or a system or the AI was busy and the run was put back to try later
Working			Time SourceLace spent running the agent: model turns and steps. Restarts and continuations after approvals add up; waiting is left out

Stage	From	To	What makes it long
Waiting for people or agents	The run paused	It carried on	People deciding its proposals, or other agents it asked. A run holds nothing while it waits
Whole journey	Queued	Finished	All three together

For each stage the page shows the **median** and the **90th percentile** over the period's runs, as "up to" bands, and the stage's **share of the whole time**. The bar draws the three shares; the table beside it has the same numbers.

What good looks like: working time in minutes, not tens of minutes; waiting that matches how fast your people decide (hours, not days); queue wait near zero. When waiting is most of the journey, the agent is fine and the process around it is slow: approve sooner, let the agent make routine changes itself, or shorten the approval window (**Days a change waits for approval** on the Limits page). When the queue wait is large, raise **Agent runs at the same time** or spread schedules.

Steps

A step is one tool call in a run: a lookup, a change, a draft, a document. The page lists every kind of step (the tool, and the connection it used), with:

- **Calls**, and how many were **OK**, **refused** (SourceLace or the system said no: access, a rule, a write that is off) or **failed** (an error);
- the **success** share, and the **median** and **90th percentile** time of a call (in bands from 0.25 s to over 2 minutes), which includes the system's own answer time;
- the **most common error types** (the type only, such as `AccessDenied`, never the message).

A step that fails often points at access or a connection; a slow step points at a wide query. The **Limits** page's agent limits say how many steps a run may take.

Approvals

Counted on the day they were decided:

Figure	Meaning
Decided	Approvals that reached an answer: approved, rejected, expired, stale or failed to apply
Approval rate	Approved out of decided
Expiry rate	Expired (nobody decided in time) out of decided
Stale	Approved, but the record had changed, so nothing was written
Time to decide, median and 90th percentile	From the moment the agent proposed the change to the decision (an expired approval counts the whole window). In bands from 1 minute to over 7 days

Because approvals count on their decision day and runs on their finish day, the approvals here can differ from the proposals counted under Meaningful work: a change proposed on Monday and approved on Wednesday is in Wednesday's approvals and in the run's own day once the run ends.

Outputs

What the runs produced: **results chats** (the run's summary and files in its results project), **files** (documents and spreadsheets), **email drafts** (never sent by SourceLace: a person reviews and sends them), **chat messages**, and **deliveries**: files attached to a draft or posted to a channel after a run, and changes a system asked SourceLace to try again later. Deliveries are counted on the day they were **sent** or **failed**, by the kind of destination (such as Slack or Google Workspace). A failed delivery is kept and can be retried from the run's page.

Credits

Credits per completed run is the AI credits of all finished runs divided by the runs that completed; **credits per change** divides by the changes applied. They are the price of a unit of work. A rising figure usually means more steps per run or bigger answers; compare agents on the same task to see which instructions are cheaper. The Usage page shows credits by agent and by day.

Trend

Runs finished and completed per day (for 7 and 30 days) or per week (for 90 days), with each day's changes applied, outputs delivered, credits and the median whole journey. **Show as table** gives the numbers. A week is the seven days ending on the date shown.

Where to improve

Plain-English hints worked out from fixed rules on the period's figures. They are advice, not alarms, and an admin can tune when each fires on the **Limits** page (the **Performance and hints** section under AI agents):

Hint	Rule (default)
Most of the time is spent waiting for people	Waiting is at least 50% of the whole journey's time
Runs wait long before they start	Queue wait is at least 30% of the whole journey's time
Only N% of runs completed	Success rate under 80%, with at least 10 finished runs
Many runs end with no change made	At least 30% of the runs that ended normally had every proposal rejected, expired or stale, with at least 10 such runs
Approvals often expire	At least 20% of decided approvals expired, with at least 5 decided
Approved changes are often stale	At least 10% of decided approvals were stale
A step fails often	A kind of step refused or failed at least 20% of its calls, with at least 10 calls (the worst one is named, with its most common error type)
Deliveries often do not arrive	At least 10% of finished deliveries failed, with at least 5
A step is slow	One in ten calls of a kind of step took longer than 60 seconds, with at least 10 calls

With fewer than 3 finished runs the page says it is too early; with nothing to point at it says nothing stands out. At most five hints show, the most important first.

Agents table and export

For the whole organization (admins), the **Agents** table lists every agent that finished a run in the period, most runs first: runs, completed, success rate, changes applied, outputs, records and credits. Open an agent for its own journey, steps, approvals and hints. An agent deleted since is shown as "Deleted agent" without a link; its past work still counts in the organization's figures.

Download CSV gives the same figures as a spreadsheet file, one block per section, for the period and agent shown. Each download is in your audit log.

A run's timeline

Every run's page has a **Timeline**: queued, started (with the queue wait), each step with how long it took and when it began, any wait for people's decisions or for agents it asked, and the end, with the whole journey, queue, working and waiting totals. It comes from the run itself, so it is there the moment the run ends. Runs that ended before SourceLace kept timelines (before 24 October 2026) show their steps' durations only.

What good looks like

- A **success rate above 90%**, with the rest explained by a few failed runs you understand.
- **Changes applied close to changes proposed**: people agree with what the agent proposes.
- **Expiry rate near zero**: approvals get decided; nobody is the bottleneck.
- A **journey where working time dominates**, and waiting matches how fast your team decides.
- **Steps that succeed above 95%**, and no step whose slowest tenth takes over a minute.
- **Credits per change that hold steady** as the agent runs more.

What is not measured

- **The value of the work** (money or hours saved): SourceLace counts the work; you put a value on it.
- **Time inside your systems** apart from the step that called them, or the AI's thinking time apart from working time.
- **Why a change was rejected, or who decided**: only the whole time to decide. The Approvals page and the audit log have the detail.
- **Runs that have not finished**: they appear once they end.
- **The waiting time of runs from before 24 October 2026**, which the page marks as estimates.
- **Rows read**: shown per step on the run's page, never summed.