

Understand approvals, guardrails, and run history

Explain how Vibework constrains sensitive actions, asks for human approval, records execution, and surfaces governance issues.

Vibework uses permissions, guardrails, approvals, and history to keep AI and automation inside organizational boundaries. These controls solve different problems: permissions decide who may access something, guardrails evaluate whether an action is allowed, approvals pause consequential work for a person, and history records what occurred.

Approvals

An approval is a recorded human decision on proposed work. A workflow or tool action may pause and show the reviewer the action, context, inputs, or expected effect. The reviewer should verify the target, scope, recipients, data, and consequence before approving. Reject or return the work when information is missing or the action is unsafe; approval should not be a routine click-through.

Guardrails

A guardrail is a rule or check that can allow, block, flag, or require review of an action. Guardrails may reflect data sensitivity, operation risk, model policy, budgets, external communication, or organization-specific requirements. A block usually means the request or configuration must change; repeatedly retrying the same action does not remove the policy.

Run and audit history

- Workflow run history records the workflow, run, steps, status, timestamps, inputs or references, outputs, and visible errors.
- Approval history records who decided, when they decided, and the resulting state.
- Integration or tool activity may record the operation and outcome without exposing the underlying credential.
- History supports troubleshooting and accountability, but retention and visibility depend on organization policy and role.

How users should respond

1. Read the full approval or governance message instead of relying on the button label alone.

2. Confirm that the action matches the original business request and uses the intended organization, resource, and connection.
3. Reject, revise, or escalate when the impact is unclear, access seems excessive, or sensitive data is unnecessary.
4. Use the run ID, decision record, timestamp, and visible policy message when asking an administrator or support for help.