

Anthropic opens a verification programme giving vetted biologists looser safeguards

Source copy kept by wheresthe.ai for the story "Anthropic opens a verification programme giving vetted biologists looser safeguards" (<https://wheresthe.ai/d/anthropic-life-sciences-verification-program-2026-09/>).

Source: <https://www.anthropic.com/news/life-sciences-verification-program>

Text saved 18 Oct 2026 (the page as published).

SHA-256 of the saved text: 4c93386633a52360ddcc9a701ed35e57bbb21e4962c58d432d6bddcca73d9b5a

Announcements Introducing the Life Sciences Verification Program Sep 17, 2026

Today, we are introducing the Life Sciences Verification Program (LSVP), which gives life science professionals access to our Mythos, Opus, and Sonnet models with a refined set of safeguards more permissive for biology-related work. We have already onboarded dozens of organizations through an early-access program, and are now opening applications to the broader life science community ([join waitlist here](#)). The program is launching in beta, initially for teams and institutions. We will continue to improve the program and expand access to individual Pro and Max plans over time. The LSVP is designed to enable life science professionals to use our models across a wide range of tasks that are currently blocked in our generally available Fable models, like drug discovery, research biology, clinical development, and manufacturing. It's built for teams of all kinds—from academic labs to startups, pharma companies, and more. Verification and access types To qualify for these grants, each applicant goes through a verification process that includes a review of their research credentials, security standards, and ethical research oversight. Once verified, teams may apply for two types of LSVP grants, "Standard Use" or "High-risk Use," depending on their access needs. These grants can be used through all our product surfaces, including Claude Science , Claude.ai , Claude Code and the API.

Standard Use grants are suitable for most life science work, including the majority of biology research and development workflows. These grants can be extended to entire teams for diverse, daily workloads, and are renewed once a year. They give those teams access to our Mythos, Opus, and Sonnet models, with refined classifiers that are more permissive for science tasks than our generally available models. Standard Use grants apply to Mythos 5.1, Opus 5, and Sonnet 5 today, and to future models as they launch. They're specifically designed to enable the full breadth of life science activities in areas spanning basic science, R&D, supply chain and manufacturing, clinical development, quality assurance, regulatory affairs, investing and diligence, and more. Although we expect Standard Use to cover the majority of access needs, some work carries a higher potential for misuse and therefore requires additional vetting. High-risk Use is an add-on grant for teams working in areas blocked under Standard Use. It removes all safeguards that block life sciences requests. This grant applies to a single research project as opposed to a full team, and must be renewed every six months. Typically, a single researcher with dual-use work would have access to one Standard Use grant for diverse, daily activities, and one or more High-risk Use grants which only apply to work on specific projects (for example, characterizing how one specific family of viral vectors is recognized by human immune pathways). High-risk grants for Claude Opus 5 and Claude Sonnet 5 are available today. We are working with the US government to make high-risk grants more broadly available for Claude Mythos, but at the time of this launch they will remain limited to a small set of entities with additional vetting. All other safeguards, such as cyber classifiers, will remain in place under LSVP grants. Enabling trusted access through shared responsibility As we've shown in our recent threat report , there are increasingly sophisticated misuse attempts happening on our platform, including attempts that could support biological weapons development. In biology, where it's often not possible to differentiate between a user doing valid work (e.g. research a viral pathogen to develop vaccines against it) and pursuing harm (e.g. trying to increase the transmissibility of a virus maliciously), the most concerning threat models are ones where valid access has been diverted or overtaken by an actor with bad intent. Indeed, insider threats and rogue-use have been major factors in significant biosafety incidents and scares. In developing the LSVP's safeguards, we aimed to protect against three concerning threat models in particular: Access compromise: Malware or account takeover diverting access to a bad actor Insider threats: Rogue or coerced employees intentionally taking malicious action or diverting their access to a bad actor Agent misuse: Agents, especially working in swarms or over long-horizon tasks, taking unintended dangerous actions In order to defend against these threats and in close collaboration with enterprise CISOs, we designed the new LSVP safeguards around the concept of shared responsibility by monitoring usage against the intended use-case for the model access. Because we vet the LSVP organizations

for their life sciences credibility and oversight, we can empower them to specify for themselves what constitutes safe usage for teams or projects within their program. Each entity's access is tied to the use cases it has specified in its grant applications, and we continuously monitor LSVP traffic to identify usage or patterns that are outside the stated safe scope. Should unauthorized activity occur, we can flag these cases to organization admins to take action within pre-agreed timeframes for triaging and remediating incidents. The use cases should include high-level descriptions of the intended work, like one would share in a job listing, and not include any sensitive information or IP. How monitoring works in LSVP Serious misuse is often spread across many requests and sessions to look disconnected and evade detection. In the LSVP, we are shifting safeguards from real-time blocking, where we reject potentially harmful access at the time of each request, to offline monitoring, which allows us to more clearly identify potential misuse across patterns of behavior. Shifting enforcement from real-time blocking to offline monitoring allows legitimate work to proceed with fewer interruptions, but it requires us to retain data associated with flagged activity for review. For LSVP traffic, we are requiring data retention for 30 days to be able to do this monitoring effectively. This data is strictly compartmentalized and cannot be used for model training or accessed by members of Anthropic's life sciences research teams. For organizations that qualify, we are also working to understand how LSVP can integrate with features from our Enterprise Frontier Safeguards (EFS) systems. What researchers are saying

Xaira is making biology more computable, generating biological data at unprecedented scale and building foundation models of cell, protein and disease biology that turn it into the next generation of life-changing medicines. We're excited to put Anthropic's frontier models to work across our drug discovery engine, and we believe pairing trusted access with intelligence is the right way to realize AI's promise in biology.

Edison's mission is to accelerate science and the discovery and development of new medicines. With the Life Sciences Verification Program, we are excited to be able to bring Anthropic's most intelligent models to bear on these problems. We look forward to collaborating with Anthropic further to end disease and improve the lives of patients everywhere.

At Manifold Bio, we're building a massively parallel interface into living systems to enable powerful AI to create medicines. We look forward to putting frontier intelligence to work safely in our engine, and we welcome Anthropic's approach of pairing access with accountability.

01 / 03

Applications and availability Organizations interested in joining the LSVP can register interest here . We expect to enroll hundreds of organizations within the first week, and to scale the program further to support the majority of the life science community in the coming weeks. Today, LSVP is available in our first-party console for API usage, as well as in Claude for Enterprise and Team plans. We do not yet support individual plans but are working to expand access for these users. It is also not yet available on third-party platforms. As a beta, LSVP is not available for BAA-enabled orgs. This means customers with PHI data should use separate non-BAA orgs with non-HIPAA. In API and Claude Science, users can switch between grants natively. In Claude.ai and Claude Code, initially only a preselected default grant applies (except while using Claude Code with API authentication). This should be fine for the vast majority of users, who will only ever require a Standard Use grant. However, we will improve support and portability of these LSVP features over time. What comes next Providing these frontier capabilities is part of our broader efforts in supporting the life sciences community in our shared mission to accelerate curing disease and improving human health. We will share more about new products, research collaborations, and improvements to the program in the coming months.

Changelog: September 30th, 2026. We clarified that the form we link to is an expression of interest, not an application.

Related content

An update on independent evaluators To evaluate the safety of our systems, Anthropic is giving employee-like access to a team from METR, a nonprofit that evaluates frontier AI models, and we're in early discussions with several other organizations, including Hugging Face. Read more Preparing for the 2026 US midterms Earlier this year, we shared an update on our election safeguards. Here, we outline how we've continued to strengthen these safeguards ahead of this year's US midterms. Read more Anthropic opens Singapore office and signs MOU with the Government of Singapore Anthropic has opened an office in Singapore and signed an MOU with the country's government to advance safe, responsible AI adoption. Read more