

How CERN Earns Trust at Scale

Understanding how one of the world's most enduring and successful large-scale science collaborations has made its work credible can help Europe build trustworthy infrastructure for artificial intelligence in science.

More than 300 feet beneath the mountain foothills, farmland, and quiet border towns between Switzerland and France, superconducting magnets in a 17-mile ring are kept at a temperature below –450 degrees Fahrenheit. Inside the ring, two beams of protons accelerate in opposite directions to nearly the speed of light before colliding head-on at four different points along the path. This machine is the Large Hadron Collider, the most powerful particle accelerator in the world, designed to recreate the universe's earliest moments and help answer fundamental questions about how it all began.

Along the beams' route, in caverns built at the collision points, are four particle detectors: ATLAS, CMS, ALICE, and LHCb. Each is its own separate experiment, designed and run by a different international collaboration. It was two of these, ATLAS and CMS, that independently confirmed the discovery of the Higgs boson in 2012. The scale of these collaborations is hard to grasp—in 2025, the Breakthrough Prize in Fundamental Physics recognized 13,508 coauthors from more than 70 countries for a decade of results across all four experiments.

While the particles race underground, control rooms above monitor the machines and collisions, coordinating the efforts of thousands of researchers in real time. These rooms, and the shared vigilance and collaboration they make possible, are fundamental to

the work done at the institution behind all this: CERN, the European Organization for Nuclear Research. A citadel for physicists and a striking example of cooperative scientific governance, CERN has famously produced radical discoveries in subatomic physics along with spillovers like the World Wide Web.

Despite operating as a coalition of 25 member states, with 11 more holding associate status, CERN is known for its stability, even in the face of global power struggles. So it's no surprise that as Europe has grappled with strategizing around artificial intelligence, one guiding metaphor for a "European way" has been "CERN for AI."

Europe's third way

Europe's push to develop its own approach to AI governance arrives as the technology has become an explicit front in the geopolitics of scientific research. In initiatives such as the United States' Genesis Mission, launched in November 2025, AI in science is framed explicitly as an issue of competitiveness and sovereignty, a way to secure scientific leadership, innovation capacity, and strategic autonomy in an era of intensifying US–China AI rivalry. Across these strategies, the central assumption is clear: Whoever controls the foundations of AI, including compute, data, and frontier models, will increasingly shape the pace and direction of discovery—and the future of science itself.



Europe has been a leader in AI-related scientific output, but it has weaker access to large-scale compute, foundational models, and platform ecosystems. In October 2025, the European Commission adopted the European Strategy for Artificial Intelligence in Science as part of its bid to ensure leadership and competitiveness. But rather than embracing an all-out race for speed and scale, the strategy anchors Europe’s approach to AI development in public values, accountability, and trustworthiness: “A European approach to a safe, sustainable, human-centric and trustworthy AI in science is a strategic opportunity in the rapidly changing geopolitical environment,” it reads.

The idea of a “CERN for AI” dates back to 2017; since then, it has circulated as a strategic metaphor, claimed by different actors to mean very different things. In February 2025, the European Commission president invoked CERN to describe Europe’s ambition to build a coordinated, public infrastructure for AI. That grew into the flagship initiative of the Strategy for AI in Science: RAISE, or the Resource for Artificial Intelligence Science in Europe, with an initial €107 million in funding from Horizon Europe.

Like CERN, RAISE is transnational and designed to do what no single nation can achieve alone. But unlike CERN, it will not take shape as a single monumental site. Instead, it is being built as a virtual institute with a distributed network, weaving together Europe’s supercomputers, data hubs, and AI factories into a shared infrastructure. In the Commission’s framing, its aim is to support Europe’s ambition to become an “AI Continent” while remaining aligned with European commitments to ethics, accountability, safety, and human rights.

RAISE and related AI-for-science initiatives are being framed in a geopolitical register, but in practice they are also institutional reforms of how science is structured and performed. Embedding AI into research infrastructures reshapes the everyday craft of science: how problems are framed, how evidence is produced, how results are validated, and how uncertainty is handled. As these new infrastructures solidify, so too do the norms and forms of authority that will govern science in the years to come, often before those arrangements are made explicit or subject to debate.

Now, while these choices are still open, is the ideal moment to really think through what it means to build a distributed AI infrastructure for science, and what must be built into that infrastructure to ensure it reinforces Europe’s stated commitment to trustworthy AI. I come to this question as an anthropologist and science, technology, and society scholar studying the future of universities and academic work, and how artificial intelligence is now reshaping the whole knowledge-making ecosystem. I spent nearly a decade visiting CERN as part of the Challenge Based Innovation program, working with student teams

tackling big societal challenges, inspired by CERN’s science and the people behind it. Seeing CERN up close, I kept returning to a question with colleagues there: What would it mean to bring not just CERN’s technology, but its ethos and methods, to bear on challenges at that scale?

Many “CERN for AI” discussions spotlight CERN’s governance and institutional design. My focus here is different: the everyday practices and shared standards through which CERN makes trust—in its machinery, and in the knowledge it produces—possible. CERN doesn’t offer a blueprint for AI, but it exemplifies elements of scientific infrastructure that are essential yet easily overlooked, including designed visibility of complex systems, shared practices for collective judgment, and careful methods for handling uncertainty.

How trust is built in practice

It is tempting to think of CERN only as a triumph of scale and engineering: a vast complex of accelerators, detectors, and tunnels built to probe the fundamental structure of matter. This image often stands in for “Big Science” more generally, reinforcing the idea that CERN’s primary lesson lies in technical ambition or concentrated resources. But this reading misses what has made CERN historically distinctive and politically durable. From its founding in 1954, in the aftermath of the Second World War, CERN was conceived not simply as a laboratory, but as a political infrastructure for rebuilding European science (and Europe) through cooperation. It was designed to pool resources and research capacity across the borders of its original 12 member states after decades of fragmentation and scientific brain drain. Its founding convention embedded a public, nonmilitary ethos from the outset, explicitly committing the organization to “have no concern with work for military requirements.”

CERN’s governance was treaty-based and intergovernmental, tying scientific collaboration to the shared funding commitments and collective decisionmaking of the CERN Council, through which member states vote to decide on scientific and administrative issues. Precisely because CERN made large-scale collaboration functional, productive, and credible, it has also become one of the most closely studied scientific infrastructures, offering unusually rich accounts of how trust is built in practice.

Literal visibility

From early on, CERN treated visibility as a central institutional task, both externally for the public and internally for its own staff and researchers. The 1953 convention establishing CERN stipulated that the results of its work “shall be published or otherwise made generally available.” At first, this constitutional openness took the form of making its knowledge visible—publishing in journals, exchanging results, collaborating across borders.

But CERN also built visibility into its administrative structure.

In 1956, a committee of the CERN Council flagged nonscientific communications and public relations activities as work “whose importance must not be underestimated,” placing it under the direct authority of CERN’s director-general. That commitment has only grown more deliberate since, and CERN has adopted the concept of transparency in a surprisingly literal way. This is no small ambition for an organization that, at first glance, has very little to see. First-time visitors often comment on the ordinariness of CERN’s core buildings—functional, plain, like an old university campus. So the institution has taken deliberate steps to engage the public and signal that some of the most complex scientific instruments ever built are running beneath visitors’ feet. For example, the “Passport to the Big Bang” bike path follows the Large Hadron Collider’s route aboveground past 10 exhibition platforms, each marking a point of interest over the machine below. And on the outside of the ATLAS building, a mural renders the detector visible, at a third of its real size, enabling visitors to marvel at the complexity of its 100 million individual sensing elements.

with shift leaders and technical staff seated at the center to speed communication, give everyone equal visibility across the stations, and let visitors watch what’s happening from the edges. The CERN Control Centre, which oversees the whole accelerator complex rather than a single experiment, has four circular islands, each dedicated to a different part of the chain—a shape that itself resulted from consolidating what were once separate, machine-specific control rooms.

Shared practices of skepticism and care

But visibility alone does not produce credibility. What a shift worker catches is only the first step in scientific judgment. Any potential finding must survive successive rounds of scrutiny, testing, and collective judgment before it can become a credible result. Different scientific communities have their own processes for this, and high-energy physics is notable for the extent to which these processes have been formalized and shared.

In her classic ethnography of high-energy physics, the anthropologist Sharon Traweek looked at the patterned practices through which the social world of high-energy physics is reproduced: how physicists present themselves, how

Europe’s push to develop its own approach to AI governance arrives as the technology has become an explicit front in the geopolitics of scientific research.

If initiatives like the mural and the bike path help make the collider visible for the public at large, CERN’s control rooms make visible the work of keeping it running and the knowledge flowing. For those working on collider experiments, onsite and around the world, control rooms are the shared spaces where researchers collectively watch experiments unfold, track anomalies, and coordinate responses in real time. Sometimes they make the decision to pause—or, in the physicists’ phrase, *to dump the beam*—and in rare cases they opt to bring the machine down altogether. The physical presence of human monitors is central to detector operations, and researchers and technicians work shifts around the clock for months at a time. Collaborators across career stages take turns working these shifts, even though remote monitoring is technically possible.

The layout of the control rooms enables the integration of interdisciplinary expertise necessary to turn collisions of particles into knowledge. ATLAS’s control room resembles NASA’s more familiar mission control, with rows of computer consoles facing large screens mounted on the walls. CMS was recently redesigned around a central circular workspace,

they talk about experiments and detectors, how newcomers learn the community’s codes, and how professional identities and hierarchies are enacted. Review committees, internal notes, and layers of approval slow down the process of claiming discovery, embedding restraint and skepticism into scientific life.

Together, these practices are integral to socializing researchers into common standards of skepticism and care. Discoveries are rarely attributed to individual brilliance. Instead, authority resides in collaborations, procedures, and collective thresholds of acceptance. Researchers refrain from publishing claims until they pass through multiple rounds of scrutiny. In the CMS experiment, for example, papers move through fixed waypoints designed so that every collaboration member who will ultimately sign the paper has the opportunity, and the obligation, to examine and improve it. Trust, in this context, emerges not only from confidence in the machines, but also from confidence in the social processes that govern how the outputs of those machines are interpreted—and, just as crucially, in how uncertainty itself is recognized and formally tracked.

Making doubt explicit

Managing uncertainty at CERN starts with the design of the instruments themselves. ATLAS and CMS, the Large Hadron Collider’s two general-purpose detectors, were built independently, so that a claim as consequential as the discovery of the Higgs boson would have to be confirmed twice by separate teams using separate equipment before it could count as a discovery. This reflects the basic fact that detectors never observe physical processes directly—they infer what happened from traces of energy and particle tracks, themselves shaped by calibration choices, reconstruction software, and models of expected background noise. Each step adds another layer of uncertainty.

Rather than eliminating that uncertainty, the discipline makes its boundaries explicit. A whole set of methods and practices has been developed just to track what remains uncertain even inside what physicists call the *fiducial region*, the specific, bounded slice of a collision the detector can actually capture. Uncertainty is broken down by source, and every result carries error bars illustrating statistical

or its consequences. Europe’s RAISE initiative is the first serious attempt to pool computing power, data, talent, and research funding into shared public infrastructure on a continental scale. To imagine how this initiative might succeed at this very difficult task, it is worth looking closely at how CERN institutionalized visibility, collective skepticism, and careful treatment of uncertainty, and then asking what it would take to replicate those practices in the context of AI infrastructure.

The first lesson is visibility, and in CERN’s case this meant surfacing a machine that sat buried deep underground, along with the people, decisions, and daily practices that keep it running. RAISE faces a more diffused version of the same problem—doubled. It exists both to advance AI itself as a field of research (“science for AI”) and to spread the use of AI as a tool across every other scientific discipline (“AI in science”), and each half of that mandate needs to be seen in its own way.

Important work is already underway on responsible AI in science—algorithmic transparency, openness, and the ethics of data collection—and calls for open models and explainable outputs have value. But audits and disclosure statements are

Embedding AI into research infrastructures reshapes the everyday craft of science: how problems are framed, how evidence is produced, how results are validated, and how uncertainty is handled.

and systematic uncertainties. Physicists also correct for the “look-elsewhere effect,” the fact that scanning enough data will always turn up something that looks like a signal by chance. In addition, many analyses are run “blind,” with the method locked-in before the real result is ever seen, so that researchers’ expectations cannot shape what they conclude they found. None of this makes uncertainty disappear. Instead, it makes uncertainty legible and shared, turning results into something the whole field can inspect, argue about, and ultimately trust.

CERN’s lessons for AI

AI has been deployed into the world with none of the deliberate, disciplined, highly social processes that are baked into particle physics. Today, a handful of companies are effectively running one of the largest uncontrolled experiments in history, and they’re doing it without disclosure; without any collective way to monitor what is being built where and at what cost; and without a way to assess either the credibility of the knowledge it generates

built to answer a narrower question: Can this particular system be inspected after the fact? They cannot supply a place where people watch together in something like real time—where they might notice that something has gone wrong as part of the work, rather than reconstructing events later.

This is where detectors and control rooms come in. Meeting RAISE’s double challenge means building a “sensing apparatus” of its own, a material, technical, social, and epistemic arrangement that makes knowledge and action possible. It would have to be built from pieces contributed by people who want to observe different things. Imagine the contributions of an engineer, a policymaker, an anthropologist, and a municipal water authority: Each would bring a different sensor, and the pieces would have to fit together. Control rooms would then be organized around what that apparatus registers, they would run around the clock, and they would be visible not only to the people who operate them but to everyone with a stake in what they do—from the researchers relying on RAISE’s infrastructure to the public being asked to trust it.

The second lesson is that control rooms will not be sufficient on their own. Processes are needed for contesting interpretation and for deciding when results are reliable enough. Science's authority has always depended on people repeating evaluations that demonstrate how claims are challenged, defended, and stabilized. This tradition started with early experiments performed in front of public audiences, and it continues in modern procedures of peer review and replication. As high-energy physics learned early on, such shared practices are central to how trust is produced at scale, embedding restraint, delay, and shared responsibility into the everyday life of research.

The third lesson concerns uncertainty. AI systems are probabilistic at their core, producing outputs shaped by statistical patterns and incomplete training data, yet presenting results in polished, authoritative language that undermines reasonable doubt. Getting things wrong is one thing, but being wrong *smoothly* is another. It produces what scholars have called an *illusion of understanding*, in which fluency stands in for explanation and users feel they grasp something they have not fully interrogated. In

Dump the beam

None of these lessons amount to a handbook for what detectors and control rooms for AI should look like, or what processes should govern them, or how the many uncertainties in AI should be tackled. That is precisely why these questions should be at the center of AI infrastructure conversations—at RAISE, the European Commission, and beyond. Working against fragmentation, toward the public good, and with AI's broader consequences in view, all must start from the recognition that AI infrastructure for science is not just machines but also practices, and the two are interconnected. Here it is worth remembering that Tim Berners-Lee conceived the World Wide Web at CERN in 1989 out of exactly such a need—helping a large number of collaborators, spread across universities and institutes around the world, share information and documents. What began as a tool for one community ended up changing the very way that humans communicate. Equally significant, in 1993 CERN relinquished all intellectual property rights to the web's code, ensuring that no single company or country could own it.

Europe's RAISE initiative is the first serious attempt to pool computing power, data, talent, and research funding into shared public infrastructure on a continental scale.

research settings, this creates a growing mismatch between the uncertainty AI systems inevitably contain, and the confidence with which their outputs, arriving without any indication of which kind of uncertainty produced them, circulate through scientific workflows. Doing science requires hard-won habits of skepticism that trace how a claim was produced, what assumptions it rests on, and what forms of uncertainty it carries.

Just as CERN institutionalized collective moments of pause, interrogation, and delay, incorporating AI in science will require taking deliberate actions to support procedural humility in the context of an industry and technology that have so far tended toward the opposite. This is where a public infrastructure like RAISE has leverage no individual laboratory, or firm, possesses: Thresholds and conventions can only be set collectively, not unilaterally by whoever happens to be building the fastest model. Amid headlines about AI agents slipping the bounds of their sandboxes, the work of building new sites and practices for collective monitoring, evaluation, and judgment becomes all the more urgent.

So far, AI is headed in the opposite direction. Built with an eye to a politics of scale that does not consider consequences, it has few of the characteristics—lawful, ethical, robust—that the European Commission specified as the marks of trustworthiness. There is much to learn from CERN's governance and institutional design, as well as the everyday practices through which its experiments are made visible and collectively monitored, claims are slowed down, uncertainty is explained, and trust is produced. Most of all, there is the capacity to stop. CERN pauses to recalibrate its detectors. It redesigns its control rooms when they no longer serve the people in them. It is honest about what it does not know, and when the evidence demands it, CERN dumps the beam. That is a capacity missing from current AI infrastructures. Europe has the chance to build it, but not through a virtual institute alone.

Kyriaki Papageorgiou is a senior researcher at the Center for Science and Society, Department of Interdisciplinary Studies of Culture, Norwegian University of Science and Technology.