The Other Side of Algorithms

Reclaiming the Body, Voice and Labor

Nicolas Gourault in conversation with Juntao Yang

Nicolas Gourault's The Eyes (2025) traces the global infrastructure of data annotation—the labor that teaches machines to see. Like the annotation process itself, the film gradually illuminates a world of digital labor, following workers who segment images for self-driving cars and content moderation systems, their voices and screens composing a portrait of an industry designed to remain invisible.

This conversation took place remotely in December 2025, following a screening organized by Jilu Commune, curated by Cissy Ye and coordinated by Han Zhang. It has been edited for clarity and length.

— Juntao Yang

Originally published in Chinese at Jilu Commune, 2025. English version by the author.

Juntao Yang

I want to start with the opening of the film. You chose to open with those data annotation interfaces—bounding boxes and labels appearing on every element of everyday images. It's a powerful visual choice. On one hand, the entire purpose of this labor is to make things visible and legible for machines and algorithms. But on the other hand, the workers doing this labor remain almost invisible to us. How did you develop this visual language? Was this tension between visibility and invisibility something you were thinking about from the beginning, or did it emerge through the process of making this film?

Nicolas Gourault

Thank you for the question. It was actually very important for me from the beginning to fuse the visual language of the film with its subject matter. The research for The Eyes grew out of an ongoing thread of investigation. Before this, I made a film called Vow, which was preliminary research on the development of self-driving cars. What caught my attention was the parametric process—how the software recognizes its surroundings and distinguishes different objects. There was an incident where a pedestrian was struck because the vehicle failed to detect them—the person wasn't walking on the crosswalk. This puzzled me deeply and pushed me to dig further: How is the machine's classification system trained in the first place? Who decides these categories? And how are these abstract categories implemented into concrete software systems? That was the starting point for The Eyes.

Through reading research and speaking with researchers, I gradually discovered a vast yet highly invisible industry: data labeling. It's extremely granular in its division of labor, spans the globe, yet for a long time almost no one talked about it. I wanted to restore visibility to these laborers. But at the same time, I didn't want to simply film them—I wanted them to become visible through the images themselves.

This is the basic conception behind the film's visual language. I used the bounding boxes you mentioned—this process is called image segmentation. It essentially traces the different elements within a given image according to a specific taxonomy. This visual language allows the audience, as they become aware of the hidden laborers, to simultaneously discover the hidden infrastructure through which the machine constructs its vision. These two threads advance in parallel, pulling each other along. I hoped that viewers would come to understand machine vision and infrastructure at the same time. For me, the images hold equal importance to the testimony.

Juntao Yang

This way of working at the interface between labor, infrastructure, and the machine is very interesting to me. Traditional documentary often aims to make the invisible visible, to expose what is hidden. But here, you show the very act of making visible, and in doing so, you reveal a second layer of invisibility—the invisibility of the laborers themselves.

This reminds me of Hito Steyerl's concept of the "proxy": in digital circulation, images always stand for something absent. The annotation workers in your film produce proxies for the real world, but they themselves also become a kind of proxy—representing a labor force that was never designed to be seen.

This also makes me want to understand further: during the filmmaking process, what changes occurred among these workers? In the middle section of the film, as the segmented, illuminated world becomes clearer and clearer, the laborers seem to gradually become aware of the divisions they're embedded in—distinctions of geography, class, ethnicity, and nationality. In your experience of filming, did you find that this awareness could translate into collective action? Or does the infrastructure of platform microwork itself block such possibilities?

Nicolas Gourault

Thank you for mentioning the concept of the proxy—that was also something I was interested in from the very beginning. I'm fascinated by the networks and objects that connect different parts of the world, different classes of people, people who are not supposed to be connected together. In the logic of the proxy, they become connected through some system of value or exchange.

Regarding organizing power—by which I mean first understanding the situation, and then taking action to change it—this was an ongoing question throughout the entire production process. During the making of the film, some workers moved from being ordinary annotators to team leaders, and then to administrators—hierarchies quietly formed within the system. I became aware of these hierarchical structures almost simultaneously with the workers themselves, gradually discovering this infrastructure. You could say that the shape of the film—as an investigation into how the system is structured—mirrors my own process of discovering that structure. In short, I was interested in understanding the entire process, while also wanting to know how much of this knowledge was shared among the workers. Did they even have time to care about these things?

Because, as you said, the entire structure of microwork and online platforms is designed to scatter the workload, and thereby scatter the workers. It's extremely difficult to get a "panoramic view" of the system—it's like having only one small piece of a puzzle. Workers don't even know who they're working for, because client names are hidden behind code names. Understanding the full scope of the structure requires enormous effort, and this difficulty is not accidental—it's by design. I believe this isn't just about cost-cutting; it's also about dispersing the workforce and preventing any form of organization, because organization brings the potential for change.

Let me extend this point. Although the workers I collaborated with primarily worked on specific online platforms, the same issues apply to those working at physical locations. A widely known case involves Kenyan workers who trained and moderated content for ChatGPT, concentrated in Sama's (formerly Samasource) office in Nairobi. Because they were physically together, they were able to form a union and openly discuss working conditions—including prolonged exposure to traumatic content, lack of psychological support, and low wages. This led to significant public outcry and media attention. But the result was that major clients, under public pressure, chose to withdraw and seek out more obscure contractors where organization was more difficult.

So yes, organizing power has happened in the past and has had effects—but it also triggers the system's "migration": changing intermediaries, outsourcing contractors, finding new proxies to bypass public awareness and worker consciousness.

This is why, in the film, I also hint at a more global organizational imagination. But I also wanted to emphasize that even if it doesn't go global, small-scale actions at the local level still matter. These are micro-actions that might not be seen as formal political action due to their scale, but they still tangibly improve workers' conditions. So I also hoped the film could serve as a kind of homage, archiving these micro-strategies—even if they don't "succeed" or become "more political" in the way we might expect.

Juntao Yang

I really appreciate the hopeful aspect your film presents: the system always works to separate, suppress, and silence, but laborers will carve out paths toward solidarity, organization, and resistance.

The Kenyan workers you mentioned remind me of something in media circulation: AI is marketed as clean and safe, but the labor that makes it clean is outsourced, hidden, proxied. The infrastructure itself—the platform, the payment system, the NDAs—is designed to prevent workers from organizing or speaking publicly.

I'm curious about a very practical question: How did you find these workers and build trust with them? How did you enter this highly enclosed system?

Nicolas Gourault

This is a question I'm often asked. Interestingly, I entered precisely through proxies—once again. I first contacted researchers and journalists who had worked on this topic for years. One was Florian Schmidt, a German researcher who wrote an article around 2018 mentioning that German car manufacturers were outsourcing image annotation to dispersed workers in Venezuela. He was describing the situation around 2018, when Venezuela's geopolitical crisis met the global demand for images from Germany. This strange connection produced by geopolitical events meant that—I don't remember the exact numbers—at that time, the majority of online microworkers came from Venezuela.

So I contacted him, and he introduced me to other journalists, including Karen Hao, who wrote a series of about five articles for MIT Technology Review about the hidden infrastructure of AI training. This was around the time ChatGPT and DALL-E were released, when we began to realize how much extractivism and resource-intensiveness AI training requires. She tracked the different sites where this extractivism could be observed—in terms of materiality, raw materials and minerals, but also human labor.

So I contacted her, and I also contacted her collaborating journalist Andrea Paola Hernández, who was a great help in reaching the teams. We wrote to workers on LinkedIn, in private Facebook groups, trying to establish as many connections as possible. Language was key: my collaborator Paula could communicate directly in Spanish; in Nairobi, my research assistant Leonardo made initial contact in Swahili, then switched to English.

Then it was just a lot of calls, slowly—one call just to get to know each other, another call to continue building the relationship, seeing if people would be willing to do multiple calls, because some were willing to do one but then got fed up. So I had to feel out who wanted to be more involved. Trust was built precisely through this gradual process.

At the same time, I constantly questioned myself: What is the framework of our conversations? What is the overall framework of this documentary? In video calls, where should the boundaries of discussion be drawn? Therefore, I always tried to keep the exchange within an explicit work framework, centered on concrete labor practices—such as narrating and demonstrating workflows.

As for more intimate relationships, I see those as something that naturally emerges in the process of contact, rather than something I deliberately pursued. I didn't try to build collaboration through emotional intimacy; instead, I let relationships grow within the context of working together.

Juntao Yang

The extensive work you put into building connections is itself like a form of resistant solidarity—very inspiring. This also makes me think of something I found very powerful in the film. At the end, returning from the digital world to physical space, you use a shaky, unstable, embodied point-of-view shot—as if deliberately leaving the clean digital image world behind to return to real rooms, real bodies, real friction.

In doing so, you actually give the theme "Eyes" multiple layers of meaning: first the machine's eye, then the eye the workers lend out, and finally the real eye facing the present and future. I'm curious how you think about the relationship between the digital and the physical—for instance, the eye as organ versus the eye as data.

Nicolas Gourault

This is a crucial question, and one I repeatedly asked myself while making this film. I often use digital images, but I don't want to be trapped by digital images themselves. Today, our connection to the world is largely mediated by digital images, by screens—including myself, who spends long hours in front of screens. So it makes sense for the screen to become part of the narrative apparatus.

But at the same time, the screen can become a cage, a trap: within it, you repeatedly see what you expect to see. This kind of loop can easily lock you into your own constructed digital "echo chamber"—a feedback loop that may prevent you from thinking outside the box.

For this reason, I consciously seek strategies to break this self-echo. For example, introducing actual footage into the film, using documentary or archival images to counter the closure of purely digital space, pulling myself—and the audience—out of that "solitary relationship with my screen." For me, digital images (even before generative AI) are more like painting: you can build a space within them and project your imagination into it. But there's also a risk of endlessly repeating your own imagination. By contrast, actual footage or archival images force you to confront a reality not entirely under your control. I enjoy the imaginative space that digital images provide, but I always remain vigilant and skeptical toward them.

During production, the first half of the film was completed relatively intensively during an art residency. Entering the second half, I clearly felt stuck. If I continued down the path of purely digital images, it felt more like a retreat. It was precisely at this juncture that I decided to introduce actual footage.

But I didn't want to use the "talking head" format. On one hand, I didn't want to expose the workers' identities; on the other, I preferred to maintain a subjective perspective. Therefore, at the film's end, I chose to return through actual footage to the workers' subjective vision—to the real spaces and physical environments they inhabit, allowing viewers to intuitively feel the spatial disparities and asymmetries within the system. The shaky, embodied reactions you mentioned are precisely meant to create tension with the machine vision and digital images that came before. After experiencing a large amount of highly rational machine images, facing this actual footage becomes moving precisely because of the contrast—even producing an unexpected beauty.

Juntao Yang

I deeply resonate with what you just said. Entering the digital world isn't entering a single-dimensional space; it requires simultaneously handling multiple levels: the perceptual layer, the digital image layer, and the materiality and real world behind them.

About two years ago, I created a film called Light and Cold Conversations: Atlantic, Algorithms and Visibility, which discusses precisely the invisible labor behind algorithms—not limited to AI itself. I drew on "Drexciya," the Afrofuturist mythology of an underwater race, to think about the material dimension of contemporary digital infrastructure, including undersea cables, data centers, and the "ghost workers" in Kenya—like those we just discussed who clean violent and traumatic content from OpenAI's training data.

In that work, I tried to approach the problem from a technical level. For instance, when an image is compressed, dark areas—especially black regions—lose the most information. In other words, a kind of technical racism is written directly into the material structure of digital images. It's manifest not only in how images are used, but in how they're encoded, how they're processed at the technical level.

Based on this longstanding concern, I want to push further: Is it possible to resist the "sanitized" and "polished" aesthetics of technical images? Not just by presenting laborers as subjects, but by having the image itself—in its structure, texture, and operation—directly carry the traces of that concealed labor?

Nicolas Gourault

The question you raise is very important. In this film, I actually only lightly touched on this dimension: the images being labeled often come from specific regions of the world, and this bias in origin itself constitutes a kind of cultural divide and cultural standardization. This is part of the larger problem of racist bias in automated systems. For automated systems to function, they must learn "patterns"; and patterns derive from trained "norms." These norms are not neutral—they're shaped by nations, cultures, and positions of power. If you're at the center of the power structure, you can also impose your norms on other countries. This opens up a huge field of questions about how to resist the misuse and biases of technology.

As for how to resist these problems, my documentary approach tries to understand technology as a broader kind of "apparatus." This includes not just the technological tools themselves, but how the film is produced: from the production process, to the choice of tools, to the use of footage, to how images are organized and fused together.

These elements collectively constitute a kind of makeshift apparatus that can offset, correct, or even form a degree of resistance to the effects of technology.

Juntao Yang

Your notion of the "makeshift apparatus" is very inspiring. It also makes me think of technologies that merge more directly with the body—like brain-computer interfaces, especially the kind Elon Musk frequently discusses. Do you think the physical and perceptual labor presented in your film has some continuity with these more invasive forms of technological integration? Or is it an entirely different problem?

Nicolas Gourault

The first thing that comes to mind is an example: there was a project connected to NFTs or blockchain that tried to use "iris scanning" as a security device for digital identity. They traveled to many so-called "emerging countries" (including Kenya), exchanging very small payments for iris data from large numbers of people—this is already a very direct, body-level invasion.

Brain-computer interfaces, brainwave reading, and similar technologies are also developing rapidly. My research isn't yet sufficient to draw a clear theoretical line between these fields, but I can strongly sense that they're accelerating in the same direction—technology is penetrating ever deeper into the body itself. This also raises a new and thornier question of resistance: when technology is already embedded in the body, how can we resist at all?

Technology reconfigures the relationship between visibility and invisibility: on one hand, we need to fight for visibility regarding workers' conditions; on the other, we must resist being over-seen, protecting a certain opacity. So I think what's needed is a new balance: you must advocate for visibility on one hand, while protecting or rebuilding opacity on the other. This is a very broad field, but also an interesting one for understanding the dynamics surrounding these new technologies.

Juntao Yang

Today, technological politics and biopolitics have become almost inseparable.

The question of technology also reminds me that you've worked with Forensic Architecture—their creative and research methods have deeply influenced me as well. But I've always been thinking about the tension within their approach. Forensic Architecture uses technology to bear witness to a world that has been harmed by technology. This is almost paradoxical in itself. How do you use these tools while avoiding inadvertently reproducing the logic of "technological authority" that you're critiquing?

Nicolas Gourault

This is indeed a core problem I've faced for a long time. Several years ago I participated in some investigative missions with Forensic Architecture, and their methodology was very inspiring to me. But I think their work must be understood within the ecosystem they operate in: it's a complex process constituted by numerous collaborators, cross-institutional cooperation, and hybrid methods. The 3D models or image assemblages that ultimately appear are really just the result of a highly collaborative, repeatedly verified process. An organization like Bellingcat is even more fundamentally collaborative in its approach.

They've developed an entire set of techniques and processes: information gathering, verification, cross-referencing, witness interviews—and they also treat "architecture" itself as a memory device, where virtual space helps people recall and relocate events.

Why do these outputs look so authoritative? Because one of their goals is to enter the courtroom: to convince legal experts, they must hide fragilities and ambiguities as much as possible.

My work won't enter a courtroom, so I instead choose to highlight these "cracks" and "fragilities"—it's precisely these unstable elements that make the issues appear more human, more complex, and more ambiguous. The tools may be similar, but because the audience and purpose differ, the logic of what gets filtered and retained is completely different.

Juntao Yang

This reminds me of Susan Schuppli's concept of the "material witness": matter itself can testify—buildings, phones, pixels can all carry evidence of violence, thereby partially shifting authority from "human experts" to the material level.

But risks remain: the person who reads and interprets these material traces still holds enormous interpretive power; and the technology used to extract and parse evidence often originates from the same military-industrial complex. Perhaps there is no "clean" way out. What we can do is perhaps maintain a high degree of self-awareness toward these already-contaminated tools, keeping the tension continuously visible in images and cultural products, rather than trying to resolve it too neatly or completely.