SOMO

English Menu

Big Tech sets unfair terms and conditions for AI data workers globally

long read  Margarida Silva 31 March 2026 reading time 11 minutes

Key insights

- Behind most AI products are thousands of workers who label, annotate, and clean data. Their work is essential but largely precarious and invisible.
- SOMO's research found Amazon, Google, Meta, Microsoft, and Nvidia combined use at least 30 intermediary companies for data work, ranging from outsourcing companies to crowdwork platforms.
- Several intermediary companies used by Big Tech are accused of failing to provide fair and safe working conditions, including paying below minimum wages, dismissing workers unfairly, preventing collective organising, and providing no social protections.
- Even when not directly employing data workers, major tech companies influence working conditions through pricing pressures and tight deadlines and by shifting contracts between vendors. Their decisions have immediate consequences for workers' pay, stability, and job security.

- → Big Tech is not just a customer. It also invests in data work vendors and integrates them into its cloud marketplaces. This increases its responsibility towards AI supply chain workers.

The AI data work market is thriving, but the workers are not

Across the world, AI data workers have exposed and challenged their precarious working conditions. Their employers, although not household names, are crucial in developing widely used AI products such as chatbots, self-driving cars, and smart devices. These companies operate as supply chain intermediaries, connecting the world's largest technology firms with a virtually invisible workforce.

These workers clean, label, and add knowledge to the large datasets used to train and refine AI models. Increasingly, they also evaluate AI systems after deployment – for example, testing chatbots to ensure they provide accurate information and generate no harmful content (e.g. instructions for violence or self-harm).

The demand for AI data work is booming. One market intelligence firm estimates the sector will reach USD 10.2 billion in revenue by 2034, with the expanding use of AI products. The market is also highly fragmented, with new companies constantly appearing. Data from Crunchbase ⓘ shows almost 500 companies active in AI data collection and labelling. While business seems to be thriving, workers often face low pay, insecure contracts, and limited protections.

The industry lacks transparency. Companies that develop AI products often hide the use of workers and rarely disclose whether work is done in-house or through intermediary companies. In 2025, Amazon, Google, and Meta refused to disclose which human annotation services they used to develop their AI models. This secrecy makes it nearly impossible to assess working conditions.

To address this gap, SOMO reviewed a wide range of publicly available sources, including media reports, trade union statements, and company publications. Our research maps the intermediary firms supplying data work to the companies leading AI development ⓘ and examines Big Tech's wider role in shaping precarious labour terms and conditions throughout the supply chain.

Data workers raise alarm over labour conditions

Recent worker testimonies and collective organising, media reports, and research such as the Data Workers' Inquiry have highlighted the risks of exploitation across the AI supply chain. They reveal a pattern of low pay, job insecurity, and exposure to harmful content that underpins the development of AI systems.

In 2024, Kenyan data workers reported very low wages, lack of job security, and exposure to unsafe content in training AI models. The same year, Intercept Brasil exposed that one of the crowdwork platforms used by Google required workers to waive their right to collective legal action, limiting their ability to challenge unfair treatment. In 2025, Equal Times found workers in the Philippines employed by Scale AI to label data for Google's self-driving cars division reporting being paid below the legal minimum wage and given no social protection such as health insurance or paid leave.

In the US, September 2025 media reports indicated workers at GlobalLogic, who were improving Google's Gemini models and AI-generated search features, were fired after raising concerns about work conditions and low pay. In January 2026, workers at Covalen's Dublin offices, who provide services for Meta including AI training, went on strike to demand union recognition, better wages, and improved leave and redundancy conditions.

Working conditions often depend on the type of intermediary involved. Some workers are employed via a business process outsourcing company; others find work through online crowdwork platforms. The latter pose unique risks as they act as intermediaries between businesses and a large mass of remote workers hired as independent contractors and paid per task – but only if the client positively evaluates the work. This leaves workers vulnerable to abuses and devoid of labour rights, reflecting problems in the wider gig economy.

Fairwork, a project of the Oxford Internet Institute and the WZB Berlin Social Science Center, surveyed crowdworkers across the world. It found workers face substantially worse conditions in the Global South than in the North, even if they work for the same client, do the same tasks, and achieve the same or better performance.

Taken together these cases point to structural issues in how AI data work is organised and how global inequalities are embedded within the AI supply chain.

Big Tech's wide network of AI data work suppliers

Amazon, Google, Meta, Microsoft, and Nvidia combined use at least 30 companies that supply human data work to build and maintain their AI systems.

Big Tech's data work suppliers

Amazon	Google	Meta	Microsoft	Nvidia
Appen	Accenture	Appen	Appen	Appen
Centific/OneForr	Centific/OneForr	Covalen	Centaur Labs	Centific/OneForr
Cogito	Defined.ai	Defined.ai	Centific/OneForr	Eaglepoint AI
Dataloop	Eaglepoint AI	Hugo AI	Dataloop	Enabled Intelligence
Defined.ai	Hugo AI	Mercor	iMerit	LXT AI/Clickworker
Eaglepoint AI	Innodata	NexData	Innodata	NexData
iMerit	NexData	Sama	Invisible Technologies	Sama
Innodata	Pareto	Scale AI	Karya	Turing
Invisible Technologies	Sama	Surge AI	LXT AI/Clickworker	
LXT AI/Clickworker	Shaip	Turing	micro1	
NexData	SuperAnnotate		Sama	
Pareto	Surge AI		Shaip	
Scale AI	TELUS Digital		SO Development	
Shaip	Turing		Surge AI	
SuperAnnotate	Welo Data		Toloka	
Toloka				
Turing				
Welo Data				

Data collated from publicly available sources (media reports, trade union statements, and company materials) which can be provided upon request

and company materials) which can be provided upon request.

Made with Flourish • Create a chart

The mapping in this table is based entirely on publicly available information and, therefore, likely incomplete. Even so, it allows us to start peeling back the secrecy in this supply chain by revealing the network of companies involved.

A clear pattern emerges of major US Big Tech companies relying on multiple external suppliers to carry out human data work. Amazon appears to lead with the largest number of providers in our sample, apparently working with 18 data firms. Google and Microsoft follow closely, each associated with 15 vendors. This reliance on multiple intermediaries highlights how extensively data work is outsourced.

The companies Big Tech uses operate different business models. Some are business process outsourcing (BPO) firms, like Appen, TELUS Digital ⓘ , and Sama, which typically employ or contract workers to deliver services to clients. Others operate as crowdwork platforms, including Clickworker, Mercor, and Scale AI, where large pools of online workers complete tasks on a per-project or per-task basis.

Although nearly all these companies are headquartered in the Global North, with the majority (20 out of 30) in the US, this does not necessarily reflect where the work is conducted. As TechEquity's Data Work Landscape shows, many such companies recruit workers globally. Much of the labour is performed in the Global South, where wages tend to be lower and labour protections weaker.

To assess the working conditions of each data work vendor requires in-depth, company-by-company analysis. One of the few initiatives that has done this systematically is the research network Fairwork, which has evaluated digital labour platforms against 'fair work' principles. Its methodology combines desk research, worker interviews and surveys, and interviews with management. Of the companies in our sample, Fairwork has assessed four, and all scored lower than 5 out of 10.

Assessment of AI data work suppliers

	Appen	Sama	Scale AI	Clickworker
	Big Tech client: Amazon; Meta; Microsoft; Nvidia	**Big Tech client**: Google; Meta; Microsoft; Nvidia	**Big Tech client**: Amazon; Meta	**Big Tech client**: Amazon; Meta; Microsoft; Nvidia
Minimum wage	X	X		
Living wage				
Safe work		X	X	
Safety net			X	
Decent contracts/terms		X		
Secure employment				
Fair treatment	X			X
Clear systems, decisions, appeals	X			
Freedom of association	X			
Democratic governance				

Source: Adapted from Fairwork AI ratings 2025, and Cloudwork 2025, client information added from SOMO's research.

Even from this limited sample, a clear pattern emerges. Only two suppliers guarantee they pay minimum wages; none provides a living wage; and only one protects workers' freedom of association.

Prior to publication, SOMO contacted Amazon, Appen, Google, Meta, Microsoft, Nvidia, and Scale AI. In response, Amazon referred to its Supply Chain Standards, which it says "apply to all suppliers of goods and services for Amazon and Amazon's subsidiaries ... are derived from internationally recognized human rights standards and frameworks and require suppliers to maintain safe working conditions, treat workers with dignity and respect, and behave ethically, among other requirements".

Appen explained that it has since introduced new pay rates to ensure "all contributors are paid at least the local minimum wage equivalent"; it provides on-demand access to wellness services for "contributors working on sensitive or high-risk projects"; and it has a zero-tolerance policy on forced labour.

Scale AI stated it had made "meaningful investments to improve the contributor experience" and argued Fairwork's "evaluation remains overly subjective in key areas". It added: "Remotasks contributors are independent contractors who choose when and how much they work and are paid on a per-task basis", and that guardrails in place ensure earnings meet or exceed local minimum wage rates. Scale AI did not confirm its relationship with Amazon.

You can read the full replies here.

Labelling AI data workers as independent contributors or contractors echoes older narratives around gig work platforms. Increasingly, however, there is recognition that, regardless of employment status, such workers may face the same conditions as other employed workers and should be treated accordingly.

Table 2 underscores another key point. While members of this workforce experience working conditions set at the level of the supplier, behind the demand for their labour, and the pressure to deliver it at scale and low cost, are some of the world's most powerful and profitable companies.

Data work intermediaries at the whim of Big Tech demands

When data workers have sought justice through the courts, major tech companies such as Google and Meta have rejected employer responsibility, instead pointing to their intermediary vendors ⓘ. Yet Big Tech firms wield substantial power over AI data work vendors.

A small number of large clients often dominate a vendor's business. Many intermediaries rely heavily on contracts with just one or a few tech companies, with a single client sometimes accounting for between 14 per cent and 48 per

cent of total revenue. This creates a clear imbalance. While vendors largely depend on Big Tech, Big Tech does not depend on any single vendor but maintains multiple contracts, which enables them to avoid vendor lock-in and shop around for the 'best deal'.

The intermediary companies acknowledge this impact. TELUS, for example, identified "client consolidation" and the potential loss of major contracts as a serious business risk, noting this may lead clients to push for lower prices. Similarly, Appen stated: "These customers can rapidly shift their spending priorities, which creates unpredictability in our revenue streams. *This dynamic requires us to remain highly adaptive and aligned with their evolving strategies*" (emphasis added).

For Appen, this is not just a hypothetical problem. In 2024, Google announced it was terminating its contract with the company. Appen workers had been training Google's AI model, now named Gemini. The decision came a year after US-based data workers on the project secured a rise from USD 10 to USD 14.50 per hour following worker complaints and organising efforts supported by the Alphabet Workers Union.

Google told the Guardian it closed the contract to improve efficiency and that the work would be transferred to "new suppliers". The Alphabet Workers Union said the move would have a "devastating impact" on subcontractors. At the time, Google accounted for approximately 30 per cent of Appen's revenue. Contract termination led to Appen having to cut costs by USD 13.5 million and reduce its workforce by at least 100 direct employees. We do not know how many 'independent contributors' were impacted.

As Big Tech moves from client to investor, its responsibility for labour conditions becomes harder to deny

In June 2025, Meta sent shockwaves through the data work industry when it announced it would buy a 49 per cent stake in Scale AI. The move drew immediate scrutiny. SOMO raised concerns, calling for antitrust regulators to review and block the deal. US public-interest organisations warned the deal could function as a "de facto vertical merger", giving Meta control over a critical data annotation supplier.

Reactions from Meta's AI competitors were swift. Google quickly announced it would cut ties with Scale AI, reportedly to prevent Meta from accessing Google's business strategy, priorities, and data. Reports claimed Google was responsible for USD 150 million of Scale AI's revenue, around 20 per cent of its total income. Other

companies followed by also closing their contracts. A month later, Scale AI announced it would reduce 14 per cent of its workforce, affecting 200 full-time employees and 500 subcontractors. Regardless of the investment's business rationale, the impact was felt most directly by workers.

Responding to SOMO, Scale AI said it "remains an independent company" and that there is "no evidence linking the Meta investment to contributor earnings or opportunities". However, the sequence of events demonstrates how shifts in ownership and client relationships can quickly translate into job losses and instability.

The prevailing AI industry narrative is that Big Tech uses intermediary companies to distance itself from the responsibility of managing workers. Meta, for instance, has argued it was not responsible for the conditions of Kenyan workers hired via an intermediary. Yet, with the acquisition of Scale AI, Meta went the other way and staked a claim within the industry. This blurs the line between contractor and owner, and with it, the question of responsibility. If tech companies invest in and shape their suppliers, it becomes even harder for them to argue they have no responsibility for suppliers' internal working conditions.

While Meta's stake in Scale AI is one of the largest such investments, it is not an isolated case. A month before, Amazon, Meta, and Nvidia had already invested in Scale AI. More broadly, at least nine other AI data work companies have received investment from Amazon, Google, Meta, Microsoft, or Nvidia. This points to a more structural level of Big Tech involvement beyond outsourcing and encompassing ownership and control.

Big Tech investments in AI data work vendors

Data work supplier	Big Tech investment
Scale AI	Amazon, Meta, Nvidia
Snorkel AI	Google Ventures
Labelbox	Gradient Ventures (Google)
Toloka	Bezos Expedition (Amazon's founder)
SuperAnnotate	Nvidia
micro1	M12 (Microsoft)
Playment	Google Launchpad Accelerator
karya	Google.org
WEimpact.AI	Google For Startups
Defined.AI	Amazon Alexa Fund

Source: Crunchbase and media reports.

Made with Flourish • Create a chart

Big Tech firms plug data labelling services straight into their clouds

Big Tech companies increasingly embed access to human labour directly into their cloud marketplaces, making data work an on-demand component of AI development.

Amazon pioneered this model. In 2005, it launched Amazon Mechanical Turk, one of the first large-scale crowdwork platforms and still available today. In 2020, Amazon went further by introducing Amazon Augmented AI (A21), which allows users of its cloud (including developers, other businesses, and public services) to integrate human labour directly into machine learning workflows.

In promoting its services, Amazon's A21 states users can draw on "an on-demand 24×7 workforce of over 500,000 independent contractors worldwide, through Amazon Mechanical Turk". If that doesn't suit, customers "can use a third-party workforce vendor through the AWS [Amazon Web Services] Marketplace. These vendors have been screened by AWS to provide high-quality reviews and follow security processes".

AWS Marketplace recommends multiple vendors, including Cogito, iMerit, and Quadrant. Promotional materials explain that Amazon screens these vendors for cybersecurity, technology controls, and whether employees sign non-disclosure agreements. There is no indication that labour conditions or workers' rights are part of the screening.

Amazon is not alone in this approach. Google's Cloud business has a partnership with Labelbox, the company behind the crowdwork platform Alignerr, to provide human evaluation services for AI. These services are integrated into Google Vertex's AI platform. Google first invested in Labelbox in 2019, securing a seat on its board. This integration allows cloud users to access services directly within Google's AI development environment.

Microsoft follows a similar model, providing access to data labelling through its infrastructure. Microsoft has for years hosted the Universal Human Relevance System (UHRS), a crowdsourcing platform that directly connects to vendors such as Appen, LXT AI/Clickworker and Centific/OneForma. In addition, Microsoft Azure's Marketplace recommends data labelling services from iSoftStone and Quadrant. The company says it "periodically reviews the list of potential labelling service providers in Azure Marketplace and may add or remove providers from the list at any time". It is not clear what criteria these reviews cover.

Each of these cloud providers recommends and directly provides access to specific data work vendors. This makes it easier for clients to access large pools of human labour on demand. At the same time, they do not appear to provide information about or screen suppliers for workers' conditions and rights. As a result, by embedding data work into their platforms, cloud providers both enable AI development and shape working conditions for its largely hidden workforce.

Responsibility for data workers must lie with Big Tech

Much AI data work is organised through layers of vendors and intermediary platforms. The people who label, annotate, and review data are rarely directly employed by the companies whose AI systems they help build. This outsourcing

model raises crucial questions about Big Tech's responsibility across the AI data supply chain, particularly for labour conditions.

Tech companies dominating AI development have an impact far beyond their own operations. Their decisions shape how vendors operate and, ultimately, how workers are treated. When prices are pushed down and contracts are abruptly terminated, workers directly bear the consequences.

Even worse, the global and digital nature of data work allows companies to quickly shift operations across borders, creating pressure on countries to weaken labour protections to attract investment. This dynamic risks triggering a 'race to the bottom', as happened in Kenya.

As Big Tech companies move beyond outsourcing to investing in data work companies and integrating their services into cloud platforms, their role becomes even more direct. They are no longer just clients but active participants in structuring the market.

"As long as Big Tech companies demand it, there will always be an intermediary company willing to provide. The responsibility for the treatment of these workers needs to be laid at the feet of Big Tech."
Karri Lybeck, a tech industry trade union adviser and organiser, told SOMO.

While most Big Tech companies have procedures in place to screen suppliers, their impact is unclear. Google, for instance, has adopted guidelines for ethical sourcing of data work. Microsoft has conducted a human rights impact assessment of its AI supply chain that identified risks for data workers, including low wages, unpaid labour, and unsafe conditions. While these are potential good steps, limited transparency around vendors, working arrangements, and oversight makes it nearly impossible to assess whether such commitments are meaningful in practice.

Big Tech firms must assume full responsibility across the AI data supply chain

Big Tech needs to ensure its data work practices do not drive a race to the bottom among suppliers. This requires prioritising stability and worker wellbeing across the supply chain, regardless of how workers are hired.

Robust human rights and labour due diligence is essential, covering the specific risks linked to crowdwork platforms and data work. Clear and accessible mechanisms must be in place for workers to report abuses and seek remedies. The

same must happen when Big Tech firms invest in data worker companies directly or integrate their services into their cloud marketplaces.

Big Tech must also enable external scrutiny by disclosing which human data work services they use and under what conditions. Without this, meaningful accountability remains out of reach.

Governments and public authorities must protect AI data supply chain workers

Amid the frenzy to 'win AI', most governments seem oblivious to workers' roles in the AI supply chain. Governments must ensure existing labour laws apply to these workers, regardless of whether work is organised through contractors, outsourcing firms, or digital platforms.

Antitrust regulators must also start to take workers into account when assessing the harms of AI market concentration. Deals such as Meta's investment in Scale AI can reshape the industry in ways that affect both competition and working conditions and should be reviewed accordingly.

Governments can also use their leverage to protect data workers, such as by pressuring companies through public procurement to be transparent and responsible. Despite having been recently watered down, due diligence rules such as the EU's Corporate Sustainability and Due Diligence Directive will also cover Big Tech and could be an avenue to seek more accountability across AI data supply chains.

The International Labour Organization is developing a binding commitment for decent work on digital platforms that could set minimum standards to ensure gig workers, including those doing AI data work via platforms, are protected. Governments must actively support this process, and workers should have a seat at the negotiating table.

Public pressure on Big Tech is needed to improve conditions for data workers

Data workers are increasingly mobilising and organising to demand change. In Kenya, for instance, they have come together to create the Data Labelers Association.

This resistance is vital, but workers face a major power imbalance between the world's largest corporations and a fragmented, remote and hidden workforce. In

the current geopolitical context and the race to dominate AI, governments increasingly align with corporate interests, prioritising competitiveness and investment over labour protections. This allows weak regulation, opacity, and the outsourcing of responsibility to persist and even deepen.

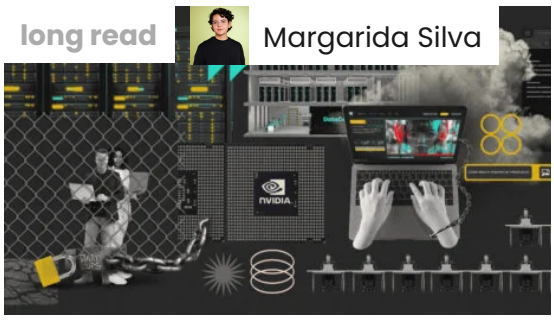
Meaningful change is therefore unlikely to come from voluntary corporate action alone or governmental willingness to act. It will depend on sustained pressure from below. Workers, trade unions, and civil society must continue to build alliances across borders, exposing supply chain abuses and demanding accountability.

At stake is not only the future of work in AI, but the kind of digital economy we are building.

Do you need more information?

Margarida Silva
Senior Tech Researcher

Related news

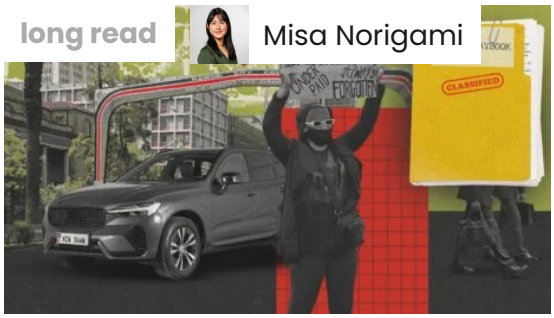


The real winners of the AI Race

7 July 2025

Big Tech's AI Shopping Spree

5 February 2026



Kenya: the human cost of Big Tech's power

12 February 2026

Latest updates →

Don't want to miss anything?

Sign up for our newsletter and always stay up to date on information and analysis on corporate power issues.

Email (Required)

Anti-Robot verification

CLOUDFLARE
Privacy • Help

Submit →

Contact

Tel: +31 (0)20 639 12 91
(Mon-Fri, 9AM-5PM)
Email: info@somo.nl

Address

Centre for Research on Multinational Corporations (SOMO)
KNSM-laan 17
1019 LA, Amsterdam

Netherlands
directions (opens in new tab)

Practical details

Bank name: Triodos
Account: NL56TRIO0198407777
BIC / SWIFT code: TRIONL2U
Chamber of Commerce: 41223308
RSIN/Fiscal number 2954631

Quick links

About us

Work at SOMO (Vacancies)

Newsletter

Donate