TECHGLOBAL INSTITUTE

Labor in the Shadows: Dignity of (Digital) Labor

Disha Verma

June 2026

Acknowledgement

We would like to thank the ten informants who have been instrumental in shaping up this report. Activists from the Data Workers' Inquiry Project and BPO Industries Employees Network (BIEN) have guided us in understanding the highly local, cultural, and contextual economies of data work across the Global Majority world.

About the Labor in the Shadows series

This is the second report in a three-part series exploring digital labor industries in Asia that powers technological platforms and artificial intelligence (AI) systems worldwide.

In the first report, we outlined how invisible labor in content moderation and data annotation industries is extracted by technology companies from cheap labor markets in the Global Majority while remaining distant and largely unaccountable to the workers who make their business a profitable reality. While both industries create opportunities for dignified and accessible work in the jurisdictions they operate in, they are also marked by a wide spectrum of harms and precarity, warranting platform accountability and policy reform.

We identified a pressing need to not only point to these structures of power within labor landscapes, but also investigate and map the socioeconomic enablers in South and Southeast Asia that make these markets fertile ground for outsourced digital labor.

Key findings from the first report are:

- Data work is outsourced transnationally to cheaper labor markets and more lenient legal environments, which in Asia are most prominently located in India and the Philippines.
- It is contracted by platforms to workers through complex and cascading structures involving one or more intermediaries—arrangements that often escape domestic legal scrutiny, sidestep labor rights, and neglect worker well-being.
- Data work industries, especially content moderation, are marked by vague hiring practices which often deliberately misrepresent the job and its requirements at the hiring stage, as well as layers of opacity between workers, management, and platforms in all manners of operations, expectations, performance metrics, as well as decisions related to contract termination.
- These industries are also thus notoriously precarious, and termination of employment often cannot be appealed against.

- Exhausting work hours beyond the usual 8-hour work-day and low wages are a persistent issue with most forms of labor extracted from the Global Majority. But data work may be uniquely performed outside formal labor contracts, exposing it to exploitation without domestic labor code protections.
- Owing to the extremely graphic and violent nature of content moderation and some forms of data annotation, psychological distress and fatigue are now central features of the job with little to no mental health support from intermediaries or platforms.
- Companies target and extract labor from socioeconomically vulnerable zones or populations, and have the freedom to relocate their business to jurisdictions with more amenable regulation, leaving workers in a perilous state.
- Despite the numerous labor rights issues in global tech supply chains, workers may find it difficult to organize or collectively bargain for their welfare due to cultural factors, remoteness of work, or skewed power dynamics with platforms.
- The present international labor law order also does not adequately protect this industry from labor harms—but with ongoing deliberations on digital platform work at the International Labor Organization (ILO), there may be a more comprehensive regulatory framework on the horizon.

This report, second in the series, delves deeper on these ideas and findings through first-hand accounts from data workers and their lived realities in the region.

Methodology

In the previous report, we noted a lack of research on Asian data work industries due to the fragmented, heterogenous, and often opaque nature of operations, pointedly in content moderation and data annotation work. To fill some of these gaps, we reached out to and interviewed 10 informants from Pakistan (4), the Philippines (4), Malaysia (1), and Venezuela (1) (see Annexure 1). A list of guiding questions for these interviews is annexed in Annexure 2.

The informants are a mix of data workers and journalists who have directly worked with data workers. Workers were contacted via journalists and field researchers who have existing relationships of trust with them. We exercised utmost precaution and care in approaching and interviewing workers in ways that do not endanger them or alert their contracting firms. Three out of ten informants chose to be anonymous but consented to us using their experiences in our analysis. Interviews were conducted between June and September 2025.

TECHGLOBAL INSTITUTE

Introduction

> Fahad* is a content moderator with Accenture in Kuala Lumpur, Malaysia. He was hired and contracted for the role of 'Content Analyst' by HR TAG, one among many of Accenture's recruitment firms.
>
> At 10:30 AM sharp, Fahad takes his seat at the production floor of Accenture's offices, and logs into a system administered by its biggest content moderation client, ByteDance. He mentally prepares for a whole day of filtering through highly graphic video content with a fine-toothed comb for ByteDance, while the platform maintains two degrees of contractual separation with him.

As soon as he logs in, an influx of "cases" begins. In no time, there is a queue of content sorted by priority. Fahad must sift through the queue in a tight timeline (often around 15 minutes) while maintaining his latency, productivity, and utilization numbers which directly impact his performance report. Utilization here means how much active time Fahad spends on each case, which is his most important performance metric.

On the menu bar on his computer screen is a communication tab where newly updated community guidelines and content policies are notified. Fahad had just finished memorizing and getting a hang of the updated policies sent last week—he now sees a new set of policies and must start from scratch. Neither Fahad nor his content moderator peers have any input or insight into how these policies are made or updated.

In the late afternoon, Fahad wants to leave his desk to read the Namaaz. However, if his system goes idle for over 10 minutes, it will gravely affect his performance report. For a few months some years ago, Fahad recalls Accenture granting them a dedicated prayer break of 30 minutes. However, once ByteDance came under fire from the US senate, Accenture felt compelled to prioritise the mental well-being of workers and replaced the 30-minute prayer breaks with mandatory "wellness breaks". Now, Fahad and his Muslim colleagues keep their prayer mats by their desks, and rush their prayers along to maintain their performance figures. Even the 30-mute wellness break proves useless—when Fahad tries to change his availability status on the system to step

out for the wellness break, it does not stop his clock. Anything beyond 10 minutes injures Fahad's metrics.

Fahad has found a small community within Accenture, where he and his colleagues crack jokes, share meals, and keep the environment light. They almost never talk about work. He has a pleasant relationship with his manager, but fears the manager may soon be replaced by an AI Team Lead, which will be impossible to negotiate and build relationships with. But Fahad firmly believes content moderators are not going anywhere. This is a job AI could never do. Even if it is gruelling, traumatic, and unrewarding at times.

He says, with an air of contentment,

> **"At the end, this is a job. If I were a content moderator in the US, I could have said something. But as South Asians, our labor is replaceable. So many are in the queue to take my job. Why should I complain?"**

Data work (used in this report series to refer to content moderation and data annotation industries specifically), as a subset of the digital platform labor economy, is a double-edged sword for Asian workers. While it creates gainful employment or avenues to earn a side income for some, it traps the others in pernicious employment arrangements with limited exit routes. These transnational, extractive structures of data work found abundantly in the Global Majority world render workers powerless and slowly eat away at their dignity.

Even though some forms of data work offer higher degrees of formality than the rest and as compared to platform-based gig labor, the lowest standards of agency and autonomy is ensured across the board. Platforms, in all such settings, build models that reduce their own risks and costs by offloading them on to their data workers. These workers rarely have the socioeconomic mobility or power to choose alternative livelihoods.

Across all conversations with workers featured in this report, it was commonly

and clearly understood that platforms model these industries solely on a profit motive. In this context, research and advocacy on these themes must serve the interests and demands of data workers occupying these roles with the goal of restoring balance to these power asymmetries.

> **One of the ways data workers can reclaim power at the workplace and within vast and largely exploitative data industries is by retaining epistemic authority and control—their challenges and grievances must become the forefront of conversations on bustling data industries and the center of any structural reform.**
>
> **The purpose of this paper series, as well as the Labor in the Shadows project, is to support this significant balancing of power.**

This report attempts to capture the realities, roadblocks, and demands of data workers through informant interviews and conversations with worker collectives. These conversations pointed to the following issues in these industries, which are explored in depth in this report:

a) That there are many interpretations of "data work", and hence, a "data worker";

b) That new and complex models of data work are emerging quicker than one can keep count;

c) That data workers are consistently relegated to the bottom of the tech supply chain in practice;

d) That a bulk of the issues explored in the first report stem from structural non-transparency in data work industries;

e) That the potential takeover of AI systems and algorithmic managers poses grave risks to workplace dynamics;

f) How gender intersects with data work in the region; and

g) Collectivizing in a fragmented industry and barriers in organizing.

Who is a data worker?

"Data work" is a colloquial umbrella term used to define workers that perform tasks related to content and data on digital platforms. The first report uses "data work" as a collective term for content moderation and data annotation, which is a context-specific use that does not reflect its breadth. Huws (2016), for instance, refers to data work as logged labor because it: (a) cuts work assignments into standard, quantifiable components; (b) engages in the continuous data-based surveillance of workers; and (c) requires individuals to be connected to an online platform to obtain work. It is also monikered cloudwork (by Fairwork), content-related digital labor, click-work, micro-work, crowdwork, ghost-work, or digital platform labor—although the last term also encompasses offline labor.

The ILO broadly divides 'digital platform labor' into two categories: *first*, work performed on online web-based platforms, which is online and remote, and allocated to a crowd (on microtask and competitive programming platforms) or to individuals (on freelance and contest-based platforms), and second, work performed on location-based platforms at a specified physical location by individuals such as taxi drivers and delivery workers.

Additionally, data work morphs into new microindustries every now and then, which become bigger and more complex with emerging technologies. From transcription, labelling, and other ostensibly simplistic tasks to the complex and psychologically distressing prompt writing and content moderation, it is difficult and fairly unfeasible to cement a definition for data work, and thus, for data workers.

Besides, data work can be interpreted differently based on cultural contexts and strides in technology. Modern interpretations of "data work" extend to the understanding that everyone who works with data—who creates, modifies, labels or processes data—can be considered a data worker. In conversation about his work for the Data Workers' Inquiry Project, I-1 stated that all users of the internet are data workers, insofar as we ***"create data and help platforms hone their systems".*** A popular workstream in data training for AI systems is labelling or organising data so that a machine can process it—work that internet users do daily in the form of CAPTCHAs, for instance. Similarly, the act of scrolling and engaging with certain content on social media to build customized 'for you' pages generates patterns of behaviours that help companies build precise consumer surveillance and targeted advertising infrastructures.

Even recreation can become unseen labor—the hugely popular Niantic-owned augmented reality game Pokémon Go, which

attracted over over 500 million players within 60 days of its launch, was recently revealed to be a massive crowdsourced data collection drive to train the parent company's AI product, Niantic Spatial. During gameplay, hundreds of millions of Pokémon Go players walked around their neighborhoods recording urban landmarks and buildings across the world, unwittingly making Niantic Spatial possible. All of this work—solving CAPTCHAs, scrolling on social media, or playing location-based games—is data labor. None of it is particularly compensated.

While this wide interpretation of data work and its extractive nature provides excellent commentary on present-day platform capitalism, it may also have the effect of undermining or subverting resources from a marginalized workforce that solely depends on contractual data labor for survival. Since data work is increasingly extracted from the Global Majority, calling everyone using social media or the internet a "data worker" may collaterally further invisiblize the workforce.

In a similar vein of data work being extracted from people against their knowledge or consent, the Data Workers' Inquiry also proposes a new category of data workers—the "zombie trainer". As explored in more depth in the previous report, this includes captive or vulnerable populations like refugees, children, prisoners, low wage workers, and applies to any arrangement where platforms exploit the socioeconomic precarity of a community for extractive labor, even if in exchange for compensation.

As a Rest of World report on how refugee camps power AI systems notes,

> **"[m]icrowork comes with no rights, security, or routine and pays a pittance — just enough to keep a person alive yet socially paralyzed… stuck in camps, slums, or under colonial occupation, workers are compelled to work simply to subsist under conditions of bare life."**

This is distinct from, say, formal employment settings where data workers opt in for this work (albeit not under the most transparent or equitable conditions, as explored in the previous report) because "zombie" work systemically overrides agency and autonomy of the worker or the community experiencing the socioeconomic vulnerability. Its existence and ubiquity hinges on such vulnerability across the Global Majority, as seen in Venezuela during its financial crisis, or in Lebanon, Jordan, and Kenya among their vast refugee populations.

Presence of crowdwork platforms like m2Work (a joint venture between Nokia and World Bank) in Palestine, introduced to alleviate high rates of unemployment in the distraught region, is an example of powerful firms leveraging socioeconomic distress (here, an ongoing genocide) to engage affected populations in unequitable labor arrangements. This is often couched in the language of economic aid or "job creation". In the Global North, AI firms have already begun using the labor of incarcerated persons to train AI systems—data services firm Vainu working with prisons in Finland is publicly hailing this arrangement as "prison reform". In the US, Amazon provides communication monitoring technologies free of cost to prisons so that it can train its AI systems on inmates' communication data. The question then becomes if these populations, who have unwillingly and unwittingly become data workers, should be included in global governance conversations on digital platform labor.

TECHGLOBAL INSTITUTE

Emerging models of data work

Public sector digitalization is creating public data workers

Building on I-1's expansive interpretation of data work, some emerging employment models suggest that data workers need not always be employed by a private entity. In Asia, flanking the proliferation of data work industries, is the rapid digitalization of various state functions and public services. India, for instance, has entered into multiple regional partnerships exporting its digital public infrastructures (DPI), which are technology frameworks that enable the digitalization of public benefits and services at societal scale. This effort, at local levels, also involves working with data—digitizing public registers, conducting digital surveys, collecting data for training indigenous, state-owned large language models, and labelling and maintaining large datasets for state use. As a result, individuals engaged in any kind of data enumeration across sectors like agriculture, health, or urban planning may be increasingly roped in to perform a variety of digital labor that fuels state-owned and -run digital platforms and systems.

A clear parallel ecosystem of state-administered data work can be seen in the community health architectures in South Asia, notably India. Healthcare is delivered and enumerated in rural pockets of India by Accredited Social Health Activists (ASHAs/ ASHA workers)—women community health workers (CHWs) who perform a litany of frontline tasks like providing primary medical care, conducting door-to-door health surveys and vaccination drives, and ensuring deep penetration of healthcare in all corners of India. This work, ever evolving and expanding in scope subject to local needs and administrative whims, forms the backbone of Indian healthcare. It is also performed informally—ASHA workers are not seen as employees of the state, and are relegated to a "volunteer" status. They have been demanding regularization and employment benefits for over a decade, alongside demanding timely and complete pay, restricting the work roster, humane work conditions and hours, and dignity rooted in social protections.

In 2020, India introduced the Ayushman Bharat Digital Mission (ABDM) as a DPI and the new epicenter of India's health infrastructure. This substantially shifted ASHA workers' priorities from community-based healthcare to data-driven, digitalized tasks. ASHAs in the Indian states of Haryana, Rajasthan, Maharashtra, among others, have come under considerable pressure to engage with digital tools for

data collection, coordination, and recording. These tasks resemble data crowdwork, are not documented in the National Health Agency's policy that anchors the ASHA ecosystem, and have little to do with the ASHA workers' primary responsibilities.

Post ABDM, ASHA workers in some states claim that for one, they now do more data work than healthcare work, and two, that this shift has distanced them from their essential work and instilled distrust between them and their communities.

Many parallels can be drawn between the ASHA infrastructure and digital platform economies in Asia insofar as they extract highly complex, intensive, and repetitive labor from vulnerable populations without formalizing for regularizing them into employment, or providing adequate recompense.

A 2022 study into digitalization of healthcare work also show that the worker (a) rarely has the bargaining power to say "no" to additional data work owing to heavily skewed power dynamics in the scheme of employment (in this case, the state being the employer itself), and (b) enjoys no real agency in deciding their workflows without it affecting their incentives and payment structures.

CHWs with a higher degree of formalization, such as Anganwadi workers in India, face similar challenges with respect to being thrust into digital labor without much agency or bargaining power.

Tasks pertaining to the management and implementation of digital health systems are commonly being posed to CHWs across the region—in Nepal, they perform data collection and information management tasks aligned with the eHealth Strategy of 2017; CHWs in Pakistan, Cambodia, and Bangladesh have started to engage with mobile health and data management platforms in their routine work. Leveraging the CHW architecture to power state-run digital platforms a microcosm of the extractive planetary digital labor market explored in the first report.

Blurring lines within digital platform labor

Similarly, the lines between the two branches of digital platform labor defined by the ILO (web-based and location-based) appear to be rapidly blurring. There is, for instance, a documented rise in app-based taxi and delivery platforms worldwide engaging registered drivers in digital labor tasks over and above their job responsibilities (driving and delivery) without separately compensating them

for it. Drivers are expected to feed data into company databases, which is used to train algorithms, which might then be used to perform managerial tasks and make decisions affecting the same driver. These drivers may not always know they are performing "data work". In 2025, Uber rolled out a paid feature for drivers to perform tasks such as voice recording, capturing and uploading images, transcription or submitting documents in certain languages—reported prompts include *"upload images of cars"* or *"record yourself speaking in your language or local dialect."* This project was piloted in 12 Indian cities before launching in the US.

These two instances demonstrate how the forms and industries of data work may be different and ever-evolving, but are consistently underpinned by:

a) Extractive, exploitative labor which affords little to no bargaining power to the worker;

b) Informalization and job precarity, especially when the worker engages in data work, despite worker demands for regularization or restricting of work rosters;

c) Subpar or inhumane working conditions;

d) Assertion of arbitrary power of the employer (which may be a state) in decisions pertaining to compensation, contractual arrangements, and termination;

e) Active efforts by the employer to distance itself from the worker, either through the employment arrangement or through opacity in operations;

f) Little to no agency or autonomy over workflows and rosters; and

g) General lack of dignity and credit for the work done.

Hence, who is a "data worker" is an increasingly blurry inquiry given the fasttracked national digitalization efforts in the region. As a loose definition, a data worker is any individual engaged in public or private digital platform labor by way of a contract or statute. A data worker may not always perform labor in exchange for payment, as evidenced by the growing use of volunteer and community apparatuses by states to delegate digital labor. The first report notes the reluctance of states to call data workers by this title, often resorting to vague terminology like "partners", "agents", and "executives", which exempts them from the application of baseline labor and social protections. While it is a complex exercise to identify and define data workers, the interpretation should necessarily be wide and such that it brings more workers under protective and prescriptive laws. This uncertainty must not be weaponized to leave data work unprotected.

Consistently relegated to the bottom of tech supply chains

When Amazon CEO Jeff Bezos launched his microwork platform Amazon Mechanical Turk (AMT) to an MIT audience in 2006, he announced: *"You've heard of software-as-a-service. Now this is human-as-a-service."* Data work, especially through such clickwork or microwork platforms, is built on steep power asymmetries to this day. As highlighted in the first report, workers do not enjoy autonomy or editorial freedoms over the tasks they perform. Content moderators, specifically, are often christened "cleaners", "janitors", or "soldiers" of the internet—a framing that minimises them to mere executors of an action, when the job demands much more.

> **I-6 outlined his experience with how content moderators have to strictly abide by concrete internal policies such as the client's community guidelines, even when logic dictates otherwise.**
>
> **Moderating in the pashto language, he is often able to predict a type of content would violate his client's content policy even if it is not culturally inappropriate.**

He explains,

> “in Pakistan and Afghanistan, a certain white flag with lettering is commonly hung outside homes as a sign of hospitality, but the content policy wrongly identifies all white flags with lettering as terrorist insignia. Even though I know better, I have to simply do my job and label any content featuring this flag as violating the community guidelines. I can’t reason with the policy—and if I use my judgment and don’t label this content as offensive, it will affect my performance metrics. I might get warnings and be terminated for repeated warnings.”

This is by design. As their work is consistently relegated to the lowest tiers of the tech industry hierarchy, data workers’ contribution to the value chain is largely confined to applying internal policy guidelines, with limited scope for meaningful input, context-sensitive interpretation, or autonomous decision-making. Due to this repeated reinforcement, data workers across the industry refrain from providing any meaningful input to the task at hand or introducing their own subjectivities, keeping the cycle intact.

Structural opacity as a leading barrier to realizing labor justice in data work

The previous report outlined how data work industries, notably content moderation and data annotation, are rife with vague and opaque hiring practices and platform design.

> Conversations with I-2, I-3, I-6, I-7, and four unnamed data workers from Asia confirm that job postings, most often broadcasted on LinkedIn, use job titles like
>
> "Analyst", "Data Analyst", "Content Writer", "Content Analyst" for what is squarely a content moderation post.
>
> Seemingly, this obfuscation is prominent in the content moderation industry, presumably due to the highly graphic and distressing nature of the work.
>
> I-2 also notes that like her, other content moderators are often absorbed into data annotation or AI prompt writing teams without their consent, depending on the client's projects and needs.
>
> Not only are these postings and transfers silent about daily responsibilities of the candidate, but the occupational risk they may be signing up for—risks that are significant, permanent, and numerous.

The ways in which digital platforms enforce structural non-transparency and deliberate vagueness can be broadly divided into (a) lack of information on clients being served, (b) lack of information, reasoning, and recourse around decisions made by a human or algorithmic manager, (c) suspicions around data worker's inputs being used for AI training, and (d) the weaponization of non-disclosure agreements to minimise bargaining power.

Non-transparency vis a vis clients served

> **While I-6, I-2, and four unnamed data workers working at well-known intermediaries in Asia claimed they were aware of the platforms and clients they were ultimately working for, data work done through crowdwork platforms is often not transparent about clients.**
>
> **I-1 noted that, at least in Venezuela, AMT did not reveal its empanelled clients to "protect them". He added,**

> **"data workers must be made aware of the companies they are working for and give them an option to opt out—actually, this should be the case with any and all kinds of work."**

The first report outlines the challenges faced by Syrian data workers in Lebanon who cannot determine if their labor is helping to build technologies that enable the very forms of violence and war crimes they fled from. In a similar vein, data work intermediary Appen, with a workforce of 1 million data workers including from Asia, reportedly helped develop surveillance aircrafts for the US military that were used in the US' violent capture of Venezuelan president Nicolas Maduro. Given Appen's presence in over 170 countries, some workers who unwittingly trained this system may have been from countries that have faced hostile attacks from US armed forces. No worker was informed, let alone asked, that they would be undertaking projects for the US military. Transparency in clients being served through data work should be a grundnorm for

crowdwork sites, and should be considered an AI ethics concern in global governance conversations.

A 2024 study on the realities of counting in data annotation work shows that AI clients, especially ones located in the Global North extracting data labor from the Global Majority, exert a high amount of control and authority over annotation processes, constituting them as reductive, standardized, and homogenous when it suits their profit motives. Clients, wielding such powers over the worker through unfair contractual arrangements and a lack of enforceable legal protections, also may refuse to pay for a task on a complete whim—and in the absence of grievance redressal systems on most platforms, the worker has no bargaining power to ask for compensation. In fact, there is a high chance the worker will not even know who the client is. Ongoing governance deliberations at the ILO, which started with the August 2025 report on decent work in the platform economy, lack adequate articulations on transparency—limiting itself to transparency around algorithmic decision-making. However, given the tripartite nature of governance at the ILO, it does not seem promising that employers will concede to the ILO prescribing transparency into their clients.

Non-transparency vis a vis managerial decision-making

> When asked what may be the biggest challenge in data work, I-1 pointed to the lack of transparency and accountability over managerial decisions like flagging an error in a task, rejecting a completed task for no discernable reason, or even unexplained worker termination.
>
> I-2 and an unnamed worker engaged in content moderation noted that they are not notified of an error until days pass, and may arbitrarily get flagged for multiple errors on a random day.
>
> I-6 explains this may be because of biases within the management pipeline, personal relationships with managers, or day-to-day office politics. In some well-known intermediary firms, as low as two errors can lead a worker to be terminated. On crowdwork sites, errors adversely affect worker rating which limits the work one can find, and repeated errors can lead to a worker getting banned.
>
> I-1 notes, not getting feedback on an error is demotivating for the worker—a feedback loop must be integrated into the crowdwork platform.

This, of course, is distinct from transparency in algorithmic decision-making—the threshold here is lower, i.e. when a human manager oversees data work, there should be a two-way channel with respect to decisions and terminations. I-2 notes that at their workplace, decisions as sensitive as how a worker is treated due to their affiliation to a BPO union are left to the personal preferences of managers and supervisors. While it may be easier to obtain feedback in a formal office setting, remote crowdwork sites make it nearly impossible for a worker to establish contact with their managers.

Potential AI training on worker inputs

A growing concern that AI content moderation will soon replace workers has been gripping the industry. A leading data work intermediary, Accenture, laid off 22,000 employees in 2025 alone, suspectedly to replace all or most of them with AI. This follows one of their 2025 earnings calls where they claimed staff unable to be reskilled to use AI will be let go off. I-6 and two unnamed content moderators at Accenture in Malaysia were a part of these mass lay-offs—they were told expressly that they are being replaced with AI moderation. Due to opaque and highly layered interfaces through which they performed content moderation, they were unable to say if the AI moderator was trained on their labor.

Similar suspicions and uncertainties plague workers engaging with chatbots. A chatroom worker reporting for Data Workers' Inquiry ruminates on this and probes further, writing, "

> **"...what if I wasn't just training an AI companion, what if I was actually impersonating one? Maybe users thought they had already purchased an AI girlfriend or boyfriend, and I was the human pretending to be the machine pretending to be human[?]"**

This is not unusual—a 2021 ILO survey of 300 online home-based workers in the Philippines concluded that about 14% of them were working as "AI virtual assistants" for clients based in Australia, Canada, the Philippines and US. It is not clear whether human workers were slowly

training the AI which was marketed to the client, and this information is difficult to extract from tightly opaque and proprietary workflows at intermediary platforms.

Non-disclosure agreements

As outlined in the first report, non-disclosure agreements (NDAs) are the most common tools to ensure secrecy and stifle advocacy in data work industries. NDAs are usually non-negotiable for the job—they are designed strategically and solely to protect the company's business model, even if it entails directly dismantling any organizing or advocacy efforts taken up by the workers.

> **I-2, I-6, and three unnamed workers brought up NDAs early into a general conversation about data work.**
>
> **According to I-2, these NDAs are continually reinforced through seminars and trainings at intermediary firms, where workers are instructed on client security and data protection measures. This, invariably, includes instructions to *"not speak about clients", "not share daily workloads on social media",* and so on.**
>
> **I-2 also noted that while content moderation teams are not barred from internally talking about these details, this routine reinforcement of secrecy discourages them from discussing work externally.**
>
> **I-6 and an unnamed worker also note that they rarely challenge the NDA arrangement and try to abide by it even in private conversations due to fears around being migrant workers and facing high job precarity in the industry.**

This is not an unsubstantiated fear. NDAs have been an active barrier in labor justice and social protections for workers across industries for years. At the webinar *'Unearthing Technology's Human Foundations: Data Work and Content Moderation in the Global South'* hosted by Tech Global Institute's labor coalition ATLAS, Filipino data worker and community organizer Renso Bajala revealed that he was fired from his role after speaking about his work with a leading regional news magazine. Sometimes, these NDAs include legally contentious clauses like non-competes and complete forfeiture of intellectual property, but these often take the form of take-it-or-leave-it contracts where workers have no negotiating rights. Data workers from Kenya and Germany put forth a joint demand in 2025 to abolish NDAs altogether in this industry, owing to their chilling effect on the freedoms of expression and association.

TECHGLOBAL INSTITUTE

The persisting challenge of algorithmic management

Digital platforms, especially crowdwork sites, are increasingly delegating managerial responsibilities for various forms of digital labor to algorithmic systems, which may or may not rely on AI to supervise human workers and make decisions directly affecting their work, wellness, and welfare. The first report noted an interest among policymakers to regulate and prescribe norms for algorithmic management on digital labor platforms, especially insofar as it affects ratings, delegation of tasks, and unappealable decisions like account suspension or termination. However, it is not the nameless algorithm that is of pressing concern to the interviewed informants or according to leading studies—it is the constant remodelling of employment arrangements in ways that increasingly isolate workers and give them less and less agency.

> I-3, I-4, I-6, and two unnamed workers have not faced algorithmic management, and do not see it as a threat. I-3 expressed concern in relation to the algorithm's inability to gauge human context, emotion, and nuance, and hence its competence as a manager.
>
> They draw on their BPO experience to note that

> "like call center workers were trained to do back in the day to hit the metrics, an AI manager will simply speak in an enthusiastic, high voice even to a grieving customer."

The concern, as explored in a 2025 study on digital platform labor, is how these systems entrench existing power imbalances and render workers with even less agency than before. The human agents shaping an algorithmic manager end up instilling their own biases in it. This is also another way for platforms to distance themselves from their workers and create smokescreens of grievance redress without instituting any.

In an environment where data workers, especially those working remotely, face challenges in building community or unionizing due to the isolating and highly confidential nature of their work, being forced to interact with algorithmic managers instead of human ones can affect their quality of work and life. Such an arrangement perpetuates "algorithmic anxiety"—a quiet social power algorithms can enjoy over human beings in specific situations. The merging of algorithms into our everyday lives, the significant decisions they are increasingly being trusted to make, and them supposedly replacing human labor altogether grimly suggests that algorithms have become a critical infrastructural element of contemporary life, and gatekeepers of power and power relations.

A 2025 study based in India notes that data workers experience various forms of anxiety in negotiating with algorithmic systems marked by opacity, randomness, and unfairness. Where algorithms allot tasks to workers at random, there may also be risks of bias (or perceived bias) on the basis of personal worker data, including social identity markers like gender, age, complexion, religion, race, and other protected characteristics.

Although there is an intent to regulate this aspect of data work, legislating has been slow domestically and internationally. The ILO is still negotiating on baseline transparency (i.e. disclosing the use of algorithmic decision-making). European Union Directive 2024/2831 and Uruguay's Law no. 20396 of February 2025, for example, contain more detailed provisions requiring transparency from platforms using automated monitoring and decision-making algorithms, and advocating for them to be prohibited in certain use cases, such as employee termination.

In India, gig work laws across the states of Rajasthan, Karnataka, Telangana, and Jharkhand do not engage with algorithms beyond encouraging the use of right to information mechanisms to probe into them. Malaysia's Gig Workers Act of 2025 also provides limited protections against algorithmic management beyond obligating the platform to inform the worker of the presence of an algorithm.

TECHGLOBAL INSTITUTE

Gender and data work

Both industries—content moderation and data annotation—interact interestingly with gender in the region. Women are less prominent in content moderation, which often subjects the worker to highly distressing, violent, and sexually aggressive imagery and content. In contrast, remote data work like data annotation is acutely popular among Asian women.

Task assignment on the basis of gender

Foiwe, an India-based data services firm, provides content moderation services to, amongst others, dating sites from across the world. Indian moderators assigned to one such client, a queer site GROWLr, receive a regular influx of pornographic and nude images. Male moderators at Foiwe accept sole responsibility to work on these, explicitly stating in a 2018 study that **"this kind of imagery would not be appropriate for a female employee to moderate".** It even became an unspoken rule to assign female Foiwe moderators clients that are "safer", such as e-commerce firms and online marketplaces. In this study, workers claim that no woman should be exposed to the amount of pornography and nude imagery that female content moderators have to sift through everyday—that it should be **"minimised or eliminated altogether where possible."** Some of them claim discomfort while having to moderate sexually explicit content in the presence of female colleagues.

> **I-6, a Muslim content moderator from Pakistan but based in Malaysia, explained the moral turpitude one goes through in accepting their role as a content moderator. Faced with an onslaught of pornographic content on his very first day, he flagged that such a job does not abide by his religious principles and beliefs—to which his manager claimed he will be *"stopping something wrong from happening [on platforms], and not supporting it".* This convinced I-6 to stay in the role for two years.**
>
> **In contrast, I-7 notes from their investigations into content moderation in Pakistan that they could not find a single female moderator to speak with.**
>
> **Religion and cultural contexts shape the way data work interacts with gender across Asia—and it is not homogenous.**

I-2, 1-3, 1-4, 1-5, and an unnamed worker stated there were no noteworthy trends in the uptake of data work by women versus men in Southeast Asia.

I-4 added that in the Philippines and Taiwan, tech factories have always had a preference for women due to their smaller, more nimble hands, and later at call centers due to a higher demand to outsource female callers.

Uptake of data work based on gender

South Asia is a global hub for data work: India accounts for roughly a third of the world's online freelance workers, followed by Bangladesh—by the early 2020s, approximately 70,000 people in India worked as data annotators (full- or part-time), in a sector valued around $250 million USD in 2023. As outlined in the first report, Nepal and Sri Lanka are projected to have larger investments into data work, with Bhutan and the Maldives tailing through arrangements with Indian vendors. The region, especially India, is understood to have a large contingent of female data workers on crowdwork platforms like AMT, engaging in freelance or gig-based data annotation and labelling work. However, their participation in the workforce is not equal. Women end up paying grey market fees to even access the platform, only to be relegated to the lowest paying microtasks and excluded from higher-paying opportunities reserved for Global North workers.

While the numbers seem low, women in India have had an interesting journey with data work. AMT and Telus saw a surge of women joining during the pandemic, as noted in the previous report, partly due to its accessible, remote nature, but also because the prestige associated with tech jobs makes it difficult for families to say no. Women's role in the success of data industries in India has been widely documented—a doctoral thesis form 2023 explores how crowdform platforms reinforce racialised and privatised logics of time, space, and worker subjectivity in the Indian labor market, similar to how the Indian state treats ASHA workers as explored in the sections above; a 2025 BehanBox report spotlights the invaluable female care and labor that drives data annotation industries; and a 2024 critically acclaimed feature film, Humans in the Loop,

tells a story of AI-enabled bias and injustice against the indigenous women who power it. Such gender-based assessments are not available for other labor markets in Asia.

It is pertinent to note that female participation in data work may also appear to be more inflated in tier 2 and 3 towns due to a recent uptick in 'tech for good' entrepreneurship in India. Start-ups travel to rural and semi rural pockets of India to advertise their job postings specifically to women—they get a "good tech" or "tech for society" label out of this, and women are able to obtain easy, semi-skilled employment. iMerit, one such India-based "fair trade" data services firm, heavily markets the fact that 52% of its workforce is female. There is not enough literature to support their fair trade and labor claims besides a social impact assessment from 2014.

Organizing in a fragmented industry

The first report explores a network of sociocultural factors impeding effective organizing and collective action in the region. While Asia is a hotbed of civic activity and organizing on digital platform labor issues, the conversation is currently heavily focused on location-based digital labor, which engages millions of informalized workers across the South and Southeast Asian region alone.

Conversations with informants suggest that organizing specifically around data work remains low—with the notable exception of the Philippines, which has built on its rich history with the BPO industry to create a robust, vast worker networks. However, as is evidenced in powerful mobilizations among factory workers in Vietnam, rural scheme workers in India, or gig workers in Indonesia—the most effective catalyst of change in Asia is mass mobilization and exerting pressure from the bottom up.

Organizing in these industries have been and will continue to be a lynchpin of worker welfare—efforts to legislate, regulate, and mediate relationships between platforms and workers can only ensure baseline protections against harm. Given the fragmented, complex, and rapidly evolving nature of digital labor, formal and informal communities and networks are much more than avenues of advocacy. This is where

workers exchange tactics for survival and success, post industry updates and job opportunities, warn each other about scams, threats, or policy changes, or simply, vent. These communities can take the form of loose groups on Facebook, Whatsapp, or Telegram. The mere existence of such spaces can build solidarity among a fragmented workforce, and help mitigate feelings of isolation and neglect that platforms and intermediaries deliberately build into these work arrangements.

> **According to I-1, a stubborn roadblock in organizing in the Global Majority is the cultural stigma around data work—that people employed in formal sectors see it as lazy, unfocused, unskilled work.**

The skillset required to perform complex AI tasks like moderation, labelling, training, testing, and prompt writing is very high—yet, due to how platforms present these job postings and relegate workers to the bottom of the tech supply chain, the public perception of the work remains low. The first report noted that the region is culturally more accommodating of tech jobs even if they are slightly gruelling—while true, this is merely a tertiary factor in Asian workers choosing not to organize.

First, and most compelling, is worker safety.

> **I-6 and two unnamed workers are Pakistani nationals living in Malaysia on work visas, working content moderation jobs at a well-known intermediary firm. They labelled their migrant status as the most significant barrier to not only fighting for their rights, but expressing themselves generally.**
>
> **In their bid for digital transformation, states in the region are turning hostile to tech worker organizing, especially in AI service industries. Speaking about work opportunities in Malaysia, I-5 noted the vast tax exemptions and visa incentives the country has started to roll out to promote a domestic AI industry.**

This attracts workers from South Asia—notably Pakistan and Bangladesh—to Malaysia, where they get subjected to constant policing, surveillance, and job precarity. Attempts to collectivise can cost them their safety without comprehensive legal recourse in their work country.

In the Philippines, BPO industry workers continue to mobilize actively in the face of union busting, red-tagging, violence, trolling, censorship, and a spectrum of harm. Mere attempts to voice one's opinions against the BPO industry attract a label of "armed rebels", which is not taken lightly by state authorities. I-3 suggests a neocolonial trade environment in the Philippines, where Filipino tech and economic policies cater generously to American clients because they continue to feel a deep cultural connection to the US. Advocacy campaigns in the Philippines' BPO Industry Employees Network (BIEN) have been setting the stage for the region, platforming open conversations on tech and labor despite termination threats from employers.

While there are hundreds of networks of location-based gig workers in South and Southeast Asia where these industries are continuing to boom, there is a lack of coordinated, formalized action on data work due to its fragmented and remote nature, structural non-transparency, and cultural leniency, as factors explored in the previous report.

Conclusion and the path ahead

What became clear through conversations with workers scattered across the region is that data workers in the Global Majority are not mere unsuspecting or powerless subjects of colonial, extractive, and exploitative labor arrangements. They are an empowered, deeply informed front working within their socioeconomic confines to ensure their own well-being and safety first, and towards better working conditions for their successors next.

The primary framing in the previous report was one of unfairness, power asymmetries, and inequity—furthered by global tech supply chain architectures and enabled by domestic policy environments apathetic to worker welfare. While all these environmental factors have been confirmed by the informants and existing scholarship on these industries, they are not the epicenter of labor justice—the workforce is. Workers' decision to join data work industries and contribute to building safer, stronger technological systems is a lawful and dignified manifestation of their right to freedom of profession and expression.

> **The informants who helped build this report reaffirm the dignity and significance of their labor, and their choice to continue working these roles. They believe that reform in these industries must begin with first acknowledging this informed choice, and then moving towards instilling equity and humanity in employment arrangements and job expectations.**

It is clear from these conversations that allyship in this context must always center and platform workers and their work as a whole—not their grievances alone. A few years ago, media outlets were fraught with personalized stories of tech labor injustice and unfair working conditions, centering the often sexually graphic and violent nature of content data workers have to sift through. This form of diagnosis attracts reader attention and intrigue. However, such attention can also fade away. Success in achieving labor justice in global tech supply chains goes beyond

merely diagnosing issues and framing them in larger sociopolitical contexts—it must include systemic, sustainable solutions like a) participating in the making and auditing of legislations, b) co-creating policy solutions and advocacy tools with data workers, and c) establishing grievance redress channels to file complaints and hold actors accountable.

Beyond this, in the lack of formal unions or organizing in data work, one can look to a) movements in other regions, and b) advocacy in adjacent industries like location-based platform labor for lessons and strategies. The Data Labelers Association (worker-led initiative designed to create a shared platform for data workers across Africa and beyond), Tech Workers Coalition (US and Africa chapters), Data Workers Inquiry are among many movements outside of Asia with direct policy engagements and successful campaigns. In location-based labor, scholarship emerging from Indonesia, India, the region, and the Global Majority on interpreting and proposing guardrails for algorithmic management is a significant avenue for joint advocacy, especially with crowdsourced data work which most closely resembles digital gig work.

The next report in this series will attempt to bridge some of these research gaps by building policy charters and checklists with inputs from data workers and groups, and aim to establish pathways for continued support and coordination with data worker groups to build a safe, dignified, and equal global tech labor ecosystem.

Annexures

Annex 1: Key informants

Alexis Chavez	Data Workers' Inquiry Project	Venezuela	I-1
Anora (alias)	Data worker (content moderation, data annotation, prompt writing)	Philippines	I-2
Christine Agan	Data worker; organizer at BPO Industries Employees Network (BIEN)	Philippines	I-3
Michael Beltran	Independent Journalist	Philippines, Taiwan	I-4
Ushar Daniele	Independent Journalist	Malaysia	I-5
Fahad (alias)	Data worker (content moderation, quality analyst)	Malaysia, Pakistan	I-6
Zuha Siddiqui	Independent Journalist	Pakistan	I-7
Unnamed	Data worker	Malaysia, Pakistan	
Unnamed	Data worker	Malaysia, Pakistan	
Unnamed	Data worker	Philippines	

Annex 2: Questionnaire for semi-structured interviews

1. Please tell me about your role and work with [intermediary/platform]. Are you working with [platform] through [intermediary]?
2. What is your contractual arrangement with [intermediary/platform]? Are you an independent contractor, and is this a full time role?
3. What attracted you to this work and what do you like the most about it?
4. What does a day in the life of a [content moderator/data annotator] look like? Take me through your day from when you log into work to when you log out, in as much detail as you can.

5. What kind of challenges do you face in your role on a daily basis?
6. Are the content guidelines provided by [intermediary/platform] adequate in your local contexts? Do moderators have a say in shaping or revising these? How often are they updated?
7. To what extent do you have to engage with an algorithm instead of a human manager? Are algorithms able to make decisions that impact you, and are you able to challenge them?
8. Dealing with graphic content has widely documented mental repercussions for [content moderator/data annotator]. What has been your experience with this?
9. Are you provided psychological support and other kinds of wellness checks like scheduled breaks by the company? Is there a counsellor? How effective are these measures?
10. What about job precarity? Is there a fear of termination, are lay-offs frequent? How is the attrition rate at [intermediary/platform], if you would know?
11. [Content moderation] With Meta laxing its content moderation policies in the US and mass lay-offs of content moderators worldwide, like Bytedance in Malaysia, what would you think the future of moderation looks like?
12. [Data annotation] Do you know what systems and projects you are training? Would you have an issue with contributing to certain projects that may not align with your ethics and values?
13. How is your relationship with your colleagues? Do you discuss work? Do you discuss any similar problems you may face?
14. Does [intermediary/platform] know if its contracted data workers join a union or otherwise collectivise? What is the response usually? Is this a barrier?
15. How do the NDAs signed by you affect your ability to seek support, collectivise, or demand reform? Have you been threatened with legal action through NDAs?
16. Now that you work for [intermediary/platform], which many consider a prestigious tech job, how do you upskill and what does progress in this industry look like for you? Is it easy to switch industries? Does working these roles limit your opportunities to shift laterally?

www.techglobalinstitute.com