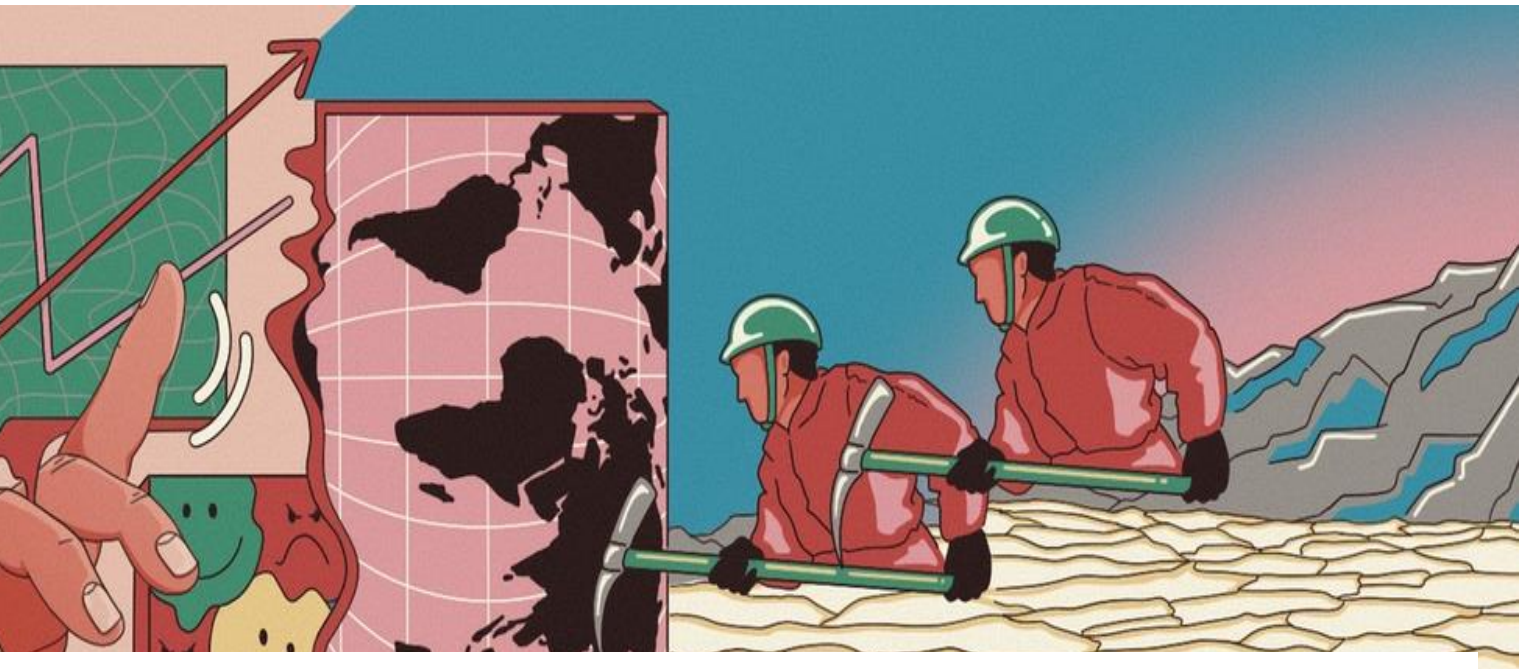

Invisible Workers, Visible Harms: Perils and Precarities of AI Labour

On behalf of
Federal Ministry for Economic Cooperation and Development

Published by
giz Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH

digital. global

Gig Economy

In cooperation with
aapti institute

IMPRINT

As a federally owned enterprise, GIZ supports the German Government in achieving its objectives in the field of international cooperation for sustainable development.

Published by:
Deutsche Gesellschaft für
Internationale Zusammenarbeit (GIZ) GmbH

Registered offices
Bonn and Eschborn, Germany

Address
Friedrich-Ebert-Allee 32+36, 53113 Bonn, Deutschland
T +49 2 28 44 60-0
F +49 2 28 44 60-17 66

E info@giz.de
I www.giz.de/en

Programme description:
Gig Economy Initiative
The Initiative advocates for fair and decent working conditions for platform workers at a global level
I www.bmz-digital.global/en/overview-of-initiatives/gig-economy
E gigeconomy@giz.de

Author:
Aapti Institute

Technical guidance:
Lukas Sonnenberg, María Paula Piñeros M.

Photo credits/sources:
Clarote & AI4Media
https://betterimagesofai.org
https://creativecommons.org/licenses/by/4.0/

On behalf of
German Federal Ministry for Economic Cooperation and Development (BMZ)
Economic and Social Development, and Digitalization Division
Bonn

Location and year of publication:
Bonn, 2026

On behalf of

Federal Ministry
for Economic Cooperation
and Development

Preface to the Series

A growing, global and largely hidden workforce rids our social media of toxic content, prepares data for AI development, and fulfils many other roles to keep our digital technologies going. These workers, called "data workers" and "content moderators," work along complex, transnational supply chains that span the globe. Despite their important contributions to technological development, they remain unrecognised and often operate under precarious working conditions. Workers tend to suffer from low pay and unpaid work to mental trauma from exposure to disturbing content.

As part of the growing discussion on the occupational hazards of data work and content moderation, Aapti Institute, in partnership with Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH, embarked on the "Exploring AI Labour in the Global South" project. We organised a series of stakeholder consultations, where people involved in labour, advocacy, research, journalism, and policy came together to discuss the problems of data work and content moderation. The primary objective of these consultations was to anchor deliberations in lived experiences and region-specific contexts, thereby moving beyond abstract or generalized discourse. Additionally, we also spoke to several practitioners involved in labour organising, research efforts, policy work, and media investigations. These conversations were centred around Sub-Saharan Africa, Southeast Asia, and South Asia – regions that are crucial to AI development and yet remain absent from policy discussion about labour welfare.

This three-part report series synthesises the findings from stakeholder consultations and secondary research, offering a distinct perspective by anchoring its analysis in the realities of the Global South. The first report explores the spectrum of precarious working conditions faced by content moderators and data workers. The second investigates the role of algorithmic management in shaping workers' lived experiences. The third and final report examines outsourcing and the transnational challenges of ensuring fair labour conditions.

Taken together, the reports are best read as a series as they address interconnected issues of global outsourcing practices, algorithmic oversight mechanisms, and on-the-ground working conditions while identifying points of intervention for state actors as well as bilateral and multilateral institutions. They also outline policy options and highlight ongoing initiatives that seek to improve working conditions. Collectively, the series provides a critical foundation for understanding the labour challenges associated with data work and content moderation in the Global South, serving as an evidence-based framework for future policy discussions and coordinated action.

Executive Summary

The global Artificial Intelligence (AI) ecosystem rests on a vast, unacknowledged and hidden workforce whose labour is indispensable yet systematically undervalued. This report examines the structural foundations of data work—the human annotation, labelling, transcription, and content moderation tasks that underpin the development, training, and maintenance of AI systems. Although often presented as automated or immaterial, AI systems rely on millions of human workers concentrated disproportionately in the Global South, where constrained labour markets, high youth unemployment, and weak regulatory safeguards create conditions conducive to the large-scale outsourcing of data-intensive tasks.

Across South Asia, Southeast Asia, and Sub-Saharan Africa, data workers operate within fragmented and opaque supply chains characterised by informal employment arrangements, limited social protections, and a profound imbalance of power between workers, vendors, and the global technology firms that ultimately benefit from their labour. Many workers are engaged through short-term or micro-contracts, lack written terms of employment, and remain outside national labour law protections due to subcontracting structures designed to externalise risk. These conditions are compounded by intensive algorithmic management systems that monitor keystrokes, errors, and productivity in real time—practices that heighten stress, restrict autonomy, and render the labour process increasingly punitive.

For content moderators and workers handling sensitive data, the risks are particularly acute. Repeated exposure to violent, abusive, or graphic material contributes to severe mental health harms, including anxiety, depression, insomnia, and symptoms consistent with post-traumatic stress. Yet access to trauma-informed occupational health support remains extremely limited, and NDAs routinely prevent workers from reporting harms or seeking recourse. Despite these structural obstacles, workers are developing new forms of solidarity and resistance.

Drawing on desk research, secondary data, and stakeholder consultations across South Asia, Southeast Asia, and Sub-Saharan Africa, the report investigates the organisation, conditions and governance of this growing workforce. The report concludes with comprehensive recommendations spanning fair employment protections, occupational health standards specific to high-risk data work, transparency requirements for supply chains, mechanisms for collective worker voice, and binding international standards to prevent regulatory arbitrage.

Acknowledgments

This project was supported by the insights and perspectives of numerous people that we consulted over the course of our research and convening efforts. We are deeply grateful to the workers, labour organisers, researchers, and journalists for sharing their experiences and time. This project could not have materialised without their openness and their contributions to ongoing efforts to change data work and content moderation for the better.

We would also like to thank the GIZ Gig Economy Initiative for creating the space for our research and for the assistance they have extended. We are grateful to Lukas Sonnenberg, María Paula Piñeros, and Maren Bernloehr for setting up this project and their continued advice and support.

Several people at Aapti came together to develop this report series and the discussions that informed it. Our thanks go to Sreya Nair for researching and developing this series' first report. We thank Sarayu Natarajan, Soujanya Sridharan, Priyam Vadaliya, and Somya Singh for their invaluable contributions to and guidance for this project.

Please see Appendix A for the list of experts consulted.

Table of Contents

1 | Introduction

Artificial intelligence (AI) now carries the promise of underpinning a wide range of digital systems— from social media and content platforms to autonomous vehicles and large-scale language models. However, beneath these technologies lies a global supply chain of human labour that remains largely invisible. This labour, commonly referred to as *data work*, encompasses the collection, curation, annotation, moderation, and evaluation of data that make AI systems functional.

Data work spans multiple stages requiring continuous human intervention. Workers collect and clean text, images, audio, and video; annotate sentiment, intent, and objects for training machine-learning models; and evaluate outputs to ensure accuracy and safety. These tasks are not peripheral, they are central: without them, AI systems cannot operate reliably or adapt to new contexts. Despite narratives of automation, the performance and integrity of AI depend on iterative human feedback loops. Here is where we see the entry of data work, a highly labour-intensive service where the focal responsibility of this labour rests in the processing of data which involves the collection, curation, annotation and evaluation of data (Muldoon et al., 2024). These tasks are not marginal; they are what enable AI systems to function.

This labour force is geographically dispersed but structurally concentrated in the Global South particularly Africa, South Asia, Southeast Asia, and Latin America. Technology firms in the Global North routinely outsource data operations to digital labour platforms and BPOs such as Appen, TELUS International, Remotasks (Scale AI), Sama, or Teleperformance, taking advantage of lower labour costs, extensive informality, and favourable regulatory environments. Workers, often motivated by economic insecurity or limited local job prospects, perform high-value digital tasks under conditions marked by low pay, algorithmic management[1], weak protections, and limited transparency.

Firms frequently treat this labour as an interchangeable computational resource rather than work deserving of rights, protections, and recognition. Scholars frame this model as a form of digital labour extraction, where technological advancement is built on the availability of low-paid human input (Verma, 2025).

This report, in congruence, positions data work as a critical component of the political economy of AI. By examining who performs this work, where they are located, and the conditions under which it occurs, the report highlights both the indispensability of data workers and the structural inequalities shaping their roles. It argues that the trajectory of AI development will depend not only on technical innovation but also on how this labour is organised, governed, and valued. The following sections outline why data work is

[1] For a detailed analysis of algorithmic management, refer to the report *Engineered precarities*, also in this series.

foundational to AI, why it remains characterised by precarity, and how global value flows reproduce longstanding patterns of uneven extraction and risk.

1.1 | Data Work as the Foundation of AI

Against the backdrop that's been established of globalised, outsourced digital labour that sustains today's AI boom, it is vital to understand the status of data work (Brandom, 2025). As mentioned earlier, the increasing proliferation of generative AI and other high-capacity systems has led to an insatiable demand for AI training corpora, such that industry analysts warn of a looming "data wall", a point at which the availability of high-quality, labelled data becomes a bottleneck for further model improvement (Villalobos et al., 2024).

While companies are experimenting with alternatives such as synthetic data generation and data-efficient training methods, these approaches are still dependent on human labour for validation, refinement, and quality assurance. Synthetic datasets often reproduce existing biases or lack the nuance and contextual accuracy that only human annotation can provide (Shanley, 2025). Moreover, human labour continues to offer a cheap, scalable, and flexible solution for producing and verifying the vast amounts of data AI systems require. As models scale, the availability of good-quality labelled data stands as an impediment to further progress. AI systems rely on vast quantities of data, making data work a central component of their development and operation. Machine learning (ML) models, thus, depend extensively on human intervention in the form of data work.

Data work typically follows a multi-stage pipeline. First, workers collect and curate large volumes of text, audio, images, and video from open or proprietary sources, ensuring datasets are complete and usable. Annotation then transforms this material into training data: workers classify sentiment and intent in text for large language models, draw boundaries around objects for autonomous driving systems, and transcribe and tag speech to refine voice technologies (Liu et al., 2021). Each annotation constitutes a micro-decision that embeds human judgement into computational systems. After model training, evaluators and verifiers assess outputs for accuracy, safety, and coherence. Because models experience performance drift or encounter new contexts, these processes are cyclical; human oversight is continuously reintroduced to recalibrate systems and maintain reliability.

This labour is carried out by a global workforce concentrated largely in the Global South, where favourable policy environments, lower labour costs, and high levels of informality encourage outsourcing. Major technology firms including social media companies, AI labs, and cloud service providers routinely contract this work through digital labour platforms (e.g., Amazon Mechanical Turk, Appen) or BPO firms (e.g., Sama, CloudFactory, Teleperformance). Smaller regional platforms such as Remotasks, Microworkers, and

TELUS International also play a significant role in supplying labour for annotation and evaluation tasks.

While essential to AI production, data workers often operate under informal conditions marred by low compensation, limited social protection, and minimal recognition of their contribution, as observed through multiple investigations and studies (Tubaro et al., 2025; Coldeway, 2024; Toxtli et al., 2025). Despite these challenges, data work is frequently framed as a pathway into the digital economy for workers facing unemployment or wage stagnation, offering access to foreign currency and supplementary income. However, without protective regulation, this work can reproduce long-standing patterns of uneven value extraction observed in other outsourced sectors (for instance, in the global apparel industry; Collins, 2003). Understanding these dynamics is essential for designing labour standards and governance frameworks that ensure fair and safe working conditions as AI systems expand. The following section situates these issues within the broader global geography of production, outlining how existing structural arrangements influence the distribution of risks and benefits in AI supply chains.

1.2 | Profit and Labour in the Global Labour Geography

As mentioned before, it is not a mere coincidence that the global geography of production is clustered in the Global South, reflecting deep structural power asymmetries in the political economy of AI (Miceli & Posada, 2022). This spatial division of labour, amplified by the legal, social, and geographic distance that firms maintain from their workforce mirrors global production systems where economic value and intellectual property accrue to technology firms, while the burdens of repetitive, poorly remunerated tasks are distributed across dispersed and often vulnerable populations (Riccio et al., 2025; Gray & Surie, 2019). This reality reflects broader structural inequalities within global digital supply chains. As observed by an expert[2], value generated by data workers in the Global South is transferred to technology firms in the Global North, concentrating and maximising profit while externalising precarity and risk. *'Heteromation'* encapsulates this dynamic form of labour; coined by Ekbia and Nardi (2017), this word refers to the capital accumulation enabled by the extraction of economic value from cheap, computer-mediated human work.

AI firms derive profit through two interrelated processes: the use of global wage differentials and the transformation of data into proprietary assets. Labour arbitrage enables companies to reduce operational costs by outsourcing data annotation and content moderation to lower-income regions characterised by limited worker protections and high informality (Tubaro et al., 2025; Casilli et al., 2024). Concurrently, the data

[2] Expert Interview, Anonymised

produced, cleaned, and classified through this labour becomes a critical input into model development and commercialisation. Yet the workers who generate this input typically remain outside the domains of ownership, recognition, and profit-sharing, creating a structural gap between value creation and value capture that underpins the contemporary political economy of AI (Muldoon et al., 2023). Lead firms hold significant power in setting contractual terms, defining work standards, and determining compensation, while workers—often dispersed, informal, or platform-mediated—have limited bargaining power. This creates a structural imbalance rendering workers subordinate in the labour relationship and dependent on the decisions of firms that control access to work and remuneration (Pilatti et al., 2023). This dynamic constitutes a contemporary form of '*accumulation by dispossession*,' (Harvey, 2003) where the intellectual and material labour of peripheral populations is extracted under conditions shaped by historical inequalities, facilitating the expansion of AI infrastructure and services without commensurate benefits to those performing the foundational work (Mudloon & Wu, 2023). The spatial and social organisation of AI labour thus mirrors broader patterns of global capital circulation, where peripheries are structurally necessary yet systematically undervalued.

Understanding AI's reliance on labour is not only an economic question but also a political and ethical one. Policy discussions often focus narrowly on the downstream effects of AI deployment such as bias, fairness, or transparency without interrogating the upstream production processes that make AI possible. To centre data work in these debates is to foreground the conditions under which AI is produced, the inequalities it reproduces, and the need for regulatory and institutional frameworks that acknowledge and protect those whose labour sustains the system.

However, in order to grasp how exploitation and risk are distributed, we must first look into contours of the workforce itself—who performs this work, where they are located, and how their labour is segmented and valued. The next section takes up this task by examining the demographic and global distribution of data workers and the organisational structures that sustain AI's hidden labour economy.

2. Landscape | Understanding the workforce

The global data workforce is demographically and geographically diffused in ways that present a pattern that reflect not only contemporary economic susceptibilities, but also historical continuities of coloniality. Age, race, and geography shape who participates in this labour market, while the North-South divide ensures that value is disproportionately extracted from the Global South—Africa, South Asia, Southeast Asia and Latin America primarily (Posada, 2022).

This section examines the global landscape of data work, mapping who performs this labour, where they are located, and how their work is organised. It begins by establishing data work as the foundation of AI (2.1), before analysing the demographic profile of data workers and the historical patterns that shape global labour arbitrage. The section then maps the geographical distribution of this workforce (2.2), identifying key regional hubs in South Asia, Southeast Asia, Sub-Saharan Africa, and Latin America, exposing the North-South divide. Finally, it provides a detailed typology of business models (2.3) that structure data work, from platform-based crowdsourcing to BPO arrangements and emerging hybrid models, examining how each moulds labour conditions, worker vulnerability, and possibilities for collective action.

2.1 | The Demographics of Data Workers

The majority of data workers globally are young, typically between the ages of 20 and 35 years old, with a significant proportion holding higher education degrees (Graham et al., 2017). This composition of data workers includes a large number of university graduates or early-career professionals unable to find employment in their local economies. In countries like India (Naseer, 2024), where youth unemployment persists at high levels despite substantial gains in educational attainment, this concentration of data workers signifies structural labour-market rigidities that limit the absorption of skilled young workers into appropriate forms of employment (Tasnim, 2025). This overrepresentation of educated youth highlights a dichotomy: skills that could be utilised in professional or knowledge-intensive industries are instead being channelled into the tedious labour of repetitive annotation, labelling, or transcription tasks.

Here, the global labour arbitrage becomes visible — young, educated populations in the Global South are systematically undervalued compared to their counterparts in the North. Major technology firms take advantage of constrained employment opportunities by relocating low-status, labour-intensive tasks. Tasks that would be rejected by workers in high-income regions such as Silicon Valley are systematically reassigned to peripheral labour markets at a fraction of the cost (Williams & Miceli, 2023). These peripheral markets form the backbone of what can be described as AI's labour periphery.

2.2 | Global South as AI's Labour Periphery

This "periphery" constitutes workers ranging from India, the Philippines, and Kenya who dominate the lower-paid segments of the digital labour market. South Asia, especially India, Pakistan and Bangladesh remain a central node (Dewan et al., 2022). India's mature Business Process Outsourcing (BPO) industry, extensive English proficiency, and abundant graduate labour supply make it a vital key to the AI value chain. In Southeast Asia, countries like the Philippines, Indonesia, Vietnam, and Malaysia have expanded from call centre work into large-scale annotation and content moderation. The Philippines hosts tens of thousands of moderators for Western platforms, a workforce that is shaped by linguistic proximity and post-colonial ties to the United States (Abuel, 2024). In Sub-Saharan Africa, Kenya, Ghana, Nigeria, and South Africa form the principal hubs (Anwar & Graham, 2022). Nairobi's "Silicon Savannah" houses major outsourcing hubs and vendors supplying global firms like Meta, Samasource, and Scale AI (Mkalama & Ragnet, 2025).

Latin America has also emerged as a significant region, with Venezuela and Argentina contributing large numbers of workers (Tubaro et al, 2024) due to persistent economic crises that pressure skilled individuals to turn to online platforms (Lenis, 2023). Geographically, an overwhelming portion of digital labour is concentrated in the Global South but also present in the United States, Canada, and the United Kingdom, though more often as supplementary income seekers. This produces a labour market where North and South workers directly compete with each other (Braesemann et al., 2022) .

Yet, structural asymmetries mean Global South workers remain price-takers while North-based workers leverage local protections and higher baseline wages (Anwar & Graham, 2022). For example, while contractors in the United States earned around $15 an hour for tasks supporting the development of ChatGPT (Ingram, 2023)—a rate already considered low in terms of Global North standards, workers in Kenya performing similar work for the same company were paid under $2 an hour (Perrigo, 2023)—revealing the drastic wage disparities embedded in the global AI supply chain.

These geographical concentrations and disparities reflect the organisational structure of data work. Digital labour platforms and Business Process Outsourcing (BPO) firms function as intermediaries that channel tasks from Global North technology companies to low-cost labour markets in the Global South. These business models shape how economic value and labour are allocated across AI production networks.

2.3 | Mapping of Business Models

Drawing mainly from Muldoon et al.'s (2024, pp. 4-8) typology, this section delves into a mapping of the key business models structuring data work across the Global South.

Across South Asia, Southeast Asia, and Africa, data work has consolidated around two dominant institutional models: platform-based systems and BPOs that represent a specific logic of organisation, control, and accountability. Insights from our stakeholder consultations reinforced this understanding, revealing how these models manifest across different regional and sectoral contexts. Drawing on academic literature, we have identified six organisational models, which can be understood along a spectrum of employment formalisation and management control—from geographically dispersed platform-based contracting to concentrated facility-based employment. Each model produces distinct patterns of labour control, worker vulnerability, and possibilities for collective organizing.

- **Platform-based models**: Under this model, digital labour platforms like Mechanical Turk, Appen, Clickworker, Microworkers, and Remotasks function as online marketplaces where AI companies post tasks such as image tagging and text annotation to be completed by geographically-dispersed workers categorised as independent contractors. These platforms have managed to garner significant participation across the Global South delivering apparent flexibility and access to foreign currency.

However, this model represents the most atomized and informal arrangement (Natarajan & Mohamed, 2021, p.9). Workers face several systemic shortcomings under this model, with payments typically being task-based and earnings that often fail to meet minimum wage standards. Most receive no training to meet AI-specific quality standards. Due to their classification as independent contractors instead of employees, these workers are further excluded from basic formal labour protections and denied social security, with most of them earning below the poverty line (Carmona, 2025) and experiencing frequent instances of wage theft with no effective recourse (Howson et al., 2025).

- **Specialised AI data platforms**: Platforms such as the ones mentioned above also offer sophisticated services, requiring workers to complete training modules before accessing higher-skilled tasks such as semantic segmentation and LiDAR annotation. These platforms offer somewhat higher pay and more consistent work, particularly attracting workers in regions such as the Philippines, who develop expertise in specific annotation types. However, vulnerabilities still persist in terms of the lack of guaranteed income, unilateral platform control over task allocation and payment, opaque performance evaluation, and no collective representation. In this case, specialisation benefits platforms and clients rather than translating into substantive worker security or protection.

- **BPO models:** Business Process Outsourcing (BPO) refers to the method of subcontracting third-party service providers to perform specific business functions on

behalf of the requester company. They act as intermediaries in the employment relationship, mediating between the clients and workers. BPOs, unlike digital labour platforms, typically employ workers in physical facilities with an established hierarchical management structure offering a wide variety of services, managing several stages of the client's data needs such as human resources (HR), logistics, and customer service along with AI data work services. Workers under this model are mostly employed on short-term contracts, receiving a marginal level of stability as compared to those working via crowdsourced platforms. As they have an organisational structure in place, workers are able to operate under direct supervision enabling employers to exercise firm control over performance, attendance, and productivity, through a tiered system of reporting and digital monitoring tools.

- **Specialised AI data BPOs**, on the other hand, focus exclusively on providing data services for ML systems, developing proficiency in specific AI applications such as computer vision, natural language processing (NLP), or autonomous vehicles. AI data BPOs like Sama, operating in Kenya and Uganda, and Cloudfactory, based in Nepal, exemplify this model. These companies provide targeted training for workers in specialised annotation processes while additionally developing proprietary platforms to manage complex projects. As opposed to platforms and generalist BPOs, AI data BPOs position themselves as suppliers of higher quality, more secure data services.

  With this structure, workers typically have access to clearer channels of communication and physical co-location allows for the formation of a peer support network creating a baseline foundation for worker solidarity and organising. While this arrangement has proven to be more advantageous than that of platforms, having more control over labour can translate into enhanced monitoring, rigid scheduling, and limited worker autonomy. The existence of multiple managerial layers such as team leaders, quality auditors, and client-facing supervisors leads to multiple layers of oversight that keeps workers subject to constant scrutiny. Hence, they are continuously observed, tracked, and assessed through both digital metrics and human supervision, essentially replicating and reproducing algorithmic management through bureaucratic management structures.

- **In-house arrangements; internal services and platform models:** Within this model, technology firms like Microsoft, Google, and Meta mirror crowdsourcing platforms but operate within controlled corporate ecosystems. These internal infrastructures, including Microsoft's Universal Human Resource System (UHRS), enable company staff to post microtasks for completion by external workers recruited through third-party vendors. Although this model offers elevated data security and confidentiality, it also institutionalises distance between firms and the individuals performing essential data work. Workers engaged through these systems remain

classified as independent contractors—confined to NDAs, and excluded from employment benefits.

- Some companies have also established **internal data service teams**, employing limited numbers of full-time or long-term contractors to perform sensitive or high-security tasks. While these arrangements permit a closer engagement with engineering teams, their scope remains restricted due to higher labour costs in company headquarters. Consequently, the vast majority of large-scale annotation and moderation continues to be outsourced to the Global South owing to the lucrative purpose it fulfils. Even within internal operations, data workers frequently occupy tangential positions in organisational hierarchies, confronting limited mobility, high performance pressures, and the same structural precarity that characterises outsourced data work.

- **Hybrid platforms (Emerging):** A recent and notable development is the convergence of location-based gig work and data work—an emerging model that complicates traditional distinctions between physical and digital labour. Uber has recently rolled out several initiatives that come at the intersection of physical and digital labour. In the U.S., Uber now pushes drivers to earn additional income (Nieva, 2025) by completing "digital tasks" that include recording voice samples, uploading images, or submitting documents through the driver app, as part of Uber AI Solutions' expansion into AI data services (Lung, 2025). In September 2025, Uber AI Solutions launched a similar pilot programme in India that enables ride-hailing drivers across 12 cities to complete digital tasks like data labelling (Swain, 2025). This experiment represents a hybridisation of data work and location-based platform work, effectively merging two precarious economies. While it may increase workers' earning potential, it also raises complex regulatory and ethical questions about data privacy, working hours, algorithmic supervision, and the blurring of boundaries between location-based work and digital work. It exemplifies how the next phase of AI's labour infrastructure may further integrate disparate forms of contingent work under a single platformised logic.

With the mapping of the dominant business models that organise data work, it becomes essential to examine how these models and their organisational logics materialise in practice. Having provided a cursory glance at some of the challenges that the workers face under these business models, the next section expands on the subject and delves deeper into how these models reflect on the lived realities of workers and the conditions that define their everyday labour.

3 | Labour challenges and working conditions of data workers

Now that we understand what classifies as data work and who manages data work, we need to direct our attention to how data work actually unfolds in the realm of reality. We know data workers perform the critical human labour that allows AI to function. Yet the conditions under which this labour is performed remain deeply perilous. Across regions such as Africa and Southeast Asia, expert interviews and stakeholder consultations point to consistent levels of precarity within data work and content moderation. Low and inconsistent pay, weak contractual clarity, limited access to support services, and routine exposure to potentially harmful or distressing material, particularly in content moderation are some of the recurring issues. These conditions contribute to financial instability, stress, and health risks for many workers. Taken together, the findings highlight structural gaps in working conditions and indicate areas where clearer standards, safeguards, and support mechanisms may be needed.

3.1 | Perils of Data Work and Content Moderation

Data workers, as mentioned earlier, are typically crowdsourced through BPO firms or digital platforms (Muldoon et al., 2024, p. 2). These vendors or intermediaries enable global technology companies to connect to local labour markets where they can access cheap labour and derive value. Workers across Southeast Asia, Africa, and South Asia report being hired with fragile contracts, on short-term agreements of less than six months, or on a per-task basis through platforms that classify them as independent contractors rather than employees (Muldoon & Wu, 2023, p. 12). For instance, in Kenya, the practice of offering fixed-duration contracts deliberately set to expire before workers become eligible for regular employment benefits categorically excluded them from minimum labour rights and social protection. Workers are often threatened with replacement, putting them under constant fear of termination, with no assurance of continued income and limited ability to plan for the future.

The consequences extend beyond job insecurity. Informally employed data workers based in the Global South receive exceedingly low wages, in some cases, the equivalent of USD 2 per hour (Arya, 2025) and in some cases USD 1.50 per hour (Gebrekidan, 2024, p.14) . Additionally, they have minimal or no non-wage benefits such as paid leave, sick leave, and social security contributions, and can be hired and fired according to firm needs without any notice or compensation.

Workers have detailed extreme levels of policing around attendance, performance, and quality of output (Equidem, 2025). They are required to complete tasks in accordance with requester specifications, with adherence influencing continued platform access and

remuneration. Consultations and expert interviews indicate that data workers often experience stress and uncertainty due to strict monitoring and frequent performance evaluations. Concerns raised by workers are typically routed through muddled managerial hierarchies, which can limit the visibility of issues at higher levels, leaving many concerns unresolved (Miceli & Posada, 2022). The labour process itself is structured around multiple productivity and efficiency targets that drive workers to process vast quantities of data at a speed that is incompatible with careful review and psychological safety. These targets are enforced through a sophisticated apparatus of monitoring and penalty. Workers operate under constant surveillance through CCTV cameras and client management tools that track every action and movement, measure their "average handling time" and automatically flag periods of inactivity (Equidem, 2025, pp. 26-27). In BPO environments, this digital oversight is reinforced by physical supervision and CCTV monitoring within office spaces. Together, these mechanisms form an intensive regime of control where every pause or deviation from performance targets is tracked and penalised.

Insights from consultations and expert interviews have revealed that the intensity of the work environment is particularly acute for content moderators and data annotators. Workers report having only seconds to process individual items within queues that often extend to hundreds or thousands of cases per shift, producing decision fatigue. Performance that falls below algorithmically determined thresholds triggers a cascade of repercussions such as wage deductions, forfeiture of bonuses that account for 50 to 70 per cent of their total compensation, mandatory unpaid overtime to meet quotas, and formal disciplinary actions followed by the ultimate act of termination (Equidem, 2025. pp. 27-28). Workers are also often mandated to maintain high rates of accuracy of 98% to 99%; failure to do so may result in job loss or the need to work additional hours as compensation.

One of the key applications of data-driven technologies is this system of monitoring and managing worker performance, which on its own carries no harm. But when an employer uses electronic monitoring and algorithms to minutely track and relentlessly push workers to achieve greater levels of efficiency, negative effects can quickly make themselves felt (Bernhardt et al, 2023). Routine practices of surveillance and oversight along with having to manoeuvre day-to-day work intensification only adds to workers' mental and physical debilitation.

This leads to arbitrary structures marked by opacity and a lack of contextual sensitivity where workers rarely have access to the rationale behind performance scores or the mechanisms that determine their evaluation (Tan & Cabato, 2023). Workers have limited visibility into how performance thresholds are set or adjusted, and even minor deflections can trigger penalties without opportunity for dialogue or contestation. Such automated rigidity creates an environment in which success is contingent less on effort

or skill and more on opaque algorithmic processes, producing a form of work tethered primarily to economic necessity rather than formal managerial directives.

The resulting pressures lead workers to engage in compensatory behaviours that compromise health, including skipping breaks, delaying meals, and working despite physical or psychological distress. These effects are aggravated with the use of team-based performance metrics, particularly prevalent in the Philippines, where entire teams are penalised—including forfeiture of off days and holidays when individual members fail to meet targets (Equidem, 2025. p. 31). This approach effectively transforms colleagues into enforcers of productivity discipline, heightening the risk of workplace bullying and emotional abuse while diminishing opportunities for peer support.

With the sector's use of non-disclosure agreements (NDAs), this climate of fear is rendered fortified (Shivakumar & Bhattacharjee, 2025). Participants from our consultations[3] noted that NDAs discourage workers from disclosing personal or professional concerns related to their working conditions or their mental and physical health. This contributes to a culture of limited transparency and reduces opportunities for collective action. The use of NDAs in this context can have the effect of minimising organisational accountability. By restricting disclosure, these mechanisms can limit external oversight and public scrutiny, enabling firms to maintain plausible deniability in situations involving workplace harm.

Regional Insights: Across the regional consultations, these specific challenges manifested differently but with a common structural pattern.

- **Africa:** Stakeholders who took part in this consultation[4] emphasised that fragmented employment relationships, often involving multiple layers of subcontracting made it difficult for workers to identify who was responsible for wages, grievances, or mental health safeguards. They noted a lack of transparency, with workers not being told what the data is for, how it will be used, or the risks involved. Annotation tasks (e.g., drawing polygons, bounding boxes) are assigned without explaining dataset purpose or downstream applications.
- **Southeast Asia:** Participants[5] highlighted that governments' prioritisation of AI expansion as an economic growth strategy often sidelines labour protections, permitting firms to operate with substantial opacity in contract structures, performance expectations, and worker rights.

[3] Stakeholder Consultation 1
[4] Stakeholder Consultation 1
[5] Stakeholder Consultation 2

- **South Asia**: Experts[6] described widespread misclassification of data workers as "independent contractors" for platform tasks, a categorisation that legally excludes them from minimum-wage protections, employee benefits, and social security.

Together, these insights reveal that while the forms of precarity appear to be similar across regions, the *regulatory blind spots* through which they operate vary significantly, shaping both worker experiences and the feasibility of redress. While organisational gaps shape the structural conditions of precarity, they also intersect with the everyday experiences of workers, influencing their psychological well-being. The next section explores these mental and emotional dimensions of digital labour.

3.2 | Psychological tolls

The high-intensity nature of AI-mediated data work produces marked physical and psychological consequences. Data workers, especially content moderators are exposed to egregiously graphic and distressing content daily, which manifests in both somatic and mental health challenges (Bhattacharjee & Shivakumar, 2025). They describe confronting acute physical symptoms including chronic fatigue, musculoskeletal strain (back and wrist pain), eye strain, hair loss, weight fluctuations, and stress-induced dermatological conditions such as psoriasis. These symptoms result from an amalgamation of prolonged computer use, inadequate rest periods, and the cumulative physiological effects of sustained periods of anxiety and secondary trauma (Gonzalez & Matias, 2024, pp. 4-5).

Psychological impacts are particularly severe. Stakeholder consultations[7] with former content moderators underscore the persistent and long-term nature of these effects, including difficulties in disengaging from traumatic material outside work hours and impaired reintegration into their personal and social lives (Dinika, 2024).

Moderators routinely face a stream of gruesome material ranging from gory road accidents, suicides to child pornography, animal cruelty, homicides and more for 12 hours and seven days a week (Siele, 2023). Such exposure has been associated with high incidences of post-traumatic stress disorder (PTSD), depression, anxiety disorders, panic attacks, insomnia caused by nightmares, emotional blunting, social withdrawal and in certain drastic cases, suicidal ideation and attempts (Wilmot & Hall, 2025). Several of the experts[8] that participated in our stakeholder consultations[9], which also comprised former content moderators, emphasised the severity of the mental health crisis faced by these workers across multiple geographies. Most of them noted workers experiencing lasting struggles of paranoia and difficulty reintegrating into society. As one expert[10] pointed out,

[6] Stakeholder Consultation 3
[7] Stakeholder Consultation 1
[8] Expert Interview, Anonymised
[9] Stakeholder Consultation 1
[10] Expert Interview, Anonymised

the nature of this form of work is such that it often spills over and seeps into the personal lives of workers, with many of them struggling to completely disengage from the material they handle, even during non-working hours.

The isolation enforced by NDAs worsens the psychological harm by preventing workers from processing their experiences through conversation with family, friends, or even mental health specialists or therapists who might not understand the specific nature of the trauma without detailed disclosure. Even when in-house counselling is formally available, workers often refrain from seeking assistance due to fear of reprisal or job loss. In Kenya, for instance, employees experiencing psychological distress risk termination or loss of income security, bolstering a cycle in which mental health challenges are both produced and punished (Booth & Kimeu, 2024; see also Reuters, 2022). In Indonesia,[11] workers who attempt to use the available mental health resources face penalties. Taking wellness breaks counts against productivity metrics or must be compensated through unpaid overtime.

Existing measures to safeguard workers from such harms have only proven to be limited and ineffective (Wilmot & Hall, 2025). For example, experts[12] from the consultations mentioned that group counselling is offered to workers in regions such as Indonesia. However, these sessions are often generic and templated, failing to address trauma or grasp unique stressors of data work, leaving workers without meaningful psychological support. The repeated exposure to disturbing materials creates specific forms of trauma, especially for women workers—as one expert[13] also highlighted who constitute most content moderators. Research establishes that content moderation and the review of graphic content have a more acute impact on women than on men, as the most violent sexual content depicts violence against women, and clinical studies demonstrate that exposure increases anxiety, depression and PTSD (Superrr Labs, 2024). This forces women to turn to therapy or taking extended breaks before they can re-enter the job market due to the extreme attrition affecting their ability to do cognitive labour, while bearing the costs of such treatments out of their own pockets (Espíndola, 2025).

Regional Insights: While psychological harm was a consistent experience across regions, consultations revealed important variations.

- **Africa**: Many reported that trauma was compounded by the absence of specialised mental health practitioners familiar with content moderation-related distress, as well as by the stigma associated with psychological conditions.

[11] Stakeholder Consultation 2
[12] Stakeholder Consultation 2
[13] Expert Interview, Anonymised

- **Southeast Asia:** Participants described employer-provided wellness programmes as largely generic, templated, and insufficiently equipped to address trauma arising from exposure to graphic material.
- **South Asia**: Experts noted that mental health is rarely recognised as an occupational concern within digital labour ecosystems, limiting access to institutionally supported counselling or trauma-informed care.

These harms are compounded by structural features of the industry that limit workers' ability to seek assistance, raise concerns, or collectively respond to risks. The following section outlines how these dynamics of silence and fragmentation unfold.

3.3 | Silencing of Data Workers

The atomisation and silencing of data workers is yet another mechanism that perpetuates exploitative conditions. NDAs that exist ostensibly to protect intellectual property function primarily to prevent workers from disclosing any details regarding their working conditions and environment, seeking external support, or organising collectively. When workers attempt to organise despite these barriers, they are faced with swift retaliation. In Kenya, 43 content moderators hired by Sama filed a lawsuit against Meta alleging that they were terminated and blacklisted for attempting to form a union (Business and Human Rights Resource Centre, 2025). Similarly, hundreds of workers under Google were fired for protesting issues such as pay and job insecurity (Bansal, 2025). Similarly, in Ghana, around 150 content moderators working for Meta contractors, including Majorel, have raised claims regarding exploitative working conditions, exposure to graphic and harmful content, inadequate mental health support, and unfair dismissals (Hall & Wilmot, 2025).

The geographic and organizational fragmentation of data work further impedes organizing. Workers are distributed across multiple sites and countries, employed through different contractors and platforms, often working remotely, with high turnover that prevents the stable relationships necessary for building collective power. One expert[14] emphasised that "many countries remain openly hostile to labour organising, especially in BPO and AI-related workforces" where "workers live under constant threats of dismissal, blacklisting, or contract non-renewal". In Kenya, under the proposed Business Laws (Amendment) Bill 2024, major technology companies such as Meta Platforms Inc. and Google LLC would be shielded from liability for labour and human-rights violations committed by their BPO contractors (BHRRC, 2025). With platforms insulated from legal action, workers are prevented from holding the ultimate beneficiary of their labour accountable. This undermines established employer-worker liability structures and limits redress.

[14] Expert Interview | Anonymised

Regional Insights:

- **Africa:** Stakeholders highlighted that NDAs and confidentiality clauses are frequently interpreted by managers to prohibit even informal conversations about distress. Workers reported that speaking about traumatic exposure, even with colleagues was discouraged, reinforcing emotional isolation. They also affirmed that attempts at collective organising are particularly vulnerable to retaliation. Participants highlighted concrete examples where workers attempting to form unions or raise grievances were dismissed, blacklisted, or subject to contract non-renewal.
- **Southeast Asia**: Experts pointed out that the political environment in several countries creates additional barriers due to restrictions on labour organising in "strategic" digital sectors, including AI and BPO operations.
- **South Asia:** Respondents noted that high turnover, dispersed work arrangements, and a dependency on platform ratings make it nearly impossible for workers to collectivise without exposing themselves to economic loss.

Together, these conditions create a work environment in which responsibility for risks is largely borne by workers, while accountability is distributed across multiple actors with limited interventions for mitigation. This configuration reflects broader systemic patterns, which are examined in the following section.

3.4 | The Human Cost of Automation

Consultation findings reinforce that exploitation reflects not isolated failures of specific employers, but the essential structural logic of how AI development is organised within the global market economy. The fragmentation of employment relationships, the pressure of algorithmic management[15], the inadequacy of health protections, the suppression of collective organisation, and the systematic undervaluation of essential labour all serve the function of enabling technology companies to easily access the human intelligence required for machine learning while displacing the costs and burdens onto the most vulnerable workers in their supply chains. As one expert[16] stated, the industry has created and entrenched "a tech oligarchy"—a handful of firms who exert global dominance while evading obligations (see Hancock, 2025).

The human cost of data work is not inherent to the nature of the work itself (Dinika, 2024). Rather, it reflects the current configuration of global AI supply chains and the absence of adequate safeguards, regulation, and institutional support. At the same time, AI systems continue to rely fundamentally on human judgment, contextual reasoning, and evaluative capacity, positioning data workers as indispensable contributors to the production and

[15] For a detailed analysis of algorithmic management, refer to the report *Engineered precarities*, also in this series.
[16] Stakeholder Consultation 2

maintenance of these technologies. Acknowledging this dependence establishes the imperative to treat their labour as both operationally critical and socially consequential.

The question, as experts[17] articulated it, is ultimately "AI for whom?"—will these technologies continue to be developed through the exploitation and disposability of data workers, or can they be reimagined as part of economic systems that value and prioritise human dignity and collective well-being as part of an infrastructure of shared prosperity?

This recognition directly informs the need for action. The recommendations that follow outline concrete measures to address the vulnerabilities of data workers: fair contracts, social protections, mental health safeguards, limits on exposure to harmful content, grievance mechanisms, and the right to collective organisation. These measures aim to correct the imbalances that allow firms to offload liability while benefiting from essential human labour, translating the understanding of AI's human dependence into enforceable standards and practices. In doing so, they seek not only to alleviate the human costs of data work but also to augment the ethical and operational integrity of AI systems, ensuring that the technologies we develop are supported by labour practices that respect dignity, fairness, and accountability.

Regional Insights: Across the Africa and Southeast Asia stakeholder consultations, participants stressed that these labour harms are not isolated incidents, but structural characteristics of how global AI supply chains operate.

- **Africa:** Stakeholders underscored that the labour dynamics of data annotation and content moderation mirror earlier outsourcing models such as call-centre work, manufacturing, and other digital service exports where weak bargaining power, fragmented contracting, and limited regulatory oversight systematically depress labour standards.
- **Southeast Asia:** Participants similarly noted that the rapid expansion of AI service industries has outpaced labour regulation, with governments prioritising digital-economy growth and foreign investment, often at the expense of occupational health, worker protections, and transparency.
- **South Asia:** Insights gathered from experts further indicated that gaps in employment classification, limited access to social protection, and opaque subcontracting arrangements have normalised a system where precarity is built into the everyday functioning of data work. Across regions, the consistent message was that unsafe and unsustainable working conditions are not aberrations, they are the predictable outcomes of a global production model that relies on dispersed labour, minimal accountability, and the externalisation of risk onto workers.

[17] Expert Interview, Anonymised

3.5 | Cross-regional findings from the Consultations and Interviews

Across all three regional consultations, participants underlined the indispensable yet undervalued role of human labour in the functioning of global AI supply chains. Despite the geographic and institutional diversity, several recurring themes emerged: the structural invisibility of data work, the opacity of algorithmic management, and the absence of binding regulation ensuring fair working conditions.

A consistent thread across regions is the extractive organisation of digital labour, where workers in the Global South provide essential inputs for AI systems developed and commercialised elsewhere. In each case, participants described forms of outsourced digital labour that replicate the inequalities of older global production systems, such as garment manufacturing or call centre outsourcing.

However, the consultations also revealed distinct regional dynamics and policy gaps:

- **Visibility and Public Discourse:** In Africa, data work and content moderation have gained significant visibility, driven by media coverage and ongoing litigation involving major tech companies. This has created a stronger policy discourse around worker rights and mental health, even as enforcement remains weak. By contrast, in India and much of South Asia, there is comparatively limited coverage or policy debate surrounding the data work industry. The labour of data annotators and content moderators largely remains invisible in both research and regulatory frameworks, folded into broader conversations on platform work.

- **Regulatory Landscape and Institutional Capacity:** African participants highlighted that existing labour and data protection laws do not adequately capture the realities of data work. Southeast Asia faces similar gaps but is further complicated by governments' promotion of AI as a growth strategy, often sidelining labour concerns. In South Asia, the discussion focused on the legal exclusion of digital workers from formal employment classifications, which restricts access to social protection. Participants emphasised the fragmented governance structure, with overlapping ministries and inconsistent mandates between technology, labour, and commerce.

- **Worker Experience and Risk**: Mental health and algorithmic surveillance surfaced as universal concerns, though their manifestations differ. In Africa and Southeast Asia, trauma from exposure to harmful content and opaque algorithmic supervision were dominant themes. Participants in South Asia, however, pointed more to legal precarity, lack of recognition, and social insecurity, reflecting a different layer of

vulnerability. Mental health concerns, though present, are less systematically addressed due the absence of industry-level recognition.

In sum, while the forms of exploitation are structurally similar, the policy visibility, institutional maturity, and public recognition of data work vary significantly. Africa and Southeast Asia are at the forefront of public debate and advocacy around the labour of AI, whereas South Asia's digital labour economy remains largely hidden, treated as an informal extension of the gig sector rather than a distinct domain of high-value digital production.

This uneven awareness not only affects the political momentum for reform but also shapes the regional strategies needed to integrate data work into broader frameworks of decent work, due diligence, and AI governance.

4 | Emerging Pathways for Worker Solidarity

Across both expert interviews and consultations, participants consistently highlighted the structural barriers that make formal unionisation extremely difficult for data workers. As one expert[18] noted, "The workforce is dispersed across micro-contracts, short-term engagements, and multiple layers of subcontracting, leaving workers without a clear employer of record and therefore outside the scope of traditional labour law protections". A stakeholder[19] further added that "despite structural precarities like fragmented employment, NDAs, repeated repression of unionisation, data workers have managed to carve out new forms of resistance, solidarity and collective action." Workers are increasingly developing networked forms of organising that incorporate digital coordination with local advocacy.

Across the Global South, data workers are utilising digital spaces to connect and share resources and experiences. In the Philippines, for instance, a stakeholder[20] noted, "workers share experiences on social media platforms and other publicly available online channels, flag unethical requests, and coordinate solutions themselves, demonstrating real agency despite formal restrictions." Similarly, an expert[21] highlighted, "across regions, these online networks have become de facto unions, filling the gaps left by formal labour protections." This also allows them to develop peer support systems and increasingly connect across borders to challenge the conditions they face collectively.

As much as stakeholders highlight the need to look for a way forward, they also emphasise the importance of unions being formed for gig workers to amplify a collective voice. One expert[22] noted, "Employment laws do not yet address the realities of the new world of work. Our efforts require engagement with multiple stakeholders and careful legislative analysis. Going to court alone may establish a precedent, but the absence of comprehensive laws will create broader challenges." A stakeholder[23] added, "Harmonised contracts across workers in different organisations are crucial. Without union support, individual cases are difficult to pursue, and resistance through legal channels is limited. Collective approaches, with standardised contracts for unionised workers, provide a stronger foundation to address legal and structural challenges."

Such efforts for formal unionisation are beginning to surface. The Alphabet Workers Union has extended representation to contract-based data workers mandating safeguards such as guaranteed remote work, access to minimum severance entitlements of six weeks and an explicit ban on the use of invasive digital monitoring tools (Eidelson,

[18] Expert Interview |Anonymised
[19] Stakeholder Consultation 2
[20] Stakeholder Consultation 2
[21] Expert Interview | Anonymised
[22] Expert Interview |Anonymised
[23] Stakeholder Consultation 1

2024) while stakeholders have highlighted the African Content Moderators Union (ACMU), formed in Kenya in 2024 has become a powerful voice demanding fair pay, mental health protections, and legal accountability from outsourcing vendors following landmark litigation against Meta's outsourcing partners (Perrigo, 2023).

Alongside unionisation and advocacy efforts, workers themselves have developed *activist technologies* that intervene directly in exploitative platform systems. Early activist technologies like Turkopticon demonstrated how workers could publicise and evaluate employer behaviour, transforming invisibility into mutual aid (Irani & Silberman, 2013). Building on this legacy, newer initiatives such as Fairwork, and Worker Info Exchange extend these efforts into coordinated collective action, platform benchmarking, and data rights advocacy. WIE helps workers access the personal data platforms collect on them (via legal rights like data-access or data-portability requests), aggregate that data through "data trusts," and use it as leverage for collective bargaining, litigation, or organizing. Meanwhile, worker-led research movements like the Data Workers' Inquiry (DWI) and Data4Mods are generating first-hand evidence of conditions across Africa and Latin America. By combining mapping, legal rights exercises, worker surveys, and public advocacy, Data4Mods has exposed the structure of the outsourced moderation industry and provided tools that strengthen transparency and worker rights.

Locally grounded organisations like the Data Labellers Association (DLA) of Kenya, has emerged as a representative body for annotators working in the AI supply chain, focusing on fair pay, mental health support and employer accountability for workplace harms. A representative[24] underlined that the formation of data-worker collectives has emerged in response to the absence of formal grievance channels and organisational mechanisms for representation. In this context, DLA has begun developing model contracts and codes of conduct to articulate baseline expectations for fair and transparent practice. These collectives were formed by annotators and moderators in Africa and Latin America to document working conditions, mental health risks, and wage exploitation in AI supply chains. Their work directly informs campaigns by unions such as the ACMU and UNI Global Union. In Southeast Asia, coalitions such as the BPO Industry Employees Network (BIEN) in the Philippines have documented risks within moderation and annotation centres. As a stakeholder[25] involved in these efforts observed, "although the right to organise remains constrained, with workers risking being labelled political threats, the network is campaigning for sector-wide wage floors, occupational safety and health frameworks, and greater job security for BPO, moderation, and data work employees." Together, these initiatives represent a shift from reactive protest to infrastructural activism, where workers not only resist exploitation but also build parallel systems of transparency and accountability within the global AI economy.

[24] Expert Interview | Anonymised
[25] Stakeholder Consultation 2

In South Asia, collective organising among gig and platform workers remains in its nascent stages. In India, the Telangana Gig and Platform Workers Union (TGPWU) has emerged since 2021 as a significant actor advocating for fair wages, social security, and algorithmic transparency for app-based workers (Siasat, 2025). Through coordinated lobbying for legislation, the union has demanded minimum-fare guarantees per trip, formal social-security inclusion, and regulatory oversight of algorithmic work allocation, asserting that in the absence of minimum-wage guarantee and welfare provisions, platform work remains precarious and exploitative.

Intergovernmental organisations like the International Labour Organization (ILO) are supporting these efforts by creating global policy trackers, convening multi-stakeholder dialogues, and developing standards for measuring and regulating digital labour.

At the regional level, the European Union's Directive 2024/2831 on Improving Working Conditions in Platform Work establishes minimum rights for platform workers, clarifies employment status, and mandates transparency and oversight of algorithmic management systems in platform work (European Parliament & Council, 2024). This framework creates enforceable standards that can cascade through global supply chains, potentially influencing AI data labour practices beyond Europe.

Together, these actions represent a growing movement to make visible and improve the conditions of AI's hidden workforce. While fragmented, these efforts demonstrate that data workers are not merely passive participants in the global AI economy but increasingly active agents reimagining its moral and political terrain. Building on these emerging efforts, the next section highlights potential directions for reform—identifying structural, legal, and institutional measures needed to transform advocacy into enforceable protections. The following section considers how these gains can be consolidated through policy reform, stronger accountability frameworks, and a rights-based approach to the governance of digital labour.

5 | Ways Forward: Improving Conditions for Data Workers

The expansion of digital supply chains has created a vast workforce of data annotators and content moderators whose labour remains precarious, yet essential to the functioning of AI systems. To build and develop fairer, more humane systems of digital work, a set of core protections and guidelines must be put in place. This section synthesises findings from the project's secondary research, stakeholder consultations, and expert interviews to inform a set of actionable recommendations.

Implementation of protections for data workers may vary across regions due to certain contextual factors. Organisational capacity for instance varies, with the strength and independence of unions, regulatory bodies, and civil society organisations differing across the region; in many contexts, worker-led initiatives and international support are essential to ensure compliance. Insights from expert interviews and stakeholder consultations indicate that unionisation and collective organisation are particularly constrained in South Asia, where digital labour remains largely invisible and poorly documented. By contrast, in parts of Sub-Saharan Africa and Southeast Asia, there is slightly greater visibility into digital workforces, enabling nascent unionisation efforts and worker-led collectives, though structural and legal barriers still persist. While the core principles of fair pay, occupational safety, mental health support, grievance mechanisms, and transparency remain consistent, implementation differs across contexts in the Global South, depending on local laws and organisational capacity.

The following recommendations are organised across three levels—supplier countries in the Global South, sourcing countries in the Global North, and international development cooperation to ensure targeted actions across the global AI data-labour ecosystem.

5. 1. | Recommendations for Supplier Countries (Global South):

1. **Mandate Fair Employment and Social Protections**: All data workers should have access to fair contracts, fair wages, paid training, and comprehensive social protections, including health insurance, sick leave, and severance pay.

Stakeholders responsible:

- **Governments:** Define fair work standards for data tasks, drawing on frameworks such as Fairwork while adapting them to informal and subcontracted environments.

- **Lead firms and vendors:** Align wages with local living-wage benchmarks; provide paid training; ensure written, transparent contracts; provide notice before termination. Insert *contractual clauses* into vendor agreements requiring: written contracts for all hired workers, minimum living-wage alignment, breakdown of pay components (base + bonus + hazard pay), and a clause requiring 30-day notice before termination of long-term assignments.
- **Donors:** Support portable-benefits pilots for informal digital workers.

2. Ensure Occupational Safety, Health, and Hazard Pay: Data work, particularly content moderation (see European Agency for Safety and Health at Work, 2021), often involves repeated exposure to violent, abusive, or disturbing material which poses serious risks to mental health.

- **Employers:** Classify tasks by *exposure risk level* (Low / Medium / High). For each level define allowed daily or weekly exposure caps, required break schedules, and mandatory rotation out of high-risk queues. Productivity metrics and KPI systems must be revised with human oversight.
    - **Mental health services**: Contract with independent trauma specialists (not in-house HR) to provide: baseline psychological screening on hire (consent-based), periodic check-ins, and on-call counselling for acute exposure.
    - **Hazard pay**: Set explicit multipliers for high-exposure tasks. Ensure hazard pay is non-discretionary and visible on payslips.
- **Governments:** Integrate digital labour into occupational safety frameworks and regulate exposure limits.
- **Donors and NGOs:** Fund psychosocial support networks in regions where clinical infrastructure is limited.

3. **Ending Silencing and Retaliation**: Remove contractual and practical barriers to disclosure and remedy. Workers must receive standardised, transparent contractual templates that guarantee workers' rights to disclosure and collective discussion of workplace harms.

- **Governments:** Prohibit NDAs and contract clauses that prevent disclosure of labour abuses or that penalise whistleblowing about working conditions; require standardised contract templates.
- **Unions and worker groups:** Unions play a key role here in monitoring workplace harms, documenting violations, and supporting worker-led reporting systems.

- **Lead firms:** Guarantee non-retaliation across all tiers of their supply chains. To enforce accountability, independent grievance mechanisms—co-designed with worker representatives and managed by impartial third parties—should be established at both domestic and transnational levels.

4. **Collective Rights**: Equally critical is the right to association and collective bargaining. Data workers, as an imperative, must be accorded the privileges of unionisation, organisation, and representation in negotiations over working conditions. Worker councils or union representatives should be formally embedded into BPO–client contracts, ensuring labour has a voice in governance. Labour organisations like the ILO and unions such as UNI Global have noted that collective bargaining and representation are among the few effective mechanisms for non-standard workers to secure safe work and fair conditions in global supply chains.

- **Governments:** Recognise data workers under national labour laws (or amend laws to include platform/digital workers). Guarantee the right to organise for data workers.
- **Lead firms and vendors:** Include elected worker councils or union representation into BPO-client contracts.
- **Unions and NGOs:** Provide legal and organisational support, especially in contexts (countries or vendor sites) where labour law is weak or enforcement is inconsistent.

5.2. | Recommendations for Sourcing Countries (Global North):

1. **Transparency and Lead Firm Accountability**: There should be visibility into the scale, conditions, and risk profiles of data-labour operations, and clear responsibility for labour standards across all tiers of AI supply chains must be established.

- **Lead firms:** Firms and their outsourcing partners must be fully transparent about the scale, conditions, and geography of their data labour supply chains (Kloiber, 2025). This includes disclosing the number of workers employed, the nature of the tasks they perform, and the risks they are exposed to.
- **Regulators:** Mandate supply chain disclosures for high-risk digital labour. Lead firms must also be held directly accountable for the conditions across their supply chains. Responsibility cannot be outsourced to middleman vendors. Regulatory frameworks should close loopholes that allow firms to evade responsibility through complex contracting arrangements.

2. **International Standards and Enforcement**: Interviews with regulators and policy experts highlighted the absence of binding international frameworks covering this labour.

- **ILO and multilateral bodies:** Recognise data annotation and moderation as high-risk work; develop global wage, safety, and exposure standards. Organisations such as the International Labour Organisation (ILO) should formally recognise data annotation and content moderation as high-risk essential labour, and push for a technical committee to draft standards on exposure limits, hazard pay, and psychosocial protections.
- **Governments:** Align national frameworks with emerging international norms and prevent regulatory arbitrage.
- **NGOs and global labour alliances:** Support monitoring and facilitate cross-border dialogue and push for inclusion of digital labour in global decent work agendas.
- **Civil society organisations**: Build transparency through worker-led monitoring, documentation, and advocacy.

5.3. Recommendations for international development cooperation:

AI data work and content moderation constitute critical nodes in the global digital economy, yet the labour sustaining these systems are structurally marginalised, underpaid and exposed to high-risk working conditions. The human costs caused by these conditions are substantial but largely invisible in existing policy and regulatory frameworks. An effective approach should combine enforcement, research, capacity-building, and cross-border collaboration to create a sustainable and durable impact.

- **Implement Tools and Partnerships for Fair and Safe Digital Labour:** To enhance the comprehensiveness and impact of interventions by global development cooperation, donor organizations can operationalise an integrated approach that combines evidence, technology, and collaboration. This includes funding longitudinal studies to track occupational health, career development, and risk exposure.
- **Fund Worker-Led Evidence Infrastructure and Independent Monitoring:** Enforcement depends on verifiable documentation of harms, yet data workers operate in contexts where violations remain systematically hidden. Finance worker-led monitoring initiatives, legal aid networks, and evidence-gathering mechanisms. Prioritise building institutional capacity within worker organisations to conduct independent audits and leverage data protection frameworks to expose exploitative practices, generating critical inputs for litigation and regulatory action.
- **Pilot Fair Labour Procurement Standards and Regional Coordination:** Leverage public procurement by embedding non-negotiable labour clauses in government AI service contracts including living wage requirements, subcontractor disclosure,

independent audit rights, and prohibition of exploitative arrangements. Simultaneously, facilitate regional coordination among Global South countries to establish harmonised minimum standards and mutual recognition agreements, reducing incentives for firms to exploit regulatory differentials through work relocation. Document and disseminate successful models as replicable templates for wider adoption.

- **Facilitate Global Policy Dialogues**: Convene inclusive, worker-centred dialogues that foreground the experiences of data workers. These forums can produce actionable policy frameworks integrating fair compensation, mental health protections, transparent task allocation, and enforceable standards for algorithmic oversight.

6 | Conclusion

The progression of Artificial Intelligence (AI) rests on the exploited labour of unseen workers. Addressing and correcting this inequity demands a fundamental rethinking of AI's political economy—recognising data work as legitimate employment, securing social protections, and enforcing accountability across borders. Without structural change, AI will deepen global inequality; with it, it has the potential to support a more equitable digital economy founded on dignity, safety, and recognition for all workers.

While this report documents the realities of data work, critical dimensions remain underexamined. Future research should rigorously evaluate what constitutes adequate psychological support for workers exposed to traumatic content, including which intervention models prove most effective in preventing long-term harm. The gendered dimensions of data work, including how care responsibilities, exposure to sexual violence in content, and workplace harassment intersect to shape women's experiences require a deeper inquiry, as do the intersections of caste and race with data work vulnerabilities. Comparative analysis across Global South regions should examine how different political economies, labour market structures, and infrastructural conditions produce both common patterns of exploitation and context-specific vulnerabilities requiring tailored policy responses. Finally, detailed research on emerging forms of collective action, transnational worker networks, and innovative organizing strategies would provide crucial knowledge for labour movements seeking to build power in this sector. Most importantly, future research must center workers' own perspectives, experiences, and demands, recognizing them as knowledgeable agents whose insights are at the centre for transforming the conditions under which AI is produced.

7 | References

- Abuel, D. D. (2025). Better in moderation: Occupational risks and hazards to the mental health of Filipino content moderators. Philippine Law Journal, *98*(1), 148-183. https://philippinelawjournal.org/wp-content/uploads/2025/01/98PLJ148_Abuel.pdf

- Anwar, M, and Graham, M. (2022), The Digital Continent: Placing Africa in Planetary Networks of Work (Oxford, 2022; online edn, Oxford Academic) https://doi.org/10.1093/oso/9780198840800.001.0001

- Arya, N. (2025). India's Data Workers: The Human Labour Making Machines Learn. The Wire. https://m.thewire.in/article/tech/indias-data-workers-the-human-labour-making-machines-learn/amp?utm=relatedarticles

- Bansal, V. (2025). Hundreds of Google AI Workers Were Fired Amid Fight Over Working Conditions. Wired. https://www.wired.com/story/hundreds-of-google-ai-workers-were-fired-amid-fight-over-working-conditions/

- Berg, J., Furrer, M., Harmon, E., Rani, U., & Silberman, M. (2018). Digital labour platforms and the future of work: Towards decent work in the online world. International Labour Organization. https://www.ilo.org/publications/digital-labour-platforms-and-future-work-towards-decent-work-online-world

- Bernhardt, A., Kresge, L., & Suleiman, R. (2022). The Data-Driven Workplace and the Case for Worker Technology Rights. ILR Review, 76(1), 3-29. https://doi.org/10.1177/00197939221131558 (Original work published 2023)

- Business and Human Rights Resource Centre. (2024). Kenya: Content moderators filed a lawsuit against Meta alleging unfair termination of subcontractor's employees and discriminatory hiring conditions. https://www.business-humanrights.org/en/latest-news/kenya-content-moderators-filed-a-lawsuit-against-meta-alleging-poor-working-conditions-including-insufficient-mental-health-support-and-low-pay/

- BHRRC. (2025). Kenya: Tech workers challenge bill that prioritises foreign investment over worker rights. https://www.business-humanrights.org/en/latest-news/kenya-tech-workers-challenge-bill-that-prioritizes-foreign-investment-over-worker-rights/

- Bhattacharjee, S and Shivakumar, N. (2025). The Hidden Human Cost of AI Moderation. Jacobin. https://jacobin.com/2025/06/ai-moderation-ndas-trauma-labor

- Booth, R and Kimeu, C. (2024). PTSD, depression and anxiety: why former Facebook moderators in Kenya are taking legal action. The Guardian. https://www.theguardian.com/world/2024/dec/18/why-former-facebook-moderators-in-kenya-are-taking-legal-action

- Brandom, R. (2025). "The billion-dollar infrastructure deals powering the AI boom". Tech Crunch. https://techcrunch.com/2025/10/10/the-billion-dollar-infrastructure-deals-powering-the-ai-boom/

- Braesemann, F, et al. (2022). The global polarisation of work. PLOS.one https://doi.org/10.1371/journal.pone.0274630

- Casilli, A. A., Tubaro, P., Cornet, M., Le Ludec, C., Torres-Cierpe, J., & Viana Braz, M. (2024). *Global inequalities in the production of artificial intelligence: A four-country study on data work* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2410.14230

- Carmona, M. (2025). A closer look at the health of Platform Workers. Equal Times. https://www.equaltimes.org/a-closer-look-at-the-health-of

- Center for Labour and a Just Economy (CLJE Lab). (2024). Worker Power and Voice in the AI Response. Harvard Law School. https://clje.law.harvard.edu/app/uploads/2024/01/Worker-Power-and-the-Voice-in-the-AI-Response-Report.pdf

- Coldeway, D. (2024). Data workers detail exploitation by the tech industry in the DAIR report. Tech Crunch. https://techcrunch.com/2024/07/08/data-workers-detail-exploitation-by-tech-industry-in-dair-report/

- Collins, J. L. (2003). *Threads: Gender, labor, and power in the global apparel industry*. University of Chicago Press. https://press.uchicago.edu/ucp/books/book/chicago/T/bo3614267.html

- Couldry, N., & Mejias, U. A. (2018). Data Colonialism: Rethinking Big Data's Relation to the Contemporary Subject. Television & New Media, 20(4), 336-349. https://doi.org/10.1177/1527476418796632 (Original work published 2019)

- Datta, N., & Chen, R. (2023). Working without borders: The promise and peril of online gig work. World Bank [Report]. https://openknowledge.worldbank.org/entities/publication/ebc4a7e2-85c6-467b-8713-e2d77e954c6c

- Dewan, S., Taneja, D., & Krishnamurthy, M. (n.d.). Digitalisation and the Indian Labour Market - Trends, Challenges, and Opportunities. GIZ.

https://justjobsnetwork.org/files/digitalisation-and-the-indian-labour-market-trends-challenges-and-opportunities_oct_2022.pdf

- Dinika, A.-A. T. (2025). Navigating labour's labyrinth: Developing a typology of platform work in Sub-Saharan Africa. Platforms & Society, 2. https://doi.org/10.1177/29768624251318084
- Dinika, A. (2024). The Human Cost of Our AI-driven Future. Noema Magazine. https://www.noemamag.com/the-human-cost-of-our-ai-driven-future/
- Dr. Natarjan, S & Mohamed, S et al. (2021). Just and Equitable Data Labelling. Towards a Responsible AI Supply Chain. Aapti Institute. https://uploads.strikinglycdn.com/files/7d492f74-a51f-423b-bf5d-65c9f88eee06/AI_Data_Labelling_Report_DIGITAL_25FEB1033.pdf
- Du, Michelle, and Okolo, C. (2025). Reimagining the future of data and AI labour in the Global South. Brookings. https://www.brookings.edu/articles/reimagining-the-future-of-data-and-ai-labor-in-the-global-south/
- Eidelson, J. (2024). Google contract staff reach a union deal banning keystroke monitoring. Business Standard. https://www.business-standard.com/world-news/google-contract-staff-reach-union-deal-banning-keystroke-monitoring-124121900118_1.html?utm
- Ekbia, H and Nardi, B. (2017). Heteromation and Other Stories of Capitalism. MIT Press Direct. https://doi.org/10.7551/mitpress/10767.001.0001
- Equidem. (2025, May 28). Scroll. Click. Suffer: The hidden human cost of content moderation and data labelling. Business & Human Rights Resource Centre. https://equidem.org/wp-content/uploads/2025/05/Equidem-Data-Workers-Report-May-29.pdf
- Espíndola, J. (2025), Against Human Content Moderation: Algorithms without Trauma. J Appl Philos, 42: 1285-1300. https://doi.org/10.1111/japp.70024
- Fairwork. (2025). Fairwork AI ratings 2024/5 Who powers AI? Assessing fairness in the AI supply chain Humans in the Loop as a case study. Global Partnership on AI. Oxford: United Kingdom. https://fair.work/wp-content/uploads/sites/17/2025/07/GPAI_Fairwork_Ratings_Report-FINAL.pdf
- Gebrekidan, F. B. 2024, 'Content moderation: The harrowing, traumatizing job that left many African data workers with mental health issues and drug dependency'. https://data-workers.org/wp-content/uploads/2024/06/Fasicas-Report-6th-review-After-Lawyer.pdf

- Gonzalez, A and Matias, N. (2025). Measuring the Mental Health of Content Reviewers, a Systematic Review. https://doi.org/10.48550/arXiv.2502.00244
- Graham, M., Hjorth, I., & Lehdonvirta, V. (2017). Digital labour and development: impacts of global digital labour platforms and the gig economy on worker livelihoods. Transfer: European Review of Labour and Research, 23(2), 135-162. https://doi.org/10.1177/1024258916687250 (Original work published 2017)
- Gray, M. L., & Suri, S. (2019). Ghost work: How to stop Silicon Valley from building a new global underclass. Houghton Mifflin Harcourt. http://sarkoups.free.fr/gray2019.pdf
- Hall, R and Wilmot, C. (2025). Meta faces Ghana lawsuits over impact of extreme content on moderators. The Guardian. https://www.theguardian.com/technology/2025/apr/27/meta-faces-ghana-lawsuits-over-impact-of-extreme-content-on-moderators
- Hancock, J. (2025). Tech Oligarchy Imperils Democratic Information Flows. Tech Policy Press. https://www.techpolicy.press/tech-oligarchy-imperils-democratic-information-flows/
- Howson, K et al. (2022). Unpaid labour and territorial extraction in digital value networks. Wiley Online Library. https://onlinelibrary.wiley.com/doi/10.1111/glob.12407
- Ingram, D. (2023). ChatGPT is powered by these contractors making $15 an hour. NBC News. https://www.nbcnews.com/tech/innovation/openai-chatgpt-ai-jobs-contractors-talk-shadow-workforce-powers-rcna81892
- Irani, L. & Silberman, M. (2013). Turkopticon: Interrupting Worker Invisibility in Amazon Mechanical Turk. https://escholarship.org/content/qt10c125z3/qt10c125z3_noSplash_17ace960c310470e904a7512eccf6ee2.pdf
- Jones, P. (2021). Work without the worker: Labour in the age of platform capitalism. Verso Books. https://onwork.edu.au/bibitem/2021-Jones,Phil-Work+Without+the+Worker+Labour+in+the+Age+of+Platform+Capitalism/
- Kloiber, J. (2025). Data workers demand safe and fair conditions. Superrr Labs. https://superrr.net/en/blog/data-workers-demand-safe-and-fair-conditions
- Lenis, H. (2023). Venezuelan migrants in delivery platform work during the COVID-19 pandemic in Buenos Aires, Argentina: Between exploitability, precariousness, and daily resistance. https://doi.org/10.1177/02633957231186908

- Lindeman, M. (2024). Data Scarcity: When Will AI Hit a Wall? Pieces. https://pieces.app/blog/data-scarcity-when-will-ai-hit-a-wall
- Liu, Zhang et al (2021). Scalable data annotation pipeline for High-Quality Large speech datasets development. https://doi.org/10.48550/arXiv.2109.01164
- Lung, N. (2025). Uber Launches Data Tasks as Option for Some US Drivers to Earn Money. Bloomberg. https://www.bloomberg.com/news/articles/2025-10-16/uber-giving-some-us-drivers-option-to-earn-money-from-tasks-like-uploading-menus?
- Miceli, Milagros & Posada, Julian. (2022). The Data-Production Dispositif. 10.48550/arXiv.2205.11963.
- Mklama, B & Ragnet, M. (2025). How Digital Colonialism Threatens Kenya's Silicon Savannah. Tech Policy Press. https://www.techpolicy.press/how-digital-colonialism-threatens-kenyas-silicon-savannah/
- Muldoon, J., Cant, C., Wu, B., & Graham, M. (2024). A typology of artificial intelligence data work. Big Data & Society, 11(1). https://doi.org/10.1177/20539517241232632 (Original work published 2024)
- Muldoon, J., Cant, C., Graham, M. et al. The poverty of ethical AI: impact sourcing and AI supply chains. AI & Soc 40, 529–543 (2025). https://doi.org/10.1007/s00146-023-01824-9
- Muldoon, J., Wu, B.A. Artificial Intelligence in the Colonial Matrix of Power. *Philos. Technol.* 36, 80 (2023). https://doi.org/10.1007/s13347-023-00687-8
- Munn, L. (2022). Automation is a Myth. Stanford University Press. https://www.lukemunn.com/2022/automation-is-a-myth/
- Naseer, T. (2024). Youth Made Up 82.9% of India's Unemployed Population in 2022: ILO. The Wire. https://thewire.in/economy/ilo-report-work-unemployment-india
- Nieva, R. (2025). Uber Is Making A Push In Data Labeling After Scale AI's Deal With Meta. Forbes. https://www.forbes.com/sites/richardnieva/2025/06/20/uber-scale-data-labeling/
- Perrigo, B. (2023). OpenAI Used Kenyan Workers on Less Than $2 Per Hour to Make ChatGPT Less Toxic. Time Magazine. https://time.com/6247678/openai-chatgpt-kenya-workers/

- Perrigo, B. (2023). 150 African Workers for ChatGPT, TikTok and Facebook Vote to Unionize at Landmark Nairobi Meeting. Time Magazine. https://time.com/6275995/chatgpt-facebook-african-workers-union/
- Pilatti, G. R., Pinheiro, F. L., & Montini, A. A. (2024). Systematic Literature Review on Gig Economy: Power Dynamics, Worker Autonomy, and the Role of Social Networks. *Administrative Sciences*, *14*(10), 267. https://doi.org/10.3390/admsci14100267
- Posada, J. (2022). The coloniality of data work: Power and Inequality in Outsourced Data production for Machine Learning. University of Toronto. https://utoronto.scholaris.ca/bitstreams/41f9a4be-7f13-48f6-9a65-28ee849aa787/download
- Reuters. (2022). Content moderator in Kenya sues Meta over working conditions. https://news.trust.org/item/20220511140054-bfsyi
- Riccio, F., Cresti, L. & Virgillito, M.E. The labour share along global value chains: perspectives and evidence from sectoral interdependence. *Rev World Econ* 161, 635–684 (2025). https://doi.org/10.1007/s10290-024-00555-3
- Shanley, D. (2025). Synthetic Data, Real Harm. Ada Lovelace Institute. https://www.adalovelaceinstitute.org/blog/synthetic-data-real-harm/
- Shivakumar, N and Bhattacharjee, S. (2025). How NDAs Became the AI Industry's Tool for Surveillance and Silence. Tech Policy Press. https://www.techpolicy.press/how-ndas-became-the-ai-industrys-tool-for-surveillance-and-silence/
- Siele, M. (2023). "It's been tough for us": Meta's Kenyan content moderators say they'll keep fighting. Rest of World. https://restofworld.org/2023/meta-content-moderators-kenya-fired-unionize/
- Swain, G. (2025). Uber turns drivers into AI data labelers in India pilot. Computer World. https://www.computerworld.com/article/4052881/uber-turns-drivers-into-ai-data-labelers-in-india-pilot.html
- Tan, R and Cabato, R. (2023). Behind the AI boom, an army of overseas workers in 'digital sweatshops'. The Washington Post. https://www.washingtonpost.com/world/2023/08/28/scale-ai-remotasks-philippines-artificial-intelligence/
- Tasnim, A. (2025). Youth unemployment is a growing concern for South Asia. The daily star. https://www.thedailystar.net/opinion/views/news/youth-unemployment-growing-concern-south-asia-3941866

- Toxtli, C et al. (2021). Quantifying the Invisible Labor in Crowd Work. Cornell University. https://doi.org/10.48550/arXiv.2110.00169
- Tubaro, P, Casilli, A, et al. (2025) The digital labour of artificial intelligence in Latin America: a comparison of Argentina, Brazil, and Venezuela. Globalizations, pp.1-16. 10.1080/14747731.2025.246517 .
- Tubaro, P., Casilli, A. A., & Coville, M. (2020). The trainer, the verifier, the imitator: Three ways in which human platform workers support artificial intelligence. Big Data & Society, 7(1). https://doi.org/10.1177/2053951720919776
- Tubaro, P., Casilli, A. A., Cornet, M., Le Ludec, C., & Torres Cierpe, J. (2025). Where does AI come from? A global case study across Europe, Africa, and Latin America. *New Political Economy*, *30*(3), 359–372. https://doi.org/10.1080/13563467.2025.2462137
- UNESCO. (2024) UNESCO Welcomes G20 Digital Economy Working Group's Focus on Artificial Intelligence for Inclusive Sustainable Development and Inequalities Reduction. https://www.unesco.org/en/articles/unesco-welcomes-g20-digital-economy-working-groups-focus-artificial-intelligence-inclusive
- United Nations. Annex 1: Global Digital Impact. Pact for the Future. https://www.un.org/pact-for-the-future/en/annex-1-global-digital-compact
- Verma, Disha (2025). Labour in the Shadows: Rights and Risks for Asia's Data Workers. Tech Global Institute. https://techglobalinstitute.com/wp-content/uploads/2025/08/Labor-in-the-Shadows-TGI.pdf
- Villalobos, P et al. (2024) 'Will we run out of data? Limits of LLM scaling based on human-generated data'. ArXiv [cs.LG], 2024. arXiv. https://arxiv.org/abs/2211.04325
- Williams, A, and Milagros, M.(2023). "Data Work and its Layers of (In)visibility." Just Tech. Social Science Research Council. DOI: doi.org/10.35650/JT.3060.d.2023
- Wilmot, C. (2025). Suicide Attempts, Sacking and a Vow of Silence: Meta's new moderators face worst conditions yet. TBIJ. https://www.thebureauinvestigates.com/stories/2025-04-27/suicide-attempts-sackings-and-a-vow-of-silence-metas-new-moderators-face-worst-conditions-yet

Appendix A | List of Experts[26]

1. **Alexandra Mateescu** | Data & Society
2. **Angela Chukunzira** | Siasaplace
3. **Benjamin Shestakofsky** | Cornell University
4. **Dunstan Allison-Hope** | Independent
5. **Ephantus Kanyugi** | Data Labelers Association
6. **Isabel Ebert** | UNHRC B-Tech
7. **Dr Jun-E Tan** | Khazanah Research Institute
8. **Joan Kinyua** | Data Labelers Association
9. **Julian Posada** | Yale University
10. **Kauna Malgwi** | African Content Moderators Unions
11. **Leslie Dwolatzky** | Research ICT Africa
12. **Martjin Arets** | GigCV and WageIndicator
13. **Milagros Miceli** | Distributed AI Research Institute (DAIR)
14. **Mohammed Amir Anwar** | University of Edinburgh and Planetary AI
15. **Mophat Okinyi** | Techworker Community Africa
16. **Rafael Grohmann** | University of Toronto
17. **Rim Melake** | SUPERRR Lab
18. **Sadhna Sanjay** | IT for Change
19. **Tim Newman**| TechEquity
20. **Wanjiru Mburu** | Qhala
21. **Wasel Bin Shadat** | University of Dhaka

[26] Although the research team consulted 38 experts, only those who explicitly agreed to have their names and affiliations published in this report are included in this list.

Appendix B | Research Methodology

This document is part of a project on AI labour in the Global South. The project's research efforts and stakeholder engagement informed three reports. The first report discusses the tech sectors' need for data and the problems arising from their engagement of human labour. Algorithmic management systems that manage data workers and content moderators are discussed in the second report. Finally, the third report discusses the transnational nature of the data work and content moderation sectors and explores existing measures for pursuing accountability and fairness in such outsourcing.

The reports were developed through a combination of 38 expert interviews, three multistakeholder discussions, and secondary research. Secondary research involved a review of scholarly literature, media investigations and coverage, reports and discussions from civil society, and policy discussions of existing and proposed measures. These sources helped understand aspects like precarious working conditions, prevalent business practices, potentially relevant regulatory developments, and worker-led initiatives.

Through secondary research and our understanding of the digital labour ecosystem, we identified several people who had expertise in topics relevant to our work. The practitioners we reached out to had insights into topics like platform work, digital labour organising, content moderation outsourcing, human rights due diligence (HRDD), and the future of work. We developed questions based on our respondents' focus areas and on the reports' themes, culminating in interviews lasting between 45 and 60 minutes.

We convened three stakeholder consultations. Each virtual event in this series focused on one of three regions: Africa, Southeast Asia, and South Asia. For each event, we invited participants by identifying them in our secondary research efforts, experts' recommendations, and through our networks. Thus, our secondary research efforts were complemented by our engagement with various practitioners and stakeholders related to data work and content moderation. The reports blend the information gathered from the stakeholder consultations and expert interviews with the ongoing research and investigations of various institutions and professionals across the world. They are intended to act as companion resources and starting points for future initiatives and interventions on data work and content moderation sectors active in the Global South.

Additional Resources:

1. Readback 1: Stakeholder Consultation on Africa
2. Readback 2: Stakeholder Consultation on Southeast Asia
3. Readback 3: Stakeholder Consultation on South Asia