AI & EQUALITY

Mainstreaming Gender Perspective in AI crowd work in the Global South: *Diagnostic, policy recommendations and smart tools for women's empowerment*

Cristina Martínez Pinto, Luz Elena González Zepeda, Tatiana Telles, Norma Elva Chávez, Maya De Los Santos, Alberto Navarrete, and Saiph Savage

December 2023

Abstract

Workers from Kenya and the Philippines were among the first to be externally employed to label data, but after 2018, 75% of the leading AI crowd work platforms' workforce was Venezuela and it is possible that other Latin American countries will follow in the context of the post covid-19 economic recovery. While crowd work presents opportunities for income and employment creation in regions where local economies are stagnant- there are not enough initiatives that address the impact of such work in the Global South through the lens of gender perspective, considering that 1 in every 5 crowd workers in the region are women. To address this knowledge gap, we conducted an experimental survey on 60 women from Latin America who use the crowd work platform Toloka to understand their personal goals, professional values, and hardships faced in their work. Key insights revealed that a majority of the women shared a desire to hear the experiences of other crowd working women; particularly to help them navigate tasks, develop technical and soft skills, and manage their finances with more ease. Additionally, 75% of the women reported completing crowd work tasks, on top of caring for their families; while over half of them confirmed they needed to negotiate their family responsibilities, in order to pursue crowd work in the first place. These findings confirmed an important component lacking from the experiences of these women was a sense of connection with one another. Based on these observations, we propose a system designed to foster community between Latin American women in crowd work to improve their personal and professional advancement.

Introduction

The production of labeled data is crucial to meet the expectations of Artificial Intelligence (AI) system development and training. The great technological promise of this century depends largely on algorithms that need millions of labeled training data to learn, recognize, and categorize information. Data labeling is the product of AI models and the manual work of people who monitor, correct and augment the predictions of the former, thus improving their accuracy. This form of labor is known as crowd work.

As crowd workers continue to make crucial contributions to the training and development of advancing AI systems, interest in ensuring they produce quality work has grown. However, previous approaches have neglected to center crowd workers in system designs and to consider their identities, motivations, and well-being. These considerations are important, because they acknowledge crowd workers are not a monolith and may have different needs and goals. Designing a system sensitive to the unique needs of a specific worker population can increase the likelihood of the tool's usefulness and, by extension, contribute to crowd workers well-being and help them produce better quality work.

In recent years, Latin American and Caribbean workers have been identified as significant crowd work contributors. The crisis conditions in these areas have left many dependent on crowd work as a steady source of income. Latin American women make up a sizable portion of these workers, using crowd work as a path to financial independence while balancing traditional caregiver responsibilities expected of them due to their strong patriarchal societies. Unfortunately, research surrounding the perspectives of Latin American crowd workers is largely incomplete as most investigations have been centered on the experiences of Western women, who comparably have greater freedom and support to prioritize their individual needs.

To address this knowledge gap, we conducted an experimental survey on 60 women from Latin America who use the crowd work platform Toloka to understand their personal goals, professional values, and hardships faced in their work. Key insights revealed a majority of the women shared a desire to hear the experiences of other crowd working women; particularly to help them navigate tasks, develop technical and soft skills, and manage their finances with

more ease. Additionally, 75% of the women reported completing crowd work tasks, on top of caring for their families; while over half of them confirmed they needed to negotiate their family responsibilities, in order to pursue crowd work in the first place. These findings confirmed an important component lacking from the experiences of these women was a sense of connection with one another.

Based on these observations, we propose a system designed to foster community between Latin American women in crowd work to improve their personal and professional advancement. By providing them with a safe platform to engage in meaningful conversation, they can begin to build an extensive foundation of knowledge for completing their work while growing their career and personal skills in the process. Moreover, since many Latin American women share the pressure of balancing their family responsibilities and reputation alongside their work, they are in the best position to relate and give advice to one another.

In order to facilitate communication between the women, and help them identify conversations related to their current interests, we would integrate an intelligent chatbot into the system. Designed to emulate the personality of Latin American heroines, the chatbot will assist users in searching for specific advice, help them textualize their interests, and be a guide for navigating the platform. Furthermore, the chatbot will be capable of recommending to users other crowd working women who may be good connections based on their interests, expertise, and experiences.

Designing a space for crowd workers to access each other's knowledge can lead to improved quality in their work by empowering them to seek support addressing problems in their working conditions. According to field practitioners and researchers, a better future in crowd work can only be fulfilled if key aspects such as transparency, fair pay, and opportunities for professional advancement are embodied in tools available to crowd workers. Having their experiences documented will not only benefit crowd workers long-term, but also provide them with the ability to support one another in identifying areas of improvement in their work environments. This can inform decisions to organize and advocate for respect, fair pay, and transparency from requesters and crowd working platforms. Despite being promoted as an opportunity to create income and employment in regions where local economies are stagnant, there are not enough initiatives that address the impact of such work in the Global

South through the lens of gender perspective, considering that 1 in every 5 crowd workers are women.

In addition to the experimental surveys conducted, we analyzed the current state of AI-related crowd work in the Latin America and the Caribbean region, studying the differentiated impact of such work opportunities for women and men, and developed a set of policy recommendations based on international best practices and use cases.

Feminist Research

Research can be considered feminist when it is grounded in a set of theoretical traditions that privilege women's issues and experiences. A key particularity of feminist research lies in its focus on power imbalance, both in the subject that is studied and the relationships between the subject and the research. Feminist research practice includes the practice of reflexivity or positionality, as a tool for producing knowledge but also an ethical method focused both in subverting unequal relationships of power in knowledge production, to consider our own position within the research as one of power, and to always take into account the vulnerability of the researched. An ethical commitment cannot happen without the step of reflexivity. Reflexivity is the deconstruction of knowledge production, by addressing academia as a place for knowledge production, where social interactions and thus social constructions happen.

Gender mainstreaming can be used as a strategy for achieving equality by integrating gender analysis in all the phases of policies, actions, plans, and budgets. Gender as a variable of analysis looks at social attributes and opportunities associated with being female and male; to the relationships between women and men, and to the relations between women and those between men. Moreover, as gender is generally understood as a social construct affected by power structures, different analytical concepts have been used to perfect the analysis of those power structures and its relations to gender, the most widely used being intersectionality. Intersectionality is used to map out the intersecting power relations in social relations by viewing categories such as race, class, gender, sexuality, nation, ability, ethnicity, among others, interrelated and mutually shaping one another.

As such, by considering the gender variable as central to our research, we go beyond sex disaggregation, and look into what the data conveys about attitudes, norms, and gaps within AI crowd work. Within the project, two survey pilots were conducted in such a way to prevent exposure of the women identified as subjects of research (crowd workers), to abide by ethical standards and fair pay, and to enable communication channels for them to open up about their lived-in experiences beyond what previous literature review on women crowd workers has revealed.

Methodology

Our team conducted two survey pilots in Spanish on the Toloka platform. Toloka is a global crowd sourcing company, founded in 2014 by Olga Megorskaya and integrated within the Yandex search engine, as an enabling environment to support data-related processes. The platform was chosen as it allowed us to filter workers geographically, to specifically analyze Latin American workers. Additionally, we contacted Toloka's Educational Program team, who expressed interest in supporting research activity through the platform.

The first version of the survey was first answered by 6 people (three men and three women); the survey pilot was divided into specific sections, and most of the questions were of Likert scale type and open-ended questions. The time to submit the survey was established to 20 minutes for women and men crowd workers that were interested in doing the task, and a filter for sex was used to get the answers of a group of women and men separately. Even though Toloka only permits binary sex desegregation, an option for more inclusive gender identity was included within our survey. As we became employers on the platforms as requesters, the payment offered for the completion of the survey as a task was calculated based on the minimum wage in the United States and the time to complete such a task The payment to submit the survey was calculated as (minimum payment x fraction of an hour). For example, if we assume a payment of USD 7.25 and the time to complete the survey was 20 minutes, the worker would receive an amount of USD 2.4 (7.25*0.333). In order to gain insight on how gender affected their work, we used different questions regarding widely recognized gendered factors such as time poverty, work-home balance, outside attitudes to the work done, care work, and direct opinions about if they had considered gender as a liability to their work. The second version of the survey was an iteration based on feedback about the need to include additional questions to better understand the characterization of women working in

crowd work in Latin America, their motivations, needs, socioeconomic context, skills, as well as hopes. The sample of the second survey included 57 women and three men.

Results of survey experiments

First survey

<u>Men respondents</u>:

Three men were part of the sample. The pilot had 3 age groups, from 18 to 55 years old. Educational profiles included associate professionals and university level participants without a degree. They found Toloka through social networks, and they do crowd work because they prefer remote work and to supplement their income.

The respondents grew up in a hierarchical, relationship-based culture. They perceive knowing their work styles, having a vision for their career, knowing how to look for job opportunities, and what strengths help them at work. They work from home with their own computer most of the time; they don't make much use of external resources to perform their tasks at Toloka (no Facebook, no WhatsApp, no Reddit). They have been working in Toloka between a month to a year; their average income ranges from .02$ to 2$ per task.

Negotiation of home care tasks is neutral for them because they can postpone them for later, or because someone else takes care of their children (if any). They consider that gender does not affect the way they perform collective work. They believe that their current schedule allows them to have enough time to rest, but not enough time for self-care.

Their families respect their work on crowd work platforms, because it supports them with household expenses; the 3 men stated spending most of their Toloka income on their families. They conducted their work without supervision from other people. They considered that remote work does not contribute so much to their financial independence, because they spend their income on their family, but that it brings them valuable skills for their CV. The respondents knew English thanks to private education, self-education and the use of translation tools. After the COVID-19 pandemic, the survey respondents considered that remote work became the new normal.

Table 1. Conclusion of qualitative answers of men sample (1)

Recurring and important topics:
- Self care: Not enough time for self care.
- Family income: All spent Toloka income on their families to some extent, men spent most of it on their families.

  Care work: Others in their proximity assumed care work to allow them time to do crowd work.
- Motivations: Valued other benefits, such as new skills.
- Alienation: Valued meeting other crowd workers. Perception that gender does not affect crowd work. Neutrality regarding crowd work

Women respondents:

The sample included 3 women from different age groups, from 18 to 45 years old. All respondents were based in their home country and worked on their own computers from home. Two of them had master's degrees. These women considered that they had identified their work culture and values as egalitarian and relationship-based. They knew their work style, goals, career vision, and how to look for opportunities, together with their main strengths. The respondents had been working on the Toloka platform since 2021-2022 and learned about it through social media. They worked with Toloka to make the most of their free time and to generate supplementary income. Their average income ranged from $0.10 to $4.00. Two people mentioned YouTube as the platform where they frequently asked for help.

They considered that their schedules allowed them to have time for self-care and to rest. They did not consider spending their extra income on their families but on themselves, and believed that crowd work contributed to their financial independence. They stated that they did not have to negotiate care-work to do collective work, because their schedules did not conflict. They were not supervised when doing collective work; their families remained neutral or supported their work on collective platforms. They perceived that the COVID-19 pandemic impacted the way they viewed collective work, because they can work more safely from home.

Table 2. Conclusion of qualitative answers of women sample (1)

Recurring and important topics:
- Self care: Work and care balance allowed them for self-care time.
- Education Capital: Youtube as a tool to search for tutorials on how to perform the tasks. Use of tools to translate English and develop English competencies.
- Independence: Crowd work is perceived as complimentary income that may promote financial independence.
- Alienation: No contact with other women crowd workers. May value contact with crowd workers. Perception that gender does not affect crowd work. Respect or neutrality regarding crowd work.

Second survey

A second survey was conducted with 61 questions in Spanish on the Toloka platform; in this iteration, a group of 60 people was considered to submit the survey. It was divided into specific sections, and most of the questions were of Likert scale type and open-ended questions. In this case, the time to submit the survey was established to 23 minutes for women and men crowd workers that were interested in doing the task, and it did not include a filter for gender. Additionally, the survey was applied specifically in the region of Latin America, using a filter for every country of the region. From the total of respondents, only 3 identified as men, and 2 preferred not to be identified.

There were four age groups represented in the sample, and 76% of respondents were 18 to 35 years old. The majority answered that their gender had little or no impact on their experience performing crowd work, while 17% believed it had some or an important impact on it. In accordance with the age group, most respondents had a professional degree, or had unfinished university studies; 13% held a master's degree.

Regarding their professional self-assessment, most of them knew their work strengths, work style, work environments, and had a long-term vision of their career. They could balance their professional and personal goals and they wanted to develop new skills. Other respondents regarded themselves as being neutral in goal balancing, building long-term vision of their careers, and identifying work environments that interest them in their career search. Few of the respondents perceived that they did not know how to search for careers relevant to their interests, and did not have a long-term vision of their careers.

They did not work in public spaces or libraries. They worked about 2 times per week from their own devices and at home, and spent about 20 minutes completing a common task, in some cases up to 40 minutes; the task they performed most is data labeling, web page testing, and surveys (only 10% performed audio transcription tasks).

Income per task was in the range of $0.01 to $0.16. 91% of the respondents conducted other economic activities in addition to Toloka. The most recurring reason for them to perform

crowd work tasks was to gain supplementary income. Additionally, a preference for working from home, and having schedule flexibility, were important factors. Few people considered it a recreational activity, and one respondent mentioned crowd work as one of few possibilities for employment. Regarding the impact of the geopolitical situation of their country of origin, participants considered the main impacts to be economic, which reduced their professional prospects and required them to perform a second economic activity; 21% remained neutral.

Table 3. Conclusion of qualitative answers of women sample (2)

Important topics towards the development of the proposed tool:
- Communitary context: 55% have not established relationships with other workers, and 68% would like to have a way to do so.
- Important Skills to improve:
  - Online marketing
  - Programming, IT, web development
  - English
  - Text interpreting, writing, reading comprehension, transcription, excel ○ AI, advanced tech tools
  - Finance
  - Soft skills, leadership
  - Habits of perseverance and concentration to execute work online
  - Effective communication, project management
  - Teamwork
- Topics of major importance for crowd workers in Toloka in the development of a tool:
  - Sharing experiences and tips to improve profitability.
  - Doing more tasks, productivity
  - Training for professional and skill development
  - Finding the best tasks
  - Clarification of instructions, ambiguities and possible mistakes
  - Financial issues; better payment on tasks

Most importantly, the respondents shared their preferred features for a tool. Explicability of tasks is important; although crowd workers in Toloka can contact the requester, multimedia resources and tutorials are useful for them to fully understand the tasks. Similarly, as 63% of respondents use search engines and translating applications as tools, prioritizing language learning features is an important skill to develop. Regarding the perceived need to share experiences with other workers, the main goal of a communication feature would be to increase task completion, profitability and performance.

Design Implications

Figure 1. Example of our intelligent platform that guides women crowd workers to build community and guides them to increase their skills with the use of AI systems.

We propose to design an AI-powered social connection and recommendation system, specifically designed to assist crowd workers in building a supportive community and developing their technical and soft skills. Figures 1, 2, and 3 provide examples of the AI based platforms and chatbot tools we have started to explore based on our survey results, understanding the unique needs and challenges facing crowd workers. The chatbot will be designed to match individuals with similar interests, goals and skill sets, allowing for the formation of meaningful connections and opportunities for skill-sharing and collaboration. The AI of a system that connects crowd workers would likely involve a number of different technologies and techniques. One key component would be natural language processing (NLP), which would allow the system to understand and analyze the language used in the survey responses and profile information provided by users. This would enable the chatbot to understand the interests, skills, and goals of each individual, and match them with others who have similar characteristics.

Another important component would be machine learning algorithms, which would be used to analyze the data collected from the survey and continuously improve the system's ability to make accurate and relevant connections. These algorithms would be trained on the survey data and would be able to identify patterns and relationships between different users, and make predictions about who would be the best match for a given individual.

The impact of such an AI-powered system would be to create a more connected and supportive community for crowd workers, by connecting them with others who have similar interests, skills and goals. This would help to increase opportunities for collaboration, skill-sharing and networking, which would be beneficial for their careers and professional development. Additionally, a system like this would help to increase representation and visibility of crowd workers in their industries, and could help to bridge the gap of inequality that they are facing, in terms of access to resources and opportunities.
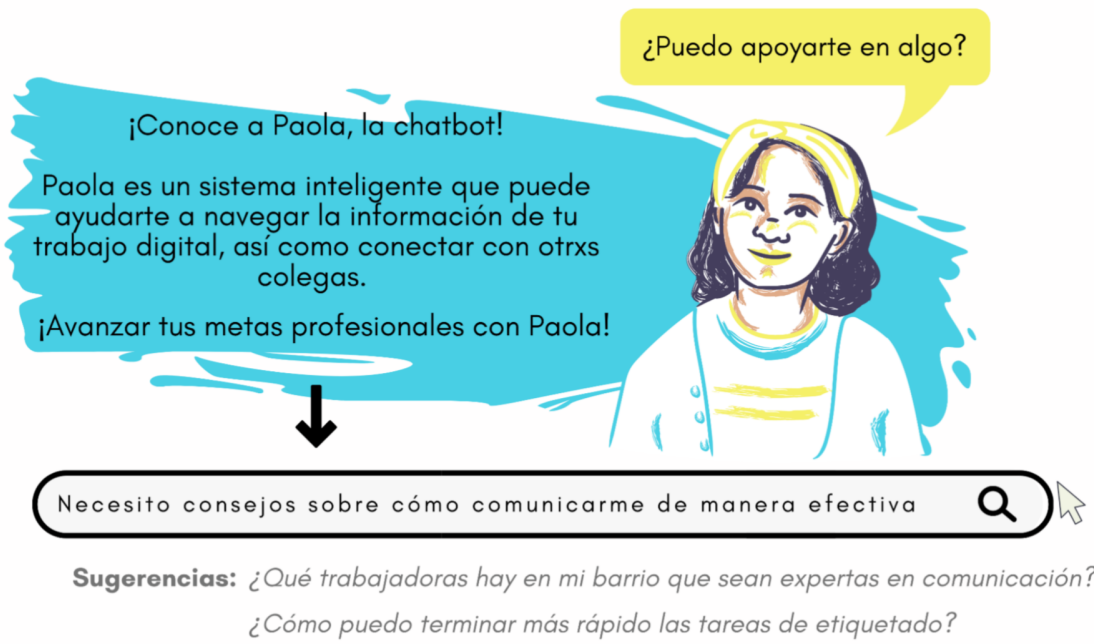
Figure 2. Example of the intelligent chatbot users would interact with to navigate the platform and assist them in finding information and other workers relevant to their questions and interests.

Figure 3. Example of the workers the intelligent chatbot recommends the user connect with based on their search query about effective communication.

The technical development of the platform would involve a combination of machine learning and natural language processing techniques. The AI would analyze the job listings on the platform and the resumes and profiles of crowd workers in order to match them with the most relevant and suitable job opportunities.

One of the key ways in which the AI would integrate the cultural background of the crowd workers would be through the use of NLP to understand the language and terminology used in job listings and resumes. For example, if a job listing used a specific term or phrase that is particularly relevant to the Latin American community, the AI would be able to identify and match that job with a crowd worker who has relevant skills and experience.

Another way in which the AI would integrate the cultural background of the crowd workers would be through the use of machine learning algorithms to analyze the data and identify patterns and trends specific to the Latin American community. For example, the AI may identify that crowd workers are underrepresented in certain types of jobs or industries, and could then use this information to prioritize job opportunities that would help to increase representation of crowd workers in those areas.

As this work is being designed with Latin American women in mind, we will refer to cultural theory to guide the integration of Latin American heroines into the design of our technology. We identified key women heroes from Latin America, such as Professor Paola Ricuarte, and utilized their stories to drive the design of our AI platform in a direction that is inspiring and encouraging for women workers. Additionally, personifying the chatbot lends to dispelling workers' concerns about interacting with AI and keeps them engaged while using the service.

Figure 4. Example of the feedback survey and option to leave a suggestion at the end of each page of the platform in order to collect information on how workers are enjoying the service.

Finally, another AI system we envision based on our survey results is one that could recommend different jobs to crowd workers based on their preferred job style and values that would involve the use of deep learning techniques. Deep learning is a type of machine learning that uses neural networks to process and analyze large amounts of data, and can be used to understand complex patterns and relationships between different data points.

The AI system would be trained on a large dataset of job listings and resumes, and would use deep learning algorithms to analyze the data and identify patterns and trends specific to the Latin American community. The system would take into account the preferred job style and values of the crowd workers, and use this information to match them with jobs that are most relevant and suitable to their skills and experience.

For example, if a crowd worker values a flexible schedule and has experience in customer service, the AI system would recommend jobs that offer flexibility and are in the customer service field. If a crowd worker values a work-life balance and has experience in marketing, the AI system would recommend jobs that offer a balance of work and life and are in the marketing field.

The AI system would also continuously learn and improve over time, by analyzing the data generated by the crowd workers, such as feedback on the jobs they applied for, and performance data. The impact of such an AI-powered system would be to increase the visibility and opportunities for crowd workers in the gig economy, by matching them with jobs that are relevant and suitable to their skills, experience, preferred job style and values. Additionally, it could help to bridge the gap of inequality that they are facing, in terms of access to resources and opportunities, and it could help to increase representation and visibility of crowd workers in their industries, giving them more opportunities for career growth and development.

Literature Review: Global South diagnostic

As a new type of labor that leverages technology and new digital platforms, crowd work is gaining popularity. Globalization and digitalization have had a significant impact on labor markets, with particular shifts in how individuals work, changing work patterns towards a new emphasis on individual agency, and responsibility. Crowd-sourcing arose as a method of sourcing labor in which individuals or organizations use digital platforms to tap into the aggregate skills, knowledge, and expertise of a vast, geographically dispersed workforce.

There are a variety of definitions for crowd work and crowd-sourcing. Estellés-Arolas and González-Ladrón de Guevara offer a definition that summarizes the key elements of the concept (2012: 197):

> "Crowdsourcing is a type of participative online activity, in which an individual, an institution, a nonprofit organization, or company proposes to a group of individuals of varying knowledge, heterogeneity, and number, via a flexible open call, the voluntary undertaking of a task. The undertaking of the task, of variable complexity and modularity, and in which the crowd should participate bringing their work, money, knowledge and/or experience, always entails mutual benefit. The user will receive the satisfaction of a given type of need, be it economic, social recognition, self-esteem, or the development of individual skills, while the crowdsourcer will obtain and utilize to their advantage that which the user has brought to the venture, whose form will depend on the type of activity undertaken".

The broad range of tasks that can be completed includes audio and video transcription, content moderation, data gathering and processing, image identification, transcription, and annotation, as well as translation. These digital platforms serve as a "middleman" between employers and workers, assisting in the supervision of the creation, submission, acceptance, and payment of the work completed. Amazon Mechanical Turk (AMT), Upwork, Clickworker, PeopleperHour, CloudFactory, CrowdFlower, Freelancer, and Microworker are a few examples of digital networks for crowd work.

Collective work is presented as an opportunity to increase the participation of women in economic activities, as the flexibility of the remote work scheme can be a viable alternative to earn primary or additional income for some women. Previous studies have demonstrated that crowd workers are driven by intrinsic motivation, as well as extrinsic motivation, such as financial reward and the task itself, which allows them to be self-employed and to feel empowered.

Workers from Kenya and the Philippines were among the first to be externally employed to label data; however, after 2018, 75% of the leading AI crowd work platforms' workforce was Venezuelan, and it is possible that other Latin American countries will follow in the context of the post COVID-19 economic recovery. These platforms offer the chance to access a sizable pool of local and international labor to increase productivity, efficiency, and market reach, lowering operating costs for requesters. Crowd workers can also actively design their work-life balance, and freely decide how many contracts to take. Despite such benefits, the regularity of labor and income, working conditions, social protection, skill utilization, freedom of association, and the right to collective bargaining are some of the challenges that crowd workers face.

Currently, the majority of crowd work is not governed by labor laws, rather, the platforms themselves determine the working conditions. The crowd worker's ability to exert control over working time is frequently constrained by the nature of the task or the necessity to maximize productivity. In order to make enough money, many crowd workers must put in hours that are significantly longer than those typical of normal employment. Additionally, they have little alternatives for legal action in cases of unfair treatment. Crowd workers may be arbitrarily deactivated as a result of algorithmic management, which results in them losing their

revenue and being prevented from using the platform. They may also be liable to additional penalties or disciplinary actions without having the chance to appeal if considered unfair. In other words, digital platforms and their algorithms have a direct impact on crowd workers, as many managerial aspects are built into algorithmic decision-making and performance and monitoring systems.

According to the Organisation for Economic Co-operation and Development (OECD), platform workers are frequently misclassified as not-being employees, because platform companies would incur costs that are 20–30% higher by doing so. By classifying workers as independent contractors, platforms are able to avoid indirect costs associated with employee rights, such as the right not to be fired unfairly; in addition to direct costs, such as minimum salaries, maximum hours, paid leave, and paid sick leave. The contracting figure under which the employment relationship is established classifies all workers as independent contractors, which exempts platforms from guaranteeing basic protections. On the other hand, the ubiquity that characterizes the platforms also works in their favor, because in addition to the fact that accountability mechanisms are scarce, it is unclear under the jurisdiction of which region or country they operate and must respond.

Moreover, the International Labor Organization's (ILO) survey results from 2015 and 2017 indicate that there are significant gender pay gaps, and that many crowd workers earn less than the local minimum wage. Depending on the platform, women were paid between 18 and 38% less on average than males. There were regional variances in average earnings as well, with Northern America ($4.70) and Europe and Central Asia ($3.00) having higher wages than other regions with pay ranging from US$1.33 (Africa) to US$2.2 (Asia and the Pacific). According to the results of the ILO survey, respondents chose to engage in crowd work despite the low pay for a variety of reasons, including their preference for working from home and the opportunity to supplement their income.

Growth of crowd work in the Global South and LAC

At the beginning of their operations, most collective work platforms worked by subcontracting people from Global North countries. However, in recent years, some structural changes within the industry have led to the diversification of the workforce. One of these changes is related to the growing demand for optimized data from the automotive industry

and the emergence of specialized data generation platforms with higher levels of precision. Unlike traditional collective work platforms that only serve as mediators between clients and workers, specialist platforms guarantee the client data with an accuracy level of 99 percent. In order to produce large volumes of information with these characteristics, new strategies have been implemented to reduce costs and increase their competitiveness; investing in the optimization of processes and quality control, use of gamification mechanisms, and recruitment of workforce in new regions and countries of the Global South, such as Latin America, where lower hourly rates can be settled.

In addition to the convenience of recruiting a lower-paid labor force, the structural and social conditions of most countries in the Global South and Latin America make the collective work scheme on platforms an attractive source of income for their employees, as summarized in Figure 5.

Global South and LAC
Economic and inequality crisis + job losses

Increased demand for optimized data from the automotive industry
Emergence of "specialist" platforms for generating data with higher levels of precision

Workforce recruitment in Global South and LAC countries
Labor force with lower pay schemes compared to other regions

Figure 5. Interaction of factors for the growth of collective work in Latin America and the Global South. Own elaboration with information from Schmidt, F. (2019).

Platforms for collective work in Venezuela

In Venezuela, the economic crisis has fueled an explosion in investment and growth of foreign collective work platforms. During the last five years, the market conditions have led to a massive recruitment by collective work platforms in the country, which represented 75% of the workforce of companies such as Hive Micro and Spare5, equivalent to about 200,000 workers in 2018. In 2019, the list of companies interested in the Venezuelan workforce increased; for instance, the company Scale launched various recruitment strategies through social media campaigns that highlighted the possibility of obtaining an attractive, stable, and long-term income. In 2020, this same company launched Remotasks Plus, a collective work platform with tasks and an interface in Spanish initially aimed exclusively at the Venezuelan population, that was later launched to the rest of the world. The COVID-19 pandemic presented itself as a new opportunity for the growth and expansion of the workforce for collective work platforms. During this period, Scale oriented its efforts on training crowd workers; they conducted boot camps in countries from different regions, including Latin America, Asia, Arabic-speaking countries and the southern region of Africa.
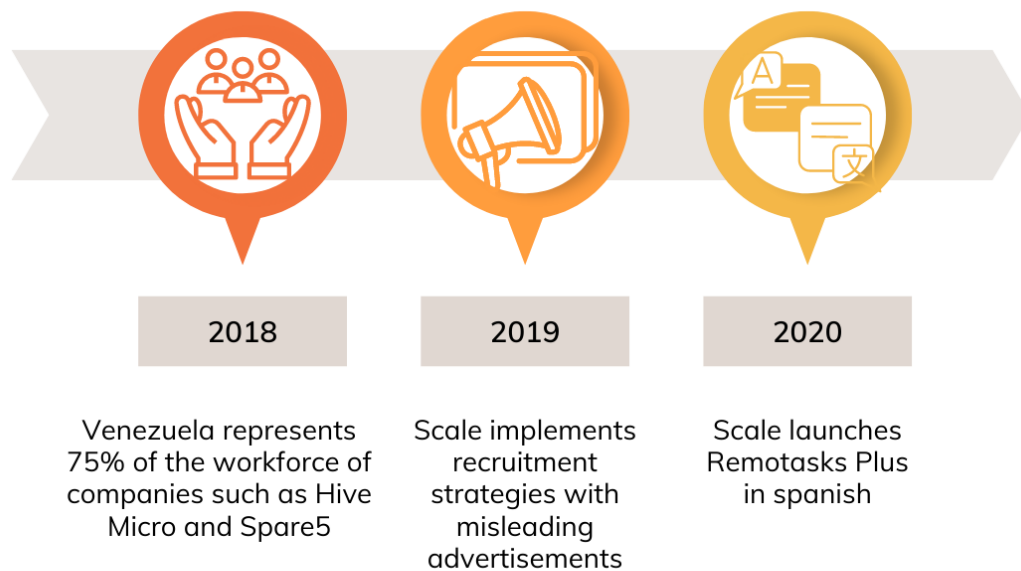
2018	2019	2020
Venezuela represents 75% of the workforce of companies such as Hive Micro and Spare5	Scale implements recruitment strategies with misleading advertisements	Scale launches Remotasks Plus in spanish

Figure 6. Growth of collective work in Venezuela. Own elaboration. Own elaboration with information from Schmidt, F. (2019)

Characteristics of the labor force in the Global South

In 2015 and 2017, ILO carried out the first editions of a survey for people engaging in collective work in some of the main English-speaking platforms, such as Amazon Mechanical Turk, Microworkers, Prolific, Clickworker and Crowdflower (see Table 1). The total sample of participants contemplated 3,500 observations of people from 75 countries. Sociodemographic data of the interviewed population were collected, as well as the main motivations for carrying out this work and some of their working conditions.

Table 1. Composition of the sample for the first and second edition of the survey of people who work in collective work carried out by the ILO. Source: Rani, U., Berg, J. & Furrer, M. (2018)

		2015 (S1)	2015 (S2)	2017
AMT	United States	686	573	231
	India	128	104	251
	Other countries	0	0	7
CrowdFlower		353		355
Clickworker				455
Prolific				495
Microworkers				556
Total		1 167	677	2 350

From the disaggregated analysis of the data presented in Work and Labor Relations in Global Platform Capitalism it is possible to identify some of the characteristics of the people who carry out crowd work in the Global South. For example, it highlighted that, considering the total population participating in both surveys, the distribution by gender is 80% men and 20% women. Regarding the educational level, 45% of the people surveyed had university education, 27% postgraduate studies, and less than 13% had high school-level education or a lesser attainment.

The population surveyed in 2017 doubled the observations of the first edition of 2015. The comparison between both periods indicates that Venezuela remained the country in Latin America with the largest population working in collective work platforms, while India

occupies the same place in relation to the rest of Global South countries (see Figures 7 and 8).
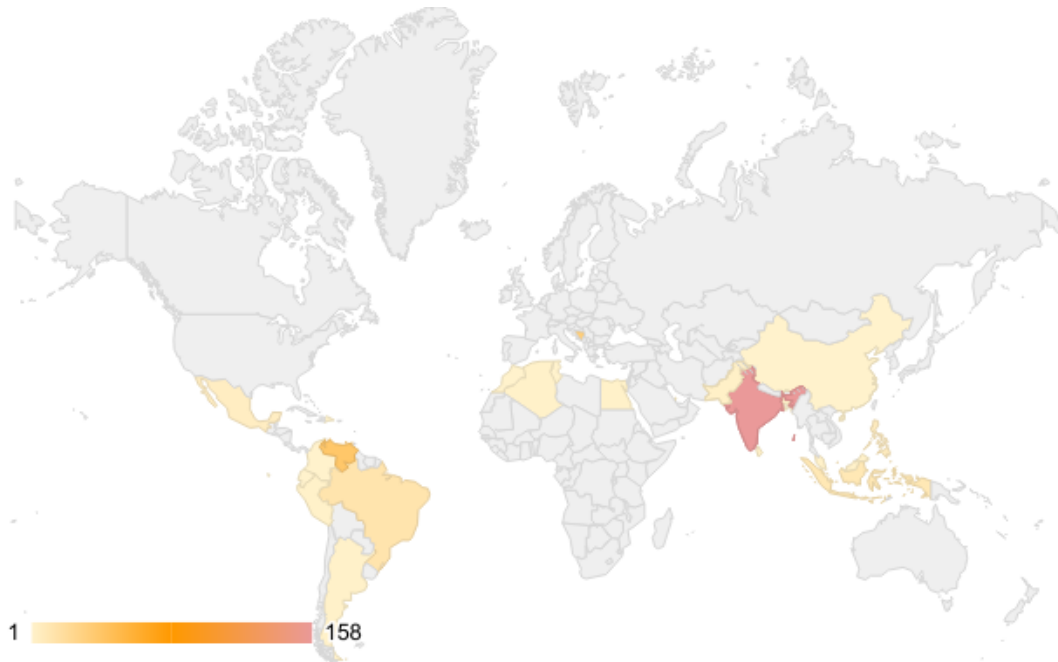

Figure 7. Country of origin of the people participating in the collective work survey carried out in 2015 by ILO. Source: Own elaboration, with data from Rani et al. (2018)
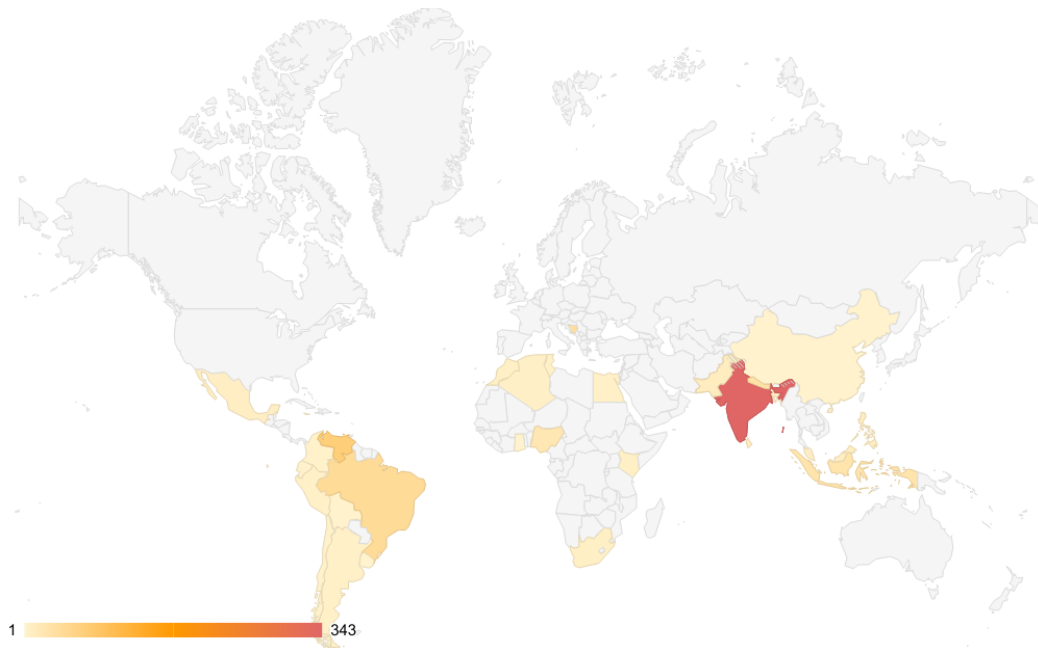

Figure 8. Country of origin of the people participating in the collective work survey carried out in 2017 by ILO. Source: Own elaboration, with data from Rani et al. (2018)

Low wages and lack of social protection and benefits that characterize these work schemes cast doubt on the optimistic perspective on collective work as an economic opportunity. Unlike other traditional schemes of job insecurity faced by thousands of workers around the world, this emerging digital workforce is not only limited in their salary expectations and professional growth in the short and long term, but also carries out their work behind anonymity, under a scheme of minimal institutionalization.

Although these characteristics predominate, it is worth noting that the emergence of new business models of specialist platforms represents changes in their operating logic, and this translates into a different experience for the people who work on them. Specialist platforms guarantee accurately tagged data to their customers, and reducing the costs of manual labor has been critical to achieving this goal. For this reason, the recruitment of people in Global South countries has increased exponentially in recent years.

Massive recruitment in multiple regions of the world has required the translation of tasks, as well as internal regulations and prior training so that people of different origins can participate in the execution of such tasks. More importantly, the role of mediator is transformed into that of direct supplier, and with this, changes are also generated in the employer-employee relationship and power asymmetries. In order to improve the accuracy of the data, the specialist platforms implement long training periods; the assignment of tasks is governed by a hierarchy in which the best paid tasks are also the most sophisticated, and there are specific requirements in terms of "expertise" or training necessary to access them. This gives crowd workers the alternative to distinguish themselves from their peers, specialize, and add value to their profiles.
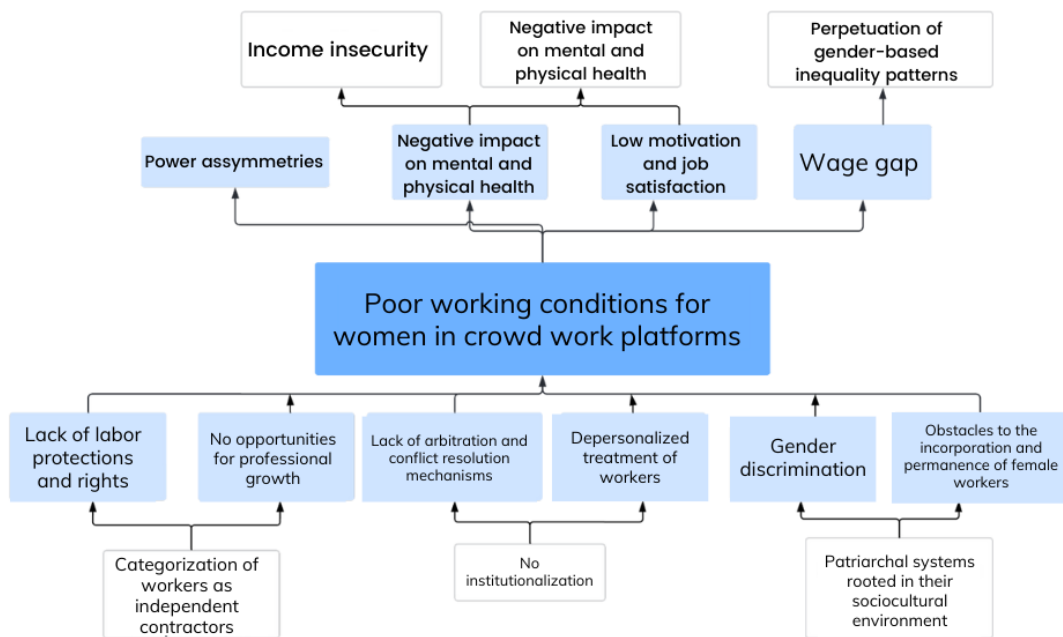
Gender perspective in crowd work

Figure 9. Problem tree based on literature review. Source: Own elaboration.

Collective work has represented an opportunity to increase the participation of women in economic activities, as the flexibility of remote work schemes can be a safe and viable alternative to earn income for many women. Women have reported feeling insecure when commuting, while some have been victims of harassment on public roads, or are dependent on their relatives to move from their home to any other point. Figures 9 and 10 show some of the motivations of women in the Global South to participate in collective work, as well as the limitations to work more hours on the platforms. While the preference and need to work from home, together with leisure activity and complimentary income are the main reasons women engage in crowd work, the category "other" points out to reasons that are less known and might not be taken into account when diagnosing the ecosystem.
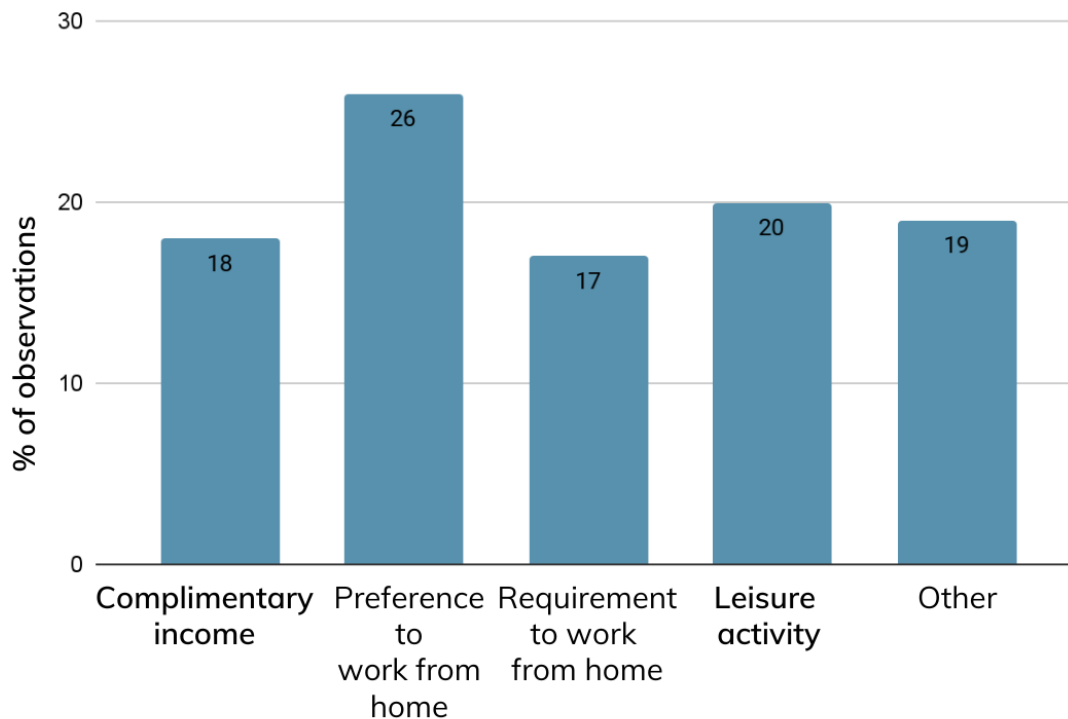
Figure 10. Motivations of women in the Global South to carry out collective work. Source: Own elaboration, with data from Rani et al. (2018)

It is important to note that experience as a platform worker is different for men and women. For example, in an experiment in 2022 with a group of 16 low-income Indian women to learn about their adaptation process and experiences working on a collective work platform through mobile devices, the exercise revealed that these women had a significantly longer adaptation process than women from other regions. One of the reasons for this was their fear of damaging their family's reputation by carrying out these types of activities. Their families did not regard online work as appropriate, mainly because it could translate into being exposed to other men and to unsafe spaces; they also had little trust in the use of platforms as a means to receive payments.

The freedom and autonomy of women working on platforms was significantly limited: the expectation for the women interviewed was to deliver their income to the male figures of their family who manage the household's resources. On the other hand, unlike the men participating in the study, all the women worked from home, in the presence of other

members of their family. It is no surprise that these women received little support from their families to work online and were questioned for prolonged cell phone use. This situation generated them anxiety, and required them to be in a state of constant vigilance, so as not to be reprimanded while doing collective work. To avoid these situations, the women interviewed tried to finish housework first, and then continued working on their cell phones. According to data retrieved internally by Karya, a crowd work platform that is designed for low-income, digitally-novice communities in India, 79% of women perform tasks at dawn, between 12:00 and 3:00 am. During the interviews, many women reported that by working on the platform, their rest periods had been reduced, and on average, they slept 5 hours a night, which compromised their sleep hygiene.

As part of the strategies that facilitated the process of adaptation and permanence of women in collective work, the relevance of the support of other women and the creation of spaces and communities through digital channels such as WhatsApp stood out. These spaces were useful to share doubts, experiences, and frustrations with other colleagues who also knew and worked on the platforms. For example, newly hired women usually looked for the guidance of more experienced women, who could share tools and useful technical knowledge to speed up the induction stage, as well as recommendations to achieve greater precision in the tasks performed.

The value and effect of support networks were also reflected in the percentage of tasks successfully completed by women (64%), compared to male participants (48%). Over time, the channels that were originally spaces to answer questions about the platform were used to share other concerns and personal experiences beyond collective work. In addition, these networks played an important role as spaces for support and emotional containment, essential for women to decide to continue on the platforms despite the many obstacles they face every day. For these women, the income generated through their work meant much more than a way to contribute to the family economy; an opportunity towards their financial autonomy. It also affected their self-perception, self-confidence and autonomy.

There are particular challenges conditioned by the social and structural circumstances of the region in which women crowd workers live. For example, in some regions of South Asia, women experience obstacles similar to those reported in India in terms of access to

electronic devices and autonomy in their use. Compared to men, women of the region are less likely to have a cell phone, and when they do have access to one, its use is usually restricted and supervised by male figures in the family.

Decolonial care and well-being centered AI development

Decolonial theory compares center and peripheral power spaces that are present as colonial continuities. Ricaurte proposes an analytical model that organizes socio-technical dimensions of present coloniality through the means of data; on the "economy" dimension of the model, data labor, and the economic value of data are mentioned as expressions of coloniality. In crowd work dynamics, technology corporations take on the role of metropolises, while workers and their organization are the power-contesting periphery.

A commonly present obstacle to developing well-being centered AI systems is that current processes can obscure asymmetrical power relations and underlying values that are not questioned and replicate colonial understandings of work; crowd work gets concealed by utilitarian approaches to technology. Ethical guidelines for AI have been proposed as a means to create common knowledge and standards for socially-responsible technologies, but as Ricaurte also includes in the model, the data epistemology that precedes many of those guidelines is yet to be structurally questioned.

The design and convening of ethical guidelines for AI started in the Global North, catalyzing dialogue in the space of autonomous weapons. The discussion on contextual values in science and technology has guided the establishment of ethical values as a minimum standard: respect for persons, beneficence and justice. While such principles are helpful to ignite ethical imagination, they are not specific enough to assess the differentiated impact of AI development in the Global South, whose communities and territories face the hardest impacts of emerging technologies. Even when predominant AI actors from the Global South engage in such dialogues, the paradox of participation "wherein inclusion can exist while structural harms persist", directly challenges the development of decolonial AI.

Marie Therese Png examines the tensions of South and North in the Inclusive AI landscape, and poses three necessary steps towards AI governance that is centered on the Global South: "To 1. Engage in a historical-geopolitical analysis of structural inequality and the coloniality of

geopolitical power asymmetries and international legal frameworks; 2. Co-construct roles for Global South actors to substantively engage in AI governance processes; and 3. Identify mechanisms and protocols that mitigate "paradoxes of participation" and redress institutional power imbalances, in order to meaningfully engage with underrepresented stakeholder groups".

Rather than enumerating more ethical principles, this paper-to-prototype process is oriented to engage and work around the tensions that arise when putting them in practice. Critical science can help foresight the prospective harms of crowd work in the Global South, and lay the foundation for a participatory, slow, reflective and cooperative development of empowerment tools.

Policy Recommendations

Crowd work platforms

- **Allow for the filtering and disaggregation of data.** This is critical to allow researchers, policymakers, together with decision makers within the private sector to better understand the characteristics of crowd workers by region, countries, and other variables of interest.
- **Establish transparent standards for assigning equitable payment**. Stipulated national and regional minimum hourly wages must be taken into account in the calculation of payment per task conducted. Following a formula as the one we used can be a first step into discussing what fair payment could entail and look like.
- **Support and promote English skill building and training to non-English speaking crowd workers.** Doing so could lead to a higher task completion and accuracy, honing crowd workers skills.
- **Design and implement Education Programs for workers.** Crowd workers can benefit from learning about the applications of the data and information they provide to requesters through education programs in STEM, the social sciences and business.
- **Establish channels, mechanisms, and processes to improve communication with crowd workers.** In addition to features that enable communication with requesters, the platform provider should establish processes and mechanisms to provide valuable information to crowd workers.

International organizations

- **Update estimates on the size of the collective labor force.** Richer data will include an intersectional perspective that takes gender, race, class, and disability as variables of interest globally, disaggregating information regionally.
- **Promote crowd workers organization.** Support initiatives that connect stakeholders from academia, civil society and policy making of the crowd work space with workers to promote organization.
- **Mitigate the paradox of participation.** Identify mechanisms that redress institutional power imbalances within initiatives that aim to include and benefit crowd workers.

Policy makers, researchers and community leaders

- **Promote equitable crowd work and governance of AI in the Global South.** Co-construct roles for Global South actors to engage in AI governance, centering on the well-being of crowd workers in varying contexts.
- **Exchange best practices.** Share insight and knowledge with platforms' policy teams, with other researchers and leading voices in the crowd work space.
- **Co-create a feminist community of practice.** There is a need to continue growing the body of crowd work research through a gender lens; to build on the existing research, the creation of a feminist community of practice can support nascent voices in the field.

Conclusion

The development of Artificial Intelligence relies heavily on the crowd-sourcing of data from international labor. While crowd work platforms such as Hive Micro, Appen, Amazon Mechanical Turk, Spare 5 and Scale in the Global South may provide employment opportunities for women, there are still issues of association opportunities, collective bargaining, transparency, fair pay, and professional development that need to be addressed. Additionally, most prior work has not focused on documenting the experience of women crowd workers in contexts like Latin America, to create technology that can truly empower them. In this research, we focused on conducting a qualitative study with more than 60 Latin American gig workers to understand the need, challenges, and opportunities they face in

crowd work. We conducted one of the first studies to uncover the unique perspectives of these populations, presenting key insights and providing design implications for creating AI that can empower women in the Global South. We also provide policy recommendations for platforms, international organizations, policymakers, researchers, and community leaders to promote more equitable and gender-responsive crowd work.

Acknowledgments

References

Adams-Prassl, A. & Berg, J. (2017), When Home Affects Pay: An Analysis of the Gender Pay Gap Among crowd workers.. http://dx.doi.org/10.2139/ssrn.3048711

Berg, J., & Rani, U. (2021). Chapter 4: Working conditions, geography and gender in global crowd work. En "Work and Labour Relations in Global Platform Capitalism" . Cheltenham, UK: Edward Elgar Publishing. https://www.elgaronline.com/view/edcoll/97818 02205121/9781802205

Bourdieu, P. (1990). The Scholastic Point of View. Cultural Anthropology 5(4): 380-391.

Collins, P. H., & Bilge, S. (2020). Intersectionality. John Wiley & Sons.

Drahokoupil, J., and Vandael, K. (2021). Labour and the Platform Economy. Edward Elgar, Cheltenham. 1-46

Durward, Blohm & Leimeiste (2020) The Nature of Crowd Work and its Effects on Individuals' Work Perception, Journal of Management Information Systems, 37:1, 66-95. https://doi.org/10.1080/07421222.2019.1705506

Estellés-Arolas, E., & González-Ladrón-de-Guevara, F. (2012). Towards an integrated crowdsourcing definition. Journal of Information Science, 38(2), 189–200
European Institute for Gender Equality (nd). Gender Mainstreaming. https://eige.europa.eu/gender-mainstreaming

European Institute for Gender Equality (nd). Gender Analysis. https://eige.europa.eu/taxonomy/term/1143

Fieseler, C., Bucher, E. & Hoffman, C. (2017). Unfairness by design? The Perceived Fairness of Digital Labor on crowd working Platforms. Journal of Business Ethics. 156.

Fredman, S., et.al. (2021). International Regulation of Platform Labor: A Proposal for Action. Weizenbaum Journal of the Digital Society, 1(1), w1.1.4.

Haro, K. & Hernández, A. P., (2022). How the AI profits from catastrophe. MIT Technology Review. https://www.technologyreview.com/2022/04/20/1 050392/ai-industry-appen-scale -data-labels/

Hastrup, Kirsten (1992). Out of anthropology: the anthropologist as an object of dramatic representation. Cultural anthropology, 8/1/1992, Vol. 7, Issue 3.

Hesse-Biber, S.N. (Ed.).(2014). Feminist Research Practice: A primer (2nd. ed.). SAGE Haraway, Donna. (1988). Situated knowledges: The science question in feminism and the privilege of partial perspective. Feminist studies 14 (3), 575-599.

Idowu, A., Elbanna, A. Digital Platforms of Work and the Crafting of Career Path: The crowd workers' Perspective. Inf Syst Front 24, 441—457 (2022). https://doi.org/10.1007/s10796-020-10036-1

International Labor Office. (2019), 'Work For a Brighter Future: Global Commission on the Future of Work'.

Kittur, A., Nickerson, J., Bernstein, E., Shaw, A., Zimmerman, J., Lease, M. & Horton, J. (2013). The future of crowd work. In Proceedings of the 2013 Conference on Computer supported cooperative work (CSCW '13). Association for Computing Machinery, New York, 1301-1318.

Mansbridge, J. J., & Okin, S. M. (2020). Feminismo: breve introducción a una ideología política. Página indómita.

Marie-Therese Png. (2022). At the Tensions of South and North: Critical Roles of Global South Stakeholders in AI Governance. In 2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22). Association for Computing Machinery, New York, NY, USA, 1434—1445. https://doi.org/10.1145/3531146.3533200

Margaryan, A. (2017). Understanding crowd workers' Learning Practices. Paper presented at 17th Biennial Conference of the European Association for Research on Learning and Instruction. EARLI 2017, Tampere, Finland

Mohamed, S., Png, MT. & Isaac, W. Decolonial AI: Decolonial Theory as Sociotechnical Foresight in Artificial Intelligence. Philos. Technol. 33, 659—684 (2020). https://doi.org/10.1007/s13347-020-00405-8

Nickerson, Jeffrey V. (March 20, 2013). Crowd Work and Collective Learning. A. Littlejohn & A. Margaryan (eds.), Technology-Enhanced

Professional Learning: Routledge, Forthcoming, Available at SSRN: https://ssrn.com/abstract=2246203

O'Higgins, N. & Pinedo Caro, L. (2022). crowd work for young people: Risks and opportunities. ILO Working paper 50. https://www.ilo.org/global/publications/working-papers/WCMS_837670/lang--en/index.html

Paolacci, Chandler, Ipeirotis (June 24, 2010). Running Experiments on Amazon Mechanical Turk. Judgment and Decision Making, Vol. 5, No. 5, 411-419, Available at SSRN: https://ssrn.com/abstract=1626226

Rama A., Divya S., Vivek S, Kalika B., & Aditya V. (2022). Feeling Proud, Feeling Embarrassed: Experiences of Low-income Women with Crowd Work. In Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22). Association for Computing Machinery, New York, NY, USA,298, 1–18. https://doi.org/10.1145/3491102.3501834

Rani, U., Berg, J., Harmon, E., Silberman, M. & Furrer, M. (2018). Digital labour platforms and the future of work: Towards decent work in the online world. International Labour Office

Ricaurte, P. (2019). Data Epistemologies, The Coloniality of Power, and Resistance. Television & New Media, 20(4), 350–365. https://doi.org/10.1177/1527476419831640

Schmidt F., A. (2019). Crowdsourced Production of AI Training Data: How human workers teach self-driving cars how to see. Forschungsförderung 155, Hans-Böckler-Stiftung, Düsseldorf.

Scott, J. W. (2008). Género e Historia. Fondo de Cultura Económica.
Sheldon, R. (2016). Tragic encounters and ordinary ethics. Manchester University Press
Toloka. (s.f.) Toloka documentation. https://toloka.ai/en/docs/

Toloka. (s.f.) Toloka documentation: Filters. https://toloka.ai/en/docs/guide/concepts/filters

Toloka. (s.f.) Toloka documentation: Messaging. https://toloka.ai/en/docs/guide/concepts/messagin

Biographies

Cristina Martínez Pinto is the Founder and CEO of PIT Policy Lab. She worked as a consultant for the Digital Development Global Practice at the World Bank, led the AI for Good Lab at C Minds and co-founded Mexico's National AI Coalition IA2030Mx. She holds a Master's degree in Public Policy from Georgetown University and a Bachelor's degree in International Relations from Tec de Monterrey.

Luz Elena González Zepeda specializes in Development Cooperation, with a focus in tech policy for the LAC region. She managed the PIT Policy consultancy with the Government of Colombia to design a Governance Model for the Data Infrastructure. She has a B.A in International Relations from Tec de Monterrey, in Mexico. She is also content curator for Women in AI Mexico, and event promoter within the MCODER.ai community for women in data science.

Tatiana Telles is Puentech Lab's Public Policy and Gender Specialist, with 10 years of experience in Government and Civil Society. Focused on problem solving. Feminist. She seeks to create more egalitarian communities.

Norma Elva Chávez is a distinguished researcher and educator based at the National Autonomous University of Mexico (UNAM). She serves as the director of both the UNAM Civic AI Lab and the Dispositivos Logicos Programmables lab. Her research primarily focuses on the design of robots and technologies aimed at empowering women, enabling them to achieve greater success. dShe has dedicated her career to understanding the biases and challenges faced by women, and using that knowledge to develop innovative technologies that support women in their diverse goals.

Maya De Los Santos is a first-generation Afro-Latina pursuing an Electrical and Computer Engineering degree in Northeastern University's Honors Program and is a research assistant in the Civic AI lab led by Dr. Saiph Savage. In her current work, Maya is dedicated to designing and researching human-centered AI systems that ensure fair work opportunities for Latina gig workers. Maya is also a Google Generation Scholarship Recipient, a Grace Hopper Celebration Scholar, and was selected for an NSF grant to participate in Carnegie Mellon University's REUSE research program.

Alberto Navarrete Hernández works as a professor at the National Autonomous University of Mexico in the Faculty of Engineering, in the Department of Computer Science, teaching digital systems design. He is also an assistant in the laboratory of Programmable Logic Devices, performing in the design of laboratory practices of the subjects. His field of expertise and main interest is digital signal processing and data analysis, using Machine Learning techniques.

Saiph Savage is an Assistant Professor at Northeastern University in the Khoury College of Computer Sciences, where she directs the Northeastern Civic A.I. Lab. The impact of Dr. Savage's research has led to her being named one of the 35 innovators under 35 by the MIT Tech Review, be awarded large grants from the National Science Foundation (NSF), and enjoy recognition from UNESCO for having globally one of the most impactful AI research projects. Her work has also won paper awards at top scientific venues in human-computer interaction, including ACM CHI, CSCW, and the Web Conference. The growing public interest in Dr. Savage's research has garnered international press coverage from the BBC, New York Times, The Economist, Deutsche Welle, Vice, Wired, Forbes, and Fortune.