Embedded Reproduction in Platform Data Work

Julian Posada^a

^aFaculty of Information & Schwartz Reisman Institute, University of Toronto, Canada

NOTICE

The Version of Record of this manuscript has been published available Information, and is inCommunication \mathscr{E} Society (2022)http://www.tandfonline.com/10.1080/1369118X.2022.2049849

ABSTRACT

This paper focuses on the experiences of Latin American data workers who annotate data for machine-learning algorithms through labor platforms. It introduces the notion of "embedded reproduction": the relationship between embeddedness, the degree to which non-economic institutions and their social environment constrain socioeconomic activity, and social reproduction, or the activities that nurture, maintain, and regenerate the workforce. The analysis of 38 interviews with platform workers suggests they are situated in a highly disembedded market due to the lack of regulations on the data production process, giving free rein to platforms to set rules to their detriment. This article explores how this disembeddedness shapes social reproduction by studying three forms of collective social support received by workers: from family members, neighbors and local communities, and online groups. The support of these networks is primarily local, depends on high levels of trust, and is gendered. These findings suggest that platform data work is unsustainable from an embedded reproductive perspective since platform intermediation leads workers and local communities to carry out the social and economic risks associated with this form of gig work. This research invites a dialogue between the embeddedness framework with social reproduction as well as a consideration of the importance of nature and natural resources in the study of social environments.

KEYWORDS

Embeddedness; social reproduction; artificial intelligence; platform labor; data work, gig economy

1. Introduction

Machine learning (ML), a subset of artificial intelligence that centers on agents that can learn with experience, often in the form of vast amounts of data (Russell & Norvig, 2020), has driven recent technological developments. For example, in the areas of facial recognition and virtual assistance. Some of the steps in the development of ML algorithms, notably the construction of datasets and the verification of algorithmic outputs, require human labor in the form of data entry, annotation, and evaluation (Tubaro et al., 2020). This "data work" is often outsourced by companies and research institutions to workers worldwide through labor platforms, hybrids of firms and markets that transact information, goods, and services between multiple actors (Casilli &

CONTACT Julian Posada. Email: julian.posada@mail.utoronto.ca. Address: Massey College, 4 Devonshire Place, Toronto, ON M5S 2E1, Canada

Posada, 2019). This platform labor is commodified or treated as an object "exchanged in a market" (Marx, 1978, 36). Workers are freelancers or independent contractors who are usually paid low wages per task; their work is subject to algorithmic management and constant surveillance (Woodcock & Graham, 2020). Although most of the requesters of outsourced data work come from advanced economies, workers come primarily from countries in the Global South (Graham et al., 2017) and carry out most of the social and economic risks associated with their work (Tubaro & Casilli, 2022).

In this paper, I explore how this commodified work shapes and is maintained by networks of people who provide for and receive social support to and from workers. To address this research question, I will decenter the productive activity of data work and the socioeconomic relationship between workers, platforms, and requesters, and will instead focus on the local and online communities that allow workers to be ready for work. These economic and social relationships will be analyzed through the lens of "embedded reproduction," a concept that explores the relationship between the *reproduction* of society and the *embeddedness* of social and economic phenomena within their environments. Embeddedness considers the degree to which economic activity — in this case labor transactions — are regulated through non-economic institutions (Polanyi, 2001), including governments, communities, and families who depend on their social environment. On the other hand, social reproduction accounts for the essential reproductive labor and institutional support that nurtures, regenerates, and maintains the workforce (Hester & Srnicek, 2017), and the importance of common goods and local organizations to mitigate the environmental damages of the disembedded economy.

This paper addresses the embedded reproduction of platform data work in Latin America by analyzing 38 semi-structured interviews with workers of three platforms, the nature of workers' personal networks, and their access to forms of institutional support and shared resources. Most of these workers are located in Venezuela, a country experiencing deeply intertwined economic, political, and social crises (ILO, 2020). The ongoing hyperinflation, at an average rate of around 50% and having depreciated the economy by 60% compared to 1999 (Singer, 2021), makes platform data work the only reliable source of income in US dollars for many workers in a moment where the COVID-19 pandemic has severely damaged the local labor market. Despite the socioeconomic and health crises, the infrastructure of the country, built primarily during previous periods of economic prosperity, has proven essential — albeit unreliable — to access the global market of data work. The social, economic, and health crises in the country and the conditions of platform labor marke this type of work highly disembed-ded.

To address the issue of disembedded labor, I will first conceptualize embedded reproduction through a review of the literature on embeddedness and social reproduction. Then, I will provide details on the empirical setting and methodology of the study. The findings will be divided into two parts: first, an account of the disembedded working conditions of platform workers; and then, how family, neighbors, and online communities support workers in these conditions. The findings will focus on instances of domestic labor and economic support that allow workers to be ready for work, worker organization through social media and the management of common online resources, and the locally managed shared natural resources that workers require for subsistence.

The qualitative analysis of these interviews suggests that data workers are not an atomized workforce, and instead, depend on support from different social groups. This support is highly gendered and relies on high levels of trust, notably with relationships outside of family and households. In the latter cases, access to shared resources, both local and virtual, proved essential for workers to increase income from platforms, get access to crucial public resources like water, and mitigate the effects of environmental damage.

This paper concludes that social reproduction when labor is commodified is unsustainable because it depends on a normative embeddedness that negatively affects the livelihoods of workers, creates dependencies, and contributes little to the longterm development of local communities. Based on these results, the paper recommends expanding the scope of research on embeddedness in platform labor outside of socioeconomic exchanges, to better account for embedded reproduction through the importance of social networks and natural and common resources for the sustainable development of local communities and technologies worldwide.

2. Theoretical Framework

This paper considers embedded reproduction from the perspective of data work intermediated through digital platforms. Contemporary data work companies and the platforms they operate are a subset of the *gig economy*, or "labor markets that are characterized by independent contracting that happens though, via, and on digital platforms" (Woodcock & Graham, 2020, 3). As independent contractors, workers bear more financial and social risks related to the economic activity of AI production than their employers (ILO, 2011; Tubaro & Casilli, 2022). Furthermore, while workers operate remotely, they are still subjected to algorithmic surveillance and control. For instance, platforms monitor workers' accuracy and the time they use to perform tasks (Wood et al., 2019). They also provide reputation and evaluation systems that reduce uncertainty from clients (Lehdonvirta et al., 2019), requiring workers to provide unpaid and unrecognized emotional labor to maintain and sustain their reputation scores (Gandini, 2018; Irani & Silberman, 2013). Their work arrangement also allows platforms to ban, terminate, or "deactivate" the accounts of workers unilaterally, without explanation, and sometimes without recourse (Gray & Suri, 2019, 94).

Recently, scholarship on the platform economy and gig work began exploring the commodification of labor and the role of social networks through the lens of embeddedness, a concept that accounts for the dependence of social phenomena on their environment, including the relationship between social actors, non-economic institutions, and nature. For instance, Wood et al. (2018) argue that despite the intermediation of platforms, the outsourced and fragmented character of their labor pushes workers to rely on personal networks of trust to counter platform power. Furthermore, Tubaro (2021) argued that the existing relationships in platform labor occur not only between individuals but also among platform firms as organizations.

This paper will contribute to these conversations by studying the support received by platform data workers using the theoretical lenses of embeddedness and social reproduction. This latter concept, developed in sociology and Marxist feminist scholarship, focuses on the reproductive dynamics of capitalism, accounting for the labor and institutional support required to nurture, regenerate, and maintain the workforce. Embedded reproduction, or the relationship between these frameworks, will be used to study the configurations of platform labor in Venezuela and how it conditions the support received by workers and their access to natural and common resources.

2.1. Embeddedness

The concept of embeddedness originates in Polanyi's book *The Great Transformation* (2001), where he uses this term in two distinct ways. First, the term refers to the activities, objects, and subjects not created by markets but exchanged as commodities. Namely labor, land, and money which in turn become disembedded from their specific normative, cultural, and legal constraints accorded by society (Polanyi, 2001). Second, the author argues against the orthodox economic model of the market, based on individual transactions, affirming that it is instead "embedded and enmeshed in institutions, economic and non-economic" (Granovetter, 1986, 250). These institutional regulations link workers with the fabric of society and its particular defined functions (Beckert, 2003). These two different types of embeddedness, which Peck (2013) denominates "hard-Polanyi" and "soft-Polanyi" respectively, and Wood et al. (2018) call *normative* and *network* embeddedness, were developed separately within the fields of sociology and geography.

Research on *network* embeddedness was popularized in economic sociology by Granovetter, who argued that beyond institutions, economic exchanges are embedded within personal networks of trust that stem from micro-level interactions (Granovetter, 1986). This structuralist approach to embeddedness stressed the importance of interpersonal networks of trust to assess the valuation of goods and improve productivity and innovation (Granovetter, 2005). This approach has been criticized, however, for over-individualizing market exchanges (Beckert, 2003) and thus ignoring Polanyi's conception of markets as "fully social institutions, reflecting a complex alchemy of politics, culture, and ideology" (Krippner, 2002, 782).

On the other hand, research on *normative* embeddedness has focused mainly on the recent wave of commodification that characterizes neoliberalism (Burawoy, 2010). Burawoy builds on Polanyi's definition of false commodities to explain that, due to their "fictitious" nature — by turning labor, land, money, and in his regard, knowledge, into market products — they lose their use-value, such as by destroying the productive capacity of labor (Burawoy, 2010, 310). In this context, commodified labor is no longer a human quality but a resource or input in the production process. Instead of being embedded in "the broader flows of cultural life and of living matter," labor becomes "restructured together under the umbrella of constitutional rights and laws founded on principles of individual rights to private property guaranteed by the state" (Harvey, 2014, 58). This conception of labor fails to recognize that, beyond this individualized and marketized vision, it is embedded "in other social institutions, such as the family, education, politics, and the healthcare sector" and "intimately related to gender, race, age" and other intersecting identities (Kalleberg, 2009), all central to the study of social reproduction.

This paper will continue to explore the configurations of *normative* and *network* embeddedness in platform data work, a type of labor that is "precarious and fractured" (Woodcock & Graham, 2020, 16), with a particular emphasis on the effects of this embeddedness in local support. Because this analysis on embedded reproduction aims not to focus on the productive aspects of platform labor but instead on how embeddedness affects and is conditioned by social reproduction, this latter theoretical framework will complement this research.

2.2. Social Reproduction

Social reproduction recognizes that "human labor is at the heart of creating or reproducing society as a whole" (Bhattacharya, 2017, 2). It focuses on the often undervalued activities which are part of a "social contract" for the benefit of capitalism (Federici, 1975), notably in the form of housework (Fortunati, 1995) and care work (Federici, 2020).

Marx introduces the concept of social reproduction when describing the productive and reproductive dynamics of capitalism. He argues that "whatever the form of the process of production in a society, it must be a continuous process... therefore, as a connected whole, and as flowing on with incessant renewal, every social process of production is, at the same time, a process of reproduction" (Marx, 1978, 401). Thus, capital requires the constant reproduction of material goods, services, social systems, and structures to access and maintain production conditions and relations.

This initial Marxian notion of social reproduction was developed in the 20th century to explore the reproduction of inequalities in society from a *structural* perspective, including how "ideological state apparatuses" maintain the existing class divisions (Althusser, 1970) and how social structures legitimize — and reproduce — the structural dominance of particular classes over others through the economic, social, and cultural capital of their social actors (Bourdieu, 1979; Bourdieu & Passeron, 1970).

In recent decades, Marxist feminist scholarship has studied the "reproductive dynamics of capitalism" in social reproduction (Bhattacharya, 2017) with a particular emphasis on *embodied* reproductive labor. In a pioneering contribution to this perspective, Dalla Costa & James (1972) discussed how the unrecognized work of women is fundamental in producing labour-power for capitalism, thus expanding the Marxist autonomist notion of the social factory, in which the sphere of production encompasses society as a whole (Tronti, 1962). This stream of social reproduction considers domestic work as one of the oldest forms of invisible labor carried in the private sphere and an essential aspect of social reproduction (Federici, 2004; Jarrett, 2014). While the conception of domestic labor recognizes that labor's invisibility precedes capitalism's development, it also highlights how the labor of women and oppressed communities becomes invisible in specific relations of production (Ferguson, 2020). Furthermore, this scholarship expands the institutional view on social reproduction, centering on the crucial role of welfare in society, for example, through education, healthcare, social security, and other institutions essential for reproducing and maintaining the workforce (Hester & Srnicek, 2017).

Based on the significance of social reproduction for capitalism's productive forces, in this paper, I will study how the *normative disembeddedness* in which outsourced data work takes place affects social reproduction, both through *embodied labor* and the renewal of *social structures*, by exploring the qualitative aspects of *network embeddedness* in the case of Latin American workers, their households, and local and online communities. To this end, I propose the term "embedded reproduction" to account for the relationship between the degree to which non-economic institutions and their social environment constrain socio-economic activity, and how this embeddedness shapes and is maintained by socially reproductive activities.

In platform labor, a focus on embedded reproduction means studying how the commodification of labor and subsequent reliance on networks of support condition and depend on activities that nurture, maintain, and regenerate the workforce. Centering on reproduction and not only on production enables the study of artificial intelligence from the perspective of the essential labor of workers in precarious conditions as well as the members of their social networks who perform reproductive labor. With this framework, I also reiterate the importance of considering worker organization and common resources as fundamental aspects of social environments in embeddedness, and account for the conditions that shape social reproduction through resource availability and management, health conditions, and community organization, which are, in turn, affected by the disembedding of productive activities from environmental regulations.

3. Research Design

This paper will explore embedded reproduction in platform data work by focusing on the types of support received by Latin American workers primarily located in Venezuela. The research will focus mainly on three main platforms here anonymized as Tasksource, Workerhub, and Clickrating. Table 1 describes the types of tasks available and the applications of this labor in the artificial intelligence market.

Platform	Primary Tasks	Applications
Clickrating	Data entry	Search engines
	Algorithmic verification	
Tasksource	2D/3D image classification	Self-driving vehicles
	2D/3D semantic segmentation	Internet of things
Workerhub	2D image, text & video classification	Content moderation E-commerce
	2D semantic segmentation	
	Text transcription	

Table 1. Studied platforms

Most interview participants were from Venezuela, a country experiencing severe economic and political crises. The country has experienced the highest levels of inflation globally in recent years, reporting around 3,000% rates in 2020 (Agence France-Presse, 2021). The coronavirus pandemic has exacerbated this crisis by pushing more Venezuelans to unemployment and dependence on informal economies for subsistence (Schmidt, 2019). In embeddedness terms, platform data work in Venezuela represents an extremely disembedded market. From an employment perspective, workers do not receive the rights commonly associated with standard employment relations, and from an institutional perspective, the political and economic crises of the Venezuelan state combined with the COVID-19 pandemic has reduced its ability to offer protections and services to workers and their communities.

Because platforms serve as intermediaries between users in markets, it is difficult to establish the demographic characteristics of workers without accessing the companies' datasets. Moreover, this research has provided evidence of users with multiple accounts and families working under a single account, suggesting that even platforms ignore the true size of the worker population. Therefore, the workers of the three studied platforms constitute a hidden population, rendering sampling difficult. Previous attempts to measure the size of the platforms' workforce used online surveys through the platforms (Difallah et al., 2018; Ludec et al., 2019). However, this approach is only compatible with generalist platforms, such as Amazon's Mechanical Turk, that allow requesters (including researchers) to access the marketplace. Instead, the three platforms specialized in data work studied here restrict their marketplace to particular AI developers and maintain a high level of secrecy about their operations.

Previous research on data work suggests that workers communicate mainly through

social media and online forums (Wood et al., 2018; Yin et al., 2016). Following these insights, I identified groups of Latin American data workers of the three studied platforms on Facebook, Discord, WhatsApp, and Reddit. These groups are diverse, and their nature will be analyzed in depth below. For instance, Tasksource uses a private Discord channel as a forum moderated by company employees to share information and coordinate with their approximately 1,500 online members. One open group for workers of Workerhub and Clickrating exists on Facebook and comprises 1,100 registered users who discuss work in these platforms without the moderation of company employees.

Based on this information, I used opportunity sampling by approaching workers in 11 open groups. First, I asked moderators permission to post a call for participants. Then, I asked participants to sign a consent form providing further information on the study and their rights as participants. A total of 38 workers participated in semi-structured interviews conducted in Spanish and collected virtually through Zoom between February and August 2021. Participants were asked about their work for the platform(s), personal networks, and access to public services. Table 2 describes their demographic characteristics. Among the participants, ten were part of four different family groups. I analyzed these qualitative data using thematic coding and analysis (Richards, 2015).

Category		Count
Country	Venezuela	36
Country	Colombia	2
Gender	Female	21
Gender	Male	17
	18-24	12
	25 - 34	13
Age	35-44	7
	45-54	3
	>54	3
	Clickrating	5
Platform Used [*]	Tasksource	13
i lationii Oseu	Workerhub	21
	Other	5
	Single	20
Marital Status	Civ. Union	9
Mailtai Status	Married	8
	Divorced	1
	Minimum	2
Household Size	Median	3
TIOUSCHOID DIZE	Average	3.84
	Maximum	8

 Table 2.
 Demographic characteristics of interview participants

*Six participants worked for more than one platform

4. Findings and Discussion

4.1. Platform Disembeddedness

Much like Wood et al. (2018) found in their research, the AI platform labor market is highly disembedded, showing a lack of economic and institutional constraints on its operations. The devalued currency and difficult access to goods and services in Venezuela since the beginning of the crisis have seen the emergence of an informal economy dependent on the US dollar. Many participants saw their salaries substantially reduced, while others experienced underemployment and increasing difficulty in finding jobs. For instance, Melba, a retired woman who used to work for *Petróleos de Venezuela* [Petroleum of Venezuela] (PDVSA), the state-owned oil company, lost most of her income to the devaluation of the country's currency:

Even though I receive a salary and a pension, and my husband as well, our paychecks don't cover anything. I wonder, how can people survive here in Venezuela when someone like me, with a [monthly] pension worth 1,800,000 bolivars [around US\$1], can't buy half a dozen eggs? You can't buy a piece of cheese or bread. (Melba, Workerhub)

The hardships of the country's economic, political, and social situation combined with the current pandemic make labor platforms one of the only sources of income available to workers. The ability to circumvent the geographical restrictions or the local labor market and earn income in US dollars were the main drivers for many workers, like Wilmar, to start working online:

The economic crisis has pushed many Venezuelans like myself to find work alternatives online. Thank God that I have a computer and internet connection. Wages here have been very low in the past four years, and many of us had to look online for external sources of income and improve our living conditions a bit. (Wilmar, Tasksource and Workerhub)

Workers of the three platforms perform data entry, data annotation, and algorithmic verification (Tubaro et al., 2020), using different criteria to select particular tasks. The most important was the payment, ranging from 30 cents per 1,000 tasks to a few dollars per task depending on the platform, the complexity of the task, location, and other sometimes undisclosed factors. Regarding difficulty, classification tasks take a single click and are often paid less, while segmentation tasks can take around one hour depending on the complexity and size of the image and are paid accordingly. Workerhub users also reported ethical and social reasons related to their task choice. For example, workers like Alcides reported disliking tasks related to pornography due to working in a family context:

There are tasks where you have to classify pornographic images, and I avoid them because it's not correct to work on that in public....I have little siblings at home, and we've had to hide the computers when some of us were working so they wouldn't see those images. (Alcides, Workerhub)

Clickrating and Tasksource workers do not often have a choice and are instead assigned tasks. They have to pay attention to the system because tasks can appear at any time, including in the early morning hours, and disappear soon after. Once a task is identified, workers must pass a qualifying exam and, in Tasksource, receive training from company coaches and pass further tests. In all cases, these periods of evaluation and training are unpaid.

Once accepted to perform the task, algorithms constantly monitor workers. A technique used by all platforms consists of making workers annotate data that the client has previously labeled. If the worker does not annotate the data according to the client's direction, the algorithm will ban the worker from the task. The platform may ask the worker to annotate the data again and, if the worker provides a different solution, the algorithm will ban them. Workers do not have any recourse against banning and are not told the reasons for this outcome. In more severe cases, the algorithm can suspend or terminate the worker's account, and, as with instances of banning, no explanation is given. Workers who face expulsion from the platform will lose all accrued income. The following testimony is from Mario, a worker who was part of a special program run by Tasksource where he was paid per hour and expelled after voicing a complaint:

I had worked 17 hours, and then I worked one night for 8 or 7 hours and, the next day, they didn't count those hours...The next day, I contacted [the moderators]. The guy, instead of giving me an answer...told me, "wait for a solution," and that was it. I didn't insist. That was in the morning, at around 10 AM. In the evening, around 7 PM, I tried to log in again to see what they would say about payments, and they had kicked me out of the group. (Mario, Tasksource)

Workers are paid differently according to the platform. All platforms pay workers in US dollars through digital wallets, such as PayPal for Workerhub and Clickrating and Airtm for Tasksource. Furthermore, platforms also pay workers in cryptocurrency, namely *Etherium* for Workerhub, and a cryptocurrency managed by Airtm and tied to the US dollar for Tasksource called *AirUSD*. Many of these digital wallets ask workers to verify their identity by matching their national identification cards with a photograph of their faces, for example, using a third-party facial recognition technology from the United Kingdom in the case of Airtm. This technology has had low accuracy for workers with darker skin, an issue widely documented in the computer science field (Buolamwini & Gebru, 2018; Gebru, 2020). In the case of Roberto, a Black worker:

I tried to verify my account. I sent a copy of my ID card and a selfie holding the ID. I had to wait for an entire month. On one occasion, I asked the moderators on Discord because they were threatening to close unverified accounts. My account wasn't verified and I was scared because that's how I made money, and I was going to practically lose my job. (Roberto, Tasksource)

The dominant informal economy also means that the government does not receive any taxes from any of the transactions of these informal markets. Aside from government aid in the form of a few dollars per month for specific individuals and a bag of groceries for most families that arrives once every few months, the presence of governmental institutions has been vastly reduced. Only two of the workers had health insurance tied to their employment or that of a close family member with a government entity. None of the children and youth in their households were receiving in-person education except for a few who attended universities in urban centers. Aside from the public services mentioned above (e.g., electricity, drinking water) the availability and quality of other services such as waste disposal have also been impacted. In sum, the presence of public institutions in the lives of the interviewed workers has been reduced and, in many cases, become non-existent.

While the government's presence has been diminished by economic hardship and high levels of corruption, giving free rein to platforms to set their own rules of the online labor market, its actions have been fundamental in allowing platforms to operate in the country. Workers rely on infrastructures that predate the current crisis. For example, many utilize computers acquired for free or at a low cost through two governmentsponsored programs. One program consisted of giving free laptops called "Canaima" (after a National Park) to school children and teachers that were either kept by families or sold and acquired in the informal market.

The work arrangement of the platforms — where workers are considered independent contractors, coupled with the government's lack of constraints and reduced support — suggests that the labor market for outsourced data work is commodified. For platforms, this disembeddedness means that there is no regulation from external parties except for the constraints derived from contracts with partners, such as third-party online wallets and clients. Therefore, platforms have almost total control over the internal

market of data work. For workers, the fragmentation and distribution of tasks, the algorithmic management, and the flexibility in which workers can create an account and work remotely, reflect an "illusion of choice" (Burawoy, 1982) that characterizes the restrictions over workers and the control exerted by platforms. This context creates a dependency on the part of the workers, manifested in an over-reliance on platforms, high economic, social, and personal risks, and a high degree of power from platforms to regulate its internal labor market.

4.2. Types of Social Support

The previous section described aspects of the labor process in Venezuelan platform data work to illustrate its degree of *normative* disembeddedness: labor is considered a commodity, workers have few rights, government intervention is minimal, and the ongoing economic and health crisis has reduced the capacity of institutions to provide support for workers.

This section will explore the network embeddedness of workers, describing the support received from (1) household and family members, (2) colleagues and online communities, and (3) neighbors and local communities, and discuss the implications of this support for social reproduction.

While many of the relationships and activities outlined here were described by data workers, they are not necessarily exclusive to the outsourced data production market. Following the focus on embedded reproduction, these findings illustrate how the disembeddedness described above shape socially reproductive activities and also, since workers depend on them for survival, how they help maintain the global production of artificial intelligence.

4.2.1. Support from Household and Family Members

Previous research has suggested that gig workers rely heavily on social networks for different types of support. For example, there is evidence of support from family members (Drahokoupil & Piasna, 2019) and online groups (Yin et al., 2016) for digital workers. This research suggests data workers from Venezuela rely heavily on personal networks to compensate for the market's difficulties and the lack of protection from institutions. These networks are primarily local and the forms of support follow traditional gendered divisions.

None of the participants lived independently, and all of them shared their households with other family members, often living in properties acquired through inheritance. In most cases, workers reported that the income received from the platforms constituted the majority or, in many cases, the entirety of their household income. The economic crisis has pushed other family members to find work in the informal economy, often selling food or taking up casual jobs. Several families chose to work together for the platform after losing their jobs due to the pandemic. For example, in one family comprised of two parents, María and Rogelio, and four children (two of whom are adults), all but the youngest worked for Workerhub using two computers at different alternating times. María explains their shared workload as follows:

Those who are working, let's say "full time," are my husband and me. When we are resting, [our children] work. They just fill in. I would tell them: "we'll rest" or "I don't feel well," or my husband would tell them: "it hurts from sitting so long." So, we stop working and they work for a while. I can't have my kids working full time, no. It's on us to work. They just fill in. For example, at noon, when we're cooking lunch, they work.

We start working again at around 2:00 or 3:00 PM. They stop, and we keep working. (María, Workerhub)

In most cases, workers relied on the domestic support of household members who do not work for the platform. Most of these were women who would perform domestic duties such as cleaning, cooking, and grocery shopping. In most cases, all household members shared some of these duties, although to a lesser extent than female members who are not platform workers:

On my days off, I buy groceries. My wife is the one who does laundry and cooks. Sometimes when I'm not busy, I cook, especially for lunch. My mom cooks for breakfast and dinner because she likes to keep herself busy. My aunt does cleaning, or sometimes divides tasks with my wife and mom when I'm busy. (Eduardo, Clickrating)

Right now, my mother-in-law supports us. Monday to Friday, she takes care of our house, she takes care of my children, she cooks lunch, dinner, breakfast. She bathes the kids. All of that. (Wilmar, Tasksource and Workerhub)

Young participants also reported preferring to live with other family members — usually their parents — to share household costs such as food and public services. Family members also support each other in case of illness in a context where health insurance is rare:

I got malaria, a very bad case. My aunt, who's always been like a mother to me, took care of me and brought me to her house. I was treated at her house. It was better with her because, in my father's house, I used to do the majority of chores that she does here: cleaning, cooking, grocery shopping....My dad helped but, you know, I was the "woman of the house" because I was the eldest [daughter]. When I came here, everything was calm; the emotional burden was less. After six months of treatment, I got cured of malaria, and I decided to stay here. (Olivia, Tasksource and Workerhub)

Besides the emotional and domestic support, family and household members also provide economic support. Notably, those living abroad contribute with remittance and those situated nearby provide access to resources necessary for work and affordable housing:

After my mother went to Spain to try to find a job and send us money, I couldn't return to my grandmother's house because she didn't have internet. I was studying at the time, and I couldn't be in a place without [access to] the internet....So, I talked to my girlfriend, we were dating only for a few months, but I didn't have a choice. The only solution was to move to her house, which is where I live right now. (Jerónimo, Clickrating)

The experiences of workers demonstrate that they are not alone and can work online annotating data for artificial intelligence thanks to the domestic, emotional, and economic support of family and household members. This support becomes more pronounced in cases of extreme disembeddedness, both when the piece-wage regime foments dependency and when government support (e.g., in healthcare provisions) becomes reduced. Furthermore, the findings confirm once more that social reproduction is gendered and carried out primarily by women and, to a certain extent, children. This division of domestic labor signals the continuation of historical exploitative working regimes related to piece work central to the work of Boris (1994) and studied in relation to platform data work by Dubal (2020) and Gray & Suri (2019).

4.2.2. Support from Colleagues and Online Communities

Because of the use of opportunity sampling to reach out to participants through social media, most workers of this study were associated with a social media group or forum. Participants who were not associated with a social media group, mostly contacted through snowball sampling, reported relying on a family member to access information shared on these groups.

Workers of all online platforms have created open groups on Facebook, WhatsApp, Telegram, and Discord for data workers. These groups range from large ones with around 12,000 members to smaller ones with a few dozen. The larger groups are usually open, have fewer restrictions, and are often hosted on Facebook and Discord. Several moderators watch over members' posts. Because of the possibility of creating several communication channels, groups on Discord tend to receive higher engagement, allowing a wide diversity of topics to be discussed among members.

For example, a Discord group for Workerhub with over 600 members has a channel for each of the tasks on the platform, in addition to channels for payments and entertainment, gaming, and other topics not related to work. Meanwhile, Facebook groups, a platform that only allows a single "wall" or channel for posts, restricts topics allowed by moderators. In the case of Tasksource, open channels on Facebook were vital for workers who voiced criticism for the platform's closure of its pay-per-hour program and the progressive reduction of bonifications. Workers also reposted content that was taken down by the moderators of the Discord channel managed by the platform. Facebook users even proposed going on strike to improve their working conditions "because in the Philippines it worked, and they earn more than in Venezuela," according to one user. However, these calls for action did not materialize because of the reported high levels of dependency on the platform.

Workers of Clickrating, which restricts newcomers due to a system that matches task availability with reputation, prefer being part of smaller private groups on Discord, WhatsApp, and Telegram. These groups have a reduced membership, rarely surpassing a hundred members, and are not free to access. Workers must pay a monthly subscription to the administrator(s) ranging from US\$3–5, and provide identity documents, and personal information, such as their home addresses. Eduardo, a member of one of these restricted groups, justifies these measures by arguing that it allows for high levels of trust among members:

In our group, they ask you for all your contact information, even a copy of your ID or passport, ID number, a link to your social media profile on Facebook, Twitter, or Instagram, and the complete address where you live....The [administrators] verify all of this and even ask for reference numbers, like telephone numbers of referees, in case something happens. (Eduardo, Clickrating)

Ensuring high levels of trust is fundamental for workers who want to find reliable currency traders because, as mentioned above, platforms pay them in US dollars or cryptocurrency. Once platforms transfer wages into online wallets such as PayPal, Airtm, Payeer, or Binance, workers rely on "buyers" of dollars and cryptocurrencies. These are local intermediaries who exchange virtual currencies into bolivars and transfer them into local bank accounts. The hyperinflation makes the exchange rate fluctuate during the day, meaning that the currency loses value so fast that workers have to check current values many times per day on social media and spend the money as soon as possible. Usually transactions are made electronically because the hyperinflation also produced a shortage of banknotes. In this context, workers do not report any savings except on rare occasions and only in virtual currency. While currency traders are present in both free and subscription groups, the groups with fees have policies to prevent fraud. Eduardo states how these groups even operate as safeguards:

Almost a year ago, someone was selling Clickrating accounts, 10 accounts for \$100. A colleague bought the accounts, transferred the money through PayPal, but didn't receive the accounts. Thank God that the group had all the information [of the seller]. They located the person. We all thought it was a scam, but it turns out that the seller had lost access to electricity, so she couldn't transfer the accounts. The money was returned, and, thank God, nothing else happened. Scams happen all the time, especially in public groups. Because private groups are protected by administrators who have all your contact information, if a seller does not respond after a payment, they take care of that. They also ensure that only trusted sellers, who are well known in the market, can be in the group. They ask for a lot of references from people who have already done business with them. (Eduardo, Clickrating)

The small Clickrating groups with paid membership also provide bots that alert workers when a task will be available. Because tasks can arrive any time of the day and last until the data has been labeled, these bots allow workers to access the tasks before they expire. José explains:

On Clickrating, to know when a task becomes available, you have to be alert at all times because tasks don't last very long. There is a bot on Discord that tells me when to access tasks. As soon as a task is available, the bot tells me, sends me a link, and I can access the task. There are even tasks that only last for a few seconds, and these are only possible to access with a bot. Telegram also has bots. (José, Clickrating)

Another use of these smaller closed groups is to share guides about the tasks that explain the instructions in Spanish and provide tips to augment productivity and reduce time spent working on the project. These guides are fundamental in a context where most of the interviewed participants have limited knowledge of the English language, which is used to write the task instructions. The guides explain the instructions in detail and provide tips and answers to the requesters' queries.

Workers' reliance on members and online social groups suggests that, while working in a highly disembedded setting from an institutional perspective, workers are embedded in networks of trust, similar to other workers in the gig economy (Qadri, 2020; Wood et al., 2018; Yin et al., 2016). These groups provide crucial information, allow workers to criticize the platform and circumvent their censorship, identify tasks, interpret requesters' instructions, create networks of trust for currency exchange, and connect with peers. Notably, they collaborate by sharing information in the form of social media content and guides to understand better how platforms and their algorithms function as well as how to circumvent the limits and constraints imposed by them. This collective endeavor in data work resonates with that of other digital workers, such as media content creators (Bishop, 2019; Cotter, 2018), gig workers (Woodcock, 2020), and those in highly digitized industrial settings (Delfanti, 2021).

4.2.3. Support from Neighbors and Local Communities

The support received from members of local communities (e.g., neighbors) relates primarily to access to public services and natural resources. In Venezuela, the quality of public services in the country has diminished considerably while still being mainly dependent on public funds. None of the interviewees paid directly for electricity, water, sewage maintenance, and garbage disposal services. However, their access was severely constrained. For instance, electricity and drinking water were frequently inaccessible:

Internet is intermittent, and water too. We can lose access at any time, sometimes for two or three days, without being alerted....We can lose access in the morning, at night, same with electricity. Some power outages are severe, especially when they last a few seconds, because they can damage your appliances. Anyway, here everything is intermittent. (Alfredo, Workerhub)

Most workers interviewed had access to water only a few days per week, and, in many cases, it was polluted. Many rely on community-managed water pumps for access and water tanks for storage:

We have water from a well, but it's not drinking water... The water comes from the sea through a pump that supplies wells in the neighborhood....It's the property of the neighborhood itself: every house block has a well, and we designate someone to operate the pump every day to fill them....We choose this person as a community; all families have a say. It's not someone who works for us, but a member of the community, a neighbor who every morning comes to your well and fills it with the pump. Sometimes we forget to open the well, and the neighbor will yell at us to open the well so we don't waste the water. (Lucas, Workerhub)

Moreover, while participants in large cities reported access to garbage collection services, most of the participants located in suburban or rural areas disposed of their garbage in local landfills:

We have a space that you can call a landfill. A sort of improvised landfill where people from the community go and throw their garbage. Every few months, we get a truck to pick up the garbage and clear the landfill. But this isn't a service, like a recurrent pickup service. There's no truck that picks up our garbage every week in the community where I live. (Enrique, Workerhub)

When workers opted to handle their waste privately, they had to burn it within their properties, reporting high pollution levels in the air, water, and soil resulting from the lack of quality public services and mismanagement of local resources. For instance, Angélica, a worker who lived close to a PDVSA oil extraction site, reported oil spillage and fires in the vicinity of her community on top of fumes produced by private garbage incineration:

People have to burn their garbage from home. It's normal seeing people burning their garbage. In my case, I avoid it because my son has asthma [and] I am allergic, I can't. The smell of smoke makes me sick, same as my son...In my case, I am against burning garbage, but then I have to accumulate garbage in my backyard...Paying a truck that picks up your garbage costs \$20, and it's not affordable. Of course, keeping your garbage attracts animals, lots of scorpions, snakes, spiders, rats. That becomes normal. (Angélica, Workerhub)

These findings suggest that, in cases where disembeddedness is also provoked by a lack of institutional support to provide public services and natural resources, local communities become fundamental to maintain minimum living standards, notably through the joint management of shared resources like water. In the case of Angélica, it demonstrates that environmental damage also has a negative effect on social reproduction, in her case, by provoking health issues that hinder her ability to work.

Thus, network and normative embeddedness should not be thought of exclusively in terms of relationships between social actors and institutions but also their place within nature. Polanyi's work spoke of the dangers of the commodification of land, "another name for nature" (Polanyi, 1945, 76). "Turning land into a commodity destroys the

community which lives on and from it" (Burawoy, 2010), foretelling a negative cycle for social reproduction in which the exploitation that manifests through disembeddedness physically, socially, and psychologically harms the same laborers required for production. While the effects of data-driven technologies like artificial intelligence have been documented to produce environmental damage to a certain extent (Bender et al., 2021; Dauvergne, 2020), my claim here is not that data work has a substantial impact on the environment, but rather that by being embedded in nature, the harms derived from pollution and other forms of environmental damage affect the livelihoods of workers like Angélica, Enrique, and Lucas.

5. Conclusion

In this paper, I examined embedded reproduction in outsourced Latin American data work used to collect and annotate data for machine learning and verifying algorithmic outputs. The internal labor markets of these gig economy platforms are highly unconstrained and deregulated. Here, workers are considered "users," ear extremely low piece wages (comparatively to the income of their employers), and lack access to social and economic protections. Most participants are located in Venezuela, where the social, political, and economic situation, exacerbated by the COVID-19 pandemic, has reduced local opportunities for work and welfare from the government.

This configuration of the internal labor market of the platform suggests a high degree of disembeddedness in terms of labor commodification. Subsequently, data workers depend on their personal networks and local and online communities to reduce the social and economic risks associated with their work and manage common resources essential to maintain adequate living standards. In sum, workers are embedded from a *network* standpoint but also disembedded from a *normative* perspective. The latter accounts for the importance of sustainability in the livelihoods of workers because a disembedded economy can lead to the degradation of ecosystems with repercussions for workers' health and living conditions.

These varying degrees of embeddedness shape and are influenced by social reproduction both from the perspective of *embodied* labor necessary for the sustenance of life and the role of *social structures* and *systems* to condition society. Both gendered labor at home, and the management of local and online communities over common resources, are fundamental for data workers to ultimately perform tasks for the platform. As such, this article has explored the support received from household and family members, colleagues and online communities, and neighbors and local communities. Many of their collective reproductive labor occurs outside of the workspace as "hidden transcripts" (Anwar & Graham, 2020), allowing workers to improve their income from the platforms without subverting the data production process significantly due to the power imbalances that benefit platforms.

The embedded reproduction framework used in this article expands the current discussion on *normative* and *network* embeddedness in the gig economy by exploring its links with *embodied* and *structural* social reproduction. This relationship stems from the importance of labor in the economy and, in this case, the development of artificial intelligence. This labor is only possible because of the institutions and social networks that resist it in precarious working conditions and where power differentials are aggravated and working conditions keep degrading, demonstrating the unsustainability of this type of work. The article also explored how the interconnectedness of the economy, society, and the environment play a role in the reproduction of the workforce, as observed in the case of Venezuelan workers. The case of data work studied here provides only a glimpse of the intricate and complex social and economic context in which artificial intelligence systems are developed and distributed. In this context, the economic gains of a few market actors come at the expense of thousands of workers and their communities.

6. Acknowledgements

This research was funded by a grant of the International Development Research Centre of Canada and a fellowship of the Schwartz Reisman Institute for Technology and Society. I want to acknowledge the workers who shared their knowledge and experience with me so that these ideas could be developed. Special thanks to Erika Biddle, Alessandro Delfanti, Josie Greenhill, Milagros Miceli, Sarah Sharma, Christine Tran, Rianka Singh, Paola Tubaro, Antonio Casilli, and the anonymous reviewers for their valuable comments and suggestions.

References

- Agence France-Presse. (2021). Venezuela reports 2020 inflation of 3,000 percent. *ABS CBN News*. https://news.abs-cbn.com/business/02/12/21/venezuela-reports-2020inflation-of-3000-percent
- Anwar, M. A., & Graham, M. (2020). Hidden transcripts of the gig economy: labour agency and the new art of resistance among African gig workers. Environment and Planning A: Economy and Space, 52(7), 1269–1291.
- Althusser, L. (1970). Idéologie et appareils idéologiques d'État. (Notes pour une recherche). La Pensée.
- Beckert, J. (2003). Economic Sociology and Embeddedness: How Shall We Conceptualize Economic Action? Journal of Economic Issues, 37(3), 769–787.
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? Conference on Fairness, Accountability, and Transparency (FAccT '21).
- Bhattacharya, T. (Ed.). (2017). Social Reproduction Theory. Remapping Class, Recentering Oppression. Pluto Press.
- Bishop, S. (2019). Managing visibility on YouTube through algorithmic gossip. New Media & Society, 21(11−12), 2589–2606.
- Boris, E. (1994). Home to Work: Motherhood and the Politics of Industrial Homework in the United States. Cambridge University Press.
- Bourdieu, P. (1979). La Distinction. Critique sociale du jugement. Les éditions de minuit.
- Bourdieu, P., & Passeron, J.-C. (1970). La Réproduction. Éléments pour une téorie du système d'enseignement. Les éditions de minuit.
- Buolamwini, J., & Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. Proceedings of the 1st Conference on Fairness, Accountability and Transparency, 81, 77–91.
- Burawoy, M. (1982). Manufacturing Consent: Changes in the Labor Process Under Monopoly Capitalism. University of Chicago Press.
- Burawoy, M. (2010). From Polanyi to Pollyanna: The False Optimism of Global Labor Studies. Global Labour Journal, 1(2).
- Burawoy, M. (2013). Marxism After Polanyi. In M. Williams & V. Satgar (Eds.), Marxisms in the 21st Century (pp. 34–52). Wits University Press.
- Casilli, A. A., & Posada, J. (2019). The Platformisation of Labor and Society. In M. Graham & W. H. Dutton (Eds.), Society and the Internet (Vol. 2). Oxford University Press.

- Casilli, A. A., Tubaro, P., Le Ludec, C., Coville, M., Besenval, M., Mouhtare, T., & Wahal, E. (2019). Le Micro-Travail en France. Derrière l'automatisation de nouvelles précarités au travail ? Projet DiPLab.
- Cotter, K. (2018). Playing the visibility game: How digital influencers and algorithms negotiate influence on Instagram. New Media & Society, 21(4), 895–913.
- Dalla Costa, M., & James, S. (1972). The Power of Women and the Subversion of Community. Falling Wall Press
- Dauvergne, P. (2020). AI in the Wild. Sustainability in the Age of Artificial Intelligence. MIT Press.
- Davis, A. (1981). Women, Race and Class. Vintage Books
- Difallah, D., Filatova, E., & Ipeirotis, P. (2018). Demographics and dynamics of Mechanical Turk workers. WSDM 2018 - Proceedings of the 11th ACM International Conference on Web Search and Data Mining, 2018-Febua(August 2017), 135–143.
- Delfanti, A. (2021). The Warehouse. Workers and Robots at Amazon. Pluto Press.
- Drahokoupil, J., & Piasna, A. (2019). Work in the platform economy: Deliveroo riders in Belgium and the SMart arrangement. http://www.bollettinoadapt.it/wp-content/uploads/2019/01/WP-2019-01-deliveroo-WEB.pdf
- Dubal, V. B. (2020). The Time Politics of Home-Based Digital Piecework. *Ethics in Context*, 56.
- Federici, S. (1975). Wages Against Housework. Falling Wall Press.
- Federici, S. (2004). Caliban and the Witch: Women, the Body and Primitive Accumulation. Autonomedia.
- Federici, S. (2020). Revolution at Point Zero: Housework, Reproduction, and Feminist Struggle (2nd ed.). Between The Lines.
- Ferguson, S. (2020). Women and Work: Feminism, Labour, and Social Reproduction. Pluto Press.
- Fortunati, L. (1995). The Arcane of Reproduction. Housework, Prostitution, Labour and Capital (J. Fleming (ed.)). Autonomedia.
- Gandini, A. (2018). Labour process theory and the gig economy. *Human Relations*, 71(9), 1–18.
- Gebru, T. (2020). Race and Gender. In M. D. Dubber, F. Pasquale, & S. Das (Eds.), Oxford Handbook on AI Ethics. Oxford University Press.
- Glenn, E. N. (2012). Forced to Care: Coercion and Caregiving in America. Harvard University Press.
- 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.
- Granovetter, M. (1986). Economic Action and Social Structure: The Problem of Embeddedness. American Journal of Sociology, 91(3), 481–510.
- Granovetter, M. (2005). The Impact of Social Structure on Economic Outcomes. The Journal of Economic Perspectives, 19(1), 33–50.
- Gray, M. L., & Suri, S. (2019). Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass. Houghton Mifflin Harcourt.
- Harvey, D. (2014). Seventeen Contradictions and the End of Capitalism. Oxford University Press.
- Hester, H., & Srnicek, N. (2017). The Crisis of Social Reproduction and the End of Work. In The Age of Perplexity: Rethinking the World We Knew. Penguin Random House Grupo Editorial. https://www.bbvaopenmind.com/en/articles/the-crisis-of-socialreproduction-and-the-end-of-work/
- ILO. (2020). World Employment And Social Outlook. International Labour Organization.
- ILO. (2011). Policies and regulations to combat precarious employment. International Labour Organization.
- Irani, L. C., & Silberman, M. S. (2013). Turkopticon. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 611–620.

- Jarrett, K. (2014). The relevance of "women's work": Social reproduction and immaterial labor in digital media. *Television and New Media*, 15(1), 14–29.
- Kalleberg, A. L. (2009). Precarious work, insecure workers: Employment relations in transition. American Sociological Review, 74(1), 1–22.
- Krippner, G. R. (2002). The elusive market: Embeddedness and the paradigm of economic sociology. *Theory and Society*, 30(6), 775–810.
- Lehdonvirta, V., Kässi, O., Hjorth, I., Barnard, H., & Graham, M. (2019). The Global Platform Economy: A New Offshoring Institution Enabling Emerging-Economy Microproviders. *Journal of Management*, 45(2), 567–599.
- Ludec, C. Le, Tubaro, P., & Casilli, A. A. (2019). How many people microwork in France? Estimating the size of a new labor force. http://arxiv.org/abs/1901.03889
- Marx, K. (1978). *Capital: A Critique of Political Economy*. Volume I (F. Engels (Ed.)). Progress Publishers.
- Peck, J. (2013). Disembedding Polanyi: Exploring Polanyian economic geographies. Environment and Planning A, 45(7), 1536–1544.
- Polanyi, K. (1945). Universal capitalism or regional planning? London Quarterly of World Affairs, 10(3), 86–91.
- Polanyi, K. (2001). The Great Transformation: The Political and Economic Origins of Our Time (2nd Ed.). Beacon Press.
- Qadri, R. (2020). Algorithmized but not Atomized? How Digital Platforms Engender New Forms of Worker Solidarity in Jakarta. Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society, 144–144.
- Richards, L. (2015). Handling Qualitative Data. A Practical Guide. Sage Publications.
- Russell, S. J., & Norvig, P. (2020). Artificial Intelligence: A Modern Approach (4th Ed.). Pearson.
- Schmidt, F. A. (2019). Crowdsourced Production of AI Training Data: How Human Workers Teach Self-Driving Cars How to See (Vol. 155, Issue August). Hans-Böckler-Stiftung.
- Singer, F. (2021). Un kilo de arroz por un salario mínimo: la hiperinflación se dispara de nuevo en Venezuela. El País. https://elpais.com/internacional/2020-12-02/un-kilo-de-arrozpor-un-salario-minimo-la-hiperinflacion-se-dispara-de-nuevo-en-venezuela.html
- Tronti, M. (1962). Factory and Society. Operaismo in English.
- 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).
- Tubaro, P. (2021). Disembedded or Deeply Embedded? A Multi-Level Network Analysis of Online Labour Platforms. Sociology.
- Tubaro, P., & Casilli, A. A. (2022). Who bears the burden of a pandemic ? Covid-19 and transfer of risk on digital platform workers. *American Behavioral Scientist*.
- Wood, A. J., Graham, M., Lehdonvirta, V., & Hjorth, I. (2019). Networked but Commodified: The (Dis)Embeddedness of Digital Labour in the Gig Economy. Sociology.
- Wood, A. J., Graham, M., Lehdonvirta, V., & Hjorth, I. (2018). Good Gig, Bad Big: Autonomy and Algorithmic Control in the Global Gig Economy. Work, Employment and Society, 00(0), 1–20.
- Woodcock, J. (2020). The algorithmic panopticon at Deliveroo: Measurement, precarity, and the illusion of control. *Ephemera. Theory & Politics in Organization*.
- Woodcock, J., & Graham, M. (2020). The Gig Economy: A Critical Introduction. Polity Press.
- Yin, M., Gray, M. L., Suri, S., & Vaughan, J. W. (2016). The Communication Network Within the Crowd. Proceedings of the 25th International Conference on World Wide Web, 1293–1303.