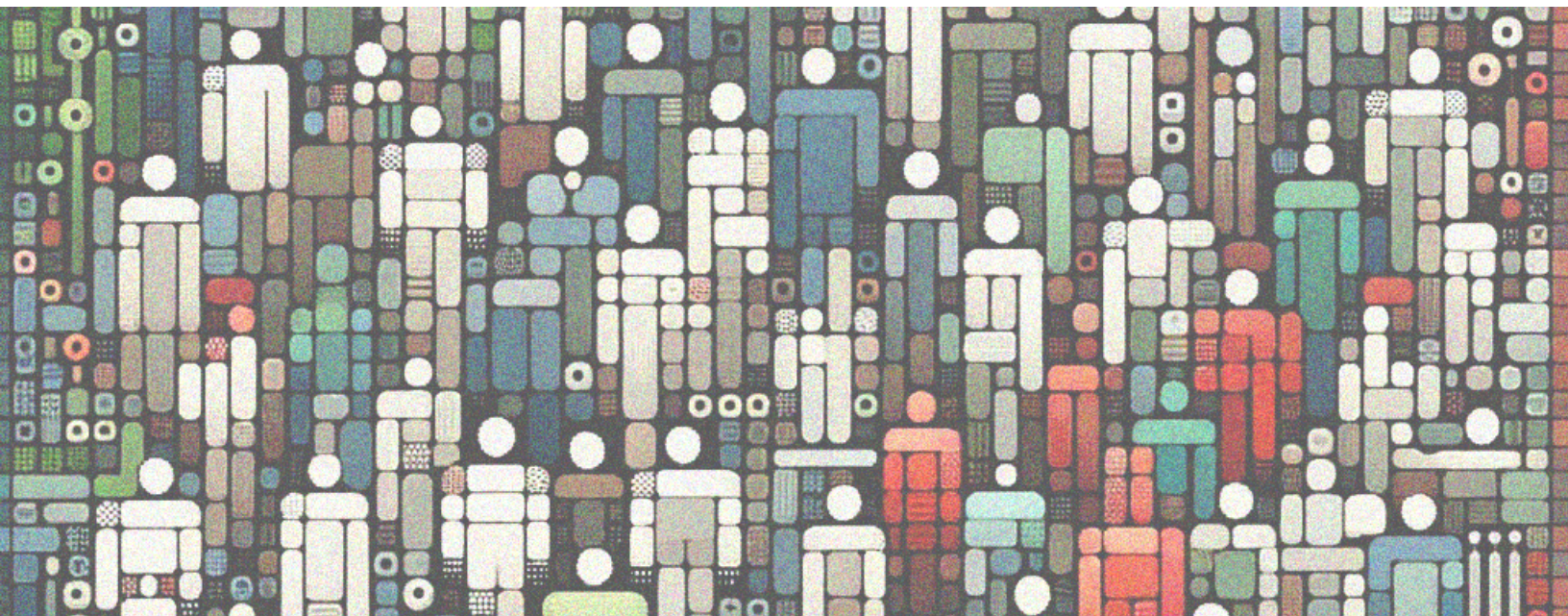# WHO TRAINS THE DATA
FOR EUROPEAN
ARTIFICIAL INTELLIGENCE?

ENCORE European Microworkers Communication and Outreach

Who Trains the Data for European Artificial Intelligence? Report of the European Microworkers Communication and Outreach Initiative (EnCOre, 2023-2024)

Milagros Miceli, Paola Tubaro, Antonio A. Casilli,
Thomas Le Bonniec, Camilla Salim Wagner, Laurenz Sachenbacher

April 16, 2024

THE LEFT

Contents

Executive Summary

Microworkers: The Forgotten Platform Workers

On February 22, 2024, a panel of microtask data workers ("microworkers" for short), assembled by the EnCOre initiative, addressed grievances at the Fourth Transnational Forum on Alternatives to Uberization, held at the European Parliament. This event convened platform workers' unions, advocacy organizations, and policymakers to advocate for the adoption of the Platform Workers Directive. Attendees at the Forum expressed unequivocal solidarity with the microworkers' representatives, leading to the integration of their specific demands into subsequent public outreach efforts and documentation.

Microwork refers to low-paid occupations, such as recording one's voice, translating short bits of text, and classifying contents in an image or webpage. These tasks are essential functions in the development of machine learning and AI because they prepare and correct algorithmic solutions. The Covid-19 pandemic, marked by lockdowns and job losses, along with technical phenomena like the advent of generative AI, has led to a significant rise in microwork across Europe.

Microwork in Europe: What We Know So Far

Based on our previous survey of more than 4,000 small and medium-sized businesses in two European countries, up to three out of four companies are aware of these platforms. Microwork platforms are increasingly utilized by companies in Europe to recruit, without committing to long-term contracts, a contingent workforce specializing in data-related tasks. Estimates of the global population of microworkers vary, ranging from overinflated platform reports claiming tens of millions of workers each, to more conservative but still alarming estimates by the World Bank, which suggest that this phenomenon accounts for between 4.4% and 12.5% of the global labor force.

Estimating the growth of microwork in Europe is challenging due to high turnover, evolving corporate practices, and trade secrets. Recent studies show a general increase in this type of platform work. Although larger EU countries have more microworkers in absolute terms, smaller countries exhibit a higher density. Additionally, significant offshoring to regions like Africa, South-East Asia, and South America may have shifted demand dynamics for European microworkers.

Recent data shows that while the majority of European businesses are aware of microwork platforms and data work deep labor networks, only a minority, mainly small businesses, admit to using them. Small businesses engage in microwork indirectly through purchased digital or AI services, with direct microwork clients being primarily technology suppliers.

The EnCOre Initiative

The EnCOre project, has been engaging platform users, workers of business process outsourcing (BPOs) firms, and technology companies through interviews, focus groups, and surveys in nine European countries: Austria, Belgium, France, Germany, Ireland, Italy, Luxembourg, Portugal, and Spain. Between March 2023 an April 2024, we have conducted exploratory interviews (N=7), focus groups (N=18), and fielded online questionnaires (N=577).

Results

Our survey reveals distinct demographic trends among microworkers, showing a significant over-representation of younger people with high education levels. The median age of participants is 27, with an average age of 30.87 years, indicating a younger demographic compared to previous international surveys. The significance of this trend can be seen in the fact that 37% of respondents possess a bachelor's degree and 27% hold a master's degree or equivalent. A high level of education among young microworkers corresponds with the recent increase in the percentage of 25-34 year-olds obtaining higher education degrees throughout the OECD. Among our respondents, women account for 41% of microworkers. In Europe, there is a relatively smaller gender gap

when it comes to the participation of women in microwork.

Microwork and Unemployment

Portuguese, Italian, and Spanish microworkers make up the largest percentages of participants among the nine targeted countries, with respectively 33%, 20%, and 14%. Notably, all of these countries have higher unemployment rates than the average for the Euro area. While 34% of respondents hold a primary, permanent salaried job, 28% are self-employed or in temporary employment. About 25% of those who engage in microwork do so because they lack other paid work. Of these, nearly 15% are unemployed and actively looking for work—double the 6.4% unemployment rate in the Euro area.

It is estimated that less than 20% of microworkers depend on platforms as their primary source of income. However, this small but active group utilizes an average of 2.46 platforms, demonstrating the critical role these platforms play in their financial ecosystems. Despite microwork being a crucial revenue stream, also for those who use it as a secondary source, 75% earn less than 50 euros per month, with median earnings ranging from 20 to 25 euros. Moreover, interviews and focus groups revealed widespread dissatisfaction among workers regarding wage theft, with payments not only being irregular and unsatisfactory but also being withheld arbitrarily.

Microwork and Migrations

The results of our investigation demonstrate the significant participation of migrants in the European microwork market. 21% of the respondents were born outside of their country of residence, and 18% hold a nationality other than that of their country of residence. There is potential evidence to indicate that migrants are engaging in online labor platforms at higher rates than previously recognized. This is a departure from earlier findings that predominantly described migrants in Europe as mainly involved in location-based platforms such as transport and deliveries rather than online freelancing and microwork.

59% of the respondents rely on a single platform for their income, suggesting a form of economic dependency, while 41% spread their efforts across an average of 3.42 platforms, indicating a strategy to diversify their income streams. It is noteworthy that almost 90% of these workers also engage in other online earning activities, internet sales, cashback operations, and participation in panels, which do not require specialized skills, and can be conducted on a temporary basis.

The above scenario contrasts with the concern expressed during focus groups that many migrants have high educational levels but are unable to find jobs or earn good wages in their new countries. Many are confined to low-paying microwork despite their qualifications. The disconnect raises important questions about the integration and utilization of competencies among migrant workers in the digital economy in Europe, where they still seem unable to find high-value online freelancing jobs requiring computer programming, design, and project management expertise. Moreover, fewer microworkers are involved in paid content creation and influence, which require significant time investments or specialized training, emphasizing the difficulty they face in leveraging their educational background into more lucrative online employment opportunities.

The Human Cost of AI

The EnCOre study reveals that despite the trend towards market polarization where high-value AI startups are located in the Global North and data workers in the South, there is strong development of a microwork ecosystem in Europe. In accordance with our framework, which distinguishes AI preparation, verification, and impersonation microtasks, we provide evidence that European microworkers are heavily engaged in AI-related tasks. During AI preparation, workers generate or annotate data for machine learning applications, such as recording voice prompts or labeling images. During AI verification, workers are required to assess the accuracy of AI-generated outputs, including audios and transcriptions. AI impersonation tasks, where human workers simulate artificial processes, are rare and typically performed internally. In these companies, workers are bound by non-disclosure agreements (NDAs), a practice that is uncommon among respondents to our survey recruited through online platforms, but ubiquitous among workers in BPO companies. The use of these agreements restricts the freedom of expression of workers, a situation made worse by the continuous monitoring of their online activities for the purpose of safeguarding proprietary information and assessing their performance.

The psychological and emotional effects of these microtasks are profound on workers. They are often referred to as merely "data tasks" but entail significant emotional engagement and pose significant socio-psychological risks. This phenomenon is particularly evident among microworkers involved in content moderation, who frequently report adverse effects on their mental health. This type of work has the potential to result in severe consequences due to its intense

nature. Among them are incidents in which workers have fainted, suffered from burnout, experienced psychotic episodes, and, tragically, in at least one instance, committed suicide. Moreover, the isolation experienced by these workers is further exacerbated because of remote work or, in the case of shared workspaces, a general distrust among colleagues and corporate management. The combination of high-stress tasks and social isolation presents a challenging environment for microworkers, highlighting the need for regulatory and support mechanisms in order to mitigate these impacts.

Policy Recommendations

As policy-oriented initiatives predominantly focus on location-based platforms, microworkers are not adequately represented. Owing to the dispersed nature of their activity, they face psychological rather than physical challenges owing to the social isolation arising from the dispersed nature of their activity, and in the case of content moderators, from exposure to all sorts of online toxicity. They are also more vulnerable to disclosures of their personal data and privacy violations. Low remunerations, limited direct communication with clients and platforms, and difficulty to challenge payment withholdings are common complaints, but the pervasive misconception that microworkers choose this type of work voluntarily discourages efforts to improve working conditions. This report highlights that microworkers are indeed platform workers, and stresses that their distinct needs must be acknowledged in discussions and legislation pertaining to platform work.

Chapter 1

Background

1.1 Microwork, data work, and platforms

In the wake of the Covid-19 pandemic, there has been a substantial increase in *microtask data work ("microwork" for the purpose of this report)* across Europe, fueled by lockdowns and ensuing job losses, leading to heightened competition and a decline in wages [TC0]. Microtasks include, for example, recording one's voice reading aloud a sentence, sorting files, watching videos, translating short bits of text, classifying contents in an image or webpage. These are essential functions in the development of machine learning and AI, from data generation and enrichment to quality controls of automated outputs. Therefore, microwork refers to low-paid occupations that (despite their varying nature) always involve annotation, tagging, and labeling of data, as well as debugging, correcting, and evaluating algorithmic and AI solutions. Because it is performed remotely and can be allocated to providers worldwide, this type of work differs from location-based platform 'gigs' such as delivery and ride-hailing. It also differs from online-only jobs for freelancers, for example in computer programming and design, insofar as its extreme segmentation and standardization allow dispersing tasks to an undefined crowd instead of a selected individual (whence the alternative denomination of 'crowdwork').

1.2 Microwork: a growing phenomenon in Europe and worldwide

Recent data collected by DiPLab in 2021 indicate [BELT21] that the majority of European businesses are aware of microwork platforms and data work deep labor networks. A total of 3,323 small and medium-sized businesses responded to the questionnaire in France, and 786 in Germany. Three out of four German companies and three out of five French companies are aware of these platforms, but few use them (one out of ten in Germany, and under one in five in France). Small businesses more often see potential for their use in the future than larger businesses. German IT companies are more likely to use microwork platforms (french data did not provide this breakdown). Several respondents mentioned legislative uncertainty, as well as opposition to platform services in principle. It is difficult to integrate platform services into the company's organization (fig. 1.1).

Microwork platforms are increasingly being utilized by companies in both France and Germany to recruit a contingent workforce who specializes in data-related tasks, allowing businesses to avoid committing to long-term contracts. However, the survey only measured how companies use microwork *directly*, by signing up to a platform or hiring an outsourcing firm on their own. Many companies use microwork *indirectly*, via suppliers and service providers that either do not advertise or don't account for the fact that a part of their service is provided by workers. In particular, small businesses purchase digital services and ready-to-use AI solutions and do not directly manage microworkers.

According to the latest market developments, local/regional platforms are emerging along with companies that offer services as part of deep labor chains, with microtasks amongst other components. New estimates, based partly on the methodology developed by DiPLab [TLLC20], reach several hundred millions.

In 2021, Kässi et al. [KLS21] estimated the number of online platform workers worldwide, including various remote work forms like microtasks and freelancing. Their analysis, extending
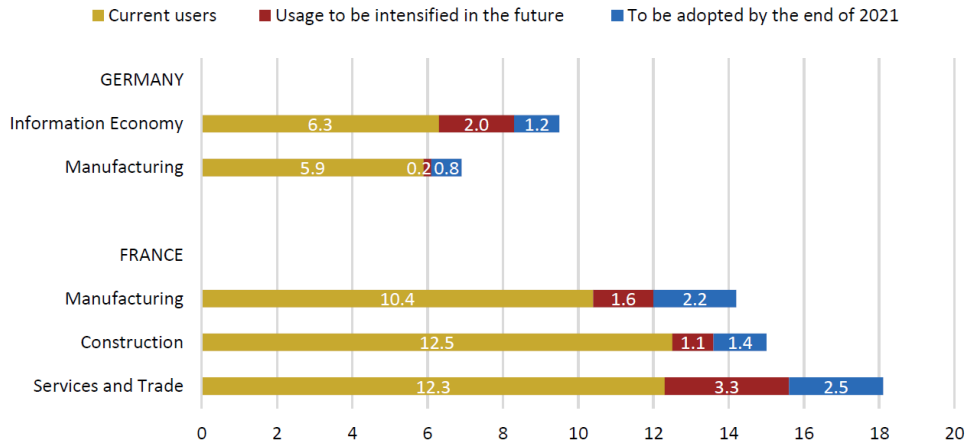
Figure 1.1: Use and future adoption of microworking by companies in France and Germany (percentage by sector). Reading guide: 8.3% of companies in the German information economy use microwork, 2% of which will intensify its usage in the future. Source: [BELT21]

beyond large platforms, revealed around 163 million registered users globally, with approximately 10% (16 million) engaged in microwork. A more recent estimate by the World Bank suggests a broader range, estimating between 154 million and 435 million online on-demand workers (freelancers and microworkers). This variation arises from including individuals who have practiced such activities, even sporadically. The lower estimate is the number of online workers, mainly freelancers, who use platforms for their primary source of income. Microworkers are more likely to be found among the rest, who use online labor platforms as a secondary or marginal source of income. The estimates show that online workers, both freelancers and microworkers, make up between 4.4 % and 12.5 % of the global labor force. [DCS+23].

These numbers may seem overinflated, but they read as conservative when compared with the platforms' own user statistics. In the mid-2010s, platform self-reported figures were already staggering: Freelancer, ZBJ, and Upwork alone claimed to have 51 million users, plus 41 million workers counted by Taskcn, Crowdsource, Microworkers, Epweike, Appen, ClickWorker, and Fiverr. Pandemic shock and lockdowns have led to a greater influx of new online workers reported by platforms. We will examine in the following sections the increasingly visible link between the growth of microwork and its reliance on both internal and external migration flows. Several of these movements are the result of geopolitical instability in Europe and the Middle-East, and the climate crisis, which has a particularly adverse effect on rural regions in the southern hemisphere.

Despite the uncertainties inherent in such estimates due to evolving practices, high turnover, and trade secrets, recent studies indicate a growing trend in online platform work, including microworkers. Assessing the extent of this growth in Europe lacks a precise estimate, but a European comparative study provides qualitative insights. In 2023, Morgan et al. [MvZtH23] used an international microtask platform user survey to estimate the density of microworkers aged 18 to 99 in each European country. The study suggests that while larger EU countries have more microworkers in absolute terms, this score does not correlate with the prevalence within their population. Smaller countries exhibit higher density, with Portugal, Croatia, Latvia, Germany, and Bulgaria ranking as the top five by microwork density. Our research findings align with these results, particularly for the countries that overlap in our study samples. It is noteworthy that Portugal has a large population of microworkers, which indicates a significant trend in this region.

These findings hint at potentially less pronounced growth in the number of microworkers in France, Germany and other more populous European countries. These results align qualitatively with observations from specific platforms and may reflect the impact of trends such as outsourcing microtasks to small businesses and complex deep labor configurations not involving the surveyed microwork platforms. Additionally, significant offshoring, particularly to Africa, South-East Asia, and South America, might have changed demand dynamics for European microworkers. In terms of demographics and economics, offshoring represents a complement to immigration, which is also an important characteristic of microwork in Europe currently. Although opposite trends, both incoming and outgoing migrations provide, for European businesses and microtask platforms, underrepresented, underpaid, and often quasi-captive labor, either on-site or remotely (for a detailed analysis of these developments, see chp. 2).

1.3 European AI production and microwork supply chains

Machine learning (ML) relies heavily on data outsourced to microwork platforms and smaller formal or informal companies. Previous studies conducted by EnCOre authors revealed outsourced BPOs and platforms in Latin America, specifically in Venezuela and in Argentina [MP22] as well as in Africa, specifically in Madagascar [LCC23]. Precarious conditions and economic dependency are rampant among these workers. Our previous findings highlight how workers' agency is restricted. The size of this workforce has increased sharply in recent years, but social progress is still lagging.

The different microtasks that are accessible to data workers can be analyzed using the framework that EnCOre authors initially designed to examine labor on the pure marketplace digital platform [TCC20], and subsequently adapted to other types of small and medium-sized businesses and BPO firms [MP22] in broader deep labor configurations (fig. 1.2).

In AI systems' production, microworkers play three important roles: "preparation," "verification," and "impersonation." As part of the framework, *AI preparation* is divided into data generation (e.g. workers are instructed to produce media content like text, images, audio, and video with their devices) as well as to data annotation (such as categorizing and labeling images of people based on gender, race, or age). *AI verification* involves the evaluation of algorithmic outputs by workers (e.g. assessing how well a search engine has responded to a user query). Finally, *AI impersonation* is the use of non-disclosed human live intervention in place of the AI solution (for instance, workers are required to dialogue with users and respond to their queries in the same manner as a chatbot). Occasionally, this is utilized to supervise or debug defective solutions in real time. Fraudulent businesses may also use this to make workers harder to distinguish from non-existent algorithms.
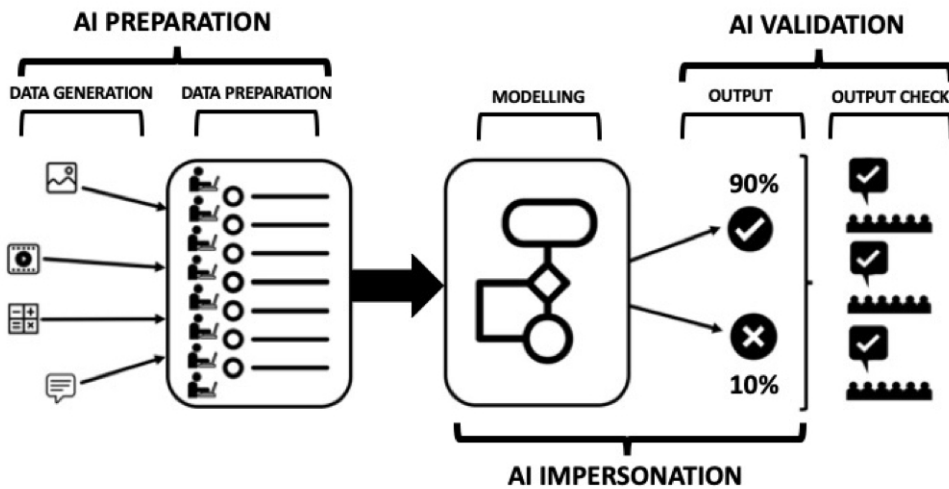
Figure 1.2: The three main functions of microwork in the development of data-intensive, machine-learning based AI solutions. Source: [TCC20].

In Europe, alongside various platforms, many companies already engaged in IT service outsourcing and offshoring have incorporated microtasks into their service portfolios. The development of AI solutions, has initiated the creation of "global value chains," leading to the recruitment of microworkers abroad. These chains often extend to countries with lower labor costs, where linguistic or cultural similarities facilitate communication and interaction with clients. Efforts to trace value chains from European AI producers have frequently led to low-income countries, often former colonies highly active in outsourcing services for former metropolises. Several formal and informal companies, either possessing platforms or using SaaS platform solutions to perform tasks in the Global South, specialize primarily in data annotation for AI. Similar to microwork in Europe, the one carried out in African, Asian, or Latin American countries remains largely invisible, charac-

terized by precarious conditions, informality, and relatively low pay compared to the cost of living, especially in sprawling megacities.

These contextual changes have two major implications. Firstly, European microworkers face intense competition, notably from foreign counterparts residing in French- or Spanish-speaking countries where wages are lower. Local European platforms do not seem to have expanded their workforce over time, and sometimes require residence in Europe or European citizenship. Thus, it becomes evident that studying microwork in Europe, cannot be confined to Europe alone. It must extend to foreign countries, especially in Africa and South America, where an increasing portion of microtasks essential to the European industry is outsourced.

1.4   Why we talk about "deep labor"

Based on these initial remarks, platform labor in Europe must be addressed in a way that encompasses more than simple marketplace platforms that act as intermediaries between clients and microtask data workers, but also the integration of digital platforms with supply chains, which include a wide range of companies, BPOs, and informal organizations.

The intersection of microwork with AI supply chains (both at local and global level) over the past decade, has created a variety of complex organizational models: intricate networks of platforms, subcontractors, and service providers that we refer to by the concept of "deep labor." (fig. 1.3)

Digital platforms serve as intermediaries between labor supply and demand. They are characterized as multi-sided markets, where two or more categories of users (e.g. couriers, customers, and restaurant owners) are algorithmically matched. As an example, Google's search engine provides information to users and advertising space to advertisers. Microwork platforms, in general, tend to have a two-sided structure, connecting clients and microworkers [CP19]. Platforms that act as intermediaries usually offer off-the-shelf solutions to companies, but allow limited customization and do not ensure quality control over the data produced.

This is why a new paradigm involving global subcontracting networks has emerged. In deep labor, several levels of platforms are arranged together. Rather than allowing clients to interact directly with microworkers in a single space, deep labor platforms provide turnkey solutions to their clients, and then recruit the necessary workers through other platforms, temp agencies, or small and medium-sized businesses, usually in countries where the labor cost is lower. When microworkers respond to an ad on a job listing app, they are directed to a site to complete the hiring process, then move to a platform to work at a piece- or hourly-rate (and sometimes even on a fixed-term contract), and then to another platform for payment.

At the international level, the trend towards complexity is more pronounced, with significant outsourcing of microtasks to lower-income countries to cut costs. Major global deep labor platforms include Telus International, Appen, and Pactera, each adopting a comprehensive approach to offer taylor-made solutions to clients. These platforms recruit the necessary workforce through various channels, redirecting microworkers to different platforms for task completion.

The fragmentation of workforce management processes is influenced by several factors, including the need for specialized labor, data confidentiality concerns, working with multinational corporation subsidiaries, and the replication of corporate hierarchies for control. This organizational complexity is not limited to one model but spans a spectrum, from pure-player platforms to small business process outsourcing (BPO) firms in hybrid deep labor chains.

Irrespective of the level of platformization, the number of subcontracting layers, outsourcing location, or status, the primary impact on workers' rights is their exclusion from the formal workforce of the client. This multiplication of intermediaries in digital labor platform value chains contributes to labor market fragmentation and reconfigures inequalities. Workers are often classified differently based on status, skills, location, and, in some cases, gender, leading to disparities in pay, access to social protection, and career opportunities [CTLL+19].

1.5   Regulation and self-regulation in European microwork

The concept of "microwork," inspired by "microcredit," initially raised expectations of social inclusion through platforms. However, the reality exposes precarious conditions and low wages, despite attempts at self-regulation through ethical codes in the tech sector. These efforts are limited in effectiveness and cover only a small number of platforms. Microworking, however, failed to live up
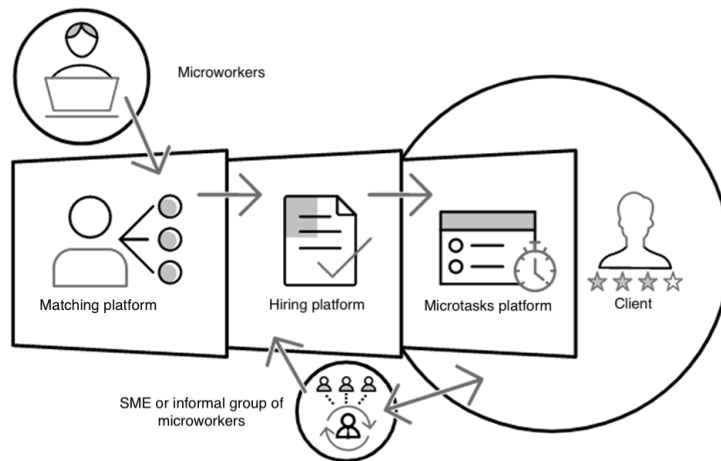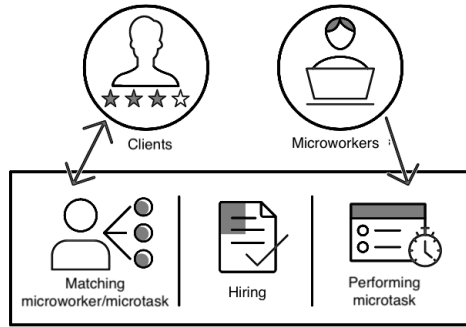
Figure 1.3: (Top) A microworking platform that is an intermediary type; (bottom) a microworking platform that is a "deep labor" type. Source: [CTLL+19].

to these promises. The dangers of precariousness and low remuneration have been highlighted in numerous studies.

Platforms have become increasingly conscious of the negative perception surrounding their operations and, over the past few years, have implemented various measures aimed at self-regulation. A special mention should be made of:

- The Australian platform Appen, which boasts over a million microworkers in 170 countries, has adopted an ethical code that commits to paying at least the minimum wage. It emphasizes inclusion, respect for privacy, and freedom of expression among workers. Its webpage, however, does not provide much information about how these principles are implemented.

- A consortium of German platforms, including Clickworker, adhered to a more precise and detailed code of ethics in 10 points: conformity with the law, legal assistance if needed, minimum salary, motivational aid (training, recognition results), respect in interaction, clarity of tasks, freedom to refuse tasks, constructive feedback, the ability to redo tasks if errors occur, protection of personal information.

- The Partnership for AI, an NGO based in the United States, developed five principles that AI producers must follow when managing their data supply chains responsibly. These principles have been incorporated into the internal operations of a large company, Deepmind, according to its website.

Despite their merits, these first efforts fall short of what is needed. These are relatively light forms of self-regulation, whose effectiveness is not systematically assessed. In addition, they only apply to a very small number of microworkers, and mainly on intermediary-like platforms.

This topic is rarely addressed in regulations, so public decision-makers need to be alerted to this issue. The European AI Act and the Directive for Platform Workers, both adopted in 2024,

do not hold producers, importers, and professionals accountable for the social conditions in which their data is labelled, annotated, filtered, enriched to implement machine learning solutions.

There are two significant legal rulings involving local platforms that demonstrate advancements in this area. Issued in 2022, these decisions, however, reflect contrasting approaches. The French microworking platform, Clic and Walk, which sported as many as 700,000 microworkers, was acquitted by the Criminal Chamber of the Cassation Court in Paris on charges of concealing employment. In contrast, the Brazilian company Ixia faced a different fate when the Regional Labor Court of the 2nd Region (São Paulo) mandated the reclassification of its users as employees.

The upcoming European Corporate Sustainability Due Diligence Directive (CSDDD), mandating parent companies to take necessary measures to prevent serious violations of fundamental social rights in their value chains, is a legislative tool that could be leveraged. The aim is to ensure that companies consider the negative impacts of their activities, including the social cost of poorly paid microwork, on human rights in their supply chain. The objective is to market AI solutions produced with human labor only when there has been a systematic effort, by the concerned companies, to identify and mitigate associated risks in the supply chain.

1.6 Report outline

In the following chapters, we examine microwork in Europe, through online surveys, interviews, and focus groups. Chapter 2 presents insights gained from an online survey exploring European microworkers' socioeconomic characteristics. With a particular focus on migrant workers, we investigate how they engage in microwork. Moreover, the chapter examines the types of tasks they perform and the reasons for their participation on such platforms.

In Chapter 3, the focus shifts to the labor conditions and concerns of these workers. Using findings from interviews and focus groups, this chapter discusses the nature of tasks, payment issues, job security, prevalent secrecy and surveillance practices, and the impact of these working conditions on workers' psychological well-being. There is also an emphasis placed on the impact of migration on these factors in this section.

Further, the report explores how these findings are related to broader labor struggles and the evolving landscape of platform work, suggesting potential future directions. It concludes by analyzing the implications of these findings for future policies, emphasizing the need for regulatory adjustments and support mechanisms.

In Annex A, we provide an overview of survey methodology, from design to data quality, assessing the robustness of our conclusions. This document serves as a resource for understanding the dynamics of microwork in Europe and shaping effective policy interventions.

Chapter 2

Who microworks in Europe, how, and why? Insights from an online survey

To provide a broad perspective on the characteristics of the European microworking population and its activities, we rely on data from an online survey fielded in Spring 2024. Data from questionnaires do not allow as much depth as the interviews and focus groups described in chapter 3, but they reach out to a broad range of respondents in order to have a better sense of the distribution and nature of the phenomenon under study.

2.1   A survey in nine European countries

We designed an online questionnaire covering a broad range of topics, including online activities and practices, access to computing devices and to the internet, practices and representations of online microwork, earnings from online microwork, educational attainment, language and computing skills, professional situation and experience, family and household, economic and financial situation, and socio-demographics. The questionnaire was rather long, as it took on average **20-25 minutes** to complete it. It was available in **English, French, German, Italian, Portuguese, and Spanish**. Allowing respondents to choose the language they know best is a way to reach out to a broader pool of respondents, possibly invisible to English-only surveys often conducted in the past. To take the maximum advantage from this methodological choice, we restricted participation to residents of the nine European countries in which one of these languages is spoken: **Austria, Belgium, France, Germany, Ireland, Italy, Luxembourg, Portugal, and Spain**.

The questionnaire was distributed in **March - April 2024** as a remunerated task through the online labour platform Prolific, and we obtained **577 complete responses**[1]. More methodological details are available in Annex A.

What follows presents the portrait of microworkers that this survey allows to draw. To put results in perspective, we also use (wherever helpful) data from two earlier online surveys which used a very similar questionnaire instrument and targeted:

- French users of the Paris-based microtasking platform Foule Factory, in May-July 2018, n = 909;

- Spanish users of the international microtasking platforms Microworkers.com (December 2020-February 2021) and Clickworker (March-April 2022), n = 447.

These two additional sources enable us to 'zoom in' on two large countries within our sample, while also taking into account cross-platform differences.

---

[1]The completion of the questionnaires corresponded to an hourly rate of €15.58 as of 9 April 2024. The selected fair compensation for participants is higher than the minimum hourly wages of all targeted countries and significantly exceeds the platform's recommended rate. Additionally, the rate is balanced to prevent potential bias in the results and undue competition among workers, resulting in an advantage for those with higher reputation scores and task counts.

2.2 Who are the European microworkers?

Let us start by drawing a socio-demographic portrait of this population.

Age. While the oldest respondent to our survey is 80 years old, **the median age is 27 and the average is 30.87 years**. By design, no one is below 18 years of age (usually the minimum required by platforms to register). Figure 2.1 illustrates the over-representation of younger people among platform microworkers, a fact that has already been noted in the literature. This tendency is particularly strong in our data: indeed, participants to the survey are somewhat younger than observed in data collected by the International Labor Organization (ILO) across five different microworking platforms and spanning the whole planet, where average age is 33.2 years [BFH+18]. The data that we collected in France (2018) and Spain (2020) are also aligned with ILO's average.
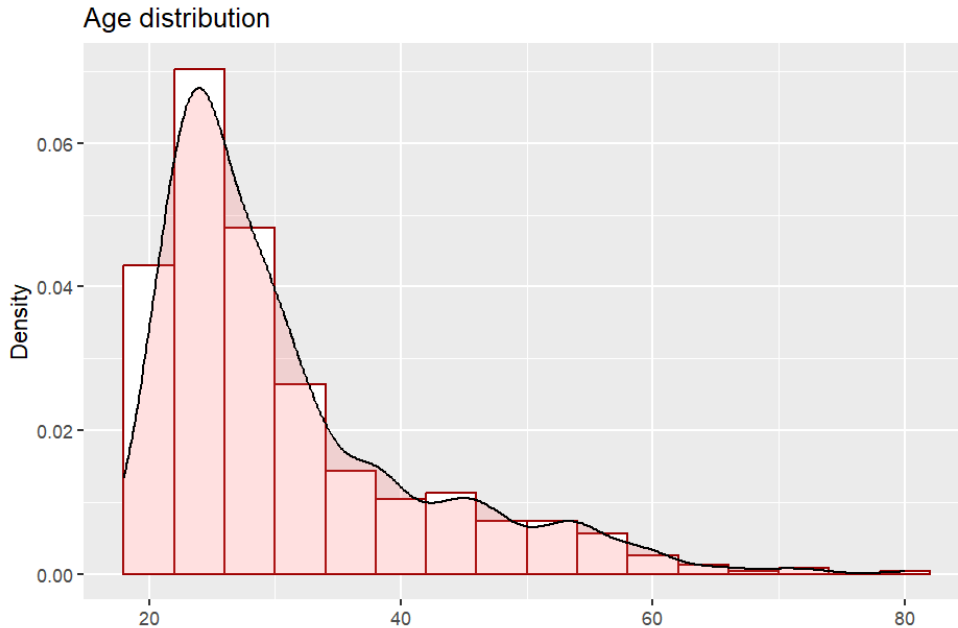

Figure 2.1: Distribution of microworkers' age: the histogram shows the counts of values in each age range, while the superposed density plot shows the proportion of values in each range. n = 576.

Gender. 41% of respondents are women[2]. As Table 2.1 shows, this ratio is about the same that we found in Spain in 2020, and is aligned with a 2017 European-Parliament study of microwork in Europe [FSJ+17]. ILO data count fewer women (one out of every three workers), a result likely driven by their more limited involvement in lower-income countries, where they constitute one out of five workers [BFH+18]). These proportions are lower than what we observed in France in 2018, where women constituted slightly over half of the microworking population [CTLL+19]. The overall impression is that women are less present than men, though the gap is less pronounced in higher-income countries in which access to labour markets is more equal.

Ethnicity. Collection of ethnicity data is complex and in many European countries, highly controversial. Our questionnaire does not include this dimension but Prolific collects users' self-classification into broad and simplified categories. 86% of those who participated in our study self-identify as White, while the others see themselves as Asian, Black, Mixed, or Other.

Education. Our data confirm that microworkers are highly educated, as already noted in previous research (for example [Int21, Tub22]). **Three quarters of them have done at least**

---

[2]As a whole, microwork platforms score poorly on inclusion and representation of trans-, queer-, and gender non-conforming people. Almost all existing studies categorize workers based on binary gender categories. According to a 2017 Amazon Mechanical Turk survey, transgender and gender queer people make up 0.6 percent and 0.9 percent of the user base, respectively [CCL+17]. According to EnCOre's questionnaire, respondents had the option to identify themselves as women, men, or others. Similarly, Prolific allows respondents to answer "prefer not to say". The number of respondents having chosen this option is both coherent with previously published data and too small (1%) to derive any meaningful conclusions.

% women	Platforms	Countries	N	Source	Year
41%	Prolific	AT, BE, DE, ES, FR, IE, IT, LU, PT	577	Authors	2024
40%	Microworkers, Clickworker	ES	447	Authors	2020-22
56%	Foulefactory	FR	909	Authors	2018
33%	Amazon Mechanical Turk, Clickworker, Crowdflower, Microworkers, Prolific	World	3,500	ILO	2018
39%	Amazon Mechanical Turk, Clickworker, CrowdFlower, Microworkers	Europe	1200	European Parliament	2017

Table 2.1: Percentage of women as observed in different studies of platform microwork covering European countries.

some University-level studies:

- 10% have a short (2-year or less) higher-education degree;

- 37% have a bachelor's degree or equivalent (3-4 years);

- 27% have a master's degree or equivalent (4-5 years);

- 3% have a doctorate.

Our previous data from Spain (2020-22) and France (2018) reveal a similar picture, though with some differences likely due to the specific educational system of each country. Of Spanish microworkers, 12% have a short tertiary degree, 46% have a bachelor's degree, 17% a master's, and 1% a doctorate; in France these percentages are, respectively, 23%, 20%, 21% and 1%. ILO data include the same proportion of workers with a bachelor's degree (37%) and 20% with a postgraduate degree [BFH+18].

> The highly-educated are clearly over-represented among microworkers, in comparison with the general population of their countries: the proportion of adults aged 25-64 with a higher education degree (short-cycle, bachelor's, master's or doctoral or equivalent) is only 38% on average in the EU25 area, according to OECD data [OEC23]. This difference may be partly driven by the high proportion of young people among microworkers, considering that the share of 25-34 year-old people with a higher education degree has increased in the past few years and across OECD countries, it is currently equal to 47.4%. Still, the proportions we observe are higher and suggest a self-selection of highly-educated people into this activity.

Of the microworkers who have a higher education degree, 51% are specialized in science and/or technology, a percentage that is very close to the 57% in ILO data [BFH+18]: specifically, 31% studied engineering, computer science, or other technology-related discipline; 10% mathematics or the natural sciences, such as physics or biology; 10% medicine or other health-related discipline such as pharmacy or nursing. The strong presence of workers with scientific and technological backgrounds is remarkable and suggests that platform activity may be partly driven by computer literacy and interest in keeping up with advances in the field. Another 22% of workers studied communication, journalism, languages, art or humanities; 14% social or political sciences or law; 13% economics, finance or business.

Students comprise 37% of respondents to our survey, and four fifths of them are currently preparing a bachelor's or a master's degree. This is a higher proportion than in our previous studies (27% in Spain 2020-22, and 10% in France 2018). The strong presence of students in this sample correlates with young age: the average age of participants who are students is 25.4 years, compared to 34 years for the others[3].

Professional status. In addition to online microwork, 34% of respondents have a main, permanent salaried job. Other professional statuses are more likely to include multi-activity. In particular, the 14% in temporary employment and the 14% self-employed often (though not always) combine both statuses. Likewise, slightly over 40% of the students (which overall, constitute

---

[3]The difference is statistically significant according to a Wilcoxon rank sum test, p-value ¡ 2.2e-16.

about a third of the whole sample as mentioned above) have another professional activity such as a fixed-term job, in addition to microwork. Note that almost 15% of respondents declare to be unemployed (and actively looking for a job), while numbers are very small in other categories (retired, inactive, member of a cooperative, unpaid worker in a family business, other informal worker).

> The survey findings reveal a notable overrepresentation of unemployed individuals among microworkers, with their percentage significantly exceeding the general unemployment rates within the Euro area. While the Euro area's seasonally-adjusted unemployment rate stood at 6.4% in November 2023 (source:[Eur24]), the proportion of unemployed microworkers was more than double this figure. Of note, the EnCOre observed unemployment rate is close to the EU youth unemployment rate, which concerns workers aged 25 years or less and was 14.9% in January 2024 (14.5% in the euro area according to Eurostat data) [Eur24]. This is not entirely driven by the young age of the respondents to this survey, only 41% of whom are under-25: in fact, the rate of unemployment is the same both for the under- and over-25. What we see is likely a manifestation of a broader **tendency for platforms to attract the unemployed, who rely on microwork** as a source of income. These findings support the conclusions of DiPLab's previous US labor market study [LVS$^+$23].

Of those who are in paid work, one third works part-time, and two thirds full-time. Taking the subgroup of those who have an employer (56% of the sample), three out of five have a permanent job contract, one in five has a temporary contract, 7% are in informal arrangements with no contract, and all others have different situations (apprenticeship, interim, paid or unpaid internship, etc.). Two thirds work for a company in the private sector, one in six works for a public-sector agency, or national or local government, and the others are employed by families or individuals, NGOs, or other types of organisations.

Country of residence. From among the 9 targeted countries, the largest numbers of participants are from **Portugal (33%), Italy (20%), Spain (14%), Germany (13%), and France (12%). Very few people participated from Austria (3%), Belgium (3%), Ireland (2%), and Luxembourg (less than 1%)**. This result confirms a 2023 insight of Morgan and co-authors [MvZtH23], according to whom some smaller countries like Portugal exhibit a high prevalence of microwork compared to the actual size of their population.

First language. While fluency in one of the six survey languages was a condition to participate in the survey, respondents may have any first language. The largest group speak Portuguese (33%), followed by Italian (18%), Spanish (14%), German (11%), English (10%), and French (8%). 7% of respondents have a different first language.

Country of birth and nationality. The number of countries in which respondents were born far exceeds the nine residence countries targeted in the study. They include both other EU member states such as the Czech Republic and Romania, and countries from Asia, Africa, the Americas, and to a lesser extent Oceania. **21% of respondents were born in a country different from the one where they currently live**. Similarly, 18% do not have the nationality of the country in which they live.

Together with first-language differences, these data point to the presence of migrants in our sample, and we discuss this aspect below in more detail.

2.2.1 Migrant microworkers

The difference between country of birth and country of current residence can be taken as a signal of a migratory experience. It is obviously not a perfect indicator, because one may be born abroad from two citizens (for example in the case of expatriate families), because it does not account for multi-country migration processes, and because it does not include descendants of migrants. It is a reasonable proxy, though, and a better one than the other variable we could leverage, current nationality, which may have changed at some point in time as a result of the migratory trajectory itself. Use of this proxy is also consistent with the principle adopted in French statistics, according to which it is the country of birth, and not the citizenship at birth that defines the geographical origin of a migrant person. This analytical choice is not intended to deny that measuring migration is a statistical challenge, especially as definitions vary from country to country and are deeply affected by political stances and decisions.

Interpreting the data in this perspective suggests that **in Europe migrants may be using online labour platforms much more than was previously believed**. If confirmed, this important finding of our research would both challenge and enrich previous scholarship. In a 2022 survey of European platform work more generally, Piasna and co-authors [PZD22] note that immigrants are especially likely to be engaged in on-location platform work (such as transport and delivery), but less likely to carry out online freelancing and microwork. Comparison with our own data suggests that things may be changing over time. While our 2020-22 survey of Spanish microworkers records a similarly high presence of immigrants, our earlier (2018) study of a French microwork platform found that they constituted only slightly less than 10% of its user base [CTLL+19].
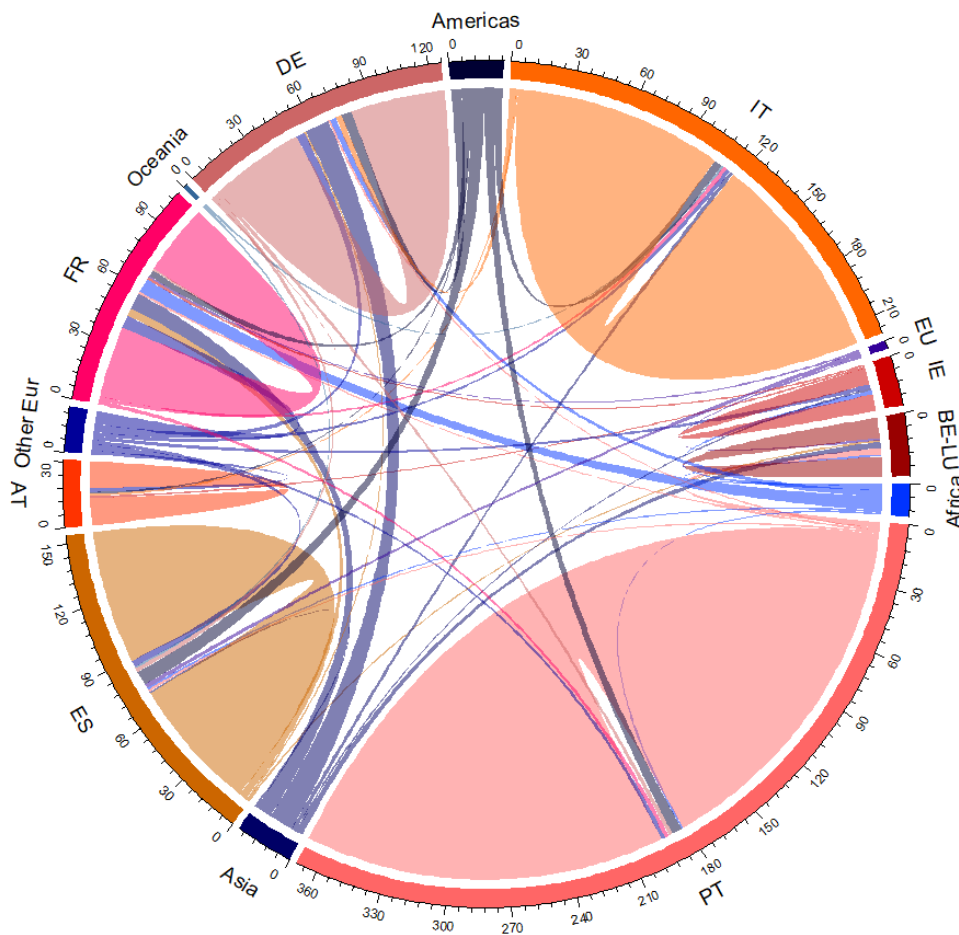

Figure 2.2: Chord diagram representing international migration among EnCOre respondents. Countries targeted by the survey (Austria, AT; Belgium, BE; France, FR; Germany, DE; Ireland, IE; Italy, IT; Luxembourg, LU; Portugal, PT; Spain, ES), and the flows originating from them, display orange-like colours. Continents migrants originate from (EU: European Union countries outside the targeted nine; Other Eur: countries located in Europe, but not EU members, for example United Kingdom; Africa, Asia, Oceania, and Americas) are coloured in nuances of blue. Each country or continent is represented by a fragment on the outer circle. Larger fragments represent larger worker populations, for example in the cases of Portugal (PT) and Italy (IT) where large numbers of respondents reside, while small fragments (for example Ireland (IE) and Belgium and Luxembourg (BE-LU), which have been grouped together here for better visibility). In the figure, arcs are drawn between locations, whose size is proportional to the importance of the migratory flow. n = 577.

Figure 2.2 shows that the involvement of migrants in microwork varies by country. It is particularly strong in France and, to a lesser extent, in Germany and Spain, while it is weak in Italy and almost nonexistent in Portugal. This analysis is an indication that in relatively lower-income, higher-unemployment Southern European countries, such as Italy and Portugal, online platform microwork is very widespread and mostly practised by locals, while in higher-income, Western European countries such as Germany and especially France, it is more often practised by immigrants.

We must put this initial finding in dialogue with the qualitative results that focus on BPOs and other types of organizations, where migrants are working on data microtasks (chp. 3). Taking only the survey data into account, platform-based microwork would seem to be largely (or almost entirely) performed by citizens in countries like Germany and Portugal. It is important to note, however, that this scenario does not take into account organizations that offer microtasks under different contractual arrangements, such as firms that specialize in the outsourcing of business processes.

2.3 What do microworkers do online, and why?

Let us now look at the activities that microworkers undertake online, and their motivations.

Engagement in microwork. The least active worker among our participants is a newcomer for whom the EnCOre survey was the first task ever, while the most experienced did 2,090 approved tasks before taking our questionnaire. These large variations largely depend on the amount of time already spent on the platform, as shown in figure 2.3: overall, it is common for those with a longer tenure on the platform to have performed more tasks. However, excluding the newcomers, there is also significant variation within each subgroup. Take for example the workers who have already been on the platform for 8-12 months: the majority of them have done a few dozen tasks, but a few have over 700. This is because workers vary significantly by their level of engagement on platforms: the very active coexist with both the regular but less frequent, and with the more occasional users.
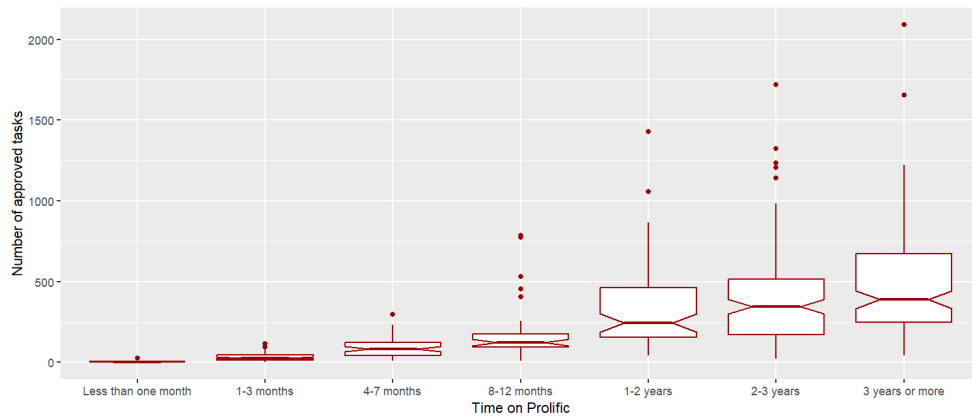

Figure 2.3: Microtasks workers have completed on Prolific, based on how long they have been signed up. The categories along the bottom (x-axis) show different lengths of time since workers joined. For each time group, there's a box that shows where most people's task numbers fall — the bottom of the box is where the lower quarter ends, and the top is where the higher quarter starts, with a special mark for the median value. As workers spend more time registered on the platform, they tend to complete more tasks, and there is also a greater difference in the number of tasks completed. The wider boxes and longer whiskers (i. e. the lines extending vertically from the boxes) in the categories representing longer times since registration indicate more variability in the number of tasks completed. The dots, representing outliers (i. e. values that are unusual compared to the rest), also become more numerous for those who have been on the platform longer, showing that there are more exceptional cases of individuals who have completed many more tasks than typical for their time on the platform. n = 576.

Multihoming. Economists refer to the condition of workers affiliating with more than one platform as multihoming. Here, we are interested in the extent to which respondents use only Prolific, or also one or more of its competitors. The data show that on average, they use (or used in the past) 1.99 online microtasking platforms, Prolific included. Among them, the 59% who have only used Prolific can be distinguished from the other 41% who have used 3.42 platforms on average (with a median of 2). These figures align with the average of 1.83 platforms per worker observed by the ILO [Int21] and by Kässi et al. 2021 [KLS21] and with the 1.27 of our previous French study [TLLC20].

If the small number of platforms used may seem surprising (wouldn't workers maximise earning opportunities by using several platforms simultaneously?), it is due to the costs that moving from

one platform to another involves: each has its own (usually non-portable) entry tests and reputation system, while payment arrangements also differ widely, involving transaction costs.

Among the workers who do use other platforms in addition to Prolific, the most common are Clickworker (16%), Fiverr (16%), Appen (15%), UHRS (12%), and Amazon Mechanical Turk (11%). The sub-sample of workers who use them is important for two reasons. First, it can give us a broader view of the market beyond the platform Prolific. Second, users active on more platforms likely have greater engagement and/or greater experience in this activity, and may qualitatively differ from the others.

Workers are much more numerous (almost 90%) to have practised other online earning activities in addition to microtasking platforms. The most commonly cited are paid survey websites (60%), discount coupon websites (45%), online sales (41%), gambling and sports bets (36%), cryptocurrency sites (31%), product and test panels (26%), freelancing platforms such as Upwork and Freelancer (24%), cashback sites such as Igraal and Poulpeo (21%), remunerative applications such as Roamler and BeMyEye (14%), and the possibility to monetise a blog, YouTube channel or other content-creation opportunity online (13%). On average, they have used 3.27 such online money-earning solutions, with a median of 3.

Overall, we observe a population that actively engages in online earning activities, is broadly aware of the opportunities available, and does not hesitate to try them. Microwork is more often practiced along activities like gambling, cashback and panels, which require internet and computing literacy but not a precise specialisation, and can be practised on a temporary basis. Participation in online freelancing platforms is more limited, likely because the latter require capacity to offer expert services such as design, computer programming, and project management. Similarly, a relatively small proportion of microworkers engage in paid content creation and influencing, which require at least a significant time investment, if not a dedicated training.

Reasons for working on online platforms. Table 2.2 summarises the top three reasons why microworkers do paid tasks on online platforms, ranked in order of importance.

	1st reason	2nd reason	3rd reason
I am unemployed	14%	6%	6%
My personal situation does not allow me to find another job	11%	7%	5%
My salary is low, or I do not have a fixed salary	23%	12%	12%
The political/economic situation in my country does not allow me to find another job	1%	2%	2%
I can acquire skills that can help me find a job	6%	6%	7%
I can choose my working hours	30%	27%	15%
I prefer to, or must, work from home	15%	24%	18%
NA	0	14%	34%

Table 2.2: Reasons for engaging in online platform microwork, ranked in order of importance, in percentage. n = 577.

Clearly, workers appreciate the time flexibility that platforms offer, which appears as the top first and second reason. The possibility to work from home also matters to them, most often due to distance, in other cases because they already telework for their main job, or because they are students and 'don't have enough time to engage in fixed out-of-home work' as one of them put it; some also mention convenience and work-life balance. In a smaller number of cases, work-from-home is preferred due to disability or illness.

The need for flexible schedules and the ability to work from home should not be construed as signs of workers being in positions of privilege in their work environments. Instead, it **is often associated with those who face significant constraints (fragmented workdays, irregular working hours, low salaries that require additional income, long commutes, etc.)**, such as a substantial portion of our surveyed group, who may already have full-time or part-time jobs. For these individuals, the need to supplement their income without jeopardizing their primary employment leads to a heightened need for flexibility in their work hours. The requirement for adaptable scheduling is more indicative of a pressured and strained job situation than of a relaxed attitude toward work.

A difficult labour market position is another common reason why many people start microworking. Low or volatile salary is the top reason for almost one in four workers, and the first, second or third reason for almost half of them. Numerous microworkers also mention lack of another job:

together, **the options 'I am unemployed' and 'My personal situation does not allow me to find another job' are chosen by one fourth of respondents as their main reason**. In contrast, few workers believe that this activity can provide them with skills that can help to find a job. Overall, we see here two different groups, one that is somehow disadvantaged on standard labour markets and needs microwork to compensate, and another that emphasises the convenience of flexible work-from-home for some, however minor, extra earnings.

Earnings from microwork. The amounts of money that can be earned while performing data microtasks on platforms varies a lot, and the distribution is very skewed. One question asked to indicate the amounts earned, in euros, in each of the three months preceding the survey. Very few workers reported absence of earnings, suggesting that most of them are quite active and at least minimally experienced on platforms. Median earnings are between 20 - 25 euros each month, while averages are between 64 - 70 per month. More generally, 75% of workers earn less than 50 euros per month. However, some individuals earn much more, with peaks of up to 2,000 euros. Figure 2.4 represents the distribution of amounts earned over three months.
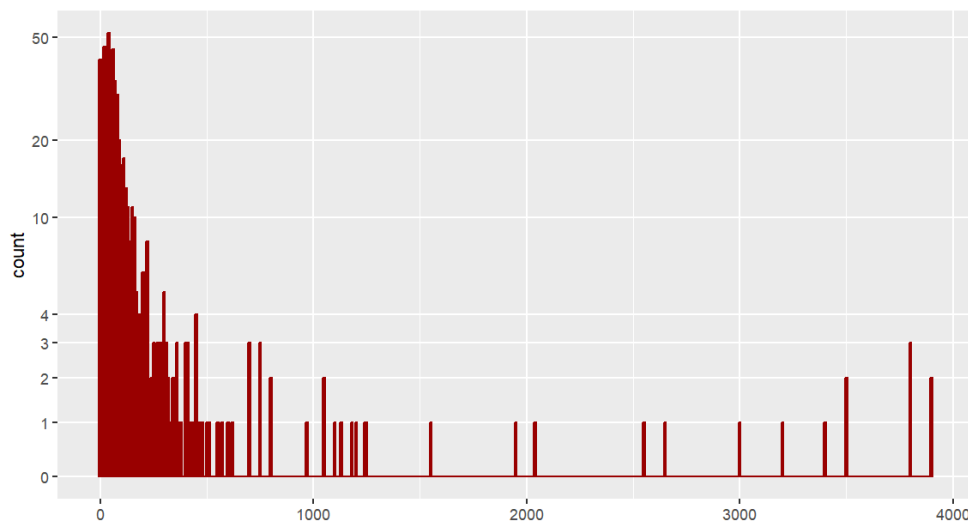

Figure 2.4: **Distribution of microworkers' earnings over three months, in euros**. n = 577.

This is largely because workers do not equally engage in platform activities. 19% of participants did fewer than 5 online tasks in the month preceding our survey, 37% did 5-10, 28% did 10-20, and only 16% did more than 20.

This European characteristic is in accordance with the existing literature on international platforms operating in North America and China. On microwork platforms, earnings follow a power law distribution, meaning that a small group of workers earn a lot, while the majority earn very little. Top earners, whose percentage can be as low as 0.4 percent of the entire user base[ZCS18], are able to earn thousands of dollars every month by getting more and better funded projects. On the other hand, microworkers in the "long trail" earn only a few dollars a month or fail to secure microtasks at all[LHGB15].

Main vs. complementary source of income. Platforms represent the main source of income of 17% of microworkers. While this group is small, it includes those who are most active in microtasking: they use on average 2.46 platforms (Prolific included), against 1.9 for those that practise this activity as a side hustle[4]. Among the 25% highest earners, they are those with the largest amounts, with an average of 1228 vs 520 euros for the others[5]. However, they do not differ by the use of other online earning websites and applications.

Microtasks. Respondents were asked to describe, in their own words, the last task performed on Prolific before our questionnaire. The majority mention questionnaires, surveys, and polls. While we do not have access to the exact formulations of these tasks and can only rely on workers' memory, the descriptions they provide suggest that in addition to marketing studies and opinion polls, there is also a significant presence of scientific research in disciplines such as psychology

---

[4]The medians are, respectively, 2 and 1, and the difference is statistically significant at 1% level
[5]The medians are, respectively, 350 and 288, and the difference is statistically significant at 5% level

and economics. This is unsurprising to the extent that Prolific's primary target are academics, rather than commercial companies. More interestingly for our purposes, some tasks are related to Artificial Intelligence (AI):

- Surveys about AI: perceptions of it, uses of it in the workplace, willingness to buy AI-designed products such as furniture.

- 'AI preparation' tasks in the terminology of [TCC20], requiring workers to generate data for machine-learning algorithms, such as voice recordings for speech recognition solutions. For example, a German-speaking worker said 'I pronounced and recorded 280 short sentences in German', and a Spanish-speaking one 'had to listen to an audio in English, state whether I agreed with what was said, and then record myself saying it'.

- 'AI verification' tasks as per [TCC20], which ask workers to assess the quality of AI-generated texts, images or videos, sometimes with the specific requirement of distinguishing an AI-generated content from a human one. Here, microworkers check the quality of algorithmic outputs and provide feedback to AI producers if needed.

The 41% of respondents who also use other microtasking platforms (such as Amazon Mechanical Turk, Appen, Clickworker, Fiverr, and UHRS as discussed above) in addition to Prolific, were also asked to describe the last task done there. AI-related tasks are much more commonly found here:

- 'AI preparation' tasks, which can be subdivided into data 'generation' and 'annotation' as per [TCC20]:
  - Data generation is most often mentioned in the form of audio recordings ('Create prompts in my mother tongue to train a voice assistant platform'), sometimes in conjunction with written text ('I had to type a name/surname/alias in Spanish, listen to the spoken form of that name produced by an AI and rate the pronunciation from 1 to 5. If it was a mispronunciation, I had to re-record that name'). In some cases, workers were also asked to provide images ('I had to take pictures of the tourist attractions in my city') or videos ('Made a video of my backyard lawn to develop the AI of a lawnmower').
  - Data annotation is the process of enriching already-available data. Microworkers label images ('I evaluated images of damaged cars to train an AI', 'The task was about recognising number plates'), classify products, sort lists of items, translate texts, and transcribe audio files. Occasionally, they also moderate contents ('Tell whether a series of images should be labelled as 'adult content'), although this type of task is relatively infrequent among this population[6].

- 'AI verification' tasks include rating AI products, for example 'choose the best of two AI-generated audios' or 'correct automated transcriptions of an audio file'. Some tasks provide services for search engine optimisation: checking the accuracy of the outputs of a search, checking the functioning of links of the suggested websites, ranking the results of different prompts, comparing different search engines.

Interestingly, although European microworkers are more expensive than their counterparts in lower-income countries, not all the needs of AI-producing companies are offshored, and we still have evidence of AI-related microwork undertaken in Europe. This is especially true of linguistic tasks that only native speakers of a given language can do. If a speech recognition producer needs 'words recorded in the Portuguese of Portugal', as one participant put it, then the task must be sent out to Portugal, not Brazil. This is one reason why, despite the growing tendency toward a polarized market where AI producers are located in the Global North and microworkers increasingly in the South, non-negligible pockets of this activity remain in Europe.

Of note, a small number of workers chose not to provide many details on their last task (whether on Prolific or elsewhere) for confidentiality reasons, although very few explicitly mentioned a Non-Disclosure Agreement (NDA). Some workers also take the opportunity to share bad or disappointing experiences: 'I registered but never received remunerations'; 'I never found available tasks'; 'I did not enjoy it at all... you only start earning something after doing more than 1000 tasks with ridiculous rewards, under 10 cents'; 'overall I made 50 dollars but then it was too much effort so I abandoned it'; 'I have no good experience with any paid survey site other than Prolific'; 'I almost never receive interesting tasks'; 'I used the site, but they never paid me'. Low pay and

---

[6]As discussed in chapter 3, content moderation is more often by business process outsourcing companies with a structured organisation, rather than just posted as a task accessible to all through a platform marketplace.

insufficient availability of tasks, largely due to swings in clients' demand, are known issues in platform microwork; and although we do not have enough information to assess whether there has been wage theft here, similar cases have been seen before. Our qualitative study goes deeper into these issues, discussed in detail in chapter 3.

Chapter 3

Labour conditions and workers' concerns in Europe

Through focus groups, surveys, and interviews, the EnCOre project endeavors to engage digital microworkers from both platforms and business process outsourcing firms (BPOs) across Europe. The insights gathered aim to foster solidarity among platform workers by facilitating communication and discussions regarding commonalities and differences across countries and work environments.

Moreover, the focus groups provided a platform for workers from various European countries to deliberate on policy recommendations and elect representatives. As part of this initiative, three microworkers representatives traveled to the European Parliament. They testified at the Forum on Alternatives to Uberization on February 22, 2024, in the context of negotiations leading up to the enactment of the Platform Workers Directive.

3.1   The Interviews

In March 2023, we conducted seven preliminary interviews with workers located in Luxembourg, Germany, France, Portugal, and the United Kingdom. With these interviews, we aimed to explore salient patterns, commonalities, and differences among countries and work settings. The insights gained from the interviews shaped the moderation guide for the focus group sessions we conducted later[1].

3.2   The Focus Groups

Between June 2023 and January 2024, we conducted a series of six focus group sessions with 18 workers in total. We organized the groups according to participants' country location and the language they felt the most comfortable speaking:

- Spain group 1 (June 12th, 2023): We conducted the session in Spanish and invited three workers currently residing in Spain, N, CH, and I. All participants started working on platforms to supplement their incomes during the pandemic or as students. For N, platform work is her primary income source. CH works whenever a good task is available, resulting in irregular workdays and hours. The participants conduct diverse tasks, including answering surveys, image labelling, algorithmic verification, testing apps, and uploading images taken by them.

- Spain group 2 (June 16th, 2023): The session was conducted in Spanish. F and V, two workers based in Spain, attended the session. F and V are active on multiple platforms, such as Toloka, Microworkers, Oneforma, UHRS, and Clickworker. V has a background in biology and works as a pharmacist in addition to working on the platforms. F has experience in marketing, business administration and customer support.

---

[1]We recruited the participants through our networks and by posting ads in different online groups and forums. The interviews were, on average, 50 minutes long. Each interview participant was compensated €30. All interviews were conducted via video call, recorded, and transcribed. The participants' names were pseudonymised to protect their identities.
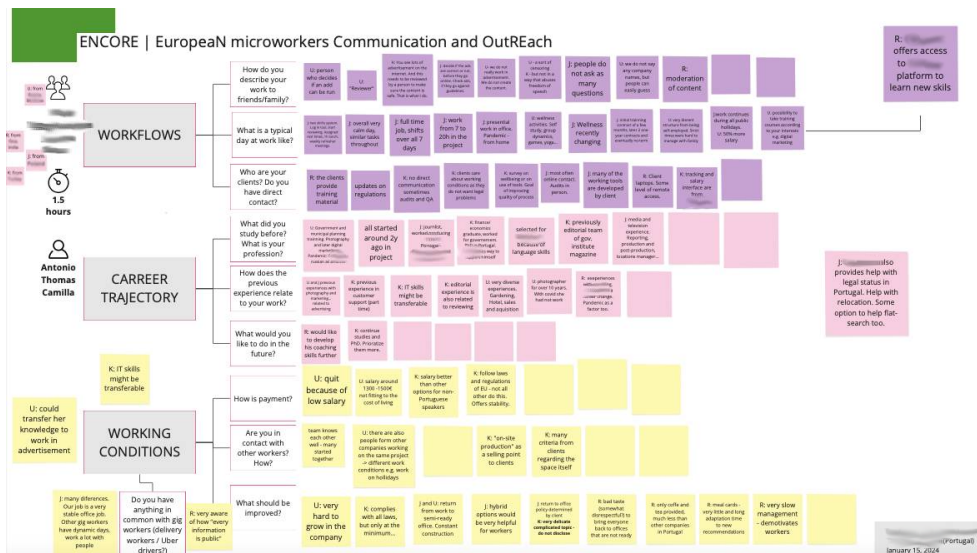
Figure 3.1: An anonymized collaborative Miro board used for an online focus group. Four participants curated and organized visual notes on themes and topics related to their workflow, career trajectory, and working conditions with the assistance of three EnCOre team members. Source: Authors' elaboration.

- Germany group 1 (June 30th, 2023): The session was conducted in English and attended by three content moderators with the BPO Telus International and one content moderator working at Accenture. All participants are currently working in part-time shifts.

- Germany group 2 (July 11th, 2024): We conducted a session in German with three content moderators with the Telus International division located in Essen, Germany. The sole client of this Telus location is the technology giant Meta. All participants got into content moderation by chance, either during their studies, after studying when no well-paying job was found, or because of strokes of fate where they had to close previously established businesses during, for example, the COVID-19 pandemic.

- UK group (July 14th, 2023): The session with two workers (D and C) was conducted in English. D lives in Northern Ireland, and C in Scotland. Both consider data work as a "side hustle." D is a construction worker, and C studies at university after migrating to the UK from Africa. They have experience working for platforms such as Fiverr, Microworkers, and Amazon Mechanical Turk. Their working hours are mainly in the mornings before their primary occupations.

- Portugal group (January 15th, 2024): We facilitated the session in English and invited four workers from a BPO in Portugal. Because of confidentiality concerns, the participants specifically requested that the company name remain anonymous. They are all full-time employees and work together as content moderators in a project, providing services for a large advertisement marketplace. All participants were immigrants to Portugal: U. is from Russia and has a background in public management and photography, J. is a journalist from Poland, K. is a PhD student of economics originally from Turkey, and R. is from India.

The sessions were conducted online via video call and recorded. We used the visual tool Miro to take visual notes that we made visible to the participants throughout the sessions. Some participants chose to participate with their cameras off to protect their identity (fig. 3.1). All names were pseudonymized in transcripts and this report[2].

---

[2]Between two and four workers attended each 1.5-hour focus group session. Paola Tubaro, Antonio Casilli, and Milagros Miceli took turns facilitating the sessions, mainly depending on availability and language proficiency. In addition, Laurenz Sachenbacher, Camilla Wagner, and Thomas Le Bonniec rotated as notetakers and facilitation support. We guided the conversation and structured our questions around three main themes: tasks and workflows, career trajectory, and working conditions.

We recruited the focus group participants through our networks, snowball sampling, and an ad posted in the platform Microworkers. We compensated each participant with €30 via bank transfer or Paypal, depending on their preference.

3.3 Salient observations

3.3.1 The tasks

The workers who participated in this study described a variety of tasks being performed by them: adding likes, comments, and views to social media content, testing software, filling out surveys, labelling images or text, moderating toxic or sensitive content on social media, validating the correct functioning of search engines, and taking and uploading pictures according to specific instructions. The tasks vary according to the platform or BPO and the needs of clients. Some platforms and BPOs specialise in one type of task.

> The participants made it clear that the platform or BPO they work for is merely a facade, an intermediary whose clients are primarily big tech giants driving the artificial intelligence industry. These giants upload task instructions, and the workers, based on their skills and availability, take on these tasks. As N from the first Spanish group explained, "the aim is to deliver high-quality work to ensure a steady flow of jobs."

Most platform workers are remunerated per task, not for the time they invest in the work. They perform these tasks on their personal computers, which they purchase with their own funds, just like they cover their internet and electricity bills, all crucial for their work. Often, they are asked to generate data in ways that infringe on their privacy: recording their voices reading various text passages (as reported in the UK focus group) or uploading selfies (as discussed in Spain group 1).

Among the tasks that require a higher level of professionalisation and a more steady workforce, and therefore are often conducted in BPOs, is content moderation. As described by the BPO workers who are content moderators for Meta and participated in this study, "the work is always the same irrespective of which shift one takes." Workflows are structured around "cues," or moderation requests, based on language. Even though content moderators have no influence on what cue is presented to them, the participants claimed that with experience, "one can anticipate more or less what types of violations are common for each language." Violations here refer to violations of Meta's terms.

3.3.2 Payment and job security

Wages vary depending on the type of task and the platform. As the participants in Spain group 1 discussed, sometimes the same task appears on different platforms and pays differently. What remains constant across platforms is the pay-per-task approach.

Upon contemplating their working conditions, all participants agreed that their payment does not reflect the toll of the work they do.

> For instance, Telus content moderators in Germany are compensated at €14.40 per hour before taxes, with a 15% increase for the night shifts. A crucial point of contention here is that workers' compensation hinges on Key Performance Indicators (KPIs), which specify that they must uphold a progress rate of 85% of tasks completed or otherwise grapple with repercussions, such as decreased payment or termination of their contract. The workers described how their team leaders harshly enforce these forms of disciplining. The Telus workers who participated in this study distinguished two types of contracts: fixed rates or commission-based payments based on KPIs. Concerning the former, they **lament forms of wage theft**, where they work up to 40 hours but are compensated only for the 32 hours specified in their contract. Concerning the latter, the problem is that there are no fixed salaries, as bonuses are calculated hourly and only if workers exceed their own personal KPIs, which creates immense pressure for already emotionally taxing tasks. A striking detail that workers reported is that breaks of 30 minutes are not paid, and the wellness breaks they are given "when you see a particularly distressing ticket" are insufficient. The content moderators also complained that their wage, only €2.40 above the minimum wage in Germany, does not reflect the mental toll their work takes on them.

The participants from the UK who had experience working for an array of platforms and conducting different types of tasks were adamant that **payment is not only irregular but unsatisfactory without real improvements in sight**. C, who migrated from Africa to Scotland to study at university, mentioned in this context that he had already worked the platforms in

his home country. Back then, he thought the payment was insufficient mostly because of the global dependencies prevalent in African economies. However, upon beginning to work in the UK, he realised that the payment was equally scant, which, coupled with the increased costs of living standards in the UK, led him to contemplate quitting this profession altogether. Both participants encountered instances where payments were withheld due to perceived inadequacies in task execution. Challenging such allegations proved nearly impossible, primarily due to the lack of communication channels between workers and clients, which platforms fail to establish.

Participants in both Spain groups also reported a lack of feedback and payment for rejected tasks and no channels of contact with clients. For instance, N recalled being disqualified halfway through a task and receiving rejections for answering too quickly. In both cases, she was not paid for the time invested in the task and received no explanation as to what she had done wrong. The participants in the Spanish group described earning between €7 and €9 per hour on average, with many tasks being compensated at much lower rates. These payment levels have remained constant despite the stark inflation rates Europe has seen in the last few years, which translates into an erosion of workers' buying power.

Piecemeal payment prevails in platforms, leading to abundant unpaid time and tasks. Workers often must study long and complicated guidelines, spend much time finding tasks, or take qualification exams, none of which are paid. Furthermore, sometimes tasks are only available for a short time and disappear in the time it takes to go over the required materials. In this sense, participants wished there was more regulation of both minimum wages and the commission and fees collected by the platforms. In addition, they expressed the need to receive precise information about how long tasks should take to better assess whether a task is fairly compensated and organise their time accordingly.

3.3.3 Secrecy and surveillance

Both platform and BPO workers highlight a pervasive culture of secrecy regarding clients, coupled with a notable lack of direct interaction with them. For example, Telus workers tasked with moderating content on social media platforms such as Facebook and Instagram report minimal contact with Meta representatives, apart from a three-week policy schooling at the beginning of their contracts and the annual audit Meta conducts on Telus' operations.

> All Telus workers are further required to sign a non-disclosure agreement (NDA), which prevents them from sharing details about their psychologically disturbing work as content moderators. All they can tell friends and family is that they make "decisions before content reaches users" or that they are "the online garbage men," as described by participants in Germany group 1. NDAs are very common in platforms, too, as reported by participants in both Spain focus groups and the UK group. This practice also hinders collectivization among workers, with an exemplary case of a Telus worker being fired in 2023 for allegedly violating the terms of his NDA after publicly speaking about his work at the German Bundestag. Accenture and Telus workers in Germany and BPO workers in Portugal reported that for security reasons, they are not allowed to touch anything outside of what is on their desks and are not allowed to take their phones or notes into the "production floor" (the large office in which they conduct content moderation tasks) during working hours.

In addition to concerns about confidentiality and corporate secrets, excessive surveillance is also a result of the establishment of performance indicators (the aforementioned KPIs). In order to "incentivize" microworkers to maintain expected levels of accuracy, production time, and productivity, **platform clients hold the power to decide if work is paid or rejected (and therefore unpaid)** based on these performance scores. Furthermore, cumulative overall ratings are crucial for workers' ability to secure further tasks, adding to the pressure of maintaining high performance levels.

The lack of communication with clients and the secrecy surrounding the utilization of participants' labour is starkly contrasted by the stringent surveillance and limited privacy many workers endure. While workers are kept in the dark about the true intentions behind their labour and its utilization by clients and platforms, various entities, including BPOs and clients, subject them to rigorous surveillance and impose terms that prohibit them from disclosing the nature of their tasks.

3.3.4 Workers' mental health

Many participants, particularly those engaged in content moderation tasks, expressed concerns about the significant impact of their jobs on their mental health. They underlined that the content they moderate fluctuates in violence and graphic explicitness, with content being especially horrifying during periods of war or after terrorist attacks.

While some companies offer psychological counselling, the counsellors often lack professional training and, according to participants, they do not seem to prioritise the well-being of the workers. The 30 minutes of "well-being time," which is just a regular break every worker is afforded, has to be taken on company grounds and cannot be structured according to each worker's psychological needs. In conjunction with the fact that the profession of content moderation is largely unknown to the German public, workers reported to be socially isolated, both in terms of structuring their daily lives but also in terms of talking with friends and family about their work and its psychological toll, also because they are inhibited from communication by NDAs.

As reported by the second Germany group's participants, Telus implemented a policy allowing workers to take "as much well-being time as necessary" following a tragic incident where a worker took his own life in 2022. However, the company maintained its KPI policy, which effectively pressured workers to exceed their limits. Consequently, the Telus workers who participated in the focus groups reported heightened distrust toward their colleagues at work and society at large. They also described struggling with sleep problems and avoiding social media, which exacerbates their sense of social isolation. To make matters worse, work in BPOs is structured in rotating shifts, weighing further on workers' psyche and complicating everyday outside-of-work tasks such as care work or attending personal appointments.

By comparison, on platforms working times are, per definition, more flexible. Nonetheless, participants report a high mental-health toll for social isolation resulting from working from home and having little or no contact with fellow microworkers. Another important source of distress is waking up and working at odd hours to catch the best-paying tasks, or being kept from meeting friends and family not to miss "good tasks" in the evening or during weekends. Furthermore, the participants in the Spain and UK groups mentioned being psychologically affected by the opacity of clients' expectations and the lack of work security and continuity. Not knowing why a task was rejected or whether their earnings will amount to an acceptable income by the end of the month "causes anxiety" among platform workers, as described by CH in the first Spain group.

3.3.5 Migration

In addition to the notable representation of migrant workers highlighted in chapter 2, it is worth mentioning migration as a significant trend observed during our qualitative study as well. Although our sampling methodology did not explicitly target migrants, it is noteworthy that all interview participants and the majority (15 out of 18) of focus group participants are migrants to the countries where they currently reside. In the exploratory interviews, five over seven respondents were migrants from Asia and Africa.

Throughout the focus group discussions, the topic of migration emerged as a salient and recurrent theme. Participants shared accounts of the myriad challenges they face when seeking employment opportunities in a new country. These challenges include navigating language barriers, as well as enduring long periods of waiting for their credentials from their home countries to be recognized and validated in their new environments.

Of particular concern, as elucidated by participants in the focus groups conducted in Germany, is **the disparity between the high levels of education workers have attained in their respective countries and the realization that this does not necessarily correlate with improved job prospects or higher wages in the new country**. Despite their academic achievements, many find themselves relegated to low-paying microwork.

Moreover, participants described a sense of vulnerability and dependency resulting from their migrant status. For many, the need to maintain employment for visa compliance often compels them to endure unfavorable working conditions, including inadequate pay, exploitative labour practices, and limited job security in platforms and BPOs.

Chapter 4

Concluding Remarks

4.1 Microwork is platform work: labour struggle and paths forward

With discussions surrounding platform work predominantly focused on location-based platforms, all participants unanimously agreed that they do not feel adequately represented. This sentiment stems not only from the inherent disparity in the nature of their respective work but also from the diverse array of challenges associated with each form of exploitation. For instance, participants identified their primary challenge as psychological rather than physical, a factor they emphatically stressed should be taken into consideration in discussions pertaining to platform work. On a positive note, the Workers Council at Telus in Essen, Germany, is actively organising and pursuing four key improvements in response to workers' concerns: They advocate for increased vacation time for all workers to alleviate mental stress, fair compensation that acknowledges the psychologically taxing nature of their tasks, access to professional mental health support without fear of reporting to management, and recognition of their work as skilled labour, which would facilitate upward professional mobility.

A significant distinction from location-based platform workers is that most participants have limited contact with others engaged in similar work, relying primarily on internet forums for connection and lacking engagement with worker organizations. One participant in the Spanish group highlighted how the pervasive misconception that microworkers choose this type of work voluntarily undermines efforts to improve working conditions. Overall, participants concluded that their social position and political agency as a workforce are currently minimal. Some of the participants attributed such lack of collective organizing and bargaining power to the dispersed nature of online microworkers, who do not share a physical space like delivery workers or drivers.

Furthermore, participants emphasized the urgent need for payment increases, direct communication with clients and platforms, and the ability to challenge unjust payment withholdings for opaque reasons. They stressed the importance of informing workers about potential privacy implications, particularly in tasks involving the submission of personal data such as selfies, which may be utilized for artificial intelligence training purposes.

On February 22, 2024, a panel of microworkers, assembled by the EnCOre initiative, addressed grievances at the Forum on Alternatives to Uberization. This event brought together platform workers unions, advocacy organizations, and political figures to deliberate on the Platform Workers Directive and advocate for its approval by the European Parliament. **The microworkers' panel effectively communicated that microworkers are indeed platform workers, stressing that their distinct needs must be acknowledged in discussions and legislation pertaining to platform work**. Attendees at the Forum unequivocally expressed solidarity with the microworkers' representatives, leading to the integration of their specific demands into subsequent public outreach efforts and documentation.

Bibliography

[BELT21]   Chiara Belletti, Daniel Erdsiek, Ulrich Laitenberger, and Paola Tubaro. Crowdworking in France and Germany. Intermediary report 21-09, ZEW, Mannheim, 2021.

[BFH⁺18]   Janine Berg, Marianne Furrer, Elliot Harmon, Uma Rani, and Mark S Silberman. *Digital labour platforms and the future of work: Towards decent work in the online world*. International Labour Organization, Geneva, 2018.

[CCL⁺17]   Logan S. Casey, Jesse Chandler, Adam Seth Levine, Andrew Proctor, and Dara Z. Strolovitch. Intertemporal differences among mturk workers: Time-based sample variations and implications for online data collection. *SAGE Open*, 7(2):215824401771277, 2017.

[CP19]   Antonio A. Casilli and Julian Posada. The Platformization of Labor and Society. In *Society and the Internet: How Networks of Information and Communication are Changing Our Lives*. Oxford University Press, 07 2019.

[CTLL⁺19]   Antonio A. Casilli, Paola Tubaro, Clément Le Ludec, Marion Coville, Maxime Besenval, Touhfat Mouhtare, and Elinor Wahal. Le micro-travail en France. derrière l'automatisation, de nouvelles précarités au travail ?, 5 2019.

[DCS⁺23]   Namita Datta, Rong Chen, Sunamika Singh, Clara Stinshoff, Nadina Iacob, Natnael Simachew Nigatu, Mpumelelo Nxumalo, and Luka Klimaviciute. *Working Without Borders: The Promise and Peril of Online Gig Work*. World Bank, Washington, DC, 2023.

[Eur24]   Eurostat. Euro area unemployment at 6.4%, 2024.

[FSJ⁺17]   Chris Forde, Mark Stuart, Simon Joyce, Liz Oliver, Danat Valizade, Gabriella Alberti, Kate Hardy, Vera Trappmann, Charles Umney, and Calum Carson. The social protection of workers in the collaborative economy. Technical report, European Parliament, Committee on Employment and Social Affairs, 2017.

[Int21]   International Labour Organization. *World Employment and Social Outlook 2021: The role of digital labour platforms in transforming the world of work*. ILO - International Labour Organization, Geneva, 2021.

[KLS21]   Otto Kässi, Vili Lehdonvirta, and Fabian Stephany. How many online workers are there in the world? A data-driven assessment [version 4; peer review: 4 approved]1. *Open Research Europe*, pages 1–53, 2021.

[LCC23]   Clément Le Ludec, Maxime Cornet, and Antonio A Casilli. The problem with annotation. human labour and outsourcing between France and Madagascar. *Big Data & Society*, 10(2):20539517231188723, 2023.

[LHGB15]   Vili Lehdonvirta, Isis Hjorth, Mark Graham, and Helena Barnard. Online labour markets and the persistence of personal networks: Evidence from workers in Southeast Asia. In *Proceedings of the American Sociological Association Conference*, Chicago, August 2015.

[LVS⁺23]   Ulrich Laitenberger, Steffen Viete, Olga Slivko, Michael Kummer, Kathrin Borchert, and Matthias Hirth. Unemployment and online labor - evidence from microtasking. *MIS Quarterly*, 47(2):771–802, 2023.

[MP22]   Milagros Miceli and Julian Posada. The data-production dispositif. *Proceedings of the ACM Human-Computer Interaction*, 6(CSCW2), nov. 2022.

[MvZtH23] Ryan A Morgan, Ward van Zoonen, and Claartje ter Hoeven. Lost in the crowd? An investigation into where microwork is conducted and classifying worker types. *European Journal of Industrial Relations*, 29(3):301–322, 2023.

[OEC23] OECD. *Education at a Glance 2023*. OECD, 2023.

[PZD22] Agnieszka Piasna, Wouter Zwysen, and Jan Drahokoupil. The platform economy in Europe. Results from the second ETUI internet and platform work survey. Technical report, 2022.

[TC0] Paola Tubaro and Antonio A. Casilli. Who bears the burden of a pandemic? covid-19 and the transfer of risk to digital platform workers. *American Behavioral Scientist*, 0(0):00027642211066027, 0.

[TCC20] Paola Tubaro, Antonio A Casilli, and Marion Coville. The trainer, the verifier, the imitator: Three ways in which human platform workers support artificial intelligence. *Big Data & Society*, 7(1):2053951720919776, 2020.

[TLLC20] Paola Tubaro, Clément Le Ludec, and Antonio A. Casilli. Counting 'micro-workers': societal and methodological challenges around new forms of labour. *Work Organisation, Labour & Globalisation*, 14(1):67–82, 2020.

[Tub22] Paola Tubaro. Learners in the loop: hidden human skills in machine intelligence. *Sociologia del lavoro*, (163):110–129, 2022.

[ZCS18] Feng Zhu, Weiru Chen, and Shirley Sun. Zbj: Building a global outsourcing platform for knowledge workers. Harvard Business School Case 618-044, Harvard Business School Publishing, January 2018. Revised October 2019.

Appendix A

Survey methodology

This annex provides methodological detail about the design of the quantitative survey, the conditions of the fieldwork, and the quality of the data collected.

A.1 Design of the online survey

The design of the online study is largely inspired by others that we used in the past (in France, 2018; in Spanish-speaking countries, 2020-22; and in Brazil, 2023) in order to facilitate comparability. The questionnaire is rather long, with **93 questions divided into 12 groups** dedicated to:

1. Introductory questions;

2. Online activities and practices;

3. Access to computing devices and internet, and literacy in use of these tools;

4. Practices and representations of online microwork;

5. Earnings from online microwork;

6. Educational attainment;

7. Proficiency in English and other languages;

8. Occupation and professional experience;

9. Family history and household composition;

10. Economic and financial situation of the household;

11. Socio-demographics.

To collect a maximum of useful information, **almost all questions are mandatory**, and clearly labelled as such. Rather than allowing respondents to skip irrelevant questions, we use **filters** extensively to remove them automatically: for example, a question about amount of time spent outside paid work is only directed to participants who have previously declared to be unemployed or inactive. For more personal questions that participants may want to skip because they are uncomfortable answering them, we offer the possibility to do so through answer options like 'I do not know / I do not wish to answer'.

We are also mindful of the length of the questionnaire and, to reduce respondents' burden, we ensure that the bulk of the survey consists in **single- or multiple-choice questions** that can be answered very quickly. We use the **'other'** answer option (and less frequently, the opportunity to add comments) to re-introduce some degree of flexibility, allowing participants to provide information in their own words whenever they do not recognize themselves in any of the predefined items. The very few open-ended questions cover issues in which we expect significant variation across respondents (for example, the exact denomination of their current profession, if they had any, excluding online micro-tasks) or when expectations are more difficult to set in advance, and predefined options could lead us to lose crucial information (for example, description of the last micro-task done online).

The survey includes **three attention checks**, meant to detect unmotivated respondents who may harm data quality by not reading questions carefully, answering randomly, or speeding through

the survey. The first, at the very beginning, tells participants about our need for high-quality data and requests a commitment to providing thoughtful answers. The second, about half-way through, is a factual single-choice question asking respondents to indicate which of five response options is a vegetable. It is obviously trivial, and its purpose is to make sure that participants are actually reading the question. The third attention check, in the second half of the survey, requires respondents to type a word. It is designed to get around the risk, that the previous single-choice check does not completely eliminate, that someone could tick the right box just by chance. Indeed, bots and random respondents are unlikely to complete a typed attention check.

We have made available our questionnaire in six languages: **English, French, German, Italian, Portuguese, Spanish**. The questions and answer options are all identical across languages, with a tiny number of exceptions needed for adaptation purposes[1]. Multi-linguism allows reaching a larger pool of potential participants regardless of their proficiency in English, and can therefore overcome a limitation of much extant literature. However, we are also limited by the languages spoken within our research group. Therefore, we are unfortunately not able to extend the same approach to Eastern European countries such as Croatia, Latvia, and Bulgaria, highly interesting because according to Morgan et al. 2023 [MvZtH23] they have high microwork prevalence.

We have programmed the questionnaire in **Limesurvey**, a widely used, professional survey software. The instance of this tool that we use is provided by the Weizenbaum Institut and hosted on its servers in Berlin, Germany, thereby complying with European data protection requirements.

A **participant information sheet**, in six languages, is provided at the beginning of the questionnaire, also including a GDPR-related notice and our contact information. Participants are asked to express their **consent** before starting.

A.2   Fielding the questionnaire

Participants are **registered users of the online labour platform Prolific**, where the questionnaire is launched as a remunerated task. Micro-work is a hardly visible activity and many of the people who practise it also have other occupations or identities: platforms are often the only setting where one can reach them. Our choice of focusing on Prolific is motivated by its vast pool of participants, its presence in Europe, and its dual specialization to serve both researchers in need of participants for surveys and experiments, and AI engineers in need of taskers to build 'human-powered datasets'[2] Prolific is also one of the best-rated platforms for online micro-work according to the Fairwork scheme[3].

We first conducted a pilot with 10 participants on 6 March 2024. After making some minor corrections based on this first attempt, we launched the final version of the questionnaire between 18 March and 9 April. We launched the questionnaire in multiple batches, on different days and at different times of the day, in order to **sample from a sufficiently diverse pool of respondents**. Indeed, different time slots may attract different people. During office hours on a weekday, one can expect more responses from people who, apart from their online microwork, are not in paid employment; while in the evening, there can be more participation from people who have a main job elsewhere and can do microtasks only when they return home. Gender differences may also play a role insofar as women, whose jobs are more often part-time and who are more often busy caring for children after school, are more likely to take up the questionnaire during office hours, while the opposite likely holds for men. Table A.1 summarizes the timing strategy of the data collection. Needless to say, each participant could respond only once.

	8:30am	10am	3pm	4pm	7pm	9:30pm
Wed 6 March						10
Mon 18 March				125		
Tue 19 March		75	121			
Mon 8 April	15				201	
Tue 9 April	30					

Table A.1: Dates and times of the data collection.

---

[1]In particular, the question 'Do you speak and understand English?' is replaced, in the English-language version of the questionnaire, by 'Are you a native speaker of English?'. In the context of online micro-work, it is important to inquire about knowledge of the English language because most tasks offered by international platforms are in English. All other cross-language differences are more minor and concern answer options (for example, to adapt mentions of educational attainments to the names of school levels and degrees in the countries concerned).

[2]See https://www.prolific.com/.

[3]See https://fair.work/en/ratings/platforms/cloudwork-prolific/?tabfor2023

To avoid having too few responses in each country, which would weaken any cross-country comparative analysis, we opened the survey only to Prolific users located in **Austria, Belgium, France, Germany, Ireland, Italy, Luxembourg, Portugal, and Spain**, that is, EU member countries where one of the questionnaire languages is spoken. We also imposed a quota requiring the same number of respondents for each of the six languages[4]. Prolific estimated that under these conditions, we could draw participants from a **pool of 13,026 people, out of its 150,271 eligible users**[5].

Each batch took no more than a couple of hours to be completed, with two minor exceptions of a worker who had to redo the survey separately at a later time, and another who had to withdraw their submission two days afterward. We checked every submission by hand before approving them and unlocking payment. The median completion time of the survey, as recorded by Prolific, was 24M 22S, and the mean was 26M 40S [6]. Of note, respondents who completed the questionnaire during office hours (between 9am and 6pm) needed more time, with a median of 25M 3S and a mean of 27M 19S, against a median of 23M 24S and a mean of 25M 51S for those who completed the questionnaire outside office hours (before 9am or after 6pm)[7].

We paid £5.43 (corresponding to about €6.34) per task, which given the median time needed to complete the survey, corresponds to an hourly rate of £13.37 (corresponding to €15.58 as of 9 April 2024)[8]. The choice of this rate responds to the need of treating participants fairly, which includes providing suitable pay for their time and effort. Fair pay is also known to give participants more motivation to answer questions mindfully and thoroughly. Indeed, it exceeds Prolific's minimum hourly rate of £6 and recommendation of £9. It is also above the hourly minimum wages of all the targeted countries: Belgium, €12.11; France, €11.65; Germany, €12.81; Ireland, €12.70; Luxembourg, €14.86; Portugal, €4.65; Spain, €8.28[9]. At the same time, the rate chosen is not conspicuously larger than the highest hourly minimum wage in the selected countries, in order to avoid introducing incentives that may unduly boost competition among workers and potentially bias results. Finally, the choice to pay the same rate to all participants avoids the risk that differentiation might become unmanageable, although we acknowledge that the incentive to participate arguably depends on comparison with local wages and therefore varies by country - a gap of which we need to be mindful. In total, including the pilot and the three batches of the main study, we obtained **577 complete responses**. 39 people started the study but then decided not to complete it, and four more were timed out[10].

A.3   Data quality

The design of the questionnaire with mostly mandatory, closed-ended questions, ensures that we have **almost no missing data**. A preliminary screening of the open-answer questions indicates that they have been answered thoroughly. Importantly, **all participants passed the attention checks**, although 2% of answers to the third attention check, which required typing a word, are misspellings: we accepted them because they all looked sufficiently close to the original word, and we did not want to unduly penalize respondents. We have thus good reasons to believe that the data are reliable.

However, the rapidity with which the questionnaire was completed suggests that **we have probably over-sampled high-intensity users**. Microworkers differ widely by their involvement in this activity: if some connect to the platform very often and are constantly searching for new tasks to do, others are regular but less frequent users, and there is also a long queue of occasional and intermittent users. This reflects the fact that microwork is not always their main earning activity or source of income, and those for whom it is only a side hustle may prioritize other activities in their daily routines. In this context, when a task is launched on a platform, it is first grabbed by those workers who happen to be connected at that particular moment in time: among them, frequent and intense users are necessarily over-represented. To reach occasional users, a

---

[4]In practice, Prolific sampled from among its users who had declared one of the six languages as primary language, aiming to get equal numbers from each group. Of note, users can declare more than one language to Prolific, and those who took our survey were free to choose the language in which they wanted to complete it.

[5]As of 18 March 2024. Figures changed slightly afterwards owing to turnover on the platform.

[6]Limesurvey's estimate is lower, with a median of 22M 19S and a mean of 24M 33S. This is because Limesurvey counts only the time spent on its website, while Prolific includes the time spent on both websites.

[7]The difference is statistically significant at 5% level.

[8]The full cost per worker was higher on our side, also including Prolific's service fees.

[9]Austria and Italy do not have an inter-professional minimum wage.

[10]Prolific excludes users who exceed the maximum amount of time allowed to complete the study, usually equal to about three times the average time.

task should offer a large number of positions so that it can remain available for a long time on the platform – so that it is still open after all the high-frequency workers have done it.

The surveys that we undertook in France (2018) and Spain (2020-22) with a very similar questionnaire instrument can be of help here. We stated in chapter 2 that they help us to 'zoom in' on two important countries in our sample, and to compare Prolific to other platforms: Foule Factory (now called Yappers.club) in France and Microworkers.com as well as Clickworker in Spain. Additionally, they help us compensate for this limitation, because in these cases, we did reach the long queue. A peak of responses was reached very early, then the pace of participation slowed down, and it took several weeks before we could reach the targeted number of responses. These comparisons help assess the extent to which our results are driven by the over-representation of high-intensity users. Of course, it will not always be possible to use them for this purpose because the fieldwork conditions, though similar, were not identical to those of the 2024 study.

For more detailed socio-demographic information, Prolific allows researchers to retrieve its own data about respondents. The variables available from the platform, in addition to time of completion of the survey and other technical details, are: Age; Sex; First language; Country of current residence; Nationality; Country of birth; Ethnicity; Student status; Employment status; and number of projects successfully undertaken on the platform.

The Authors

Milagros Miceli is a sociologist and computer scientist, leading the research group Data, Algorithmic Systems, and Ethics at the Weizenbaum-Institut. She is also an associated researcher with the Distributed AI Research Institute (DAIR) and a research partner at the DiPLab. Her research is focused on labor conditions and power asymmetries in outsourced data work, examining their impact on machine learning datasets. With a background in ethnographic fieldwork, interviews, and workshops with data workers globally, she actively engages communities of precarized data workers from the Global South. Her broader interests include questions of legitimacy and knowledge production in data, worker-led research, labor organizing, and the intersection of critical data studies and data activism.

Paola Tubaro is research professor (*Directrice de Recherche*) in sociology and technology at the National Centre for Scientific Research (CNRS, in French Centre National de la Recherche Scientifique). She practices interdisciplinary research that leverages synergies between economic sociology, social and organisational network analysis, and artificial intelligence. She is currently studying the role of platforms in the digital economy, the place of human work in the global production networks of artificial intelligence, and social inequalities in online labour. She has also extensively published in the fields of data methodologies and ethics in data science and technology. She is the co-founder of the research initiative DiPLab.

Antonio A. Casilli is full professor of sociology at the Telecommunications school (Telecom Paris) of the Institute Polytechnique de Paris. He is a by-fellow at Churchill College, University of Cambridge and an associate researcher of the "Data, algorithmic Systems and Ethics" research group at the Weizenbaum-Institut, Berlin, directed by Milagros Miceli. With Paola Tubaro, he co-founded the DiPLab (Digital Platform Labor) research initiative and the INDL (International Network on Digital Labor). In addition to several scientific publications in French, English, Spanish, and Italian, he is the author of the book Waiting for Robots. The Hired Hands of Automation, University of Chicago Press, 2024, and the editorial adviser for Invisibles – Les travailleurs du clic [Invisible Clickworkers], a documentary series for France Televisions, based on his work.

Thomas Le Bonniec is a Sociology PhD researcher at the Institut Polytechnique de Paris. He is working on the recognition of AI workers by data protection authorities and digital commons. He obtained a master's degree in sociology in 2018 at Sorbonne-Université. In 2019, he helped uncover how Siri, Apple's vocal assistant, collected users' audio recordings without their knowledge. He has worked on the digital industry from a critical perspective since then. He participates in advising public decision makers, civil society, NGOs, and trade unions. He also writes news articles on surveillance, platforms, digital labor, the environmental cost of digital economies and regulation.

Camilla Salim Wagner is a a political scientist, currently persuing her masters degree. She works as student researcher in the group "Data, Algorithmic Systems and Ethics" of the Weizenbaum Institute in Berlin. She is involved in projects related to the working conditions of data workers who provide the essential basis for artificial inteligence systems, including the EnCOre (European Microworkers Communication and Outreach) and the Data Workers' Inquiry.

Laurenz Sachenbacher has been working as a student assistant in the research group "Data, Algorithmic Systems and Ethics" at the Weizenbaum Institute since June 2023. He is studying philosophy of technology in his master's program at the Technische Universität Berlin. He completed his bachelor's degree in cognitive science at University College Maastricht. In 2020 and 2021, he was a student assistant at the Freiburg Institute for Advanced Studies (FRIAS). As part of the project "Responsible AI: Normative aspects of the interaction of humans and intelligent systems", he conducted research on socio-political consequences of algorithmic systems, with a special focus on the care sector and supported the organization of an interdisciplinary conference on "Responsible Artificial Intelligence".