

Beyond 'platformisation'

Designing a mixed-methods approach
to inspect (digital) working conditions through
organisational systems

*Maxime Cornet, Clément Le Ludec, Elinor Wahal
and Mandie Joulin*

Maxime Cornet is a PhD candidate at Télécom Paris/
Institut Polytechnique de Paris, France

Clément Le Ludec is a PhD candidate at Télécom Paris/
Institut Polytechnique de Paris, France

Elinor Wahal is a PhD candidate at the École Normale
Supérieure, Paris-Saclay, France

Mandie Joulin is an MA student at the University of
Rennes, France

ABSTRACT

The transformations brought by the digitisation of work and the emergence of platform labour have deep implications for working conditions. However, researchers face difficulties studying platformised work. Workers' invisibilisation and the lack of physical co-presence renders field access difficult. How can the variety of platforms, their organisational systems and the working conditions they offer be accounted for? In this article, we propose a mixed-methods methodology to study platforms in all their diversity by articulating the macro level – the market structure revealed through a multiple correspondence analysis – and the micro level – detailed studies of targeted platforms carried out using desk research. We apply this method in two projects, and, in each case, a typology emerges that supports the need for a diversification of the concept of 'platform labour' when related to working conditions.

KEY WORDS

Platform work, virtual work, labour organisation, mixed-method research, multiple correspondence analysis

Introduction

The increasingly widespread use of digital technologies is causing rapid transformations to the global economy and leading to deep changes in the labour market. With between 1% and 5% of the European population having had a revenue from a Digital Labour Platform (DLP) (Forde et al., 2017), platform labour represents a key aspect of such shifts. The ‘virtualisation of work organisation’ is accompanied, notably, by a growing digitisation of workplaces. The emergence of DLPs is a symbolic example of this evolution, as it is based on four trends linked to work digitisation: ‘the elaboration of Global Value Chains’ in the service sector; the ‘evolution of freelance Labour Markets’, with the replacement of ‘word-of-mouth’ traditional ways of finding jobs by the use of internet mediated service; the ‘growth of teleworking’; and the increase in ‘standardisation and performance monitoring’ through digital tools (Huws, 2017).

A better understanding of platform work is thus of the utmost importance for researchers focusing on labour issues, as it a vital factor in workspace flexibilisation and fragmentation processes (Weil, 2014) and has major implications for working conditions. Until now, researchers have mainly focused on specific DLPs. The most well-known DLPs can be divided between ‘geographically sticky’ (Ojanperä et al., 2017), ‘location-based’ (Berg et al., 2018) urban transportation and food delivery services, such as Uber or Deliveroo, freelancing platforms such as Upwork, and a specific subset of DLPs known as ‘micro-work’ platforms, notably involved in the AI production chain (Tubaro, Casilli & Coville, 2020). Among these, Amazon Mechanical Turk (AMT) has been the object of the majority of studies. However, because AMT only provides monetary compensation for workers located in the USA and India, generalising findings about the situation of micro-workers from a study based only on AMT de facto excludes workers from other countries. Similarly, neither Deliveroo nor Uber are representative of the whole set of practices in their respective subset of DLPs.

Furthermore, the current set of methods deployed to study platform work displays several limitations, preventing the generalisation and contextualisation of their findings. In other words, we lack an established and efficient method to map the current DLP market and to systematise the study of the variations that can be found in the organisational structure of DLPs. This article offers a methodological approach to this challenge. Such a methodological approach could be used to define targeted and hypothesis-related social maps of DLPs as well as linking the organisational characteristics of platforms to workers’ working conditions. This article also validates the proposed approach by applying it to two distinct research projects led by the authors.

The challenges of studying platform work

The precariousness of platform work

Labour market transformations brought about by workspace digitalisation are rooted in a pre-internet era that saw the development of occupations related to the information society, in particular professions grouped under the label of ‘clerical workers’, most of whom were women (Machlup, 1962; Gardey, 2001; Hicks, 2017). From the 1960s to the

1990s, the digitisation of the economy created a need for data processing workers, especially to create digitised databases, with outsourcing chains going from the global North to the global South (Posthuma, 1987; Soares, Mitter & Pearson, 1992; Freeman, 2000). From the 1990s, researchers became increasingly interested in the development of telework (Huws, Korte & Robinson, 1993; Huws Jagger & O'Regan, 1999; Huws, 2007), in some cases with a focus on whether it represented an alternative form of employment, most notably for women (Gurstein, 2001).¹ Nowadays, attention is being increasingly drawn to the question of a possible platformisation of work (International Labour Organisation, 2021). These trends point to a future of work in which digitalisation allows the increased externalisation and distancing of service jobs, raising several questions about how to define these new occupations and changes in work, and to classify platform workers.

A first set of DLPs and social media platforms has been investigated through the lens of virtual work. The scientific literature has particularly distinguished three activities: crowdsourcing (micro-work), game-related work activities (goldfarming) and value extracted from online communities (Cherry, 2010; Holts, 2013). Some researchers, such as Pamela Meil (2018), support a definition of virtual work that excludes workers performing localised activities.² On the other side, 'digital labour', researchers tend to include those workers among virtual workers, since they are producing data used not only by these websites to display ads, but also to enhance their services, as well as to train machine learning algorithms (Scholz, 2012; Fuchs, 2012; Casilli, 2019).

Despite the heterogeneity of services performed and sectors involved, all DLPs present key common features. They allegedly act as intermediaries and not employers, mediating single tasks or services rather than more comprehensive positions. Within these platforms, the management of workers is carried out by applications using algorithms which automate various managerial tasks, such as work-shift allocation and work monitoring, while job quality evaluation is outsourced to consumers (Mateescu & Nguyen, 2019).

The precariousness of work performed within DLPs is well-established, especially within local-based platforms. Workers often fall outside standard forms of employment, and as a result have limited or non-existent access to social protection (e.g. Berg, 2016; Aloisi & De Stefano, 2018; Eurofound, 2018; Adams-Prassl & Risak, 2016; Prassl, 2018; Daugareilh, Degryse & Pochet, 2019). Work performed through DLPs has also been found to be remunerated significantly less than other, more traditional, forms of employment (Berg, 2015; Leimester, Durward & Zogaj, 2016; De

1 With a very mitigated result (p. 197): 'For low-skilled workers, such as data processors, the degree of autonomy remains low; their status does not change because of a change in work venue. As recent studies illustrate, low-skilled workers are the most vulnerable and the most in need of protection (Pearson and Mitter, 1992). The disparate nature of contract work and piecework, coupled with domestic obligations for this predominately female population, renders it virtually impossible for workers to organize themselves effectively, making them vulnerable to exploitation and willing to accept substandard working conditions.'

2 However Meil remarks that: 'There is no such thing as purely virtual work unless perhaps if it is being performed by algorithms, which were nonetheless written, implemented and managed by real workers.'

Stefano, 2016; Forde et al., 2017). 'Working hours flexibility' imposed on the workers, which forms the basis of many DLP managerial models, also implies incentive systems that can lead to intense time pressure (Huws et al., 2018; Eurofound, 2018).

DLP workers moreover often find themselves defending their rights without being able to form workers' unions, or gain sufficient knowledge about their rights (Kilhoffer et al., 2017; De Groen et al., 2018; Johnston & Land-Kazlauskas, 2018). This lack of representation is particularly noticeable in the case of remote workers, as the lack of physical co-presence makes them especially vulnerable to the invisibilisation of their social struggles.

The post-colonial functioning of micro-work, deploying outsourcing mechanisms originating from the global North toward the global South, reinforces the invisibilisation and the precariousness of those workers (Gray & Suri, 2019). Even if some literature shows that micro-work and online freelancing can improve the living conditions of workers in developing countries, the risks of discrimination against workers and the lack of social protection are two important pitfalls (Graham, Hjorth & Lehdonvirta, 2017; Berg et al., 2018).

Existing methodological strategies to study platform work

To study this virtual, precarious and *in fine* invisible workforce, researchers have developed several methodological strategies. Field access is particularly difficult, due to the lack of physical co-presence, and the relative reluctance displayed by platforms to allow access to internal information.

Despite this, ethnography has been an instrumental method in the analysis of virtual work. A notable example is the work of Alex Rosenblat in her volume *Uberland: How Algorithms are Rewriting the Rules of Work* (2018). She presents the results of a four-year-long ethnographic study on the functioning of Uber and technology's role in shaping the contemporary economy. Methodologically, the study was carried out in 24 cities and online, and included the analysis of Uber drivers' online forums, participant-observations in over 400 rides as a passenger as well as formal interviews with Uber drivers.

Another technique commonly adopted in the study of virtual work is the study of workers' online forums. Conversations and messages in these forums are treated and analysed as archival data, which are used to study a number of different social phenomena. The use of such techniques has been particularly widespread in the study of online sociability and self-organisation among communities of platform workers. One example can be found in a 2018 study by Wang and colleagues, which addressed the self-organisation practices of Amazon Mechanical Turk workers into online communities as well as the impact of such organisation on workers and clients.

The digital dimensions of DLPs also allow researchers to use webscraping methods. A well-known application of these techniques is the 'Online Labour Index' whose purpose is to track and measures the utilisation of DLPs across occupations and countries (Kässi & Lehdonvirta, 2018).

Among qualitative techniques, interviews are also widely used. Mark Graham and colleagues (Graham & Anwar, 2019), investigating the geographical dimension of

digital work and the emergence of a 'planetary labour market', for instance, carried out 65 semi-structured interviews with online platform workers in South Africa, Kenya, Nigeria, Ghana and Uganda, recruited from an online labour platform and via social media groups.

Most studies on virtual work make use of a combination of these and other research techniques. In most cases, mixed-methods research includes survey questionnaires that are often administered to workers directly through DLPs. This is the case for a large number of policy reports and international studies on the subject (see e.g.; Piasna & Drahoukoupil, 2019; Wood et al., 2019; ILO, 2021). In other cases, random population surveys have been used (e.g. Huws et al., 2019)

Like all research techniques, the aforementioned methods have both advantages and disadvantages. The analysis of archival material from forums, Reddit threads and other online sources allows researchers to witness interactions among workers that would have otherwise remained inaccessible, where key topics are often addressed and collectively negotiated, such as self-organisation strategies and common lines of action towards clients and platforms. On the other hand, studies have shown that only a small minority of platform workers are active on forums (Yin et al., 2016), and that among such minorities, activist approaches are over-represented, which could lead to an overestimation of collective action phenomena on DLPs. This holds true in the case of qualitative interviews as well, where workers are often recruited from online forums and social media groups.

Another important epistemological consideration should also be made relating to the use of DLPs for the recruitment of interview respondents and the administration of survey questionnaires. While such use of DLPs represents an efficient tool for researchers, as it grants a wide pool of study participants in a short period of time, it also brings important constraints. In particular, it does not allow for a wide diversity of DLPs to be included, thus hindering study representativeness. This is because many DLPs only operate on a B2B basis and do not allow researchers to upload interview- or survey-related tasks.

As a result, the presented methods make it difficult to get a broad picture of the variety present in the DLP space, especially regarding work-related factors.

Several DLP typologies exist in the literature but they are generally informed by theory and built on the skill level required, on how the service is delivered (online/offline) and on the work allocation process (Drahoukoupil, 2016; Eurofound, 2018; Kilhoffer et al., 2020). Few researchers try to divide DLPs according to 'types of service provision or relationships between the platform worker and the client being facilitated by the platform' (Hauben et al., 2020).

Apart from specific particularities, companies that fall under the DLP name are far from a unified category. Some are fully integrated in the already touched-on 'post-colonial functioning of micro-work', recruiting workers in the global South to carry out remote work for Western companies. Some recruit workers in targeted areas, including the EU, for specific reasons (data security, local culture knowledge, etc.). Others, such as food delivery platforms, deploy workers in the public space. All of these factors may affect working conditions and legal statuses differently. It thus becomes clear that a fragmentation of the concept of 'platform work', along technical

lines but also along more traditional organisational lines, might be a useful approach to analyse the implications for labour-related outcomes. In the words of Lynne Pettinger (2019:137):

I don't accept the assumptions . . . about the effect of IT, that IT is a unified, coherent and consistent thing and has a consistent effect on other practices . . . I find the fetish of the present and future to be especially disturbing in historically and technologically naïve discussions of algorithms and platforms.

In this sense, finding a way to identify variation, and, eventually, to be able to develop targeted DLPs typologies, becomes paramount to acquire a better understanding of what those workers face.

Since platform work is not a unified field, typologies should be hypothesis driven, and considered valid only as long as they prove useful to discuss any given hypothesis. The methodological method proposed in this article does not aim at defining a global DLP typology, but rather at defining a set of steps that can be mobilised to approach DLPs from a specific angle, and produce a research question-oriented classification. The two DLP typologies through which we illustrate our proposition later in this article should be read in this light. They are both only valid insofar as they provide material useful to address our initial research topics: platform workers' dependence to the platforms and legal status (for the Swirl project), and platforms' outsourcing chains (for the Hush project).

Empirical approach

As part of the Digital Platform Labour (DIPLab) research programme, we conducted two projects that illustrate online research methods tailored to the study of DLPs. The first of these, Slash Workers and Industrial relations (SWIRL) was a project aiming to assess the contingent work phenomenon, in particular slash work in its different forms, and to analyse its characteristics and dynamics in European countries. SWIRL was inscribed in a larger corpus of research aiming to better understand the relationship between technical innovations (digitisation of work) and the labour market. One of the goals of the study was to map the European DLP space and to document the way different platforms structures impact both workers' legal status and their ability to negotiate working conditions (through collective bargaining processes).

The second project, HUman Supply cHain behind smart technologies (HUSH) project addressed the global value chains of data labour within the Artificial Intelligence (AI) industry, which has been mainly studied from the perspective of platform micro-workers. For the moment, the issue of the social ethics of artificial intelligence is increasingly being studied through the lens of platform work and the outsourcing of menial tasks like data generation and annotation, particularly to countries in the global South (Gray & Suri, 2019). In line with the work of Kate Crawford and Vladan Joler (2018) on the 'extractive activities' required to operate an AI system and from the perspective of the study of global value chains, the HUSH study aimed to contribute to the understanding of the value chain of business AI and the modalities of outsourcing 'menial' AI tasks.

Multivariate analysis and case studies to investigate working conditions beyond organisational structures

In both cases, a database comprised of secondary data extracted from information publicly available online was collected, describing platforms rather than workers. Each time, these databases were complemented with qualitative case studies allowing to cross-validate, deepen and contextualise statistical findings. The rationale behind the use of secondary data is to approximate the state of the DLP market and the data annotation industry, with the idea that a 'market map' allows us to capture the variation of statuses, situations and working conditions faced by workers on these platforms, and most importantly, to draw and analyse the structure emerging from those variations.

Studying platform work by combining multivariate analysis and case studies

The use of case studies is well-established in sociology, whether in mixed-methods or qualitative surveys. This type of method has found itself at the core of the opposition between quantitativists and qualitativists, with the former pointing out the limits of these methods regarding the generalisation of results, i.e. the limits of objectivity and representativeness. Jacques Hamel (1993) proposes possible responses to these criticisms, indicating that the explanation of methodological strategies in the definition of the case study responds to the imperative of objectivity and by distinguishing between statistical representativeness and sociological representativeness. This second concept makes it possible to determine the value of the case study with regard to what it says about the phenomenon studied.

Ultimately, at the heart of the issue is the idea of statistical inference. How can we generalise results and infer statistical findings from discrete and unique qualitative case studies? Part of the answer lies in the proposal of Charles Ragin's (1987/2014) qualitative comparative analysis (QCA) which is a case-based approach that 'involves defining patterns of variables within each case' to then create case typologies before comparing the cases. This approach can be thought of as a formalisation of variables methodically drawn from qualitative material. Defining 'synthetic variables' that can be obtained from desk study analysis or interviews allows us to identify and study repeating patterns of variables.

Once the variables are defined, we still have to deal with the issue of extracting useful information from non-representative datasets. For our projects, we reviewed several geometrical analysis methods which make it possible to visualise and analyse categorical data. One of the widely used approaches in social sciences is Multiple Correspondence Analysis (MCA). MCA is a statistical analysis method that can be used to systematically describe information contained in large datasets. It offers the advantage of displaying the different variables and observations used to run the analysis in an orthogonal geometrical space visually.

Husson and Josse (2013) describe correspondence analysis as 'a geometric and algebraic point of view that enables a simultaneous visual display of rows and columns of a contingency table' (Lebart & Saporta, 2014, p.35). In their view, MCA is an extension of CA 'from the case of a contingency table to the case of a complete

disjunctive binary table' (Lebart & Saporta, 2014, p.35). With this method, we are not trying to infer a true distribution of DLPs, or to model parameters effects on a dependent variable. As Jean Paul Benzécri writes, 'Philosophically, correspondence analysis, which simultaneously deals with a large set of facts and confronts them to reveal their global order, is more a matter of synthesis . . . and induction than of analysis and deduction' (Benzécri, 1977:10). In other words, the method is data-driven, and reveals hidden structures based on categories that co-occur in a dataset. In this sense, it allows us to support typologies, or 'profiles' (Benzécri, 1977:12) depending on the characteristics of DLPs, even without statistically representative datasets.

In both of our studies, our statistical unit was 'companies operating in the AI industry and in the platform economy'. To study their activities and how they organise and manage their workforce, we proposed to articulate MCA, geometrical clustering and case studies to draw typologies from own-defined variables of companies.

We chose to complement the MCA with a k-means clustering led on the retained axis coordinates to develop a discrete partition of platforms. The resulting partition is, by design, strongly linked with the 'social space' drawn by the MCA, which served as our analytical framework. Moreover, using axes coordinates, even if it reduced information by constraining variation, allowed us to get around the limitations of our datasets resulting from being composed exclusively of categorical variables. We could then avoid dealing with the non-trivial issue of computing distance metrics on non-numerical data. The usage of MCA coordinates to compute geometrical clusters is a well-known practice in social sciences (for application examples see Bourdieu, 1999; Alvarez, Becue & Lanero, 2000), and the implications have been thoroughly discussed since the inception of the method (Benzécri, 1977).

Presentation of methodological strategies

Both projects adopted similar data analysis plans, allowing us to cross-validate our mixed method approach to study DLPs and data workers, with slight design variations.

On the HUSH project, we first conducted case studies on 22 AI companies. These companies were selected on the basis of two market reports dedicated to the data annotation market. Our first exploratory phase objective was to identify the activities required for artificial intelligence by analysing the business propositions of these companies and their use of data annotators (micro-workers) for AI.³ This first set of case studies allowed us to compare companies seemingly operating in the same industry (AI), but in fact carrying out sometimes quite varied activities.

On the basis of this first exploratory phase, we built a database of 127 companies in the AI sector using company registries such as Crunchbase.com and Owler. The latter allowed us to look for companies' competitors. Between June and July 2020, we systematically included competitors and competitors of competitors in our database until we reached data saturation. To mitigate possible biases, this database was

3 We defined variables grouped into three categories of information that are present in these case studies: the profile of the company ('company name', 'founder' or 'company location'); the products or services they sell; and their revenues.

supplemented by a database of AI companies rather focused on annotation from data collected in May 2020 using the snowball method. Based on our case studies, we then manually binary coded (i.e. yes or no) our database regarding companies' characteristics⁴ and selected only those doing AI activities.

For SWIRL, the added difficulty of coordinating research teams based in six European countries during the data collection phase demanded a slightly different methodological approach. A first theoretical DLP typology was drawn from existing literature. We then designed a two-step lax quota selection system. We asked our research partners to freely add a first round of platforms to our DLP list, and then hand sorted those according to our theoretical typology.⁵ To mitigate selection biases we also performed a second round of triage, by asking partners to go through the overall list, and highlight which platforms added by another partner also operated in their country.

Using the resulting sample, we collected various information from the platform websites, as well as any relevant resource found via online documentary sources. This resulted in a database describing 26 variables, sorted into two groups.

The first group comprised contextual variables: information needed either to fill the database, as general meta-data surrounding the collection process or as transitional variables used to limit the sample. (For example, is the platform operational or discontinued? What is the platform URL? In which partner country does the platform operate?)

The second was made up of empirical variables: information collected directly from the platform websites: payment systems; geographical scale of operation; industrial sectors; services provided by the platforms, etc.

Within the framework of these two projects, we considered these public sources as archival data likely to provide information on the organisation of production and work within these digital factories. In spite of the invisibility of workers, the platforms were providing marketing and regulatory discourses, whether through their websites or their terms of use. Far from the subjective analysis of the working conditions allowed by interviews, these public archive data become opportunities to address these issues by trying to objectify them at the level of a sector or an industry (McDonald, Williams & Mayes, 2020).

In both projects, once a stable database had been obtained, we conducted a multiple correspondence analysis (MCA) to classify platforms according to the activities they offer (i.e. a typology). Missing data were imputed using the MissMDA R package, which tried to find a local optimum through an iterative, EM type, MCA process (Josse & Husson, 2016). As is always the case with imputation, this relies on the assumption that missing data distribution is random (that no observed or unobserved variable can explain the fact that the value is missing). To try to

4 A first set of variables describes the type of activities proposed by the company: 'data collection'; 'annotation tools'; 'development tools'; 'models development'; 'computing power'; 'models optimisation'; 'data hosting'. The second set of variables is focused on the organisation of the annotation workforce: 'workforce'; 'impact sourcing'; 'crowd-sourced'; 'outsourced'; 'in-house'.

5 We subsequently, we asked each SWIRL partner to compare their initial lists with those of the other countries, and to either complement it or to provide a detailed explanation of the observed discrepancies.

minimise biases, we used the same set of active and supplementary variables for the imputation process and the subsequent analysis. Active variables were used to build the n-dimensions MCA space, on which supplementary variables were then projected. Variables with a rate of missing values above 30% or with a category under 20 observations and contextual variables were always used as supplementaries (see supplementary tables 1 and 2). We ran the MCAs using the FactoMineR R package (Lê, Josse & Husson, 2008).

One important advantage of the MCA lies in its geometrical properties. Each individual is projected onto an independent axis (synthetic dimensions), and therefore is attributed a numerical value describing a position on each axis. To derive discrete structural categories from our independent axis, we therefore ran geometrical clustering methods (k-means) on DLPs coordinates. Conceptually, this gave us a data-driven typology of DLP/AI companies.

The number of axes relevant to our analysis, as well as the number of clusters for the partition were chosen using the inflexion point method, based respectively on the percentage of explained variance for the MCA, and the scree plot for the k-means clustering (see supplementary figures 1 and 2, 4 and 5). For both projects, the k-means scree plot suggested a cut between three and four clusters. Partition decisions were ultimately made in accordance with the analytical relevance of the clusters.

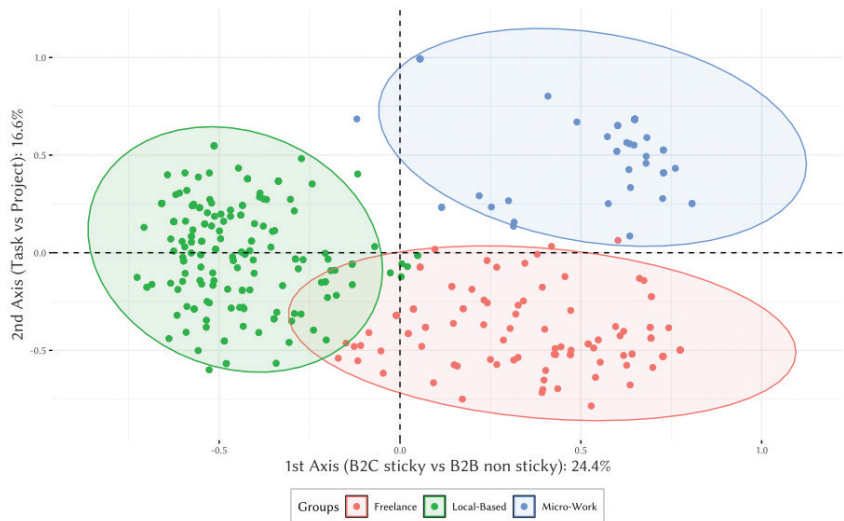


Figure 1: SWIRL: Digital Labor Platforms typology: Platforms projection on the 1st and 2nd MCA axis

In both cases, a cross-comparison between the ‘preliminary’ typology (derived from cases studies or theoretical framing) and the one derived from the MCA gave important results, allowing us to validate the relevance of our methodological approach. For

SWIRL, we then chose 15 DLPs belonging to different clusters (or ‘types’), on which to base the case studies.

Articulating studies of structure and individuals: using mixed methods to reveal the organisation of production

The challenge of this section is to specify the contributions of each method, individually and in combination with each other. We will see that it is the articulation between the study of micro and macro levels that has allowed us to reveal the structure of the organisation of production.

In the HUSH study, we chose to use initial case studies as a basis for understanding the AI value chain. A further objective was to understand how client companies (large groups) or AI companies access annotation resources. A main hypothesis regarding the organisation of annotation work was that, despite the fact that the literature is largely focused on platformised micro-work, other forms of organisation could exist in this field.

The first step was to identify activities related to artificial intelligence which had not necessarily been considered by previous researchers. In this first list of companies, we found big tech companies offering computing power or data hosting services (i.e. data centres) and pre-trained, generic models (for instance natural language processing neural networks trained on web data). We also found that some companies offered integrated annotation services. A second contribution was to underline the variety of ways to outsource annotation. For instance, within the same annotation platform, there sometimes existed distinctions between workforces regarding task complexity, time management and/or data security, as illustrated by the following workforce offering advertised by one company:

First, an on-demand crowd, i.e. annotators available 24/7 for low intensity micro-tasks;

Second, a ‘remote crowd’, i.e. qualified annotators selected according to customer needs;

Third, a ‘secure crowd’, i.e. certified and ‘identified’ annotators working for the company from ‘secure premises’;

Fourth, an ‘on-site crowd’, i.e. qualified annotators who work from the customer’s premises, but are managed by the company; and

Fifth, ‘international customer resources’, who are the employees of customers using the company’s platform.

This work also allowed us to define variables from which we derived our theory-informed database of artificial intelligence companies on which we ran a k-means clustering method. We found the social space of AI companies to be divided along two axes, the first composed of companies using an annotation workforce, contrasted with a second made up of those selling computing power services. We identified three clusters,

Cluster 1 was made up of full stack AI companies offering all the services (large digital companies like IBM or Amazon). Some have their own annotation workforce; others offer their partners’ workforce.

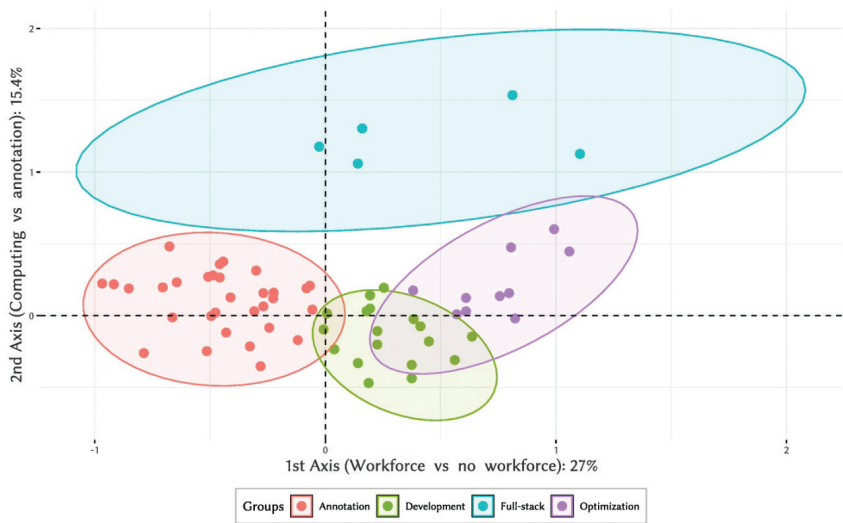


Figure 2: HUSH: AI value chain from a labour-process perspective: platforms projection on the 1st and 2nd MCA axis

Cluster 2 was composed of companies offering computing power, particularly in the form of hardware (e.g. a specialised AI processor from an Intel subsidiary) and model optimisation companies (computing optimisation companies).

Cluster 3 comprised AI software companies which mainly offered development tools, pre-trained AI models that customers could adapt to their needs and data science services.

Cluster 4 consisted of annotation companies offering two types of service associated with AI: annotation tools (software); and an annotators workforce. Many annotation companies offer these two kinds of service direct.

We then delved deeper into some of the case studies. This work in progress consists in analysing the different modes of outsourcing for data annotation services. One of the first interests of this phase was to nuance the distinction between workforce activities and computing activities within the ‘social space’ of the AI industry. For example, we found that some of the big companies in the first cluster were in fact integrating an annotation workforce into their services. The same phenomenon can also be observed for cluster 3, with companies specialising in AI model integration that are often based on models from big companies in cluster 1 and rely on annotation services to adapt these models to their customers’ use cases.

We also note that data labour is not as widely platformised as the very large part of the literature devoted to Amazon Mechanical Turk suggests. On the contrary, the logic of workforce stabilisation and data security is now pushing AI companies to propose outsourcing modes closer to what is observed in the Business Process Outsourcing (BPO) sector with, as a result, distinct working conditions and types of contract. For example, Amazon offers enterprise customers of Amazon Web Services (AWS) the

choice either to use the Amazon Mechanical Turk platforms with self-employed individuals or to use enhanced services (a stable and secured crowd) through a marketplace of artificial intelligence services, with employees working for BPO-like companies in low-income countries.

For SWIRL, the MCA allowed us to derive clusters from our self-collected variables, and to compare how this data-driven classification related to the theoretical ones studied in the literature. From the MCA/k-means tandem, we identified three clusters largely overlapping with our theoretical categories.

The first of these clusters is made up of local consumer-oriented services (49.60% of the observations). These are generally business to consumer platforms, such as on-demand drivers or food delivery platforms as well as on-demand child or pet care platforms, etc.

The second consists of Crowdwork services (19.89%): platforms offering remote services (such as AI training and micro-translations) as well as local micro-tasks (such as retail-intelligence and taking photographs of specific products in various stores).

The third category is made up of self-employed workers working remotely (30.51%). These platforms offer high-skilled, complex jobs, that require some level of project-management skills, such as freelance graphic design or software engineering platforms.

Our social DLP space seemed moreover to be divided, according to the MCA, mainly along two axes: one opposing spatial variables (local versus remote) and client types (business-to-business versus business-to-consumer) and one made up of platforms offering skilled, qualified work opposed to a second one offering more 'taskified' work. While underlining the importance of the spatial division of labour, this divide also seems to imply a power imbalance in the platform-work division along the second axis. The main factors differentiating freelancers from crowdworkers, if we do not account for geographical stickiness or work fragmentation, seem to be the skill level expected to complete a job, the fact that the platform offers training to its workers and who sets the rates of the transaction (workers, platform or clients?). This divide raises questions as to the persistence of traditional class-based cleavages at play within the platform workforce.

Detailed case studies on 15 platforms selected from the three clusters allowed us to better characterise variations in DLP structure regarding working conditions. Notably, for remote workers, the divide between crowdworkers and freelancers seems to have more to do with individual workers' ability to build and leverage a reliable client network, thus mitigating algorithmic influence. This skillset partially depends on the nature of the work being traded on the platforms, as longer and more complex projects demand at least some level of client-worker interaction. This finding is mitigated by the fact that it is also highly dependent on technical choices made by the platforms, as well as on specific modes of platform governance. Inequalities among workers are more prominent on big on-demand, remote work platforms such as Fiverr or Upwork, where crowdworkers are rendered invisible by informal intermediaries and middle persons.

Discussion: strengths and limitations of the method

Research on platform workers is a rapidly expanding field of study in digital sociology and the sociology of work. At the same time, this academic effort is combined with the mobilisation of workers to defend their rights and with the regulatory efforts of public

authorities. This can be explained by the fact that the 'platform model' has spread into more and more sectors, affecting more and more workers. From our point of view, this interest in this new 'object', as a typical model of organisation, masks the fact that platforms do not constitute a homogeneous group of organisations, regarding sector of operation, legal status definitions or other organisational structure traits.

Overall, those organisational characteristics have real implications for working conditions, workers' precariousness and health and safety at work. The type of job that the platform offers, its definition regarding its legal relation to the workers (are workers considered as end users of the platform, subjected to terms of services (ToS?); are they employees or contractors?), as well as the space in which the work is carried out, all show significant variation depending on the platform.

Existing methods employed to approach 'platform work' as a sociological object face several limits that make them poor choices for studying the diversity in the organisational characteristics of platforms.

Targeted qualitative methods like web ethnography are often limited in scale and focus on 'visible workers'. As such, they are subject to selection biases, as discussed above, with workers active in online interaction spaces being more likely to be heard than those who are more involved in the platform. Quantitative methods, while generally more scalable, also face specific difficulties depending on study design. Questionnaire diffusion through the platform interface implies targeting platforms with an interface designed to allow some form of client/worker interaction. Similarly, striking a deal with platforms to gain access to their internal data, while tremendously useful to grasp an internal view of the structure, also means being constrained by the willingness of the platforms to allow researchers to analyse their data. Moreover, internal data shared by the platforms are, by design, not research-focused. The quality of the data, information collected and data formats are defined according to the platforms' internal needs, and not in accordance with research hypotheses.

Focusing on platforms where deploying the presented methods is possible potentially introduces serious bias in the analysis if the results are generalised as descriptive of all platform work. There is evidence that a non-trivial share of what could be considered platform work is undertaken on platforms offering only a business to business interface, as in the case of Appen, for example, described by Tubaro (2021), rendering surveys or workers' interviews unfeasible.

Data scraping, based either on workers' communication tools or on data displayed on the platforms websites, is conceptually closer to the method proposed here. Automated data collection, however, relies on the ability of the researcher to systematise ways to access the said data. Technical tools (websites, forums and direct messaging tools) are not standardised and change widely depending on the platform. This makes large-scale automated data scraping on numerous platforms tedious. Researchers using data scraping tools thus usually focus their analysis on selected cases.

We propose in this article a series of methodological steps that make it possible to describe, rationalise, classify and study the impact of a range of characteristics. We assert that studying the aggregation of organisational variables on a broad level allows us to draw useful conclusions about working conditions. Such rationalisation allows us

to go beyond a unified view of DLPs and to study the different roles they play in framing the so-called ‘future of work’.

In both of the projects presented here, we faced the issue of a lack of available, representative and structured data describing the organisational characteristics of DLPs. The lack of any extensive list of DLPs prevented us from using random draws to build a representative sample. Datasets had to be hand-tailored by the researchers using publicly available data (mainly extracted from the platform websites or second-hand sources) obtained through search engine queries.

To mitigate the limits of representativeness, we chose to systematise the study of platforms’ organisational characteristics using multiple component analyses. We then successfully partitioned our DLP datasets in usable data-driven typologies. It should be noted still that using non-inferential statistical methods is not a silver bullet to avoid bias when analysing non-random samples. Even if here we do not intend to present a meaningful descriptive count of our DLP typology, under-representation of certain organisational traits under a certain count threshold can still threaten the quality of the analysis. MCA is subject to the same restrictions as chi-square tests: categories presenting too low a count can affect the reliability of the analysis. As is also traditionally the case, one can never fully assume that all of the existing category combinations are observable using the chosen database constitution approach, and have been effectively observed.

The best way to mitigate both limitations is, in our opinion, along with basic statistical data health checks (especially useful to detect low count potential issues), to supplement the quantitative analysis with more qualitative methods.

Findings in both of the presented projects illustrate the way in which mixed methods can be complementary and extremely useful to avoid partial or even entirely mistaken interpretation of seemingly self-sustaining statistical results.

For instance, for SWIRL, the MCA second axis, as well as the cluster classification, tended to suggest a division along ‘power and work management’ lines. However, detailed cases studies carried out on platforms selected from the clusters shed a different light on this interpretation. Using extracts from workers’ forums and other secondary sources (mainly press interviews), it appeared that what we first identified as a ‘power divide’ was in fact caused by the workers’ ability to directly interact with clients, and to retain an ‘out of platform client network’, therefore lessening their dependencies on a single DLP. Our MCA dimensions therefore described an opposition between platforms that structurally allowed workers to do this and platforms that entirely controlled the client–worker relationship. On HUSH, the study of AI at the industry level allowed us to show that the evolution of AI business uses pushes companies to consider other outsourcing models than using DLPs, with the result that different types of workforces coexisted with various degrees of integration within the companies that composed our typology.

Conclusion

The diversity of DLPs is a barrier to the appreciation of working conditions within organisations which supposedly share common features. The pervasiveness of these firms across the economy, enabling client firms on the one hand to access a variety

of occupations through their services and on the other hand to use AI services powered by data workers, makes it essential to understand the place of these firms and their workforces in contemporary value chains. As a whole, in both projects, our methods allowed us to map the diversity of platforms in relation to specific research topics. One of the strengths of this method is to broaden the existing attempts at general platform typology-building observed in the literature on platforms (International Labour Organisation, 2021) and the literature on AI extractive activities (Crawford & Joler, 2018) by offering a way to build more operationalisable hypothesis-based, data-driven typologies.

In the SWIRL project, the combination of case studies and MCA allowed us to approximate a 'precarity scale' in which working conditions could be seen as dependent on the organisational characteristics of platforms. In the HUSH project, our mixed-methods approach, extended by semi-structured interviews with workers, allowed us to identify the variety of value chain organisation within the annotation industry and *in fine* the variety of contractualisation and working conditions. In both cases, the approach we adopted allowed us to better understand the interactions between the organisation of production and working conditions at an industry level. In this article, we have underlined the fact that platforms do not constitute a distinct, unified and homogeneous part of the service industry.

In many ways, variables differentiating the types of work and conditions offered on DLPs are reminiscent of existing divides observed in traditional forms of employment. The data-driven typologies presented here enabled us to deepen the study of variations in characteristics across Digital Labour Platforms.

Here, our findings have allowed us to open several leads toward the study of platform embeddedness in already-established industrial sectors, and their relationship with other more traditional companies operating in their field.

Such work could have multiple implications for a better understanding of the integration of DLPs into their respective industrial sectors. This would, for instance, permit policy-makers and regulators to strengthen workers' protection by subjecting platforms to the same obligations that are faced by other companies operating in the same industrial sector.

© Maxime Cornet, Clément Le Ludec, Elinor Wahal and Mandie Joulin, 2022

REFERENCES

- Adams-Prassl, J. & M. Risak (2016) 'Uber, Taskrabbit, & Co.: Platforms as employers? Rethinking the legal analysis of crowdwork' (SSRN Scholarly Paper ID 2733003), Social Science Research Network. <https://papers.ssrn.com/abstract=2733003>
- Akgüç, M., M. Beblavý, E. Cirule & Z. Kilhoffer (2018). *Industrial Relations and Social Dialogue in the Age of Collaborative Economy (IRSDACE)*. National Report France, 21.
- Aloisi, A. & V. De Stefano (2018) *European Legal Framework for 'Digital Labour Platforms'*, Brussels: European Commission and Joint Research Centre. http://publications.europa.eu/publication/manifestation_identifier/PUB_KJ0418843ENN
- Alvarez, R., M. Becue & J. J. Lanero (2000) 'Le vocabulaire gouvernemental espagnol (1979–1996)', *Mots*, 62:30–64.
- Barton, D. C. (2020) 'Impacts of the COVID-19 pandemic on field instruction and remote teaching alternatives: Results from a survey of instructors', *Ecology and Evolution*, 10 (22):12499–12507.

- Benzécri, J. P. (1977) 'Histoire et préhistoire de l'analyse des données. Partie V L'analyse des correspondances', *Les cahiers de l'analyse des données*, 1 (2):9–40.
- Berg, J. (2015) 'Income security in the on-demand economy: Findings and policy lessons from a survey of crowdworkers', *Comparative Labour Law & Policy Journal*, 37:543.
- Berg, J., M. Furrer, E. Harmon, U. Rani & M. S. Silberman (2018) *Digital Labour Platforms and the Future of Work: Towards Decent Work in the Online World*, Geneva: ILO.
- Bourdieu, P. (1999) 'Une révolution conservatrice dans l'édition', *Actes de la Recherche en Sciences Sociales*, 130:4–28.
- Borchert, K., M. Hirth, M. E. Kummer, U. Laitenberger, O. Slivko & S. Viète (2018) 'Unemployment and online labor', *ZEW-Centre for European Economic Research Discussion Paper*, 18-023.
- Casilli, A. A. (2019) *En attendant les robots: Enquête sur le travail du clic*, Paris: Le Seuil.
- Casilli, A., P. Tubaro, C. L. Ludec, M. Coville, M. Besenval, T. Mouhtare & E. Wahal (n.d.) *Le Micro-Travail en France: Derrière l'automatisation, de nouvelles précarités au travail?*, [Research report], DipLAB Research Project.
- Cherry, M. A. (2010) 'A taxonomy of virtual work' (SSRN Scholarly Paper ID 1649055), Social Science Research Network. <https://papers.ssrn.com/abstract=1649055>
- Christie, N. & H. Ward (2018) *The Emerging Issues for Management of Occupational Road Risk in a Changing Economy: A Survey of Gig Economy Drivers, Riders and their Managers*, London: UCL Centre for Transport Studies.
- Crawford, K., & Joler, V. (2018) 'Anatomy of an AI system'. Retrieved 18 September 2018.
- Daugareilh, I., C. Degryse & P. Pochet (2019) 'The platform economy and social law: Key issues in comparative perspective', *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3432441>
- De Groen, W. P., Z. Kilhoffer, K. Lenaerts & I. Mandl (2018) *Employment and Working Conditions of Selected Types of Platform Work*, Luxembourg: Publications Office of the European Union.
- De Stefano, V. (2016) 'Introduction: Crowdsourcing, the gig-economy and the law', *Comparative Labor Law & Policy Journal*, 37 (3).
- Drahokoupil, J., & B. Fabo (2016) 'The platform economy and the disruption of the employment relationship', *ETUI Research Paper-Policy Brief*, 5.
- Eurofound (2018) *Employment and Working Conditions of Selected Types of Platform Work*, Luxembourg: Publications Office of the European Union.
- Forde, C., M. Stuart, S. Joyce, L. Oliver, D. Valizade, G. Alberti, K. Hardy, V. Trappmann, C. Umney & C. Carson (2017) *The Social Protection of Workers in the Collaborative Economy*, Brussels: European Parliament.
- Freeman, C. (2000) *High Tech and High Heels in the Global Economy: Women, Work, and Pink-Collar Identities in the Caribbean*, Durham, NC: Duke University Press.
- Frey, C. B. & M. A. Osborne (2017) 'The future of employment: How susceptible are jobs to computerisation?', *Technological Forecasting and Social Change*, 114:254–280. <https://doi.org/10.1016/j.techfore.2016.08.019>
- Fuchs, C. (2012) 'Dallas Smythe today – the audience commodity, the digital labour debate, Marxist political economy and critical theory', *TripleC: Communication, Capitalism & Critique. Open Access Journal for a Global Sustainable Information Society*, 10 (2):692–740. <https://doi.org/10.31269/triplec.v10i2.443>
- Gardey, D. (2001) *La dactylographe et l'expéditionnaire: Histoire des employés de bureau, 1890–1930*, Paris: Belin.
- Graham, M. & M. Anwar (2019) 'The global gig economy: Towards a planetary labour market?', *First Monday*, 24 (4).
- Graham, M., I. Hjorth & V. Lehdonvirta (2017) 'Digital labour and development : Impacts of global digital labour platforms and the gig economy on worker livelihoods', *Transfer: European Review of Labour and Research*, 23 (2):135–162. <https://doi.org/10.1177/1024258916687250>
- Gray, M. L. & S. Suri (2019) *Ghost Work : How to Stop Silicon Valley from Building a New Global Underclass*, Boston, MA: Houghton Mifflin Harcourt.

- Gurstein, P. (2001) *Wired to the World, Chained to the Home: Telework in Daily Life*, Vancouver: University of British Columbia Press and Eurospan.
- Hamel, J. (1993) 'The case study in sociology: The contribution of methodological research in the French language', *Canadian Review of Sociology/Revue Canadienne de Sociologie*, 30 (4):488–509. <https://doi.org/10.1111/j.1755-618X.1993.tb00651.x>
- Hamel, J., S. Dufour & D. Fortin (1993) *Case Study Methods*, Thousand Oaks, CA: Sage Publications.
- Hauben, H., K. Lenaerts & W. Wayaert (2020). *The Platform Economy and Precarious Work. Study*. Luxembourg: European Union.
- Hicks, M. (2017) *Programmed Inequality: How Britain Discarded Women Technologists and Lost its Edge in Computing*, Cambridge, MA: MIT Press.
- Holts, K. (2013) 'Towards a taxonomy of virtual work', *Work Organisation, Labour & Globalisation*, 7 (1): 31–50. <https://doi.org/10.13169/workorgalaboglob.7.1.0031>
- Husson, F. & J. Josse (2013) 'Multiple correspondence analysis' in J. Blasius & M. Greenacre (eds) *Visualization and Verbalization of Data*, Boca Raton, FL: CRC Press:205–223.
- Huws, U. (2007) 'The emergence of EMERGENCE: The challenge of designing research on the new international division of labour', *Work Organisation, Labour & Globalisation*, 1 (2):20–35.
- Huws, U. (2017) 'Where did online platforms come from? The virtualization of work organization and the new policy challenges it raises', in P. Meil & V. Kirov (eds) *Policy Implications of Virtual Work*, Cham: Springer International Publishing:29–48. https://doi.org/10.1007/978-3-319-52057-5_2
- Huws, U., N. Jagger & S. O'Regan (1999) *Teleworking and Globalisation: Towards a Methodology for Mapping and Measuring the Emerging Global Division of Labour in the Information Economy*, Grantham: Grantham Book Services, Ltd.
- Huws, U., W. B. Korte & S. Robinson (1993) *Telework: Towards the Elusive Office* (reprinted), Chichester: Wiley.
- Huws, U., N. Spencer, M. Coates, D. Sverre Syrdal & K. Holts (2019) 'The platformisation of work in Europe: Results from research in 13 European countries'. <https://doi.org/10.18745/ds.21600>
- International Labour Organisation (2021) *The Role of Digital Labour Platforms in Transforming the World of Work* [Report]. http://www.ilo.org/global/research/global-reports/weso/2021/WCMS_771749/lang--en/index.htm
- Johnston, H. & C. Land-Kazlauskas (2018) *Organizing On-Demand: Representation, Voice, and Collective Bargaining in the Gig Economy*, Conditions of Work and Employment Series, 94, Geneva; ILO.
- Josse, J. & F. Husson (2016) 'MissMDA: A package for handling missing values in multivariate data analysis', *Journal of Statistical Software*, 70 (1):1–31.
- Kässi, O. & V. Lehdonvirta (2018) 'Do digital skill certificates help new workers enter the market? Evidence from an online labour platform' (No. 7810), CESifo Working Paper.
- Kilhoffer, Z., K. Lenaerts & M. Beblavý (2017) *The Platform Economy and Industrial Relations: Applying the Old Framework to the New Reality*.
- Kilhoffer, Z., W. P. De Groen, K. Lenaerts, I. Smits, H. Hauben, W. Waeyaert, E. Giacumacatos, J.-P. Lhernould & S. Robin-Olivier (2020). *Study to Gather Evidence on the Working Conditions of Platform Workers*, Brussels: European Commission. <https://doi.org/10.2767/26582>
- Krause, P., O. Szekely, M. Bloom, F. Christia, S. Z. Daly, C. Lawson ... & A. Zakayo (2021) 'COVID-19 and fieldwork: Challenges and solutions', *Political Science & Politics*, 1–6.
- Lê, S., J. Josse & F. Husson (2008) 'FactoMineR: An R package for multivariate analysis', *Journal of Statistical Software*, 25 (1):1–18.
- Lebart, L. & G. Saporta (2014) 'Historical elements of correspondence analysis and multiple correspondence analysis', *Visualization and Verbalization of Data*, 31–44.
- Leimeister, J. M., D. Durward & S. Zogaj (2016) *Crowd Worker in Deutschland Eine empirische Studie zum Arbeitsumfeld auf externen Crowdsourcing-Plattformen*. <https://nbn-resolving.org/urn:nbn:de:101:1-2018072711081159366471>

- Machlup, F. (1962) *The Production and Distribution of Knowledge in the United States* (1st paperback ed.), Princeton, NJ: Princeton University Press.
- Mateescu, A. & A. Nguyen (2019) 'Algorithmic management in the workplace', *Data & Society*, 1–15.
- McDonald, P., P. Williams & R. Mayes (2020) 'Means of control in the organization of digitally intermediated care work', *Work, Employment and Society*, 0950017020969107. <https://doi.org/10.1177/0950017020969107>
- Meil, P. (2018) 'Spinning the web: The contradictions of researching and regulating digital work and labour' in P. Bilić, J. Primorac & B. Valtýsson (eds) *Technologies of Labour and the Politics of Contradiction*, Cham: Springer International Publishing: 271–290. https://doi.org/10.1007/978-3-319-76279-1_15
- Mitter, S. & R. Pearson (1992) 'Global information processing: The emergence of software services and data entry jobs in selected developing countries', *ILO Working Papers*, 992897713402676. International Labour Organisation. <https://ideas.repec.org/p/ilo/ilowps/992897713402676.html>
- Montgomery, T. & S. Baglioni (2020) 'Defining the gig economy: Platform capitalism and the reinvention of precarious work', *International Journal of Sociology and Social Policy*. <https://doi.org/10.1108/IJSSP-08-2020-0400>
- Ojanperä S., M. Graham, R. Straumann, S. De Sabbata & M. Zook (2017) 'Engagement in the knowledge economy: Regional patterns of content creation with a focus on Sub-Saharan Africa', *Information Technologies and International Development*, 13:33–51.
- Pettinger, L. (2019) *What's Wrong with Work?* Bristol: Policy Press.
- Piasna, A. & J. Drahokoupil (2019) 'Digital labour in central and eastern Europe: Evidence from the ETUI Internet and Platform Work Survey', *ETUI Research Paper-Working Paper*.
- Posthuma, A. (1987) *The Internationalisation of Clerical Work: A Study of Offshore Office Services in the Caribbean*, Brighton: Science Policy Research Unit, University of Sussex.
- Prassl, J. (2018) *Humans as a Service: The Promise and Perils of Work in the Gig Economy* (1st ed.), Oxford: Oxford University Press.
- Price, J., R. A. Straus & J. R. Breese (eds) (2009) *Doing Sociology: Case Studies in Sociological Practice*, Lanham, MD: Lexington Books.
- Ragin, C. C. (2014) *The Comparative Method*, Berkeley, CA: University of California Press.
- Ragin, C. C. & H. S. Becker (eds) (1992) *What is a Case? Exploring the Foundations of Social Inquiry*, Cambridge: Cambridge University Press.
- Scholz, T. (2012) *Digital Labor: The Internet as Playground and Factory*, London: Routledge.
- Sismondo, S. (2010) *An Introduction to Science and Technology Studies* (2nd ed.), Chichester: Wiley-Blackwell.
- Soares, A. S. (1991) 'Work organization in Brazilian data processing centres: Consent and resistance', *Labour, Capital and Society / Travail, capital et société*, 24 (2):154–183.
- Thompson, P. & C. Smith (2010) *Working Life: Renewing Labour Process Analysis*, Basingstoke: Palgrave Macmillan.
- Tassinari, A. & V. Maccarrone (2020) 'Riders on the storm: Workplace solidarity among gig economy couriers in Italy and the UK', *Work, Employment and Society*, 34 (1):35–54.
- Tubaro, P. (2021) 'Disembedded or deeply embedded? A multi-level network analysis of online labour platforms', *Sociology*, 55 (5).
- Tubaro, P., A. A. Casilli & M. Coville (2020) 'The trainer, the verifier, the imitator: Three ways in which human platform workers support artificial intelligence', *Big Data & Society*, 7 (1).
- Wang, Y. (2018) 'Building an honest and capable crowd workforce', *SIGMIS-CPR: Proc. ACM SIGMIS Conf. Comput. People Res.*, 111-18. Scopus. <https://doi.org/10.1145/3209626.3209713>
- Weil, D. (2014) *The Fissured Workplace*, Cambridge, MA: Harvard University Press.
- Wood, A. J., M. Graham, V. Lehtonvirta & I. Hjorth (2019) 'Networked but commodified: The (dis) embeddedness of digital labour in the gig economy', *Sociology*, 53 (5):931–950.
- Yin, R. K. (2018) *Case Study Research and Applications: Design and Methods* (6th ed.), Thousand Oaks, CA: SAGE.
- Yin, M., M. L. Gray, S. Suri & J. W. Vaughan (2016) 'The communication network within the crowd', *Proceedings of the 25th International Conference on World Wide Web*:1293–1303.

ACKNOWLEDGMENTS

The SWIRL project is funded by DG Employment, Social Affairs and Inclusion - EMPL.A – Employment and Social Governance “Improving expertise in the field of industrial relations - VP/2018/004” Agreement number: VS/2019/0076. The authors also acknowledge the support of the French Agence Nationale de la Recherche (ANR), under grant ANR-19-CE10-0012 for their funding contribution to the HUSH project. We also thank Antonio Casilli, Paola Tubaro and Ulrich Laitenberger as well as all the members of the ‘DIPLab’ research project for their valuable comments on this article. Finally, we thank the anonymous reviewers who also helped us improve this article.