

The platform belongs to those who work on it! Co-designing worker-centric task distribution models

David Rozas
drozas@ucm.es
Universidad Complutense de Madrid
Madrid, Spain

Jorge Saldivar
jasaldivar@ucm.es
Universidad Complutense de Madrid
Madrid, Spain

Eve Zelickson
eve@datasociety.net
Data & Society
New York City, United States of
America

ABSTRACT

Today, digital platforms are increasingly mediating our day-to-day work and crowdsourced forms of labour are progressively gaining importance (e.g. Amazon Mechanical Turk, Universal Human Relevance System, TaskRabbit). In many popular cases of crowdsourcing, a volatile, diverse, and globally distributed crowd of workers compete among themselves to find their next paid task. The logic behind the allocation of these tasks typically operates on a “First-Come, First-Served” basis. This logic generates a competitive dynamic in which workers are constantly forced to check for new tasks.

This article draws on findings from ongoing collaborative research in which we co-design, with crowdsourcing workers, three alternative models of task allocation beyond “First-Come, First-Served”, namely (1) round-robin, (2) reputation-based, and (3) content-based. We argue that these models could create fairer and more collaborative forms of crowd labour.

We draw on Amara On Demand, a remuneration-based crowdsourcing platform for video subtitling and translation, as the case study for this research. Using a multi-modal qualitative approach that combines data from 10 months of participant observation, 25 semi-structured interviews, two focus groups, and documentary analysis, we observed and co-designed alternative forms of task allocation in Amara on Demand. The identified models help envision alternatives towards more worker-centric crowdsourcing platforms, understanding that platforms depend on their workers, and thus ultimately they should hold power within them.

CCS CONCEPTS

• **Human-centered computing** → *Collaborative and social computing*; **Empirical studies in collaborative and social computing**;

KEYWORDS

crowdsourcing, digital labour, distribution of value, future of work, human computation, platform economy, task allocation, worker-centric platforms

ACM Reference Format:

David Rozas, Jorge Saldivar, and Eve Zelickson. 2021. The platform belongs to those who work on it! Co-designing worker-centric task distribution

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

OpenSym 2021, September 15–17, 2021, Online, Spain

© 2021 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-8500-8/21/09.

<https://doi.org/10.1145/3479986.3479987>

models. In *17th International Symposium on Open Collaboration (OpenSym 2021)*, September 15–17, 2021, Online, Spain. ACM, New York, NY, USA, 12 pages. <https://doi.org/10.1145/3479986.3479987>

1 INTRODUCTION

Contemporary working practices are changing and digital platforms are increasingly mediating our day-to-day work. In this scenario, crowdsourced forms of labour are progressively gaining importance, and large corporations such as Amazon and Microsoft are entering the field. Amazon Mechanical Turk (AMT), Universal Human Relevance System (UHRS), and TaskRabbit are all examples of market-driven crowdsourcing platforms. These platforms operate as labour marketplaces for businesses to outsource work to globally distributed and diverse workers [47].

In crowdsourcing platforms, work is “taskified” [5]. Entering receipts into expense reports, curating data to train an Artificial Intelligence (AI) model, translating texts and tagging words and images, are examples of work that can be easily “taskified”. These tasks are carried out by an invisible workforce of “humans in the loop” [18]. Through platforms such as AMT, in this volatile and globally distributed crowd of workers, with varying degrees of expertise and backgrounds [47], people compete against each other to find tasks to work on. The logic behind the allocation of these tasks typically operates on a “First-Come, First-Served” (FCFS) basis [19, 59].

Previous research has argued that FCFS is a convenient method of task allocation because of its simplicity and capacity to decrease task completion time [30]. The approach, however, creates a competitive dynamic in which workers are forced to be constantly alert for new tasks to appear, producing a sense of anxiety and frustration in case they cannot obtain the work [18]. Additionally, FCFS disadvantages workers who do not have access to a reliable Internet connection or those who work in time zones different from the requesters. Thus, it can create inequitable work distribution by relying on circumstances that are often beyond workers’ control.

Alternatives of work distribution have been proposed in crowdsourcing literature to optimize worker-task matching, maximising the task-requesters’ benefits, and improving results (e.g., [10, 23, 31, 58]), yet, besides some empirical studies that examine crowdsourcing issues from the workers’ perspective (e.g., [7, 14, 40]), there is a lack of proposals which aim to improve the working conditions and well-being of workers. This article reports the preliminary results of an interactive design approach where workers have been involved throughout a research process that includes a variety of methods and which aims to identify and validate alternative task allocation logics defined and agreed upon by the workers themselves.

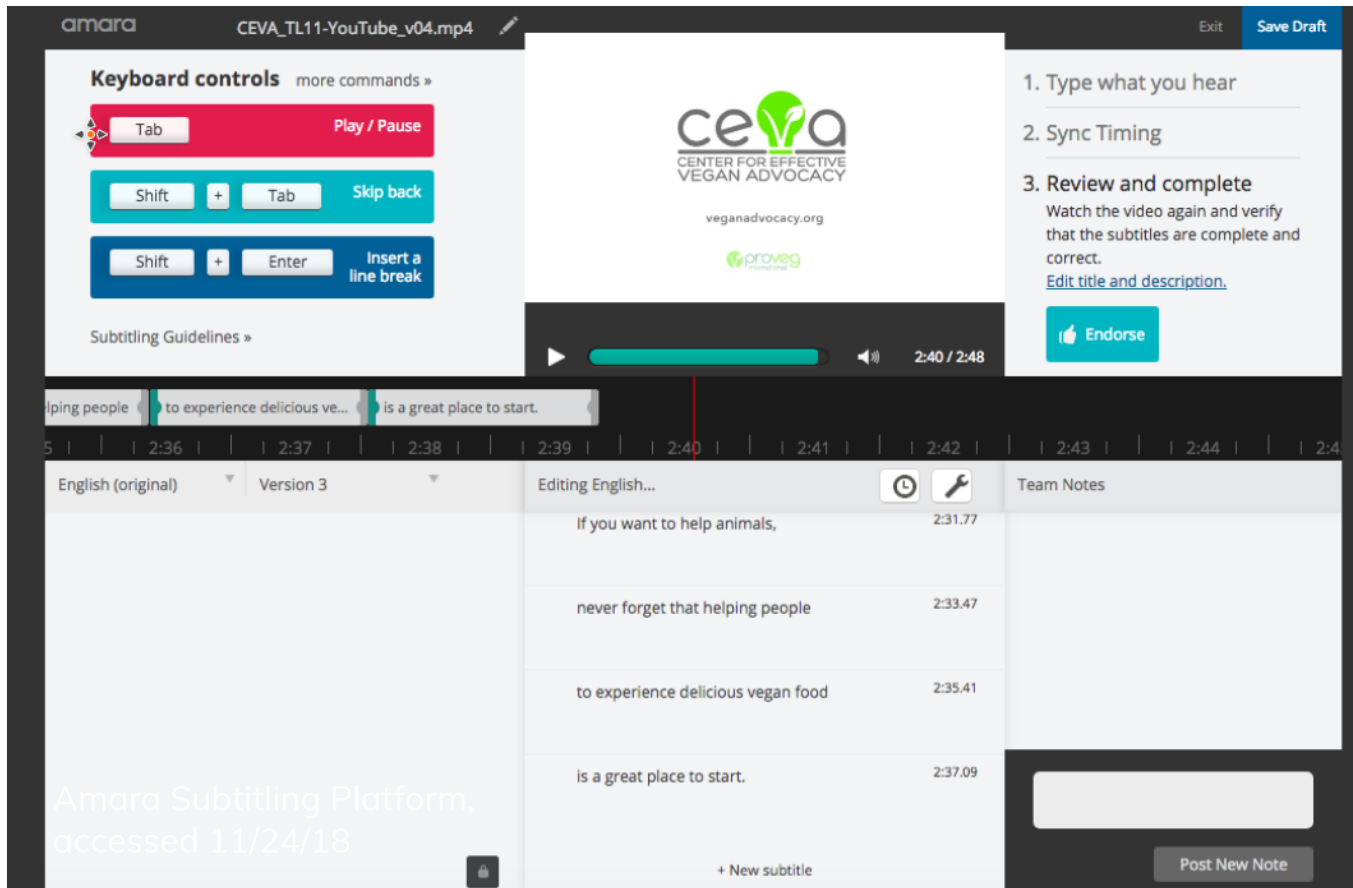


Figure 1: Screenshot of AOD’s subtitling platform, captured on 24th November 2018.

Our vision towards more worker-centric crowdsourcing platforms is summarised by the motto: “*the platform belongs to those who work on it*”¹, which aims at empowering workers to define the rules that govern the distribution of value in crowdsourcing platforms.

The remainder of the paper proceeds as follows. Next, we present our case study, followed by a description of the theoretical concepts that frame the work and a review of related works. Section 5 introduces the methods employed in the study. Later, Section 6 describes the results. A general discussion about the implications of the results is provided in Section 7. We close the paper by presenting conclusions in Section 8.

2 CASE STUDY: AMARA ON DEMAND

Amara is a project which sustains an open and collaborative platform for the creation of subtitles [27]. Examples of organisations employing Amara’s platform to create subtitles drawing on volunteer engagement include Khan Academy, Scientific American, and the California Academy of Science [4]. More specifically, our focus in this research is placed on the use of Amara’s platform for the

creation of subtitles as an on-demand and paid service: Amara On Demand (AOD). AOD was launched in 2013 [60] as a result of the success [27] of Amara’s platform (Figure 1 shows a screenshot of AOD’s subtitling tool). AOD is organised as a non-profit organisation, under the umbrella of the Participatory Culture Foundation. AOD is inspired by cooperative and commoning practices [18], presenting a remarkable contrast when compared with the market-based logic of other crowdsourcing platforms. While AOD grew into its own enterprise within *Amara.org*, it adopted the values of the original volunteer community [3].

Over the past years, AOD moved from a few linguists to more than nine hundred at the time of writing². The work of linguists in AOD is remunerated and they are organised on a per-language direction basis and in which English operates as the master language. For example, if a customer requires a set of videos in German to be subtitled into Spanish, this will involve the groups German->English and English->Spanish. In order to join AOD, linguists are required to submit a resume, two examples of captioned or translated work, and pass an online interview as well as a test. The test is intended to ensure linguists understand AOD’s guidelines, maintaining quality and thus, client satisfaction.

¹We have adapted this motto inspired by Teodoro Flores’s phrase “la tierra es para quien la trabaja” (“the land belongs to those who work it”), which captures the revolutionaries’ vision for land reform in the context of the Mexican Revolution (1910-1920) [41].

²As self-reported by key members of AOD’s core team during the interviews.

An essential part of AOD is the core team that facilitates and oversees the whole production process, coordinating and sustaining the infrastructure required for the successful creation of subtitles and captioning. The core team operates as a central node in AOD, although their members are globally distributed. The core team also monitors linguists' compliance with the rules. In AOD, there are explicit rules, practices, and guidelines to govern participation and foster professionalism. For linguists, this means completing tasks by deadlines, not assigning themselves more than one video "at the same time", and adhering to project-specific rules. Linguists are expected to adhere to them in order to receive payment.

3 THEORETICAL FRAMEWORK

Crowdsourcing has many definitions, but can be captured by the idea of an open call for anyone to participate in an online task [6, 12, 24] by contributing information, knowledge, or skills. The 'crowd' refers to the group of people who participate in the crowdsourcing initiative online. The crowd can, in theory, emerge from anyone online or specific subsets of people. Participation is either voluntary (uncompensated) or for money (financially incentivised). An instance of voluntary crowdsourcing can be found in crowd-sourced journalism [1] or crowdsourcing in crisis management [52]. In paid crowdsourcing, participants are compensated per task, as in microtasking on digital labour market-places such as Amazon Mechanical Turk (AMT) [53] or based on performance as in innovation challenges [28].

For this research, we frame our case study as part of the aforementioned broader phenomenon of crowdsourcing. More specifically, we draw on Hansson et al.'s [20, 21] categorisation of the different modes of production in crowdsourcing platforms to frame AOD as a case of *human computing*. *Human computing* crowdsourcing platforms, such as Amara and AMT, are those in which "users do micro-tasks that do not require much expertise, such as transcribing audio and video files, translating texts, or tagging maps [...] and in which] individual crowd members usually undertake tasks independently of one another, sometimes even competing for work on this market [...]" [21]. In this study, Hansson et al. [21] draw on Marx's [39] theory of alienation to understand the relationships between participants in crowdsourcing and the role of the platforms employed to mediate in the activities. Marx [39] described four types of relationships: (1) between producer and consumer, (2) between the producer and product, (3) the producers' relationship to themselves, and (4) their relationships to other producers. Applying Marx's theory of alienation to crowdsourcing, Hansson et al. [21] developed a typology of alienation that reveals significant differences between the cases studied. Figure 2, adapted from their work, depicts the cases of Amara and AMT. For example, with regards to the relationship between an individual producer with the rest of the producers, the position of Amara being closer towards the inner circle means there are stronger bonds between producers (linguists, in the case of Amara). For the case of AMT, which is in the fourth outer circle, this position represents a lack of bonds between producers. This typology is not to be understood as mutually exclusive: these concepts and different modes sometimes co-exist within the same platforms and processes. However, this typology is "useful as a way to discuss how participation in crowdsourcing is

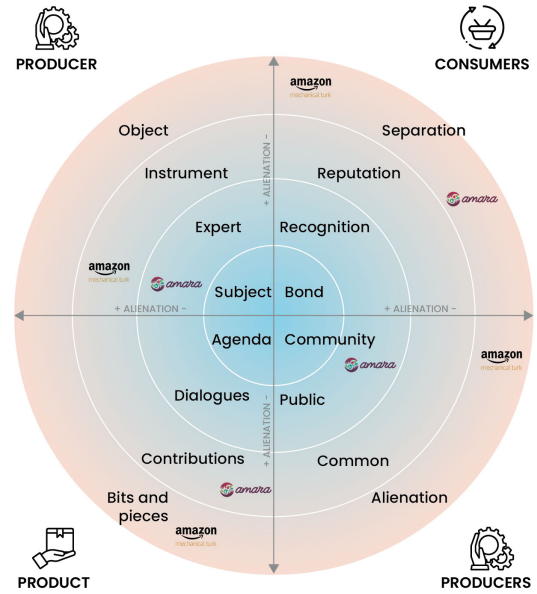


Figure 2: A graphic representation of Hansson et al.'s [21] typology of alienation, according to Marx's [39] four types of relationship. The further from the centre, the higher the degree of alienation. We have adapted Hansson et al.'s [21] Figures 1 and 2 in order to merge the categories and the position of the two key cases (from the 21 studied by them) which we employ to establish comparisons: Amara (our case study) and Amazon Mechanical Turk. See Table 4 on [21] for further details and a summary of the relationships with corresponding modes of productions and the categories employed.

motivated and to develop tools with a better awareness of different types of relationships and how these modes of productions produce different types of knowledge".

Drawing on Hansson et al.' typology [21] and to further our understanding of how crowdsourcing platforms might support social relationships in these contexts, instead of merely capitalising them [21], we decided to explore the following research question: *can we identify alternative models for the distribution of tasks in crowdsourcing that consider the needs of the workers?*

To this aim, we establish a collaboration with Amara, whose strong cooperative values [18] offer an opportunity to design models of task distribution in which the producer is also the owner of the means of production and the products created are an expression of self-realisation [21]. Since, following the concepts from Hansson et al.' typology [21] depicted in Figure 2, Amara contrasts with platforms such as AMT [18], which understand workers as "instruments" from which "bits and pieces" can be sourced [21].

4 RELATED WORKS

Various alternatives have been proposed in the literature to allocate crowdsourcing tasks [8, 10, 16, 22, 29, 30, 35, 36, 59, 61] as well as to re-think the conditions of crowdsourced labour more generally [17, 32]. Some authors have suggested implementing the power of AI techniques to assign tasks to workers, while other scholars have introduced reputation schemes to delegate tasks. Workers' background, expertise, and social connections have also been considered in approaches presented to improve the assignment of tasks in crowdsourcing platforms.

In Machado et al. [35], the authors suggest using AI planning to help choose the best delegation strategy based on parameters that configure the crowdsourcing environment, such as task duration and workers' skills. A machine learning-based approach that combines supervised learning with reinforcement learning to infer task allocation strategies that best fit the available rewards and workers' reputation is proposed by Cui and colleagues [8]. A content-based recommendation method is introduced in Mao et al. [36] to match crowdsourced development tasks to developers automatically. The system learns from historical activities to favour appropriate workers. Ho et al. [22] present an algorithm to allocate tasks in situations of heterogeneous tasks or a diverse, skilled workforce.

Difallah and colleagues [10] employ information available in the workers' social network profiles, such as their interests, to automatically assign workers to tasks aligned to them. The matching between workers and tasks is based on a taxonomy derived from categories extracted from workers' interests and descriptions of tasks. Likewise, the construction of workers' profiles using historical data of their performance and data extracted from social networks is suggested in Kamel et al. [30]. With these data, the authors propose the development of a machine learning model to recommend relevant tasks to workers based on their built profiles. In a similar way, Zhao et al. [61] discuss a model that considers the relationship between workers to assign tasks in crowdsourcing. The proposal is to use social networking sites to learn about the social connections between workers and therefore allocate them and their friends the same or similar tasks.

The allocation of tasks to groups or teams of workers instead of individuals is explored in [29]. In group-oriented crowdsourcing, members of naturally existing groups of workers cooperate to perform tasks. Jiang et al. introduce, in this article, the concept of contextual crowdsourcing value, which determines the priority of a group of workers being allocated a task. The contextual crowdsourcing value measures the capacity of the group of workers to complete a given task in coordination with other groups that complement the missing skills of the group's members.

Reputation models for task allocation have been studied by [59]. Here, workers' reputation is estimated based on the workers' past performance, considering the quality of previous work and meeting deadlines. Increasing fairness while reducing costs is proposed by Fu and Liu [16] who introduce a task allocation model (F-Aware) to create fairer crowdsourcing workflows. The proposed approach monitors the execution of workflows, adjusting the operation of the allocation algorithm to achieve a fairer distribution of labour among workers.

Our work contributes a novel perspective of task allocation on crowdsourcing platforms. Instead of proposing an approach that targets cost reduction, budget balance, quality assurance, or timely completion optimisation, as in the reviewed literature, we report on alternative models that have the potential to help allocate tasks in a fairer way drawing on co-designing techniques which allow workers themselves to define task allocation models that improve their welfare in the platform. Previous research has also explored collaborations with workers of the platforms to explore alternatives to change the nature of crowdsourcing work. In response to concerns from AMT workers over a lack of employer accountability, Irani et al. [25] developed "Turkopticon." Turkopticon is a platform and browser extension where workers can share experiences about employers, allowing for greater transparency and communication among workers [25]. As part of the tool, Turkopticon reveals workers' views of their task lists with information others have written about employers. AMT workers have also employed generic platforms, such as Reddit, to share advice and experiences of working on AMT [38, 62]. Additionally, researchers in collaboration with AMT workers created a platform called Dynamo to support collective action [51].

However, our study differs from theirs in co-designing directly with AOD workers after establishing a collaboration with the core team that controls and sustains the platform and its code. As discussed in Section 3, AMT and AOD represent different forms of crowdsourcing platforms, showing us how platforms can increase the alienation of workers, but they can also help to reduce it [21]. In this sense, the core team of AOD is willing to experiment with alternative logics that provide more power to the workers themselves and integrate them into the primary platform. In the aforementioned studies [25, 51], on AMT workers co-designing alternative platforms, the platforms developed represent a form of counter-power, rather than control over the main platform. Next, we provide an overview of the methods employed to follow this co-designing approach.

5 METHODS

This study employs a multi-modal qualitative approach that combines data collected from 25 online and face-to-face (F2F) semi-structured interviews, ten months of participant observation, focus groups, and documentary analysis of 55 documents, mainly internal AOD documents provided to linguists and official blog posts from *blog.amara.org*. Table 1 and Table 2 provides an overview of the main characteristics of the participants with whom we conducted semi-structured interviews and organised focus groups. The table includes their gender, main role³ in AOD, number of years in AOD, location and language groups they belong to (only for linguists), among others.

The collected data were coded following an ethnographic content analysis approach [2], which involved a continuous process of discovery and comparison of key categories emerging from the

³This refers to the main tasks carried out by the participant in AOD. For example, as discussed in Section 2, whether they are part of the core team. The term DQA refers to Designated Quality Assurer. DQAs are responsible for managing large, active projects where clients often request special instructions. Further details are discussed in Section 6.2.

Table 1: The participants' main characteristics.

Participant ID	Gender	Main role in AOD	Main translation languages	Country of residence	Years in AOD	Related research method
P ₁	Male	Linguist	German	Germany	2	Online semi-structured interview
P ₂	Male	Linguist and DQA	Marathi and Hindi	India	3	Online semi-structured interview
P ₃	Female	Linguist	Greek	United States of America	7	Online semi-structured interview
P ₄	Female	Linguist	Russian and English	United States of America	3	Online semi-structured interview
P ₅	Male	Linguist and DQA	Greek	Greece	4	Online semi-structured interview
P ₆	Male	Linguist	Arabic	Egypt	2	Online semi-structured interview
P ₇	Male	Linguist	Dutch and Spanish	Chile	6	Online semi-structured interview
P ₈	Female	Linguist and DQA	Hindi and English	India	2	Online semi-structured interview
P ₉	Female	Linguist	Simplified Chinese, Traditional Chinese and Malay	Malaysia	2	Online semi-structured interview
P ₁₀	Female	Linguist and DQA	English	United States of America	5	Online semi-structured interview
P ₁₁	Female	Linguist	French and Romanian	Brazil	1	Online semi-structured interview
P ₁₂	Female	Linguist	Greek	Netherlands	3	Online semi-structured interview
P ₁₃	Female	Linguist	Russian	Russia	2	Online semi-structured interview
P ₁₄	Male	Linguist	Portuguese-Portugal	Portugal	6	Online semi-structured interview
P ₁₅	Male	Linguist	Polish	Poland	2	Online semi-structured interview
P ₁₆	Male	Linguist	Swedish	Sweden	4	Online semi-structured interview
P ₁₇	Female	Accountant (core)	N/A	United States of America	5	F2F semi-structured interview
P ₁₈	Female	Project Manager (core)	N/A	United States of America	7	F2F semi-structured interview
P ₁₉	Male	Project Leader (core)	N/A	United States of America	8	F2F semi-structured interview
P ₂₀	Female	Project Leader (core)	N/A	United States of America	3	F2F semi-structured interview

Continued on Table 2

data. The various analytical tasks were supported by the Computer-Assisted Qualitative Data Analysis Software NVivo 12.

5.1 Participant observation and interviews

Online participant observation was carried out over six months (October 2018 - March 2019) to engage with the day-to-day practices of AOD linguists: from the recruitment and onboarding processes to the execution of regular tasks, such as captioning. In addition, 17

semi-structured interviews (see P₁ - P₁₆ in Table 1 and P₂₅ in Table 2) were conducted following a purposive sampling [43] intended to gather the diversity of linguists in terms of language group, experience level, and degree of engagement. The data collected provided us with a rich picture of the experiences, needs and vision of the workflow of an AOD linguist. The primary outcomes of this part of the research were the mapping of the workflow of AOD and the identification of an initial set of communitarian needs which

Table 2: The participants' main characteristics (continuation).

Participant ID	Gender	Main role in AOD	Main translation languages	Country of residence	Years in AOD	Related research method
Continued from Table 1						
P ₂₁	Non-binary	Developer (core)	N/A	United States of America	2	F2F semi-structured interview
P ₂₂	Female	Project Manager (core)	N/A	Brazil	6	Online semi-structured interview
P ₂₃	Non-binary	Recruiter (core)	N/A	United States of America	4	Online semi-structured interview
P ₂₄	Female	Customer service (core)	N/A	Spain	6	Online semi-structured interview
P ₂₅	Female	Linguist and DQA	Portuguese-Brazilian, English and Spanish	Spain	5	F2F semi-structured interview and focus group
P ₂₆	Female	Linguist and DQA	Portuguese-Brazilian	Brazil	4	Focus group
P ₂₇	Male	Linguist	Portuguese-Brazilian	Brazil	7	Focus group
P ₂₈	Male	Linguist and DQA	Portuguese-Brazilian	Brazil	5	Focus group
P ₂₉	Female	Linguist and DQA	Portuguese-Brazilian	Brazil	3	Focus group
P ₃₀	Female	Linguist	Portuguese-Brazilian	Brazil	5	Focus group

led us to discover several intervention points as potential areas to experiment with the development of worker-centric tools to support crowdsourced labour.

A similar approach was conducted but this time with core members of AOD. It involved four months of online participant observation (April 2019 - July 2019), eight semi-structured interviews (see P₁₇ - P₂₄ in Table 1 and Table 2) and documentary analysis of materials generated and posted in the official channels of AOD. As well as with the linguists, the semi-structured interviews were conducted following a purposive sampling [43] with key members of AOD's core team considering the diverse roles in AOD, i.e., project managers, developers, members of the finance team, and project leaders, among others. The data analysis carried out here allowed us to further our understanding of the organisational processes of the workflow and the changes experienced in it over time. The aim was to include all of the different perspectives of the actors involved in the platform, to supplement the information gathered from the linguists. More importantly, the analysis of these data led us to select our point of intervention: task allocation. Task allocation is a necessary precursor to working. As a result, task allocation represents a suitable starting point for envisioning more cooperative labour processes.

5.2 Focus Groups

Interviews and participant observations were followed up with an online two-day workshop that included several focus group sessions (organised in June 2020). A call for participation was disseminated through the official AOD channels, including a short survey to show interest in involvement. From all of the linguistic groups in AOD, we chose the Portuguese-Brazilian due to its high degree of organisational complexity. We selected six linguists (see P₂₅ - P₃₀ in Table 2) according to their different degrees of experience,

since we aimed to have a variety of backgrounds. These focus group sessions allowed us, together with the linguistics, to identify alternative models for allocating tasks. The identified models were subsequently validated by the AOD's core team.

5.3 Ethical considerations

The ethical principles described by the European Research Council [13] were followed, as well as the recommendations from the Association of Internet Researchers [37]. Drawing on these guidelines, we constantly reassessed so that the discovery of any new issues resulted in remedial action. These actions include anonymising participants and references to customers in field notes and transcripts, in addition to the use of information sheets and consent forms to participate in the interviews and the focus groups.

6 RESULTS

Next, we describe the series of problems regarding the current logic of task allocation and the main categories surrounding it, which emerged as key from our analysis: (1) first-come, first-served logic, (2) competitiveness, (3) constantly checking for work, and (4) inconsistent workload. Subsequently, we provide an overview of the three alternative models for task allocation identified in this study beyond FCFS.

6.1 Behind the First-Come, First-Served logic

The "First-Come, First-Served" logic embedded in the platform is the main component of task allocation in AOD. This logic creates a competitive dynamic between linguists to assign tasks to themselves. The following quote, from an interview with P₄, depicts the competitive nature associated with FCFS logic:

“What I realised very fast is that the competition is absolutely awful, the competition is huge, you need to learn how to get the work.”

The competitiveness embedded in this logic becomes even more problematic in a global environment, as in the case of AOD. For example, in theory all linguists belonging to the same language group should have the same opportunity to assign themselves a specific task. However, the reality is that some of them might be in time zones that are less convenient concerning the times in which the tasks are usually posted for assignation. Overall, FCFS encompasses a need for workers to check to find more work continuously, an issue identified also in other crowdsourcing projects [18]. For instance, P₄ explained:

“I learnt how to be fast and not sleep with my computer, but [to] wake up with my computer right next to me. There is also a difference in the time zones, [and] I think we are in the worst position. [...] It is competitive even just to grab it [a task]. I need at least 7 hours of sleep. [...] If you really want to get this work, you need to be next to your computer for hours.”

Furthermore, this FCFS logic needs to be understood in an environment in which the workload is typically inconsistent for most of the language groups, as P₁₂ explains:

“Last year, we only had text work like one month, and the rest of the months we had really short videos, like one minute, two minutes, like an advertisement. So that was it: last year it was poor. But the year before, it was a very, very good year. [...] I would like to make a full wage out of Amara, but I don’t have the chance. Maybe later on, if the organisation expands. Because it’s cool that you can work from wherever you are, on your own timing.”

Some linguists also suggested that the competitiveness embedded in the platform’s task allocation method undermines the sense of community. P₇, for example, explained it in the following way:

“In general, I believe we have a very neutral attitude towards each other because... yes, we share the language. But we are also competing to get the jobs [...]. Especially, in the last couple of years, due to the decline in job orders that I have won, I can tell you that it [competition] is growing.”

As we introduce in Section 2, one of the key changes implemented by AOD’s core was to limit to “one at a time” the number of tasks that linguists can assign themselves simultaneously. This change in the logic was a counter-measure to avoid platform vandalism (e.g. it was found that some participants implemented computer scripts to assign themselves tasks as soon as they were published) and as a first attempt to distribute work more equally. Nevertheless, as we have seen, this has not been sufficient to avoid competition between linguists. FCFS is not the most efficient way in terms of productivity either, as the members of the core explained. P₁₇, a core member of Amara and one of the key workers responsible for managing the overall organisational processes in AOD, explains

the need to increase the throughput and reduce the time provided to linguists to fulfil deadlines:

“If I’m a linguist, the way it works now is: I’ll get a caption [...] whether or not it takes me three days to finish the job, I know in advance it’s doable in three days, so I can wait until day three to do it. So I can assign myself on day one and wait until day three and do it. From my perspective and my job [as a manager], I see that as a disadvantage for the company. Because, first of all, the client is going to get it later. [...] also, maybe there was another linguist that could’ve done it on day one. So, in a way, we do [work ‘on-demand’] [...], but not in a way in which Uber is on-demand. [...] And we don’t have as many people.”

Considering all of the problems discussed, we organised a co-designing workshop with AOD linguists that allowed us to incorporate their perspectives into the tools that mediate their day-to-day practices. Next, we discuss the three main models which emerged from this initiative.

6.2 Exploring and identifying alternative models for tasks allocation

We identified three alternative models for task allocation beyond FCFS: (1) round-robin, (2) reputation-based, and (3) content-based.

6.2.1 Round-robin. Round-robin (RR) refers, in computer science, to an algorithm proposed in the context of operating systems [54] to decide how to schedule multiple processes competing simultaneously for CPU time. In RR, computational processes are assigned similar amounts of computing time circularly. It is one of the simplest and most straightforward solutions to avoid *starvation* in-process execution and it is known as one of the fairest scheduling algorithms [55].

Within the context of the focus group, a parallel with RR emerged when discussing the need to split the work equally between linguists, as P₂₈ suggested:

“What I was thinking was a way that all translators could, uh, work on tasks on Amara so that we could split the work equally between translators so that there would be a similar monthly workload for everyone.”

Rather than in the form of a “pure RR”, the model was discussed as a starting point that could be customised according to the context to find alternatives in which a more balanced assignation of tasks is achieved. A key aspect related to this model was the “pre-assignation of tasks”, as depicted by the following excerpt in which participants P₂₇ and P₂₈ intervene:

“P₂₇: I was thinking about... about (sic) her idea of pre-assignment. Like, (sic) every day the linguists would come, and we’d have their inbox, uh, the tasks for that day. [...] it could be that on a certain day, uh, we wouldn’t have the time needed for that particular task. So it would be necessary to consider, uh, something like the option to accept or not that particular task. [...]”

P₂₈: [...] So the translator could be contacted like, uh, they will be given 24 business hours to respond and take a task [...] And the priority would be for someone who's behind this monthly workload. [...]"

The main advantage discussed by the linguists regarding this model is that it tackles the competitive character discussed in section 6.1, as P₂₈ explained:

"I really liked P₂₆'s idea of the pre-assignment of tasks because this takes away the competitiveness aspect of task allocation. [...] This could also be integrated [...] with a spreadsheet that would rank: this translator has worked on this many minutes this month. So the priority would be for a translator who has not worked that many minutes. So that we could, um, reach a fair amount of work for everyone."

Indeed, we found that some linguists in Amara already use similar informal practices in their day-to-day operating. While most tasks related to translating and captioning are allocated in AOD following a FCFS logic, reviewing (another type of task) escapes this logic in some cases. This alternative allocation occurs particularly for active projects in which clients often request special instructions. These reviewing tasks are carried out by linguists with more experience and selected by Project Managers. This special role in AOD is known as Designated Quality Assurer (DQA⁴). As the quote by P₂₈ below depicts, within this specific scope, linguists themselves employ a similar RR logic to balance the workload between themselves:

"I thought about this because [working as a DQA] I developed a spreadsheet that would sum up all the videos that were available for us to work on, uh, so that we could know which video to allocate to whom like, uh, there's a new video. So if *Emma* was about, I don't know, 20 minutes, uh, shorter than I was, then the video would be given to her. And then the next one would be given to me. And we would find a balance between this workload."

As we shall discuss in Section 7, the challenge of this model lies in identifying the specific parameters to encode in these forms of RR allocation and in providing the linguists with mechanisms that enable them to reach consensus among themselves. Furthermore, the parameters of this model could be combined with those from other models, such as the reputation-based system (presented next). Reflecting on these issues, P₂₇ and P₃₀ explained:

P₂₇: [...] there isn't always a new video to work on. So I think the second part we might have to work on, uh, and I think this could also tie in with my suggestion to split the work equally. So based on the, um, background and the ratings of, uh, translators, they would be pre-assigned to new tasks.[...]
P₃₀: [...] I think attention to deadline is important, and it doesn't matter if you send like a spreadsheet, tell me how many hours, can you work [referring to the calendar idea]. They put like a thousand hours, and

then in the day-by-day, you see that they can't, uh, comply to that."

Next, we discuss the model of allocation by reputation to which the linguists refer.

6.2.2 Reputation-based. Reputation systems have been proposed to build trust among Internet users. They are based on collected and aggregated feedback about users' past feedback and help to foster trustworthy behaviours, assess credibility, and discourage dishonest participation [45]. Usually in crowdsourcing platforms, e-commerce websites, and Q&A forums, feedback on users' actions is instrumented through textual comments, numerical rating scores such as one-to-five scales, and boolean evaluations (e.g., yes/no, like/dislike) [46]. Once built, users' reputation is represented as badges, stars, points, or average scores attached to their screen names [33, 44, 57].

In the context of AOD, this category emerged as a model in which tasks are offered to linguists according to the quality of their previous work, based on feedback received by their peers, and depending on the characteristics of the tasks themselves. During the workshop, this model emerged as a *points system*. P₂₇, for example, proposed the following idea for a model:

"[...] a system that would be able to rank productivity of linguists [...], something like a points system in which the points would be earned, um, with reference to the volume that was processed before, and also based on the quality of previous work [...]"

The main problem tackled by this model, according to the linguists, is that it would help to increase the level of transparency within AOD. Currently, in AOD, several levels operate according to the linguist's experience. The transition between these levels, however, currently lacks clarity and explicit mechanisms, as several linguists pointed out. The following excerpt, from an interview with P₂, illustrates this:

"[...] About the transition in levels. Translators should know how to reach next levels from the current levels. What's the criteria for that? [...] [we need] transparency on how to move up or just [some] guidelines."

Indeed, as with the previous case of informal practices in RR, the transition between the different levels already operates, although it does so without explicit parameters, as P₃₀ explained:

"[...] We had knowledge of the previous work that was done by certain linguists that had displayed more quality, more commitment. Uh, they have processed more volume. So [they are promoted] based on all this, but [it was] not, not (sic) quantified, so it was more like a qualitative selection."

The challenge of a reputation-based model, as that proposed by the linguists, is arriving at agreements on what to consider within the system. However, within the focus group, linguists found a preliminary consensus regarding its application in the context of the current *AOD levels system*. The system proposed was based on tiers, as the excerpt below by P₂₈ illustrates:

"[...] building a tier system based on the [amount of] minutes of videos that translators have worked on. So, um, that could be, for example, three tiers: novice,

⁴As discussed in Section 5, DQAs are exclusively responsible for the whole reviewing process for large projects. This contrasts with the usual workflow, in which multiple AOD members review videos within a project following a FCFS logic.

intermediate, and veteran. So the novice [translators] would get shorter non-technical videos, and veteran translators would be given the opportunity to work on longer technical videos. And the intermediate [level] would be a balance between the two.[...]"

The number of possibilities is such that we concluded that specific sessions would be required to fully explore the parameters of this reputation-based model. However, we identified two key characteristics to be incorporated into the design. Firstly, linguists agreed on not only including the final quality of the translation into the system through reviews made by their peers, but the system should also consider if deadlines were met and how well the changes suggested by the reviewer were implemented. The following excerpt by P₃₀, illustrates this:

"[...] I just wanted to add that not only the quality of work should be considered in the scoring system, but also the attention to the deadline. In the past, we had a lot of problems with translators that were good, but they were terrible with deadlines. That actually made us lose some clients.[...]"

Secondly, the system should promote inclusiveness. For example, linguists expressed their concerns regarding the possible barriers which a reputation-based model could generate. The following quote, from a comment by P₂₈, illustrates this concern with regards to the barriers for newcomers:

"[...] My only worry is that we should be careful not to exclude newcomers. Uh, for example, if, if (sic) a new person received the low rating, uh, we need to ensure this doesn't compromise how much work they get, because this could be, this could turn into a vicious, vicious (sic) circle as the ones who need to practice the most would not be given enough work to improve on. Uh, so we need to be careful about that. [...]"

Furthermore, linguists envisioned and proposed ways to tackle such challenges in a reputation-based model. For example, they suggested that linguists' degree of experience should be considered to facilitate the allocation of simpler tasks to newcomers to tackle this type of barrier. Other proposals suggested considering a fixed number of recent tasks carried out:

"P₃₀: [...] So for the newcomers, we should give them shorter videos and simpler subjects. And for the expert translators, the larger videos, longer videos and high profile projects, and high profile clients. [...]"

P₂₈: [...] I also thought about rating based on a fixed number of tasks. Like the five or ten latest videos would be taken into account in this rating system, so that upon working on new projects, your ranking could also improve like so that you don't get affected by the first videos. [...]"

In sum, a model based on reputation would help to tackle the need to constantly check for work and increase the degree of transparency of promotions within the platform. However, the model poses a myriad of challenges regarding inclusiveness or the generation of different, and perhaps more challenging, forms of competitiveness. As the previous excerpt illustrates, the model could not

be purely based on reputation. It could instead be combined with content-based assignment characteristics, which could help tackle some of these challenges. The next section explores precisely this in our third model: content-based allocation.

6.2.3 Content-based. Recommendation systems are the cornerstones of modern online services. In social media, they are used to suggest publications [42], in e-commerce sites to offer products [34], in video-streaming applications to recommend multimedia materials [9], and in crowdsourcing platforms to suggest tasks to workers, as we saw in Section 4. Content-based is one of the most widely used techniques employed in recommendation systems. It focuses on matching the characteristics of the artefact to recommend (e.g., topic of publication, movie genre or task description) with attributes of users based on their profiles and historical data [26].

For our case study, this model emerged as one in which tasks are pre-assigned according to two different types of possible matching logics: (1) either the linguist's skills and/or personal preferences regarding specific areas of knowledge, and (2) the linguist's previous experience in the platform concerning the complexity and/or size of the task.

The former initially emerged from discussions on how to ensure the quality of translations, as the following excerpt by P₂₆ depicts:

"[...] if people work based on their backgrounds, they're much more used to [the] terminology and that, in the end, increases the quality."

The initial ideas revolved around attempts to match the linguist's skills to the content of the video to be subtitled. For example, some of the participants in the focus group possessed a degree in Law. Therefore, some argued that the model should prioritise them to complete tasks involving videos concerning Law. This proposal, however, did not reach a consensus. As other linguists argued, sometimes they prefer to work on contents that are not part of their official background:

"P₂₇: [...] not working only on what we are already specialised in, but having the chance to learn something new. Sometimes that's even the main motivation: entering a new field, uh, learning a new subject, dealing with a completely different area, different worlds. Like we come from from (sic) Law. And we are working as linguists precisely because we didn't like Law that much (laugh)."

P₃₀: [...] I think not only [the] backgrounds of the translators should be considered, but also [their] interests. You know, people have hobbies, people like some sort of movies and [some] kinds of stuff more than others. So I think we should also be considering not only like, you know, university backgrounds, but also personal interests."

As with the case of the reputation-based models discussed in subsection 6.2.2, we concluded specific focus groups would be required in order to explore the details of this model. Nevertheless, an initial consensus emerged about considering the contents of previous tasks as a possible avenue to explore and implement this type of model:

P₂₇: [...] we could define this background and this specialisation of interests based on the most recent work [...] their most recent work in the previous weeks or the previous months [...]

P₂₅: [...] maybe considering also the size [...] like the five last big videos that the person has worked with would define their, (sic) their interests or their, (sic) [areas of] specialisation.”

The session concluded by discussing logics which could, therefore, be potentially merged with those from reputation-based models, in which the complexity and the size of the task would also be considered when carrying out this content-based assignation, as the excerpt below including an intervention by P₂₆ depicts:

“[...] I think that may be one way to do that. Eh, like new people would get more priority on smaller jobs, and then people with more experience get priority with larger jobs. And then people, as they get experience, they start working on longer videos. [...] we know that if a person has had little experience before, then [it] is a bit complicated. So maybe the way would be prioritising smaller videos to newcomers, so everybody would have a chance to improve and to learn.”

As with the previous cases, content-based models have the potential to be helpful in overcoming some of the problems for workers derived from the predominant FCFS logic, as those described in Section 6.1. The three identified models presented in this section are, however, to be understood as ideal types [56]. These models are not necessarily a description of reality as such, but valuable concepts when employed as methodological tools to systematise and consider facts that enable us to analyse and intervene within this specific social context. Therefore, identifying these models provides us with a helpful starting point to structure and guide our research, enabling the possibility of establishing comparisons between them and identifying specific characteristics that facilitate merging features of one model with the characteristics of another. Next, we discuss our results concerning the previous literature in the area and provide an overview of future avenues for research with the models mentioned above.

7 DISCUSSION

Conceptualising more egalitarian working conditions in crowdsourcing in collaboration with workers is not unique to this study [17, 32]. Nevertheless, our focus on alternate forms of task allocation *within* a cooperative crowdsourcing organisation — as opposed to platforms like AMT — is. Most large crowdsourcing platforms are less cooperative than AOD, making it challenging to alter aspects of the platform. This study illustrates the potential of co-designing hand-in-hand with workers and the significance of task allocation in dictating the nature of work.

Previous research has criticised algorithmic task allocation — typical in ridesharing platforms like Uber and Lyft — for its opacity and control over workers’ profiles, routes, customers, and wages. As reported by Gray and Suri in [18], and confirmed in this research, current task allocation methods used in crowdsourcing platforms,

which are primarily governed by an FCFS logic, generate competitive dynamics that result in an unfair and unequal distribution of labour and a sense of frustration among workers. We also found that the inherent competitiveness imposed by FCFS also harms the relationship between workers, negatively impacting the sense of community.

Our research in collaboration with AOD’s workers allowed us to define alternative methods to FCFS for task allocation. Through a multi-modal methodology that included interviews, participant observation, focus groups and documentary analysis, we identified together with AOD’s workers three models that could allocate tasks considering their needs.

The review of the literature demonstrates that most task allocation approaches are focused on optimising the needs of task requesters. These task-requester-centric solutions are intended to reduce costs [23], maximise matching [35], and increase the quality of results [59]. Fairness in task allocation has been discussed by Fu and Liu [16]. However, their solution focuses on creating fair crowdsourcing workflows, i.e., a logically related series of tasks, by minimising costs. In contrast, our work aims to shift the current FCFS logic of task allocation for models designed by workers to create cooperative and more equitable working conditions.

In this sense, round-robin (RR), one of the identified models, was proposed by the participants to create a more balanced workload. In RR, tasks are pre-assigned to workers in rounds, reducing the competitive dynamics of the current FCFS allocation practice. Although there was consensus among participants that having a model that pre-assigns tasks in rounds could improve their experience in the platform, there are still essential implementation details in question. In the forthcoming focus groups with linguists, we will examine the rules and parameters that will define the operation of the model. For example, how does the availability of workers impact the model? How does the complexity of tasks affect the assignation of labour? How does worker expertise influence the distribution of tasks?

In discussing workers’ expertise, the participants expressed that a reputation-based model might be a valuable complement to RR. In this context, workers’ reputation would be captured through a system that reflects workers’ performance based on historical feedback. Although performance-based reputation systems are already part of crowdsourcing platforms (e.g., Waze, TaskRabbit, AMT), participants see this model as applicable to improve transparency in the promotion mechanisms used in AOD. Here, workers need to build a “career path” by moving through different stages to reach higher levels of responsibility within the organisation. The significant concern with the reputation model, according to participants, is discrimination against newcomers and low-rated workers as well as task concentration by highly-rated workers. Brawley and Pruy have also mentioned the harmful impacts of reputation systems in crowdsourcing platforms in [7]. The participants suggest distributing more straightforward tasks to novices to promote inclusiveness, ensuring that they have assigned tasks when joining the platform. As with the case of RR, there are implementation details to be addressed, and we plan to organise specific focus groups with linguists to explore them. It is still unclear, for example, what aspects of workers’ history should be factored into one’s reputation. Apart from workers’ level and their historical performance, some participants suggested other elements such as attention to

past deadlines, caring work [48] (e.g., welcoming and tutoring newcomers), and consideration of colleagues' feedback. Gamification techniques based on artefacts that represent participants' reputation (e.g., badges, points, ranking) [15] were also discussed as potential features to equip reputation-based models.

A task allocation model that includes features of the tasks (e.g., complexity, size, topic, length) and workers' skills, background, and preferences was also seen by the participants as a suitable complement to reputation systems. Apart from complementing the reputation model, this method has been found advantageous to match workers with their expertise and skills resulting in high-quality tasks. Alternatively, some participants indicated that assigning tasks that are not necessarily related to workers' abilities and background might favour learning opportunities for workers. The parametrisation of this content-based model was also left for the next round of focus groups with the participants. It remains unclear how workers' domain of expertise should be included in the model and how to operationalise workers' preferences and interests.

The results presented in this article cannot, however, be generalised and should be understood within the particular case of AOD and similar crowdsourcing platforms. Moreover, due to our qualitative approach, the results cannot be generalised within AOD, considering other groups of linguists may relay a significantly different context, and therefore other models might be more suitable for them. In order to tackle this, we are currently carrying out a longitudinal quantitative analysis of the relationship between users and their activities, drawing on the data already available on AOD's platform. Drawing on this data, we plan to carry out a similar research process with linguists of other language groups (e.g., English-Japanese, English-Italian, English-Spanish).

Furthermore, we want to explore how to develop tools which allow crowdsourcing workers to decide on the models to use in different contexts. Similarly, these tools to support decision-making could be employed to determine collectively how a model could be parameterised. In doing so, we plan to develop collaborative decision-making tools that leverage the affordances [49, 50] of distributed-ledger technologies, such as blockchains, to allow crowdsourcing workers to prioritise parameters within the models which better suit their needs, as well as to decide collectively between the models themselves. In the context of our case study, a blockchain-based solution might enable linguists belonging to different language groups to self-organise in Decentralised Autonomous Organizations (DAOs) and resolve task allocation mechanisms that satisfy their requirements⁵. Given that the results are part of a work-in-progress research endeavour, whether the identified models help to reduce competitiveness and improve working conditions is to be validated.

8 CONCLUSION

This article reports on an endeavour to engage crowdsourcing workers in a multi-modal user-centred research process to identify alternative models of value distribution in crowdsourcing platforms. Three models have emerged as a result of the process, namely (i) round-robin, (ii) content-based, and (iii) reputation-based. Although

the proposed models have the potential to improve the workers' experience in crowdsourcing platforms by distributing tasks more fairly, implementation details need to be discussed in subsequent research. Our aim is for this line of research to impact beyond this case study to broader areas of the platform economy. Ultimately, our goal is to foster workers' participation in and ownership of the platforms that mediate their work. Similar platforms owned by cooperatives of workers, distributing tasks and value according to the agreements reached by them, could be envisioned for a variety of areas. An example could be a cooperative of taxi drivers whose organisation is mediated by a platform which they control. The platform could distribute and monitor rides and payments according to the rules defined by the workers within their specific context.

In sum, despite the inherent limitations due to the ongoing nature of this research, the models identified throughout the collaboration with AOD show us that it is possible to envision more cooperative models to distribute work, in which the producers progressively become owners of the means of production and the fruits of their labour expressions of self-realisation [21] and, as a result, *platforms might belong to those who work on them*.

ACKNOWLEDGMENTS

This work was partially supported by the project P2P Models (<https://p2pmodels.eu>) funded by the European Research Council ERC-2017-STG (grant no.: 759207) and by the project Chain Community funded by the Spanish Ministry of Science, Innovation and Universities (grant no.: RTI2018-096820-A-100). We thank all the members of the AOD community: their help made this research possible. A special mention goes to all the linguists who participated in the focus group and the interviews. We also thank the anonymous reviewers for their suggestions. We are also grateful to all of our friends and colleagues of P2PModels for their helpful comments, guidance and advice. We are particularly thankful to Elena Martínez-Vicente and Paulo Colombo for their help organising the focus group and their overall work in the case study. Finally, we would like to thank Tabitha Whittall for her help in copyediting and proofreading this article.

REFERENCES

- [1] Tanja Aitamurto. 2015. Motivation factors in crowdsourced journalism: Social impact, social change, and peer learning. *Social Change, and Peer Learning* (October 16, 2015). *International Journal of Communication* 9 (2015), 3523–3543.
- [2] David Altheide. 1987. Reflections: Ethnographic content analysis. *Qualitative Sociology* 10, 1 (1987), 65–77. <https://doi.org/10.1007/BF00988269>
- [3] Amara.org. [n.d.]. *Amara - Award-winning Subtitle Editor and Enterprise Offerings*. <https://amara.org/en/>
- [4] Amara.org. [n.d.]. *Caption, Subtitle and Translate Video*. <https://amara.org/>
- [5] Sihem Amer-Yahia and Senjuti Roy. 2016. Toward worker-centric crowdsourcing. *Bulletin of the Technical Committee on Data Engineering* (2016).
- [6] Daren C Brabham. 2013. *Crowdsourcing*. MIT Press.
- [7] Alice M Brawley and Cynthia LS Pury. 2016. Work experiences on MTurk: Job satisfaction, turnover, and information sharing. *Computers in Human Behavior* 54 (2016), 531–546.
- [8] Lizhen Cui, Xudong Zhao, Lei Liu, Han Yu, and Yuan Miao. 2017. Complex crowdsourcing task allocation strategies employing supervised and reinforcement learning. *International Journal of Crowd Science* (2017).
- [9] James Davidson, Benjamin Liebold, Junning Liu, Palash Nandy, Taylor Van Vleet, Ullas Gargi, Sujoy Gupta, Yu He, Mike Lambert, Blake Livingston, et al. 2010. The YouTube video recommendation system. In *Proceedings of the fourth ACM conference on Recommender systems*. 293–296.
- [10] Djellel Eddine Difallah, Gianluca Demartini, and Philippe Cudré-Mauroux. 2013. Pick-a-crowd: tell me what you like, and i'll tell you what to do. In *Proceedings of the 22nd international conference on World Wide Web*. 367–374.

⁵See [11] for an overview of how organisations have been using blockchain technologies with the aim to decentralise governance.

- [11] Youssef El Faqir, Javier Arroyo, and Samer Hassan. 2020. An overview of decentralized autonomous organizations on the blockchain. In *Proceedings of the 16th International Symposium on Open Collaboration*. 1–8.
- [12] Enrique Estellés-Arolas and Fernando González-Ladrón-de Guevara. 2012. Towards an integrated crowdsourcing definition. *Journal of Information science* 38, 2 (2012), 189–200.
- [13] European Research Council. 2018. Ethics Self-Assessment step by step. <https://erc.europa.eu/content/ethics-self-assessment-step-step>
- [14] Rita Faullant, Johann Fueller, and Katja Hutter. 2013. Fair play: perceived fairness in crowdsourcing communities and its behavioral consequences. In *Academy of Management Proceedings*, Vol. 2013. Academy of Management Briarcliff Manor, NY 10510, 15433.
- [15] Yuanyue Feng, Hua Jonathan Ye, Ying Yu, Congcong Yang, and Tingru Cui. 2018. Gamification artifacts and crowdsourcing participation: Examining the mediating role of intrinsic motivations. *Computers in Human Behavior* 81 (2018), 124–136.
- [16] Donglai Fu and Yanhua Liu. 2021. Fairness of Task Allocation in Crowdsourcing Workflows. *Mathematical Problems in Engineering* 2021 (2021).
- [17] Mark Graham and Jamie Woodcock. 2018. Towards a fairer platform economy: Introducing the fairwork foundation. *Alternate Routes* 29 (2018).
- [18] Mary L Gray and Siddharth Suri. 2019. *Ghost work: How to stop Silicon Valley from building a new global underclass*. Eamon Dolan Books.
- [19] Lei Han, Kevin Roitero, Ujwal Gadiraju, Cristina Sarasua, Alessandro Checchio, Eddy Maddalena, and Gianluca Demartini. 2019. All those wasted hours: On task abandonment in crowdsourcing. In *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining*. 321–329.
- [20] Karin Hansson, Tanja Aitamurto, and Thomas Ludwig. [n.d.]. From alienation to relation: Examining the modes of production in crowdsourcing. ([n.d.]). <https://doi.org/10.18420/ecscw2017-13>
- [21] Karin Hansson, Thomas Ludwig, and Tanja Aitamurto. [n.d.]. Capitalizing Relationships: Modes of Participation in Crowdsourcing. ([n.d.]). <https://doi.org/10.1007/s10606-018-9341-1>
- [22] Chien-Ju Ho, Shahin Jabbari, and Jennifer Wortman Vaughan. 2013. Adaptive task assignment for crowdsourced classification. In *International Conference on Machine Learning*. PMLR, 534–542.
- [23] Chien-Ju Ho and Jennifer Vaughan. 2012. Online task assignment in crowdsourcing markets. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 26.
- [24] Jeff Howe. 2008. *Crowdsourcing: How the power of the crowd is driving the future of business*. Random House.
- [25] Lilly C Irani and M Six Silberman. 2013. Turkopticon: Interrupting worker invisibility in amazon mechanical turk. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 611–620.
- [26] Folasade Olubusola Isinkaye, YO Folajimi, and Bolande Adefowoke Ojokoh. 2015. Recommendation systems: Principles, methods and evaluation. *Egyptian Informatics Journal* 16, 3 (2015), 261–273.
- [27] Dean Jansen, Aleli Alcalá, and Francisco Guzman. [n.d.]. Amara: A Sustainable, Global Solution for Accessibility, Powered by Communities of Volunteers. In *Universal Access in Human-Computer Interaction. Design for All and Accessibility Practice* (Cham, 2014), Constantine Stephanidis and Margherita Antona (Eds.). Springer International Publishing, 401–411.
- [28] Lars Bo Jeppesen and Karim R Lakhani. 2010. Marginality and problem-solving effectiveness in broadcast search. *Organization science* 21, 5 (2010), 1016–1033.
- [29] Jiuchuan Jiang, Bo An, Yichuan Jiang, Chenyan Zhang, Zhan Bu, and Jie Cao. 2019. Group-oriented task allocation for crowdsourcing in social networks. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* (2019).
- [30] Menna Maged Kamel, Alberto Gil-Solla, and Manuel Ramos-Carber. 2020. Tasks Recommendation in Crowdsourcing based on Workers' Implicit Profiles and Performance History. In *Proceedings of the 2020 9th International Conference on Software and Information Engineering (ICSIE)*. 51–55.
- [31] David R Karger, Sewoong Oh, and Devavrat Shah. 2014. Budget-optimal task allocation for reliable crowdsourcing systems. *Operations Research* 62, 1 (2014), 1–24.
- [32] Aniket Kittur, Jeffrey V Nickerson, Michael Bernstein, Elizabeth Gerber, Aaron Shaw, John Zimmerman, Matt Lease, and John Horton. 2013. The future of crowd work. In *Proceedings of the 2013 conference on Computer supported cooperative work*. 1301–1318.
- [33] Robert E Kraut and Paul Resnick. 2012. *Building successful online communities: Evidence-based social design*. MIT Press.
- [34] Seth Siyuan Li and Elena Karahanna. 2015. Online recommendation systems in a B2C E-commerce context: a review and future directions. *Journal of the Association for Information Systems* 16, 2 (2015), 2.
- [35] Leticia Machado, Rafael Prikladnicki, Felipe Meneguzzi, Cleidson RB de Souza, and Erran Carmel. 2016. Task allocation for crowdsourcing using AI planning. In *Proceedings of the 3rd International Workshop on CrowdSourcing in Software Engineering*. 36–40.
- [36] Ke Mao, Ye Yang, Qing Wang, Yue Jia, and Mark Harman. 2015. Developer recommendation for crowdsourced software development tasks. In *2015 IEEE Symposium on Service-Oriented System Engineering*. IEEE, 347–356.
- [37] Annette Markham and Elizabeth Buchanan. 2012. Ethical Decision-Making and Internet Research: Recommendations from the AoIR Ethics Working Committee. <http://aoir.org/reports/ethics2.pdf>
- [38] David Martin, Benjamin V Hanrahan, Jacki O'Neill, and Neha Gupta. 2014. Being a turker. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*. 224–235.
- [39] Karl Marx. 1959. *Economica* 26, 104 (1959), 379–379. <http://www.jstor.org/stable/2550890>
- [40] Brian McInnis, Dan Cosley, Chaebong Nam, and Gilly Leshed. 2016. Taking a HIT: Designing around rejection, mistrust, risk, and workers' experiences in Amazon Mechanical Turk. In *Proceedings of the 2016 CHI conference on human factors in computing systems*. 2271–2282.
- [41] John H McNeely. 1966. Origins of the Zapata revolt in Morelos. *Hispanic American Historical Review* 46, 2 (1966), 153–169.
- [42] Alan Menk, Laura Sebastia, and Rebeca Ferreira. 2019. Recommendation systems for tourism based on social networks: A survey. *arXiv preprint arXiv:1903.12099* (2019).
- [43] Ted Palys. 2008. Purposive sampling. In *The SAGE Encyclopedia of Qualitative Research Methods*, Lisa M Given (Ed.). Vol. 2. Sage, 697–698.
- [44] Maria Papoutsoglou, Georgia M Kapitsaki, and Lefteris Angelis. 2020. Modeling the effect of the badges gamification mechanism on personality traits of Stack Overflow users. *Simulation Modelling Practice and Theory* 105 (2020), 102157.
- [45] Paul Resnick, Ko Kuwabara, Richard Zeckhauser, and Eric Friedman. 2000. Reputation systems. *Commun. ACM* 43, 12 (2000), 45–48.
- [46] Paul Resnick, Richard Zeckhauser, John Swanson, and Kate Lockwood. 2006. The value of reputation on eBay: A controlled experiment. *Experimental economics* 9, 2 (2006), 79–101.
- [47] Joel Ross, Lilly Irani, M. Six Silberman, Andrew Zaldivar, and Bill Tomlinson. 2010. Who Are the Crowdworkers? Shifting Demographics in Mechanical Turk. In *CHI '10 Extended Abstracts on Human Factors in Computing Systems* (Atlanta, Georgia, USA) (CHI EA '10). Association for Computing Machinery, New York, NY, USA, 2863–2872. <https://doi.org/10.1145/1753846.1753873>
- [48] David Rozas, Nigel Gilbert, Paul Hodgkinson, and Samer Hassan. 2021. Talk Is Silver, Code Is Gold? Beyond Traditional Notions of Contribution in Peer Production: The Case of Drupal. *Frontiers in Human Dynamics* 3 (2021), 12. <https://doi.org/10.3389/fhumd.2021.618207>
- [49] David Rozas, Antonio Tenorio-Fornés, Silvia Díaz-Molina, and Samer Hassan. 2021. When Ostrom Meets Blockchain: Exploring the Potentials of Blockchain for Commons Governance. *SAGE Open* 11, 1 (2021), 21582440211002526. <https://doi.org/10.1177/21582440211002526>
- [50] David Rozas, Antonio Tenorio-Fornés, and Samer Hassan. 2021. Analysis of the Potentials of Blockchain for the Governance of Global Digital Commons. *Frontiers in Blockchain* 4 (2021), 15. <https://doi.org/10.3389/fbloc.2021.577680>
- [51] Niloufar Salehi, Lilly C Irani, Michael S Bernstein, Ali Alkhatib, Eva Ogbe, and Kristy Milland. 2015. We are dynamo: Overcoming stalling and friction in collective action for crowd workers. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. 1621–1630.
- [52] Gerald Schimak, Denis Havlik, and Jasmin Pielorz. 2015. Crowdsourcing in crisis and disaster management—challenges and considerations. In *International symposium on environmental software systems*. Springer, 56–70.
- [53] Daniel B Shank. 2016. Using crowdsourcing websites for sociological research: The case of Amazon Mechanical Turk. *The American Sociologist* 47, 1 (2016), 47–55.
- [54] Avi Silberschatz, Peter Baer Galvin, and Greg Gagne. 1999. *Applied operating system concepts*. John Wiley & Sons, Inc.
- [55] Andrew S Tanenbaum and Herbert Bos. 2015. *Modern operating systems*. Pearson.
- [56] Max Weber. 1904. Die "Objektivität" sozialwissenschaftlicher und sozialpolitischer Erkenntnis. *Archiv für sozialwissenschaft und sozialpolitik* 19, 1 (1904), 22–87.
- [57] Jurgen Willems, Carolin J Waldner, and John C Ronquillo. 2019. Reputation Star Society: Are star ratings consulted as substitute or complementary information? *Decision Support Systems* 124 (2019), 113080.
- [58] Xiaoyan Yin, Yanjiao Chen, and Baohun Li. 2017. Task assignment with guaranteed quality for crowdsourcing platforms. In *2017 IEEE/ACM 25th International Symposium on Quality of Service (IWQoS)*. IEEE, 1–10.
- [59] Han Yu, Zhiqi Shen, and Cyril Leung. 2013. Bringing reputation-awareness into crowdsourcing. In *2013 9th International Conference on Information, Communications & Signal Processing*. IEEE, 1–5.
- [60] Eve Zelickson. 2019. A Balancing Act: Hybrid Organizations in the Collaborative Economy Empirical Evidence from Amara on Demand.
- [61] Bingxu Zhao, Yingjie Wang, Yingshu Li, Yang Gao, and Xiangrong Tong. 2019. Task allocation model based on worker friend relationship for mobile crowdsourcing. *Sensors* 19, 4 (2019), 921.
- [62] Kathryn Zyskowski and Kristy Milland. 2018. A crowded future: Working against abstraction on Turker Nation. *Catalyst: Feminism, Theory, Technoscience* 4, 2 (2018), 1–30.