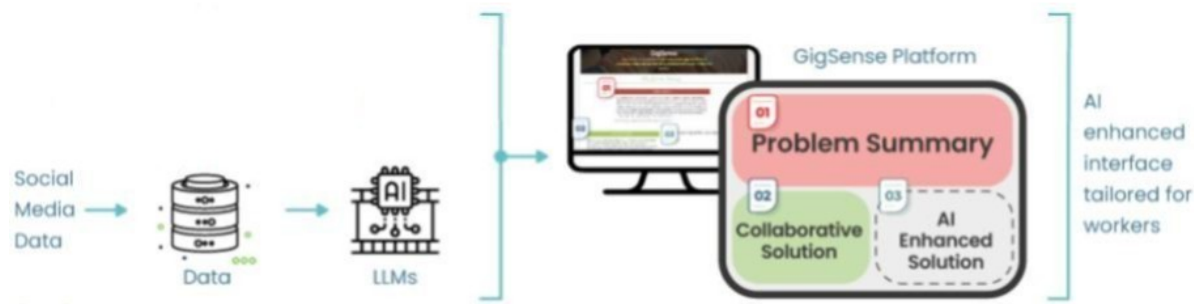
GigSense:An Intelligent Tool for Supporting Gig Worker Collective Action

KASHIF IMTEYAZ, Northeastern University, USA

CLAUDIA FLORES SAVIAGA, Northeastern University, USA

SAIPH SAVAGE, Northeastern University, USA



automatically organizes and summarizes workers' problems. Within short timeframes, workers can zoom in and out to analyze and understand their collective issues. The tool also offers a collaborative space for brainstorming and selecting optimal solutions to address their problems. Moreover, GigSense's intelligent agent provides AI-enhanced solutions to further support workers' brainstorming and planning endeavors.

Collective action among gig workers holds promise for advocating better labor conditions. Nevertheless, despite efforts from researchers, practitioners, and workers themselves to design tools to support collective action among gig workers, enabling collective action continues to present challenges. Existing tools often assume that gig workers have ample time and equal skills for problem analysis and solution proposals, which is often not the case. To address this, we introduce GigSense, a tool integrating large language models and collective action and sensemaking theories. It assists gig workers to swiftly comprehend collective challenges and devise effective solutions, regardless of their backgrounds. In a between-subject user study (N=24), GigSense users outperformed those using a state-of-the-art control interface, producing faster and higher-quality solutions. They also reported enhanced usability experiences. GigSense opens new design possibilities, showcasing how interfaces with large language models can empower gig workers in collective action efforts.

CCS Concepts: • Human-centered computing → User studies; Collaborative and social computing.

ACM Reference Format:
Kashif Imteyaz, Claudia Flores Saviaga, and Saiph Savage. 2023. GigSense: An Intelligent Tool for Supporting Gig Worker Collective Action. In . ACM, New York, NY, USA, 25 pages. https://doi.org/XXXXXXX.XXXXXXX

Permission to make digital or hard copies of all or part 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 components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
© 2023 Association for Computing Machinery.

1 INTRODUCTION

Collective action by gig workers can be a powerful method for improving labor conditions on platforms such as Upwork, Amazon Mechanical Turk, and Uber [18, 42, 100, 101]. Examples of successful gig worker collective action include the remarkable efforts of ride-sharing drivers affiliated with companies like Uber and Lyft. These drivers were able to organize strikes and protests to advocate for higher wages and fairer treatment [30]. Through their collective action, they achieved significant improvements in gig workers' base pay rates and compensation structures [8]. For example, within certain cities, workers were able to negotiate for improved working conditions and better regulation [84]. The power of gig worker collective action has been demonstrated also when it led to the implementation of "pro-worker" policies and led to collective bargaining with the gig platforms [58].

However, despite the occasional success stories, gig worker collective action is rare [8]. As a result, most worker problems remain unresolved. We contend that two key issues are impeding the progress of collective action among gig workers. Note that these challenges pertain to the initial stages of collective action, encompassing the tasks of comprehending collective issues and formulating effective solutions to tackle these problems [107]. The first challenge stems from the scarcity of available technologies designed to assist gig workers in problem-solving [41, 57, 69, 108, 124]. For example, technologies like Dynamo or Coworker.org allow workers to share their problems, upvote the most crucial ones, as well as propose and upvote solutions to those problems [26, 105]. These technologies typically employ list-based interfaces to present the problems and solutions, and the lists are usually sorted based on the upvotes. However, list-based interfaces can restrict the exploration of problems and solutions from various perspectives, impeding a deep dive into their complexities, as well as obtaining a grasp of the entire problem-solution landscape [16, 60, 73, 82, 97, 118]. This limitation can hinder gig workers' ability to effectively prioritize critical issues like identify optimal solutions [119]. The absence of interfaces that facilitate comprehensive problem-solving analysis can lead to workers focusing on trivial problems or irrelevant solutions, ultimately hindering workers' potential to drive meaningful change.

The second challenge affecting gig worker collective action stems from the diversity of skills possessed by gig workers [36, 45, 46]. For instance, some gig workers may have IT expertise, while others could be video editors and other data labelers. With workers skilled in everything from IT to video editing to data labeling, gig economy workforces encompass a wide range of expertise [3]. However, this very diversity that defines the gig workforce also hinders unified problem-solving efforts. The diversity of skills and backgrounds among gig workers poses challenges for collective action [121]. For example, a video editor skilled in their craft may lack the software expertise needed to address transparency issues on gig platforms. Without knowledge of the underlying technology, they cannot advocate for technical remedies. This lack of a unifying work experience and shared skillset hinders collective action, as gig workers lack common ground to mobilize around shared goals. The diversity that defines the gig workforce fractures it into discrete roles rather than bringing workers together [106].

Furthermore, even though many workers might have an interest in participating in collective action, their time commitments could be a limiting factor. This highlights the crucial need for inclusive tools that can aid all gig workers in collaboratively addressing challenges, irrespective of their expertise or availability constraints [92, 93]. These tools could empower *all* workers to conduct more strategic problem-solving and create better futures [73].

To initiate collective action among gig workers, we introduce GigSense—a novel platform designed to ignite collaborative problem-solving within the gig worker community. Leveraging Sense Making Theory [99] and the inherent capabilities of large-language models (LLMs) in text generation and processing, we integrate LLMs with interactive interfaces within GigSense. This synergy empowers gig workers to meticulously dissect their challenges and potential

52

solutions, examining them from diverse dimensions and viewpoints. The GigSense platform also empowers gig workers to zoom in for detailed scrutiny or zoom out for a panoramic understanding of their problems. This sets it apart from existing technologies [105], which mainly present workers with an extensive list of problems and solutions, lacking mechanisms for exploring different facets of the problems or achieving a comprehensive understanding of workplace dynamics and crucial matters. In addition, GigSense employs LLMs to aid workers in brainstorming solutions, enabling all workers to collectively address their challenges. Our design challenges revolve around creating interfaces that facilitate workers' comprehension and discussion of workplace matters and solutions, ensuring ease of participation for all. Gigsense works with social media data, particularly focusing on Reddit. Additionally, it employs datasets containing reviews for gig work platforms that have been compiled from app stores accessible on both Android and iOS devices. In this paper, we contribute 1) a system supporting gig worker problem-solving to initiate collective action; 2) AI-enhanced interactions that facilitate in-depth analysis of problems and solution generation for gig workers; and 3) a between-subjects experiment demonstrating that GigSense makes problem-solving easier, faster, and results in the generation of higher quality solutions by gig workers.

2 RELATED WORK

Gig Work and Collective Action. "Gig" or platform-based work represents one of the most recent labor market trends [116]. The popularity of gig work can be attributed to the increased demand for flexibility on the part of employers [80], the desire for greater flexibility on the part of workers [25] and the fact that it is work facilitated through technology and digital markets, on-demand [58]. Even though gig work can offer economic advantages to socially disadvantaged populations like the unemployed, those in remote areas, and refugees [25]; it can also have adverse effects on workers, such as unpredictable schedules, fluctuating income, and unreliable long-term employment prospects [25, 95]. Collective action activities (i.e. negotiations, strikes, unionization campaigns) are ways for workers to advocate for better working conditions [61]. Gig workers have participated in collective endeavors to enhance their work conditions; however, the results have been mixed. The reasons vary, from gig platforms not facilitating in-app communication among workers [41] to the geographical dispersal of gig workers [58]. As they move in and out of various short-term "gigs" across different industries [113], this inevitably hinders their ability to form a sense of community and identify common interests among themselves [121]. These characteristics create obstacles for collective organizing efforts, as workers can be challenging to locate, difficult to reach, and hard to engage [117, 121]. Researchers acknowledge that participating in collective action increases a worker's ability to choose better jobs with clear expectations and higher remuneration rates [7]. Even though gig workers have achieved a degree of success in coordinating collaborative efforts to enhance their working conditions [30], systems designed to assist gig workers in organizing collective actions are scarce [58]. Considering the potential growth in this sector, it becomes crucial to develop solutions that support worker organizing and collective bargaining [58]. In this paper, we focus on creating a tool to help gig workers in the sensemaking of their challenges and initiating collective action.

Tools to Support Workers in Collective Action. Over the years different solutions designed to help workers initiate collective action have emerged. For instance, Catalyst [20] was the first attempt at supporting collective action based on "activation thresholds". However, while Catalyst is useful for predetermined events, it lacks coordinating features to allow users to coordinate and reach a consensus. This is an important limitation as researchers emphasize the importance of collaboration among workers for the future of gig work, especially in the context of creative tasks [63]. A more recent solution that aims to allow workers to collaborate is Turkopticon [54] is a browser extension and website designed to enhance transparency and empower workers in the Amazon Mechanical Turk (MTurk) marketplace. Turkopticon allows MTurk workers to review and rate the requesters on the platform. The platform also enables workers to share their experiences, voice concerns, and warn others about potentially problematic requesters. Platforms in the form of forums have emerged as well, aiming to foster the sharing of valuable insights and experiences among workers, thereby nurturing a sense of community and collaborative learning. Notably, examples like Turker Nation and Turkerview have arisen in response to the policy adopted by Amazon Mechanical Turk (AMT), where requesters have the discretion to pay only for tasks they deem subjectively satisfactory. However, it's worth noting that Turkopticon, Turker Nation and Turkerview are tailored exclusively for Amazon Mechanical Turk, rendering them inaccessible to gig workers operating on different platforms.

In this paper we use a more recent initiative as a baseline: "We Are Dynamo," a forum that was specifically created to facilitate collective action for crowd workers [105]. Through this forum, workers come together, identified common interests, and collaboratively crafted a list of best practices. However, it is important to highlight that "We Are Dynamo," does not facilitate communication, collaboration, and sensemaking. These aspects need to be coordinated outside the platform, which introduces several barriers [58]. First, the need for external coordination can result in fragmented communication, making it challenging to ensure all relevant parties are engaged in discussions. Second, navigating between platforms can be cumbersome and time-consuming, potentially deterring workers from actively participating. Lastly, the lack of integrated communication tools might inhibit the spontaneous exchange of ideas and hinder the swift organization of collective actions and sensemaking. Our work builds upon "We Are Dynamo,' as we incorporate collaborative functionalities, harness the power of AI to enhance idea generation and tailor our system design to facilitate gig workers' sensemaking process to find solutions to their challenges.

Sensemaking and Collective Action. Comprehending the challenges faced by gig workers can be perceived as an act of sensemaking, involving the collection and analysis of diverse and unstructured data in order to reach a conclusion. Pirolli and Card [97] define sensemaking as a series of iterative steps. For instance, it starts with the initial gathering of relevant data (*"Step: Search and Filter"*), akin to brainstorming gig workers' problems. Subsequently, it involves extracting valuable information (*"Step: Read and Extract"*), akin to selecting the most pertinent issues. Further, it encompasses summarizing and schematizing the information (*"Step: Schematize"*), akin to the manual procedure of condensing and structuring of the brainstormed ideas. Then, it involves generating hypotheses from various perspectives (*"Step: Build Case"*), resembling the development of viable solutions. Lastly, it culminates in decision-making to determine the best solution (*"Step: Tell Story")*. Significant research endeavors are currently directed toward the development of tools that facilitate collaborative sensemaking in different domains, such as literature review[126], web search and organization of results [48, 65, 94], organizing academic literature [102], solving mysteries [74] and tackling disinformation [35]. In this paper, we introduce GigSense, a system tailored to aid gig workers in their collaborative endeavors, automating segments of the sensemaking process pipeline. Furthermore, we leverage the power of Large Language Models (LLMs) to aid workers in the sensemaking process and idea generation.

LLMs for Idea Generation. In recent years, significant advancements in large language models (LLMs) have positioned them as a promising tool for facilitating a diverse array of writing tasks, such as story generation [1, 21, 85, 122], academic writing [37], question-answering [14] and idea generation [39]. However, despite their impressive utility, LLMs have faced criticism for generating text that, while appearing logically and grammatically coherent, may actually contain factual inaccuracies or lack meaningful coherence (referred to as hallucinations) [55]. Nevertheless, researchers suggest that this weakness can be reframed as a strength [39], and have started to use LLMs for idea suggestions[122]. In creative writing and problem-solving, possessing a range of idea quality and quantity holds more value than maintaining unwavering consistency [40]. To achieve a wide range of ideas with varying levels of quality, most ideation research advises generating numerous ideas first and delaying their evaluation [40]. LLMs are designed to do exactly this— quickly generate many somewhat plausible solutions [39]. In this research, we align with the principle that an unrestrained influx of ideas can often pave the way for innovative solutions. Therefore, we harness the capabilities of LLMs by integrating them into Gigsense to assist gig workers in rapidly producing multiple reasonably viable ideas for their most pressing challenges.

Collaborative Human Centered Systems. Our system design is inspired by prior human-centered design systems that foster collaboration. First, we draw inspiration from work that utilized sensemaking to facilitate collaborations among strangers. For instance, systems like GroundTruth [115] integrated sensemaking to support collective efforts in detecting visual disinformation. Likewise, the system Crowdia employed sensemaking to solve mysteries among strangers [74]. We examine these designs to determine how to integrate the sensemaking process into our own system. We also connect with prior work on interface designs aimed at enhancing idea generation [17, 51, 52, 111]. For instance, AnalogiLead explored interactive interfaces to encourage analogical innovations [111]. Likewise, we propose interfaces to assist gig workers in exploring various problem aspects, leading to more viable solutions.

3 GIGSENSE

Gigsense modularizes the sensemaking process allowing gig workers to formulate solutions for their challenges. We present a scenario in which Gigsense can be utilized, followed by an overview of the system.

User Scenario. Maria, an Upwork freelancer, has noticed an influx of vague job ads and unusual high-pay job ads, possibly generated by AI. She wants to tap into the new job opportunities, but she is also unsure if they are real. She applied for such gigs, but it turned out to be fake postings. On GigSense, she rapidly discovers numerous peers facing the same issue by looking at Gigsense's data-viz module. Utilizing GigSense's zoomed-in problem view, she identifies a split—some ads seek free labor, while others are legitimate jobs with high pay. Switching to GigSense's Collaborative Solution module, Maria feels unsure about how to solve the problem. Turning to GigSense's AI- enhanced solution, she quickly finds an initial direction: uniting workers to create a spreadsheet listing scammers across the gig work platforms. From this idea, Maria then envisions a web plugin that could help workers automatically flag suspicious clients and fake job postings. However, she lacks the software expertise needed to develop such a web plugin. She inputs the idea into GigSense's Collaborative Solution Module. Kai and Ana who happen to experience a similar problem with fake job postings then spot Maria's idea and discuss it asynchronously. They concur on a solution involving a web plugin that flags problematic clients, using the spreadsheet list, they create a database of scammers who are posting fake job ads. Then they further develop the web plugin that can use AI to predict clients that could potentially be problematic based on their job ads. Maria's idea and Kai and Ana's tech skills led to the implementation of the web plugin, effectively addressing the issue and benefiting the entire community.

System Description. Our work is driven by Pirolli and Card sense-making and collective action theories [27, 97, 108, 118], translating it into the different modules of the system: "Data Gathering Module", "Data Viz Module", "Problem Summary Module", "Collaborative Solution" "AI-Enhanced Solution" modules. In this section, we introduce GigSense, a platform showcasing the potential of these modules for sensemaking, along with interactive interface components. We also provide a comprehensive explanation of each individual module. Fig ?? presents an overview of GigSense.

3.1 GigSense's Modules.

GigSense has a series of modules that focus on automating the sense-making process, as well as enabling universal participation, empowering everyone to devise creative collective solutions. In doing so, it facilitates the initial phases of collective action: (1) identification of collective problems; and (2) proposal of action plans (solutions) to address those problems [108]. GigSense recognizes the time constraints faced by gig workers when engaging in collective action and provides interfaces that facilitate rapid sensemaking of problems and proposals of solutions.

To address this, GigSense utilizes Large Language Models (LLMs) to provide concise summaries of issues found on gig platforms, enabling workers to quickly understand prevalent problems. By leveraging the OpenAI GPT-4 API [91], GigSense categorizes data and generates these summaries, aiding workers in gaining insights and brainstorming solutions. GigSense also offers an interactive interface, allowing users to "zoom in" and explore specific discussions related to a problem. While LLMs supports the creation of problem summaries, GigSense's interface design empowers workers to independently assess all the information. To overcome the challenges identified in recent research regarding the design of effective prompts for LLMs [123], GigSense also includes carefully crafted prompts in its backend. These prompts and their outputs underwent iterative refinement with gig workers to ensure satisfactory results. You can find the prompts GigSense uses in our appendix.

Next, we describe each GigSense module:

Data Gathering Module. Gig workers supply Gigsense with a roster of subreddits from which they intend to pinpoint potential problems and datasets containing assessments for gig work platforms (This is the "Step: Search and Filter" in Pirolli et al.'s sensemaking loop [97]). Next, Gigsense connects to the Reddit API to read and extract all the posts from the subreddits that gig workers initially provide. GigSense additionally uses a web scrapper to extract data from reviews left on Apple's and Google's app stores by gig workers. Note that our data gathering module only collects reviews that have between one and three-star ratings. The module considers that these review data would represent complaints and problems that gig workers are experiencing. Gigsense also lets workers manually enter issues into the system if they choose to do so ("Step: Read and Extract" in the sensemaking loop). Using the real-world gig worker's complaint datasets (actual gig workers' subreddits and complaints) in our system design aims to bring inclusiveness about gig worker concerns and complaints. The show button allows workers to see a list of all the data sources that are going to be used within Gigsense.

Problem Summary Module. Acknowledging the potential enormity of the data gathered through the Data Gathering Module and its potential complexity for human interpretation, this module centers its efforts on summarizing the data. For this purpose, Gigsense uses LLMs to categorize and summarize large sets of data into problem categories. This module also has two buttons to allow workers to navigate to the Data Visualization module, which offers a "zoom-in" view for meticulous examination of the data or a "zoom-out" view for a panoramic grasp of their issues (See Fig ?? C).

Collaborative Solution Module. This module further facilitates the sensemaking process and focuses on helping gig workers to develop concrete solutions to address the problem analyzed ("Step: Build Case' " in the sensemaking loop). It incorporates sub-modules such as the: "Sensemaking Chat", "Shared Document" and "Collaborative Solution Space". The Sensemaking Chat submodule allows workers to engage in conversations to discuss and investigate the problems they encounter in their work. They can communicate through asynchronous text messages to accommodate different schedules. The "Shared Document" enables workers to understand the existing problems and collaboratively create action plans (solutions) to address them. Workers can collectively review and approve their proposed solutions. Finally, the "Collaborative Solution Space", just like the sensemaking process, features a space where workers can showcase the final solution they mutually agreed upon [27, 118] ("Step: Tell Story" in the sense-making loop).

AI-Enhanced Solution Module. Recognizing the diverse backgrounds of gig workers, it is important to acknowledge that for some workers it may be hard to propose solutions, or effectively plan and organize efforts to tackle specific problems [92, 110]. To address this, the module leverages LLMs to offer workers suggestions on potential solutions and concrete collaboration plans, providing inspiration and initial guidance. However, given our values of prioritizing human connections among workers, AI-generated solutions are presented with lower priority in GigSense's interface. This module serves as an inspirational resource for gig workers, but it is important to note that GigSense emphasizes that it is not the definitive solution and explicitly states that this solution was generated using Generative AI. Furthermore, workers are informed about the possibility of errors in AI-generated solutions via a disclaimer. This helps not only to ensure that solutions generated by humans receive priority, but also foster ethical, transparent, and responsible use of Generative AI while still utilizing AI to assist and inspire workers in their problem-solving endeavors. This module aims to support: "Step: Build Case" and "Step: Tell Story" in the sensemaking loop.

Data Viz Module. Grounded in Olson's collective action theory [90] which suggests the conciseness of the problem has to be developed in the first place in order for the collective action to begin. We introduced The zoomed-out view of this module presents an interactive chart of all the problems ("Step: Schematize" in the sensemaking loop) where workers can grasp a high-level understanding of the specific types of problems (e.g. Payment, Platform Policy, Scam, Customer Support, Usability, etc) faced by them. The interactive chart also serves as a powerful tool for raising awareness among gig workers about the shared challenges they face.

To let the worker delve deeper into the individual problem faced by a worker and suggest the solution to an individual problem, Gigsense offers a zoom-in perspective. This view gives workers the opportunity to upvote on individual problems. By visually displaying the prevalence of various issues (number of upvotes), it creates a sense of solidarity. Workers can see that they are not alone in their struggles, fostering a collective identity and shared purpose. It provides a static depiction of the number of workers grappling with problems across different gig work platforms, acting as a catalyst in building communities of gig workers facing similar challenges to encourage collective action.

4 GIG SENSE: EVALUATION

The evaluation of GigSense aims to address key research questions: 1) *Speed:* Can GigSense facilitate more rapid sensemaking, allowing gig workers to more seamlessly integrate it into their workflow? 2) *Contribution:* Can GigSense effectively support the generation of more feasible solutions for gig workers? 3) *Usability:* Does GigSense's AI-enhanced interactive interface bring better user experiences?

Procedure. To study the above questions, we conducted an IRB-approved between-subject user study with 24 participants. We divided the participants into intervention (GigSense condition) and control condition. Participants in both groups were asked to complete the same tasks linked to the initial phases of collective action: pinpointing collective issues and suggesting solutions [124]. Participants in the Gigsense condition used our Gigsense platform to complete the tasks (see Fig ?? and ??, while participants in the control condition used the "We Are Dynamo" interface. We built "We Are Dynamo" to simulate the general functionality of the original system which is no longer available for use. In this interface, users can do the typical things they would do when using the original version of "We Are
Dynamo" (such as posting their ideas for action and general upvoting of the ideas), see Fig ??.

Next, we compared the latency and quality of solutions that were generated using GigSense and the control interface "We Are Dynamo". We also studied the usability of GigSense in comparison to the control interface. Note that both systems (GigSense and the control condition) used identical datasets encompassing gig worker problems, which were taken from social media posts (subreddits) and reviews on the Google and Apple app stores. GigSense received the data and leveraged its backend with LLMs and interactive interfaces to offer gig workers a multi-level analysis of the problem space. Similarly, the control interface organized problems based on upvote count, akin to Dynamo's original design where workers can post and upvote short ideas for action. If an idea gets enough upvotes, it becomes an active campaign.

We arranged the study under the assumption of gig workers operating asynchronously in their collective efforts. This asynchronous setup is crucial due to the varied schedules of gig workers [68], which might hinder synchronous collaboration. Our aim was thus to ensure effective asynchronous utilization of our tool for seamless completion of collective action tasks. Participants in both conditions engaged with their respective assigned systems and fulfilled the following tasks, drawn from existing literature concerning activities associated with the initial phases of collective action [6, 23, 108, 124].

(1) Provide a summary of one specific problem encountered by gig workers.
(2) Provide a summary of three different problems faced by gig workers.
(3) Enumerate three problems that demand attention due to the adverse impact they inflict on workers.
(4) Explain why it is important to address those problems.
(5) Propose solutions to the three problems you identified that were crucial to be addressed.
(6) Propose a solution to any other problem faced by gig workers.
(7) Propose three solutions to any other problem faced by gig workers.

In our study setup, we tracked and recorded how participants completed these different tasks.

Participant Recruitment. To recruit participants, we generated a job listing on Upwork, extending an invitation to gig workers to join our study. Our selection criteria for participation in the study were workers who: (a) were aged 18 or above; (b) possessed at least one year of gig work experience (to ensure familiarity with the challenges faced by workers); and (c) demonstrated proficiency in spoken, written, and comprehended English (to facilitate effective communication with participants). From this, we recruited 24 participants (8 females, 16 males, Median age=27, SD=7.186). After recruitment, we randomly assigned participants to the control and GigSense conditions using the block randomization technique [31]. In the end, 12 participants were assigned to the GigSense condition, and 12 were assigned to the control condition. Participants in our user study were compensated $10/hr for their participation.

Measures and Data Analysis

We adopted a mixed-method approach, enabling us to harness the strengths of both quantitative and qualitative techniques. Alongside the collection of essential sociodemographic data, we harnessed a range of quantitative metrics to address our three research questions related to speed, contribution, and usability. To provide a deeper perspective and also contrast against state-of-the-art tools, we complemented our quantitative findings by conducting exit interviews with participants in both conditions. Through this approach, we gained invaluable insights into their impressions and experiences.

Metric: Speed. In both the control and GigSense conditions, participants used a button to signal task start and completion. The systems recorded timestamps for each button press, enabling precise tracking of task durations per participant.

ID	Age	Gender	Race	Work Exp.	Area of Expertise
P1	48	Male	White	4 Years	Content writer, Graphic designer
P2	30	Female	Asian	2 Years	HR, Data Entry, Survey
P3	19	Male	Black	1 Years+	Book Editing, Copywriting
P4	32	Male	Black	6 Years	Web Developer, Digital Marketing, Virtual Assistance
P5	38	Male	White	7 Years	Career Coaching, Content Writing, SEO, Human Resource
P6	22	Female	South Asian	2 Years	Content Writer
P7	24	Male	South Asian	3 Years	Content Writing, Online Education
P8	35	Male	White	7 Years	Scriptwriter, Book Editing, Content Writing
P9	31	Male	South Asian	4 Years	Software, IT, Graphic Design
P10	26	Female	Black	5 Years	Virtual Assistant, Digital Marketing
P11	28	Female	Black	3 Years	Data labelling, Content Writing
P12	36	Female	Asian	4 Years	Content Writing, Digital Marketing
P13	27	Female	White	4 Years	Book Editing, Copywriting, Content Writing
P14	25	Male	Black	6 Years	Human Resource, Career Coaching
P15	23	Male	South Asian	5 Years	Software, IT
P16	40	Male	Black	10 Years	Academic Research, Audio Production
P17	19	Male	White	2 years	Software
P18	22	Male	South Asian	3 years	Digital Marketing, Software
P19	22	Female	South Asian	2 years	Content Writing, Graphic Design
P20	27	Female	Black	5 years	Book Editing, Content Writing
P21	25	Male	South Asian	2 years	Software, Digital Marketing
P22	22	Male	South Asian	3 years	Online Education
P23	28	Male	Asian	6 years	Human Resource, Digital Marketing
P24	34	Male	White	8 years	Software, Career Coaching

Table 1. Participant Demographics in our study

Metric: Contribution. To assess GigSense's impact on participant problem-solving (i.e., on the contributions they made using our system), we gathered the solutions participants generated in both conditions. Via Upwork, we then hired three English-speaking, college-educated gig workers as independent raters. They evaluated each solution's feasibility on a 7-point Likert scale, considering the problem and solution. For the remaining solutions, the third rater's assessment was sought to resolve disagreements. A "majority rule" was applied to determine scores for these solutions.

Metric: Usability. To assess participants' views on GigSense's usability and compare it with the control, we employed the System Usability Score (SUS) [13], a validated metric. The SUS is comprised of 10 questions on a five-point Likert scale. It is widely used for measuring usability and comparing systems [70, 96]. It offers valuable insights into users' subjective experiences with a given technology. Following participants' interaction with their assigned system (control or GigSense), they received the SUS questionnaires. We then computed the SUS scores reported by gig workers for their respective systems. After this step, for each participant, we had the usability score they gave their assigned technologies.

Qualitative Study. To augment our quantitative data, we conducted exit interviews for richer insights into participants' user experience, adding a qualitative dimension. These interviews captured participants' feedback and thoughts about the technologies they used. Interview data was transcribed, coupled with study notes and memos, and subjected to open coding [83]. This process involved deriving initial concepts from the data through a mix of bottom-up and top-down theme extraction. Two of the paper's authors independently conducted bottom-up coding, resulting in 13 axial codes, which were then applied top-down to all interview transcripts. Out of these 13 axial codes, four themes emerged, structuring the key insights from our semi-structured interviews.

5 RESULTS

Our user study encompassed a cohort of 24 gig workers, and it was structured with two conditions: the GigSense condition, designated as the intervention group, and the control group. Next, we proceed to unveil both the quantitative and qualitative results that surfaced during the course of our investigation.

5.1 Quantitative Results.

5.1.1 Time. Time can pose a challenge for gig workers aiming to engage in collective action, as not all workers enjoy the luxury of allocating extensive time to this pursuit [92]. To address this concern, we assessed the duration participants required to accomplish the different problem-solving tasks defined in our study, which constitute the initial phases of collective action [90]. Figure ??.a) provides a comprehensive depiction of the median time taken by participants to complete the entire set of tasks in both conditions. (Figure ??.c) depicts the box plot for both groups. The results of our study indicate that participants in the GigSense group exhibited faster task completion times (Mean 264.08 seconds, Median=170 seconds, SD= 175.45 seconds) compared to the control condition (Mean= 862.5 seconds, Median=779 seconds, SD= 313.93 seconds). To study whether these differences between the GigSense condition and control were significant, we conducted appropriate statistical tests. First, since our data did not meet the assumption of normality we employed the Mann-Whitney U test, a non-parametric test specifically designed to compare the medians of task completion times between the intervention group (GigSense) and the control group. The Mann-Whitney U test revealed a statistically significant difference between the two groups in our study, with a p-value of 0.002. This p-value indicates that there is a statistically significant difference between the two groups in our study. This implies that gig workers were significantly quicker in their problem-solving tasks when utilizing GigSense compared to traditional interfaces. Overall, the data provides evidence that GigSense offers a promising approach (RQ1) to improve task completion times in problem-solving tasks related to collective action.

5.1.2 Contribution (Evaluating Gig Workers' Solutions). To study whether GigSense effectively supports the generation of more feasible solutions for gig workers, we conducted an expert evaluation of the solutions that participants in both groups proposed. We found that gig workers in the GigSense group produced in general more feasible solutions (Me- dian=7 ["Very Feasible"], Mean=5.76 [somewhat feasible], SD=1.8) than workers using the control interface (Median=3 [Slightly Unfeasible], Mean=3.58 [Slightly Unfeasible], SD=2.1). We plotted a box plot graph (??.b) to better visualize the differences in the solutions each group contributed. Next, we wanted to identify whether the differences in the feasibility of solutions were significant. Through our analysis, we first identified that the distribution of feasibility scores did not meet the assumption of normality. Consequently, we again performed the Mann-Whitney U test. The results of this test indicated a statistically significant difference between the two groups, with a p-value of 0.248. This suggests the presence of a significant distinction between the feasibility of the solutions that gig workers contributed in the GigSense condition and the control condition. In conclusion, our findings reveal that GigSense facilitates the contribution of more feasible solutions (RQ2) by gig workers, as evidenced by the significant difference in the expert evaluation scores between the two groups.

5.1.3 System Usability Scale. Utilizing the System Usability Scale (SUS) [4, 13, 96], we studied the reported usability levels of GigSense among gig workers and drew a comparison with those reported for the control condition. Figure ??.d presents the boxplots for the System Usability Scale scores of GigSense and the control condition. Our findings revealed a notable trend: the median SUS score for GigSense (Mean=86.25, Median=86, (adjectival rating: Excellent), SD=11.6) was higher than the median SUS score for the control condition (Mean=20.41, Median=14, (adjectival rating: Poor), SD=18.7). Building upon this observation, our subsequent focus was to determine the significance of this disparity. We first found that the SUS scores did not meet the assumption of normality. We therefore opted for the Mann-Whitney U test once again to compare the medians of SUS scores between the GigSense condition and the control condition. The analysis revealed a statistically significant difference between the two groups in our study(RQ3), with a p-value of 0.001. This indicates that the difference in SUS scores between GigSense and the control group is unlikely to have occurred by chance alone.

These higher scores indicate increased user satisfaction, ease of use, and overall acceptance of GigSense. These findings offer strong evidence for GigSense's effectiveness in both quicker task completion and enhanced usability perception. This blend of faster tasks and better usability makes GigSense a promising intervention.

5.2 Qualitative Results

To analyze, the interview response, the transcripts were open-coded individually by two researchers and further axial coding was performed. From 13 top-down axial codes, we developed four themes that organized the main insights from our exit interviews. We labeled participants' responses within the "GigSense group" as "*PGS*" and those in the control group as "*PC*".

5.2.1 ***Assisting Workers in Solution Generation.*** Participants appreciated how GigSense effortlessly enabled solution generation. They particularly liked the fusion of the AI-enhanced interface with the interactive collaborative space, as it seamlessly facilitated the combination of existing solutions into more fortified ones: *"I like this part [pointing to GigSense's AI-enhanced solution space and the button for generating new solutions]. I love that I can interchange the suggestions that I want anytime [Here, they highlighted the collaborative solution space]"* (PGS 5).

Participants also valued how GigSense empowered them to formulate solutions for less familiar problems, broadening their capacity to participate in problem-solving across areas where they might not usually contribute ideas: *"Often I see a problem that I am not familiar with. The AI suggestion kind of suggests you somewhat relatable thing, so in a situation where you might not have anything to contribute, you still have something to suggest, I think that's really cool."* (PGS 11). Given the diverse backgrounds and regions of gig workers, AI assistance played a crucial role in helping participants overcome language barriers, facilitating the generation and seamless sharing of solutions among them: *"The AI is very useful for me because it gives me the many ideas needed, and I also struggle with English, and many times I can't put clearly what I am trying to say"* (PGS 6).

5.2.2 ***Collaborative Problem-Solving****.* Our system's interface was well-received for its ability to help people to come together and drive collaborative problem-solving efforts among gig workers: *"forums [traditional tools] can just be people complaining at each other and nothing happening, whereas this [GigSense] kind of gathers data and proposes solutions. It is just more of a solution based, I think it helps visualize problems [...] then the solutions can be given or can be received by somebody that says: "Oh, let me make something [a solution]" [...] It seems like a space for the solutions to kind of come together"* (PGS 8). Participants also underscored the potential of GigSense to empower them in joining forces, collaboratively suggesting solutions that encompass technological enhancements for gig platforms and necessary policy changes. Participants believed that these combined efforts could cultivate a more nurturing environment for gig workers: *"So, with these solutions [the ones generated collectively through GigSense] either that is, somebody creates a new and better platform or* *o generated collectively through GigSense] and updates their policies and changes things to n make it more friendly for the Freelancers. I think it's like it could be space for creating*

e

o
f

t
h

e

p
l
a
t
f
o
r
m
s

r
e

a

d

s

i
t

[
s
o
l
u

520

t
i
o

n

s

676

better paths for both freelancers and platforms" (PGS 12). Participants also expressed that our tool facilitated stronger connections among fellow workers, fostering a sense of empathy for shared challenges. This, in turn, with the help of the interactive interface, bolstered their collaborative problem-solving efforts: "*I resonate with what is being said [workers sharing their problems] as somebody that's been doing gig work for a while. I highly resonate with it [...] And this interface makes me able to put in my words for [creating] the solution*" (PGS 8).

5.2.3 ***AI, Humans, and Interface Design.*** We crafted GigSense's interface to prioritize the placement of human- generated inputs—solutions proposed by gig workers—above AI-generated suggestions. Participants mentioned that they appreciated this design, highlighting that they liked seeing solutions from both human and AI sources, but especially appreciated the emphasis on those contributed by fellow humans: *"It is good to get the ideas [solutions] from the AI. But I liked that some live answers were there [solutions proposed by gig workers], and some live people, real human beings answered it [gig workers provided solutions]"* (PGS 4).

On the other hand, LLMs occasionally produce content with errors. Our concern revolved around gig workers unquestioningly adopting the solutions provided by the LLM. In our interface design, we proactively introduced an additional measure by including a disclaimer regarding AI-generated suggestions. This strategic move aimed to encourage workers to approach the AI's recommendations thoughtfully, rather than hastily adopting them. Our primary objective was to avert any unforeseen repercussions stemming from thoughtless adherence to AI advice. Participants valued the implementation of this design approach, which underscores our commitment to responsible AI utilization: *"The solutions, (AI suggestions) can be helpful to guide you, but cannot provide you with a perfect solution. That's already there in the disclaimer, So you can get ideas. But you still have to use your brain and experience to answer"* (PGS 2). Participants in general expressed that GigSense's interface gave them a sense of autonomy in their decision-making, as they were not constrained to unquestioningly adhering to the AI-generated solutions: *"I like the fact that the system is suggesting a solution, not completely telling me this is exactly the solution for this problem"* (PGS 5). Participants also conveyed a sense of resonance with GigSense's AI-generated solutions, emphasizing that they were not out of place within GigSense's interface: *"The AI suggestions continuously synced with me, while I'm thinking of my things to write. The answers all seem super aligned with my thoughts and were really helpful suggestions*" (PGS 3).

Workers also valued how GigSense's intelligent interface empowered them to methodically structure and analyze problems in diverse ways: *"..the bar chart [bar chart that was automatically generated by GigSense to show number of worker messages generated about a particular problem] that like immediately told me what the stats were, and then it was quite straightforward to go through each complaint [workers' messages about the problem] and understand all that stuff behind the problem"* (PG 4). This type of interface capability was absent in the control condition and participants complained about it: *"I keep on scrolling, and there are different problems. I am not able to categorize it [categorize worker messages about a particular problem]. So it is time- consuming. I have to scroll through each and every message"* (PC 9). Similarly, some of the workers on the control group imagined better futures where they could have access to some of the interface features of GigSense, such as problem categorization: *"I think it would have been better if the categorization of problems was there [on the control interface]. It breaks down things easier*"(PC7). Other workers in control also desired some of GigSense's interface features, particularly those about concise summaries and the ability to conduct sensemaking in shorter periods of time: *"So it will be very difficult to read those 1,000 words line by line. So it's specifically if you use some artificial intelligence or something, the whole* *thing can be summarized in 200 words. The lesser it is the the less time it will consume, and the user can easily understand."* (PC9). In contrast, participants in the GigSense condition highlighted how they enjoyed that GigSense's interface helped them to minimize sensemaking costs: *"I can* *read the issues that gig workers are facing from here [The worker pointed to the summary module of GigSense], and I don't have to open them [worker messages about a particular problem] one by one. I think it's helpful. So, since gig workers are very busy, having a summary here is very helpful [...] I can read all of this in under a minute* (PG2).

5.2.4 ***Collective Action.*** GigSense played a pivotal role in stimulating participants to embark on the path of collective action. The interface offered by GigSense facilitated a comprehensive exploration of issues from various perspectives, empowering participants to meticulously analyze problems. This, in turn, led to a heightened awareness of the severity of specific challenges: *"I think it [GigSense] can bring people a sense of unity and frustration. So like, it's nice to be able to see, especially all these [pointing at GigSense's graphs showcasing the magnitude of problems faced by workers], because you can see how bad it [a specific problem] is. Then here, [the worker clicked a problem and zoomed into the specific things other workers complained about the problem], all of these reviews [workers' complaints about a given problem], show this problem is super common [...] the reviews [workers' complaints about a given problem] are like written in honest frustration. And it seems kind of like a recognition of: "Oh, everybody's really got this problem" and frustrated by these things. I found this interface like a space for people to gather and create change"* (PGS 8). Additionally, GigSense showed workers their challenges weren't just experienced by themselves. This realization spurred action, as participants often hesitated to address issues or engage with peers because they were uncertain of their problem's significance: *"There could be a time like, you might be facing a problem. But you might be thinking that that's something that you are only facing. And it could be a problem related to usability or payment. You do nothing about it. But when I see this thing on your website [GigSense], you see, other people are also facing it. So this is not just you since there are others like you [facing similar problems]. You get to talk to each other through your shared concerns"* (PGS 10).

6 DISCUSSION

Our user study showcased that GigSense users generated solutions for collective issues significantly faster, with a significant increase in perceived usability, and a significant enhancement in the feasibility of these solutions. Here, we discuss ongoing challenges and prospects for sensemaking tools in collective action, with a focus on GigSense.

Catalyzing Inclusive Problem-Solving for Collective Action. GigSense is designed to facilitate gig workers' participation in problem-solving for collective action. We factored into GigSense's design that it should address time constraints, recognizing the limited availability of gig workers for extensive collective action. The aim was thus to ensure quick sensemaking, enabling more rapid production of solutions for collective issues. Balancing this aspiration with the production of feasible solutions presented a challenging task for GigSense. Our user study demonstrated that GigSense indeed yielded more feasible solutions compared to the control interface. A likely contributing factor was that GigSense's interface empowered workers to swiftly assess zoomed-in and zoom-out dynamics of their problems. This likely led workers to have a more profound understanding and thus generate more attainable solutions, compared to list-based interfaces. However, acknowledging that not all workers might prioritize in-depth problem exploration is essential. To address this, incorporating informative messages within GigSense could enlighten users about the benefits of investing slightly more time in analyzing and comprehending problems. Moreover, envisioning interfaces that highlight potential solution gaps could encourage workers to contemplate solutions more deeply. In this context, drawing insights from data visualization research could provide valuable d instance, exploring how visual encodings can assist workers in identifying overlooked solutions that have not been i taken into account during the process [59]. Another factor to consider is that after identifying an optimal solution, r workers must communicate it effectively to encourage others to join their efforts and complete the collective e action [38, 108]. Creating a narrative

c
t
i
o

n

f
o
r

G
i
g
S
e
n
s
e
’
s

e
n
h
a
n
c
e
m
e
n
t
.

F
o
r

around the problem being addressed and explaining the effectiveness of the solution can be beneficial. Drawing from research on data visualizations that convert data into shared visual stories could aid in integrating storytelling to enhance the completion of collective action [66].

From Slacktivism to Activism: Unraveling the Engagement Spectrum. GigSense's intelligent interface promotes solutions with LLM assistance, yet there can be a concern that workers might opt for copying suggestions instead of contributing thoughtfully [67]. To enhance collective action, we suggest a nuanced strategy: crafting systems that employ LLMs to cater to diverse engagement levels [2, 29]. Certain gig workers may lean towards robust collaboration, while others could lean towards limited involvement, or even face limitations stemming from disabilities or specific constraints [72]. Intelligent interfaces could be tailored to facilitate deep reflections for committed individuals and enable less engaged ones to send quick supportive messages to their more dedicated counterparts. Additionally, LLM-integrated interfaces could support specific tasks and roles, similar to Wikipedia's SuggestBot [22].

Collaborative Problem Solving with Human-AI Interaction. Our system introduced a collaborative problem-solving process that integrated human-AI interactions. This approach complements prior research on LLMs' assistance in enhancing human's creativity [51, 64, 109], and highlights the design of interfaces to empower non-experts to utilize LLM technology for collective problem-solving. Our results reveal that LLMs supported gig workers (non-experts in the technology) in generating solutions, but our human-AI design ensured workers did not rely solely on the LLM output. Instead, workers used it to complement human-generated content, considering LLM suggestions as one of many sources they could incorporate. For this purpose, we strategically positioned LLM outputs below human-generated content and provided disclaimers about their reliability. Unlike previous studies [47], our participants welcomed LLM suggestions, incorporating it into their sensemaking process for creating solutions that improved their collective action plans. Nonetheless, unexpected LLM outcomes could potentially hinder workers' sensemaking and solution production. Future research could explore new human-AI interfaces for addressing problematic LLM outcomes, as well as study interface designs that prioritize different types of solutions based on workers' needs, e.g., novel solutions vs feasible solutions. Notice that the design of the human-AI interactions could influence the nature of generated solutions. Future research should consider recent studies on designing interactive interfaces to explain large language model responses [52, 56]. This transparency can enhance collaboration between end-users and AI-generated solutions.

Another direction is to investigate online, hybrid, and in-person collaborations, as well as asynchronous and synchronous interactions for problem-solving through Human-AI Interactions. Inspired by previous research [89, 120], future studies could investigate how in-person vs. online collaborations, coupled with human-centered AI, impact solution quality[104]. In the realm of collaborative problem-solving, researchers might explore physical robots integrating large language models to assist in brainstorming and addressing issues within the local physical space [11, 12, 87].

Uncertainty in Problem Solving with AI. The intrinsic uncertainty accompanying AI processes, stemming from their inherently probabilistic nature [43, 86], is a critical factor to consider within the design process. Designing interfaces capable of conveying this uncertainty to end-users is important for cultivating better interactions with the interface [112], and in the case of Gigsense, facilitating the production of enhanced solutions.

In our study, gig workers employed a Language Model (LM) which was not fine-tuned to the specific context of summarizing workers' issues or proposing solutions. This likely yielded solutions that were less than optimal, and could lead to unforeseen (uncertain) outcomes. Yet, these very uncertainties could also act as catalysts for fostering innovation, empowering workers to devise novel solutions that transcend conventional boundaries [15, 32]. Uncertainty has the potential to amplify the creativity capacities of workers [53]. Leveraging "imperfect AI models" in conjunction with interactive interfaces to enhance human creativity is a compelling avenue for future exploration.

Embracing uncertainty into the interface design has the potential to yield distinctive and unparalleled solutions [81]. However, it is unclear whether this approach enhances the production of feasible solutions by gig workers. Striking an optimal balance between feasibility and innovation poses a noteworthy challenge [10]. Future research could delve into the effective harnessing of uncertainty, aiming to foster both innovation and feasibility in the generation of solutions.

6.0.1 Problems and Biases of Interfaces Powered with LLMs. Employing LLMs demands substantial computational resources, leading to energy-intensive processes [50]. For instance, the training of GPT-3 with 175B parameter consumed considerable compute during pre-training [71], far exceeding the needs of a 1.5B parameter GPT-2 model [9, 125]. This necessitates a thorough evaluation of cost and efficiency when integrating LLMs into the human-centered AI interfaces we design [5, 34]. It is crucial to assess the resources needed for training LLMs and the distribution of such resources across the model's lifecycle, as well as assessing the amount of computation that is needed for fine-tuning the LLMs models, if needed [88]. However, it is important to recognize that despite the significant training costs, large pre-trained models like GPT-3 have demonstrated being highly efficient post- training [76]. For example, generating 100 pages of content from a trained GPT-3 model, incurs a cost of around
0.4 kW-hr, translating to just a few cents in energy expenses [14]. Furthermore, techniques like "'model distillation" can further economize the costs [24, 75]. Notice that model distillation, in the context of LLMs, involves transferring knowledge and information from a larger, complex LLM to a smaller, more efficient model [127]. This empowers the smaller model to execute tasks with fewer computational resources and memory while still capitalizing on the expertise embedded within the larger model. Given the advancements in image recognition and neural networks, we expect that algorithmic progress will improve LLM efficiency over time [49, 79].

It is also pivotal to acknowledge that while LLMs excel in generating solutions and summarizing problems, they are not exempt from limitations. The solutions and summaries that LLMs generate stem from learned patterns within existing data, potentially perpetuating biases or flawed assumptions present in their training data [33, 62, 103]. Recent research underscores the need for active intervention to address biases in LLMs [19, 114]. Just identification of the biases is not sufficient. Substantial work remains, particularly in effectively communicating potential biases [44]. We advocate for a holistic approach to bias mitigation rather than a singular focus on metric-based "elimination," as this approach has inherent limitations [28, 78]. Future endeavors could explore transparent interfaces communicating biases to stakeholders, aiding in informed decision-making and ultimately enhancing users' success in integrating LLMs into their workflows [77, 98].

7 CONCLUSION

In this paper, we presented GigSense, a system that facilitates sensemaking to support gig workers to initiate collective action. At the heart of GigSense's intellectual merit is the creation of systems that enable workers to gain a comprehensive understanding of the challenges they collectively face, assess their severity, and devise effective plans to address them together. Our between-subject evaluation study revealed that GigSense helped participants generate more feasible solutions for various problems in less time as compared to our control condition group. Participants in the GigSense group also reported significantly enhanced usability experience while using the system. We utilized a mixed-method approach in our study design. Our qualitative feedback revealed that the various features and functionalities of our system were well received by participants. Participants valued how GigSense enhanced the sensemaking of problems they faced, enabled solution generation, drove a collaborative problem-solving approach, and prioritized human-generated solutions, while still integrating the AI solution to assist workers.

Limitations. Our study has limitations: We recruited actual gig workers as participants who have diverse skills and come from diverse geographical backgrounds, which introduces potential individual differences that could impact generalizability. To mitigate these differences, we established inclusion criteria, requiring participants to have a minimum of one year of gig platform experience. This ensured they possessed an understanding of the challenges in gig work, enabling meaningful engagement and evaluation of our system. Furthermore, we integrated real-world complaint data about gig platforms and actual conversations of gig workers to contextualize the challenges faced by gig workers more broadly than what our specific participants experienced. It may also not be feasible to address all gig worker challenges, even with the help of systems like GigSense. Future research could categorize these issues into those with and without clear solutions. It could explore the potential of using LLMs and interactive interfaces to improve problem-solving, particularly for complex issues. A longitudinal study with prolonged participant interaction could have offered more behavioral data on social interactions facilitated by our system. However, due to the limited support for gig worker collective action, our system and study remain significant. Upon publication, we will open-source our platform, allowing the scientific community to conduct longitudinal studies on intelligent systems for collective action, potentially benefiting other actors in their collective endeavors. We also acknowledge large language models (LLMs) can generate erroneous solutions which can lead to unintended consequences. To address this, our system prioritizes human-generated solutions over AI suggestions. We also incorporated a disclaimer warning users of potential bias and error. The LLM responses serve merely as hints to guide workers, who remain the final decision-makers. Democratizing LLM tools for gig workers represents a significant step toward leveling the playing field in the gig economy. By providing gig workers with access to valuable resources, decision-making support, and collective action capabilities, these tools empower them to thrive in their independent roles while advocating for fairer working conditions and greater autonomy. Similarly, GigSense employs LLMs to democratize these new technologies and empower workers.

Acknowledgements This work was partially supported by the NSF CAREER award #2339443.

REFERENCES

[1] Amal Alabdulkarim, Siyan Li, and Xiangyu Peng. 2021. Automatic Story Generation: Challenges and Attempts. *NAACL HLT 2021* (2021), 72.

[2] Mariam Asad and Christopher A Le Dantec. 2015. Illegitimate civic participation: supporting community activists on the ground. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. 1694–1703.

[3] Susan J Ashford, Brianna Barker Caza, and Erin M Reid. 2018. From surviving to thriving in the gig economy: A research agenda for individuals in the new world of work. *Research in Organizational Behavior* 38 (2018), 23–41.

[4] Aaron Bangor, Philip T Kortum, and James T Miller. 2008. An empirical evaluation of the system usability scale. *Intl. Journal of Human–Computer Interaction* 24, 6 (2008), 574–594.

[5] Emily M Bender, Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. 2021. On the dangers of stochastic parrots: Can language models be too big?. In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*. 610–623.

[6] Yochai Benkler. 2017. Peer production, the commons, and the future of the firm. *Strategic Organization* 15, 2 (2017), 264–274.

[7] Janine Berg. 2015. Income security in the on-demand economy: Findings and policy lessons from a survey of crowdworkers. *Comp. Lab. L. & Pol'y J.* 37 (2015), 543.

[8] Ioulia Bessa, Simon Joyce, Denis Neumann, Mark Stuart, Vera Trappmann, and Charles Umney. 2022. *A global analysis of worker protest in digital labour platforms*. Vol. 70. International Labour Organization.

[9] Stella Biderman, Hailey Schoelkopf, Quentin Gregory Anthony, Herbie Bradley, Kyle O'Brien, Eric Hallahan, Mohammad Aflah Khan, Shivanshu Purohit, USVSN Sai Prashanth, Edward Raff, et al. 2023. Pythia: A suite for analyzing large language models across training and scaling. In *International Conference on Machine Learning*. PMLR, 2397–2430.

[10] Volker Bilgram and Felix Laarmann. 2023. Accelerating Innovation with Generative AI: AI-augmented Digital Prototyping and Innovation Methods. *IEEE Engineering Management Review* (2023).

[11] Andriana Boudouraki, Joel E Fischer, Stuart Reeves, and Sean Rintel. 2021. " I can't get round" Recruiting Assistance in Mobile Robotic Telepresence. *Proceedings of the ACM on Human-computer Interaction* 4, CSCW3 (2021), 1–21.

[12] Andriana Boudouraki, Joel E Fischer, Stuart Reeves, and Sean Rintel. 2023. " Being in on the Action" in Mobile Robotic Telepresence: Rethinking Presence in Hybrid Participation. In *Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*. 63–71.

[13] John Brooke. 1996. Sus: a "quick and dirty'usability. *Usability evaluation in industry* 189, 3 (1996), 189–194.

[14] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. 2020. Language models are few-shot learners. *Advances in neural information processing systems* 33 (2020), 1877–1901.

[15] Rodrigo F Cádiz, Agustín Macaya, Manuel Cartagena, and Denis Parra. 2021. Creativity in generative musical networks: evidence from two case studies. *Frontiers in Robotics and AI* 8 (2021), 680586.

[16] Dan Calacci. 2022. Organizing in the End of Employment: Information Sharing, Data Stewardship, and Digital Workerism. In *2022 Symposium on Human-Computer Interaction for Work* (Durham, NH, USA) *(CHIWORK 2022)*. Association for Computing Machinery, New York, NY, USA, Article 14, 9 pages. https://doi.org/10.1145/3533406.3533424

[17] Joel Chan, Steven Dang, and Steven P Dow. 2016. Improving crowd innovation with expert facilitation. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. ACM, 1223–1235.

[18] Julie Yujie Chen, Alessandro Delfanti, and Michelle Phan. 2023. Worker Resistance in Digital Capitalism| Worker Resistance in Digital Capitalism—Introduction. *International Journal of Communication* 17 (2023), 8.

[19] Mark Chen, Jerry Tworek, Heewoo Jun, Qiming Yuan, Henrique Ponde de Oliveira Pinto, Jared Kaplan, Harri Edwards, Yuri Burda, Nicholas Joseph, Greg Brockman, et al. 2021. Evaluating large language models trained on code. *arXiv preprint arXiv:2107.03374* (2021).

[20] Justin Cheng and Michael Bernstein. 2014. Catalyst: triggering collective action with thresholds. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*. ACM, 1211–1221.

[21] John Joon Young Chung, Wooseok Kim, Kang Min Yoo, Hwaran Lee, Eytan Adar, and Minsuk Chang. 2022. TaleBrush: Sketching stories with generative pretrained language models. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–19.

[22] Dan Cosley, Dan Frankowski, Loren Terveen, and John Riedl. 2007. SuggestBot: using intelligent task routing to help people find work in wikipedia. In *Proceedings of the 12th international conference on Intelligent user interfaces*. 32–41.

[23] Rob Cross and Lee Sproull. 2004. More than an answer: Information relationships for actionable knowledge. *Organization science* 15, 4 (2004), 446–462.

[24] Sayantan Dasgupta, Trevor Cohn, and Timothy Baldwin. 2023. Cost-effective distillation of large language models. In *Findings of the Association for Computational Linguistics: ACL 2023*. 7346–7354.

[25] Valerio De Stefano. 2016. The rise of the" just-in time workforce": on demand work, crowdwork, and labor protection in the" gig economy". *Comparative labor law and policy journal* 37, 3 (2016), 461–471.

[26] Alfred T DeMaria. 2016. The Future of Internet Organizing. *Management Report for Nonunion Organizations* 39, 2 (2016), 3–4.

[27] Brenda Dervin. 1983. An overview of sense-making research: Concepts, methods, and results to date. (1983).

[28] Jwala Dhamala, Tony Sun, Varun Kumar, Satyapriya Krishna, Yada Pruksachatkun, Kai-Wei Chang, and Rahul Gupta. 2021. Bold: Dataset and metrics for measuring biases in open-ended language generation. In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*. 862–872.

[29] Jill P Dimond, Michaelanne Dye, Daphne LaRose, and Amy S Bruckman. 2013. Hollaback! The role of storytelling online in a social movement organization. In *Proceedings of the 2013 conference on Computer supported cooperative work*. 477–490.

[30] V.B. Dubal. 2017. The Drive to Precarity: A Political History of Work, Regulation, Labor Advocacy in San Francisco's Taxi Uber Economies. *Berkeley Journal of Employment and Labor Law* 38, 1 (2017), 73–135. http://www.jstor.org/stable/26356922

[31] Jimmy Efird. 2011. Blocked randomization with randomly selected block sizes. *International journal of environmental research and public health* 8, 1 (2011), 15–20.

[32] Jason K Eshraghian. 2020. Human ownership of artificial creativity. *Nature Machine Intelligence* 2, 3 (2020), 157–160.

[33] Emilio Ferrara. 2023. Should chatgpt be biased? challenges and risks of bias in large language models. *arXiv preprint arXiv:2304.03738* (2023).

[34] Chelsea Finn, Paul Christiano, Pieter Abbeel, and Sergey Levine. 2016. A connection between generative adversarial networks, inverse reinforcement learning, and energy-based models. *arXiv preprint arXiv:1611.03852* (2016).

[35] Claudia Flores-Saviaga, Shangbin Feng, and Saiph Savage. 2022. Datavoidant: An AI System for Addressing Political Data Voids on Social Media. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (2022), 1–29.

[36] Claudia Flores-Saviaga, Yuwen Li, Benjamin Hanrahan, Jeffrey Bigham, and Saiph Savage. 2020. The challenges of crowd workers in rural and urban America. In *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing*, Vol. 8. 159–162.

[37] Katy Ilonka Gero, Vivian Liu, and Lydia Chilton. 2022. Sparks: Inspiration for science writing using language models. In *Designing interactive systems conference*. 1002–1019.

[38] Sucheta Ghoshal and Amy Bruckman. 2019. The role of social computing technologies in grassroots movement building. *ACM Transactions on Computer-Human Interaction (TOCHI)* 26, 3 (2019), 1–36.

[39] Karan Girotra, Lennart Meincke, Christian Terwiesch, and Karl T Ulrich. 2023. Ideas are Dimes a Dozen: Large Language Models for Idea Generation in Innovation. *Available at SSRN 4526071* (2023).

[40] Karan Girotra, Christian Terwiesch, and Karl T Ulrich. 2010. Idea generation and the quality of the best idea. *Management science* 56, 4 (2010), 591–605.

[41] Mary L Gray and Siddharth Suri. 2019. *Ghost work: How to stop Silicon Valley from building a new global underclass*. Eamon Dolan Books.

[42] Rafael Grohmann, Mateus Mendonça, and Jamie Woodcock. 2023. Worker Resistance in Digital Capitalism| Communication and Work From Below: The Role of Communication in Organizing Delivery Platform Workers. *International Journal of Communication* 17 (2023), 19.

[43] Aditya Grover. 2023. Generative Decision Making Under Uncertainty. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 37. 15440–15440.

[44] Felix Hamborg. 2022. Frame Analysis. In *Revealing Media Bias in News Articles: NLP Techniques for Automated Frame Analysis*. Springer, 109–140.

[45] Kotaro Hara, Abi Adams, Kristy Milland, Saiph Savage, Chris Callison-Burch, and Jeffrey P. Bigham. 2018. A Data-Driven Analysis of Workers' Earnings on Amazon Mechanical Turk. In *SIGCHI Conference on Human Factors in Computing Systems* (Montreal, Canada) *(CHI '18)*. ACM, New York, NY, USA, 14 pages. https://arxiv.org/abs/1712.05796

[46] Kotaro Hara, Abigail Adams, Kristy Milland, Saiph Savage, Benjamin V Hanrahan, Jeffrey P Bigham, and Chris Callison-Burch. 2019. Worker demographics and earnings on amazon mechanical turk: An exploratory analysis. In *Extended abstracts of the 2019 chi conference on human factors in computing systems*. 1–6.

[47] Stefan Harrer. 2023. Attention is not all you need: the complicated case of ethically using large language models in healthcare and medicine. *EBioMedicine* 90 (2023).

[48] Marti A Hearst and Duane Degler. 2013. Sewing the seams of sensemaking: A practical interface for tagging and organizing saved search results. In *Proceedings of the symposium on human-computer interaction and information retrieval*. 1–10.

[49] Danny Hernandez and Tom B Brown. 2020. Measuring the algorithmic efficiency of neural networks. *arXiv preprint arXiv:2005.04305* (2020).

[50] Jordan Hoffmann, Sebastian Borgeaud, Arthur Mensch, Elena Buchatskaya, Trevor Cai, Eliza Rutherford, Diego de Las Casas, Lisa Anne Hendricks, Johannes Welbl, Aidan Clark, et al. 2022. Training compute-optimal large language models. *arXiv preprint arXiv:2203.15556* (2022).

[51] Tom Hope, Ronen Tamari, Daniel Hershcovich, Hyeonsu B Kang, Joel Chan, Aniket Kittur, and Dafna Shahaf. 2022. Scaling Creative Inspiration with Fine-Grained Functional Aspects of Ideas. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–15.

[52] Ziheng Huang, Kexin Quan, Joel Chan, and Stephen MacNeil. 2023. CausalMapper: Challenging designers to think in systems with Causal Maps and Large Language Model. In *Proceedings of the 15th Conference on Creativity and Cognition*. 325–329.

[53] Angel Hsing-Chi Hwang. 2022. Too late to be creative? AI-empowered tools in creative processes. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts*. 1–9.

[54] Lilly C Irani and M Silberman. 2013. Turkopticon: Interrupting worker invisibility in amazon mechanical turk. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 611–620.

[55] Ziwei Ji, Nayeon Lee, Rita Frieske, Tiezheng Yu, Dan Su, Yan Xu, Etsuko Ishii, Ye Jin Bang, Andrea Madotto, and Pascale Fung. 2023. Survey of hallucination in natural language generation. *Comput. Surveys* 55, 12 (2023), 1–38.

[56] Peiling Jiang, Jude Rayan, Steven P Dow, and Haijun Xia. 2023. Graphologue: Exploring Large Language Model Responses with Interactive Diagrams. *arXiv preprint arXiv:2305.11473* (2023).

[57] Hannah Johnston. 2020. Labour geographies of the platform economy: Understanding collective organizing strategies in the context of digitally mediated work. *International Labour Review* 159, 1 (2020), 25–45.

[58] Hannah Johnston and Chris Land-Kazlauskas. 2018. Organizing on-demand: Representation, voice, and collective bargaining in the gig economy. (2018).

[59] Sean Kandel, Jeffrey Heer, Catherine Plaisant, Jessie Kennedy, Frank Van Ham, Nathalie Henry Riche, Chris Weaver, Bongshin Lee, Dominique Brodbeck, and Paolo Buono. 2011. Research directions in data wrangling: Visualizations and transformations for usable and credible data. *Information Visualization* 10, 4 (2011), 271–288.

[60] Vera Khovanskaya, Lynn Dombrowski, Jeffrey Rzeszotarski, and Phoebe Sengers. 2019. The Tools of Management: Adapting Historical Union Tactics to Platform-Mediated Labor. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 208 (nov 2019), 22 pages. https://doi.org/10.1145/3359310

[61] Vera Khovanskaya, Phoebe Sengers, and Lynn Dombrowski. 2020. Bottom-Up organizing with tools from on high: Understanding the data practices of labor organizers. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–13.

[62] Hannah Rose Kirk, Yennie Jun, Filippo Volpin, Haider Iqbal, Elias Benussi, Frederic Dreyer, Aleksandar Shtedritski, and Yuki Asano. 2021. Bias out-of-the-box: An empirical analysis of intersectional occupational biases in popular generative language models. *Advances in neural information processing systems* 34 (2021), 2611–2624.

[63] Aniket Kittur. 2010. Crowdsourcing, collaboration and creativity. *XRDS: crossroads, the ACM magazine for students* 17, 2 (2010), 22–26.

[64] Chinmay Kulkarni, Steven P Dow, and Scott R Klemmer. 2013. Early and repeated exposure to examples improves creative work. In *Design thinking research: Building innovation eco-systems*. Springer, 49–62.

[65] Andrew Kuznetsov, Joseph Chee Chang, Nathan Hahn, Napol Rachatasumrit, Bradley Breneisen, Julina Coupland, and Aniket Kittur. 2022. Fuse: In-Situ Sensemaking Support in the Browser. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology*. 1–15.

[66] Bongshin Lee, Nathalie Henry Riche, Petra Isenberg, and Sheelagh Carpendale. 2015. More than telling a story: Transforming data into visually

shared stories. *IEEE computer graphics and applica*

015), 84–90.

t
i
o
n
s
3
5
,
5
(
2

[67] Yu-Hao Lee and Gary Hsieh. 2013. Does slacktivism hurt activism? The effects of moral balancing and consistency in online activism. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 811–820.

[68] Vili Lehdonvirta. 2018. Flexibility in the gig economy: managing time on three online piecework platforms. *New Technology, Work and Employment* 33, 1 (2018), 13–29.

[69] Vili Lehdonvirta. 2022. *Cloud empires: How digital platforms are overtaking the state and how we can regain control*. MIT Press.

[70] James R Lewis. 2018. The system usability scale: past, present, and future. *International Journal of Human–Computer Interaction* 34, 7 (2018), 577–590.

[71] Feng Li, Jingxian Chen, and Xuejun Zhang. 2023. A Survey of Non-Autoregressive Neural Machine Translation. *Electronics* 12, 13 (2023), 2980.

[72] Hanlin Li, Disha Bora, Sagar Salvi, and Erin Brady. 2018. Slacktivists or activists? Identity work in the virtual disability march. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–13.

[73] Hanlin Li, Nicholas Vincent, Stevie Chancellor, and Brent Hecht. 2023. The Dimensions of Data Labor: A Road Map for Researchers, Activists, and Policymakers to Empower Data Producers. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*. 1151–1161.

[74] Tianyi Li, Kurt Luther, and Chris North. 2018. Crowdia: Solving mysteries with crowdsourced sensemaking. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 1–29.

[75] Kevin J Liang, Weituo Hao, Dinghan Shen, Yufan Zhou, Weizhu Chen, Changyou Chen, and Lawrence Carin. 2020. Mixkd: Towards efficient distillation of large-scale language models. *arXiv preprint arXiv:2011.00593* (2020).

[76] Percy Liang, Rishi Bommasani, Tony Lee, Dimitris Tsipras, Dilara Soylu, Michihiro Yasunaga, Yian Zhang, Deepak Narayanan, Yuhuai Wu, Ananya Kumar, et al. 2022. Holistic evaluation of language models. *arXiv preprint arXiv:2211.09110* (2022).

[77] Q Vera Liao and Jennifer Wortman Vaughan. 2023. AI Transparency in the Age of LLMs: A Human-Centered Research Roadmap. *arXiv preprint arXiv:2306.01941* (2023).

[78] Ruibo Liu, Chenyan Jia, Jason Wei, Guangxuan Xu, Lili Wang, and Soroush Vosoughi. 2021. Mitigating political bias in language models through reinforced calibration. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 35. 14857–14866.

[79] Xiaolong Liu, Zhidong Deng, and Yuhan Yang. 2019. Recent progress in semantic image segmentation. *Artificial Intelligence Review* 52 (2019), 1089–1106.

[80] I Mandl, M Curtarelli, S Riso, O Vargas Llave, and E Gerogiannis. 2015. New forms of employment. Luxembourg: Publications Office of the European Union. 160 p.

[81] Marcello Mariani. 2022. Generative Artificial Intelligence and Innovation: Conceptual Foundations. *Available at SSRN 4249382* (2022).

[82] Lars Gunnar Mattsson, Daniela Corsaro, and Carla Ramos. 2015. Sense-making in business markets–the interplay between cognition, action and outcomes. *Industrial Marketing Management* 48 (2015), 4–11.

[83] Paul Mihas. 2019. Qualitative data analysis. In *Oxford research encyclopedia of education*.

[84] Kate Minter. 2017. Negotiating labour standards in the gig economy: Airtasker and Unions New South Wales. *The Economic and Labour Relations Review* 28, 3 (2017), 438–454.

[85] Piotr Mirowski, Kory W Mathewson, Jaylen Pittman, and Richard Evans. 2023. Co-Writing Screenplays and Theatre Scripts with Language Models: Evaluation by Industry Professionals. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–34.

[86] Michael Muller, Lydia B Chilton, Anna Kantosalo, Charles Patrick Martin, and Greg Walsh. 2022. GenAICHI: generative AI and HCI. In *CHI conference on human factors in computing systems extended abstracts*. 1–7.

[87] Prasanth Murali, Ian Steenstra, Hye Sun Yun, Ameneh Shamekhi, and Timothy Bickmore. 2023. Improving multiparty interactions with a robot using large language models. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–8.

[88] MZ Naser. 2023. Do We Need Exotic Models? Engineering Metrics to Enable Green Machine Learning from Tackling Accuracy-Energy Trade-offs. *Journal of Cleaner Production* 382 (2023), 135334.

[89] Thomas Neumayr, Banu Saatci, Sean Rintel, Clemens Nylandsted Klokmose, and Mirjam Augstein. 2021. What was hybrid? a systematic review of hybrid collaboration and meetings research. *arXiv preprint arXiv:2111.06172* (2021).

[90] Mancur Olson. 1965. The logic of collective action [1965]. *Contemporary Sociological Theory* 124 (1965).

[91] OpenAI. 2023. OpenAI Platform — platform.openai.com. https://platform.openai.com/docs/api-reference. [Accessed 13-09-2023].

[92] Carol Packham. 2008. *Active citizenship and community learning*. Learning Matters.

[93] Maria-Carmen Pantea. 2013. The changing nature of volunteering and the cross-border mobility: where does learning come from? *Studies in Continuing Education* 35, 1 (2013), 49–64.

[94] Sharoda A Paul and Meredith Ringel Morris. 2009. CoSense: enhancing sensemaking for collaborative web search. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 1771–1780.

[95] Jamie Peck and Nik Theodore. 2012. Politicizing contingent work: Countering neoliberal labor market regulation... from the bottom up? *South Atlantic Quarterly* 111, 4 (2012), 741–761.

[96] S Camille Peres, Tri Pham, and Ronald Phillips. 2013. Validation of the system usability scale (SUS) SUS in the wild. In *Proceedings of the human factors and ergonomics society annual meeting*, Vol. 57. SAGE Publications Sage CA: Los Angeles, CA, 192–196.

[97] Pirolli Peter and Stuart Card. 2005. The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. *Proceedings of international conference on intelligence analysis* (2005).

[98] Aleksandra Piktus, Christopher Akiki, Paulo Villegas, Hugo Laurençon, Gérard Dupont, Alexandra Sasha Luccioni, Yacine Jernite, and Anna Rogers. 2023. The roots search tool: Data transparency for llms. *arXiv preprint arXiv:2302.14035* (2023).

[99] Peter Pirolli and Stuart Card. 2005. The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. In *Proceedings of international conference on intelligence analysis*, Vol. 5. McLean, VA, USA, 2–4.

[100] Frances Fox Piven and Richard Cloward. 2012. *Poor people's movements: Why they succeed, how they fail*. Vintage.

[101] Rida Qadri. 2023. Worker Resistance in Digital Capitalism| Algorithmized Not Atomized: The Distributed Solidarity of Jakarta's Gig Workers. *International Journal of Communication* 17 (2023), 20.

[102] Yolanda Jacobs Reimer, Matthew Hagedal, Peter Wolf, and Bradley Bahls. 2011. Turning the desktop inside-out: Evaluating information access and management through a single interface. *Journal of the American Society for Information Science and Technology* 62, 12 (2011), 2327–2346.

[103] David Rozado. 2023. The political biases of chatgpt. *Social Sciences* 12, 3 (2023), 148.

[104] Banu Saatçi, Roman Rädle, Sean Rintel, Kenton O'Hara, and Clemens Nylandsted Klokmose. 2019. Hybrid meetings in the modern workplace: stories of success and failure. In *Collaboration Technologies and Social Computing: 25th International Conference, CRIWG+ CollabTech 2019, Kyoto, Japan, September 4–6, 2019, Proceedings 25*. Springer, 45–61.

[105] Niloufar Salehi, Lilly C Irani, Michael S Bernstein, Ali Alkhatib, Eva Ogbe, Kristy Milland, and Clickhappier. 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.

[106] Peter Kalum Schou and Eliane Bucher. 2022. Divided we fall: The breakdown of gig worker solidarity in online communities. *New Technology, Work and Employment* (2022).

[107] Cyrille Schwellnus, Assaf Geva, Mathilde Pak, and Rafael Veiel. 2019. Gig economy platforms: Boon or Bane? (2019).

[108] Aaron Shaw, Haoqi Zhang, Andres Monroy-Hernandez, Sean Munson, Benjamin Mako Hill, Elizabeth Gerber, Peter Kinnaird, and Patrick Minder. 2014. Computer supported collective action. *interactions* 21, 2 (2014), 74–77.

[109] Pao Siangliulue, Kenneth C Arnold, Krzysztof Z Gajos, and Steven P Dow. 2015. Toward collaborative ideation at scale: Leveraging ideas from others to generate more creative and diverse ideas. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. 937–945.

[110] Jack W Snook and Dan C Olsen. 2006. *Recruiting, training, and maintaining volunteer fire fighters*. Jones & Bartlett Learning.

[111] Arvind Srinivasan and Joel Chan. 2023. AnalogiLead: Improving Selection of Analogical Inspirations with Chunking and Recombination. In *Proceedings of the 15th Conference on Creativity and Cognition* (Virtual Event, USA) *(C&C '23)*. Association for Computing Machinery, New York, NY, USA, 338–341. https://doi.org/10.1145/3591196.3596817

[112] Jiao Sun, Q Vera Liao, Michael Muller, Mayank Agarwal, Stephanie Houde, Kartik Talamadupula, and Justin D Weisz. 2022. Investigating explainability of generative AI for code through scenario-based design. In *27th International Conference on Intelligent User Interfaces*. 212–228.

[113] Will Sutherland and Mohammad Hossein Jarrahi. 2017. The gig economy and information infrastructure: The case of the digital nomad community. *Proceedings of the ACM on human-computer interaction* 1, CSCW (2017), 1–24.

[114] Himanshu Thakur, Atishay Jain, Praneetha Vaddamanu, Paul Pu Liang, and Louis-Philippe Morency. 2023. Language Models Get a Gender Makeover: Mitigating Gender Bias with Few-Shot Data Interventions. *arXiv preprint arXiv:2306.04597* (2023).

[115] Sukrit Venkatagiri, Jacob Thebault-Spieker, Rachel Kohler, John Purviance, Rifat Sabbir Mansur, and Kurt Luther. 2019. GroundTruth: Augmenting expert image geolocation with crowdsourcing and shared representations. *Proceedings of CSCW'19* 3 (2019), 1–30.

[116] Gwendolyn Paige Watson, Lauren D Kistler, Baylor A Graham, and Robert R Sinclair. 2021. Looking at the gig picture: Defining gig work and explaining profile differences in gig workers' job demands and resources. *Group & Organization Management* 46, 2 (2021), 327– 361.

[117] Juliet Webster. 2016. Microworkers of the gig economy: Separate and precarious. In *New Labor Forum*, Vol. 25. SAGE Publications Sage CA: Los Angeles, CA, 56–64.

[118] KE Weick. 1995. Sensemaking in organizations Sage Publications. *Thousand Oaks, CA* (1995).

[119] Alex Wood and Vili Lehdonvirta. 2019. Platform labour and structured antagonism: Understanding the origins of protest in the gig economy. *Available at SSRN 3357804* (2019).

[120] Tianna Xu, Advait Sarkar, and Sean Rintel. 2023. Is a Return To Office a Return To Creativity? Requiring Fixed Time In Office To Enable Brainstorms and Watercooler Talk May Not Foster Research Creativity. (2023).

[121] Zheng Yao, Silas Weden, Lea Emerlyn, Haiyi Zhu, and Robert E Kraut. 2021. Together but alone: Atomization and peer support among gig workers. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), 1–29.

[122] Ann Yuan, Andy Coenen, Emily Reif, and Daphne Ippolito. 2022. Wordcraft: story writing with large language models. In *27th International Conference on Intelligent User Interfaces*. 841–852.

[123] JD Zamfirescu-Pereira, Richmond Y Wong, Bjoern Hartmann, and Qian Yang. 2023. Why Johnny can't prompt: how non-AI experts try (and fail) to design LLM prompts. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–21.

[124] Haoqi Zhang, Andrés Monroy-Hernández, Aaron D Shaw, Sean A Munson, Elizabeth M Gerber, Benjamin Mako Hill, Peter Kinnaird, Shelly Diane Farnham, and Patrick Minder. 2014. WeDo: End-To-End Computer Supported Collective Action.. In *ICWSM*.

[125] Shengyu Zhang, Linfeng Dong, Xiaoya Li, Sen Zhang, Xiaofei Sun, Shuhe Wang, Jiwei Li, Runyi Hu, Tianwei Zhang, Fei Wu, et al. 2023. Instruction Tuning for Large Language Models: A Survey. *arXiv preprint arXiv:2308.10792* (2023).

[126] Xiaolong Zhang, Yan Qu, C Lee Giles, and Piyou Song. 2008. CiteSense: supporting sensemaking of research literature. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 677–680.
[127] Sanqiang Zhao, Raghav Gupta, Yang Song, and Denny Zhou. 2019. Extreme language model compression with optimal subwords and shared projections. (2019).

A APPENDIX

We present the prompts we encoded in our system to generate the: (1) categorization of gig workers' problems; (2) summaries of their problems; and (3) solutions generated with OpenAI's GPT-4 API.

A.1 Prompts for categorizing Gig Workers' Problems

Our data-gathering module focuses on collecting data about different potential problems faced by gig workers. It has webscrappers to collect data from different gig worker subreddits (r/Freelancers, r/UpWork), as well as the reviews that people left on Apple's and Google's app store data about major gig platforms (particularly: Amazon Mechanical Turk, Fiver, UpWork, Taloka, and Freelancer). Note that our data gathering module only collects reviews that have between one and three-star ratings. The module considers that the review data would represent complaints and problems that gig workers were experiencing. Next, our module connects with OpenAI's API in the backend of our system to first categorize the data into different categories of gig work challenges. To obtain such categorization, We prompted:

A.2 Prompt for Summarization

"Please categorize the following data into specific problem categories. The goal is to sort the large chuck of data into distinct categories for better comprehension. You should have a problem category with a name, and a list of statements related to the problem. If you find a category or theme that is not a problem, do not include it"

A.3 Prompt for generating solutions

The categorized data was then again fed into the backend of our system where we had encoded the following prompt: "Given the following dataset of complaints related to X (where X represents an identified problem), please provide a concise summary of the problem to facilitate a comprehensive understanding."

A.4 LLM Response

Summary of Problem-related to Platform Policy The common theme among the listed 'Platform Policy' problems faced by gig workers is dissatisfaction with the platform's fees and policies. Gig workers are unhappy with the high fees, especially for newbies who cannot afford them, and the arbitrary increases in fees for more experienced freelancers. There are complaints about the cost of connects, commission charges, and withdrawal policies. Additionally, some gig workers are unhappy with the platform's lack of transparency regarding proposal views and analytics. Overall, these problems suggest that the platform prioritizes profits over the well-being of freelancers.

Solution suggested by LLM

(1) Negotiate platform fees: Gig workers can come together and negotiate with the platform for lower fees or a fairer fee structure. They can explain their challenges and how high fees affect their ability to earn a living. With a collective voice, they may be able to persuade the platform to revise its policies.

(2) Seek alternative platforms: Gig workers can explore other platforms that charge lower fees or offer more benefits. They can research and compare platforms to find the one that best suits their needs. Switching platforms may help gig workers find better-paying jobs and reduce the impact of high fees.
(3) Advocate for fair policies: Gig workers can form advocacy groups or join existing ones to lobby for fairer policies for all gig workers. They can work with policymakers to push for regulations that protect gig workers' rights and ensure a level playing field.
(4) Share knowledge and resources: Gig workers can collaborate and share knowledge and resources to improve their chances of finding better-paying jobs. They can form communities and support groups where they share tips, advice, and best practices for finding and landing jobs. By sharing their experiences, gig workers can help each other overcome common challenges.
(5) Encourage transparency - Gig platforms should encourage clients to be transparent in their job postings, payment processes, and communication with gig workers. Clients should disclose all the necessary details related to the job upfront, and payment processes should be clear and easy to understand. This will help gig workers to identify any red flags and avoid working with fraudulent clients.

Summary of Problem related Usability

This category of problems is related to the usability of gig work platforms. Gig workers are facing issues with the web version and mobile apps. Some workers are complaining about missing features or outdated versions of the app, leading to delays and incorrect data. Others report problems logging in, receiving error messages, and being unable to download or install the app. Several users have encountered issues with the messaging system, including delays in loading messages, formatting issues, and glitches that make it difficult to send and receive messages. Many workers have also reported difficulty applying for jobs, with some issues including the page refreshing after attaching files, lost text inputs, and difficulty retracting proposals. Additionally, there are reports of notifications being delayed or not working, as well as issues with the app crashing or freezing.

Solution suggested by LLM

(1) Collaborate with other gig workers: Gig workers can form groups or communities where they share information and experiences about different gig platforms. By doing this, they can help each other find the best- paying gigs and avoid low-paying ones. This collaboration can also help them negotiate better rates and resolve conflicts with clients.
(2) Leave feedback and reviews: Gig workers can leave honest feedback and reviews about the gig platforms they have used. This can help other gig workers make informed decisions when choosing which platform to work with. It can also encourage platforms to improve their services and address any issues that are negatively affecting workers.
(3) Advocate for better working conditions: Gig workers can come together and advocate for better working conditions, such as fair pay, better benefits, and protections. This can be done through social media campaigns, contacting policymakers, or even starting a union. By advocating for their rights, gig workers can help create positive change in the gig economy.
(4) Stay informed about platform policies: Gig workers should regularly check the policies and terms of service of the platforms they use. This can help them understand their rights and responsibilities, as well as any changes

that may affect their work. By staying informed, gig workers can avoid any negative surprises and make informed decisions.

(5) Diversify income sources: Gig workers can reduce their reliance on a single gig platform by diversifying their income sources. This can be done by working with multiple platforms, finding offiine work opportunities, or even starting their own business. By diversifying their income, gig workers can mitigate the risk of relying on a single platform and increase their overall earnings.

Summary of Problem-related Payment

This category includes payment-related problems faced by gig workers on a gig work platform. These issues include fraudulent clients who don't pay, difficulties withdrawing funds, high fees, delayed payment, and issues with pay- ment methods, such as exchange rates, tax information, and payment restrictions. Additionally, some freelancers are concerned about a change in policy that requires payment for connects or receiving a 1099-K form from the platform. Furthermore, some freelancers have trouble getting paid in their currency or complying with billing require- ments. The lack of effective customer support and communication channels from the platform also adds to the difficulties.

Solution suggested by LLM

(1) Communication is key: Gig workers should communicate with their clients before starting any project and clarify all the payment terms and conditions beforehand. They should ask for any external expenses and ensure that there are no surprises later on.
(2) Use secure payment methods: Gig workers should use secure payment methods that are trusted and reliable. They should also be aware of the exchange rates and any fees associated with withdrawing money to avoid any issues later on.
(3) Research the platform: Gig workers should research the gig work platform they intend to use and understand their fees and policies. They should also read reviews from other gig workers to ensure that they are using a reliable and trustworthy platform.
(4) Keep track of earnings and tax information: Gig workers should keep track of their earnings and ensure that their tax information is up-to-date and accurate. They should also ensure that their bank account information matches the name on their gig work platform profile to avoid any issues with withdrawals.
(5) Report fraudulent behavior: Gig workers should report any fraudulent behavior to the gig work platform's support team immediately. They should also keep all communication and evidence related to the project and payment to help the support team investigate and resolve the issue quickly.

Summary of Problem-related Poor Customer Support

This category is about the poor customer support experienced by gig workers on the gig work platform. The problems include fraud clients who don't pay for work and unhelpful dispute teams, difficulties with withdrawing funds, errors in the account creation process, incorrect account bans, lack of access to live chat support, and unresponsive customer service. There are also problems with account reinstatement, delayed project approvals, and difficulty in logging in due to system errors. In addition, the high fees and long payment processing times, as well as the requirement to pay for Connects, are frustrating for freelancers. The customer support team is also unhelpful in addressing the concerns of freelancers whose gigs are denied, removed, or flagged as inappropriate.

Solution suggested by LLM

(1) Protect yourself from fraud clients by always having a contract in place, with clearly defined milestones and deadlines. Use escrow payments or invoicing services to ensure that you receive payment for your work. In case of any dispute, provide clear evidence of the work done and communicate with the client politely, but firmly. If the dispute team is of no use, consider filing a complaint with the gig work platform or legal authorities.
(2) If you face technical issues while working on the gig platform, such as login problems or error messages, reach out to customer support immediately. If you do not receive a response within a reasonable time, try different channels of communication, such as email, phone, or social media. Provide detailed information about the issue, including screenshots and error messages, and follow up regularly until the issue is resolved.
(3) If the live chat feature is not available or you are unable to get a response from customer support, try to connect with other gig workers on online forums or social media groups. They may have faced similar issues and can provide valuable advice and support.
(4) If you are facing issues with reinstating your account or creating a new one, try to contact customer support through different channels and provide all the necessary information, such as your email address and phone number. If the support team is not helpful, consider switching to a different gig platform that has more user-friendly policies and procedures.
(5) If you are unhappy with the fees charged by the gig platform or the quality of customer support, consider exploring other gig platforms that offer better rates or customer service. Read reviews and compare the features and policies of different platforms before making a decision. You can also join a professional association or community of gig workers to stay updated on industry trends and best practices.

Summary of Problem-related Scam

Gig workers on gig work platforms face several scam-related problems, including fraud clients who don't pay, poor dispute resolution, and difficulty withdrawing funds. Other problems include receiving job offers outside the platform, downloading harmful files, fake job posts, unfair contract termination, and clients demanding high commissions after offering high wages. Furthermore, some clients use poor rating tactics to exploit workers, while others demand work at lower rates than the platform rate. Some clients also refuse to pay the required service fees and keep workers' earnings, while platforms charge excessive fees for services and tips.

Solution suggested by LLM

(1) Create a scam alert system - The gig platform can have an AI-based scam alert system that can track the behavior and patterns of the clients on the platform. The system can identify any suspicious behavior by analyzing the clients' work history, payment patterns, and other activities on the platform. If a client is identified as a potential scammer, the system can alert the gig workers before they start working with the client, saving them from any fraudulent activities.
(2) Improved dispute resolution - Gig platforms should have a more efficient and responsive dispute resolution mechanism. A 24/7 support system with dedicated dispute resolution teams should be available to the gig workers. This will help them to resolve any payment or work-related issues quickly, ensuring that they are not

scammed out of their hard work.

(3) Introduce a verification system - Platforms can verify the client's identity and payment information to ensure that they are genuine. This can be done by verifying their phone numbers, email addresses, and other personal information. The platform can also ask for a deposit from the client, which can be refunded after the gig is completed. This will help to prevent any fraudulent activities and give the gig workers a sense of security.
(4) Educate gig workers about potential scams - Gig platforms should provide educational materials to gig workers on how to identify and avoid potential scams. These materials can include videos, articles, and webinars that cover various types of scams and how to avoid them. This will empower gig workers with knowledge and make them more aware of potential scams.
(5) Encourage transparency - Gig platforms should encourage clients to be transparent in their job postings, payment processes, and communication with gig workers. Clients should disclose all the necessary details related to the job upfront, and payment processes should be clear and easy to understand. This will help gig workers to identify any red flags and avoid working with fraudulent clients.

A.5 System Usability Scale Template

Here are the questionaries for the System Usability Scale. It is measured on a scale of (1-5) where *1=Strongly Disagree, 2=Disagree, 3=Neutral, 4=Agree, 5=Strongly Agree*.

(1) I think that I would like to use this system frequently.
(2) I found the system unnecessarily complex.
(3) I thought the system was easy to use.
(4) I think that I would need the support of a technical person to be able to use this system.
(5) I found the various functions in this system were well-integrated.
(6) I thought there was too much inconsistency in this system.
(7) I would imagine that most people would learn to use this system very quickly.
(8) I found the system very cumbersome to use.
(9) I felt very confident using the system.
(10) I needed to learn a lot of things before I could get going with this system.

After calculating the composite scores of SUS (System Usability Scale) scores, they are compared to the following adjectival rating table [4].

SUS Score	Grade	Adjectival Rating
>80.3	A	Excellent
68-80.3	B	Good
68	C	Okay
51-68	D	Awful
<51	D	Poor

Table 2. Adjectival Rating for SUS scores.