Avances en Interacción Humano-Computadora, 2024, Volumen 9, No. 1, pp. 135-145
DOI: 10.47756/aihc.y9i1.159

RESEARCH PAPER

GigSense: An LLM-Infused Tool for Workers' Collective Intelligence

Kashif Imteyaz, Claudia Flores Saviaga, Saiph Savage

Published: 30 November 2024

Abstract
Collective intelligence among gig workers yields considerable advantages, including improved information exchange, deeper social bonds, and stronger advocacy for better labor conditions. Especially as it enables workers to collaboratively pinpoint shared challenges and devise optimal strategies for addressing these issues. However, enabling collective intelligence remains challenging, as existing tools often overestimate gig workers' available time and uniformity in analytical reasoning. To overcome this, we introduce GigSense, a tool that leverages large language models alongside theories of collective intelligence and sensemaking. GigSense enables gig workers to rapidly understand and address shared challenges effectively, irrespective of their diverse backgrounds. GigSense not only empowers gig workers but also opens new possibilities for supporting workers more broadly, demonstrating the potential of large language model interfaces to enhance collective intelligence efforts in the evolving workplace.

Keywords:
Human- centered computing; Human computer interaction (HCI); Interactive systems and tools, LLM, Empirical studies in collaborative and social computing

1. Introduction
Harnessing collective intelligence among gig workers can significantly enhance their ability to improve labor conditions [5, 13, 27, 28]. This approach allows workers to jointly address challenges, develop solutions, and implement action plans [3, 9] However, despite occasional successes, collective intelligence among gig workers is uncommon, leaving many issues unresolved [3, 10]. This scarcity largely stems from the lack of technologies designed for facilitating collective problem-solving among gig workers [12, 14, 17, 34, 4 5]. Platforms like Dynamo and Coworker.org, which allow for sharing, prioritizing, and solving issues through upvote-based lists, fail to provide a comprehensive view of problems and solutions [7, 29]. Such list-based interfaces limit the exploration of issues from multiple angles and understanding the full scope of potential solutions [4, 15, 18, 20, 26, 41], often leading workers to prioritize less critical concerns [42]. This impedes meaningful progress in addressing the fundamental challenges of gig work, especially during the initial phases of collective intelligence focused on recognizing shared problems and formulating solutions [33]. Another hurdle is that the diverse and significant time commitment workers may limit their ability to contribute to collective intelligence, underscoring the importance of developing inclusive tools that can accommodate workers' varying schedules, reducing the barriers to collective intelligence and enabling easier participation in collaboratively tackling challenges [24, 25, 99]. To address these challenges, we introduce GigSense, a novel platform for fostering collective intelligence among gig workers by enabling collaborative problem-solving. By integrating Sense-Making Theory and leveraging the capabilities of large-language models (LLMs), GigSense features an LLM-enhanced interface that allows work- ers to deeply analyze their challenges and solutions from multiple perspectives [26]. Unlike existing platforms that offer a simple list- based view of issues [29], GigSense enables a detailed exploration of problems, offering both a close-up and a broad overview to understand workplace dynamics better. Additionally, GigSense uses LLMs to facilitate collective brainstorming, helping workers develop solutions together. Fig 1 presents an overview of GigSense. As a significant advancement in fostering collective intelligence among gig workers, GigSense addresses the unique challenges of this workforce, which typically operates in isolation without a common platform for problem identification and solution generation [4, 43]. GigSense bridges this gap by offering a communal space for gig workers to collectively understand and address their issue. Leveraging shared digital resources to build community [1, 32], GigSense emerges as a vital tool for a fragmented workforce, potentially catalyzing the formation of supportive gig worker communities [6, 38, 40]. Beyond organization, it could reveal systemic issues, encouraging workers to see problems like client disputes and payment delays as collective challenges. This

Imteyaz K., Flores Saviaga C.
Northeastern University
Boston, MA, USA.
Email: {imteyaz.k, floressaviaga.c}@northeastern.edu

Savage S.
Northeastern University
Boston, MA, USA.
Email: s.savage@northeastern.edu

recognition can promote solidarity, crucial for collective intelligence, and empower workers to develop effective negotiation strategies, potentially improving their working conditions. In this paper, we highlight a system that integrates LLMs with an interactive interface to support gig workers in sensemaking for problem-solving and kickstart collective intelligence.

2. Related Work

Gig Work and Platform. “Gig" or platform-based work is a significant trend in the labor market, driven by the demand for flexibility from both employers and workers [105], and facilitated by digital technology [21, 46, 62]. While offering economic benefits to disadvantaged groups, gig work also presents challenges like unstable schedules, income variability, and uncertain long-term job security [21, 75]. Workers have engaged in collective actions (e.g., negotiations, strikes, unionization) to improve conditions, yet face obstacles due to platform constraints, geographic dispersion, and a lack of community and common interest, making organizing difficult [34, 46, 97, 106, 110]. Despite these challenges, some success has been noted in collective efforts [25], but tools to support organizing are limited [46]. This paper proposes a tool to aid gig workers in understanding their challenges and initiating collective intelligence.

Sensemaking and Collective Intelligence. 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 to reach a conclusion. Pirolli and Card [77] 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”). Smith [93] defines collective intelligence as a group of individuals working together on tasks, where the collective group itself demonstrates coherence, intelligence, and constant improvement, enabling more effective mobilization of skills than any single individual working independently. Significant research explores collaborative sensemaking tools for domains like literature review [114], web search [74], organizing academic literature [82], solving mysteries [57], and tackling disinformation [30]. We introduce GigSense to aid gig workers’ collaborative sensemaking, automating parts of the process pipeline. We also leverage Large Language Models to enhance collective intelligence.

Interfaces for Visualizing Collective Problems and Solutions. Other research that has inspired our work includes interfaces and systems designed to assist individuals in visualizing both problems and solutions. For example, MacNeil et al. [60], presented a design gallery for visualizing problems and the stakeholders involved in those problems. Similarly, Huang et al. [40] introduced a novel system to aid designers in comprehending various issues within space and exploring diverse solutions. Their research focused on mitigating design fixation by encouraging a broader consideration of context and essential relationships during the design process. Siangliulue et al.’s system [92] innovatively merged crowdsourcing with machine learning to create a semantic solution space model, which helped to foster user creativity and diversity of ideas. Our research is inspired by the principles of these systems, concentrating on assisting gig workers, who may lack expertise in design and problem-solving, in exploring and comprehending various aspects of problems and having support to provide solutions.

3. GigSense

GigSense is an AI platform to aid gig workers in sensemaking and collective intelligence for problem identification and solution proposal [34]. It incorporates sensemaking and collective intelligence theories in its design modules to automate the sensemaking process with workers steering it. GigSense has the following modules:

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 [77]). 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 scraper 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 workers’ complaint datasets (actual gig workers’ subreddits and complaints) in our system design aims to bring inclusiveness about gig worker concerns and complaints. Finally, the data is stored and sent to the backend, where prompt engineering techniques are applied for use in subsequent modules of our system.

Problem Summary Module: This module acting as “Step: Schematize” loop of sensemaking, summarizes the gathered data to aid exploration. GigSense uses LLMs to categorize and summarize the large complaint datasets into problem categories [23]. It has buttons to navigate to the Data Visualization module for more detailed or high-level views.

Data Viz Module: This uses semantic zooming to visualize problems at varying abstraction levels [2, 36, 37]. The zoomed-out view shows an interactive chart categorizing problems (e.g. Payment, Policy, Scams) using LLMs. Workers can hover over sections to see complaint volumes, raising awareness of shared challenges. The zoom-in view shows full text of complaint posts for deep analysis and upvoting, fostering collective identity [16, 31]. This is “Step: Schematize” in the sensemaking loop.

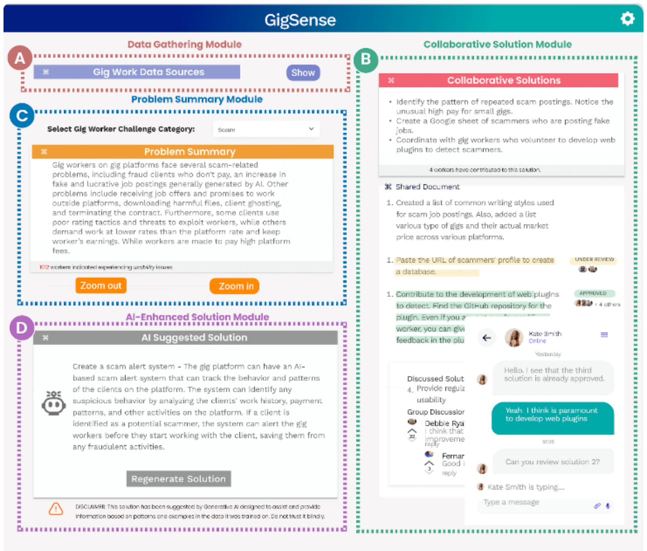

Figure 1: Overview of GigSense's problem-solving interface with the: A) "Data Gathering Module"; B) "Collaborative Solution Module" where workers can collaborate to create solutions to specific problems; C) "Problem Summary Module" where workers obtain summaries of specific problems to help their sensemaking; and D) "AI-Enhanced Solution Module" where workers obtain solution suggestions from LLMs

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 existing problems and create action plans (solutions) to address them. GigSense also includes a functionality that allows users to make annotations in the shared document. This way, 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 [23, 107] ("Step: Tell Story"in the sense-making loop). Further the system utilizes prompt engineering in the backend to summarize the approved solutions only via LLM. Note that this module supports collaborative work that initiates early stages of collective intelligence.[90].

AI-Enhanced Solution Module: It is important to acknowledge that for some workers it may be hard to propose solutions [7, 94]. 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. The system also provides nudge via disclaimers warning users of potential AI errors, promoting responsible AI use while still leveraging assistance. This module is designed to enhance the "Step: Build Case" and "Step: Tell Story"s tage of the sensemaking loop.

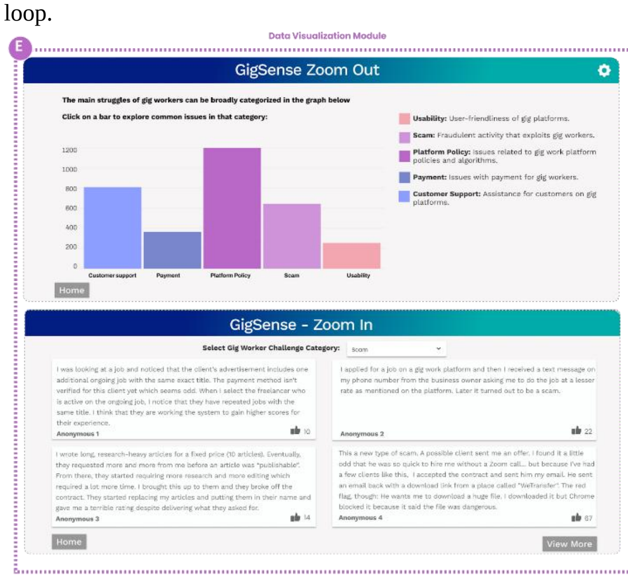

Figure 2: DataViz Module (Gigsense's zoom-out and zoom-in view)

4. Evaluation

The evaluation of GigSense aims to address key research questions:
1) Speed: Can GigSense facilitate rapid sensemaking, allowing gig workers to seamlessly contribute to the collective intelligence?
2) Contribution: Can GigSense amplify workers' contributions in sensemaking by enhancing problem identification and solution generation?
3) Usability: Does GigSense's AI-enhanced interactive interface bring better user experiences?

Participant Recruitment. To recruit participants, we generated a job listing on Upwork [103], 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). Our selection criteria ensured we have participants who understood the challenges in gig work and can communicate well in our study. 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 [26]. In the end, 12 participants were assigned to the control condition, (P1-P12), and 12 were assigned to the GigSense condition (P13-P24). Participants in our user study were compensated $10/hr for their participation.

Procedure. 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 collective intelligence: pinpointing collective issues and suggesting solutions. Participants in the Gigsense condition used our Gigsense platform to complete the tasks (see Fig 2 and 3, while participants in the control condition used an interface resembling gig platform community forums and features resembling "We Are

Dynamo” interface [83, 103]. We built “We Are Dynamo” to simulate the general functionality of the original system which is no longer available for use. In the control interface, users can engage in the features they would in the original version of ’We Are Dynamo’ (such as posting ideas for action and upvoting others’ ideas), see Fig. 3. Next, we compared the quality of solutions that were generated using GigSense and the control interface. We also studied the usability of GigSense in comparison to the control interface. It is also essential to recognize that human-AI interfaces do not automatically surpass traditional list-based ones in effectiveness. Therefore, it is not clear that the control interface will lead to inferior outcomes. The perceived superiority of AI-enhanced interfaces, often attributed to their informative nature, does not always translate to practical advantage. In fact, interfaces with less information can be preferable, as they help avoid cognitive overload and reduce complexity [3, 48, 98]. Their simplicity, coupled with lower training requirements and ease of use, can make list-based interfaces particularly beneficial in dynamic, time-sensitive work environments [42, 102]. Therefore, in certain scenarios, opting for a streamlined and less informative interface can be a more practical and efficient choice [18, 61]. Acknowledging the uncertainty surrounding the most effective design for this scenario, we initiated our user study. 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 the idea gets enough upvotes, it turns into a 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 [52], which might hinder synchronous collaboration. Our aim was thus to ensure effective asynchronous utilization of our tool for seamless completion of collective intelligence 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 intelligence [7, 19, 90, 113]. (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 on workers. (4) Propose solutions to the three problems you identified that were crucial to be addressed. (5) Propose a solution to a problem raised by another gig worker. (6) Propose three solutions to problems raised by other gig workers.

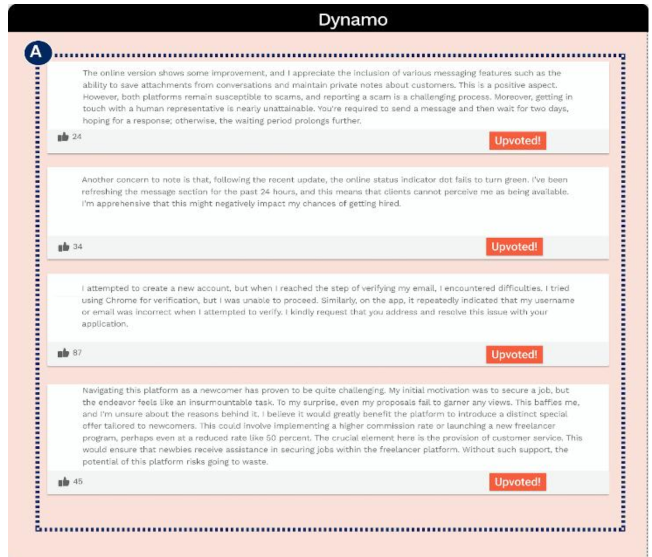

Figure 3: Overview of the control interface.

Measures and Data Analysis

Alongside the sociodemographic data, we collected a range of quantitative metrics to address our three research questions related to speed, contribution, and usability.

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.

Metric: Contribution. To determine if GigSense improves workers’ contributions to sensemaking and collective intelligence, we evaluated its impact on helping workers identify problems and develop various and feasible solutions. For feasibility evaluation, we gathered all problems and solutions from each condition and hired three English-speaking, college-educated raters (gig workers) through Upwork. They independently evaluated the feasibility of each solution proposed by our study participants, assessing how well each addressed its corresponding problem using a 7-point Likert scale. Feasibility was assessed based on the viability of the proposed solution in effectively addressing the problems related to gig work. Inter-rater reliability for each group was measured using a two-way mixed Intraclass Correlation Coefficient (ICC), revealing excellent agreement among raters: 0.93 for GigSense and 0.94 for the control group

Metric: Usability. To assess participants’ views on GigSense’s usability and compare it with the control, we employed the System Usability Score (SUS) [11], 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 [54 , 76]. SUS scores above 80.3 indicate excellent usability (Grade A), while 68-80.3 represent good usability (Grade B). A score of 68 is average (Grade C), suggesting a functional system with room for improvement. Scores between 51-68 indicate poor usability (Grade D) with significant issues, and below 51 is considered awful (Grade F), signaling major usability problems that severely impact user satisfaction and system effectiveness. Following participants’ interaction with their assigned system (control or GigSense), they received the SUS questionnaires. We then computed the SUS scores reported by participants for their respective systems. Following this, each participant provided a usability score for the system they used in their assigned condition.

5. Results

5.1 Time. Time can pose a challenge for gig workers aiming to engage in collective intelligence, as not all workers enjoy the luxury of allocating extensive time to this pursuit [71]. To address this concern, we assessed the duration participants required to accomplish the different problem-solving tasks defined in our study. Figure 4.a) provides a comprehensive depiction of the median time taken by participants to complete the entire set of tasks in both conditions. (Figure 4.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 when we plotted a histogram, 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 control condition. Overall, the data provides evidence that GigSense offers a promising approach (RQ1) to improve task completion times in problem-solving tasks related to collective intelligence.

5.2 Contribution. (Evaluating Gig Workers' Solutions). To evaluate GigSense's effectiveness in enhancing workers' contributions to sensemaking and collective intelligence, for both groups, we assessed the number of identified problems, the number of proposed solutions, and their feasibility. First, we observed that the GigSense group identified more problems compared to the control group. Specifically, the GigSense group had a median of 10 problems identified (mean 9.58) versus 6 (mean 6.33) for the control. Sinch our data didn't meet the assumption of normality when we plotted a histogram. We opted for a Mann-Whitney U test, which revealed this was a statistically significant difference (p = .01047). Similarly, in assessing the solutions proposed, the GigSense group suggested a higher number of solutions.

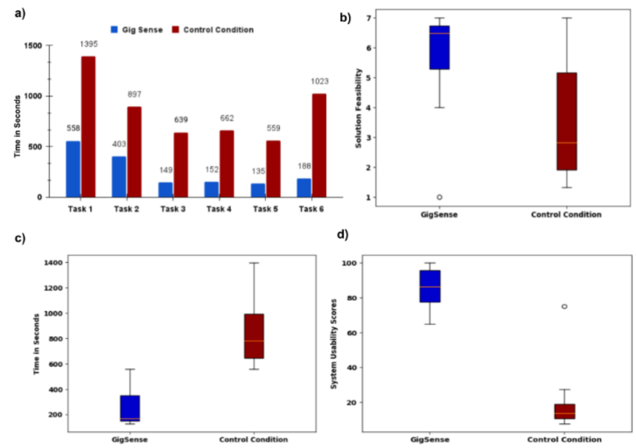

Figure 4: Overview of: a) Median amount of time gig workers in each condition took to complete the different tasks; b)Box plot showing the evaluation of solutions proposed by gig workers in both groups based on their feasibility (on a 7-point Likert scale); c) Box plot showing task completion time in both groups; d) Box plot showing System Usability Scores in both groups

In particular, they had proposed a median of 10 solutions (mean 9.41) compared to 6 (mean 6.33) for control - a statistically significant difference (p = .01379). 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 (Median=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 (5.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 difference 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. GigSense also enhances sensemaking capabilities, given the GigSense group identified more problems, and proposed more feasible and plentiful solutions, compared to the control group.

5.3 System Usability Scale. Utilizing the System Usability Scale (SUS) [4, 11, 76], we studied the reported usability levels of GigSense among gig workers and drew a comparison with those reported for the control condition. Figure 5.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. As the SUS scores did not meet the assumption of normality while plotting a histogram. 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 usability of two groups in our study (RQ3), with a p-value of 0.001.

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. These outcomes provide valuable insights into the role of LLMs in supporting sensemaking processes within collective intelligence. Here, we discuss ongoing challenges and prospects for sensemaking tools for problem-solving with gig workers.

Powering Collective Intelligence. In pursuit of enhanced problem-solving, GigSense strategically incorporated the advanced capabilities of Large Language Models (LLMs) to streamline and optimize the sensemaking process. For instance, participants 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. This correlates with earlier studies that have indicated LLMs' ability to generate notably superior ideas compared to humans [32]. GigSense, aided by LLMs, allowed workers to explore and generate solutions for less familiar problems, broadening the scope of solutions generated, analogous to the process of searching for relevant information in diverse data sources in the sensemaking phase ("Step: Search and Filter"). Moreover, participants appreciated how GigSense effortlessly enabled solution generation. This ease of generating solutions emphasizes the efficiency and user-friendly nature of GigSense's interface, which was further enhanced by the integration of LLMs. Additionally, LLMs contributed to the rapid extraction and presentation of solutions, aligning with the sensemaking step of extracting valuable information, analogous to the ("Step: Read and Extract") in the sensemaking process. In fact, participants expressed that they found it easier to derive solutions swiftly from the information provided within the GigSense platform, facilitated by LLM-driven capabilities. Moreover, Gigsense use of LLMs facilitated the summarization of problems, ("Step: Schematize") in the sense-making process, resulting in enhanced comprehension and consequently enabling the proposal of better solutions. For instance, participants using GigSense expressed their satisfaction with how the platform's interface facilitated their comprehension of gig workers' challenges about the topics presented on the platform. They appreciated the "Problem Summary Module", which spared them from the tedious process of opening and reading through the long list of individual worker complaints about specific problems. This improvement significantly aided workers in streamlining this task and in making information gathering and summarization much more efficient. In fact, participants expressed that they found it easier to derive solutions swiftly from the information provided within the GigSense platform, facilitated by LLM-driven capabilities. These user experiences resonate with prior studies demonstrating how LLMs can enhance users' information analysis capabilities [95, 96] In its current iteration, GigSense did not integrate Large Language Models (LLMs) directly within the specific steps of ("Step: Build Case") and ("Step: Tell Story") in the sensemaking process. Instead, it supports these steps via its Collaborative Solution Module, allowing users to interact and discuss, and upvote the different solutions generated by them and the AI-suggested solutions. Future work could explore the incorporation of LLMs into these pivotal sensemaking steps to further optimize the process. For instance, integrating LLMs within the ("Step: Build Case") could involve utilizing their language generation capabilities to assist in constructing a comprehensive case or argument based on gathered data. In the ("Step: Tell Story"), LLMs might aid in synthesizing and articulating narratives or insights drawn from the information collected, enhancing the storytelling aspect of sensemaking.

Catalyzing Inclusive Problem-Solving. GigSense is designed to facilitate gig workers' participation in problem- solving. In designing GigSense, we prioritized the unique time constraints faced by gig workers, acknowledging their often-limited availability due to potential financial hardships [34, 39, 109]. Our 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 the zoomed-in and zoomed-out dynamics of their problems. This likely led workers to have a better 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 [99, 100]. To address this, incorporating informative messages within GigSense could enlighten users about the benefits of investing slightly more time in analyzing and comprehending problems. However, it is also important to highlight that GigSense's design does aim to counter the risk of overlooking nuances in problem-solving by featuring both zoomed-in and zoomed-out interfaces, ensuring a comprehensive analysis. The zoomed-out view provides a comprehensive perspective, crucial for understanding how singular issues are interwoven into broader systemic patterns. This helps workers in recognizing overarching trends and contexts. On the other hand, the zoomed-in interface facilitates a detailed examination of specific problem aspects, allowing for a thorough analysis of individual components. This dual-mode approach effectively balances a macro and micro perspective, ensuring that complex issues are not oversimplified. Consequently, by enabling workers to effortlessly toggle between these views, GigSense enhances their ability to engage in more effective sensemaking and problem-solving. Our user study underscored this, showing that workers using Gig Sense identified significantly more and varied problems, and proposed more diverse and useful solutions, compared to the control group. GigSense's design also raises a thought-provoking discussion about its potential role in promoting "techno solutionism," the idea that technology can swiftly resolve complex design issues without deeply engaging with their intricacies [66]. This perception can stem from GigSense's design goals to expedite solution-finding, particularly under the constraints of workers who often lack the luxury of time for problem-solving due to their need to focus on livelihood-sustaining activities [34 , 39, 109]. To start to address this conflict, our GigSense design was inspired by collective intelligence research [68], highlighting the critical role of interfaces that support focused collaborative problem-solving. Consequently, GigSense's interface was crafted to enable workers to concentrate on specific issues, offering the tools to examine these problems from various deep perspectives. This strategy can hopefully enable workers to delve into their problems with greater depth and less superficiality. Based on these ideas, GigSense focuses on orienting workers towards solution-driven actions, actively engaging workers in the process of change, and reinforcing their sense of agency - a core tenet for promoting collective intelligence [68]. Contrastingly, interfaces with a problem-focused approach might induce in workers feelings of helplessness or passivity [1, 88], thus diminishing workers' agency [78, 89]. This is why we chose not to limit ourselves to a sole problem-focused interface. Nevertheless, GigSense does not disregard the importance of understanding problems in depth. Its interface, designed for both zooming in for detailed problem analysis and zooming out for a broader perspective, supports a more nuanced engagement with problems. This functionality proved effective in our studies, where workers using GigSense identified a wider array and greater number of problems compared to those using the control interface. To summarize, GigSense not only simplifies the journey towards finding solutions, but it also promotes a comprehension of the issues at hand.

Collaborative Problem Solving with Human-AI Interaction. Our system introduced a collaborative problem- solving process that integrated human-AI interactions, with a primary objective of enhancing human creativity by leveraging LLMs to empower workers to devise creative solutions to their challenges. This approach complements prior research on LLMs' assistance in enhancing human creativity [13, 38, 49, 91], emphasizing their supportive role rather than substituting human involvement[13]. GigSense demonstrates the value of designing interfaces that harness the power of LLMs to augment and streamline the sensemaking process to empower non-experts to utilize LLM technology for collective problem-solving. Our results reveal that LLMs supported gig workers (non-experts in 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 [36], our participants. welcomed LLM suggestions, incorporating it into their sensemaking process for creating solutions that improved their collective intelligence. 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 [41, 44]. This transparency can enhance collaboration between end-users and AI-generated solutions.

7. Limitations

Our study has several limitations, including a diverse but small sample size of gig workers with varying skills and geographic backgrounds, potentially affecting generalizability. We addressed this by requiring participants to have at least a year of gig work experience, ensuring they could meaningfully evaluate GigSense. Despite its smaller scale, our mixed-methods approach allowed for a rich, qualitative understanding of gig workers' interactions and experiences with the system. Future research could aim to validate our findings across broader samples and investigate the standalone impact of LLMs and interactive interfaces on solution generation for workers. We also integrated real-world gig worker data to broaden the context of challenges faced, acknowledging that not all problems can be solved even with tools like GigSense. Future studies could differentiate between solvable and complex challenges and assess the potential of LLMs and interactive interfaces for complex problem-solving. Although our study offers in-depth insights, a longitudinal study could provide additional data on social network analysis facilitated by GigSense. Upon publication, we will open-source GigSense, inviting further research on AI for collective intelligence. We recognize that LLMs can produce erroneous solutions, and we have aimed to implement safeguards by prioritizing human input and warn of potential errors. However, the introduction of LLMs might suggest a broader solution capability than feasible, indicating the need for future studies to refine user guidance on LLM limitations.

8. Conclusion

This paper presents GigSense, an AI-enhanced tool designed to help gig workers collectively understand and tackle their challenges. GigSense enables rapid sensemaking, reducing the barriers to collective intelligence, and fostering effective problem-solving and solution generation among workers. Our study revealed that users of GigSense identified more problems, and proposed solutions that were quicker, more plentiful, and more feasible than those in the control group. Users also reported enhanced usability experiences, valuing GigSense's support in problem understanding, collaborative solution crafting, and its integration of AI with a focus on human-driven solutions. The favorable feedback and functionality of GigSense highlight its capacity to transform gig workers' approach to challenges, turning individual struggles into collective successes.

9. References

[1] Lyn Y Abramson, Martin E Seligman, and John D Teasdale. 1978. Learned helplessness in humans: critique and reformulation. Journal of abnormal psychology 87, 1 (1978), 49.

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

[3] B Appleyard. 2004. The Paradox of Choice: why more is less by Barry Schwartz. NEW STATESMAN-LONDON- 810 (2004), 49–50.

[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] Stephen R Barley. 1986. Technology as an occasion for structuring: Evidence from observations of CT scanners and the social order of radiology departments. Administrative science quarterly (1986), 78–108.

[6] Benjamin B Bederson and James D Hollan. 1994. Pad++ a zooming graphical interface for exploring alternate interface physics. In Proceedings of the 7th annual ACM symposium on User interface software and technology. 17–26.

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

[8] 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.

[9] Janine Berg, Uma Rani, and others. 2021. Working conditions, geography and gender in global crowdwork. Work and Labour Relations in Global Platform Capitalism. Cheltenham: Edward Elgar (2021), 93–110.

[10] 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.

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

[12] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, and others. 2020. Language models are few-shot learners. Advances in neural information processing systems 33 (2020), 1877–1901. 18 GigSense:

[13] Alice Cai, Steven R Rick, Jennifer L Heyman, Yanxia Zhang, Alexandre Filipowicz, Matthew Hong, Matt Klenk, and Thomas Malone. 2023. DesignAID: Using Generative AI and Semantic Diversity for Design Inspiration. In Proceedings of The ACM Collective Intelligence Conference. 1–11.

[14] 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 (CHIWORK 2022). Association for Computing Machinery, New York, NY, USA, Article 14, 9 pages. DOI: http://dx.doi.org/10.1145/3533406.3533424

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

[16] 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.

[17] Philip R Cohen, Jerry Morgan, and Martha E Pollack. 2003. Collective Intentions and Actions. (2003).

[18] Giles Colborne. 2017. Simple and usable web, mobile, and interaction design. New Riders.

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

[20] 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.

[21] 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.

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

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

[24] 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.

[25] 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

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

[27] Rebecca S Etz, Martha M Gonzalez, Aimee R Eden, and Jodi Winship. 2018. Rapid sense making: A feasible, efficient approach for analyzing large data sets of open-ended comments. International Journal of Qualitative Methods 17, 1 (2018), 1609406918765509.

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

[29] 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).

[30] 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.

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

[32] 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).

[33] 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.

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

[35] 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.

[36] 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).

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

[38] 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.

[39] Jane Hsieh, Miranda Karger, Lucas Zagal, and Haiyi Zhu. 2023. Co-Designing Alternatives for the Future of Gig Worker Well-Being: Navigating Multi-Stakeholder Incentives and Preferences. In Proceedings of the 2023 ACM Designing Interactive Systems Conference. 664–687.

[40] Ziheng Huang and Stephen MacNeil. 2023. DesignNet: a knowledge graph representation of the conceptual design space. In Proceedings of the 15th Conference on Creativity and Cognition. 375–377.

[41] 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.

[42] James Garrett Jesse. 2011. The elements of user experience: User-centered design for the web and beyond. (2011).

[43] 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.

[44] 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).

[45] 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.

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

[47] 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. DOI:http://dx.doi.org/10.1145/3359310

[48] Steve Krug. 2000. Don't make me think!: a common sense approach to Web usability. Pearson Education India.

[49] 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.

[50] Chinmay Kulkarni, Tongshuang Wu, Kenneth Holstein, Q Vera Liao, Min Kyung Lee, Mina Lee, and Hariharan Subramonyam. 2023. LLMs and the Infrastructure of CSCW. In Companion Publication of the 2023 Conference on Computer Supported Cooperative Work and Social Computing. 408–410.

[51] 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 applications 35, 5 (2015), 84–90.

[52] 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.

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

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

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

[56] 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.

[57] 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.

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

[59] Michael Xieyang Liu, Tongshuang Wu, Tianying Chen, Franklin Mingzhe Li, Aniket Kittur, and Brad A Myers. 2023. Selenite: Scaffolding Decision Making with Comprehensive Overviews Elicited from Large Language Models. arXiv preprint arXiv:2310.02161 (2023).

[60] Stephen MacNeil, Zijian Ding, Kexin Quan, Ziheng Huang, Kenneth Chen, and Steven P Dow. 2021. ProbMap: Automatically constructing design galleries through feature extraction and semantic clustering. In Adjunct Proceedings of the 34th Annual ACM Symposium on User Interface Software and Technology. 134–136.

[61] John Maeda. 2006. The laws of simplicity. MIT press.

[62] 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. (2015).

[63] 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.

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

[65] 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.

[66] Evgeny Morozov. 2013. To save everything, click here: The folly of technological solutionism. PublicAffairs.

[67] Geoff Mulgan. 2018a. Big mind: How collective intelligence can change our world. Princeton University Press.

[68] Geoff Mulgan. 2018b. Big mind: How collective intelligence can change our world. Princeton University Press.

[69] 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.

[70] OpenAI. 2024. OpenAI Platform — platform.openai.com. https://platform.openai.com/docs/api-reference. (2024). [Accessed 13-01-2024].

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

[72] Srishti Palani, Aakanksha Naik, Doug Downey, Amy X Zhang, Jonathan Bragg, and Joseph Chee Chang. 2023. Relatedly: Scaffolding Literature Reviews with Existing Related Work Sections. In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems. 1–20.

[73] 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.

[74] 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.

[75] 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. 20

[76] 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.

[77] 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).

[78] Christopher Peterson and Martin EP Seligman. 1983. Learned helplessness and victimization. Journal of Social Issues 39, 2 (1983), 103–116.

[79] 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).

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

[81] 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.

[82] 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.

[83] Niloufar Salehi, Lilly C Irani, and Michael S Bernstein. 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. ACM, 1621–1630.

[84] Jörgen Sandberg and Haridimos Tsoukas. 2015. Making sense of the sensemaking perspective: Its constituents, limitations, and opportunities for further development. Journal of organizational behavior 36, S1 (2015), S6–S32.

[85] Saiph Savage and Mohammad H Jarrahi. 2020. Solidarity and AI for Transitioning to Crowd Work during COVID-19. In The New Future of Work Symposium 2020.

[86] Steve Sawyer, Kevin Crowston, and Rolf T Wigand. 2014. Digital assemblages: evidence and theorising from the computerisation of the US residential real estate industry. New Technology, Work and Employment 29, 1 (2014), 40–56.

[87] Cyrille Schwellnus, Assaf Geva, Mathilde Pak, and Rafael Veiel. 2019. Gig economy platforms: Boon or Bane? (2019). [88] Martin EP Seligman. 1972. Learned helplessness. Annual review of medicine 23, 1 (1972), 407–412.

[89] Martin EP Seligman. 2011. Building resilience. Harvard business review 89, 4 (2011), 100–106.

[90] 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.

[91] 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.

[92] Pao Siangliulue, Joel Chan, Steven P Dow, and Krzysztof Z Gajos. 2016. IdeaHound: improving large-scale collaborative ideation with crowd-powered real-time semantic modeling. In Proceedings of the 29th Annual Symposium on User Interface Software and Technology. 609–624.

[93] John B Smith. 1994. Collective intelligence in computer-based collaboration. CRC Press.

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

[95] Sangho Suh, Meng Chen, Bryan Min, Toby Jia-Jun Li, and Haijun Xia. 2023a. Structured Generation and Exploration of Design Space with Large Language Models for Human-AI Co-Creation. arXiv preprint arXiv:2310.12953 (2023).

[96] Sangho Suh, Bryan Min, Srishti Palani, and Haijun Xia. 2023b. Sensecape: Enabling Multilevel Exploration and Sensemaking with Large Language Models. arXiv preprint arXiv:2305.11483 (2023).

[97] 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.

[98] John Sweller. 1994. Cognitive load theory, learning difficulty, and instructional design. Learning and instruction 4, 4 (1994), 295–312.

[99] Henri Tajfel and John C Turner. 2004. The social identity theory of intergroup behavior. In Political psychology. Psychology Press, 276–293.

[100] Henri Tajfel, John C Turner, William G Austin, and Stephen Worchel. 1979. An integrative theory of intergroup conflict. Organizational identity: A reader 56, 65 (1979), 9780203505984–16.

[101] 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).

[102] Jenifer Tidwell. 2010. Designing interfaces: Patterns for effective interaction design. " O'Reilly Media, Inc.".

[103] Upwork. 2024. Upwork Community Forum. https://community.upwork.com/. (2024). Accessed: 2024-02-28.

[104] Martijn Van Zomeren, Tom Postmes, and Russell Spears. 2008. Toward an integrative social identity model of collective action: a quantitative research synthesis of three socio-psychological perspectives. Psychological bulletin 134, 4 (2008), 504.

[105] 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.

[106] 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.

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

[108] 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).

[109] Alex J Wood, Mark Graham, Vili Lehdonvirta, and Isis Hjorth. 2019. Good gig, bad gig: autonomy and algorithmic control in the global gig economy. Work, employment and society 33, 1 (2019), 56–75.

[110] 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.

[111] 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.

[112] 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. '

[113] 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.

[114] 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.