

Reimagining Global Crowdsourcing for Better Human-AI Collaboration

ANDY ALORWU, University of Oulu, Finland

SAIPH SAVAGE, Northeastern University, USA

NIELS VAN BERKEL, Aalborg University, Denmark

DMITRY USTALOV, Yandex, Russia

ALEXEY DRUTSA, Yandex, Russia

JONAS OPPENLAENDER, University of Oulu, Finland

OLIVER BATES, Lancaster University, UK

DANULA HETTIACHCHI, RMIT University, Australia

UJWAL GADIRAJU, TU Delft, Netherlands

JORGE GONCALVES, University of Melbourne, Australia

SIMO HOSIO, University of Oulu, Finland

Crowdworkers silently enable much of today's AI-based products, with several online platforms offering a myriad of data labelling and content moderation tasks through convenient labour marketplaces. The HCI community has been increasingly interested in investigating the worker-centric issues inherent in the current model and seeking for potential improvements that could be implemented in the future. This workshop explores how a reimagined perspective on crowdsourcing platforms could provide a more equitable, fair, and rewarding experience. This includes not only the workers but also the platforms, who could benefit e.g. from better processes for worker onboarding, skills-development, and growth. We invite visionary takes in various formats on this topic to spread awareness of worker-centric research and developments to the CHI community. As a result of interactive ideation work in the workshop, we articulate a future direction roadmap for research centred around crowdsourcing platforms. Finally, as a specific interest area, the workshop seeks to study crowdwork from the context of the Global South, which has been arising as an important but critically understudied crowdsourcing market in recent years.

ACM Reference Format:

Andy Alorwu, Saiph Savage, Niels van Berkel, Dmitry Ustalov, Alexey Drutsa, Jonas Oppenlaender, Oliver Bates, Danula Hettiachchi, Ujwal Gadiraju, Jorge Goncalves, and Simo Hosio. 2022. Reimagining Global Crowdsourcing for Better Human-AI Collaboration. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts (CHI '22 Extended Abstracts)*, April 29-May 5, 2022, New Orleans, LA, USA. ACM, New York, NY, USA, 10 pages. <https://doi.org/10.1145/3491101.3503725>

1 BACKGROUND

Crowdsourcing refers to the division of work into smaller parallel tasks distributed to a workforce, with results being aggregated in some meaningful way. This often takes place via online platforms that provide the functionality needed to interact between the requesters of work and the available online workforce. Paid online crowdsourcing has rapidly become the backbone of industry applications in data enrichment and machine learning, and platforms such as

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

© 2022 Copyright held by the owner/author(s).

Manuscript submitted to ACM

53 Amazon's MTurk¹, Appen², or Toloka³ are now broadly used to fuel the numerous AI-based innovations that depend
54 on high-quality training data. Collecting and enriching training data typically includes tasks such as image labelling,
55 sentiment analysis, and different types of classification tasks.

56 The practice of crowdsourcing is not without issues. Workers on crowdsourcing platforms are often left without
57 the attention needed in order to build a fair, sustainable future model for crowdsourcing. There are real people behind
58 the unique participant identification numbers assigned by the platforms, with real-life needs and worries. To this
59 end, HCI scholars have investigated aspects of worker mental health [24], invisible labour [29, 47], data management
60 and reuse [1, 2], dystopian future forms of crowdwork [4], and worker attitudes toward e.g. creative work and skills-
61 development [42]. Further, HCI researchers have recently become interested in leveraging the potential of these
62 platforms for more nuanced and complex use cases, which might be accompanied by a more rewarding work experience.
63 Here, examples include collaboration and work that requires multiple iterations [5, 46], creative work [42], or digital
64 health and well-being solutions [6, 28].

65 Our aim with this workshop is to collectively explore what the future of these platforms could look like 10–20 years
66 from now. With this goal, we will not only gather a community of researchers and industry experts around notions
67 of work, but also build knowledge together. Some overarching questions that we will seek to address by drawing on
68 the vast experiences and perspectives present at the workshop include: What are the platform affordances that could
69 best support worker well-being and professional development, fundamental work rights, legal issues? An example
70 starting point here could be e.g. the Fairwork⁴ principles on Gig Work and/or Cloudwork or through the lens of
71 stopping Silicon Valley building a new global underclass of workers through crowdwork [26]. Would it be possible to
72 envision data management models to help workers retain some degree of ownership of their personal data while still
73 also benefiting the requesters? We encourage thinking about these questions in the context of identifying how the
74 future of crowdsourcing platforms could look like especially in the Global South, particularly Latin America, where
75 crowdsourcing platforms have started to become important key spaces to find labour⁵ [31, 50]. We acknowledge,
76 however, that many issues to be discussed might be generic and not necessarily specific to this context, and submissions
77 discussing more generic viewpoints are welcome as well.

78 To this end, related workshops across HCI and relevant venues have been organized, including many by the organizers
79 of this one. Workshops in CHI, CSCW, and UbiComp have examined crowdsourcing from various different angles. For
80 example, a CSCW '17 workshop explored the role of crowdsourcing in the context of law and policy [38]. At UbiComp
81 '18, researchers explored smartphones and other mobile devices as a tool to support crowdworkers and increase data
82 quality [54]. A CHI '19 workshop highlighted opportunities to support creativity in crowdsourcing platforms [41].
83 NeurIPS '20 and VLDB '21 workshops on trust, ethics and data excellence in crowdsourcing [52, 53]. Finally, at CSCW
84 '21, a workshop was organised around the notion of the hidden labour that powers modern day's AI systems [7].

85 Building on such past workshops, we are now specifically interested in investigating the affordances of crowdsourcing
86 platforms concerning the following thus far underrepresented or largely ignored topics, including, but not limited to:

- 87 • Worker onboarding into the platform ecosystems: How to facilitate a smooth start in a distributed, largely
88 anonymous working environment?
- 89 • Offering meaningful work or enriching tasks in meaningful ways.

90 ¹<https://www.mturk.com/>

91 ²<https://appen.com/>

92 ³<https://toloka.ai/>

93 ⁴<https://fair.work/en/fw/homepage/>

94 ⁵<https://restofworld.org/2021/refugees-machine-learning-big-tech/>

- 105 • Skills-development and personal growth opportunities of workers as a byproduct of work.
- 106 • Psychological support and feedback for workers.
- 107 • Trust between workers and requesters and how to engage the workers in quality assurance, while keeping the
- 108 requesters in loop with them.
- 109 • Tooling-related support measures for workers, as offered by the platforms themselves (as opposed to being
- 110 orchestrated externally).
- 111 • Novel data management solutions that can, for example, lead to novel incentive mechanisms through retained
- 112 data ownership.
- 113 • New use cases for platform workers far beyond what is feasible currently, leveraging *e.g.* XR technologies,
- 114 metaverse concepts, or any other futuristic technologies.
- 115 • Ensuring the fundamental work rights of crowd workers, for example, the legality of their contracts, their access
- 116 to workplace benefits (e.g. sick leave, parental leave, pensions).
- 117
- 118
- 119

120 To foster viewpoint diversity we seek to accept visions and considerations from any other aspect relevant to the future
 121 potential of paid crowdsourcing, and from the perspective of any relevant stakeholders.

122 **Impact statement:** We are sensitive to the fact that often CHI workshops happen ‘by researchers, for researchers’,
 123 and it may be unrealistic to expect results that contribute concrete functionalities or instantly implementable policies.
 124 With this in mind, our intended impact is two-fold: 1) Spread awareness of worker-centric research and developments
 125 on the current platforms and 2) articulate a concrete research roadmap for crowdsourcing that integrates the identified
 126 global challenges. Doing so, we contribute our small part in redirecting the community’s attention toward the workers’
 127 benefit – to a degree dictated ultimately by the workshop attendance and post-activities.

128 We will offer a low barrier of entry concerning submission types, accepting *e.g.* small-scale empirical studies, written
 129 essays, concept videos, and pictorials. We encourage creativity and futurism in form of visionary work. All types of
 130 visions are welcome: utilitarian, dystopian, utopian, or logical and practical. The organizers will use the submissions
 131 to curate an initial set of possible future affordances for the platforms to be briefly presented during the workshop
 132 and used during the workshop as a discussion stimulant. As a concrete output of the collaborative workshop, we will
 133 contribute a curated set of reflections, open research questions, and a concrete roadmap to ‘*what can the future of paid*
 134 *crowdsourcing platforms look like 10–20 years from now?*’

135 **Integrating Critical Theory to Design the Future of Global Crowdsourcing.** For CHI it can be important to
 136 question whether new proposed designs for the future of crowdsourcing are truly engaging and addressing critical
 137 societal problems that current crowdsourcing platforms are generating, or if the new designs are simply starting to
 138 touch on the mere symptoms of a problem [3]. The latter case can result when we create designs that make a societal
 139 problem more bearable. Without addressing the root of the problem, it is, however, possible that our proposed designs
 140 could actually reinforce critical structural issues, causing more abuse and harm to workers.

141 To address this problem, our workshop will draw from theorist Hebert Marcuse from the Frankfurt School of Critical
 142 Theory [30], who has argued that it is extremely challenging to engage in critical analysis of the structures and processes
 143 that exist in society [3, 37]. But a way to start practising more critical analysis is via artistic creativity [36]. According
 144 to Marcuse, artistic creativity facilitates creating designs that can truly challenge the current reality of what is possible
 145 and allows us to consider directions we might have been blind to consider otherwise. Based on these ideas, we plan
 146 to engage participants of our workshop in ‘creative artistic co-design sessions’ where they use fictional narratives to
 147 design ‘alternative realities’ to contemporary digital labour platforms and tools [20, 35].

157 **Designing Crowdsourcing Tools and Platforms for Wider Audiences.** Our goal is to focus one part of the
158 design of solutions for the future of crowdsourcing within the context of Latin America as a prominent example. We
159 focus on this population and culture given that studies have shown that a majority of the A.I. industry is relying
160 on a global workforce located in the Global South, particularly Latin America, to complete their data labelling tasks
161 [8, 27, 31, 32, 44, 45, 55]. For example more than 75% of the labelled training data that Tesla used for its self-driving car
162 came from Venezuelan crowd workers [9, 50]. Additionally, we take advantage of the fact that some of the organisers of
163 the workshop are leaders designing for Latin America [10, 21, 22, 39, 48].

166 Note that within our workshop, we do acknowledge that it can be difficult for participants to be able to analyse
167 how certain crowdsourcing designs might be affecting Latin America and design adequate solutions, especially if
168 these individuals are not from the region or familiar with the related culture [12, 49]. For this purpose, we will include
169 workshop activities where we help participants to adopt the “*circuit of culture framework*” [18]. This framework provides
170 culture probes that can help outsiders to Latin America to understand how a particular technology is currently being
171 interpreted and understood by people of that region. Such cultural understanding facilitates creating better interventions
172 and positive design changes for Latin America [33]. We will also include hybrid activities in our workshop that
173 include outsiders and insiders to Latin America, e.g., co-design sessions with locals from Latin America and workshop
174 participants from other parts of the world. Through these co-design sessions and by adopting the circuit of culture
175 framework, we will be able to empower all participants to better understand the type of representation and identity that
176 Latin Americans attribute to crowdsourcing platforms, as well as the type of production, consumption, and regulation
177 that locals relate to crowd work [25, 51]. Through this, we will all start to collectively understand how Latin America
178 is encoding meaning and lifestyles into crowdsourcing platforms. We will then use the understanding to collectively
179 design better artefacts that power a better future of work for the region.

183 **Designing the Future of Crowdsourcing with Workers, Requesters, and Platform Owners** For crowdsourc-
184 ing platforms to truly bring positive futures, we need the platforms to enable success for the different stakeholders: those
185 providing the labour, those requesting it, and the platforms who benefit from productive commerce. In our workshop
186 we will focus on having conversations on how can we imagine, design, and structure the data supply ecosystem to
187 benefit all three parties. We adopt value-sensitive design as a principled framework to integrate the diverse needs of
188 multiple stakeholders in the overall data labelling ecosystem [23]. To put this into practice in our workshop, we will
189 have guided activities using the value sensitive design framework to help participants identify the different needs of
190 the different stakeholders, as well as points where there are conflicts between needs. Through this, we will empower
191 participants to start to design solutions that aim to empower all stakeholders [14]. We will aim to include also all
192 stakeholders as participants in our workshop. Currently the workshop organisers already encompass some of these
193 different stakeholders. Notice that this approach also differentiates our workshop from previous efforts that focused
194 primarily on workers’ needs instead of the whole ecosystem.

200 2 ORGANISERS

201 The organisers of the workshop come from a variety of backgrounds and all have expertise and have been involved in
202 the field of crowdsourcing, data management, and human computation over the years. With representation also from the
203 microwork industry, the team is in a good position to disseminate and put forward the workshop’s contributions. Many
204 of the organizers have also been involved in the organization of a series of workshops and tutorials on crowdsourcing
205 and human computation over the years [13, 15–17, 41, 43, 52–54, 56].

209 **Andy Alorwu** is a Doctoral Researcher in the Crowd Computing research group at the Center for Ubiquitous Com-
210 puting, University of Oulu, Finland. His research interests include personal data management, m-health, and mobile
211 computing.
212

213 **Saiph Savage** is an Assistant Professor at the College of Computer Science at Northeastern University in Boston. Dr.
214 Savage is also a Research Collaborator at the Civic Innovation Lab at the Universidad Nacional Autonoma de Mexico
215 (UNAM) in Mexico City. Dr. Savage has vast experience designing tools to empower crowd workers. She also has
216 experience designing technology for the global south, especially Latin America.
217

218 **Niels van Berkel** is an Associate Professor at the Human-Centered Computing Group at Aalborg University. His
219 research interests lay in Human-Computer Interaction, Social Computing, and Ubiquitous Computing.
220

221 **Dmitry Ustalov** is the Head of Research at the Toloka crowdsourcing platform. His research interests are focused on
222 Crowdsourcing, Natural Language Processing, and Evaluation.
223

224 **Alexey Drutsa** is the Head of Efficiency and Growth Division at the Toloka crowdsourcing platform. His research
225 interests are focused on Machine Learning, Data Analysis, Auction Theory.
226

227 **Jonas Oppenlaender** is an alumni of the Crowd Computing research group at the Center for Ubiquitous Computing,
228 University of Oulu, Finland. His research interests include crowdsourcing, crowd feedback systems, and crowd-powered
229 creativity support systems.
230

231 **Oliver Bates** is a Research Fellow at the School of Computing and Communications, Lancaster University, UK. He is
232 broadly interested in labour rights especially in the gig economy.
233

234 **Danula Hettiachchi** Research Fellow at the ARC Centre of Excellence on Automated Decision Making and Society
235 and RMIT School of Computing, Australia. His research interests include crowdsourcing, social computing and human-
236 computer interaction.
237

238 **Ujwal Gadiraju** is an Assistant Professor at the Web Information Systems (WIS) group of the Software Technology
239 department at TU Delft in the Netherlands. He is a Director of the Delft Design@Scale AI Lab and leads a research line
240 on Human-Centered AI and Crowd Computing at the WIS group.
241

242 **Jorge Goncalves** is a Senior Lecturer in the School of Computing and Information Systems at the University of
243 Melbourne. He has conducted extensive research on improving crowd data quality, and bringing crowdsourcing beyond
244 the desktop by using ubiquitous technologies. He has also served as Workshops Co-Chair for CHI'19 and CHI'20.
245

246 **Simo Hosio** is an Associate Professor at the Center for Ubiquitous Computing, University of Oulu, Finland. He leads
247 the Crowd Computing Research Group and is affiliated with the Center for Life Course Health Research. His research
248 interests include crowdsourced wellbeing solutions, social computing and ubiquitous computing.
249

250

251

252

253

254

255

256

257

258

259

260

3 LINK TO WEBSITE

The workshop's website will be hosted on <https://crowdscience.ai/>, as part of the Crowd Science Initiative, and will be finalised upon acceptance of this proposal. The website will be used to spread the word about the workshop and act

261 as a central repository for all important information to attendees. This includes the submission date, workshop date,
 262 submission modality, and links to related material so that candidates can get familiar with the scope of the subject and
 263 the goals of the workshop. A detailed schedule of the workshop and requirements for participants will be provided, as
 264 well as, the background of each organiser. Further, the website will host the primary output of the workshop; the future
 265 vision statement and the related Miro board with condensed versions of the teams' work, and accepted position papers
 266 will be available for download on the dedicated website.
 267
 268

269 4 PRE-WORKSHOP PLANS

270
 271 We will first finalise the workshop website, to act as a simple information hub about the workshop and, later on, as a
 272 hub that links to all workshop submissions which are stored online.
 273

274 We require submissions from all workshop participants in the format of short visions that may deal with one or
 275 more of the topics of the workshop, as listed earlier. Prior to the workshop, we mostly focus on spreading aware-
 276 ness of the workshop and distributing the CFP across all relevant academic as well as industry mailing lists (e.g.,
 277 chi-announcements). We will further use our own distribution lists (based on a variety of workshops and conferences
 278 we have held in the past). Workshop organisers will also use social media channels at their disposal to publicise the CFP.
 279

280 Each submission will be checked for topicality and reviewed by at least two of the organisers. While we are not
 281 planning to reject any submissions based on empirical contribution depth, we reserve the right to reject submissions
 282 with no connection to any of the workshop themes.
 283

284 The submissions will be curated into a publicly available Notion dashboard (<https://www.notion.so/>) that will be made
 285 available to the workshop participants a week before the workshop date. This way, participants can already familiarise
 286 themselves with other participants as well as their interests and positions concerning the future of crowdsourcing
 287 platforms.
 288

289 5 WORKSHOP STRUCTURE

291 First Session	
292 30 min	Workshop introduction and keynote
293 60 min	Break out in teams, start co-designing new futures (value-sensitive methods)
294 30 min	Coffee break, informal panel discussion
295 60 min	Teamwork continues, focusing on further conceptualising new platform affordances
296 60 min	Lunch break
297 Second Session	
298 60 min	Team presentations and discussion
299 15 min	Teams create a synthesis of their work in a shared workspace online
300 30 min	Panel discussion focusing on the created synthesis
301 15 min	Final thoughts, closing the workshop

302
 303
 304
 305
 306 The workshop emphasises discussion, teamwork, and collaboration in teams of 2–4 members. We aim for 15–20
 307 participants (excluding the organisers). As we have a fairly diverse and numerous organiser board, we hope to attract a
 308 lot of expertise around the subject matter and from various different backgrounds to facilitate different viewpoints.
 309 There will be no traditional paper presentations; instead we will focus on ideation and collaborative activities that are
 310 supported by the potentially informal and hopefully provocative submissions by participants.
 311
 312

313 We start the workshop with a brief keynote presentation by an industry representative, to be recruited upon workshop
314 approval and dates confirmation, followed by rapidly moving into break out groups. All participants have access to the
315 workshop submissions through a shared Notion workspace online. In these breakout discussions and informal groupings
316 in person participants will receive fiction prompts to start co-designing with workers, academics, and industry actors
317 new crowdsourcing interfaces. In specific, we use the context and concepts of design fictions to enable workshop
318 participants to conceptualise, explore, and critique new design ideas for crowdsourcing platforms [34, 35, 57]. Here we
319 will build off our research studying and organising hybrid events [19],⁶ especially with minorities, to now leverage the
320 crowd in different ways (the people in-person could build and hack things, while the audience online could search for
321 information, such as needs and facts of the region to help motivate the design. We will also utilise the hybrid aspect to
322 facilitate the participation from workers and researchers from Latin America, who might not be able to travel to the
323 conference due to costs.

326 Following a 30-minute coffee break during which we will organize an informal panel discussion, the participants
327 continue the morning's tasks in groups, focusing specifically on novel platform affordances. Doing so, we hope the
328 different groups can bring along their unique expertise and viewpoints to each of the themes that were derived from
329 the initial ideas. Each group will work in a shared Miro board.

331 The session after lunch starts with team presentations and discussion. We generously reserve time for this activity,
332 as experience has shown that HCI workshops typically run out of time when participants get to present and discuss
333 newly-created work and ideas. This will also help participants refine and rethink some of their ideas from the earlier
334 stage. In the final stage, we ask for the groups to condense the key objectives, research questions and other ideas to the
335 main Miro board of the workshop. This way, the main board acts as a live memo for the organisers who will use it to
336 create the overarching vision statement on the workshop website. Finally, we will have one panel discussion around
337 the key shared concepts emerging during the day and close the workshop.

341 6 POST-WORKSHOP PLANS

342 One of the primary goals of the workshop is to create a vision with a set of research questions and engineering challenges
343 – a roadmap for future. These will be summarised on the workshop website, as a central repository of research ideas
344 to explore. The workshop results will be communicated to a larger audience through the extensive network of the
345 organizers. We will also ask for volunteers among the organisers, as well as from the participating platform workers
346 and industry participants, to act as a task force to publish the workshop results in an academic format in a suitable
347 venue (*e.g.*, HCOMP or related workshops).

351 7 REMOTE/ONSITE/ASYNCHRONOUS PLANS

352 We will organise this workshop as a hybrid event, with most participants encouraged to participate physically in the
353 event as part of the CHI conference. Selected virtual attendees will be requested to fill out an online form regarding
354 their timezones, platform restrictions, and accessibility requirements, and interests in topics related to crowdsourcing.
355 This information will be used to schedule the workshop, arranging an online communication format, and creation
356 of discussion groups. Most importantly, all practical work takes place on a shared central Miro board, where we can
357 facilitate also asynchronous work. Short presentations can be posted there as video messages using a free tool such as
358 Loom, but for the most part the ideation work is shaped and edited in written form, uploads of sketches, and diagrams

362 ⁶<https://www.milenio.com/cultura/hackaton-une-mujeres-crear-casas-inteligentes>, [https://www.criptonoticias.com/educacion/unam-anfitrion-](https://www.criptonoticias.com/educacion/unam-anfitrion-hackaton-blockchain-mexico/)
363 [hackaton-blockchain-mexico/](https://devday4w.com/blog/mujeres-conectadas/), <https://devday4w.com/blog/mujeres-conectadas/>

365 directly in Miro. Based on current research in hybrid events that has shown that shorter, more focused sessions keep
 366 attendees' attention better [11, 40]. We will organise different activities within the workshop to keep delegates moving
 367 between activities and create a flow for the day. As mentioned in the workshop structure, some of the activities will
 368 include: panels with global south crowd workers, academics, industry actors who design the crowdsourcing platforms.
 369 We will also have collective discussions on the type of labour conditions and power dynamics present in current
 370 crowdsourcing platforms. The audience of the panel will participate via fireside chats.

371
 372 All materials will be distributed as fully digital using Miro and Notion, which will further help remote participation.
 373 The presentations will be streamed online for remote participants to view by using a tabletop microphone and a web
 374 camera.
 375

376 377 **8 250-WORD CALL FOR PARTICIPATION**

378
 379 This hybrid workshop reimagines crowdsourcing to better answer to the modern global worker-centric challenges and
 380 specific issues in the Global South.

381 Following a juried process, we invite anyone interested in the future of crowdsourcing to submit vision statements
 382 with an overarching question of what could these platforms look like in 10-20 years from now. To this end, we focus on
 383 the following topics (for the full list of topics, please see the workshop website):
 384

- 385 • Platform affordances: onboarding issues, skills development, personal growth opportunities
- 386 • Psychological support and feedback for workers
- 387 • Trust between workers and requesters: Fair quality assurance and good communication
- 388 • Tooling-related support measures for workers
- 389 • Novel data management solutions that can, for example, lead to new incentive mechanisms
- 390 • New use cases for platforms beyond what is feasible currently, leveraging *e.g.* XR technologies, metaverse
 391 concepts, or other futuristic technologies
- 392 • Ensuring the fundamental work rights of crowd workers, for example, the legality of their contracts, access to
 393 workplace benefits (*e.g.* sick leave, parental leave, pensions)

394
 395 We solicit submissions through the workshop website in the ACM Primary Article Template, with links to *e.g.* online
 396 videos or other necessary resources. We do not impose page limits, but we encourage the authors to keep submis-
 397 sions as concise as possible. At least one author of each accepted submission must attend the workshop and that all
 398 participants must register for both the workshop and for at least one day of the conference. For details please visit
 399 https://crowdsience.ai/conference_events/chi22.
 400
 401
 402

403 REFERENCES

- 404
 405 [1] Andy Alorwu, Saba Kheirinejad, Niels van Berkel, Marianne Kinnula, Denzil Ferreira, Aku Visuri, and Simo Hosio. 2021. Assessing MyData
 406 Scenarios: Ethics, Concerns, and the Promise. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–11.
 407 [2] Andy Alorwu, Aku Visuri, Niels van Berkel, and Simo Johannes Hosio. 2021. (Re) using Crowdsourced Health Data: Perceptions of Data Contributors.
 408 *IEEE Software* (2021).
 409 [3] Jeffrey Bardzell, Shaowen Bardzell, and Mark A Blythe. 2018. *Critical theory and interaction design*. MIT Press.
 410 [4] Oliver Bates, Christian Remy, Callum Nash, and Ben Kirman. 2019. The future of techno-disruption in gig economy workforces: challenging the
 411 dialogue with fictional abstracts. In *Proceedings of the Halfway to the Future Symposium 2019*. 1–4.
 412 [5] Michael S. Bernstein, Greg Little, Robert C. Miller, Björn Hartmann, Mark S. Ackerman, David R. Karger, David Crowell, and Katrina Panovich. 2010.
 413 Soylent: A Word Processor with a Crowd Inside. In *Proceedings of the 23rd Annual ACM Symposium on User Interface Software and Technology* (New
 414 York, New York, USA) (*UIST '10*). Association for Computing Machinery, New York, NY, USA, 313–322. <https://doi.org/10.1145/1866029.1866078>
 415 [6] Jeffrey P. Bigham, Chandrika Jayant, Hanjie Ji, Greg Little, Andrew Miller, Robert C. Miller, Robin Miller, Aubrey Tatarowicz, Brandyn White,
 416 Samuel White, and Tom Yeh. 2010. VizWiz: Nearly Real-Time Answers to Visual Questions. In *Proceedings of the 23rd Annual ACM Symposium on*

- 417 *User Interface Software and Technology* (New York, New York, USA) (*UIST '10*). Association for Computing Machinery, New York, NY, USA, 333–342.
 418 <https://doi.org/10.1145/1866029.1866080>
- 419 [7] Benedetta Catanzariti, Srrayva Chandhiramowuli, Suha Mohamed, Sarayu Natarajan, Shantanu Prabhat, Noopur Raval, Alex Taylor, and Ding
 420 Wang. 2021. The Global Labours of AI and Data Intensive Systems. In *Companion of the 2021 ACM Conference on Computer Supported Cooperative
 421 Work and Social Computing (CSCW '21 Companion)*. Association for Computing Machinery, New York, NY, USA, to appear.
- 422 [8] Alan Chan, Chinasa T Okolo, Zachary Turner, and Angelina Wang. 2021. The Limits of Global Inclusion in AI Development. *arXiv preprint
 423 arXiv:2102.01265* (2021).
- 424 [9] Angela Chen. 2021. *Desperate Venezuelans are making money by training AI for self-driving cars*. Mit Tech Review.
- 425 [10] Chun-Wei Chiang, Eber Betanzos, and Saiph Savage. 2018. Exploring blockchain for trustful collaborations between immigrants and governments.
 426 In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–6.
- 427 [11] Ben Chodor. 2020. *Transitioning to Virtual and Hybrid Events: How to Create, Adapt, and Market an Engaging Online Experience*. John Wiley & Sons.
- 428 [12] Andy Crabtree, Terry Hemmings, Tom Rodden, Keith Cheverst, Karen Clarke, Guy Dewsbury, John Hughes, and Mark Rouncefield. 2003. Designing
 429 with care: Adapting cultural probes to inform design in sensitive settings. In *Proceedings of the 2004 Australasian Conference on Computer-Human
 430 Interaction (OZCHI2004)*. 4–13.
- 431 [13] Hettiachchi Danula, Sanderson Mark, Goncalves Jorge, Hosio Simo, Kazai Gabriella, Lease Matthew, Schaekermann Mike, and Emine Yilmaz. 2021.
 432 Investigating and Mitigating Biases in Crowdsourced Data. In *2021 Conference on Computer Supported Cooperative Work and Social Computing
 433 (CSCW '21 Companion)*. Association for Computing Machinery, New York, NY, USA, 4 pages. <https://doi.org/10.1145/3462204.3481729>
- 434 [14] Xuefei Deng, Kshiti D Joshi, and Robert D Galliers. 2016. The duality of empowerment and marginalization in microtask crowdsourcing: Giving
 435 voice to the less powerful through value sensitive design. *Mis Quarterly* 40, 2 (2016), 279–302.
- 436 [15] Alexey Drutsa, Valentina Fedorova, Dmitry Ustalov, Olga Megorskaya, Evfrosiniya Zerminova, and Daria Baidakova. 2020. Crowdsourcing Practice
 437 for Efficient Data Labeling: Aggregation, Incremental Relabeling, and Pricing. In *Proceedings of the 2020 ACM SIGMOD International Conference on
 438 Management of Data (SIGMOD '20)*. Association for Computing Machinery, Portland, OR, USA, 2623–2627. <https://doi.org/10.1145/3318464.3383127>
- 439 [16] Alexey Drutsa, Valentina Fedorova, Dmitry Ustalov, Olga Megorskaya, Evfrosiniya Zerminova, and Daria Baidakova. 2020. Practice of Efficient Data
 440 Collection via Crowdsourcing: Aggregation, Incremental Relabelling, and Pricing. In *Proceedings of the 13th International Conference on Web Search
 441 and Data Mining (WSDM '20)*. Association for Computing Machinery, Houston, TX, USA, 873–876. <https://doi.org/10.1145/3336191.3371875>
- 442 [17] Alexey Drutsa, Dmitry Ustalov, Valentina Fedorova, Olga Megorskaya, and Daria Baidakova. 2021. Crowdsourcing Natural Language Data at Scale:
 443 A Hands-On Tutorial. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human
 444 Language Technologies: Tutorials (NAACL-HLT 2021)*. Association for Computational Linguistics, Online, 25–30. <https://doi.org/10.18653/v1/2021.naacl-tutorials.6>
- 445 [18] Paul Du Gay, Stuart Hall, Linda Janes, Anders Koed Madsen, Hugh Mackay, and Keith Negus. 2013. *Doing cultural studies: The story of the Sony
 446 Walkman*. Sage.
- 447 [19] Rosta Farzan, Saiph Savage, and Claudia Flores Saviaga. 2016. Bring on board new enthusiasts! A case study of impact of Wikipedia art+ feminism
 448 edit-a-thon events on newcomers. In *International Conference on Social Informatics*. Springer, 24–40.
- 449 [20] Casey Fiesler. 2018. Owning the Servers: A Design Fiction Exploring the Transformation of Fandom into ‘Our Own’. *Transformative Works and
 450 Cultures* 28 (2018).
- 451 [21] Claudia Flores-Saviaga and Saiph Savage. 2021. Fighting disaster misinformation in Latin America: the# 19S Mexican earthquake case study. *Personal
 452 and Ubiquitous Computing* 25, 2 (2021), 353–373.
- 453 [22] Michelle Forelle, Philip N Howard, Andrés Monroy-Hernández, and Saiph Savage. 2015. Political bots and the manipulation of public opinion in
 454 Venezuela. *Available at SSRN 2635800* (2015).
- 455 [23] Batya Friedman. 1996. Value-sensitive design. *interactions* 3, 6 (1996), 16–23.
- 456 [24] Ujwal Gadiraju and Gianluca Demartini. 2019. Understanding Worker Moods and Reactions to Rejection in Crowdsourcing. In *Proceedings of the
 457 30th ACM Conference on Hypertext and Social Media (Hof, Germany) (HT '19)*. Association for Computing Machinery, New York, NY, USA, 211–220.
 458 <https://doi.org/10.1145/3342220.3343644>
- 459 [25] Frank W Geels. 2004. From sectoral systems of innovation to socio-technical systems: Insights about dynamics and change from sociology and
 460 institutional theory. *Research policy* 33, 6-7 (2004), 897–920.
- 461 [26] Mary L Gray and Siddharth Suri. 2019. *Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass*. Eamon Dolan Books.
- 462 [27] Rafael Grohmann and Willian F ARAÚJO. 2021. Beyond Mechanical Turk: the work of Brazilians on global AI platforms. *VERDEGEM, Pieter. AI for
 463 everyone* (2021), 247–266.
- 464 [28] Simo Johannes Hosio, Niels van Berkel, Jonas Oppenlaender, and Jorge Goncalves. 2020. Crowdsourcing personalized weight loss diets. *Computer
 465 53, 1* (2020), 63–71. <https://doi.org/10.1109/MC.2019.2902542>
- 466 [29] Lilly C. Irani and M. Six Silberman. 2013. Turkopticon: Interrupting Worker Invisibility in Amazon Mechanical Turk. In *Proceedings of the SIGCHI
 467 Conference on Human Factors in Computing Systems (Paris, France) (CHI '13)*. Association for Computing Machinery, New York, NY, USA, 611–620.
 468 <https://doi.org/10.1145/2470654.2470742>
- [30] Martin Jay. 1996. *The dialectical imagination: A history of the Frankfurt School and the Institute of Social Research, 1923-1950*. Vol. 10. Univ of California Press.
- [31] Phil Jones. 2021. *Work Without the Worker: Labour in the Age of Platform Capitalism*. Verso Books.

- 469 [32] Zachary Kilhoffer. 2020. Platform Work and Participation: Disentangling the Rhetoric. In *Using New Media for Citizen Engagement and Participation*.
470 IGI Global, 1–15.
- 471 [33] Stephen Kline, Nick Dyer-Witthford, Greig De Peuter, et al. 2003. *Digital play: The interaction of technology, culture, and marketing*. McGill-Queen's
472 Press-MQUP.
- 473 [34] Joseph Lindley and Paul Coulton. 2015. Back to the future: 10 years of design fiction. In *Proceedings of the 2015 British HCI Conference*. 210–211.
- 474 [35] Conor Linehan, Ben J Kirman, Stuart Reeves, Mark A Blythe, Theresa Jean Tanenbaum, Audrey Desjardins, and Ron Wakkary. 2014. Alternate
475 endings: using fiction to explore design futures. In *CHI'14 Extended Abstracts on Human Factors in Computing Systems*. 45–48.
- 476 [36] Herbert Marcuse. 1964. *One-dimensional man: Studies in the ideology of advanced industrial society*. Beacon Press.
- 477 [37] Herbert Marcuse. 1966. The individual in the great society. *Collected Papers of Herbert Marcuse Vol. 2: Towards a Critical Theory of Society* (1966),
59–80.
- 478 [38] Brian McInnis, Alissa Centivany, Juho Kim, Marta Poblet, Karen Levy, and Gilly Leshed. 2017. Crowdsourcing Law and Policy: A Design-Thinking
479 Approach to Crowd-Civic Systems. In *Companion of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (Portland,
480 Oregon, USA) (*CSCW '17 Companion*). Association for Computing Machinery, New York, NY, USA, 355–361. <https://doi.org/10.1145/3022198.3022656>
- 481 [39] Vishwajeet Narwal, Mohamed Hashim Salih, Jose Angel Lopez, Angel Ortega, John O'Donovan, Tobias Höllerer, and Saiph Savage. 2017. Automated
482 assistants to identify and prompt action on visual news bias. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in*
483 *Computing Systems*. 2796–2801.
- 484 [40] Linda Nilsson. 2020. Hybrid Events Breaking the Borders: Transferring your hybrid event into an engaging and inclusive experience for different
485 audiences and stakeholders. (2020).
- 486 [41] Jonas Oppenlaender, Maximilian Mackeprang, Abderrahmane Khiat, Maja Vuković, Jorge Goncalves, and Simo Hosio. 2019. DC²S²: Designing
487 Crowd-Powered Creativity Support Systems. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow,
488 Scotland Uk) (*CHI EA '19*). Association for Computing Machinery, New York, NY, USA, 1–8. <https://doi.org/10.1145/3290607.3299027>
- 489 [42] Jonas Oppenlaender, Kristy Milland, Aku Visuri, Panos Ipeirotis, and Simo Hosio. 2020. Creativity on paid crowdsourcing platforms. In *Proceedings*
490 *of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. ACM, New York, NY, USA, Article 548, 14 pages. [https://doi.org/10.](https://doi.org/10.1145/3313831.3376677)
491 [1145/3313831.3376677](https://doi.org/10.1145/3313831.3376677)
- 492 [43] Jonas Oppenlaender, Naghmi Shireen, Maximilian Mackeprang, Halil Erhan, Jorge Goncalves, and Simo Hosio. 2019. Crowd-Powered Interfaces
493 for Creative Design Thinking. In *Proceedings of the 2019 on Creativity and Cognition* (San Diego, CA, USA) (*C&C '19*). Association for Computing
494 Machinery, New York, NY, USA, 722–729. <https://doi.org/10.1145/3325480.3326553>
- 495 [44] Julian Posada. 2021. The Coloniality of Data Work in Latin America. *arXiv preprint arXiv:2105.06262* (2021).
- 496 [45] Uma Rani and Marianne Furrer. 2021. Digital labour platforms and new forms of flexible work in developing countries: Algorithmic management of
497 work and workers. *Competition & Change* 25, 2 (2021), 212–236.
- 498 [46] Daniela Retelny, Sébastien Robaszekiewicz, Alexandra To, Walter S. Lasecki, Jay Patel, Negar Rahmati, Tulsee Doshi, Melissa Valentine, and Michael S.
499 Bernstein. 2014. Expert Crowdsourcing with Flash Teams. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology*
500 *(Honolulu, Hawaii, USA) (UIST '14)*. Association for Computing Machinery, New York, NY, USA, 75–85. <https://doi.org/10.1145/2642918.2647409>
- 501 [47] Shruti Sannon and Dan Cosley. 2019. Privacy, Power, and Invisible Labor on Amazon Mechanical Turk. In *Proceedings of the 2019 CHI Conference*
502 *on Human Factors in Computing Systems* (Glasgow, Scotland Uk) (*CHI '19*). Association for Computing Machinery, New York, NY, USA, 1–12.
503 <https://doi.org/10.1145/3290605.3300512>
- 504 [48] Saiph Savage, Andres Monroy-Hernandez, and Tobias Höllerer. 2016. Botivist: Calling volunteers to action using online bots. In *Proceedings of the*
505 *19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. 813–822.
- 506 [49] Eric Schatzberg. 2018. *Technology: critical history of a concept*. University of Chicago Press.
- 507 [50] Florian Alexander Schmidt. 2019. *Crowdsourced production of AI training data: how human workers teach self-driving cars how to see*. Technical
508 Report 155. Hans-Böckler-Stiftung, Düsseldorf, Germany. <http://nbn-resolving.de/urn:nbn:de:101:1-2019102414513364122668>
- 509 [51] John Storey. 2018. *Cultural theory and popular culture: An introduction*. Routledge.
- 510 [52] Dmitry Ustalov, Fabio Casati, Alexey Drutsa, and Daria Baidakova (Eds.). 2020. *Crowd Science Workshop: Remoteness, Fairness, and Mechanisms as*
511 *Challenges of Data Supply by Humans for Automation*. Vancouver, BC, Canada (Online). <http://ceur-ws.org/Vol-2736/>
- 512 [53] Dmitry Ustalov, Fabio Casati, Alexey Drutsa, Ivan Stelmakh, Nikita Pavlichenko, and Daria Baidakova (Eds.). 2021. *Crowd Science Workshop: Trust,*
513 *Ethics, and Excellence in Crowdsourced Data Management at Scale*. Copenhagen, Denmark. <http://ceur-ws.org/Vol-2932/>
- 514 [54] Niels van Berkel, Simo Hosio, Jorge Goncalves, Katarzyna Wac, Vassilis Kostakos, and Anna Cox. 2018. MHC '18: International Workshop on
515 Mobile Human Contributions: Opportunities and Challenges. In *Proceedings of the 2018 ACM International Joint Conference and 2018 International*
516 *Symposium on Pervasive and Ubiquitous Computing and Wearable Computers* (Singapore, Singapore) (*UbiComp '18*). Association for Computing
517 Machinery, New York, NY, USA, 590–593. <https://doi.org/10.1145/3267305.3274134>
- 518 [55] PJ Wall, Deepak Saxena, and Suzana Brown. 2021. Artificial Intelligence in the Global South (AI4D): Potential and Risks. *arXiv preprint arXiv:2108.10093*
519 (2021).
- 520 [56] Robert West et al. 2021. Summary of Tutorials at The Web Conference 2021. In *Companion Proceedings of The Web Conference 2021 (WWW '21)*.
521 Association for Computing Machinery, Ljubljana, Slovenia, 727–733. <https://doi.org/10.1145/3442442.3453701>
- 522 [57] Richmond Y Wong, Ellen Van Wyk, and James Pierce. 2017. Real-fictional entanglements: Using science fiction and design fiction to interrogate
523 sensing technologies. In *Proceedings of the 2017 Conference on Designing Interactive Systems*. 567–579.