PARTNERSHIP ON AI

WHITE PAPER

Responsible Sourcing of Data Enrichment Services

JUNE 17, 2021



Contents

Executive Summary 4

1 Introduction 6

2 Methodology 7

3 Definitions 7

4 Models of Engagement for Data Enrichment Services 9

5 Worker-Oriented Considerations for Clients 10

6 Conclusion and Future Work 23

Annexes 25

Annex 1. Existing Tools and Resources for Clients 25

Annex 2. Considerations for Service Provider Selection 28

Annex 3. Existing Guidance for Human Rights, Workers Rights, Supply Chain, and Platforms 33

Endnotes 40

Acknowledgements

This white paper draws heavily on insights and input gathered during semi-structured interviews with members of the AI enrichment ecosystem conducted throughout 2020 as well as during a five-part workshop series held in the fall of 2020. The workshop series brought together more than 30 professionals from different areas of the data enrichment ecosystem, including representatives from data enrichment providers, researchers and product managers at AI companies, and leaders of civil society and labor organizations. We'd like to thank all of them for their engaged participation and for reviewing and providing valuable feedback on the drafts of this white paper.

We'd like to thank Elonnai Hickok for serving as the lead researcher on the project and Heather Gadonniex for her committed support and championship. This work would not be possible without the invaluable guidance, expertise, and generosity of Mary Gray. Many PAI staff members, current and former, contributed directly and indirectly to this work. In particular, Sonam Jindal is leading the project and served as the lead editor of the white paper, B Cavello led the workshop series, Katya Klinova oversaw the development of the project from its inception.

Participants of the Responsible Sourcing of Data Enrichment Services Workshop Series who contributed their expertise, insights, and suggestions include Aleli Alcala (Chief Operations & Sustainability Officer, Amara and PCF), Bama Athreya (Labor Rights Specialist), Olivia Blanchard (Researcher, Digital Future Society Think Tank), Jessica Custer (Manager, BSR), Bryan Dai (CEO, Daivergent), Orinola Gbadebo- Smith (Co-Founder & CEO, Hugo), Iva Gumnishka (Founder and CEO, Humans in the Loop), Dean Jansen (Executive Director, Participatory Culture Foundation), Kristen Itani Koue (Senior Impact Manager, Sama), Ivan Lee (CEO, Datasaur.ai), Milagros Miceli (TU Berlin/Weizenbaum Institute), Pamela Mishkin (Policy, OpenAI), Graeme Phillipson (Senior Project Engineer, BBC R&D), and Mark Sears (Founder & CEO, CloudFactory).

While individuals representing many of PAI's Partner organizations contributed text or suggestions to this document, it should not be read as representing the views of any specific member of the Partnership. Additionally, contributions from individuals do not necessarily reflect the views of their employers.

Please cite as following: *Partnership on AI, 2021. Responsible Sourcing of Data Enrichment Services.*

Available at: partnershiponai.org/responsible-sourcing

Executive Summary

As AI becomes increasingly pervasive, there has been growing and warranted concern over the effects of this technology on society. To fully understand these effects, however, one must closely examine the AI development process itself, which impacts the world both directly and through the models it creates. This white paper addresses an often overlooked aspect of the development process and what AI practitioners can do to help improve it: the working conditions of data enrichment professionals, without whom the value being generated by AI would be impossible.

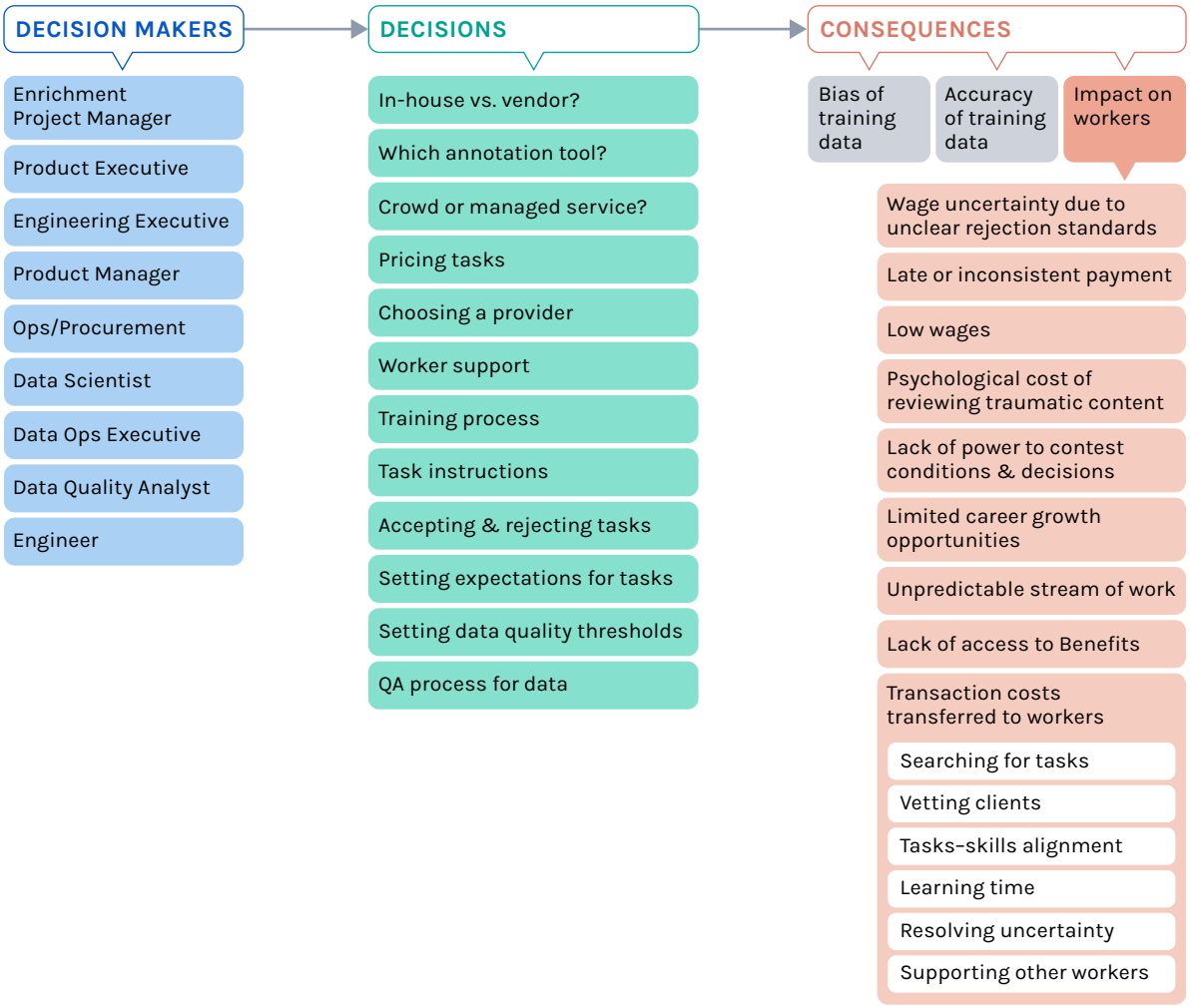
High-precision AI models are dependent on clean and labeled datasets. While obtaining and enriching data so it can be used to train models is sometimes perceived as a simple means to an end, this process is highly labor-intensive and often requires data enrichment workers to review, classify, and otherwise manage massive amounts of data. Given that this process of labeling and enriching data inherently embeds human judgement and lived experiences into data, AI's intelligence is highly dependent on human intelligence. Despite the foundational role played by these data enrichment professionals, a growing body of research reveals the precarious working conditions these essential, but largely unseen, workers face.[1] There is, however, an opportunity to make a difference. The decisions AI developers make while procuring enriched data have a meaningful impact on the working conditions of data enrichment professionals. This paper focuses on how these decisions during the procurement process impact workers and proposes avenues for AI developers to meaningfully improve these working conditions.

The decisions AI developers make while procuring enriched data have a meaningful impact on the working conditions of data enrichment professionals.

This paper draws upon existing literature on the experience of data enrichment professionals and insights gathered from AI developers and key stakeholders through conversations and a series of workshops. Acknowledging the existing complexity and lack of standards around how to build equitable data supply chains, we aim to critically evaluate the impact of the industry's current practices on workers, explore practices the industry can adopt to improve worker well-being, and advance the discourse around the future of data enrichment work and the indispensable role it plays in AI development. While more work and research is needed, we have outlined key worker-oriented considerations that practitioners can use as a starting point to raise conversations with internal teams and vendors. Specifically, this paper covers worker-centric considerations for AI companies making decisions in: selecting data enrichment providers, running pilots, designing data enrichment tasks and writing instructions, assigning tasks, defining payment terms and pricing, establishing a communication cadence with workers, conducting quality assurance, and offboarding workers from a project.

Our intention with this paper is to aid the industry in accounting for well-being when making decisions about data enrichment and to set the stage for further conversations within and across AI organizations. Recognizing the critical role that data enrichment professionals play in building AI is imperative, both for ensuring that their work is fairly recognized and compensated and for understanding that the resulting models are a product of human

intelligence. We hope this paper serves as a step forward, bringing us closer to a world where data enrichment professionals are recognized and rewarded by the industry for their central role in enabling AI advancement.



1. Introduction

The development and deployment of Artificial Intelligence (AI) systems relies on the cognition of human workers whose judgment and intelligence are widely employed to build the datasets used to train and validate models and ensure reliable real-time performance. This work ranges from preparing, cleaning, and labeling training data to providing human review of algorithmic outputs such as low-confidence predictions. For the purpose of this white paper, we refer to all of these tasks as "data enrichment work."*

* Please see the Definitions section for a working definition of "data enrichment" work.

The increase in AI development has given rise to a parallel industry in data enrichment work which serves as a growing source of jobs, particularly in the Global South.* Existing research on data enrichment professionals reveals the precarious working conditions they operate under. Workers often face inconsistent and inappropriate pricings for their work, unclear instructions, lack of recognition, and emotional and physical stress related to long and ad-hoc working hours and exposure to graphic content.[2] Some of these challenges are inherent to the work itself while others are shaped by company architectures, software used to mediate the work, business models, and client and vendor behavior.

* According to a Cognilytica report, the market for AI and machine learning preparation solutions has been estimated to grow to $1.2B by the end of 2023.

As the AI industry and the data enrichment workforce it relies on continue to grow, it is increasingly important to critically evaluate the conditions under which this work is being done. In particular, ensuring that these jobs are of a decent quality and provide for a decent level of worker well-being is crucial. Though there are many stakeholders in the industry that can and should play a role in ensuring favorable working conditions in the data enrichment industry—including policymakers, labor unions, civil society, investors, and company executives—this white paper will focus on the role of the immediate clients of data enrichment services. Clients making the day-to-day decisions related to sourcing data enrichment work for AI projects (such as product and program managers, AI developers, and data scientists) often shape the working conditions of data enrichment professionals and thus are in a position to directly make improvements.

Today, the data enrichment ecosystem is complex and unstandardized with few resources that clients can turn to for guidance on how to take concern for worker well-being into account when making sourcing decisions and how to incorporate practices that benefit workers. This has created a situation where, even if a client wants to make decisions that are mindful of their impact on workers' experiences, it is not easy for them to do so.

This white paper aims to make it simpler for clients to navigate this complex ecosystem, critically evaluate how their decisions may be impacting worker experience, and position themselves to develop better practices that benefit workers. The paper offers considerations for clients as they navigate the full process of sourcing and managing data enrichment work, from selecting a data enrichment service provider to writing instructions, setting up payment terms, and finally offboarding workers.

2. Methodology

This white paper draws heavily on existing research, media articles, international best practices, examples of company practice as found in company policy, and informal interviews with suppliers and clients of data enrichment work conducted by the Partnership on AI throughout 2020. The paper also draws heavily on comments and insights received during a five-week workshop series held in the fall of 2020, which brought together more than 30 professionals from different areas of the data enrichment ecosystem, including representatives from data enrichment providers, researchers and product managers at AI companies, and leaders of civil society and labor organizations.

The white paper recommendations are informed by:

- An analysis of practices and challenges that data enrichment service providers and their clients face, as identified through informal interviews with clients and providers in the ecosystem conducted throughout 2020
- A review of challenges data enrichment workers face and the positive and negative impact on workers as a result of that data enrichment work, associated business models, and client practices as identified in existing research and literature
- Regulations and guidelines covering business and human rights, supply chain and sourcing practices, and workers rights.

3. Definitions

Data enrichment work: Data curation for the purposes of machine learning model development that requires human judgment and intelligence. This can include data preparation, cleaning, labeling, and human review of algorithmic outputs, sometimes performed in real time. Examples of data enrichment work:

- Data preparation, annotation, cleaning, and validation:
    1. Intent recognition
    2. Sentiment tagging
    3. Image labeling
- Human review (sometimes referred to as “human in the loop”):
    1. Content moderation
    2. Validating low confidence algorithmic predictions
    3. Speech to text error correction

For the purposes of this white paper we refer to all these types of work as data enrichment work. The term “data enrichment” has been used by multiple companies in the industry to describe these services offered.* Other terms that have been used to refer to this work have

* For example: imerit CloudFactory Effect Force

included “data labelling,” “data annotation,” and “data curation.”

Sourcing data enrichment work: A process that requires a number of steps including, but not limited to, defining the enrichment goal, choosing the enrichment provider, defining the enrichment tools, defining the technical requirements, writing instructions, ensuring that instructions make sense, setting worker hours, determining time spent on a particular task, communicating with enrichment workers, rejecting or accepting work, defining a project budget, determining workers’ payment, checking work quality, and providing performance feedback.

Clients: For the purposes of this white paper we refer to professionals sourcing data enrichment work as “clients.” People in a number of different roles can be involved in sourcing data enrichment work: See section 3.1, “Mapping the Ecosystem for Sourcing Data Enrichment Work,” for more details.

Workers: For the purposes of this white paper we refer to individuals completing data enrichment as “workers.” In doing so, we recognize the variety of employment statuses that can exist in the data enrichment industry, including independent contractors on self-service crowdsourcing platforms, subcontractors of data enrichment providers, and full-time employees.

3.1 Mapping the Ecosystem for Sourcing Data Enrichment Work

There are a number of decisions made by both clients and service providers over the course of sourcing data enrichment work. These decisions can involve coordination across a range of roles and can be done in- house, in collaboration with a service provider, in collaboration with a third-party partner like an academic institution, or in some combination of these. Based on feedback received during the Responsible Sourcing workshop series held by PAI, it is clear that there is a wide range of actors and choices involved in sourcing data enrichment work. Sourcing data enrichment work involves decisions around defining the work, selecting a service provider, and engaging with a service provider, which are made by a variety of stakeholders—from data scientists to company executives—across the hierarchy of an organization. This large range of stakeholders suggests a similarly large range of people that have the potential to step in to improve conditions for data enrichment workers.

4. Models of Engagement for Data Enrichment Services

Typically data enrichment work is offered via four different engagement models which often get combined. These models are:

1. **In-house data enrichment:** Clients may have an in-house team to carry out data enrichment work. Such a team might be staffed by full-time employees or contractors brought on to carry out data enrichment work. Contractors may be located on the premise of the client, but might not be treated as full-time employees. Clients may build their own tools or leverage existing annotation tools to manage their data enrichment work.
2. **Managed data enrichment service provider:** Clients may choose to work with managed data enrichment service providers that find, train, and manage workers to enrich data according to the clients' specifications. Managed service providers can work in a variety of configurations including employing an in-house team, working with a set of subcontractors, or even setting up tasks on crowdworking platforms on behalf of clients. Clients of managed service providers do not always have full visibility into the specific employment configurations used by service providers. Depending on the configuration of the service, workers can be full-time employees, consultants, or independent contractors. Managed service providers typically support their clients in developing and refining instructions and task design, monitoring quality, and determining the price for the work.
3. **Self-service crowdsourcing platform:** Crowdsourcing platforms act as an intermediary for task-based work, connecting clients and workers. Policies and practices vary platform to platform and clients can be faced with different tools and processes for developing and assigning tasks, ensuring quality, setting prices, making payments, and engaging with workers. Crowdsourcing platforms can have curated workforces or may be open for anyone to join. Some platforms provide clients with the ability to work with a "private crowd" specifically assembled for the duration of the project. Others provide application programming interfaces (APIs) which allow clients to customise the platform's core functionality to meet their unique needs. Workers on crowdsourcing platforms are typically considered to be independent contractors. This model can be considered as a sub-segment of what is often referred to as the "gig economy," or "platform economy."* While there are platforms that are fully dedicated to providing data enrichment work, tasks such as data labelling are also frequently done on platforms that offer other kinds of task-based work.
4. Automated and synthetic: Software can be used to carry out data enrichment work such as labelling, annotating, and tagging features in data sets and can be used to create new data sets that contain necessary attributes.[3] Automated and synthetic

* For example, Arne Kalleberg and Michael Dunn describe four categories of work platforms in the gig economy: crowdwork platforms, transportation platforms, delivery/home task platforms, and online freelance platforms.

methods are typically used to supplement data enrichment work already being carried out by workers.

Each of the above models differ in terms of security requirements, cost, quality, flexibility, efficiency, and scalability. In-house services can provide the highest quality and security, but can be resource-intensive and less scalable.[4] As in other sectors,[5] an area of concern for managed services and crowdsourcing platforms could be unauthorized subcontracting.

5. Worker-Oriented Considerations for Clients

Upon identifying a need for data enrichment, it can be challenging to figure out the logistics of setting up a full data pipeline. There are numerous decisions to be made and often little guidance or established best practices. This section is meant to make it easier for those setting up data enrichment workflows to integrate workers' needs into the decision-making matrix from the outset. This section can also be used by companies with existing data enrichment workflows to critically analyze how their own practices may be impacting worker well-being and make changes. Collaborating with organizations that can bring in workers' perspectives and have a strong grounding in workers needs and rights can help achieve this and ensure that considerations of worker well-being are embedded within data enrichment workflows.

Drawing on the critical discourse around ethical supply chains and sourcing practices, labor rights, different dimensions of work via crowdsourcing platforms, and working conditions of data enrichment professionals, this section seeks to equip clients with key considerations necessary to make decisions that positively benefit workers. It also highlights how even decisions seemingly disconnected from workers can inadvertently impact them.

Once a client has determined the requirements of a data enrichment project, there are a series of steps and decisions that follow for selecting a service model and provider, defining the terms of engagement, and managing the entire data enrichment workflow. The rest of this section highlights choices made during the data enrichment process where clients should incorporate key worker-oriented considerations. As direct customers of enrichment services, clients' actions have a tangible impact on worker experience.

As direct customers of enrichment services, clients' actions have a tangible impact on worker experience. By incorporating the below considerations into their decisions, clients have the ability to positively influence workers' livelihoods and well-being.

We will highlight conscious practices clients should incorporate into their decision making during the following points of the data enrichment sourcing process:

- Selecting a service provider
- Managing data enrichment workflows, including:
    1. Running a pilot
    2. Designing tasks, developing instructions, creating training materials, and setting timeframes
    3. Assigning tasks
    4. Defining payment terms and pricing
    5. Establishing communication cadence
    6. Assuring quality
    7. Closure and offboarding

The sections below explain how each of the above decision points impact worker experience and provide recommendations for how clients can promote better practices. We recognize project requirements often guide clients' decisions around selection and management of the data enrichment work (e.g. timeframe, scale, data security needs, and available budget). By specifically highlighting how these decisions impact labor conditions, we hope to empower clients with the tools they need to incorporate the consideration of worker well-being into the decision-making process.

5.1. Service Provider Selection

After a client decides that they will need data enrichment work for the development of their AI solution, they will need to choose a service model and likely a vendor or company to engage with. As noted in Models of Engagement for Data Enrichment Work, the most common models by which clients take on data enrichment work are in-house data enrichment, managed service, crowdsourcing platforms, automated software, or some combination of these. In addition to taking into account how different solutions may meet the company's objectives with respect to scale, cost, security, and quality, it is also important to take into consideration the impact on working conditions for data enrichment professionals. Annex 1: Crowdsourcing Platform Comparison lists a few existing resources that compare crowdsourcing platforms on criteria such as transparency of terms of service, commitment to fair wages, etc. However, these resources do not fully account for different models for sourcing data enrichment work. While it may be difficult to provide a comprehensive guide that accounts for the full range of engagement models, we intend to work towards addressing this gap by building off of company-specific commitments that are emerging.* We have identified nine worker-oriented considerations that can be used by clients to guide decisions around selecting a data enrichment service provider. To the extent possible, we have adapted these to apply broadly to different models of sourcing data enrichment work. The considerations include:

1. What commitments to labor standards, models of ethical sourcing, and social missions are in place?
2. What worker-oriented protections and considerations are incorporated within the

* For example, Appen has developed a "Crowd Code of Ethics" which includes a commitment to fair pay, inclusion, crowdvoice, privacy and confidentiality, communication, and well-being.

terms of service, privacy and security policies, and redress mechanisms? How are workers' interests represented in these policies?

3. What benefits and support programs are in place for workers?
4. What information is available with respect to pricing models and base wages for workers?
5. What opportunities for recognition and reputation building are available to workers?
6. What work space, equipment, and forms of communication are available to workers?
7. What training is provided to workers and how is training time compensated?
8. What processes and mechanisms are in place to encourage diversity?
9. What policies and support structures are in place to address potential exposure to graphic or age-sensitive content?

Recognizing the differences between service models (i.e. managed service vs. crowdsourcing platform), we have explained how each of these considerations applies to various service models and have supported each consideration with an explanation of the implications for workers. The full list of considerations and explanations can be found in Annex 2: Considerations for Service Provider Selection.

5.2. Management of Data Enrichment Work

Clients undertake a series of steps related to the engagement and management of the data enrichment work. These include defining the enrichment tooling, defining technical requirements, designing enrichment tasks and writing instructions, ensuring instructions make sense to the workers, assigning tasks, setting payments and timeframes, rejecting or accepting work, checking quality, communicating with workers, and closing and offboarding of the project. While managed service providers may work alongside clients to manage some or all of these steps, clients directly using crowdsourcing platforms often manage these steps on their own. Various platforms offer differing levels of support, policies, processes, interfaces, and tooling for enrichment work. Management of workers and tasks can be partially or fully automated.* The below sections provide considerations for how clients can approach the various steps of managing data enrichment work and outline existing tools and guidance, with the aim of enabling positive outcomes for workers.

* There is a body of research that examines the use of algorithms for managing work on crowdsourcing platforms and the impact of the same on workers. For example, see:

"Good Gig, Bad Gig: Autonomy and Algorithmic Control in the Global Gig Economy"

"Algorithms that Divide and Unite: Delocalisation, Identity and Collective Action in 'Microwork'".

RUNNING A PILOT

As with most product development, developing AI solutions is often an iterative process that requires flexibility. As a result, it can be difficult to set realistic expectations on timing and cost for a data enrichment project at the beginning. By running a pilot with a smaller subset of data prior to implementing a data enrichment project, clients can establish a more realistic baseline for time and cost, refine task design, establish clear acceptance and rejection criteria for tasks, and assess impact of potential guidelines on workers.* More specifically, a pilot can help with:

- **Setting Timeframes:** During a pilot, clients can collect data and establish a baseline for the amount of time needed to complete all activities related to completing a task, including reading the instructions, reviewing examples, reviewing documentation, completing and submitting a task, and more. This can inform the final timeframe that is set for each task and the project.
- **Defining Per-Task Payment:** While there is additional complexity around defining the parameters of "per- task" work and what is included in a given "task," pilots can help companies develop baselines for how much time a task will realistically take for workers. The time required to complete a task is highly variable for different projects depending on the state of the data, amount of training necessary for workers, how long it takes workers to get used to a task, difficulty of the task, whether the task will require consulting outside sources, and more. In order to estimate the amount of time necessary for a given task, clients can deliberately evaluate the distribution of amounts of time it took workers to complete a task during a pilot. Using this distribution and a living wage base (usually based on hourly living wage for a given location), clients can calculate what would be a reasonable per-task payment. When such benchmarks are available, it is a good practice to compare that number to fair payments for similar tasks performed at larger scale.
- **Writing Good Instructions, Designing Tasks, and Ensuring Tool Usability:** During a pilot, clients can "test" their instructions, worker experience of completing the tasks based on task design, and usability of the enrichment tool. They can do this by collecting direct feedback from workers through surveys, regular check-in sessions, and worker interviews. This feedback can inform any necessary improvements before scaling the enrichment process.

* For example, researchers at Stanford University have looked to improve the quality of task designs through "prototype tasks," a strategy that requires all new tasks to be run through a rapid sample run where workers have the ability to provide feedback on the task design. The research found that running a prototype with a small sample of workers resulted in better outcomes for both clients and workers.

Running a pilot to understand what would be a proportionate payment has also been recommended by researchers.

Similar to any product pilot, it is important to follow research best practices to limit skewed results. That being said, being mindful of how a data enrichment project is being set up and running a pilot to gather workers' feedback can meaningfully help clients to set realistic timeframes, improve task instructions and tools, improve worker experience, and establish a fair price for each task.

DESIGNING TASKS, DEVELOPING INSTRUCTIONS, CREATING TRAINING MATERIALS, AND SETTING TIMEFRAMES

Designing tasks, developing instructions, developing training materials, outlining performance expectations, and establishing clear timeframes are essential components of setting up a data enrichment project. In most cases, clients are involved in each of these decisions which have a meaningful influence on both the workers' experience with fulfilling the tasks and on the quality of enriched data. If the client has engaged with a managed service provider, they may work together to develop the project parameters, instructions, training materials, etc. In these circumstances, clients may stipulate the tool or platform that the managed service should use or defer to the managed service's preferred tools.

Though crafting effective instructions can be challenging, investing in this process is critical. Taking the time to translate data enrichment needs into clear and concise instructions can save time in the long run by creating less confusion around guidelines and therefore less back-and-forth when executed tasks do not meet the necessary standards. By investing up front in explicit instructions with clearly communicated expectations, clients can decrease the chances of having to redo work, thereby making it more likely to meet budgets and timelines. As stated above, there is value in testing these instructions with workers during a pilot or in early feedback sessions with a smaller subset of workers. Intentionally prioritizing this process can not only save time and money, but positively shape worker experience. In addition to incorrect and delayed work for a project, research has highlighted that unclear instructions can impact workers ability to succeed[6] and can result in multiple iterations of a task, rejected and uncompensated work, or tasks that timeout.[7] On crowdsourcing platforms, multiple iterations and rejected work have a significant impact on a worker's ratings and can result in nonpayment for a task depending on how the platform is designed. Penalizing a worker's rating because they did not meet vaguely specified expectations can unfairly preclude them from getting future tasks. This places an unreasonable burden on workers when instructions are not clear and threatens their source of income. Furthermore, unclear instructions on crowdsourcing platforms create a situation where even workers who are putting in their best effort and investing time in completing tasks may have their work rejected and unpaid.

There is a body of research that has examined the challenge of task design and writing instructions on crowdsourcing platforms and has sought to develop solutions. Some companies have also published guidance on how to develop effective instructions for data enrichment work. These resources can be found in Annex 1: Task Design and Writing Instructions. Here, we highlight a few practices that can improve outcomes for workers:

Designing Tasks and Developing Instructions:

1. Define clear and consistent rules for what constitutes a well-executed task. Test them internally, as well as with data enrichment workers prior to implementing a task. Ensure continuous communication with workers if questions or issues arise.
2. Incorporate worker feedback into the instructions, particularly with respect to any unclear aspects of the tasks. This is important for both improving processes and empowering workers.
3. Keep in mind the audience when crafting instructions. Depending on the project, the team crafting instructions may have more extensive domain knowledge than the workers conducting the data enrichment work under tight timelines. Make sure that instructions provide enough context to enable workers to complete the task in the expected amount of time. As addressed earlier, testing the instructions with the workers can help to ensure their perspectives and questions are incorporated into the final instructions. Providing examples of correct and incorrect work can also go a long way in establishing clear expectations for workers.

4. To the extent possible, communicate the purpose of a task and how it connects to a larger project or objective.*

5. Ensure that consent forms and confidentiality agreements provide workers with the necessary context on how the enrichment work results will be used.*

6. Ensure that tasks are designed with the tool in mind in order to produce a clear and intuitive user experience. Additionally, ensure that instructions address how to navigate and use the tool to complete tasks efficiently and how to address any technical difficulties that may arise.

* Research has explored different ways that training can be provided to workers on crowdsourcing platforms. Among other things, the research found that offering feedback and the purpose of a task positively impacted worker motivation.

Developing Training Materials:

7. Analyze what type of background knowledge and training is needed in order to effectively complete the relevant tasks. Design and provide any necessary training to workers which will make the work easier for them and improve the quality of work.* Ensure that the training time is compensated.

* This has been recommended in guidelines on the use of crowdsourcing platforms from a number of Universities.

Setting Timeframes:

8. When setting a timeframe, take into consideration the time needed to go through any preparatory work, complete the tasks themselves, and review work prior to submission. Preparatory work may involve reviewing instructions, reviewing consent forms and other associated documentation, and going through any required training.* Timeframes should also account for the amount of time it may take for workers to get acquainted with the user flow necessary to complete the task and some buffer time to address potential technical issues. One way to approach this would be to collect more granular data during the pilot to establish a baseline that takes these factors into account. In suggesting this, we recommend working closely with the workers during the pilot to get an accurate accounting of how much time was needed for the pre-task, task, and post-task activities. Another way may be to add a generous buffer to the slowest time from your pilot results. To the extent possible, clients should verify if the initial time estimates to complete a task were accurate and use this information to make adjustments.

* For example, research has found that providing training to workers when necessary has been found to be an effective method of quality assurance.

9. If the workers on your project are not exclusively working for your team, design tasks and timeframes in a way that allows workers and contributors flexibility in how and when they complete a task so they can plan for their other work obligations.[8]

* This approach has been recommended by the University of Waterloo in guidance on the use of crowdsourcing platforms.

ASSIGNING TASKS

Depending on the engagement models, the mechanisms by which tasks are assigned to workers may differ. Clients and managed service providers can be more directly involved with assignments when working with contractors or full-time employees. Crowdsourcing platforms typically provide recommendations and filters to support task assignment based on parameters like skill set, area of expertise, qualifications, rating, performance statistics, work histories, test scores, rejection rates and others.[9] These matching algorithms have been controversial and raised questions about fairness, the bias that these systems may inadvertently bring in, the impact of these on worker autonomy, and the way in which workers may

unfairly lose access to platforms they rely on to make a living.[10] It is important that clients exercise caution and incorporate a deliberate consideration of worker well-being when using automated matching algorithms and filters.

Finding the right worker for a task is important and can be challenging. As researchers have noted, a mismatch between the skills and knowledge required for a task can result in delayed projects, inaccurate work, and rejected work.* Rejected work is particularly costly to workers who may have committed to a task before being shown the full details: not only do workers usually get penalized in their ratings for rejected work, they may not get paid for the critical time they have already invested in completing the task to the best of their ability. In other employment models, employers invest significant energy in finding the right fit between workers and skills required. When an imperfect algorithm is used for matching, the transaction costs of finding the right fit usually falls on workers. Researchers have explored ways in which platforms can improve matching tasks and workers by taking into consideration nuanced characteristics of both.[11] Other work has explored solutions that can help clients navigate worker selection on crowdsourcing platforms by bringing together different dimensions related to pricing, task difficulty, and worker skill.[12]

* For example, research has explored the creation of human-oriented frameworks for crowdsourcing towards addressing issues like unfair compensation, incompatible task assignments, and unintended amplification of human biases.

Considerations related to task assignment that can improve the outcome for workers and the quality of data include:

- **Redundancy:** To improve enriched data accuracy as well as reduce biases, consider assigning multiple workers to the same tasks to confirm the results are the same.
- **Workforce Consistency:** Given the importance of consistency across related datasets, the level of skill that is needed for nuanced complex data enrichment work, and the value of strong relationships in ensuring smooth work processes, clients should consider engaging with the same workers on a crowdsourcing platform or team at a managed service provider. This can also allow workers to build reputations, relationships, and skill sets which can all be leveraged to find opportunities in the future. Engaging with the same workers can also help ensure high-quality enriched datasets.[13]
- **Attention to Diversity:** Given the subjective nature of classifying and labeling data, it is important to pay attention to the diversity of data enrichment workers. A lack of diversity could be a source of bias in the labeled dataset.
- **Mindful Screening Criteria:** When selecting how to screen workers, consider the impact that each applied filter may have on the workers, and if the filter accurately captures the skills or qualities needed. For example, research has recommended that workers should not be screened based on non-payment rates as non-payment does not necessarily reflect quality and that workers should not be penalized (through poor ratings or other actions) for refusing to accept a task.*

* Worker evaluations and ratings should not be based on non-payment rates and workers should be given reasons for any negative ratings. For more information see: Digital Labour Platforms and the Future of Work.

DEFINING PAYMENT TERMS AND PRICING

Clients set or negotiate payment for data enrichment work. There are four predominant pricing models for data enrichment services:

- **Per Task:** This is a common payment model on crowdsourcing platforms. If data enrichment workers are being paid by the task, the price per task is often set by the client, sometimes with input from the service or platform they are working with* or through a bidding process facilitated by the platform.[14] Given the international makeup of workers on crowdsourcing platforms, a bidding process in which workers bid for tasks drives prices per task down. Adding to worker precarity, many platforms allow clients to reject work without payment after workers have already completed the task.[15] There are no standard rules or guidelines to protect workers from unpredictable payment rates.
- **Per Hour:** If workers are being paid by the hour, they are typically paid for the total amount of time spent completing the necessary tasks. Time can be tracked either by the tool being used or through manual time-recording through timesheets. If a tool is recording time, how that time is measured is critical to ensuring that workers are being paid fairly.
- **Per Tier of Service:** Clients may also pay managed service providers a fixed fee for a given tier of service and use per task payment if there is a need to go beyond what is included in the base price. Under this service model, the managed service provider would typically determine the workers' wages.
- **Flat Fee:** Clients may also pay managed service providers a flat amount per project delivered which is negotiated specifically for each individual project.

* For example, research that undertook a comparison of platforms has noted that platforms such as LeadGenius automatically set the price per task based on estimations of completed time, effort, and local hourly wages. The prices can be adjusted if the task takes longer than estimated. For more information see: Beyond Mechanical Turk: An Analysis of Paid Crowd Work Platforms.

Independent of the pricing model used, running a pilot and collecting data about the amount of time it takes workers to complete all activities related to a given task is critical to estimating how a task should be priced. The amount of time it takes to do the task and any necessary pre- and post-completion activities should all be used to estimate a worker's hourly compensation and accurately assess if the per hour total compensation is fair.

When using a crowdsourcing platform directly or working with a managed service that uses a crowdsourcing platform, it is critical to be attentive to how the platform structures its payments with workers. Clients engaging with crowdsourcing platforms can make a difference in workers' lives by watching out for the following:

- **Unreasonably Low-Priced Tasks:** A wide range of tasks and task types, a lack of standards for how tasks should be priced, and platform designs which force workers to bid for tasks all contribute to tasks being priced down. In particular, as mentioned above, when platforms force workers to compete for tasks with workers from around the world, there is a race to the bottom. Without the client's active consideration, workers may end up getting paid unreasonably low amounts. When choosing a platform to work with or working with a managed service using

a crowdsourcing platform, clients should make sure they have the ability to pay workers a fair amount for their contributions. The next section covers ways to calculate this.

- **Lost Wages:** Another important payment term consideration when choosing a platform or vendor is how the platform handles accepting and rejecting completed tasks. While investing in a pilot that helps craft effective instructions and training materials can limit rejections later in the project, some work may still need to get rejected occasionally. If completed work is rejected without explanation and the worker's ratings go down, this can make it harder for them to get work later, result in lost wages for the work they have already put in, and can create a power imbalance if workers do not have meaningful avenues for redress.[16] Some of this can be mitigated by ensuring that the crowdsourcing platform being used provides workers with transparency over why a task is rejected and workers have the ability to contest rejected tasks. Furthermore, the impact of rejections on lost pay can be decreased by reviewing work promptly and providing fair and detailed justifications for rejections, so workers can learn from the client's acceptance criteria and identify early on if their skill set does not match the needs of the project and step down before investing too much time. Clients should still pay for rejected work if the rejection occurred for reasons outside of the worker's control.
- **Non-Monetary Payments:** Clients should ensure that the service or crowdsourcing platform they use is paying workers in cash as opposed to vouchers or rewards. Furthermore, it is worth asking the platform or managed service about their payment cadence to ensure they are regularly paying their workers.
- **Additional and Hidden Costs Borne by Workers:** In the absence of a formal management team, many of the functions that a "traditional" employer usually takes on as a part of managing their workforce come to land on workers' shoulders.[17] This is particularly true of crowdsourcing platforms. Workers take on the costs associated with tracking their own progress, searching for tasks, vetting clients, learning how to do tasks, resolving uncertainty when there is no one to answer questions, and oftentimes take on equipment costs.

  Furthermore, they do not have access to benefits, pay local taxes, and often pay a transaction fee per task to the platform* or an overarching fee to access the platform.* When choosing a crowdsourcing platform to use, either directly or through a managed service, clients should evaluate the additional, uncovered costs workers bear on that platform and consider how the pricing strategy should reflect those additional costs.

* On Freelancer, for example, the platform takes a 3% cut from employers and a 10% cut from freelancers.

* For example, according to Fair Crowd Work, CrowdFlower charges $1,500 p/m for access to the platform.

As data enrichment becomes an increasingly common job,[18] it is crucial to address the low wages that characterize this sector. Clients have reported the challenges associated with identifying fair prices for data enrichment.[19] For crowdsourcing platforms, many best practices have highlighted the importance of paying at least a minimum wage in a jurisdiction (of

the employer and employee), a fair wage, or a living wage.* Other solutions have explored the possibility of clients committing to a "wage pledge" to pay a minimum or living wage.[20] While the global labor market is increasingly complex, below is a summary of a few approaches clients can take to estimate a fair price for this labor.

* See table comparing crowdsourcing principles and best practices in Annex 3.

- **Calculate and Pay at Least a Living Wage:** When working with a managed service provider, clients can demand transparency in the wage structure and ensure contracts with workers include a guaranteed living wage or at least the local minimum wage. There are a number of resources that can be used to calculate a living wage which are listed in Annex 1: Calculating a Living and Minimum Wage. When using these tools or setting a price, it is important to recognize the difference between a living wage (amount an individual needs to cover basic costs) and the minimum wage (legal minimum pay per hour), to consider where workers are located and living costs associated with the location, and to critically analyze the tool's inputs, how the information it uses was obtained, and how frequently it is updated, if at all.
- **Account for Additional Costs:** When relevant (especially when using crowd-sourcing platforms), account for costs that workers may bear including time spent searching for tasks, time spent training and learning the parameters of a task, time spent reviewing task samples, platform fees workers may pay to access the platform, local taxes, equipment costs, and cost of basic benefits, such as healthcare and sick leave. In order to account for these additional costs, research suggests offering individuals categorized as self- employed a multiple of the minimum wage based on the worker's location or a multiple of the median local wage earned by employed individuals.[21] Recognizing that it can be challenging for clients to know the location of a worker, it can be useful for clients to use the maximum minimum wage for OECD and non-OECD countries.*
- **Leverage Data on Similar Tasks:** While pilots can help refine task pricing, referring to similar tasks in the past can also inform initial estimations.* That being said, prices should be adjusted if workers or data suggest that a task is taking more time than initially expected. If this is discovered after the fact, clients can compensate workers with bonuses.[22] Prices should also be adjusted based on difficulty of task.*
- **Track Task Completion Time:** Clients should confirm whether the actual completion times per task match the initial estimate. Depending on the data enrichment tool, this can be completed via the tool or can be ascertained via random sampling. The results can be used to offer a wage adjustment to workers and inform future iterations of the project.
- **Compensate for Changes That Occur on the Client Side:** Pay workers for lost time if there were workflow problems due to a lapse on the client's side, due to changes in the instructions or the scope of work, or unforeseen technical issues on the client or platform side.[23]

* See wage data for OECD countries

* This approach was suggested by the University of Waterloo in guidance on the use of crowdsourcing services for research.

* This approach was recommended in guidance to MIT researchers.

ESTABLISHING COMMUNICATION CADENCE

Clear communication is critical to ensuring that workers have the information they need to effectively and efficiently complete tasks. Research has shown miscommunication as an area of frustration for both workers and clients.[24] Having clear communication can positively impact worker experience, decrease the total amount of time to resolve uncertainties, and mitigate the risk of lost pay or missed deadlines due to misalignment. Depending on the engagement model, the communication methods and responsibilities may differ. Regardless of who is carrying out the communication, it is important to make sure there is an established process to communicate task assignments, training procedures, expectations around acceptable tasks, how to resolve uncertainty, and who to reach out to if any problems arise.

Communication practices that clients can use to improve the outcomes for workers include:

- **Seek out feedback on communication from workers and make necessary adjustments.** If you are using a pilot, ensure that the communication procedures and materials intended to be used during the project are also tested during the pilot. Another way to test communication procedures may be to have a targeted feedback session with a representative group of workers before widespread dissemination of communication materials and procedures. Getting this feedback earlier on can save time in the long run as it will help avoid misalignment.
- **Use services that allow for clear communication with workers.** When working with any service, particularly if a crowdsourcing platform is involved, consider which mechanisms are available to facilitate communication between workers and those overseeing the data enrichment process (either from the client or vendor side).
- **Clarify expectations around communication** prior to the start of the project and ensure workers know how to resolve uncertainty or who to reach out to should any issues arise. Additionally, make sure you have procedures in place to provide a quick turnaround for workers regarding any questions or concerns.*
- **Provide early and consistent feedback.** Establish regular milestones and checkpoints to review work and provide feedback to ensure that workers are aligned with the project's expectations. Regular feedback allows workers to make adjustments early on and help avoid project delays or extra costs arising from having to redo work.[25]
- **Establish clear mechanisms for workers to contest rejections** if they have reason to believe their work should not have been rejected. As mentioned earlier, providing feedback around when work meets or does not meet the acceptance criteria can allow workers to resolve issues earlier and make sure they are aligned with the project's guidelines. In addition to making sure that the acceptance and rejection process is accurate and fair, allowing workers to contest rejections can improve the overall quality of the data and provide workers with an avenue to recover ratings or lost wages.
- **Build trust with workers** by identifying yourself or the company you represent in the task name or description.[26]

* For example, guidance for MIT researchers recommends that requesters respond to emails from workers within seven working days.

QUALITY ASSURANCE

As a critical input into AI models, the enriched data needs to be of a high quality. Once tasks are complete, clients or managed service providers typically undertake a final evaluation of the work and accept, reject, or request modifications. Evaluating the quality of enriched data can be challenging and raises questions about how to identify and handle inaccurate data, how to assess biased data, and how to handle payment for data that may not meet the necessary standards.

While setting clear expectations prior to the start of the project can mitigate some of the quality risks, it is also important to design the quality assurance process with workers in mind. Depending on the engagement model, the quality assurance process may look different. While managed service providers often take on an active role in supporting clients with conducting quality assurance, crowdsourcing platforms vary in terms of the support and mechanisms they provide.

There are many resources that explore different methods for quality assurance on crowdsourcing platforms. A few common methods include assigning test tasks to workers, tracking historical accuracy,[27] having workers go through training projects, and having multiple workers do the same tasks and reviewing tasks with disagreements.[28] Some of these approaches can be time-intensive and difficult to scale. Quality is highly dependent on a number of factors including the working conditions of data enrichment professionals, the quality of training provided to workers, the way tasks are designed, and the processes and mechanisms established by a particular platform.[29] It is worth calling out that some tasks require subjective judgement and can be labeled differently depending on the background of the workers and some tasks may require workers with a specific background or skill set. Being able to recruit a worker base with the specialized knowledge and relevant diverse backgrounds to complete tasks can determine the quality of the enriched dataset. Resources that dive into best practices around establishing quality assurance practices are further outlined in Annex 1: Quality Assurance.

In addition to influencing dataset accuracy and bias, quality assurance practices also have implications for workers. On a crowdsourcing platform, clients exercise their discretion in rejecting work they find to be inaccurate. For workers, this can result in unpaid labor, lower ratings, and lack of access to future work.[30]

Worker-minded practices to consider when establishing a quality assurance routine include:

- Ensure that the raw, unenriched data is of a high quality so it is easy to decipher for workers. This will result in higher quality enrichment and less inaccuracies.
- Clearly communicate acceptance and rejection criteria to workers prior to the start of the enrichment project.* Clarify how quality will be measured ahead of time in order to ensure that expectations are clear to all parties involved.
- Provide workers with examples of correct and incorrect work. This will also help workers familiarize themselves with the project's expectations.

* This approach was recommended to researchers at Berkeley.

- Provide workers with an opportunity to complete a sample of the work either in a test environment or a test project and confirm its accuracy prior to the start of the project.
- Provide mechanisms for workers to correct work upon receiving early feedback.[31] This will minimize the overall quantity of incorrect work done.
- Ensure that workers have enough time to complete a given task and are paid appropriately for their time so they do not feel pressured to do tasks too quickly.
- If work is rejected for a reason not connected to quality, this should be clearly indicated. In these circumstances, make sure workers' ratings are not negatively impacted.
- If work is rejected and not compensated, include an explanation of why and provide a commitment to the worker that the work will not be used.[32]
- If a negative review is given to a worker, the reasons should be fully supported and communicated to the worker.[33]

CLOSURE AND OFFBOARDING

It is important that workers are recognized and have the ability to provide feedback as part of the closure and offboarding process. These steps can help to bring recognition, power, and voice to workers.

Practices that can improve the outcome for workers during project offboarding include:

- Actively seek feedback from workers in order to make improvements in the future.*
- Recognize workers' contributions in context of the larger project. This will allow workers to build a reputation and portfolio of work which can help them develop their careers as data enrichment professionals. Recognizing the central role played by these professionals also helps push the industry towards greater transparency around the AI development process, which is critical to treating these workers with the respect they deserve and ensuring that their work is not hidden.

* For example, guidance for researchers from the University of Massachusetts recommends the sharing of a debriefing form with workers.

6. Conclusion and Future Work

As the complexity of AI systems continues to increase, so too will the demand for data enrichment work. It is important that workers at the heart of this growth are respected, supported, and fairly compensated for their contributions. Though there are a number of stakeholders that play a key role in shaping the working conditions of individuals in the data enrichment sector, this paper focuses on recommendations for the sector's clients given their involvement in everyday decisions which directly or indirectly impact workers during the data enrichment process. We hope that equipping clients with an understanding of how their choices impact workers will empower them to make decisions that prioritize worker well-being.

Dedicating attention to worker well-being is all the more important due to the complexity of data supply chains and the lack of standardization in data enrichment practices across the industry. While we acknowledge that there is more work to be done to transform industry practice, this white paper highlights key junctures during the data enrichment process where clients should incorporate worker-oriented thinking into their decision-making process. Specifically, we have examined how clients can incorporate impact on worker well-being into their decision-making process around data enrichment provider selection, running a pilot, task design and writing instructions, task assignment, pricing, communication, quality assurance, and closure and offboarding. We highlight existing research and best practices that clients can incorporate into their decisions.

The data enrichment process has always been an essential part of AI development. Recognizing that AI advances are contingent on data enrichment labor and taking steps to invest in supporting this labor force are similarly essential to the future of this industry. As AI becomes more complex, data enrichment needs and the demands being placed on workers are also growing in complexity. This complexity also creates a need for more workers with specialized knowledge to complete data enrichment work. For the AI industry to grow sustainably, creating the infrastructure to transform data enrichment work into decent jobs is imperative.

Future research and work are needed to address a number of questions:

1. **Systems of Recognition:** Data enrichment work remains largely hidden and done in informal work arrangements. How can workers be recognized for their contributions to the products that rely on data enrichment work? What circumstances are leading to this labor being underrecognized and undervalued today?
2. **Measuring and Evaluating Working Conditions:** How can we work towards a reliable system of evaluation to assess and verify working conditions in the data enrichment sector? What metrics and targets should be adopted around responsible data enrichment sourcing?

3. **Models for Payment and Ensuring Fair Wages:** Is there a way to create a standardized approach to determine a fair price for data enrichment work? How can fair wages be ensured for workers supplying their labor on crowdsourcing platforms?
4. **Transparency:** Given the complex and opaque nature of the data enrichment ecosystem, how can we create greater transparency around data supply chains and the practices undertaken in the industry to source and enrich data?
5. **Standardizing Instructions:** Considering the importance of and difficulty in creating high-quality instructions for data enrichment work, how can we drive towards standard practices that will result in clearer instructions? Are there ways to standardize approaches based on the data enrichment technique being used? For example, are there best practices that can be standardized and used across instructions for image annotation or text sentiment analysis?
6. **Portable Benefits System:** Are existing proposals for portable benefits attainable? Can a portable benefits system function internationally? How can we ensure all data enrichment workers have access to basic benefits like healthcare and paid sick leave?
7. **Evaluating Data Enrichment Work:** On crowdsourcing platforms, are there objective mechanisms that can be developed for evaluating data enrichment work beyond rating systems?
8. **Developing a Deeper Understanding the Data Enrichment Ecosystem:** What forces are currently shaping the data labor market? What interventions are needed to produce better outcomes for workers?

While these questions may be specific to data enrichment work and labor in the context of AI, this work has broader implications for the way labor is organized in society. Much of the precarity that characterizes data enrichment jobs can be attributed to how this labor is algorithmically managed and broken into task-sized chunks, making this critical work appear less like our traditional conceptualization of a "job."[34] As researchers have pointed out, many other jobs are at risk of being fragmented in the same way and might be converted into task-based work in the near future. How task-based work is classified and recognized in society has implications for how labor and "knowledge" is valued within the context of the expanding "knowledge economy." While some data enrichment tasks may seem simple, they require sustained concentration, and often nuanced, domain-specific knowledge and experience. As AI gets more advanced, more people's expertise will be needed to support additional use cases. Recognizing the importance of data enrichment workers and building infrastructure to support them and improve their working conditions is critical to the future of the AI industry.

ANNEX 1

Existing Tools and Resources for Clients

Crowdsourcing Platform Comparison and Related Resources

Existing tools that can be used by clients to learn more about different practices on crowdsourcing platforms and review research comparing crowdsourcing platforms include:

- Fair Crowd Work: Fair Crowd Work provides reviews of crowdsourcing platforms based on the criteria of pay, communication, evaluation, tasks, technology, ability to refuse payment, terms of service, reviews from workers, and reviews from clients.
- Beyond Mechanical Turk: An Analysis of Paid Crowd Work Platforms: Researchers at the University of Texas at Austin undertook a cross-platform analysis to see how platforms compare across a number of attributes, including workforce composition, demographics and worker identities represented, mechanisms for tracking qualifications and reputation, management structures, incentive mechanisms, support to ensure quality assurance and control, accessibility of the tool and types of services offered, support for specialized and complex tasks, and promotion of ethics and sustainability.
- The Online Labour Index: Developed by the iLabour Project at the Oxford Internet Institute, this index quantifies key measures describing the online gig economy. It tracks labor markets moderated over the internet across countries, primarily through online platforms.

Calculating a Living Wage

Calculating wages for data enrichment workers is complicated due to the international nature of this labor market and due to this work typically being done on a per-task basis. Though we recognize these challenges and acknowledge more work needs to be done to create accessible standards, below are existing tools and resources that can be used to determine a living wage and/or minimum wage:

- MIT Living Wage Calculator: Provides living wages in locations across the United States.
- Anker Methodology: The Anker methodology calculates a living wage by estimating the cost of a decent lifestyle and taking into consideration housing, healthcare, education, groceries, and transportation as well as deductions, benefits etc.
- Living Wage Foundation: Provides accreditation to UK employers who commit to pay a living wage to direct employees and contractors and meet the necessary requirements.

- Living Wage: This is an open source tool by OpenUp that will tell a user if they are paying a living wage to domestic workers in South Africa.
- WageIndicator: The foundation provides resources and information about living wages for more than 110 countries.
- Global Living Wage Coalition: This coalition provides information about living wages in 27 countries across the globe using the Anker Methodology. The coalition defines a living wage as: "The remuneration received for a standard workweek by a worker in a particular place sufficient to afford a decent standard of living for the worker and her or his family. Elements of a decent standard of living include food, water, housing, education, healthcare, transportation, clothing and other essential needs including provision for unexpected events."
- Fair Wage Guide: Developed by Good World Solutions, this tool calculates wages and compares them to local and international standards. As a note, the tool only calculates minimum wage and international poverty lines and does not calculate a living wage.
- Fair Work for Amazon Mechanical Turk: A tool developed by Stanford researchers that can be used by clients on Amazon Mechanical Turk to "ensure that workers are paid at least a minimum wage."
- Crowd-Workers: A browser extension developed by the University of Pennsylvania that enables workers to sort HITs based on an hourly rate.

Task Design and Writing Instructions

Below are existing tools and research that can guide clients in designing tasks and writing instructions:

- Google Cloud: Provides guidance on designing and developing instructions for human labelers.
- Sprout: An open source tool developed by researchers at the University of Washington and the Indian Institute of Technology that helps clients improve task design on crowdsourcing platforms by collecting feedback from workers, synthesizing this feedback for clients, and providing suggestions to clients to improve task design.
- TurKit: A toolkit designed by researchers at MIT "for deploying iterative tasks on mechanical Turk."
- CrowdWeaver: Developed by researchers at Carnegie Mellon University, CrowWeaver provides graphical tools to help manage and track worker progress.
- Structured Labeling: Researchers at Oregon State University and Microsoft Research have proposed structured labeling solutions to facilitate consistent labeling.

- Revolt: Developed by researchers at Carnegie Mellon University and Microsoft Research, Revolt seeks to leverage disagreements to achieve higher label accuracy and create reusable structures with a more nuanced range of labels.
- Turkomatic: Developed by researchers at the University of California, Berkeley and Stanford University, Turkomatic seeks to create a collaborative process for workflow design by leveraging input from crowd workers to help clients design and carry out complex tasks.
- Fantasktic: Developed by researchers at the University of Berkeley, Fantasktic seeks to improve task design by providing an interface with guidelines and recommendations that can be used by clients, a preview interface that allows clients to see the task from the perspective of the worker or contributor, and automatically generated task tutorials.
- WingIt: Developed by researchers at Purdue University, WingIt proposes a system that enables workers to resolve ambiguities in task instructions by enabling workers to ask questions, propose edits to tasks, and discuss ambiguities with other workers. The work calls out three main types of ambiguity in instructions relating to input, process, and output.

Quality Assurance

Existing tools and research that can guide clients in establishing quality assurance practices on crowdsourcing platforms are outlined below:

- Quality Management on Amazon Mechanical Turk: Researchers from New York University developed an algorithm that generates an evaluation of quality that takes into account worker bias and error rate for a given worker.
- Quality Control in Crowdsourcing: This research provides a comprehensive taxonomy and overview of existing quality control aspects and techniques, a review of quality assurance mechanisms on fourteen different platforms, and proposes a model for crowdsourcing platforms built around people.
- How Many Workers to Ask?: Adaptive Exploration for Collecting High Quality Labels: Research that explores the question of how many workers are needed to complete a task to ensure statistically significant results through the development of an algorithm that draws upon the quality score of a worker and the difficulty of a hit.
- Quality Control in Crowdsourcing Systems: Issues and Directions: Researchers have provided a taxonomy of quality and quality assurance techniques where quality is characterized by task design and worker profiles.
- Quality Management in Crowdsourcing using Gold Judges Behavior: Research that explores the effectiveness of embedding known answers as a method to ensure quality. This research ultimately concludes that embedding gold-standard known “answers” is a useful technique to improve quality.

ANNEX 2

Considerations for Service Provider Selection

Below are worker-oriented considerations that clients should weigh when selecting a service provider for data enrichment work.

1. Commitment to Labor Standards, Worker-Oriented Policies, Ethical Sourcing, and Sustainability

Criteria:

Clients should consider and analyze what commitments a prospective service provider has made to adhere to labor standards and ethical sourcing. While some service providers may have certifications demonstrating they have met the standards of an independent agency, others may articulate commitments on their own. It is important to consider both the commitments they have made and how they can be held accountable for those commitments. If a service provider does not explicitly mention any worker oriented policies or standards, explicitly asking them can help push them to provide greater transparency over their practices and provide a powerful signal that this is an important consideration for clients.

In evaluating service providers' commitments, clients can refer to the recommendations made in this paper to see how potential providers' practices compare and consult a number of applicable international instruments and standards. A few standards to refer to are the UN Guiding Principles on Business and Human Rights, key labor rights and principles defined by the International Labor Organization, and the Global Impact Sourcing Coalition's "Impact Sourcing Standard." When it comes to assessing crowdsourcing platforms, there are additional factors that need to be evaluated for their impact on workers, including: the code of conduct, the privacy policy, security policy, terms of service (particularly in terms of how they handle workers' accounts being suspended or terminated), and policy for resolving disputes (such as handling disputes over rejected work, nonpayments, and ratings).

For a full list of relevant tools, refer to Annex 2: International Guidance, Standards, Certifications, and Codes of Conduct for Supply Chains, and Principles and Best Practices for Crowdsourcing Platforms.

Explanation:

The policies that govern service providers and crowdsourcing platforms can significantly impact outcomes for workers by shaping workers' abilities to own/access the information that is provided and generated by them,[35] understand and control how such information is used by the service,[36] navigate systems of management and recognition including the ability to contest an action taken against them, and provide feedback on their experiences.

2. Clarity on Benefits

Criteria:

It is important to assess the benefits and support programs available to workers engaged with a particular service, such as shared benefits like micro-insurance or portable benefits.

Explanation:

Individuals working on crowdsourcing platforms are often categorized as independent contractors and work on short-term contracts. This is also the case for some workers working with managed services. Existing labor laws do not always adequately cover independent contractors which can result in job insecurity, lack of benefits, and lack of access to organizing opportunities and labor unions.[37]

Micro-insurance* and portable benefits* have been explored as potential solutions to ensure workers on crowdsourced services receive the benefits they need. However, it is important to call out that this is an area that needs further investigation, research, and progress. Furthermore, solutions should not be limited to portable benefits. Addressing inconsistent pay, low wages, transaction costs, and lack of benefits are all challenges that need to be addressed in order to make these jobs sustainable for workers.[38] Regulators and policymakers have started exploring challenges around employment statuses for crowdworkers. An outline of emerging law and policy that is applicable to crowdsourcing platforms can be found in Annex 2: National Legislation and Policy.

* For example see: Micro insurance, Next Billion, Insuring the Gig Economy, and Freelancers Union.

* Defined by the Aspen Institute as an arrangement where "benefits are connected to an individual, rather than a particular employer, and so they can be taken from job to job without interruption in coverage or loss of funding." Also see New Economy, New Social Contract

3. Commitment to Transparent Pricing and Base Wage

Criteria:

Reviewing the service providers' pricing methodology is an essential part of assessing their approach to worker well-being. If the pricing and wages for workers are unclear, explicitly asking for clarity over how service providers pay workers can push them to provide more transparency overall. In addition to pricing and wages, obtaining more information from service providers about the workforce itself, including geographic locations, are crucial in order to evaluate if the wages are acceptable. Other relevant practices to ask for clarity around include methods and forms of payment, bonuses, promotions, access to work, regularity of work provided to workers, handling of equipment costs, if wages grow over time, and approach to wage negotiations.[39] Such information can be indicated within the engagement contract or may be established organization-wide through accreditation from bodies like the Living Wage Foundation.[40]

Explanation:

While data enrichment work can be a source of additional income and economic opportunity for some individuals, it is a primary source of income for others. Many pricing and payment models adopted by platforms have resulted in low prices.[41] Increasing transparency over pricing is an essential step towards pushing for better wages for data enrichment professions. Clients are in a unique opportunity to help increase transparency in the industry

by explicitly asking about how wages are set when working with a service provider. Some suggestions for improving pricing practices have included having companies publish data on rates and wages to enable comparative analysis, classifying freelancers as employees, instituting a minimum task rate based on a minimum wage, and developing a wage range that is based on the experience of the independent contractor.[42]

Additional information about how clients can approach pricing is outlined in the section Defining Payment Terms and Pricing.

4. Opportunities for Recognition and Reputation Visibility

Criteria:
Clients should also consider what opportunities are in place for personal and professional development including recognition and career mobility. A few ways to measure service providers' actions in this area include whether they have a portablue reputation system or comparable solution, how they measure qualifications, their process for evaluating and rewarding workers, their approach to promotions, and whether they provide training or educational opportunities for workers to obtain new skills.

Explanation:
While data enrichment work is getting increasingly sophisticated and requires more specialized skills, many have pointed out how this work is often undefined, unrecognized, and underappreciated.[43] This can partly be attributed to the way AI is marketed as a technological advance that can be more efficient than humans. Acknowledging the high labor costs necessary to train AI models runs counter to this narrative.[44] However, a lack of recognition and meaningful career advancement opportunities is problematic for workers.

Workers on crowdsourcing platforms are often dependent on clients to provide ratings and statistics about the type of tasks they have completed. Yet research has noted that ratings on crowdsourcing platforms can be inaccurate and undependable.[45] Some crowdsourcing platforms use "badge" systems to distinguish workers' skills and some have processes in place to promote workers to different levels (leadership, expert, trainer, etc.).[46] Crowdsourcing platforms can also claim ownership over information on the platform including reviews, ratings, and feedback.*

This restricts workers ability to move their work and profiles from one service to another and "locks" them[47] into a particular platform because they would have to start from scratch if they started working on another platform.[48] This puts workers at a disadvantage in terms of their negotiating power. It also puts them in a risky situation in the event their account is closed and they lose the reputation they have developed on a particular platform.[49] In response to this challenge, research has recommended that workers have the ability to export human and machine readable work histories in order to empower workers to continue work relationships, independent of any platform.[50] Recommendations

* For example, the Terms of Service for Freelancer state: "You acknowledge that you transfer copyright of any feedback, reputation or reviews you leave consisting of comments and any rating(s) (e.g. quality, communication etc.) together with any composite rating by us. You acknowledge that such feedback, reputation and reviews belong solely to us, notwithstanding that we permit you to use it on our Website while you remain a User. You must not use, or deal with, such feedback, reputation and reviews in any way inconsistent with our policies as posted on the Website from time to time without our prior written permission."

for strong "portable reputation" systems include being worker-controlled, transparent, reparable with improved work, able to incorporate input and reviews from multiple companies, resistant to bias and prejudice, fair in how they distribute rewards, and equipped with a grievance process.[51] Some features of portable reputation systems that are being researched include incorporating personal references, being publicly hosted, providing profile verification, having decentralised open data standards, and having a centralised data holder.[52]

5. Provision of Work Space and Communication

Criteria:

When engaging with a service, it can be useful for a client to consider what physical and/or virtual workspaces are available to workers and if the workers work out of an ISO certified facility.* Additional considerations include the type of environment being provided for workers at physical facilities such as amount of space per person, lighting, equipment, air-quality, etc. If engaging with a crowdsourcing platform, it can be useful to consider what policies, systems, or forums are in place to allow workers to communicate with other workers, clients, and platform administrators.

* For example, see: https://www.iso.org/standard/68021.html

Explanation:

Research has noted the impact that remote work and disaggregated tasks can have on workers.[53] For crowdsourcing platforms in particular, researchers have pointed out that improvements can be made in facilitating communication between workers, clients, and administrators.[54] As mentioned in the paper, better communication can benefit workers and clients by making sure there are efficient ways to resolve uncertainties, which can in turn impact the project timeline. While some crowdsourcing platforms provide methods for workers to communicate with each other, this is rare enough that a number of organic communities and networks outside of platforms have formed as spaces for workers to share experiences and help each other navigate this work. For example, forums like Turkopticon have emerged to provide workers with a means of evaluating and navigating clients.* Researchers have underscored the need for these types of networks and spaces for workers to build community, particularly when unions are not present.[55]

* Turkopticon is a community and tool to enable workers to rate requesters on factors such as communicativity, generosity, fairness, and promptness.

6. Systems for Diversity

Criteria:

Clients should also assess if there are processes and mechanisms in place to support workforce diversity in terms of skill set, background, and geography. Particularly when engaging with a crowdsourcing platform, it is useful to consider what mechanisms are available to incorporate diversity or a targeted background. For example, analyze which filters are available to select a workforce for a given project. This also requires having an understanding of what diversity means, the type of diversity needed for a particular project, and how diversity is measured.

Explanation:

Given the nature of data enrichment work, where individual perspectives and experiences can lead to various workers interpreting tasks differently, it is important to take diversity or composition of the workforce into consideration.[56] Research has spotlighted the important role played by data quality, diversity, and accuracy in shaping machine learning models.[57] Depending on the type of data that is being curated and the technique being used, there is also potential for bias to be introduced into enriched datasets. For example, research has found that tasks like sentiment analysis or content moderation involve more subjective determinations and are subject to human bias.[58]

Content Policies

Criteria:

For projects that involve working with data or content that is either violent or age-sensitive, clients should consider what policies and support are in place to protect workers.

Explanation:

Some tasks may involve being exposed to graphic or violent content that needs to be viewed, annotated, cleaned, or otherwise handled for extended periods of time. Research has documented the negative impact that extended exposure to such content can have on workers.[59] To address this, researchers have stressed the importance of providing support for workers and adopting policies that take the risks of this exposure into account.[60]

ANNEX 3

Existing Guidance for Human Rights, Workers Rights, Supply Chain, and Platforms

This section outlines key international guidelines and tools related to workers rights, human rights, and ethical sourcing. It also covers emerging legislation around the "platform economy" and best practices for crowdsourcing platforms. The intention of this section is to provide clients with a brief overview of each resource as a starting point to undertake a deeper investigation into emerging standards relevant to ethical data enrichment.

International Guidance

There are a number of international standards for ensuring human rights and worker rights. These include:

ILO Declaration on Fundamental Principles and Rights at Work	As defined by the International Labor Organization, fundamental principles and rights at work include: freedom of association and the right to collective bargaining, elimination of forced or compulsory labor, abolition of child labor, and elimination of discrimination. The ILO has also developed International Labor Standards around basic human rights, occupational safety and health, wages, working time, employment policy and promotion, vocational guidance and training, skills development, specific categories of workers, labor administration and inspection, maternity protection and social security, indigenous and tribal people, and migrant workers.
United Nations Guiding Principles on Business and Human Rights	The UN Guiding Principles present a framework based on states' duties to protect human rights, the corporate responsibility to respect human rights, and states' obligation to provide access to remedy in cases of human rights abuses. The principles outlined here are meant to apply to organizations across different industries, sizes, regions, operational contexts, and ownership structures. The principles refer to the International Bill of Human Rights and the rights set out by the ILO in order to provide guidance to businesses on how to embed human rights due diligence into their processes. In doing so, the principles place a positive obligation on businesses to avoid, prevent, mitigate, and address adverse impacts to human rights. Key recommendations for companies include adopting policy commitments articulating a responsibility to respect human rights, a due diligence process to address their impacts on human rights, and mechanisms to provide remedy for any adverse impacts to human rights.
The United Nations Global Compact	The UN Global Compact articulates 10 principles to guide organizations in responsible business practice in the areas of human rights, labor, environment, and anti-corruption. These principles are grounded in the UDHR, the ILO principles, the Rio Declaration on Environment and Development, and the United Nations Convention Against Corruption.
United Nations Sustainable Development Goals	These outline key goals related to poverty, inequality, climate change and more. Sustainable Development Goal 1 pertains to ending poverty in all its forms everywhere and Goal 8 pertains to decent work and economic growth.

ILO Tripartite Declaration of Principles for Multinational Enterprises and Social Policy	The declaration provides guidance to national and multinational companies on how to develop and implement inclusive, responsible, and sustainable workplace practices which create decent work for all, and enable economic and social progress. The principles focus on employment, training, conditions of work and life, and industrial relations. More specifically, the principles outline the importance of promoting employment opportunities, providing social security, eliminating forced or compulsory labor, abolishing child labor, promoting equal opportunity and treatment, providing secure employment, providing relevant training opportunities, ensuring fair wages and benefits, providing a safe and healthy environment, protecting the freedom to associate and organize, ensuring workers can collectively bargain, and guaranteeing workers have access to remedy, and facilitating processes to settle industrial disputes.
Universal Declaration of Human Rights (UDHR)	The UDHR articulates human rights including right to equality, freedom from discrimination, right to life and liberty, right to personal security, freedom from slavery, freedom from torture and degrading treatment, right to recognition as a person before the law, right to equality, right to remedy by a competent tribunal, freedom from arbitrary arrest and exile, right to a fair public hearing, right to be considered innocent until proven guilty, freedom from interference with privacy, family, home, and correspondence, right to free movement in and out of the country, right to asylum in other countries from persecution, right to a nationality and freedom to change it, right to marriage and family, right to own property, freedom of belief and religion, freedom of opinion and information, right of peaceful assembly and association, right to participate in government and free elections, right to social security, right to desirable work and to join trade unions, right to rest and leisure, right to adequate living standard, right to education, and right to participate in cultural life of community.
International Covenant on Civil and Political Rights (ICCPR)	Adopted in 1966, the ICCPR articulates a commitment for members to respect the civil and political rights of individuals including the right to life, freedom of religion, freedom of speech, freedom of assembly, electoral rights, and right to due process.

Standards, Certifications, and Codes of Conduct for Supply Chains

Ethical supply chain practices have traditionally been articulated and audited against a number of different standards and certifications. These have been developed to address concerns around professionalism, working conditions, ethical business practices, transparency and accountability, due diligence, and legal compliance. Companies have also committed to individual codes of conduct and ethical sourcing frameworks. In some sectors, such as textiles and manufacturing, such practices and standards have evolved to be comprehensive. While these are still developing for data enrichment, there are some examples of standards and frameworks that are potentially relevant to companies involved in data enrichment services:

Impact Sourcing Standard	This model outlines minimum requirements necessary for businesses to ensure employees earn an equitable or living wage while meeting business objectives. The Global Impact Sourcing Coalition is a network of individuals and organizations working to build inclusive global supply chains through the adoption of Impact Sourcing. The Coalition offers a standard and a self-assessment tool that is built around five pillars: commitment to impact sourcing, recruiting and hiring, remuneration and benefits, training and career development, and management systems for impact sourcing.
Social Accountability International SA800	This is a social certification program by Social Accountability International that is based on the UDHR and ILO conventions. This standards covers areas such as child labor, forced labor, health and safety, freedom of association and collective bargaining rights, discrimination, discplinary practices, working hours, remuneration, and management system.
Ethical Labor Sourcing Standard, BES, 6002	This framework for ethical labor sourcing governance standard was developed by the BRE group. The framework covers company structure, management policies, management systems, assurance auditing, HR, immigration, supply chain management, bribery, L&D, forums, and reporting. While the framework can be used by a company to assess their own practices, BRE also provides an Ethical Labor Sourcing verification based on this framework.

Principles and Best Practices for Crowdsourcing Platforms

There are emerging standards and codes of conduct that speak directly to crowdsourcing platform work. Examples of principles and best practices that clients can use to understand best practices when engaging with a crowdsourcing platform include:

Crowdsourcing Code of Conduct	Developed in 2017, the Code of Conduct outlines ground rules towards enabling fair and prosperous cooperation between crowdsourcing companies and crowdworkers that can be adopted by platforms and regulators.
Manifesto for the Gig Economy	Developed by Antonio Aloisi, Valerio De Stefano, and Six Silberman, this manifesto articulates a set of goals for platforms, policymakers, and unions to ensure a "healthy digital transition."
Frankfurt Declaration on Platform Based Work	Developed in 2016, the Declaration articulates seven commitments to govern digital labor platforms which can be adopted by platforms and regulators. These commitments address fair working conditions and worker participation.
Fairwork Foundation Principles for Online Work	The Fairwork Foundation aims to enable voluntary scoring of platforms, facilitate ethical choices by stakeholders in the ecosystem, and improve working conditions for those partaking in the digital platform economy. The framework allows for the evaluation of platforms based on principles of pay, conditions, contracts, management, and representation. It also includes specific principles in these areas that apply to online work and gig work.
Model Rules on Online Platforms	Developed by the European Law Institute, Model Rules on Online Platforms evaluates the relationship between platform operators and users, addresses questions of platform liability, sets minimum requirements for fairness and transparency, considers designs of reputation systems, and explores structures for the right to portability.

These principles and codes are compared in the table below.

Comparative Table of Principles and Best Practice for Crowdsourcing Platforms

PRINCIPLE/CODE	CATEGORISATION	RIGHTS/BENEFITS	WAGES AND PAYMENT	FRAMEWORK FOR WORK
Crowdsourcing Code of Conduct	N/A	N/A	• Commitment to provide fair payment • Advise requesters on how to calculate fair wages by taking into consideration factors like task complexity, qualifications, local wage standards, and timeframe for tasks	• Ensure work offered on a platform is legal and indicate when content might be age specific • Clarify applicable regulations • Offer motivating and good work • Enable respectful interactions • Provide clear tasks and reasonable timeframes • Enable freedom and flexibility of workers, including removing penalties for workers who refuse work
Manifesto for the Gig Economy	• Recognize platform "regulars" as employees • Policymakers to regulate gig economy like a form of casual employment	• Support forms of unionizing and organizing • Ensure high working standards and conditions including protection against discrimination, access to collective bargaining, health and safety measures, a living wage, training opportunities, and ownership over one's own work	• Ensure a living wage • Ensure transparency in payment rules • Clarify minimum levels of payment	• Provide code of conduct that clarifies payment, rating criteria, and transparency of internal processes • Provide a dispute resolutions mechanism • Provide good working standards to all platform contributors • Provide a portable rating system
Frankfurt Declaration on Platform Based Work	• Clarify employment status of platform based workers	• Ensure all non-self-employed workers have the right to organize • Ensure access to social protection	• Ensure at least a minimum wage in their jurisdiction	• Develop dispute resolution mechanisms • Increase transparency in platform work • Ensure legal compliance with national laws and international instruments

Continued on next page

PRINCIPLE/CODE	CATEGORISATION	RIGHTS/BENEFITS	WAGES AND PAYMENT	FRAMEWORK FOR WORK
Model Rules on Online Platforms		• Platforms have a duty to protect users • Address misleading information given by users • Provide reporting mechanisms and redress mechanisms		• Ensure that contract terms are clear, machine readable, and available to platform users at all stages of the engagement—users should be notified of any changes • Information about the parameters determining ranking should be provided to users • Users should be informed if the result of a search query has been influenced by financial or corporate ties between a platform and the supplier • Users should be informed about how and what information is used in reputation systems • Platforms must ensure reviews meet the standards of professional diligence • Reviews should be portable • Facilitate communication between customers and suppliers as needed
Fairwork Foundation Principles for Online Work		• Platforms should have policies in place to protect workers and promote the health and safety of workers • Platforms should provide a process through which workers can express their thoughts and organise	• Workers should earn a decent income relative to their home jurisdiction—this should take into consideration work-related costs and active hours worked • Workers should be paid on time • Workers should be paid for all work completed	• Platform terms of service should be transparent and accessible to workers • The party contracting with the worker should be subject to local laws and identified in the contract • Changes to terms of services should be clearly communicated to workers • Workers should be permitted to seek redress for grievances • Workers should have access to documented due process and mechanisms for appeal • Contracts should be consistent with terms of workers' engagement on the platform • Any use of algorithms should be transparent and result in equitable outcomes • A policy should ensure equity in the management of workers on the platform

National Legislation and Policy

Policymakers, governments, and courts around the world are beginning to take steps to provide regulatory clarity around protections for individuals working as freelancers and on temporary contracts, particularly those working through digital labor platforms. This includes clarity in aspects such as the categorization of workers, wage requirements, eligibility for benefits and entitlements, and acceptable conditions for work. While some of these efforts have been met with appreciation, others have been met with criticism and concern of unintended consequences such as limiting the ability of individuals to work as freelancers if they choose and increased difficulty in finding a job if organizations extend benefits to a larger group of workers.* This demonstrates that best practices are still emerging. A nuanced approach is needed to craft regulation governing the digital economy in order to improve working conditions for workers, while minimizing unintended consequences. At the moment, there appear to me more legislative developments focused on specific types of work in the gig economy, such as through transportation platforms and apps.

The table below summarizes some of the developments and regulatory trends that are emerging globally. Please note that the table is not exhaustive:

* For example, the discourse around AB5 has recognized it as having both positives and negatives. For more information see The Washington Post, "Can California rein in tech's gig platforms? A primer on the bold state law that will try."

Comparative Table of Regulatory Developments for the Gig Economy

LEGISLATION/FEATURE	SCOPE/CATEGORIZATION	RIGHTS	NOTES
EU Rules for Gig Economy 202091	Workers on "atypical contracts" and working +12hrs p/m	• Entitled to receive information about the essential aspects of a job within a week • Entitled to receive compensation if there is late cancellation of work • Access to free mandatory training • Limit of six month probationary period • Ban on "exclusivity clauses" for workers	In April 2020, as part of the EU's social policy, the European Parliament passed Rules to ensure rights and set a standard for working conditions for individuals working on 'atypical' contracts. (European Parliament, Gig Economy: EU law to improve workers rights)
California Gig Economy Law AB592	Categorized as an employee unless the worker: (a) is free from the control and direction of the hiring entity in connection with the performance of the work. (b) performs work that is outside the usual course of the hiring entity's business. (c) is engaged in a similarly but independently established trade, occupation, or business of the same nature.	If categorized as employee, guaranteed minimum wage, workers' compensation if they are injured on the job, unemployment insurance, paid sick leave, and paid family leave.	Coming into effect on January 1st 2020, AB5 expands the definition of what constitutes an employee. For more information see: AB-5 Worker status: employees and independent contractors

Indian Code on Social Security 201993	Expands the definition of the unorganized sector to include gig, platform, contract, migrant, and domestic workers	Social schemes for unorganized workers which can include benefits such as life and disability cover, health and maternity benefits, old age protection, education, housing, PF, employment injury benefit, housing, child care, skilling, funeral assistance, and old age homes.	See: PRS Legislative Research, The Code on Social Security, 2019
Protecting the Right to Organize Act 201994		Give workers more power in work related disputes, penalize companies violating labor law, protect against misclassification of freelance workers and ensure that workers have access to collective bargaining rights.	See: Congress.gov, H.R.2474 - Protecting the Right to Organize Act of 2019

Endnotes

1 Gray, M.L and Suri. S. 2019. Ghost Work: How to Stop Silicon Valley From Building a New Global Underclass. Houghton Mifflin Harcourt. Google-Books-ID: 8AmXDwAAQBAJ

2 Metz, Cade. A.I. Is Learning From Humans. Many Humans. The New York Times. August 16th 2019. Accessed August 12th 2020. https://www.nytimes.com/2019/08/16/technology/ai-humans.html

3 Krig, Scott. Ground Truth Data, Content, Metrics, and Analysis. Computer Vision Metrics. Apress. May 26th 2014. Accessed August 12th 2020. https://link.springer.com/chapter/10.1007/978-1-4302-5930-5_7.

4 Lee, Ivan. Data Labeling for Natural Language Processing. A Comprehensive Guide. datasaur.ai. September 42020. Accessed February 13th 2021. https://medium.com/datasaur/data-labeling-for-natural-language-processing-a-comprehensive-guide-741343fea20e

5 Deloitte. Responsible Supply Chain Tools: Understanding the Market Opportunity. April 2019. Accessed August 12th 2020. https://www2.deloitte.com/content/dam/Deloitte/us/Documents/about-deloitte/us-about-deloitte-humanity-united-responsible- supply-chain-tools.pdf

6 Gadiraju, Ujwal, Yang, Jie, Bozzon, Alessandro. Clarity Is a Worthwhile Quality: On the Role of Task Clarity in Microtask Crowdsourcing. HT'7: Proceedings of the 28th ACM Conference on Hypertext and Social Media. July 2017. Accessed September 5th 2020. https://doi.org/10.1145/3078714.3078715

7 Semuels, Alana. The Internet Is Enabling a New Kind of Poorly Paid Hell. The Atlantic. January 23rd, 2018. Accessed September 5th 2020. https://www.theatlantic.com/business/archive/2018/01/amazon-mechanical-turk/551192/

8 Yin, Ming. Suri, Siddharth, Gray, M.L. Running Out of Time: The Impact and Value of Flexibility in On-Demand Crowdwork. CHI'18: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. April 2018. Accessed September 5th 2020. https://dl.acm.org/doi/abs/10.1145/3173574.3174004

9 https://www.ischool.utexas.edu/~ml/papers/donna-iconf15.pdf

10 https://dl.acm.org/doi/abs/10.5555/3398761.3398923, https://journals.aom.org/doi/10.5465/annals.2018.0174, https://journals.sagepub.com/doi/full/10.1177/0950017018785616

11 Schnitzer, Steffen, Rensing, Christoph. Demands on Task Recommendation in Crowdsourcing Platforms - the Worker's Perspective. CrowdRec. September 19th 2015. Accessed September 5th 2020. https://www.semanticscholar.org/paper/Demands-on-task- recommendation-in-crowdsourcing-Schnitzer-Rensing/e411b1e635698a47a6e82acf2e67718780f0579c

12 Rajpal, Shreya, Goel, Karan, Mausam. POMDP-Based Worker Pool Selection for Crowdsourcing. Proceedings of the 32nd International Conference on Machine Learning. Lille France, 2015, Accessed September 10th 2020. http://www.crowdml.cc/icml2015/papers/CrowdML-Paper19.pdf

13 Daniel, Florian, Kucherbaev, Pavel, Cappiello, Cinzia, Benatalla, Boualem, Allahbakhsh, Mohammad. Quality Control in Crowdsourcing: A Survey of Quality Attributes Assessment Techniques and Assurance Actions. ACM Comput. Surv. Article 7 January 2018. Accessed September 10th 2020. https://arxiv.org/pdf/1801.02546.pdf

14 Royer, Alexandrine. The Urgent Need for Regulating Global Ghost Work. Brookings, Brookings, 9 Feb. 2021, Accessed May 18, 2021 https://www.brookings.edu/techstream/the-urgent-need-for-regulating-global-ghost-work/

15 Pury, Cynthia, Brawley, Alice. Work Experiences on MTurk: Job Satisfaction, Turnover, and Information Sharing. Computers in Human Behavior 54 p. 531-546, September 11th 2015. Accessed September 5th 2020. http://crowdsourcing-class.org/readings/downloads/ethics/mturk-job-satisfaction.pdf

16 Berg, Janine, Furrer, Marianne, Harmon Ellie, Rani, Uma, Silberman M. Six. Income Security in the On-Demand Economy: Findings and Policy Lessons from a Survey of Crowdworkers. International Labour Office. 2016. Accessed September 5th 2020. https://www.ilo.org/wcmsp5/groups/public/---ed_protect/---protrav/---travail/documents/publication/wcms_479693.pdf

17 Caitlin Lustig, Sean Rintel, Liane Scult, and Siddharth Suri. 2020. Stuck in the middle with you: The transaction costs of corporate employees hiring freelancers. Proc. ACM Hum.-Comput. Interact. 4,CSCW1, Article 37 (May 2020), 28 pages. https://doi.org/10.1145/3392842

18 Silberman, M.S, Tomlinson, B, LaPlante, R, Ross. J, Irani. L, Zaldivar, A. Responsible Research with Crowds: Pay Crowdworkers at Least Minimum Wage. Communications of the ACM, March 2018, Vol, 61. No. 3. Accessed September 10th 2020. https://cacm.acm.org/magazines/2018/3/225476-responsible-research-with-crowds/fulltext

19 Caitlin Lustig, Sean Rintel, Liane Scult, and Siddharth Suri. 2020. Stuck in the Middle with You: The Transaction Costs of Corporate Employees Hiring Freelancers. Proc. ACM Hum.-Comput. Interact. 4, CSCW1, Article 37 (May 2020), 28 pages. https://doi.org/10.1145/3392842

20 Berg, Janine, Furrer, Marianne, Harmon Ellie, Rani, Uma, Silberman M. Six. Income Security in the On-Demand Economy: Findings and Policy Lessons From a Survey of Crowdworkers. International Labour Office. 2016. Accessed September 5th 2020. https://www.ilo.org/wcmsp5/groups/public/---ed_protect/---protrav/---travail/documents/publication/wcms_479693.pdf

21 Berg, Janine, Furrer, Marianne, Harmon Ellie, Rani, Uma, Silberman M. Six. Income Security in the On-Demand Economy: Findings and Policy Lessons From a Survey of Crowdworkers. International Labour Office. 2016. Accessed September 5th 2020. https://www.ilo.org/wcmsp5/groups/public/---ed_protect/---protrav/---travail/documents/publication/wcms_479693.pdf

22 M. S. Silberman, B. Tomlinson, R. LaPlante, J. Ross, L. Irani, A. Zaldivar. Responsible Research With Crowds: Pay Crowdworkers at Least Minimum Wage. Communications of the ACM, March 2018, Vol. 61 No. 3. Accessed September 5th 2020. https://cacm.acm.org/magazines/2018/3/225476-responsible-research-with-crowds/fulltext

23 Berg, Janine, Furrer, Marianne, Harmon Ellie, Rani, Uma, Silberman M. Six. Income Security in the On-Demand Economy: Findings and Policy Lessons From a Survey of Crowdworkers. International Labour Office. 2016. Accessed September 5th 2020. https://www.ilo.org/wcmsp5/groups/public/---ed_protect/---protrav/---travail/documents/publication/wcms_479693.pdf

24 Berg, Janine, Furrer, Marianne, Harmon Ellie, Rani, Uma, Silberman M. Six. Income Security in the On-Demand Economy: Findings and Policy Lessons From a Survey of Crowdworkers. International Labour Office. 2016. Accessed September 5th 2020. https://www.ilo.org/wcmsp5/groups/public/---ed_protect/---protrav/---travail/documents/publication/wcms_479693.pdf and Irani, Lilly, Silberman, M. Six. Turkopticon: Interrupting Worker Invisibility in Amazon Mechanical Turk. CHI 2013: Changing Perspectives, Paris, France. Accessed September 10th 2020. http://crowdsourcing-class.org/readings/downloads/ethics/turkopticon.pdf

25 Papoutsaki, Alexandra, Hua Guo, Kakavouli, Danai, Gramazio, Connor, Rasley, Jeff, Xie, Wenting, Wang, Guan, Huang, Jeff. Crowdsourcing from Scratch: A Pragmatic Experiment in Data Collection by Novice Requesters. Proceedings, the Third AAAI Conference on Human Computation and Crowdsourcing. 2015. Accessed September 10th 2020. https://www.aaai.org/ocs/index.php/HCOMP/HCOMP15/paper/viewFile/11582/11436 and Berg, Janine, Furrer, Marianne, Harmon Ellie, Rani, Uma, Silberman M. Six. Income Security in the On-Demand Economy: Findings and Policy Lessons From a Survey of Crowdworkers. International Labour Office. 2016. Accessed September 5th 2020. https://www.ilo.org/wcmsp5/groups/public/---ed_protect/---protrav/---travail/documents/publication/wcms_479693.pdf

26 Berg, Janine, Furrer, Marianne, Harmon Ellie, Rani, Uma, Silberman M. Six. Income Security in the On-Demand Economy: Findings and Policy Lessons From a Survey of Crowdworkers. International Labour Office. 2016. Accessed September 5th 2020. https://www.ilo.org/wcmsp5/groups/public/---ed_protect/---protrav/---travail/documents/publication/wcms_479693.pdf

27 Barbosa, Nata, Chen, Monchu. Rehumanized Crowdsourcing: A Labeling Framework Addressing Bias and Ethics in Machine Learning. In CHI Conference on Human Factors in Computing Systems Proceedings (CHI 2019), May 4–9, 2019, Glasgow, Scotland UK. ACM, New York, NY, USA. https://www.natabarbosa.com/chi_rehumanized_crowdsourcing.pdf

28 https://dsg.tuwien.ac.at/Staff/sd/papers/Zeitschriftenartikel%20-%20Quality%20Control%20SD%202013.pdf

29 https://arxiv.org/pdf/1801.02546.pdf

30 Berg, Janine, Furrer, Marianne, Harmon Ellie, Rani, Uma, Silberman M. Six. Income Security in the On-Demand Economy: Findings and Policy Lessons From a Survey of Crowdworkers. International Labour Office. 2016. Accessed September 5th 2020. https://www.ilo.org/wcmsp5/groups/public/---ed_protect/---protrav/---travail/documents/publication/wcms_479693.pdf and Silberman, M.S ,Tomlinson, B, Laplante, R, Ross, J, Irani, L, Zaldivar, A. Responsible Research with Crowds: Pay Crowds: Pay Crowd Workers At Least Minimum Wage. Communications of the ACM. March 2018. Vol. 61, No. 3. Accessed September 10th 2020. Available at: https://cacm.acm.org/magazines/2018/3/225476-responsible-research-with-crowds/fulltext

31 McInnis, Brian, Cosley, Dan, Nam, Chaebong. Taking a HIT: Designing around Rejection, Mistrust, Risk, and Workers' Experiences in Amazon Mechanical Turk. CHI'16: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. May 2016. Accessed September 5th 2020. https://doi.org/10.1145/2858036.2858539

32 Berg, Janine, Furrer, Marianne, Harmon Ellie, Rani, Uma, Silberman M. Six. Income Security in the On-Demand Economy: Findings and Policy Lessons From a Survey of Crowdworkers. International Labour Office. 2016. Accessed September 5th 2020. https://www.ilo.org/wcmsp5/groups/public/---ed_protect/---protrav/---travail/documents/publication/wcms_479693.pdf

33 Berg, Janine, Furrer, Marianne, Harmon Ellie, Rani, Uma, Silberman M. Six. Income Security in the On-Demand Economy: Findings and Policy Lessons From a Survey of Crowdworkers. International Labour Office. 2016. Accessed September 5th 2020. https://www.ilo.org/wcmsp5/groups/public/---ed_protect/---protrav/---travail/documents/publication/wcms_479693.pdf

34 Gray, M.L and Suri. S. 2019. Ghost Work: How to Stop Silicon Valley From Building a New Global Underclass. Houghton Mifflin Harcourt. Google-Books-ID: 8AmXDwAAQBAJ

35 O'Conner, Sarah. Let Gig Workers Control Their Data Too. Financial Times. April 3rd 2018. Accessed September 5th 2020. https://www.ft.com/content/a72f7e56-3724-11e8-8b98-2f31af407cc8

36 Sergison, Danica. Privacy risks for customers and workers in the gig economy. Privacy News Online. September 2nd 2018. Accessed September 5th 2020. https://www.privateinternetaccess.com/blog/privacy-risks-for-customers-and-workers-in-the-gig-economy/

37 Towards a Fairer Gig Economy. Meatspace Press 2017. Accessed September 20th 2020. https://ora.ox.ac.uk/objects/uuid:de091436-0482-4818-8c87-ff89707f8339/download_file?file_format=pdf&safe_filename=Towards_A_Fairer_Gig_Economy.pdf&type_of_work=Book and Robertson, Pete. How the Gig Economy Creates Job Insecurity. BBC. September 18th 2017. Accessed September 5th 2020. https://www.bbc.com/worklife/article/20170918-how-the-gig-economy-creates-job-insecurity, and Kapoor, Astha, Natarajan, Sarayu. Productivity Vs. Well-Being: The Promise of Tech Mediated Work and Its Implications on Society. Observer Research Foundation. October 5th 2019. Accessed September 5th 2020. https://www.orfonline.org/expert-speak/productivity-vs-well- being-the-promise-of-tech-mediat-

ed-work-and-its-implications-on-society-56962/

38 Berg, Janine, Furrer, Marianne, Harmon Ellie, Rani, Uma, Silberman M. Six. Income Security in the On-Demand Economy: Findings and Policy Lessons from a Survey of Crowdworkers. International Labour Office. 2016. Accessed September 5th 2020. https://www.ilo.org/wcmsp5/groups/public/---ed_protect/---protrav/---travail/documents/publication/wcms_479693.pdf

39 Vakharia, Donna. Lease, Mathew. Beyond Mechanical Turk: An Analysis of Paid Crowd Work Platforms. University of Texas- iConference. 2015. Accessed September 5th 2020. https://www.ischool.utexas.edu/~ml/papers/donna-iconf15.pdf

40 The Living Wage Foundation. Accessed September 5th 2020. https://www.livingwage.org.uk/

41 Berg, Janine, Furrer, Marianne, Harmon Ellie, Rani, Uma, Silberman M. Six. Income Security in the On-Demand Economy: Findings and Policy Lessons from a Survey of Crowdworkers. International Labour Office. 2016. Accessed September 5th 2020. https://www.ilo.org/wcmsp5/groups/public/---ed_protect/---protrav/---travail/documents/publication/wcms_479693.pdf and Metz, Cade. A.I. Is Learning from Humans. Many Humans. New York Times. August 16th 2019. Accessed September 5th 2020. https://www.nytimes.com/2019/08/16/technology/ai-humans.html

42 Caitlin Lustig, Sean Rintel, Liane Scult, and Siddharth Suri. 2020. Stuck in the Middle With You: The Transaction Costs of Corporate Employees Hiring Freelancers. Proc. ACM Hum.-Comput. Interact. 4, CSCW1, Article 37 (May 2020), 28 pages. https://doi.org/10.1145/3392842

43 Whittaker, Meredith. Crawford, Kate. AI Now Report 2018. AI Now Institute. December 2018. Accessed September 5th 2020. https://ainowinstitute.org/AI_Now_2018_Report.pdf

44 Gray, M.L and Suri. S. 2019. Ghost Work: How to Stop Silicon Valley From Building a New Global Underclass. Houghton Mifflin Harcourt. Google-Books-ID: 8AmXDwAAQBAJ

45 Whiting, Mark. Gamage, Dilrukshi. Crowd Guilds: Worker- led Reputation and Feedback on Crowdsourcing Platforms. https://arxiv.org/pdf/1611.01572.pdf

46 Vakharia, Donna. Lease, Mathew. Beyond Mechanical Turk: An Analysis of Paid Crowd Work Platforms. University of Texas- iConference. 2015. Accessed September 5th 2020. https://www.ischool.utexas.edu/~ml/papers/donna-iconf15.pdf

47 https://magazine.seats2meet.com/research-on-platform-based-reputation-scores-contributes-to-an-inclusive-labor-market/

48 Sipp, Kati. Ratings in the Gig Economy Are a Mess. Here's How to Fix Them. Wired. December 27th 2017. Accessed September 5th 2020. https://www.wired.com/story/how-to-fix-ratings-in-the-gig-economy/

49 Robinson, Carrie. Exploring Portable Ratings for Gig Workers. Medium. February 2nd 2020. Accessed September 5th 2020. https://medium.com/doteveryone/exploring-portable-ratings-for-gig-workers-5632fd9b262e

50 Berg, Janine, Furrer, Marianne. Digital Labour Platforms and the Future of Work. International Labour Organization. 2018. Accessed September 5th 2020. https://www.ilo.org/wcmsp5/groups/public/---dgreports/---dcomm/---publ/documents/publication/wcms_645337.pdf

51 Sipp, Kati. Ratings in the Gig Economy Are a Mess. Here's How to Fix Them. Wired. December 27th 2017. Accessed September 5th 2020. https://www.wired.com/story/how-to-fix-ratings-in-the-gig-economy/

52 Robinson, Carrie. Exploring Portable Ratings for Gig Workers. Medium. February 2nd 2020. Accessed September 5th 2020. https://medium.com/doteveryone/exploring-portable-ratings-for-gig-workers-5632fd9b262e

53 Hui, Julie, Cranshaw, Justin, Kotturi, Yasmine, Kulkarni, Chinmay. The Future of Work(places): Creating a Sense of Place for On- Demand Work. CSCW'19: Conference Companion Publication of the 2019 on Computer Supported Cooperative Work and Social Computing. November 2019. Accessed September 5th 2020. https://doi.org/10.1145/3311957.3359432

54 Berg, Janine, Furrer, Marianne, Harmon Ellie, Rani, Uma, Silberman M. Six. Income Security in the On-Demand Economy: Findings and Policy Lessons from a Survey of Crowdworkers. International Labour Office. 2016. Accessed September 5th 2020. https://www.ilo.org/wcmsp5/groups/public/---ed_protect/---protrav/---travail/documents/publication/wcms_479693.pdf

55 Wood, Alex, Lehdonvirta, Vili, Graham, Mark. Workers of the Internet Unite? Online Freelancer Organisation Among Remote Gig Economy Workers in Six Asian and African Countries. New Technology, Work and Employment. Volume, Issue 2. July 10th 2018. Accessed September 5th 2020. https://doi.org/10.1111/ntwe.12112

56 Barbosa, Nata, Chen, Monchu. Rehumanized Crowdsourcing: A Labeling Framework Addressing Bias and Ethics in Machine Learning. In CHI Conference on Human Factors in Computing Systems Proceedings (CHI 2019), May 4–9, 2019, Glasgow, Scotland UK. ACM, New York, NY, USA. https://www.natabarbosa.com/chi_rehumanized_crowdsourcing.pdf

57 Redman, Thomas. If Your Data Is Bad, Your Machine Learning Tools Are Useless. Harvard Business Review. April 2nd 2018. Accessed September 5th 2020. https://hbr.org/2018/04/if-your-data-is-bad-your-machine-learning-tools-are-useless

58 Barbosa, Nata, Chen, Monchu. Rehumanized Crowdsourcing: A Labeling Framework Addressing Bias and Ethics in Machine Learning. In CHI Conference on Human Factors in Computing Systems Proceedings (CHI 2019), May 4–9, 2019, Glasgow, Scotland UK. ACM, New York, NY, USA. https://www.natabarbosa.com/chi_rehumanized_crowdsourcing.pdf

59 Roberts, Sarah T. Content Moderation in the Shadows of Social Media. Yale University Press. June 5th 2019.

60 Berg, Janine, Furrer, Marianne, Harmon Ellie, Rani, Uma, Silberman M. Six. Income Security in the On-Demand Economy: Findings and Policy Lessons from a Survey of Crowdworkers. International Labour Office. 2016. Accessed September 5th 2020. https://www.ilo.org/wcmsp5/

groups/public/---ed_protect/---protrav/---travail/documents/publication/wcms_479693.pdf