

## **Service description for Purchasing Market Survey Document**

The Center for Humans and Machines (CHM from here onwards) needs a **scalable, on demand crowdsourcing service** website. CHM will often use human participants for behavioral data collection, prototyping, experiments, surveys and feedback. The service must be well-suited to take on simple and repetitive tasks in our workflows which need to be handled manually by human volunteers. CHM needs to outsource microtasks that get done quickly. Many experiments conducted by CHM will require groups of participants working together in large groups. The experiment must then be visible to a large pool of participants at the same time so to ensure that enough people show up synchronously on the experiment platform. Most projects will require high volumes of respondents.

The most important factors that CHM prioritizes are listed below.

### *Size of participation pool*

1. Participant pool of around 70.000 or more.
2. Around 10.000 workers at any given time.
3. Accessibility to active users is a strong advantage

### *Scalability*

1. The service needs to allow access to a global, on-demand, 24x7 workforce. It needs to enable businesses and organizations to get work done easily and quickly when they need it – without the difficulty associated with dynamically scaling in-house workforce.
2. The service also needs to provide a messaging service on the platform, that allows sending emails to multiple participants easily
3. Specialization in academic studies is a bonus

### *Flexibility*

1. By using the service, CHM should be able to create and submit virtually any task that can be completed using a computing device connected to the Internet.
2. The service should provide integration with both custom-based templates and applications as well as third party services like Qualtrics or SurveyMonkey.
3. The service should provide the possibility to create separate batches of workers so to divide the workflow according to CHM needs.

### *Speed of recruitment*

1. Low latency between crowd task posting and execution (under 1 hour).
2. Efficient task design through GUI or API, and that does not require third parties involvement.

#### *Heterogeneity of participation pool*

1. Participation pool must include people from a large range of recorded demographics, including age, gender, nationality, education, language and background.
2. Failure to access a heterogeneous population will jeopardize the ecological validity of the results and the generalization of the conclusions.
3. Demographics of participant pool should be transparent and easily accessible
4. A reasonably naive participant pool is a strong advantage

#### *Scope*

1. Flexibility in changing filters so to micro-target different crowdworking populations of interest, based on demographics, geography, past accomplishments, language, age group, quality of previous work etc., at no additional cost
  - a. Ability to pre-screen participants who agreed to do deception studies is a plus
  - b. Ability to blacklist participants e.g. who did not make enough effort in the study or answered randomly.
2. Crowdworking base is expected to be mainly based in English-speaking countries (Canada, US, UK, India, Europe etc).
3. Access to worldwide as well as national population.
4. Reachability and acceptance of service in targeted groups.

#### *Multiple data collection streams*

1. The service should provide multiple streams of data collection and be flexible to the experimental design used.
2. Ability to post tasks with different methods, including surveys, stimuli presentation, tweets, audiovisual material, custom based applications, text, pictures, videos etc.
3. built-in tools for multi-party studies, enabling continuous running of follow up studies on the same sample of participants without having to create a new study and trace them manually.

#### *Ease of integration with JavaScript and custom developed web applications*

1. The service should provide an easy integration with JS based and other applications.

2. Many applications will be custom-built by CHM. Many of them, e.g. Empirica, offer functionalities that are optimized for use with AMT participants and procedures.

### *Efficiency*

1. Many platforms for online data collection often take large cuts from transactions, meaning that experimenters spend a lot and participants earn very little.
2. The prospective service should offer very transparent fee policies.
3. The prospective service should take a small cut compared to other platforms, ensuring that most of what experimenters pay end up compensating participants for their time and effort.
4. Reduction of fixed costs. The overhead and fixed costs associated with hiring and managing a temporary workforce can often be significant. The service should employ a “paying only for what you use” structure to significantly lower costs.

### *Ease of payment of participants*

1. Participants will be paid both for time spent and performance achieved
2. Participation fees should be easily paid via the prospective service without the need to know participants' contact details
3. Bonuses based on performance should equally be easy to assign via GUI or API.
4. Receipts and Payment summaries should be easily accessible and should include a clear overview

### *Easy management of crowdworkers*

1. The service should provide CHM with the possibility to screen, filter and review workers in a simple and intuitive way, at no extra cost.
2. It should offer the possibility to experimenters to assign workers Qualifications to have fine grain control over which Workers are allowed to view and work on the experimenters' tasks.
3. It should offer the possibility to create own custom Qualification Types, to qualify and manage the workforce.
4. Expectations to reference of employers and their accomplishments.

### *Support*

1. Timely support provided by platform staff in case of questions
2. Support team that comes with extensive training and experience in scientific research is a strong plus

### *Ethics and Participant's Safety*

1. Data should be anonymous, Participant's IDs should not reveal personal information
2. Participants privacy should be valued and their identity should be not disclosed

3. Participants should have the right to complain to the experimenter directly (best via private message) to solve misunderstandings and problems fast
4. Participants should be paid according to minimum wage standards, high enough to ensure participants motivation as well as diversity over different socioeconomic groups
5. Similarly, the researcher's privacy and identity should not be disclosed. Easy communication with study participants safely with in-app messaging system is a plus. This helps protect the privacy of participants on the platform, without taking away the ability to communicate with them.

### *Easy workflow*

1. The service must offer a way to effectively manage labor and overhead costs associated with hiring and managing a temporary workforce.
2. It needs to leverage the skills of distributed Workers on a pay-per-task model, to significantly lower costs while achieving results that might not have been possible with just a dedicated team.
3. flexible user interface or direct integration with a simple API

### *Fully developed review system*

1. Apparatus aimed at improving the quality of the experiments offered and the work produced by paid volunteers.
2. Supporting online communities, allowing to increase the visibility of fair experiments and down weight the visibility of unfair ones. At the same time these communities offer direct feedback from the workers to the experimenters. This ensures that over time, the experimenter can improve the quality of their experiments to the benefit of all parties involved.
3. The platform should provide a simple way to reject low quality submissions
  - a. The dispute process should be clear and easily manageable

### *Existing Records of data reliability and quality*

1. As experiments move away from traditional lab settings into online (yet controlled) settings, being able to estimate the reliability of these new procedures is extremely important.
2. A large body of literature is emerging aimed at quantifying the reliability of existing online experimentation platforms.
3. The service must provide evidence to be equivalent to traditional lab based experiments with the added virtues of scalability and scope
4. Data quality should have assessed in peer reviewed academic publications, investigating attention to instructions and naive to common research tasks

### *Comparability with existing literature*

1. A large body of literature in social science, behavioral science, computer science makes use of behavioral experiments with human participants.
2. It is important for new studies to be comparable to the existing literature so to place results on the same scale and build a comprehensive and reliable description of human behavior.
3. The service must be among the most used platforms used by behavioral scientists across countries and disciplines. Using the service means that results can be interpreted in the light of previous research done using the same modalities as well as the same pool of participants.

### *Data Processing and Compliance with Data Protection Laws*

1. Online recruitment platforms that are GDPR compliant will be favored