mWork

Re-imagining the Future of Mobile Work for the Masses

Lea Coligado | Design + Development

Andrea Sy | Managment + Design

Allen Yu | Documentation + User Testing

John Yang-Sammataro | Development + Management

October 11th, 2014 | CS147 HCI + D

Value Proposition

mWork helps complete tasks and alleviate poverty by connecting ***clients*** with micro-tasks that require human intelligence to previously inaccessible pools of mobile device carrying ***workers***, especially in the developing world, who can perform them to supplement their livelihoods in the process.

Team

Our team is comprised of Lea Coligado, Andrea Sy, Allen Yu, and John Yang-Sammataro. Each member contributes across the board on all aspects of the project. Each member also has responsibility for the primary and secondary roles respectively below:

Lea Coligado - Design and Development
Andrea Sy - Management and Design
Allen Yu - Documentation and User Testing
John Yang-Sammataro - Development and Management

Problem and Solution Overview

Poverty and underemployment are two of the biggest global problems in our day and age. One of the starkest examples is what we call the "micro-task gap": On one side, companies and individuals are willing to pay to complete millions of small tasks - such as determining the content of a picture - that still can only be performed well with human intelligence. On the other side, over 25 million people[1] in the United States and over 202 million people[2] around the globe are unemployed and over 3 billion people live on less than $2.50 a day.[3] These people could make multiples of their current income by completing micro-tasks. However, existing solutions such as Amazon Mechanical Turk and Samasource only allow workers with full computers to bridge this gap and pass over the increasing number of global smartphone users in all levels of society.[4] mWork is a proposed solution to fill this micro-task gap. The application enables "clients" to create micro-tasks that need to be completed and receive the results of those human intelligence micro-tasks by paying a small fee. The mobile component of the application will provide an interface that will allow anyone with a smartphone, especially those in the developing world, to complete those tasks in return for small payments that can supplement their earnings.

Contextual Inquiry

Overview and Methodology

Exploring contextual inquiry and task analysis for our mico-work provided the challenge of interviewing two sides of the market for a potential application: **clients** who might have micro-tasks that needed to be completed and **workers** who might complete those tasks. When recruiting subjects, we balanced subjects that we thought might be potential clients and workers with the transportation constraints of living on campus. We researched local labs and companies using micro-tasks for clients. For workers, we found students and staff who might want to supplement their income and who we could interview in person.

[1] Source: http://data.worldbank.org/country/united-states
[2] Source: http://www.theguardian.com/business/2013/jan/22/ilo-unemployment-numbers-rise-2013
[3] Source: https://www.dosomething.org/facts/11-facts-about-global-poverty
[4] Source: http://readwrite.com/2013/05/13/mobile-is-taking-over-the-world

We also conducted Skype call (beyond the three required interviews) to get a better understanding of the lives of potential workers in the developing world.

When interviewing subjects, we used the master apprentice model and asked participants to walk us through their daily work and lives. We did this *in situ* at their place of work when we could. After getting our initial bearings, we guided the conversation towards relevant cases of micro-work in their lives. For those relevant problems and challenges, we delved deeper into how the approached their tasks and challenges and framed the questions in the context of how they would train us to do their work. After completing each session, we would discuss a quick summary of the points we had talked about, make any clarifications, and then used the remaining minutes to answer their questions and hear feedback on proposed changes to their current workflows. We sent chocolates and thank yous to all participants.

Customer Interviewees

Client 1: C. B. - Human Subjects Lab Administrator

Contextual Inquiry Customer
C. B. is a part-time mother, part-time employee in the Stanford Psychology department. She manages the Psychology SONA system, a database that stores information about studies and participants and oversees the execution of the studies. Individuals sign up to participate in SONA and it has both a paid pool and a credit pool. She fits well into our target clients because completing surveys is one human intelligence task we want to learn more about. She was contacted via the Psychology Department Human Studies website and the interview was scheduled in her office cubicle. The environment was quiet and spacious but slightly isolated. Her responsibilities mainly included adding/deleting participants into the system, troubleshooting for participants, managing the Psychology department's social media, and other administrative tasks. She prefers managing social media because it is more creative and gives her more decision-making abilities. Some of the most important skills required are marketing, task management, and communication.

Contextual Inquiry Results
From her descriptions, we identified three main challenges C. B. faced. First, she held the responsibility of keeping her participants happy, and the major factor is spam. The Psychology department sends emails to qualifying participants to notify them of available studies. Maintaining the balance between finding enough participants and not spamming their inboxes is something Cindy has to wrestle with. Second, she finds the SONA user interface difficult to use. One particular example she cited was that deleting a participant deletes all their past studies' information. Instead she can only deactivate a participant when they are no longer on campus or wish to participate. Small details such as this can be easily fixed by SONA. When we further inquired why they do not use another system other than SONA, she informed us that "there was no other option other than SONA." This is an interesting point as this monopoly may be an opportunity for new competitors to enter the market. Third, there is a lot of demand for participants in the summertime when researchers are freer to perform studies. However, there are significantly less students and professors to participate because of vacation absences. This is an interesting phenomenon that can be more closely examined.
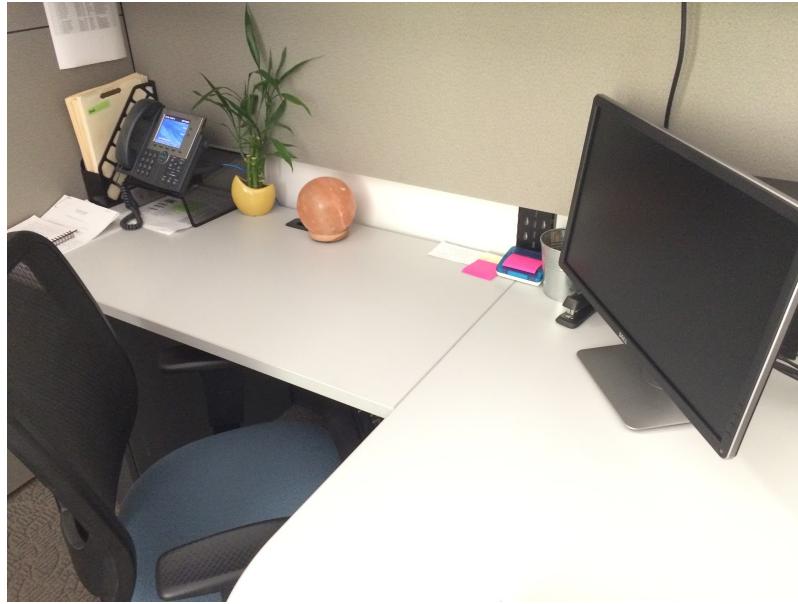
Figure 1. C. B.'s work station

Client 2: A. M. - Social Sciences Lab Manager

Contextual Inquiry Customer

A. M. is the manager of Stanford's Mind, Culture, and Society Lab and oversees approximately 26 assistants. She graduated with a B. A. in Psychology at Tufts. She has her own office and the environment is pleasant to work in. On a daily basis she submits Institutional Review Board (IRB) protocols to get research approved, operates studies in-person and online, maintains lab space, and perform other administrative tasks. The studies usually have a couple hundred participants. We asked her to list the steps in how she would help someone organize a study. First the IRB must approve the studies and the consent forms and this process can be done through an online portal. The researcher would then create a survey on qualtrics (a survey website). They would discuss target demographics and pilot test questions. The study would be uploaded to either the paid pool, the credit pool, or Mechanical Turk.

Contextual Inquiry Results

Since one of our ideas is similar to Mechanical Turk, we inquired more details regarding that area. They would create a pre-test that paid 10 cents for a 1 minute survey regarding demographics. Frequently the survey would be US-only which has proved to be more reliable. Some of the advantages of Mechanical Turk is that the survey population is diverse and the survey is completed quickly. The price is also relatively cheap at around $3-$4 for a 15 minute survey. One disadvantage is the language barrier. Some answers aren't in English and some people who do not speak English will just copy past nonsensical English. This may actually pose more work for researchers since they have to manually parse through the answers and determine which ones are usable. Another interesting point was that the pay cannot be too high or else there would be a lot of scammers who are only there for the money. It seems like there is a certain pay range that yields more accurate answers. One method of preventing people from arbitrarily answering is to throw in attention check questions such as "Click '10' if you read this question" to automatically filter through obvious junk surveys.

One major challenge in her work are manually reading through survey answers for spam / inappropriate results. She would like to have a red flag detection system. A potential improvement she suggested for the SONA system was to link demographic information to participants so they do not have to repeatedly input it.

Figure 2. Interview with A. M.

Client 3: J. C. - Tech Company Trust & Safety Analyst

Contextual Inquiry Customer

J. C. is a Senior Product Specialist in Trust & Safety at LinkedIn. He is responsible for receiving, processing, investigating, and resolving customer disputes, fraud cases, and legal compliance issues. Report and reconciliation of disputes as well as minimizing risk and loss is essential in this position. One of our team members met him during her internship this summer and we believed he would be the perfect person to ask about industrial needs for human intelligence tasks. He worked at eBay prior to LinkedIn and started a start-up, Locket, at Stanford 3 years ago. His day-to-day tasks include meetings with product managers and engineers, working on proposals, and reviewing new products or features that may be risky.

Context Inquiry Results

The biggest challenge he encounters is content that identification that cannot be done by computers. Every social media site needs to identify illegal activities such as human trafficking and child porn and inappropriate content such as nude or explicit photos. We identified these tasks as potential human intelligence tasks that can be outsourced, but the problem is that people may have different standards of nudity and it would be difficult to maintain a level of consistency. Another issue is the legitimacy of the work. Scams and people trying to game the system would damage their system's consistency and reliability.

One interesting idea J. C. suggested for micro tasking was to have really simple tasks on the unlock screen. If it only requires a few seconds it wouldn't be a huge hassle and it may be a convenient way for people to make some extra disposable income. An example would be to show 4 images and ask them which one is of a mountain.

Figure 3. Interview with J. C.

Worker 1: M. E. - Part Time Filipino Cook

Contextual Inquiry Customer

      M. E. is a part-time cook for a family in the Philippines. She is a 40-year-old woman who is the second of 8 children but has no immediate family. She stays with the host family from 10AM – 8PM to cook lunch and dinner 6 times a week. She gets paid 8 hours a day from 10AM – 2PM and 4PM – 8PM. This means that she has downtime in the morning, afternoon, and evenings. She worked as a cook before at a fast food restaurant but didn't enjoy it because of the long hours, unfriendly work environment, and lack of "real" cooking. Her current work environment is much nicer as she get to interact with other people, the family treats her well, and she cooks food that she enjoy. She is an ideal candidate for interviewing as she represents an individual in the developing nations who could perform human intelligence tasks. The pay that may not attract US users may be more attractive to her. We contacted her through family in the Philippines.

Contextual Inquiry Results

      M. E. currently uses her downtime to study, read books, listen to the radio, and use her phone. It is interesting to note that the majority of people in Philippines have more than one phone and SIM cards since they are so cheap there. M. E. uses her phone mostly for texting, calling, Facebook, and YouTube. She likes to call her sisters while cooking. The biggest motivation factor is to make enough money to send back home to her family. She says she would not be opposed to making some extra income from doing tasks on her phone since her schedule is quite flexible. She also speaks English, which makes her an ideal employee.
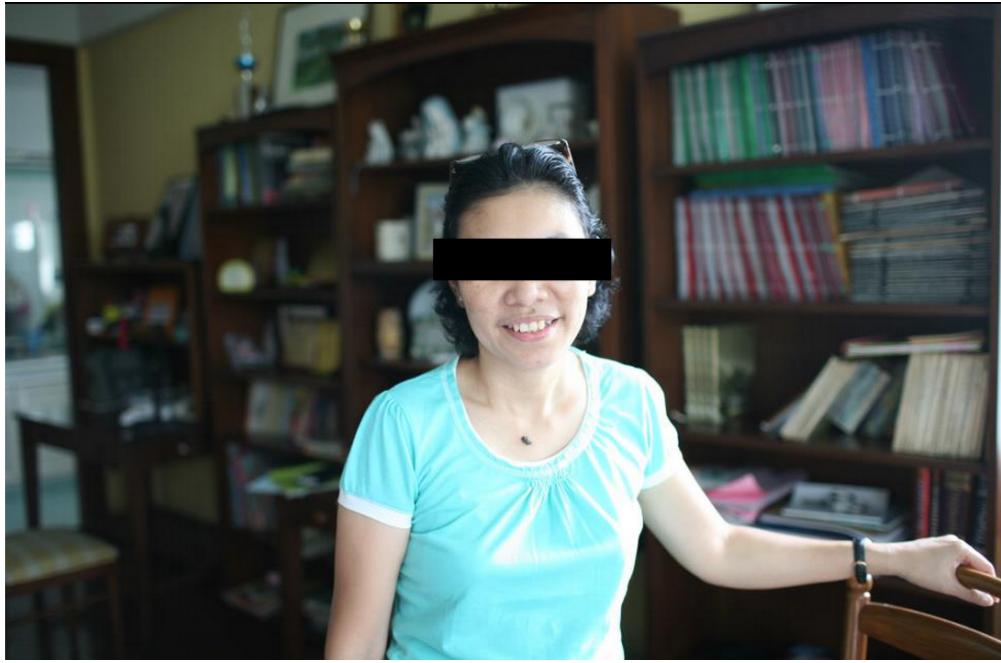
Figure 4. M. E. in the Philippines

Worker 2: D. C. - Part Time Student Worker

Contextual Inquiry Customer

D. C. is a Junior majoring in Economics and a TA for Econ 50. He is a student advisor for the Bing Overseas Studies Program, Kyoto and he serves as a resource for students interested in studying abroad in Kyoto. He also acts as an advisor to BOSP Kyoto and hosts info sessions. He works 2 hours a week and does it because he wants to share the amazing experience he had at Kyoto. The workload usually congregates in the first few weeks of the quarter during application time. He represents a typical college student in US and we discovered him tabling outside of White Plaza.

Contextual Inquiry Results

Money is not the primary goal for D. C.; he works mainly for his enjoyment. David often does online surveys to earn some quick extra cash. He says the advantages are that it doesn't require much thinking and unlike many jobs there are no negative repercussions such as affecting sleep schedule or social life. When asked about their downtime, they suggested an interesting point about killing time while sitting on the toilet. If the tasks are not too dull and the pay is higher, D. C. would be willing to do them in his downtime. A potential direction we could take would be to make the tasks extremely simple and parallel game designs.

Figure 5. Interview with D. C. and his friend

Key Contextual Inquiry Results and Takeaways

There were some clear common themes on the client and worker sides. Workers want higher pay so that it is worth the time, while clients' main constraining factor is budget. An equilibrium would have to be reached that may be determined through natural supply and demand. Outside factors such as minimum pay requirements or subsidies from the platform may improve the situation for both parties.

Another main concern is the reliability of the workers and the results they turn in. It is difficult to prevent people from just clicking or gaming the system to make money. Clients may have to manually weave through the spam and remove unusable answers which may require more work. Potential solutions could be developing algorithms that automatically detect spam and adding in more attention questions. Sometimes workers simply make mistakes or the quality of their work may not be on par to what the client aims for. Perhaps the tasks could be modified so that it is less objective and the goal is geared more towards crowdsourcing than simply manual labor.

One attractive feature of outsourcing human intelligence tasks is the turnaround time. This is a huge issue as most worker populations are usually limited by geographical location or certain demographics. Outsourcing it globally would mean that somebody could do them all 24 hours of the day. This would also provide more diversity if that is a quality desired.

Task Analysis

Who is going to use the system?

Our Contextual Inquiry is comprised of participants from a variety of occupational backgrounds, including A. M. and C. B. from the Stanford Psych Department, J. C. from LinkedIn's Trust & Safety team,

Filipino house cook M. E., and Stanford BOSP Kyoto student advisor D. C. Despite the vocational and demographic diversity of the participants, we see each player relies on, or at least heavily uses, the Internet.

What tasks do they now perform?

Our observed clients perform a variety of tasks for their respective careers, including but not limited to: searching for random and demographically representative survey takers for studies; maintaining the Psychology department's database of studies and participants SONA; determining if images or messages are inappropriate in content (child pornography, human trafficking, etc.); fulfilling directives immediately from upper management; and reviewing security of software products and features.

On the workers' end, tasks include: serving as a personal family cook; occasionally filling out online, optional surveys from school research departments for money; and asking people for referrals to get jobs.

What tasks are desired?

In speaking with our participants, we found that clients desire a number of tasks, respective to their jobs: finding representative survey participants for Psychology studies and obtaining those results; automating the process parsing images and messages for inappropriate content; and identifying when LinkedIn users are taking advantage of invitation requests to send spam.

On the other hand, we found that workers desired to make money through performance of menial tasks and make that money from anywhere.

How are the tasks learned?

Clients typically learned their tasks through professional training. For example, A. M. and C. B. underwent professional training through Stanford's Psychology department to acclimate their experience with maintaining the lab space and coordinating the study-participant database respectively.

Workers learn their tasks in variable ways conditioned on the ease of the task. D. C., for example, required no significant training (only reading simple instructions) to fulfill research surveys due to their straightforwardness, while M. E. learned to prepare food through years of experience cooking.

Where are the tasks performed?

For almost all participants, tasks are performed primarily online. Clients typically work from their workplaces, whether it be the Stanford Psychology lab or LinkedIn headquarters in Mountain View. Users display more variability in where they perform tasks, mostly because they choose where they are when performing them. For example, D. C. completes Psychology study surveys from the comfort of his dorm room, "on the toilet," but holds office hours as a student advisor for Kyoto in the BOSP office. As a family cook, M. E. works entirely in the home of her patrons.

What's the relationship between customer & data?

To generalize, users are givers of data, providing answers to Qualtrics surveys, administering advice to students interested in studying abroad in Kyoto, or proffering cooking service to a family. Clients, on the other hand, have a much more intimate relationship with that data the users provide, directly using it, even relying on it to some end. For example, A. M. may draw conclusions about psychological behavior from the

results of the Psych surveys she administers, and J. C. uses customer feedback to find security loopholes in LinkedIn's software.

What other tools does the customer have?

Clients have multiple resources for obtaining the data they want. Besides mechanicalTurk, A. M. advertises studies within the Psych Department's Credit (Stanford students paid for participation) and Paid (general, non-Stanford public also paid for participation) pools. J. C. relies on fellow employees and the LinkedIn customer-base at large to report product issues.

On the other hand the tools that workers have are diverse and dependent on a number of factors, including their socio-economic position and occupation. For example, Stanford student D. C. has a number of other avenues outside of student advising for making money via other student jobs, even other research surveys, although making money is not a pressing incentive for him in finding work. However M. E.'s options for creating income seem limited to domestic labor, and she relies on her occupation to sustain not only her, but her family's, livelihood.

How do users communicate with each other?

As we have two categories of participants, clients and workers, there arise two types of communication: client-worker and client-client. Clients communicate to workers online, by administering Qualtrics surveys through the Psych department's different pools, by seeking user feedback via LinkedIn online reporting methods, even by directly speaking with participants to guide them through the SONA database. Clients may further communicate with each other to discuss the data they are seeking and set ground rules. For example, A. M. consults with her 26 research assistants to determine which surveys they receive are authentic and maintain the Mind, Culture, and Society lab space in general. Workers in turn communicate to clients via online platforms, as in survey responses or LinkedIn bug reports.

How often are the tasks performed?

Clients perform tasks at a frequency conditioned on their occupation and objective. For example, A. M. aims to distribute roughly 5 to 10 surveys a week through the Psych department, and LinkedIn Trust & Safety team iterates through the process of reviewing specific products multiple times a day.

Similarly, workers perform tasks in quite variably because the significance of the tasks to their lives differs quite largely. Stanford student D. C. responds to surveys from the Psych department at his convenience, roughly once a month, while cook Merlita prepares food daily as part of her livelihood.

What are the time constraints on the tasks?

Again, time constraints are dependent on the task for both clients and users. A. M. prefers to to receive all cumulative data for a given Psych survey within a month of its release. LinkedIn product specialist J. C. determines the amount of time a task takes based on the priority of its "ticket."

As for workers, Stanford student D. C. faces no time constraints; he chooses to take online surveys, and their completion is entirely optional. M. E. alots regular times every day to the preparation of her food and the interstices to her downtime.

What happens when things go wrong?

When things go awry for clients, the repercussions involve manual, human labor. When the Psych department receives scammy survey data, A. M. and her assistants must go through each response to determine validity. When members of SONA input illogical or misplaced data, C. B. must go into the database herself and fix it herself. When bugs arise in LinkedIn features, product specialists must delve into the code or profiles of individual LinkedIn customers and figure out the issue.

On the other hand, worker D. C. faces no negative repercussions, as his participation in online surveys is anonymous and optional. M. E. faces the threat of job loss if she fails to full her daily cooking duties.

Representative Tasks

Representative Task 1: Determining if content (i.e. images, messages) are valid or appropriate
Type: Simple

	Client	User
Frequency	High	Low
Importance	Medium	Low

Rationale:
Many small and large companies need abstract information, in the form of images and messages, parsed into meaningful discrete data (e.g. appropriate or inappropriate). Companies such as LinkedIn need to understand whether images on their platform are "professional" photos and which need to be flagged for "inappropriate" content. Many research labs also need assistance in parsing the large quantities of data that they receive. This task is simple to complete as it can be done through a simple tap or swipe on the phone.

Representative Task 2: Getting hard to access information that are location specific. One specific example is finding the price of mangoes in the Philippines versus finding the price of mangoes in Malaysia for import/export purposes.
Type: Moderate

	Client	User
Frequency	Medium	Low
Importance	High	Low

Rationale:
This task was chosen because it was a general enough task that multiple companies would need this type of work done. Companies who need tasks like this done are those in information driven markets, where an additional piece of information can change company wide decisions. Examples include hedge funds, financial firms, consulting firms, and other corporations who have worldwide operations. This task has

moderate difficulty as the client needs to ensure that their question has enough specificity in the details they are looking for. There is also an added complexity in finding the individuals in the correct location.

Representative Task 3: Matching the user or tasker with the right micro-tasks to help them earn additional disposable income
Type: Complex

	Client	User
Frequency	High	Low
Importance	High	High

Rationale:
        This task is the basis of all outsourced micro-tasks and can be an incredibly complex process. The client needs to provide information on the required skills, experience, location, reliability and other details of their ideal tasker. Some sort of backend processing will need to be done to accurately match the tasker that best fits the criteria. One huge question is how to screen taskers and clients. This task is extremely important for both the client and user side since the client wants to ensure quality work and the user wants to ensure that they are accurately prepared to complete the task.

Application Ideas

Application 1: World mapping application
        Despite advances in satellite and GPS technology, there are still many parts of the world that are unmapped. There is an opportunity to create a game and/or provide income to encourage smart phone users to help map out uncharted territory. After downloading and opening the app, smartphone users can identify locations that have limited topographical information. They then either earn game points or money as they walk around these locations and send topographical and land based information to a main server. This can be used by individuals with a lot of down time, explorers, and thrill seekers to earn extra income. Sample clients would include large companies such as Google and Apple who need mapping technologies.

Significant: Yes, Feasible: Maybe, Interest: Yes

Application 2: Crowdsourced predictions
        This application will crowdsource information gathering from different individuals in order to make predictions on events around the world. We'll use an example to illustrate how it can be applied. For instance, a shipping company wants to know what the Ukraine sentiment is surrounding the Russian conflict in order to predict the stability of the economy. They can provide parameters of the information they need that can answer their question (e.g. number of military personnel in the streets, number of political rallies, negatives vs positive comments from individuals). Users can use a Google Glass app to record situations that fulfill parameter requirements to help "predict" events.

Significant: Yes, Feasible: Maybe, Interest: Yes

Application 3: MTurk mobile
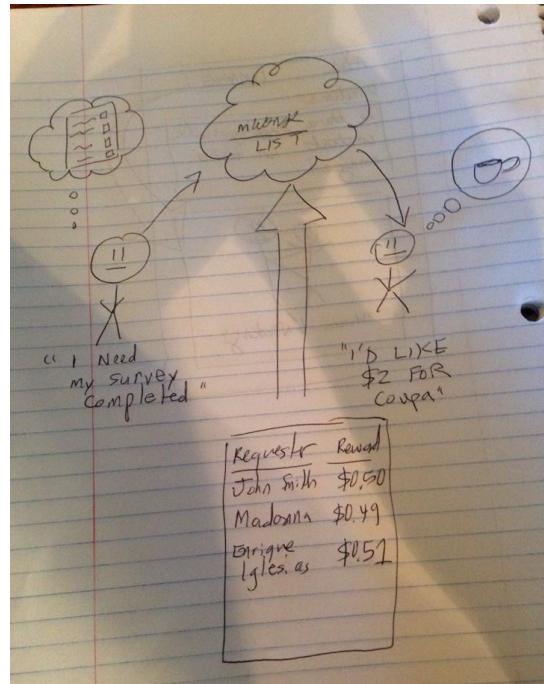
The MTurk mobile app is based off the regular Amazon MTurk platform. However, this mobile app will leverage off existing social media channels to increase user reach. Many individuals in developing nations (aka target users) have smart phones and are consistently on social media channels such as Facebook and Tumblr. This app would display easy to fulfill tasks on these social media channels to decrease the friction between a user's daily life and finding appropriate tasks. This will easily integrate money making opportunities with users' daily Facebook check ins.

***Significant: Yes, Feasible: Yes, Interest: Yes***
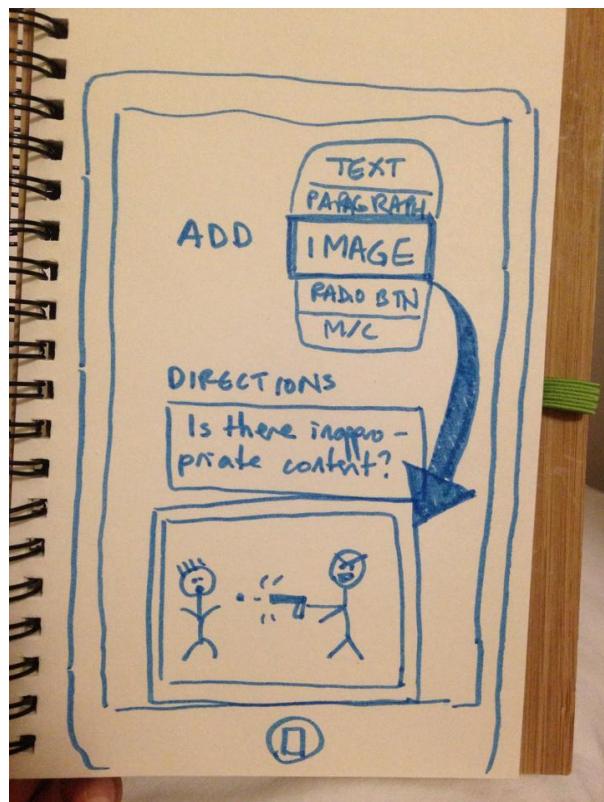
The team decided to pursue this application since it was the most feasible application among the three. It leverages existing social media channels and technologies and has a lower barrier to entry. Most importantly, it has immediate positive impact on both the client and user base. Furthermore, once this platform is developed there is opportunity to build on top this technology to pursue our first two application ideas.

|	World Mapping Application	Crowdsourced Predictions	MTurk Mobile
Significance	Yes	Yes	Yes
Feasibility	Maybe	Maybe	Yes
Interest	Yes	Yes	Yes
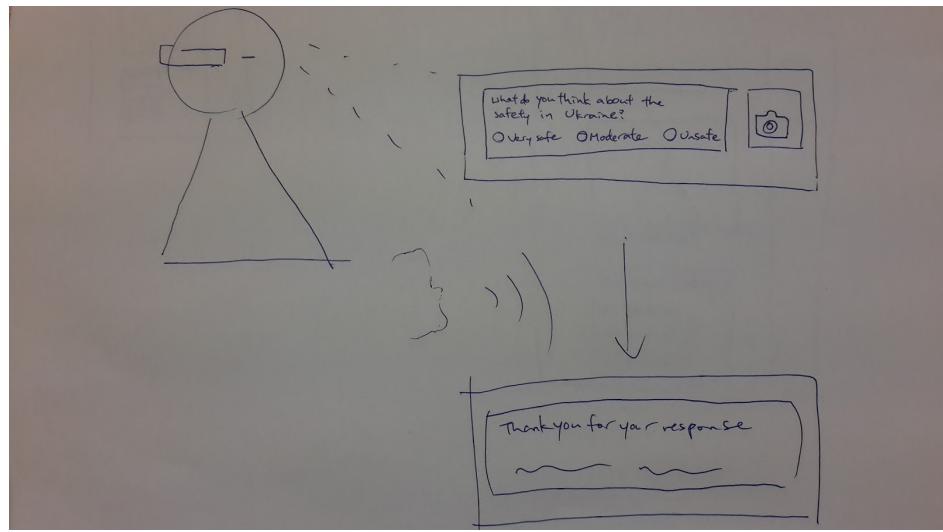
Conceptual Sketches
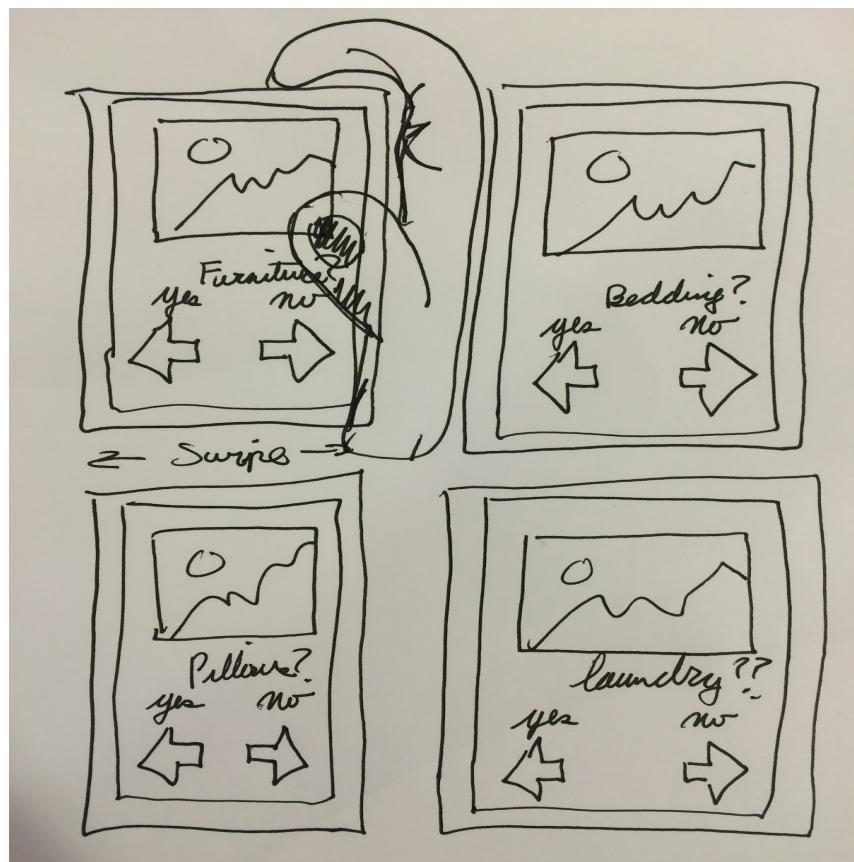
By Andrea Sy

By Lea Coligado

By Allen Yu

By John Yang-Sammataro