Peer-to-Peer Human Computation & "*Help Me Decide*":
Enabling search users to help other users make purchase decisions

Raman Chandrasekar
Microsoft Research
One Microsoft Way
Redmond, WA 98052
+1 (425) 705 2883
ramanc@microsoft.com

Kamal Jain
Microsoft Research
One Microsoft Way
Redmond, WA 98052
+1 (425) 705 9765
kamalj@microsoft.com

ABSTRACT
In this paper, we propose the notion of Peer-to-Peer Human Computation, where users help each other, mediated by a platform that connects people who request help to people who offer help, acts as a repository of acquired information, and learns from users' questions and advice. We compare this idea to related ideas such as human computation and social search.

As an example, we describe a peer-to-peer human computation system named Help Me Decide, designed to assist search users make complex purchase decisions. We list design goals for such a system. We describe the architecture of this system, and detail a prototype of a subset of the Help Me Decide system that was developed and deployed to several hundred users in Microsoft. We discuss the incentive system we developed. We summarize the statistics collected on the usage of the Help Me Decide prototype and the results from a survey of the users. We contrast this system with other related systems. We conclude with some thoughts about next steps.

Keywords
Purchase decisions, shopping, incentives, peer-to-peer human computation, search engines.

1. INTRODUCTION
In today's world, there is a lot of choice in the online retail market, but there are many confused users. When faced with complex purchase decisions, people use search and other channels to find help. Current search engines provide special product or shopping search features to address this need. However, if a person has a sophisticated decision to make, search engines may not always be useful and may sometimes cause frustration [1].

To solve problems like this, we propose an approach we call Peer-to-Peer Human Computation (PPHC). The idea in PPHC is to provide a channel for users to help each other directly, primarily for the end-users' benefit. PPHC is related to but different in many ways from human computation, crowdsourcing and social search.

To make the idea clear, we apply it specifically to purchasing decisions. Consider the following scenario: Katie is a software developer who is considering buying a large screen TV. She has heard of plasma TVs, LCD TVs and recently LED TVs. She has no idea what features to look for, or what brands are well-known. She knows her budget and a rough idea of the size of the screen she wants in her TV.

The system we describe in this paper, named Help Me Decide, is designed to assist search users like Katie make complex purchase decisions. Help Me Decide connects users who are exploring what to buy to people who help them decide what they need. When a user (an 'asker') posts a request for advice to the system, other users with expertise in this area ('helpers') respond with advice and purchase links, based on their perspective and judgment. This is similar to going to one or more retail stores and talking to experienced salespeople, or going to knowledgeable friends -- they understand your actual intent and tailor their responses to your intent, instead of responding at a superficial level to the (possibly incomplete or inconsistent) articulation of your intent. When a purchase is made, all the participants who worked on the request benefit: askers get advice and pointers, helpers get reputation points and other rewards, and advertisers get branding and virtual sales-force benefits. Because of the purchase-based incentives offered, and because of the transparent nature of the system, we expect the system to be an efficient, self-correcting marketplace for ideas and advice.

In Section 2 of this paper, we define the notion of a peer-to-peer human computation and differentiate it from related ideas. We outline design goals for such a system. In Section 3, we outline the architecture of a PPHC system named Help Me Decide designed to help with purchase decisions. We then (in Section 4) describe a prototype of Help Me Decide that was developed and deployed to several hundred users in Microsoft. We also discuss the incentive system we developed, aimed at encouraging people to use the system as they would use it in real life. In Section 5, we summarize the statistics we collected from this deployment and the results from a survey of users. In Section 6, we describe other related systems, with a focus of community question-answering systems. We conclude in Section 7 with some thoughts about next steps.

2. PEER-TO-PEER HUMAN COMPUTATION
Human computation [6, 7] is a relatively new area that studies the utilization of Internet users to perform tasks or provide data towards solving difficult problems for which there are no known efficient computer algorithms. Work on human computation has thus far concentrated on getting vast numbers of users to solve problems or collate data, typically for the benefit of a company or a group requiring data. There can be significant effort involved in

Web Science Conf. 2010, April 26-27, 2010, Raleigh, NC, USA.

defining the problem or task to be solved. Crowdsourcing [4] is similar in that problems are broadcast openly to a group of people and these people (the 'crowd') offer solutions; sometimes the crowd is also utilized to rank the solutions offered. As with human computation problems, crowdsourcing problems are defined by a company or similar entity, and distributed to users. The solutions are aggregated and owned by the problem-defining entity, and typically not shared with the community. In both cases, users may be given some monetary rewards or some other form of recognition.

Social search covers a number of ideas, ranging from the use of page connectivity, tagging or bookmarking sites to rank results to searching over user contributed content and community driven question-answering sites. See [5] for a quick overview of social search. For example Yahoo! Answers (http://answers.yahoo.com) allow users to post questions on a wide range of topics, and to answer other users' questions. Aardvark [3] directs user's questions to people in their own social network who are most likely to provide useful responses. While some systems get users to rank the answers provided, there is often no feedback on the quality of the answers, and no aggregated learning from the information garnered.

In contrast to these approaches, we propose a model we call *Peer-to-Peer Human Computation* (PPHC). We view a decision as a computation which a decision maker (asker) is trying to perform. Askers submit requests for help to a PPHC system and other users (helpers) respond with their advice, based on which the askers take decisions. In the first stage of this approach, the system builds up a repository of the requests, the advice provided and the decisions reached.

Once the repository reaches some critical mass, these data are used to build up a model of purchase decision making. In the second stage, when there are new requests, the user's request is matched against this learned model. If a matching decision is found, the user is informed. If, however, a matching solution is not found, then this step is used to help the asker refine her request. The refined request is then solved in a peer-to-peer manner. The interaction and the eventual decisions, if available, are fed back to update the learned model.

To summarize, in this PPHC approach:

- Users are provided a platform to help each other directly
- The system is primarily for the users' benefit.
- All requests are initiated by users.
- Responses are sought not just from friends but from a larger set of users.
- The system has a focused purpose – to take a decision in the domain of interest. Askers use the responses they get from other users or from learned models to make these decisions.
- The decisions taken by the askers are fed back to the system so that the system learns with every request and response.
- All the knowledge generated is shared with the community of users.

The PPHC system's role is primarily to connect users to advice givers, to encourage and mediate helping, and to keep track of and learn from past requests and responses to help users. While these systems can be stand-alone, they will typically be part of larger systems (e.g. search engines).

2.1 Design Goals

There are many ways to design PPHC mechanisms. To create a successful PPHC, any such mechanism should preferably:

- Provide end-to-end support and advice to help users go from information-seeking to decision.
- Allow for multiple perspectives and answers, so that users get a range of options for decision-making.
- Make the assistance meaningful. Show clearly that the people seeking assistance acted on the advice and benefited by it.
- Make it rewarding and exciting for people to help others. The success of the system is based on active participation by askers and helpers. In a study of network newsgroups, Welser et al [8] point out that only a very small fraction of users (less than 2%) actively reply to newsgroup postings. In order to get more users to participate in our PPHC system, we should incorporate suitable incentives, which may be social or material.
- Build a community around the people who help – this can reinforce and refine helping.
- Make appropriate use of computers and people. In general, the system must use computers and automation to narrow down choices, and utilize users' experience and judgment to go beyond known requests.

3. HELP ME DECIDE: ARCHITECTURE

When search user Katie goes to a search engine and issues a product-oriented search, say for "LCD TVs", she may get several hundreds of results, which can be quite bewildering. This is the point when most of us turn to a friend who knows more about the area, and ask for advice. Help Me Decide (henceforth HMD) enables just this kind of reaching out, without the user having to personally know someone with expertise in the area.

In HMD, we focus on purchase decisions, whether it involves buying products like a camera, PC, cars, clothes, etc., or services like roofing and catering. In this section we outline the information flow in a PPHC oriented towards purchase-decisions, considering askers first and helpers next.

3.1 Information Flow for Askers

If a search query has a purchase-related intent, when the search engine returns product-oriented results, the engine also asks Katie if she wishes to ask for advice from people about her purchase. So Katie uses the simple HMD *Ask for Help* UI to ask for help in choosing a TV, entering a category ("Home electronics/Entertainment") and the request:

> "I'm thinking of buying a 48 inch LCD television, costing $1600 or less. I live in Seattle, WA. What brands should I consider, and what features should I look out for?"

In Stage 1, HMD's *Related Thread Finder* displays related requests, if any. If one of the related requests is directly related to (and answers) the user's request, this request for help is satisfied. However, if no similar request for advice has been made, this

request gets posted on a HMD *Help Others* message board. It is also sent to all users subscribed to receive this class of requests. Optionally, HMD's *Responder Selector* also identifies other users who may be able to respond to this request, and the request is sent to these users as well.

In Stage 2, instead of looking for related requests, HMD's *Model Identifier* classifies Katie's request to see if there is a learned model for this query. If there is a model, the system executes a decision tree (or equivalent) using inputs from Katie, and computes a decision. If the model does not exactly fit the user's requirements, or if there is no such model, HMD treats the request as a new request which is a variant of a known request.

3.2 Information Flow for Helpers

People can subscribe to Help Me Decide if they wish to help and provide advice. These helpers can choose to be alerted about requests that match their profile/category of interest. Alternatively, they can look at HMD UI and search for relevant requests. Additionally, the system may itself forward requests to helpers likely to have useful inputs.

Assume that helper Beth gets the request and decides she has a suggestion for Katie's request. Beth enters it into the HMD *Help Others* UI. To get related information and links for specific products she has in mind, Beth may use HMD's *Product Search*. This component returns product overviews, specifications, and images, and where available, feature comparisons, price comparisons etc. Beth can look through the results from this search and use her expertise to decide which products to recommend or comment upon. This search tool allows for simple search and copy of details to a message which can be sent to the asker and other respondents. The copied text includes a product description, usually a photograph of the product, links for further information and a purchase link with the price.

Instead of looking up specific products, Beth may provide free-form input about systems she has researched or tried out, something like:

> "I have had a Silo 48-inch LCD television for a year and it works great. Got it for $1400 from the Silo Store"

Beth can optionally add a link from which this set may be ordered. Alternatively, Beth can ask for clarifications or refinement of the original set of requirements, and then use that to suggest products. Assume that, in addition to Beth, helpers Charlie, Diana and others may also add their suggestions, comments on others' suggestions, advice and purchase links.

Every time there is a new response, the asker, all helpers for this request and all other subscribed helpers get email updates. They can click on an embedded link to go to HMD thread. On that page, they see the user's request and suggestions from helpers, and can add their own inputs.

We may assume that all the suggestions are made in good faith, and based on information available to users. As in any such system, there is the possibility of spam. Spam may be kept in control by using reputation points or users' comments as a filter, or by using community policing. This is something that will have to be dealt with, along with issues such as trust and authority of advice provided.

3.3 Purchase Decisions in HMD

At some point, Katie looks at all the responses received so far, and chooses one that she likes. Assume that she likes Beth's suggestion. Katie goes to the purchase link provided by Beth and completes a purchase transaction, thus declaring Beth's suggestion the winner. Katie can also rate or comment other responses as well. The *Purchase Module* ensures that the asker (Katie) and the winning helper (Beth) get rewarded with discounts/coupons. Beth's reputation points are boosted and updated. She may also get added to a product previewers list, from which manufacturers may choose people to preview new products.

If Katie wishes to complete the purchase transaction in a retail store (not online), she gets a coupon to a nearby store of her choice. Once the coupon is redeemed offline, all relevant information is updated.

As this stage, the thread with the Katie's request and the responses is annotated with the details of the Katie's purchase, so that all users who view the thread can see that the system works.

Other helpers who make useful suggestion either recommending buying or avoiding a product on reasonable grounds gets coupons and/or reputation points.

The purchase information and the decision reached are used by the *Model Update Module* to update the user profiles and the learning model. This thread with the request and the suggestions are archived. However, ratings and reviews can be modified at any time. For example, if Katie buys the system Beth suggests, and has some comments on it or on Beth's advice, she may add these to the discussion at a later date.

3.4 Advantages of this model

There are several advantages to the model as outlined above.

Search users get advice and incentives: Katie and users like her who request help get peer-advice in an area in which they are not expert, and get a monetary incentive to ask for and accept help within the search engine eco-system. The incentives in the system will grow not only askers but also loyal ('sticky') helpers, who will help users while growing their reputation.

Users get a broad range of suggestions: For users, Help Me Decide is better than going to any single store. When you go to a retail store, the sales people do not tell you about deals in other stores. But here, since any informed person can be a helper, buyers get a spectrum of suggestions and prices, including from smaller stores, making their purchase decision more informed. Choice is good if it is clear and substantiated with reasons.

Helpers get reputation points and natural incentives: If Katie chooses Beth's suggestion, Beth gets a reward for her expertise and willingness to help, again within the search engine's eco-system. In fact every helper who provides sensible comments/suggestions gets a boost in their reputation for helpful advice. In addition, they may get a token reward as well.

There are many natural incentives for helpers that could motivate them to compete and provide the best advice they can. For example, for helpers in the area of cool gadgets, companies may make new or prototype gadgets available to proven gadget geek helpers. Retailers may offer jobs to a qualified gadget geek helpers based on their reputation and responses. Manufacturers may hire qualified helpers to evangelize the benefits of their products.

Retailers and manufacturers could offer performance based incentives to such gadget geeks, instead of employing them. Any such incentive has to be transparent, so that there is no issue of bias.

Manufacturers/retailers get a virtual sales-force: Now every internet user can act as a virtual salesperson for any company, vastly extending their reach. This has a tremendous significance for retailers. The next section (Section 3.5) details how Help Me Decide can use responses directly from advertisers/retailers. In addition, the entire conversation generates brand awareness, targeted advertising value and potentially sales (if coupons are provided) for all the companies mentioned in the original request and in the responses. Because there is more value generated, we can get the positively cited companies to pay for the rewards in the system; the system also provides for negatively-cited companies to get customer feedback and information about pain points.

Help Me Decide lets helpers feel the vicarious thrill of spending other's people money for them – and that is entertainment by itself! This may motivate more sticky users. In addition, Help Me Decide has network effects, and so as it grows, its benefits grow.

Lastly, useful data can be garnered. Because the incentive system in Help Me Decide depends on actual purchases, the advice that helpers give and the purchase data we get are both likely to be clean. We can use this data to model the decision-making process.

3.5 Companies Providing Advice

When the Help Me Decide service gets to scale, companies may participate through an assisted intelligence agent. Their assisted intelligence agent would consist of two stages, an automated filtering stage and a human stage.

When a user posts a request, it is forwarded to all interested helpers. Assume one of the helpers is a company. The company's filtering stage analyzes the request to see if it is relevant to their area. If it is, then the company may send the request to a salesperson representing the company, who then responds to the request. The company response could also have additional details embedded, e.g., an IM link, a phone number, recommended place to buy the suggested product, any promotional deals etc.

This situation is not very different from today's marketplaces. When you go to a retail store, a salesperson provides this service. When the marketplace is online, this can be done even more efficiently. For example, several retail websites have a "May I help you?" link to initiate a chat with a real human; these human operators help the user with information and clarifications. In our scenario, the company is responding to targeted requests from users ready to purchase, and it is sensible to assist these users.

When companies provide advice and suggestions, bias is a distinct possibility. But since everything in the Help Me Decide marketplace is public, the community can easily see if the company is misrepresenting its goods and services .In a variation of this idea, the company may create sales incentives for (volunteer) helpers. To avoid any question of bias, we would expect these paid helpers to declare this in their profile.

4. PILOT STUDY

A version of the Help Me Decide system was developed and deployed to a small group of users within Microsoft. We ran this study from Jan 8th, 2010 through Jan 25th, 2010. Users on a few mailing lists were invited to try out the system to help them ask for and offer help on purchase decisions. Users had to be logged on to the company's intranet to use it, and that meant that home usage was inhibited. Users joined the study from all over the world, not just from technical groups but also from sales, support and other groups.

In this version, our goals were to see if users would ask questions, and if others would respond to these requests. We want to explore some incentives. We also wanted to find out what users would be interested in, and how they phrased their queries, and how responses were constructed. We did not ask for or use user's identities (all users were anonymized), or their social graphs. This version covered only Stage 1 and did not include a learning component. Users could join a mailing list to get email alerts about requests as they were submitted.

4.1 Incentives Provided

To motivate usage resembling real-life use, we awarded points to users for asking, helping and showing intent to purchase and even for signing in the first time as follows:

Signing up: 200 points to each user for signing up to use HMD.

First request: 100 points for the first request the asker posted on the site. We requested askers to post meaningful questions about purchases they were interested in making.

Answering other's requests: For each response, 50 points to the helper for answering someone's request, and 10 points for the corresponding asker. This was designed to avoid people asking meaningless questions with no answer, while encouraging people to answer others' requests.

Purchase reward: When a helper suggested a purchase, it usually came with a dollar value of the item and a "I would purchase this" link. If the asker clicked on this link, points based on the dollar value of the item (1 point for each dollar, subject to a maximum of 1000 points) were distributed in the ratio of 60:40 between the asker and the winning helper. If another user viewed someone else's question and found a relevant answer, they could record their purchase intent, and get the asker's quota of points.

Users accrued points throughout the study. The system had a leaderboard displaying the anonymous ids of the top 10 users in terms of points. At the end of the study, we had a raffle drawing where each user got chances proportional to the points they accrued. We randomly drew five names, and each of these users got a $100 gift card.

4.2 Help Me Decide details

The HMD prototype that was released to users was developed using ASP.Net and C# on an IIS server, with a SQL backend. Users went to an intranet URL to use the system. All visits to the site were recorded.

Once the user goes to the HMD site, she has the option of asking for help, looking at her own requests or helping others.

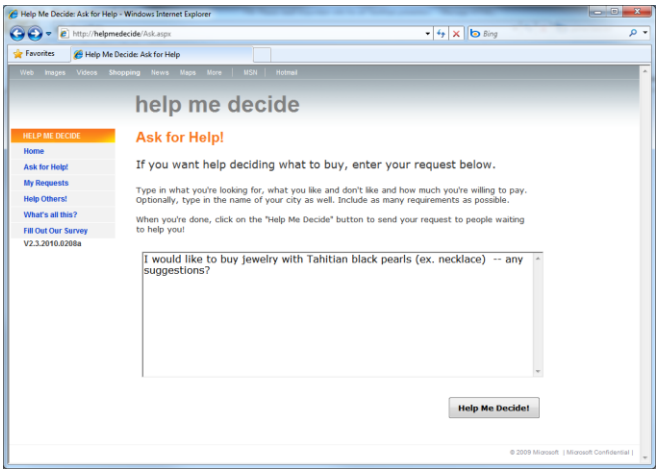

Figure 1. The "Ask for Help" page, with a sample request

The "Ask for Help" screen is shown in Figure 1. This is a real user question, reading "I would like to buy jewelry with Tahitian black pearls (ex. necklace) – any suggestions?" When the asker submitted this, this request was sent to all users on a local opt-in mailing list. (In principle we could have filtered it to users corresponding to a profile or category.) Other users preferred to view requests on the HMD UI.

Figure 2 shows some requests from users extracted from the "Help Others" page. Each row on that page shows the time of the request, the number of people who viewed the request details and the number who responded to the request. Users had the option of searching by keyword for requests of interest to them; they could also search for requests with no responses as yet. Unanswered queries are not deleted from the system at any point. When users responded to a request, they had the option of searching for products from the Bing Shopping site.

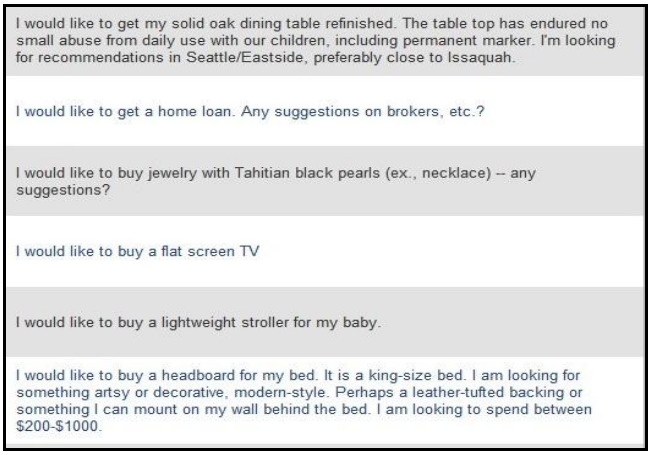

Figure 2. Sample Requests on the "Help Others" page

Figure 3 shows the Tahitian pearl request with a portion of the first response. There were 3 responses to this request, two of them with details of specific items that the user could buy. In the third response, the user has provided her subjective opinion on a place to buy such jewelry, along with reasons why. The company name has been anonymized in this response.

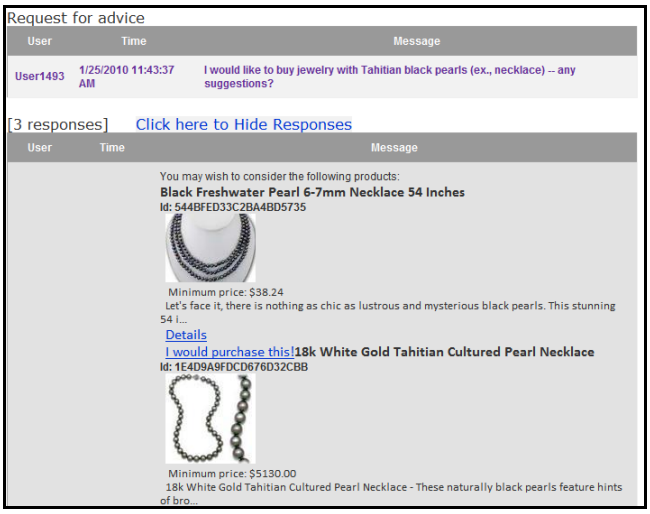

Figure 3. Help Me Decide thread with a request and part of the first response to the request

The third response read as follows:

User1505 1/25/2010 3:01:42 PM

*My husband gave me a lovely necklace of Tahitian pearls a few years back, and matching earrings. He worked with ABC Company to get the design and size of pearls he had in mind, for a reasonable price. I *love* the necklace, the design is lovely and the pearls are gorgeous.*

ABC Company has some lovely pieces, but they are great about working with you if you're looking for something different than what you see in the display cases.

5. EVALUATION

We evaluated the system based on the actual usage. We also conducted a survey of users. We report results from both of these in this section.

5.1 Usage Statistics and Analysis

In the few weeks that we deployed the prototype, we had 1498 unique visitors to the site. We had 387 requests for help, and 702 responses, after removing test requests. A total of 3578 threads (a request with associated responses) were viewed by users. 44 queries had zero responses. A few of these zero response queries were duplicate entries.

We observed that people did ask for help and others responded. Because these were internal company users, spam and fraud were not issues we had to deal with.

Rapidity of responses: We did not design the system for rapid responses by using, say, Instant Messenger to alert users. Even so, the response time is quite reasonable, as shown in Figure 4. Of the 343 requests with at least one response, there were responses for about 9% of them in less than 5 minutes, 28% in less than 30 minutes, 39% in less than an hour and 74% in one day. This time to respond may be slightly high because home use was inhibited by the requirement to use the system while logged in within the intranet. The median response time was about 170 minutes. If (as in Yahoo! Answers) we deleted old, unanswered queries, this duration would be smaller. With more users, the response time is likely to go down as well.

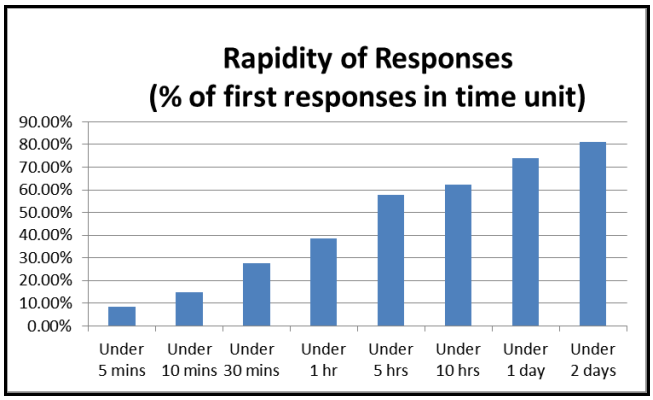

Figure 4. Rapidity of first responses to requests

Length of queries and responses: Compared to search engine queries, where average query length is between 2 and 3 words, the average length of HMD requests is 26.2 words (146 characters). The average length of responses in HMD is 70.6 words (434.7 characters).

There are some standard phrases that are added to the response (e.g. "You may wish to consider the following products:") when the product search and cut and paste functions are used. However in each of these responses, we add less than 20 words. In any case, only 31.3% of the responses use product search, and this would not contribute too much to the length of the responses.

Overall, it is clear that both askers and helpers are trying to provide a fair amount of detail

Categories of Requests: We did a manual categorization of the first 324 requests (see Table 1) and found about 61% of them relate to electronics. We see at least two reasons for this. Because the users are from a software company, their interests are likely to be biased towards electronics and related items. In addition, electronics is a fast-changing field, and users are more likely to have questions about electronics than, say, clothing.

We had imagined that HMD users would ask only about products. Surprisingly, about 5% of the requests related to products. This figure, along with the survey finding (see next section) showing user interest in information on services, suggests a need for features around services.

5.2 Survey and Survey Results

We developed and sent out a survey to all HMD users. Ninety-two users completed the survey.

The top level interpretation of the results seems to be that users have questions, and there are experts all around, ready to give advice. People seem to be using HMD at the beginning of their shopping decision process, rather than at the end. They find it useful to get information on issues, what to buy and also where to buy from.

Of the 92 respondents, 77 had used HMD. Of them 51 had asked for help and they found 27 responses were useful, based on which 12 purchases were made. 44 of the 77 users responded to requests for help, and 54 were willing and 17 may be willing to help. Interestingly, 43 users benefited from someone else's request.

Of the 51 askers and 44 helpers, 18 got to know about features to look out for, 24 got good advice about *what* to buy. 16 got good advice about *where* to buy from, and 39 felt good giving advice.

Category	Number of requests	% requests
Electronics	199	61.4%
Household (incl. cars)	43	13.3%
Services	16	4.9%
Furniture	11	3.4%
Clothing	9	2.8%
Housing	7	2.2%
Accessory	5	1.5%
Kitchen	2	0.6%
Appliance	2	0.6%
Test	2	0.6%
Others	28	8.6%
Total	**324**	**100.0%**

Table 1. Categories of Requests

65 people thought HMD would help shoppers new to certain categories.

We asked the 54 people willing to give advice what reward they expect. 13% wanted to be considered an 'expert', about 20% wanted a small payment per response, and 18.5% would like to have priority in buying early units of gadgets. 13% were happy with decals calling them expert on the site, and 18.5% did not care what the rewards were. 16.7% were not interested in any rewards.

There was considerable (71%) interest shown in getting information on services (in addition to products). In addition, 58% thought it was fine to get advice from companies and 54% thought it was OK to get relevant advertisements. Only 33% prefer to get advice only from friends. 78% said they would "prefer advice from unknown experts who seem to be knowledgeable or friends (who may/may not know the area)", and 19.5% were not sure.

On the whole, Help Me Decide seems to have been of use. The survey gave us valuable insights into features that would be useful to add.

6. RELATED SYSTEMS

At this point, it is useful to discuss related systems. These include purchase helper sites and community question answering sites. In this section, we list a few such sites, and discuss how HMD differs from them.

6.1 Purchase Helpers

In certain product areas, there are programs to help people decide what to buy. For example, About.com has (at http://compreviews.about.com/od/buyers/l/blcompqstart.htm) a program to help people buy PCs and laptops. Some electronic retailers and manufacturers have similar sites. These programs ask a series of questions and then propose a solution based on the user's inputs. These systems typically do not have a feedback path to improve their suggestions. In these systems (and others that have a faceted search option to help you refine your purchase decisions), it is also difficult to discern cost-benefit tradeoffs when choosing a

value for a facet. HMD gets around these problems by using human judgment to provide the flexibility.

General review and opinion sites such as *Epinions* (http://epinions.com) and *Amazon.com* as well as domain-specific review sites (such as http://dpReview.com for digital photography) are a great source of information and opinions. But they are more for people who know a lot about the domain, and more useful if you know exactly what product you're looking for, down to the model number. In our scenario, the user typically does not have this expertise. However, these sites may be useful references for HMD helpers.

Engines such as http://hunch.com seek to help people to decide, and improves itself with use, using machine learning methods. Hunch poses a number of questions about the purchase, and proposes a solution. The system is good in that it recognizes some mutually exclusive constraints. In practice, some of the questions posed by hunch.com seem unrelated to the purchase (why would a political stance affect a choice of cameras, for instance), and that reduces user confidence in the system. The goal in Stage 2 of HMD would be to avoid these pitfalls.

6.2 Community Question-Answering Systems

HMD is similar in some ways to Community Question-Answering (CQA) systems like *Yahoo! Answers* (http://answers.yahoo.com) or *Aardvark* (http://vark.com). In all these systems, users ask questions in unrestricted natural language. Responses are got from users who choose to help out, and hence it is important to provide a good user experience; this will motivate users to visit the site often to ask and to answer questions.

Aardvark focuses on finding the right person to ask rather than the right document. Aardvark depends on a social graph and several good ideas to identify people with the requisite expertise who are likely to answer. Questions in Aardvark are not restricted to any particular domain. In [3], the authors say "Real-time responses from socially proximal responders tend to elicit (and work well for) highly contextualized and subjective queries". Further, they say "The goal of the (connectedness) scoring is to optimize the degree to which the asker and the answerer feel kinship and trust…" However, several of their examples are of factual queries or helpdesk queries, for which they need neither social proximity from responders nor (arguably) real-time responses. Friends may not know about specific products or services, even if they know what color suits you. It may be better to send requests outside your social graph, as HMD does, to get a broader set of responses. Finally there are classes of personal questions (as reported in [5]) in areas such as health, finance, religion and dating which people do not feel comfortable asking their social networks about.

Yahoo! Answers is an open forum where anyone can answer questions. Since questions are not restricted to any domain, there are requests for recommendations mixed with discussion-provoking questions and factual questions. As Gyongi et al [2] point out, "Yahoo! Answers was not designed with discussions in mind… Users cannot answer their own questions, thus cannot participate in a discussion. Similarly, a user may only post at most one answer to a question."

The inability to have a threaded conversation is a big limitation. HMD allows users (askers or helpers) to keep adding to their thread to clarify, add details or ask for clarifications. Purchasing consists of multiple decisions, e.g., what to buy and from where to buy. There is significant scope for expert helpers in HMD to collaborate with each other. Collaborations bring with them spam resistance, and also strengthen the community.

While HMD allows for threaded conversations, it is focused on purchase decisions and not on discussions. Because of this focus, HMD can have performance verification and incentives based on the very tangible feedback of purchases made. Verifiable performance measures improve any such system.

Further, Gyongi et al assert that a large fraction of best answers (answers judged as such by people asking questions) "are of questionable value … Most answerers are not domain experts … but rather individuals with diverse interests…"

It is not clear if we need responders to be experts. In one question we saw in HMD about where to get cheap HDMI cables, four different users recommended the site http://monoprice.com for all cables. One user said "Universal truth: cables = monoprice.com" and another said "Go to monoprice.com. They ship fast and are very inexpensive. Cable quality is also good." In this situation, it matters little if the responders are experts – the fact that they all agree should boost our confidence in the answer; this is the wisdom of the crowds in action.

Gyongi et al [2] suggest "…, Ideally, [a question answering system, such as Yahoo! Answers] should facilitate users finding answers to their information needs and experts providing answers." With the HMD platform and the product search feature built into it, HMD provides for both of these.

7. CONCLUSION

In this paper, we proposed a model called Peer-to-Peer Human Computation (PPHC), and defined its characteristics. We listed our design goals for such systems. We described the architecture of Help Me Decide, a PPHC system to help people with making purchase decisions.

We implemented a prototype of Help Me Decide, and deployed it to a few hundred people within Microsoft. We described the version we deployed and analyzed the usage of the system in a pilot experiment. The system we deployed sticks closely to our design goals. It helps users go from request to decision and purchase, allowing multiple people to provide responses. When people use advice given on the system, all the participants in the advice-giving benefit, and are seen to benefit. A system of incentives makes it rewarding to help others. Because of the purchase-based incentives offered, and because of the transparent nature of the system, we expect the system to be an efficient, self-correcting marketplace for ideas and advice.

We conducted a survey of Help Me Decide users. Results from the survey are included in this paper.

We also provided an overview of related systems, with a focus on community question-answering systems, along with brief differentiating comments to show how HMD circumvents problems of related systems.

Our deployment has shown that this is an important area of work, that Help Me decide is a useful system, and that people are willing

to ask for and offer help. Our next steps are to deploy the system to a larger group of (internet) users, and to implement Stage 2 of PPHC, incorporating learning.

To prove that the PPHC approach is relevant to decision making in general, we hope to create systems that will help people with other complex decisions such as planning travel, or choosing colleges.

8. ACKNOWLEDGMENTS

Our thanks goes to all the users of Help Me Decide who gave us valuable data and feedback with their use of the system. We thank our anonymous reviewers for their useful comments. We also thank Ryen White and Matt Richardson for their help and comments at various stages on the Help Me Decide system and this paper.

9. REFERENCES

[1] J. Battelle. Search Frustration: It's Still Hit Or Miss On Complex Decisions. John Batelle's Searchblog, Sep 4th, 2009, http://battellemedia.com/archives/004999.php.

[2] Z. Gyongyi, G. Koutrika, J. Pedersen, and H. Garcia-Molina. Questioning Yahoo! Answers. In Proceedings of the First WWW Workshop on Question Answering on the Web, 2008.

[3] D. Horowitz, S. D. Kamvar. The Anatomy of a Large-Scale *Social* Search Engine. In Proceedings of WWW, 2010.

[4] J. Howe. *Crowdsourcing: Why the power of the crowd is driving the future of business.* Three Rivers Press, NY, USA, 2008.

[5] M. R. Morris, J. Teevan, and K. Panovich. What do people ask their social networks, and why? A survey study of status message Q&A behavior. In Proceedings of CHI, 2010.

[6] L. von Ahn and L. Dabbish. Labeling images with a computer game. In CHI '04, pages 319-326, New York, NY, USA, 2004. ACM.

[7] L. von Ahn and L. Dabbish. Designing games with a purpose. Communications of the ACM, 51(8):58-67, 2008.

[8] H. Welser, E. Gleave, D. Fisher, M.A. Smith. Visualizing the Signatures of Social Roles in Online Discussion Groups, Jnl of Social Structure, Vol. 8.