Cloud and Crowd Powered Personal Knowledge Management

Yefeng LIU

Major in Computer Science and Engineering
Research on Distributed Systems

Graduate School of Fundamental Science and Engineering
Waseda University

A thesis submitted for the degree of

Doctor of Engineering in Computer Science

July 2013

To my grandfather.

He taught me what is important

in this noisy world.

Acknowledgements

Many people helped enormously in the developing of this PhD thesis.

My supervisor, Prof. Tatsuo Nakajima, offered me the opportunity to pursue my passion, guided me into the research world, supported me with priceless advises and encourages. This thesis would not have been possible without him and without the freedom and inspiration he has given me over the past three years. I am a extremely lucky postgraduate student to have a supervisor like him.

I would like to give special thanks to Prof. Todorka Alexandrova, at Waseda University, and Vili Lehdonvirta, at Oxford Internet Institute for their extensive and unconditional advice, comments, and support on my research projects.

I would like to thank my friends and colleagues at Diggerlab inc., Xiaoguang Han, Gang Chen, Shengyu Huang, Qidi Zheng, Jida Zhou, Ying Jin, and others. They helped me understood the meaning of "*life is about the people you meet, and the things you create with them*".

I would also like to thank my colleagues Khalid Mahmood, Wei Fan, at Tokyo Institute of Technology, Tetsuo Yamabe, Hiroaki Kimura, Ning Li, Akihito Yoshii, Ping-hui Lin, Han Chen, Rui Tao, Connsynn Chye, Mizuki Sakamoto, Francisco Lepe, and all other members at Distributed and Ubiquitous Computing Lab., Waseda University, who have provided an immense amount of help in both research and daily life. I have to admit I would not have been possible to survive in Tokyo with my broken Japanese without them. I would also like to thank my mentors at Microsoft, Darren Edge, Koji Yatani, and at HIIT, Herkko Hietanen.

Lastly, my wife, Jing Wen, has provided both emotional and intellectual support without which this thesis would not have been possible.

I was supported by a Japan Society for the Promotion of Science's Global-COE Fellowship, a Microsoft Research Graduate Fellowship, and a Japanese Government MEXT Scholarship. You all make graduate school possible — thank you.

Abstract

Personal knowledge management is a collection of processes that a person uses to seek, share, and create knowledge in her daily life. Recent technology advances with the Internet have strongly impacted individual's knowledge management behaviors. The new design challenges have been identified. On the one hand, in the Big Data era a large amount of re-usable data is generated every day, but there is few knowledge creation tools support highly efficient remix of such valuable materials. On the other hand, in the Social Web era people are sharing various content publicly, but there is rarely any expert-sourcing system could make good sense of the shared information for identifying the right experts for specific problems.

I introduce a general framework that combines models of re-use existing knowledge in the *cloud database* with models of retrieve non-existing context-awareness knowledge via *social web crowds*. This thesis develops this framework through a series of prototype systems. The first, *OMR*, is a model for supply creative workers with ability to access legal media sources and provide automatic credits. The second, *SidePoint*, is a peripheral panel that supports presentation authoring by embedding implicit search. The third, *UbiAsk*, is a Q&A service designed for assisting foreigners by involving the local crowds to answer their image based questions. The last, *MoboQ*, is a social Q&A that the user can ask time and geo-sensitive questions and receive answers that crowdsourced from strangers on social web.

The results of the thesis are currently in use by over 100,000 people. Overall these systems point to a future where the social web crowdsourcing and the big data in the cloud are central elements of personal knowledge management.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In 2012, Twitter[1] send 340 million tweets a day, Tumblr[2] publish 27,000 new posts a minute, and Foursquare[3] perform 2,000 check-ins every 60 seconds [3]: big user base, big data, big noise. A problem for individuals to seek, create and share valuable knowledge? No, an opportunity.

1.1 Background

Personal Knowledge Management (PKM) [4] is a sub-domain of Knowledge Management (KM). While KM normally focuses on how to manage organizational knowledge, PKM emphasizes the crucial important of the individual in every knowledge processes and focuses on the individual's quest to learn, work efficiently, or socialize in the processes. Therefore, the management of personal knowledge comprises the practice and study of the daily activities a person performs in order to acquire, create, organize, search, maintain, retrieve, use, and distribute the knowledge needed to meet goals and carry out roles and responsibilities [56; 108].

Although initially PKM was approached as a framework to organize the knowledge that is important for individuals [4], it has evolved over time and now involved how a (large) group of individuals collaboratively organizing personal

[1]http://www.twitter.com/
[2]http://www.tumblr.com/
[3]http://www.foursquare.com/

data, making sense, sharing, exchanging, recommending, remixing and interacting [66; 110] of information. Such collaborative knowledge management activities has also evolved from simple person-to-person exchange, where one end simply shares the knowledge to another; to massively publishing the knowledge — with printed and digitalized books; till most recently, the Web 2.0 technologies became a revolutionary power that impact PKM [67; 74]. Especially with the widely spread of emerging social web services, cloud computing infrastructure, and mobile devices, for the first time in human history we came to the We-Media era [21], where almost every ordinary people could playing an active role in the process of collecting, creating, analyzing, and disseminating information and knowledge regardless to the time or location.

With the impact of Web technologies on managing personal knowledge, new opportunities and challenges arise for designing future PKM tools. Using the social computing systems, people do not just passively consume knowledge. Rather, they also act as active contributors generating new knowledge. While the on-line systems moved to cloud infrastructure, the huge amount of user-generated content is usually stored in the cloud database (that is, the Big Data), and potentially available to be re-used by other users. However, current Big Data related applications are generally only focus on data analysis, but bad at supporting intelligent data re-use and re-mix for creating new knowledge content. On the other hand, the Wisdom of the Crowd concept [101] argues that large group of people are smarter than an elite few. No matter how intelligent they are, large group of people are better at collecting knowledge and coming to wise decisions for supporting individual knowledge management activities. However, for tasks that require expert knowledge, the quality of outcomes is still quite depend on the member of the crowd [18]. Current PKM tools normally can only utilize a general crowd that consists of random members, but poorly at identify the *right* knowledge workers for forming the *right* crowd for complex tasks.

This thesis presents a general framework **(Chapter 2)** for building future PKM tools that address the mentioned challenges. This framework is applicable for systems that are designed to support personal knowledge exchanging, managing and creating in the era of Big Data [57], Crowd Computing [105], and Social Computing [109]. This framework aims to support users the ability of in-

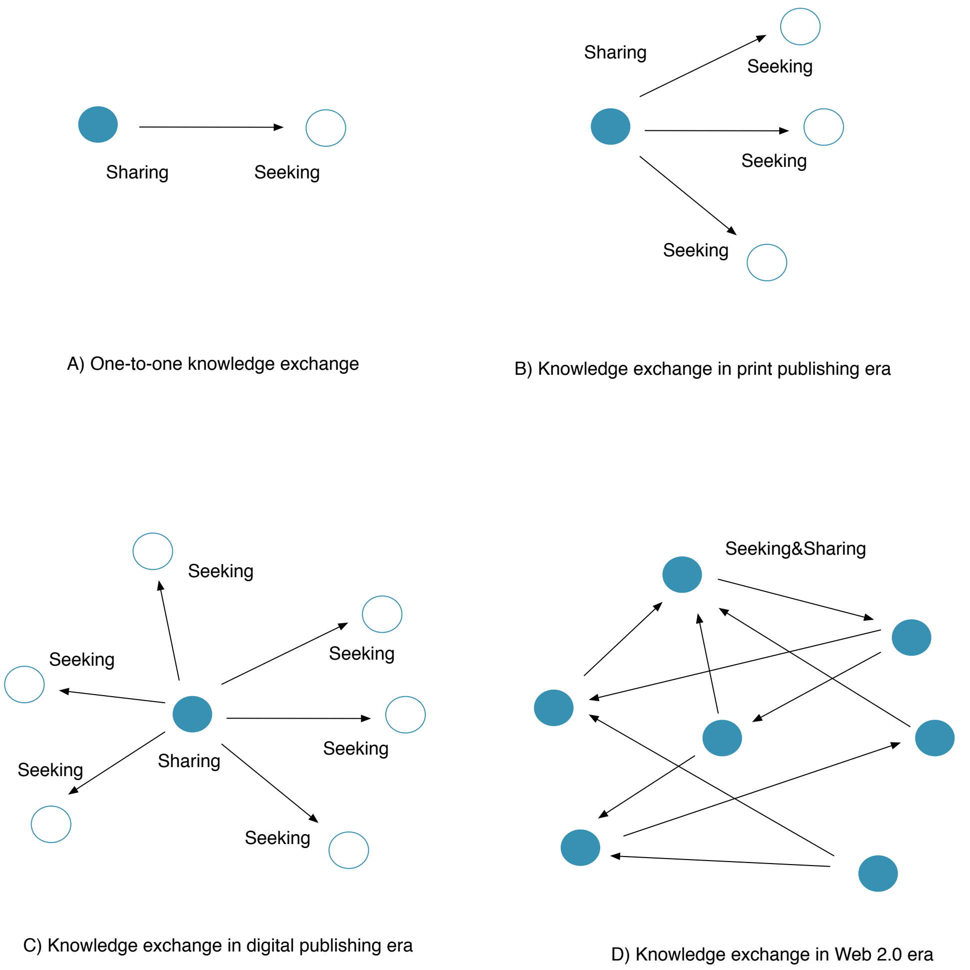


Figure 1.1: The evolving of knowledge exchange

teracting with not only the knowledge that is existed in cloud databases, but also the knowledge that is not existing before whereas could be generated by specific knowledge workers. Hence, the framework combined models of seeking and sharing existing re-usable knowledge with models of searching and using non-existing contextual-sensitive information.

1.2 Seeking Existing Re-usable Knowledge

Knowledge creation applications cater poorly to one very common usage: situations in which the users need material that they do not own and for which they are unwilling to pay. Finding and using externally produced knowledge material is currently a cumbersome process. Often, users locate the content using a search engine, copy it into their work, cross their fingers, and hope they do not infringe on any copyrights. While the authors have shared hundreds of millions of media content with permissive licenses (e.g., Creative Commons licenses), the license terms are too complicated for other users to follow. We therefore introduce an **Open Media Retrieval model (Chapter 3)** to remedy this problem and supplement it with prototypes that access various legal media sources directly within the creative work flow and provide automatic credits to the original authors.

To evaluate the concept, we developed two search functionalities for open content retrieval. **Open Content Ribbon** is a Microsoft PowerPoint add-in that enables users to search and retrieve Creative Commons-licensed images in the sidebar of the programs window. The application automatically attributes the authors according to the license terms. **AudioImager** is a video editor that can be used from a web browser. It inserts still images into the audio track based on users' search queries. These studies were focused on Creative Commons-licensed image use, but the lessons and the Media Retrieval model apply to other copyrighted media as well.

After addressing legal issues of knowledge remixing, we built end-user tools for further improving the productivity and efficiency of knowledge creation process. An example of such tools is **SidePoint (Chapter 4)**. SidePoint is a peripheral panel that panel that supports presentation authoring by showing concise knowledge items relevant to the slide content. The basic idea is integrate implicit web search and peripheral may reduce the effort to satisfy not just active needs the author is aware of, but also latent needs that she is not aware of until she encounter content of the perceived value.

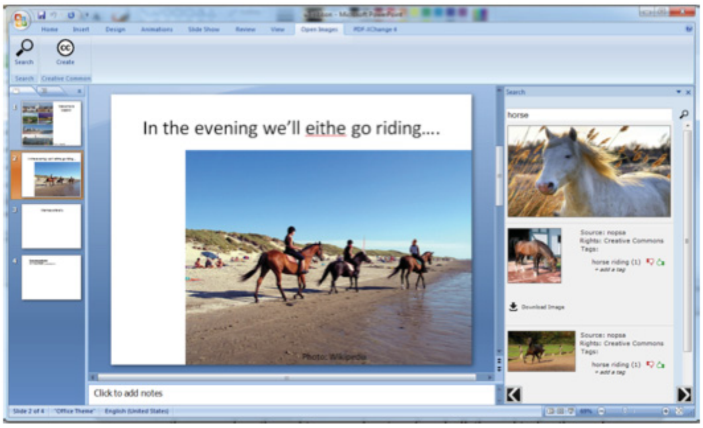


Figure 1.2: Open Content Ribbon user interface

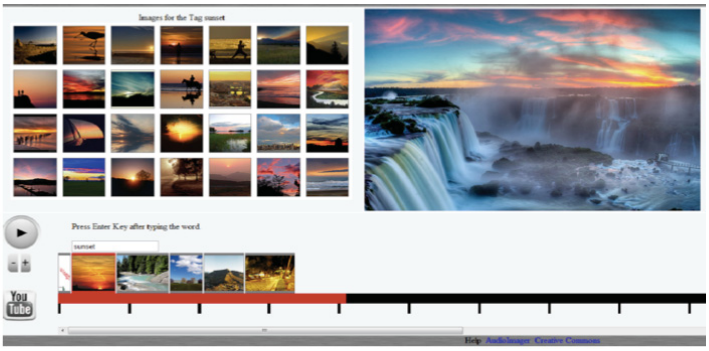


Figure 1.3: AudioImager user interface

1.3 Seeking Non-existing Knowledge

Recently, crowdsourcing is becoming a popular approach for completing knowledge works [52; 91], because there are still a large number of knowledge tasks that current computers could not do but we humans can easily handle. To our

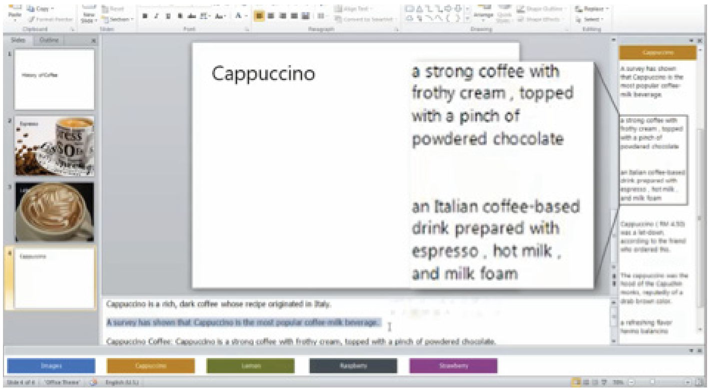


Figure 1.4: SidePoint user interface displays into presentation authoring tools

Q: Why there are so many kimono grils on the street today?
LE1: They were attending the "seijin-shiki", but I dont know its eng name, sorry.
LE2: Today is the Coming-of-age day, youth are 20 years of age need to attend a very special ceremony.

Figure 1.5: Example Q&A of UbiAsk

best knowledge, most of such systems utilize the paid crowdsourcing model [51], which basically means the task assigners should pay certain amount of money to the task workers for completing every piece of work. However, we argue that, for some types of task, existing social networking sites can potentially provide a free and on-demand knowledge worker pool for a real-time human-powered solution. In this thesis we introduce **UbiAsk (Chapter 5)** as an example. UbiAsk is a mobile crowdsourcing platform that is built on top of existing social networking infrastructure. It is designed for assisting foreign visitors by involving the local crowds to answer their image-based questions at hand in a timely fashion. Existing social media platforms are used to rapidly allocate micro-tasks to a wide network of local residents.

Figure 1.6: iPhone Application user interface of MoboQ

Since the birth of the Ubiquitous Computing (UbiComp) vision, one of the key challenges is how to extract context information from the physical environment. Without such information, UbiComp applications cannot provide genuine context-aware services. The ordinary approach is using sensors and sensor networks. However, The capacity of these machine sensors is still limited to gathering somewhat low-level physical environmental data, e.g., speed, temperature, and pressure, but for context-aware applications, it is important to also have higher-level information, such as local knowledge, human activity, social environments, and so on. Moreover, this knowledge is often highly time oriented, which means its re-usability is very low, and the value of the knowledge decrease as the time pass.

We extended the original idea of *Human Computation* or *Human as Processor* [105] to *Human as Sensor*, where human users of existing social media services become part of sensor networks and report local information around them using portable devices. Using such an approach, a system can collect information that is difficult (if not impossible) to obtain by machine sensors, thus offering the ability to generate a richer contextual model. In this thesis, a location-based real-time question answering service called **MoboQ (Chapter 6)** is presented as one study case. Using MoboQ, people can ask temporal and geo-sensitive questions, such as how long is the waiting line at a popular business right now, and they then receive answers that are crowdsourced from other users in a timely fashion. To obtain answers for questions, MoboQ utilizes a "strangersourcing" approach by analyzing the live stream from public microblogging services to identify people who are likely to currently be at the place that is associated with a question and sends them the unsolicited question through the microblogging service from which they were identified.

1.4 Contributions

This thesis provides evidence that PKM systems can draw on crowd intelligence and big data to support a wide range of user goals and activities. The systems in this dissertation demonstrate that cloud- and crowd-powered systems re-open in many areas of PKM design. For example, MoboQ introduces new social computing techniques such as social web based expert searching, and Open Content Ribbon demonstrates new interaction such as make effective use of legal open images in presentation slides authoring tools.

The contribution of this thesis, therefore, is a general way of think about the convergence of crowd computing, open content movement, social computing, and big data for comprehensively supporting individuals to seek, share, and create not only existing re-usable knowledge but also non-existing contextual knowledge. Whereas previous models have not covered a number of critical new user requirements such as legal usage issues, serendipity knowledge discovery, knowledge re-mix, and contextual-oriented knowledge seeking.

Chapter 2

General Framework

This chapter lays out a general framework for designing online personal knowledge management systems. For the clarity of discussion this chapter are divided into three parts:

- **Design requirements**. We present what are the design requirements that are derived from potential users' demands. In addition, we discuss the design factors and newly available technologies that we need to consider while building a system to meet the requirements.
- **A framework for developing Internet-based personal knowledge exchange systems**. We propose a general method for designing applications to solve open problems. We also discuss some of the design trade-offs in this section.
- **Limitations**. We discuss a list of limitations of this approach.

2.1 Related Work

There have been some earlier efforts that were related to personal knowledge management, cloud powered content remix tools, and crowd powered systems. However, as far as we know there is no previous research discussed how to combine their advantages together to support personal knowledge management activities.

In this section we present some of the previous research that is relevant to the framework that is proposed in this paper.

2.1.1 Personal Knowledge Management

Traditionally, personal knowledge management is handled by non-electronic tools, such as todo list, calendar, desk diary, appointment book, and so on [60]. Those tools only support authoring but not searching, reminding, re-use, or collaboration. With the wide adoption of IT technologies, recent researches have been conducted to situate personal knowledge management in the Internet and Web 2.0, and investigated the potential role of information technologies for harnessing and managing personal knowledge.

Kirby Wright proposed a personal knowledge management framework in [110]. The model involves four interrelated domains: 1) analytical, 2) information, 3) social, and 4) learning. The analytical domain involves competencies such as interpretation, envisioning, application, creation, and contextualization. The information dimension comprises the sourcing, assessment, organization, aggregation, and communication of information. The social dimension involves finding and collaborating with people, development of both close networks and extended networks, and dialogue. The learning dimension entails expanding pattern recognition and sensemaking capabilities, reflection, development of new knowledge, improvement of skills, and extension to others. This model stresses the importance of both bonding and bridging networks.

Moreover, Ikujiro Nonaka and Hirotaka Takeuchi proposed a more general model for Knowledge Management known as *SECI model* [85; 86]. In SECI model, Nonaka defined four modes of knowledge conversion in knowledge creation process:

- **Tacit to Tacit (Socialization)**. This dimension explains Social interaction as tacit to tacit knowledge transfer,sharing tacit knowledge through face-to-face or share knowledge through experiences. For example, meetings and brainstorm can support this kind of interaction. Since tacit knowledge is difficult to formalize and often time and space specific, tacit knowledge can be acquired only through shared experience, such as spending time together or living in the same environment. Socialization typically occurs in a traditional apprenticeship, where apprentices learn the tacit knowledge needed in their craft through hands-on experience, rather than from written manuals or textbooks

- **Tacit to Explicit (Externalization)**. Between tacit and explicit knowledge by Externalization (publishing, articulating knowledge),developing factors, which embed the combined tacit knowledge which enable its communication. For example, concepts, images, and written documents can support this kind of interaction.When tacit knowledge is made explicit, knowledge is crystallized, thus allowing it to be shared by others, and it becomes the basis of new knowledge. Concept creation in new product development is an example of this conversion process

- **Explicit to Explicit (Combination)**. Explicit to explicit by Combination (organizing, integrating knowledge), combining different types of explicit knowledge, for example building prototypes. The creative use of computerized communication networks and large-scale databases can support this mode of knowledge conversion. Explicit knowledge is collected from inside or outside the organisation and then combined, edited or processed to form new knowledge. The new explicit knowledge is then disseminated among the members of the organization.

- **Explicit to Tacit (Internalization)**. Explicit to tacit by Internalization (knowledge receiving and application by an individual), enclosed by learning by doing; on the other hand, explicit knowledge becomes part of an individual's knowledge and will be assets for an organization. Internalization is also a process of continuous individual and collective reflection and the ability to see connections and recognize patterns and the capacity to make sense between fields, ideas, and concepts.

Although these previous works have studied and discussed modes of how people seeking and sharing knowledge, those frameworks might have difficult to directly applied to the future information-rich environment, for example, increasing amount of open content knowledge data is available in the cloud databases that could be ubiquitously accessed through various kinds of devices; moreover, a growing number of new generation user are seeking for context-aware knowledge rather than existing knowledge stored in a knowledge repository; in addition, provide serendipity information discovery for satisfying customers' latent needs is becoming as important as search-based knowledge seeking for meeting customers' active needs.

2.1.2 Knowledge Remix Tools & Legal Issues

Lawrence Lessig [69] has criticized our legal system for criminalizing the new remix generation. The web-savvy generation uses remixing as a self-expression tool involving borrowing and recombining works for which they do not own the copyrights. Studies on content reuse indicate unanimously that consumers' respect for copyrights leaves room for improvement.

Kantar Media's copyright survey [80] revealed that 73% of the respondents were not sure what is legal or illegal under the copyright laws. However, around half of the subjects who shared files illegally had also used legal online services. End users may therefore value convenience and availability over legality in their media use activities.

Diakopoulos et al. [35] found that users actually had a positive attitude toward the attribution of copyright contents. However, Seneviratne et al. [87] estimated that 70%–90% of Flickr's CC-licensed images have been reused in an improper fashion. They contend that users may be ignorant of the meaning of the licenses; do not read the license terms; use incorrect licenses in their derivative works, which violates the original content creator's intention; or even intentionally ignore the license due to their own interests. Similarly, most websites lack adequate functionality for finding licensed content, remixing, and attributing sources for remixes [95]. As a result, end users lack opportunities for effortless and appropriate media reuse.

Because of this pressing need, applications have been developed for easier attribution management in open content reuse. ThinkFree[1] lets users directly pick images from Flickr and inject license metadata automatically with them

[1] http://www.thinkfree.com

into a document. Zemanta[1] is a browser add-on for searching and recommending reusable media while working with the content. AuthorSTREAM[2] enables users to search images and videos from several online sources within PowerPoint and features an automatic attribution. However, it does not limit the search to legally reusable content. OpenOfficeOrg Addin[3] lets presentation creators set sharing and reuse permissions to their work and insert legally reusable images from Flickr, Wikimedia Commons, and OpenClipart.org. Firefox browser plugins such as OpenAttribute[4] and Semantic ClipBoard[5] simplify attribution by parsing the webpage contents and displaying the license and attribution information separately. Finally, developers have been active in helping CC authors to enforce their rights. Many prototypes and algorithms have been developed to monitor the Web for illegal re-use of open content [62; 87] and to inform the original contributors automatically.

Although the media retrieval process is simplified with such tools, Monroy-Hernandez et al. [82] have shown that automatic attribution also has drawbacks. It does not increase the content author's social value as much as a credit provided by a human user. However, our assumption is that automated attributions are a major improvement to the current situation where users fail to properly attribute altogether.

2.1.3 Crowdsourcing Systems

Crowdsourcing and human computation are approaches that seek to use human intelligence to overcome the limitations of computing technology. Jeff Howe introduced this term in 2006 to draw comparisons to outsourcing [52]: where outsourcing model pays a few external contractors, crowdsourcing model makes an open call to public for help.

[1]http://www.zemanta.com
[2]http://www.authorstream.com
[3]http://wiki.creativecommons.org/OpenOfficeOrg_Addin
[4]http://openattribute.com
[5]http://dig.csail.mit.edu/2009/Clipboard/

Marketplace Based Crowdsourcing

There are paid crowdsourcing online marketplaces allow task assigners to post tasks and provide momentary incentives. In many cases, the payment can be as cheap as several cents for each task. Workers browse these task lists to take a task. One of the most successful online crowdsourcing marketplace is Amazon Mechanical Turkr[1]. In 2010, Mechanical Turk workers have completed over five million tasks a year [55]. Other examples include Freelancer[2], oDesk[3], and mClerk[4].

Many systems have been built on top of the paid crowdsourcing platforms. Common usage of paid crowd included gathering labeled data, solving difficult artificial intelligence problems, or conduct user studies. In computer vision domain, paid crowds annotate people and objects in images [99] and identifying depth layers in graphs [41]. In Natural Language Processing, crowds help with tasks like affect and word sense disambiguation [25; 78]. In Speech, paid crowds are used to generate large speech corpses for spoken language research [25; 78]. In Human Computer Interaction, many researchers started to use Mechanical Turk for user studies [48; 63; 77].

Community Based Crowdsourcing

Many popular crowdsourcing systems are powered by volunteered community members or members with other non-monetary incentives for participation. Luis von Ahn proposed Game with a Purpose model [104] to design online games for completing difficult artificial intelligent problems such as labeling images [106], rating photos [45], or building 3D models [103], while users are playing.

Other examples of non-monetary incentives included Duolingo system[5], an online education platform where people learning new language while translating the web into other languages as a side effect; or reCAPTCHAs [107], a web security measure requires user to complete an AI task to identify themselves as

[1]http://www.mturk.com
[2]http://www.freelancer.com
[3]http://www.odesk.com
[4]http://www.mclerk.com
[5]http://www.duolingo.com

human being rather than bots.

2.1.4 Social and Expert Search

Typical crowdsourcing and human computation systems treat contributors as exchangeable "cogs in the machine". In contrast, services that can be categorized as social search systems seek to leverage each contributors unique skills and knowledge. They help users identify and connect with relevant experts who can offer solutions in a timely and human manner. Macdonald and Ounis [75] proposed a data fusion-based approach for predicting and ranking candidate expertise with regard to a given question. The effectiveness of using adapted data fusion techniques was demonstrated by their evaluation results. Aardvark [50] is currently one of the most popular community-based social search engines. Aardvard users ask questions similar to other on-line question and answer sits, but the Aardvard server "routes the question to the person in the requesters extended social network most likely to be able to answer the question".

2.1.5 Conclusion

Tools for supporting remix legal content to help knowledge creators build new works were previously out of reach. Simultaneously, crowdsourcing are used for aiding computation by gathering training data, evaluating systems, and providing on-demand cognition. This thesis synthesizes these ideas, drawing on cloud computing and crowd intelligence to support future personal knowledge management systems.

2.2 Design Requirements

The nature of Internet based knowledge management, networked society, and potential usage scenarios, pose the following three design requirements that need to be addressed for designing a new framework for building tools for supporting individuals' knowledge seeking, sharing, and creation activities.

- **Support knowledge re-usage and remix**. Current on-line services users, especially Web 2.0 services, are generating a huge amount of content everyday. The content types are range from simple plain text to various multimedia. Since the sever side infrastructure has moved to the cloud, now those User Generated Content data is often stored in the cloud database. Therefore the UGC data is potentially available for re-using by other users. On the other hand, in the recent years we have witnessed the raising of Remix Culture, which refers to the a "global activity consisting of the creative and efficient exchange of information made possible by digital technologies" and is supported by the practice of cut/copy/paste [83]. The term remix can be defined as "separating and recombining many types of media, including images, video, literary text, and vide game assets" [36]. Consequently, remixer are those users who require different types of media in creative processes and thus the suitable content in the cloud could become an integral part of their remix creation processes.

Legal ownership is an important issue regards to the content remix behaviors. Normally the remixers do not own the copyright of the media they need and unwilling to pay. However, the Open Content movement [28; 71] is growing and creating a pool of openly licensed knowledge content by making hundreds of millions of works available free of charge using Creative Commons (CC) [10] licenses. But the Creative Commons licensed content all has some restrictions on its use. Therefore, compliance to the license terms has become another important requirement of legal remixing.

- **Support serendipitous knowledge discovery**. The term of *serendipity* is generally defined as the art of making an unsought finding [13]. It describes the moment when people meet fortunate discoveries by chance, or in other words, the accident of finding something interesting or useful without looking exactly for it. Such magic moments often consist of two different but equally important characteristics: the item or the information should be totally unexpected to the person that comes upon it, but she would feel it is valuable and interesting once presented to her. In other words, it represents the user's latent needs. While the active needs means the typical of what user would attempt to satisfy through proactive information search.

Design for serendipity has used to be a challenge for urban planners, because a in a free and peace society the understanding between people from different social class is often achieved by discussions and communications [88], and now it has also become a design requirement for digital designers of Internet services. Therefore, an Internet based knowledge creation application should be able to help users find not only informational content (to meet active needs) but also inspirational content (to satisfy latent needs).

- **Support contextual knowledge seeking**. Nowadays, personal gadgets and other handset devices communicate wirelessly with each other and also with the various services in the cloud. This phenomenon results in a rising demands of context-awareness applications. This movement changes the traditional knowledge exchange processes as well, user not only seek for the existing knowledge that is stored in a database, but also the contextual knowledge, such as time oriented knowledge (real-time knowledge where the value of information is strongly affected by time) and location oriented knowledge (local knowledge that is only valid in specific location). On the other hand, in many situations machine sensors and sensor networks have met difficult problem to extract high layer contextual knowledge from environment, such as human activities, social environment, and group emotions. Therefore, in addition to searching data on the Web or databases, a knowledge seeking application needs to have a new model for collecting such dynamic and sensor un-extractable information.

2.3 A General Framework for Personal Knowledge Management

The framework is defined with three main components:

- **Model of re-usable knowledge seeking**. Proxies that search and collects suitable knowledge items source from various knowledge sources in the cloud.

- **Model of non-existing knowledge seeking**. Proxies that identify specific knowledge workers who can potentially provide the required knowledge using social web crowdsourcing.
- **End user applications**. Various client side programs that user could use to access the services. They can be independent applications as well as add-ons of existing knowledge creation tools. The output of the services may be stored in the databases for future re-use.

2.3.1 Problem Identification

There are numbers of knowledge management activities that could be supported by digital technologies. The proposed method works best for certain types of problems, so the first step in developing a new personal knowledge exchange tool is identifying whether the problem in question is a good candidate.

The problem should meet three general requirements. Firstly, the activity should involves heavy knowledge exchange process, thus improving the process could significantly affect the performance and/or user experience of the activity. Secondly, the knowledge items are public shared common knowledge, thus the process would not require special domain knowledge that is only stored in a few close source databases or could provided by only very specific experts. Finally, a single task towards solving the problem should be lightweight and uncostly.

Take our example of the developed prototypes. We chose activities such as presentation authoring, music video re-mixing, local knowledge seeking, and image based translation.

2.3.2 Knowledge Sources

The tools collect lightweight micro-tasks via end-user applications. Different approaches are available for solving the problems based on the specific requirements: one is search in the existing database or Web for direct answer, the other one is search in the social web to find human experts to solve the issue. Likewisely, there are two different proxy modules for the two approaches, each proxy connects to a number of online service platforms. The *Social Search Proxy* links to social

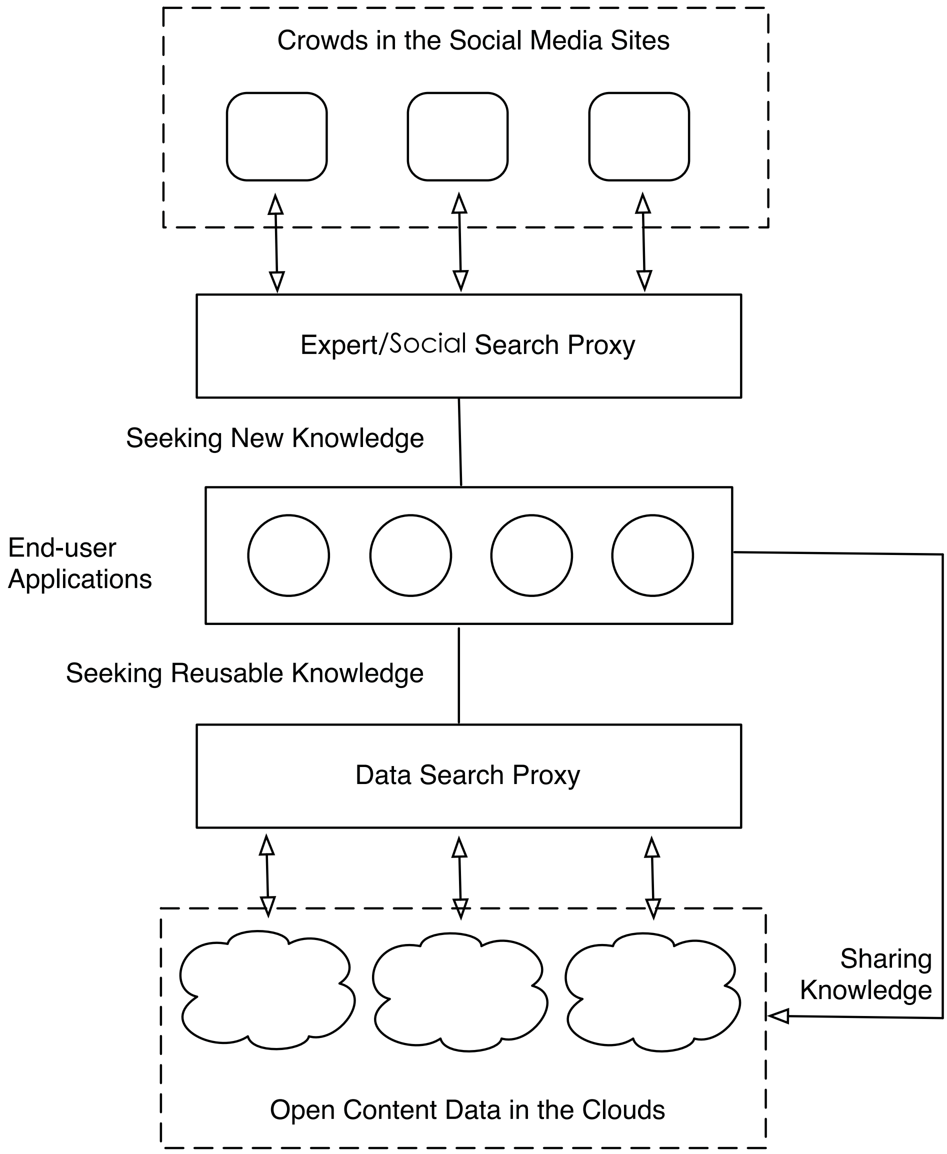


Figure 2.1: The general model

web services such as Twitter, Sina Weibo, and Facebook via their open APIs, search for the right knowledge workers for the task on the social web platforms, and assigns the micro-tasks to them through the social media services which they were found. The *Data Search Proxy* might connect to Web search engines such as Google or Bing Search, knowledge database such as Wikipedia or NeedleSeek [1], and online media sharing services such as Flickr or Youtube, through their APIs, and search for the necessary materials.

2.3.3 Incentivizing Contributions

Since social web crowdsourcing approach is utilized in the of the framework, how to motivate individuals to contribute work becomes one of the most important challenges.

Basically two motivation methods are applied in the developed prototypes: the social incentives and the game-based incentives. Monetary incentives are not discussed here because once with money being involved, quality control becomes a major issue due to the anonymous and distributed nature of crowdworkers. Although the quantity of work performed by participants can be increased, the quality cannot, crowdworkers may tend to cheat the system in order to increase their overall rate of pay. Another drawback with economic incentives is that they can destroy pre-existing intrinsic motivations in a process known as "crowding out" [34], so they are better used when other motivations are not likely to exist. In addition, previous study has shown for lightweight simple tasks the performance difference between paid mechanism and free mechanism was significantly less than complex tasks [39].

Social Psychological Incentives

One widely harnessed set of non-monetary approaches to promoting increased contributions in digital services can be found in the literature of social psychology. Social facilitation and social loafing are, perhaps, the most commonly cited effects. Social facilitation effect [112] refers to the tendency of people perform better on simple tasks while under someone else's watching, rather than while they are

[1] http://needleseek.msra.cn/

alone or when they are working alongside other people. On the other hand, the social loafing effect [17] is the phenomenon of people making less effort to achieve a goal when they work in a group than when they work alone, since they feel their contributions do not count, are not evaluated or valued. This is seen as one of the main reasons group are less productive than the combined performance of members working alone.

Ways of taking the advantage of the positive social facilitation and avoiding the negative social loafing in on-line data collection systems were suggested in [14]: individuals' efforts should be prominently displayed, individuals should know that others can easily evaluate their work, and the unique value of each individual's contribution should be emphasized. Cheshire et al. [31] conducted a series of quantitative field experiments to examine the effects of social psychological incentives and their results demonstrated social psychological incentives like historical reminders of past behavior or ranking of contributions can significantly increase repeat contributions.

Game Based Incentives

More recently, digital designers have begun to adopt ideas from game design to seek to incentivize desirable user behaviors. The idea of taking entertaining and engaging elements from computer games and using them to incentivize participation in other contexts is increasingly studied in a variety of fields. In education, the approach is known as serious game [115] and in human computing it is sometimes called games with a purpose [104]. Most recently, digital marketing and social media practitioners have adopted this approach under the term gamification [114]. The idea is to make a task entertaining, like an on-line game, thus making it possible to engage people to conscientiously perform tasks. The valuable output data are actually generated as a byproduct by the game. The main game mechanics that are used to gamify a system include a points system, rewards, a leaderboard, levels, and an epic meaning. The idea is to create a so-called gamication loop (see Figure 2.2) in the non-game system: the iteration starts with a clear goal or challenge that has a specific winning condition. Every time that the user achieves a small goal, some rewards are given accordingly, which is supported by the point system (score, virtual currency, experience point). Based on the point system and the achievements history, a leaderboard (global or partially) and badges are provided to players for motivating competitiveness, which eventually results in a change in the players' virtual status in their social network or in the system. In the meanwhile, the user interface should be interesting, and the interaction with the system should be fun (e.g., add animated transitions, well-designed avatars, funny sound effects, and more). In short, the gamified system should have a game-like surface to provide the user a more fun experience.

The ESP game [1] is one of the most famous examples of this kind. Players are paired in the game and they need to give relevant descriptions for a given image. If the description matched with other player's answer, players win and score the points, otherwise lose. The real purpose of the game is to rapidly collect annotations for a large number of images. Thereafter, various games have been devised in the style of ESP game. Arase et al. [15] proposed a web-based multi-player game to collect knowledge on the geographical relevance of images, in order to better represent certain images' geographical context for searching and browsing. Other than the casual games, Krause et al. [65] implemented a relatively complex action game OnToGalaxy in the context of human computation. In their design, the task is hidden in the game play that it is no longer perceived as a dominant element of the game.

2.3.4 Temporality and Quality

Depending on the usage cases, different applications have different requirements on how quickly and how good the system needs to finish the work. Usually system designers need to face the trade-off between time and quality or cost. However, by having the Crowd and Cloud models working concurrently, it is possible to improve the trade-off. By utilizing re-usable knowledge in the cloud, system could provide an almost instant response to the end users. In the cases that the data in the cloud cannot satisfy the request, the social search model could try to involve other human experts to support the request. Such expert search function might

[1] http://www.espgame.org/gwap/

Figure 2.2: Gamification Loop

take some time, but complement the data search model naturally.

Therefore, based to the application's requirements, system designers should choose to apply one of the models or both during implementation. For the presented example prototypes in this thesis, the applications only uses one model, but it is not hard to foresee the value of using both models into one application.

2.4 Limitations

To complement systems research, we also must understand trade-offs and fundamental limits of the crowd and cloud based framework. This subsection offers some reflections on the limitations of the proposed models.

2.4.1 Size of Crowds

Crowd model is designed to target contextual knowledge seeking problem, where one basic scenario is location and time oriented knowledge seeking. In this scenario, system tries to find the "local experts" — social web users that are located at the specific location around the specific time. The assumption is that with the fast growing user number of Web 2.0 users, the social web users will be our ubiquitously distributed sensors. However, the assumption could be limited to urban area in the countries with a reasonable large amount of population.

We deployed our Location based Social Q&A prototype MoboQ in China and integrated with Sina Weibo[1], a local social media service with more than 500 million registered users in 2012. The Weibo users are basically speak same language and live in mainland China. With such an user background, the performance of MoboQ is desirable. However, if we consider places like north European countries where the territorial area is large while the population is rather small, this approach might not work as well as it in China. On the other hands, from another point of view, the places with large number of people surrounded should naturally become the most cared places of other people, and places with less people also get less interest from other users. In this sense, it would be very interesting to deploy similar service in less crowded countries and see the actually performance.

2.4.2 Service Level Guarantee

The output of crowd model is unstable, this issue directly affects user experience. When apply non-paid social web crowdsourcing approach, the micro-task workers are somewhat “volunteered” to contribute to the service without any strict contract or agreement. Which result in an inconsistent overall performance. Although it is possible to generally predict a rough distribution of the performance, for example, 70% of the questions can receive the first response within 10 minutes, for any individual user we can hardly guarantee her that hers question would surely get a response within 10 minutes.

Although currently we do not have a clear solution, some workarounds may exist. For example, it could be helpful to provide an estimated processing time for each task based on historical data. Additionally, a notification should be given when it is highly likely that the user’s expectation for the response time would be higher than what the service could offer, for example, in the case of seeking for contextual knowledge about an uncommon location or at midnight. An appropriate guide should be provided to suggest possible alternative tasks or

[1]http://www.weibo.com/

choices.

2.4.3 Scalability

This thesis shows examples of push-based method to allocate micro-tasks to social web crowds, but crowdsourcing model generally have limitations regards to scalability. If the active user number of our application grows to dozens of millions, it might be very difficult to find an enough amount of micro task workers. However, compared to traditional crowdsourcing systems that hire workers from online marketplaces such as Amazon Mechanical Turk[1], use existing social web sites with hundreds of millions registered users as potential worker candidate pool seemingly a more promising and scalable solution.

When considering use social web crowdsourcing method, there are at least two factors need to take into account. Firstly, we should well understand that whether the social web service has friendly policies for its third party applications. Normally the social web site's open platform has some limitations of the API access amount per day or per hour. Therefore, when your application scales, you need to be able to get higher API access authority from the social web service as well, otherwise the capacity of the API access might be outstripped. Secondly, although use non-paid crowdsourcing increases the size of crowd-worker pool for scaling, we should be aware of that certain tasks are not suitable for non-paid approach. In generally, repeat, simple, and boring tasks that require low innovate and intelligence, such as word translation, could have a better performance and be better-motivated using paid crowdsourcing [16].

[1] http://www.mturk.com/

Chapter 3

Open Media Retrieval: Support Legal Usage of Open Knowledge

In our networked society, users are inserting more visual elements into their everyday communications. People create and share presentations online, post blogs, and upload videos more than ever. Our Information Society has become what Lessig calls a read-write society [69] where culture is no longer consumed only by viewing it, but also as a source for remixed creativity. Remix culture can be defined as a "global activity consisting of the creative and efficient exchange of information made possible by digital technologies" and is supported by the practice of cut/copy/paste [83]. The term *remix* refers to "separating and recombining many types of media, including images, video, literary text, and video game assets" [36]. Consequently, remixers are those users who use different types of media in creating their works. They need external content to fuel their creative processes and thus the search for suitable content is an integral part of their remix creation processes.

With growing demand for free and remixable content, millions of creators have realized that in the remix culture, exclusive copyrights can be a serious hindrance. They have helped to create a pool of openly licensed works by making hundreds of millions of works available free of charge using Creative Commons (CC) licenses. CC images are royalty free and available to anyone. However, they all have some restrictions on their use. Compliance to the license terms has become another important part of legal remixing. Unfortunately, this is an area where the marriage of remix culture and the CC licensed content faces a problem. For example, every CC license requires the licensee to give proper credit to the original author. Recent studies suggest that users fail to properly attribute CC licensed works [94]. This mismatch of media use motivated us to address a number of research issues including:

- Do people use open content according to the license terms provided with the content? If there are infringements, how prevalent are they?
- Why and how do people fail in using open content appropriately?
- How can we improve this situation?
- Can an improved media management concept be more generally applied to remix applications and different media formats?

This section describes an Open Media Retrieval (OMR) model in which the use of legally available media is automated. This is exemplified with two media creation systems, their pilot studies, and our analysis of the implications. We started our research by studying how users find and use media. This led to many insights into the problems of users' behavior with existing tools and how the needs of users can be facilitated with improved tools. This enabled us to envision improved interaction principles to facilitate the use of open content in a media creation process. To evaluate the model, we developed two search functionalities for open content retrieval.

Open Content Ribbon is a Microsoft PowerPoint add-on that enables users to search and retrieve CC-licensed images in the sidebar of the program's window. The application automatically attributes the authors according to the license terms. AudioImager is a video editor that can be used from a web browser. It inserts still images into the audio track based on users' search queries. These studies were focused on CC-licensed image use, but the lessons and the OMR model apply to other copyrighted media as well.

3.1 The Cost of the Media Retrieval Process

Remixers do not always have the exact media file they need in their collections. This triggers a media retrieval process. At first, the user needs to reach a source where a suitable media can be found. The user's subsequent actions are directed at finding particular piece of media that satisfies the search criteria [22]. This is followed by insertion of the media file from the Internet into user's media creation application. However, media retrieval and usage come with legal costs.

3.1.1 Legal Costs

In principle, legal restrictions are one of the biggest obstacles in using online media. Default search engine results present mostly copyright-protected media. These media files can be copied and used, but without permission from the rights owner, this use can easily infringe on copyright law. However, in practice, some users decide to ignore copyrights and infringe anyway. They purposefully choose to risk the legal repercussion of their infringement.

Lessig has criticized excessive copyrighting for hurting modern creativity, where remixing is the modusoperandi [69]. Copyrighting creates an obstacle for the use of media files that were not created by the user. It grants the author an exclusive right to reproduce and publicly display or perform images, videos, and other creative works. Fortunately, authors and service providers have created markets for commercial and free exchanges of works, thereby helping to match supply to demand. These markets carry various direct and indirect costs.

Users can overcome copyright's exclusive nature by getting permission from the rights owner. However, finding the rights owner does not guarantee his/her willingness to license the content. Individual licensing is rarely efficient and several rights owners have a policy of non-cooperation, which means that he/she is not willing to license work to individuals, even for a fee [68]. Luckily the digital content markets are highly competitive and offer several options for obtaining content from content brokers like stock content services. However, rights owners typically set a price to license their otherwise exclusive rights and for many users even the smallest direct license cost can limit the use of the work.

Licensing also includes indirect transaction costs, for example, the search costs of finding the content, the licensor and negotiation costs to reach an agreement in terms of use, and enforcement costs to make sure that the use complies with the licensing terms [44]. Many commercial stock media services have reduced transaction costs by providing tools that simplify the licensing and retrieval processes. There are several add-ons[1] that plug and integrate the search, payment, and downloading of stock photos and other media for media creation applications. However, many users are either unable or unwilling to pay direct licensing costs for commercial content.

Fortunately, there are several sources for getting remixable media free of charge. In 2007, Wikimedia Commons[2] adopted by far the most popular open licensing scheme — Creative Commons — that is currently used to license millions of media files. The files are crowd-selected, annotated, and structured in a hierarchical ontology. Another source is Flickr[3] with its 200 million CC-licensed images. As a result, during the past few years, we have witnessed an increasing number of users sharing, promoting, and spreading their works and ideas [29; 30; 73], under CC licenses. CC-licensed material provides an attractive means for end-users to overcome the direct and indirect licensing costs because the licensed content is royalty free and getting the license does not involve laborious negotiations with the rights owners. Also, it seems that the quality of even the least restrictive CC-licensed images is comparable to Flickr's "all rights reserved" images [49]. The CC license's attribution requirements, however, lead to a design challenge, which we describe below.

3.1.2 Open Content Transaction Costs

A growing pool of open content is licensed with various public licenses (such as CC) that permit sharing and often remixing of the works. While these licenses reduce legal restrictions and are often gratis, they include a search cost for finding the content, a technical cost for downloading the content, and a compliance cost for making sure that the intended use falls within the license's scope. For a user,

[1] E.g., iStock Office (http://www.istockphoto.com/istockphoto-office-ribbon) and Getty Images Photoshop (http://www. gettyimages.com/creative/frontdoor/adobeplugin) plug-ins/

[2] http://commons.wikimedia.org/

[3] http://www.flickr.com/creativecommons/

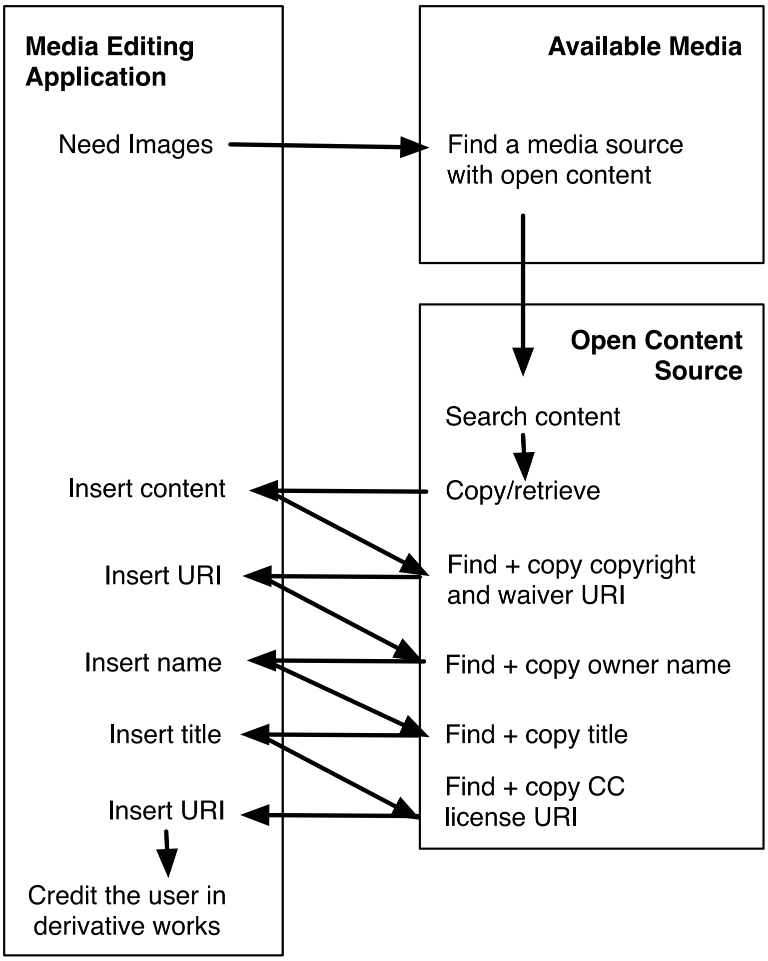


Figure 3.1: Media retrieval process with CC content

a proper licensed use of works involves numerous steps.

Figure 3.1 describes the required steps for using CC-licensed content. It illustrates the number of effort-related actions that the licenses impose on the users. Most of these steps are aimed at ensuring the correct attribution alongside the actual content. The attribution requirement is defined in detail in a long legal document that few licensees have ever read. CC licenses require the proper attribution to include 1) the Uniform Resource Identifier (URI) of the CC license used, 2) the name of the author, 3) the title of the work, 4) the copyright and the disclaimer of warranties' URIs if provided by the licensor, and 5) a credit identifying the use of the licensed work in the derivative work.

At least a dozen separate tasks are needed to ensure the compliance with the attribution terms. Such a procedure is prone to errors. Unfortunately, none of the major media creation applications are optimized to support a retrieval process in which remixers search and use open content. Therefore, remixers have the responsibility to study the details of the license terms and follow all of them. Failing to do so can lead to a grave situation as a CC license states that any breach of the license terms will terminate the license.[1] As any non-compliant attribution will terminate the license, the licensee can be left defenseless. The licensor could bring an infringement lawsuit against the licensee. Had the licensee followed the license terms to the letter, the licensee would have a legal defense against such claims.

3.2 Pre-studies

We carried out two pre-studies to answer our research questions:

- How prevalent is the inappropriate use of works?
- Why are users failing to follow the terms of open content reuse?
- How can we improve the situation by identifying the difficulties?

[1] Creative Commons Attribution 2.0. Available at http://creativecommons.org/licenses/by/2.0/legalcode/

3.2.1 Image Reuse in SlideShare.net

To get an idea of how image copyrights are currently respected, we analyzed image reuse in presentations. SlideShare[1] is the largest on-line presentation slide hosting and sharing service with 16 million unique visitors a month and about 100 million registered users. Our goal was to examine the image sources that people use when preparing presentations and to study whether they knew how to correctly credit the original authors. We analyzed 51 popular presentations on SlideShare that were randomly selected from the first five pages of results among the most viewed, most downloaded, and featured works. The presentations had an average of 300,000 views. All but one of the presentations were in English.

We reviewed the images in these slides and developed a binary classification tree (see Figure 3.2) to categorize the presentations into seven copyright treatment categories: No photos, own images, stock images (i.e., purchased content), images under fair use (e.g., screenshots from a website, see [70]), infringing images (either non-CC-licensed photos used without permission or CC-licensed images with improper attribution), and CC-licensed images with proper attribution.

For each image, we started by searching for a URL to the image source. If one was found, we checked whether the image was CC-licensed and continued by checking the status of proper attribution. If the URL was not available, we used the image recognition feature in Google Image Search[2] to find the source. If the search led to stock photos, we assumed that the image had been appropriately purchased. However, it was difficult to tell stock photos from non-CC-licensed photos used without permission. We had to mark seven cases as "Unknown permission status."

As a result, we found that, although 10 authors used CC- licensed images and tried to attribute, not one had managed to do so correctly. The most common mistakes included missing the URL for the specific CC license and the lack of an original image title. In particular, one of the infringing presentations was about the appropriate use of CC-licensed images in presentations! Eleven presentations had more obvious infringements by containing non-CC-licensed images without the authors' permissions.

[1]http://www.slideshare.com/
[2]http://images.google.com/

The users we studied had released 16 of the presentations with a CC license, five of which clearly contained stock images or non-CC-licensed images, causing a likely copyright infringement. We believe that most of the creators of these infringing presentations had, in fact, released the presentations under the CC license with good intentions. This shows that while CC licenses are familiar to some users, the details of the licenses can be challenging to understand and internalize even for experienced users.

3.2.2 Image Reuse Processes

To find out in more detail why users find it hard to respect copyrights, we conducted a study on users' image search strategies. We were interested in how users initiate their image searches, whether they pay attention to copyrights during their searches, whether they know which images they are allowed to use, and how they credit the author of the image when they include it in a presentation.

The test resembled a traditional usability study. We presented a story of an upcoming academic conference for which the participants were asked to prepare a three-slide presentation. The slides had to be prepared so that they could be uploaded to the Internet without any copyright infringements. The three slides had to contain images suitable for a welcoming speech to our home city and for two social events which were a horseback-riding excursion and a trip on a pirates' ship. For each slide, the participants had to find one image to illustrate their topic. We asked the participants to focus on finding suitable images and ignore all fine-tuning of the slides. We videotaped all the sessions so that only hands, keyboards, mice, and the displays were visible, and we asked the participants to verbalize their processes by thinking aloud. We let each participant find at least one image. If the image was found and included in the presentation in less than five minutes, we let the participant move onto the next slide and interrupted him or her when five minutes had passed.

Twelve lecturers (five males, average age 40) from our home university participated in the study. They came mostly from cartography (3) and computer science departments (5). On average, they gave eight lectures or other presentations per month, always using some kind of presentation software.

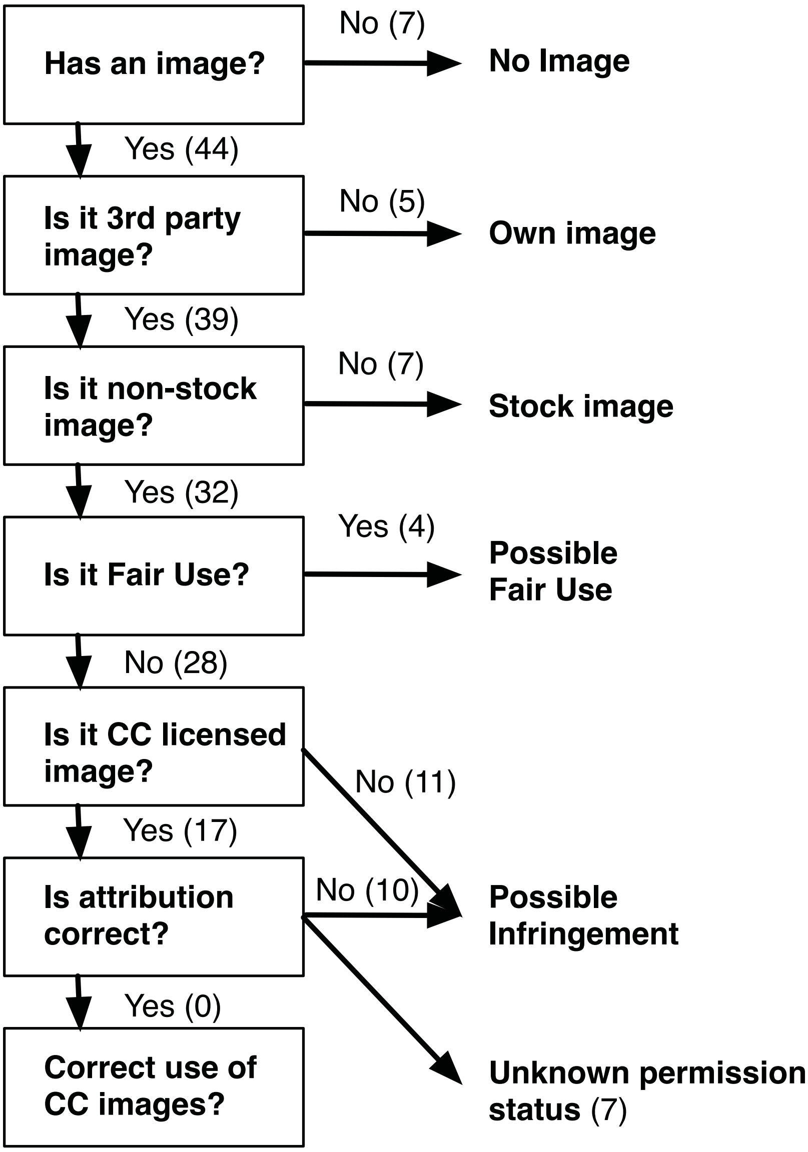


Figure 3.2: Categorization of images in SlideShare presentations

Initially, all participants remembered the requirement that the presentation should respect copyrights. As expected, based on the SlideShare study, none of the participants succeeded in creating slides that would respect copyrights. When analyzed using the schematic task sequence of necessary steps presented in Figure 3.2, our participants failed in performing these steps:

- Use of a media source with open images: 6/12. One participant attempted to find open images but failed because of the choice of keyword ("non-commercial"). Five participants made no attempt to use open media sources and performed normal Google or Flickr searches.
- Insert copyright and waiver URI: 0/12
- Insert author name: 0/12. Three users incorrectly provided the name of the service (e.g., Wikimedia).
- Insert title of the work: 0/12
- Insert CC license URI: 0/12

In short, while seven participants out of 12 tried to respect image authors' copyrights by starting to search for open images, they all failed in the process in multiple ways.

None of the participants stated that they did not care about the copyright-related instructions in our task, and therefore all the observed infringements were unintentional. Inspection of the videotapes showed some salient reasons for this. First, five participants seemed to forget the need to use open images altogether and made no attempt to direct their searches to correct sources. This was possibly because the search for a suitable image caught their attention, making them forget the related requirements.

Second, the participants were unaware of advanced search features. Of the seven participants who attempted to find open images, five used Google or Flickr (among other services), but only two used the advanced search feature that limited the results to CC-licensed images. Third, the participants did not possess the knowledge of the steps required by the CC license for image use. While six participants finally chose an image that would have not infringed on any copyrights, none of them performed the attribution-related steps successfully. Three of the attempts were erroneous because they only attributed the media source (e.g., Wikimedia), and not the author, and they did not include a link to the specific CC license.

The study therefore showed that those content users who would like to respect the copyrights would need assistance in all of the required steps. The primary reasons for infringements seemed to be related to forgetting the latter steps when a suitable image was finally found and the lack of knowledge about the required steps.

3.3 Summary of Pre-studies

Throughout the studies, we confirmed that appropriate remixing of digital media is a challenging task. We observed failures in all parts of the image retrieval process presented in Figure 3.2: in media source selection, in content searching, and in content attribution.

Nowadays, users commonly have the skill to look for images via search engines or other popular on-line image databases. However, it seems that many users copy/paste the images without attention to copyright issues. In some slides analyzed in the SlideShare study, the images even contained watermarks of a copyright notice, suggesting the users' ignorant attitude towards digital copyrights.

One noticeable thing in the studies was the fact that many users were aware of CC licenses and were able to find CC-licensed images. These users had at least basic awareness of copyrights, had an intention to respect them, and tried to show their respect to the original authors. However, even with their respectful attitude, all participants still failed in the attribution task. It appears that the current CC licenses are too difficult and confusing even for fairly advanced users. Moreover, even if they understand the requirements, the transaction costs are high: for every attribution, they have to move between different programs and repeat copy- and-paste actions several times.

3.4 Open Media Retrieval Model

To solve the issues in open content remixing, we have developed a solution that we call the Open Media Retrieval (OMR) model. The model integrates searches into the user's creative work-flow and automates the attribution process. It treats the remix process as a collection of tightly coupled searches. The media creation application's role is to act as a composing platform through which the user controls how the search results are presented to the audience. For example, presentation software is used to control the slide order and timing, and video-editing software controls the synchronization of clips and images with audio.

Our goal has been to minimize the tasks a user needs to perform to import CC works into the media creation application. The OMR model reduces the task load from twelve to two tasks: 1) users enter a query and 2) users select a media file to be attached. The second task can be avoided if the system enables the automatic insertion of suggested media from the best-matching search results. Optionally, the user can refine the search to include only certain media sources or search for more media. Figure 3.3 presents the model.

Developers can integrate this search into their media creation applications and call image repository APIs directly or they can rely on a separate back-end service to which the application can delegate the queries. In our tests, we reduced the complexity of the application by creating a back-end system that provided the query process as a web service. The majority of the task load is delegated to the back-end system, which has a list of suitable CC media repositories. The features of one possible implementation of the OMR model can be listed as follows:

- Repository contains CC-licensed content.
- Repository is aware of all the information necessary for proper attribution.
- Repository has a policy that permits connections to the stored works and their meta-data.
- Repository has an API that provides structured access to all required attribution metadata: title of the work, author and rights owners' names, licensing information and the URL of additional copyright information.

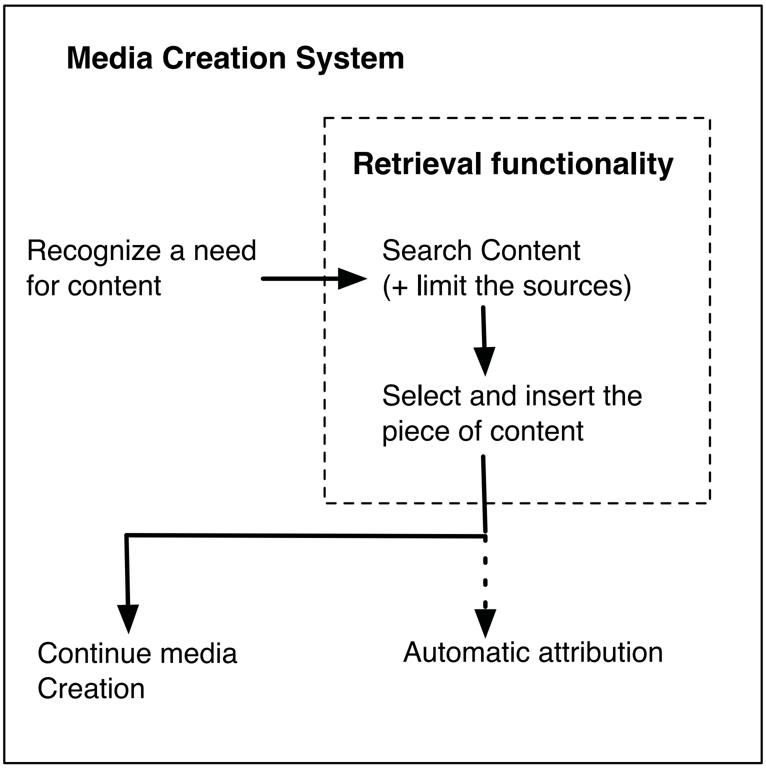


Figure 3.3: Open media retrieval model

For the end user, the media creation application has to provide integrated retrieval functionality or allow third- party functionality (e.g., a plug-in) to be connected with it for automatic content searches and insertions. This functionality will communicate with the back-end service that connects to a repository with the above-listed features. By using the back-end's API, the application either gets a URL to the original source of the media or retrieves the content from the back-end's database. It then presents the results to the user and takes care of the proper attribution.

The media creation application also needs to be able to read and analyze the author and copyright metadata and translate both into a suitable attribution format that can be embedded in the remixed content. Here the model can take advantage of the CC licensing structure, which was designed to automate the tedious license compliance issues. CC licenses come with three versions that serve different needs. 1) The "lawyer readable" license is a long legal document that includes the terms of the license written in detailed court enforceable language. 2) The "human readable" version is a simple summary of the central concepts of the license and is intended to make the license more approachable. 3) The "machine readable" version uses a rights description language to describe the license terms. Because the CC licenses come in a machine-readable form, the system can automatically help users with correct attributions. The OMR systems can use the machine-readable feature when individual works are imported to the system.

The OMR model is independent of the media format that is remixed. Depending on the media format, the content needs to be embedded differently. In our implementation, we focused on CC-licensed images. While Google and Wikimedia Commons provide ways to search CC-licensed content, Flickr is the only service that provides the means for making the required meta-data requests. Because of the speed limitations of the Flickr's API, in our implementation, we downloaded a three million image subset of Flickr's CC-licensed images along with their meta-data to our own server. Next we present applications that we created that implement the OMR model.

3.5 Study I: Open Image Ribbon for PowerPoint

Open Image Ribbon is our implementation of the OMR model for one of the most popular remix platforms: Microsoft PowerPoint. It runs as an add-on, providing a sidebar so the user can perform queries for CC-licensed images (see Figure 3.4). Integrating search and retrieval into PowerPoint saves the user from leaving the program and performing the multiple copy-and-paste tasks presented in Figure 3.1. Following the OMR model, the Ribbon can do these tasks automatically, letting the user focus on the choice of keywords and browsing the results. Every time the user inserts an image that the plugin has retrieved into the presentation, the Ribbon automatically adds an attribution to a separate attribution slide to the end of the presentation.

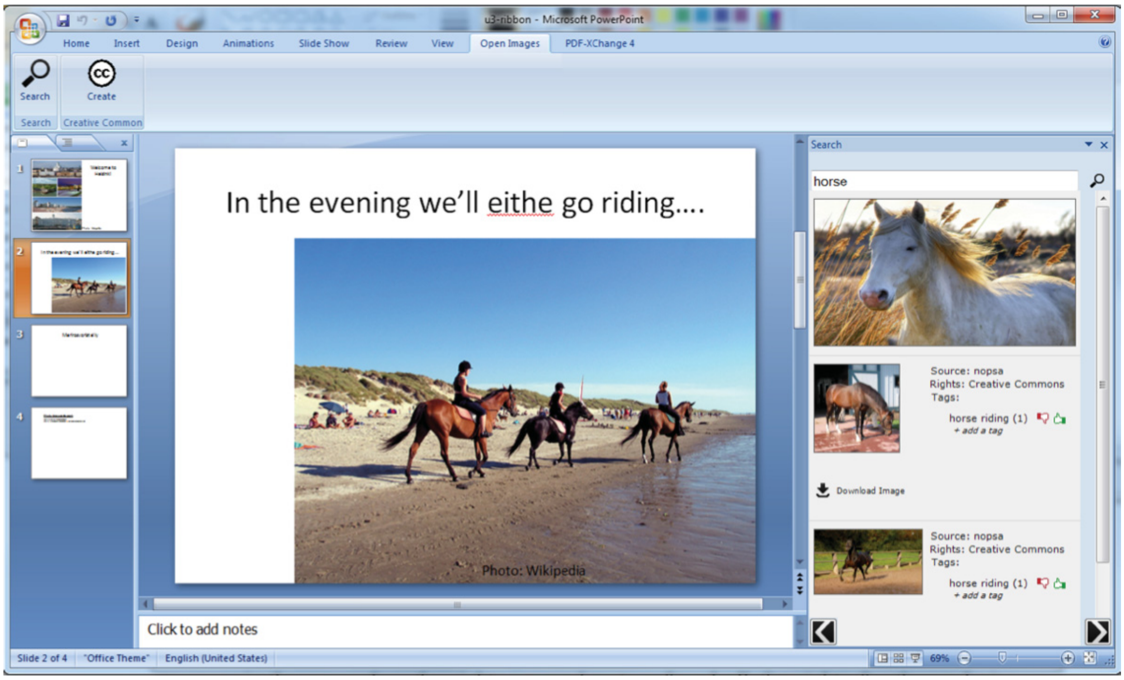


Figure 3.4: Screenshot from PowerPoint with Open Image Ribbon's Sidebar at right

We carried out a lightweight comparative evaluation of Open Image Ribbon along with our study of image search processes (presented above). In the same sessions, we asked our participants to create the same presentations using our

	Ribbon	Baseline	Significance of difference[1]
N	8	8	
Raw NASA-TLX, mean	2.5 (sd.90)	2.5 (sd.75)	Wilk's $\lambda = .86$, $F(1,7) = 1.152$, $p = .32$
Slide creation speed, slides/5 minutes	1.8 (sd.88)	1.4 (sd.72)	Wilk's $\lambda = .94$, $F(1,7) = 0.484$, $p = .51$

Table 3.1: Summary of results from Study 1

Ribbon tool. We measured the difference between the Ribbon and the Baseline conditions (i.e., the image search processes study) using the following two measures:

- Task load: After each condition, the participants filled in the Raw NASA-TLX questionnaire [24; 47] that measures the participants subjective task load using six statements on a seven-point semantic differential scale (e.g., "How hard did you have to work to accomplish your level of performance?").
- Speed: How many slides (03) did the participants create within the first five minutes of the task?

In the analysis, we had to drop the number of participants to N=8 because of a bug that caused very slow image retrievals for four participants. All the participants carried out the tasks in both conditions. We switched the order of the two tasks every time for the next participant. The results, analyzed using a within-subjects repeated-measures ANOVA, are presented in Table 3.1, showing that while the averages were slightly in favor of the Ribbon in both measures, these differences were not significant on either of the measures, neither on the task load nor on the slide creation speed.

While these statistical results are tentative, they point to a conclusion that using Ribbon was not more difficult than creating the slides using the familiar method. When participants spent a long time using Ribbon, it was because they were dissatisfied with the retrieved images. This happened fairly often and was due to the limited number of images in our database, most of which are snapshots taken by Flickr users without explicit intentions for remixing. Also, our database did not have as sophisticated rating algorithm as Google does. Therefore, the best images are not presented first in a set of retrieved images. The Ribbon was therefore handicapped in terms of image quality. We still find the results promising in this light.

Obviously, all the presentations respected the copyrights, since the Ribbon actively enforced that. Failing in this aspect could only happen if the participants had deleted the presentation's last slide or its information on purpose.

3.6 Study II: AudioImager

Video creation is another application class where the OMR model can be applied. To study this scenario, we created a prototype video creation application called AudioImager. We treated the process of video creation as a chain of image searches synchronized to audio.

AudioImager helps users turn audio files into videos. Users create videos by typing keywords while they listen to the audio file. When a user types a keyword and presses "Enter", the application automatically retrieves CC-licensed image that matches the keyword and inserts it on the timeline.

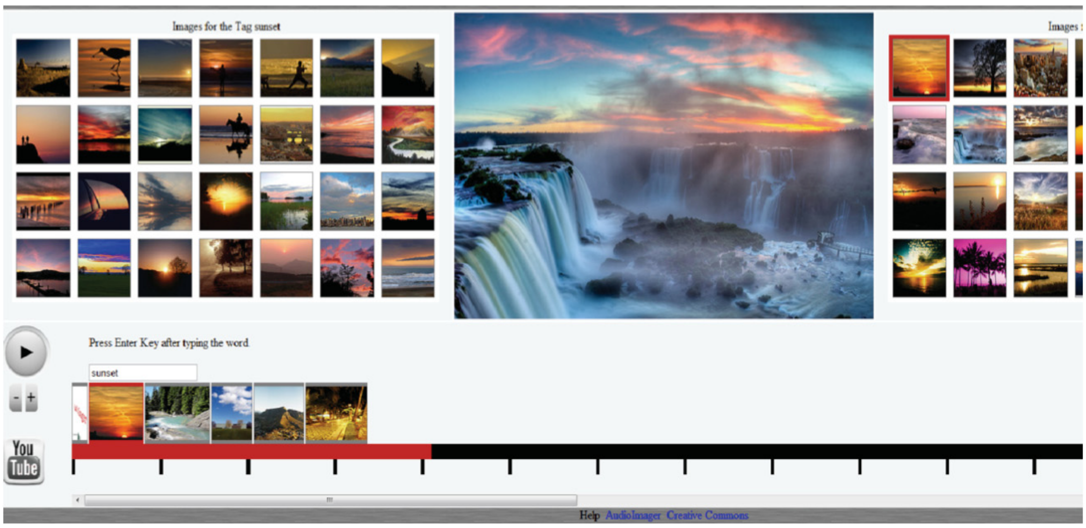


Figure 3.5: Screenshot of AudioImager. The search box is always presented at the point where video editing task place.

After the initial round of insertions, the user can change the images as needed from the selections shown on both sides of the video window. Also the exact transition times of the images are adaptable through image border dragging. When the user chooses to publish the video, the system creates the required attribution information as end credits.

In the AudioImager's design, we strove for an integrated image search and subsequently adapted the OMR model to follow a principle of a within-workspace search. By this, we mean that the image search task is integrated into the primary work area of the application. Thus, while Open Image Ribbon provided an image search in the sidebar, Figure 3.5 shows how the users could search for images in AudioImager inside the workspace. In terms of the image retrieval process and the OMR model, the primary difference between AudioImager's within-workspace design and the sidebar-based design featured in the Open Content Ribbon is that the search and the selection tasks are restructured: First, the search queries are entered one after another for all images, and then the user can start selecting the final images.

We conducted a study on AudioImager to analyze whether the within-workspace search model is suitable for the OMR model. We installed the system on workstations at a public library for one evening during an art festival. Our participants (N=16) constituted a self-selected sample of people with ages ranging from five to 70 years. First, the participants watched a one-minute demo video of the use of the AudioImager. Then we let them select a song and start creating videos. We measured the video creation durations (from the start time to the press of a publish button) and the numbers of images that the participants added to the timeline during this process. We asked the participants to work on three CC-licensed song clips, each having one- minute duration.

It took 5.7 minutes on average ($sd = 2.8min$) from the participants to create a 60-second video. On average, the videos had 11.3 ($sd = 6.2$) images. All participants succeeded in video creation, and the reorganized task sequence, consisting of rapid image searches followed by a more relaxed period of detailed image selection, ensured that the timeline of the video was filled with content that was leisurely improved. We found that the within-workspace design was suitable for this task.

In terms of the evaluation of the OMR model, the within- workspace search feature portrayed an alternative, more compressed way to implement an image retrieval functionality that follows the OMR model. The successful design of two systems — Open Image Ribbon and AudioImager — implemented using two different interaction principles, indicates that the OMR model is more generally applicable to a class of open content retrieval applications.

3.7 Chapter Discussion

This chapter has examined the difficulties that remixers have with the reuse of open content, specifically CC-licensed images. The results have been difficult with regard to the primary intentions of CC licensing: Following the licenses' rules requires a multitude of steps from the users; users lack the necessary knowledge about the requirements that the licenses put on them; they misuse content with good intensions; and they sometimes forget to search content under the CC license because they get carried away in the process of searching suitable images from sources where all content is not CC-licensed.

The chapter has presented a series of small studies, following a research approach that prioritized exploration and breadth over depth. This approach helped us examine the scope of the problem in a wide range of different content remixing activities. However, the studies involved only a small number of participants and were based mostly on quantitative measures. For example, we could not prove the superiority of our Open Image Ribbon over the baseline solution in terms of usability. A more finely tuned interaction design, larger user sample, and a wider range of measures could have probably done that, but now that has to be left for future research. However, our results have provided a promising start in that direction.

With the scale of participants used and the research methods applied, the results should not be treated as "proof" that copyright infringements permeate all remix cultures. Likewise, the results do not present direct design solutions that would remedy such infringements. Instead, they should be understood as an indication that legal reuse of digital content involves solvable issues that have social, motivational, and design-oriented dimensions. These dimensions should be studied in more detail using qualitative methods and targeted interaction design studies.

Qualitative studies would be critically needed to understand the social and motivational basis of law-respecting remixing. We believe that these two dimensions are closely interlinked. Our assumption is, pending verification, those users, when they remix others' material, need a feeling that they are doing it in the right way and take pride in doing so. They understand that remixing is reusing the material of authors very much like themselves. The motivation would therefore have a social basis, arising from peer loyalty and wishes to be perceived as reputable creators.

Regarding the design dimension, designers of media editing software can make use of this aspect of pride in at least two ways. As the user edits content, the user interface may continuously confirm and indicate points where the program knows that copyrights have been properly managed. This will increase the user's confidence and erase doubts of unintended infringement. Another suggestion is that the program adds a logo or a note to the final output file (e.g., the last slide in the slide set) indicating that a program was used to ensure that copyrights were managed when preparing the content. When other users see creations having that sign, they can then choose to adopt the same procedure in their creations. This increases awareness of copyright issues and points to a location where the work and more information can be found. In this way, the users can learn the correct methods from each other. Currently the opposite is true: Users mimic content reuse practices with good intentions from others and inadvertently sustain the infringing practices. However, eventually, when there are enough works with proper attributions, the users can spot the infringing forms of attribution. Also, with proper attributions, the licensors can use search engines to follow where and how their works are being used. This can form social pressure that starts correcting false practices.

As a general interaction design strategy, we find that users should only have to care about licensing to the level that they need to know that copyrights exist and that they are being respected. The system should hide the functionality as much as possible without forgetting the license requirements that make it mandatory. By correctly attributing the reused works, users know that they are giving the credit in a proper way to those who deserve it.

On the other hand, we do not find it fruitful that systems could force correct attributions. This would irritate users and backfire with intentional infringements. In addition, in many media formats, especially images and text, enforcements can be easily bypassed with copy/pasting and screenshots. We believe that positive encouragement is likely to lead to better results.

The OMR model presented in the chapter serves as a guideline for implementing automatic attributions that save users from error-prone manual attribution tasks. The space for design explorations is large. The two sample systems — Open Image Ribbon and AudioImager — addressed the seamlessness of search functionality: While the Ribbon presented a search box and the results on the sidebar, AudioImager featured a search directly within the workspace. Longer-term studies in real use are needed to better understand how users will perceive and understand the OMR model's content retrieval paradigm, which is tightly integrated into the content creation software. Studies should also examine more extensively how such designs compare to the current practices that often involve separate web searches, downloads, and insertion phases.

Wide-scale adoption of copyright-compliant creations could also have positive second-order effects. For example, proper attribution information could be used for improving search engine results. Engines could use the copyright meta-data as hints of content that users have found valuable and which is also legally reusable, meaning that it should be ranked high in retrieval results. Also, the content creators would get more credit for their work.

We find that application developers and vendors should start taking into account those users who want to use open content in their creations. Providing powerful open content retrieval tools within media creation applications would increase users' expressive capabilities for content creation. Powerful tools would lead to more attractive-looking contents. These, in turn, would receive attention from other users, and thereby increase the interest toward those applications. Also the Creative Commons organization and the rights owners should consider how to react to the attribution problem. While the license terms can be hard to understand, they serve valid purposes and cannot be changed much. One of the tasks for Creative Commons has been to educate users in copyright issues.

Based on this study, the organization could think how to include also application developers into their education efforts.

While the current situation with license compliance looks bleak, we believe that the problem can be addressed with changes to existing software by following the OMR model or its derivatives. The demand for remix features and the growing number of remixers should present an easy business case for software companies. For the open content movement, the correct attribution could also create more visibility for CC-licensed content and in time help to improve the quality of the open content in the long run.

Chapter 4

SidePoint: Peripheral Knowledge Panel for Presentation Slide Authoring

Preparing for presentations is of critical importance for many people, but often complicated by the need to collect necessary information during the authoring of slides. To understand how this affects user behavior, we conducted an exploratory study with 8 students, each spending an hour to prepare for a 5-minute presentation of their research. We observed participants spending a substantial amount of time using the Web to find both informational and inspirational content, with constant switching between the Web browser and the presentation environment. Presentation authors thus have both active needs (typical of what they would attempt to satisfy through proactive Web search), and latent needs (previously unanticipated by the author and often identified and satisfied simultaneously when browsing).

Our exploratory study also showed that the unstructured nature of Web search results and the density of Web pages often made it hard to isolate relevant and usable information. As a result, participants often had to spend more time in the Web browser than in PowerPoint. We can interpret this through the lens of Information Foraging Theory [90], which suggests people naturally adapt their environments to maximize the gain of valuable information per unit cost of time and attention. We therefore identified a design opportunity to embed implicit search into the presentation authoring environment, reducing the cost of information finding and supporting both kinds of information needs.

Figure 4.1 shows our SidePoint add-in for PowerPoint — a peripheral panel that shows concise knowledge items relevant to the content of the current slide as it is being created. We source these knowledge items from NeedleSeek [6; 97; 113] — which offers semantically relevant facts and descriptive sentences before processing and displaying them in a concise, browsable format. These items can provide value in two ways. The first is directly by left-clicking an item to copy to the notes section of the slide. The second is indirectly through item information-scent [90]: right-clicking on an item loads a Web browser showing literal search results for that item to explore details of the information. Future improvements in knowledge base technologies can directly transfer to interfaces like SidePoint.

Using SidePoint as a technology probe [54], we examine the potential benefits and issues associated with peripheral knowledge panels for presentation slide authoring. This work contributes to the understanding of the needs of presentation authors, demonstrates the feasibility of peripheral knowledge panels, and offers future design directions for such systems.

4.1 Design and Implementation

SidePoint is a peripheral interactive panel that displays knowledge relevant to current slide content. Our design allows SidePoint to be integrated into existing tools (e.g., Microsoft PowerPoint) without interfering with the main workspace.

4.1.1 Content Text Parsing

SidePoint infers what topic the user is considering for her presentation based on her text input. When the user finishes entering words or sentences in a textbox, SidePoint analyzes them with the Splat parser [8]. This parser returns a dependency tree (a tree representing a grammatical structure), and the system extracts nouns and noun phrases (e.g., adjective + noun, or noun + preposition + noun) as a keyword set. SidePoint then retrieves knowledge about each keyword set

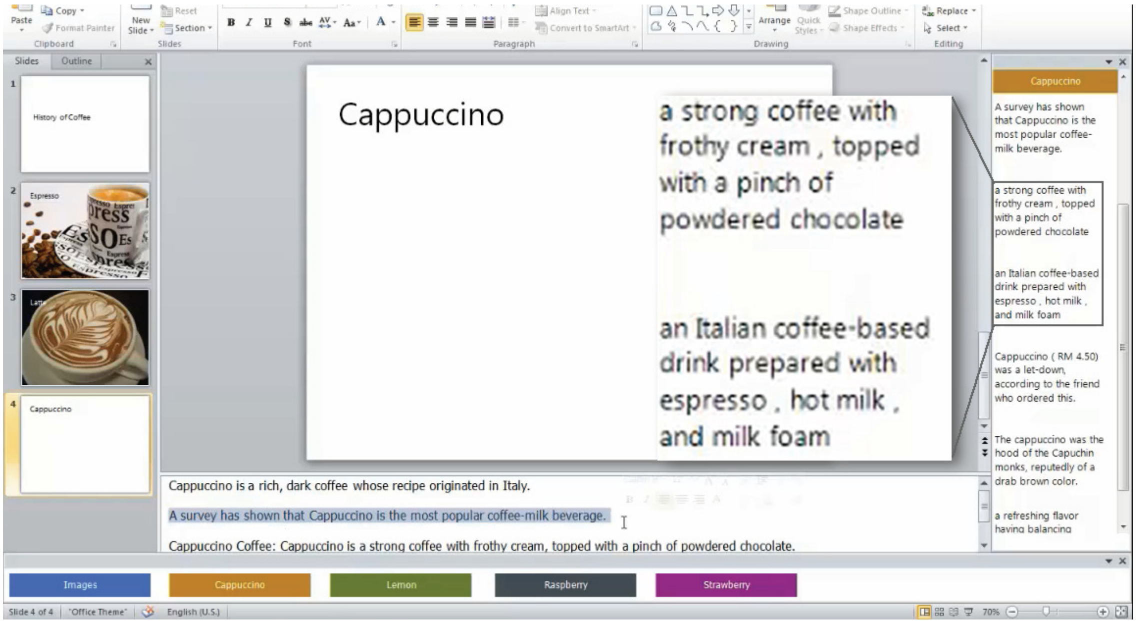


Figure 4.1: SidePoint interface showing concise knowledge items relevant to the slide content ("Cappuccino"). We use SidePoint as a technology probe to examine how peripheral knowledge panels can support presentation authoring

through NeedleSeek [6] (technical details of NeedleSeek are available in [97; 113]).

When the keyword is too specific or too long, NeedleSeek may not offer any result; in this case, SidePoint uses Web search results. Although these results are not necessarily optimal, our pilot studies indicated that people would still want to see some information about such keywords.

4.1.2 Knowledge Panel

After SidePoint collects knowledge from NeedleSeek, it shows the results in a knowledge panel on the right side of the PowerPoint window. These are derived from key sentences describing the keywords. These sentences in general have a structure of [keywords]-[Be verb]-[description]. SidePoint parses the sentences with Splat to extract the subject complement of the keywords as a concise description. It then orders all descriptions by increasing length. In this manner, SidePoint prioritizes descriptions that are quick to browse. Some NeedleSeek results are short facts related to the keyword (e.g., demographic information of a city), which we list first.

Left-clicking on an item in the panel imports that fact or description into the note section of the current slide. Items can thus be stored for reference or copied onto the slide itself. Right-clicking on an item navigates the user to the Web browser and automatically shows literal Web search results for that item, revealing its source. In this way, the system offers a smooth transition from glanceable facts and descriptions to deeper exploration on the Web. When NeedleSeek returns no results, SidePoint simply lists Web search result summaries in the knowledge panel. They link to search results in the browser in the same way.

4.1.3 Concept Panel

In addition to the knowledge panel on the right side of the interface, SidePoint also incorporates a concept panel that runs along the bottom of the interface and shows alternative concepts that can populate the knowledge panel if selected. This panel displays three different kinds of concept. The first is images: clicking on the image concept appropriates the space of the knowledge panel to display categories of image search results (photos, monotones, graphics, and people). Clicking on any of these shows results of that type; clicking on a desired image automatically adds it to the current slide. Semantically related concepts retrieved from NeedleSeek and alternative nouns and noun phrases from Splat make up the final two concept types. Clicking on any of these updates the knowledge panel results accordingly.

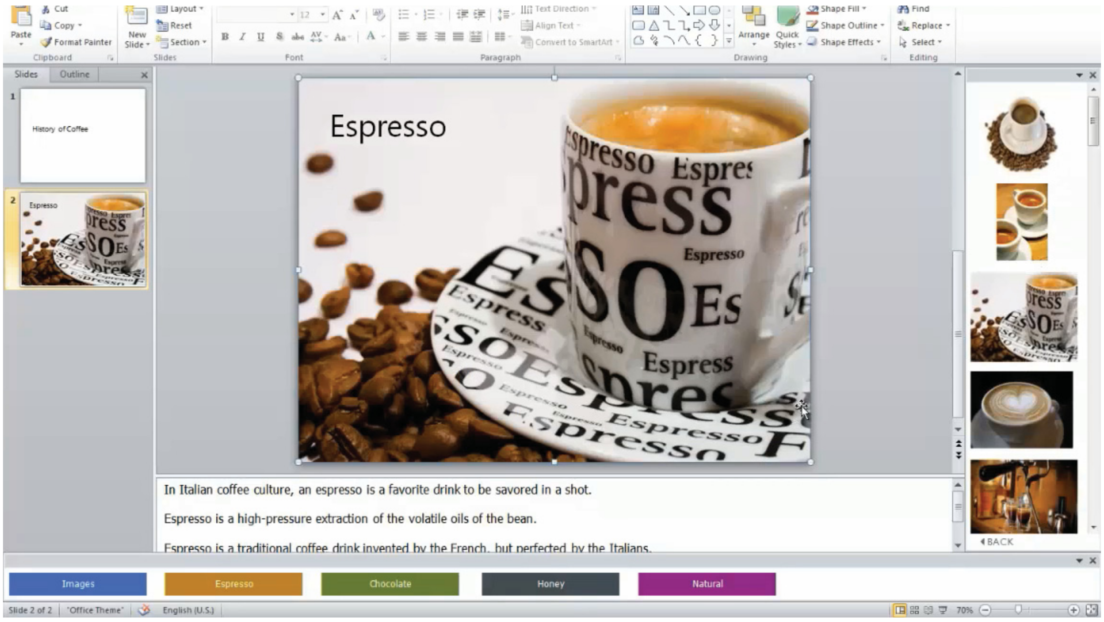


Figure 4.2: Concept Panel for the term “espresso”. The user can simply click an image to import to the slide.

4.2 User Study

We examined the benefits and problems of SidePoint in the context of presentation authoring in particular, how they could satisfy active and latent needs. Our integration of NeedleSeek knowledge allowed us to investigate the satisfaction of both need types and observe how it would support slide authoring. We recruited 12 participants (8 male, 4 female, $P1-P12$) of age 21-44 (mean 24), with 5+ years PowerPoint experience and backgrounds in engineering, design, and finance, from our laboratory. None of them knew the SidePoint system before the experiment.

4.2.1 Procedure

The main task in our experiment was to build a 5-minute presentation about their life (e.g., places lived, schools or universities attended, and countries visited). The experimenter first guided the participant through an example presentation showing SidePoint functionality, before participants progressed through three stages:

- **Plan**. Plan key points to present by typing them in the notes of empty slides; e.g., in a slide on Paris, these might be food, culture, landmarks, or examples thereof.
- **Build**. Use SidePoint while authoring slides. As our focus was to observe the user experience of SidePoint, we did not allow Web search or free browsing.
- **Discuss**. Describe how SidePoint supported both planned and unplanned key points, and discuss how SidePoint affected the overall authoring experience.

4.3 Findings

On average, participants planned 9 key points per 5-minute talk, found supporting information from SidePoint for 5 of them, and incorporated 2 additional items. However, there were large individual differences, shown in Figure 4.3.

$P1$, $P3$, $P5$, and $P12$ had detailed preliminary ideas about what key points to discuss (e.g., weather, culture), finding supporting information for about half of those points. On the other hand, $P7$ and $P9$ decided immediately what to say and wrote detailed scripts up front, and did not use SidePoint often. Uniquely, $P8$ found greater inspiration from suggested items than in her initial planning. We now explain these behaviors by grouping into four themes.

4.3.1 Saving Author's Time vs. Using Authoring Space

Participants positively mentioned the time-saving nature of peripheral knowledge panels. For example, $P11$ commented if you want to add information or add images, it really saves time, referring to how she was pleased to find and use directly

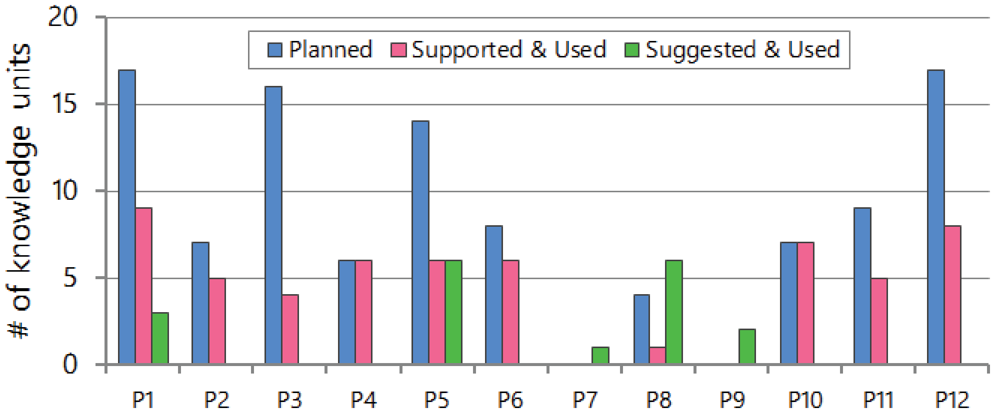


Figure 4.3: Numbers of planned key points and used knowledge items during the user study.

an image of her favorite building. This process was "*really convenient, much better than searching on the Internet, saving to my local desktop, and dragging onto slides*". Along with other participants, she also described the images and text as inspirational: "*this kind of text (description points) is helpful for my presentation. It helps me to explore the topic. I got inspiration from this*".

SidePoint was also described as a beneficial use of the authoring space. For *P*5, "*this design is not distracting, the image search is convenient, and automatically adding a reference into notes is good too*". For *P*8, "*the ambient display design is good, it only takes up a very small space. I don't feel it is distracting*". These descriptions suggest that relatively small embedded knowledge panels can save time and support fluid authoring when they reduce application switching costs and automate otherwise manual actions.

4.3.2 Triggering Implicit vs. Explicit Search

Several participants remarked that SidePoint felt like search without searching. For *P*11, "*when I want to get some information, I dont need to go to the Web browser and search, I can just input something here I dont even need to input really*". She continued that "*If I just describe places I have been, I don't want to go and search them on the Internet. It's too boring. So if you show them* *on this place (the panel), it gives me inspiration, I can see more about it*". In particular, she was inspired by the description of her hometown as "*a beautiful and romantic city with countless stories and fascinating scenery*" and used it on her slide. In another example, *P*2 noted "*I dont have particular goals or expected topics in mind. I just want to input the words and see what information returns, then choose the ones I feel are good for this slide, and then think about what else I may need*". This approach created opportunities to satisfy latent needs, e.g., for *P*3, "*the 'surprising result' moment happened several times, which is really good for me*". This demonstrates the potential value of implicit search to support thought processes without losing attention to slides.

Although participants agreed that implicit search helped their presentation slide authoring, some requested greater control over what would be searched for and when, e.g., for *P*3, "*I want to have a button or something to trigger search when I need it, and to stop or refresh the search as I want*". Satisfying both active and latent needs is difficult, but one behavior in particular linked the two and demonstrated our intended crafting of information scent: "*I didnt find what I wanted from the results, but I did find them by right-clicking and jumping to web sites... hmm... just by right click I could arrive at the content, in that sense it is quite useful*". For example, the description "*Tianjin dining culture is renowned throughout China, and not only the traditional Tianjin snacks, but the cuisines from other regions of China can also be found*" linked to a website listing all of the different kinds of such regional snacks.

4.3.3 Finding the Right Facts vs. Descriptions

Some participants appreciated knowledge presented as facts; e.g., for *P*5, "*the fact data is good... it basically provides most of the information I want to introduce my city*". For others like *P*6, however, "*the facts are not interesting. I don't think people will be interested in the population or GDP of my hometown*". This highlights how the value of such facts can be subjective and relative to the specific purpose of the talk, whether to inform or entertain.

Many participants found more value in the description points. For *P*8, "*the descriptions arent bad and did suggest some useful aspects I didnt think of*".

During the study, he read through the text points while repeatedly commenting that "*this is good*", "*this is right*", and "*this is very true*". He particularly liked the precise and succinct descriptions of his research area, e.g., "*the study of how knowledge about the world can be represented and what kinds of reasoning can be done with that knowledge*".

4.3.4 Exploring Related Concepts vs. Attributes

Some participants appreciated the display of related concepts in the concept panel, e.g., for *P*1, "*the related concepts are useful. I want to show places to travel when I introduce my city, and it suggested nearby cities*". *P*9 remarked how the system was "*smart*", offering Graphic Design after she had entered Industrial Design on her slide.

However, participants also found that the related concepts offered by NeedleSeek were often too tangential to be relevant in a talk about a particular topic. A suggested alternative to related concepts is to reveal more of the attributes of the current concept. For *P*11, "*it's better if they have any kind of filter, like if I input a place, I can chose people or scenery*".

4.4 Design Considerations

Based on our results, we propose four design considerations for peripheral knowledge panels that support slide authoring:

- Trade interface space for time saving until the complexity of the transformed task requires tool switching, providing information-scent-rich entry points for tool switching.
- Show implicit knowledge results for visual and verbal inspiration, but appropriate the display space for search results when information needs are expressed explicitly.
- Show a variety of knowledge types at the beginning of the task (e.g., facts, descriptions), but improve relevance over time by adapting to author behavior and content.

- Support attribute-driven idea refinement through both structured exploration of known attributes and dynamic combination of cues to test for possible attribute relationships (e.g., by searching "*bullet of title*").

Together, these directions characterize the ways in which peripheral knowledge panels can become pivotal in their support of presentation authoring helping authors to pivot between different applications, search modes, result types, and knowledge levels in the increasingly semantic Web.

4.5 Chapter Discussion

The core design problem of the information age is to increase the relevant information encountered by a user as a function of the amount of time they invest in interaction [90]. Our SidePoint prototype aims to maximize such exposure to relevant content during time already committed to the primary task of authoring presentation slides. Through our study of SidePoint as a technology probe, we have shown that peripheral knowledge panels have the potential to satisfy both active and latent information needs in ways that transform presentation authoring for the better.

Future work should investigate peripheral knowledge panels in broader contexts and over longer timescales. Comparative studies with other intelligent search tools would also help identify the characteristic benefits and issues of peripheral knowledge panels. In addition, the present study only examines the effect of peripheral knowledge panels on specific presentation topics. Our criterion for the topics was something the user was familiar with but does not necessarily remember all details about. This type of topic frequently generated both active and latent information needs as seen in our experiment. Future studies should examine how frequently other kinds of topic raise active and latent needs and how this affects the use and value of peripheral knowledge panels in general.

Chapter 5

UbiAsk: Crowd Powered Image Based Mobile Social Search

Mixing mobile platforms and the crowdsourcing model potentially offer vast resources for computation. For instance, today, in most urban areas it is common that people allot a large amount of time for commuting or waiting for various events. Their time is actually fragmented into numerous small pieces of time and most of them are occupied with meaningless activities. We believe that, in this case, the crowdsourcing model provides a winwin solution for better use of this time by engaging people through their networked and ubiquitous mobile devices. However, a majority of current human computation and crowdsourcing systems (e.g., Amazon Mechanical Turk) are passive services that are using worker-pull strategy to allocate tasks, and require relatively complex operation to create a new task. Consequently, such systems fail to adapt to the mobile context where users require simple input process and rapid response.

We explore crowdsourcing platform designed for mobile users. Our design principle is simple: build the system around existing social media platforms the popular services that inherently have the culture of sharing and participation, and are already well known and used by a large and growing number of users. In this chapter, we practice the platform in the context of one specific application area image-based mobile translation/search. This refers to camera phone applications that attempt to solve the problem of translating text written in an unfamiliar script. This kind of a system is particularly useful for travelers and short-term residents in a foreign country. Ordinary digital pocket translators are useless if the user is unable to input the text they see (e.g., Japanese or Chinese text for Western visitors). Image-based mobile translation/search systems typically employ optical character recognition (OCR) algorithms to extract text out from image, and apply machine translation (MT) technologies to translate the text. Several systems like this have been proposed during the past decade [46; 92]. However, only a handful of them have gone beyond pilot development. Even the state-of-the-art in this field, such as Word Lens[1] or Google Goggles[2], demonstrate very limited performance in real-world situations (i.e., complex background, dark environment, blurred photos, irregular fonts, size or formats, etc.), especially for non-Latin scripts like Japanese and Chinese.

To solve the problem, we present UbiAsk, a social media crowdsourcing application built on top of existing social networking infrastructure. UbiAsk provides translation services and situational advice to mobile users in unfamiliar environments. Instead of applying machine algorithms, we draw on the power of ordinary people in the cloud via social networks to solve the difficult computational problems such as image recognition and text translation. Since the workload of each task in the image-based mobile translation/search service is lightweight enough to be described as a micro-task, the tasks are perfectly suitable to be distributed to large groups of casual workers.

In UbiAsk, users can issue requests via several channels that use a common API. Native mobile applications and email are the currently implemented channels. The requested task is pushed to a community of voluntary local experts in the form of an open call via different social media platforms (Twitter, Facebook, etc.) and email. The crowdsourced result data are not only returned to requesters but also visualized on location-based social mapping and augmented reality (AR) platforms (e.g., Sekai Camera[3] and Ushahidi[4]). This gradually results in an information pool that constitutes a public good.

[1] http://questvisual.com/
[2] http://www.google.com/mobile/goggles/
[3] http://www.sekaicamera.com/
[4] http://www.ushahidi.com/

5.1 User Centered Design

In this section, we present the design of the UbiAsk platform. Parallel to the technical development, we have studied the application possibilities of UbiAsk by conducting a formative user study in the form of proof-of-concept simulation. Future UbiAsk services have been illustrated to the potential users. The users' feedback has been analyzed to identity the users' requirements to the UbiAsk system. Thus, we have been able to influence the technical developments well before the actual application development stage. Such a user-centered design process ensured that the forthcoming platform could support the features and characteristics needed by the end users.

5.1.1 Initial Design

The platform makes use of client-server architecture. There are two different types of clients: the users who make requests are called requesters, and the crowdworkers are named local experts. This feature differentiates UbiAsk from other typical server-client systems. The UbiAsk server plays a proxy role. It receives requests from client users, assigns these tasks to appropriate local experts, and finally forwards the answers to the original requesters. Figure 5.1 illustrates the basic workflow of the proposed translation model and a detailed description of it is given below:

- A client user makes a request by taking a picture using a mobile phone's camera, and submits the image to the server. In addition, a short text message should be attached to the photo in order to clarify what exactly the requester wants to know. This short message extends the application possibilities, since users can now ask questions with no direct relationship with any object in the image. Location information and time will be automatically attached to the request, although the availability of such context information may depend on the client's terminal's functionality (e.g., if the locationing module is embedded). Such context information, together with a work user's profile, is useful for assigning tasks appropriately.

- For enhancing the response time, each task is assigned to multiple local experts simultaneously.
- The original request is sent to translators via email. Translators are encouraged to reply in “key words” or “short message” style.
- For the response time considerations, the first answer to be received from local experts is forwarded to the requester immediately. For the rest of the replies, the client user can set a maximum waiting period and reject any responses received after that.
- Eventually, the client user receives the results.

5.1.2 Formative Study

For this platform to work, we have to know if the translators are able to deliver the desired responses to the requesters. Therefore, we designed a qualitative experiment in laboratory conditions to answer the research question of how should the users questions be presented to the local expert in order to provide the preferred results for the user? We held a series of meetings for discussing the experiment design. Participants included Japanese and foreigners (that can be seen as our potential users) from different background areas such as technology, design, economics, user experience, and psychology. The original intention of the project has been to design a human-based image translator. However, through discussions with potential users we found out that what their requirements are more than a simple translator. Instead of just knowing the semantic meaning of the words, users are much more interested in a service that can answer their questions related to the photo, which is more like an image-based mobile social search across languages and cultures. Based on this finding, in the later meetings we have improved our concept idea and extended the design from simple translation to a mobile social search service. Eventually, the experiment workflow has been realized as follows.

First, we collected around 100 sample images (took by mobile phone) with corresponding questions from five foreigners who were currently staying in Tokyo.

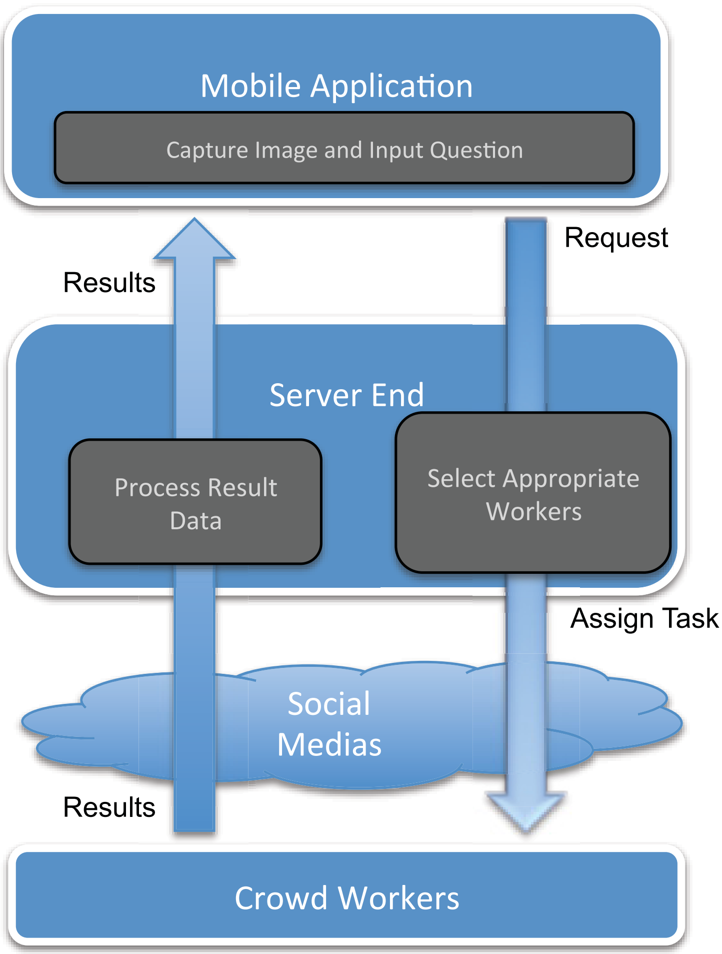


Figure 5.1: System basic workflow.

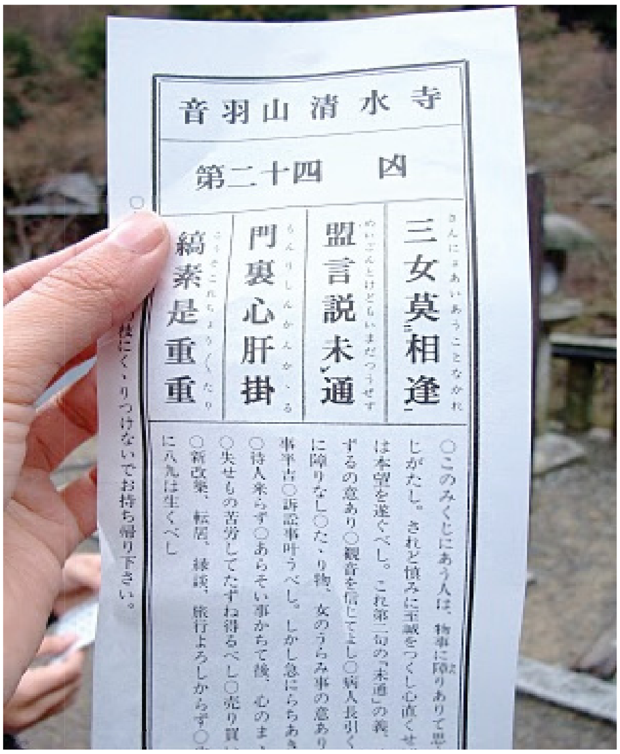
Together with end users, we selected 15 characteristic cases from different types of requesters (or we could also say, from users who had different needs). Then, we interviewed the photos providers, questioning in what situation they took the photo and what kind of answers they were expecting. In the next stage we sent the requests to a group of seven invited local experts (all male, mostly students from School of Science and Engineering, Waseda University, Japan). After receiving the answers, we interviewed the translators for collecting their feedback on the usage of the service. Finally, we compared the results from the translators with the expected answers of the requesters, examined whether they matched and discussed the possible reasons for mismatching.

Figure 5.2 shows one example that was observed in the pilot study. From the collected questions, we found out that in many cases (i.e., approximately 47%) requests were actually driven from curiosity rather than real problems or troubles the visitors were facing (i.e., approximately 53%). A typical situation is when the users obtain partial information for something interesting but are unable to figure out the whole information that they want to know. For example, this requester knows what this piece of paper is, but cannot understand the exact information in the content.

5.1.3 Impression of Pilot Study and Redesign

The quality of the results in crowdsourcing is a hot issue, and has been discussed widely recently [11; 53]. In this study, we found that the way to input requests and answers is an important factor, which affects not only the usability but also the quality of the outcome. The simplest way of making a request is just sending the picture directly, but it can hardly be an option because translators can be easily confused about the real purpose of the request. In other words, client users must clarify what exactly they want to ask by adding more information. On the other hand, the translators often use English as their second or third language and thus, they may misunderstand the question if the text is too complicated.

There are different ways to lower the possibility of misunderstanding. One way is to limit the complexity of the messages by e.g., instructing a client user, setting maximum size of message, etc. Moreover, depending on the question,



Q: ***What divination result I got?***

A1: the oracle says your fortune is 'bad luck' Knock on wood nearby as soon as you can.
A2: bad fortune :(there is a description at the bottom of the picture. But it takes long time to translate :(

Figure 5.2: One sample of formative study.

the local experts may need to reply in different ways in addition to the text, e.g., for questions like "*which button should I click?*" or "*which one will you recommend?*", it is better to provide them an image editor thus they can simply circle the corresponding part in the picture rather than giving a description by words. However, a local expert (as a human) always makes mistakes no matter how perfect the instructions are. In the study, we observed that sometimes even if the question was clear and simple enough, "bad" replies still appeared due to the fact that some translators began answering before finishing reading the request message. Even worse, we also have to consider the possible existing of a malicious reply. As single reply can hardly be trusted, another possibility is to provide multiple results to the client user. The users can compare the different replies by themselves, and make their own decision (e.g., choose the majority answer). Nevertheless, if we consider the response time, this approach might be expensive.

There is a third solution, which can be seen as a com- promise between the above-mentioned methods. We can add a proofreader (as Figure 5.3 shows) to verify the correctness of the answer and to prevent from malicious replies. Moreover, the task of classifying/tagging images can also be assigned to the proofreader, for maintaining a more valuable results database. Depending on the different clients types, there is a trade-off between the accuracy and the response time of the reply. For requests that need immediate answer, timeliness is the key factor concerning clients' quality of experiences. We may want to skip intermediate stages and directly forward the answer from the first translator to the user. On the other hand, for waitable type of requests, the proofreading or multiple answers should be a strict requirement. In general, we advocate the appropriate use of different request processing strategies, depending on the users' request types.

Through the experiments, we also confirmed the inherent disadvantage of the computer translation. Even if we assume the machine-based image recognition and text translation technologies can always provide perfect outcome, they still have no chance of offering the desired answer for such kind of services that demand higher-layer information.

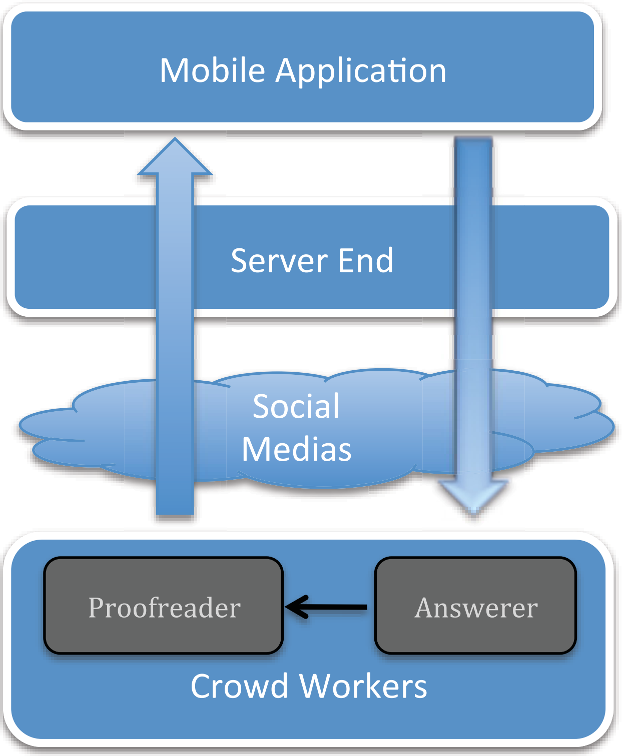


Figure 5.3: System basic workflow with additional proofreading phase

5.2 Prototype implementation

After several rounds of design iteration including formative evaluation with potential users in [72], we implemented a prototype version of UbiAsk. Figure 5.4 illustrates the basic system structure of UbiAsk. Requesters can quick create a task and upload to server by utilizing diverse mobile applications. Proxy server pushes task to appropriate workers via regular emails as well as different social media platforms. When a task accomplished, we argue that the questionanswers pairs are not only valuable for original requester but also be able to beneficial to the broader public, thereby, the location-based result data are also visualized with social mapping tool or social augmented reality services (Figure 5.5).

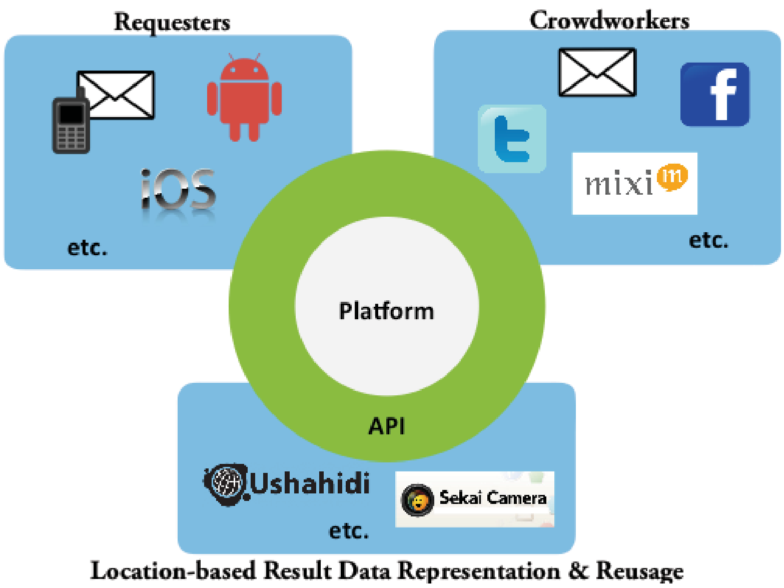


Figure 5.4: UbiAsk system overview

The proxy server was built as a REST style web service with open API to clients. A task assigner was developed as the component with responsibility for assigning task to and receiving answers from local experts. The task assigner also connects to the social media platforms APIs. All back-end programs are implemented in Java. Requester can use an iPhone application (see Figure 5.6) or mobile email to make a request. The main interface of the iPhone application consists of three main parts: an interface to show a list of existing requests, an camera display view to take picture of what requester want to ask, and a text editor to input the short question. The uploaded image was saved at the server end. The link of photo and corresponding question was pushed to local experts via email and Twitter. The local experts can submit their answer by simply replying the email or tweet. The maximum living period of a request was 12 h and server will reject any later answers. All result data were visualized on an Ushahidi-based interactive map (see Figure 5.7).

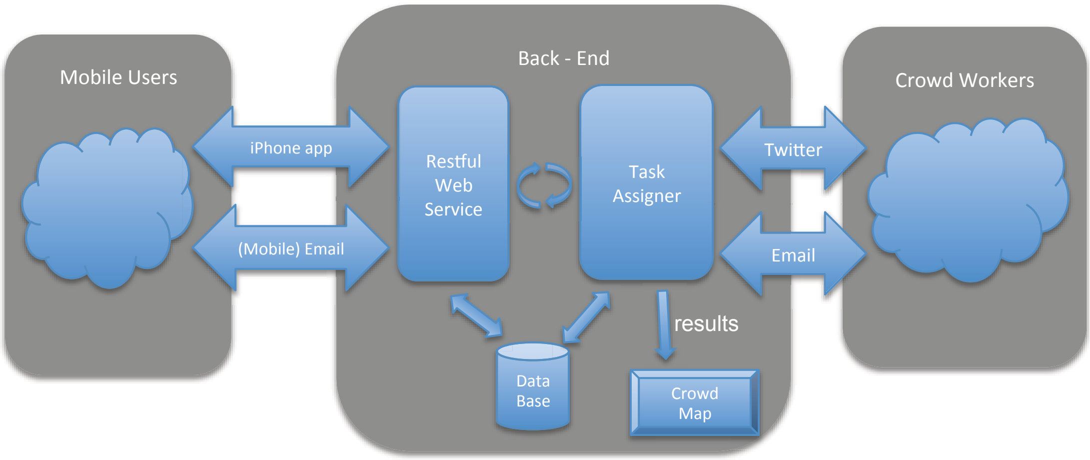


Figure 5.5: UbiAsk prototype

5.3 Field User Study

After UbiAsk was implemented, we launched a field user study with 55 end-users (19 requesters and 36 local experts) for 6 weeks[1]. On the requester end, the user study is designed to find out if UbiAsk could provide answers of satisfactory speed and quantity; on the crowdworker end, the relevance of the activating participation to different motivational mechanisms is examined. Also, the usage data are gathered to study the typical usage pattern.

[1]From January to March 2011

Figure 5.6: iPhone application for requesters

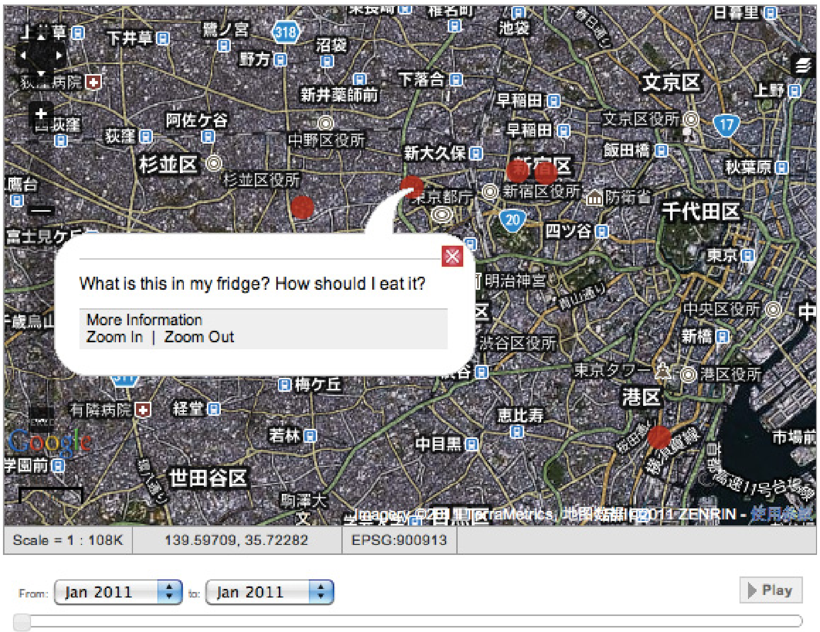


Figure 5.7: Ushahidi-based interactive map for visualizing results

5.3.1 Participants

The 19 foreign requesters were mainly recruited from a total of 13 popular international travel forums such as Lonely Planet thorn tree[2], Tripadvisor[3], Fodor's Travel Talk[1], etc. The majority of the requesters were short-term travelers, the exceptions included Japanese returnees, foreign students, foreign employees, foreign housewife and visiting scholars. All, but one, of them were from Western countries (i.e., United States, European countries, and Australia), the only exception was from Southeast Asia. Of the 36 local experts, 17 were recruited from Internet by twitter adverts or posts on local forums such as Yahoo!Chiebukuro[2] (The Japanese version of Yahoo!Answer) - English questions. The rest of the local experts were mainly under- graduate students, graduate students, and staffs from Waseda University that were recruited by emails.

However, truly active participants in the numbers listed above were signifi-

[2]http://www.lonelyplanet.com/thorntre/e
[3]http://www.tripadvisor.com/
[1]http://www.fodors.com/community/
[2]http://chiebukuro.yahoo.co.jp/

cantly less. Initially, 23 travelers claimed that they were interested in the study and agreed to participate, however, we lost contact with 4 of them before starting the user study, and 8 of them never used the service during their visiting period. We re-contacted the missing requesters during the experiment in order to identify the reasons of their absences. Not so surprisingly, most of them said it was because overseas 3G/4G usage fee is too expensive, and they could not find any free Wi-Fi spot when they want to use the service. Although alternatively it was possible to save the photo and send it later when there was Wi-Fi available (e.g., in their hotels), the users did not do so. Thirteen local expert testers were likewise the "lookyloos", who never reply any question during the user study. One possible interpretation is the experiment duration happened to be during the exam season for the Japanese schools and the student participants might have been caught up in examination preparation and final reports writing. Overall, this user study reported 37.5% no show rate, which was considerably higher than the quantitative user studys average number of 11% [7].

5.3.2 Procedure

For collecting demographic information, all participants completed a pre-test questionnaire before the requesters left their countries for Japan. The iPhone application with the usage instructions and the information of the user study were emailed to the requesters before their departures. When requesters arrived in Japan, they were told to use the service freely.

5.3.3 Overall Performance

In the pre-test questionnaire, two important questions were asked: the users expected response time, and the users bearable response time. Based on the answers, we can further identify three time periods with regard to user satisfaction: we assume a user will be satisfied if the first answer comes before his/her expected response time; a user will be unsatisfied if the first answer comes after his/her bearable response time; and it is acceptable if the first answer comes between the two time points (see Figure 5.8 for an example). The survey result is shown in Figure 5.9, if the response time is less than 10 minutes, we can satisfy all users.

It is beyond our expectations that even if it takes 30 min to have the response, only 25% of user will be unsatisfied. But if the waiting time becomes more than 1 hour, it will be unacceptable for most of the users.

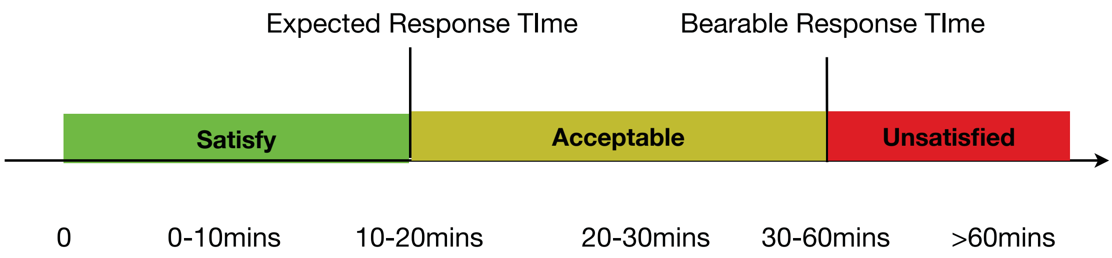


Figure 5.8: Sample of one requester's requirements on response time

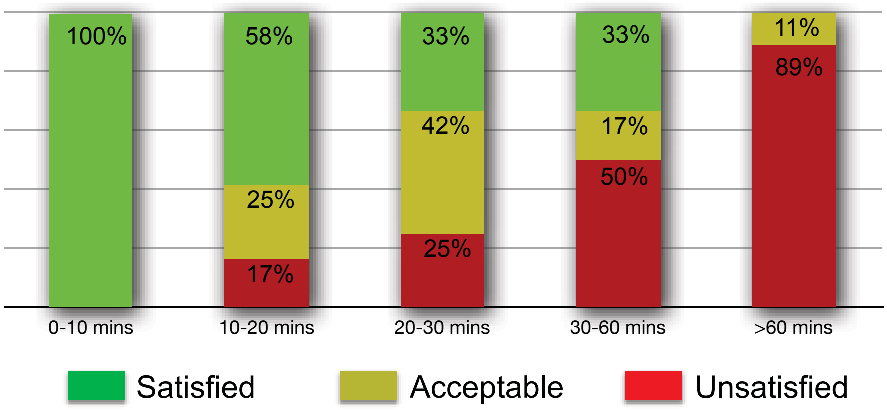


Figure 5.9: Requesters' overall requirements on response time results

5.3.4 Incentive Mechanisms Comparison

In terms of participation motivation mechanism design, local experts were randomly placed in three incentive groups, among which there were two experimental groups and one control group.

- **Group A** Derived from the Game-With-A-Purpose concept, a location-based mobile social game was designed and implemented as participation incentive. The main interface of the game (see Figure 5.10) was a Google map-based real world map, which was divided into non-overlapping hexagons. The goal of the game is to conquer territories. Every hexagon has one owner, who is the player with the highest number of "task-done" in the area. In other words, crowdworkers need to compete with each other to get the "local expert" title of an actual location on the map.

- **Group B**. A simple feedback mechanism, social psycho- logical incentive that was known to be effective, was applied: when a local expert provided an answer, the system will rapidly reply him the number of existing answers of that question and previous answers content.

- **Group C**. Control group, no additional motivational method was applied.

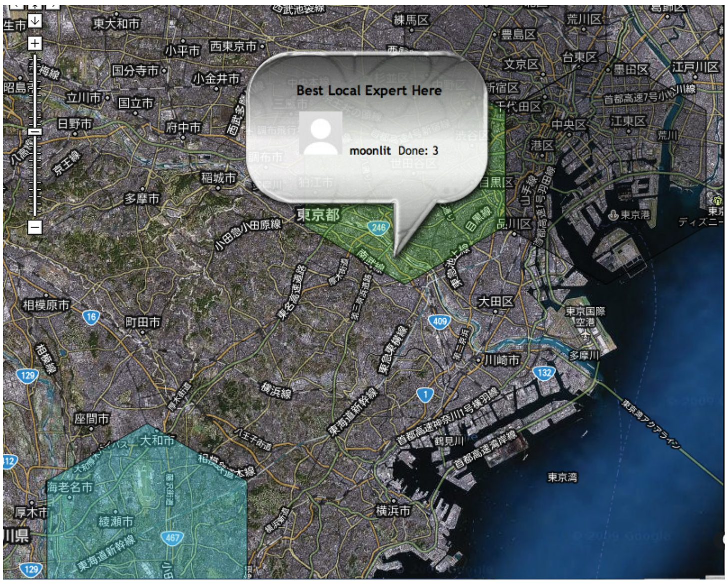


Figure 5.10: The interface of the location-based ranking game

For performance measures, we instrumented our prototype to log the timestamp and the question/replies ratio. To measure user satisfaction, ease of use, and overall experiences, we administer a post-test questionnaire with Likert scales. The questions covered typical usage patterns, content's quality, features preferences, likes, dislikes, and suggestions.

5.4 Results

In this subsection, we present the main findings from the user study.

5.4.1 Performance Results

In all, the system recorded 180 interactions, covering 33 questions and 147 answers. We expected to see a much greater number of requests. However, eventually there were only 11 (58%) requesters that submitted their questions to the system. On the other end, 23 (64%) local experts answered at least once question. Of the 147 answers, the local experts that recruited from Internet provided 93 of them. The seven most active local experts accounted for nearly 70% of the answers. The requests are relatively equally distributed across the day expect early in the morning and mid-night.

Figure 5.11 shows the overview of the response time (the first answer). Approximately 50% of the questions were answered within 10 min, and 27% of them were responded within just 5 min. Three-quarters of the first answers arrived within half an hour. Only 6% of the answers were arrived after 60 min, and most of them were asked during midnight or in the morning. The 9% of the requests were never being answered; the main reasons were bad timing (e.g., local experts were busy) and boring question's content (e.g., translation or explanation of a long text).

According to the pre-test questionnaires result, we set a strict rule for the response time: if the first answer of a question arrives after 60 min, we consider this question has no answers. Based on this rule, we show a time-of-the-day breakdown of the average response speed, average response rate, and answers per request in Figure 5.12. Generally speaking, the overall performance is reliable: except mid-night and early morning, nearly 88% of the requests could be answered in approximately 10 minutes, with more the 4 answers per request.

It is true that results data were collected under an experiment setting and the end-users may behave differently if the service was deployed in the real world environment. But it is also necessary to point out that the system was intentionally designed without any additional incentive mechanism due to the fact that the incentives comparison is one of our research questions. A number of the fundamental components of the Q&A systems motivation module were not provided, e.g., global view of all questions status, contribution histories, others contributions, etc. In addition, the above-mentioned no-show-rate also indicates, this experiment was not intended to access the quality of data that can be achieved in optimal (strictly coordinated) conditions.

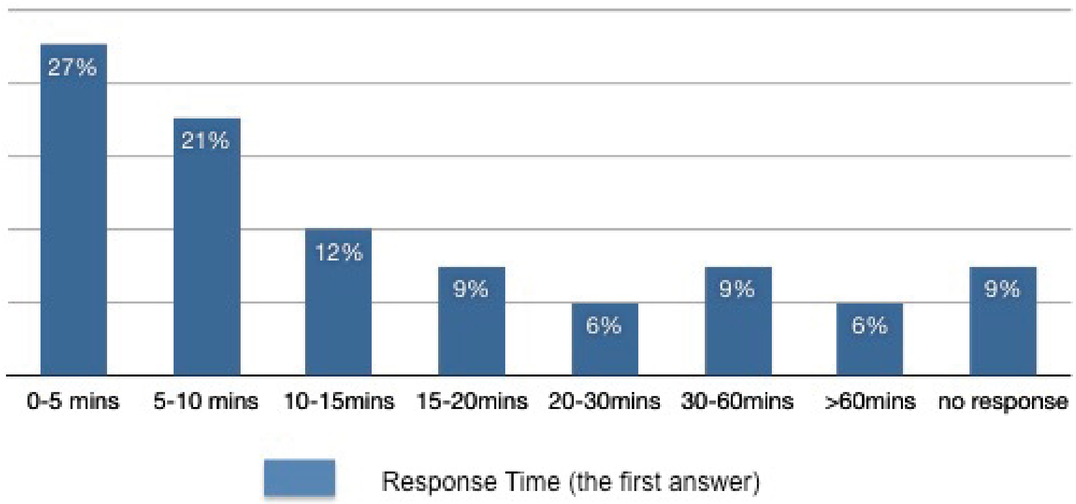


Figure 5.11: The performance results: response time

5.4.2 Incentives Comparison Results

The incentives comparison results were more complicated. Figure 5.13 (I and II) shows the response time and number of answers by groups. Local experts from Group A demonstrated a much more active engagement of the system, 67% of the first answers were provided by them. In terms of quantity, more than half of the total number of answers were likewise produced by people from group A. In the meanwhile, the control group produced a better outcome than group B. However, the web access logs did not support these findings. The access history reveals that the page of the location-based ranking only drew participants attention in the beginning stage of the study. The access number shows an obvious decrease over time, and eventually dropped to zero per day in the last phase of the experiment.

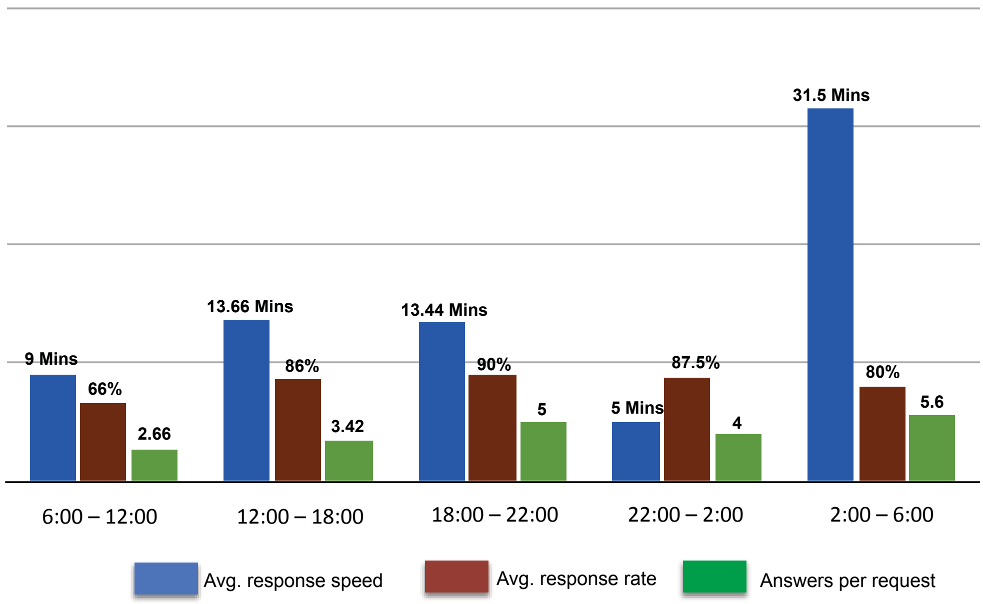


Figure 5.12: The time-of-the-day breakdown of the average response speed, response rate and answers per request

Moreover, in the post-test questionnaire, there was no participant that claimed that he/she answered those questions because of the location-based game.

To better interpret this phenomenon, we further analyzed the top active users. We found four of the top five most frequent local experts were from group A, and the exception is from group C. Based on their replies in the questionnaire, all of them have demonstrated strong intrinsic motivations, i.e., "*I want to help the people in trouble*" or "*I want to introduce Japanese cultural to foreigners*", which may have a stronger impact than the extrinsic motivations we provided.

On the hand, however, if we measure the user's participation as the number of users in the system (i.e., users who answered at least one question), and give equal weight to every user no matter the amount of the answers they provided (i.e., minority highly active users data will have less power of influence to the final comparison result), the result will be more explainable. In this case (see Figure 5.13 III), we observed the same number (38%) of active users were from group A and B, and fewer users (24%) were from the control group.

Overall, we argue that although the most frequent participants might be more motived by intrinsic incentives, the effectiveness of the designed extrinsic incentives to the rest of less self-motived users was still verified. Nevertheless, based on the results, we could hardly come to the conclusion that the proposed game-based incentive has a greater impact than the social psychological incentives. We believe that it is not only because the game may not very interesting for the local experts, but also because the expected fierce competition did not happen due to insufficient requests' number.

5.4.3 Types of Requests

Three main types of questions were identified: translation, explanation, and cultural (see Figure 5.14). The number of answers of each type of questions is also shown in same figure. As most of the requesters cannot speak Japanese, the majority of the requests were translation related, including pure translation questions (42%) and requests for to explanation of a thing or a piece of text in English (36%). But as for local experts, results clearly shows they were more interested in answering explanation questions (44%) than pure translation questions

(39%), because pure translation can be extremely boring work, e.g., translating the meaning of every button on a controller or a washing machine. In contrast, providing higher layer information is much more interesting. When asked the main reason for not answering a question, one of the top five active local experts also said:

> "*Sometimes the entries were not question, actually they are translation request.*"

Cultural related questions were also popular. It may be quite difficult since normally such question requires answerers with some extra domain or local knowledge, but it is more meaningful for both parties, sometimes it can even be a quite interesting interaction:

> "*Q: Why all toilets in Japan are with brand TOTO?*
>
> *A: I searched the topic on the Internet. I found that TOTO used a beautiful woman as their advertisement while the competition company used a gorilla. That's why TOTO toilet is widely used.*"

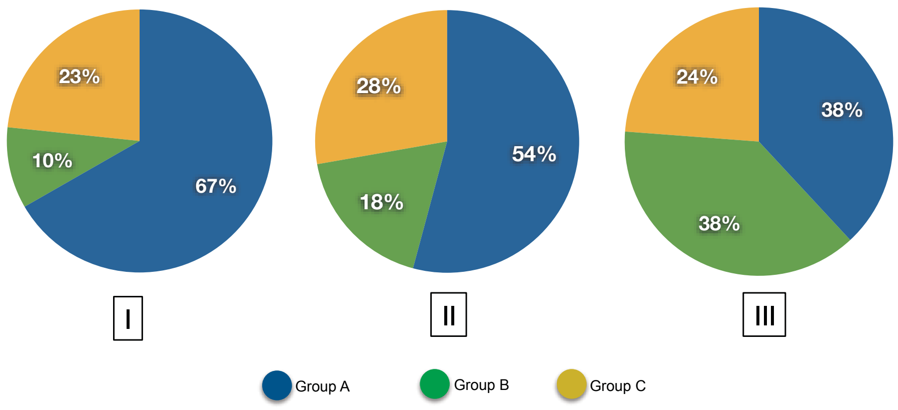


Figure 5.13: I Response time by groups; II number of answers by groups; III number of active users by groups

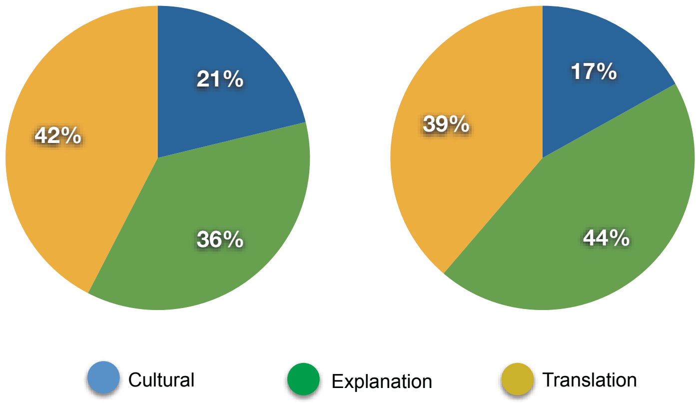


Figure 5.14: Left: number of question by types. Right: number of answers by types

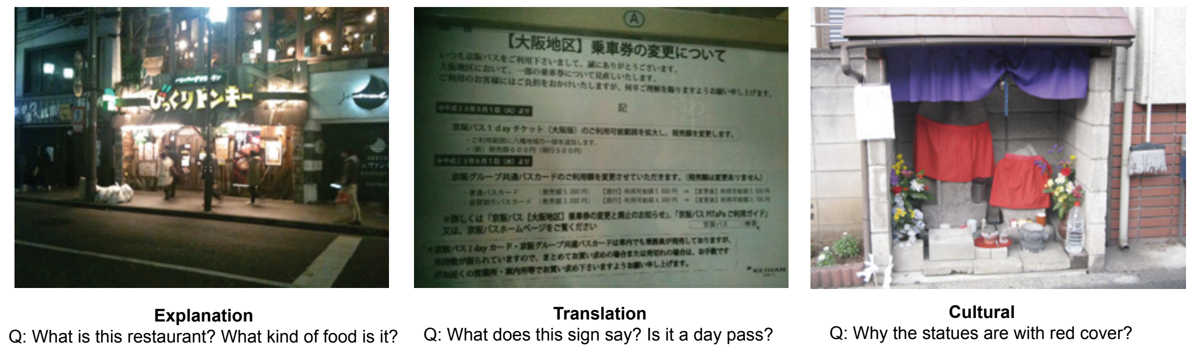


Figure 5.15: Examples of different types of requests

5.5 Implications

In this section we discuss the main implications of the study.

5.5.1 Users Desire Rich Information Instead of Translation

Real-time image translation is one of the classic application ideas for the mobile argument reality systems. But how useful it really is in the real world situations? For example, one popular usage scenarios of such system are the translating of important signs, but one of our pilot testers has said:

> "*No, I can almost always make a reasonable guess of the meaning of such signs, just based the contextual information, or simply follow what other (local people) do.*"

Based on the user study results, we confirmed that the majority of the users were actually more interested in higher-level information rather than literature translation. That suggests even if the OCR and MT technologies can eventually always provide 100% correct results, they still cannot meet most of the users needs. Travelers without language skill normally tend to ignore most of the information surrounding them, and are only concerned about the information directly related to their activities. Most of the needed information is a rich explanation (or the real meaning) of a text rather than simple translation of the semantic meaning, e.g., in the case of a dish, "*pork, spicy, famous Chinese food*", is a better answer for the requester than "*Huiguo rou*" (the actual name of the dish) as far as understandability is concerned.

5.5.2 Image Q&A versus Text Q&A

Participants were asked their preference on picture-based questions and text-based questions in the questionnaires. A quite even number of users voted to "*depends on the situation*" and "*picture-based question*", only one answerer preferred "*text-based question*". This results well support the common setting of the mainstream on-line question answering services, have uploading picture as an optional choice.

In the context of mobile application, we believe that the picture-based question is even more useful and important. The questions people asked in the mobile context are often directly related to their current situations and the surroundings, which mean there is almost always something for them to "picture", and a picture can normally express a situation a lot more easier than explaining it in a text. Besides, via mobile devices interface it is normally harder to input text compared to computer keyboard, and thus even if user can describe the context in text, it is still easier to just attach an image.

Moreover, our application involves users from different countries using different languages. In many cases, both the parties communicated using their second or third languages. Hence the picture also becomes a very important component of the quality assurance mechanism to reduce the misunderstanding between the asker and answerer.

5.5.3 User—User Communication

Our current design does not involve any means for establishing a direct link between translators and requesters, but the necessity of such a communication link is worthy to be discussed. From the study results, we noticed a trend of requiring worker-to-requester communication. When local experts cannot comprehend the meaning of the questions, some of them wish to confirm what they have understood with the requester. On the other end, the requesters may need to ask further questions related to an answer they received. In the user study design, we actually took this issue into consideration. We provided the answerers email address or Twitter id together with their answer, but did not instruct the requesters to use the contact information or embassy on this design. Via post-test questionnaire and unobtrusive observation, we found none of the requesters further contacted the answerers, although in the questionnaire many of the requesters expressed their desires to ask further questions related to the answer and to thank the answerer. But there is one answerer tried to communicate with requester via his answer:

"*I think that attached picture is wrong. Please attach again.*"

However, building a communication link brings obvious drawbacks as well. Serial and continuous questions heavily increase one local experts workload, which is against our original intentions to outsource micro-weight tasks to a large number of work users. When single task becomes heavier, active user participation and engagement likewise become much more difficult to achieve.

5.6 Chapter Discussion

In this chapter, we introduced a mobile crowdsourcing platform built on top of existing social and mobile computing infrastructural. UbiAsk, a mobile system for sup- porting foreign traveller by involving local user in the cloud to answer their image-based queries in a timely fashion, have presented as one case study. We have described the user center design circle of UbiAsk and the findings of a qualitative formative experiment. Moreover, we presented the results of six weeks quantitative field study of UbiAsk.

The user study results demonstrated a reliable overall performance on response speed and response quantity: half of the requests were answered within 10 min, 75% of requests were answered within 30 minutes, and on average every request had 4.2 answers. Especially in the afternoon, evening and night, nearly 88% requests can be answered in average approximately 10 minutes, with more the 4 answers per request. We also investigated the participation motivation of crowd-workers. We found the top active users were more driven by intrinsic motivations rather than any of the extrinsic incentives (i.e., game-based incentives and social incentives) we provided. But the effectiveness of designed extrinsic incentives to the less self-motived users was still verified. However, based on this study we could not come to any conclusion that can suggest the comparison result between the game-based incentives and the social psychological incentives. In this section, we discuss the implications of our findings.

Existing machine-based image search or translation applications (e.g., [40; 64], Word Lens and Goggles) provide very limited performance in real world conditions (such as complex background, dark environment, blurred photos, irregular fonts/handwritings), and cannot satisfy users who want to ask concrete questions. Whereas the proposed crowdsourcing-based approach showed the potential to deliver reliable service under above-mentioned situations and answer users concrete questions on demand. Also, most of the existing social search or human search system (like [50; 76]) are not able to gratify mobile user whom requires an immediate response. UbiAsk system, on the other hand, is designed for mobile context thus can deliver rapid answers in (almost) real time. Moreover, in contrast to paid mobile crowdsouring systems such as [37; 111], UbiAsk demonstrated a reasonable good performance by engaging with (free) active crowd workers from existing social media.

Chapter 6

MoboQ: Temporal and Geo Sensitive Question Answering via Social Media

Since the birth of the Ubiquitous Computing (UbiComp) vision, one of the key challenges is how to extract context information from the physical environment. Without such information, UbiComp applications cannot provide genuine context-aware services. Recently, the design concept of a human-in-the-loop has been introduced to environment sensing, which has resulted in novel concepts such as Human Probe Sensing, People-centric Sensing, or Participatory Sensing [23; 26]. These concepts refer to the design patterns of utilizing sensor-equipped mobile phones for sensing and monitoring urban areas, which takes advantage of the rapid growth and ubiquity of today's smartphone users. The capacity of these machine sensors is still limited to gathering somewhat low-level physical environmental data, e.g., speed, temperature, and pressure, but for context-aware applications, it is important to also have higher-level information, for example, human activity, group emotions, and social environments.

Inspired by the notion of Human Computation [105], we extended the original idea of Human as Processor to Human as Sensor, where human users become part of sensor networks and report local information around them using portable devices. Using such an approach, a system can collect information that is difficult (if not impossible) to obtain by machine sensors, thus offering the ability to generate a richer contextual model. Additionally, we argue that existing social networking sites can potentially provide a free candidate pool for such human-powered sensing. Recently, social media platforms, including microblogging services, are opening their resources to third parties via Application Programming Interfaces (APIs). As a result, other applications, such as Location-Based Services (LBS) systems, have started sharing their users' activities through these social media platforms. Consequently, microblogging services have become an information hub of users daily experiences, where people publicly share their activities such as where they are, what they are doing, and whom they are with. In China[1], for example, the major LBS applications Jiepang[2] and Dianping[3] both synchronize their users check-in data with the microblogging service Sina Weibo[4]. This phenomenon creates a unique opportunity for identifying who is likely to be at a given location at any specific moment by analyzing only their publicly shared information.

We present MoboQ, the location-based real-time question answering service that is built on top of a microblogging platform. In MoboQ, end users can ask location- and time-sensitive questions, such as whether a restaurant is crowded, whether a bank has a long waiting line, or if any tickets are left for an upcoming movie at the local cinema; these are questions that are difficult to answer with ordinary Q&A services. MoboQ analyses the real-time stream of microblogging service Sina Weibo, searches for the Weibo users who are most likely to be at the given location at this moment based on the content of their microblog posts, and pushes the question to those strangers. Note that the answerers in this system are Sina Weibo users, not MoboQ users, and might not even be aware of the existence of MoboQ. This design takes advantage of the popularity and furious growth rate of Weibo[5], which provides us with the confidence to foresee that microblogging users will be regularly available at any reasonably popular Point-of-Interest (POI)

[1]The system presented in this chapter was deployed in China
[2]http://www.jiepang.com/
[3]http://www.dianping.com/
[4]http://www.weibo.com/
[5]According to companies official report, Sina Weibo gathered more than 300 million users in 30 months. Tencent Weibo reached 372 million users within 24 months.

in the near future. The real-time nature of microblogging platforms also makes it possible to expect a faster response time than traditional Q&A systems. To some extent, MoboQ utilizes the strangers on Weibo as local human sensors and allows a question asker to extract context information on any given location by asking the local "human sensors" about what is happening right now around them. To the best of our knowledge, there is no other similar system that has been deployed in the field before; thus, we are the first to evaluate real-world usage of the fundamental concept of stranger-sourcing.

Figure 6.1 shows a sample inquiry and its answers from our database[1]. In this example, an asker was checking whether there was a free seat available at a university library. MoboQ successfully yielded 4 answers. Two of the answers provided possibly helpful and useful information to the asker. At least three of the answerers were very likely at the library at that moment, or just left for a few minutes.

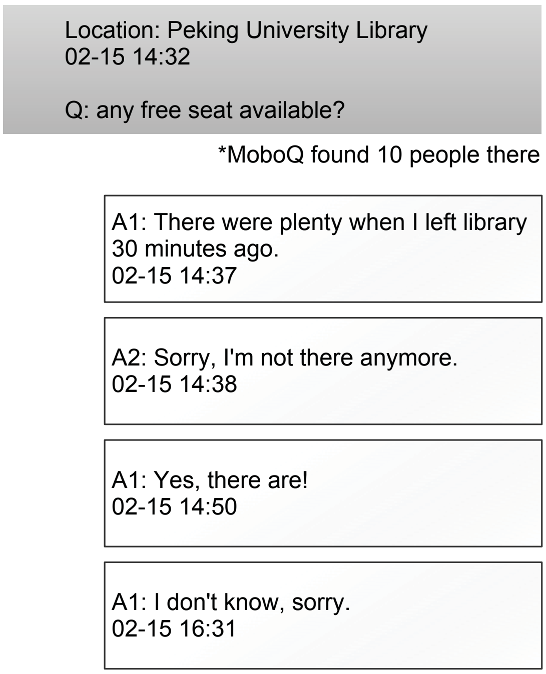


Figure 6.1: Example question and answers

[1]Translated from Chinese into English

6.1 Background: Microblogging and Location Based Services in China

Sina Weibo is a Chinese microblogging service. Sina Corporation[2] launched this service in August, 2009. The basic functionalities of Sina Weibo are similar to those of Twitter[3]: messaging, re-posting, direct messaging, mention (@), hashtag (#), and more. In addition, Weibo users can comment on a post, which means that the comments will be inserted into the original post rather than into their own timelines. This feature allows different Weibo users to have a discussion with each other under one topic (post). A screenshot of Weibo UI is given in Figure 6.2. Moreover, Weibo officially supports posting images, video, and music via local files or on-line URLs and locations, and build-in location check-in. Sina Weibo allows up to 140 Chinese characters per post, which means that one Sina Weibo message could contain more information than in Twitter. Furthermore, unlike Twitter, Chinese microblogging sites normally support a large amount of "verified accounts" under their real name systems (i.e., users must link their national identification number to a microblogging account), for people to declare their professional identities and titles. By the middle of 2012, Sina Weibo reached 368 million registered users, which makes it the second largest microblogging service in China.

There are a number of Foursquare[1]-like location-based sharing services in China, including Jiepang, Dianping, and Sina Place[2]. They each work in a similar way: people check in when they arrive at a location and earn badges or other rewards (sometimes coupons from local shops). Users can also give comments on the location or upload multimedia content of the location. According to an industry report from EnfoDesk[3], the total number of location-based services' users reached 30 million by the end of 2011 and was growing quickly. MoboQ can take advantage of both Sina Weibo and various location check-in services because almost all of the check-in services provide their users with the functionality of

[2]http://www.sina.com/
[3]http://www.twitter.com/
[1]http://www.foursquare.com/
[2]http://place.weibo.com/
[3]http://www.enfodesk.com/

broadcasting their check-in activities on social media platforms such as Weibo.

Figure 6.2: Screenshot of the main feed user interface of Sina Weibo (up:repost; down: post with image and its comments

6.2 Design and Implementation

In this section, we present the design of the MoboQ system: the overview of the main components, the basic workflow of the usage, the core algorithm for selecting and ranking the potential stranger answerers from Weibo, and the user interface design.

6.2.1 Main Components

The main components of MoboQ are:

- **Communication Module**: A REST Web Service, with an open API to client applications and the Sina Weibo API. This module handles the communications between an asker from MoboQ and an answerer from Sina Weibo.

- **Ranking Engine**: Searches and selects the best candidates on the Weibo platform to answer a question.

- **Client Applications**: Includes a Web site, mobile web, and a native mobile application to present the questions and answers to the user in an accessible and interactive form.

The MoboQ server, which comprises the communication module and the ranking engine, is implemented in Ruby on Rails. The mobile web is implemented using HTML5 technologies.

6.2.2 The life of a Query

Before discussing MoboQ in detail, it is useful to describe what happens behind the scenes when a user asks a question (Figure 6.3).

- A user begins by asking a question, which ideally is both location- and time-sensitive. To ask a question, the user must search and select a specific location or POI (Point-of-Interest); the location data are available via Sina location API. Then, the user inputs the question about the POI. This input can be accomplished through different client applications.

- Once the new question is delivered to the server, the Communication Module calls the Sina Weibo Search API to obtain a raw result data set of real-time conversations about the location (i.e., Weibo posts having this location keyword in the content). These raw data are then sent to the Ranking Engine. Raw results are re-ordered based on a series of principles in the Ranking Engine. For example, a high voting weight is given to a post that has the following characteristics: a check-in message, posted recently, from a Weibo user whose profile shows he/she lives in the same area of the question.

- Based on the ranking result, the Communication Module selects at most fifteen Weibo users to answer the question. The original question is then pushed to each user via the Sina Weibo Messaging API as a normal Weibo post. The post body describes that this request is a real request from a real person who needs help rather than a spam message. Links to the asker's MoboQ profile and the question page are also given in the content. The features are designed for helping Weibo users to understand why they have received such a message from a stranger. The Weibo users are notified by the mention function (i.e., @), and they can repost or comment on the question message to respond to the request. Either way, the answer will be caught by the Communication Module and will be delivered back to the MoboQ asker.

- Optionally, after receiving an answer, the asker can "thank" or start further conversation with the answerer. A one-to-one communication channel is established via Sina Weibo to support this possibility.

- After 24 hours, a question is marked as "closed", and no further responses are collected. Such a "deadline" is added because of the real-time character of the MoboQ service and the limited access rights of Weibo API.

6.2.3 Ranking Algorithm

Ranking in MoboQ is performed by the Ranking Engine, which determines an ordered list of candidate answerers who should be contacted to answer a question, from the raw data returned from the Sina Weibo search API (i.e., the Weibo posts that have the location keyword in the content).

At the beginning, all of the raw candidates are given a zero score. The main factors that determine this ranking of users are the *Current Location*, the *Expertise*, and the *Availability*:

- **Current Location**. First, the Ranking Engine finds the subset of users who have recently checked-in at a target location: those users that have posted content that matches the canned format of the check-in information

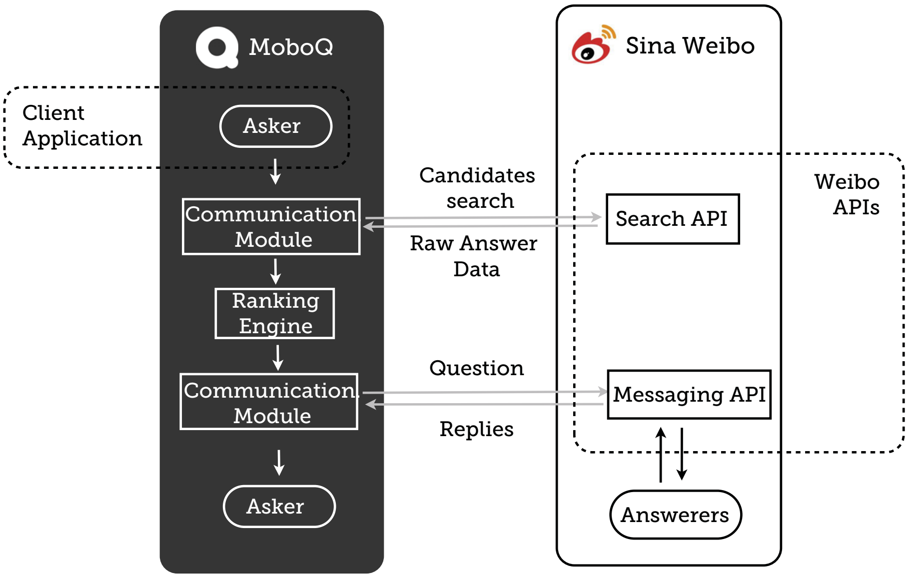


Figure 6.3: The life of a query in MoboQ

that is shared by the other LBS applications, such as “I’m here: #location#”. The Ranking Engine gives a weight (a positive score) to all of these users. The posting times of the tweets are also considered; we give newer posts higher scores and older posts lower scores, and the check-in messages that are posted more than 24 hours ago are ignored.

- **Location Expertise**. Second, the Ranking Engine matches the users’ profile with the given location to verify whether they are local residents. If they are, then the algorithm adds their weight by giving additional scores. In the next stage, the tag of the user and the classification of the location will be considered. In Sina Weibo, profile users can set up to seven tags for themselves, and the LBS services also assign a category to POIs, to allow us to match some popular pairs (such as a “foodie” tag and the “restaurant” location) to weight a user. Moreover, recent historical posts of the user will be analyzed to determine whether this user frequently checks in at the same location. The overall goal is to score each user according to the degree to which he/she is a good “match” for this location information query.
- **Availability**. Third, the Ranking Engine prioritizes candidate answerers

in such a way as to optimize the possibility that the present question will be answered. This task involves factors such as prioritizing users who are the most likely to be currently online (e.g., have posted a new message in the last 5 minutes); who are active Weibo user (e.g., have reasonable amounts of Weibo posts); or who posted the last message using a mobile or tablet device (because a Weibo app on those devices has a notification function).

- **Filter**. Given this ordered list of candidate answerers, the Ranking Engine then filters out Weibo users whose weight should be reduced. These filters operate largely as a set of rules, such as give a negative score to users who have less than 80 followers as well as who have more than 1200 followers, and totally filter out users who have less than 30 or more than 3000 followers. This rule is used because such users most likely are inactive users or have vast numbers of mention (@) messages every day and, thus, are unlikely to respond. Another rule is to reduce the weight of the people who are without a profile picture or use a default display name (e.g., "Weibo User 1672832") because there is a chance that their accounts are used infrequently or are automated robots. The Ranking Engine gives a negative score if the user's recent posts are shared from other social networks (several sites can share their user's status or activities to the user's Weibo account) because it is possible that the user is not really using Weibo. The engine also filters out the posts that are re-post messages and only considers candidates who originally mentioned the given location.

Finally, the top candidate answerers (at most ten in the beginning, which is increased to fifteen later) who have survived this filtering process are sent to the Communication Module. This model then proceeds with opening channels via the Weibo Messaging API to each of the candidates, serially inquiring whether they would like to help the MoboQ user by answering the present question; a polling service checks regularly for new notifications and returns any answer or conversation to the asker.

We should note that the ranking factors (and the corresponding weights) optimization is a continuous process; the list is neither final nor an optimal version and instead is (at most) only a starting point.

6.2.4 Support Trust Building Between Strangers

Because MoboQ is a Q&A system between people who are likely to be complete strangers; it is a challenge to motivate the potential answerers from Weibo to answer the strangers' questions. From a design point of view, we concentrate on two aspects simultaneously: 1) how to establish trust among the answerer, the MoboQ platform, and the asker, and 2) how to provide appropriate incentives to the candidate answerers.

Early studies suggest that a lack of trust can be identified as one of the greatest barriers that inhibit Internet transactions [38; 61]. For supporting online trust building, the content of a MoboQ query message has been specially designed. This message includes the reason why the candidate answerer has been selected (e.g., "Hi, we found that you just visited #Location#); the URL of the asker's profile page on the MoboQ site to show that the asker is a real person; and another URL to the question's page on MoboQ to help the answerer learn about the service. All of the necessary information is open to the candidate answerers, and we hope that this option helps the answerer to understand that this question is not a spam message but instead is a real question from a real person who is seeking help.

We also utilize findings from social psychology as incentives in our system. *Social Facilitation* [112] and *Social Loafing* [17] are two commonly cited behaviors that can affect contributions on social sites. The Social facilitation effect refers to the tendency of people to perform better on simple tasks while someone else is watching, rather than while they are alone or when they are working alongside other people. The Social loafing effect is the phenomenon of people making less effort to achieve a goal when they work in a group than when they work alone because they feel that their contributions do not count or are not evaluated or valued as much. This consideration is seen as one of the main reasons that groups are less productive than the combined performance of members working alone. Different mechanisms are employed in MoboQ to take advantage of positive social facilitation and to avoid negative social loafing:

- A public thank you message is provided; thus individuals' efforts can be prominently displayed.

- The query is a public message; thus, individuals know that others can easily evaluate their work.

- Every question is sent to up to 15 people in a separate message, to allow the unique value of each individual's contribution to be emphasized.

Some other popular incentive mechanisms, such as monetary incentives or game mechanisms, are not implemented in the current version of MoboQ. The reason is that the relationship between the answerers and the MoboQ platform is different than in previous systems: the answerers (who need to be incentivized to perform) are not yet registered users; thus, it is difficult to offer points, currency, or scores directly to them.

6.2.5 End-user Applications

MoboQ provides various client-side programs that people use to ask questions: Web site, mobile web (adopted by major mobile devices), and an iOS native application. The screenshot of an iOS App is shown in Figure 6.4. The functionalities of the desktop version and the mobile version application are slightly different. Other than the basic functions, such as ask, answer, comment, and check the latest questions in the city, the mobile version also provides the user the ability to browse nearby questions and answers, to follow nearby POIs or people, and to report the current situation of a nearby POI. Then, any future activity about the POIs or users would push to the user's feeds.

Moreover, we foresee that the future user interface will ask a question on MoboQ in its simplest form, which can be any type of lightweight text input mechanism along with a mechanism for displaying textual messages returned from MoboQ, e.g., SMS, Microblogging, Instant messaging, E-mail.

In this initial version, the MoboQ production system has no strong design for supporting in-community question answering. Although current MoboQ users can answer questions in the system, their answers are specially marked to notify the original asker that these answers are from a different source other than from Weibo.

Figure 6.4: Screenshot of the MoboQ iPhone application. In the picture is the page of recent questions

6.2.6 How we Advertise MoboQ

We mainly use free channels to promote MoboQ services. These channels include inviting friends to test the application, on-line social media marketing (i.e., creating social networking site accounts, posting related content and stories, and actively interacting with potential users), being featured in the media (blogs, technology news sites, magazines), and arranging campaigns on special dates such as festivals or large events. In general, the overall marketing activities are very lightweight; thus, it is likely that a majority of registered users get to know MoboQ through word-of-mouth recommendations.

6.3 Analysis

The Web version of MoboQ was deployed and made available publicly in January 2012, and the iOS version App was published in the iTunes store in March 2012. By the beginning of October 2012, there were a total of 35,214 users registered on MoboQ, and a total of 15,224 questions were asked. Those questions were sent to 162,954 strangers on Weibo, which resulted in 29,491 answers and 3,408 follow-up conversations. Most of the questions are about locations in an urban area. We provide an overview of the data in the following.

6.3.1 Demographics

Aside from data stored in the server, we also use Google Analytics[1] to monitor the traffic of the MoboQ Web site, and we use the HTML5 App. Not spuriously, 94.06% of the traffic was from Mainland China, followed by U.S. (2.19%) and Japan (1.38%). Moreover, most of the accesses were from the major Chinese cities, namely Shanghai (20.09%), Beijing (13.71%), and Guangzhou (6.81%).

6.3.2 Temporal Patterns and Answer Rates

The MoboQ users were more active during the daytime, from 10 am to 6 pm. Compared to major social media sites such as Facebook (10 am to 10 pm) [1] or

[1] http://www.google.com/analytics/

Twitter (9 am to 8 pm) [9], users appeared to use MoboQ more during typical working hours, and the usage faded quickly after 6 pm. Figure 6.5 provides an overview of the hourly distribution of the questions.

Of the 15,224 questions asked in MoboQ, 1383 questions (9% of the total questions) could not be sent out because the server failed to find any candidate. The main reasons were an uncommon POI/City and a mistyping of the location name. For the remaining questions, the average question found 8.17 candidate answerers ($SD = 3.34$), and 10,319 of the questions eventually received at least one answer, which resulted in a 74.6% overall answering rate.

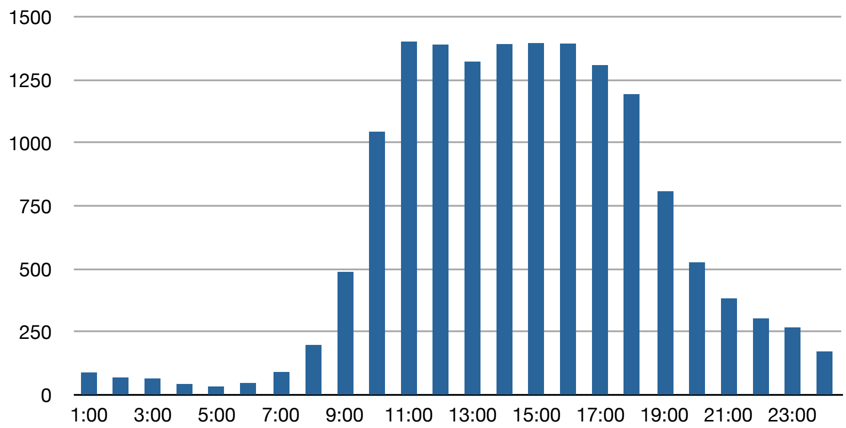


Figure 6.5: The temporal patterns associated with the questions

6.3.3 Desktop and Mobile Usage

Accesses from mobile devices (i.e., using the mobile Web and/or an iPhone App) accounted for 25.3% of the total traffic. This number is lower than our expectation. Because the majority of the usage scenarios that we pictured were for mobile users, we believe that the microblogging-style short query and answer is more useful on a phone, and the tasks are more likely to have time constraints when people are in a mobile context.

We assume that the absence of an Android App may be part of the reason because a recent report [2] shows that Android phones have already achieved more than 90% of the smartphone market share in China. In fact, even without a native App, more than 10% of the mobile traffic was from Android devices.

6.3.4 Question Classifications

We randomly selected 500 questions that were sent to MoboQ between September and October of 2012, and we analyzed all of the questions' contents via human coding. Eventually, we classified the questions into 6 major topics (see Table I for an overview). Reviews, advice, comments and recommendation queries regarding restaurants and bars are very popular (for example, "*Is the food good? Can you recommend some nice dishes?*"). A large number of questions are looking for insights, recommendations, and advice from local residents/experts (for example, "*Can you recommend any good place nearby this university that is suitable for a business meeting?*"). Some of the questions are socially oriented (e.g., "*Hi, is there anyone there who wants to have a cafe with me?*"), where users were trying to attempt a conversation rather than asking a question. We also provide the response rate for each category in table 6.1, but because there are many potential factors (e.g., the location, time, date, question content), these factors could affect the values, and we cannot give further conclusions about the relationships.

Surprisingly, we found that MoboQ users did not ask many time-sensitive questions. MoboQ was originally designed for situations in which a user is in the middle of an on- going activity and requires an answer to continue. The user interface, sample question, and instructions are all designed to emphasize this unique feature as well. However, only 27 (5.4% of the sample questions) of the questions actually required real-time information at the given location and, accordingly, demanded a rapid response. Most of such queries ask about traffic conditions or parking. We consider that one problem could be that there is a general-purpose system design without a clear and helpful usage scenario, which failed to guide new users to better understand the new concept of MoboQ.

6.3.5 Answer Quality

For a system that comprises human and machine sensors, the delay between two sensors is one of the most important problems. Furthermore, how fast human

Category	Questions	Answer Rate
Restaurants	48.6%	69.1%
Travel and Local Information	32.6%	83.4%
Transportation	6.6%	90.9%
Vote and Opinion	2.2%	100%
Social	1.6%	90.9%
Miscellaneous	7.2%	72.2%

Table 6.1: Categories of 500 Questions

participants can provide an answer largely determines the asker's experiences.

With regard to the response time of MoboQ (see Figure 6.6), we found that 2868 first answers (28%) arrived within 10 minutes, 5,366 (51%) arrived within 20 minutes, and 7494 (72%) arrived within an hour. In contrast, for asking questions using Facebook or in a friend-sourcing manner, only 24.3% of the queries received a response in 30 minutes, and 42.8% received a response in one hour or less [20].

If there is any type of additional process, it will delay the response time in MoboQ; for the current version, we do not emphasis quality assurance. This constraint means that all of the answers were delivered to askers immediately, for them to make a judgment, i.e., to select the alternatives that most likely are the correct answers. As a result, the quantity of answers also indirectly determines the quality of the answer.

Of the 10,307 answered questions, an average of 2.86 answers were received per question ($SD = 27.57$). As Figure 6.7 shows, approximately 69% of the questions receive multiple answers, and 10.2% of the questions receive more than 5 replies.

The original askers sent in a total of 6206 "thank you" messages back to the answerers. This result could suggest that the askers explicitly acknowledged the quality of at least 21.04% of the answers. Moreover, the original askers generated 3,408 follow-up conversations after receiving answers. Most of the cases were the askers attempting to obtain more information to improve the quality of the original answers. Some of the other cases were askers who do not want to use the default "thank you" message (or, perhaps they did not notice that there is such a function) and are willing to express their gratefulness in their own language.

Because of the push-based nature, MoboQ has fewer threads that result from spammer issues. Most of the meaningless answers are "*I dont know*" or "*I'm not familiar*", rather than malicious answers (i.e., answers that intentionally provide wrong answers for their own good, such as advertising their own business).

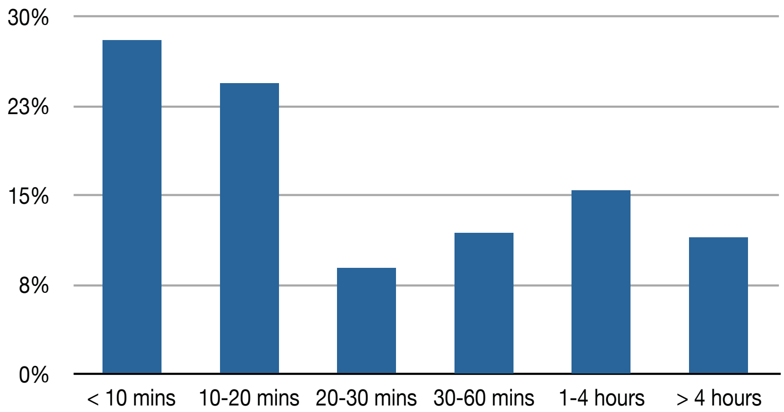


Figure 6.6: Response time (the first-answers)

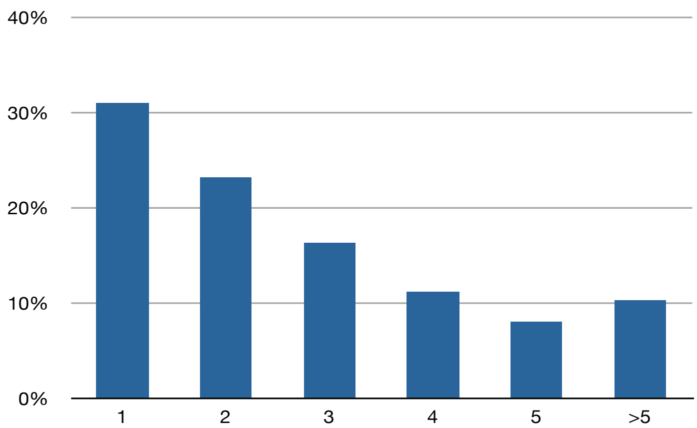


Figure 6.7: Number of answers per question

6.3.6 User Behaviors

We found that only 4909 registered users (13.94%) posted at least one question. On average, these active users posted 3.101 questions ($SD = 17.95$), with one extreme outlier who posted 295 questions. The distribution of questions per active user is shown in Figure 6.8; a large number of users have asked a question only once. From the data, we could see that the askers can be divided into two clear classes, a small group of highly active users who have contributed on average more than 50 questions and a distinctively larger class of users with limited activity who have asked fewer than 10 questions. On the other hand, it is normal for an on-line community to have a power law distribution, with only a small minority contributing. For example, [12] reported that, in the Peer-to-Peer music sharing service Gnutella, 10% of the users provided 87% of the music. In the open source software community, 4% of the participants likewise contributed 88% of the new code [81]. Moreover, MoboQ is designed for a specific type of question, which is location- and/or time-oriented questions; this type of problem usually does not arise very frequently. At the same time, in the user interface, the instructions and tips in which we explain the questions that would push real people who are willing to offer to help total strangers, and we ask to please not send meaningless questions to break this trust. Hence, unlike in traditional Q&A sites, we did not see many "test" questions, which often test out only the users' initial curiosity about the service.

We found that, of the 29,487 answers, 2051 MoboQ registered users (5.8%) contributed 4731 responses (16%). This observation might also arise because the users typically are able to answer questions only about the locations that they are familiar with. When the number of questions is relatively small, the users might hardly ever see a question that they could answer. As a comparison, 27% of the Yahoo! Answers users have ever provided a single response [43]. Figure 6.9 shows the user behaviors, it is interesting to note that we found that only 193 registered users have posted both the question and answer, which is 3.9% of the total users who ever asked a question and 9.4% of the users who ever answered a question. Additionally, a total of 1110 registered users posted at least one comment in the answers; from our observations, most of these comments are further questions or discussion about the answers.

Surprisingly, even if 22,882 Weibo strangers have answered at least one question, less than 0.5% of them joined MoboQ afterward. Every question contains two URL links to a MoboQ site. One link is the asker's page and the other link is the questions detail. However, we tracked very few incoming accesses from those links. After receiving answers, 6,206 thank you messages were sent by askers (i.e., 21% of the answers received a thank-you message that also contains an implicit invitation to join the community). This result might indicate that the strangers could be willing to help by sending a simple reply sentence within Weibo, but clicking an additional URL to leave for another site might be too costly or risky for them.

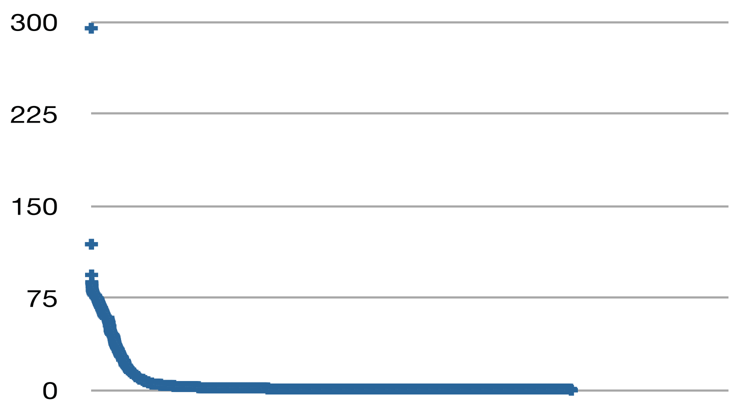


Figure 6.8: Distribution of questions per user

6.4 Implications

In this section, we focus on the implications of the analysis of the usage data, and we discuss the findings.

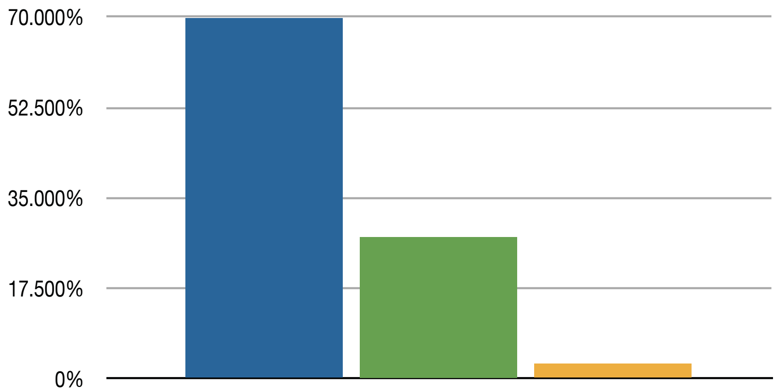


Figure 6.9: Blue: user who asked at least one question; Green: user who answered at least one question; Yellow: user who both asked and answered questions

6.4.1 Privacy

The privacy issue has been widely discussed ever since the birth of social media, and the stranger Q&A feature of MoboQ pushes further on social media users understanding of on-line privacy. In [98], the authors propose a number of guidelines for creating a location-sharing service. In [102], the authors discuss how a socially driven sharing (e.g., the location-based service Foursquare) is fundamentally different from purpose-driven sharing (e.g., Google Maps[1]).

In MoboQ, we define several fundamental rules for protecting a user's private information; for example, the collected data that are used to predict a user's location must be publicly shared by the user himself/herself. However, we had the concerns that a Weibo user's understanding of the service might be different from the actual model that is used by the site, for example, users might not be aware that their check-in message is publicly posted. Therefore, we maintain our attention on any user feedback that is useful for analyzing how much of an issue this concern is for the MoboQ system. Overall, based on the feedback data, there were only two answerers who complained that a sudden question was annoying to them and that it violated their privacy. However, we do not have data on

[1] http://maps.google.com/

whether users have marked the question messages as spam in Weibo. From the community management point of view, it is also important to guarantee that all of the questions are real and sincere. Spam questions should be strictly filtered out before being sent to social media users; thus, a comprehensive monitoring system should be established in the next stage.

6.4.2 Quality Control

As explained in the previous section, there is no explicit quality control method that is implemented in the current version of MoboQ. However, we observed people who were not at the location or did not understand the MoboQ service rather than malicious users who generated a large number of irrelevant answers. We believe that the push-based question assignment mechanism also helped to reduce the possibility of having malicious answerers because all of the answerers were selected by the algorithms and, thus, spammers had less opportunity to insert themselves into the process.

To further improve the answer quality, additional components might be included. In the future, we are interested in building classifiers that could detect extremely bad answers and filter out obviously useless messages and spam messages using Natural Language Processing technologies.

6.4.3 Augmenting MoboQ through Social Media Mining

Systems such as the Social Telescope [96] collect location-related information via *Social Network Mining*. Although many questions that are asked in MoboQ could not be satisfied through such an implicit way of collecting a review, such an approach could significantly enhance the response time for the party of the queries.

When introducing a human into the loop, one of the drawbacks is that humans increase the time-cost of the whole process. Hence, data mining in the social media could potentially augment the weakness of MoboQ.

Therefore, we consider a combined Social Network Mining approach, which presents the results to an asker as reference data; this approach could be another interesting future direction. Simultaneously, we would consider adding a new feature of “report” into MoboQ, with which the user can “report” what activities are occurring at his current position and share these interesting stories with the community. By accomplishing these changes, the system could generate a “potential answers pool” before anyone is actually asking a question, which could be used as a reference for the potential askers.

6.4.4 Trust and Incentives

When discussing the stranger-sourcing concept, many people would first think of what incentive to provide to the answerers; however, we found that perhaps a more important question is how to build trust between strangers via the Internet. In MoboQ, a point system and a simple overall leaderboard are implemented to motivate community members to answer questions, but neither of these approaches appears to be effective because few MoboQ users contributed answers. However, the Weibo users (to whom we can hardly provide any strong incentives) actively responded to the requests.

We believe that this result might demonstrate that once the strangers consider that the message is coming from a real person who is in trouble, they would very likely offer their help. However, without further study, we cannot provide a conclusion as to which factor might drive such altruistic activity.

Nevertheless, we designed the system as openly as possible, to support online trust between strangers: non-registered strangers can easily verify the asker’s profile and history as well as the MoboQ’s introduction information. All of the messages between MoboQ and Weibo are sent in public channels rather than private messaging, to allow all of the information to be open to everyone. Furthermore, the Weibo user’s effort to help the asker is minimized to the task of simply replying to the tweet.

6.4.5 Support Region-Oriented Questions

We found that such a location-based system should be able to search for a larger region instead of only a specific POI. Current MoboQ is designed for asking questions to very specific POIs, e.g., a shop, a restaurant, or a bar. One of the reasons for this constraint is that current location-based sharing services users normally check in at POIs rather than at regions.

However, we found that a MoboQ user often sets a whole area as a target location and, then, asks a question about one specific place in that area. For example, if Dom Coffee is one shop in a large shopping mall called Nordstroms, then many users might ask "#City, Nordstrom shopping mall# is Dom Coffee crowded right now?" rather than "#City, Dom Coffee# is it crowded over there". This action is logical and is an understandable way of thinking; however, it adds new requirements when searching for appropriate answer candidates. This type of query might suggest that, in a future version of MoboQ, we should implement an additional semantic parser to retrieve possible POI names from the question content and consider these POIs to be another factor in the ranking algorithm.

6.4.6 User Expectations and Demands

In traditional Q&A sites such as Yahoo! Answers, there is already a large database of questions and answers; thus, a user can search for the answers immediately. On the other hand, MoboQ attempts to solve real-time questions, which means that, even for the same question at the same location, the answer would most likely be different depending on the time and the context; thus, the previous questions and answers are much less useful than in other Q&A services. As a result, compared to other Q&A sites, MoboQ does not have a well-supported function to display and search historic questions and answers. However, such a design leaves a user almost nothing to do other than asking/answering questions and waiting. From the users' feedback, we received several suggestions about providing more intelligent questions and answers that contain only the locations that they are interested in, which implies that there would be a demand for browsing historical data.

We found that another core issue is that users have different expectations of the response time, with regard to the type of question or the context. However, MoboQ cannot guarantee the waiting time to users while adjusting these expectations. In the future, we believe that it could be very helpful to give an estimated waiting time for each question based on historical data. Additionally, a notification should be given when it is highly likely that the user's expectation for the response time would be higher than what the service could offer, for example, in the case of asking a question about an uncommon location or at midnight. An appropriate guide should be provided to suggest possible alternative tasks or choices.

6.5 Applications for More Specific Scenarios

MoboQ is designed to be a common platform for supporting location-related real-time question answering. In our initial deployment, we focused on a wide range of questions, but the results show that it is difficult for a new user to understand such a general-purpose application; thus, user adoption is slow. We believe that it is important to repurpose our general infrastructure to support more focused applications that are difficult to be implemented by traditional methods. Nonetheless, from the asker's usage data and the classifications, we can see the potential of narrowing down the usage scenario and building ubiquitous computing applications for more specific purposes on top of the MoboQ platform. Moreover, it is reasonable to assume that such application performances could be better than the current MoboQ system if the ranking algorithm and the data sources can be adjusted for the specific application domain.

The most successful applications built on top of the MoboQ would need to have the following properties:

- Utilize information that cannot be collected via traditional sensor technologies.
- Be able to tolerate a reasonable delay length (e.g., five minutes to half an hour) because human sensors cannot provide a response as fast as machine sensors can.
- Location and time-sensitivity should be required components of the solution.
- The problem targeted is difficult for existing social search systems to solve.

With the above guidelines and the question classification results in mind, we can at least foresee the following potential applications directions:

6.5.1 Tourism Information

Visitors might want to obtain immediate support from local residents, e.g., recommendations for restaurants, advice for transportation, questions about directions, and weather conditions. This possibility is different from existing users' reviews and local information search services such as Yelp or Dianping, because people can have asynchronous conversations with other people to ask information with certain requirements and to obtain specific answers for their demands. For example, "how cold is Beijing, what should I wear, is a jacket enough" instead of having a number for the temperature, or "is there a cafe suitable for a business meeting" instead of checking comments on Yelp to find out the answer.

6.5.2 Emergency Management Information

A stranger-source system could also be used for information collection during emergency situations, such as detecting the current state of a given location after a natural disaster. During natural disasters, such as an earthquake, it is common for the telephone voice network to be destroyed or inoperative. Some of the authors witnessed the situation that occurred at the nearest major cities to the epicenters of the 2011 Tohoku Ms9.0 earthquake: the telephone network was completely down after the first shock waves, but the Internet connection survived, although the speed was slow. During the important first 24 hours, many people relied on social networks to report, communicate and receive the latest news. Previous research[32; 42; 59; 93] studied how to utilize Twitter feeds to detect, in real-time, crisis events such as earthquakes; however, there was no method similar to MoboQ to push questions to a specific location for extracting the needed information. Thus, the information could be greatly helpful for assessing the extent of the disaster and then for effectively allocating the valuable rescue and relief resources to the hardest hit areas.

6.5.3 Shopping Information

Real-time (or so-called semi-real time) information about shops, restaurants, and entertainment facilities can be helpful for supporting customers' decision-making processes. Often, people need such additional information for making plans to visit a location. For example, "which department store has the larger sale today", "is there a ticket left for the movie", or "which restaurant has no line waiting outside." We discovered that this type of question accounted for a large proportion of the questions that were asked in the current MoboQ system, which suggests that it is an especially useful application of this technology.

6.5.4 Event Report

Check-ins sometimes occur at an event rather than at a location. It is, therefore, possible to find people who are currently at large on-going events, such as conferences, sports games, music concerts, and so forth. Although there are alternative studies [27; 84] about mining social media data to summaries ongoing or past events, such as football games, for people with specific questions or purposes, a question answering type of service could be more useful. Moreover, for small events, because it is less likely that there will be sufficient information to generate a summary, a Q&A approach could be a good alternative.

6.5.5 Context Acquisition Systems

Stranger-sourcing technicalities could also support an existing context-sensing system for extracting data that cannot be collected otherwise. For example, a computer system can automatically and autonomously generate questions, translate system needs into human language, and extract data from answer content (i.e., natural language). Thus, such a system can constantly fetch high- level context data such as a social environment or human activities and can store the information that is needed to build a rich context-information database.

6.6 Chapter Discussion

We investigated the concept of strangersourcing on Social Media via a deployed application called MoboQ. Using MoboQ, a user can ask temporal and geo-sensitive questions, such as how long is the line at a popular business right now, and then receive answers that are crowdsourced from other users in a timely fashion. To obtain answers for questions, the system analyses the live stream from the public microblogging service Sina Weibo to identify people who are likely to be currently at the place that is associated with a question and sends them the unsolicited question through the microblogging service from which they were identified.

Through nine months of real-world operation data, we verified the feasibility and advantages of building real-time social question-answering services on top of an existing microblogging platform, we analyzed users' behaviors and preferences, we discussed lesson learned, and we proposed a list of future applications that could be built on top of the MoboQ platform. We foresee that there is great potential for MoboQ to become an important infrastructure component for building future context-aware applications.

Chapter 7

Discussion

The previous chapters introduces four case studies upon the general framework. Through the case studies we have learned experiences and practical knowledge on the cloud and crowd powered personal knowledge management systems. Based on the lesson learned of the independent case study, some important and interesting tradoffs are further identified and discussed in this chapter. The first design tradeoff between use *one-way communication* and *two-ways communication* impacts the ecosystem of the system, the second design tradeoff between *browse based information seeking* and *search based information seeking* impacts how information flows among people and the information structure of our society, the third tradeoff between *controlled lab experiment* and *field user study* impacts what we could learned from the evaluations. We also discuss how to evaluate the systems and trade-off between performance and cost.

7.1 One-way and Two-ways Communication

The cloud powered systems described in the thesis generally use a one-way communication model, where machine and people are collaboratively find and use the re-mixable information in the cloud for the knowledge creation tasks. Here *one-way* means the interaction in system is the information seeker pulls existing information from database.

It is said that the great creator Pablo Picasso once said "*good artists copy great* *artists steal*". Remix of old thing is already well acknowledged as one of most important methods for creating new things. The proposed re-usable knowledge seeking module is designed for support remix activity. Because with the rise of cloud infrastructure and big data movement, increasing amount of digital works is generated everyday; and the Open Content and Open Resource licenses provide legal support for re-use other people's digital work. However, even if the re-usable data is already available, there is only a few digital creation tools with easy, cheap, and convenience remix support. Throughout our prototypes, we found it is particularly interesting to consider how to combine implicit search and open content databases with the existing knowledge creation applications. We have tested the concept with some popular tools such as PowerPoint and Video creator, but there is still a large under explored area.

On the other hand, not every demanded information could be found in the open database. But some people with specific knowledge or experience might be able to provide such information. The two-way communication model is useful in such situation to reinforce the one-way communication model, where information seekers can actively find and communicate with the people with expert knowledge for getting the specific information they want. In the meanwhile, this model could hardly provide a general algorithm for searching experts. Different factors and features should be applied in the algorithms according to the specific application. In our case studies, we investigated the area of local expert search and confirmed the stranger local experts could be identified based on their digital footprint, and they could provide a nearly-real-time and on-demand support without money in return. But for other application domain, the answerer recommendation algorithms need to be re-designed.

There is a trade-off when deciding which model(s) is suitable for the a problem the application is facing. A combination solution is powerful but expensive, while different model offers the system completely different abilities. System designers should analysis carefully about: what kind of information is needed, where to find the information, how to access it, and how what are the user requirements on response time and quality.

7.2 Browse-based and Search-based Information Seeking

In city design, urban planner tried to improve the qualities of "third places" [88] — the informal public places on neutral ground where people can gather and interact. Different from home (the "first place") and work (the "second place"), third places allow people to "put aside their concerns and simply enjoy the company and conversation around them". It is important because they are the places where people meet different people and information, while in a free and peaceful society the understanding between people from different social classes is often achieved by discussions and communications. Recently, in the digital world designers also try to take such "third places"[100] in to consideration when design on-line systems. However, the Information Foraging Theory [89] suggests that humans are inherently trending to consume as few energy as possible to get the information they are seeking. Based on the theory (and, of course, with the increasing amount of digital data), the on-line systems trended to provide search-based and push-based (i.e., recommendation) information seeking assistant rather than a browser-based solution.

We have already witnessed the significant impact brought by search engines and recommendation engines on people's information accessing behaviors. However, it is difficult for current algorithms to provide serendipitous information that is truly "unexpected" for people, since the recommended items are mainly selected based on user's historical on-line behaviors. Social media is also an emerging information accessing channel. On social media people basically receive information that is shared by their social relationships and is mostly related to the topics their friends are interested in. However, as the Homophily Principle indicates, "birds of a feather flock together" [79], people naturally like to be friends with people that have similar minds. Thus, people with a relationship on social networks normally have interests in topics that closely resemble each other. The situations might eventually cause the Echo Chamber problem [58] in the digital world: having on-line (virtual) communities where people's existing beliefs are reinforced and they are not exposed to different voices.

Digital designers need to consider the trade-off between providing *browse based* and *search based* information seeking. Browsing takes time, but offers the opportunities of serendipitous information discovery; searching or recommendation is convenient and might be preferred by users, but the may arise the concern of Echo Chamber problem. It is challenging to integrate the two methods to provide a more balanced solution for not only good individual experiences but also positive social impact.

7.3 Controlled and Field User Study

Evaluating human-computer interaction systems is never an easy work [5; 19; 33]. Generally speaking, there are two major ways: controlled lab user study and larger scale on-field study. Normally, controlled user studies have a rather small number of participants (usually less than 30). Which is suitable for testing usability or other qualitative results. However, such lab experiment based evaluation is challenged by the social computing applications. By design, social computing systems usually require a relatively larger number of user to create an appropriate eco-system. It might be difficult to study and understand large-scale phenomenons in the eco-system using a small number of experiment participants. Therefore, for social computing systems, deploy the prototype into a real world environment and conduct field user study becomes a more appropriate evaluation method to study user behavior, verify feasibility, and proof concept. Additionally, the outcomes and lesson learned of a field study is often more convincing and appreciated by other scholars and designers in the community.

On the other hand, conducting large scale on field user study can be painful, especially for students in a academic lab. It often requires a large amount of time. Taking MoboQ system as an example, it costed over 4 months developing the initial version and kept iterating for nearly ten months to attract user. For a successful large scale deployment, researchers often need to deliver a product level quality prototype, provide continuous QA support, and flight against spammers. As a result, researchers need to function as an entire start-up to handle different tasks including design, marketing, customer support, user experience study and engineering. Moreover, sometimes the research projects need to compete with industry companies for attention and users, which is quite challenging.

The trade-off between different study methods is not simply a pro/con decision, it is related to how much time you have, what kind of resources you can access, and most importantly, what is your research goal. Large scale study might be time consuming and possibly have low pay back. But for most of the HCI systems, without data collected from real world users, you can hardly even convince yourself that this design could be impactful. While the possibilities of impact the future products and services is probably the most important goal of HCI systems research, because unlike other domains, HCI systems is often not about complex design or technical challenges.

7.4 Evaluation Metrics

When a crowd and cloud powered system is created for a problem, how can it be evaluated? In the presented projects we introduce a set of metrics for evaluating. Mostly, the metrics are *user centered* and focus on Human Factors rather than *system centered factors*.

7.4.1 Efficiency

Efficiency is one of the classical metrics for evaluate algorithms. For standard algorithms, the efficiency is measured by counting atomic steps. For example, QuickSort is said run in $O(nlogn)$ time on average, that means the algorithm sorts a list of n elements in approximately $nlogn$ computation steps. However, in the case of social media crowdsourcing algorithms, it is difficult to measure the efficiency using computation step. Therefore, we selected more user experiences related metrics that can potentially indicate how user's exceptions could be satisfied.

First, we use *response rate* to define the quality of expert search algorithms. Differ from paid crowdsourcing, the tasks in social crowdsourcing systems are pushed to a group of strangers. The strangers normally have no mutual agreements with the service providers about whether they would like to participate. Therefore, how many strangers in the group will response to questions is an important aspect when measuring the system's performance.

Depending on the type of tasks, how long must requesters wait for a response can be a critical metric. In a mobile computing environment, sometimes users are in the middle of an activity, and need the response from service to support their decision-making process to continue the activity. In such situations, the *response time* of the first answer is important metric to measure the quality of service.

In this thesis, we mainly focus on the response time and response rate thus the presented prototypes did not provide much effort for analyzing the output (basically, any additional process will affect the response time). However, the *quality of the response* should be considered and measured in the next stage of this line of research. How to evaluate response's quality can be different, for instance, in UbiAsk the quality of answer could be verified via additional proof-reading process; but in MoboQ system it is difficult to use proof-reading method, whereas approaches like majority voting and reputation of answerers are more suitable alternatives.

7.4.2 Performance and Cost

To improve the overall performance, we often need to face different trade-offs. For example, additional process can improve quality of output, but communication between humans are time costly. Involving more crowd members and let them working simultaneously could improve the response time, but it could also arise the attention of spam filters at the social media platforms. Getting higher third-party developer priority can have less limitation when interacting with the social media platforms' APIs, but normally the price is quite expensive.

The strategies of how to decide these trade-offs are based on the user's requirements in the particular usage scenarios. Under different conditions, user may have different preferences on which performance metrics are more important than others, system should have the flexibility to adopt user's needs and update the strategies automatically.

Chapter 8

Conclusion and Future Direction

This thesis presents a general framework that is applicable for systems that are designed to support personal knowledge exchanging, managing and creating activities that current technologies cannot yet handle. This framework combined models of seeking and sharing existing re-usable knowledge in the cloud database with models of seeking and sharing non-existing contextual-sensitive information via crowdsourcing.

8.1 Vision

It is time to begin defining long-term goals for the framework. This section articulates such millennium goals for cloud and crowd powered systems.

Figure 8.1 layouts the future vision of the proposed framework. Generally speaking, the crowd model and cloud model will work more closely to support each other, the crowd workers will collaborated with each other to perform the tasks, and more intelligent and accurate experts search algorithms will be built top of data analysis results. In such systems, the computers should try to provide the right information to the right crowd workers in order to have the right solutions for the current problems.

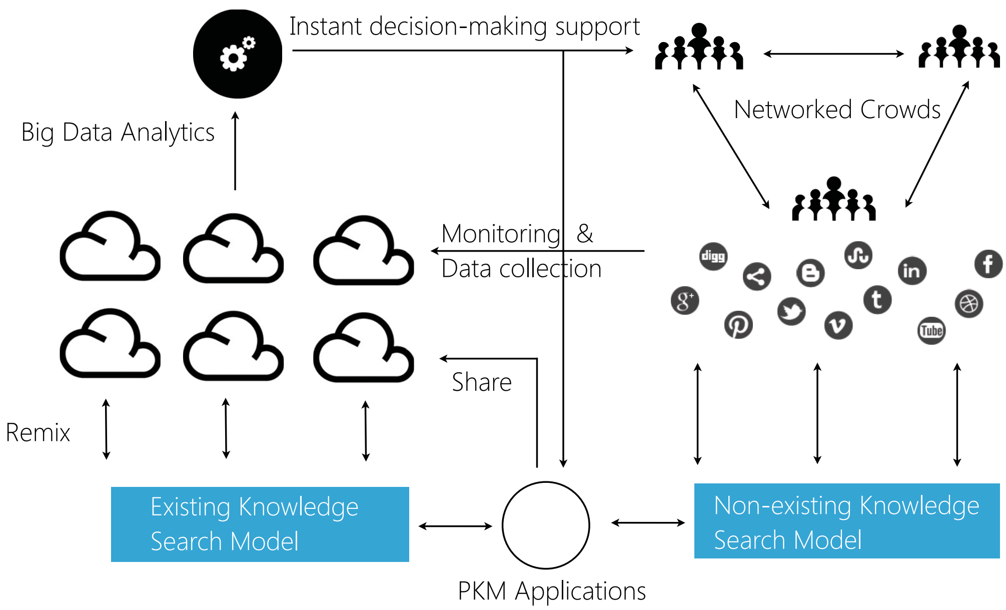


Figure 8.1: The vision

8.1.1 Combine Machine Intelligence to Aid Crowds

The systems presented in this thesis do not deeply integrate cloud and crowd models together. However, the two models also complement each other naturally. On the one hand, the re-usable data could not only used by information seeker to meet their needs, but might also valuable for providing to crowd workers to help them to develop their answers. On the other hand, the result data provided by trusted crowd members could be used to train the machine learning or data mining algorithms, and improve the performance of searching and recommending re-usable data. As a result, a meta-algorithm may want to balance between the speed and cost of the crowd model and the quality of the cloud model.

8.1.2 Networked Workers for Complex Tasks

Most targeted tasks in this thesis are relatively simple and direct tasks. It would be interesting to consider more complex tasks where require massive scale collaboration between the different components, where multiple crowd members need to work concurrently and collaborate with each other as a team on the same task. A number of new challenges could be raised, such as how to design working procedure, how to define priority of participants and tasks, and how to provide communication support.

8.1.3 Identify Right Re-usable Data and Experts

The presented systems in this thesis mostly contain recommendation model for activity searching the right information items or providers. Since these systems are the pioneer prototypes in this new research domain, the performance and efficiency of the recommendation algorithms would definitely have plenty of room for improvement. For example, MoboQ system achieved a 71% overall response rate, but since each question is pushed to more than 10 people, the response rate for individual is only lower than 20%.

We believe there is an immense amount of work to be done in the continued development of this framework. We hope researchers will use and improve upon the methods presented in the thesis for development and evaluation of cloud and crowd powered systems. We believe that this tight integration of user, crowd, and cloud will be powerful model for future personal knowledge management systems.

References

[1] Anatomy of a facebook post http://www.vitrue.com/educational-resources. 79

[2] Android market share in china 2012 http://www.techinasia.com/android-market-share-china-2012/. 80

[3] Data on the big data http://marciaconner.com/blog/data-on-big-data/. 1

[4] Frand, j. and hixon, c., (1999) personal knowledge management: who, what, why, when, where, how? http://www.anderson.ucla.edu/faculty/jason.frand/researcher/speeches/educom98pkm/. 1

[5] Landay, j.a. 2009. i give up on chi/uist. blog. http://dubfuture.blogspot.com/2009/11/i-give-up-onchiuist.html. 96

[6] Needleseek http://needleseek.msra.cn/. 42, 43

[7] Nielsen, j.: Recruiting test participants for usability studies. http://www.useit.com/alertbox/20030120.html. 59

[8] Splat http://msrsplatdemo.cloudapp.net/. 43

[9] Time is on your side http://blog.bitly.com/post/22663850994/time-is-on-your-side. 80

[10] *CC, Creative Commons*. Free and Open Source Software, 2007. 14

[11] Lada A. Adamic, Jun Zhang, Eytan Bakshy, and Mark S. Ackerman. Knowledge sharing and yahoo answers: everyone knows something. In *Proceedings of the 17th international conference on World Wide Web*, WWW '08, New York, NY, USA, 2008. ACM. 55

[12] Eytan Adar and Bernardo A. Huberman. Free riding on gnutella. *First Monday*, **5**[10], 2010. 83

[13] P. V. André. Anatomy of the unsought finding: serendipity: origin, history, domains, traditions, appearances, patterns and programmability. *British Journal for the Philosophy of Science*, **45**[2]:631–648, 1994. 14

[14] J. Antin. Designing social psychological incentives for online collective action. In *Directions and Implications of Advanced Computing; Conference on Online Deliberation, p. ()*, DIAC '08, 2008. 18

[15] Y. Arase, X. Xie, M. Duan, T. Hara, and S. Nishio. A game based approach to assign geographical relevance to web images. In *Proceedings of the 18th International Conference on World Wide Web*, WWW '09, 2009. 19

[16] Dan Ariely. *Predictably irrational, revised and expanded edition: The hidden forces that shape our decisions.* Harper, 2009. 22

[17] K.W. B Latane and S. Harkins. Many hands make light the work: the causes and consequences of social loafing. *Pers. Soc. Psychol.*, **37**[6]:822–832, 1979. 18, 77

[18] M. S. Bernstein. Crowd-powered systems. *KI-Künstliche Intelligenz*, **27**[1]:69–73, 2013. 2

[19] Michael S. Bernstein, Mark S. Ackerman, Ed H. Chi, and Robert C. Miller. The trouble with social computing systems research. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '11, pages 389–398, New York, NY, USA, 2011. ACM. 96

[20] Michael S. Bernstein, Desney Tan, Greg Smith, Mary Czerwinski, and Eric Horvitz. Personalization via friendsourcing. *ACM Trans. Comput.-Hum. Interact.*, **17**[2]:6:1–6:28, 2010. 82

[21] S. Bowman and C. Willis. *We Media: How Audiences are Shaping the Future of News and Information*. The Media Center at the American Press Institute., 2003. 2

[22] Andrei Broder. A taxonomy of web search. *SIGIR Forum*, **36**[2]:3–10, 2002. 25

[23] J. Burke, D. Estrin, M. Hansen, A. Parker, N. Ramanathan, S. Reddy, and M. B. Srivastava. Participatory sensing. In *Workshop on World-Sensor-Web: Mobile Device Centric Sensor Networks and Applications. WSW'06*. ACM Press, 2006. 69

[24] J. C. Byers, A. C. Bittner, and S. G. Hill. Traditional and raw task load index (tlx) correlations: Are paired comparisons necessary? *Advances in Industrial Ergonomics and Safety*, **1**:481–485, 1989. 34

[25] Chris Callison-Burch and Mark Dredze. Creating speech and language data with amazon's mechanical turk. In *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk*, pages 1–12. Association for Computational Linguistics, 2010. 12

[26] Andrew T. Campbell, Shane B. Eisenman, Nicholas D. Lane, Emiliano Miluzzo, Ronald A. Peterson, Hong Lu, Xiao Zheng, Mirco Musolesi, Kristóf Fodor, and Gahng-Seop Ahn. The rise of people-centric sensing. *IEEE Internet Computing*, **12**[4]:12–21, jul 2008. 69

[27] Deepayan Chakrabarti and Kunal Punera. Event summarization using tweets. In *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*, ICWSM '11, pages 66–73, Palo Alto, California, 2011. AAAI Press. 90

[28] Schweik Charles, Tom Evans, and J. Morgan Grove. Open source and open content: A framework for global collaboration in social-ecological research. *Ecology and Society*, **10**[1]:33, 2005. 14

[29] G. Cheliotis. From open source to open content: Organization, licensing and decision processes in open cultural production. *Decision Support Systems*, **47**[3]:229–244, 2009. 26

[30] Giorgos Cheliotis and Jude Yew. An analysis of the social structure of remix culture. In *Proceedings of the fourth international conference on Communities and technologies*, C&T '09, pages 165–174, New York, NY, USA, 2009. ACM. 26

[31] C. Cheshire and J. Antin. The social psychological effects of feedback on the production of internet information pools. *Com- put. Mediat. Commun.*, **13**:705–727, 2008. 18

[32] Andrew Crooks, Arie Croitoru, Anthony Stefanidis, and Jacek Radzikowski. Earthquake: Twitter as a distributed sensor system. *Transactions in GIS*, 2013. 90

[33] Jr. Dan R. Olsen. Evaluating user interface systems research. In *In Proceedings of the 20th annual ACM symposium on User interface software and technology*, UIST '07, NY, USA, 2007. ACM, New York,. 96

[34] Flaste R. Deci, E. *Why We Do What We Do: Understanding Self-Motivation.* Penguin, London, 1996. 17

[35] N. Diakopoulos, K. Luther, Y.E. Medynskiy, and I. Essa. The evolution of authorship in a remix society. In *In Proceedings of the eighteenth conference on Hypertext and hypermedia*, HT 2007, pages 133–136. ACM Press, 2007. 10

[36] Nicholas Diakopoulos, Kurt Luther, Yevgeniy (Eugene) Medynskiy, and Irfan Essa. The evolution of authorship in a remix society. In *Proceedings of the eighteenth conference on Hypertext and hypermedia*, HT '07, pages 133–136, 2007. 14, 23

[37] NATHAN EAGLE. txteagle: Mobile crowdsourcing. In *Proceedings of the 3rd International Conference on Internationalization, Design and Global Development: Held as Part of HCI International 2009*, IDGD '09, pages 447–456, Berlin, Heidelberg, 2009. Springer-Verlag. 68

[38] J. FENG, J. LAZAR, AND J. PREECE. Empathy and online interpersonal trust: A fragile relationship. *Behaviour and Information Technology*, **23**[2]:97–106, 2004. 77

[39] B. FREI. Paid crowdsourcing: Current state and progress toward mainstream business use. Technical report, Whitepaper Produced by Smartsheet.com, 2009. 17

[40] HIROMICHI FUJISAWA. Forty years of research in character and document recognition-an industrial perspective. *Pattern Recogn.*, **41**[8]:2435–2446, 2008. 68

[41] YOTAM GINGOLD, ARIEL SHAMIR, AND DANIEL COHEN-OR. Micro perceptual human computation for visual tasks. *ACM Transactions on Graphics (TOG)*, **31**[5]:119, 2012. 12

[42] MICHELLE GUY, PAUL EARLE, CHRIS OSTRUM, KENNY GRUCHALLA, AND SCOTT HORVATH. Integration and dissemination of citizen reported and seismically derived earthquake information via social network technologies. In *Proceedings of the 9th international conference on Advances in Intelligent Data Analysis*, IDA '10, Berlin, Heidelberg, 2010. Springer-Verlag. 90

[43] ZOLTAN GYONGYI, GEORGIA KOUTRIKA, JAN PEDERSEN, AND HECTOR GARCIA-MOLINA. Questioning yahoo! answers. Technical Report 2007-35, Stanford InfoLab, 2007. 83

[44] HIETANEN. H. *The Pursuit of Efficient Copyright Licensing — How Some Rights Reserved Attempts to Solve the Problems of All Rights Reserved.* PhD thesis, Acta Universitatis Lappeenrantaensis 325, Lappeenranta, Finland, 2008. 26

[45] SEVERIN HACKER AND LUIS VON AHN. Matchin: eliciting user preferences with an online game. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, pages 1207–1216. ACM, 2009. 12

[46] ISMAIL HARITAOGLU. Infoscope: Link from real world to digital information space. In *Proceedings of the 3rd international conference on Ubiquitous Computing*, UbiComp '01, pages 247–255, London, UK, UK, 2001. Springer-Verlag. 52

[47] S.G. HART AND L.E. STAVELAND. Development of nasa- tlx (task load index): Results of empirical and theoretical research. *Human mental workload, Advances in psychology,*, pages 139–183, 1988. 34

[48] JEFFREY HEER AND MICHAEL BOSTOCK. Crowdsourcing graphical perception: using mechanical turk to assess visualization design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 203–212. ACM, 2010. 12

[49] HERKKO HIETANEN, KUMARIPABA ATHUKORALA, AND ANTTI SALOVAARA. What's with the free images?: a study of flickr's creative commons attribution images. In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*, MindTrek '11, New York, NY, USA, 2011. ACM. 26

[50] DAMON HOROWITZ AND SEPANDAR D. KAMVAR. The anatomy of a large-scale social search engine. In *Proceedings of the 19th international conference on World wide web*, WWW '10, pages 431–440, New York, NY, USA, 2010. ACM. 13, 68

[51] JOHN JOSEPH HORTON AND LYDIA B. CHILTON. The labor economics of paid crowdsourcing. In *In Proceedings of the 11th ACM conference on Electronic commerce*, EC '10, pages 209–218., NY, USA, 2010. ACM, New York. 4

[52] JEFF HOWE. The rise of crowdsourcing. *Wired magazine*, **14**[6]:1–4, 2006. 4, 11

[53] Pei-Yun Hsueh, Prem Melville, and Vikas Sindhwani. Data quality from crowdsourcing: a study of annotation selection criteria. In *Proceedings of the NAACL HLT 2009Workshop on Active Learning for Natural Language Processing*, HLT '09, pages 27–35, Stroudsburg, PA, USA, 2009. Association for Computational Linguistics. 55

[54] Hilary Hutchinson, Wendy Mackay, Bo Westerlund, Benjamin B. Bederson, Allison Druin, Catherine Plaisant, Michel Beaudouin-Lafon, Stéphane Conversy, Helen Evans, Heiko Hansen, Nicolas Roussel, and Björn Eiderbäck. Technology probes: inspiring design for and with families. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '03, pages 17–24, New York, NY, USA, 2003. ACM. 42

[55] Panagiotis G. Ipeirotis. Analyzing the amazon mechanical turk marketplace. *XRDS: Crossroads, The ACM Magazine for Students*, **17**[2]:16–21, 2010. 12

[56] Grundspenkis J. Agent based approach for organization and personal knowledge modelling: knowledge management perspective. *Journal of Intelligent Manufacturing*, **18**[4]:451–457, 2007. 1

[57] Manyika James, Michael Chui, Brad Brown, Jacques Bughin, Richard Dobbs, Charles Roxburgh, and Angela Hung Byers. Big data: The next frontier for innovation, competition, and productivity. Technical report, McKinsey Global Institute, 2011. 3

[58] K. H. Jamieson and Joseph N. C. *Echo Chamber: Rush Limbaugh and the Conservative Media Establishment*. Oxford University Press, 2010. 95

[59] M. Jiang and W. L. McGill. Human-centered sensing for crisis response and management analysis campaigns. In *Proceeding of the 2010 Information Systems for Crisis Response and Management conference*, 2010. 90

[60] Stephen Jones and Peter Thomas. Empirical assessment of individual's personal information management systems. *Journal of Behavior and Information*, **16**[3]:158–160, 1997. 8

[61] Hee-Woong Kim, Yunjie Xu, and Joon Koh. A comparison of online trust building factors between potential customers and repeat customers. *ournal of the Association for Information Systems*, **5**[10]:392–420, October 2004. 77

[62] Jong Wook Kim, K Selçuk Candan, and Junichi Tatemura. Efficient overlap and content reuse detection in blogs and online news articles. In *Proceedings of the 18th international conference on World wide web*, WWW '09, pages 81–90. ACM, 2009. 11

[63] Aniket Kittur, Ed H Chi, and Bongwon Suh. Crowdsourcing user studies with mechanical turk. In *Proceedings of the SIGCHI conference on human factors in computing systems*, CHI '08, pages 453–456. ACM, 2008. 12

[64] Masashi Koga, Ryuji Mine, Tatsuya Kameyama, Toshikazu Takahashi, Masahiro Yamazaki, and Teruyuki Yamaguchi. Camera-based kanji ocr for mobile-phones: Practical issues. In *Proceedings of the Eighth International Conference on Document Analysis and Recognition*, ICDAR '05, pages 635–639, Washington, DC, USA, 2005. IEEE Computer Society. 68

[65] M. Krause, A. Takhtamysheva, M. Wittstock, and R. Malaka. Frontiers of a paradigm: exploring human computation with digital games. In *Proceedings of the ACM SIG- KDD Workshop on Human Computation*, HCOMP '10, 2010. 19

[66] Efimova L. Discovering the iceberg of knowledge work: A weblog case. In *In Proceedings of the Fifth European Conference on Organisational Knowledg, Learing and Capabilities*, 2004. 2

[67] Razmerita L, Kirchner K, and Sudzina F. Personal knowledge management: The role of web 2.0 tools for managing knowledge at individual and organisational levels. *Online Information Review*, **33**[6]:1021–1039, 2009. 2

[68] J.D. Lasica. When the studios won't give permission, darknet: Hollywood's war against the digital generation. www.darknet.com/2005/07/ when$_{t}he_{s}tudio.html$, 072005. 25

[69] L. Lessig. *Making Art and Commerce Thrive in the Hybrid Economy*. Penguin, New York, USA, 2008. 10, 23, 25

[70] P.N. Leval. Toward a fair use standard. *Harvard Law Review*, **103**[5]:1105–1136, 1990. 28

[71] Lawrence Liang. Guide to open content licenses. Technical report, Piet Zwart Institute, Institute for Postgraduate Studies and Research, Willem de Kooning Academy, 2005. 14

[72] Yefeng Liu, Vili Lehdonvirta, Mieke Kleppe, Todorka Alexandrova, Hiroaki Kimura, and Tatsuo Nakajima. A crowdsourcing based mobile image translation and knowledge sharing service. In *Proceedings of the 9th International Conference on Mobile and Ubiquitous Multimedia*, MUM '10, pages 6:1–6:9, New York, NY, USA, 2010. ACM. 57

[73] L.P. Loren. Building a reliable semicommons of creative works: Enforcement of creative commons licenses and limited abandonment of copyright. *George Mason Law Review*, **14**:271, 2007. 26

[74] Levy M. Web 2.0 implications on knowledge management. *Journal of knowledge management*, **13**[1]:120–134, 2009. 2

[75] Craig Macdonald and Iadh Ounis. Voting for candidates: adapting data fusion techniques for an expert search task. In *Proceedings of the 15th ACM international conference on Information and knowledge management*, pages 387–396. ACM, 2006. 13

[76] Craig Macdonald and Iadh Ounis. Voting for candidates: adapting data fusion techniques for an expert search task. In *Proceedings of the 15th ACM international conference on Information and knowledge management*, CIKM '06, pages 387–396, New York, NY, USA, 2006. ACM. 68

[77] Winter Mason and Siddharth Suri. Conducting behavioral research on amazon's mechanical turk. *Behavior research methods*, **44**[1]:1–23, 2012. 12

[78] Ian McGraw, Chia-ying Lee, I Lee Hetherington, Stephanie Seneff, and Jim Glass. Collecting voices from the cloud. In *LREC*, 2010. 12

[79] M. McPherson, L. Smith-Lovin, and J. M. Cook. Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, **27**:415–444, 2001. 95

[80] Kantar Media. Time to change the tune: Consumer research briefings on copyright. Technical report, Technical Report, Consumer Focus., London, UK., 2009. 10

[81] Audris Mockus, Roy T. Fielding, and James D. Herbsleb. Two case studies of open source software development: Apache and mozilla. *ACM Transactions on Software Engineering and Methodology*, **11**[3]:309–346, jul 2002. 83

[82] Andrés Monroy-Hernández, Benjamin Mako Hill, Jazmin Gonzalez-Rivero, et al. Computers can't give credit: How automatic attribution falls short in an online remixing community. In *Proceedings of the 2011 annual conference on Human factors in computing systems*, CHI '11, pages 3421–3430. ACM, 2011. 11

[83] E. Navas. *Regressive and Reflexive Mashups in Sampling Culture.* Number 157-177. Springer, 2010. 14, 23

[84] Jeffrey Nichols, Jalal Mahmud, and Clemens Drews. Summarizing sporting events using twitter. In *Proceedings of the 2012 ACM international conference on Intelligent User Interfaces*, IUI '12, New York, NY, USA, 2012. ACM. 90

[85] Ikujiro Nonaka and Hirotaka Takeuchi. *The knowledge creating company: how Japanese companies create the dynamics of innovation.* New York: Oxford University Press, 1995. 8

[86] Toyama R. Nonaka, I. and N. Konno. Seci, ba, and leadership: a unified model of dynamic knowledge creation'. *Long Range Planning*, **33**:5–34, 2000. 8

[87] Seneviratne O, Kagal L, and Berners-Lee T. Policy-aware content reuse on the web. *The Semantic Web-ISWC.*, pages 553–568., 2009. 10, 11

[88] R. Oldenburg. *The Great Good Place: Cafes, Coffee Shops, Bookstores, Bars, Hair Salons, and Other Hangouts at the Heart of a Community.* Da Capo Press, 1999. 15, 95

[89] P. Pirolli and S. Card. Information foraging. *Psychological Review*, **106**[4]:643–675, 1999. 95

[90] P. L. Pirolli. Exploring and finding information. *HCI Models, Theories, and Frameworks: Toward a Multidisciplinary Science.*, pages 157–191, 2003. 41, 42, 49

[91] Alexander J. Quinn and Benjamin B. Bederson. Human computation: a survey and taxonomy of a growing field. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 1403–1412, New York, NY, USA, 2011. ACM. 4

[92] J. Rekimoto. *Navicam: A palmtop device approach to augmented reality.* Laurence Erl- baum Associates, New Jersey, 2001. 52

[93] Takeshi Sakaki, Makoto Okazaki, and Yutaka Matsuo. Earthquake shakes twitter users: real-time event detection by social sensors. In *Proceedings of the 19th international conference on World wide web*, WWW '10, New York, NY, USA, 2010. ACM. 90

[94] O. Seneviratne, L. Kagal, and T. Berners-Lee. Policy- aware content reuse on the web. in. , (2009), 553-568. In *Proceedings of the 8th International Semantic Web Conference*, ISCW '09, pages 553–568, 2009. 24

[95] Oshani Seneviratne and Andres Monroy-Hernandez. Remix culture on the web: A survey of content reuse on different user-generated content websites. *Web Science (Raleigh, NC*, 2010. 10

[96] Pravin Shankar, Woody Huang, Paul Castro, Badri Nath, and Liviu Iftode. Crowds replace experts: Building better location-based services using

mobile social network interactions. In *2012 IEEE International Conference on Pervasive Computing and Communications*, PerCom 2012. IEEE, 2012. 86

[97] Shuming Shi, Huibin Zhang, Xiaojie Yuan, and Ji-Rong Wen. Corpus-based semantic class mining: distributional vs. pattern-based approaches. In *Proceedings of the 23rd International Conference on Computational Linguistics*, COLING '10, pages 993–1001, Stroudsburg, PA, USA, 2010. Association for Computational Linguistics. 42, 43

[98] Ian Smith, Sunny Consolvo, Anthony Lamarca, Jeffrey Hightower, James Scott, Timothy Sohn, Jeff Hughes, Giovanni Iachello, and Gregory D. Abowd. Social disclosure of place: from location technology to communication practices. In *Proceedings of the Third international conference on Pervasive Computing*, PERVASIVE '05, pages 134–151, Berlin, Heidelberg, 2005. Springer-Verlag. 84

[99] Alexander Sorokin and David Forsyth. Utility data annotation with amazon mechanical turk. In *Computer Vision and Pattern Recognition Workshops, 2008. CVPRW'08. IEEE Computer Society Conference on*, pages 1–8. IEEE, 2008. 12

[100] C. Soukup. Computer-mediated communication as a virtual third place: building oldenburg's great good places on the world wide web. *NEW MEDIA AND SOCIETY*, **8**[3]:421–440, 2006. 95

[101] J. Surowiecki. *The Wisdom of the Crowds*. Anchor Books, New York, NY, USA, 2005. 2

[102] Karen P. Tang, Jialiu Lin, Jason I. Hong, Daniel P. Siewiorek, and Norman Sadeh. Rethinking location sharing: exploring the implications of social-driven vs. purpose-driven location sharing. In *Proceedings of the 12th ACM international conference on Ubiquitous computing*, Ubicomp '10, New York, NY, USA, 2010. ACM. 85

[103] Kathleen Tuite, Noah Snavely, Dun-yu Hsiao, Nadine Tabing, and Zoran Popovic. Photocity: training experts at large-scale image acquisition

through a competitive game. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 1383–1392. ACM, 2011. 12

[104] L. von Ahn. Games with a purpose. *IEEE Comput. Mag.*, **39**[6]:92–94, 2006. 12, 18

[105] Luis von Ahn. Human computation. In *Proceedings of the 4th international conference on Knowledge capture*, K-CAP '07, 2007. 3, 5, 69

[106] Luis Von Ahn and Laura Dabbish. Labeling images with a computer game. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '04, pages 319–326. ACM, 2004. 12

[107] Luis Von Ahn, Benjamin Maurer, Colin McMillen, David Abraham, and Manuel Blum. recaptcha: Human-based character recognition via web security measures. *Science*, **321**[5859]:1465–1468, 2008. 12

[108] Jones W. *Keeping Found Things Found: The Study and Practice of Personal Information Management: The Study and Practice of Personal Information Management*. Morgan Kaufman, 2010. 1

[109] Fei-Yue Wang, Kathleen M. Carley, Daniel Zeng, and Wenji Mao. Social computing: From social informatics to social intelligence. *Intelligent Systems, IEEE*, **22**[2]:79–83, 2007. 3

[110] Kirby Wright. Personal knowledge management: supporting individual knowledge worker performance. *Knowledge Management Research and Practice*, **3**[3]:156–16, 2005. 2, 8

[111] Tingxin Yan, Vikas Kumar, and Deepak Ganesan. Crowdsearch: exploiting crowds for accurate real-time image search on mobile phones. In *Proceedings of the 8th international conference on Mobile systems, applications, and services*, MobiSys '10, pages 77–90, New York, NY, USA, 2010. ACM. 68

[112] R.B. Zajonc. Social facilitation. *Science*, **149**[3681]:269–274, 1965. 18, 77

[113] Fan Zhang, Shuming Shi, Jing Liu, Shuqi Sun, and Chin-Yew Lin. Nonlinear evidence fusion and propagation for hyponymy relation mining. In

Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies - Volume 1, HLT '11, pages 1159–1168, Stroudsburg, PA, USA, 2011. Association for Computational Linguistics. 42, 43

[114] G. Zichermann and J. Linder. *Game-Based Marketing: Inspire Customer Loyalty Through Rewards, Challenges, and Contests.* Wiley, London, 2010. 18

[115] M. Zyda. From visual simulation to virtual reality to games. *Computer*, **38**[8]:25—32, 2005. 18

Publication List

Journal Papers

1. Yefeng Liu, Vili Lehdonvirta, Todorka Alexandrova, Tatsuo Nakajima. Drawing on mobile crowds via social media. *Multimedia systems*, 18[1]:53-67, 2012.

2. Xiaodong Lu, Yefeng Liu, Tatsuya Tsuda, Kinji Mori. Autonomous Node Allocation Technology for Assuring Heterogeneous Streaming Service under the Dynamic Environment. *IEICE TRANSACTIONS on Communications*, E94-B[1]:30-36, 2011.

Conference Papers

1. Yefeng Liu, Todorka Alexandrova, and Tatsuo Nakajima. 2013. Using stranger as sensors: temporal and geo-sensitive question answering via social media. In *Proceedings of the 22nd international conference on World Wide Web* (WWW '13). International World Wide Web Conferences Steering Committee, 803-814. [Best Student Paper Nominee]

2. Yefeng Liu, Darren Edge, and Koji Yatani. 2013. SidePoint: a peripheral knowledge panel for presentation slide authoring. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '13). ACM, New York, NY, USA, 681-684.

3. Yefeng Liu, Todorka Alexandrova, Satoshi Hirade, and Tatsuo Nakajima. 2013. Facilitating natural flow of information among "taste-based" groups. In *Proceedings of the 2013 ACM annual conference extended abstracts on Human Factors in Computing Systems Extended Abstracts* (CHI EA '13). ACM, New York, NY, USA, 871-876.

4. Mizuki Sakamoto, Tatsuo Nakajima, Yefeng Liu, and Todorka Alexandrova. 2013. Achieving sustainable society through micro-level crowdfunding. In *Proceedings of the 2013 ACM annual conference extended abstracts on Human Factors in Computing Systems Extended Abstracts* (CHI EA '13). ACM, New York, NY, USA, 1443-1448. .

5. Herkko Hietanen, Antti Salovaara, Kumaripaba Athukorala, and Yefeng Liu. 2012. ¡insert image¿: Helping in the legal use of open images. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '12). ACM, New York, NY, USA, 599-608.

6. Francisco Lepe Salazar, Tetsuo Yamabe, Todorka Alexandrova, Yefeng Liu, and Tatsuo Nakajima. 2012. Family interaction for responsible natural resource consumption. In *Proceedings of the 2012 ACM annual conference extended abstracts on Human Factors in Computing Systems Extended Abstracts* (CHI EA '12). ACM, New York, NY, USA, 2105-2110.

7. Yefeng Liu, Vili Lehdonvirta, Mieke Kleppe, Todorka Alexandrova, Hiroaki Kimura, and Tatsuo Nakajima. 2010. A crowdsourcing based mobile image translation and knowledge sharing service. In *Proceedings of the 9th International Conference on Mobile and Ubiquitous Multimedia* (MUM '10). ACM, New York, NY, USA, 9 pages

Workshop Papers & Posters

1. Satoshi Hirade, Ping-Hui Lin, Han Chen, Yefeng Liu, Todorka Alexandrova, and Tatsuo Nakajima. Case Study on Crowdsourcing for Serendipity. *IE-*

ICE Technical Report, Artificial Intelligence and Knowledge-based Processing, 111[447]:1-4, 2012.

2. Yefeng Liu, Todorka Alexandrova, and Tatsuo Nakajima. 2011. Gamifying intelligent environments. In *Proceedings of the 2011 international ACM workshop on Ubiquitous meta user interfaces* (Ubi-MUI '11). ACM, New York, NY, USA, 7-12.

3. Yefeng Liu, Vili Lehdonvirta, Todorka Alexandrova, and Tatsuo Nakajima. Real-time Crowd Computing using Social Media. Poster Session at *The 2nd Crowdsourcing Conference* (CrowdConf '11)

4. Yefeng Liu, Todorka Alexandrova, Tatsuo Nakajima, and Vili Lehdonvirta. 2011. Mobile Image Search via Local Crowd: A User Study. In *Proceedings of the 2011 IEEE 17th International Conference on Embedded and Real-Time Computing Systems and Applications - Volume 02* (RTCSA '11), Vol. 2. IEEE Computer Society, Washington, DC, USA, 109-112.

5. Yefeng Liu, Vili Lehdonvirta, Todorka Alexandrova and Tatsuo Nakajima. Real-time Crowd Computing. Poster Session at *The 8th Global COE International Symposium on Ambient SoC* [Best Poster Award]

6. Yefeng Liu. A User Study of Crowd Powered Mobile Image Search. Poster Session at the *Symposium on Interaction with Smart Artifacts, Bilateral DFG-Symposium between Japan and Germany*, March 7-9, 2011, Tokyo, Japan.

7. Yefeng Liu, Vili Lehdonvirta, T. A. M. L., and Nakajima, T. Engaging socal medias - case mobile crowdsourcing. In *SoME 2011: International Workshop on Social Media Engagement*. Workshop at International World Wide Web Conference 2011 (WWW 2011).

Monographs

1. Yefeng Liu. *Service Composition Workspace - a Semantic Web Approach*. ISBN 978-3-639-37740-8, VDM Publishing, Saarbrcken Germany, August 24th, 2011