MarkIt: Community Play and Computation to Generate Rich Location Descriptions through a Mobile Phone Game

Ketan Patel, Mohamed Ismail, Sara Motahari, David J. Rosenbaum, Stephen T. Ricken, Sukeshini A. Grandhi, Richard P. Schuler and Quentin Jones
{krp8, mxi8616, sg262, djr8, sr82, sg49, rps22, quentin.jones}@NJIT.edu
New Jersey Institute of Technology

Abstract

The capture and description of the numerous 'places' found in urban environments could enable the deployment of seamless mobile and location-aware social navigation tools. Unfortunately, we are limited in our approaches to systematically capture place data at the requisite level of granularity or how to obtain labeling consensus. We explored this issue through 'research through design', developing and deploying the MarkIt mobile phone game, which aimed to engage a community of users in place description. In this paper, we present our design rationale and the game's design. We then highlight how key innovations instantiated in MarkIt resulted in rapid data generation of both images and location related labels.

1. Introduction

Developing mobile applications that improve people's navigation of their social environment so that they can more easily engage in activities of interest, meet new people and coordinate interaction with others, is a seemingly straightforward way in which pervasive computing could improve social connectivity.

However actualizing this design task requires the overcoming of some enormous challenges, such as reliably leveraging rich information about the social geography - the people, the places and the connections between them - of any environment users wish to traverse. Key among these challenges is capturing detailed place information as it is not simply a matter of collecting geo-coded images, but understanding and representing how locations are used and understood by the community. For example, the green patch in a University campus may be viewed as simply a lawn, or alternatively as a personal hangout place by a group of friends, an open air theater by a theatrical club, or even weekly ultimate Frisbee challenge field. It is not hard to imagine how such different interpretations of locations could then be utilized by mobile navigation tools to inform individuals about where to find and engage in social activities of interest. Consequently, place identification and description is an important issue for ubiquitous computing researchers to resolve.

In this paper we describe our efforts aimed at addressing the challenge of place description for a large-scale mobile social computing test-bed through the development and deployment of the MarkIt mobile phone game. We adopted a *research through design* approach due to the nature of the problem of place description, which can be considered 'wicked' [27], with 'no single true solution'. The paper begins with an overview of this research approach, before presenting the related background literature upon which we based MarkIt and a discussion of related recent work on mobile multiplayer games. We then present MarkIt's design rationale, interaction design and results from our field deployments. We conclude with a discussion of the lessons learned about the design challenges and tradeoffs.

2. Research through Design

Building on Frayling [9], Zimmerman, Forlizzi and Evenson [34] formalized the *research through design* paradigm. They argued that significant contributions to HCI could be made by designers producing software applications that integrate in a novel fashion previous research in an attempt to make "a product that transforms the world from its current state to a preferred state". This differs from the more formalized notion of *design science in information systems* [19], which is fundamentally an organizational "problem solving paradigm" (page 76). Rather than placing an emphasis on 'science' and 'scientific design', *research through design* focuses on reflection upon the natural process associated with artifact design and deployment.

The need for *research through design* follows from the arguments made by Rittel and Weber [27] that many problems are 'wicked' in that they cannot be accurately modeled or solved through hypothesis testing or engineering approaches. This is because the problem being studied is under-defined, with many possible and contradictory solutions. This provides an opportunity for *research through design* to provide complementary knowledge by proposing 'better' as

IEEE computer society
978-0-7695-3869-3/10 $26.00

opposed to ‘true’ solutions (scientifically validated) to complex real world problems.

The challenge of designing and deploying an application that identifies and describes the places in an urban environment is fundamentally under constrained and therefore an ideal candidate for *research through design*. As the following background sections show, there are a diverse set of solutions / approaches to capture place data, each with its pros and cons. As Rittel and Weber note [27], wicked problem solutions are ‘one-shot efforts’, which in the HCI context can be translated as design solutions for a particular time and place, with unique environmental characteristics. This makes notions of scientific replication impractical.

Considering the nature of the *research through design* paradigm, an alternative evaluation lens for such work is also required. Zimmerman et al. suggest (i) process (design rationale), (ii) invention (innovative combination of features), (iii) relevance (addressing real world problems), and (iv) extensibility (exemplars for future research and design). As a result we focus in this paper on these aspects of MarkIt games design, development and deployment.

3. Background

MarkIt was designed and developed by building upon previous work on place identification and description [1][6] and user generated content that can be linked to place [24][30]. We will briefly review these two areas before relating it to three mobile multiplayer games that try to address the issues of place identification, description and user generated content but differ in their design and implementation.

3.1 The Challenge of Place Description

People through their activities both individually and collectively, create “places”, locations imbued with social meaning. Some of these places are intimate and personal (e.g. “the ‘gamers’ hangout”, or “Mike’s dorm room”) while others are understood by communities in terms of ‘place-types’ – coffee shop, museum, school, etc. Environmental psychologists [8], sociologists [10], and more recently computer scientists [14], have recognized the extent to which place and place-type labels inform individuals about appropriate social communication and activities. As a result, the effective digitization (computerized data capture) of place information requires the engagement of those people who occupy the relevant environments.

The most obvious way to enable place digitization is through mobile-device enabled *in situ* data capture. This is now possible with the wide spread use of technologies such as camera cell phones [25], GPS [17] and Wi-Fi [13]. In addition, we would ideally want to leverage people’s movement through the physical world to prompt the identification of places they find personally important. Previous work in this area has focused on the analysis of individual’s mobility traces/histories to identify places of personal [17] or group relevance [13]. Interestingly, the results of such efforts show that while people traverse many places, they only rate and label a handful of locations as personally meaningful [16]. A number of commercial and research [3][4] applications have been deployed that support the capture of user generated place data. These have focused on either location-linked tagging of user generated content [22][12], or the analysis of user-location histories [34].

3.2 User Generated Content Related to Place

Various web sites that allow users to upload and manage their pictures (e.g., flickr.com) have managed to collect significant place information through the geo-coding of images and picture tagging. Unfortunately, the tagging of pictures taken by consumers for personal use typically results in data sets that lack geographic density. For example, web sites such as Flickr.com, which allow users to tag photos with place labels, have generated significant content but rarely describe more than a small proportion of the places for any urban locale. This is because routine picture taking does not typically result in the capture of information about everyday places, instead, images and text about places tend to focus on well-known landmarks [26][21]. In addition, descriptions of images tend towards personally significant tags [12] (e.g. Mike's party) that support self-organization and recall [24]. This limits the utility of tags as general place descriptions. Another challenge with the generation of detailed place descriptions from consumer use of picture tagging sites is that users typically tag and describe images hours or days after their capture. As a result, many images are never tagged and those that are suffer issues of recall accuracy. A number of research applications have explored in-situ picture tagging [1] [25] finding that it allows for more accurate and useful tags.

Similarly, previous experimental systems that have utilized individual’s mobility traces (location histories) generated by various people tracking technologies to prompt identification places has resulted in the sparse identification of locations. This outcome results from the fact that individuals only consider a handful of locations they traverse each day to be meaningful enough to warrant description [17].

3.3 Mobile Multiplayer Games Enabling User Generated Place Content

Gaming has been applied to the systematic generation of relevant user content. One well known example, the ESP game [32], was able to describe images more efficiently than alternative computerized approaches by having pairs of players to guess each other's picture labels. This process of using people's natural abilities to carry out tasks that computers cannot do easily is referred to 'human computation' [11][32]. In a sense, the emergent properties of large scale community activities such as tagging on Flickr are a form of human computation. However, the gaming approach is interesting in its ability to customize and optimize the type of content generated while making it enjoyable and motivating for content providers.

Recently, human computation has been put to use to generate place related information in two mobile multiplayer games: CityExporer [23] and EyeSpy [3]. Both these games utilize player's common sense and local knowledge of a place to generate location/place descriptions however they differ in their design and hence the nature of data generated.

CityExplorer allows players to defined place categories (e.g., food, art) and then create and place markers in those categories. A marker is a photograph that is given a name (tag), category and location during game play. The setting of a marker involves taking a photo, approaching the location as close as possible and then selecting on one of the pre-defined location categories (such as "food"). Gamers then in pairs confirm the accuracy of markers placed in place categories through an associated website. Players are motivated by points in two ways 1) the number of markers they place and 2) the correctness of the placement as judged by a maximum of two fellow players. Fourteen players in Bamberg collected a total of 772 markers over a period of 20 days. It is not clear how many of the markers were confirmed as the authors note that the online part of CityExplorer was not perceived as entertaining and that the reviewing step was not well received by players.

EyeSpy is another location-aware game in which players mark geographic locations with photos and text. The accuracy of player content could be confirmed through the game play of one other player who was physically present at the location in question. Players are motivated by points for 1) the number of pictures/text tags created, 2) number of confirmations made on other's tags and 3) number of confirmations they receive on tags they created. Each player randomly received 5 photo and 5 text tags to confirm per day. Each image could be confirmed once. The game was run over two separate rounds. In the first round 9 participants played for one week, with the subsequent 9 participants playing a second round lasting two weeks. Over the three weeks 453 tags were collected, out of which 257 were photos and 196 were text tags. Pictures collected through EyeSpy were overlaid on maps by the developers and then effectively used to help visitors with navigation.

While both these games resulted in user-generated place content in terms of geo-coded pictures, text and tags, their designs constrained the quantity and quality in three fundamental ways. First, each of the games limited the content that could easily be generated by players: CityExplorer's maps reached saturation point after only a small number of player actions in close proximity; EyeSpy game players were limited to confirming 5 photo and 5 text tags each day; CityExplorer limited place categories to short list provided by developers; and both games required the player's physical presence in a particular location for the generation of user-content.

Second, the quality of text descriptions generated by players fell short of providing useful place descriptions. This was noted by the EyeSpy team who stated, "*we found that the way in which text tags were crafted by players, such as the creation of riddle tags and lack of high enough density to provide appropriate search terms, resulted in less than 20% of the closet text tags to images providing relevant descriptions.*"

Finally, the user content-validation process was severely constrained in both these games. In CityExplorer the placing of the markers in a place category was considered correct when a fellow player judged it correct or incorrect if two players deem it incorrect. Similarly, in EyeSpy, only one other player was required for validation (limited to a maximum of 10 per user, per day).

These two systems demonstrate the potential ability of mobile gaming applications to produce useful user content. However, alternative designs need to be presented that address the limitations they impose on content generation and validation.

4. Design Goals and Rationale

Previous *research through design* approaches to providing a solution to the problem of capturing large scale meaningful 'place' data suggests four design elements that would provide a novel solution if instantiated in a single application. These are:

1. Allow all players to provide multiple pictures and multiple textual descriptions of places. This should help generate large volumes of the place content.
2. Game-play rules should encourage players to provide textual descriptions that can be agreed

upon by a large number of other players. This should help generate meaningful descriptions.

3. Ensure that the process of generating both content and validating content is entertaining and fun through creative game design and game-play rules.
4. Provide players with a sense of community that will encourage content generation and validation.

The context for this *research though design* project was to partially address the problem of place description for a large-scale university mobile social-computing test-bed, designed for the study of People-to-People-to-Places (P3) services [16]. The central feature of the test-bed is a suite of mobile social computing applications, supported by a unified middleware that share and support a single rich model of people and places within the New Jersey Institute of Technology (NJIT) campus. Situated in downtown Newark, the NJIT campus is geographically compact, enabling blanket Wi-Fi coverage for connectivity and locatability.

Beyond the goal of large volumes of rich place content around the NJIT campus the applications we deployed had to run on the hundreds of mobile smartphones that we were deploying in 2007-2008 to create our test-bed (primarily HTC designed AT&T model 8925 and 8525 PocketPC phones, with touch screens, cameras, Wi-Fi and keyboard, running Windows Mobile 5.x/6.x). Further, the application running on our cell phone had to be very easy to use, which in turn resulted in the following self imposed design constraints:

a. Main user interaction had to be achieved through one-handed actions (based on our early user studies of our HTC devices). This lead us to develop a thumb-based, rather than stylus-based interface with large buttons; and
b. Many of the previously developed interfaces for mobile picture taking and tagging typically take five or six steps to finish the process. In order to unburden our users, we decided that each action would have to have a maximum of two to three steps for completion.

5. The MarkIt Game

Based on the design goals and rationale we developed 'MarkIt' to instantiate our solution to the wicked problem of place description for our campus. MarkIt is a multiplayer mobile phone game that makes it fun for users to digitally describe their physical surroundings. Players take pictures (SnapIts), which are uploaded to the MarkIt server and displayed on each player's device. Players view all pictures and have one chance with each image to perform the following actions: TagIt (label it with multiple "tag" words and MapIt (map it to where they feel the picture was taken).

The goal of the game is for the player to guess where each image was taken and what tags were placed on the image by other players. Matching tags and MapIts with other players results in users gaining points. The points system has three components:

Action Points: These points are given to the players any time they SnapIt, TagIt, or MapIt, giving an instant reward as an incentive for participation.

Matching points: These points are given to players who enter the same tag or MapIt as a previous player for a particular image.

Royalty Points: Ownership is awarded to a player in three situations: a) snapping a picture; b) being the first to submit a unique tag; or c) being the first to place an image at a particular map location. Owners gain royalty points when other players match the owner with respect to TagIts or MapIts.

Our objective as developers in employing this points system was to maximize the likelihood of rich descriptions of locations around campus that are agreed up on by the player community.

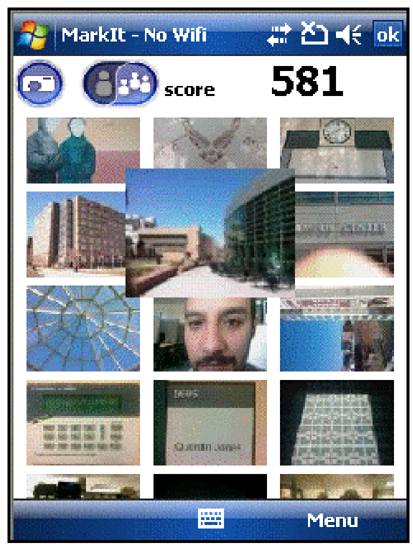

Figure 1: MarkIt landing page

5.1 User Interfaces and Game Play

MarkIt game activity revolves around a dynamic and intuitive landing page (Figure 1) consisting of recent images generated by the player community. The full landing page contains sixteen three-image rows (48 images in total), four rows of which (12 images) can be viewed on screen at a time. The images are sorted by creation date/time with the most recent images appearing at the beginning (top) of the landing page.

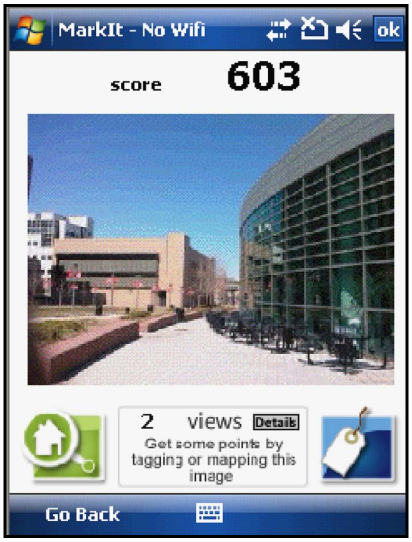

Figure 2: MarkIt image page

An image will disappear off the landing page when it has been both tagged and mapped by the player or when new material generated by the player community pushes older images off the bottom of the page. Once an image is tagged, mapped, and removed from the landing page, an older untagged or unmapped image will be added at the end to replace it. The landing page is thumb-scrollable and the images it contains can be enlarged by a single tap. Double-tapping on an image moves the player to the image page (Figure 2) where they then can decide to tag and/or map the image.

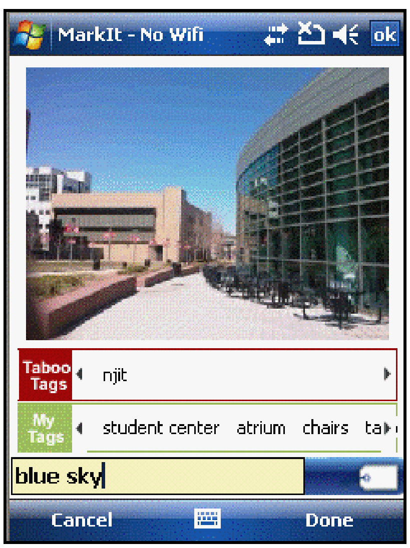

Figure 3: MarkIt TagIt page

The landing page achieves two key design objectives. First, drawing users into game play: Like *Tetris*, where the flow of shapes down the screen encourages play, the flow of images down the MarkIt landing page encourages users to act on images quickly and often. Second, it provides users with a sense of campus community. Various studies of online communities highlight how the sharing of content builds bonds between users. In the case of MarkIt, players share images of the campus, providing the "community content" for game play. In addition, the resulting stream of images from around the campus on the landing page provides players with a visualization of a "community in motion." The emergent qualities of mobile picture image streams has been noted by others, sometimes referred to as "constant photo eye" [6], or media "co-presence" [15], however in the case of MarkIt, the geographical focus on a single downtown campus clearly connected all the players within our campus community.

Tagging of images is carried out through the TagIt user interface (Figure 3). Two types of tags were disallowed: 1) terms that might be considered offensive to other players such as swear words and 2) "taboo" terms that we felt would be universal to our images such as the University name and words such as "image", "picture", etc.

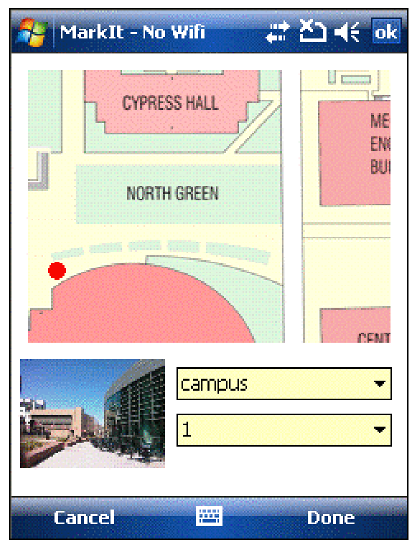

Figure 4: MarkIt MapIt page

The players associated images with location through the MapIt user interface (Figure 4), which supported various levels of resolution and drill down ranging from a campus landscape map to detailed floor plans of each campus building.

When the MarkIt game is open, pressing the camera button on the phone opens the camera utility. Taking a picture results in a SnapIt action that directly uploads the image to the game and loads the image page (Figure 2) of the picture taken. From the image page players can then easily TagIt or MapIt.

The MarkIt game uniquely combines the following set of features: 1) mobile in-situ data capture of user generated content (primarily images and location-activity data captured by our location engine that notes where users are playing and stored in the MarkIt middleware); 2) multiplayer actions on user content; 3) "community in motion" visualization; 4) asynchronous yet dynamic game play (tagging and mapping can occur anywhere and anytime a player is on campus); and 5) a focus on human computation for user content that not only motivates game play but also aids in generation of significant content. Large-scale play of MarkIt aims to product user generated: i) images; ii) tags of images; iii) user estimates of where images should be located; iv) consensus regarding image tags through matching; v) consensus measures of image locations through matching; and vi) activity-location data for game play from our location engine.

6. MarkIt Pilot Deployment and Game Rounds

After initial testing of the core set of MarkIt user interfaces with a small group of users, during March-April 2008 we explored large-scale game play through 3 rounds of deployment (see Table 1.). The aim of our pilot deployment was to:

1) Ensure that our middleware was able to simultaneously locate a large number of wireless users and reliably capture user actions;
2) Assess the overall game design through play on a realistic scale. Key game features such as the "community in motion" landing page and matching data were expected to emerge only from large-scale play with a large number of users; and
3) Establish a MarkIt game player community. The MarkIt game assumes a critical mass of players who are able to act on other players' content. We therefore had to ensure that we would have the requisite number of committed players for our game rounds.

Pilot study gamers were part of a larger pool of student-volunteers involved in various our test-bed usability and beta-test opportunities. Over a 5-day work week, 150 students were given PocketPC phones with an alpha version of MarkIt installed and were allowed to play the game until the last night of the pilot deployment week. Individuals piloting the game knew that the game scores and content would be reset at the end of the pilot phase. Pilot data informed the point structure of the subsequent game rounds. During our pilot deployment we were at first delighted to see how quickly the landing page came to visualize our campus community in motion, with the rapid streaming of a large number of images. However, it soon became apparent that the speed with which images were being forced off the landing screen greatly reduced the potential for matched tags. As a result, we changed our point structure for Round 1, increasing matching and Royalty points and reducing SnapIt points.

Pilot Deployment (1 week)			
150 Testers/Players	SnapIt	TagIt	MapIt
Action Points	5	5	10
Matching Points	N/A	1	2
Ownership Royalty Points	1	1	1
Game Round 1 (2 weeks)			
108 individuals installed software	SnapIt	TagIt	MapIt
Action Points	1	5	15
Matching Points	N/A	10	15
Ownership Royalty Points	2	3	4
Game Round 2 (1 week)			
96 Players from Round 1	**SnapIt**	**TagIt**	**MapIt**
Action Points	1	5	15
Matching Points	N/A	5	15
Ownership Royalty Points	2	3	4

Table 1. Deployment Lengths and Points Structure

In Round 1, the participants were given the opportunity to play the game for two weeks, which included a prize giveaway for the high scorers. One hundred and eight beta-testers opted into Round 1 by independently downloading and installing the MarkIt beta software onto their phone. While this round generated higher proportions of matching actions with respect to tags, we wanted our players to also map a higher proportion of images to help us more effectively characterize places. As a result, for the final game Round 2 which ran for 1 week we experimented with lowering the points gained for matching tags.

For each of the rounds players were told that if any form of cheating was found, they would not be able to continue playing. To encourage privacy, gamers were also told that taking photos of people rather than locations was against the rules.

6.1 Observations from the Field

During the pilot deployment approximately 50% of testers actively played the game for more than one day. This outcome was expected because of our subject recruitment method (giving the application to all students who expressed a general interest in our test-bed as opposed to the MarkIt game in particular) and the technical problems with our software (including numerous bugs which only became apparent with large numbers of users and the need for players to install updates during the pilot phase). In each of the two game rounds the percentage of active users was around 65%. For each of the three deployment phases the top 60 players accounted for 95% of user activity. As we were interested in the actions of those individuals who chose to play our game and were willing to work with us to overcome teething problems we choose to report on the activities of the top 60 players of the pilot deployment and two game rounds.

On examining the data from the field, we found that our brief and moderate deployments produced a large amount of valuable content. From the top 60 players in each of the three rounds, we collected 70,077 tags from 10,686 TagsIts, of which 48,186 were unique image tags, 21,891 were matching tags, with 9,806 unique tags for the game as a whole. Users snapped 3,122 images, with 18,739 image views. There were 7,077 MapIts, of which 1,408 were matching.

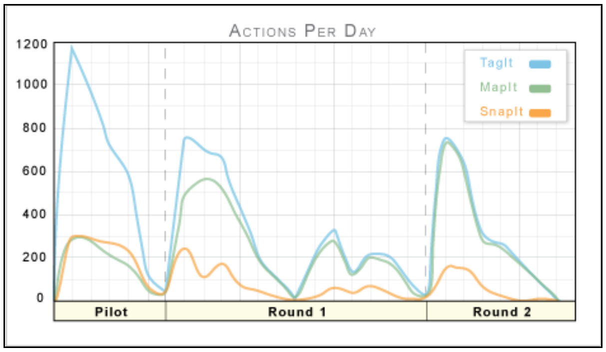

Figure 5. User TagIts, MapIts and SnapIt Actions Per Day for each Deployment

Figure 5 illustrates the impact of debugging problems we had with the MapIt user interface during our pilot deployment and the impact of the initial point structure which emphasized the snapping and tagging of images. In the two game rounds the relationship between SnapIts, MapIts and TagIts remained fairly constant despite our manipulation of the point structure.

In the pilot deployment, 24.3% of users' tags were matching as compared with round 1 where 35.5% were matching tags ($n=113$, $t=-4.1$ $df=111$, $p < 0.001$). This difference can be explained by the slow rate of images being added to the landing page, which gave players more time to view each image (4.47 views per image in the pilot in contrast with 8.54 views per image in round 1) and subsequently tag them (2.94 TagIts per image in round 1 vs 4.68 per image in round 1). Our minor changes in the point structure between rounds 1 and 2 did not produce statistically significant differences in action rates although the trend was in the desired direction. As the proportion of MapIt actions increased, the TagIt actions decreased. In round 1, 46.8% of all actions were TagIt actions while in round 2, 40.9% were TagIt actions ($n=85$, $df=83$, $t=2.0$, $p=0.053$).

On examining the type of images being captured, we found that public locations were more likely to be snapped, mapped and tagged. This is clearly illustrated in the MapIt data where we found that the vast majority of images were mapped to either our campus landscape map (31.3% for rounds 1 and 58.9% for round 2) and our large campus student-center containing multiple eating and social venues (36.4% for round 1 and 22.8% for round 2). This was not surprising as players were more likely to share knowledge of public spaces. This result also highlights the challenges associated with ensuring a critical mass of players with shared knowledge of various 'nooks and crannies' who can richly describe the paths less traveled.

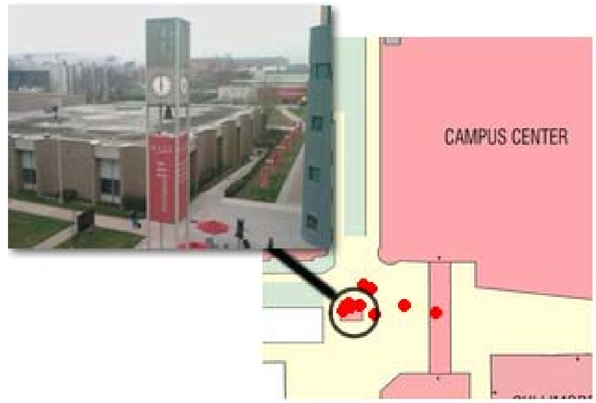

Figure 6. MapIt Image Cluster of the Most Tagged Image

When we looked at the tags of the images that were consistently mapped to specific locations, we found what can best be described as 'location tags'. Typically if we looked at the most tagged image in a location, its most popular tags generally matched the most popular location tags. Figure 6 and Table 2 illustrates this for "campus center bell tower" which was the most tagged image in the most mapped campus location.

Tags for Figure 6 SnapIts		Tags of Images Mapped in Figure 6	
Tag	**Count**	**Tag**	**Count**
clock	12	clock	44
time	8	clock tower	35
clock tower	8	tower	24
tower	7	bell	21
campus	5	bell tower	17
bell	4	campus	16
campus center	4	campus center	16

Table 2. Frequent Tags and Associated Location Tags of Most Tagged Image

Despite the quality of information that can be extracted from computing location tags, in general tags described images rather than locations. Looking at the 25 most frequent tags, we found that only six referred to a building or place name while the rest referred to colors, furniture, and fixtures. It should be noted that the "campus center", the busiest place on campus, was the most frequent tag, provided over 2000 times by players.

7. Post-Game Survey

The Markit game was designed in a way that multiple participants could confirm one participant's TagIt and MapIt action by matching their answer. The more that TagIts or MapIts were matched the greater the community consensus as to that tag or image location. We conjectured that higher consensus provides higher quality data and hypothesized that:

1. Images with higher mapping consensus better represents a specific locale than images without mapping consensus.
2. Popular tags of a singular image with high tagging and mapping consensus will better describe the place it belongs in than popular tags of images with high tagging, but low mapping consensus.

7.1. Method

We tested the above hypotheses in a post-game survey of participants at the end of the rounds of game play.

7.1.1. Subjects: Seventy nine game players responded to our email survey request and 76 of these completed relevant survey questions at the end of the second game round. Fifty eight respondents were between 18-21 years of age, the rest were over 21, and 16 of the respondents were women.

7.1.2. Procedure: As there were thousands of TagIts, MapIts and SnapsIts which we could use for this study, we decided focus on frequently tagged and mapped images, and frequently tagged locations. Adopting this approach we hoped would allow us to more robustly demonstrate any effect observed. We compared top tagged and mapped images, and top tagged locations (those with the highest 5% of matching TagIts and MapIts) with highly tagged and mapped images, and highly tagged locations (those whose frequency of matching TagIts and MapIts was in the top 10-15%). We selected at random two of the top scoring images in the three categories of high and low MapIts and TagIts matching. Our survey responses were obtained through the following steps:

- The subjects were shown an official place title and a 'top MapIt image' and were asked:
  "Does this picture belong in this place?" (y,n)
- The subjects were shown an official place title and one 'high MapIt image' and were asked:
  "Does this picture belong in this place?" (y,n)
- The subjects were shown an official place title and the 6 most frequent image tags for 'top MapIt image":
  "Here are labels used to describe this place. How well do these labels as a whole describe this place?" (They ranked the tag set on a 1-7 scale)
- The subjects were shown an official place title and the 6 most frequent image tags for 'high MapIt image":
  "Here are labels used to describe this place. How well do these labels as a whole describe this place?" (They ranked the tag set on a 1-7 scale)

7.2. Results

Analysis of the responses to our survey confirmed both our hypotheses. Images with higher mapping consensus better represented a specific locale than images with lowing mapping consensus (Wilcoxon signed–rank test, $n=69$, $Z=-2.91$, $p=0.021$). Popular tags of an image with high tagging and mapping consensus better describe the place it belongs in than popular tags of images with high tagging, but low mapping consensus (Wilcoxon signed-rank test, $n=57$, $Z=-3.279$, $p=0.001$).

8. Research through Design Outcomes

Overall our brief MarkIt game deployments were highly successful, producing a large amount of valuable content. From the top 60 players in each of the three rounds, we collected 70,077 tags, 21,891 matching tags, 3,122 images, with 18,739 image views, 7,077 MapIts, and 1,408 MapIt matches. The result contrasts greatly with EyeSpy and CityExplore in that each were able to generate only a few hundred tags.

Of course user generated location data was only one of outcomes of MarkIt game deployment. Our post game survey responses suggest that a large number of our players made new friends on campus as a direct result of game play (this occurred both among players and between players and other curious non-players of the campus community). In addition, game deployment resulted in the positive branding of our test-bed.

We were also able to show that higher rates of mapping and tagging matching better represented a specific location than those with only slightly lower frequency of matches. This again highlights the comparative utility of our design to other more recently developed mobile games which restricted the extent to which users could confirm the validity of user-generated content.

In the process of prototyping and deploying MarkIt, our research artifact, we gained understanding of a number of design challenges and tradeoffs.

An ongoing design challenge is how to structure the game points system to maximize the mix of content generation and player satisfaction. We found that awarding more points for SnapIts relative to matching and royalty points provided incentive to taking of pictures while the reverse resulted in the players focusing on the tagging and mapping of images.

The picture generation rate is related to an inherent conflict between the 'community in motion' visualization and rate of community matching. This is because the faster the flow of images on the community landing page, the more the UI provides a real time sense of community. Conversely, the number of players that act on an image will relate to the amount of time an image was on different player's landing screens and the opportunity for TagIts and MapIts. One solution to this tradeoff is to provide an alternative method for viewing and acting on images outside of current game parameters.

The current design also raises challenges in the interpretation of content generated. The community landing page presents the most recently uploaded, untagged and/or unmapped images to users. As a result, we are unable to gain an understanding of how images age and how the tagging of older images could inform our location labeling. One solution could be to mix old and new images on the landing page but this would effectively remove the 'community in motion" aspect.

There are also fundamental issues regarding the idea of matching being centered on image TagIts and MapIts. This may result in the descriptions of the objects that are repeatedly in a location, rather than usable place descriptions. Although at present our descriptions of locations is so limited that this narrower outcome may itself be an enormous step forward.

Finally, there is the issue of the social norms regarding picture taking that we need to embody in game play. It is important that people take photos of locations rather than people and understand that the tagging should be focused on place. In this regard with have the advantage of deployment for use on a university campus with university students who should respond to technology use policies/directives.

It is possible with each of the challenges presented above to imagine usability studies that provide insight into the value of various alternative designs.

Our research through design process allowed us to create an innovative mobile phone game that uniquely combined the following set of features:

- mobile in-situ data capture of user generated content;
- multiplayer actions on user content;
- "community in motion" visualizations;
- asynchronous yet dynamic game play (tagging and mapping can occur anywhere and anytime a player is on campus); and
- a game point structure that focus user action on content production that not only makes game play fun but provides an opportunity for deployers to manipulate the types and rates of content generated.

In so doing we were rapidly able to generate large numbers of geo-coded images and tags whose quality and relevance was assessed by the player community.

10. Acknowledgements

This research is partially supported by the National Science Foundation Grant NSF CNS 0454081 and DST 0534520. The opinions expressed are those of the authors and may not reflect those of the NSF.

11. References

[1] Ahern, S. Ahern, M. Davis, D. Eckles, S. King, M. Naaman, R. Nair, M. Spasojevic, J. Hui-I Yang. Zonetag: Designing context-aware mobile media capture to increase participation, In *Proceedings of PICS'06*, 2006, 3.

[2] Ashbrook, D. and Starner, T. Learning Significant Locations and Predicting User Movement with GPS, *In Proceedings of ISWC02,* 2002, 101-108.

[3] Bell, M., Reeves, S., Brown, B., Sherwood, S., MacMillan, D., Ferguson, J., and Chalmers, M. EyeSpy: supporting navigation through play, In *Proceedings of Conference on Human Factors in Computing Systems* Boston, MA, USA, 2009.

[4] Bellucci, A., Ghiron, S. L., Aedo, I., and Malizia, A. Visual tag authoring: picture extraction via localized, collaborative tagging, In *Proceedings of the working conference on Advanced visual interfaces (AVI '08)*, 2008.

[5] Bentley, F., Metcalf, C. Flexible Views: Annotating and finding context-tagged mobile content, In *Proceedings from the Eighth International Conference on Ubiquitous Computing*, 2006.

[6] Davis, M. Mobile Media Metadata: Metadata Creation System for Mobile Images, In *Proceedings of the 12th annual ACM international conference on Multimedia*, 2004, 396-397.

[7] Davis, M., Van House, N. A., Burgener, C., Perkel, D., King, S., Towle, J., Ahern, S., Finn, M., Viswanathan, V., and Rothenberg, M., MMM2: Mobile Media Metadata for Media Sharing, In *CHI '05 Extended Abstracts on Human Factors in Computing Systems*, 2005, 1335-1338.

[8] Franck, K. H., Schneekloth, L. *Ordering Space: Types in Architecture and Design*, Van Nostrand Reinhold 1994.

[9] Frayling, C. Research in Art and Design, *Royal College of Art Research Papers*, 1993, 1-5.

[10] Genereux, R., Ward, L., and Russell, J. The behavioral component in the meaning of places, *Journal of Environmental Psychology*, 3, 1983, 43-55.

[11] Gentry, C., Ramzan, Z., and Stubblebine, S. Secure distributed human computation, *EC '05: Proceedings of the 6th ACM conference on Electronic commerce*, 2005.

[12] Golder, S., Huberman, B., Usage patterns of collaborative tagging systems, *JIS'06*, 2006, 198-208.

[13] Gupta A., Paul S., Jones Q., and Borcea C., Automatic Identification of Informal Social Groups and Places for Geo-Social Recommendations, *International Journal of Mobile Network Design and Innovation, Special Issue on Interdisciplinary Issues in Telecom*, 2, 2007, 159-171.

[14] Harrison, S., Dourish, P. Re-place-ing space: the roles of place and space in collaborative systems, In *Proceedings of the 1996 ACM conference on Computer supported cooperative work*, 1996, 67-76.

[15] Ito, M. Intimate Visual Co-Presence. Pervasive Image, Capturing and Sharing Workshop, Ubicomp. Tokyo, Japan. September, 2005.

[16] Jones Q., Grandhi S.A, Terveen L., Whittaker S., and Chivakula K.. Putting Systems into Place: A Qualitative Study of Design Requirements for Location-Aware Community Systems, In *Proceedings of the ACM's Conference on Computer Supported Cooperative Work*, 2004.

[17] Jones, Q., Grandhi, S.A., Karam S., Whittaker S., Zhu C., and Terveen L. Geographic 'Place' and 'Community Information' Preferences, *Journal of Computer Supported Cooperative Work*, 17, 2008, 137-167.

[18] Joshi, D., Luo, J. Inferring Generic Activities and Events from Image Content and Bags of Geo-Tags, In *Proceedings of the international conference on Content-based image and video retrieval*, 2008, 37-46.

[19] Hevner, A., S. March, J. Park, and S. Ram. Design Science in Information Systems Research, MIS Quarterly (28:1), 2004, 75-105.

[20] Kennedy, L., Naaman, M., Ahern, S., Nair, R., Rattenbury, T., How flickr helps us make sense of the world: context and content in community-contributed media collections, In *Proceedings of the 15th International Conference on Multimedia*, 2007, 631-640.

[21] Kindberg, T., Spasojevic, M., Fleck, R. and Sellen, A. The Ubiquitous Camera: An In-Depth Study of Camera Phone Use, *IEEE Pervasive Computing* 4, 2, 2005, 42-50.

[22] Maarek, Y., Marmasse, N., Navon, Y., Soroka, V., Tagging the physical world, *Collaborative Web Tagging Workshop at WWW2006,* Edinburgh, 2006.

[23] Matyas, S., Matyas, C., Schlieder, C., Kiefer, P., Mitarai, H., and Kamata, M. Designing location-based mobile games with a purpose: collecting geospatial data with CityExplorer, In *Proceedings of the international Conference on Advances in Computer Entertainment Technology* (Yokohama, Japan). ACE '08, vol. 352. ACM, New York, NY, 2008, 244-247.

[24] Marlow, C., Naaman M., Davis M., and Boyd d., Why We Tag: Motivations for Annotation in Mobile and Online Media, In *Proceedings of the seventeenth ACM conference on Hypertext and hypermedia,* 2006, 971-980. New York: ACM Press.

[25] Naaman, M., Song, Y. J., Paepcke, A., and Garcia-Molina, H. Automatic organization for digital photographs with geographic coordinates, *In Proceedings of the 4th ACM/IEEE-CS Joint Conference on Digital Libraries*, 2004.

[26] Rader, E., Wash, R., Influences on Tag Choices in del.icio.us, In *Proceedings of the ACM 2008 conference on Computer supported cooperative work*, 2008, 239-248.

[27] Rittel, H.W.J., Webber, M.M. Dilemmas in a General Theory of Planning, *Policy Sciences* 4, 2, 1973, 155-66.

[28] Sarvas, R., Vihavainen, S. A Holistic View on Future Snapshot Media, *Pervasive Image, Capturing and Sharing Workshop,* Ubicomp, 2005.

[29] *Schuler, R. P., Laws, N., Bajaj, S., Grandhi, S., Jones, Q.* Finding your way with CampusWiki: a location-aware wiki, *CHI '07 extended abstracts on Human factors in computing systems*, 2007, 2639-2644.

[30] Simon, I., Snavely, N., and Seitz, S.M. Scene summarization for online image collections, In *ICCV '07: Proceedings of the 11th IEEE international Conference on Computer Vision*, 2007.

[31] Van House, N., & Davis, M. The Social Life of Cameraphone Images, *Pervasive Image, Capturing and Sharing Workshop,* Ubicomp, *2005*.

[32] Von Ahn, L. and Dabbish, L., Labeling Images with a Computer Game, In *Proceedings of ACM Conference on Human Factors in Computing*, 2004, 319-326.

[33] Zhou, C., Frankowski, D., Ludford, P., Shekar, S. and Terveen, L., Discovering Personal Gazetteers: An interactive clustering approach. In *Proceedings of ACM GIS 2004*, 2004, 266-273

[34] Zimmerman, J., Forlizzi, J., and Evenson, S. Research Through Design as a Method for Interaction Design Research in HCI, In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 2007, 492-502.