Chapter 2

Collective Intelligence
Contribution Systems

2.1 Taxonomy of Collective Intelligence

On daily basis users are contributing to the on-line content with millions of comments, tags, ratings, searches and other in the form of comments, tags, ratings, searches, and type of input.

Cook [2008] built a taxonomy of user contributions in which he distinguished between passive contributions, in which people provide their input unintentionally (running queries to improve Google searching algorithm or purchasing items which determines the popularity of the products), and active contribution in which users intentionally share content with a purpose of bringing value to others.

2.1.1 Human Computation Tasks Types

Yuen et al. [2011] propose the following categorization of tasks that can be fulfilled through collective intelligence contribution systems:

- **Geometric reasoning**: humans are better at interpreting contextual object interpretation and in some cases object position optimization [Jagadeesan et al., 2009]. This skills have been put to use in projects like object discrimination in Peekaboom game [Von Ahn et al., 2006a], sketch interpretation [Engel et al., 2012], combinatorial problem solving [Mecanica et al., 2010].
• **Entity annotation**: annotation of images ([Sorokin and Forsyth, 2008], [Law and Von Ahn, 2009]), music and sounds [Law et al., 2007], [Mandel and Ellis, 2008], [Turnbull et al., 2007]), difficult cases of speech transcription [Williams et al., 2011], and video [Siorda and Simperl, 2011] etc.;

• **Opinions**: human subject information collection through an open call [Kriplean et al., 2012], [Krause et al., 2012];

• **Common sense extraction**: [Von Ahn et al., 2006b, Stork et al., 2000] and relevance evaluation [Alonso et al., 2008];

• **Natural language annotation**: sentiment analysis [Rafelsberger and Scharl, 2009], word sense disambiguation [Akkaya et al., 2010], Biemann and Nygaard [2010], [Chamberlain et al., 2008, Lawson and Easticke, 2010, Finin et al.];

• **Security and Spam identification**: [Von Ahn et al., 2008] (Recaptcha);

• **Open collective information sharing**: Wikipedia\(^1\), del.icio.us\(^2\), Yahoo Answers\(^3\), Pinterest\(^1\), Mendeley\(^5\), DBPedia\(^6\) and crowd assisted translation (DuoLingo\(^7\), Google Translator\(^8\))

### 2.1.2 Forms of Motivation

If passive human computation contributions consists of browsing and activities log data [Cook, 2008] collected automatically, active contribution requires open call format. Open call contributions will actually take place if users are sufficiently motivated to participate. Current collective intelligence contribution systems utilize various motivational mechanisms ranging from intrinsinc to extrinsic in order to induce user participation.

Deci and Ryan [1985] define extrinsic and intrinsic motivation as following: "Intrinsic motivation refers to doing something because it is inherently interesting or enjoyable, and extrinsic motivation refers to doing something because it leads to a separable outcome."

There are however, collective contributions which motivate users both intrinsically and extrinsically. In fact, Deci and Ryan [2000] suggests that there is a gradual shift from extrinsic motivators to intrinsic through the degree of internal involvement as depicted in

---

\(^1\)http://wikipedia.com

\(^2\)http://del.icio.us

\(^3\)http://answers.yahoo.com/

\(^4\)http://pinterest.com

\(^5\)http://mendeley.com

\(^6\)http://dbpedia.com

\(^7\)http://duolingo.com

\(^8\)http://googletranslate.com
figure 2.1. As the figure suggests, the line between the intrinsic and extrinsic motivations may at some point become unclear.

**Figure 2.1: Extrinsic and Intrinsic Motivations.**

Quinn and Bederson [2011] define the position of known collective intelligence systems in terms of their appeal to extrinsic, personal and intrinsic motivators based on definitions found in literature (figure 2.2).

**Human Computation** is defined as "a paradigm for utilizing human processing power to solve problems that computers cannot yet solve" [Von Ahn, 2005]. **Crowdsourcing work** is referred to as "an act of taking a job traditionally performed by a designated agent (usually an employee) and outsourcing it to an undefined, generally large group of people in the form of an open call" [Howe, 2006]. **Social computing or content-sharing** refers to blogs, wikis, on-line communities and was defined by Parameswaran and Whinston [2007] as "applications and services that facilitate collective action and social interaction online with rich exchange of multimedia information and evolution of aggregate knowledge".

Furthermore Quinn and Bederson [2011] suggest that while human factor is the common denominator in the scheme, there are plenty of other elements that shape the variety of approaches of utilizing the human factor, such as: *motivation, quality control, aggregation, required skills, process order, task-request cardinality.*

Considering these variety of these factors, the following subsections will give the collective intelligence contribution platforms mentioned above a deeper look.
2.2 Crowdsourcing Platforms from the Perspective of HumanComputation tasks

Referring the figure 2.2, one of the most well known extrinsically motivating crowdsourcing platforms is Amazon Mechanical Turk (AMT)\(^1\). The tasks that are normally posted as jobs (HITTs) require sufficient English language and computer proficiency and can include such tasks as annotation, word sense disambiguation, information search, affective text analysis (see section 4.1.1).

There are similar platforms like Crowdflower\(^2\), StandardMinds\(^3\) that are, in fact, backed by AMT platform but are either focused on a specific task type or conversely provide extended capabilities. Namely Crowdflower allows filtering of workers by country, and extending employer base to the rest of the world (AMT is open directly only to US employers) and offers more sophisticated quality control services. The StandardMinds offers advanced proofreading services with thorough content editing services.

Snow et al. [2008], Feng et al. [2009] show that the quality of returned jobs on Amazon Mechanical Turk is acceptably high. Experiments of Paolacci et al. [2010] showed that

---

\(^1\)https://www.mturk.com
\(^2\)http://crowdflower.com
\(^3\)http://standardminds.com
Chapter 2. Collective Intelligence Contribution Systems

non-expert Mechanical Turk workers provided data quality equal to one expert worker and higher than results collected from users on Internet discussion boards.

At the same time, Horton and Chilton [2010] still estimated that the smallest median wage an Amazon Mechanical Turk worker is willing to accept for a task is as low as 1.38 dollars per hour and Paolacci et al. [2010] observes that 90 percent of tasks cost on average 10 cents. Interestingly, Mason and Watts [2010] have proven through experiments that higher pay does not motivate workers to work harder because they assumed that the task complexity is worth the value. Higher pay, however, increased the quantity of the accomplished tasks.

In terms of demographics, studies by Paolacci et al. [2010] revealed that Mechanical Turk has two major worker segments: well educated Americans (with 65 percent females, 35 percent males, with average age of 36 years old). 13 percent of survey participants reported that the work on AMT is a primary income resource, with the rest participating because they find it a fruitful way of spending free time, earning on average less than 20 dollars per week. Similar tendencies are established among 2nd largest segment of amazon mechanical turk workers are Indian workers, but with slightly younger average age, larger number of male workers and 26 percent relying on Mechanical Turk as major income resource [Wang et al., 2011]. In general, there is a trend indicating that the segment of Indian workers is growing.

Indeed, the low wages seem to be most fit for crowd workers from financially disadvantaged counties where the cost of life is significantly lower than in most developed countries. Such crowdsourcing projects as Samasource\(^1\), txtEagle Eagle [2009], and MobileWorks\(^2\) are targeting worker pools in developing countries living on less than 2 dollars per day and therefore, benefiting most noticeably from the earnings that Amazon Mechanical Turk can offer. However, lower qualification often results in what Viurens [2009] labels as sloppy work. Moreover, honest worker may perform poor due to misunderstandings and even lack of qualification. The unintentionally poor quality coupled with even bigger issue of massive spamming create a need for a solid quality control, which so far is is represented by majority voting approach [Raykar et al., 2010], as well as other spam filtering algorithms of Viurens [2009], Ipeirotis et al. [2010], Callison-Burch [2009].

More versatile and comprehensive tools like TurkKit [Little et al., 2010], or Turkalytics [Heymann and García-Molina, 2011] allow to process, report, and visualize in real-time logging events (workers’ searching behavior, locations, browser environments, activity information) can help to make conclusions regarding task attractiveness, execution time

---

1\(\text{http://samasource.com}\)
2\(\text{http://mobileworks.com}\)
and consequently help setting better pricing, track quality and reputation of specific worker IPs.

The Turakities tool partially addresses another serious issue of crowdsourcing, which is the limited information about employees’ demographic data (concealed due to privacy concerns). It allows pre-screening, and recontacting specific Worker IDs or coding that can exempt specific groups of workers from viewing the task.

Another concern of monetized crowdsourcing platforms is related to the protection of employees that lack immediate performance feedback and often feel "cheated" without ability to prove so due to the fact that feedback is given only after the work is submitted. Dow and Klemmer [2011] posit that immediate feedback has impact on future task performance and propose to develop synchronous feedback giving systems that unlike of current asynchronous system (including AMT), can provide parallel peer review. The authors further present a system called Shepherd, which manages the parallel feedback of the crowd. Soylent [Bernstein et al., 2010], another project that manages real-life word processing system, in which it requests Amazon Mechanical Turk workers to shorten, proofread texts and perform other complex tasks that normally require programming with macros, also provides nearly real-life synchronous feedback. A synchronous crowdsourcing application VizWiz [Bigham et al., 2010] intends to help blind people to be aware of objects that surround them using crowdworkers or their social network friends.

Kaufmann and Schulze [2011] indicate that while extrinsic motivation categories are important for Amazon Turk contributors (i.e., immediate payoffs) intrinsic motivation aspects also present significant equal importance (task autonomy and variety). At the same time, while Joyce [2007] finds that under certain conditions users are capable of providing creative and versatile ideas, Dontcheva et al. [2011] indicate that crowdsourcing platforms with strong intrinsic task quantity-money correlation do not motivate creativity. In their study, participants spent on average 6 minutes on each task, while creative solutions normally requires more time. In general, most crowdsourcing platforms contend with standard performance and have no means, tools or mechanics to encourage singular performance.

Conversely, recently emerged platforms like Brainrack1, Jovoto2, Innocentive3, Innovent4 and many more let employers make open calls, collect and reward best ideas using competitive reward system [Archak and Sundararajan, 2009]. However, referencing them as crowdsourcing platforms is dubious since there is little or no peer collaboration or

---

1http://Brainrack.com
2http://Jovoto.com
3http://Innocentive.com
4http://Innovent.com
interdependency involved [Howe, 2006] and the efforts of non-winners are left behind as in any traditional competitions.

The future of monetary based crowdsourcing systems lies in improving the ability to coordinate more complex human computation tasks, decomposing larger, more complex tasks into smaller subtask [Zhang et al., 2011]; involving third party coordination; and using more sophisticated algorithms of crowdsourced data validation.

### 2.3 Human Computation Tasks on Content Sharing Platforms

The next category of human computation contribution platforms is represented by content sharing platforms. Typically, they involve no financial remuneration or direct personal interest (personal use or career advantages as found to be important for open software developers [Ghosh, 2005]), but provide self-esteem, expert status and recognition, as well as feeling of usefulness and importance of the contribution as indicated in figure 2.1. Indeed, an experiment conducted on video sharing platforms by Huberman et al. [2008] demonstrated that lack of attention and acknowledgment from public lead to a decrease in the number of uploaded videos and in some cases to no uploads at all.

Classical examples of the largest content sharing systems are Wikipedia and Yahoo Answers (content creation), Reddit\(^1\) (content submission and ranking), Imgur\(^2\) (image sharing), YouTube and Facebook [Smith, 2008].

### 2.4 Playful Human Computation Applications

In order to entice participants, content sharing platforms often utilize playful elements (including metaphorical processes, funny avatars), and make use of rich graphics to facilitate the process of collaboration, appeal to users, and inspire exploratory mood. Rao [2008] argues that *playfulness*, (unlike *play* which is opposite of *serious* process), can easily be integrated in serious settings.

Cailllois [2001a] described all forms of play as being positioned on an axis between Paidia and Ludus, where Paidia is a primary power of improvisation and joy compared to its opposite, Ludus - defined by strict rules. Bateman and Boon [2005] points out that the first experience of playing a new game involves playfulness, until the user learns the

---

\(^1\) [http://reddit.com]

\(^2\) [http://imgur.com]
Ludus part (rules) and starts following the structured patterns of play. Thus, playfulness is linked to exploration. Indeed, Ghani and Deshpande [1994] note that the state of flow, characterized by intense concentration and enjoyment is significantly linked with exploratory use behavior. Woszczynski et al. [2002] discuss the difference between the flow state and playfulness, stating that they are to be measured differently, since state of flow is an internal psychological state, and playful behavior is something to be externally observed.


To take up this notion of playfulness through analogy, Yahoo’s ”Thumb-Up/Thumb-down” button is used to measure the popularity of the posted content or comments [Ghosh and Hummel, 2011].

Playful image tagging application Links\(^1\) is designed to allow uploading images albums and suggests tags that are relevant to that description. It also allows to rate albums and post updates on Facebook. Application like SlaveryFootPrint\(^2\), Wahl-O-Mat\(^3\), ConsiderIt [Kriplean et al., 2012], Bake your Personality! and Who’s got it? [Krause et al., 2012], and a variety of Facebook applications (including in-built quizz making plug-in) use relaxing, humorous, simulative context and interactions to provide pleasurable experience while fulfilling some more serious goal like educating/informing users on specific topics or collecting useful input.

Ferrari et al. [2009] uses manually generated drawings on top of photos to help reconstruct the shape of objects on the photos. With similar intentions, Russell et al. [2007] developed LabelMe tool that helps to annotate specific objects identified in images using graphic tools, also providing a way to facilitate collaboration with the Amazon Mechanical Turk workers. These tools, however, lacked motivational positioning offering neither intrinsic nor extrinsic benefits for contribution.

Noticeably more user-friendly are the series of Zooniverse\(^3\) projects that use crowd-sourcing to filter through enormous space, earth and underwater world images to locate desired objects Gal [2009]. For example, MoonZoo project requires contributors to locate craters, indicate their type and shape using special graphical tools provided by

---

\(^1\)http://www.inseminates-links.net/
\(^2\)http://slaveryfootprint.org
\(^3\)http://www.wahl-o-mat.de
\(^3\)https://www.zooniverse.org
the web program interface. The tasks are relatively easy to comprehend and execute. Similar in its approach is the Boodleian Project\(^4\) requests contributors to describe the covers of more than four thousands of digitized music scores using specifically designed web interface.

A newly introduced application that provides benefits to public and simultaneously allows users to benefit from their own contributions is Duolingo\(^2\) language learning environment. It allows users to learn languages for free in a playful environment while helping to translate websites and other documents that match the skill level of users.

Following playful content sharing applications, the next on the way to intrinsically motivating applications that collect human input (figure 4.3) are human computation games and gamified applications, which are discussed in section below.

### 2.5 Collective Intelligence and Games

The development of web 2.0 spurred the rise of gamification, defined by Deterding [2011] as "the use of game design elements in non-game contexts." Today there are platforms that integrate playful gamification in every day challenges like healthy eating, exercising, performing, geotagging (Noah Project\(^1\)), ecological living (social network game Ecotopia\(^2\)), and other routine but useful tasks.

**Human computation games**, as part of gamification wave, emerged in 2005 in the form of Games with a Purpose (GWAP)\(^3\) and were defined by their founder Luis von Ahn as an attempt to "... use human effort to perform tasks that computers cannot yet perform, in an enjoyable manner."

So far, game designers applied human computation games to the following domains:

- Ontology creation
- Human-computer interaction research
- Spatial reasoning
- Geotagging

\(^4\)http://whats-the-score.org
\(^2\) http://duolingo.com/
\(^1\) http://www.projectnoah.org/users/NeilDazet/patches
\(^3\)http://gwap.com
In addition, human computation games vary in terms of their setup parameters (mechanics):

1. **Number of players:**
   - a) single-player
   - b) two-players
   - c) three or more

2. **Agreement system:**
   - a) sync. agreement
   - b) majority vote
   - c) system

3. **Location:**
   - a) non-location based
   - b) location based

4. **Game genre:**
   - a) gamified tasks
   - b) cooperative
   - c) game genres

Games can be designed for single (1a), two (1b) or more players (1c). The rewards can be gained by reaching instant agreement with co-players (2a), by matching the decisions made by majority (2b) or by creating input that remains within borders set up by the system (2c). In some cases, when games involve coordination of real life actions, players have to be at the same location (3b); most games, however, let the users to keep their privacy (3a). In addition, some games will contain goals and missions related only to the serious tasks although using metaphoric analogies or playful interaction (4a), in others, game designers appeal to existing game genres and incorporate playful goals and interaction patterns (that are typical for a selected game genre) into the game mechanics (4c), or game mechanics follow the two-player cooperative setup (4b).

Below we give an overview of various human computation games that are classified according to their domain of application and the above stated criteria.

### 2.5.1 Ontology and Data Labelling

- **Two players/majority vote/non-location based/cooperative:**

The first widely known games with a purpose were developed by Luis von Ahn. Their mechanics consisted of 2-player synchronous cooperative gameplay where both players have to agree on a certain decision with each other in order to progress through the game. GWAP’s most successful tasks included image tagging *(ESP Game [Von Ahn and Dabbish, 2004])*), common sence knowledge collection with games like *Verbosity* [Von Ahn et al., 2006b] presented in figure 4.3, *Squeegle and Peekaboom* [Law and Von Ahn, 2009]. The latter two games allowed object tagging within an image using two different methods.

The game setup provided a successful mechanism of data quality filtering. In addition, only when agreement on the same tag was reached 3 times would the tag be retained in the database, which ensured very good quality control. The success of the game spurred the appearance of the games that intended to improve the original setup or use it for different purposes. *Matchin* game offered users to guess the preferences of the other player [Hacker and Ahn, 2009]. *Takhtamysheva et al. [2009]* used two-player synchronous game
to generate a variety of phrases that describe the same physical actions. Thumbs-Up game [Dasdan et al., 2009] used the same set-up to rank search results. MaSweep game [Chang et al., 2011] was designed to collect mutual exclusive concepts like “food” and “furniture” to contribute to Never-Ending Language Loop project (NEL) where mutually exclusive concepts have to be entered manually [Carlson et al., 2010]. Conversely, Curator game [Walsh and Golbeck, 2010] uses a standard two-player output agreement setup but increases the number of possible matches by proposing a more flexible approach where users compare a collection of items, instead of just 2 options. SiOrpaes and Simperl [2011] presented their series of simple 2-player human computation games under OntoGames umbrella: OntoPronto for semantic annotation of texts; TubeLink and its predecessor SpotTheLink for aligning ontology; and OntoTube for video tagging.

**Three or more players/synchronous agreement or majority vote/non-location based/cooperative:**

Ho et al. [2009] slightly modified the original 2-player GWAP setup bringing in a third player (KissKissBan game). The role of the 3rd player was to push further the border of complexity of the task that two players can agree on. Another three-or-more player setup was introduced in the game HeardIt [Barrington et al., 2009], which aimed at collection of audio annotations and in which players were awarded points for reaching majority consensus.

**Two players/synchronous agreement/non-location-based/traditional game genre:**

Vickrey et al. [2008] presented three human computation games (Categorilla, Categor Godzilla and Free Association) that were inspired by GWAP; a number of traditional
non-serious games like *Scattergories* and *Taboo* games; and *Open Mind Common Sense System Collection* [Chklovski and Rey, 2003], one of the earliest web based massively collaborative ontology extention systems.

*Categorilla* and *Catedizzle* ask players to supply phrases or words which fit specific category, i.e. "Things that fly" or "Type of vehicle". The 3rd game *Catedizzle* slightly increases the challenge by limiting users to words that start from certain letters, which, however, made her less popular than the other two games. The data the games collect were used to augment the database of hyponyms of WordNet [Fellbaum, 1998]. For example, WordNet had a hyponym "burglary" and "fraud" for word "crime", with the games one had an opportunity to add "homicide" and "murder" [Vickrey et al., 2008].

- **Three-or-more players/synchronous agreement or majority vote/non-location based/traditional game genre:**

  *Doodling* [Kumaran et al., 2012] game is a based on popular board sketch-and-convey game that uses graphic drawing tool where one player draws series of objects and the other guesses the phrase behind drawn objects.

  Another game that is built using an existing game genre (*Snap*) is *PhotoSlap* [Ho et al., 2007], that imitates a multiplayer (two and more) card game. Instead of cards, players use images. The players score up when other game players show no disagreement with the player’s photo recognition results.

- **Single-player/majority vote/non-location based/gamified tasks:**

  The most recent *OntoGames* such as *TubeLink* [Thaler et al., 2011a] and *SeaFish* [Thaler et al., 2011b] moved away from two-player setup and were transformed into more flexible single player games that do not stall the game or force to resort to bots when 2nd player is not available. Also the games have noticeably improved graphics, playful interfaces and better game dynamics, unlike early games that could have been regarded as two player quiz-like applications.

  *PhraseDetectives*, [Chamberlain et al., 2008] is another text-based human computation game where players have to annotate text blocks by indicating whether certain phrase or object has already been referenced in the text. The game utilizes single-player format, and the points are awarded based on the consensus with majority. The game introduces two modes: annotation generation and confirmation.
Sentiment Analysis game [Rafelsberger and Scharl, 2009] aims at creation of sentiment lexicons - sentiment terms with corresponding sentiment value. The game allows players to rate words or phrases on a scale varying from negative to positive and awards them with points based on the consensus with the majority. The deployment of Sentiment Analysis game on Facebook provides relatively high user involvement and satisfactory number of votes.

- Single-player/majority vote/non-location based/traditional game genre:

OntoGalaxy game [Krause et al., 2010] attempted to reach casual shooter game’s fans, by introducing more dramatic game mechanics and dynamics as well as richer graphics in a single-shooter human computation game.

Also, Semantic Web Research Group from Leipzig University experimented with casual game design by introducing VeriLinks\(^1\), a single-player verification-based human computation game that allows to simultaneously play simple shooter game and earn bullets by fulfilling human computation task (comparison of content of two images).

DigitalKoot, another single-player asynchronous agreement game [Chronis and Sundell, 2011a] is a playful version of ReCaptcha [Von Ahn et al., 2008], which entices players to decode archival documents, by interpreting phrases printed in unintelligible for computers way. In one sub game players enter their text interpretations, and in the other subgame - the tags are verified. The game’s mission is to save a mole (eng.)/Maulwurf (ger.) whose life depends on the rightfulness and promptness of user’s input.

### 2.5.2 Spatial Reasoning and Cognition

- Single-player/majority vote/non-location based/traditional game genre:

Another single player human computation game yet applied to spatial reasoning is Plumnings [Terry et al., 2009]. Users have to minimize the network of ”pipes” to increase "oxigen" pressure and thus increase its supply to the colony of Plumms. What players actually do while playing the game is designing optimized critical paths between field programmable gate array clusters.

- Single-player/agreement with system/non-location based/gamified tasks

\(^1\)http://aksw.org/Projects/VeriLinks
Well-known human computation bio-medicine games that also use single player set up with asynchronous consensus but focus game mechanics on task gamification are *FoldIt* [Cooper et al., 2010], *Phylo* [Kawrykow et al., 2012] and recent *Eterna*.¹ The games’ tasks complexity requires the use of tutorials. The purpose of *FoldIt* is to refold proteins in the most optimal manner, something that requires unreasonably large computational power of computers. Similar to *FoldIt*, *Phylo* [Kawrykow et al., 2012] requests users to reposition genes of living creatures in the most optimal way. *Eterna*, a Carnegie Mellon University’s human computation game invites users to create a large-scale library of synthetic RNA design that play important role in controlling cells with disease-causing viruses.

### 2.5.3 Human-Computer Interaction Research

- Single-player/majority vote/non-location based/gamified tasks or use of a traditional game genre:

A series of games that aimed at researching patterns of human-computer interactions in order to improve software or interfaces has emerged. Thus, for example, an android based mobile human computation game *TypeIt* [Nie, 2012] allows to track users’ typing patterns, which are then used to improve mobile keyboard interfaces. Yan and Yu [2009] presented a game *Magic Bullet* which is intended to train people’s typing skills to aid the development of better algorithms for handwriting recognition.

A social game *Passion Fruit* was run on social network in order to study group cooperation within on-line games [Kirman, 2010].

With their sketching games *Picture-Phone* and *Stellasketch*, Johnson and Do [2009] addressed the issue of sketch image interpretation and aimed at collecting data about how people make and describe hand-made drawings. Another project in similar direction of research presented by Wang and Yu [2012] is based on *DrawThat*, which in its turn, is a digitalized version of a traditional same-place group guessing game. It suggests that players guess the object behind the depicted images created by one of the game participants, which allows to collect information to be used for machine learning regarding sketching optimization.

Talton et al. [2009] explore 3d modeling patterns using multiple input in collaborative design spaces.

Ma et al. [2009] propose *PageHunt* game, the data of which can be used to improve search results by providing meta data for pages, page ranks and ways to refine queries.

---

¹[http://eterna.cmu.edu/](http://eterna.cmu.edu/)
Clemesha launched series of Wikipedia hypertextual multiplayer games also available as mobile applications: *SpeedRace*, *Least Clicks*, *Six Degrees of Wikipedia*, *Five Clicks to Jesus*, *No United States*. The goal of the games is to find a search path from given starting and finishing points, which can be used to study user navigation in non-directional search patterns.

Another project, a *Sentence Recall Game* [Wang and Yu, 2010] is a single player game that allows to study the usage of language patterns among non-native speakers.

### 2.5.4 Geotagging

Improvement in mobile location-tracking technology gave rise to series of early location based mission games like *Gopher Game* [Casey et al., 2007], *MobiMission* [Grant et al., 2007], *Indogator* [Lee et al., 2010] where users create human computation missions and fulfill those created by other players. Cooperative human computation geo-tagging games like *EyeSpy* [Bell et al., 2009], *GeoSnake* [Matyas et al., 2011] allow to produce images, create geo-spatial tags and verify those using the wisdom of the crowd and cooperative mechanisms.

A more advanced *UrbanMatch* game [Celino et al., 2012] is a mobile location-based human computation game, which choses photos that are most representative for the point of interest and selects those best ranked by multiple participants.

A project by Tuite et al. [2010] called *PhotoCity* is a hybrid alternative reality mobile game with a purpose, which is played outdoors with a goal of taking pictures of buildings/objects from various angles and positions in order to recreate their 3D models. Another alternative reality game, *The World Without Oil*[^3], which had a social goal introduced an alternative reality simulation available for massive collaboration of network users who helped visualize a world disaster and offer multiple solutions.

### 2.6 Players Response to the Playful Human Computation Platforms

The results of a study of Nichols and Kang [2012] on human collaboration involving Twitter questions and answers revealed no significant connection between the usefulness of the request and the number of contributions. This finding also supports the assumption that human computation games are played for the variety of personal and fun reasons.

[^1]: [http://thewikigame.com]
[^3]: [http://www.worldwithoutoil.org]
rather than for the purpose of contributing to the good. Indeed, the findings of Lee et al. [2010] produced as a result of evaluation of their human computation mobile application Indogator revealed that perceived gratification factors, such as information discovery, entertainment, information quality, socialization, and relationship maintenance, were most significant predictors of their application use.

Chrons and Sundell [2011b] reported that players accomplished 2.5 million microtasks, and spent in total 2,740 hours playing two DigitalKoot minigames (generation and approval). From 2003 to 2007 ESP game provided 10 million of image captions [Russell et al., 2007]. [Vickrey et al., 2008] pointed out that within a year their games have collected 800,000 instances and at least 24 guesses per category. Within 2 years more than 200,000 volunteers had made more than 100 million galaxy classifications for Galaxy Zoo project Gal [2009].

At the same time, as indicated in section 3.4.3 Social Player Attitudes and Behavioral Patterns, users display pattern of playing that fits the power law [Kirman, 2010], which says that only small part of players displays hard-core player behavior. This goes along with findings of Chrons and Sundell [2011b] who reported that most hardworking 1 percent of gamers of 55,000 players of DigitalKoot, a human computation game distributed on Facebook, accomplished almost third of all tasks.

Chien-Ju and Kuan-Ta [2009] demonstrates that in systems with social verification (vs. machine based verification) "sequential verification makes users to give a more diverse and descriptive set of outcomes than simultaneous verification, though the latter is stronger in ensuring the correctness of verified answers". Yet, in spite of more solid simultaneous verification, the requirement of having at least two players willing to start a particular game at the same point of time poses a serious challenge to the 2-player game setup. Vickrey et al. [2008] points out the necessity to use bots, which use past guesses or make imaginary guesses with a pre-calculated probability of generating a word that matches the guess of the player. However, once the users realize that they play with bot, their interest to the game decreases. To counteract this challenge, Chamberlain et al. [2008], Chrons and Sundell [2011b] propose division of input generation and confirmation tasks into separate single-player mini-games.

Terry et al. [2009] of Plummings game suggest that their game will perform better if their spatial reasoning task will be divided into smaller subtasks, which, games like Eterna and FoldIt have achieved. In addition, the latter games have introduced interactive mini games tutorials, which increased the understanding of task requirement and game rules and provided handling of novice players as presented in section 3.3.1 Maturity stages of players. Indeed, the necessity of providing tutorials when appealing to large non-sophisticated audience is also reflected in social network games, which offer tutorials
blended into game dynamics as described in Chapter 3 subsection *Maturity Stages of Players* and is indispensable for human computation games with complex tasks.

Finally, compliance with user experience and interface development guidelines is another area that improves game effectiveness, reduces gamers confusion, and consequently positively influences the input quality. Ho et al. [2007] point out that improving interface design of their game *PhotoSlap* increased the precision of users’ input by 10 percent.

### 2.7 Conclusions

In the background of global connectedness Web 2.0 technologies brought an opportunity to bring together global workers and global employers making crowdsourcing and human computation indisputably feasible; and provide opportunities for further evolution of the platforms that harness energy of netizens.

In this chapter we gave an overview of different methods of input solicitation from the perspective of user motivation (section 2.1) ranging from purely extrinsic to intrinsic factors.

In section 2.2 we covered systems that use extrinsic motivation of user contributions (crowdsourcing) and reviewed the challenges that these systems currently face. While the variety of tasks that can be fulfilled by these systems is extensive, their execution depend on immediate presence of financial resources, and a budget that increases proportionately to the number of required responses.

Section 2.3 presented content sharing systems, which use playful approaches to motivate sharing, commenting and other contributions. While users are motivated to participate by the tasks attractiveness and its personal relevance to the user, the type of input and its amount are limited and not regulated or solicited as much as by more complex playful systems presented in subsection 2.3.1 and games in section 2.4.

Subsection 2.3.1 and section 2.4 cover more challenging approaches that propose to appeal to intrinsic motivation of users to fulfill large quantity of more complex tasks that are of a little personal relevance to users. This approach assumes high fixed costs for system development, but relatively low variable costs that unlike crowdsourcing system do not increase with the number of tasks and can serve research projects, which require unlimited amount of data and when quality of results improves with the growth of input data amount (ontology, linguistic databases, algorithm training data).
At the same time, most of the human computation game projects reviewed in section 2.4 display sporadic compliance with theoretical principles of game development guidelines and do not make the full use of the potential that good mechanics, dynamics and aesthetics have to offer. As a result, only few games reached satisfactory amount of user input with many more games remaining in the prototype stage of development. Therefore, human computation games would greatly benefit from using game development approaches and design principles presented in chapter 3 and 4, and review the game design process from the point of user motivation and fine-tune the proportion and quality of serious and playful tasks.

The next chapter presents the theories behind efficient game development, focus on player’s motivation, guidelines for constructing intuitive gaming environments.
Bibliography


David G Stork, Sand Hill, and Road Suite. Open data collection for training intelligent software in the open mind initiative. *Mind*, (June):1–7,


Jeffrey P Bigham, Chandrika Jayant, Hanjie Ji, Greg Little, Andrew Miller, Robert C Miller, Robin Miller, Aubrey Tatarowicz, Brandyn White, Samual White, and et al.


Stuart E Dreyfus and Hubert L Dreyfus. A five-stage model of the mental activities involved in directed skill acquisition. *Distribution*, (February):1–18.


Brian E Mennecke, David McNeill, Matthew Ganis, Edward M Roche, David A Bray, Benn Konsynski, Anthony M Townsend, and John Lester. Second life and other


Amy Jo Kim. Smart gamification: Seven core concepts for creating compelling experiences, 2011b. URL http://www.youtube.com/watch?v=F4YP-hGZTuA.


N. Berry. Who is playing games on facebook?, 2011.


