Mechanism Design for Designing Annotation Tools

Roberta Cuel, Oksana Tokarchuk, Marco Zamarian

Dept. of Computer and Management Sciences
University of Trento
Trento, Italy

Email: (roberta.cuel, oksana.tokarchuk, marco.zamarian)@unitn.it

Abstract—The Web 2.0 is increasingly considered as a phenomenon that affects the way people interact, search, post and share information on the Internet. Namely it affects the daily life of any Web user, expert or company that works on the network daily. One of the dominant traits of Web 2.0 applications is the capability of co-opting end-users in endeavors which traditionally have been considered as top-down activities and exploiting user-based networks. Through Web 2.0 applications users add content and annotation in order to describe and share pictures, videos, files, etc. Apart from some of the most well known applications (e.g., Facebook, Twitter, Flickr and the like), many Web 2.0 tools are not good at attracting a critical mass of individuals. In fact, many studies have shown that a common outcome for online communities is an 'onion' structure were only a few core individuals actively participate discussing and contributing to the common content, whereas others are considered as peripheral users who observe the community and simply use the content. Participation and willingness to contribute, thus, become two of the critical issues that companies and software developers should take into account when creating Web 2.0 applications. In other words, we claim that understanding and analyzing appropriate sets of incentives that might motivate users to contribute are critical steps in the design of Web 2.0 applications. In this paper, we describe how theories and techniques that are well known and used by scientists in economics and management studies can be used to develop incentive-compatible Web 2.0 tools. Specifically, we will provide an example of an application of mechanism design and applied experimental economics in the development of an annotation tool.

Keywords—Incentives; Mechanism Design; Web 2.0; Content Creation and Annotation.

I. INTRODUCTION

The collaborative, social way of generating, organizing, and managing knowledge has been growingly considered as a trigger of creativity and innovation in several applied fields. This phenomenon has been heavily shaped by the advent of social based technologies such as grid computing, peer to peer file sharing, collaborative authorship of digital content, social networks, and, more in general, Web 2.0 applications [7], [18], [28].

In this scenario, the increasing popularity of Web 2.0 applications dramatically changes the way people interact, have fun, communicate and consume. The common trait of Web 2.0 applications is the empowerment of end-users by co-opting individuals in endeavors which traditionally have been considered as top-down activities, and the exploitation of user-based networks of relationships [20], [23]. A glaring example of such a trend is the new emerging crowd-sourcing phenomenon [12], [14], [17]. Nowadays there are many Web 2.0 applications that can be considered as concrete examples of the crowdsourcing phenomenon: SourceForge (sourceforge.net), Wikipedia (en.wikipedia.org), Galaxyzoo (www.galaxyzoo.org), Crowdflower (crowdflower.com), Innocentive (www2.innocentive.com). All these applications take advantage of communities of experts and/or users who proactively contribute to the creation of a common good.

In economics, a good or service is called a 'common good' [21] when it has contemporaneously the properties of non-exclusivity, and non-rivalry. Non-exclusivity means that it is impossible or costly to exclude any person from the use of the public good. Non-rivalry means that each person can consume simultaneously the public good, without constraining the use for others. In all of the above-mentioned examples, users act according to their own personal purposes and, at the same time, provide information and knowledge that others can share and, in turn, use.

Many studies on online communities and peer to peer collaboration identified the motivations that drive people to participate in a rather large set of heterogeneous and context dependent elements [2], [13], [15], [25], [26], [27], [29] (e.g., reputation, altruism, competition, self-esteem, money, reciprocity, fun, etc.).

Despite these many motives to contribute and the popularity of many Web 2.0 applications, some studies observe how a significant number of online communities and social network applications fail because of under-contribution of participants. For instance, an analysis of the P2P file sharing site Gnutella show that in 2000, only 25% of users shared 98% of the content while 66% of users shared nothing [1].

More specifically, insights concentrating on the patterns that characterize annotation efforts in Web 2.0 communities found that annotations are characterized by power law distributions, both in the relationship between number of tags and number of posts [5] and number of tags and number of contributors [10], indicating that few people contribute disproportionately more than others.

It follows that not all Web 2.0 tools - and specifically the ones that take advantage of social networks - can become
killer applications à la Facebook or Twitter. This is due to the fact that users are not motivated enough to spend time interacting and contributing to a common good. Thus, motivating users to contribute to this kind of collective effort is essential to reach critical mass and ensure a sustainable growth for these crowdsourcing applications.

Since we consider the whole software development process very relevant, we claim that success of an online community-based application requires a blend of well-designed software (i.e., usability) and carefully crafted policies aimed at achieving participation. In this paper, we focus our attention on the social aspects of software development and deployment and we offer an example on how these aspects can be incorporated into the development of annotation and semantic content creation tools. Specifically we will adopt a set of methods and techniques, often referred to as ‘mechanism design’ in the field of economics, that can be used to develop incentives which can be embedded into online applications. This choice does not reflect a disregard for the technical aspects of software design, but is meant to underline features of the process that are oftentimes neglected within the community of developers.

We will focus on the so called sociability design and in particular on the first two phases of the software development process. These are the analysis of the use scenario prior to application design and the fine tuning process of the incentive structure. These phases should be seen as a continuous improvement process that enables designers to adjust the software according to the social needs emerging from the users’ experiences.

The paper is structured as follows: Section 2 describes some basic notions of motivations and mechanism design, Section 3 sketches out some techniques for analysis and design of incentivized applications, Section 4 describes the analysis of an annotation prototype, and finally Section 5 draws our conclusions.

II. BASIC NOTIONS AND DEFINITIONS

Bouman et al. [4] argue that designers of social software have to design software and carefully craft social policies such as: enabling practice, mimicking reality, building identity and actualizing self. In order to effectively design social software we focus our analysis on motivations in the context of Web 2.0 and on tools that enable us to identify key incentives that can be embedded in the software.

A. Motivations

Several studies on the motivation to participate in knowledge sharing indicate that people participate because they want to be part of a ‘community’, and engage in the exchange of ideas and solutions [26]. Similarly, Forte and Bruckman find that peer recognition plays a role in Wikipedia which is similar to the dynamics shaping up scientific collaboration [9]. Wang and Fesenmaier [24] demonstrate that efficacy is a major factor affecting members’ active contribution to online communities. The study also indicates that the possibility of future reciprocation (expectancy) is another major motivation driving an individual’s contribution. Beenen et al. show that challenging goals are powerful motivators of online contributions [2], while Wasko and Faraj found that people contribute when they perceive that this enhances their professional reputation [27]. Kuznetsov argues that the motivations of Wikipedians to contribute are grounded in the values of reputation, community, reciprocity, altruism and autonomy [15], [25]. Wiertz and de Ruiter found that a customer’s online interaction propensity, commitment to the community, and the informational value s/he perceives in the community are the strongest drivers of knowledge contribution [29]. Bock and colleagues suggest the provision of appropriate feedback to employees engaged in (or not engaged in) knowledge sharing [3]. These actions follow from two considerations. On the one hand, they leverage on the importance of pressure exerted from a person’s reference group (e.g., peers, supervisors, senior managers, etc.) to engage in knowledge-sharing behavior; on the other, they underline the importance of enhancing the individual’s sense of self-esteem.

Studies concentrating on the patterns that characterize annotation efforts in Web 2.0 communities found an interesting correlation between a set of emerging social roles and tagging behaviors. These behaviors seem to be spurred by the attempt to create a community, the awareness of one’s audience and a perceived need to communicate with a small group [22]. Analogously, Chen and colleagues found that social comparisons help explain the tendency to contribute more (or less) in a social experiment involving the MovieLens community [6]. Joinson identified these unique uses and gratifications in the context of Facebook: social connection, shared identities, content, social investigation, social network surfing and status updating [13].

From these examples, we can clearly see that motivation can be produced by heterogeneous motives, and might derive from incentives that are assigned to the performer or from an intrinsic desire. Motivation is intrinsic if the performer enjoys the act of performing the task per se. In all other cases, a set of extrinsic incentives can be provided in order to make an individual/team perform. Incentives are a set of instruments (e.g., money, reputation, rewards, prices, credit points, medals) assigned by an external ‘judge’ typically according to some sort of evaluation of the effort exercised by the performer. In principle, these can be totally uncorrelated to the nature of the task.

B. Mechanism design theory

Mechanism design is a field of game theory developed in economics that studies the effective design of rules for human behavior. If individuals follow these rules, they will reach the outcome desired by the game designer. The
underlying hypothesis is that individuals act according to their own private interests and only a careful development of appropriate incentives can enable the alignment of individual and social interests.

To develop a set of incentives from the mechanism design point of view, the first thing developers should do is to understand the social environment (the context) and codify its constraints from the point of view of game theory. In these terms, the game is defined by the following features [19]: the players, the rules, the outcomes and payoffs. In the case of Web 2.0 applications the players are the users (both contributors and readers of the content produced), the rules refer to how the players interact among each other, the outcomes are constituted by the public good produced by means of the application, and finally the payoffs are the values players attribute to the outcomes.

In the analysis of the context, the designer should focus on the system of individual inner interests and the motivations embedded in the structure they interact with (the tool for instance). These interests are affected by various elements which are:

- **The goal**: people interact to communicate and participate.
- **The nature of good produced**: a stylized description, in game-theoretical terms, of the relationship between what good is produced and who consumes it.
- **The tasks**: an ordered collection of tasks into which the contributions can be broken down.
- **The skills**: competences and abilities required to carry on the tasks.
- **The social structure**: a stylized and simplified set of social relationships among the subjects participating in the exercise.

After the analysis of the context it is necessary to define the desirable outcome that the designers want to achieve. Based on the given definition of players, desirable outcome, and the context it is possible to define a set of rules and payoffs that permit to achieve the desirable outcome.

Looking at Web 2.0 applications users spend time and effort to produce - and at the same time to consume - a public good. All individuals benefit from the outcomes that others produce and the application provides. The possibility to access the content without necessarily contributing to its creation leads to the phenomenon of free riding: expecting that others will spend their time and effort to create the content while dedicating own time to other rewarding activities. Mechanism design tools allow us to analyze the context corresponding to each Web 2.0 application and to design some case-specific rules that lead to a reduction of this kind of behavior within the desired context.

### III. How to Design an Incentivized Application

Mechanism design enables to analyze the social structure of the scenario prior to application design, to fine tune the incentive structure of an application, and to provide a set of requirements for a first prototype of the incentivized tool. The lab experiments, then, allow us to evaluate such a tool and to fine tune the incentive structure that is embedded into it.

Ideally the process of design and development of an incentivized application (see Figure 1) should start from the analysis of the concrete situation.

In the first phase, the field analysis is crucial to identify the motivations of both individuals and the social groups which they belong to. Direct observations, interviews and questionnaires are very effective techniques that can be used to unveil and better define the crucial elements described in the previous section.

In the second phase, the raw knowledge is then analyzed in terms of the above mentioned elements. Mechanism design, as a set of techniques, allows the modeling of the situation by using game theoretical predictions about the behaviors of the actors described in the model. Given a set of goals, this model enables the analysts to design a set of incentive schemes that would spur users to behave in line with the desired outcomes.

The third phase is the creation of the prototype which should be the simplest possible solution that can effectively support the users.

In the fourth phase, the resulting prototype is tested, better if the test is conducted in a controlled environment, such as a laboratory. The laboratory allows the experimenter to test an hypothesis with artificially controlled conditions, manipulating or eliminating extraneous factors. As soon as the previous hypotheses are confirmed a sequence of experiments can be organized to fine tune the set of incentives that are embedded in the tool. The design of each experiment may depend on the results of previous ones.
In the fifth phase, the field experiments, the tool is tested in the field, mimicking the situation in the lab. This fine tuning process should increase the complexity of the trial since the experiment get closer to the reality. For instance, add new realistic components such as real actors (i.e., community members), tasks (a daily activity that actors usually carry on), situations (the field and the social structure which actors belong to). Since the control over the ability to manipulate variables reduces, the experiments gain awareness of interaction among variables. This might continuously address new changes to the tool that finally is introduced in the field.

This process is continuously repeated since the tool is finalized, sixth phase.

IV. THE ANALYSIS OF AN ANNOTATION PROTOTYPE DEVELOPED BY T+ID

Founded in 1988, Telefónica Investigación y Desarrollo (TID) [16] is nowadays the largest private research and development centre in Spain as regards activity and resources, and is the most active company in Europe in terms of European research projects in the Information and Communication Technology sector. TID is currently developing a prototype for the annotation of the internal portal of Telefónica. The tremendous amount of available information hinders the access to the right pieces of information required by a person for a specific goal, affecting the company workflows. This is why the company decided to develop an annotation tool to the corporate portal. Adding a layer of annotations, helps obtain many advantages, such as more efficient resource retrieval and navigation, real integration of heterogeneous sources of information (linked data), personalization based on context and roles.

In the following section, we will go through the steps we adopted to support the design of TID’s annotation tool. It will serve as an illustration of the general process we described in the previous section.

A. The first phase: the field analysis

We started our analysis collecting data on people’s motives and motivation drivers interviewing 11 employees of TID representing the community at large (heads of division, senior project managers, project managers, developers, computer engineers, and consultants). Each semi-structured interview was conducted by two interviewers, took 60 to 90 minutes and was recorded on audio tape. These recordings have been transcribed and analyzed descriptively according to ex-post categories. Additionally, a focus group discussion with 6 TID employees was conducted, focusing on usage problems of the existing system and on innovative solutions that might overcome these problems. Since the number of interviewees was not very high, the TID interviews were decisive to provide starting data for the desk analysis phase.

More in depth requirements are identified in the following phases.

B. The second phase: the desk analysis

We analyzed the interviews and classified TID portal’s features according to the variables described above.

The goal of the annotation tool is to improve the search and navigation experience in a corporate knowledge base. TID providers of annotations might have two different motivations to annotate. Users - by annotating resources - can show their areas of competence and interest to the community. They improve navigation, searching and syndicating capabilities of the enterprise portal by using annotations.

The nature of good produced by annotating is public. This scenario is an almost straight out of the textbook case of public good provision in which providers and consumers are the same people. The part where things become problematic is the problem shared by many knowledge management systems: there is a huge incentive to keep strategic knowledge private so one can leverage on it when dealing/negotiating with others.

The task is a typical annotation task. This means that it is very repetitive and lacks a fun element.

The required skills of the agents to complete the annotation task are very basic.

Finally, the social structure is quite complex and various dynamics coexist. Employees work in teams and communities of experts, but also work in the company with a strong hierarchical structure. Visibility, reputation, career development, and money are all part of the mix of motivations driving the behavior of employees. An important issue to deal with is the large number of employees. As underlined in many studies, group size plays an important role in modifying behavior of individual contributors. If, in principle, employees perceive that their contribution is vital for the success of the group we could expect a higher probability of contribution from each employee. In other words, the reputation mechanism might be developed at group or project level.

In this paper, we do not focus our analysis on the tool development, therefore we do not discuss the third phase: the prototype development process.

C. The fourth phase: the lab experiments

In the evaluation of the first prototype we run a laboratory experiment with the goal of identifying the more appropriate set of incentives to spur people to annotate resources.

The incentive structures we tested are focusing on two different rewarding systems:

• pay per click models: each participant get a fixed amount of money for each annotation provided (0.03 Euros per tag up to 3 Euro). For payment requirements we had to round rewards in pay per tag treatment to the highest 50 cents.
• winner takes all model: only the actor that provide the higher number of tags and annotation wins 20 Euros.

Participants were 36 students with no experience in the annotation task nor in the tool that they tested. At the beginning participants read a clear set of instructions and performed a first training session in order to understand the rules of the game and the basic features of the annotation tool. Each individual has been assigned to one of two 'treatments' of images annotation: the first using the pay per click model and the second using the winner takes all models. They performed the task under time pressure (8 minutes for each treatment) with the goal to produce the maximum amount of tags in the allotted time on a random set of images.

We obtain the following results:

• The mean number of tags produced are: 47.42 with the pay per tag model and 62.76 with the winner takes all model. We can measure a 32% increase in the winner takes all model.

• The average amount of money participants perceived are: 6.66 Euros with the pay per tag model and 6.18 Euros with the winner takes all model.

• The average cost per tag is 0.1404 Euros with the pay per tag model and 0.098407 Euros with the winner takes all model.

• The budget has been 31.5 Euro for the pay per tag model and 20 Euros for the winner takes all model.

In the experiment there are some biases. Students are volunteers who are used to participate in experiments. Also, students didn’t find any problem in the annotation task because they are strong Web users and game players.

D. The fifth phase: the field experiments

Despite the biases that affect the experiment results, it clearly emerged that the winner takes all model is dramatically more effective than the pay per click one, in this context. These results constitute a baseline for the implementation of the tool within TID’s corporate environment. The obvious next step is that of running a few other lab experiments in order to move from simple tagging to more complex tasks closely mirroring what happens in TID.

Since annotation is used to retrieve, organize and exchange information in the company portal, in the next experiments annotations and the resulting tags obtained in a first stage of the experiment will be used as an input to retrieve information in the fifth phase: the field experiments.

In the TID, the prototype and the set of incentives will be tested first with a small group of users in order to fine tune the adequate set of requirements. Eventually, the finalized version of the incentivized application will be applied to the whole TID.

V. CONCLUSION

As mentioned above, one of the dominant traits of the Web 2.0 applications is the capability of co-opting end-users in creating new content and sharing it across their networks. The tremendous growth of data, pictures, images, videos that are shared and copied by individuals in the networks, render the access to the right piece of information more difficult. As a consequence many Web 2.0 applications, such as Facebook and others, introduce annotation tools enabling users to add, modify or remove information about a Web resource without modifying the resource itself.

In this paper we demonstrate that, since the whole software development process is very relevant, the success of social software tools requires carefully crafted social design policies aimed at fostering participation and willingness to contribute.

For this purpose, mechanism design and laboratory/field experiments seem promising methods to enable designers to analyze the motivations of users and embed this information into their software.

These phases should be seen as a continuous improvement process that enables designers to adjust the software according to the social needs emerging from the users’ experiences.

As we have just seen, we have proposed a methodology that encompasses the ability to a) analyze any work environment of a social nature in which one aims to introduce annotating tasks; b) conduct a design process for the tools that serve as vehicle for the tasks themselves taking the previous analysis of the social and technological context as an input. This approach stresses a value-chain outlook on the design process, clearly distinguishing problems regarding motivation/participation in the design process strictu sensu, on the one hand, and motivation/participation of users once the tool is in place. The most glaring advantage of this approach is the ability to consolidate theory, method and applications in distinct units, avoiding redundancy.

Finally, since the methodology we propose seems promising and accepted by developers, the real benefits of the incentivized application will be measured only when it will be adopted by TID.

ACKNOWLEDGMENT

This work was supported by the EU funded project INSEMTIVES - Incentives for Semantics (www.insemtives.eu, FP7-ICT-2007-3, Contract Number 231181).

REFERENCES


