Incentive-Centered Design for User-Contributed Content

Lian Jian
Jeffrey K. MacKie-Mason

1. WHAT IS UCC?

We are seeing the rapid growth of an unusual form of production for information resources on the Internet. The defining characteristics are that much of the information provided by a producer is *donated to* the producer, by people *not employed* by the producer. There are several names for this production technology; we favor *user-contributed content* (typically there is no boundary between contributors and users of the content).[1]

Many information producers now significantly (though rarely exclusively) rely on a user-contributed production model. Many are successful and quite socially valuable. We survey several different categories of such information services in this chapter. Well-known examples include Wikipedia articles; Amazon product reviews; Flickr photos; Digg news stories; del.icio.us Web bookmarks; CiteULike scholarly citations; open-source software projects such as Linux and the Apache web server; Peer2Patent comments on patent applications; the Stanford Encyclopedia of Philosophy, and WorldCat library cataloguing. Of course, for every success, there are numerous failures.

Providing product created and donated by volunteers seems like a great idea; why didn't anyone think of this before? Of course, the idea is not new: Tom Sawyer induced his friends and neighbors to whitewash Aunt Polly's fence for him.[2] More generally, user-contributed content is clearly related to the *private provision of public goods* [Bergstrom et al., 1986].[3] User-contributed content on the Internet exhibits some interesting features, however. For example, most prior examples involved monetary donations (e.g., public radio fundraising), not raw material or finished production (information content). Likewise, prior examples were typically produced or managed by a non-profit or government agency; much donated Internet content is given to for-profit firms. Further, when the donation is content rather than money, quality is a crucial concern. In fact, Tom Sawyer's example is quite apt: in-kind donations to a private producer, with a significant concern for quality.

In this chapter we review what economists — and to some extent, social psychologists — have to say about solving two fundamental problems facing a for-profit or not-for-profit firm that wants to rely on UCC: how to motivate contributors to *get (good) stuff in*, and motivate miscreants to *keep bad stuff out.* Getting stuff in: why would volunteers be motivated to contribute information

[1] Other common names for this phenomenon are user-generated content, consumer-generated media, and user-created content.

[2] They even paid him for the privilege of providing volunteer labor: "an apple, a dead rat on a string, part of a jews-harp, a piece of blue bottle-glass to look through, a spool cannon, a key that wouldn't unlock anything, a fragment of chalk, a glass stopper of a decanter, a tin soldier, a couple of tadpoles, six fire-crackers, a kitten with only one eye, a brass door-knob, a dog-collar – but no dog – the handle of a knife, four pieces of orange-peel, and a dilapidated old window sash" [Twain, 1876, pp. 32–33].

[3] We focus on the private provision of public goods, rather than philanthropy in general, because information resources typically are public goods: they are *non-rivalrous*, or the use of the resource by one person does not reduce the ability of any other person to use the resource [Samuelson, 1954].

content? Why spend time proofreading and correcting Wikipedia articles? Further, for many if not most information resources, the quality and variety of contributions is crucial for the social value of the resource. Thus, the problem is not just getting people to contribute, but getting the right people to contribute, or getting the people who contribute to devote effort to providing desirable quality or variety. Keeping bad stuff out: user-contributed content systems are, generally, open-access publishing platforms. They provide an opportunity for users to contribute content that is in the interest of the contributor, but outside the scope, and possibly harmful to the information service. Spam is a primary example: it shows up in blog comments, book reviews, encyclopedia articles, comments on photos, announcements to Facebook users, and so on. Manipulation is another example of bad stuff: self-interested users trying to manipulate outcomes of, say, rating and recommendation systems to favor attention to (and often sales of) their books, movies, music, and other self-interests. These are the problems we address in sections 4 – 6.

We begin with a brief survey of various types of UCC:

Information sharing systems There are many online knowledge, experience, and media sharing systems, e.g., Usenet, forums, weblogs, Amazon book reviews, Yahoo!Answers, YouTube, Flickr. In such systems, there is no specific goal beyond the collection and sharing of the information itself. Little structure is imposed on users' contributions, although their formats or type may vary. For example, while most forums use threads to organize conversations, question-answer systems such as Yahoo!Answers specify a question-answer format which limits the extent of discussions. Some systems only facilitate posting text, i.e., forums and Usenet; others allow images, video, or audio, e.g., Flickr, YouTube, and blogs.

Open, coordinated creation Some UCC systems have clear goals, which in turn break down into tasks for individual contributors. For example, Wikipedia was founded to produce an online free encyclopedia, which requires contributors willing and able to create, write, style, edit, and proofread individual entries. An open source software project usually breaks down to smaller modules written by different programmers. This division of labor makes coordination among contributors necessary, leading to the creation of policies, guidelines, or licenses. For example, Wikipedia has established a number of policies to regulate the nature of the content produced, such as *neutral point of view*, and *no original research*.[4]

Rating systems People often care about the relative quality of a product or service. Online rating systems aggregate and publish ratings contributed by volunteers. Rating systems work well for products or services for which consumers have similar tastes (they agree on what is high or low quality, albeit with various intensities). For example, most items listed on Amazon have a

[4] http://en.wikipedia.org/wiki/Wikipedia:List_of_policies, retrieved on Mar 29, 2008.

rating that indicates quality. LawyerRatingz.com allows users to rate their lawyers.

Recommender systems People's tastes may differ sufficiently for certain products or services, such that they cannot agree on a particular ranking of these products or services. For example, some care only for romantic comedies, while others strongly prefer horror films. In these cases, the ideal ratings need to be individually customized. A recommender system recommends items to best suit an individual's tastes, while relying on a dataset of opinions contributed by users. Typical recommendation systems employ some form of collaborative filtering: a technology that automatically predicts the preference of a user based on many others' tastes.[5] Two standard types of collaborative filtering techniques are person-to-person and item-to-item recommendations. Person-to-person recommender systems use users' past evaluations of items to identify pairs of people with similar tastes. The system then recommends to users items liked by others with similar tastes. There are many variations of this method (see, e.g., Breese et al. [1998]). An item-to-item recommender system relies on a similarity measure between pairs of items, instead of pairs of people [Linden et al., 2003]. When a user shows interest in an item, a set of items with high-similarity scores are recommended to her. NetFlix — a movie rental and recommendation site — uses an algorithm of this type [Bennet, 2006].

Social tagging systems Search and retrieval from many online resources, such as bookmarks, images, and videos, makes use of metadata (information about information, such as keywords, creation date, or creator identity). To a certain extent, metadata can be generated automatically using computer algorithms, such as by parsing an image's file name. However, machine semantic comprehension is limited, and information contributors often do not provide useful data from which to infer metadata (e.g., "img003.jpg" as a file name). As a result, metadata are generated, for the most part, by expensive humans. Another problem is that the person constructing a query has to guess what terms and concepts the metadata providers have associated to the object. This is hard: inter-user agreement on unrestricted labels applied to describe an information object tends to be only about 20 percent [Furnas et al., 1987, Rader and Wash, 2006].

Social tagging is a UCC approach to these problems. By this we mean resources that consist of user-contributed metadata; the content to which the metadata refer may or may not be UCC. Social tagging systems enable multiple volunteers to label objects in their own words, and to share these labels. The advantages of human over automated labeling are retained, while volunteer labor reduces the cost and the vocabulary matching problem may be reduced by

[5]Recommender system can also be built without collaborative filtering technology. For example, Pandora Radio (`http://pandora.com`) generate recommendations based on a prior manual analysis of the music files. Since we focus on user-contributed content, this type of recommender system is outside of our scope.

aggregating the labels applied by many users. Successful social tagging systems include del.icio.us (for web bookmarks), Flickr.com (for pictures), CiteULike and Connotea (for scientific references), YouTube (for videos), the Google image labeling system (based on the ESP game, `espgame.org`), and Technorati (for blog posts). These systems differ in a number of implementation features [Marlow et al., 2006].

2. WHAT IS INCENTIVE-CENTERED DESIGN?

The challenges facing UCC systems — getting good stuff in and keeping bad stuff out — are the result of behavioral choices. Though people are autonomous, their choices are often influenced, if not fully determined, by their preferences and the configuration of motivators they face. Of course choices need not be fully consistent with axioms of rationality, as we typically assume in economic theory models. As long as there is some systematic and predictable tendency in behavior responses to motivators, system designers can provide or manage incentives to induce more, and more valuable, participation. We call this approach *incentive-centered design* (ICD).

Our approach to incentive-centered design relies heavily on the economics literature, where it usually is called "mechanism design"; see, e.g., [Laffont and Martimort, 2002] for an introduction. However, the economics literature tends to focus somewhat narrowly on either monetized or direct-revelation mechanisms, and to assume full rationality. We have found that many incentive design ideas that relax this limitations can be found in the social psychology and other literatures, which do not use the term "mechanism design." Further, most UCC-specific design work to date, both in practice and in the scholarly literature, comes from noneconomists. Therefore, we adopt the more inclusive "incentive-centered design."

The scope for incentive-centered design is large. For example, designing constraints to limit or shape human behaviors (sometimes called "affordances" in the design theory literature) is usually a matter of creating incentives. A law prohibiting certain behavior, for example, can be seen as the imposition of some sort of cost (fine, jail time, social approbation) on a particular behavior, providing a disincentive to choosing that behavior, but as we know, not always preventing that behavior. A technological barrier, such as a password, can also be thought of as a (dis)incentive design: for example, password-protected systems are not impossible to enter unauthorized, but well-designed password schemes make unauthorized entry quite costly (adding, perhaps, to the legal costs imposed on unauthorized trespass if one is caught and successfully prosecuted); see, e.g., [Wash and MacKie-Mason, 2007, Herley, 2009]. In later sections 4 – 6, we offer an overview of families of incentive-centered designs and what we know about their usefulness in various UCC settings.

3. WHY NOT MARKETS?

Before we elaborate on ICD methods for UCC, we briefly discuss a natural question: why not markets? After all, we are talking about obtaining (information) inputs to a production process, and the second fundamental theorem of welfare economics asserts that any efficient allocation can be supported by a competitive (market) equilibrium. More prosaically, the Encyclopedia Britannica and New York Times pay writers: why don't Wikipedia and Amazon?

Of course, there are times when market solutions can work for UCC problems. However, we find that the fundamental market failures we discuss in the chapter are so pervasive for UCC that it is more useful to focus on non-market incentive design methods.

There are two straightforward and fundamental reasons that market solutions may not provide efficient outcomes for UCC systems: externalities and asymmetric information. Consider first the problem of keeping the bad stuff out: for example, preventing spammers (e.g., unsolicited advertisers) from monopolizing the content in a UCC resource. The problem, in part, is analogous to pollution: in pursuing its own self-interest (advertising), the spammer generates detritus that reduces the utility of others. In addition, undesirable content can be characterized as an asymmetric information problem of the hidden information (or adverse selection) type: spammers know who they are *ex ante*, but UCC system managers and content users do not know until after they bear the cost of seeing and processing the spam. If we knew spammers' type *ex ante*, we simply would deny them permission to use the UCC platform.

The externality problem for getting good stuff in generally goes the other way: contributed information often is a public good, and thus potential contributors benefit from the contributions of others, and have a strong incentive to free ride. Asymmetric information is also a problem. For example, individuals' willingness to contribute (say, due to altruism or social-regarding preferences) generally is hidden information, as is their talent or differential ability to produce quality content. Creating high-quality content also requires effort, not just talent, and that effort is often hard to monitor, leading to a hidden action (moral hazard) problem of asymmetric information.

Of course, information content and processing effort are often obtained through labor market solutions, even on the Internet. For example, MySpace has an army of paid employees who review all user-posted photos, in order to delete those that are inappropriate. The externalities are monetized (through advertisers who pay for the free-riding visitors to the web pages) and the hidden action problems have, apparently, been adequately resolved through standard contracting and monitoring methods. So why do so many services rely on unpaid volunteers?

We believe that another problem favors the use of volunteers over employees in many UCC settings: transaction costs. Successful UCC systems might have hundreds or thousands of contributors; in Wikipedia's case, millions. These volunteers are sometimes allowed to participate without transacting at all, providing anonymous contributions without even creating an account. In most cases, there is only the minimal, automated transaction of creating an (often pseudonymous) account. The producer avoids the costs of verifying identity and immigration status; does not need to collect and report any employment data to the tax and government pension authorities; need not keep time or effort records, nor process and document payments; and so on. These transactions costs generally are small enough when hiring full-time employees, and even part-time employees who work 10 or more hours most weeks. However, in UCC systems many volunteers might contribute only once, or quite infrequently, and the transactions of contractually employing them would swamp their value.[6]

4. GETTING STUFF IN

A main reason that the problem of getting stuff in cannot be solved by markets efficiently, is the externality in UCC participants' contributions (see section 3). Most UCC systems make their content publicly accessible: everybody can read any Wikipedia article, browse any bookmarks on del.icio.us, and watch any video on YouTube. These content exhibit characteristics of *public goods*: they are *nonrivalrous* — user A's consumption of the good does not diminish user B's. In general, such type of goods are under-provided if provided voluntarily [Samuelson, 1954]. Why contribute if I could consume it for free? In UCC in particular, the under-provision problem of public goods is especially pernicious as the content provided usually only benefits others, not the provider himself.[7] It would appear that individuals have little incentive to share their private information or knowledge with others.

For the under-provision problem of public goods, economists have proposed many theoretical solutions, mostly focused on settings in which consumers collectively fund some public goods. Some of these solutions achieve the socially efficient levels of provision [Groves and Ledyard, 1977, Walker, 1981, Bagnoli and Lipman, 1989, Admati and Perry, 1991], and some improve on voluntary contribution [Morgan, 2000]. These solutions offer valuable insights into what economic mechanisms can achieve in a setting where side-payments are permissible among the participants. However, in most UCC systems, side payments are either not possible (e.g., due to transaction costs) or not desired by the members of the system (e.g., due to social or psychological costs). It is unclear how these classical mechanisms could be implemented in UCC systems that do not support side payments.

These classical solutions also miss the novel opportunities offered by information technology to motivate content contribution. In particular, information

[6]Mechanical Turk, at `https://www.mturk.com/`, is one attempt to create a low-transaction marketplace for piece-rate online labor. Mechanical Turk was unveiled in 2005, but is still in beta and currently offers only about 2000 different employment tasks on a typical day.

[7]That is, the contributor already has the information and the benefits of its use. Contributing it to the public resource only provides additional benefits to the contributor if it is enhanced or improved as a consequence.

technology makes it relatively inexpensive to collect, process, aggregate, and disseminate information. What these capabilities offer to the UCC systems is a vast open area to be explored by both researchers and UCC system designers. Here, we discuss a few areas in which information technology has shown its promise.

First, technologies enable UCC systems to collect information generated by their users in the process of enjoying the functionalities provided by the site. Such information can then be aggregated and made public. For instance, del.icio.us, a bookmark sharing website, provides its users with a personal online repository for their bookmarks, along with certain valuable features, including the convenience of accessing bookmarks from any networked computer, and a tagging system for organizing them. Then, as individuals submit and annotate bookmarks for their own information management purposes, del.icio.us automatically aggregates these privately motivated contributions, and makes them public by default[8]. At least in part, the publicly available bookmarks are a side effect of the individual, private activities. In section 4.1 we discuss various ways of motivating user contribution by offering them useful functionality.

Second, information technology allows UCC system designers to use exclusion ("If you do not contribute, you do not get access") to motivate contribution. Using exclusion as a motivating instrument effectively converts pure public goods into impure (nonrivalrous but excludable) public goods. Exclusion is not new: club goods — congestible public goods that are excludable, such as private parks and swimming pools — are funded at least in part through their excludability. Information technology significantly enhances UCC systems' ability to monitor and evaluate individuals' contributions. Based on each individual's contribution level, the system can dynamically customize her level of consumption. For example, one peer-to-peer (P2P) file sharing community, *ilovetorrents.com*, uses a simple exclusion rule: it requires its members to maintain a minimum upload/download ratio to be eligible for continued downloading.[9] In section 4.2, we review the literature on using exclusion to motivate contribution in the context of UCC.

Last, by facilitating social interactions, information technology provides UCC system designers yet another lever for motivating contribution: social motivators (mechanisms that make use of incentives that are not traditionally considered economic, but which have received attention in, e.g., social psychology). People respond to a range of motivators such as their social identity, social norms, and whether people are rewarded equally for their efforts. In section 4.3 we discuss studies that have used these insights to motivate contribution into UCC systems.

[8]Users can opt-out of sharing their bookmarks, but they have to do it for each individual bookmark.

[9]See `http://www.ilovetorrents.com/rules.php`, retrieved on Dec 11, 2008.

4.1. Side-effect public goods

Some UCC systems offer useful functionalities to their users from which they privately benefit. These privately-beneficial functions might be provided in exchange for, or as a result of contributions. Then, with user permission (either opt-in or opt-out), the contributions can be aggregated and provided to the public. We characterize these as side-effect public goods systems, and identify four categories of functionalities that we have seen in use to motivate user contributions.

First, providing directly consumable products offers robust incentives for user contribution. We have mentioned Del.icio.us, which offers its users a personal bookmark management system, as well as a well-tagged public repository of bookmarks as a spill-over [Wash and Rader, 2007]. Similarly, movie recommender systems offer their users personalized movie recommendations. A user rates movies in order to receive good recommendations for herself. In the mean time, her ratings benefit other users as the recommending algorithm takes her input into account while generating recommendations for other users. Indeed, users on MovieLens[10], a movie recommendation website, self-reported that a main motivation for them to rate movies was to improve the quality of their own movie recommendations [Harper et al., 2005]. In these settings, users may not realize, nor care that they are contributing to a public good. Even if they do care, the distinct private benefits they receive may be necessary (and perhaps sufficient) to motivate their contributions.

Second, many users find participating in UCC systems intrinsically enjoyable. Both fun and ideological reasons attract contributors to Wikipedia, but it is fun that causes a high level of contributions [Nov, 2007]. MovieLens users rate movies for fun as well [Harper et al., 2005]. von Ahn and Dabbish [2004] created a game in which pairs of players try to guess each other's choice of tags to the same images. While tagging images is a boring task, this matching game makes it entertaining. Each player consumes her entertainment, and at the same time contributes metadata, i.e., tags, to those images. These tags are valuable metadata for performing online image searches, or tuning image recognition algorithms. Google uses this system as the Google Image Labeler with a goal of tagging all publicly accessible images on the Web to make them more findable by searchers.[11]

Third, sometimes contributing to UCC systems is an opportunity to learn valuable skills. For example, learning programming in a team environment employing industry standard group engineering tools is an important reason for open source programmers to participate in the projects [Raymond, 1999, Hars, 2002, Lakhani and von Hippel, 2003].

Last, UCC systems offer a good place for people to receive feedback, or establish a reputation for their ability. The contributions themselves, and feedback on them evidence an individual's ability, which might lead to future career opportunities [Lerner and Tirole, 2002]. Feedback may also have a warm-glow

[10]See http://movielens.umn.edu, retrieved on Sept 21, 2010.
[11]See `http://images.google.com/imagelabeler/`, retrieved on Oct 26, 2010.

effect, making the receiver feel that she has "done her bit" [Andreoni, 1995]. Sometimes feedback is in and of itself rewarding: people want to know how competent they are compared to others [Ryan and Deci, 2000].

4.2. Exclusion-based mechanisms

Using exclusion to motivate contribution is not a new idea. Clubs solve the getting-stuff-in problem by charging membership and/or use fees, and excluding those who do not pay.[12] Clubs for sharing excludable public goods have been studied and are often modeled as cost-sharing games [Moulin, 1994, Deb and Razzolini, 1999, Young, 1998, Bag and Winter, 1999, Moldovanu, 1996, Dearden, 1997]. We rule out monetary payments in the definition of UCC, but this does not mean that excludability is unhelpful. Consider barter, as in a babysitting co-op: parents are motivated to donate their time to babysit for others because they get access to an in-kind good ("free" babysitting for their children). In other words, some form of exchange may be available to motivate provision, but for UCC systems we limit our consideration to *non-monetary exclusion mechanisms*.

The advancement of information technology makes it easy to exclude any individual based on her contribution level, at any time. For networked computational information systems, access control is relatively cheap, with varying degrees of identification possible using passwords, Internet Protocol (IP) addresses, or cookies. Some metrics on contributions can be automatically monitored (such as number by contributor, and total bytes), creating the possibility of algorithmic exclusion based on contribution. Contribution quality can be evaluated and quantified either by human raters, or algorithmically in some cases. Examples of human ratings include user ratings of book reviews on Amazon, of answers on Yahoo!Answers, and of comments posted on Slashdot.org. As an algorithmic example, the value of each edit on Wikipedia can now be measured by how long it survives subsequent deletions or modifications [Adler and de Alfaro, 2007].

Although information technology offers some opportunities for using exclusion-based mechanisms to motivate content contribution, systematic exploration of exclusion-based mechanisms is yet to begin. A natural starting point is to adapt the mechanisms proposed in the cost-sharing literature [Moulin, 1994, Deb and Razzolini, 1999, Young, 1998, Bag and Winter, 1999, Moldovanu, 1996, Dearden, 1997] to the context of UCC. Such adaptation is not straightforward. First of all, the cost-sharing mechanisms were proposed to motivate monetary contribution, but a UCC system designer needs to elicit contribution in the form of time and effort (see section 5 for a detailed discussion on this difference). Second,

[12] We do not delve into the large body of literature on club goods (see Sandler and Tschirhart [1980], Wildasin [1986], Starrett [1988], and Scotchmer [2008] for reviews of the literature), as most of these studies treat club goods as impure public goods. Thus the goods are to a certain extent rivalrous. Barring special cases under which a member of a UCC system might derive negative utility from an additional member (e.g., a large user population reduces one's sense of community), we treat the content on UCC as non-rivalrous goods.

even if we set aside the issues implied by the difference between monetary and non-monetary contribution, it might still be hard to implement the cost-sharing mechanisms in UCC systems, due to various reasons. For example, Dearden [1997] and Deb and Razzolini [1999] propose mechanisms to share the cost of an indivisible public good (e.g., a new bridge), assuming each participant knows her private valuation *a priori*, regardless of others' valuation. In UCC systems, however, the amount of the public good produced is not known in advance. The goals of an effective mechanism are to determine both how much content to produce and how much each individual contributes. Other mechanisms, e.g., the serial cost sharing mechanism [Moulin, 1994] and the subscription based mechanism Bag and Winter [1999], require complex communication or coordination among the participants, which is not feasible for many UCC systems with a large number of participants.

In practice, however, simple and effective exclusion-based non-monetary mechanisms have been used in popular UCC applications. Some peer-to-peer (P2P) file sharing communities, e.g., ilovetorrents.com, use a simple exclusion rule: it requires its members to maintain a minimum upload/download ratio to be eligible for continuing to download files.[13] As another example, Glassdoor.com provides free information on salaries of a wide range of professionals, if you sign up and report your own salary.

Although these simple mechanisms may not always be Pareto efficient or incentive compatible, the fact that they are used in practice indicates that it is useful to examine their properties in a systematic manner.[14] A number of human-subject laboratory experiments have been conducted to investigate these mechanisms, under the assumption that individuals have homogeneous preferences. For example, Croson et al. [2008] found that simply excluding the person with the lowest level of contribution can be sufficient to motivate high levels of contribution. Swope [2002] found the minimum threshold mechanism (MTM) — one only gets to access the full content if her contribution level is above a pre-specified threshold — can in some situations achieve full cooperation, given that the gain from a fully coordinated cooperation is sufficiently high. When the gain from full cooperation is relatively low, though still higher than no cooperation, people have trouble coordinating on the full cooperation equilibrium. Alternatively, instead of specifying a threshold exogenously, Kocher et al. [2005] let the subjects themselves determine the threshold, and found that such an endogenously determined threshold also helps reducing free-riding, and that the time at which the announcement of the threshold is made affects the outcome.

Two recent analytical studies consider the performance of an MTM and a ratio mechanism when individuals have heterogeneous preferences. The ratio

[13]See `http://www.ilovetorrents.com/rules.php`, retrieved on Oct 13, 2010.

[14]As a reminder: "incentive compatibility" in the economics mechanism design literature refers to a direct revelation mechanism in which it is rational for a participant to reveal his or her private information truthfully. Incentive compatibility is often imposed as a constraint in theoretical derivations of optimal mechanisms, because the revelation principle assures us that we lose no generality by restricting consideration to incentive-compatible direct revelation mechanisms. However, incentive-compatibility may not be per se desirable, and in practice many, perhaps most, revelation mechanisms are in fact not incentive compatible.

mechanism permits a participant to consume no more than an amount that is a fixed proportion to her own contribution. Assuming that all the individuals can be ordered by their marginal net utility of the public goods, Wash and MacKie-Mason [2008] show that MTM can improve the efficiency compared to a voluntary contribution mechanism. Assuming individuals have homogeneous cost functions but heterogenous valuation of the public goods, Jian [2010] found that MTM has some advantages over the ratio mechanism.

In summary, although some mechanisms to motivate contribution to excludable public goods have been studied before, their applicability to UCC systems needs systematic examination. Meanwhile, new mechanisms emerge in the field. Further research is required to understand the properties of these new mechanisms, to enable systematic empirical comparisons and to derive generalizable design recommendations.

4.3. Social motivators

People derive happiness not only from consumption-related activities, e.g., enjoying a movie, but also from non-consumption related activities, e.g., being altruist (see [Fehr and Fischbacher, 2003], for a review) or helping society achieving fairness [Knoch et al., 2006]. Thus, in the context of UCC, users might care intrinsically about how much they themselves contribute (independent of any extrinsic motivators provided for contribution), and not just how much they consume. Here we introduce a few, non-exhaustive, examples of ways to motivate contribution, drawing on both economics and social psychology literature.

Collective effort model Karau and Williams [1993] developed the collective effort model (CEM) to explain individual effort in performing collective tasks, and used this model to identify factors that help reducing social loafing, or free-riding. Examples of CEM's predictions are that individuals work harder when they believe their own contributions lead to identifiable values in the aggregated outcome, when they work in relatively small groups, when they like their teammates or identify with the group, or when they consider the group tasks meaningful.

A number of CEM predictions were recently tested in a UCC system, MovieLens. On MovieLens users rate movies, receive movie recommendations from others with similar tastes to themselves, and participate in discussions. Having a sufficient number of movie ratings is essential for the recommendation algorithm to produce accurate predictions for its users. In fact, more than 20 percent of the movies do not have enough ratings to make predictions about them to the users. The system designers would also like to encourage people to post messages to the discussion forums, in order to make the site more social. The researchers [Ling et al., 2005] decided to use insights suggested by the CEM to motivate movie rating contributions. One particular treatment they implemented was based on the insight that people work harder if they think their contributions are unique. They divided the subjects into a treatment and a control group. The participants in the uniqueness treatment group received weekly email messages to remind them how unique their movie ratings are, and to invite them to participate in an online discussion. The non-uniqueness control group only received an invitation to an online discussion. They found supporting evidence for the uniqueness hypothesis: people who were reminded of the uniqueness of their contributions rated more movies and posted more often in the discussion forums.

The CEM also predicts that people are more likely to perform a task for their group if they know the value of doing so. This prediction is confirmed by another field experiment conducted on the MovieLens community. Rashid et al. [2006] showed each user a list of movies each with smilies indicating how much value she would have contributed to the community if she rated this movie.[15] They found that users were more likely to rate a movie if there were smilies indicating the potential value of rating this movie.

Social identity The premise of the social identity theory is that a person derives an identity of self (social identity) from his or her membership in social groups, such as being a woman, or a caucasian [Tajfel and Turner, 1979]. One predicted effect of social identity is that it creates in-group favoritism: people make choices that favor their in-group members and have higher opinions of their in-group members than those out-group. These predictions are supported by a large number of experimental studies in social psychology [see, e.g., the following literature surveys: Tajfel and Turner, 1986, Deaux, 1996, Hogg, 2003, McDermott, 2009]. Building on the existing literature, Ren et al. [2007] predict that in online communities on which many UCC systems rely, in-group favoritism may take the following forms: group cohesion and commitment, more focused conversations than off-topic chat, increased levels of contribution to public goods and less social loafing, conforming to group norms, welcoming new members, and engaging in generalized reciprocity (when one's giving is reciprocated by a third party not the receiver). A group with a strong identity may also attract new members to join and contribute to the group's goal. For example, Bryant et al. [2005] and Nov [2007] both document that Wikipedia attracts new editors due to people's belief in its goal of producing a useful online knowledge repository.

There are multiple ways to induce social identity in an online setting [Ren et al., 2007]. One can categorize people into groups based on their existing social categories, such as organizational membership, ethnicity, or their political values. One can also simply assign people into arbitrary groups [Tajfel and Turner, 1986, Postmes et al., 2002, Worchel et al., 1998]. In addition, group identity can be induced by emphasizing in-group homogeneity, such as that all members share a joint task [Worchel et al., 1998, Cartwright and Zander, 1968, Sherif et al., 1961], a common purpose [Postmes et al., 2001], or a common fate [Michinov et al., 2004]. Group identity can also be made salient by doing inter-group comparisons, and emphasizing the differences from the out-

[15] The value of the ratings are calculated using algorithms that take into account the similarity of this movie to other movies in the database. See [Rashid et al., 2006] for details.

groups [Postmes et al., 2001, Rogers and Lea, 2005, Worchel et al., 1998]. For example, Wikipedia editors frequently compare their product to Encyclopedia Britannica[16], and open source software project (e.g., Linux) developers often compare their product against commercial proprietary products (e.g., Microsoft Windows).

Using social identity theory to motivate contributions is an active research area. Researchers have made progress in incorporating social identity into economic theories [for example, Akerlof and Kranton, 2000, 2002, 2005, Benabou and Tirole, forthcoming], and in measuring and quantifying the effects of social identity on people's social preferences (how much people care about others' earnings) [Chen and Li, 2009]. Towards the implementation of social identity theory in UCC systems, open questions remain. For example, Ren et al. [2007] proposed many design principles for fostering group identity in online communities, such as encouraging newcomers to observe the group norm before contributing content, and restricting off-topic discussions among contributors. The effectiveness of these proposals need to be put to test, especially by field experiments, before precise design suggestions can be made to practitioners.

Social comparison, norms, and social preference Cyber-infrastructure facilitates collecting, aggregating, and disseminating social information in real time. Information about what others do can affect individuals' behavior in a number of ways. Comparison with superior others may trigger self-improvement motives [see Wood, 1989, for a review]. People have the tendency to conform to social norms [see Cialdini and Goldstein, 2004, for a review]. People may also fundamentally care about other people's welfare, for example exhibiting inequality aversion: people do not like to earn different amounts from others and would like to reduce the difference if they can [Bolton and Ockenfels, 2000, Fehr and Schmidt, 1999]. Some people even try to maximize social welfare, just like a social planner [Charness and Rabin, 2002].

In light of the predictions of these theories, UCC system designers can strategically deliver information about what other participants do to UCC users in order to induce increased contributions. In fact, some UCC systems are already using social information to motivate contribution. For example, leader boards for Yahoo!Answers and Amazon book reviews are presumably intended to motivate some users to aspire to contribute as much or more than the leaders.

There have been a few empirical studies on using social information to motivate contribution to UCC systems. For example, while discussing the collective effort model we mentioned MovieLens, a site that relies on user-contributed ratings to make personalized movie recommendations. In this system under-contribution is common. Based on insights from the social comparison theory, Chen et al. [forthcoming] designed a field experiment to test the use of social information on MovieLens contributions. After receiving information about the median user's total number of movie ratings, users below the median exhibit a 530 percent increase in the number of monthly movie ratings, while those above

[16]See http://en.wikipedia.org/wiki/Reliability_of_Wikipedia.

the median do not necessarily decrease their ratings.

5. CONTRIBUTION QUALITY

When designing for *content* contributions, the variable quality of information content is an unavoidable issue. A large literature on public goods contribution focuses on *monetary* donations; this literature can ignore quality because money is homogeneous [Andreoni, 2006]. But information is not: however its quantity is measured, its value per unit will vary across users, or across units of content, or both. Some content may even have negative value (to at least some users) because it offends or psychologically harms (some would say obscenity falls in this category), or because the clutter it causes — and corresponding increase in search or filtering costs — outweigh the value to the recipient (spam email is an obvious example). We address negative-valued content ("bad stuff") in section 6. In this section we address variations in the quality of contributions with non-negative value.

The differences between contributed content and contributed money — "crowd-sourcing" versus "crowd-funding"— are fundamental. First, monetary contributions from different individuals are substitutable and additive. Content quality is heterogeneous, however, and so contributions are not always substitutable; when they are, the marginal rate of substitution may vary across users or across different units of content (two different articles on user-contributed content will not be equally good substitutes for this chapter). Second, if content of varying quality is not treated the same (e.g., if incentives are created to induce provision of desirable quality mixes), then the cost of *evaluating* quality must also be considered in system design. The fundamental differences between content and monetary contributions require fundamentally different solutions.

The quality of user-contributed content generally must be characterized in multiple dimensions. Economists have long partitioned the space into two sets of quality dimensions: "vertical" and "horizontal." A vertical quality characteristic is one about which (essentially) all agents agree on a rank ordering (a Porsche is better than a Pinto) [e.g., Mussa and Rosen, 1978]. A horizontal quality characteristic is a feature over which agents have heterogeneous preferences ("Alice likes red cars; Bob likes blue") [e.g., Lancaster, 1966]. UCC system designers can, and often do, enable and encourage the provision of multiple levels of vertical quality, and varieties of horizontal quality. For example, Slashdot.org asks readers to evaluate comments posted by other users on a numeric (vertical) scale, and also to designate various adjectives that horizontally characterize the comment (such as informative, funny, insightful, and off topic). A photo uploaded to Flickr has numerous attributes (that users may or may not provide) including the photo groups in which it is included, tags assigned to it, and the geographic location at which the photo was taken. At Amazon, readers are encouraged to contribute ratings of products: these are subjective and clearly horizontal (or a mix of horizontal and vertical); Amazon then makes public an aggregate evaluation score (that is, a vertical numeric indicator between zero and five), as well as the heterogeneous ratings by individuals (and their even more heterogeneous written comments).

Depending on the particular type of content that a UCC system designer aims at producing, she will have preferences over the mix of horizontal and vertical quality. She then designs mechanisms to motivate the users to create content that provide desirable quality characteristics (taking into account the cost of inducing those qualities). Next we discuss mechanisms that can, in general, be employed to obtain various mixes of horizontal and vertical quality, though we can expect that different mechanisms will work better or worse for different desired quality mixes, and in different contexts. There is very little empirical literature to date on the effectiveness of these mechanisms for eliciting desired contributed-content quality.

To elicit content with her desired mix of vertical and horizontal quality characteristics, a UCC system designer must address two incentives problems: hidden information (also known as adverse selection), and hidden action (also known as moral hazard). The first concerns heterogeneity across users. UCC contributors vary in their abilities and the UCC system designers usually have little information about individual contributor abilities. The abilities (often referred to in the economics literature as "types") are thus private, or hidden, information. The system designer needs mechanisms to elicit more contributions from users with desirable types, and less (or none) from those with undesirable types. The second problem — hidden action — concerns variability in a given contributor's effort to produce desirable quality characteristics. Whatever a contributor's abilities, it is generally the case that he can produce more desirable quality with more effort. But the amount of effort invested by a contributing user is generally unobservable (or at least, unverifiable). The problem of hidden action is to motivate a desirable level of effort.

In the following three subsections, we discuss three commonly used mechanisms: moderation, rewards, and reputation, while keeping in mind how they each address the hidden information and/or hidden action problems.

5.1. Metadata

Metadata, or information about information, can be a vehicle carrying an intrinsic incentive payload in various applications. The nature of the user-contributed content application will determine which types of metadata mechanisms are most useful. For example, ratings are a type of metadata most directly suited for vertical quality dimensions; recommender systems tend to be designed for horizontal quality dimensions ("people who liked John Irving's most recent novel often also like works by Nicholas DelBanco").[17] [18]

[17] Of course, both might be used in either setting, and more generally, combinations of them might be used. For example, Netflix collects (vertical) ratings from movie viewers, but then uses a collaborative filtering technique to summarize the ratings of "viewers who like what you like" in order to provide recommendation to potential viewers.

[18] Not all metadata have an obvious use for the provision of intrinsic motivation. For example, Library of Congress classification numbers for books are valuable metadata, but it is

In many UCC systems the process by which ratings are assigned to user postings is called moderation. Moderators can be designated community members. For instance, the Encyclopedia of Life (www.eol.org) assigns a curator for each of its pages [Encyclopedia of Life, 2010]. Peer-moderation, sometimes by randomly selected peers, is also popular. For example, Plastic.com and Slashdot.org both invite randomly selected users to moderate others' postings. Slashdot additionally employs a meta-moderation mechanism: a meta-moderator checks a moderator's ratings.

An immediate use of a moderation system is filtering. The system, or individual users, can choose to filter out low-rated content. Further, to the extent that content contributors value their readership, filtering based on moderation can induce contributors to make increased effort to provide higher-quality content. Using game-theoretic analysis, Chen et al. [2007] show that when the probability of one's post being moderated is high enough, an opportunistic contributor who tries to maximize his/her readership and minimize efforts will exert high effort. When this probability of moderation decreases, an opportunistic contributor either uses a mixed strategy of exerting high and low effort, or always exerts low effort.

A moderation system might be designed to elicit hidden information. MacKie-Mason [2010] models a system in which content contributors post a bond which is revoked if user ratings do not exceed a fixed threshold. He shows that such a system can satisfy the conditions for a screening mechanism with a separating equilibrium (contributors of poor content self-select out, and do not contribute). The key conditions, as usual for screening mechanism, are that the cost of contributing be higher for the low type, and that the cost be increasing faster for the low type (the Spence-Mirrlees condition).

Of course, a peer-moderation system relies on UCC, itself. Thus its success depends on solving the same problems of contribution quantity and quality, this time from the peer moderators. Do the moderators do their job when they are given the chance? By analyzing Slashdot's server log data, Lampe and Resnick [2004] found that its moderation system works fairly well: many users moderate and/or meta-moderate posted comments. Another question we might ask of a vertical ranking system, which is one of Slashdot's mechanisms, is whether moderator opinions about particular posts converge, as what we expect with vertically-distributed consumer tastes? Lampe and Resnick found that moderators do reach consensus: 92 percent of all meta-moderations show agreement with the evaluated moderation, suggesting that readers' tastes are close to vertically distributed (in the "overall quality" dimension applied to Slashdot articles), and that Slashdot's use of a vertical ranking mechanism may be effective.

---

hard to imagine that they have much effect on the motivation of writers. In this section we are interested in metadata that at least in part provides motivation. Metadata used for other purposes (such as finding and selecting information resources, e.g., search keywords and tags at `dmoz.org`, `citeulike.org`, and `del.icio.us`) may be generated as user-contributed content, through systems with their own incentives problems to manage; many of our examples in earlier sections were user-contributed metadata.

5.2. Rewards

For some content contributors, the act itself of contributing is rewarding, be it psychologically, socially, or even career-wise, as we discussed in section 4. For others, explicit rewards might help motivating them to exert higher effort.

For example, Wikipedia puts up its "featured articles" on its front page, as a way to showcase its best work. These articles are nominated and selected via voting among peer Wikipedia editors. When an article is featured, it remains in the "Today's featured article" section on the front page for a day. In addition, a golden star appears on the top of the article's page thereafter to indicate its "featured" status. Although articles on Wikipedia are usually products of collaborative work by multiple editors, some major contributors take ownership. For example, many editors put links to the articles they worked on onto their personal Wikipedia user pages [Bryant et al., 2005]. Bryant et al. quote from an interview of an editor to illustrate how the potential of having their articles featured motivates editors to make high efforts:

> Recently I've been working on the article ... as a featured article candidate. If my article is accepted as a featured article, it will appear on the main page with a multi-paragraph excerpt and photo. Featured articles stay on the front page for a day, and then they're swapped for another, so I'm really just trying for bragging rights with this one.

Many explicit rewards for high-quality contributions in UCC systems are in the form of "bragging rights." However, some use direct financial reward. YouTube (youtube.com) invites heavily-followed content contributors to share the (advertisting) profits generated by their videos. For the successful few, income from YouTube is sufficient for making a living [Stelter, 2008].

YouTube relies on its managers to hand pick individuals (called "partners") whose video postings qualify for profit sharing. This partner-selection process could be automated. If the value of a piece of content will be revealed shortly after its creation, automatic profit-sharing could be directly based on the value of the content that a user has created [Bhattacharjee and Goel, 2007]. In the event that the value of a piece content would not be revealed in the near future, alternative evaluation methods are required. In the context of rewarding honest reporting of ratings, e.g, ratings for movies, music, or academic articles, Miller et al. [2005] propose to reward a user's agreement with other users: if a user's rating agrees with others', there is a higher chance that she is right. Miller et al. show that, under some general conditions, the Nash equilibrium of this mechanism is for every rater to report truthfully. Of course, successful implementation of this proposed mechanism requires that certain assumptions be maintained, including in this case that users are not able to manipulate the system.

5.3. Reputation

Reputation systems can solve the problem of hidden contributor ability information (contributors know their own abilities, but the UCC system designer does not). The essence of a successful reputation system is that it satisfies a set of sufficient conditions for a signaling mechanism with a separating equilibrium: that the cost of obtaining reputation satisfies the single crossing property [Spence, 1973, Rothschild and Stiglitz, 1976]. That is, low-ability users' cost of attaining reputation has to be higher than that of high-ability users'.

A reputation system can also be used to address the moral hazard problem, that is, to motivate contributors to exert desirable levels of (hidden) effort. Consider a typical system that accumulates a reputation score for participants: the score, in general, increases with the amount of effort a user puts in, thus providing an observable factor correlated with the unobservable effort. If the observable — reputation score — is desirable, or can be made so, then contributors can be motivated to exert more desirable effort. When the value of higher-quality contributions can be monetized by the UCC system, contributors can be offered monetary awards correlated with their reputation score.[19] Of course, a high reputation score itself might be desirable. For example, research documents that humans have basic psychological needs to feel competent [Ryan and Deci, 2000].

A reputation score can be generated by aggregating the ratings that an individual has received in the past. For example, Wikipedia editors gave each other barnstars — placing an image of a star on the receiver's profile page, usually with text explaining why the barnstar is given — to reward their "hard work and due diligence" [Wikipedia, 2010b]. Ever since the barnstar system was introduced in 2003, various kinds have been introduced to reward editors performing all kinds of specific tasks. The "photographer's barnstar" is to reward editors who "tirelessly improve the Wikipedia with their photographic skills." And the "anti-vandalism barnstar" is to reward editors who "show great contributions to protecting and reverting attacks of vandalism on Wikipedia."

Besides aggregating explicit peer ratings, in some UCC systems, one can automatically generate reputation scores for the users based on implicit reactions of their peers [Yang et al., 2008, Chen and Singh, 2001, Adler and de Alfaro, 2007]. For example, Adler and de Alfaro propose an algorithm to construct reputation scores for Wikipedia editors based on the amount of their contribution that survives other editors' subsequent edits. That is, if one's edits are of high quality, others will also preserve them on Wikipedia instead of quickly deleting them.

[19] Of course, monetary awards suffer from the problems we described in section 3, but one we think is often the most severe — transaction costs — may be reduced if rewards are only paid to high-scoring participants, and only after they have accumulated a substantial number of contributions (rather than a micro-payment per contribution).

6. KEEPING BAD STUFF OUT

In section 5 we focused a contributor who is motivated to provide positive value to the UCC system, but due to variations in effort or capability may provide higher or lower quality. A different type of contributor is motivated by a goal distinct from the objectives of the UCC system, and thus may *desire* to contribute negative quality (from the platform's perspective).

There are many different types of negative-value content contributions. One of the most familiar is email spam and its cousins. Another is writers using pseudonyms to write great reviews for their own books on Amazon [e.g., Harmon, 2004]. Another is intentional manipulation, typically by exploiting information processing algorithms on which a UCC system relies, such as a movie distributor that might manipulate the ratings and recommendations made by Netflix. Some content providers use UCC systems to distribute malware, such as software carrying a disguised payload (trojan horse) that when installed on a user's machine creates a "bot" to collect and transmit personal (e.g., financial) information back to the perpetrator, or to take control of the computer's resources to participate in denial of service attacks, or as spam-forwarding mail hosts. It is beyond our scope to offer a comprehensive discussion of all security threats, of course.[20] In this section we focus on threats that come in the form of bad content, and some of the incentive-centered design approaches available for keeping the bad stuff out.

6.1. Pollution: spam and negative externalities

A producer soliciting user-contributed content often tries to lower the costs of contribution by offering a simple, easy access content collection system. Further, in many applications, the producer will spend little or nothing on content selection or editorial functions, instead making most contributed content available to consumers.[21] For example, anyone can post a book or product review to Amazon, and this content will appear to all users shortly thereafter [Amazon.com, 2010]. Amazon is providing an open access publishing platform (albeit with limited functionality), and some might want to publish information other than book or product reviews. For example, a publisher might wish to post a pseudonymous "book review" that is in fact just an advertisement for the publisher's products. Email is another relatively open-access publishing platform, and in that context we refer to unsolicited and unwanted advertising as spam. The phenomenon is widespread, and has led people to coin terms for it in other information product or service contexts, such as splog or blam (unsolicited advertisements in blog comments), spim (instant messaging), spamdexing (online indices), sping (blog pings), m-spam (mobile phones), spit (voice-over-internet telephony).

[20] See [MacKie-Mason, 2009] for on overview of the application of incentive-centered design to security problems.

[21] But, see Section 5.1 for our discussion of the use of user-contributed metadata for filtering and selecting.

As mentioned earlier, this is a common problem for UCC because two conditions often hold: some agents want to contribute content that has negative value to the producer, and the producer offers an open publishing platform. When these obtain, the producer faces a special instance of the quality management problem: how to keep out content that has negative value. This is closely related to the classic pollution or negative externality problem: some agent wants to engage in an activity productive for itself (here, for example, advertising), but in so doing imposes costs on others (here, the producer providing the publishing platform) without bearing (internalizing) those costs.

For traditional pollution problems, the Coasian efficient bargaining solution generally fails due to undefined property rights and/or high transaction costs [Coase, 1960]. Therefore, many mechanisms focus on these problems [MacKie-Mason, 2000]. Property right ambiguity is not a central problem for user-contributed content in most applications. Ownership of the producer's publishing platform is usually well-established. Likewise, the content supplier's copyright over content is usually clear enough; in any case, uncertainty about the content creator's copyright does not explain why production is organized to rely on content inputs donated by volunteers.

Transaction, or bargaining costs, on the other hand, are often significant. As we discussed earlier, a defining characteristic of production based on user-contributed content is that there are many input suppliers, most of whom are providing a small quantity of the producer's inputs. Negotiating, contracting, monitoring and transacting payments with hundreds or thousands of micro-suppliers to internalize externalities they impose typically will be too costly to support. We may need look no further than transaction cost economics to understand why supplier contracting is not employed to reduce or eliminate polluters.

Nonetheless, transaction costs are not the only, and perhaps not even the most important problem for producers who want to limit user-contributed pollution. With user-contributed content the pollution problem is complicated by the coincident presence of the private provision of public goods problem (*getting stuff in*, see Section 4). That is, the producer wants to encourage content *contributions*, but simultaneously wants to discourage *polluting* contributions. Mechanisms implemented to discourage pollution must be sufficiently discerning that they do not overly discourage good contributors as well; conversely, mechanisms to encourage good contributions must not too greatly encourage polluting contributions.

Providing different incentives for different types of behavior is straightforward unless it is difficult to distinguish between the two types of behavior. But this is precisely the problem with much user-contributed content: quality is not costlessly observable, so there is an evaluation problem: how to distinguish good from bad content at low cost, and before the pollution has imposed a cost on the producer? Because information is an experience good, information pollution typically is accompanied by a hidden information problem: the polluter is better informed *ex ante* about whether his content is polluting or desirable to different classes of affected users. For example, Amazon has rules characterizing permissible content contributed in the form of product reviews and will remove postings it discovers that violate these terms of use. But there have been several documented (and surely a large number of undocumented) instances of interested parties falsely posting "arm's length reviews" to promote purchases of their own, friends or clients' products [Dellarocas, 2006, Harmon, 2004, Lea, 2010]. Evidently the cost of detecting this kind of misrepresentation is sufficiently high that Amazon chooses not to spend the monitoring resources that would be needed to keep all pollution of this sort from its site.

Thus, the incentive problem facing the designer of a UCC system that provides some open-access publishing capabilities is two-fold: induce contributors to reveal whether they are contributing pollution (content of negative value), and if identified, induce them to stop providing it. The latter problem is usually the less challenging: once polluting content is identified, it is relatively inexpensive in digital information systems to block it or purge it. Therefore, we focus on incentives for the former: how do we induce contributors to reveal that their contributions are polluting? Of course, sometimes the two problems collapse into one: with some incentives, polluters may reveal themselves by the act of *not contributing*.

6.1.1. *Spam email*

Many authors have proposed designs to reduce or better manage spam email. These ideas provide a fertile example of the unavoidable necessity of addressing pollution as an *incentives* problem. Most of the ideas also apply, at least qualitatively, to other types of UCC pollution.

The most familiar designs for combating spam email are usually presented as technological, not incentive mechanisms. Such proposals include filtering (rule-based, Bayesian, and community or "collaborative"), disposable identities using extended email addresses [Bleichenbacher et al., 1998], DomainKeys Identified Mail [Perez, 2005], Sender ID or Sender Policy Framework [Crocker, 2006] [22], challenge-response [Dwork and Naor, 1993, Laurie and Clayton, 2004], whitelists, and blacklists. See Cranor and LaMacchia [1998] for an overview of many such ideas.

Although these systems generally do not involve extrinsic motivations (explicit exchanges of valuable resources), they fundamentally *are* incentive designs. The polluters (spammers) are, at root, humans who are polluting in order to achieve some objective (e.g., for many of them, advertising their products to attract purchasers). There is a cost to sending spam; the spammers receive a benefit. They will keep sending spam as long as the benefits exceed the costs. The designs mentioned earlier are technological, but they work to the extent that they raise the cost of spamming higher than the benefits: that is, they are providing (dis)incentives by raising costs. For example, authenticated sending mechanisms aim to raise cost by making it more difficult for spammers to

[22] As of now, spam-sending domains are ironically the biggest users of SPF tags [MXLogic, 2005]

hide their identity (which for spam is often the central piece of hidden information needed by the UCC system provider). Wash and MacKie-Mason [2007] have demonstrated that challenge-response technologies are incentive mechanisms (of the *screening* type). To succeed, all such technological mechanisms must get the balance of incentives correct. In the formal literature, the main sufficient conditions known for a screen to work (stated informally) are that (a) the cost of passing the screen must be sufficiently higher for polluters than for desirable participants, and (b) the rate at which the cost increases with the quantity of contribution must be sufficiently higher for polluters than for desirable participants (the Spence-Mirrlees condition).

Some recent work has addressed anti-spam design more directly as an incentive problem, suggesting both extrinsic and intrinsic motivators. An old and familiar idea in this genre is imposing some form of e-mail stamp, or more generally, a fee on content contributors in a UCC setting [e.g., Kraut et al., 2005, Hermalin and Katz, 2004]. Only contributors who value delivery sufficiently will pay the fee; if undesirable contributors don't value their contributions much, they won't contribute. This is a crude mechanism that will not work well if there is substantial heterogeneity in the value contributors place on providing their content, specifically if there is substantial (density-weighted) overlap in the value distribution supports of the two types of contributors.[23] When such proposals have been made for email, the resulting outcry has focused on the many low-wealth providers of desirable content (whether it be email from grandparents or from socially conscious nonprofit organizations), and the likelihood that commercials advertisers (spammers) will be willing to pay the fee. Yahoo! and AOL announced that they were going to introduce a type of stamp system (provided by Goodmail) in 2005. This system is not mandatory for all senders: rather, senders can elect to purchase a stamp. If they do, their email is guaranteed to pass through spam filters unscathed, and recipients see a certificate of authenticity attached to the mail. As of May 2007, Goodmail claimed over 300 senders were paying for its stamps, but it appears that most of them are commercial senders of unsolicited email: spammers [Goodmail, 2007].[24]

Loder et al. [2006] proposed a signaling device to reduce unwanted bulk email advertising (spam), which they called an "attention bond mechanism." Email senders are required to post a revocable bond or their content will not be delivered. Given a greater likelihood that a recipient will claim a spammer's bond, and appropriately selecting the size of the bond, the cost for a good sender to signal its belief that the recipient wants to see its mail will be sufficiently lower that they will be willing to comply and send email, whereas bad senders will be discouraged.

Chiao and MacKie-Mason [2006] addressed the email spam pollution problem from a somewhat different perspective. They show conditions under which

[23] Hermalin and Katz [2004] also consider payments by recipients, and a mixed system with payments by both senders and recipients.

[24] Goodmail markets its stamps to commercial senders who want to distinguish themselves as "high quality" by virtue of having bought a stamp, in contrast to low quality, but nonetheless some if not most of the mail they send using the stamps is unsolicited bulk commercial mail.

offering an alternative, even lower-cost (and higher-quality) platform for advertisers to distribute their spam might induce enough diversion of advertising to the alternative channel, simultaneously lowering the value of the regular email channel for other spammers, so that in equilibrium most or all spam will move to the advertising ("open") channel. The inducement to spammers is two-fold: in the open channel they can provide higher-quality information to potential customers (because there are no filters to evade), and as high-quality advertising leaves the regular email channel, it will be easier to filter the remaining low-quality spam, thus reducing the value (the "hit rate") of sending to the regular channel in the first place.

6.2. Manipulation

We already discussed one type of bad stuff: pollution. Pollution is a side-effect: a content contributor is trying to accomplish something (say, promote her products), with an indirect cost that she is cluttering a UCC resource (like a blog, or an inbox) with content outside the intended scope of that resource. The polluter doesn't gain from the side effect itself. In this section we turn to a different type of bad stuff: manipulation. Content intended to manipulate is the flip side of pollution: the contributor usually doesn't care about her contribution itself, but makes the contribution precisely to cause the harm.

Thus, in general terms, by manipulation in a UCC setting we mean contribution of content of the type for which the system is intended, but for the purpose of having a deceptive or misleading effect on the choices made by other users. Of course, based on the particular type of UCC, manipulation takes a different form. Here we discuss two main types: misleading textual content and numeric manipulations.

6.2.1. *Misleading textual content*

Many examples of manipulation have drawn attention in the past few years. Some of the most striking have been book reviews presented as the views of independent third parties, but in fact published by someone (under a pseudonym) with a direct financial interest with the intent of manipulating book sales. That such manipulation occurs has been known since at least 2004, when a software error on Amazon's Canadian site temporarily revealed the true authors of reviews [Dellarocas, 2006, Harmon, 2004]. Most reported cases have involved authors who wrote favorable reviews of their own books, but a recent case involved negative manipulation. Orlando Figes, a highly regarded historian at the University of London, recently blamed his wife for pseudonymously penning poisonous reviews of books by other historians on topics related to his (while also writing fatuous paeans to Figes's works). A week later he recanted, acknowledging that he in fact wrote the false reviews; later he agreed to pay monetary damages to two of the authors he attacked [Lea, 2010, Topping, 2010].

In another example, two founders of GiveWell admitted that they posted false-name recommendations steering people to their organization. This case is particularly piquant because GiveWell is a nonprofit to help people evaluate the quality (presumably, including reliability!) of nonprofit charitable organizations, and GiveWell itself is supported by charitable donations. One of the manipulations used a simple method: the executive posted a question on the Ask Metafilter service in which he sought recommendations for information about charities, then using a second pseudonym he posted a response recommending his own service [Strom, 2008a,b].

Sites that aggregate news stories are a natural target for manipulation. Digg.com is one of the more successful. Users post links to news stories found elsewhere on the web. Other users then vote on the interest or significance of the story (the "digg" it, or not), and those stories most highly rated are presented on the front page of the site (or the front page of several topical areas). In 2006 a story that Google was going to acquire Sun Microsystems was promoted to the front page. It is dubious whether the rumor was true (in any case, Google never made a public offer for Sun), and some investigators concluded that the story was a ruse, with a group collaborating to promote it in order to manipulate stock prices for a short-term trading gain [Gray, 2006]. Others have found evidence that a small group of frequent participants have regularly manipulated which stories are promoted to the front page, though no one has admitted publicly to doing so [Hayward, 2006].

Many observers express concern about the reliability of Wikipedia: students and others rely heavily on this encyclopedia as a first (and often only) source of facts, yet because it is a wiki, any user can change any content at any time, with no guarantee that the changes are correct. There have been several instances of intentional manipulation by users violating a rule against first-person editing (that is, editing entries about oneself). The most amusing were multiple revisions by Wikipedia's own founder, Jimmy Wales, of the entry about him, to reduce or eliminate discussion about his earlier business activities in online pornography and revise the recounting of the founding of Wikipedia (these incidents, of course, are now documented in Wikipedia) [Hansen, 2005, Wikipedia, 2010a]. The staff of several U.S. senators and representatives have been caught modifying the biographies of their employers [Davis, 2006].

6.2.2. *Numeric manipulations*

Manipulation can also occur in systems that collect, aggregate, and publish numeric ratings, such as those on Epinions, Netflix or Amazon's product recommendation system. Such systems can be manipulated by injecting dishonest ratings into the system. Google's search is a special kind of rating system (in which one user's Web links to the Web pages of another are the (indirectly) contributed ratings). It collects votes in a binary format which are susceptible to manipulations Altman and Tennenholtz [2005], Cheng and Friedman [2005]. The same is true for social tagging systems, e.g., Del.icio.us [Ramezani et al., 2008]. These types of manipulations are known as "shilling" or "sybil" attacks, because they use a large number of fake user identities, i.e., sybils, to provide misleading ratings.[25]

Recommender systems such as MovieLens or NetFlix are prone to sybil attacks as well [O'Mahony et al., 2002, Lam and Riedl, 2004]. An attacker first use the sybils to rate a number of movies to establish similarity in tastes with some users in the system. Then she can use all the sybils to rate one target movie highly, so as to influence the system to recommend this movie to all the users who share similar tastes with the sybils. Attacks using these profiles can increase the chance of a target item appearing on any user's recommendation list by 19 times [Lam and Riedl, 2004]. Solutions have been proposed to make recommender systems resistant to sybil attacks. For example, Chirita et al. [2005], O'Mahony et al. [2006], Mehta et al. [2007] demonstrate that statistical metrics can be quite effective in detecting, hence eliminating shills [26]. The problem of this approach though, is that the effects are temporary; that is, it leads to an arms-race. Resnick and Sami have developed recommender systems that are provably resistant to sybil attacks using an influence-limiter method [Resnick and Sami, 2007b, 2008].

6.2.3. *Reputation systems to reduce manipulation*

Reputation systems can be effective in reducing the influence of manipulators in at least two ways. One can use the reputation information of the content provider to filter out potential illegitimate content: a contributor with a low reputation score might not be trustworthy. Further, the negative effect of low reputation scores might discourage potential manipulation, as one has to build a sufficiently good reputation profile in order to exert influence.

A well-functioning reputation system requires persistent user identities. Some UCC systems either encourage or require real-world identity. For example, Citizendium (citizendium.org), a user-contributed online encyclopedia, requires and verifies that its voluntary editors use their real names. When comparing itself to Wikipedia it asserts that its real-name policy provides its articles "a kind of real-world credibility" [Citizendium, 2010]. Nonetheless, in many UCC systems it is relatively inexpensive to create new identities, which enables manipulators to whitewash their identities. Though widespread use of cheap pseudonyms undercut the usefulness of a reputation system, there may be an equilibrium in which reputation still has some value. Friedman and Resnick showed situations in which a user has to "pay her dues" before she can enjoy the benefits of participation on a UCC platform [Friedman and Resnick, 2001].

One natural way to limit manipulation is to only trust "word-of-mouth", that is, opinions of trusted acquaintances. A UCC system can allow each consumer to hand-pick the contributors that she trusts as friends, and then provide aggregate ratings that give higher weight to these friends. Such a circle of friends can also extend to friends' friends, forming a "web of trust" [Massa and Avesani, 2005,

[25] These attacks are named after the famous case of the woman known as Sybil, who suffered from multiple-personality disorder, [Schreiber, 1989].

[26] See Mobasher et al. [2007] for a literature survey of statistical solutions to defend recommender systems from shilling attacks.

Massa and Bhattacharjee, 2004, Massa and Avesani, 2005, Guha et al., 2004]. For example, Epinions.com uses a "web of trust" system to make recommender systems robust to malicious manipulation[27]. As long as the manipulators do not gain trust from any of the legitimate users, such a system can be effective in preventing manipulations.

Alternatively, rather than relying on user inputs to construct a web of trust, Resnick and Sami [2007a] propose generating a trust graph based on users' rating histories. For recommendee A, other users each have a reputation score indicating how much they have contributed to recommending good items to A in the past. These users' reputation scores are used to weigh their influence on the future recommendations to A. Since this reputation system takes into account the order in which ratings are submitted, it rewards raters contributing unique and timely opinions, which is something shills cannot do well. Such a reputation system can put an upper bound on how much influence shills can have.

Making rating, recommendation, and reputation systems manipulation-resistant is an active research area. A detailed discussion of the proposals made in various systems is certainly beyond the scope of this article. Here we only discussed a few approaches that apply across a general set of rating, recommendation, and reputation systems. We refer our readers to Friedman et al. [2007] for a detailed discussion.

7. SUMMARY

In this chapter we reviewed incentive-centered design for user-contributed content (UCC) systems or services on the Internet. UCC systems, produced (in part) through voluntary contributions made by nonemployees, face fundamental incentives problems. In particular, to succeed, users need to be motivated to contribute in the first place ("getting stuff in"). Further, given heterogeneity in content quality and variety, the degree of success will depend on incentives to contribute a desirable mix of quality and variety ("getting *good* stuff in"). Third, because UCC systems generally function as open-access publishing platforms, there is a need to prevent or reduce the amount of negative value (polluting or manipulating) content.

We have taken a somewhat multidisciplinary approach to the questions, though we draw mostly on the economics literature in asymmetric information and mechanism design. We point out some useful theories and findings from other areas, in particular social psychology, but we think there is much valuable work to be done to bring the economics of rational individual motivation and the psychology of social-regarding motivation together, particularly to work on problems such as UCC, which necessarily involves productive activity in social publics, an activity that often requires cooperation or collaboration.

Our review of the theoretical and empirical literatures bearing on the incentives problems facing UCC reveal an important finding: the work done to date

[27]See the Web of Trust FAQ `http://www.epinions.com/help/faq/?show=faq_wot`.

on these problems is limited, and uneven in coverage. Much of the empirical research concerns specific settings and does not provide readily generalizable results. And, although there are well-developed theoretical literatures on, for example, the private provision of public goods (the "getting stuff in" problem), this literature is only applicable to UCC in a limited way because it focuses on contributions of (homogeneous) money, and thus does not address the many problems associated with heterogeneous information content contributions (the "getting *good* stuff in" problem). We believe that our review of the literature has identified more open questions for research than it has pointed to known results.

User-contributed content is a fascinating, and clearly important phenomenon enabled by the development of the Internet and tools for user participation generally known as Web 2.0. How successful UCC ventures will be — whether as for-profit, or as non-profit organizations aiming to improve social welfare — will depend critically on the ability to solve the fundamental incentives challenges they face.

References

B. Thomas Adler and Luca de Alfaro. A content-driven reputation system for the wikipedia. In *WWW '07: Proceedings of the 16th International Conference on World Wde Web*, pages 261–270, New York, NY, USA, 2007.

Anat R. Admati and Motty Perry. Joint projects without commitment. *Review of Economic Studies*, 58:259–276, Jan 1991.

George A. Akerlof and Rachel E. Kranton. Economics and identity. *Quarterly Journal of Economics*, 115(3):715–753, Aug 2000.

George A. Akerlof and Rachel E. Kranton. Identity and schooling: Some lessons for the economics of education. *Journal of Economic Literature*, 40(4):1167–1201, 2002.

George A. Akerlof and Rachel E. Kranton. Identity and the economics of organizations. *Journal of Economic Perspective*, 19(1):9–32, 2005.

Alon Altman and Moshe Tennenholtz. Incentive comaptible ranking systems. In *The Proceedings of the 2007 International Conference on Autonomous Agents and Multiagent Systems*, Honolulu, Hawaii, USA, 2005.

Amazon.com. General review creation guidelines, 2010. URL `http://www.amazon.com/gp/help/customer/display.html?ie=UTF8&nodeId=14279631#reviewguidelines`.

James Andreoni. Warm-glow versus cold-prickle: The effects of positive and negative framing on cooperation in experiments. *Quarterly Journal of Economcis*, 110(1):1–21, 1995.

James Andreoni. Philanthropy. In Serge-Christophe Kolm and Jean Mercier Ythier, editors, *Handbook on the Economics of Giving, Reciprocity and Altruism*, volume 2, chapter 18, pages 1201–1269. Elsevier, 2006.

Parimal Kanti Bag and Eyal Winter. Simple subscription mechanisms for excludable public goods. *Journal of Economic Theory*, 87:72–94, 1999.

Mark Bagnoli and Barton L. Lipman. Provision of public goods: Fully implementing the core through private contributions. *Review of Economic Studies*, 56:583–601, Jan 1989.

Roland Benabou and Jean Tirole. Identity, dignity and taboos: Beliefs as assets. *Quarterly Journal of Economics*, forthcoming.

James Bennet. The Cinematch system: Operation, scale coverage, accuracy impact. Summer School on the Present and Future of Recommender Systems, 2006.

Ted Bergstrom, Larry Blume, and Hal Varian. On the private provision of public goods. *Journal of Public Economics*, 29:25–49, 1986.

Rajat Bhattacharjee and Ashish Goel. Algorithms and incentives for robust ranking. In *SODA '07: Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 425–433, Philadelphia, PA, USA, 2007.

D. Bleichenbacher, Eran Gabber, Markus Jakobsson, Yossi Matias, and Alain J. Mayer. Curbing junk e-mail via secure classification. In *FC '98: Proceedings of the Second International Conference on Financial Cryptography*, pages 198–213, London, UK, 1998.

Gary E. Bolton and Axel Ockenfels. ERC: A theory of equity, reciprocity, and competition. *The American Economic Review*, 90(1):166–193, March 2000.

John S. Breese, David Heckerman, and Carl Kadie. Empirical analysis of predictive algorithms for collaborative filtering. In *Proceedings of the Fourteenth Annual Conference on Uncertainty in Artificial Intelligence*, pages 43–52, 1998.

Susan L. Bryant, Andrea Forte, and Amy Bruckman. Becoming Wikipedian: transformation of participation in a collaborative online encyclopedia. In *GROUP '05: Proceedings of the 2005 International ACM SIGGROUP Conference on Supporting Group Work*, pages 1–10, New York, NY, USA, 2005.

D. Cartwright and A. Zander. Group cohesiveness: Introduction. In D. Cartwright and A. Zander, editors, *Group Dynamics: Research and Theory*. Row Peterson, Evanston, IL, 1968.

Gary Charness and Matthew Rabin. Understanding social preferences with simple tests. *Quarterly Journal of Economics*, 117(3):817–869, 2002.

Jianqing Chen, Hong Xu, and Andrew B. Whinston. Moderated online communities. Working paper, 2007.

Mao Chen and Jaswinder Pal Singh. Computing and using reputations for internet ratings. In *EC '01: Proceedings of the 3rd ACM conference on Electronic Commerce*, pages 154–162, New York, NY, USA, 2001.

Yan Chen and Sherry Xin Li. Group identity and social preferences. *American Economic Review*, 99(1):431–457, March 2009.

Yan Chen, Max Harper, Joseph Konstan, and Sherry Xin Li. Social comparisons and contributions to online communities: A field experiment on movielens. *American Economic Review*, forthcoming.

Alice Cheng and Eric Friedman. Sybilproof reputation mechanisms. In *P2PECON '05: Proceedings of the 2005 ACM SIGCOMM Workshop on Economics of Peer-to-Peer Systems*, pages 128–132, New York, NY, USA, 2005.

Benjamin Chiao and Jeffrey MacKie-Mason. Using uncensored communication channels to divert spam traffic. In *Proceedings of the 34th Research Conference on Communication, Information and Internet Policy (TPRC)*, 2006.

Paul-Alexandru Chirita, Wolfgang Nejdl, and Cristian Zamfir. Preventing shilling attacks in online recommender systems. In *WIDM '05: Proceedings of the 7th annual ACM International Workshop on Web Information and Data Management*, pages 67–74, New York, NY, USA, 2005.

Robert B. Cialdini and Noah J. Goldstein. Social influence: Compliance and conformity. *Annual Review of Psychology*, 55:591–621, 2004.

Citizendium. Cz:why citizendium?, 2010. URL `http://en.citizendium.org/wiki/CZ:Why_Citizendium%3F`.

Ronald Coase. The problem of social cost. *Journal of Law and Economics*, 3 (1):1–44, 1960.

Lorrie Faith Cranor and Brian A. LaMacchia. Spam! *Communications of the ACM*, 41(8):74–83, 1998.

Dave Crocker. Challenges in anti-spam efforts. *The Internet Protocol Journal*, 8(4), 2006.

Rachel Croson, Enrique Fatás, and Tibor Neugebauer. The effect of excludability on team production. Working paper, 2008.

Matthew Davis. Congress 'made wikipedia changes', February 9 2006. Online; accessed 25-October-2010.

James A. Dearden. Efficiency and exclusion in collective action allocations. *Mathematical Social Science*, 34:153–174, 1997.

Kay Deaux. Social identification. In E. Tory Higgins and Arie W. Kruglanski, editors, *Social Psychology: Handbook of Basic Principles*. The Guilford Press, New York, 1996.

Rajat Deb and Laura Razzolini. Auction-like mechanisms for pricing excludable public goods. *Journal of Economic Theory*, 88(2):340 – 368, 1999.

Chrysanthos Dellarocas. Strategic manipulation of internet opinion forums: Implications for consumers and firms. *Management Science*, 52(10):1577–1593, 2006.

Cynthia Dwork and Moni Naor. Pricing via processing or combatting junk mail. In *CRYPTO '92: Proceedings of the 12th Annual International Cryptology Conference on Advances in Cryptology*, Lecture Notes in Computer Science, pages 139–147, London, UK, 1993.

Encyclopedia of Life. Help build EOL, 2010. URL `http://www.eol.org/content/page/help_build_eol#curators`.

Ernst Fehr and Urs Fischbacher. The nature of human altruism. *Nature*, 425 (6960):785–791, October 2003.

Ernst Fehr and Klaus M. Schmidt. A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, 114(3):817–868, 1999.

Eric Friedman and Paul Resnick. The social cost of cheap pseudonyms. *Journal of Economics and Management Strategy*, 10(2):173–199, 2001.

Eric Friedman, Paul Resnick, and Rahul Sami. *Manipulation-Resistant Reputation Systems*, pages 677–698. Cambridge University Press, New York, NY, 2007.

G. W. Furnas, T. K. Landauer, L. M. Gomez, and S. T. Dumais. The vocabulary problem in human-system communication. *Communications of the ACM*, 30 (11):964–971, 1987.

Goodmail. Goodmail systems launches international expansion, 17 May 2007. URL `http://tinyurl.com/3y35ut`.

Michael (Graywolf) Gray. Digg, sun, and google the tale of market manipulation. *Threadwatch.org*, March 17 2006. URL `http://www.threadwatch.org/node/5921`.

Theodore Groves and John Ledyard. Optimal allocation of public goods: A solution to the "free rider" problem. *Econometrica*, 45(4):783–809, May 1977.

Ratan Guha, Ravi Kumar, Prabhakar Raghavan, and Andrew Tomkins. Propagation of trust and distrust. In *WWW '04: Proceedings of the 13th international conference on World Wide Web*, pages 403–412, New York, NY, USA, 2004.

Evan Hansen. Wikipedia founder edits own bio. *Wired*, December 19 2005. URL http://www.wired.com/culture/lifestyle/news/2005/12/69880.

Amy Harmon. Amazon glitch unmasks war of reviewers. *New York Times*, February 14 2004. URL http://www.nytimes.com/2004/02/14/us/amazon-glitch-unmasks-war-of-reviewers.html?pagewanted=all&src=pm.

Maxwell Harper, Xin Li, Yan Chen, and Joseph Konstan. An economic model of user rating in an online recommender system. In *UM 2005 User Modeling: Proceedings of the Tenth International Conference*, Lecture Notes in Artificial Intelligence. Springer, 2005.

Alexander Hars. Working for free? motivations for participating in open-source projects. *International Journal of Electronic Commerce*, 6(3):25–39, Spring 2002.

Taylor Hayward. How the digg system is being gamed by a small number of users, September 2006. URL http://taylorhayward.org/digg_gaming.html.

Cormac Herley. So long, and no thanks for the externalities: the rational rejection of security advice by users. In *NSPW '09: Proceedings of the 2009 Workshop on New Security Paradigms*, pages 133–144, New York, NY, USA, 2009. ACM.

Ben Hermalin and Michael Katz. Sender or receiver: Who should pay to exchange an electronic message? *RAND Journal of Economics*, 35(3):423–447, 2004.

Michael A. Hogg. Social categorization, depersonalization and group behavior. In Michael A. Hogg and Scott Tindale, editors, *Blackwell Handbook of Social Psychology: Group Processes*. Blackwell Publisher Ltd., Malden, MA, 2003.

Lian Jian. *Three essays on the economics of information systems*. PhD thesis, University of Michigan, 2010. Chapter 4.

Steven J. Karau and Kipling D. Williams. Social loafing: A meta-analytic review and theoretical integration. *Journal of Personality and Social Psychology*, pages 681–706, 1993.

Daria Knoch, Alvaro Pascual-Leone, Kaspar Meyer, Valerie Treyer, and Ernst Fehr. Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science*, Jan 2006.

Martin Kocher, Matthias Sutter, and Verena Waldner. Exclusion from public goods as an incentive system–an experimental examination of different institutions. Working paper, 2005.

Robert E. Kraut, Shyam Sunder, Rahul Telang, and James Morris. Pricing electronic mail to solve the problem of spam. *Human-Computer Interaction*, 20(1-2):195–223, 2005.

Jean-Jacques Laffont and David Martimort. *The Theory of Incentives*. Princeton University Press, 2002.

Karim R. Lakhani and Eric von Hippel. How open source software works: "free" user-to-user assistance. *Research Policy*, 32(6):923–943, June 2003.

Shyong K. Lam and John Riedl. Shilling recommender systems for fun and profit. In *WWW '04: Proceedings of the 13th International Conference on World Wide Web*, pages 393–402, New York, NY, USA, 2004.

Cliff Lampe and Paul Resnick. Slash(dot) and burn: distributed moderation in a large online conversation space. In *CHI '04: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 543–550, New York, NY, USA, 2004.

Kelvin Lancaster. A new approach to consumer theory. *The Journal of Political Economy*, 74(2):132–157, 1966.

Ben Laurie and Richard Clayton. Proof of work proves not to work. In *Workshop on the Economics of Information Security*, 2004.

Richard Lea. Poison pen reviews were mine, confesses historian orlando figes. *The Guardian*, April 23 2010.

Joshua Lerner and Jean Tirole. Some simple economics of open source. *Journal of Industrial Economics*, 50(2), June 2002.

Greg Linden, Brent Smith, and Jeremy York. Amazon.com recommendations: item-to-item collaborative filtering. *IEEE Internet Computing*, 7(1):76–80, Jan/Feb 2003.

Kimberly Ling, Gerard Beenen, Pamela Ludford, Xiaoqing Wang, Klarissa Chang, Xin Li, Dan Cosley, Dan Frankowski, Loren Terveen, Al Mamunur Rashid, Paul Resnick, and Robert Kraut. Using social psychology to motivate contributions to online communities. *Journal of Computer-Mediated Communication*, 10(4), 2005.

Thede Loder, Marshall van Alstyne, and Rick Wash. An economic response to unsolicited communication. *Advances in Economic Analysis and Policy*, 6(1), 2006.

Jeffrey K. MacKie-Mason. *Economic Incentives*. Marshall Cavendish, Tarrytown, NY, 2000.

Jeffrey K. MacKie-Mason. Incentive-centered design for security. *IEEE Security and Privacy*, pages 72–75, July/August 2009.

Jeffrey K. MacKie-Mason. Tom Sawyer production on the Internet: Getting the good stuff in, keeping the bad stuff out. October 2010. URL `http://hdl.handle.net/2027.42/78183`.

Cameron Marlow, Mor Naaman, Danah Boyd, and Marc Davis. HT06, tagging paper, taxonomy, Flickr, academic article, to read. In *HYPERTEXT '06: Proceedings of the Seventeenth Conference on Hypertext and Hypermedia*, pages 31–40, New York, NY, USA, 2006.

Paolo Massa and Paolo Avesani. Controversial users demand local trust metrics: An experimental study on epinions.com community. In *Proceedsings of the 25th American Association for Artificial Intelligence Conference*, pages 121–126, 2005.

Paolo Massa and Bobby Bhattacharjee. Using trust in recommender systems: An experimental analysis. In *Proceedings of the 2nd International Conference on Trust Management*, 2004.

Rose McDermott. Psychological approaches to identity: Definitions, measurement and experimentation. In R. Abdelal, Y. Herrera, A. I. Johnson, and Rose McDermott, editors, *Measuring Identity: A Guide for Social Science Research*. Cambridge University Press, Cambridge, U.K., 2009.

Bhaskar Mehta, Thomas Hofmann, and Peter Fankhauser. Lies and propaganda: detecting spam users in collaborative filtering. In *IUI '07: Proceedings of the 12th International Conference on Intelligent User Interfaces*, pages 14–21, New York, NY, USA, 2007.

N. Michinov, E. Michinov, and M. C. Toczek-Capelle. Social identity, group processes, and performance in synchronous computer-mediated communication. *Group Dynamics: Theory Research and Practice*, 8(1):27–39, 2004.

Nolan Miller, Paul Resnick, and Richard Zeckhauser. Eliciting honest feedback: The peer prediction method. *Management Science*, 51(9):1359–1373, September 2005.

Bamshad Mobasher, Robin Burke, Runa Bhaumik, and Chad Williams. Toward trustworthy recommender systems: An analysis of attack models and algorithm robustness. *ACM Trans. Inter. Tech.*, 7(4):23, 2007.

Benny Moldovanu. The production and cost-sharing of an excludable public good. *Economic Theory*, 7:531–539, 1996.

John Morgan. Financing public goods by means of lotteries. *Review of Economic Studies*, 67(4), 2000.

Herve Moulin. Serial cost-sharing of excludable public goods. *Review of Economic Studies*, 61(2):305–325, April 1994.

Michael Mussa and Sherwin Rosen. Monopoly and product quality. *Journal of Economic Theory*, 18(2):301–317, August 1978.

MXLogic. Mxlogic reports spam accounts for 67 percent of all email in 2005. Press Release, 22 September 2005.

Oded Nov. What motivates wikipedians. *Communications of the ACM*, 50(11): 60–64, November 2007.

Michael P. O'Mahony, Neil J. Hurley, and Guenole C.M. Silvestre. Promoting recommendations: An attack on collaborative filtering. In *Proceedings of the 13th International Conference on Database and Expert System Applications*, pages 494–503, 2002.

Michael P. O'Mahony, Neil J. Hurley, and Guenole C.M. Silvestre. Detecting noise in recommender system databases. In *Proceedings of the 2006 International Conference on Intelligent User Interfaces*, pages 109–115, 2006.

Juan Carlos Perez. Yahoo and Cisco to submit e-mail ID spec to IETF. *NetworkWorld*, 11 July 2005.

T. Postmes, R. Spears, K. sakhel, and D. de Groot. Social influence in computer-mediated communication: The effects of anonymity on group behavior. *Personality and Social Psychology Bulletin*, 27(10):1243–1254, 2001.

T. Postmes, R. Spears, and M. Lea. Intergroup differentiation in computer-mediated communication: Effects of depersonalization. *Group Dynamics: Theory Research and Practice*, 6(1):3–16, 2002.

Emilee Rader and Rick Wash. Tagging with del.icio.us: Social or selfish? In *CSCW '06: Conference on Computer Supported Cooperative Work*, Banff, Alberta, CA, Nov 2006.

Maryam Ramezani, Runa Bhaumik J. J. Sandvig, Tom Schimoler, Robin Burke, and Bamshad Mobasher. Exploring the impact of profile injection attacks in social tagging systems. In *10th Kdd Workshop on Web Mining and Web Usage Analysis*, Las Vegas, Nevada, August 2008.

Al M. Rashid, Kimberly Ling, Regina D. Tassone, Paul Resnick, Robert Kraut, and John Riedl. Motivating participation by displaying the value of contribution. In *CHI '06: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 955–958, New York, NY, USA, 2006.

Eric S. Raymond. *The Cathedral and the Bazaar: Musings on Linux and Open Source by an Accidental Revolutionary (O'Reilly Linux)*. O'Reilly, October 1999.

Yuqing Ren, Robert Kraut, and Sara Kiesler. Applying common identity and bond theory to design of online communities. *Organization Studies*, Jan 2007.

Paul Resnick and Rahul Sami. Making recommender systems provably manipulation-resistant through influence limits. In *Recommender Systems 2007*, Minneapolis, Minnesota, USA, October 2007a.

Paul Resnick and Rahul Sami. The influence-limiter: Provably manipulation-resistant recommender systems. In *Recommender Systems*, Minneapolis, MN, October 2007b. ACM.

Paul Resnick and Rahul Sami. The information cost of manipulation resistance in recommender systems. In *ACM Electronic Commerce (EC '08)*, Chicago, IL, July 2008. ACM.

P. Rogers and M. Lea. Social presence in distributed group environments: The role of social identity. *Behaviour and Information Technology*, 24(2):151–158, 2005.

Michael Rothschild and Joseph E Stiglitz. Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. *The Quarterly Journal of Economics*, 90(4):630–49, November 1976.

Richard Ryan and Edward L. Deci. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1):68–78, 2000.

Paul A Samuelson. The pure theory of public expenditure. *Review of Economics and Statistics*, 36(4):387–389, November 1954.

Todd Sandler and John T. Tschirhart. The economic theory of clubs: An evaluative survey. *Journal of Economic Literature*, 18(4):1481–1521, Dec 1980.

Flora Rheta Schreiber. *Sybil: The true and extraordinary story of a woman possessed by sixteen separate personalities*. Grand Central Publishing, New York, 1989.

Suzanne Scotchmer. Clubs. In L. Blume and S. Durlauf, editors, *New Palgrave Dictionary of Economics*. MacMillan, 2008.

M. Sherif, L. J. Harvey, B. J. White, W. R. Hood, and C. W. Sherif. *Intergroup conflict and cooperation: The robbers' cave experiment*. Wesleyan University Press, Middletown, CT, 1961. Reprinted in 1988.

Michael Spence. Job market signaling. *The Quarterly Journal of Economics*, 87(3):355–374, August 1973.

David A. Starrett. *Foundations of Public Economics*. Cambridge University Press, New York, 1988.

Brian Stelter. Youtube videos pull in real money. *New York Times*, December 10 2008. URL `http://www.nytimes.com/2008/12/11/business/media/11youtube.html`.

Stephanie Strom. Founder of a nonprofit is punished by its board for engaging in an internet ruse. *New York Times*, January 8 2008a. URL http://www.nytimes.com/2008/01/08/us/08givewell.html?scp=3&sq=givewell&st=cse.

Stephanie Strom. Nonprofit punishes a 2nd founder for ruse. *New York Times*, January 15 2008b. URL http://www.nytimes.com/2008/01/15/us/15givewell.html?scp=5&sq=givewell&st=cse.

Kurtis J. Swope. An experimental investigation of excludable public goods. *Experimental Economics*, 5:209–222, 2002.

Henri Tajfel and John Turner. An integrative theory of intergroup conflict. In Stephen Worchel and William Austin, editors, *The Social Psychology of Intergroup Relations*. Brooks/Cole, Monterey, CA, 1979.

Henri Tajfel and John Turner. The social identity theory of intergroup behavior. In Stephen Worchel and William Austin, editors, *The Social Psychology of Intergroup Relations*. Nelson- Hall, Chicago, 1986.

Alexandra Topping. Historian Orlando Figes agrees to pay damages for fake reviews. *The Guardian*, July 16 2010.

Mark Twain. *The Adventures of Tom Sawyer*. Harper and Brothers, New York, 1903 author's national edition, 1876.

Luis von Ahn and Laura Dabbish. Labeling images with a computer game. In *CHI '04: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2004.

Mark Walker. A simple incentive compatible scheme for attaining lindahl allocations. *Econometrica*, Jan 1981.

Rick Wash and Jeffrey K. MacKie-Mason. Security when people matter: Structuring incentives for user behavior. In *International Conference on Electronic Commerce (ICEC'07)*, Minneapolis, MN, 19-22 August 2007.

Rick Wash and Jeffrey K. MacKie-Mason. A social mechanism for home computer security. In *The Workshop on Information Systems and Economics (WISE)*, Dec 2008.

Rick Wash and Emilee Rader. Public bookmarks and private benefits: An analysis of incentives in social computing. In *Proceedings of the American Society for Information Science and Technology*, 2007.

Wikipedia. Jimmy wales — Wikipedia, the free encyclopedia, 2010a. URL http://en.wikipedia.org/wiki/Jimmy_wales#Controversy. Online; accessed 25-October-2010.

Wikipedia. Wikipedia:Barnstars, 2010b. URL http://en.wikipedia.org/wiki/Wikipedia:Barnstars.

David E. Wildasin. *Urban Public Finance*. Harwood, Chur, Switzerland, 1986.

Joanne V. Wood. Theory and research concerning social comparisons of personal attributes. *Psychological Bulletin*, 106:231–248, 1989.

S. Worchel, H. Rothgerber, E. A. Day, D. Hart, and J. Butemeyer. Social identity and individual productivity within groups. *British Journal of Social Psychology*, 37:389–413, 1998.

Jiang Yang, Lada Adamic, and Mark Ackerman. Competing to share expertise: the taskcn knowledge sharing community. In *International Conference on Weblogs and Social Media*, 2008.

H. Peyton Young. Cost allocation, demand revelation, and core implementation. *Mathematical Social Sciences*, 36:213–228, 1998.