Tom Sawyer Production on the Internet: Getting the Good Stuff In, Keeping the Bad Stuff Out*

Jeffrey MacKie-Mason†

DRAFT: PRELIMINARY AND INCOMPLETE

April 8, 2009

Abstract

User-contributed content as an input to the production of information services is not new, but it is growing rapidly in significance and prevalence. Open-source software, Wikipedia, and Flickr are but a few examples of the variety of information products and services relying on user-contributed content. I propose a characterization of user-contributed content, and identify contributor behavior issues critical for success. From the perspective of an information service provider, or the economy as a whole, these issues predict underprovision of content, inefficient mixes of quality and variety, and undesirable levels of content pollution. How might we design information services or systems to ameliorate these problems? Given the centrality of autonomous,

*This paper grew out of an invited plenary presentation at the Southern Economic Association Annual Meeting in November 2006; I am grateful to Barry Hirsch and the SEA Board for that opportunity. I appreciate comments from Michael Wellman, Paul Resnick, Tilman Börgers, Yan Chen, Dan Grosu and Brian Noble, my doctoral students in the Incentive-Centered Design Lab, Andrea McVittie for her quick and careful research assistance, and especially Lian Jian for her comments, insights and collaboration in the larger project of which this is a part. I gratefully acknowledge financial support from NSF awards CNS-0716196 and DGE-0654014.

†Dept. of Economics, and School of Information, Univ. of Michigan, Ann Arbor, MI 48109, `jmm` *at* `umich` *dot* `edu`

motivated human behavior in user-contributed content problems, I argue this is a problem for *incentive-centered design*: how to configure economic, social and psychological incentives to induce contribution, discourage pollution, and motivate sufficient effort to generate quality? To illustrate, for a content pollution problem loosely based on a popular Web site's experience, I offer a stylized mechanism that relies on user-contributed (meta)content to screen out polluting contributions.

> At this dark and hopeless moment an inspiration burst upon him! Nothing less than a great, magnificent inspiration .... There was no lack of material; boys happened along every little while; they came to jeer, but remained to whitewash ....And when the middle of the afternoon came, from being a poor poverty-stricken boy in the morning, Tom was literally rolling in wealth.
>
> Mark Twain, The Adventures of Tom Sawyer

1 Let's whitewash the fence!

We are seeing the rapid growth of an unusual form of information production 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; I favor *user-contributed content* (the content creators are often users of the resulting resource).[1]

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

Many information producers now significantly (though rarely exclusively) rely on the user-contributed production model. Many are successful and quite socially valuable. Well-known examples include (the user contribution is highlighted in parentheses) Google (`www.google.com`, links to Web resources are the data for the PageRank algorithm); Wikipedia (`www.wikipedia.org`, encyclopedia content); Amazon (`www.amazon.com`, product reviews); Flickr (`www.flickr.com`, photos); Digg (`www.digg.com`, editorial selection of news stories); del.icio.us (`del.icio.us`, recommended Web sites); CiteULike and Bibsonomy (`www.citeulike.org` and `www.bibsonomy.org`, scholarly citations), and, of course open-source software projects such as Linux and the Apache web server. Less familiar but important examples include the new Peer2Patent (`http://www.peertopatent.org`, patent commenting system), the Stanford Encyclopedia of Philosophy (`http://plato.stanford.edu/`), and the WorldCat (`http://www.worldcat.org/`, world-wide library cataloguing). Of course, far more projects do not generate revenues greater than their costs (even with so much of the production cost donated by volunteers) and thus are short-lived.

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 entirely new. Tom Sawyer's friends and neighbors whitewashed Aunt Polly's fence for him.[2] Public radio is produced with resources donated by listeners. User-contributed content is clearly related to the *private provision of public goods* [Bergstrom et al., 1986]. The current burst of user-contributed content on the Internet exhibits some interesting features, however. For example, most prior examples involved monetary donations, not raw material or finished production (information content). Likewise, prior examples were typically produced or managed by a non-profit or government agency; much, perhaps most 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 model is quite apt: in-kind donations to a private producer, with a significant concern for quality.

[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].

> She found the entire fence whitewashed, *and not only white-washed but elaborately coated and recoated*
>
> Mark Twain, The Adventures of Tom Sawyer (emphasis added)

Evidently, the success of a content provider who relies on user-contributed content requires solving at least two problems: *getting the good stuff in*, and *keeping the bad stuff out*.

2 What is user-contributed content?

To understand the economic issues invoked by the rise of *user-contributed content* on the Web, we need to define the concept. I use: information content as a production input, provided at least in part by people without *direct extrinsic* compensation (as would be conventional employees or commercial suppliers). These input providers are often, but need not be, consumers of the final product, thus the sometimes appropriate label "user-contributed". By *extrinsic* I mean measurable, transferable compensation, for example money or an in-kind payment (barter). By *direct* I mean a more or less contractual specification of the relationship between the amount of input provided and the compensation. A typical example of direct compensation would be payment of a fixed fee per 1000 words of content (such as might be the compensation to a free-lance writer).

Why do we need such a detailed definition of the type of compensation? In most economic articles we might just write "without compensation". I qualify compensation with "extrinsic" because we cannot rule out intrinsic compensation. If we limited ourselves to contributions made without *any* compensation (in the sense of personal gain received), we would only consider contributions made accidentally or out of pure altruism: this is surely too great a restriction, ruling out much (perhaps most) user-contributed content. I allow for the possibility that contributors experience some form of intrinsic compensation, which might include "warm glow" from (impure) altruism, or something else (such as social respect or acquisition of a professional reputation).

In addition, and especially important in some settings, we want to consider cases in which contributors may receive *indirect* yet extrinsic compensation: some benefit from the final product that is not contractually (even informally) tied to their contribution but that they believe does depend in some informal manner on their contribution. For example, contributors may believe that they have a marginal impact on the survival or quality of a product which they value as a consumer. One often observes contributors describing themselves as members of a community and suggesting that contribution by themselves (or all members of the community) is essential if the community is to survive. Such statements suggest an equilibrium based on *generalized reciprocity*.[3]

Another common type of indirect compensation occurs when content contribution plays (at least) two productive functions: one as an input to the producer's effort, one as an input to the contributor's utility function. That is, the contribution, while costly for the contributor to make, may also provide some personal benefits. This is the case, for example, with certain types of shared indexing services, such as `del.icio.us` (a cross-referenced index of user-contributed Web bookmarks). When a user posts bookmarks to `del.icio.us`, she is contributing to that firm's product, but she is also storing her bookmarks in a database that she can access from any network-connected device with a browser, which many users value as an alternative to a bookmark database that is locally stored on a single device [Rader and Wash, 2006]. This contributor is extrinsically (in-kind) but not directly compensated by `del.icio.us`; rather, she benefits indirectly through the personal use she can make of her own stored bookmarks.

Thus, to admit a wide range of interesting examples, and focus on what distinguishes them from other arrangements for input supply, I shall limit *user-contributed content* to those inputs that would be contributed even if the only compensation is either not extrinsic or not direct (or neither).

We need not rule out the possibility that a producer would choose to provide direct extrinsic compensation, but I am concerned primarily with applications in which the producer usually chooses not to do so. For example, the Encyclopedia Britannica pays its authors, but it is not necessary to do so in order to create a high-quality encyclopedia: Wikipedia does not. Both producers need content as an input; Wikipedia could have chosen to pay for

[3]I do not here explore whether there is a rational basis for a contributor's perception that she receives indirect benefits, but see, for example, [Jian and MacKie-Mason, 2006].

content. Evidently, Wikipedia at least implicitly made an economic decision to rely on user-contributed content rather than paying, but I will not be concerned with that decision in this article.[4] Rather, I will consider producers who for some unexamined reason choose to rely on user-contributed content.

Before proceeding, let me note that I focus on *information content* ("content" for short) as the contributed input because there is an interesting and growing set of applications specific to information content. In other settings "users" could contribute other inputs to production (e.g., raw natural resource materials, semi-finished goods, etc.), suggesting the possibility of a more general theory of *contributed factors of production*, but I do not address that possibility here. There are two other common types of contribution — monetary and labor — that I shall briefly discuss to focus attention on the distinct characteristics of contributed content.

Monetary contributions are a familiar phenomenon, usually referred to as charitable giving. Some have much in common with user-contributed content applications. For instance, donations to a public radio station result in an open-access information product, and contributors benefit indirectly from the consumption of the final product. At least two economic questions relevant for user-contributed content are studied for monetary contributions: why contribute, and how much? I will refer to this literature below. One typical distinction, however, is that monetary contributions usually are to non-profit organizations, whereas information content is often contributed to profit-seeking entities. This may be significant because of the role of intrinsic compensation to motivate contribution. Another important distinction is that money is (largely) homogeneous, whereas content contributions usually are heterogeneous in their quality.

Labor contributions are also familiar. Like money, they are often observed in charitable contexts. For example, many people with valuable skills donated time and energy to assist victims of Hurricane Katrina with medical, shelter and other needs. Thus, as with money, labor contributions to non-profit activities are distinguished from content contributions to for-profit activities. There is another way in which labor and content contributions are

[4]There might be a variety of economic rationales for such a decision. For example, contracting, accounting and payment costs may be too large if a large number of individuals are contributing small amounts of content. There is some evidence (and considerable popular conjecture) that providing direct extrinsic compensation for certain efforts affects the motivations of the provider and may have an adverse impact on the quantity or quality provided.

quite similar however: they both can be heterogeneous in quality (for labor, how skilled is the worker? how hard does she work?). Indeed, in one sense it makes sense to think of content contributions as labor contributions: much or most of the resources used to create the contributed content may be the labor it took to create, find, or otherwise produce the content. Nonetheless, that the labor was used to produce *information* distinguishes it from labor supplied to other production functions because of at least two peculiar characteristics exhibited by information content. First, it is *nonrivalrous*. Once created, information generally can be used by many without diminishing its value to any. Second, it is approximately true for most information content that it can be costlessly reproduced after the original instance has been created. When contributed labor is used to construct a house by Habitat for Humanity, on the other hand, the resulting house can generally be used by only one (appropriately-sized) family, and a second house will require approximately as many resources to create as did the first. Thus, user-contributed information content typically is distinguished from most labor contributions when the information is contributed to for-profit entities, and when the information is nonrivalrous or (approximately) costless to reproduce.

3 A stroll through Web 2.0

Tim O'Reilly coined the expression "Web 2.0" in 2003 to refer not to a new version of Internet protocols, but to a new style of application development and a new mode of user interaction on the Web [O'Reilly, 2005]. Web 2.0 is not well-defined, but one common use refers to the family of emerging information services that rely heavily on user-contributed content.

Reliance on user-contributed content is common, but also diverse: it is the foundation for a wide variety of firms.[5] The "life-cycle of information" (see Figure 1) is a common characterization of the many different types of information producers. Some create information (novelists, rock bands, scholars); others publish and distribute (Random House, MGM); still others maintain inventories and provide access to end users (libraries, video stores), and so forth. User-contributed content supports businesses operating in each

[5]I use "firm" loosely to refer to any non-governmental organization producing a product or service for use or consumption by others. Such firms may be for-profit or not-for-profit. The form of organization may be relevant to some decisions about user-contributed content as a production input, but I do not address it in this article.

of these different markets. See Table 1) for illustrative examples.

Well-known user-contributed content successes such as Wikipedia, Linux and Flickr demonstrate breakthrough contributions to social welfare. Lesser-known services suggest that this form of production may have a wide and deep impact on knowledge production and use. For example, most scholars now rely heavily on at least some forms of online collaboration [Finholt, 2003]. arXiv.org is an early and successful example, started as a pre-print server at Los Alamos National Laboratories, and now hosted by Cornell. This archive served initially the high-energy physics community, and now hosts over 450,000 e-prints in physics, mathematics, computer science, computational biology and statistics. The Social Sciences Research Network (SSRN) provides a similar service to economics, law, accounting and related areas, with over 130,000 e-prints.

Other types of scholarly production services are emerging. For example, CiteULike, Bibsonomy, and Connotea are social bibliographic services to which scholars upload and share their bibliographic databases. To illustrate, I uploaded some of my bibliographic data files in BibTeX format to BibSonomy (see Figure 2). I can search my own citation database or those of others by user, popularity of the cite, or user-supplied tags (see Figure 3). I can then select and download cites in (currently) more than 15 formats (see Figure 4). Crucial determinants of the value of this service, and thus its long-term success, are the quantity of references provided by the users (so that I can easily find citations I need), their quality (whether they are properly formatted and accurate), and the effort users make to provide useful tags for searching.

Another emerging area with tremendous potential is in the provision of user-contributed metadata. With the flood of information now accessible, there is a crucial need for information about the information (metadata) that facilitates organizing, finding, retrieving, annotating (and other functions that comprise an idealized "semantic web" Shadbolt et al. [2006]). The Library of Congress without a classification system would be largely useless; efforts to impose a hierarchical, controlled vocabulary metadata schema on Internet-accessible content have proved futile. For example, few web page creators provide the simple Dublin Core [Weibel and Koch, 2000], PICS [Lessig and Resnick, 1999] or P3P [Cranor, Nov.-Dec. 2003] metadata necessary for standardized and widely-respected efforts to create controlled vocabularies. Uncontrolled vocabulary efforts (folksonomies) such as tagging on information sharing sites (e.g., del.icio.us, Flickr, LibraryThing) may support the

Table 1: UCC throughout the information life-cycle industries

	URL	Product
Authoring		
Wikipedia	en.wikipedia.org	encyclopedia
Wikibooks	en.wikibooks.org	textbooks
Stanford Encycl. of Philosophy	plato.stanford.edu	encyclopedia
Linux	www.linux.org	operating system
Publishing		
Public Library of Science	www.plos.org	Scholarly research
Selecting		
Digg	digg.com	news
Cloudmark	www.cloudmark.com	spam filtering
YouTube	www.youtube.com	video finding
Indexing		
CiteULike	www.citeulike.org	citation records
del.icio.us	del.icio.us	URL bookmarking
dmoz	www.dmoz.org	Web directory
Storing		
Flickr	www.flickr.com	photos
Gmail	mail.google.com	email, files
Xdrive	www.xdrive.com	files
WalkerTracker	walkertracker.com	pedometer records
Finding		
Amazon	www.amazon.com	books, merchandise
iTunes	www.apple.com/itunes	music finding, distribution
LinkedIn	www.linkedin.com	business contacts
Using		
Swivel	www.swivel.com	share, use datasets
RateMyProfessor	www.ratemyprofessors.com	aggregate evaluations
MediaPredict	mediapredict.com	prediction market

semantic web but only with effective incentives for users to contribute, and for the community to develop a semantically-workable vocabulary [Rader and Wash, 2006].

4 What are the economic issues?

A firm that wishes to obtain user-contributed content as a production input faces at least three fundamental economic problems: obtaining *contributions*, and managing their *quality*, and handling contributions of inappropriate or harmful content (such as unsolicited advertising, the most common type of "spam"). The first two I refer to as "getting the good stuff in": since the firm seeks content in exchange for little or no monetary compensation, it is not assured of obtaining enough (or any) input. The third problem I refer to as "keeping the bad stuff out".[6]

4.1 Quantity

The first, and perhaps most obvious economic challenge, is inducing people to contribute their content creation efforts with no direct or extrinsic compensation. However, having recognized that user-contributed content is just another input factor for production, why is this a special challenge? Why is this not simply the usual task of negotiating a price with an input supplier? Every producer would like to get a zero price (holding quality constant).

We might want to give special attention to the problem of inducing content contribution with zero payment if socially valuable products or services will become available only when the content is contributed. For example,

[6]Taxonomies are never perfect, and this one is no exception. It is reasonable to argue that there are really only two types of problems: obtaining contributions, and managing their quality. If we think simplistically of quality falling along a gradient from high positive values to low negative values, inappropriate or harmful content is simply lower down the gradient, and need not be seen as a separate problem. However, much of the existing literature in economics tends to treat the problem of markets for a gradient of positive-valued qualities (e.g., "good", "better", "best") separately from the threshold problem of discouraging or eliminating negative qualities (e.g., pollution). From a modeling viewpoint, the most useful distinction is probably between gradient problems and threshold problems, and I would probably adopt this taxonomy in a more formal analysis. However, that is not the only distrinction: there are also intersecting issues concerning information asymmetries: who knows the quality of a contribution, and when?

Wikipedia is produced by the Wikipedia Foundation, a non-profit that does not publish advertising, and that to date has had (donated) revenues sufficient to employ no more than ten people. Without contributed content, there is little doubt that a minimally-funded foundation could produce an encyclopedia with more than 2 million English entries, and versions in over 253 separate languages [Wikipedia contributors, 2007].

However, I do not find this argument by itself very compelling. First, it is not obvious that any or many socially valuable products or services could not recover the costs of their production. Assuming the social benefits from use of Wikipedia are larger than the social costs (for the most part, of the time spent by writers and editors), it is reasonable to think that generally there will be ways to extract enough of the surplus to cover the costs of production. In the Wikipedia example, one obvious mechanism to consider is site advertising, such as that found on Google or Yahoo! search result pages. Certainly the many for-profit firms obtaining user-contributed content (e.g., Amazon) have means for extracting value from users of the content.

I think the circumstance that usually distinguishes procurement of user-contributed content from other production inputs is the transaction cost of providing direct, extrinsic compensation to contributors. Most prominent examples of user-contributed content involve relatively small contributions by a relatively large number of contributors.[7] Providing direct, extrinsic compensation requires establishing a reliable business relationship, typically including a contract, a monitoring and accounting process, and a payment system. This is more costly than the already difficult problems that challenge micropayments systems: before making payment it is necessary to agree on terms and condition, and to measure the amount (and quality) of the provided content. It seems reasonable that the costs of setting up supplier relationships and managing them with all of the individuals who provide book reviews might exhaust the benefits to Amazon (or Netflix, or Flickr, &c.) from the contributed content.

Thus, when the number of contributors is large and the value of their contributions relatively small, the producer is faced not with a question of how much to pay content providers, but whether it is possible to induce contributions without compensation, or do without the content altogether.

[7]Wikipedia reports over 5.8 million registered writers and editors, and also permits anonymous, unregistered users to edit [Nov, 2007]. The number who have edited at least once appears to be about 4 million [Almeida et al., 2007].

Finally, before proceeding to quality issues, it might seem that in some cases there is no quantity problem, that is, that producers do not need to do anything to induce contributions. For example, scholars want their research articles to be published, distributed and read, and generally will create the content and offer it (along with copyright) to publishers. However, as long as there is some competition among publishers, they still have a quantity problem: a given publisher needs to induce content creators to contribute to it, rather than to a competitor.

4.2 Quality

Quality is a vague label that encompasses a complex array of characteristics, whether the factor be information or any other product or service. We often distinguish between "vertical" and "horizontal" qualities. Vertical refers to characteristics on which everyone agrees in their preference ordering: a monotonic gradient from worst to best [Mussa and Rosen, 1978]. An example for information might be its degree of objective accuracy. Horizontal refers to characteristics on which preference orderings vary: some people prefer red, some blue, and a few chartreuse [e.g., Lancaster, 1966].[8] One central problem for the producer is to induce contributors to provide a desirable mix of horizontally- and vertically-differentiated qualities.

The best method for managing the mix of contribution qualities critically depends on observability (or, if quality is to be contracted upon, verifiability). If the producer costlessly can observe input quality when provided, then the problem is largely one of production planning: what mix of qualities is desirable (e.g., will support the most profitable output mix)? For example, what is the distribution of article accuracy a newspaper publisher wants to maintain, and what diversity of viewpoints does she want to express? If the qualities are identified, the problem is reduced to multiple contribution quantity problems. To see this, label information of different qualities (vertical or horizontal) as different goods: the producer needs to induce contributors to donate the desired quantity of different types of information inputs.

Suppose, however, that quality characteristics are not (costlessly) observable (*ex ante*, referring to the point at which a transaction is consummated) or verifiable (*ex post*). That is, suppose there is asymmetric information of

[8]Still others care deeply about the difference between chartreuse green and chartreuse yellow, see `en.wikipedia.org/wiki/Chartreuse_(color)`.

the hidden characteristics type [Stiglitz, 2000] about the quality dimensions of contributed content. The problem of obtaining the desired quantities of different quality types becomes more difficult, since the producer also needs to learn what the qualities are. This is the problem of "evaluating the stuff."[9]

A related but somewhat different problem also is important when there is asymmetric information about quality. Offering a user-contributed content product can be much the same as offering a free publication and distribution vehicle. As long as the producer can observe quality costlessly, the use of this vehicle can be limited to the publishing and distribution goals of the producer. When quality is unobservable, content contributors have an opportunity to hijack the platform for their own publishing ends. For example, when a blog writer allows readers to freely comment on posts, the writer soon discovers that many comments posted are commercial advertisements for unrelated products, that is, spam. Thus, with asymmetric information the quality problem is not just one of obtaining enough of the desired quality mix, but also of preventing (or otherwise managing) contribution of undesirable content. This I call the problem of "keeping the bad stuff out."

Let me summarize. Producers of information products or services may choose to obtain some inputs to production in the form of user-contributed content. This falls into the familiar domain of obtaining factors of production. By seeking this content without directly compensating the contributors, however, the producer faces several challenging problems. First, she needs to obtain the desired quantity of inputs (with the understanding that content of different observable qualities can be thought of as different types of input). Second, if quality is not costlessly observable, she needs to evaluate contributions, either directly (to then select which contributions to use), or indirectly (by inducing contributors to provide a desirable mix of qualities). Third, because her product is attractive to others as a low-cost publishing platform, she needs to discourage or prevent contributions of material that is harmful to her goals. Collectively, I summarize these challenges as "getting the good stuff in" and "keeping the bad stuff out".

In the remainder of this article, I will discuss two of these three problems: getting the good stuff in and keeping the bad stuff out. These problems

[9]The evaluation performed may be direct, by the producer, or indirect, if the producer induces contributors to self-select the quality they provide, or otherwise honestly reveal their quality information. Many of the mechanisms economists design for heterogeneous quality problems assume that monitoring is impossible or too costly to be worthwhlie, and thus focus on inducing the contributors to self-evaluate.

naturally arise because the direct rewards for contributing, and (largely to minimize *disincentives* to contribute) the protective walls are also low. Evaluating is a complex and wide-ranging problem better treated elsewhere.[10]

4.3 Getting the good stuff in

If the individuals are self-interested, then there is an intrinsic incentive to free-ride. Why and when would autonomous individuals take actions that benefit others without being (directly or explicitly) compensated for providing the benefit? Why, for example, does someone take the time to write a detailed book review and post it to the Amazon site?[11] Why make the effort to label and then share bookmarks with others (`del.icio.us`, `Markaboo`, `furl`, etc.) [Rader and Wash, 2006]? In other literatures this is known as the problem of community contribution: how to get users to actively participate in, for example, wikis (open-access group written documents), or open source software development projects [Finholt, 2002, Olson and Olson, 2003]? Economists refer to this as the private provision of public goods [Bergstrom et al., 1986]

The theory of the private provision of public goods depends on participants receiving some degree of intrinsic personal benefit (e.g., garnering reputation or experience), which sometimes may be sufficient incentive. Lerner and Tirole [2002] propose this motive to explain programming effort donated to open source software projects: programmers obtain experience in a professional group development setting, and their work is documented so they can show potential employers their accomplishments. Resnick et al. [2006] found through a human-subject controlled field experiment on eBay that in public, semi-anonymous Internet settings reputation-garnering provides meaningful incentives for socially beneficial behavior.[12]

[10]The topic of evaluation invokes economic research on reputation, recommendation, and signaling and screening, among other topics. I am not aware of economics literature that discusses these issues as they apply to user-contributed content, but for related work that starts the conversations, see Riley [2001], Resnick and Varian [1997], Resnick et al. [2000]. I discuss signaling and screening below as they apply to keeping out undesirable content.

[11]Harriet Klausner has contributed over 15,000 reviews to Amazon. Quite a few others have contributed more than 1000 [Amazon.com, 2007b].

[12]This was not an experiment in the private provision of a public good in the sense of contributing inputs to joint production. Rather, the authors were testing a reputation system's effectiveness at generating motivation for private behavior. See also Resnick and

Another example in which users derive some direct personal use benefit from their contributions is the photo storage and sharing site, Flickr (owned by for-profit Yahoo!). The number of photos taken about which people do not have proprietary concerns (for either perceived marketable value or privacy) appears to be astronomical (over two billion to date, since inception in 2004, Reuters [2007]. One direct user benefit is that Flickr provides free storage; another, perhaps more important (since local hard disk space is getting very inexpensive) is that this storage is accessible from any network-attached computer (so users can, for instance, see their photo collections while they are at work, or while traveling). A related benefit is that Flickr makes it easy to share photos with friends and family (which I consider to be a direct personal benefit, different from the benefit that unknown users receive from access to one's photos). Flickr provides a variety of organizational tools to help users manage their collections (such as tagging, which is simpler and faster than categorizing according to a controlled vocabulary). By making contribution and metadata provision easy, Flickr keeps the net contribution cost sufficiently low that vast numbers of amateur photographers are willing to contribute content. Rader and Wash [2006] document that direct private benefits are an important motivation for `del.icio.us` contributors, who use it to maintain a portable, tagged index of web site for their own use.

Attention to the cost of contributing, as well as to the benefits, likely is important for inducing user contributions. MacKie-Mason et al. [2000] found that the costs of using a standard, relatively simple interface to an online scholarly journal service can be as important as typical intellectual property fees for the service. Particularly in settings in which it is difficult to provide direct extrinsic incentives such as monetary transfers, it is natural to attend more to lowering the costs of participating (or of providing higher quality).

Other motivations for contribution likely are important as well. Social psychologists attend to the role that individual conceptions of social identity and norms of collective effort play as intrinsic motivators. For example, in the CommunityLab project (`communitylab.org`) a team of economists, social psychologists and computer scientists has been testing mixtures of economic mechanisms, computer algorithms, and user interfaces to provide improved incentives for participation and quality, both by increasing intrinsic motivations and by lowering offsetting costs of and barriers to participation

Zeckhauser [2002], Dellarocas et al. [2004].

[Ling et al., 2005, Harper et al., 2005, Rashid et al., 2006].

Economic theory suggests that relying solely on intrinsically-motivated private provision of public goods will result in socially-inefficient underprovision, and thus mechanisms to induce greater contributions to the public good will increase social welfare. A large body of theoretical and experimental research specifically addresses the design of extrinsic motivation mechanisms to induce greater philanthropic contributions [see, e.g., Andreoni, 2006]. There is little work yet on the transferability of these results to settings in which the contributions are in content, not cash (and thus in which quality characteristics are also important), and even less work on ways to increase intrinsic motivations sufficiently to maximize social welfare.

4.4 Keeping the bad stuff out

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.[13] For example, anyone can post a book or product review to Amazon, and this content will appear to all users shortly thereafter immediately [Amazon.com, 2007a]. This system provides 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 lead 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).

As mentioned above, this is a common problem for user-contributed content because two conditions often hold: some agents want to contribute con-

[13] One of the interesting responses to low-quality content is to solicit user-contributed *meta-content*. For example, Slashdot and many other discussion or commentary fora do not select or edit content, but let users rate contributions so that readers can filter and see only content that other users have rated above a chosen threshold. I apply this idea in my example model below, Section 5.

tent 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 high transaction costs, so 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 their creators.

Bargaining costs, on the other hand, are often significant. As I discussed above, 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. Contracting, monitoring and transacting payments with hundreds or thousands of microsuppliers to internalize their externalities typically will be too costly to support. We may not need to look beyond 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. That is, the producer wants to induce productive content contributions but simultaneously 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 not a problem, 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 prob-

lem: how to distinguish good from bad content at low cost and before the pollution has imposed a cost on the producer?

For example, one way to reduce the amount of contributed pollution would be to charge contributors a fee for publishing their content, much as some governments tax polluters per ton of effluent. However, to succeed such a system would necessarily need to monitor content and distinguish between pollution and good content, because the producer does not want to charge suppliers of good content.[14] If monitoring is low cost, such a system (or even simply blocking polluting content) might work, but in many settings the quantity of information contributed is so large and the qualities so diverse that monitoring is too costly to fully solve the problem. 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]. Evidently the cost of detecting this kind of misrepresentation is sufficiently high that Amazon chooses not to monitor enough to keep all pollution of this sort from its site.

Thus, 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.

Loder et al. [2006] proposed a signaling device to reduce unwanted bulk email advertising (spam), which they called an "attention bond mechanism". The identifying characteristics of costly-to-observe pollution are present: recipients incur a cost (reading time, security risk) to determine whether the message is desired, and the sender generally has a better prior estimate of the message's expected value. Email senders are required to post a revocable bond. Given a greater likelihood that a recipient will claim a spammer's bond, and appropriate 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 suf-

[14]Of course, the producer would be delighted if it could obtain desired quantities and qualities of its content inputs while simultaneously charging the suppliers for providing the inputs, but rarely is a business sustainable if it charges its suppliers for the privilege of providing inputs. Vanity presses are one (small) niche in which this model is sometimes successful. Of course, fictional Tom Sawyer was paid handsomely by friends who wanted to whitewash his fence; see footnote 2, above.

ficiently lower that they will be willing to comply and send email, whereas bad senders will be discouraged.

Ben Chiao and I also addressed the email spam pollution problem, but with a somewhat novel economic mechanism [Chiao and MacKie-Mason, 2006]. We show conditions under which offering an alternative, even lower cost (and higher quality) platform for advertisers to distribute their spam might induce enough to divert their advertising to the alternative channel, simultaneously lowering the value of the regular email channel for other spammers, until all spam has moved elsewhere.

Rick Wash and I showed that password authentication to information resources works to the extent the passwords satisfy the necessary economic condition for effective screening mechanisms [Wash and MacKie-Mason, 2007] . We also demonstrated that the increasingly popular CAPTCHA system is a screening device, and its success depends on satisfying the necessary economic conditions for screening.

When user-contributed pollution is a problem in repeated interaction environments, a *reputation mechanism* might be an appropriate design response. A reputation mechanism works like an implicit bonding system: in an equilibrium in which a better reputation leads to higher value transactions (on average), the cost of foregoing those value premia in the future by sacrificing reputation now with a polluting action may be higher than the benefit of the pollution. Of course, for a reputation or other implicit-bonding mechanism to work in a repeated interaction environment, it is necessary to have persistent (pseudonymous) identifiers for contributors [Friedman Eric J., 2001].

5 Example: Keeping the bad stuff out

I will illustrate the application of incentive-centered design concepts to the problem of keeping undesirable content out of an information service that relies on user-contributed content. For concreteness, I situate this application in the context of `Digg.com`, which is a news aggregation service.[15] Users submit news stories, and then readers vote for those they think are worthy of attention. The stories with the top vote counts during a time interval are promoted to the "front page" which users see first. Display on the front page

[15]There are several similar services. `Slashdot.org` is perhaps the best known in the information technology community; `Reddit` is also popular.

typically leads thousands to go read the referenced news story (the "Digg effect").[16]

One of the interesting features of Digg is that it obtains as production inputs two levels of user-contributed content: the posted news stories, and the voting metadata. Over the past two years there have been numerous assertions that some users on Digg have been polluting or manipulating the voting system in order to manipulate the promotion of their preferred stories to the front page. One of the more interesting examples concerns a story that Google was about to announce its acquisition of Sun Microsystems [Green, 2006]. Whether there were serious discussions of this possibility at one or both companies I do not know, I am unaware of any confirming stories, and in the event, such an acquisition did not occur. It is not clear if this was a stock price manipulation scheme, but after reaching the front page of Digg, the story was picked up by mainstream media with greater reach, and it appeared there was a stock price response to the rumor. Whether or not this was a successful manipulation of Digg through voting pollution, it has been widely discussed as a possible problem with Digg reliability, along with several other incidents mostly involving efforts to publish self-promoting materials for personal commercial gain [Sandoval, 2006].

Digg makes money by selling site advertising, and depends on being able to draw users to see the ads, which in turn no doubt depends on (among other things) the extent to which users have confidence in the veracity, quality and independence of highly rated news stories. I will now present a simple model of Digg's problem with an incentive-centered system design which in this overly simplified, illustrative setting, provides a reasonable solution for Digg. Many of the specific concerns about Digg have focused on the voting system through which users elevate the visibility of stories, but for this example I focus on the underlying issue: the quality of the stories submitted to Digg in the first place (before user voting).

5.1 A model

Suppose there are two types of content suppliers: Truthtellers and Liars, which I will indicate with subscripts: $\tau \in \{T, L\}$. Truthtellers always post true stories, and they have no personal stake in the content of the story. Liars

[16] See., e.g., Fadial [2007], who documents an increase from about 200 daily visits to 2000 hourly visits to his web site after his article was promoted to the top of the Digg front page.

always post stories that are less desirable (to Digg, and to its readers) than stories posted by Truthtellers. Suppose that in the population of potential suppliers there is a fixed and known fraction of Truthtellers, denoted by $0 \leq p \leq 1$. When suppliers contribute an article, they experience some benefit (indirect, non-extrinsic motivation for contributing); for simplicity I assume all Truthtellers experience a value of v_T from contributing a story, whereas Liars experience a value of v_L. Digg benefits, presumably, by drawing readers to the web site, and then earning advertising revenues; I assume that in equilibrium, stories posted by Truthtellers generate more viewing traffic and thus more profit for Digg (say, because some users who discover they are reading low quality stories posted by Liars do not return to Digg in the future). The expected net value to Digg of a posted story is π_τ, with $\pi_T > \pi_L$. Depending on the actual quality of stories posted by Liars, and on the severity of the reaction by Digg's readers, π_L could be positive or negative; for the analysis below it is only necessary that it be less valued by Digg than Truthteller contributions.[17] I will show below that whether π_L is positive or negative has interesting implications for the *interpretation* of the results.

Now suppose that Digg implements the following practice: to contribute (post) a story, a supplier must deposit a fee per story of ϕ into a Digg-managed account. Digg then encourages readers to cast a binary vote on the "truthfulness" of each story they read: "truthful" or "not".[18] I denote the ith vote by

$$\alpha_i = \begin{cases} 1 & \text{if "truthful"}, \\ 0 & \text{if not.} \end{cases}$$

Digg then calculates a score, $\rho(|\alpha|, \sum \alpha_i)$, for each contributed story, with $|\alpha|$ the cardinality (count) of the set of votes, and $\sum \alpha_i$ the number of votes

[17] It makes sense to allow stories contributed by Liars to have positive value most importantly because "Liar" was chosen merely as a catchy label: the stories may have lower quality, but not actually be lies. Of course, even stories that are lies (or that are supplier-self-serving advertisements) may have positive value to Digg: news that is false may attract repeat readers as long as it is entertaining or controversial enough, as we know from the longevity of various tabloids.

[18] For this model, it does not matter whether the content of a story is actually truthful or not; simply that there are two vertically-differentiated qualities, and so that Digg values the L quality less than the T quality.

for "truthful":

$$\rho = \begin{cases} 1 & \text{if } |\alpha| < 20, \\ \frac{\sum \alpha_i}{|\alpha|} & \text{if } |\alpha| \geq 20. \end{cases}$$

Every story starts with a score of one ("truthful") until some threshold number of votes are submitted (set arbitrarily to 20 in this example). At that point, the score is simply the ratio of positive votes to total votes. After some specified period of time, Digg returns the deposit (ϕ) if the score is above a threshold, $\rho > \bar{\rho}$, with $\bar{\rho}$ determined and announced in advance by Digg.

Given this mechanism, Digg's incentive design problem is to choose $\{\phi, \bar{\rho}\}$ to maximize its expected value function. To calculate this, we need to model the probability distribution of scores. For a story contributed by a supplier of type T, there will be some probability that a reader correctly will rate the story as truthful, a residual probability that the reader incorrectly will rate the story as not; similarly for stories contributed by type L suppliers. From these primitive probabilities, Digg can arrive at an estimate of the probability distribution that generates an observed score (when the number of votes is above the threshold). If the votes are drawn from i.i.d. distributions, then the score distribution will have the form $\rho_\tau \sim f(\mu_\tau, \sigma_\tau^2(|\alpha|))$, in which the variance is decreasing in the number of votes ($|\alpha|$). To simplify, I henceforth condition on the number of votes made as of the moment of the analysis, and suppress that argument. I assume that the votes are weakly informative signals and are positively correlated with the actual value of the story, so on average the score for a story contributed by a Truthteller will be higher than the score for a Liar: $\mu_T > \mu_L$. The cumulative distribution function is $\Pr(\rho \leq \bar{\rho}|\tau) \equiv F_\tau(\bar{\rho})$, which is the probability that a supplier loses its deposit (that is, that the score is below threshold). I assume that not only is the mean score higher for Truthteller contributions, but that the distribution of scores for Truthtellers first-order stochastically dominates the distribution for Liars, so that for any threshold $\bar{\rho}$ the probability of a score below threshold is greater for contributions posted by Liars: $F_L(\bar{\rho}) \geq F_T(\bar{\rho})$.

Now I can write down Digg's design problem. Content suppliers receive expected utility $EU_\tau = v_\tau - \phi F_\tau(\bar{\rho})$, and Digg solves

$$\max_{\bar{\rho},\phi} p(\pi_T + \phi F_T(\bar{\rho})) + (1-p)(\pi_L + \phi F_L(\bar{\rho}))$$

$$\begin{aligned} \text{s.t.} \quad v_T - \phi F_T(\bar{\rho}) &\geq 0 \quad \text{(Participation Constraints)} \\ v_L - \phi F_L(\bar{\rho}) &\geq 0 \end{aligned}$$

if both types $\tau \in \{T, L\}$ contribute. The maximand is a straightforward statement of Digg's expected value per contributor, given a fraction p of Truthtellers in the population, the value from stories of types T and L, and the fee (ϕ) and deposit return likelihood ($F_\tau(\bar{\rho})$). The participation constraints reflect the fact that suppliers can decline to contribute if their expected utility is less than zero.[19] The participation constraint corresponding to that type does not need to be satisfied if Digg is content with that type not participating.

5.2 Results from introducing user-contributed (meta)content

The problem above is a fairly standard incentive or mechanism design problem, and can be solved straightforwardly by calculating the first- and second-order conditions and manipulating them. I will summarize the results here, and sketch their intuition (full proofs available upon request). Let me make one more restrictive assumption; I will discuss its relaxation after presenting the main results. Assume, for Results 1–3, that the intrinsic value to a supplier of contributing an article is the same for Truthtellers and Liars: $v_T = v_L$.

Result 1 *The participation constraint is binding for Liars:* $v_L = \phi F_L(\bar{\rho})$.

This is a standard result in hidden characteristic mechanism design problems. From the Participation Constraints above, and given the assumptions of first-order stochastic dominance and $v_T = v_L$, we see that the expected utility of a T type is always at least as large as that of an L type. Thus, if the participation constraint for Liars was not binding, Digg could, for example, raise the deposit ϕ until it does bind, without changing the participation decision by either type, yet increasing Digg's expected value (since Digg keeps the forfeited fees). In the mechanism design literature, this result is often referred to as "no information rent" for the less desirable type: if Digg is to share surplus with either type, it should be with the type that provides the more valuable content.

Result 2 *The participation constraint is not binding for Truthtellers:* $v_T > \phi F_T(\bar{\rho})$.

[19] The reservation utility, or "outside option" available to suppliers may be worth something different than zero, but we can normalize to zero without loss of generality by adding or subtracting a constant to or from v_τ on the left side of the participation constraint.

Truthtellers keep some rent; that is, Digg shares more of the value created by Truthteller stories than is necessary to induce Truthtellers to contribute. This is a standard result in mechanism design, and can be explained with the following intuition: Whether a particular supplier is a Truthteller or a Liar is *ex ante* unknown to Digg. Therefore, Digg must treat all contributors the same (*non-discrimination*); in particular, ϕ and $\bar{\rho}$ cannot depend on the supplier's type. Thus, from the Participation Constraints, if Liars get zero expected utility (Result 1, Truthtellers must get positive expected utility. Digg can only extract more value from Truthtellers by putting them on a different contract, and but since Truthtellers can pretend to be Liars, they would not admit they are Truthtellers and sign a different contract unless it yielded as much expected utility to them as does being on the same contract as Liars. In other words, Digg would have to pay them this amount of rent to get Truthtellers to admit their type anyway, so there is no gain to Digg from trying to differentiate between them.

Result 3 *If $\pi_L + v_L < 0$ then set $\phi\bar{\rho}$ marginally higher to exclude participation by Liars.*

I mentioned above that the model is flexible enough to handle content contributions from Liars that have positive value for Digg, as long as the value is less than for Truthteller contributions. But it also accommodates situations in which the value from Liar contributions is negative ($\pi_L < 0$): what I above called "user-contributed pollution". If the value is negative enough, so that $\pi_L + v_L < 0$, then Digg will actually want to exclude the Liar's content altogether.[20] It can do so simply by either increasing the fee (ϕ), or the score threshold below which the supplier forfeits the fee ($\bar{\rho}$); it turns out that these two parameters are only determined up to a multiplicative constant, so that Digg is effectively choosing their product ($\phi\bar{\rho}$).

Why is the critical test for whether to let Liars contribute whether $\pi_L + v_L < 0$? Put another way, why ever let Liars participate if the value they create for Digg is negative, that is, whenever $\pi_L < 0$? First, it is straightforward to see why the result emerges: by Result 1, $\phi F_L(\bar{\rho}) = v_L$. Make that substitution into Digg's maximand. The only place that π_L and v_L appear is as a sum in the second expression, and indeed this expression becomes a

[20]Below I explain why $\pi_L < 0$ is insufficient for Digg to want to exclude Liar contributions.

constant. Since $(1-p)$ is non-negative, Digg is better off by having Liars participate as long as $\pi_L + v_L > 0$.

The intuition is also straightforward, and revealing. The expression $\phi F_L(\bar{\rho})$ is the amount of its deposit that the Liar expects to forfeit to Digg. This has the obvious interpretation of a payment from Liars to Digg. As long as $-\pi_L < v_L = \phi F_L(\bar{\rho})$, the Liar is in expectation paying Digg more than the cost that it imposes with its negative-valued content contribution.

We have arrived at a sensible, intuitive result, one we normally call *advertising*! The Liar type gets some benefit from publishing its content to DIgg. Digg itself may get negative value directly from the content, but will be willing to publish it if the price is right. Indeed, this is what Digg does already with some content: it names a price to buy advertising space, and will display material that typically lowers the value of the site for Digg's users on behalf of advertisers willing to pay the price (as does Google, etc.).

What is different from the current design of Digg and similar information products or services? Although Digg does offer advertising space for sale, it also offers an open, free publishing platform to anyone who wants to submit "news stories". Just as our inboxes are open publishing platforms and product sellers stuff them with their ads rather than limit themselves to paid channels like newspaper advertising, so do some contributors to Digg choose to publish their advertisements (or other self-serving content) rather than pay for advertising space. They can do this because Digg, apparently, has concluded that hiring employees to "evaluate the stuff" that is contributed, in order to screen out ads, is too costly to justify. The mechanism described above provides an evaluation method that is less expensive. In fact, it is a method that relies on user-contributed (meta)content as an input to production of the monitoring: the votes submitted by content users. Combining user-contributed (meta)content with direct extrinsic *penalties* (rather than positive compensation) offers the promise of a low-cost, robust system for monitoring quality, and enables Digg to either exclude the bad stuff or, if if the suppliers of bad stuff value publishing it sufficiently, to charge a price for publishing the material as ads.[21]

[21] If the current "truthfulness" vote score is displayed alongside the article, much as the current "Digg" score is displayed to indicate a story's popularity, then potential readers will have an indicator of whether the content they are considering is advertising or not.

5.3 Extending the model

This is a very simple model, with many issues to be addressed before before implementing such a scheme. One that is now quite obvious: adding only a second dimension of quality distinction to Digg's own "up or down" (popularity) voting does not cover all of the quality issues likely to concern the producer. I labeled this dimension as "truthfulness", somewhat playfully but also inspired by the Google-Sun merger "news" story posted on Digg. The particular mechanism, however, distinguishes material that readers vote down on some (ambiguous) criterion ("truthful", "not self-serving", &c.), which the producer then publishes if the contributor is willing to pay more than the average cost to the producer of including content of this type. Of course, a producer might be willing to publish commercial advertising for a price, but still prefer not to publish advertising that is false or misleading, or racially insensitive, &c. Thus, we can see opportunities to entrain users to contribute meta-content that ameliorates some of the problems of an open publishing platform, but more sophisticated solutions for multiple dimensions of quality become increasingly costly and unwieldy. User-contributed editorial services may be an important element in future information services and publications, but they are not a panacea.

I have also set aside obvious questions about the motivations of the users. In the model above, I assume that a sufficient number of readers will be motivated to submit votes. I further assume that those who vote will for the most part be motivated to vote honestly.[22] Without some motivation for voting, and for honesty, the score will not be a very good (perhaps even perverse) signal of the underlying type of the contributed content, and the mechanism will not work well. Thus, there is another incentive-centered design problem to be addressed: how to motivate contribution and quality by the user contributors of meta-content ("truthfulness" votes). Gazzale [2005] provides results for a related problem in reputation systems: how to motivate consumers to rate (and rate honestly) service provided by sellers?

Imagine there are enough motivated voters, but they do not reliably vote their true opinion. For example, let there be a sizable number of naysayers, such that the score for "truthful" articles is expected to fall below the threshold 50% of the time. It might be quite hard to get good contributors

[22] A closely related question is how accurate the readers are in their assessments of the intrinsic "truthfulness" of the posted story: giving honest reports that are too often incorrect will be a problem as well.

to provide stories if they believe that they will have to pay a fee of $\phi/2$ on average for posting. Of course, the threshold can be adjusted if there is a systematic bias towards nay-saying, but the fundamental problem remains: the signal must be sufficiently discriminatory (in statistical decision parlance, have low enough Type I error *and* low enough Type II error) to be able to keep the expected cost of posting low for good articles, and sufficiently high for bad articles.

In fact, I assumed that all Truthtellers had the same preferences as each other, and likewise for Liars. In this simplification, we need only find a vote scoring scheme and expected penalty that separates the two groups: all Truthtellers are willing to participate, and all Liars either are not willing, or are willing to pay the (advertising) fee to participate. In a real system we will expect different Truthtellers to have different willingness to participate (indicated in the model by v_T), and likewise for Liars; then the more challenging task will be to construct a combination of vote scoring rules and fee level that generates a desirable mix of Truthteller and Liar content (that is, not discouraging too much of the former, and not overlooking too much of the latter).

I made another simplifying assumption just before stating the results above: that the benefit from posting an article is the same for Truthtellers and Liars ($v_T = V_L$). Just as I allowed in this discussion above that different Truthtellers might value posting differently, so, of course, may Truthtellers and Liars differ in their desire to post. If the willingness to post differs (I will just discuss the simpler case in which that willingness is common across all contributors of a given type), nothing material changes as long as Truthtellers care more about posting than do Liars: $v_T > v_L$. The problem becomes somewhat easier for the producer, in fact, because now a given fee (will relatively less likely to dissuade Truthtellers and more likely to dissuade Liars. However, things can get more complicated if Liars have a higher willingness to post, which we might expect to often be the case (if, for example, Liars stand to gain commercially from posting their advertisements): $v_L > v_T$. If the gap is not large, but results above still hold, but if the gap gets large enough then the producer faces the problem known as "countervailing incentives".[23] For intermediate gaps between v_L and v_T, it becomes desirable to raise the expected fee (through some combination of raising ϕ and low-

[23]Wash and MacKie-Mason [2007] discuss the implications of countervailing incentives in a related application of incentive-centered design to security systems.

ering the threshold for a rebate, $\bar{\rho}$), which reduces the content provided by Truthtellers (in extreme cases, perhaps to zero), thus reducing the value of the overall system. As the gap gets large enough, this system will not work very well and the producer will choose to set parameters so that more and more undesirable content gets in (rather than shut out all desirable content). However, at that point, if the value to bad contributors of getting in is high enough, it can become possible to identify them and use other designs to reduce the value to them of participating (e.g., isolating their stories to a special quarantine that suggests a pejorative connotation and reduces their financial gain from posting the stories).

6 Incentive-centered design for user-contributed content

User-contributed content as an input to the production of informations services or goods is not a new phenomenon, but it is growing rapidly in significance and prevalence. In this article I proposed an economic characterization of user-contributed content, and identified a variety of economic issues arising from its characteristics. In a descriptive analysis, these are issues that would lead us to expect socially inefficient under- or overprovision of user-contributed content, or inefficient mixes of quality and variety. I turn to a normative, design perspective: given that basic market forces do not automatically lead to efficient quantity or quality when a producer relies on user-contributed content, in what ways might we design information services or goods to ameliorate these problems? The latter is the problem of *incentive-centered design*; in this article I focused on economic incentives, but I hasten to add that social and psychological incentives may be equally or more useful to the designer.

The characterization of user-contributed content I use is that it is information content provided as an input to product (generally of some information good or service), provided at least in substantial part by people who receive no *direct extrinsic* compensation: for example, no contract specifying quantities and qualities (or even hours of creative effort), and no cash payment per unit of quantity or quality.

A producer who wishes to obtain user-contributed content inputs — that is, to engage in "Tom Sawyer production" — will need to face a variety

of economic problems. First, how to obtain the desired quantity of input, when supplier are not (directly and extrinsically) paid for their supply efforts? Second, given that some content is contributed, how to manage the mix of qualities, especially given the pervasive asymmetric information problem known as *hidden characteristics*: the suppliers know both the horizontal and vertical quality dimensions of what they are contributing better than does the producer obtaining this content as an input. Third, given that services relying on user-contributed content often function as open publishing platforms for the contributed content, how is the producer to deal with pollution, such as advertising content that has zero or even negative value to the producer, but positive value to the content supplier? Collectively, I refer to these problems as *getting the good stuff in*, and *keeping the bad stuff out*.

I worked out an example to illustrate how incentive-centered design might address some of these problems in a specific context. I focused on a problem of "keeping bad stuff out", that is, information pollution. In so doing, I allowed extrinsic payments to be sometimes made (though perhaps never or rarely in equilibrium), in the form of posting a bond that would be returned if other users collectively determined that the contributed content was not polluting. The particularly interesting feature of the proposal is that the producer solicits yet a second type of user-contributed (meta)content — ratings on a particular quality dimension of user content contributions — and uses this meta-content as an input to the editorial function of selecting or categorizing information inputs. Thus, the potentially enormous task of screening content (often supplied by thousands or even millions of suppliers in rather small chunks) may be outsourced to volunteers: user contributors of editorial screening.

The example addresses but one of the several design problems I identified, and only a special case of that problem (pollution), and in a very stylized setting. There are considerable opportunities for further work on incentive-centered design for user-contributed content services.

References

R.B. Almeida, B. Mozafari, and J. Cho. On the evolution of wikipedia. In *Proc. of the Int. Conf. on Weblogs and Social Media*, 2007.

Amazon.com. General review creation guidelines, 2007a. URL

http://www.amazon.com/gp/customer-reviews/guidelines/review-guidelines.html.

Amazon.com. Top reviewers, 2007b.

James Andreoni. Philanthropy. In S-C. Kolm and J. Mercier Ythier, editors, *Handbook of Giving, Reciprocity and Altruism*, pages 1201–1269. North Holland, Amsterdam, 2006. URL http://econ.ucsd.edu/~jandreon/WorkingPapers/Philanthropy.pdf.

Theodore Bergstrom, Lawrence Edward Blume, and Hal Varian. On the private provision of public goods. *Journal of Public Economics*, 29 (1):25–49, January 1986. URL http://repositories.cdlib.org/cgi/viewcontent.cgi%3Farticle%3D1024%26context%3Dpostprints.

Benjamin Chiao and Jeffrey MacKie-Mason. Using uncensored communication channels to divert spam traffic. In *Conf on Communications, Information and Internet Policy (TPRC)*, 2006.

L.F. Cranor. P3p: making privacy policies more useful. *Security & Privacy Magazine, IEEE*, 1(6):50–55, Nov.-Dec. 2003. ISSN 1540-7993.

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

Chrysanthos Dellarocas, Neveen Awad, and X. Michael Zhang. xploring the value of online reviews to organizations: Implications for revenue forecasting and planning. In *Proceedings of the 24th International Conference on Information Systems*, 2004.

Tom Fadial. Erratic wisdom, November 2007.

Thomas A. Finholt. Collaboratories. In B. Cronin, editor, *Annual Review of Information Science and Technology*, volume 36, pages 74–107, Washington, D.C., 2002. American Society for Information Science.

Thomas A. Finholt. Collaboratories as a new form of scientific organization. *Economics of Innovation and New Technology*, 12:5–25, 2003.

Resnick Paul Friedman Eric J. The social cost of cheap pseudonyms. *Journal of Economics and Management Strategy*, 10(2):173–199, June 2001. URL http://ideas.repec.org/a/bla/jemstr/v10y2001i2p173-199.html.

Robert S. Gazzale. Giving gossips their due: Information provision in games with private monitoring. Dept. of Economics, Williams College, 2005.

Heather Green. Digg argues it has ways to prevent manipulation, 2006.

Amy Harmon. Amazon glitch unmasks war of reviewers, February 2004.

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

Lian Jian and Jeffrey MacKie-Mason. Why share in peer-to-peer networks? In *Proceedings of Workshop on the Economics of Networked Systems (NetEcon06)*, Ann Arbor, MI, June 2006. URL http://www.cs.duke.edu/nicl/netecon06/. Available from http://www.cs.duke.edu/nicl/netecon06/.

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

Joshua Lerner and Jean Tirole. Some simple economics of open source. *Journal of Industrial Economics*, 50(2), June 2002. URL http://ideas.repec.org/a/bla/jindec/v50y2002i2p197-234.html.

Lawrence Lessig and Paul Resnick. Zoning speech on the Internet: A legal and technical model. *Michigan Law Review*, 98(2):395–431, 1999.

K. Ling, G. Beenen, P. Ludford, X. Wang, K. Chang, D. Cosley, D. Frankowski, L. Terveen, A. M. Rashid, P. Resnick, and R Kraut. Using social psychology to motivate contributions to online communities. *Journal of Computer-Mediated Communication*, 10(4), 2005. URL http://jcmc.indiana.edu/vol10/issue4/ling.html.

Thede Loder, Marshall Van Alstyne, and Rick Wash. An economic solution to unsolicited communication. *Advances in Economic Analysis and Policy*, 6(1), 2006. Available from http://www.bepress.com/bejeap/advances/vol6/iss1/art2.

Jeffery K. MacKie-Mason, Juan F. Riveros, and Robert S. Gazzale. Pricing and bundling electronic information goods: Experimental evidence. In B. Compaine and I. Vogelsang, editors, *The Internet Upheaval: Raising Questions, Seeking Answers in Communications Policy*. MIT Press, 2000.

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

Michael Mussa and Sherwin Rosen. Monopoly and product quality. *Journal of Economic Theory*, 18(2):301–317, August 1978. URL `http://ideas.repec.org/a/eee/jetheo/v18y1978i2p301-317.html`.

Oded Nov. What motivates wikipedians, 2007. URL `http://opensource.mit.edu/papers/Nov_Wikipedia_motivations_opensource.mit.edu.pdf`. Accepted for publication in the Communications of the ACM.

Gary Olson and Judy Olson. Groupware and computer-supported cooperative work. In J. A. Jacko and A. Sears, editors, *The human-computer interaction handbook: fundamentals, evolving technologies and emerging applications*, pages 583–595. Lawrence Erlbaum Associates, Inc., Mahwah, NJ, USA, 2003. ISBN 0-8058-3838-4.

Tim O'Reilly. What is web 2.0?, 2005.

Emilee Rader and Rick Wash. Tagging with del.icio.us: Social or selfish? In *Computer Support Cooperative Work*, 2006. Poster Session.

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. ACM Press. URL `http://doi.acm.org/10.1145/1124772.1124915`.

P. Resnick, R. Zeckhauser, J. Swanson, and K. Lockwood. The value of reputation on ebay: A controlled experiment. *Experimental Economics*, 9 (2):79–101, 2006.

Paul Resnick and Hal R. Varian. Recommender systems. *Commun. ACM*, 40 (3):56–58, 1997. URL `http://doi.acm.org.proxy.lib.umich.edu/10.1145/245108.245121`.

Paul Resnick and Richard Zeckhauser. Trust among strangers in internet transactions: Empirical analysis of ebay's reputation system, the economics of the internet and e-commerce. In Michael R. Baye, editor, *Advances in Applied Microeconomics*, volume 11. Elsevier Science, Amsterdam, 2002.

Paul Resnick, Richard Zeckhauser, Eric Friedman, and Ko Kuwabara. Reputation systems. *Communications of the ACM*, 43(12):45–48, 2000.

Reuters. Flickr to map the world's latest photo hot spots, November 2007.

John G. Riley. Silver signals: Twenty-five years of screening and signaling. *Journal of Economic Literature*, 39(2):432–478, 2001.

Greg Sandoval. Digg continues to battle phony stories, 2006.

Nigel Shadbolt, Tim Berners-Lee, and Wendy Hall. The semantic web revisited. *IEEE Intelligent Systems*, 21(3):96–101, May/June 2006.

Joseph E. Stiglitz. The contributions of the economics of information to twentieth century economics. *QJE*, pages 1441–1478, Nov 2000.

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

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, August 2007. ACM.

Stuart L. Weibel and Traugott Koch. The dublin core metadata initiative: Mission, current activities, and future directions. *D-Lib Magazine*, 6(12), 2000. ISSN 1082-9873.

Wikipedia contributors. Wikipedia, 2007.

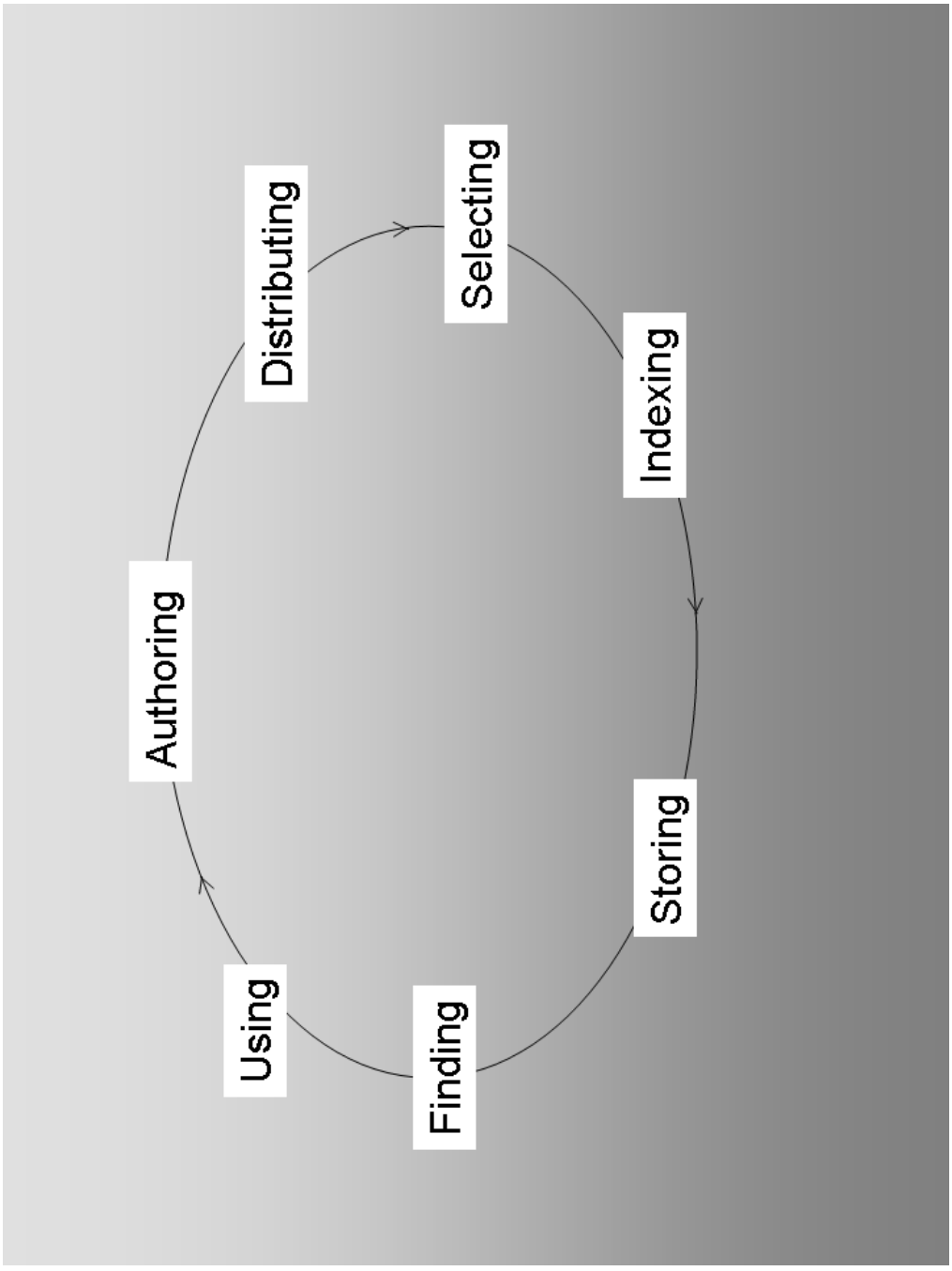

Figure 1: User-contributed content in the life-cycle of information.

Figure 2: Upload bibliographic data to BibSonomy.

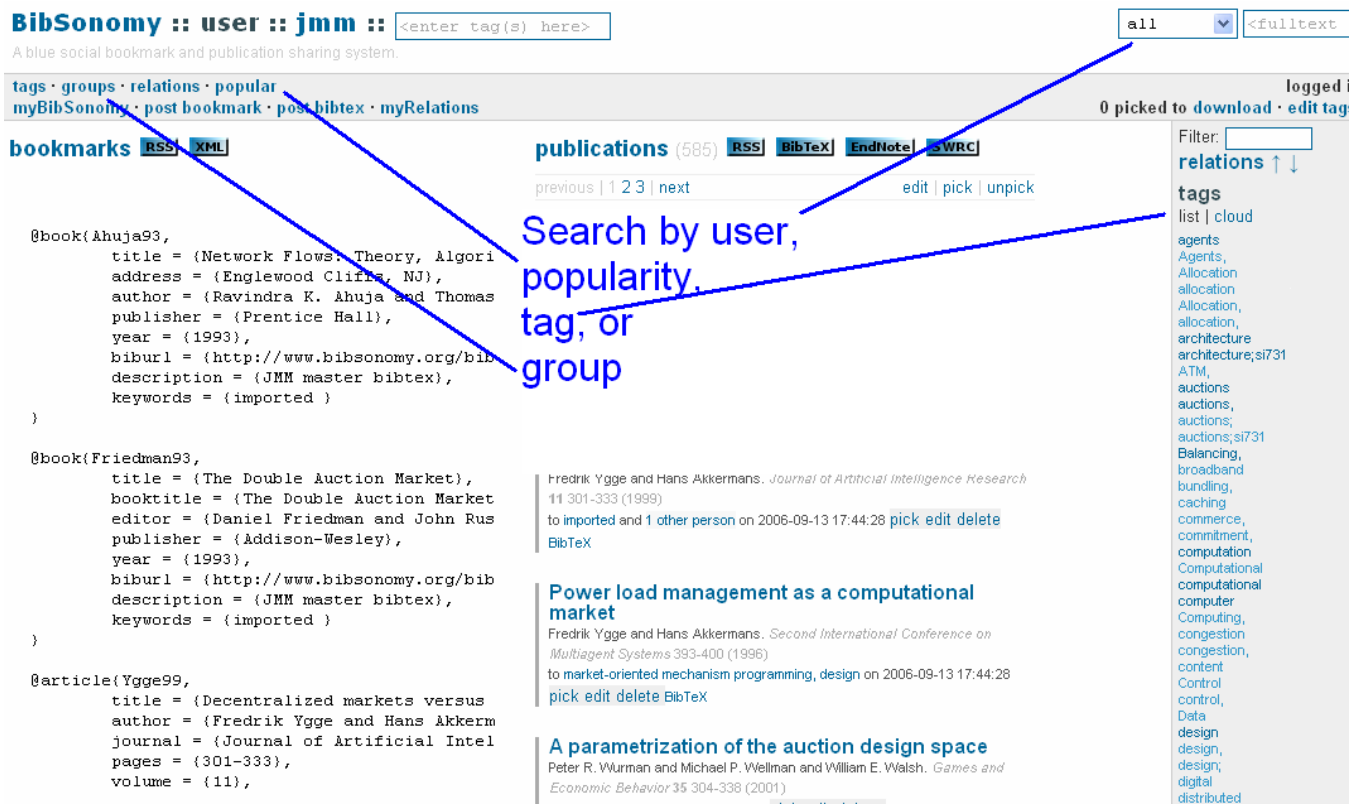

Figure 3: Search BibSonomy.

Figure 4: Select and download from BibSonomy in preferred format.